Index of abundance exploration and development by the
Bluefish Working Group's Fishery Independent Data Group (FIG)

M Celestino, K Drew, S Truesdell, A Tyrell

August 2022

Introduction

The Atlantic States Marine Fisheries Commission (ASMFC) held a meeting of the Bluefish Technical Committee on 3 November 2021 to review state and agency (e.g., Virginia Institute of Marine Science) collected bluefish data (e.g., fishery independent datasets, biological datasets, volunteer angler survey programs) that could be of use in the 2022 bluefish research track assessment. The bluefish working group (WG) subsequently held several meetings to review these data to determine which datasets might be suitable for index of abundance development as part of the 2022 assessment (Table 1). The initial screening criteria for surveys was that to retain a survey, it must have a duration of at least 7 years, 10% positive tows for seine surveys, and 5% positive tows for trawls. We set different thresholds because there were many YOY seine surveys and they tended to have high interannual and intra-annual variation, while the age 0+ trawl surveys were less common and so we lowered the positive tow threshold to retain more of them. The WG formed a subgroup (fishery independent data group, hereafter referred to simply as the FIG) to examine these datasets and develop indices of abundance from them.

Methods

The FIG developed a set of standardized R codes to generate GLM, geometric mean, and stratified mean indices of abundance. For every survey, the FIG developed a nominal (either geometric mean or stratified arithmetic mean depending on how the survey was designed) and a standardized (GLM) index. GLMs required non-missing covariates for every catch record, and so records with missing covariates were dropped from model fitting; since this was not required for geometric or stratified means, all catch records were retained in calculation of those means. The FIG also documented all index exploration and development in a series of R markdown documents (Appendices I and II). Very briefly, the FIG developed a workflow that included (see also Figure 1):

- Data processing: screen for outliers, relevant covariates, highly correlated covariates, relevant months, years with zero positive observations, etc.; standardize continuous covariates to mean 0 and sd =1.
- Model building:
  - Fit full model using all covariates (always include year)
  - Check variance inflation factor and drop any terms > 3 (a measure of collinearity)
  - Include effort as offset (if documented and needed)
  - Fit full model using negative binomial, zero inflated, and zero altered GLMs using the glmmTMB function in R (Mollie et al. 2017; RCT 2022)
  - Some discretion applied with 2-part models (i.e., which terms were included in the conditional and count components)
- Model selection:
  - Fit a suite of models using stepwise procedure to drop covariates by AIC selection (or fit all possible models and compare AICs if not too many terms). Repeat for each model

structure (negative binomial error distribution, zero inflated and zero altered model structures).
    - Convergence issues with zero inflated/altered models resolved by, for example, increasing the maximum allowable iterations of optimizer, dropping covariates that were non-significant in the negative binomial model, dropping all covariates except for year, adjusting terms in just the conditional or just the count portion of the model.
    - Select the best model using AIC, contingent on acceptable diagnostics.
- Diagnostics:
    - Used DHARMa package for residual diagnostics to resolve interpretation of conventional residuals for GLMs; more information on the DHARMa library is available on the DHARMa website.
    - If DHARMa residuals indicated major patterning in the partial residuals of particular variables, drop them from the model; other remedial measures included replacing very large catches with 95th percentile of time series abundance; see also Figure 1. Models with diagnostics viewed as good to sufficient (Figures 2 and 3) were considered final; models with problematic diagnostics (Figure 4) were reconfigured until diagnostics were considered good to sufficient (or excluded from consideration).
    - Check on QQ plot equivalents, residual vs predicted equivalents, and test of uniformity within categorical variables.
- Index preparation: develop index and CV for each survey year.
    - Marginal means, standard errors, and confidence intervals were calculated using a “by-hand” method used by other ASMFC stock assessment committees (e.g., *Brevoortia tyrannus*) that included:
        - Create prediction dataset by expanding a grid of all years, all levels of categorical variables, and the means of continuous variables.
        - For 2-part models, calculate the predicted values for the link and zero-inflated model components of the index for each level of factor combinations using the predict.glmmTMB function. For NB models, calculate predicted values for the link model.
        - Calculate the predicted value of each year, averaged over the other factors (for each component of 2-part models, or link for NB models).
        - Calculate the final index on the response scale.
        - Bootstrap the standard errors and confidence intervals for each year by resampling the annual data (with replacement), re-run GLM, calculate marginal means (as described above), back-transform, then calculate standard deviation and the 2.5th and 97.5th percentiles of bootstrapped runs.
- Document all work in R Markdown documents (Appendices I and II)

For more details on survey-specific index development, see Appendices I (age 0+ surveys) and II (YOY surveys). For a very brief summary of GLM details, see Table 2.

Results and discussion

Please see the main assessment document for brief summaries of each survey used in the present assessment – to minimize redundancy, they are not included in this working paper. Please see the ASMFC Technical Committee state and agency data review report (ASMFC 2022) for even more details related to surveys and their design for all surveys considered (retained as well as those not used) in the present assessment. A chart of the spatial distribution of surveys used in the final model is provided in Figure 5, below.

Figure 6 shows the nominal and standardized indices scaled to their respective means, with 95% confidence intervals. This figure shows that stratified means resulted in trends substantially similar to the GLM indices, with confidence intervals that overlap. Actual index values for the standardized (GLM) and nominal indices are provided in Tables 3 and 4, respectively.

Conclusion

The FIG did not in general observe notable differences between nominal and standardized (GLM) indices and supported use of nominal indices for inclusion in the base model run. An exception to this was the SEAMAP index where index standardization seemed to stabilize some interannual variation in the index of abundance, and included some variables thought to be important (e.g., latitude), especially in recent years where the survey has been truncated due to funding. The WG supported inclusion of all other GLM-based indices as a sensitivity run as part of the present assessment.

Biases

No surveys were designed with the primary goal of catching bluefish; all surveys were designed to either sample multiple or to target species other than bluefish. To select the surveys that were most suitable for bluefish, the FIG set a minimum for % positive tows and minimum for consecutive years of sampling (to eliminate intermittent sampling), consistent with other assessed species (e.g., *Centropristis striata*, *Brevoortia tyrannus*, *Tautoga onitis*). In several instances, indices were standardized to attempt to reduce bias (e.g., SEAMAP), but biases could persist if important factors that affect standardization were not included. In general, the standardized index and the nominal index resulted in nearly identical trends, indicating that the surveys collected fairly representative samples of bluefish and that nominal indices should suffice for the assessment.

Literature cited

ASMFC. 2022. Report from the 2021 Bluefish State Data Review Workshop. Arlington, VA. 115p.

Mollie E. Brooks, Kasper Kristensen, Koen J. van Benthem, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Maechler and Benjamin M. Bolker. 2017. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. The R Journal, 9(2), 378-400.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Table 1. Table of available fishery independent dataset for use in the 2022 research track assessment.

State	Index	Used in 2015?	Use in 2022?	Comments
ME	ME-NH Inshore survey	No	No	Low % pos tows
NH	NH Seine Survey	Yes (in Conn)	Yes (in Conn)	
MA	MA Inshore Trawl Survey	No	No	Explored as part of Conn trawl index
RI	Trawl-Weekly (URI/GSO)	No	No	
RI	Beach seine (Narragansett Bay)	Yes (in Conn)	Yes (in Conn)	
RI	Beach seine (coastal ponds)	No	No	Low %positive, similar area to other RI survey
RI	Trawl - Seasonal	No	No	Explored as part of Conn trawl index
CT	Long Island Sound Trawl Survey	Yes	No	
CT	Narrows	No	No	Short duration
CT	LIST Summer Survey	No	No	Short duration (~3 yrs)
CT	UCONN	No	No	Short duration (~2 yrs)
CT	EPA Survey	No	No	Short duration (~5 yrs)
CT	Estuary seine survey	No	No	Low encounter rate
NY	Peconic Bay Trawl	No	No	Explored as part of Conn trawl index
NY	WLIS Seine Survey	Yes (in Conn)	Yes (in Conn)	
NJ	NJ Ocean Trawl	Yes	No	Explored as part of Conn trawl index; omitted from 0+ index due to low catch of 1+ fish
NJ	DE R. Seine Survey	Yes (in Conn)	Yes (in Conn)	Revised from 2015 index (region 1 only)
DE	30' Trawl	No	No	Additional data needs
DE	16' Trawl	No	No	Low %positive; same area as 30' trawl
MD	Striped Bass Seine Survey	Yes (in Conn)	Yes (in Conn)	
VA	NEAMAP	Yes	Yes	
VA	ChesMMAP	No	Yes	Age 1 index
VA	Juvenile Finfish Trawl	No	No	redundant with Juv SB seine
VA	Juv. Striped Bass Seine	Yes (in Conn)	Yes (in Conn)	
NC	PSIGNS	Yes	Yes	
NC	P195 (trawl)	No	No	Explored as part of Conn trawl index
NC	IGNS (gillnet)	No	No	needed metadata
SC	SEAMAP	Yes (age 0)	Yes	Age-0 & Age-1 indices
SC	Electrofishing	No	No	Low %positive
SC	Trammel Net	No	No	Low %positive
SC	Stop Net	No	No	Low %positive; ends in 1993
SC	Rotenone	No	No	Low %positive; ends in 1994
GA	EMTS	No	No	Low %positive
FL	FIM 21.3-m bag seine	No	No	Low %positive
FL	FIM 6.2-m otter trawl	No	No	Low %positive

State	Index	Used in 2015?	Use in 2022?	Comments
FL	FIM 183-m haul seine	No	No	Low %positive

Table 2. General details of standardized indices for surveys used in final assessment model. Final model family: NB = negative binomial, ZANB = zero-altered negative binomial, ZINB = zero inflated negative binomial. Nominal index: GM = geometric mean, SAM = stratified arithmetic mean.

State	Survey	Initial (full) model	Final (reduced) model	Final model family	Nominal index	Note(s)
Maryland	Chesapeake Bay Seine Survey	FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS	FREQ ~ YEAR + TEMPWATR + BOTTYPE1	NB	GM	
Virginia	ChesMMAP	FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)	FREQ ~ Year + Bottom.wt + Bottom.do + Stratum + offset(lnSweptArea)	NB	SAM	
Virginia	NEAMAP	"FREQ ~ Year + Bottom.wt + Bottom.do + Stratum + offset(lnSweptArea),	FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea), ziformula = ~ Year + Stratum + Bottom.wt + Bottom.do	ZANB	SAM	
Virginia	Striped Bass Seine Survey	za_model = ""~ Year + Bottom.wt + Bottom.do + Stratum"""	FREQ ~ Year + Bottom.sa + Month + Bottom.wt	NB	GM	
New York	WLIS seine	FREQ ~ Year + Month + Lat + Bottom.wt + Bottom.sa	CPUE ~ YEAR + AREA, ziformula = ~ YEAR + MONTH + AREA + STEMP + SDO	ZANB	GM	
New Hampshire	NH Seine	CPUE ~ YEAR + MONTH + AREA + STEMP + SDO, ziformula = ~ YEAR + MONTH + AREA + STEMP + SDO	FREQ ~ YEAR + MONTH + BOTTOM.TYPE + SURFACE.TEMP + SALINITY	NB	GM	Attempted poisson, negative binomial, zero inflated and hurdle models
Rhode Island	RI Seine	FREQ ~ YEAR + MONTH + BOTTOM.TYPE + SURFACE.TEMP + SALINITY	FREQ ~ YEAR + BOTTOM.TEMP + SALINITY	ZANB	GM	Attempted poisson, negative binomial, zero inflated and hurdle models
NEFSC	Albatross	FREQ ~ YEAR + BOTTOM.TEMP + SALINITY	CPUE ~ MONTH + BTEMP + DEPTH, ziformula = ~ YEAR + MONTH + BTEMP + DEPTH, data = dat,offset=lnEffort	ZINB	SAM	Nominal index uses continuity strata; final GLM uses additional strata that were initially considered by WG, but ultimately dropped.
NEFSC	Bigelow	CPUE ~ YEAR + MONTH + BTEMP + DEPTH, ziformula = ~ YEAR + MONTH + BTEMP + DEPTH, data = dat,offset=lnEffort	CPUE ~ YEAR + BTEMP + DEPTH + SOG,offset=lnEffort, ziformula = ~ YEAR + BTEMP + DEPTH + SOG	ZINB	SAM	Nominal index uses continuity strata; final GLM uses additional strata that were initially considered by WG, but ultimately dropped.
North Carolina	PSIGNS	CPUE ~ YEAR + BTEMP + DEPTH + SOG,offset=lnEffort, ziformula = ~ YEAR + BTEMP + DEPTH + SOG	FREQ ~ YEAR + STRATUM + offset(lnEffort), ziformula = ~ YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO	ZINB	SAM	

State	Survey	Initial (full) model	Final (reduced) model	Final model family	Nominal index	Note(s)
South Carolina	SEAMAP A0	CPUE~YEAR+STRATUM+BOTTOM.TEMP+BOTTOM.SALINITY+BOTTOM.DO+offset(lnEffort)	CPUE~YEAR+LATITUDE+BOTTOM.TEMP+BOTTOM.SALINITY	ZINB	SAM	
South Carolina	SEAMAP A1	CPUE~YEAR+LATITUDE+BOTTOM.TEMP+BOTTOM.SALINITY	CPUE~YEAR+LATITUDE+BOTTOM.TEMP+BOTTOM.SALINITY	NB	SAM	

Table 3. Standardized (GLM) indices for each survey used in the final assessment model.

Year	Albatross	CV	Bigelow	CV	NEAMAP	CV	SEAMAP-A0	CV	PSIGNS	CV	Conn-YOY	CV	ChesMMAP	CV	SEAMAP-A1	CV
1985	17.763	0.6									2.158	0.71				
1986	40.897	1.74									1.171	0.74				
1987	7.446	1.27									1.552	0.41				
1988	30.468	2.35									1.069	0.3				
1989	91.379	0.77					12.80	0.54			1.612	0.3				
1990	8.738	0.91					0.28	0.41			1.301	0.29			20.55	0.29
1991	15.839	0.96					4.96	0.49			1.743	0.28			9.66	0.35
1992	17.486	0.28					2.07	0.63			0.631	0.33			20.09	0.27
1993	1.874	0.86					0.63	0.22			0.492	0.31			1.80	0.29
1994	12.379	0.48					9.86	0.30			0.635	0.32			4.26	0.30
1995	9.317	0.78					0.81	0.18			0.561	0.31			25.74	0.23
1996	10.861	0.96					2.79	0.20			0.928	0.29			5.22	0.18
1997	5.695	0.64					0.32	0.76			1.413	0.29			14.99	0.19
1998	10.516	1.31					0.24	0.29			0.768	0.31			6.96	0.18
1999	24.57	1.32					4.74	0.54			1.633	0.38			8.40	0.36
2000	4.591	0.96					1.61	0.39			0.601	0.3			6.81	0.26
2001	17.549	0.6					5.53	0.28	5.284	0.107	1.448	0.29			2.81	0.16
2002	18.566	0.23					7.08	0.29	4.296	0.132	0.649	0.23	0.159	0.53	5.07	0.46
2003	35.721	0.69					3.92	0.43	4.788	0.124	0.634	0.24	0.513	0.66	4.10	0.26
2004	10.617	0.56					1.27	0.41	5.201	0.14	0.714	0.23	0.102	0.42	34.37	0.24
2005	35.655	0.27					22.74	0.20	6.379	0.137	1.45	0.23	0.474	0.52	4.72	0.31
2006	25.358	0.59					0.92	0.62	6.067	0.16	0.994	0.23	0.087	0.41	14.00	0.36
2007	14.912	0.51			30.993	0.3	0.44	0.47	8.45	0.124	0.95	0.24	0.26	0.38	10.66	0.28
2008	13.237	0.72			46.467	0.3	3.57	0.21	6.716	0.11	1.057	0.23	0.277	0.98	7.61	0.28
2009			12.009	0.7	31.995	0.3	29.79	0.59	6.332	0.112	0.821	0.25	0.043	0.54	36.81	0.63
2010			8.428	0.7	17.707	0.3	33.64	0.27	5.694	0.101	0.956	0.23	0.678	1.12	5.31	0.22
2011			10.198	0.7	24.941	0.3	2.27	0.57	5.977	0.154	0.801	0.24	0.034	0.69	10.76	0.32
2012			6.764	0.7	41.365	0.3	3.10	0.32	4.058	0.128	0.727	0.25	0.085	0.6	3.01	0.18
2013			3.565	0.7	21.066	0.3	0.61	0.41	5.719	0.151	1.154	0.22	0.129	0.42	5.09	0.23
2014			4.694	0.7	23.597	0.3	0.16	0.26	4.8	0.123	1.265	0.22	0.147	0.5	12.54	0.26
2015			1.79	0.7	20.894	0.3	2.99	0.29	2.768	0.114	0.586	0.24	0.467	1.07	13.53	0.37
2016			2.403	0.7	25.548	0.3	0.86	0.20	4.164	0.174	0.467	0.25	0.124	0.48	4.31	0.30
2017					23.07	0.3	0.63	0.27	3.606	0.138	0.706	0.26	0.143	0.58	2.77	0.62
2018			4.037	0.7	16.749	0.3	1.99	0.22	4.036	0.157	0.401	0.28	0.329	0.59	2.43	0.39
2019			1.132	0.7	10.589	0.3	5.80	0.54	8.646	0.124	0.665	0.26			7.17	0.40
2020					14.741	0.3					0.785	0.29				
2021			1.54	0.7	23.183	0.3	0.06	0.97	4.951	0.143	0.795	0.27				

Table 4. Nominal (geometric mean or stratified arithmetic mean) indices for each survey used in the final assessment model.

Year	Albatross	CV	Bigelow	CV	NEAMAP	CV	SEAMAP-A0	CV	PSIGNS	CV	Conn-YOY	CV	ChesMMAP	CV	SEAMAP-A1	CV
1985	17.763	0.6									2.158	0.7				
1986	40.897	1.7									1.171	0.7				
1987	7.446	1.3									1.552	0.4				
1988	30.468	2.4									1.069	0.3				
1989	91.379	0.8					12.80	0.5			1.612	0.3				
1990	8.738	0.9					0.28	0.4			1.301	0.3			20.55	0.3
1991	15.839	1.0					4.96	0.5			1.743	0.3			9.66	0.4
1992	17.486	0.3					2.07	0.6			0.631	0.3			20.09	0.3
1993	1.874	0.9					0.63	0.2			0.492	0.3			1.80	0.3
1994	12.379	0.5					9.86	0.3			0.635	0.3			4.26	0.3
1995	9.317	0.8					0.81	0.2			0.561	0.3			25.74	0.2
1996	10.861	1.0					2.79	0.2			0.928	0.3			5.22	0.2
1997	5.695	0.6					0.32	0.8			1.413	0.3			14.99	0.2
1998	10.516	1.3					0.24	0.3			0.768	0.3			6.96	0.2
1999	24.57	1.3					4.74	0.5			1.633	0.4			8.40	0.4
2000	4.591	1.0					1.61	0.4			0.601	0.3			6.81	0.3
2001	17.549	0.6					5.53	0.3	5.284	0.1	1.448	0.3			2.81	0.2
2002	18.566	0.2					7.08	0.3	4.296	0.1	0.649	0.2	0.159	0.5	5.07	0.5
2003	35.721	0.7					3.92	0.4	4.788	0.1	0.634	0.2	0.513	0.7	4.10	0.3
2004	10.617	0.6					1.27	0.4	5.201	0.1	0.714	0.2	0.102	0.4	34.37	0.2
2005	35.655	0.3					22.74	0.2	6.379	0.1	1.45	0.2	0.474	0.5	4.72	0.3
2006	25.358	0.6					0.92	0.6	6.067	0.2	0.994	0.2	0.087	0.4	14.00	0.4
2007	14.912	0.5			30.993	0.3	0.44	0.5	8.45	0.1	0.95	0.2	0.26	0.4	10.66	0.3
2008	13.237	0.7			46.467	0.3	3.57	0.2	6.716	0.1	1.057	0.2	0.277	1.0	7.61	0.3
2009			12.009	0.7	31.995	0.3	29.79	0.6	6.332	0.1	0.821	0.3	0.043	0.5	36.81	0.6
2010			8.428	0.7	17.707	0.3	33.64	0.3	5.694	0.1	0.956	0.2	0.678	1.1	5.31	0.2
2011			10.198	0.7	24.941	0.3	2.27	0.6	5.977	0.2	0.801	0.2	0.034	0.7	10.76	0.3
2012			6.764	0.7	41.365	0.3	3.10	0.3	4.058	0.1	0.727	0.3	0.085	0.6	3.01	0.2
2013			3.565	0.7	21.066	0.3	0.61	0.4	5.719	0.2	1.154	0.2	0.129	0.4	5.09	0.2
2014			4.694	0.7	23.597	0.3	0.16	0.3	4.8	0.1	1.265	0.2	0.147	0.5	12.54	0.3
2015			1.79	0.7	20.894	0.3	2.99	0.3	2.768	0.1	0.586	0.2	0.467	1.1	13.53	0.4
2016			2.403	0.7	25.548	0.3	0.86	0.2	4.164	0.2	0.467	0.3	0.124	0.5	4.31	0.3
2017					23.07	0.3	0.63	0.3	3.606	0.1	0.706	0.3	0.143	0.6	2.77	0.6
2018			4.037	0.7	16.749	0.3	1.99	0.2	4.036	0.2	0.401	0.3	0.329	0.6	2.43	0.4
2019			1.132	0.7	10.589	0.3	5.80	0.5	8.646	0.1	0.665	0.3			7.17	0.4
2020					14.741	0.3					0.785	0.3				
2021			1.54	0.7	23.183	0.3	0.06	1.0	4.951	0.1	0.795	0.3				

Figure 1. FIG workflow details.

Process data

If present, **remove/do not consider**: haul 2, stratum area, any covariate that has a lot of missing data, correlated covariates

If present, **remove**: months with low bluefish counts, years with 0 bluefish observations

If present, **remove or correct** outliers or obvious incorrect information

Standardize numeric data (but not latitude and longitude)

Build the full set of model parameters

Build full set of covariates for model: check VIF for the full lm(). Drop covariates if the VIF is above 3.

Include effort (swept area, time, etc.) as an offset

Model selection

Fit negative binomial, zero-inflated, and zero-adjusted models with the full set of covariates

If zero models do not fit, iteratively drop covariates that were not significant in the NB model. If still not successful, try fitting null models with only year and effort as a covariate.

Once all model families have been fit, select the best model from each family using stepwise AIC comparison

Compare the final models from all 3 families with AIC comparison. Select the winner.

DHARMa simulations, refine model

If DHARMa indicates that there are problems…

drop data points at ecosystem conditions unlikely to be favorable for bluefish (low temperature, low salinity)? Creaser and Perkins 1994 - minimum salinity of 9.5, minimum temperature 16. EFH - minimum temperature 10.

Bin variables into categories

Drop covariate from model

Figure 2. Example DHARMa diagnostics judged by the FIG to be 'good.'

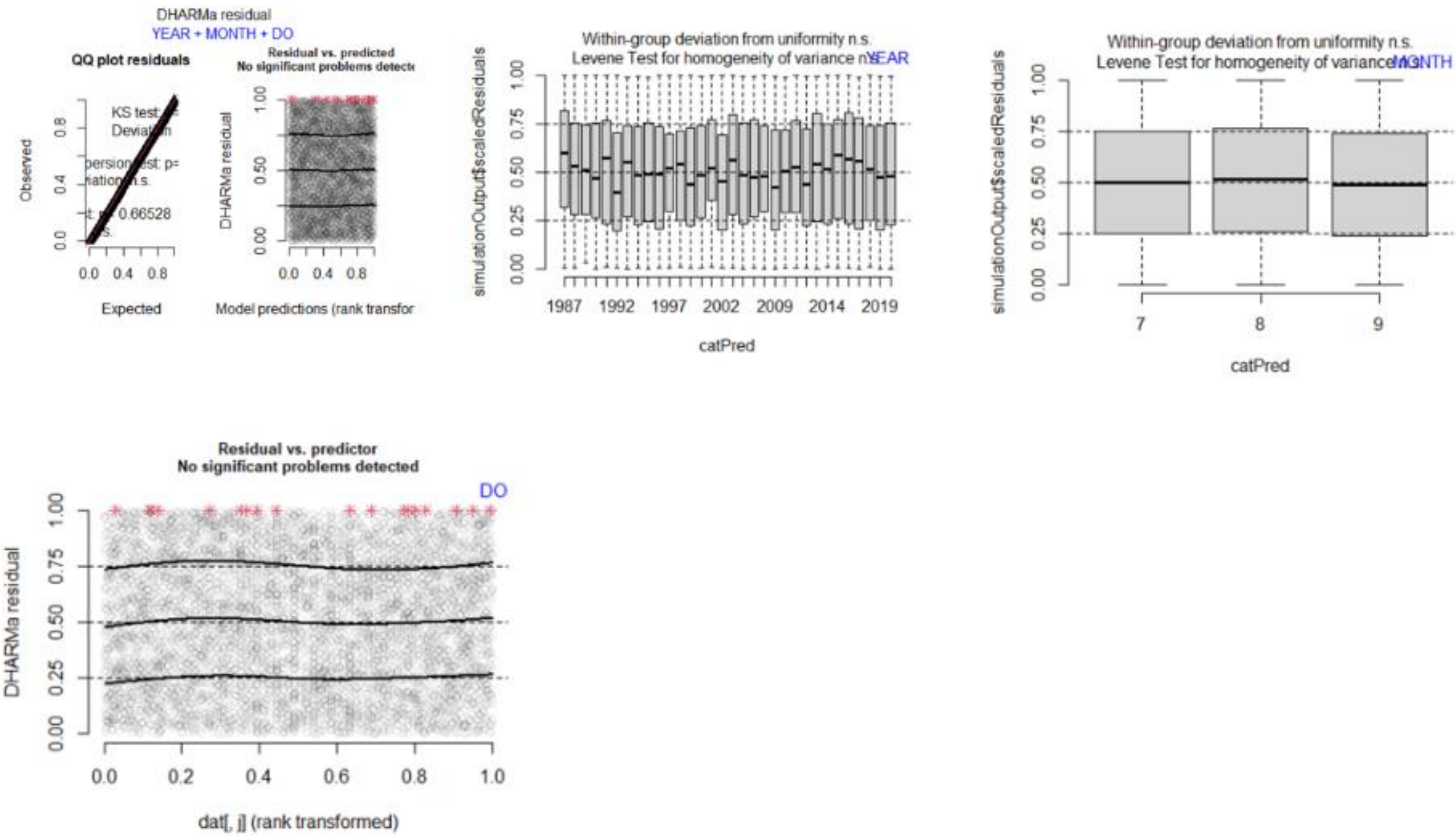


Figure 3. Example DHARMa diagnostics judged by the FIG to be sufficient, though not ideal.

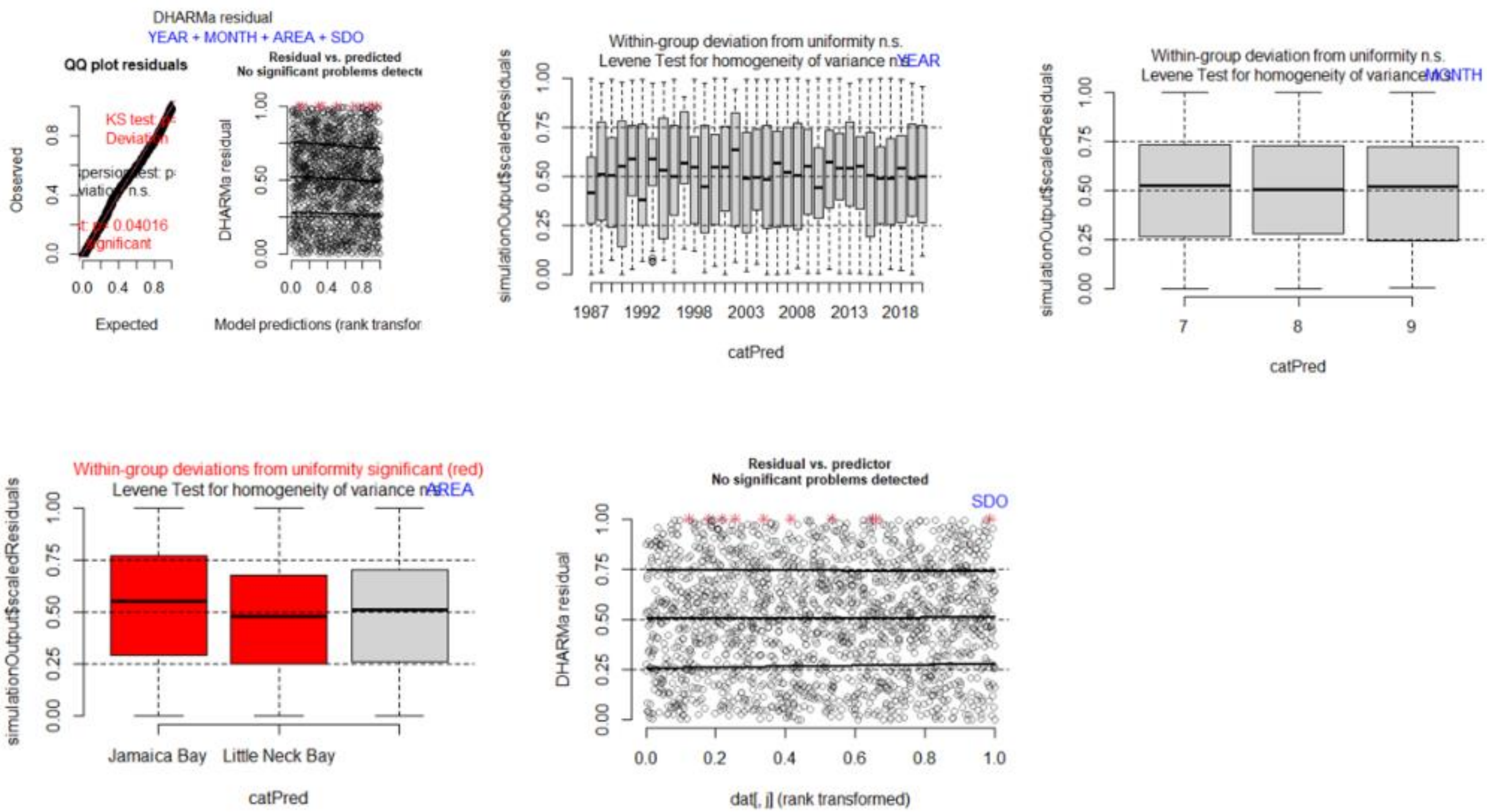


Figure 4. Example DHARMa diagnostics judged by the FIG to be problematic.

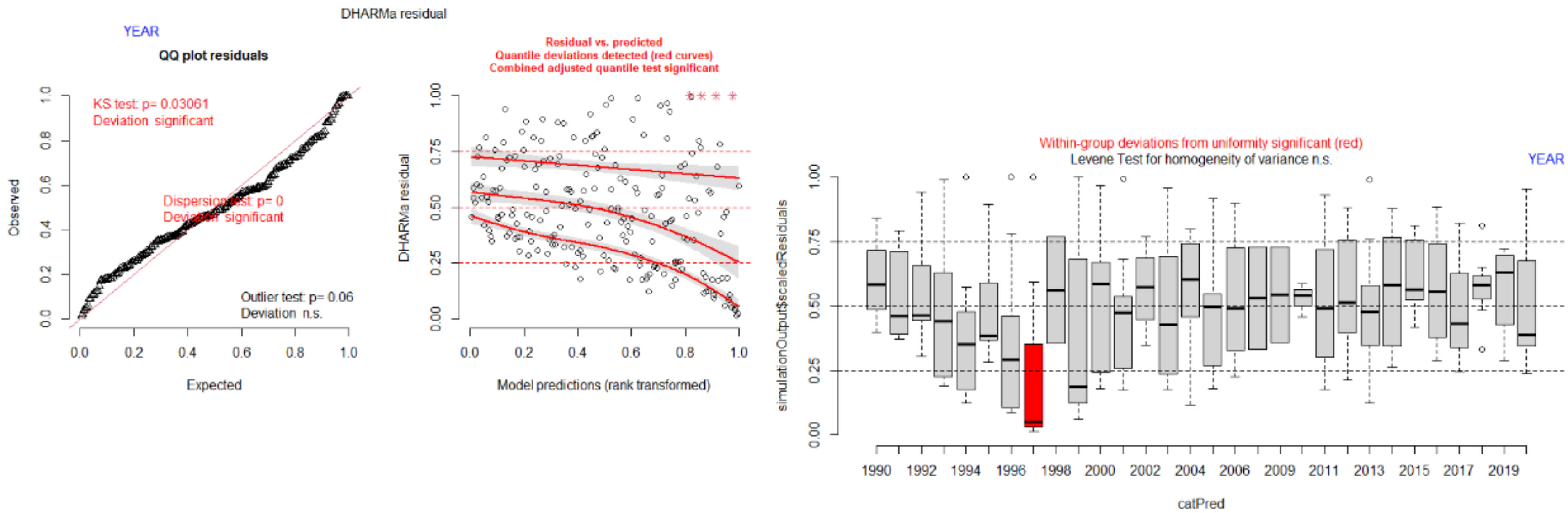


Figure 5. Map of the east coast of the United States showing the approximate locations of each of the surveys used in the final assessment model.

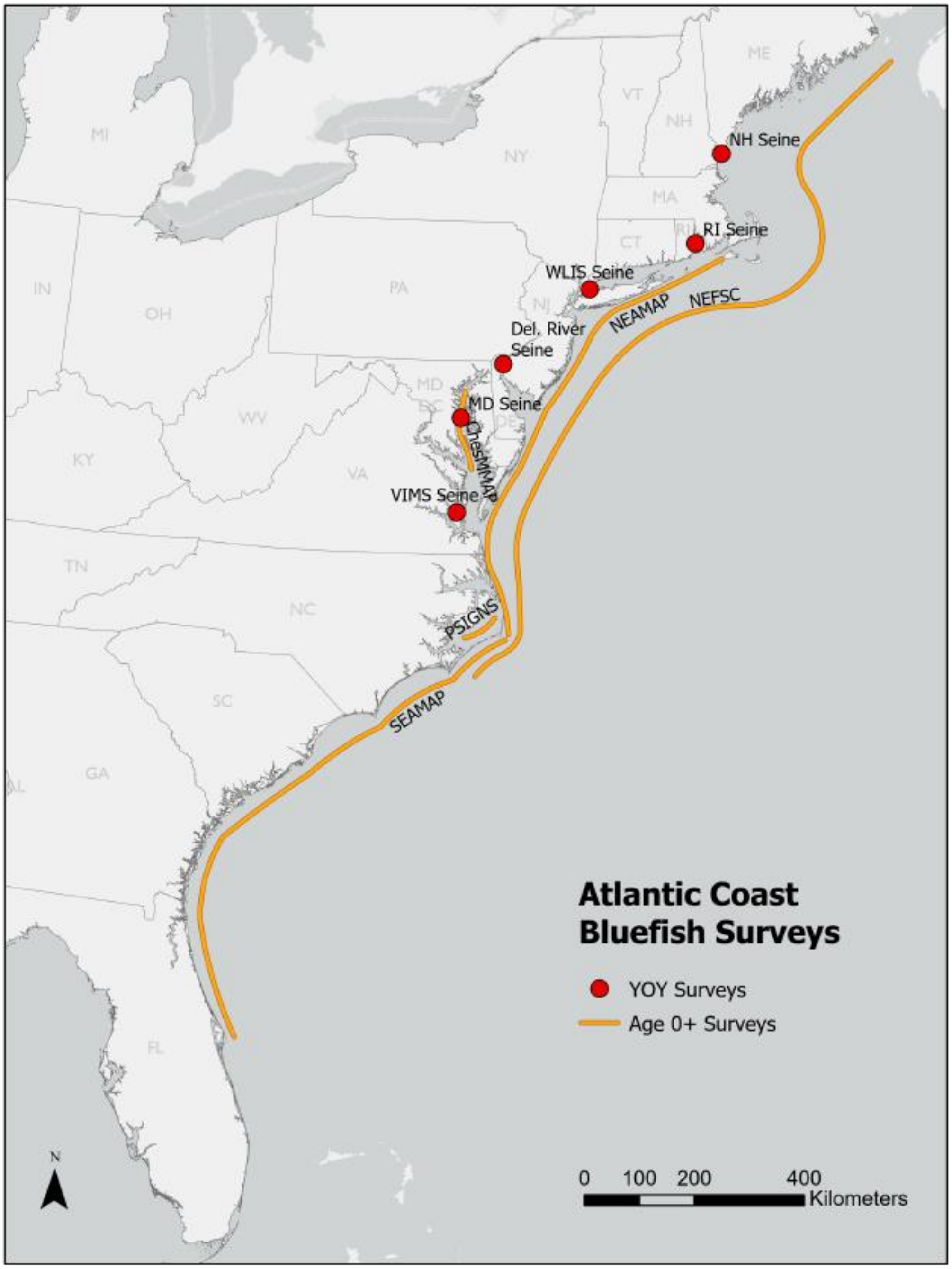
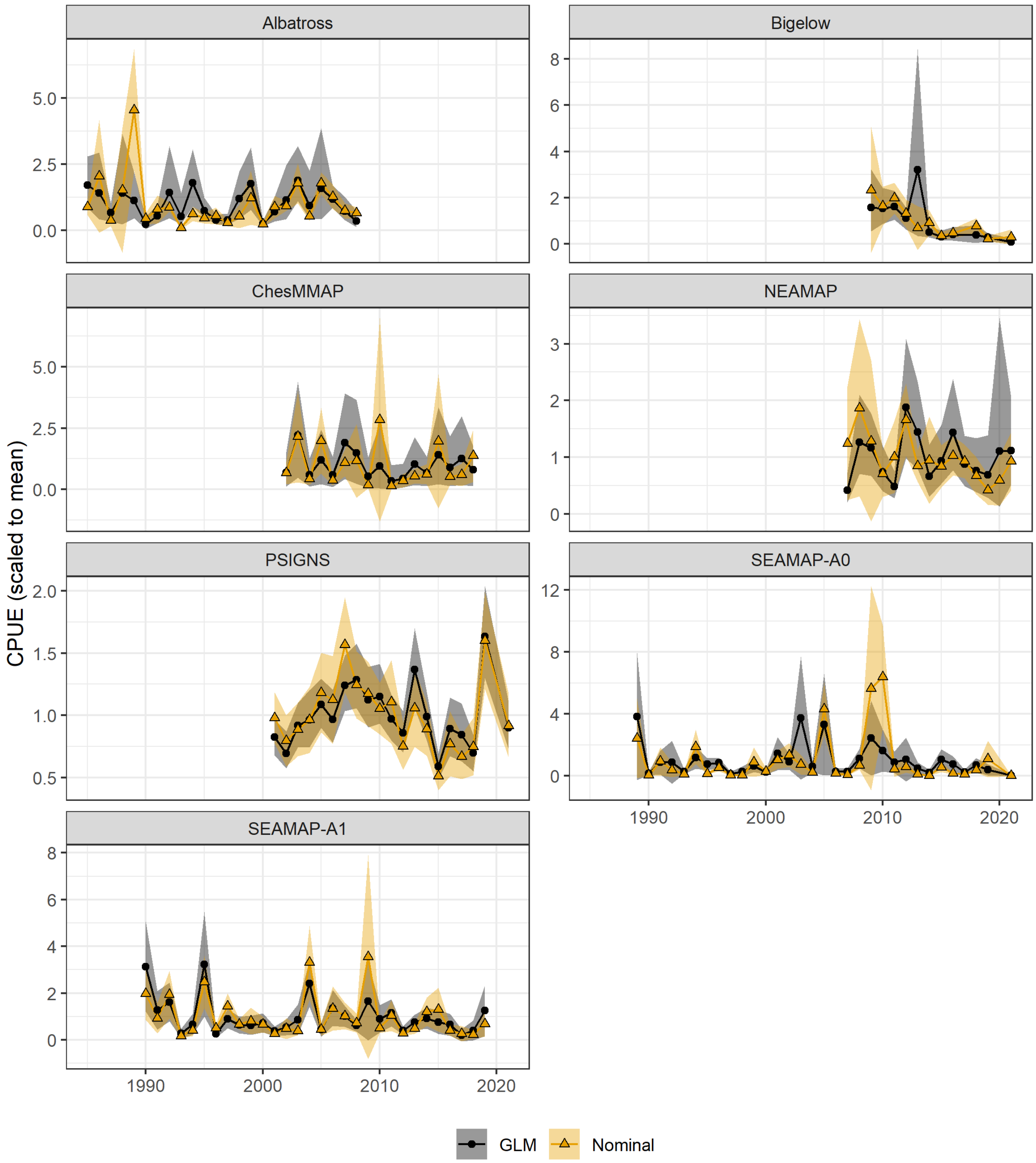


Figure 6. Standardized (GLM) and nominal indices, with 95% confidence intervals, used in the final assessment model.



Bluefish index standardization - NEFSC trawl (Albatross & Delaware) with new strata

M Celestino

2022-06-10 11:45:34

Bluefish index standardization

Step 1: Data processing

Load functions & libraries

```
source("f:\\BF_WG\\FIG\\figFuns.R")
source("f:\\BF_WG\\FIG\\originalCode\\KDrew2\\bootstrap_functions_AT2.R")

library(car, quietly = TRUE, verbose=FALSE)

## Warning: package 'car' was built under R version 4.1.2

## Warning: package 'carData' was built under R version 4.1.2

library(lmtest, quietly = TRUE, verbose=FALSE)

## Warning: package 'lmtest' was built under R version 4.1.2

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

library(glmmTMB, quietly = TRUE, verbose=FALSE)
library(DHARMa, quietly = TRUE, verbose=FALSE)

## Warning: package 'DHARMa' was built under R version 4.1.2

## This is DHARMa 0.4.5. For overview type '?DHARMa'. For recent changes,
type news(package = 'DHARMa')

library(bbmle, quietly = TRUE, verbose=FALSE)
library(ggeffects, quietly = TRUE, verbose=FALSE)
library(ggplot2, quietly = TRUE, verbose=FALSE)

## Warning: package 'ggplot2' was built under R version 4.1.2

library(readxl, quietly = TRUE, verbose=FALSE)
library(lattice, quietly = TRUE, verbose=FALSE)
```

```
## Warning: package 'lattice' was built under R version 4.1.2
library(emmeans, quietly = TRUE, verbose=FALSE)
## Warning: package 'emmeans' was built under R version 4.1.2
library(rmarkdown, quietly = TRUE, verbose=FALSE)
## Warning: package 'rmarkdown' was built under R version 4.1.2
library(knitr, quietly = TRUE, verbose=FALSE)
## Warning: package 'knitr' was built under R version 4.1.2
library(reshape2, quietly=TRUE, verbose=FALSE)
## Warning: package 'reshape2' was built under R version 4.1.2
library(doBy, quietly=TRUE, verbose=FALSE)
## Warning: package 'doBy' was built under R version 4.1.2
library(dplyr, quietly=TRUE, verbose=FALSE,warn.conflicts=FALSE)
```

Read in data

```
indata <-

read_xlsx(path="f:\\BF_WG\\FIG\\NEFSC\\NEFSC_Fall_Survey_Data_for_Std_V3.xlsx
",
   sheet="BLUEFISH_FALL_ALB_RT2022_85_08",range="A125:AJ3662",na="NA")
head(indata)

## # A tibble: 6 x 36
##   CRUISE6 STRATUM   TOW STATION   SHG SVVESSEL SVGEAR EST_YEAR EST_MONTH
EST_DAY
##     <dbl>   <dbl> <dbl>   <dbl> <dbl> <chr>     <dbl>    <dbl>     <dbl>
<dbl>
## 1  198508    1010     1     109   111 DE           11     1985        10
8
## 2  198508    1010     2     123   111 DE           11     1985        10
9
## 3  198508    1010     3     108   111 DE           11     1985        10
8
## 4  198508    1010     4     121   111 DE           11     1985        10
9
## 5  198508    1010     5     110   123 DE           11     1985        10
8
## 6  198508    1010     6     124   111 DE           11     1985        10
9
## # ... with 26 more variables: TIME <dbl>, TOWDUR <dbl>, DOPDISTB <dbl>,
## #   DOPDISTW <dbl>, AVGDEPTH <dbl>, STATYPE <dbl>, HAUL <dbl>, GEARCOND
<dbl>,
```

```
## #   TYPE_CODE <lgl>, OPERATION_CODE <lgl>, GEAR_CODE <lgl>,
## #   ACQUISITION_CODE <lgl>, AREA <dbl>, BOTTEMP <dbl>, BEGLAT <dbl>,
## #   BEGLON <dbl>, STRATUM_AREA <dbl>, STRGRP_DESC <chr>, STRATUM_NAME
<chr>,
## #   AREA_SWEPT_WINGS_MEAN_KM2 <lgl>, AREA_SWEPT_WINGS_MEAN_NM2 <dbl>,
## #   SWEPT_AREA_RATIO <dbl>, SVSPP <dbl>, CATCHSEX <dbl>, EXPCATCHWT <dbl>,
...

str(indata)

## tibble [3,537 x 36] (S3: tbl_df/tbl/data.frame)
##  $ CRUISE6                   : num [1:3537] 198508 198508 198508 198508
198508 ...
##  $ STRATUM                   : num [1:3537] 1010 1010 1010 1010 1010 1010
1010 1050 1050 1050 ...
##  $ TOW                       : num [1:3537] 1 2 3 4 5 6 7 1 3 4 ...
##  $ STATION                   : num [1:3537] 109 123 108 121 110 124 122 147
171 149 ...
##  $ SHG                       : num [1:3537] 111 111 111 111 123 111 111 111
111 111 ...
##  $ SVVESSEL                  : chr [1:3537] "DE" "DE" "DE" "DE" ...
##  $ SVGEAR                    : num [1:3537] 11 11 11 11 11 11 11 11 11 11
...
##  $ EST_YEAR                  : num [1:3537] 1985 1985 1985 1985 1985 ...
##  $ EST_MONTH                 : num [1:3537] 10 10 10 10 10 10 10 10 10 10
...
##  $ EST_DAY                   : num [1:3537] 8 9 8 9 8 9 9 15 18 16 ...
##  $ TIME                      : num [1:3537] 705 1342 431 1007 949 ...
##  $ TOWDUR                    : num [1:3537] 30 30 30 30 30 30 30 30 30 30
...
##  $ DOPDISTB                  : num [1:3537] 1.77 1.74 1.74 1.78 1.74 1.72
1.72 1.84 1.88 1.69 ...
##  $ DOPDISTW                  : num [1:3537] 0 0 0 0 0 0 0 0 0 0 ...
##  $ AVGDEPTH                  : num [1:3537] 37 45 37 55 52 40 44 40 52 46
...
##  $ STATYPE                   : num [1:3537] 1 1 1 1 1 1 1 1 1 1 ...
##  $ HAUL                      : num [1:3537] 1 1 1 1 2 1 1 1 1 1 ...
##  $ GEARCOND                  : num [1:3537] 1 1 1 1 3 1 1 1 1 1 ...
##  $ TYPE_CODE                 : logi [1:3537] NA NA NA NA NA NA ...
##  $ OPERATION_CODE            : logi [1:3537] NA NA NA NA NA NA ...
##  $ GEAR_CODE                 : logi [1:3537] NA NA NA NA NA NA ...
##  $ ACQUISITION_CODE          : logi [1:3537] NA NA NA NA NA NA ...
##  $ AREA                      : num [1:3537] 615 613 612 613 615 612 613 537
537 539 ...
##  $ BOTTEMP                   : num [1:3537] 14.5 NA 13.3 NA 11.5 16.5 NA NA
NA NA ...
##  $ BEGLAT                    : num [1:3537] 39.9 40.3 40.2 40.1 39.7 ...
##  $ BEGLON                    : num [1:3537] -73.5 -73 -73.6 -72.7 -73 ...
##  $ STRATUM_AREA              : num [1:3537] 2516 2516 2516 2516 2516 ...
##  $ STRGRP_DESC               : chr [1:3537] "OFFSHORE REGULAR" "OFFSHORE
```

```
REGULAR" "OFFSHORE REGULAR" "OFFSHORE REGULAR" ...
##  $ STRATUM_NAME             : chr [1:3537] "SNE HUDSON CANYN" "SNE HUDSON
CANYN" "SNE HUDSON CANYN" "SNE HUDSON CANYN" ...
##  $ AREA_SWEPT_WINGS_MEAN_KM2: logi [1:3537] NA NA NA NA NA NA ...
##  $ AREA_SWEPT_WINGS_MEAN_NM2: num [1:3537] 0.01 0.01 0.01 0.01 0.01 0.01
0.01 0.01 0.01 0.01 ...
##  $ SWEPT_AREA_RATIO         : num [1:3537] 1 1 1 1 1 1 1 1 1 1 ...
##  $ SVSPP                    : num [1:3537] 135 NA NA NA 135 135 NA NA NA
135 ...
##  $ CATCHSEX                 : num [1:3537] 0 0 0 0 0 0 0 0 0 0 ...
##  $ EXPCATCHWT               : num [1:3537] 8.5 0 0 0 8.5 6.4 0 0 0 9.5 ...
##  $ EXPCATCHNUM              : num [1:3537] 8 0 0 0 2 7 0 0 0 3 ...
```

```
summary(indata)
```

```
##     CRUISE6          STRATUM          TOW            STATION          SHG
##  Min.   :198508   Min.   :1010   Min.   : 1.00   Min.   :  1   Min.
:111.0
##  1st Qu.:199105   1st Qu.:1190   1st Qu.: 1.00   1st Qu.: 69   1st
Qu.:111.0
##  Median :199706   Median :3020   Median : 2.00   Median :119   Median
:111.0
##  Mean   :199670   Mean   :2275   Mean   : 2.65   Mean   :120   Mean
:112.7
##  3rd Qu.:200306   3rd Qu.:3250   3rd Qu.: 4.00   3rd Qu.:162   3rd
Qu.:111.0
##  Max.   :200807   Max.   :3460   Max.   :11.00   Max.   :375   Max.
:136.0
##
##    SVVESSEL             SVGEAR      EST_YEAR      EST_MONTH
##  Length:3537        Min.   :11   Min.   :1985   Min.   : 9.000
##  Class :character   1st Qu.:11   1st Qu.:1991   1st Qu.: 9.000
##  Mode  :character   Median :11   Median :1997   Median : 9.000
##                     Mean   :11   Mean   :1997   Mean   : 9.271
##                     3rd Qu.:11   3rd Qu.:2003   3rd Qu.:10.000
##                     Max.   :11   Max.   :2008   Max.   :11.000
##
##     EST_DAY           TIME          TOWDUR         DOPDISTB
##  Min.   : 1.00   Min.   :   0   Min.   :19.83   Min.   :0.170
##  1st Qu.:10.00   1st Qu.: 607   1st Qu.:30.00   1st Qu.:1.740
##  Median :17.00   Median :1214   Median :30.00   Median :1.819
##  Mean   :16.38   Mean   :1195   Mean   :29.91   Mean   :1.821
##  3rd Qu.:23.00   3rd Qu.:1817   3rd Qu.:30.00   3rd Qu.:1.910
##  Max.   :30.00   Max.   :2359   Max.   :31.43   Max.   :2.161
##                                                 NA's   :6
##     DOPDISTW         AVGDEPTH         STATYPE       HAUL          GEARCOND
##  Min.   :0.0000   Min.   :  5.00   Min.   :1   Min.   :1.000   Min.
:1.000
##  1st Qu.:0.0000   1st Qu.: 18.00   1st Qu.:1   1st Qu.:1.000   1st
Qu.:1.000
```

```
##  Median :0.0000   Median : 27.00   Median :1   Median :1.000   Median
:1.000
##  Mean   :0.0049   Mean   : 33.43   Mean   :1   Mean   :1.143   Mean
:1.232
##  3rd Qu.:0.0000   3rd Qu.: 45.00   3rd Qu.:1   3rd Qu.:1.000   3rd
Qu.:1.000
##  Max.   :1.9500   Max.   :114.00   Max.   :1   Max.   :3.000   Max.
:6.000
##  NA's   :2384     NA's   :1
##  TYPE_CODE      OPERATION_CODE GEAR_CODE      ACQUISITION_CODE      AREA
##  Mode:logical   Mode:logical   Mode:logical   Mode:logical     Min.
:521.0
##  TRUE:113       TRUE:113       TRUE:113       TRUE:113         1st
Qu.:539.0
##  NA's:3424      NA's:3424      NA's:3424      NA's:3424        Median
:613.0
##                                                                 Mean
:594.3
##                                                                 3rd
Qu.:622.0
##                                                                 Max.
:636.0
##                                                                 NA's   :1
##     BOTTEMP          BEGLAT          BEGLON        STRATUM_AREA
##  Min.   : 5.88   Min.   :35.20   Min.   :-76.07   Min.   :  13
##  1st Qu.:13.26   1st Qu.:38.08   1st Qu.:-74.98   1st Qu.: 154
##  Median :16.94   Median :40.13   Median :-73.67   Median : 383
##  Mean   :16.91   Mean   :39.41   Mean   :-72.84   Mean   :1130
##  3rd Qu.:20.79   3rd Qu.:40.81   3rd Qu.:-71.12   3rd Qu.:2454
##  Max.   :28.12   Max.   :42.08   Max.   :-67.02   Max.   :2832
##  NA's   :717
##  STRGRP_DESC        STRATUM_NAME       AREA_SWEPT_WINGS_MEAN_KM2
##  Length:3537        Length:3537        Mode:logical
##  Class :character   Class :character   NA's:3537
##  Mode  :character   Mode  :character
##
##
##
##
##  AREA_SWEPT_WINGS_MEAN_NM2 SWEPT_AREA_RATIO     SVSPP         CATCHSEX
##  Min.   :0.01              Min.   :1        Min.   :135    Min.   :0
##  1st Qu.:0.01              1st Qu.:1        1st Qu.:135    1st Qu.:0
##  Median :0.01              Median :1        Median :135    Median :0
##  Mean   :0.01              Mean   :1        Mean   :135    Mean   :0
##  3rd Qu.:0.01              3rd Qu.:1        3rd Qu.:135    3rd Qu.:0
##  Max.   :0.01              Max.   :1        Max.   :135    Max.   :0
##                                             NA's   :2054
##    EXPCATCHWT       EXPCATCHNUM
##  Min.   :  0.000   Min.   :   0.00
##  1st Qu.:  0.000   1st Qu.:   0.00
```

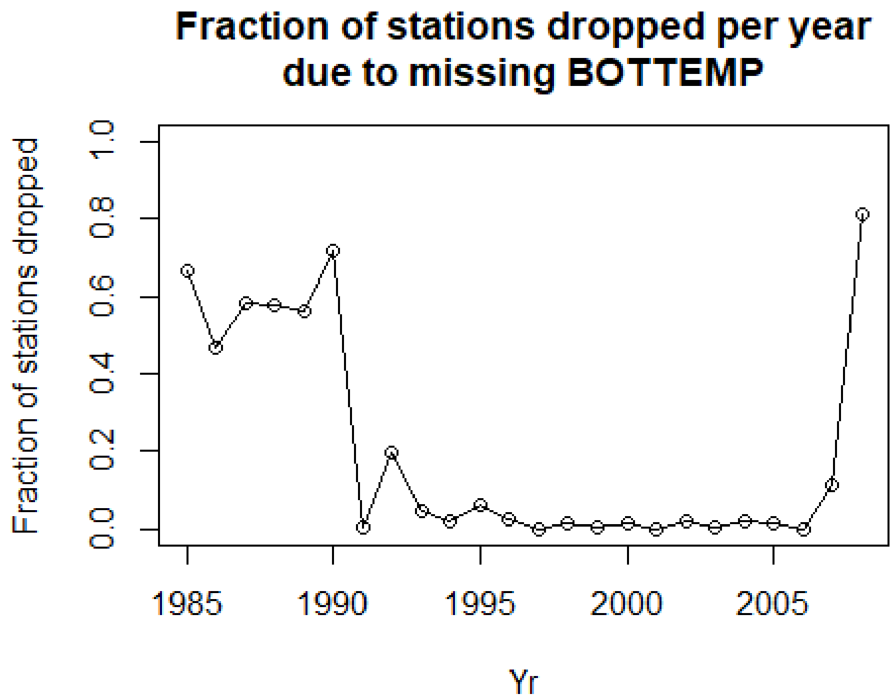
```
##  Median :  0.000   Median :   0.00  
##  Mean   :  2.162   Mean   :  12.43  
##  3rd Qu.:  1.980   3rd Qu.:   3.00  
##  Max.   :150.700   Max.   :2743.00  
## 
```

Some data exploration to determine which years and covariates to keep; then remove missing values

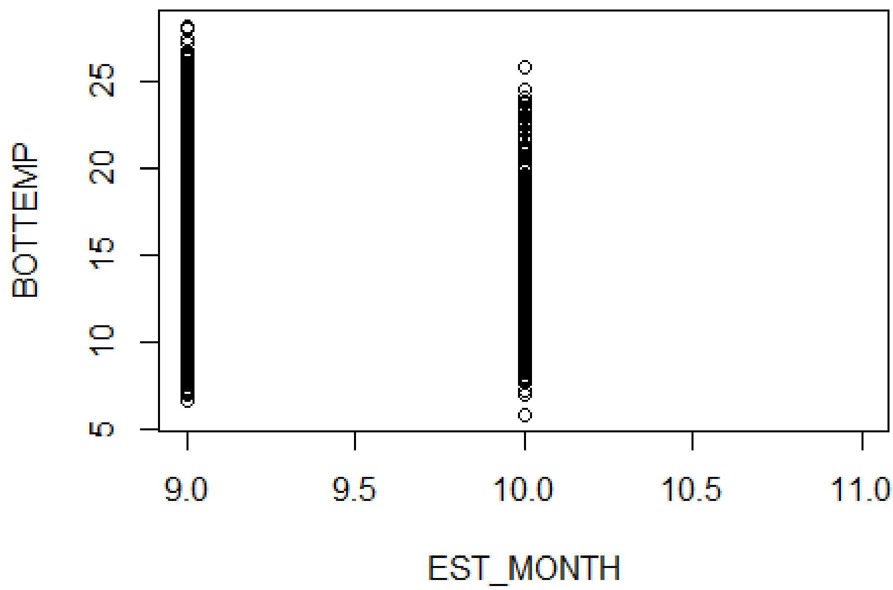
```
keepVars <-
c("CRUISE6","STRATUM","STATION","SVVESSEL","EST_YEAR","EST_MONTH","TOWDUR",
   "DOPDISTB","AVGDEPTH","AREA","BOTTEMP","EXPCATCHNUM")
#    "HAUL","GEARCOND","ACQUISITION_CODE",

# how many records would we lose to NAs:
apply(indata[,keepVars], 2, function(x) sum(is.na(x)))
```

```
##     CRUISE6     STRATUM     STATION    SVVESSEL    EST_YEAR   EST_MONTH 
##           0           0           0           0           0           0 
##      TOWDUR    DOPDISTB    AVGDEPTH        AREA     BOTTEMP EXPCATCHNUM 
##           0           6           1           1         717           0
```

```
# how many records will we lose per year due to NA BOTTEMP
plot(unique(indata$EST_YEAR),
apply(
do.call("rbind",with(indata[,c("EST_YEAR","BOTTEMP")],
   tapply(BOTTEMP,EST_YEAR,function(x) c(sum(is.na(x)),length(x) )))),1,
   function(x) {x[1]/x[2]}),
   xlab="Yr",ylab="Fraction of stations dropped",
   ylim=c(0,1),type="o",
   main="Fraction of stations dropped per year\ndue to missing BOTTEMP")
```



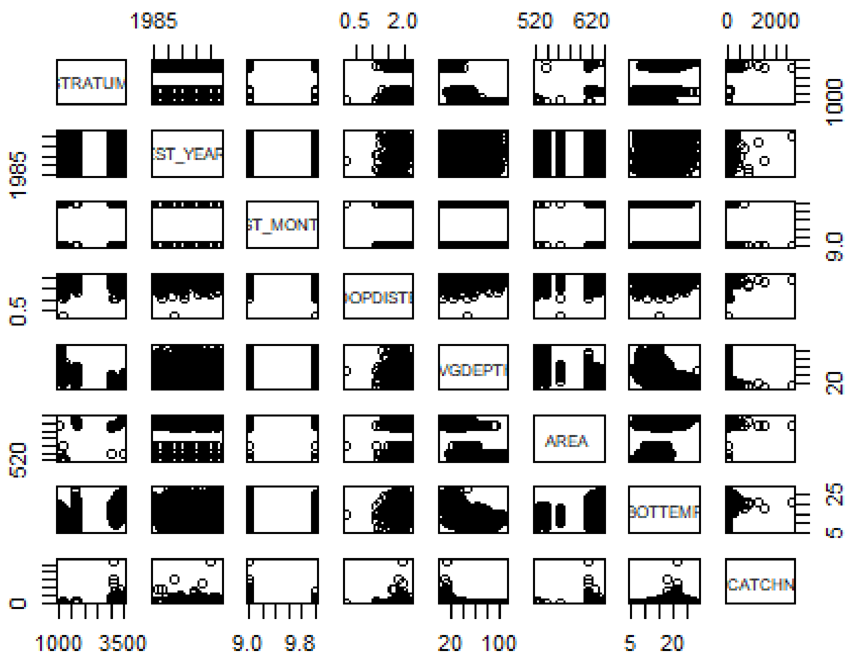

```
# Month and BOTTEMP are weakly negatively correlated
with(indata, {
   plot(EST_MONTH,BOTTEMP)
   cor(EST_MONTH,BOTTEMP,use="pairwise.complete.obs")
   }
)
```



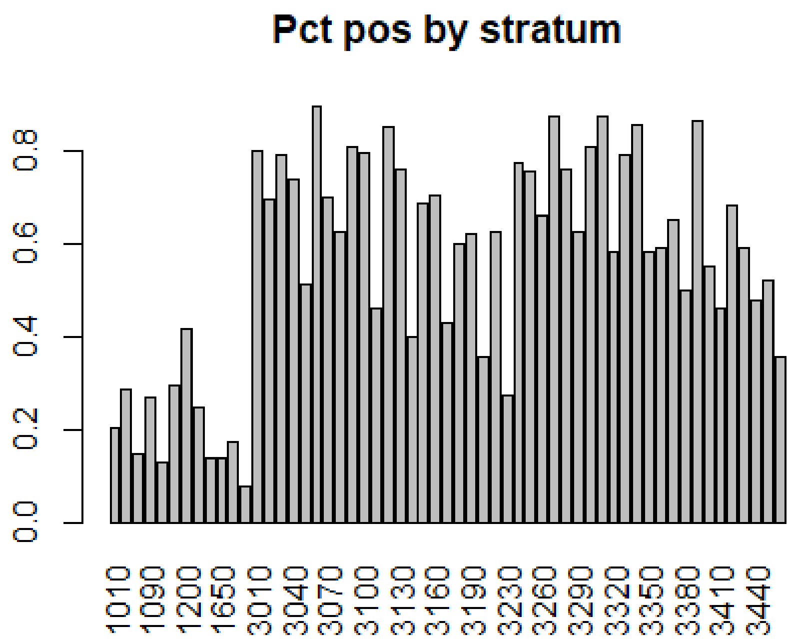

```
## [1] -0.2645855

indata <- indata[,keepVars]
table(indata$SVVESSEL,indata$EST_YEAR)

##
##      1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998
1999
##   AL    6   75  135  128    0    0    0  146    0  152  146  147  148  150
148
##   DE  138   70    0    5  149  150  147    0  148    0    0    0    0    0
0
##
##      2000 2001 2002 2003 2004 2005 2006 2007 2008
##   AL  149  150  148  150  145  148  160  150  149
##   DE    0    0    0    0    0    0    0    0    0

indata2 <- indata[complete.cases(indata),]

summary(indata2)

##     CRUISE6          STRATUM        STATION         SVVESSEL
##  Min.   :198508   Min.   :1010   Min.   :  1.0   Length:2815
##  1st Qu.:199306   1st Qu.:1190   1st Qu.: 69.0   Class :character
##  Median :199804   Median :1730   Median :120.0   Mode  :character
##  Mean   :199784   Mean   :2255   Mean   :120.4
##  3rd Qu.:200306   3rd Qu.:3245   3rd Qu.:163.0
```

```
##  Max.   :200807   Max.   :3460   Max.   :366.0
##     EST_YEAR      EST_MONTH         TOWDUR          DOPDISTB
##  Min.   :1985   Min.   : 9.000   Min.   :19.83   Min.   :0.170
##  1st Qu.:1993   1st Qu.: 9.000   1st Qu.:30.00   1st Qu.:1.760
##  Median :1998   Median : 9.000   Median :30.00   Median :1.864
##  Mean   :1998   Mean   : 9.254   Mean   :29.93   Mean   :1.836
##  3rd Qu.:2003   3rd Qu.:10.000   3rd Qu.:30.02   3rd Qu.:1.920
##  Max.   :2008   Max.   :10.000   Max.   :31.43   Max.   :2.161
##     AVGDEPTH          AREA         BOTTEMP        EXPCATCHNUM
##  Min.   :  5.00   Min.   :521   Min.   : 5.88   Min.   :   0.0
##  1st Qu.: 18.00   1st Qu.:539   1st Qu.:13.27   1st Qu.:   0.0
##  Median : 27.00   Median :613   Median :16.97   Median :   0.0
##  Mean   : 33.61   Mean   :594   Mean   :16.92   Mean   :  11.7
##  3rd Qu.: 46.00   3rd Qu.:621   3rd Qu.:20.80   3rd Qu.:   2.0
##  Max.   :112.00   Max.   :636   Max.   :28.12   Max.   :2743.0
```

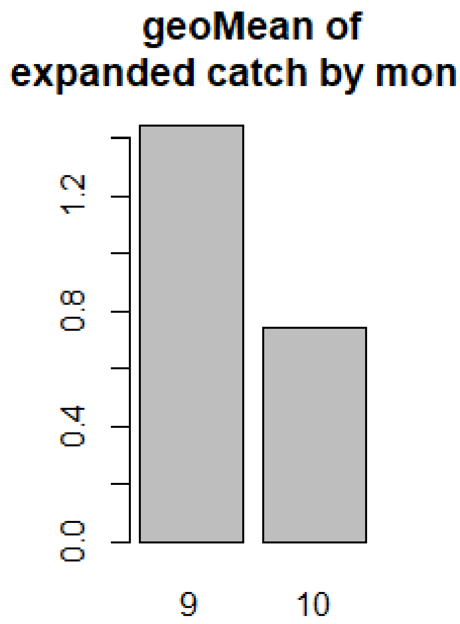
```
# The data I received were already subset to specific strata and months, so
not too much more to do here
pairs(indata2[,-c(1,3,4,7)])
```

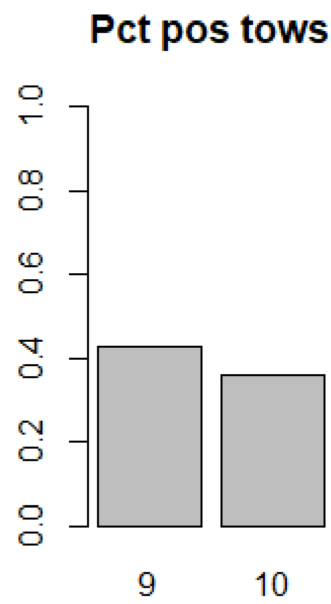



```
barplot(with(indata, tapply(EXPCATCHNUM, STRATUM, pctPos)),las=3,main="Pct
pos by stratum")
```




```
op <- par(mfrow=c(1,2))
barplot(tapply(indata2$EXPCATCHNUM, indata2$EST_MONTH,
geoMean,alpha=NULL,warn=FALSE),main="geoMean of\nexpanded catch by month")
barplot(tapply(indata2$EXPCATCHNUM, indata2$EST_MONTH, pctPos),main="Pct pos
tows",ylim=c(0,1))
```

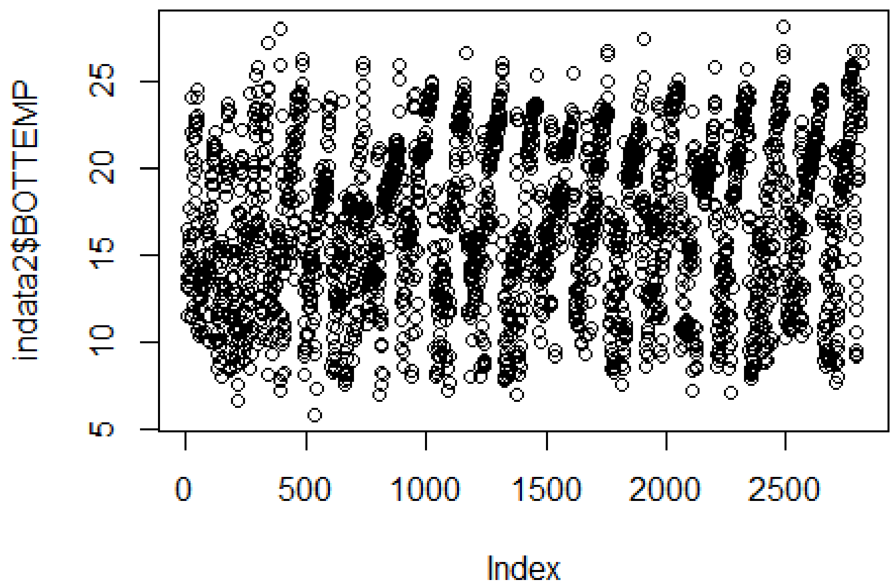





```
par(op)
```

Check for outliers

```
# Check for outliers or weirdo values
plot(indata2$BOTTEMP)
```

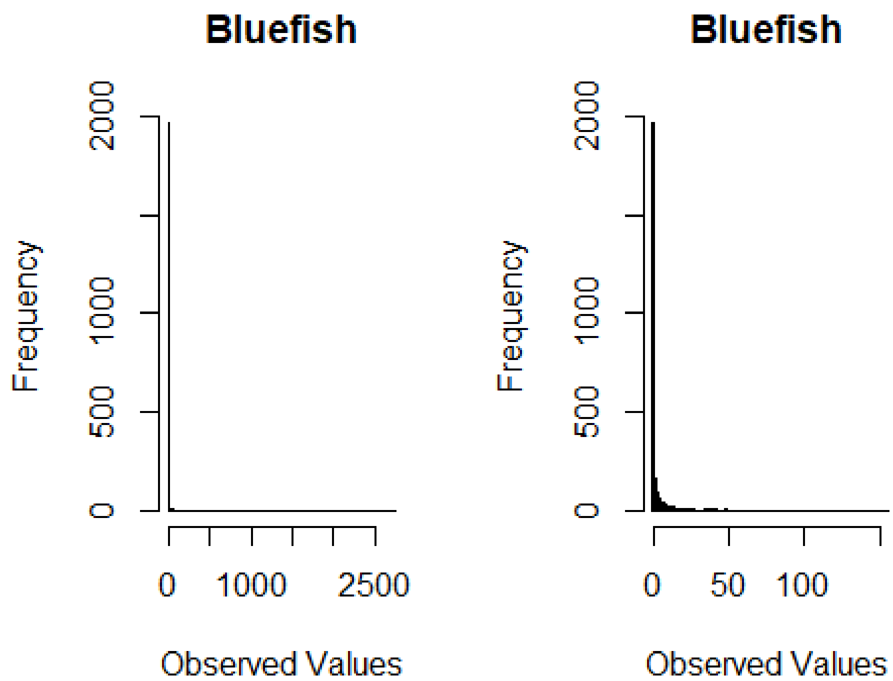


Plot distribution of catches:

```
op <- par(mfrow=c(1,2))
hist(indata2$EXPCATCHNUM,main="Bluefish",xlab="Observed Values",
   ylab="Frequency", breaks=seq(0,max(indata2$EXPCATCHNUM),1))
sort(indata2$EXPCATCHNUM,decreasing=TRUE)[1:50]
```

```
##  [1] 2743 1531 1286  867  860  701  640  531  524  483  410  395  369  368
363
## [16]  303  296  290  287  283  254  247  238  235  234  230  229  222  222
220
## [31]  217  214  213  211  207  207  201  192  184  181  173  168  166  164
160
## [46]  155  151  150  145  133
```

```
hist(indata2$EXPCATCHNUM,main="Bluefish",xlab="Observed Values",
   ylab="Frequency", breaks=seq(0,max(indata2$EXPCATCHNUM),1),xlim=c(0,150))
```




```
par(op)
```

Finally, check to see if there are any years with zero catch that need to be removed:

```
diff(as.numeric(names(xtabs(EXPCATCHNUM~EST_YEAR, indata2))))

##  [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

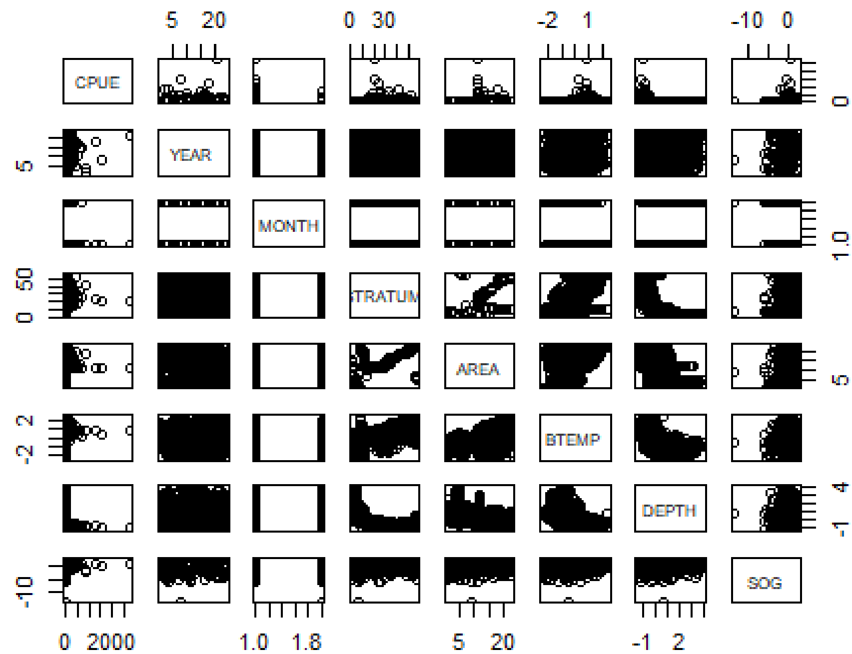
Create a dataframe with all the factors we need (z-score all continuous variables; all others as.factors)

```
dat = data.frame(CPUE = indata2$EXPCATCHNUM,
                 YEAR = as.factor(indata2$EST_YEAR),
                 MONTH = as.factor(indata2$EST_MONTH),
                 STRATUM = as.factor(indata2$STRATUM),
                 AREA=as.factor(indata2$AREA),
                 VESSEL=as.factor(indata2$SVVESSEL),
                 BTEMP = Z.scr(indata2$BOTTEMP),
                 DEPTH = Z.scr(indata2$AVGDEPTH),
                 SOG=Z.scr(indata2$DOPDISTB),
                 EFFORT = indata2$TOWDUR)

# Calculate the proportion of positive tows/sets/hauls
dat$PosTow <- ifelse(dat$CPUE > 0, 1, 0)
mean(dat$PosTow) # 62%

## [1] 0.410302
```

```
dat$lnEffort <- log(dat$EFFORT)

head(dat)

##   CPUE YEAR MONTH STRATUM AREA VESSEL       BTEMP     DEPTH        SOG
EFFORT
## 1    8 1985    10    1010  615     DE -0.52324323 0.1715041 -0.5290705
30
## 2    0 1985    10    1010  612     DE -0.78298941 0.1715041 -0.7713055
30
## 3    2 1985    10    1010  615     DE -1.17260868 0.9294862 -0.7713055
30
## 4    7 1985    10    1010  612     DE -0.09033292 0.3231005 -0.9327955
30
## 5    0 1985    10    1060  537     DE -0.80463492 1.4348076 -0.5290705
30
## 6    3 1985    10    1060  613     DE -0.67476183 1.8390647 -0.6098155
30
##   PosTow lnEffort
## 1      1 3.401197
## 2      0 3.401197
## 3      1 3.401197
## 4      1 3.401197
## 5      0 3.401197
## 6      1 3.401197
```

Check collinearity

```
pairs(~CPUE+YEAR+MONTH+STRATUM+AREA+BTEMP+DEPTH+SOG,data=dat)
```




```
round(cor(with(dat,
data.frame(CPUE,unfactor(YEAR),unfactor(MONTH),unfactor(STRATUM),unfactor(ARE
A),BTEMP,DEPTH,SOG))),2)

##                      CPUE  unfactor.YEAR.  unfactor.MONTH.  unfactor.STRATUM.
## CPUE                 1.00            0.02            -0.04               0.13
## unfactor.YEAR.       0.02            1.00            -0.07               0.04
## unfactor.MONTH.     -0.04           -0.07             1.00              -0.38
## unfactor.STRATUM.    0.13            0.04            -0.38               1.00
## unfactor.AREA.       0.08            0.01            -0.55               0.60
## BTEMP                0.11            0.10            -0.26               0.73
## DEPTH               -0.14           -0.04             0.33              -0.79
## SOG                  0.02            0.58            -0.02               0.04
##                     unfactor.AREA. BTEMP DEPTH   SOG
## CPUE                          0.08  0.11 -0.14  0.02
## unfactor.YEAR.                0.01  0.10 -0.04  0.58
## unfactor.MONTH.              -0.55 -0.26  0.33 -0.02
## unfactor.STRATUM.             0.60  0.73 -0.79  0.04
## unfactor.AREA.                1.00  0.50 -0.64  0.06
## BTEMP                         0.50  1.00 -0.71  0.09
## DEPTH                        -0.64 -0.71  1.00 -0.03
## SOG                           0.06  0.09 -0.03  1.00

# Check the variance inflation factor for a more statistical check.
# You want GVIF to be less than ~3.
mod = lm(CPUE ~ YEAR + MONTH + STRATUM + VESSEL + BTEMP + DEPTH + SOG, data =
```

```
dat) # won't run with STRATUM and AREA, so removing AREA
vif(mod)

##              GVIF Df GVIF^(1/(2*Df))
## YEAR    68.838435 23        1.096359
## MONTH    2.523890  1        1.588675
## STRATUM 32.422360 56        1.031549
## VESSEL  19.125506  1        4.373272
## BTEMP    4.408093  1        2.099546
## DEPTH    8.734620  1        2.955439
## SOG      2.082344  1        1.443033

# remove VESSEL
mod = lm(CPUE ~ YEAR + MONTH + STRATUM + BTEMP + DEPTH + SOG, data = dat)
vif(mod)

##              GVIF Df GVIF^(1/(2*Df))
## YEAR     4.105193 23        1.031177
## MONTH    2.392791  1        1.546865
## STRATUM 30.495793 56        1.030985
## BTEMP    4.382076  1        2.093341
## DEPTH    8.726082  1        2.953994
## SOG      2.082260  1        1.443004

# You want GVIF to be less than ~3. Drop STRATUM
mod = lm(CPUE ~ YEAR + MONTH + BTEMP + DEPTH + SOG, data = dat)
vif(mod)

##           GVIF Df GVIF^(1/(2*Df))
## YEAR  2.800579 23        1.022640
## MONTH 1.412240  1        1.188377
## BTEMP 2.250778  1        1.500259
## DEPTH 2.241660  1        1.497217
## SOG   2.017183  1        1.420276

# You want GVIF to be less than ~3. Drop SOG (strongly correlated with year,
which is interesting
mod = lm(CPUE ~ YEAR + MONTH + BTEMP + DEPTH, data = dat)
vif(mod)

##           GVIF Df GVIF^(1/(2*Df))
## YEAR  1.404039 23        1.007405
## MONTH 1.405551  1        1.185560
## BTEMP 2.250510  1        1.500170
## DEPTH 2.240843  1        1.496945
```

Step 2: Model selection

Compare negative binom, zero-infl neg binom, zero-altered neg binom

Model 1: Negative Binomial

```
tmb1.NB <- glmmTMB(CPUE ~ YEAR + MONTH + BTEMP + DEPTH+offset(lnEffort), data
= dat, family = nbinom2)

# Check the Std. Error of the estimates; high SEs indicate problems with the
fit.
summary(tmb1.NB)

##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + BTEMP + DEPTH + offset(lnEffort)
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  11026.2  11192.6  -5485.1  10970.2     2787
##
##
## Dispersion parameter for nbinom2 family (): 0.184
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.7767520  0.3933421  -4.517 6.27e-06 ***
## YEAR1986    -0.1329080  0.4692543  -0.283 0.776999
## YEAR1987    -0.7771543  0.5235169  -1.484 0.137680
## YEAR1988    -0.2373454  0.5127685  -0.463 0.643457
## YEAR1989    -0.4584620  0.4949824  -0.926 0.354332
## YEAR1990    -1.9399629  0.5629118  -3.446 0.000568 ***
## YEAR1991    -1.1391753  0.4413152  -2.581 0.009842 **
## YEAR1992     0.1945058  0.4479037   0.434 0.664101
## YEAR1993    -0.9654816  0.4662636  -2.071 0.038389 *
## YEAR1994     0.2712168  0.4353402   0.623 0.533285
## YEAR1995    -0.5625545  0.4578331  -1.229 0.219172
## YEAR1996    -1.3979756  0.4485599  -3.117 0.001830 **
## YEAR1997    -1.3894209  0.4376364  -3.175 0.001499 **
## YEAR1998    -0.1882151  0.4240465  -0.444 0.657147
## YEAR1999     0.0495669  0.4157929   0.119 0.905109
## YEAR2000    -1.9728119  0.4510506  -4.374 1.22e-05 ***
## YEAR2001    -0.8252248  0.4414848  -1.869 0.061595 .
## YEAR2002    -0.3820539  0.4384592  -0.871 0.383560
## YEAR2003    -0.0683666  0.4221482  -0.162 0.871346
## YEAR2004    -0.3251374  0.4613469  -0.705 0.480962
## YEAR2005    -0.2094150  0.4349726  -0.481 0.630201
## YEAR2006    -0.1755306  0.4368450  -0.402 0.687821
## YEAR2007    -0.8423378  0.4421865  -1.905 0.056788 .
## YEAR2008    -1.1802693  0.6499494  -1.816 0.069380 .
## MONTH10     -0.0008976  0.1595320  -0.006 0.995511
## BTEMP        1.1297564  0.0974224  11.596  < 2e-16 ***
## DEPTH       -1.0387825  0.0849445 -12.229  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Calculate dispersion for a glmmTMB model
disp3(tmb1.NB)

## df.resid
## 3.307467
```

Model 2: ZINB

```r
ZINB <- glmmTMB(CPUE ~ YEAR + MONTH + BTEMP + DEPTH,offset=lnEffort,
ziformula = ~ YEAR + MONTH + BTEMP + DEPTH, data = dat, family=nbinom2)
summary(ZINB)

##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + BTEMP + DEPTH
## Zero inflation:        ~YEAR + MONTH + BTEMP + DEPTH
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##  10931.8  11258.6  -5410.9  10821.8     2760
##
##
## Dispersion parameter for nbinom2 family (): 0.211
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.02586    0.38694  -2.651 0.008021 **
## YEAR1986    -0.27688    0.46268  -0.598 0.549556
## YEAR1987    -1.20525    0.50881  -2.369 0.017847 *
## YEAR1988    -0.60095    0.51889  -1.158 0.246810
## YEAR1989    -0.64596    0.47871  -1.349 0.177210
## YEAR1990    -2.11185    0.58157  -3.631 0.000282 ***
## YEAR1991    -1.30282    0.43030  -3.028 0.002464 **
## YEAR1992    -0.42927    0.43781  -0.980 0.326841
## YEAR1993    -1.55059    0.47567  -3.260 0.001115 **
## YEAR1994    -0.14922    0.42583  -0.350 0.726019
## YEAR1995    -1.15454    0.43286  -2.667 0.007648 **
## YEAR1996    -1.69075    0.43397  -3.896 9.78e-05 ***
## YEAR1997    -1.54076    0.42958  -3.587 0.000335 ***
## YEAR1998    -0.17606    0.42666  -0.413 0.679858
## YEAR1999     0.00213    0.39424   0.005 0.995690
## YEAR2000    -2.06906    0.44086  -4.693 2.69e-06 ***
## YEAR2001    -1.07350    0.43319  -2.478 0.013206 *
## YEAR2002    -0.60456    0.42004  -1.439 0.150067
## YEAR2003     0.14808    0.40839   0.363 0.716900
## YEAR2004    -0.90987    0.44066  -2.065 0.038945 *
## YEAR2005    -0.24343    0.42664  -0.571 0.568284
## YEAR2006    -0.67377    0.41632  -1.618 0.105575
## YEAR2007    -1.04159    0.42466  -2.453 0.014177 *
## YEAR2008    -1.79265    0.60839  -2.947 0.003213 **
## MONTH10     -0.77253    0.17842  -4.330 1.49e-05 ***
```

```
## BTEMP        0.46655    0.11758   3.968 7.24e-05 ***
## DEPTH       -1.19583    0.08709 -13.730  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.7463     2.9985   0.249   0.8034
## YEAR1986     -1.6641     3.2458  -0.513   0.6082
## YEAR1987     -5.6863     3.4021  -1.671   0.0946 .
## YEAR1988     -6.6740     3.6153  -1.846   0.0649 .
## YEAR1989     -4.7921     4.1584  -1.152   0.2492
## YEAR1990     -1.7394     3.3819  -0.514   0.6070
## YEAR1991     -3.1497     3.2183  -0.979   0.3277
## YEAR1992     -4.6668     3.2781  -1.424   0.1546
## YEAR1993     -5.0177     3.2967  -1.522   0.1280
## YEAR1994     -3.4663     3.2345  -1.072   0.2839
## YEAR1995    -19.8109   676.9360  -0.029   0.9767
## YEAR1996     -3.7422     3.2130  -1.165   0.2441
## YEAR1997     -1.4674     3.1479  -0.466   0.6411
## YEAR1998      2.1735     3.0575   0.711   0.4772
## YEAR1999     -0.8506     3.2054  -0.265   0.7907
## YEAR2000     -2.3676     3.1344  -0.755   0.4500
## YEAR2001     -3.5519     3.1814  -1.116   0.2642
## YEAR2002     -4.2717     3.2781  -1.303   0.1925
## YEAR2003     -1.4648     3.0702  -0.477   0.6333
## YEAR2004     -6.0228     3.3211  -1.814   0.0698 .
## YEAR2005     -3.5606     3.1958  -1.114   0.2652
## YEAR2006     -6.8019     4.9436  -1.376   0.1689
## YEAR2007     -4.0865     3.5084  -1.165   0.2441
## YEAR2008    -15.9121    62.7241  -0.254   0.7997
## MONTH10      -9.2706     2.0330  -4.560 5.11e-06 ***
## BTEMP        -4.2199     0.8853  -4.767 1.87e-06 ***
## DEPTH        -1.2515     0.5562  -2.250   0.0244 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# large SEs, so removing YEAR from ZI part of model:
ZINB.2 <- glmmTMB(CPUE ~ YEAR + MONTH + BTEMP + DEPTH,offset=lnEffort,
ziformula = ~ MONTH + BTEMP + DEPTH, data = dat, family=nbinom2)
summary(ZINB.2)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + BTEMP + DEPTH
## Zero inflation:        ~MONTH + BTEMP + DEPTH
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##  10958.9  11149.1  -5447.5  10894.9     2783
```

```
##
##
## Dispersion parameter for nbinom2 family (): 0.205
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.28983    0.38875  -3.318 0.000907 ***
## YEAR1986    -0.18769    0.46015  -0.408 0.683355
## YEAR1987    -0.94944    0.51935  -1.828 0.067531 .
## YEAR1988    -0.19239    0.54586  -0.352 0.724490
## YEAR1989    -0.42038    0.48239  -0.871 0.383502
## YEAR1990    -2.01957    0.54967  -3.674 0.000239 ***
## YEAR1991    -1.15041    0.42812  -2.687 0.007207 **
## YEAR1992    -0.17964    0.44322  -0.405 0.685256
## YEAR1993    -1.18311    0.48042  -2.463 0.013791 *
## YEAR1994     0.04833    0.42607   0.113 0.909690
## YEAR1995    -0.83040    0.43916  -1.891 0.058638 .
## YEAR1996    -1.47723    0.43373  -3.406 0.000660 ***
## YEAR1997    -1.48935    0.42400  -3.513 0.000444 ***
## YEAR1998    -0.35829    0.41324  -0.867 0.385930
## YEAR1999     0.03153    0.39640   0.080 0.936606
## YEAR2000    -1.96894    0.43292  -4.548 5.42e-06 ***
## YEAR2001    -0.89472    0.43202  -2.071 0.038358 *
## YEAR2002    -0.39461    0.42262  -0.934 0.350444
## YEAR2003     0.09867    0.40514   0.244 0.807583
## YEAR2004    -0.59708    0.44796  -1.333 0.182568
## YEAR2005    -0.07223    0.42926  -0.168 0.866374
## YEAR2006    -0.37609    0.42072  -0.894 0.371364
## YEAR2007    -0.85923    0.42794  -2.008 0.044662 *
## YEAR2008    -1.57169    0.61692  -2.548 0.010845 *
## MONTH10     -0.53349    0.17713  -3.012 0.002597 **
## BTEMP        0.48465    0.12384   3.914 9.09e-05 ***
## DEPTH       -1.23724    0.08846 -13.987  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.6629     0.4753  -5.603 2.10e-08 ***
## MONTH10      -2.8390     1.0152  -2.796  0.00517 **
## BTEMP        -3.2321     0.4506  -7.173 7.34e-13 ***
## DEPTH        -1.0804     0.3408  -3.170  0.00152 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tmbZI.resids(ZINB.2, y.dat=dat$CPUE)
```

```
## df.resid
## 2.704786
```

Model 3: ZANB (aka hurdle model)

```
ZANB <- glmmTMB(CPUE ~ YEAR + MONTH + BTEMP + DEPTH, ziformula = ~ YEAR +
MONTH + BTEMP + DEPTH, data = dat,offset=lnEffort,family =
truncated_nbinom2(link = "log"))
summary(ZANB)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + BTEMP + DEPTH
## Zero inflation:        ~YEAR + MONTH + BTEMP + DEPTH
## Data: dat
##  Offset: lnEffort
## 
##      AIC      BIC   logLik deviance df.resid
##  10800.0  11126.8  -5345.0  10690.0     2760
## 
## 
## Dispersion parameter for truncated_nbinom2 family (): 4.16e-09
## 
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -17.88998 1329.85567  -0.013 0.989267
## YEAR1986      -0.46021    0.65756  -0.700 0.484000
## YEAR1987      -1.69495    0.68578  -2.472 0.013452 *
## YEAR1988      -0.73474    0.78353  -0.938 0.348383
## YEAR1989      -1.23432    0.67373  -1.832 0.066939 .
## YEAR1990      -2.51409    0.83374  -3.015 0.002566 **
## YEAR1991      -1.36168    0.64705  -2.104 0.035339 *
## YEAR1992      -0.73614    0.61209  -1.203 0.229100
## YEAR1993      -2.19737    0.65064  -3.377 0.000732 ***
## YEAR1994      -0.40124    0.59579  -0.673 0.500653
## YEAR1995      -1.92231    0.59475  -3.232 0.001229 **
## YEAR1996      -2.21653    0.61151  -3.625 0.000289 ***
## YEAR1997      -2.01166    0.59871  -3.360 0.000779 ***
## YEAR1998       0.07174    0.59437   0.121 0.903928
## YEAR1999       0.22696    0.55108   0.412 0.680449
## YEAR2000      -2.82272    0.60738  -4.647 3.36e-06 ***
## YEAR2001      -1.69810    0.59531  -2.852 0.004338 **
## YEAR2002      -1.15846    0.58447  -1.982 0.047472 *
## YEAR2003       0.23140    0.56817   0.407 0.683810
## YEAR2004      -1.10073    0.64946  -1.695 0.090104 .
## YEAR2005      -0.51246    0.61456  -0.834 0.404360
## YEAR2006      -1.25900    0.57457  -2.191 0.028438 *
## YEAR2007      -1.60626    0.58853  -2.729 0.006347 **
## YEAR2008      -2.46563    0.83675  -2.947 0.003212 **
## MONTH10       -1.38437    0.26524  -5.219 1.80e-07 ***
## BTEMP          0.57960    0.17625   3.288 0.001007 **
## DEPTH         -1.14196    0.12882  -8.865  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
```

```
## Zero-inflation model:
##                Estimate Std. Error z value Pr(>|z|)    
## (Intercept)  4.249e-01  3.518e-01   1.208  0.22709    
## YEAR1986     3.014e-01  4.255e-01   0.708  0.47879    
## YEAR1987     1.597e-02  4.609e-01   0.035  0.97236    
## YEAR1988     2.619e-01  4.834e-01   0.542  0.58794    
## YEAR1989     7.494e-02  4.391e-01   0.171  0.86448    
## YEAR1990     9.566e-01  5.057e-01   1.892  0.05854 .  
## YEAR1991     1.107e+00  4.018e-01   2.755  0.00588 ** 
## YEAR1992     1.825e-01  4.003e-01   0.456  0.64840    
## YEAR1993     3.891e-01  4.038e-01   0.964  0.33515    
## YEAR1994     7.227e-05  3.885e-01   0.000  0.99985    
## YEAR1995    -2.102e-01  3.978e-01  -0.528  0.59729    
## YEAR1996     4.583e-01  3.985e-01   1.150  0.25015    
## YEAR1997     8.149e-01  3.920e-01   2.079  0.03765 *  
## YEAR1998     8.670e-01  3.908e-01   2.219  0.02652 *  
## YEAR1999     2.751e-01  3.809e-01   0.722  0.47003    
## YEAR2000     6.247e-01  3.951e-01   1.581  0.11380    
## YEAR2001     1.711e-01  3.941e-01   0.434  0.66426    
## YEAR2002    -3.384e-02  3.906e-01  -0.087  0.93096    
## YEAR2003     3.941e-01  3.884e-01   1.015  0.31033    
## YEAR2004     1.188e-01  3.929e-01   0.302  0.76242    
## YEAR2005     3.451e-01  3.939e-01   0.876  0.38096    
## YEAR2006    -7.018e-01  3.970e-01  -1.768  0.07710 .  
## YEAR2007     3.067e-01  3.947e-01   0.777  0.43722    
## YEAR2008     5.563e-01  5.492e-01   1.013  0.31109    
## MONTH10     -7.379e-01  1.249e-01  -5.910 3.42e-09 ***
## BTEMP       -6.685e-01  7.203e-02  -9.282  < 2e-16 ***
## DEPTH        6.979e-01  8.327e-02   8.381  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# large SE on intercept in conditional model, so removing YEAR
ZANB.1 <- glmmTMB(CPUE ~ MONTH + BTEMP + DEPTH, ziformula = ~ YEAR + MONTH +
BTEMP + DEPTH, data = dat,offset=lnEffort,family = truncated_nbinom2(link =
"log"))
summary(ZANB.1)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ MONTH + BTEMP + DEPTH
## Zero inflation:        ~YEAR + MONTH + BTEMP + DEPTH
## Data: dat
##  Offset: lnEffort
## 
##      AIC      BIC   logLik deviance df.resid 
##  10848.6  11038.8  -5392.3  10784.6     2783 
## 
## 
## Dispersion parameter for truncated_nbinom2 family (): 3.55e-09 
## 
```

```
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -19.3566  1172.5498  -0.017  0.98683
## MONTH10       -0.3842     0.1761  -2.181  0.02917 *
## BTEMP          0.4301     0.1579   2.723  0.00647 **
## DEPTH         -1.4644     0.1135 -12.901  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  4.250e-01  3.518e-01   1.208  0.22704
## YEAR1986     3.013e-01  4.255e-01   0.708  0.47884
## YEAR1987     1.592e-02  4.609e-01   0.035  0.97245
## YEAR1988     2.619e-01  4.834e-01   0.542  0.58798
## YEAR1989     7.489e-02  4.391e-01   0.171  0.86458
## YEAR1990     9.566e-01  5.057e-01   1.892  0.05854 .
## YEAR1991     1.107e+00  4.018e-01   2.755  0.00588 **
## YEAR1992     1.825e-01  4.003e-01   0.456  0.64847
## YEAR1993     3.891e-01  4.038e-01   0.964  0.33521
## YEAR1994     2.745e-05  3.885e-01   0.000  0.99994
## YEAR1995    -2.102e-01  3.978e-01  -0.528  0.59722
## YEAR1996     4.582e-01  3.985e-01   1.150  0.25020
## YEAR1997     8.149e-01  3.920e-01   2.079  0.03766 *
## YEAR1998     8.669e-01  3.908e-01   2.218  0.02653 *
## YEAR1999     2.751e-01  3.809e-01   0.722  0.47010
## YEAR2000     6.247e-01  3.951e-01   1.581  0.11383
## YEAR2001     1.710e-01  3.941e-01   0.434  0.66434
## YEAR2002    -3.387e-02  3.906e-01  -0.087  0.93088
## YEAR2003     3.940e-01  3.884e-01   1.014  0.31037
## YEAR2004     1.187e-01  3.929e-01   0.302  0.76251
## YEAR2005     3.450e-01  3.939e-01   0.876  0.38103
## YEAR2006    -7.018e-01  3.970e-01  -1.768  0.07708 .
## YEAR2007     3.066e-01  3.947e-01   0.777  0.43728
## YEAR2008     5.563e-01  5.492e-01   1.013  0.31113
## MONTH10     -7.380e-01  1.249e-01  -5.910 3.42e-09 ***
## BTEMP       -6.685e-01  7.203e-02  -9.282  < 2e-16 ***
## DEPTH        6.979e-01  8.327e-02   8.381  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# still large SE on intercept in conditional model, so removing non-sig
factors. Remove BTEMP
ZANB.2 <- glmmTMB(CPUE ~ MONTH + DEPTH, ziformula = ~ YEAR + MONTH + BTEMP +
DEPTH, data = dat,offset=lnEffort,family = truncated_nbinom2(link = "log"))
summary(ZANB.2)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ MONTH + DEPTH
## Zero inflation:        ~YEAR + MONTH + BTEMP + DEPTH
```

```
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##  10854.0  11038.3  -5396.0  10792.0     2784
##
##
## Dispersion parameter for truncated_nbinom2 family (): 3.92e-09
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -19.10151 1125.17493  -0.017  0.98646
## MONTH10       -0.54828    0.16634  -3.296  0.00098 ***
## DEPTH         -1.66853    0.08711 -19.155  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  4.249e-01  3.518e-01   1.208  0.22707
## YEAR1986     3.014e-01  4.255e-01   0.708  0.47881
## YEAR1987     1.594e-02  4.609e-01   0.035  0.97241
## YEAR1988     2.619e-01  4.834e-01   0.542  0.58796
## YEAR1989     7.491e-02  4.391e-01   0.171  0.86454
## YEAR1990     9.566e-01  5.057e-01   1.892  0.05854 .
## YEAR1991     1.107e+00  4.018e-01   2.755  0.00588 **
## YEAR1992     1.825e-01  4.003e-01   0.456  0.64844
## YEAR1993     3.891e-01  4.038e-01   0.964  0.33519
## YEAR1994     5.501e-05  3.885e-01   0.000  0.99989
## YEAR1995    -2.102e-01  3.978e-01  -0.528  0.59725
## YEAR1996     4.583e-01  3.985e-01   1.150  0.25018
## YEAR1997     8.149e-01  3.920e-01   2.079  0.03765 *
## YEAR1998     8.669e-01  3.908e-01   2.218  0.02652 *
## YEAR1999     2.751e-01  3.809e-01   0.722  0.47006
## YEAR2000     6.247e-01  3.951e-01   1.581  0.11382
## YEAR2001     1.710e-01  3.941e-01   0.434  0.66430
## YEAR2002    -3.385e-02  3.906e-01  -0.087  0.93093
## YEAR2003     3.941e-01  3.884e-01   1.014  0.31035
## YEAR2004     1.188e-01  3.929e-01   0.302  0.76247
## YEAR2005     3.451e-01  3.939e-01   0.876  0.38100
## YEAR2006    -7.018e-01  3.970e-01  -1.768  0.07709 .
## YEAR2007     3.066e-01  3.947e-01   0.777  0.43725
## YEAR2008     5.563e-01  5.492e-01   1.013  0.31112
## MONTH10     -7.380e-01  1.249e-01  -5.910 3.42e-09 ***
## BTEMP       -6.685e-01  7.202e-02  -9.282  < 2e-16 ***
## DEPTH        6.979e-01  8.327e-02   8.381  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# still large SE on intercept in conditional model, so removing non-sig
factors. Remove MONTH
ZANB.3 <- glmmTMB(CPUE ~ DEPTH, ziformula = ~ YEAR + MONTH + BTEMP + DEPTH,
data = dat,offset=lnEffort,family = truncated_nbinom2(link = "log"))
summary(ZANB.3)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ DEPTH
## Zero inflation:        ~YEAR + MONTH + BTEMP + DEPTH
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##  10861.9  11040.2  -5401.0  10801.9     2785
##
##
## Dispersion parameter for truncated_nbinom2 family (): 3.39e-09
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -19.40932 1225.55415  -0.016    0.987
## DEPTH         -1.73532    0.08762 -19.806   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  4.250e-01  3.518e-01   1.208  0.22704
## YEAR1986     3.013e-01  4.255e-01   0.708  0.47886
## YEAR1987     1.591e-02  4.609e-01   0.035  0.97246
## YEAR1988     2.619e-01  4.834e-01   0.542  0.58801
## YEAR1989     7.487e-02  4.391e-01   0.171  0.86460
## YEAR1990     9.566e-01  5.057e-01   1.892  0.05855 .
## YEAR1991     1.107e+00  4.018e-01   2.755  0.00588 **
## YEAR1992     1.825e-01  4.003e-01   0.456  0.64848
## YEAR1993     3.891e-01  4.038e-01   0.964  0.33522
## YEAR1994     2.391e-05  3.885e-01   0.000  0.99995
## YEAR1995    -2.102e-01  3.978e-01  -0.528  0.59720
## YEAR1996     4.582e-01  3.985e-01   1.150  0.25020
## YEAR1997     8.149e-01  3.920e-01   2.079  0.03766 *
## YEAR1998     8.669e-01  3.908e-01   2.218  0.02653 *
## YEAR1999     2.751e-01  3.809e-01   0.722  0.47011
## YEAR2000     6.247e-01  3.951e-01   1.581  0.11383
## YEAR2001     1.710e-01  3.941e-01   0.434  0.66436
## YEAR2002    -3.388e-02  3.906e-01  -0.087  0.93088
## YEAR2003     3.940e-01  3.884e-01   1.014  0.31038
## YEAR2004     1.187e-01  3.929e-01   0.302  0.76253
## YEAR2005     3.450e-01  3.939e-01   0.876  0.38103
## YEAR2006    -7.018e-01  3.970e-01  -1.768  0.07708 .
## YEAR2007     3.066e-01  3.947e-01   0.777  0.43729
```

```
## YEAR2008     5.562e-01  5.492e-01   1.013  0.31115
## MONTH10     -7.380e-01  1.249e-01  -5.910 3.42e-09 ***
## BTEMP       -6.685e-01  7.202e-02  -9.282  < 2e-16 ***
## DEPTH        6.979e-01  8.327e-02   8.381  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# still have high SE on intercept, so can't use hurdle model
```

Preliminary model comparisons

```
AICtab(tmb1.NB, ZINB.2)

##         dAIC df
## ZINB.2   0.0 32
## tmb1.NB 67.3 28

# For now, ZINB better than NB, but model selection below fine tunes this
```

Step 3: Factor selection (by model)

Recall that I have removed ZANB from consideration due to high standard errors and lack of a conditional model

Negative Binomial

```
NB1 <- glmmTMB(CPUE ~ YEAR + MONTH + BTEMP + DEPTH+offset(lnEffort), data =
dat, family = nbinom2)
summary(NB1)

##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + BTEMP + DEPTH + offset(lnEffort)
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  11026.2  11192.6  -5485.1  10970.2     2787
##
##
## Dispersion parameter for nbinom2 family (): 0.184
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.7767520  0.3933421  -4.517 6.27e-06 ***
## YEAR1986    -0.1329080  0.4692543  -0.283 0.776999
## YEAR1987    -0.7771543  0.5235169  -1.484 0.137680
## YEAR1988    -0.2373454  0.5127685  -0.463 0.643457
## YEAR1989    -0.4584620  0.4949824  -0.926 0.354332
## YEAR1990    -1.9399629  0.5629118  -3.446 0.000568 ***
## YEAR1991    -1.1391753  0.4413152  -2.581 0.009842 **
## YEAR1992     0.1945058  0.4479037   0.434 0.664101
## YEAR1993    -0.9654816  0.4662636  -2.071 0.038389 *
## YEAR1994     0.2712168  0.4353402   0.623 0.533285
## YEAR1995    -0.5625545  0.4578331  -1.229 0.219172
```

```
## YEAR1996    -1.3979756  0.4485599  -3.117 0.001830 **
## YEAR1997    -1.3894209  0.4376364  -3.175 0.001499 **
## YEAR1998    -0.1882151  0.4240465  -0.444 0.657147
## YEAR1999     0.0495669  0.4157929   0.119 0.905109
## YEAR2000    -1.9728119  0.4510506  -4.374 1.22e-05 ***
## YEAR2001    -0.8252248  0.4414848  -1.869 0.061595 .
## YEAR2002    -0.3820539  0.4384592  -0.871 0.383560
## YEAR2003    -0.0683666  0.4221482  -0.162 0.871346
## YEAR2004    -0.3251374  0.4613469  -0.705 0.480962
## YEAR2005    -0.2094150  0.4349726  -0.481 0.630201
## YEAR2006    -0.1755306  0.4368450  -0.402 0.687821
## YEAR2007    -0.8423378  0.4421865  -1.905 0.056788 .
## YEAR2008    -1.1802693  0.6499494  -1.816 0.069380 .
## MONTH10     -0.0008976  0.1595320  -0.006 0.995511
## BTEMP        1.1297564  0.0974224  11.596  < 2e-16 ***
## DEPTH       -1.0387825  0.0849445 -12.229  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(NB1,test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + BTEMP + DEPTH + offset(lnEffort)
##        Df   AIC    LRT  Pr(>Chi)
## <none>    11026
## YEAR   23 11104 124.22 7.192e-16 ***
## MONTH   1 11024   0.00    0.9955
## BTEMP   1 11160 135.33 < 2.2e-16 ***
## DEPTH   1 11175 150.38 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# drop MONTH
NB2 <- glmmTMB(CPUE ~ YEAR + BTEMP + DEPTH+offset(lnEffort), data = dat,
family = nbinom2)
summary(NB2)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + BTEMP + DEPTH + offset(lnEffort)
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  11024.2  11184.7  -5485.1  10970.2     2788
##
##
## Dispersion parameter for nbinom2 family (): 0.184
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept) -1.77759    0.36332  -4.893 9.95e-07 ***
## YEAR1986    -0.13243    0.46174  -0.287 0.774258
## YEAR1987    -0.77666    0.51551  -1.507 0.131918
## YEAR1988    -0.23663    0.49747  -0.476 0.634307
## YEAR1989    -0.45786    0.48291  -0.948 0.343071
## YEAR1990    -1.93918    0.54525  -3.557 0.000376 ***
## YEAR1991    -1.13844    0.42375  -2.687 0.007218 **
## YEAR1992     0.19512    0.43204   0.452 0.651542
## YEAR1993    -0.96475    0.44487  -2.169 0.030114 *
## YEAR1994     0.27186    0.41885   0.649 0.516301
## YEAR1995    -0.56173    0.43360  -1.296 0.195147
## YEAR1996    -1.39723    0.42846  -3.261 0.001110 **
## YEAR1997    -1.38890    0.42671  -3.255 0.001134 **
## YEAR1998    -0.18816    0.42396  -0.444 0.657180
## YEAR1999     0.04963    0.41544   0.119 0.904915
## YEAR2000    -1.97215    0.43392  -4.545 5.49e-06 ***
## YEAR2001    -0.82451    0.42138  -1.957 0.050384 .
## YEAR2002    -0.38138    0.42199  -0.904 0.366116
## YEAR2003    -0.06801    0.41688  -0.163 0.870410
## YEAR2004    -0.32434    0.43991  -0.737 0.460945
## YEAR2005    -0.20879    0.41903  -0.498 0.618294
## YEAR2006    -0.17478    0.41493  -0.421 0.673589
## YEAR2007    -0.84181    0.42978  -1.959 0.050149 .
## YEAR2008    -1.17974    0.64188  -1.838 0.066069 .
## BTEMP        1.12981    0.09692  11.658  < 2e-16 ***
## DEPTH       -1.03890    0.08295 -12.525  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(NB2, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + BTEMP + DEPTH + offset(lnEffort)
##        Df   AIC    LRT  Pr(>Chi)
## <none>    11024
## YEAR   23 11107 128.84 < 2.2e-16 ***
## BTEMP   1 11160 137.52 < 2.2e-16 ***
## DEPTH   1 11177 154.72 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Let's compare to a model with year only.
NB0 <- glmmTMB(CPUE ~ YEAR+offset(lnEffort), data = dat, family = nbinom2)

# compare models with AIC:
AICtab(NB1, NB2, NB0)
```

```
##     dAIC  df
## NB2   0.0 27
```

```
## NB1   2.0 28
## NB0 825.4 25
```

```
lrtest(NB0, NB2, NB1)
```

```
## Likelihood ratio test
##
## Model 1: CPUE ~ YEAR + offset(lnEffort)
## Model 2: CPUE ~ YEAR + BTEMP + DEPTH + offset(lnEffort)
## Model 3: CPUE ~ YEAR + MONTH + BTEMP + DEPTH + offset(lnEffort)
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1  25 -5899.8
## 2  27 -5485.1  2 829.37     <2e-16 ***
## 3  28 -5485.1  1   0.00     0.9955
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sapply(list(NB2,NB1,NB0),disp3)
```

```
## df.resid df.resid df.resid
## 3.307005 3.307467 2.455413
```

ZINB

```
ZINB.1 = glmmTMB(CPUE ~ YEAR + MONTH + BTEMP + DEPTH,offset=lnEffort,
ziformula = ~ MONTH + BTEMP + DEPTH, data = dat, family=nbinom2)
summary(ZINB.1)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + BTEMP + DEPTH
## Zero inflation:        ~MONTH + BTEMP + DEPTH
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##  10958.9  11149.1  -5447.5  10894.9     2783
##
##
## Dispersion parameter for nbinom2 family (): 0.205
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.28983    0.38875  -3.318 0.000907 ***
## YEAR1986    -0.18769    0.46015  -0.408 0.683355
## YEAR1987    -0.94944    0.51935  -1.828 0.067531 .
## YEAR1988    -0.19239    0.54586  -0.352 0.724490
## YEAR1989    -0.42038    0.48239  -0.871 0.383502
## YEAR1990    -2.01957    0.54967  -3.674 0.000239 ***
## YEAR1991    -1.15041    0.42812  -2.687 0.007207 **
## YEAR1992    -0.17964    0.44322  -0.405 0.685256
## YEAR1993    -1.18311    0.48042  -2.463 0.013791 *
## YEAR1994     0.04833    0.42607   0.113 0.909690
```

```
## YEAR1995    -0.83040    0.43916  -1.891 0.058638 .
## YEAR1996    -1.47723    0.43373  -3.406 0.000660 ***
## YEAR1997    -1.48935    0.42400  -3.513 0.000444 ***
## YEAR1998    -0.35829    0.41324  -0.867 0.385930
## YEAR1999     0.03153    0.39640   0.080 0.936606
## YEAR2000    -1.96894    0.43292  -4.548 5.42e-06 ***
## YEAR2001    -0.89472    0.43202  -2.071 0.038358 *
## YEAR2002    -0.39461    0.42262  -0.934 0.350444
## YEAR2003     0.09867    0.40514   0.244 0.807583
## YEAR2004    -0.59708    0.44796  -1.333 0.182568
## YEAR2005    -0.07223    0.42926  -0.168 0.866374
## YEAR2006    -0.37609    0.42072  -0.894 0.371364
## YEAR2007    -0.85923    0.42794  -2.008 0.044662 *
## YEAR2008    -1.57169    0.61692  -2.548 0.010845 *
## MONTH10     -0.53349    0.17713  -3.012 0.002597 **
## BTEMP        0.48465    0.12384   3.914 9.09e-05 ***
## DEPTH       -1.23724    0.08846 -13.987  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.6629     0.4753  -5.603 2.10e-08 ***
## MONTH10      -2.8390     1.0152  -2.796  0.00517 **
## BTEMP        -3.2321     0.4506  -7.173 7.34e-13 ***
## DEPTH        -1.0804     0.3408  -3.170  0.00152 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZINB.1, test="Chi") # all covars significant in main model, so skipping
this subsetting
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + BTEMP + DEPTH
##        Df   AIC     LRT  Pr(>Chi)
## <none>    10959
## YEAR   23 11031 118.434 7.952e-15 ***
## MONTH   1 10966   9.071  0.002596 **
## BTEMP   1 10972  15.220 9.568e-05 ***
## DEPTH   1 11146 188.600 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# everything is significant, so let's try a year-only model
ZINB.0 <- glmmTMB(CPUE ~ YEAR,offset=lnEffort, data = dat, family=nbinom2)
summary(ZINB.0)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR
```

```
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##  11849.6  11998.2  -5899.8  11799.6     2790
##
##
## Dispersion parameter for nbinom2 family (): 0.112
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.25616    0.43442  -2.892  0.00383 **
## YEAR1986     0.60541    0.55288   1.095  0.27351
## YEAR1987    -0.54623    0.59329  -0.921  0.35722
## YEAR1988     0.75438    0.59353   1.271  0.20373
## YEAR1989     0.52645    0.57214   0.920  0.35750
## YEAR1990    -1.27739    0.64160  -1.991  0.04649 *
## YEAR1991     0.08881    0.50070   0.177  0.85922
## YEAR1992     0.81693    0.51600   1.583  0.11337
## YEAR1993    -1.56259    0.50607  -3.088  0.00202 **
## YEAR1994     0.57589    0.49959   1.153  0.24902
## YEAR1995    -0.38219    0.50522  -0.756  0.44936
## YEAR1996    -0.35509    0.50246  -0.707  0.47975
## YEAR1997    -0.82504    0.50094  -1.647  0.09956 .
## YEAR1998    -0.11293    0.50086  -0.226  0.82161
## YEAR1999     0.74355    0.49993   1.487  0.13694
## YEAR2000    -1.05780    0.50184  -2.108  0.03504 *
## YEAR2001     0.22951    0.49895   0.460  0.64552
## YEAR2002     0.70846    0.50078   1.415  0.15716
## YEAR2003     0.97099    0.49903   1.946  0.05169 .
## YEAR2004    -0.08474    0.50258  -0.169  0.86611
## YEAR2005     1.10350    0.50023   2.206  0.02739 *
## YEAR2006     0.67745    0.49495   1.369  0.17109
## YEAR2007     0.16929    0.50666   0.334  0.73828
## YEAR2008    -0.52407    0.71776  -0.730  0.46530
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
AICtab(ZINB.1,ZINB.0)
```

```
##        dAIC  df
## ZINB.1   0.0 32
## ZINB.0 890.7 25
```

```
lrtest(ZINB.0,ZINB.1)
```

```
## Likelihood ratio test
##
## Model 1: CPUE ~ YEAR
## Model 2: CPUE ~ YEAR + MONTH + BTEMP + DEPTH
##   #Df  LogLik Df  Chisq Pr(>Chisq)
```

```
## 1  25 -5899.8
## 2  32 -5447.5  7 904.68  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sapply(list(ZINB.1,ZINB.0),tmbZI.resids,y.dat=dat$CPUE)
```

```
## Warning in sqrt(pvar): NaNs produced
```

```
## df.resid df.resid
## 2.704786      NaN
```

```
tmbZI.resids(ZINB.1,dat$CPUE)
```

```
## df.resid
## 2.704786
```

Model comparison:

```
AICtab(NB2,ZINB.1) # ZINB.1 wins
```

```
##        dAIC df
## ZINB.1  0.0 32
## NB2    65.3 27
```

Model diagnostics

Negative binomial & ZINB

```
# remove depth:
ZINB.1b = glmmTMB(CPUE ~ YEAR + MONTH + BTEMP,offset=lnEffort, ziformula = ~
MONTH + BTEMP, data = dat, family=nbinom2)
summary(ZINB.1b)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + BTEMP
## Zero inflation:        ~MONTH + BTEMP
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##  11143.6  11321.9  -5541.8  11083.6     2785
##
##
## Dispersion parameter for nbinom2 family (): 0.176
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.7840     0.4136  -1.895 0.058035 .
## YEAR1986     -0.3808     0.4814  -0.791 0.428884
## YEAR1987     -1.0557     0.5466  -1.931 0.053424 .
## YEAR1988     -0.1050     0.5759  -0.182 0.855269
## YEAR1989     -0.9757     0.5145  -1.896 0.057899 .
## YEAR1990     -2.5623     0.5920  -4.328 1.50e-05 ***
```

```
## YEAR1991     -1.4462     0.4594  -3.148 0.001646 **
## YEAR1992      0.1987     0.4685   0.424 0.671516
## YEAR1993     -1.4174     0.4796  -2.955 0.003123 **
## YEAR1994     -0.2505     0.4497  -0.557 0.577539
## YEAR1995     -1.7052     0.4624  -3.687 0.000227 ***
## YEAR1996     -1.8360     0.4653  -3.946 7.96e-05 ***
## YEAR1997     -1.8837     0.4476  -4.208 2.57e-05 ***
## YEAR1998      0.4341     0.4291   1.011 0.311781
## YEAR1999      0.5026     0.4213   1.193 0.232914
## YEAR2000     -2.6202     0.4591  -5.707 1.15e-08 ***
## YEAR2001     -1.1238     0.4602  -2.442 0.014612 *
## YEAR2002     -0.8973     0.4516  -1.987 0.046930 *
## YEAR2003      0.1721     0.4386   0.392 0.694728
## YEAR2004     -1.2138     0.4602  -2.637 0.008356 **
## YEAR2005     -0.4123     0.4588  -0.899 0.368868
## YEAR2006     -0.6475     0.4517  -1.434 0.151712
## YEAR2007     -1.2041     0.4555  -2.643 0.008209 **
## YEAR2008     -1.4258     0.6551  -2.177 0.029517 *
## MONTH10      -0.9284     0.2007  -4.626 3.73e-06 ***
## BTEMP         1.5626     0.1177  13.279  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.6846     0.5901  -4.549 5.39e-06 ***
## MONTH10      -3.0807     2.1018  -1.466    0.143
## BTEMP        -2.4950     0.4056  -6.151 7.69e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

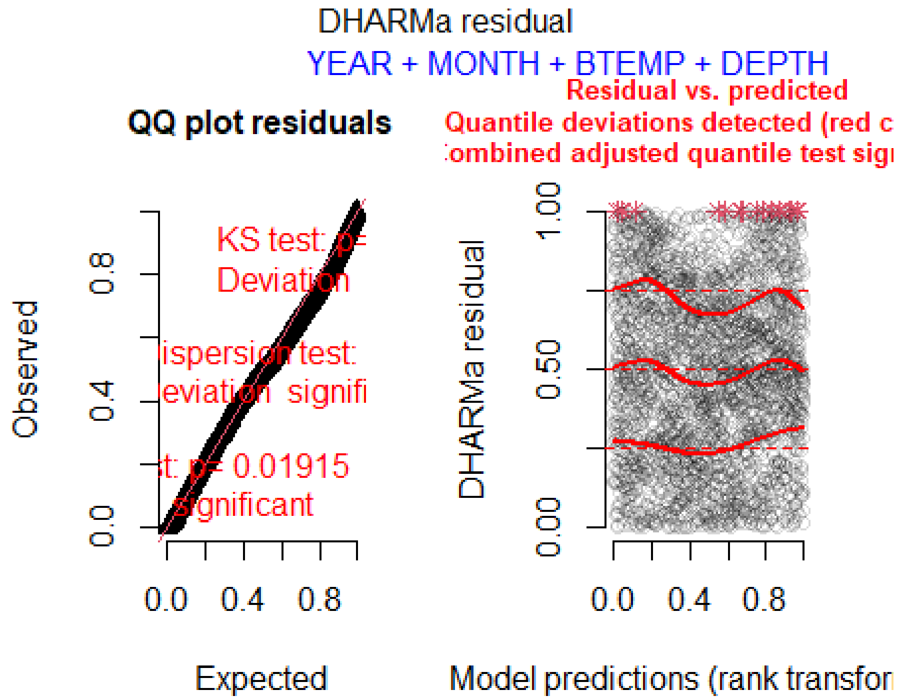
```
drop1(ZINB.1b, test="Chi") # month in ZI no longer significant
```

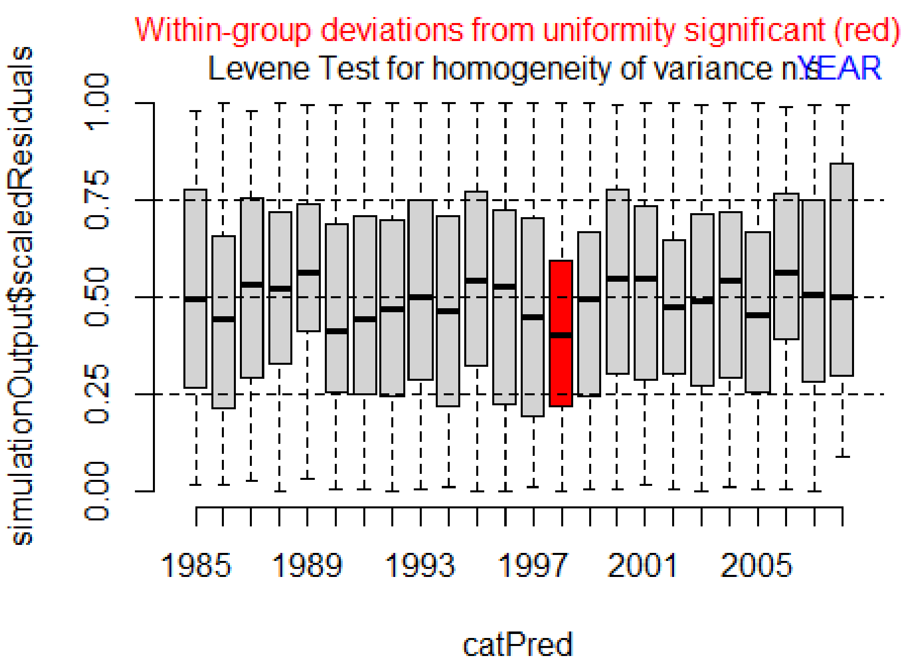
```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + BTEMP
##        Df   AIC     LRT  Pr(>Chi)
## <none>    11144
## YEAR   23 11299 201.432 < 2.2e-16 ***
## MONTH   1 11162  20.827 5.026e-06 ***
## BTEMP   1 11294 152.410 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

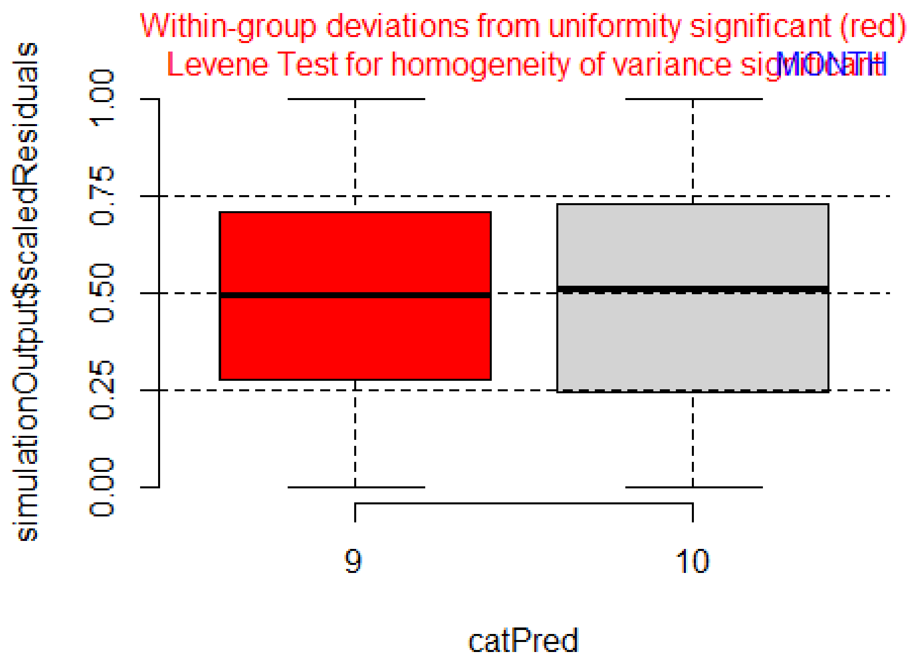
```
# drop MONTH from zi model:
ZINB.1c = glmmTMB(CPUE ~ YEAR + MONTH + BTEMP,offset=lnEffort, ziformula = ~
BTEMP, data = dat, family=nbinom2)
summary(ZINB.1c)
```

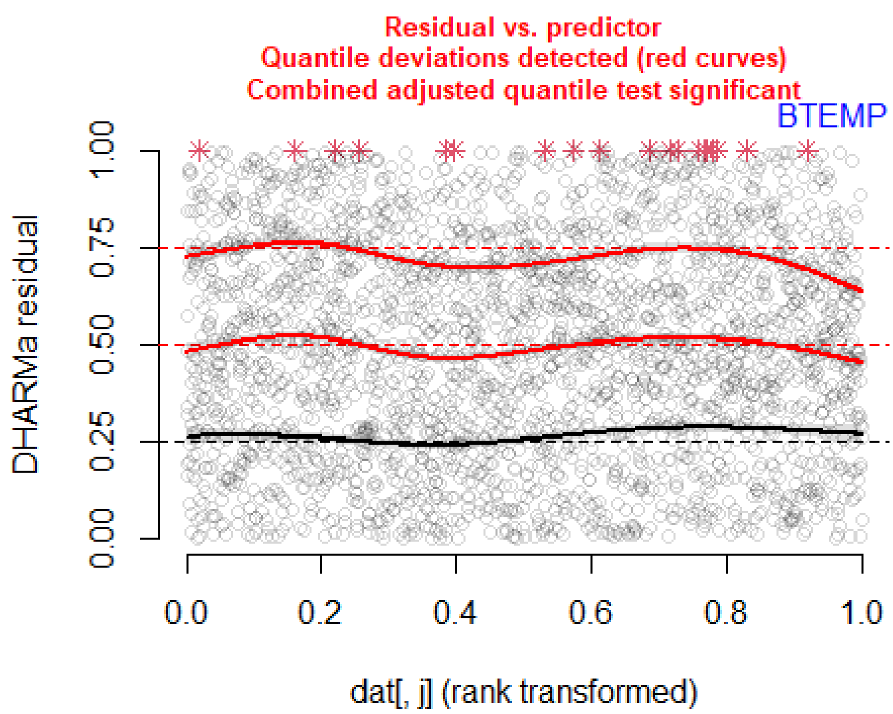
```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + BTEMP
## Zero inflation:        ~BTEMP
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##  11156.2  11328.5  -5549.1  11098.2     2786
##
##
## Dispersion parameter for nbinom2 family (): 0.175
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.02082    0.42224  -2.418 0.015621 *
## YEAR1986    -0.22140    0.50192  -0.441 0.659143
## YEAR1987    -0.91729    0.56105  -1.635 0.102057
## YEAR1988     0.04146    0.58071   0.071 0.943081
## YEAR1989    -0.78266    0.53176  -1.472 0.141062
## YEAR1990    -2.38379    0.59604  -3.999 6.35e-05 ***
## YEAR1991    -1.27579    0.47204  -2.703 0.006877 **
## YEAR1992     0.42215    0.48127   0.877 0.380405
## YEAR1993    -1.24537    0.48672  -2.559 0.010507 *
## YEAR1994    -0.07352    0.46215  -0.159 0.873601
## YEAR1995    -1.51251    0.47374  -3.193 0.001409 **
## YEAR1996    -1.65796    0.47705  -3.475 0.000510 ***
## YEAR1997    -1.74181    0.46064  -3.781 0.000156 ***
## YEAR1998     0.49877    0.44597   1.118 0.263399
## YEAR1999     0.49729    0.43630   1.140 0.254370
## YEAR2000    -2.47711    0.47058  -5.264 1.41e-07 ***
## YEAR2001    -0.97251    0.47157  -2.062 0.039179 *
## YEAR2002    -0.73052    0.46615  -1.567 0.117088
## YEAR2003     0.16145    0.45153   0.358 0.720670
## YEAR2004    -1.01197    0.47186  -2.145 0.031981 *
## YEAR2005    -0.22935    0.47536  -0.482 0.629465
## YEAR2006    -0.48270    0.46329  -1.042 0.297461
## YEAR2007    -1.03749    0.47059  -2.205 0.027478 *
## YEAR2008    -1.25006    0.66314  -1.885 0.059422 .
## MONTH10     -0.61817    0.19043  -3.246 0.001170 **
## BTEMP        1.62346    0.11592  14.004  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -3.3225     0.6373  -5.213 1.86e-07 ***
## BTEMP        -2.5137     0.4181  -6.013 1.82e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

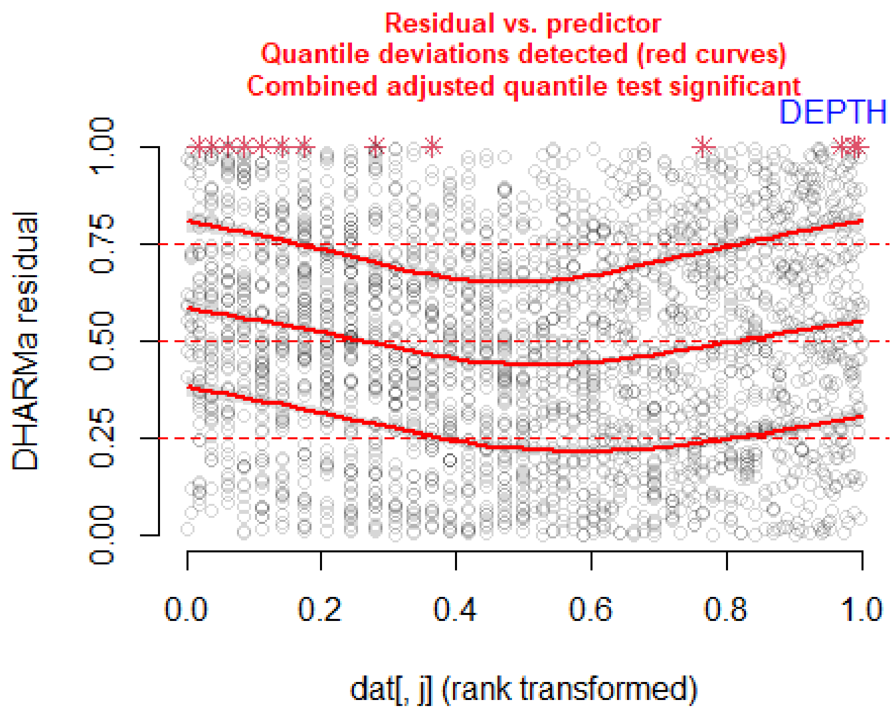
```r
drop1(ZINB.1c, test="Chi")

## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + BTEMP
##        Df   AIC     LRT  Pr(>Chi)
## <none>    11156
## YEAR   23 11304 193.510 < 2.2e-16 ***
## MONTH   1 11165  10.358  0.001289 **
## BTEMP   1 11323 168.906 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Remedial measures had worse diagnostics than original model, and original
model diagnostics were good enough.

models <- list(ZINB.1,ZINB.1b,ZINB.1c,NB2)# ZINB.1 diagnostics appear to be
'good enough'
for(i in 1:length(models)) {
   NB.sim <- simulateResiduals(models[[i]])
   this <- names(dat)[names(dat) %in% names(models[[i]]$frame) & names(dat)
!= "CPUE"]
   print(this)
   plot(NB.sim,quantreg=TRUE); mtext(paste(this,collapse=" +
"),side=3,adj=0,outer=FALSE,line=1,col="blue")

   for(j in this) {
      plotResiduals(NB.sim, form=dat[,j],quantreg=TRUE);
mtext(j,3,col="blue",adj=1)
   }
}

## [1] "YEAR"  "MONTH" "BTEMP" "DEPTH"

## DHARMa:testOutliers with type = binomial may have inflated Type I error
rates for integer-valued distributions. To get a more exact result, it is
recommended to re-run testOutliers with type = 'bootstrap'. See ?testOutliers
for details
```


DHARMa residual
YEAR + MONTH + BTEMP + DEPTH
QQ plot residuals
KS test: p=
Deviation
ispersion test:
eviation signifi
t: p= 0.01915
significant
Observed
Expected
0.0
0.4
0.8
Residual vs. predicted
Quantile deviations detected (red c
:ombined adjusted quantile test sig
DHARMa residual
0.00
0.50
1.00
Model predictions (rank transfor

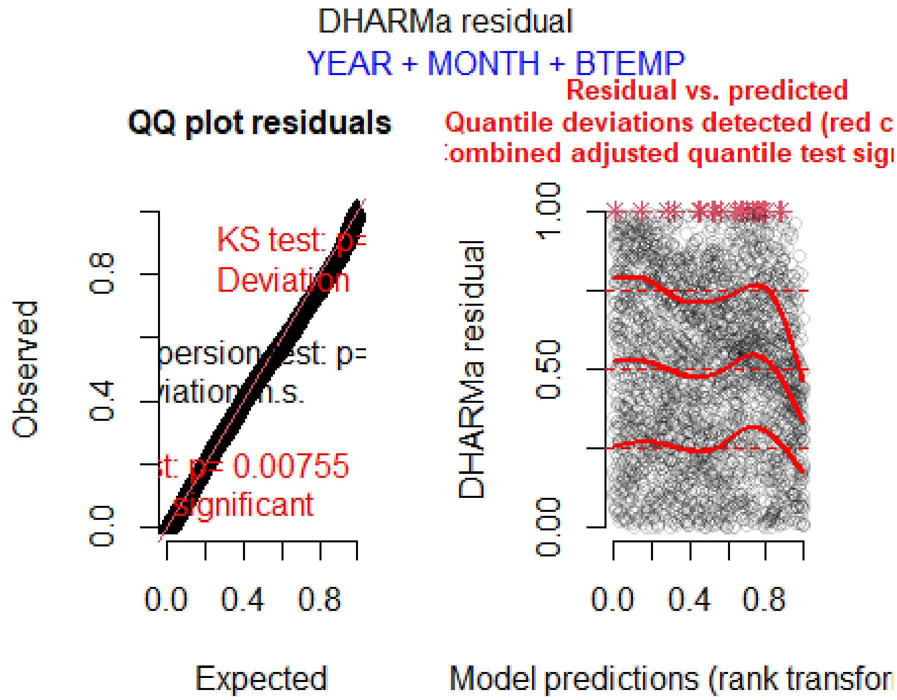

Within-group deviations from uniformity significant (red)
Levene Test for homogeneity of variance n.s.
YEAR
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1985
1989
1993
1997
2001
2005
catPred


Within-group deviations from uniformity significant (red)
Levene Test for homogeneity of variance significant
MONTH
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
9
10
catPred

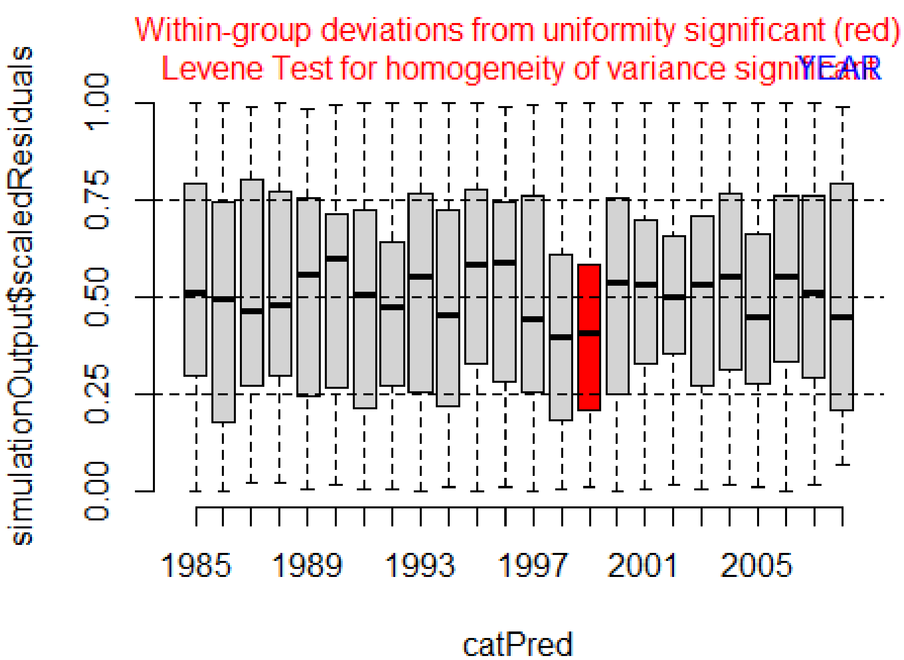

Residual vs. predictor
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
BTEMP
DHARMa residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
dat[, j] (rank transformed)



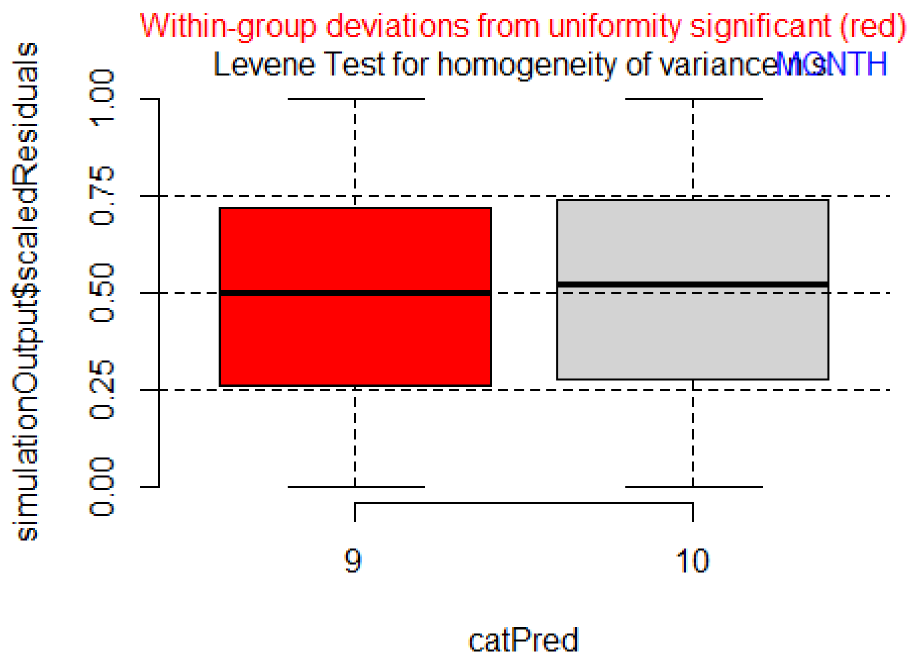

```
## [1] "YEAR"  "MONTH" "BTEMP"

## DHARMa:testOutliers with type = binomial may have inflated Type I error
rates for integer-valued distributions. To get a more exact result, it is
recommended to re-run testOutliers with type = 'bootstrap'. See ?testOutliers
for details
```

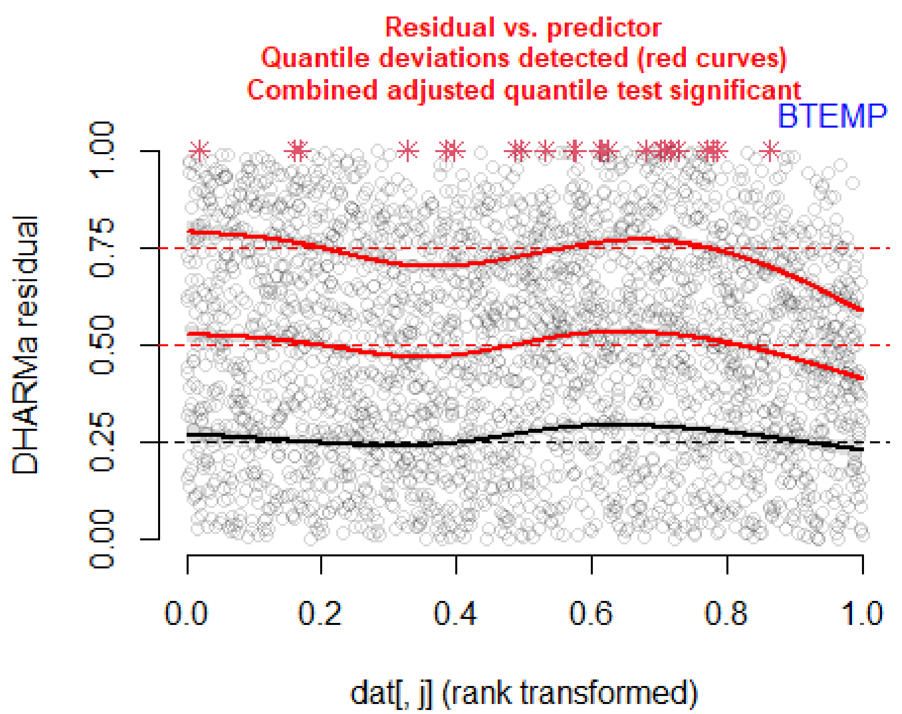


DHARMa residual
YEAR + MONTH + BTEMP
QQ plot residuals
KS test: p=
Deviation
persion test: p=
viation n.s.
t: p= 0.00755
significant
Observed
0.0
0.4
0.8
Expected
Residual vs. predicted
Quantile deviations detected (red c
ombined adjusted quantile test sig
DHARMa residual
0.00
0.50
1.00
Model predictions (rank transfor



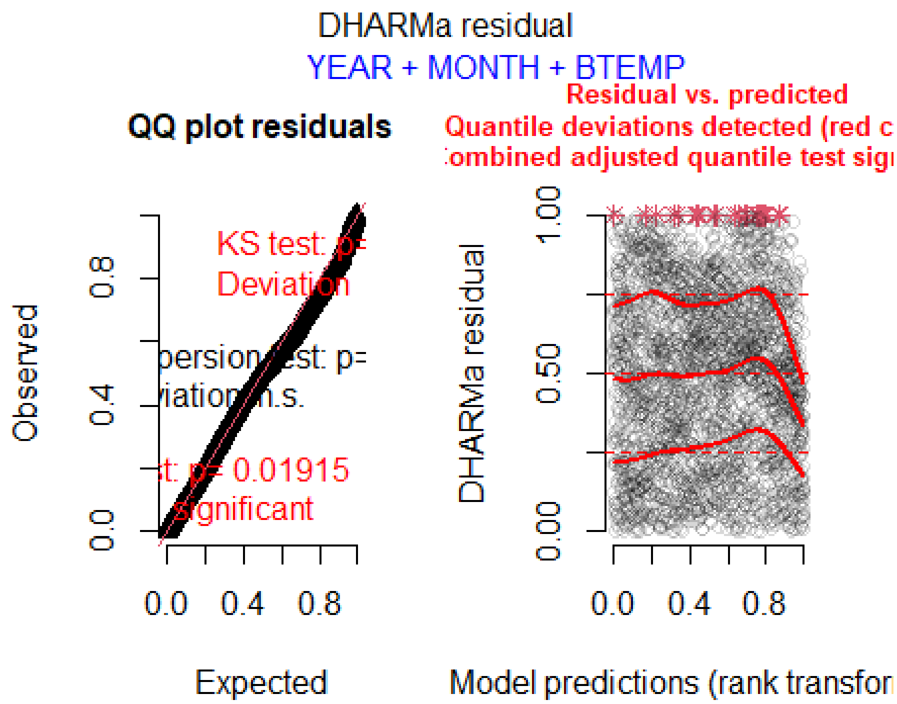
Within-group deviations from uniformity significant (red)
Levene Test for homogeneity of variance significant
YEAR
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1985
1989
1993
1997
2001
2005
catPred

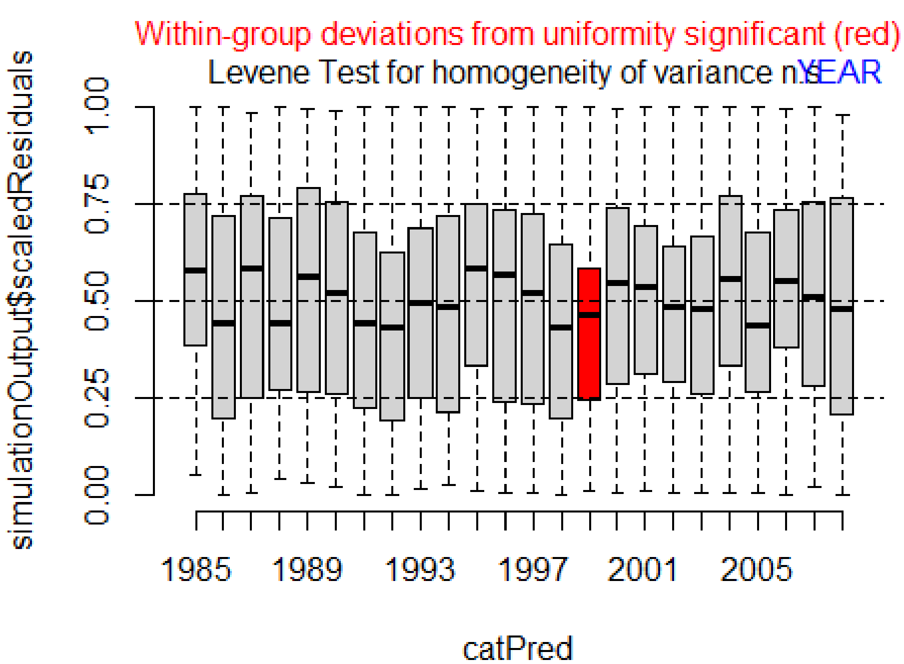


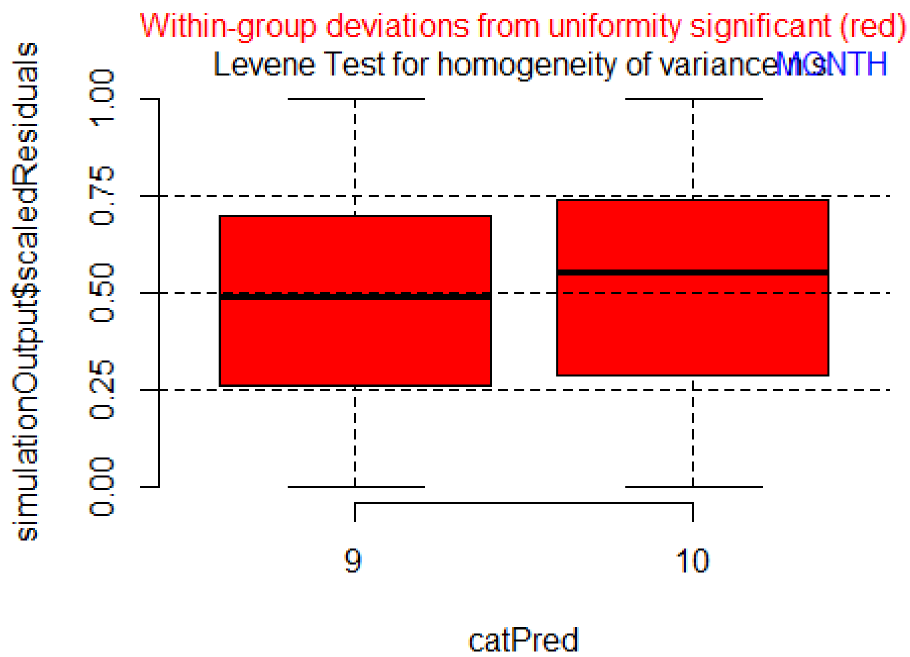


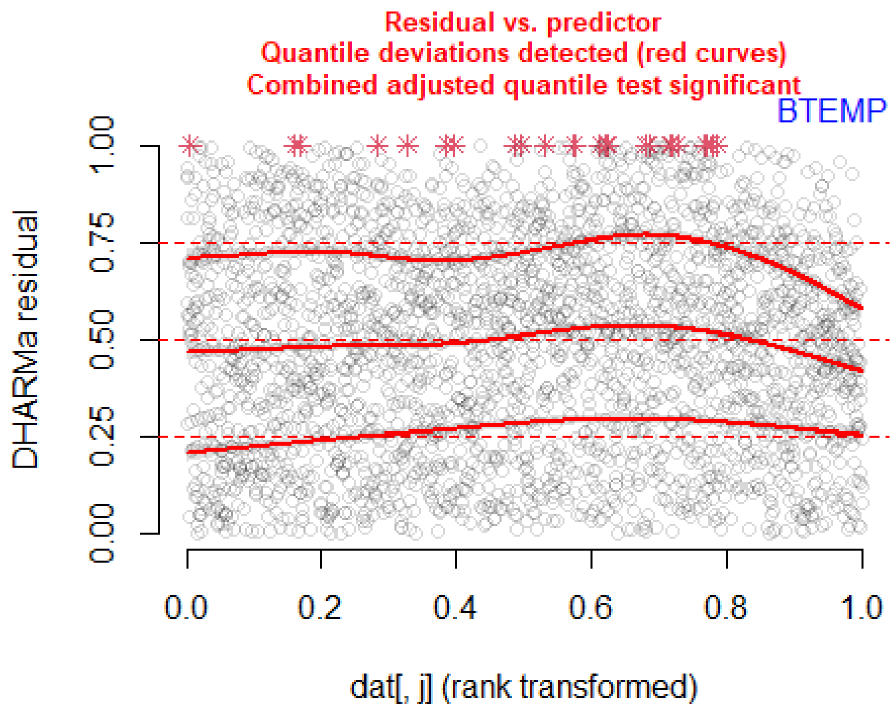

```
## [1] "YEAR"  "MONTH" "BTEMP"

## DHARMa:testOutliers with type = binomial may have inflated Type I error
rates for integer-valued distributions. To get a more exact result, it is
```

```
recommended to re-run testOutliers with type = 'bootstrap'. See ?testOutliers
for details
```



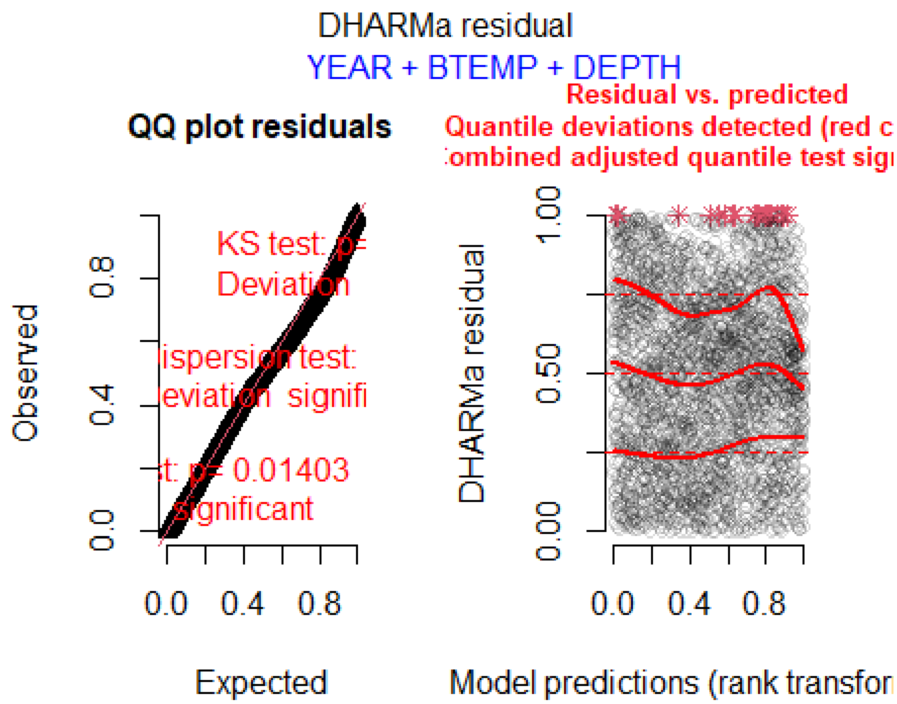


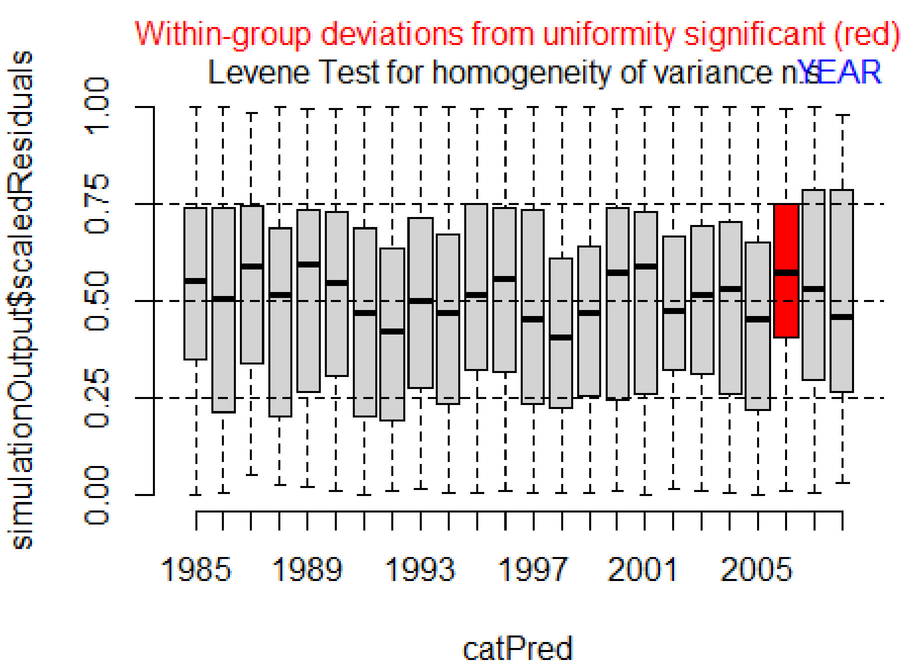


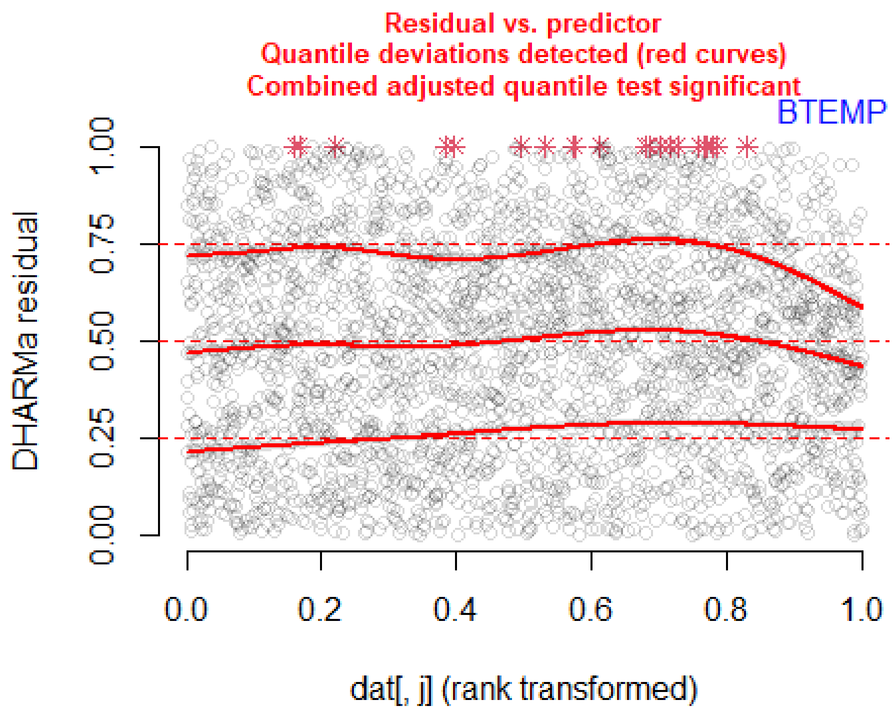


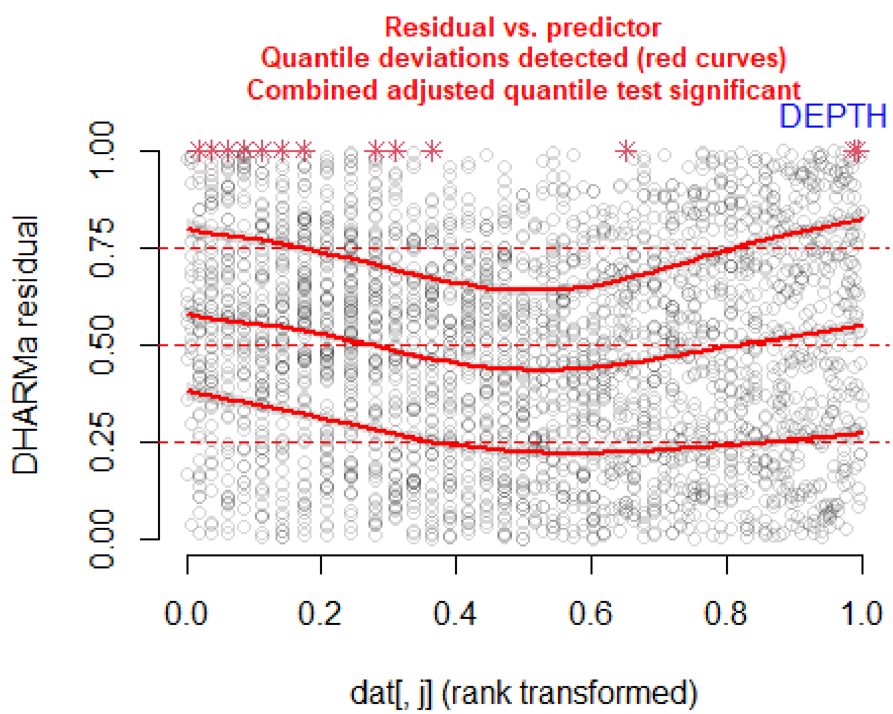

```
## [1] "YEAR"  "BTEMP" "DEPTH"

## DHARMa:testOutliers with type = binomial may have inflated Type I error
rates for integer-valued distributions. To get a more exact result, it is
```

```
recommended to re-run testOutliers with type = 'bootstrap'. See ?testOutliers
for details
```









Conclusion: ZINB.1 wins on AIC, and had diagnostics that were 'good enough.' Remedial measures did not improve diagnostics (in fact, appeared to make things worse). ZANB models had high SEs so not included in final model selection.

Generate and compare indices

```
# by-hand with bootstrapping :
# requires final dataset object called 'dat'
out <- byHand2(ZINB.1,nBoot=1000,Method="ZINB.1")

# calc stratified mean (using all data, not just the data used in dat, which
has some records removed due to missing env data):
# stratified mean:

a <-

read_xlsx(path="f:\\BF_WG\\FIG\\NEFSC\\NEFSC_Fall_Survey_Data_for_Std_V3.xlsx
",
   sheet="BLUEFISH_FALL_ALB_RT2022_85_08",range="A125:AJ3662",na="NA")
   #BLUEFISH_FALL_ALB_RT2022_85_08",range="A125:AJ3662
# for Albatross/DE, use this sheet and cell range
   #BLUEFISH_FALL_BIG_RT2022_09_21",range="A75:                          #
for Bigelow, use this sheet and cell range

head(a)

## # A tibble: 6 x 36
##   CRUISE6 STRATUM   TOW STATION   SHG SVVESSEL SVGEAR EST_YEAR EST_MONTH
EST_DAY
##     <dbl>   <dbl> <dbl>   <dbl> <dbl> <chr>     <dbl>    <dbl>     <dbl>
<dbl>
## 1  198508    1010     1     109   111 DE           11     1985        10
8
## 2  198508    1010     2     123   111 DE           11     1985        10
9
## 3  198508    1010     3     108   111 DE           11     1985        10
8
## 4  198508    1010     4     121   111 DE           11     1985        10
9
## 5  198508    1010     5     110   123 DE           11     1985        10
8
## 6  198508    1010     6     124   111 DE           11     1985        10
9
## # ... with 26 more variables: TIME <dbl>, TOWDUR <dbl>, DOPDISTB <dbl>,
## #   DOPDISTW <dbl>, AVGDEPTH <dbl>, STATYPE <dbl>, HAUL <dbl>, GEARCOND
<dbl>,
## #   TYPE_CODE <lgl>, OPERATION_CODE <lgl>, GEAR_CODE <lgl>,
## #   ACQUISITION_CODE <lgl>, AREA <dbl>, BOTTEMP <dbl>, BEGLAT <dbl>,
## #   BEGLON <dbl>, STRATUM_AREA <dbl>, STRGRP_DESC <chr>, STRATUM_NAME
<chr>,
## #   AREA_SWEPT_WINGS_MEAN_KM2 <lgl>, AREA_SWEPT_WINGS_MEAN_NM2 <dbl>,
## #   SWEPT_AREA_RATIO <dbl>, SVSPP <dbl>, CATCHSEX <dbl>, EXPCATCHWT <dbl>,
...

str(a)
```

```
## tibble [3,537 x 36] (S3: tbl_df/tbl/data.frame)
##  $ CRUISE6                  : num [1:3537] 198508 198508 198508 198508
198508 ...
##  $ STRATUM                  : num [1:3537] 1010 1010 1010 1010 1010 1010
1010 1050 1050 1050 ...
##  $ TOW                      : num [1:3537] 1 2 3 4 5 6 7 1 3 4 ...
##  $ STATION                  : num [1:3537] 109 123 108 121 110 124 122 147
171 149 ...
##  $ SHG                      : num [1:3537] 111 111 111 111 123 111 111 111
111 111 ...
##  $ SVVESSEL                 : chr [1:3537] "DE" "DE" "DE" "DE" ...
##  $ SVGEAR                   : num [1:3537] 11 11 11 11 11 11 11 11 11 11
...
##  $ EST_YEAR                 : num [1:3537] 1985 1985 1985 1985 1985 ...
##  $ EST_MONTH                : num [1:3537] 10 10 10 10 10 10 10 10 10 10
...
##  $ EST_DAY                  : num [1:3537] 8 9 8 9 8 9 9 15 18 16 ...
##  $ TIME                     : num [1:3537] 705 1342 431 1007 949 ...
##  $ TOWDUR                   : num [1:3537] 30 30 30 30 30 30 30 30 30 30
...
##  $ DOPDISTB                 : num [1:3537] 1.77 1.74 1.74 1.78 1.74 1.72
1.72 1.84 1.88 1.69 ...
##  $ DOPDISTW                 : num [1:3537] 0 0 0 0 0 0 0 0 0 0 ...
##  $ AVGDEPTH                 : num [1:3537] 37 45 37 55 52 40 44 40 52 46
...
##  $ STATYPE                  : num [1:3537] 1 1 1 1 1 1 1 1 1 1 ...
##  $ HAUL                     : num [1:3537] 1 1 1 1 2 1 1 1 1 1 ...
##  $ GEARCOND                 : num [1:3537] 1 1 1 1 3 1 1 1 1 1 ...
##  $ TYPE_CODE                : logi [1:3537] NA NA NA NA NA NA ...
##  $ OPERATION_CODE           : logi [1:3537] NA NA NA NA NA NA ...
##  $ GEAR_CODE                : logi [1:3537] NA NA NA NA NA NA ...
##  $ ACQUISITION_CODE         : logi [1:3537] NA NA NA NA NA NA ...
##  $ AREA                     : num [1:3537] 615 613 612 613 615 612 613 537
537 539 ...
##  $ BOTTEMP                  : num [1:3537] 14.5 NA 13.3 NA 11.5 16.5 NA NA
NA NA ...
##  $ BEGLAT                   : num [1:3537] 39.9 40.3 40.2 40.1 39.7 ...
##  $ BEGLON                   : num [1:3537] -73.5 -73 -73.6 -72.7 -73 ...
##  $ STRATUM_AREA             : num [1:3537] 2516 2516 2516 2516 2516 ...
##  $ STRGRP_DESC              : chr [1:3537] "OFFSHORE REGULAR" "OFFSHORE
REGULAR" "OFFSHORE REGULAR" "OFFSHORE REGULAR" ...
##  $ STRATUM_NAME             : chr [1:3537] "SNE HUDSON CANYN" "SNE HUDSON
CANYN" "SNE HUDSON CANYN" "SNE HUDSON CANYN" ...
##  $ AREA_SWEPT_WINGS_MEAN_KM2: logi [1:3537] NA NA NA NA NA NA ...
##  $ AREA_SWEPT_WINGS_MEAN_NM2: num [1:3537] 0.01 0.01 0.01 0.01 0.01 0.01
0.01 0.01 0.01 0.01 ...
##  $ SWEPT_AREA_RATIO         : num [1:3537] 1 1 1 1 1 1 1 1 1 1 ...
##  $ SVSPP                    : num [1:3537] 135 NA NA NA 135 135 NA NA NA
135 ...
##  $ CATCHSEX                 : num [1:3537] 0 0 0 0 0 0 0 0 0 0 ...
```

```
##  $ EXPCATCHWT              : num [1:3537] 8.5 0 0 0 8.5 6.4 0 0 0 9.5 ...
##  $ EXPCATCHNUM             : num [1:3537] 8 0 0 0 2 7 0 0 0 3 ...
```

```
summary(a)
```

```
##     CRUISE6         STRATUM          TOW           STATION         SHG
##  Min.   :198508   Min.   :1010   Min.   : 1.00   Min.   :  1   Min.
:111.0
##  1st Qu.:199105   1st Qu.:1190   1st Qu.: 1.00   1st Qu.: 69   1st
Qu.:111.0
##  Median :199706   Median :3020   Median : 2.00   Median :119   Median
:111.0
##  Mean   :199670   Mean   :2275   Mean   : 2.65   Mean   :120   Mean
:112.7
##  3rd Qu.:200306   3rd Qu.:3250   3rd Qu.: 4.00   3rd Qu.:162   3rd
Qu.:111.0
##  Max.   :200807   Max.   :3460   Max.   :11.00   Max.   :375   Max.
:136.0
##
##    SVVESSEL             SVGEAR      EST_YEAR      EST_MONTH
##  Length:3537        Min.   :11   Min.   :1985   Min.   : 9.000
##  Class :character   1st Qu.:11   1st Qu.:1991   1st Qu.: 9.000
##  Mode  :character   Median :11   Median :1997   Median : 9.000
##                     Mean   :11   Mean   :1997   Mean   : 9.271
##                     3rd Qu.:11   3rd Qu.:2003   3rd Qu.:10.000
##                     Max.   :11   Max.   :2008   Max.   :11.000
##
##     EST_DAY          TIME          TOWDUR         DOPDISTB
##  Min.   : 1.00   Min.   :   0   Min.   :19.83   Min.   :0.170
##  1st Qu.:10.00   1st Qu.: 607   1st Qu.:30.00   1st Qu.:1.740
##  Median :17.00   Median :1214   Median :30.00   Median :1.819
##  Mean   :16.38   Mean   :1195   Mean   :29.91   Mean   :1.821
##  3rd Qu.:23.00   3rd Qu.:1817   3rd Qu.:30.00   3rd Qu.:1.910
##  Max.   :30.00   Max.   :2359   Max.   :31.43   Max.   :2.161
##                                                 NA's   :6
##     DOPDISTW         AVGDEPTH         STATYPE       HAUL          GEARCOND
##  Min.   :0.0000   Min.   :  5.00   Min.   :1   Min.   :1.000   Min.
:1.000
##  1st Qu.:0.0000   1st Qu.: 18.00   1st Qu.:1   1st Qu.:1.000   1st
Qu.:1.000
##  Median :0.0000   Median : 27.00   Median :1   Median :1.000   Median
:1.000
##  Mean   :0.0049   Mean   : 33.43   Mean   :1   Mean   :1.143   Mean
:1.232
##  3rd Qu.:0.0000   3rd Qu.: 45.00   3rd Qu.:1   3rd Qu.:1.000   3rd
Qu.:1.000
##  Max.   :1.9500   Max.   :114.00   Max.   :1   Max.   :3.000   Max.
:6.000
##  NA's   :2384     NA's   :1
##  TYPE_CODE      OPERATION_CODE GEAR_CODE      ACQUISITION_CODE      AREA
```

```
##  Mode:logical   Mode:logical   Mode:logical   Mode:logical     Min.
:521.0
##  TRUE:113       TRUE:113       TRUE:113       TRUE:113         1st
Qu.:539.0
##  NA's:3424      NA's:3424      NA's:3424      NA's:3424        Median
:613.0
##                                                                Mean
:594.3
##                                                                3rd
Qu.:622.0
##                                                                Max.
:636.0
##                                                                NA's   :1
##     BOTTEMP          BEGLAT          BEGLON        STRATUM_AREA
##  Min.   : 5.88   Min.   :35.20   Min.   :-76.07   Min.   :  13
##  1st Qu.:13.26   1st Qu.:38.08   1st Qu.:-74.98   1st Qu.: 154
##  Median :16.94   Median :40.13   Median :-73.67   Median : 383
##  Mean   :16.91   Mean   :39.41   Mean   :-72.84   Mean   :1130
##  3rd Qu.:20.79   3rd Qu.:40.81   3rd Qu.:-71.12   3rd Qu.:2454
##  Max.   :28.12   Max.   :42.08   Max.   :-67.02   Max.   :2832
##  NA's   :717
##  STRGRP_DESC        STRATUM_NAME       AREA_SWEPT_WINGS_MEAN_KM2
##  Length:3537        Length:3537        Mode:logical
##  Class :character   Class :character   NA's:3537
##  Mode  :character   Mode  :character
##
##
##
##
##  AREA_SWEPT_WINGS_MEAN_NM2 SWEPT_AREA_RATIO     SVSPP         CATCHSEX
##  Min.   :0.01              Min.   :1        Min.   :135    Min.   :0
##  1st Qu.:0.01              1st Qu.:1        1st Qu.:135    1st Qu.:0
##  Median :0.01              Median :1        Median :135    Median :0
##  Mean   :0.01              Mean   :1        Mean   :135    Mean   :0
##  3rd Qu.:0.01              3rd Qu.:1        3rd Qu.:135    3rd Qu.:0
##  Max.   :0.01              Max.   :1        Max.   :135    Max.   :0
##                                             NA's   :2054
##    EXPCATCHWT       EXPCATCHNUM
##  Min.   :  0.000   Min.   :   0.00
##  1st Qu.:  0.000   1st Qu.:   0.00
##  Median :  0.000   Median :   0.00
##  Mean   :  2.162   Mean   :  12.43
##  3rd Qu.:  1.980   3rd Qu.:   3.00
##  Max.   :150.700   Max.   :2743.00
##
```

```
# Standardize catch for 30 minute tow (the standard unit of effort of NEFSC)
a$EXPCATCHNUM <- round((a$EXPCATCHNUM*30)/a$TOWDUR,0)#a$EXPCATCHNUM
                                              #^^^^^^^^^^^^^^^ NEFSC uses un-
adjusted catches; the FIG recommended adjusting to a standard tow (which for
```

```
NEFSC is 30 min for Albatross, and 20 min for Bigelow).

# strata wts:
lookup <- with(a, tapply(STRATUM_AREA,STRATUM,median))

lookup <- data.frame("STRATUM_AREA"=lookup, "STRATUM"=rownames(lookup))
lookup$wts <- lookup$STRATUM_AREA/sum(lookup$STRATUM_AREA)

lookup <- lookup[order(lookup$STRATUM),] # should already be in increasing
order, but just to be certain

   # mean:
   M <- with(a, tapply(EXPCATCHNUM ,list(STRATUM,EST_YEAR),mean,na.rm=TRUE))
   # variance:
   V <- with(a, tapply(EXPCATCHNUM ,list(STRATUM,EST_YEAR),var,na.rm=TRUE))

   # matrix of stratum wts:
   wts <- ((!is.na(M))*1) * lookup$STRATUM_AREA
   wts <- t(t(wts)/colSums(wts))


   # Possible workaround if a year-stratum combination has NA variance (i.e.,
from n= 1), pool mean variance from entire year:
   needVar <- which(is.na(V))
   if(length(needVar)>0) {
      allYears <- sort(unique(a$EST_YEAR))
      # years of interest:
      yoi <- rep(allYears,each=nrow(lookup))[needVar]
      lookupTable <-
data.frame(year=allYears,.var=apply(V,2,mean,na.rm=TRUE))
         # The special case of all NAs in a given year:
         # If NaN, then pool timeseries mean variance
         if(any(is.nan(lookupTable$.var))) {
            mod <- which(is.nan(lookupTable$.var))
            lookupTable$.var[mod] <- mean(lookupTable$.var[-mod])
         }
      ans <-
merge(lookupTable,data.frame(year=yoi,tmp=NA),by="year",sort=FALSE)
      V[which(is.na(V))] <- ans$.var
   }

      # sample size (per stratum, per year)
      N <- with(a, tapply(EXPCATCHNUM,list(STRATUM,EST_YEAR),length))


   # UCL
   UCL <-
   colSums(M*wts,na.rm=TRUE)+
      ((sqrt(colSums(wts^2*V/N,na.rm=TRUE)))*qnorm((1-0.95)/2,lower=FALSE))
```

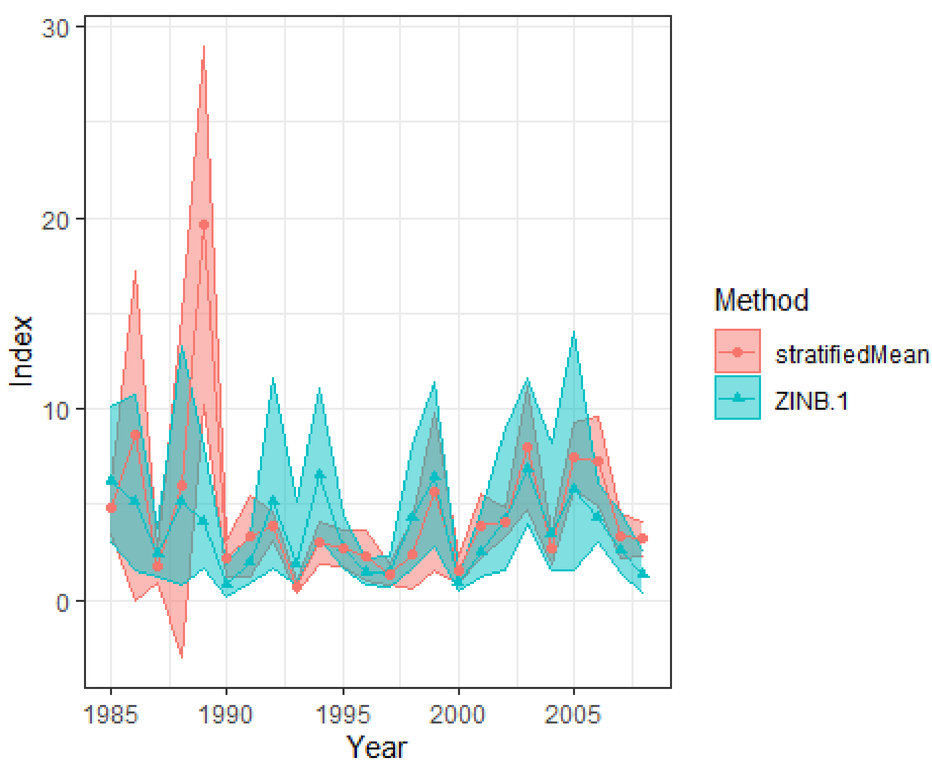
```r
  # stratified mean:
  .m <- colSums(M*wts,na.rm=TRUE)

  # LCL
  LCL <-
  colSums(M*wts,na.rm=TRUE)-
     ((sqrt(colSums(wts^2*V/N,na.rm=TRUE)))*qnorm((1-0.95)/2,lower=FALSE))

  SE <- sqrt(colSums(wts^2*V/N,na.rm=TRUE))
  CV <- SE/.m

  nom <- data.frame("Year"=sort(unique(a$EST_YEAR)),"Index"=.m,SE,
     "LCI"=LCL,"UCI"=UCL,CV,"Method"="stratifiedMean")



tmp <- rbind(out,nom)
tmp
```

```
##         Year     Index        SE         LCI        UCI         CV
## 1985    1985 6.2009111 1.8725439  3.06103628 10.1611436 0.30197883
## 1986    1986 5.1397514 2.3398590  1.50324232 10.7221593 0.45524750
## 1987    1987 2.3994885 0.6578926  1.19581750  3.7438215 0.27418035
## 1988    1988 5.1156348 3.5091546  0.79142440 13.2600091 0.68596660
## 1989    1989 4.0727317 1.6516403  1.62854230  8.0046221 0.40553624
## 1990    1990 0.8229412 0.4290959  0.19805713  1.8331874 0.52141753
## 1991    1991 1.9626337 0.6417207  0.94812494  3.4924431 0.32696916
## 1992    1992 5.1813224 2.7892252  1.65939957 11.6195400 0.53832304
## 1993    1993 1.8994948 1.1551090  0.81237489  5.1273585 0.60811380
## 1994    1994 6.5079561 2.0614094  3.24581557 11.1175823 0.31675219
## 1995    1995 2.7028231 0.7058253  1.69064640  4.4634550 0.26114372
## 1996    1996 1.4154741 0.3385814  0.83985256  2.1737017 0.23920000
## 1997    1997 1.3984234 0.3743343  0.74796286  2.2640207 0.26768306
## 1998    1998 4.3336438 1.7517253  1.61388825  8.0934642 0.40421533
## 1999    1999 6.3995323 2.1701844  2.79851559 11.4161424 0.33911610
## 2000    2000 0.8656798 0.2175075  0.53266621  1.3597141 0.25125631
## 2001    2001 2.5344504 0.8812726  1.18279133  4.6256752 0.34771744
## 2002    2002 4.1790484 2.0119837  1.49940724  8.9871220 0.48144542
## 2003    2003 6.8439582 1.8531078  4.01434095 11.6257463 0.27076551
## 2004    2004 3.4130921 1.7236579  1.58073929  8.1778067 0.50501359
## 2005    2005 5.7688163 3.7213440  1.50941410 14.0563863 0.64507930
## 2006    2006 4.2571694 0.7796913  3.08145819  6.1402737 0.18314783
## 2007    2007 2.6260223 0.8188926  1.41933090  4.5813538 0.31183765
## 2008    2008 1.2878886 0.5662311  0.41742068  2.6525500 0.43965848
## 19851 1985 4.7821878 0.6961162  3.41782503  6.1465505 0.14556439
## 19861 1986 8.6295843 4.3990646  0.00757608 17.2515925 0.50976553
## 19871 1987 1.7458585 0.4404302  0.88263116  2.6090859 0.25227143
## 19881 1988 5.9497658 4.5578782 -2.98351133 14.8830430 0.76606010
```

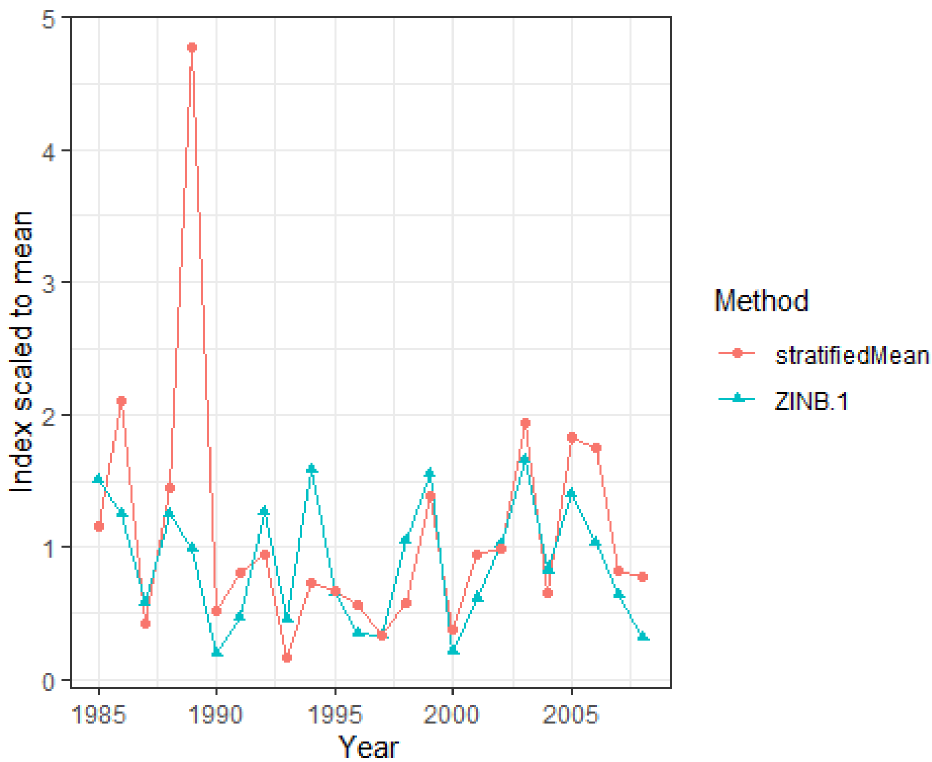
```
## 19891 1989 19.6048153 4.7663990 10.26284502 28.9467856 0.24312389
## 19901 1990  2.1553593 0.4862226  1.20238046  3.1083382 0.22558774
## 19911 1991  3.3422838 1.0930593  1.19992687  5.4846406 0.32703965
## 19921 1992  3.9103749 0.3797467  3.16608517  4.6546647 0.09711260
## 19931 1993  0.6976376 0.1474589  0.40862353  0.9866516 0.21136885
## 19941 1994  2.9775617 0.5437194  1.91189132  4.0432321 0.18260558
## 19951 1995  2.7249156 0.4655808  1.81239392  3.6374372 0.17086064
## 19961 1996  2.3390837 0.6505137  1.06410031  3.6140671 0.27810620
## 19971 1997  1.3649270 0.2629899  0.84947614  1.8803778 0.19267694
## 19981 1998  2.4033439 0.9122355  0.61539524  4.1912925 0.37956926
## 19991 1999  5.6793004 2.1118497  1.54015110  9.8184498 0.37185032
## 20001 2000  1.5629944 0.3445815  0.88762697  2.2383618 0.22046243
## 20011 2001  3.9091548 0.8586567  2.22621852  5.5920910 0.21965278
## 20021 2002  4.0599740 0.3771263  3.32081991  4.7991280 0.09288886
## 20031 2003  7.9567391 1.6489344  4.72488712 11.1885912 0.20723746
## 20041 2004  2.7166603 0.4255574  1.88258317  3.5507375 0.15664726
## 20051 2005  7.4961017 0.8880085  5.75563709  9.2365664 0.11846270
## 20061 2006  7.2299985 1.1974288  4.88308121  9.5769158 0.16561951
## 20071 2007  3.3637664 0.5827516  2.22159424  4.5059387 0.17324378
## 20081 2008  3.2137171 0.4701164  2.29230600  4.1351283 0.14628430
##               Method
## 1985          ZINB.1
## 1986          ZINB.1
## 1987          ZINB.1
## 1988          ZINB.1
## 1989          ZINB.1
## 1990          ZINB.1
## 1991          ZINB.1
## 1992          ZINB.1
## 1993          ZINB.1
## 1994          ZINB.1
## 1995          ZINB.1
## 1996          ZINB.1
## 1997          ZINB.1
## 1998          ZINB.1
## 1999          ZINB.1
## 2000          ZINB.1
## 2001          ZINB.1
## 2002          ZINB.1
## 2003          ZINB.1
## 2004          ZINB.1
## 2005          ZINB.1
## 2006          ZINB.1
## 2007          ZINB.1
## 2008          ZINB.1
## 19851 stratifiedMean
## 19861 stratifiedMean
## 19871 stratifiedMean
## 19881 stratifiedMean
## 19891 stratifiedMean
```

```
## 19901 stratifiedMean
## 19911 stratifiedMean
## 19921 stratifiedMean
## 19931 stratifiedMean
## 19941 stratifiedMean
## 19951 stratifiedMean
## 19961 stratifiedMean
## 19971 stratifiedMean
## 19981 stratifiedMean
## 19991 stratifiedMean
## 20001 stratifiedMean
## 20011 stratifiedMean
## 20021 stratifiedMean
## 20031 stratifiedMean
## 20041 stratifiedMean
## 20051 stratifiedMean
## 20061 stratifiedMean
## 20071 stratifiedMean
## 20081 stratifiedMean
```

```
ggplot(tmp, aes(x=as.numeric(Year), y=Index, color=Method, shape=Method)) +
  geom_ribbon(aes(x=as.numeric(Year), ymin=LCI, ymax=UCI, fill=Method),
alpha=0.5) +
  geom_line() + geom_point() +
  xlab("Year") + theme_bw()
```



```r
# scale indices to respective means:
ggplot(tmp, aes(x=as.numeric(Year), y=scaleToMean(Index), color=Method,
shape=Method)) +
  geom_line() + geom_point() +
  xlab("Year") + theme_bw()+ylab("Index scaled to mean")
```




```r
round(cor(sapply(split(tmp,tmp$Method),function(x) x$Index)),2)

##                stratifiedMean ZINB.1
## stratifiedMean           1.00   0.44
## ZINB.1                   0.44   1.00

round(cor(sapply(split(tmp,tmp$Method),function(x)
x$Index),method="spearman"),2)

##                stratifiedMean ZINB.1
## stratifiedMean            1.0    0.7
## ZINB.1                    0.7    1.0

Sys.time()

## [1] "2022-06-10 12:06:50 EDT"
```

Summary:

- Note that index standardization was done with continuity run strata, but the baserun will use some additional offshore strata, so I re-examined all model selsection (on first cut, I did cursory exploration, but revisited and did more thorough examination).

- I can’t use the summary values in the spreadsheet Tony provided to calc CIs (the CIs from summary data are too narrow; though the summary values do appear to provide realistic CVs), so I used the raw data in the spreadsheet to cal CIs (and SEs and CVs). I standardized all catch data to a 30 min tow (which the summary data does not reflect) since it is consistent with the FIG protocols. The CVs from Tony’s summary spreadsheet information and from scratch/by-hand are very similar.

Bluefish index standardization - NEFSC trawl (Bigelow) with new strata

M Celestino

2022-06-10 10:51:59

Bluefish index standardization

Step 1: Data processing

Load functions & libraries

```
source("f:\\BF_WG\\FIG\\figFuns.R")
source("f:\\BF_WG\\FIG\\originalCode\\KDrew2\\bootstrap_functions_AT2.R")

library(car, quietly = TRUE, verbose=FALSE)

## Warning: package 'car' was built under R version 4.1.2

## Warning: package 'carData' was built under R version 4.1.2

library(lmtest, quietly = TRUE, verbose=FALSE)

## Warning: package 'lmtest' was built under R version 4.1.2

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

library(glmmTMB, quietly = TRUE, verbose=FALSE)
library(DHARMa, quietly = TRUE, verbose=FALSE)

## Warning: package 'DHARMa' was built under R version 4.1.2

## This is DHARMa 0.4.5. For overview type '?DHARMa'. For recent changes,
type news(package = 'DHARMa')

library(bbmle, quietly = TRUE, verbose=FALSE)
library(ggeffects, quietly = TRUE, verbose=FALSE)
library(ggplot2, quietly = TRUE, verbose=FALSE)

## Warning: package 'ggplot2' was built under R version 4.1.2

library(readxl, quietly = TRUE, verbose=FALSE)
library(lattice, quietly = TRUE, verbose=FALSE)
```

```
## Warning: package 'lattice' was built under R version 4.1.2

library(emmeans, quietly = TRUE, verbose=FALSE)

## Warning: package 'emmeans' was built under R version 4.1.2

library(rmarkdown, quietly = TRUE, verbose=FALSE)

## Warning: package 'rmarkdown' was built under R version 4.1.2

library(knitr, quietly = TRUE, verbose=FALSE)

## Warning: package 'knitr' was built under R version 4.1.2

library(reshape2, quietly=TRUE, verbose=FALSE)

## Warning: package 'reshape2' was built under R version 4.1.2

library(doBy, quietly=TRUE, verbose=FALSE)

## Warning: package 'doBy' was built under R version 4.1.2

library(dplyr, quietly=TRUE, verbose=FALSE,warn.conflicts=FALSE)
```

Read in data

```
indata <-

read_xlsx(path="f:\\BF_WG\\FIG\\NEFSC\\NEFSC_Fall_Survey_Data_for_Std_V3.xlsx
",
   sheet="BLUEFISH_FALL_BIG_RT2022_09_21",range="A75:AJ1513",na="NA")
head(indata)

## # A tibble: 6 x 36
##   CRUISE6 STRATUM   TOW STATION   SHG SVVESSEL SVGEAR EST_YEAR EST_MONTH
EST_DAY
##     <dbl>   <dbl> <dbl>   <dbl> <dbl> <chr>     <dbl>    <dbl>     <dbl>
<dbl>
## 1  200904    1010     1     142   111 HB           10     2009        10
1
## 2  200904    1010     2     126   111 HB           10     2009         9
29
## 3  200904    1010     3     132   122 HB           10     2009         9
30
## 4  200904    1010     4     130   111 HB           10     2009         9
30
## 5  200904    1010     5     125   111 HB           10     2009         9
29
## 6  200904    1010     6     131   111 HB           10     2009         9
30
## # ... with 26 more variables: TIME <dbl>, TOWDUR <dbl>, DOPDISTB <dbl>,
## #   DOPDISTW <lgl>, AVGDEPTH <dbl>, STATYPE <dbl>, HAUL <dbl>, GEARCOND
<dbl>,
```

```
## #   TYPE_CODE <dbl>, OPERATION_CODE <dbl>, GEAR_CODE <dbl>,
## #   ACQUISITION_CODE <dbl>, AREA <dbl>, BOTTEMP <dbl>, BEGLAT <dbl>,
## #   BEGLON <dbl>, STRATUM_AREA <dbl>, STRGRP_DESC <chr>, STRATUM_NAME
<chr>,
## #   AREA_SWEPT_WINGS_MEAN_KM2 <dbl>, AREA_SWEPT_WINGS_MEAN_NM2 <dbl>,
## #   SWEPT_AREA_RATIO <dbl>, SVSPP <dbl>, CATCHSEX <dbl>, EXPCATCHWT <dbl>,
...
```

```
str(indata)
```

```
## tibble [1,438 x 36] (S3: tbl_df/tbl/data.frame)
##  $ CRUISE6                  : num [1:1438] 200904 200904 200904 200904
200904 ...
##  $ STRATUM                  : num [1:1438] 1010 1010 1010 1010 1010 1010
1010 1010 1050 1050 ...
##  $ TOW                      : num [1:1438] 1 2 3 4 5 6 7 8 1 2 ...
##  $ STATION                  : num [1:1438] 142 126 132 130 125 131 141 124
119 120 ...
##  $ SHG                      : num [1:1438] 111 111 122 111 111 111 111 111
111 111 ...
##  $ SVVESSEL                 : chr [1:1438] "HB" "HB" "HB" "HB" ...
##  $ SVGEAR                   : num [1:1438] 10 10 10 10 10 10 10 10 10 10
...
##  $ EST_YEAR                 : num [1:1438] 2009 2009 2009 2009 2009 ...
##  $ EST_MONTH                : num [1:1438] 10 9 9 9 9 9 9 9 9 9 ...
##  $ EST_DAY                  : num [1:1438] 1 29 30 30 29 30 30 29 29 29
...
##  $ TIME                     : num [1:1438] 228 1853 445 138 1649 ...
##  $ TOWDUR                   : num [1:1438] 20 20 20 20 20 ...
##  $ DOPDISTB                 : num [1:1438] 0.958 0.99 1.024 1.009 0.955
...
##  $ DOPDISTW                 : logi [1:1438] NA NA NA NA NA NA ...
##  $ AVGDEPTH                 : num [1:1438] 54 35 45 44 37 42 46 55 42 53
...
##  $ STATYPE                  : num [1:1438] 1 1 1 1 1 1 1 1 1 1 ...
##  $ HAUL                     : num [1:1438] 1 1 2 1 1 1 1 1 1 1 ...
##  $ GEARCOND                 : num [1:1438] 1 1 2 1 1 1 1 1 1 1 ...
##  $ TYPE_CODE                : num [1:1438] 1 1 1 1 1 1 1 1 1 1 ...
##  $ OPERATION_CODE           : num [1:1438] 1 1 1 3 3 1 1 1 1 1 ...
##  $ GEAR_CODE                : num [1:1438] 1 1 2 1 1 1 1 1 1 1 ...
##  $ ACQUISITION_CODE         : num [1:1438] 1 1 1 1 2 1 2 2 2 1 ...
##  $ AREA                     : num [1:1438] 616 613 612 613 613 613 615 613
613 613 ...
##  $ BOTTEMP                  : num [1:1438] 13.6 19.3 18.9 NA 13.7 ...
##  $ BEGLAT                   : num [1:1438] 39.9 40.6 40.2 40.4 40.7 ...
##  $ BEGLON                   : num [1:1438] -72.8 -72.7 -73.1 -72.9 -72.4
...
##  $ STRATUM_AREA             : num [1:1438] 2516 2516 2516 2516 2516 ...
##  $ STRGRP_DESC              : chr [1:1438] "OFFSHORE REGULAR" "OFFSHORE
REGULAR" "OFFSHORE REGULAR" "OFFSHORE REGULAR" ...
```

```
##  $ STRATUM_NAME            : chr [1:1438] "SNE HUDSON CANYN" "SNE HUDSON
CANYN" "SNE HUDSON CANYN" "SNE HUDSON CANYN" ...
##  $ AREA_SWEPT_WINGS_MEAN_KM2: num [1:1438] 0.0227 0.0216 0.0228 0.0228
0.0202 ...
##  $ AREA_SWEPT_WINGS_MEAN_NM2: num [1:1438] 0.00662 0.00631 0.00664 0.00665
0.00588 ...
##  $ SWEPT_AREA_RATIO        : num [1:1438] 1.06 1.11 1.05 1.05 1.19 ...
##  $ SVSPP                   : num [1:1438] NA 135 135 135 NA NA NA NA NA
NA ...
##  $ CATCHSEX                : num [1:1438] 0 0 0 0 0 0 0 0 0 0 ...
##  $ EXPCATCHWT              : num [1:1438] 0 2.04 4.68 5.82 0 ...
##  $ EXPCATCHNUM             : num [1:1438] 0 1.11 1.05 2.11 0 ...
```

```
summary(indata)
```

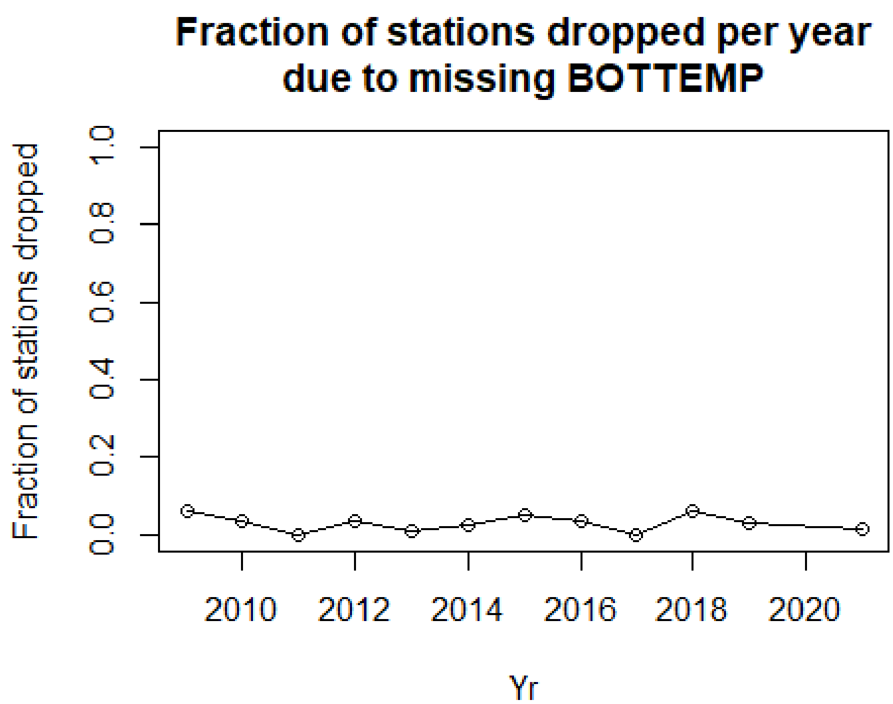
```
##     CRUISE6          STRATUM          TOW             STATION
##  Min.   :200904   Min.   :1010   Min.   : 1.000   Min.   :  7.0
##  1st Qu.:201105   1st Qu.:1100   1st Qu.: 2.000   1st Qu.: 93.0
##  Median :201404   Median :1690   Median : 3.000   Median :136.0
##  Mean   :201437   Mean   :2113   Mean   : 3.423   Mean   :146.7
##  3rd Qu.:201804   3rd Qu.:3222   3rd Qu.: 5.000   3rd Qu.:196.0
##  Max.   :202104   Max.   :3460   Max.   :12.000   Max.   :412.0
##
##       SHG          SVVESSEL             SVGEAR      EST_YEAR
##  Min.   :111.0   Length:1438        Min.   :10   Min.   :2009
##  1st Qu.:111.0   Class :character   1st Qu.:10   1st Qu.:2011
##  Median :111.0   Mode  :character   Median :10   Median :2014
##  Mean   :112.7                      Mean   :10   Mean   :2014
##  3rd Qu.:111.0                      3rd Qu.:10   3rd Qu.:2018
##  Max.   :222.0                      Max.   :10   Max.   :2021
##                                     NA's   :11
##    EST_MONTH         EST_DAY            TIME            TOWDUR
##  Min.   : 9.000   Min.   : 1.00   Min.   :   1.0   Min.   :16.56
##  1st Qu.: 9.000   1st Qu.:12.00   1st Qu.: 632.2   1st Qu.:20.00
##  Median : 9.000   Median :17.00   Median :1239.0   Median :20.02
##  Mean   : 9.295   Mean   :17.16   Mean   :1221.7   Mean   :20.01
##  3rd Qu.:10.000   3rd Qu.:24.00   3rd Qu.:1833.0   3rd Qu.:20.03
##  Max.   :11.000   Max.   :31.00   Max.   :2359.0   Max.   :35.13
##
##     DOPDISTB     DOPDISTW         AVGDEPTH         STATYPE
##  Min.   :0.259   Mode:logical   Min.   : 16.00   Min.   :1.000
##  1st Qu.:0.979   NA's:1438      1st Qu.: 26.00   1st Qu.:1.000
##  Median :0.995                  Median : 34.00   Median :1.000
##  Mean   :1.000                  Mean   : 39.92   Mean   :1.006
##  3rd Qu.:1.014                  3rd Qu.: 50.00   3rd Qu.:1.000
##  Max.   :1.539                  Max.   :110.00   Max.   :2.000
##                                 NA's   :1
##       HAUL          GEARCOND       TYPE_CODE OPERATION_CODE   GEAR_CODE
##  Min.   :1.000   Min.   :1.000   Min.   :1   Min.   :1.00   Min.   :1.000
##  1st Qu.:1.000   1st Qu.:1.000   1st Qu.:1   1st Qu.:1.00   1st Qu.:1.000
```

```
##  Median :1.000   Median :1.000   Median :1    Median :1.00   Median :1.000  
##  Mean   :1.099   Mean   :1.172   Mean   :1    Mean   :1.64   Mean   :1.056  
##  3rd Qu.:1.000   3rd Qu.:1.000   3rd Qu.:1    3rd Qu.:2.00   3rd Qu.:1.000  
##  Max.   :4.000   Max.   :9.000   Max.   :1    Max.   :3.00   Max.   :2.000  
##                                                                             
##  ACQUISITION_CODE      AREA          BOTTEMP          BEGLAT     
##  Min.   :1.000    Min.   :521.0   Min.   : 8.37   Min.   :35.15  
##  1st Qu.:1.000    1st Qu.:537.0   1st Qu.:13.71   1st Qu.:38.21  
##  Median :1.000    Median :613.0   Median :16.74   Median :40.25  
##  Mean   :1.389    Mean   :589.6   Mean   :16.95   Mean   :39.47  
##  3rd Qu.:1.000    3rd Qu.:621.0   3rd Qu.:20.08   3rd Qu.:40.88  
##  Max.   :4.000    Max.   :636.0   Max.   :29.85   Max.   :42.08  
##                   NA's   :2       NA's   :47                     
##      BEGLON        STRATUM_AREA  STRGRP_DESC        STRATUM_NAME      
##  Min.   :-75.72   Min.   :  62   Length:1438        Length:1438       
##  1st Qu.:-74.83   1st Qu.: 238   Class :character   Class :character  
##  Median :-73.33   Median :1318   Mode  :character   Mode  :character  
##  Mean   :-72.55   Mean   :1343                                        
##  3rd Qu.:-70.57   3rd Qu.:2454                                        
##  Max.   :-67.04   Max.   :2832                                        
##                                                                       
##  AREA_SWEPT_WINGS_MEAN_KM2 AREA_SWEPT_WINGS_MEAN_NM2 SWEPT_AREA_RATIO
##  Min.   :0.00499           Min.   :0.001455          Min.   :0.6582  
##  1st Qu.:0.01999           1st Qu.:0.005866          1st Qu.:1.0363  
##  Median :0.02137           Median :0.006289          Median :1.1144  
##  Mean   :0.02143           Mean   :0.006290          Mean   :1.1296  
##  3rd Qu.:0.02287           3rd Qu.:0.006757          3rd Qu.:1.1967  
##  Max.   :0.03648           Max.   :0.010636          Max.   :4.8115  
##  NA's   :81                                                          
##      SVSPP        CATCHSEX   EXPCATCHWT        EXPCATCHNUM     
##  Min.   :135   Min.   :0   Min.   :  0.0000   Min.   :  0.000  
##  1st Qu.:135   1st Qu.:0   1st Qu.:  0.0000   1st Qu.:  0.000  
##  Median :135   Median :0   Median :  0.0000   Median :  0.000  
##  Mean   :135   Mean   :0   Mean   :  1.3224   Mean   :  3.054  
##  3rd Qu.:135   3rd Qu.:0   3rd Qu.:  0.5225   3rd Qu.:  1.195  
##  Max.   :135   Max.   :0   Max.   :197.5953   Max.   :394.217  
##  NA's   :979
```

Some data exploration to determine which years and covariates to keep; then remove missing values

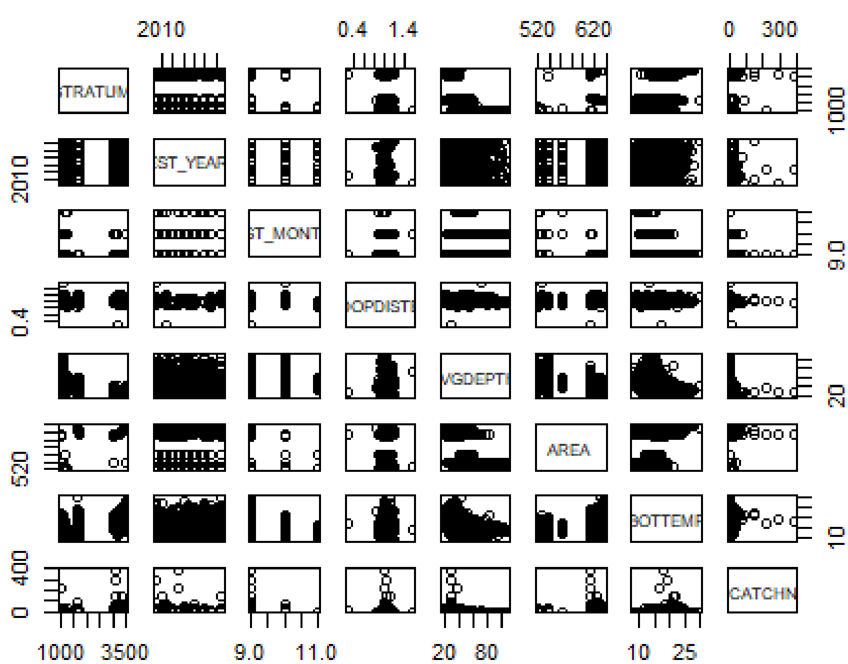
```
keepVars <-
c("CRUISE6","STRATUM","STATION","SVVESSEL","EST_YEAR","EST_MONTH","TOWDUR",
   "DOPDISTB","AVGDEPTH","AREA","BOTTEMP","EXPCATCHNUM")
#   "HAUL","GEARCOND","ACQUISITION_CODE",

# how many records would we lose to NAs:
apply(indata[,keepVars], 2, function(x) sum(is.na(x)))
```

```
##      CRUISE6      STRATUM      STATION     SVVESSEL     EST_YEAR    EST_MONTH
##            0            0            0            0            0            0
##       TOWDUR     DOPDISTB     AVGDEPTH         AREA      BOTTEMP  EXPCATCHNUM
##            0            0            1            2           47            0

# how many records will we lose per year due to NA BOTTEMP
plot(unique(indata$EST_YEAR),
apply(
do.call("rbind",with(indata[,c("EST_YEAR","BOTTEMP")],
   tapply(BOTTEMP,EST_YEAR,function(x) c(sum(is.na(x)),length(x) )))),1,
   function(x) {x[1]/x[2]}),
   xlab="Yr",ylab="Fraction of stations dropped",
   ylim=c(0,1),type="o",
   main="Fraction of stations dropped per year\ndue to missing BOTTEMP")
```

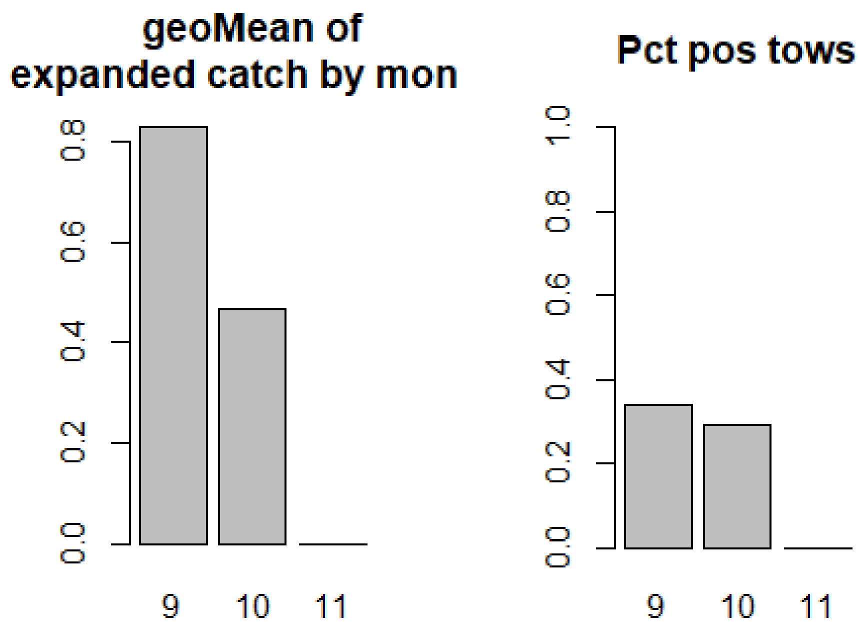



```
indata <- indata[,keepVars]
table(indata$SVVESSEL,indata$EST_YEAR)

##
##       2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2021
##    HB  133  136  125  132  134  129  130  132    0  116  130  130
##    PC    0    0    0    0    0    0    0    0   11    0    0    0

indata2 <- indata[complete.cases(indata),]

summary(indata2)
```

```
##      CRUISE6           STRATUM         STATION         SVVESSEL
##  Min.   :200904   Min.   :1010   Min.   :  7.0   Length:1388
##  1st Qu.:201105   1st Qu.:1100   1st Qu.: 92.0   Class :character
##  Median :201404   Median :1690   Median :136.0   Mode  :character
##  Mean   :201438   Mean   :2121   Mean   :146.6
##  3rd Qu.:201804   3rd Qu.:3230   3rd Qu.:195.0
##  Max.   :202104   Max.   :3460   Max.   :412.0
##     EST_YEAR      EST_MONTH          TOWDUR         DOPDISTB
##  Min.   :2009   Min.   : 9.000   Min.   :16.56   Min.   :0.259
##  1st Qu.:2011   1st Qu.: 9.000   1st Qu.:20.00   1st Qu.:0.979
##  Median :2014   Median : 9.000   Median :20.02   Median :0.995
##  Mean   :2014   Mean   : 9.295   Mean   :20.01   Mean   :1.000
##  3rd Qu.:2018   3rd Qu.:10.000   3rd Qu.:20.03   3rd Qu.:1.014
##  Max.   :2021   Max.   :11.000   Max.   :35.13   Max.   :1.539
##     AVGDEPTH           AREA          BOTTEMP       EXPCATCHNUM
##  Min.   : 16.00   Min.   :521   Min.   : 8.37   Min.   :  0.000
##  1st Qu.: 26.00   1st Qu.:537   1st Qu.:13.72   1st Qu.:  0.000
##  Median : 34.00   Median :613   Median :16.73   Median :  0.000
##  Mean   : 39.72   Mean   :590   Mean   :16.95   Mean   :  3.126
##  3rd Qu.: 50.00   3rd Qu.:621   3rd Qu.:20.08   3rd Qu.:  1.201
##  Max.   :110.00   Max.   :636   Max.   :29.85   Max.   :394.217

# The data I received were already subset to specific strata and months, so
not too much more to do...

pairs(indata2[,-c(1,3,4,7)])
```



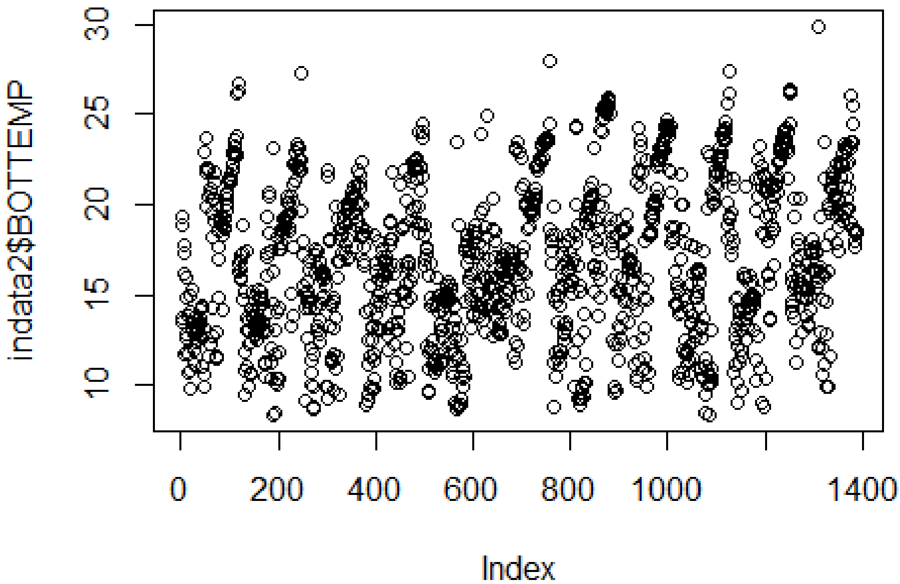
```
op <- par(mfrow=c(1,2))
barplot(tapply(indata2$EXPCATCHNUM, indata2$EST_MONTH,
geoMean,alpha=NULL,warn=FALSE),main="geoMean of\nexpanded catch by month")
barplot(tapply(indata2$EXPCATCHNUM, indata2$EST_MONTH, pctPos),main="Pct pos
tows",ylim=c(0,1))
```




```
par(op)
```

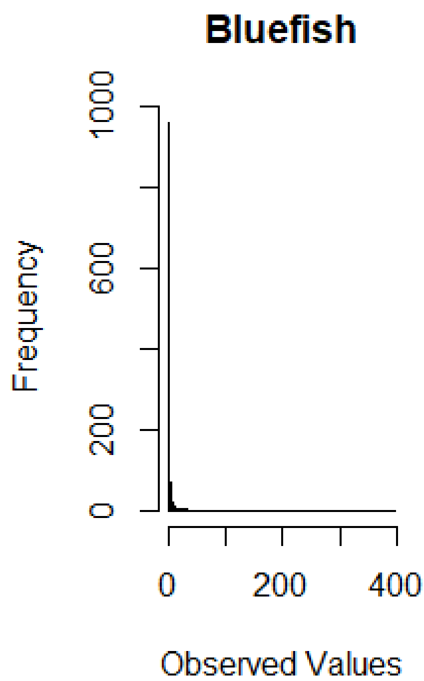
Check for outliers

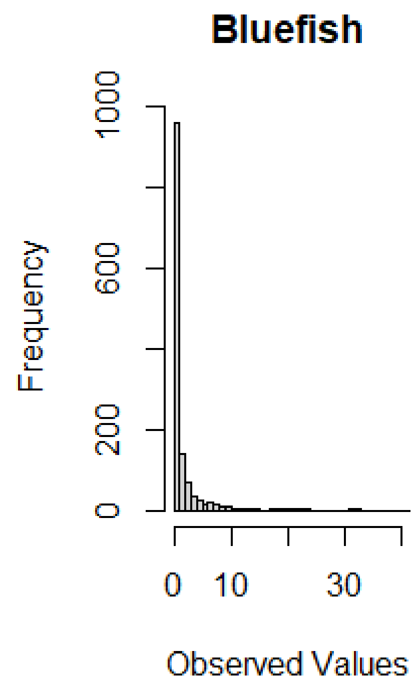
```
# Check for outliers or weirdo values
plot(indata2$BOTTEMP)
```



Plot distribution of catches:

```
op <- par(mfrow=c(1,2))
hist(indata2$EXPCATCHNUM,main="Bluefish",xlab="Observed Values",
   ylab="Frequency", breaks=seq(0,max(indata2$EXPCATCHNUM)+1,1))
#sort(indata2$EXPCATCHNUM,decreasing=TRUE)[1:50]
hist(indata2$EXPCATCHNUM,main="Bluefish",xlab="Observed Values",
   ylab="Frequency", breaks=seq(0,max(indata2$EXPCATCHNUM)+1,1),xlim=c(0,40))
```






```
par(op)
```

Finally, check to see if there are any years with zero catch that need to be removed:

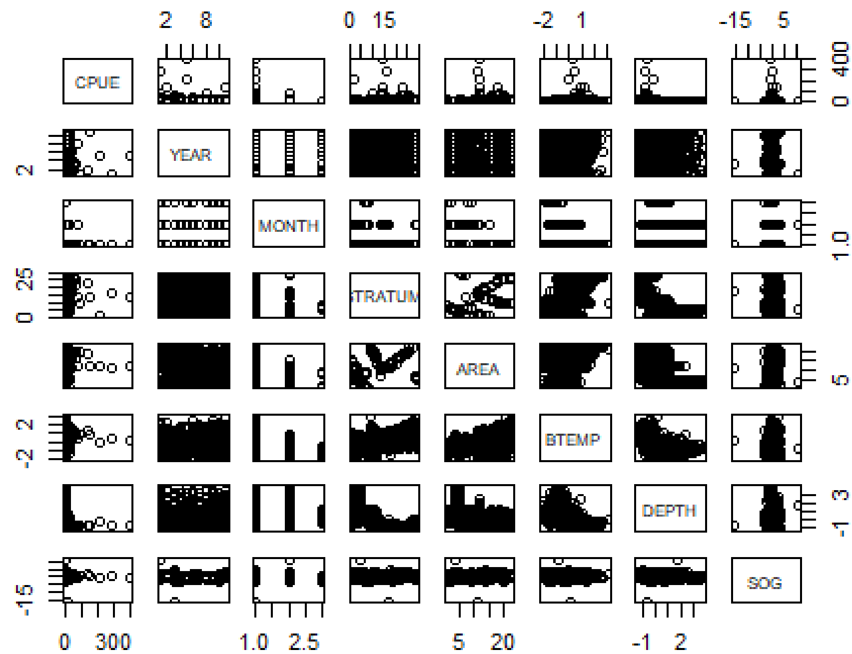
```
diff(as.numeric(names(xtabs(EXPCATCHNUM~EST_YEAR, indata2))))

##  [1] 1 1 1 1 1 1 1 1 1 1 2

table(indata2$EST_YEAR) # 2017 omitted due to not sampling southern strata;
2020 omitted due to covid

##
## 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2021
##  125  131  123  126  133  126  123  127   11  109  126  128

indata2 <- indata2[indata2$EST_YEAR!=2017,]
diff(as.numeric(names(xtabs(EXPCATCHNUM~EST_YEAR, indata2))))

##  [1] 1 1 1 1 1 1 1 2 1 2

table(indata2$EST_YEAR)

##
## 2009 2010 2011 2012 2013 2014 2015 2016 2018 2019 2021
##  125  131  123  126  133  126  123  127  109  126  128
```

Create a dataframe with all the factors we need (z-score all continuous variables; all others as.factors)

```
dat = data.frame(CPUE = round(indata2$EXPCATCHNUM),# needed to round catch to
0 decimal places (need integer only values for discrete distribution models)
                 YEAR = as.factor(indata2$EST_YEAR),
                 MONTH = as.factor(indata2$EST_MONTH),
                 STRATUM = as.factor(indata2$STRATUM),
                 AREA=as.factor(indata2$AREA),
                 VESSEL=as.factor(indata2$SVVESSEL),
                 BTEMP = Z.scr(indata2$BOTTEMP),
                 DEPTH = Z.scr(indata2$AVGDEPTH),
                 SOG=Z.scr(indata2$DOPDISTB),
                 EFFORT = indata2$TOWDUR)

# Calculate the proportion of positive tows/sets/hauls
dat$PosTow <- ifelse(dat$CPUE > 0, 1, 0)
mean(dat$PosTow) # 56%

## [1] 0.325345

dat$lnEffort <- log(dat$EFFORT)

# was lnEffort
head(dat)

##   CPUE YEAR MONTH STRATUM AREA VESSEL       BTEMP      DEPTH        SOG
EFFORT
## 1    0 2009    10    1010  616     HB -0.81491852  0.7738047 -0.8781448
20.02
## 2    1 2009     9    1010  613     HB  0.54675478 -0.2517805 -0.2121109
20.02
## 3    1 2009     9    1010  612     HB  0.46567441  0.2880012  0.4955501
20.03
## 4    0 2009     9    1010  613     HB -0.78153249 -0.1438242 -0.9405855
20.03
## 5    0 2009     9    1010  613     HB  0.19620141  0.1260667 -0.1080431
20.02
## 6    0 2009     9    1010  615     HB  0.08650443  0.3419794  0.9534484
20.05
##   PosTow lnEffort
## 1      0 2.996732
## 2      1 2.996732
## 3      1 2.997231
## 4      0 2.997231
## 5      0 2.996732
## 6      0 2.998229
```

Check collinearity

```
pairs(~CPUE+YEAR+MONTH+STRATUM+AREA+BTEMP+DEPTH+SOG,data=dat)
```




```
round(cor(with(dat,
data.frame(CPUE,unfactor(YEAR),unfactor(MONTH),unfactor(STRATUM),unfactor(ARE
A),BTEMP,DEPTH,SOG))),2)

##                     CPUE  unfactor.YEAR.  unfactor.MONTH.  unfactor.STRATUM.
## CPUE                1.00           -0.09            -0.07               0.13
## unfactor.YEAR.     -0.09            1.00            -0.10               0.01
## unfactor.MONTH.    -0.07           -0.10             1.00              -0.36
## unfactor.STRATUM.   0.13            0.01            -0.36               1.00
## unfactor.AREA.      0.09           -0.01            -0.68               0.48
## BTEMP               0.12            0.13            -0.26               0.64
## DEPTH              -0.13           -0.01             0.37              -0.72
## SOG                -0.01            0.26            -0.04               0.03
##                    unfactor.AREA.  BTEMP  DEPTH    SOG
## CPUE                         0.09   0.12  -0.13  -0.01
## unfactor.YEAR.              -0.01   0.13  -0.01   0.26
## unfactor.MONTH.             -0.68  -0.26   0.37  -0.04
## unfactor.STRATUM.            0.48   0.64  -0.72   0.03
## unfactor.AREA.               1.00   0.40  -0.59   0.05
## BTEMP                        0.40   1.00  -0.61   0.10
## DEPTH                       -0.59  -0.61   1.00  -0.04
## SOG                          0.05   0.10  -0.04   1.00

# You want GVIF to be less than ~3.
# Check the variance inflation factor for a more statistical check.
mod = lm(CPUE ~ YEAR + MONTH + STRATUM + BTEMP + DEPTH + SOG, data = dat)
vif(mod) # won't run with STRATUM and AREA, so removing AREA
```

```
##              GVIF Df GVIF^(1/(2*Df))
## YEAR     2.060997 10        1.036821
## MONTH    3.794906  2        1.395726
## STRATUM 32.495909 28        1.064136
## BTEMP    3.101363  1        1.761069
## DEPTH    6.976262  1        2.641261
## SOG      1.507148  1        1.227660
```

```
# remove STRATUM
mod = lm(CPUE ~ YEAR + MONTH + BTEMP + DEPTH + SOG, data = dat)
vif(mod)
```

```
##           GVIF Df GVIF^(1/(2*Df))
## YEAR  1.722761 10        1.027570
## MONTH 1.292601  2        1.066267
## BTEMP 1.813212  1        1.346556
## DEPTH 1.834388  1        1.354396
## SOG   1.445360  1        1.202231
```

```
# need to remove month because November is sampled too infrequently and
caused model convergence problems:
with(dat, table(YEAR,MONTH))
```

```
##       MONTH
## YEAR     9  10  11
##   2009  79  42   4
##   2010  75  53   3
##   2011  85  37   1
##   2012 101  25   0
##   2013 107  21   5
##   2014  84  41   1
##   2015 108  15   0
##   2016  78  48   1
##   2018  97  11   1
##   2019  96  28   2
##   2021  95  31   2
```

```
round(cor(with(dat, data.frame(CPUE,unfactor(YEAR),BTEMP,DEPTH,SOG))),2)
```

```
##                  CPUE unfactor.YEAR. BTEMP DEPTH   SOG
## CPUE             1.00          -0.09  0.12 -0.13 -0.01
## unfactor.YEAR. -0.09           1.00  0.13 -0.01  0.26
## BTEMP            0.12           0.13  1.00 -0.61  0.10
## DEPTH           -0.13          -0.01 -0.61  1.00 -0.04
## SOG             -0.01           0.26  0.10 -0.04  1.00
```

Step 2: Model selection

Compare negative binom, zero-infl neg binom, zero-altered neg binom

Model 1: Negative Binomial

```
tmb1.NB <- glmmTMB(CPUE ~ YEAR + BTEMP + DEPTH + SOG+offset(lnEffort), data =
dat, family = nbinom2)

# Check the Std. Error of the estimates; high SEs indicate problems with the
fit.
summary(tmb1.NB)

##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + BTEMP + DEPTH + SOG + offset(lnEffort)
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   3997.9   4076.3  -1984.0   3967.9     1362
##
##
## Dispersion parameter for nbinom2 family (): 0.188
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.1068     0.2318  -9.090  < 2e-16 ***
## YEAR2010      0.1318     0.3187   0.414 0.679128
## YEAR2011      0.1018     0.3210   0.317 0.751137
## YEAR2012     -0.2790     0.3481  -0.802 0.422821
## YEAR2013      0.3251     0.3205   1.014 0.310484
## YEAR2014     -1.1010     0.3288  -3.348 0.000814 ***
## YEAR2015     -1.5839     0.3438  -4.607 4.09e-06 ***
## YEAR2016     -1.2866     0.3317  -3.879 0.000105 ***
## YEAR2018     -1.3055     0.3541  -3.686 0.000227 ***
## YEAR2019     -1.5275     0.3426  -4.459 8.23e-06 ***
## YEAR2021     -1.4003     0.3332  -4.203 2.63e-05 ***
## BTEMP         0.8355     0.1202   6.951 3.62e-12 ***
## DEPTH        -0.9354     0.1212  -7.720 1.16e-14 ***
## SOG          -0.3019     0.1138  -2.653 0.007979 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Calculate dispersion for a glmmTMB model
disp3(tmb1.NB)

## df.resid
## 2.622119
```

Model 2: ZINB

```
ZINB = glmmTMB(CPUE ~ YEAR + BTEMP + DEPTH + SOG,offset=lnEffort, ziformula =
~ YEAR + BTEMP + DEPTH + SOG, data = dat, family=nbinom2)
summary(ZINB)

##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + BTEMP + DEPTH + SOG
## Zero inflation:        ~YEAR + BTEMP + DEPTH + SOG
```

```
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##   3885.0   4036.6  -1913.5   3827.0     1348
##
##
## Dispersion parameter for nbinom2 family (): 0.273
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.69544    0.23398  -7.246 4.29e-13 ***
## YEAR2010    -0.09394    0.29595  -0.317 0.750921
## YEAR2011    -0.01343    0.30512  -0.044 0.964887
## YEAR2012    -0.39645    0.33303  -1.190 0.233874
## YEAR2013     0.97389    0.39398   2.472 0.013439 *
## YEAR2014    -1.22220    0.29543  -4.137 3.52e-05 ***
## YEAR2015    -1.27377    0.35340  -3.604 0.000313 ***
## YEAR2016    -1.09777    0.33298  -3.297 0.000978 ***
## YEAR2018    -0.69512    0.38314  -1.814 0.069637 .
## YEAR2019    -1.75831    0.32122  -5.474 4.40e-08 ***
## YEAR2021    -0.26391    0.45987  -0.574 0.566054
## BTEMP        0.07728    0.12904   0.599 0.549228
## DEPTH       -1.12830    0.12952  -8.711  < 2e-16 ***
## SOG         -0.06363    0.10298  -0.618 0.536678
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.5274     0.7717  -3.275  0.00106 **
## YEAR2010      -2.3563     1.0841  -2.174  0.02974 *
## YEAR2011      -0.9613     0.9189  -1.046  0.29546
## YEAR2012      -0.8605     0.9521  -0.904  0.36610
## YEAR2013       1.6096     0.7382   2.180  0.02923 *
## YEAR2014     -17.9758  2053.1213  -0.009  0.99301
## YEAR2015       1.9120     0.9415   2.031  0.04228 *
## YEAR2016       1.6654     1.1520   1.446  0.14825
## YEAR2018       2.7195     1.1287   2.409  0.01598 *
## YEAR2019      -1.0115     1.3676  -0.740  0.45953
## YEAR2021       5.3215     1.0565   5.037 4.73e-07 ***
## BTEMP         -3.5075     0.6352  -5.522 3.35e-08 ***
## DEPTH         -0.7421     0.3915  -1.896  0.05802 .
## SOG            0.1063     0.2661   0.399  0.68955
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 3: ZANB (aka hurdle model)

```
ZANB <- glmmTMB(CPUE ~ YEAR + BTEMP + DEPTH + SOG, ziformula = ~ YEAR + BTEMP
+ DEPTH + SOG, data = dat,offset=lnEffort,family = truncated_nbinom2(link =
```

```
"log"))
summary(ZANB)

##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + BTEMP + DEPTH + SOG
## Zero inflation:        ~YEAR + BTEMP + DEPTH + SOG
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##   3864.8   4016.4  -1903.4   3806.8     1348
##
##
## Dispersion parameter for truncated_nbinom2 family (): 7.86e-09
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -18.56154 2468.20396  -0.008  0.99400
## YEAR2010      -0.14196    0.42851  -0.331  0.74043
## YEAR2011      -0.27201    0.42195  -0.645  0.51915
## YEAR2012      -0.52758    0.46912  -1.125  0.26075
## YEAR2013       1.26960    0.58829   2.158  0.03092 *
## YEAR2014      -1.22225    0.44856  -2.725  0.00643 **
## YEAR2015      -1.47511    0.49981  -2.951  0.00316 **
## YEAR2016      -0.95031    0.49471  -1.921  0.05474 .
## YEAR2018      -0.96408    0.51191  -1.883  0.05966 .
## YEAR2019      -1.38237    0.51962  -2.660  0.00781 **
## YEAR2021       0.16968    0.72727   0.233  0.81552
## BTEMP          0.26643    0.18968   1.405  0.16013
## DEPTH         -0.93931    0.20086  -4.676 2.92e-06 ***
## SOG           -0.05305    0.14452  -0.367  0.71357
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.48590    0.21157   2.297 0.021639 *
## YEAR2010    -0.36901    0.29332  -1.258 0.208379
## YEAR2011    -0.74764    0.29856  -2.504 0.012275 *
## YEAR2012    -0.23127    0.30506  -0.758 0.448380
## YEAR2013     0.87452    0.32617   2.681 0.007337 **
## YEAR2014     0.49916    0.29599   1.686 0.091716 .
## YEAR2015     1.02825    0.31134   3.303 0.000958 ***
## YEAR2016     0.95363    0.30729   3.103 0.001913 **
## YEAR2018     0.68084    0.31994   2.128 0.033335 *
## YEAR2019     1.33217    0.31753   4.195 2.72e-05 ***
## YEAR2021     2.26624    0.38507   5.885 3.98e-09 ***
## BTEMP       -0.61397    0.09779  -6.278 3.42e-10 ***
## DEPTH        0.76170    0.12288   6.199 5.69e-10 ***
## SOG          0.02970    0.08496   0.350 0.726674
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Preliminary model comparisons

```
AICtab(tmb1.NB, ZINB, ZANB)

##         dAIC  df
## ZANB      0.0 29
## ZINB     20.2 29
## tmb1.NB 133.1 15

# For now, ZANB wins, but model selection below fine tunes this
```

Step 3: Factor selection (by model)

Negative Binomial

```
NB1 <- glmmTMB(CPUE ~ YEAR + BTEMP + DEPTH + SOG+offset(lnEffort), data =
dat, family = nbinom2)
summary(NB1)

##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + BTEMP + DEPTH + SOG + offset(lnEffort)
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   3997.9   4076.3  -1984.0   3967.9     1362
##
##
## Dispersion parameter for nbinom2 family (): 0.188
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.1068     0.2318  -9.090  < 2e-16 ***
## YEAR2010      0.1318     0.3187   0.414 0.679128
## YEAR2011      0.1018     0.3210   0.317 0.751137
## YEAR2012     -0.2790     0.3481  -0.802 0.422821
## YEAR2013      0.3251     0.3205   1.014 0.310484
## YEAR2014     -1.1010     0.3288  -3.348 0.000814 ***
## YEAR2015     -1.5839     0.3438  -4.607 4.09e-06 ***
## YEAR2016     -1.2866     0.3317  -3.879 0.000105 ***
## YEAR2018     -1.3055     0.3541  -3.686 0.000227 ***
## YEAR2019     -1.5275     0.3426  -4.459 8.23e-06 ***
## YEAR2021     -1.4003     0.3332  -4.203 2.63e-05 ***
## BTEMP         0.8355     0.1202   6.951 3.62e-12 ***
## DEPTH        -0.9354     0.1212  -7.720 1.16e-14 ***
## SOG          -0.3019     0.1138  -2.653 0.007979 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

drop1(NB1,test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + BTEMP + DEPTH + SOG + offset(lnEffort)
##        Df    AIC    LRT  Pr(>Chi)
## <none>    3997.9
## YEAR   10 4077.6 99.639 < 2.2e-16 ***
## BTEMP   1 4044.6 48.733 2.933e-12 ***
## DEPTH   1 4057.9 62.021 3.398e-15 ***
## SOG     1 4003.5  7.598  0.005844 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# all significant, so keep all covars


# Let's compare to a model with year only.
NB0 <- glmmTMB(CPUE ~ YEAR+offset(lnEffort), data = dat, family = nbinom2)

# Let's compare the models using AIC to see if adding more factors
# improves how much deviance is explained.

AICtab(NB1, NB0)
```

```
##     dAIC  df
## NB1   0.0 15
## NB0 259.4 12
```

```
lrtest(NB0, NB1)
```

```
## Likelihood ratio test
##
## Model 1: CPUE ~ YEAR + offset(lnEffort)
## Model 2: CPUE ~ YEAR + BTEMP + DEPTH + SOG + offset(lnEffort)
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1  12 -2116.7
## 2  15 -1984.0  3 265.39  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sapply(list(NB1,NB0),disp3)
```

```
## df.resid df.resid
## 2.622119 2.360246
```

ZINB

```
ZINB1 <- glmmTMB(CPUE ~ YEAR + BTEMP + DEPTH + SOG,offset=lnEffort, ziformula
= ~ YEAR + BTEMP + DEPTH + SOG, data = dat, family=nbinom2)
summary(ZINB1)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + BTEMP + DEPTH + SOG
## Zero inflation:        ~YEAR + BTEMP + DEPTH + SOG
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##   3885.0   4036.6  -1913.5   3827.0     1348
##
##
## Dispersion parameter for nbinom2 family (): 0.273
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.69544    0.23398  -7.246 4.29e-13 ***
## YEAR2010    -0.09394    0.29595  -0.317 0.750921
## YEAR2011    -0.01343    0.30512  -0.044 0.964887
## YEAR2012    -0.39645    0.33303  -1.190 0.233874
## YEAR2013     0.97389    0.39398   2.472 0.013439 *
## YEAR2014    -1.22220    0.29543  -4.137 3.52e-05 ***
## YEAR2015    -1.27377    0.35340  -3.604 0.000313 ***
## YEAR2016    -1.09777    0.33298  -3.297 0.000978 ***
## YEAR2018    -0.69512    0.38314  -1.814 0.069637 .
## YEAR2019    -1.75831    0.32122  -5.474 4.40e-08 ***
## YEAR2021    -0.26391    0.45987  -0.574 0.566054
## BTEMP        0.07728    0.12904   0.599 0.549228
## DEPTH       -1.12830    0.12952  -8.711  < 2e-16 ***
## SOG         -0.06363    0.10298  -0.618 0.536678
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.5274     0.7717  -3.275  0.00106 **
## YEAR2010      -2.3563     1.0841  -2.174  0.02974 *
## YEAR2011      -0.9613     0.9189  -1.046  0.29546
## YEAR2012      -0.8605     0.9521  -0.904  0.36610
## YEAR2013       1.6096     0.7382   2.180  0.02923 *
## YEAR2014     -17.9758  2053.1213  -0.009  0.99301
## YEAR2015       1.9120     0.9415   2.031  0.04228 *
## YEAR2016       1.6654     1.1520   1.446  0.14825
## YEAR2018       2.7195     1.1287   2.409  0.01598 *
## YEAR2019      -1.0115     1.3676  -0.740  0.45953
## YEAR2021       5.3215     1.0565   5.037 4.73e-07 ***
## BTEMP         -3.5075     0.6352  -5.522 3.35e-08 ***
## DEPTH         -0.7421     0.3915  -1.896  0.05802 .
## SOG            0.1063     0.2661   0.399  0.68955
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
drop1(ZINB1, test="Chi")

## Single term deletions
##
## Model:
## CPUE ~ YEAR + BTEMP + DEPTH + SOG
##         Df    AIC    LRT  Pr(>Chi)
## <none>     3885.0
## YEAR    10 3953.4 88.466 1.079e-14 ***
## BTEMP    1 3883.3  0.358    0.5494
## DEPTH    1 3943.1 60.094 9.043e-15 ***
## SOG      1 3883.4  0.429    0.5124
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# drop YEAR in zero-infl model:
ZINB2 <- glmmTMB(CPUE ~ YEAR + BTEMP + DEPTH + SOG,offset=lnEffort, ziformula
= ~ BTEMP + DEPTH + SOG, data = dat, family=nbinom2)
summary(ZINB2)

##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + BTEMP + DEPTH + SOG
## Zero inflation:        ~BTEMP + DEPTH + SOG
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##   3954.3   4053.6  -1958.1   3916.3     1358
##
##
## Dispersion parameter for nbinom2 family (): 0.229
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.75897    0.24752  -7.106 1.19e-12 ***
## YEAR2010     0.08036    0.32080   0.251 0.802200
## YEAR2011     0.01294    0.32201   0.040 0.967956
## YEAR2012    -0.46069    0.35274  -1.306 0.191545
## YEAR2013     0.44709    0.35663   1.254 0.209973
## YEAR2014    -1.20815    0.31851  -3.793 0.000149 ***
## YEAR2015    -1.55871    0.34139  -4.566 4.98e-06 ***
## YEAR2016    -1.32848    0.32773  -4.054 5.05e-05 ***
## YEAR2018    -1.16766    0.35347  -3.303 0.000955 ***
## YEAR2019    -1.73532    0.33787  -5.136 2.81e-07 ***
## YEAR2021    -1.34265    0.32262  -4.162 3.16e-05 ***
## BTEMP        0.28166    0.14310   1.968 0.049038 *
## DEPTH       -1.10359    0.13264  -8.320  < 2e-16 ***
## SOG         -0.26620    0.11153  -2.387 0.016994 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Zero-inflation model:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.7515919  0.8062862  -3.413 0.000643 ***
## BTEMP       -3.1542864  0.6782481  -4.651 3.31e-06 ***
## DEPTH       -0.6259394  0.3808558  -1.644 0.100278
## SOG          0.0002208  0.2235632   0.001 0.999212
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZINB2, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + BTEMP + DEPTH + SOG
##        Df    AIC    LRT  Pr(>Chi)
## <none>    3954.3
## YEAR   10 4033.2 98.889 < 2.2e-16 ***
## BTEMP   1 3956.2  3.894   0.04845 *
## DEPTH   1 4008.1 55.774 8.128e-14 ***
## SOG     1 3957.9  5.608   0.01787 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# drop SOG in zero-infl model:
ZINB3 <- glmmTMB(CPUE ~ YEAR + BTEMP + DEPTH + SOG,offset=lnEffort, ziformula
= ~ BTEMP + DEPTH, data = dat, family=nbinom2)
summary(ZINB3)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + BTEMP + DEPTH + SOG
## Zero inflation:        ~BTEMP + DEPTH
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##   3952.3   4046.4  -1958.1   3916.3     1359
##
##
## Dispersion parameter for nbinom2 family (): 0.229
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.75898    0.24714  -7.117 1.10e-12 ***
## YEAR2010     0.08037    0.32077   0.251 0.802165
## YEAR2011     0.01293    0.32201   0.040 0.967962
## YEAR2012    -0.46068    0.35271  -1.306 0.191524
## YEAR2013     0.44709    0.35645   1.254 0.209744
## YEAR2014    -1.20815    0.31847  -3.794 0.000148 ***
## YEAR2015    -1.55872    0.34132  -4.567 4.95e-06 ***
```

```
## YEAR2016    -1.32847    0.32761  -4.055 5.01e-05 ***
## YEAR2018    -1.16767    0.35301  -3.308 0.000941 ***
## YEAR2019    -1.73533    0.33787  -5.136 2.81e-07 ***
## YEAR2021    -1.34268    0.32228  -4.166 3.10e-05 ***
## BTEMP        0.28167    0.14144   1.991 0.046430 *
## DEPTH       -1.10359    0.13261  -8.322  < 2e-16 ***
## SOG         -0.26622    0.10905  -2.441 0.014635 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.7518     0.7655  -3.595 0.000325 ***
## BTEMP        -3.1545     0.6608  -4.774 1.81e-06 ***
## DEPTH        -0.6260     0.3791  -1.651 0.098672 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZINB3, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + BTEMP + DEPTH + SOG
##        Df    AIC    LRT  Pr(>Chi)
## <none>    3952.3
## YEAR   10 4031.8 99.500 < 2.2e-16 ***
## BTEMP   1 3954.3  4.003   0.04542 *
## DEPTH   1 4007.1 56.770 4.898e-14 ***
## SOG     1 3956.6  6.349   0.01175 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# drop depth in zero-infl model:
ZINB4 <- glmmTMB(CPUE ~ YEAR + BTEMP + DEPTH + SOG,offset=lnEffort, ziformula
= ~ BTEMP, data = dat, family=nbinom2)
summary(ZINB4)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + BTEMP + DEPTH + SOG
## Zero inflation:        ~BTEMP
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##   3954.2   4043.0  -1960.1   3920.2     1360
##
##
## Dispersion parameter for nbinom2 family (): 0.229
##
## Conditional model:
```

```
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.69314    0.24948  -6.787 1.15e-11 ***
## YEAR2010     0.05713    0.32284   0.177 0.859548
## YEAR2011    -0.01222    0.32433  -0.038 0.969950
## YEAR2012    -0.49240    0.35397  -1.391 0.164205
## YEAR2013     0.41943    0.36019   1.164 0.244233
## YEAR2014    -1.21880    0.32135  -3.793 0.000149 ***
## YEAR2015    -1.59760    0.34190  -4.673 2.97e-06 ***
## YEAR2016    -1.34136    0.33052  -4.058 4.94e-05 ***
## YEAR2018    -1.18586    0.35731  -3.319 0.000904 ***
## YEAR2019    -1.75446    0.34028  -5.156 2.52e-07 ***
## YEAR2021    -1.35244    0.32523  -4.158 3.20e-05 ***
## BTEMP        0.31091    0.14236   2.184 0.028966 *
## DEPTH       -1.00135    0.12687  -7.893 2.96e-15 ***
## SOG         -0.26843    0.11016  -2.437 0.014824 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.3473     0.6706  -3.500 0.000465 ***
## BTEMP        -2.6204     0.5196  -5.043 4.58e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZINB4, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + BTEMP + DEPTH + SOG
##        Df    AIC    LRT  Pr(>Chi)
## <none>    3954.2
## YEAR   10 4032.3 98.174 < 2.2e-16 ***
## BTEMP   1 3957.0  4.809   0.02831 *
## DEPTH   1 4012.3 60.106  8.99e-15 ***
## SOG     1 3958.5  6.288   0.01215 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# everything is significant, so let's try a year-only model
ZINB0 <- glmmTMB(CPUE ~ YEAR,offset=lnEffort, data = dat, family=nbinom2)
summary(ZINB0)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##   4257.3   4320.0  -2116.7   4233.3     1365
```

```
##
##
## Dispersion parameter for nbinom2 family (): 0.125
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.32430    0.25624  -5.168 2.36e-07 ***
## YEAR2010     0.01703    0.35818   0.048  0.96208
## YEAR2011    -0.03011    0.36392  -0.083  0.93405
## YEAR2012    -0.40829    0.36272  -1.126  0.26033
## YEAR2013    -0.01629    0.35693  -0.046  0.96360
## YEAR2014    -0.94836    0.36491  -2.599  0.00935 **
## YEAR2015    -1.58027    0.37185  -4.250 2.14e-05 ***
## YEAR2016    -1.12318    0.36518  -3.076  0.00210 **
## YEAR2018    -0.93473    0.37898  -2.466  0.01365 *
## YEAR2019    -1.77972    0.37170  -4.788 1.68e-06 ***
## YEAR2021    -1.19018    0.36490  -3.262  0.00111 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
AICtab(ZINB1,ZINB2,ZINB3,ZINB4,ZINB0) #ZINB1-3 had convergence problems
```

```
##       dAIC  df
## ZINB1   0.0 29
## ZINB3  67.3 18
## ZINB4  69.2 17
## ZINB2  69.3 19
## ZINB0 372.3 12
```

```
lrtest(ZINB0,ZINB4,ZINB3,ZINB2,ZINB1)
```

```
## Likelihood ratio test
##
## Model 1: CPUE ~ YEAR
## Model 2: CPUE ~ YEAR + BTEMP + DEPTH + SOG
## Model 3: CPUE ~ YEAR + BTEMP + DEPTH + SOG
## Model 4: CPUE ~ YEAR + BTEMP + DEPTH + SOG
## Model 5: CPUE ~ YEAR + BTEMP + DEPTH + SOG
##   #Df  LogLik Df    Chisq Pr(>Chisq)
## 1  12 -2116.7
## 2  17 -1960.1  5 313.1420  < 2.2e-16 ***
## 3  18 -1958.2  1   3.8661    0.04927 *
## 4  19 -1958.2  1   0.0000    0.99922
## 5  29 -1913.5 10  89.3176  7.313e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sapply(list(ZINB1,ZINB4,ZINB3,ZINB2,ZINB0),tmbZI.resids,y.dat=dat$CPUE)
```

```
## Warning in sqrt(pvar): NaNs produced
```

```
## df.resid df.resid df.resid df.resid df.resid
## 1.482754 2.014014 2.040002 2.041459      NaN
```

ZANB

```
ZANB1 <- glmmTMB(CPUE ~ YEAR + BTEMP + DEPTH + SOG,
         ziformula = ~ YEAR + BTEMP + DEPTH + SOG,
         data = dat,offset=lnEffort,family = truncated_nbinom2(link = "log"))
summary(ZANB1)

##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + BTEMP + DEPTH + SOG
## Zero inflation:        ~YEAR + BTEMP + DEPTH + SOG
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##   3864.8   4016.4  -1903.4   3806.8     1348
##
##
## Dispersion parameter for truncated_nbinom2 family (): 7.86e-09
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -18.56154 2468.20396  -0.008  0.99400
## YEAR2010      -0.14196    0.42851  -0.331  0.74043
## YEAR2011      -0.27201    0.42195  -0.645  0.51915
## YEAR2012      -0.52758    0.46912  -1.125  0.26075
## YEAR2013       1.26960    0.58829   2.158  0.03092 *
## YEAR2014      -1.22225    0.44856  -2.725  0.00643 **
## YEAR2015      -1.47511    0.49981  -2.951  0.00316 **
## YEAR2016      -0.95031    0.49471  -1.921  0.05474 .
## YEAR2018      -0.96408    0.51191  -1.883  0.05966 .
## YEAR2019      -1.38237    0.51962  -2.660  0.00781 **
## YEAR2021       0.16968    0.72727   0.233  0.81552
## BTEMP          0.26643    0.18968   1.405  0.16013
## DEPTH         -0.93931    0.20086  -4.676 2.92e-06 ***
## SOG           -0.05305    0.14452  -0.367  0.71357
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.48590    0.21157   2.297 0.021639 *
## YEAR2010    -0.36901    0.29332  -1.258 0.208379
## YEAR2011    -0.74764    0.29856  -2.504 0.012275 *
## YEAR2012    -0.23127    0.30506  -0.758 0.448380
## YEAR2013     0.87452    0.32617   2.681 0.007337 **
## YEAR2014     0.49916    0.29599   1.686 0.091716 .
## YEAR2015     1.02825    0.31134   3.303 0.000958 ***
## YEAR2016     0.95363    0.30729   3.103 0.001913 **
```

```
## YEAR2018     0.68084    0.31994   2.128 0.033335 *
## YEAR2019     1.33217    0.31753   4.195 2.72e-05 ***
## YEAR2021     2.26624    0.38507   5.885 3.98e-09 ***
## BTEMP       -0.61397    0.09779  -6.278 3.42e-10 ***
## DEPTH        0.76170    0.12288   6.199 5.69e-10 ***
## SOG          0.02970    0.08496   0.350 0.726674
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZANB1, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + BTEMP + DEPTH + SOG
##        Df    AIC    LRT  Pr(>Chi)
## <none>    3864.8
## YEAR   10 3890.2 45.413 1.832e-06 ***
## BTEMP   1 3864.8  1.965    0.1610
## DEPTH   1 3881.2 18.389 1.801e-05 ***
## SOG     1 3862.9  0.158    0.6911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# large intercept in conditional model; haphazardly remove YEAR from
conditional model:
ZANB2 <- glmmTMB(CPUE ~ BTEMP + DEPTH + SOG,
         ziformula = ~ YEAR + BTEMP + DEPTH + SOG,
         data = dat,offset=lnEffort,family = truncated_nbinom2(link = "log"))
summary(ZANB2)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ BTEMP + DEPTH + SOG
## Zero inflation:        ~YEAR + BTEMP + DEPTH + SOG
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##   3890.2   3989.5  -1926.1   3852.2     1358
##
##
## Dispersion parameter for truncated_nbinom2 family (): 3.6e-09
##
## Conditional model:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -19.58164 2065.90983  -0.009    0.992
## BTEMP          0.06997    0.20704   0.338    0.735
## DEPTH         -1.03431    0.20981  -4.930 8.23e-07 ***
## SOG           -0.03939    0.12966  -0.304    0.761
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.48590    0.21157   2.297 0.021638 *
## YEAR2010    -0.36901    0.29332  -1.258 0.208374
## YEAR2011    -0.74764    0.29856  -2.504 0.012275 *
## YEAR2012    -0.23127    0.30506  -0.758 0.448376
## YEAR2013     0.87452    0.32617   2.681 0.007337 **
## YEAR2014     0.49916    0.29599   1.686 0.091717 .
## YEAR2015     1.02825    0.31134   3.303 0.000958 ***
## YEAR2016     0.95362    0.30729   3.103 0.001913 **
## YEAR2018     0.68084    0.31994   2.128 0.033336 *
## YEAR2019     1.33217    0.31753   4.195 2.72e-05 ***
## YEAR2021     2.26624    0.38507   5.885 3.98e-09 ***
## BTEMP       -0.61397    0.09779  -6.278 3.42e-10 ***
## DEPTH        0.76170    0.12288   6.199 5.69e-10 ***
## SOG          0.02970    0.08496   0.350 0.726673
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZANB2, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ BTEMP + DEPTH + SOG
##        Df    AIC     LRT  Pr(>Chi)
## <none>    3890.2
## BTEMP   1 3888.3  0.1145    0.7350
## DEPTH   1 3908.8 20.6382 5.548e-06 ***
## SOG     1 3888.3  0.1021    0.7494
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# large intercept in conditional model; haphazardly remove SOG from
conditional model:
ZANB3 <- glmmTMB(CPUE ~ BTEMP + DEPTH,
         ziformula = ~ YEAR + BTEMP + DEPTH + SOG,
         data = dat,offset=lnEffort,family = truncated_nbinom2(link = "log"))
summary(ZANB3)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ BTEMP + DEPTH
## Zero inflation:        ~YEAR + BTEMP + DEPTH + SOG
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##   3888.3   3982.4  -1926.2   3852.3     1359
##
##
```

```
## Dispersion parameter for truncated_nbinom2 family (): 3.42e-09
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -19.62248 2124.54357  -0.009    0.993
## BTEMP          0.05247    0.19857   0.264    0.792
## DEPTH         -1.03915    0.20870  -4.979 6.39e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.48590    0.21157   2.297 0.021639 *
## YEAR2010    -0.36901    0.29332  -1.258 0.208378
## YEAR2011    -0.74764    0.29856  -2.504 0.012275 *
## YEAR2012    -0.23127    0.30506  -0.758 0.448379
## YEAR2013     0.87452    0.32617   2.681 0.007337 **
## YEAR2014     0.49916    0.29599   1.686 0.091716 .
## YEAR2015     1.02825    0.31134   3.303 0.000958 ***
## YEAR2016     0.95363    0.30729   3.103 0.001913 **
## YEAR2018     0.68084    0.31994   2.128 0.033335 *
## YEAR2019     1.33217    0.31753   4.195 2.72e-05 ***
## YEAR2021     2.26624    0.38507   5.885 3.98e-09 ***
## BTEMP       -0.61397    0.09779  -6.278 3.42e-10 ***
## DEPTH        0.76170    0.12288   6.199 5.69e-10 ***
## SOG          0.02970    0.08496   0.350 0.726674
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZANB3, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ BTEMP + DEPTH
##       Df    AIC     LRT  Pr(>Chi)
## <none>    3888.3
## BTEMP   1 3886.4  0.0698    0.7916
## DEPTH   1 3907.3 20.9730 4.658e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# large intercept in conditional model; haphazardly remove SOGBTEMP from
conditional model:
ZANB4 <- glmmTMB(CPUE ~ DEPTH,
         ziformula = ~ YEAR + BTEMP + DEPTH + SOG,
         data = dat,offset=lnEffort,family = truncated_nbinom2(link = "log"))
summary(ZANB4)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ DEPTH
```

```
## Zero inflation:        ~YEAR + BTEMP + DEPTH + SOG
## Data: dat
##  Offset: lnEffort
##
##      AIC      BIC   logLik deviance df.resid
##   3886.4   3975.2  -1926.2   3852.4     1360
##
##
## Dispersion parameter for truncated_nbinom2 family (): 3.44e-09
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -19.608   2119.435  -0.009    0.993
## DEPTH         -1.071      0.170  -6.302 2.94e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.48590    0.21157   2.297 0.021638 *
## YEAR2010    -0.36901    0.29332  -1.258 0.208374
## YEAR2011    -0.74764    0.29856  -2.504 0.012275 *
## YEAR2012    -0.23127    0.30506  -0.758 0.448379
## YEAR2013     0.87452    0.32617   2.681 0.007337 **
## YEAR2014     0.49916    0.29599   1.686 0.091716 .
## YEAR2015     1.02825    0.31134   3.303 0.000958 ***
## YEAR2016     0.95363    0.30729   3.103 0.001913 **
## YEAR2018     0.68084    0.31994   2.128 0.033336 *
## YEAR2019     1.33217    0.31753   4.195 2.72e-05 ***
## YEAR2021     2.26624    0.38507   5.885 3.98e-09 ***
## BTEMP       -0.61397    0.09779  -6.278 3.42e-10 ***
## DEPTH        0.76170    0.12288   6.199 5.69e-10 ***
## SOG          0.02970    0.08496   0.350 0.726677
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

drop1(ZANB4, test="Chi")

## Single term deletions
##
## Model:
## CPUE ~ DEPTH
##        Df    AIC    LRT  Pr(>Chi)
## <none>    3886.4
## DEPTH   1 3913.7 29.348 6.047e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# still have large SE in intercept, so can't use ZANB
```

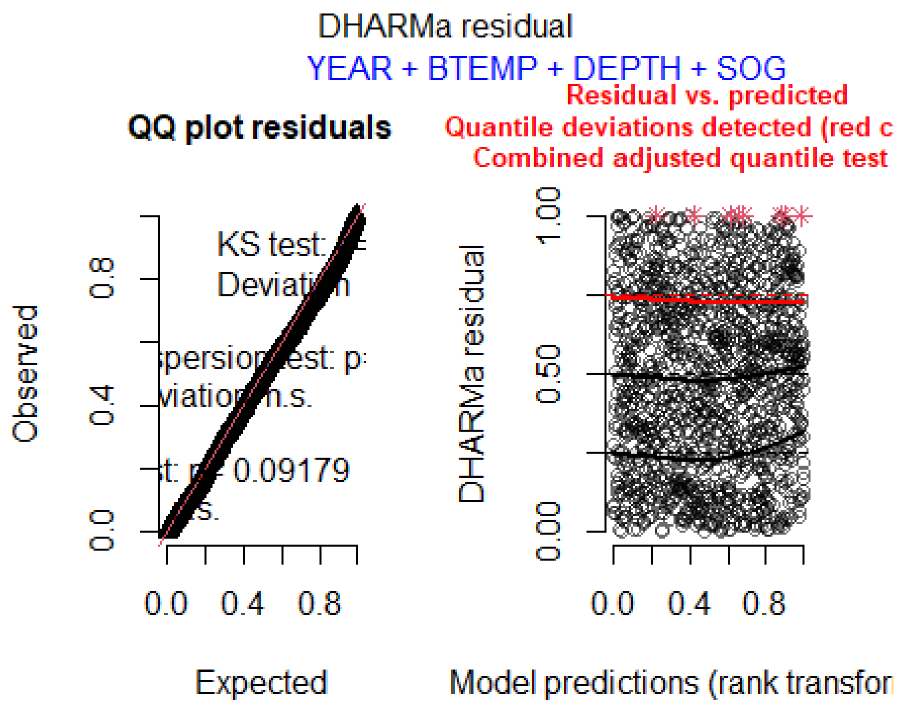
Model comparison:

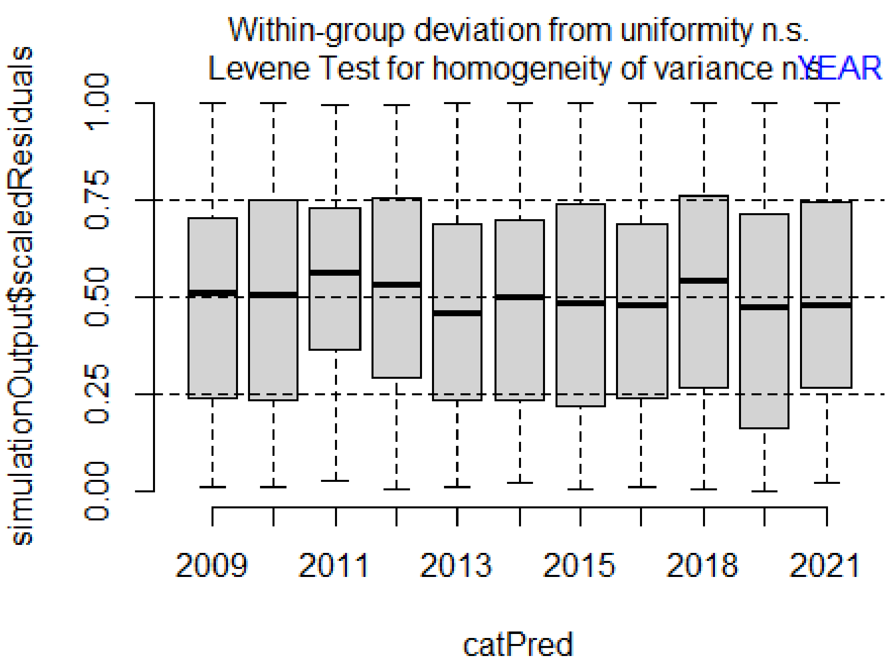
```
AICtab(NB1,ZINB1)

##        dAIC  df
## ZINB1    0.0 29
## NB1    112.9 15
```

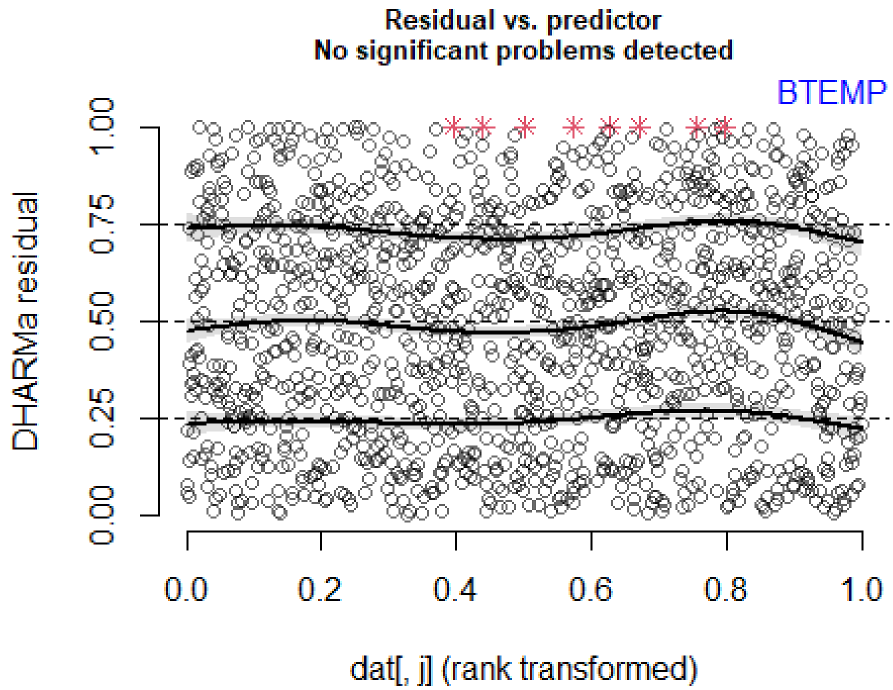
Model diagnostics

Negative binomial & ZINB

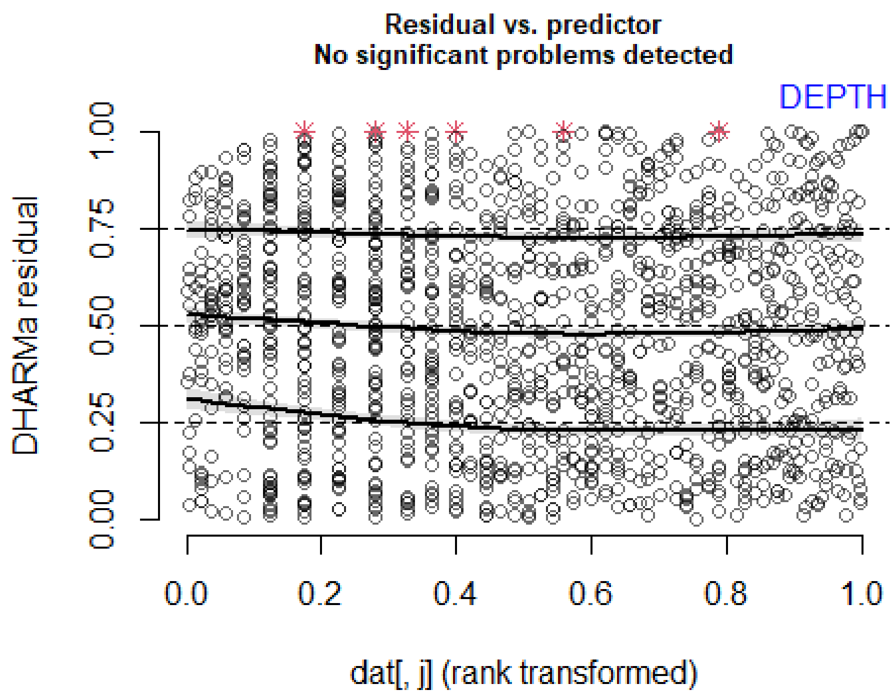
```
models <- list(ZINB1)# pretty good diagnostics
for(i in 1:length(models)) {
   NB.sim <- simulateResiduals(models[[i]])
   this <- names(dat)[names(dat) %in% names(models[[i]]$frame) & names(dat)
!= "CPUE"]
   print(this)
   plot(NB.sim,quantreg=TRUE); mtext(paste(this,collapse=" +
"),side=3,adj=0,outer=FALSE,line=1,col="blue")
   #mtext(formula(models[[i]]),side=3,adj=0,outer=FALSE,line=1,col="blue")

   for(j in this) {
      plotResiduals(NB.sim, form=dat[,j],quantreg=TRUE);
mtext(j,3,col="blue",adj=1)
   }
}

## [1] "YEAR"  "BTEMP" "DEPTH" "SOG"
```


DHARMa residual
YEAR + BTEMP + DEPTH + SOG
QQ plot residuals
Residual vs. predicted
Quantile deviations detected (red c
Combined adjusted quantile test
KS test:
Deviati n
persio est: p
viatio n.s.
t: p = 0.09179
Observed
Expected
DHARMa residual
Model predictions (rank transfor
0.0
0.4
0.8
0.00
0.50
1.00

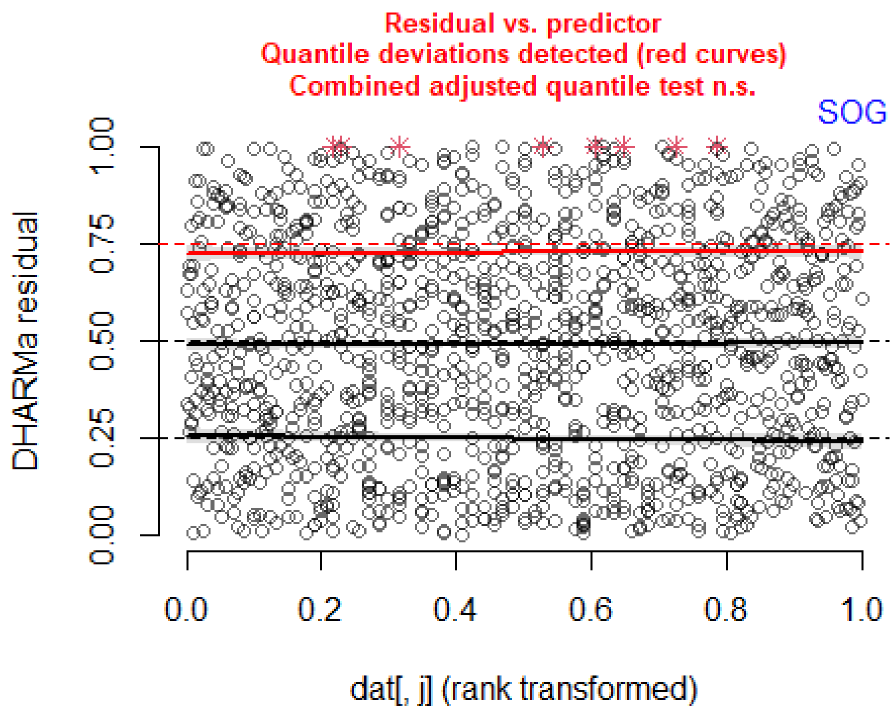

Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
YEAR
simulationOutput$scaledResiduals
catPred
0.00
0.25
0.50
0.75
1.00
2009
2011
2013
2015
2018
2021



Residual vs. predictor
No significant problems detected
BTEMP
DHARMa residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
dat[, j] (rank transformed)



Residual vs. predictor
No significant problems detected
DEPTH
DHARMa residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
dat[, j] (rank transformed)



Conclusion: ZANB7 is the best model based on AIC, and has very good model diagnostics.

Generate and compare indices

```
# by-hand with bootstrapping :
# requires final dataset object called 'dat'
out <- byHand2(ZINB1,nBoot=1000,Method="ZINB1")

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

# calc stratified mean (using all data, not just the data used in dat, which
has some records removed due to missing env data):
# stratified mean:
# Standardize catch for 30 minute tow (the standard unit of effort of NEFSC)
a <-

as.data.frame(read_xlsx(path="f:\\BF_WG\\FIG\\NEFSC\\NEFSC_Fall_Survey_Data_f
or_Std_V3.xlsx",
   sheet="BLUEFISH_FALL_BIG_RT2022_09_21",range="A75:AJ1513",na="NA"))
head(a)

##   CRUISE6 STRATUM TOW STATION SHG SVVESSEL SVGEAR EST_YEAR EST_MONTH
EST_DAY
## 1  200904    1010   1     142 111       HB     10     2009        10
1
## 2  200904    1010   2     126 111       HB     10     2009         9
```

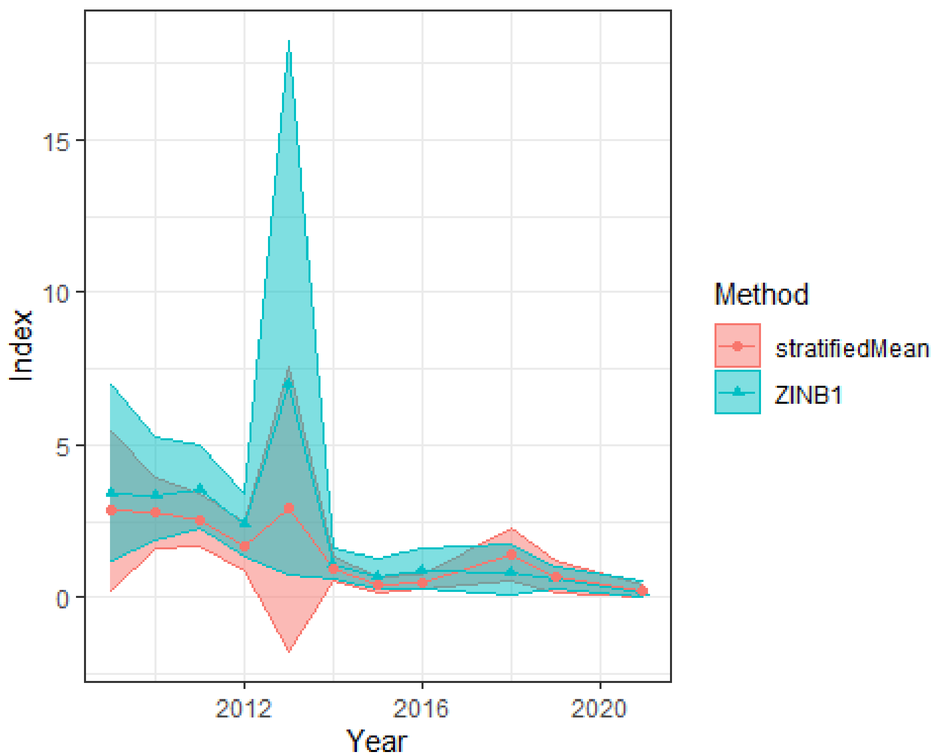
```
29
## 3  200904    1010   3     132 122       HB     10     2009         9
30
## 4  200904    1010   4     130 111       HB     10     2009         9
30
## 5  200904    1010   5     125 111       HB     10     2009         9
29
## 6  200904    1010   6     131 111       HB     10     2009         9
30
##   TIME TOWDUR DOPDISTB DOPDISTW AVGDEPTH STATYPE HAUL GEARCOND TYPE_CODE
## 1  228  20.02    0.958       NA       54       1    1        1         1
## 2 1853  20.02    0.990       NA       35       1    1        1         1
## 3  445  20.03    1.024       NA       45       1    2        2         1
## 4  138  20.03    1.009       NA       44       1    1        1         1
## 5 1649  20.03    0.955       NA       37       1    1        1         1
## 6  252  20.02    0.995       NA       42       1    1        1         1
##   OPERATION_CODE GEAR_CODE ACQUISITION_CODE AREA BOTTEMP   BEGLAT
BEGLON
## 1              1         1                1  616   13.56 39.92030 -
72.84734
## 2              1         1                1  613   19.27 40.64319 -
72.65719
## 3              1         2                1  612   18.93 40.22850 -
73.08143
## 4              3         1                1  613      NA 40.39845 -
72.85305
## 5              3         1                2  613   13.70 40.68921 -
72.44950
## 6              1         1                1  613   17.80 40.39484 -
72.90423
##   STRATUM_AREA      STRGRP_DESC     STRATUM_NAME AREA_SWEPT_WINGS_MEAN_KM2
## 1         2516 OFFSHORE REGULAR SNE HUDSON CANYN                   0.02271
## 2         2516 OFFSHORE REGULAR SNE HUDSON CANYN                   0.02164
## 3         2516 OFFSHORE REGULAR SNE HUDSON CANYN                   0.02276
## 4         2516 OFFSHORE REGULAR SNE HUDSON CANYN                   0.02280
## 5         2516 OFFSHORE REGULAR SNE HUDSON CANYN                   0.02016
## 6         2516 OFFSHORE REGULAR SNE HUDSON CANYN                   0.02174
##   AREA_SWEPT_WINGS_MEAN_NM2 SWEPT_AREA_RATIO SVSPP CATCHSEX EXPCATCHWT
## 1               0.006621181         1.057213    NA        0   0.000000
## 2               0.006309219         1.109488   135        0   2.041457
## 3               0.006635759         1.054891   135        0   4.683714
## 4               0.006647421         1.053040   135        0   5.823311
## 5               0.005877720         1.190938    NA        0   0.000000
## 6               0.006338374         1.104384    NA        0   0.000000
##   EXPCATCHNUM
## 1    0.000000
## 2    1.109488
## 3    1.054891
## 4    2.106080
```

```
## 5    0.000000
## 6    0.000000

a <- a[a$EST_YEAR!=2017,] # omit 2017 due to sparse sampling


# Standardize catch for 20 minute tow (the standard unit of effort of NEFSC)
a$EXPCATCHNUM <- round((a$EXPCATCHNUM*20)/a$TOWDUR,0)#a$EXPCATCHNUM
                                                         #^^^^^^^^^^^^^^^ NEFSC uses un-
adjusted catches; the FIG recommended adjusting to a standard tow (which for
NEFSC is 30 min for Albatross, and 20 min for Bigelow).

# strata wts:
lookup <- with(a, tapply(STRATUM_AREA,STRATUM,median))

lookup <- data.frame("STRATUM_AREA"=lookup, "STRATUM"=rownames(lookup))
lookup$wts <- lookup$STRATUM_AREA/sum(lookup$STRATUM_AREA)

lookup <- lookup[order(lookup$STRATUM),] # should already be in increasing
order, but just to be certain

   # mean:
   M <- with(a, tapply(EXPCATCHNUM ,list(STRATUM,EST_YEAR),mean,na.rm=TRUE))
   # variance:
   V <- with(a, tapply(EXPCATCHNUM ,list(STRATUM,EST_YEAR),var,na.rm=TRUE))

   # matrix of stratum wts:
   wts <- ((!is.na(M))*1) * lookup$STRATUM_AREA
   wts <- t(t(wts)/colSums(wts))


   # Possible workaround if a year-stratum combination has NA variance (i.e.,
from n= 1), pool mean variance from entire year:
   needVar <- which(is.na(V))
   if(length(needVar)>0) {
      allYears <- sort(unique(a$EST_YEAR))
      # years of interest:
      yoi <- rep(allYears,each=nrow(lookup))[needVar]
      lookupTable <-
data.frame(year=allYears,.var=apply(V,2,mean,na.rm=TRUE))
         # The special case of all NAs in a given year:
         # If NaN, then pool timeseries mean variance
         if(any(is.nan(lookupTable$.var))) {
            mod <- which(is.nan(lookupTable$.var))
            lookupTable$.var[mod] <- mean(lookupTable$.var[-mod])
         }
      ans <-
merge(lookupTable,data.frame(year=yoi,tmp=NA),by="year",sort=FALSE)
      V[which(is.na(V))] <- ans$.var
```

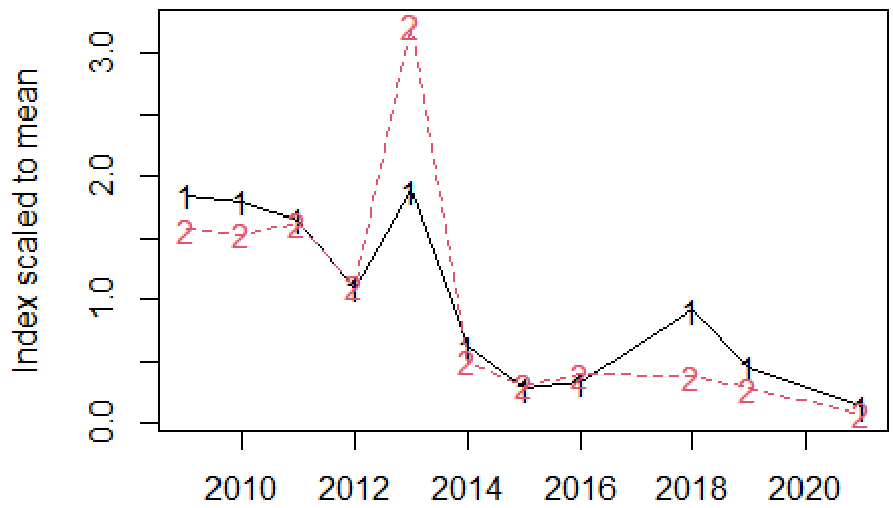
```
  }

     # sample size (per stratum, per year)
     N <- with(a, tapply(EXPCATCHNUM,list(STRATUM,EST_YEAR),length))


  # UCL
  UCL <-
  colSums(M*wts,na.rm=TRUE)+
     ((sqrt(colSums(wts^2*V/N,na.rm=TRUE)))*qnorm((1-0.95)/2,lower=FALSE))

  # stratified mean:
  .m <- colSums(M*wts,na.rm=TRUE)

  # LCL
  LCL <-
  colSums(M*wts,na.rm=TRUE)-
     ((sqrt(colSums(wts^2*V/N,na.rm=TRUE)))*qnorm((1-0.95)/2,lower=FALSE))

  SE <- sqrt(colSums(wts^2*V/N,na.rm=TRUE))
  CV <- SE/.m

  nom <- data.frame("Year"=sort(unique(a$EST_YEAR)),"Index"=.m,SE,
     "LCI"=LCL,"UCI"=UCL,CV,"Method"="stratifiedMean")


tmp <- rbind(out,nom)
tmp

##         Year     Index        SE          LCI        UCI        CV
Method
## 2009   2009 3.3998299 1.5470618  1.203493360  6.9758987 0.4550409
ZINB1
## 2010   2010 3.3170658 0.8655845  1.866683971  5.2405061 0.2609488
ZINB1
## 2011   2011 3.5150361 0.6888713  2.281810066  4.9753351 0.1959784
ZINB1
## 2012   2012 2.3890450 0.5294157  1.353883652  3.3979203 0.2216014
ZINB1
## 2013   2013 6.9477561 4.6425683  0.733553535 18.2323566 0.6682112
ZINB1
## 2014   2014 1.0815218 0.2499971  0.603440903  1.6026940 0.2311531
ZINB1
## 2015   2015 0.6667906 0.2424692  0.318304123  1.2660964 0.3636362
ZINB1
## 2016   2016 0.8611394 0.3336165  0.319317905  1.5932982 0.3874128
ZINB1
## 2018   2018 0.8282965 0.4385221  0.116537658  1.7604135 0.5294265
ZINB1
```

```
## 2019  2019 0.6148474 0.1987637  0.268550762  1.0241561 0.3232732
ZINB1
## 2021  2021 0.1625319 0.1757970  0.003450342  0.5418350 1.0816152
ZINB1
## 20091 2009 2.8433400 1.3323651  0.231952285  5.4547277 0.4685916
stratifiedMean
## 20101 2010 2.7763567 0.5850538  1.629672229  3.9230411 0.2107272
stratifiedMean
## 20111 2011 2.5508164 0.4335892  1.700997169  3.4006356 0.1699806
stratifiedMean
## 20121 2012 1.6801595 0.3912599  0.913304246  2.4470148 0.2328707
stratifiedMean
## 20131 2013 2.9090408 2.3786187 -1.752966110  7.5710477 0.8176642
stratifiedMean
## 20141 2014 0.9705161 0.2029900  0.572663095  1.3683691 0.2091567
stratifiedMean
## 20151 2015 0.4327516 0.1253329  0.187103591  0.6783996 0.2896186
stratifiedMean
## 20161 2016 0.5086776 0.1154291  0.282440709  0.7349145 0.2269200
stratifiedMean
## 20181 2018 1.4188919 0.4475638  0.541682918  2.2961010 0.3154319
stratifiedMean
## 20191 2019 0.7034634 0.2693079  0.175629720  1.2312972 0.3828314
stratifiedMean
## 20211 2021 0.2097204 0.1059093  0.002141998  0.4172987 0.5050023
stratifiedMean
```

```
ggplot(tmp, aes(x=as.numeric(Year), y=Index, color=Method, shape=Method)) +
  geom_ribbon(aes(x=as.numeric(Year), ymin=LCI, ymax=UCI, fill=Method),
alpha=0.5) +
  geom_line() + geom_point() +
  xlab("Year") + theme_bw()
```




```
# scale indices to respective means:
matplot(unique(out$Year),
sapply(split(tmp[!is.element(tmp$Year,c(2017,2020)),],
   tmp[!is.element(tmp$Year,c(2017,2020)),"Method"]), function(x) {
      scaleToMean(x$Index)
   }
),type="o",xlab="",ylab="Index scaled to mean"
)
```




```
round(cor(sapply(split(tmp[!is.element(tmp$Year,c(2017,2020)),],
   tmp[!is.element(tmp$Year,c(2017,2020)),"Method"]),function(x) x$Index)),2)

##                stratifiedMean ZINB1
## stratifiedMean           1.00  0.87
## ZINB1                    0.87  1.00

round(cor(sapply(split(tmp[!is.element(tmp$Year,c(2017,2020)),],
   tmp[!is.element(tmp$Year,c(2017,2020)),"Method"]),function(x)
x$Index),method="spearman"),2)

##                stratifiedMean ZINB1
## stratifiedMean           1.00  0.91
## ZINB1                    0.91  1.00

Sys.time()

## [1] "2022-06-10 11:04:32 EDT"
```

Summary:

- When I re-ran code with the addition of the newly added strata I realized that I had MONTH as a covariate in the runs with the continuity strata, and since Nov was a month with very little sampling, that was causing model convergence problems. So I did full model selection for all models, etc. (As opposed to Albatross, where I did expedited model runs).

- I can’t use the summary values in the spreadsheet Tony provided to calc CIs (the CIs from summary data are too narrow; though the summary values do appear to provide realistic CVs), so I used the raw data in the spreadsheet to cal CIs (and SEs and CVs). I standardized all catch data to a 30 min tow (which the summary data does not reflect) since it is consistent with the FIG protocols. The CVs from Tony’s summary spreadsheet information and from scratch/by-hand are very similar.

Bluefish index standardization - NEAMAP Survey

Abigail Tyrell

2022-02-11

Develop standardized index

Step 1: Data processing

Load functions

```
source(here::here("scripts/functions.R"))

## Loading required package: carData

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Warning: package 'glmmTMB' was built under R version 4.0.5

## Registered S3 methods overwritten by 'lme4':
##   method                          from
##   cooks.distance.influence.merMod car
##   influence.merMod                car
##   dfbeta.influence.merMod         car
##   dfbetas.influence.merMod        car

## Warning: package 'DHARMa' was built under R version 4.0.5

## This is DHARMa 0.4.4. For overview type '?DHARMa'. For recent changes,
type news(package = 'DHARMa')

## Warning: package 'bbmle' was built under R version 4.0.5

## Loading required package: stats4

## Warning: package 'ggeffects' was built under R version 4.0.5

## Warning: package 'ggplot2' was built under R version 4.0.5
```

Read in data

```
indata <- readxl::read_excel(here::here("data/2021_bluefish_data_VA.xlsx"),
                             sheet = "Survey Data_NEAMAP"
```

```
                             ) %>%
  dplyr::rename(Stratum_area = Wgt.Factor)
head(indata)

## # A tibble: 6 x 21
##   SampleID      State Program Gear         Year  Season Date
##   <chr>         <chr> <chr>   <chr>        <chr> <chr>  <dttm>
## 1 NM20070901001 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-28
00:00:00
## 2 NM20070901002 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-28
00:00:00
## 3 NM20070901003 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-27
00:00:00
## 4 NM20070901004 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-27
00:00:00
## 5 NM20070901005 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-28
00:00:00
## 6 NM20070901006 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-27
00:00:00
## # ... with 14 more variables: Time <dttm>, Areasw <dbl>, Bf.count <dbl>,
## #   Bf.totwght <dbl>, Area <chr>, Region <chr>, Depthstrat <chr>, Depth
<dbl>,
## #   Stratum_area <dbl>, Lat <dbl>, Lon <dbl>, Bottom.wt <dbl>, Bottom.sa
<dbl>,
## #   Bottom.do <dbl>

bio_data <- readxl::read_excel(here::here("data/2021_bluefish_data_VA.xlsx"),
                             sheet = "BioSamples_NEAMAP"
                             ) %>%
  dplyr::rename(Length_expansion = Wgt.Factor,
                Season = Month) %>%
  dplyr::mutate(old_sampleID = SampleID,
                SampleID = SampleID %>%
                  stringr::str_trunc(side = "right", width = 13, ellipsis =
""))
head(bio_data)

## # A tibble: 6 x 17
##   SampleID  Station   Specimid State Year  Season Source   Program Gear
Area
##   <chr>     <chr>        <dbl> <chr> <chr> <chr>  <chr>    <chr>   <chr>
<chr>
## 1 NM200709~ NM200709~        1 NY    2007  FALL   Fisheri~ NEAMAP  Botto~
Coast~
## 2 NM200709~ NM200709~        2 NY    2007  FALL   Fisheri~ NEAMAP  Botto~
Coast~
## 3 NM200709~ NM200709~        3 NY    2007  FALL   Fisheri~ NEAMAP  Botto~
Coast~
## 4 NM200709~ NM200709~        4 NY    2007  FALL   Fisheri~ NEAMAP  Botto~
Coast~
```

```
## 5 NM200709~ NM200709~        5 NY    2007  FALL   Fisheri~ NEAMAP  Botto~
Coast~
## 6 NM200709~ NM200709~        1 NY    2007  FALL   Fisheri~ NEAMAP  Botto~
Coast~
## # ... with 7 more variables: Oto.Age <dbl>, Fork.Length <dbl>, Weight
<dbl>,
## #   Sex <chr>, Maturity <chr>, Length_expansion <dbl>, old_sampleID <chr>

all_data <- dplyr::full_join(indata, bio_data,
                             by = c("SampleID", "State", "Program",
                                    "Gear", "Year", "Season"))
head(all_data)

## # A tibble: 6 x 32
##   SampleID      State Program Gear         Year  Season Date
##   <chr>         <chr> <chr>   <chr>        <chr> <chr>  <dttm>
## 1 NM20070901001 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-28
00:00:00
## 2 NM20070901002 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-28
00:00:00
## 3 NM20070901002 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-28
00:00:00
## 4 NM20070901002 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-28
00:00:00
## 5 NM20070901002 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-28
00:00:00
## 6 NM20070901002 NY    NEAMAP  Bottom Trawl 2007  FALL   2007-09-28
00:00:00
## # ... with 25 more variables: Time <dttm>, Areasw <dbl>, Bf.count <dbl>,
## #   Bf.totwght <dbl>, Area.x <chr>, Region <chr>, Depthstrat <chr>,
## #   Depth <dbl>, Stratum_area <dbl>, Lat <dbl>, Lon <dbl>, Bottom.wt
<dbl>,
## #   Bottom.sa <dbl>, Bottom.do <dbl>, Station <chr>, Specimid <dbl>,
## #   Source <chr>, Area.y <chr>, Oto.Age <dbl>, Fork.Length <dbl>, Weight
<dbl>,
## #   Sex <chr>, Maturity <chr>, Length_expansion <dbl>, old_sampleID <chr>

# not sure how to use length data in standardization??

# modify data formatting - not using all_data
indata <- indata %>%
  dplyr::rename(FREQ = Bf.count) %>%
  dplyr::select(-c(Program, Gear, Area)) # same value for all data
head(indata)

## # A tibble: 6 x 18
##   SampleID     State Year  Season Date                Time
Areasw
##   <chr>        <chr> <chr> <chr>  <dttm>              <dttm>
<dbl>
## 1 NM200709010~ NY    2007  FALL   2007-09-28 00:00:00 2007-09-28 10:21:00
```

```
27800.
## 2 NM200709010~ NY    2007  FALL   2007-09-28 00:00:00 2007-09-28 07:50:00
25146.
## 3 NM200709010~ NY    2007  FALL   2007-09-27 00:00:00 2007-09-27 18:03:00
31078.
## 4 NM200709010~ NY    2007  FALL   2007-09-27 00:00:00 2007-09-27 15:22:00
22701.
## 5 NM200709010~ NY    2007  FALL   2007-09-28 00:00:00 2007-09-28 06:29:00
23059.
## 6 NM200709010~ NY    2007  FALL   2007-09-27 00:00:00 2007-09-27 13:41:00
22826.
## # ... with 11 more variables: FREQ <dbl>, Bf.totwght <dbl>, Region <chr>,
## #   Depthstrat <chr>, Depth <dbl>, Stratum_area <dbl>, Lat <dbl>, Lon
<dbl>,
## #   Bottom.wt <dbl>, Bottom.sa <dbl>, Bottom.do <dbl>
```

Remove missing data

```
new_data <- remove_missing(indata,
                           f_col = "FREQ",
                           t_col = "Year")
```

```
## Data before removing missing values:

##    SampleID           State              Year              Season
##  Length:3982        Length:3982        Length:3982        Length:3982
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##       Date                           Time                           Areasw
##  Min.   :2007-09-27 00:00:00   Min.   :2007-09-27 07:17:00   Min.   :11852
##  1st Qu.:2010-10-13 00:00:00   1st Qu.:2010-10-13 13:17:00   1st Qu.:23685
##  Median :2014-04-27 00:00:00   Median :2014-04-27 17:46:43   Median :24984
##  Mean   :2014-04-08 16:23:59   Mean   :2014-04-09 05:03:35   Mean   :24913
##  3rd Qu.:2017-10-17 00:00:00   3rd Qu.:2017-10-17 17:36:54   3rd Qu.:26263
##  Max.   :2021-05-24 00:00:00   Max.   :2021-05-24 15:00:53   Max.   :35910
##
##       FREQ            Bf.totwght          Region           Depthstrat
##  Min.   :    0.00   Min.   :   0.000   Length:3982        Length:3982
##  1st Qu.:    0.00   1st Qu.:   0.000   Class :character   Class :character
##  Median :    0.00   Median :   0.000   Mode  :character   Mode  :character
##  Mean   :   22.93   Mean   :   2.262
##  3rd Qu.:    4.00   3rd Qu.:   0.678
##  Max.   :15575.00   Max.   :2118.498
##
##      Depth          Stratum_area            Lat             Lon
##  Min.   : 3.566   Min.   : 31301304   Min.   :35.16   Min.   :-76.00
##  1st Qu.:11.278   1st Qu.:267450194   1st Qu.:37.55   1st Qu.:-75.46
```

```
##  Median :14.722   Median :523849238   Median :38.85   Median :-74.77
##  Mean   :16.815   Mean   :482887057   Mean   :38.94   Mean   :-74.17
##  3rd Qu.:18.898   3rd Qu.:697642474   3rd Qu.:40.65   3rd Qu.:-73.06
##  Max.   :60.716   Max.   :993986179   Max.   :41.44   Max.   :-70.87
##
##    Bottom.wt       Bottom.sa       Bottom.do
##  Min.   : 4.92   Min.   :22.66   Min.   :-128.670
##  1st Qu.:10.96   1st Qu.:31.07   1st Qu.:   6.820
##  Median :16.27   Median :32.01   Median :   7.550
##  Mean   :14.82   Mean   :32.20   Mean   :   7.135
##  3rd Qu.:18.31   3rd Qu.:33.10   3rd Qu.:   8.530
##  Max.   :24.52   Max.   :37.50   Max.   :  60.000
##  NA's   :50      NA's   :194     NA's   :97
```

```
## Data after removing missing values:
```

```
##    SampleID            State               Year              Season
##  Length:3697        Length:3697        Length:3697        Length:3697
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##       Date                             Time                            Areasw
##  Min.   :2007-09-27 00:00:00   Min.   :2007-09-27 07:17:00   Min.   :11852
##  1st Qu.:2010-10-08 00:00:00   1st Qu.:2010-10-08 13:34:00   1st Qu.:23729
##  Median :2013-10-27 00:00:00   Median :2013-10-27 16:26:47   Median :25027
##  Mean   :2014-02-06 09:57:30   Mean   :2014-02-06 22:38:13   Mean   :24951
##  3rd Qu.:2017-06-23 00:00:00   3rd Qu.:2017-06-23 12:32:59   3rd Qu.:26315
##  Max.   :2021-05-24 00:00:00   Max.   :2021-05-24 15:00:53   Max.   :35910
##       FREQ             Bf.totwght          Region           Depthstrat
##  Min.   :    0.00   Min.   :   0.000   Length:3697        Length:3697
##  1st Qu.:    0.00   1st Qu.:   0.000   Class :character   Class :character
##  Median :    0.00   Median :   0.000   Mode  :character   Mode  :character
##  Mean   :   23.46   Mean   :   2.308
##  3rd Qu.:    4.00   3rd Qu.:   0.664
##  Max.   :15575.00   Max.   :2118.498
##      Depth         Stratum_area            Lat             Lon
##  Min.   : 3.566   Min.   : 31301304   Min.   :35.16   Min.   :-76.00
##  1st Qu.:11.278   1st Qu.:267450194   1st Qu.:37.52   1st Qu.:-75.47
##  Median :14.722   Median :523849238   Median :38.83   Median :-74.80
##  Mean   :16.773   Mean   :484195382   Mean   :38.92   Mean   :-74.18
##  3rd Qu.:18.898   3rd Qu.:697642474   3rd Qu.:40.66   3rd Qu.:-73.04
##  Max.   :60.716   Max.   :993986179   Max.   :41.44   Max.   :-70.87
##    Bottom.wt       Bottom.sa       Bottom.do
##  Min.   : 4.92   Min.   :22.66   Min.   :-128.670
##  1st Qu.:10.89   1st Qu.:31.07   1st Qu.:   6.820
##  Median :15.72   Median :31.99   Median :   7.550
##  Mean   :14.71   Mean   :32.19   Mean   :   7.124
```

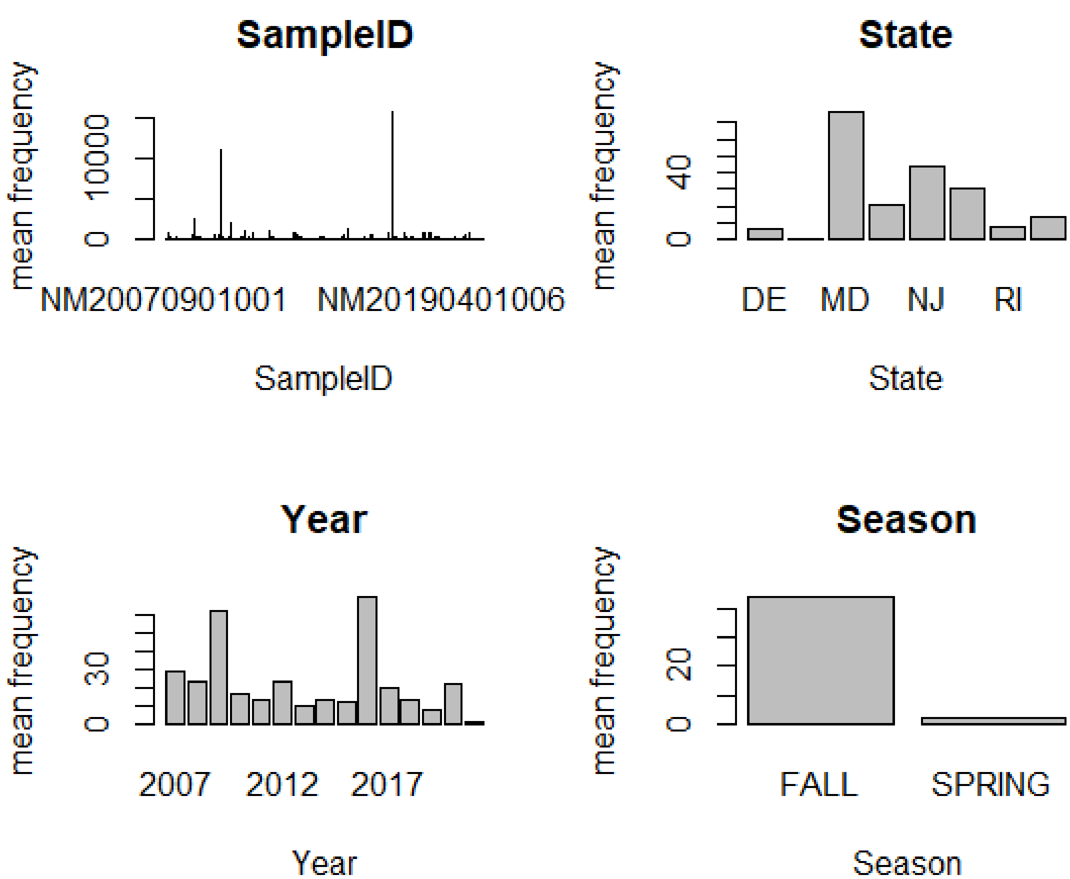
```
##   3rd Qu.:18.26    3rd Qu.:33.07    3rd Qu.:   8.530
##   Max.   :24.52    Max.   :37.50    Max.   :  60.000

## 285 datapoints removed (7.2% of data removed)

## All years had positive catch
```

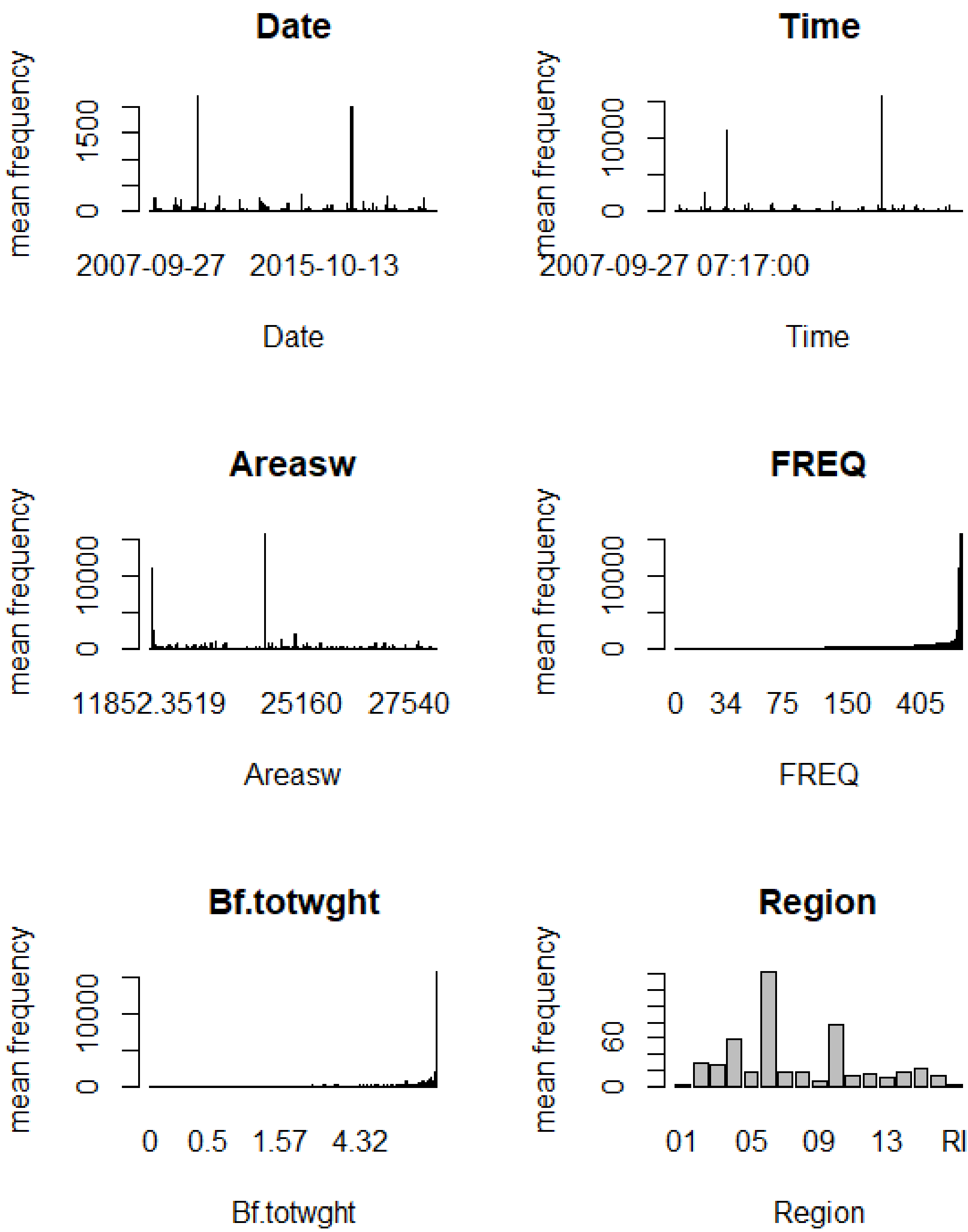
Add Month column

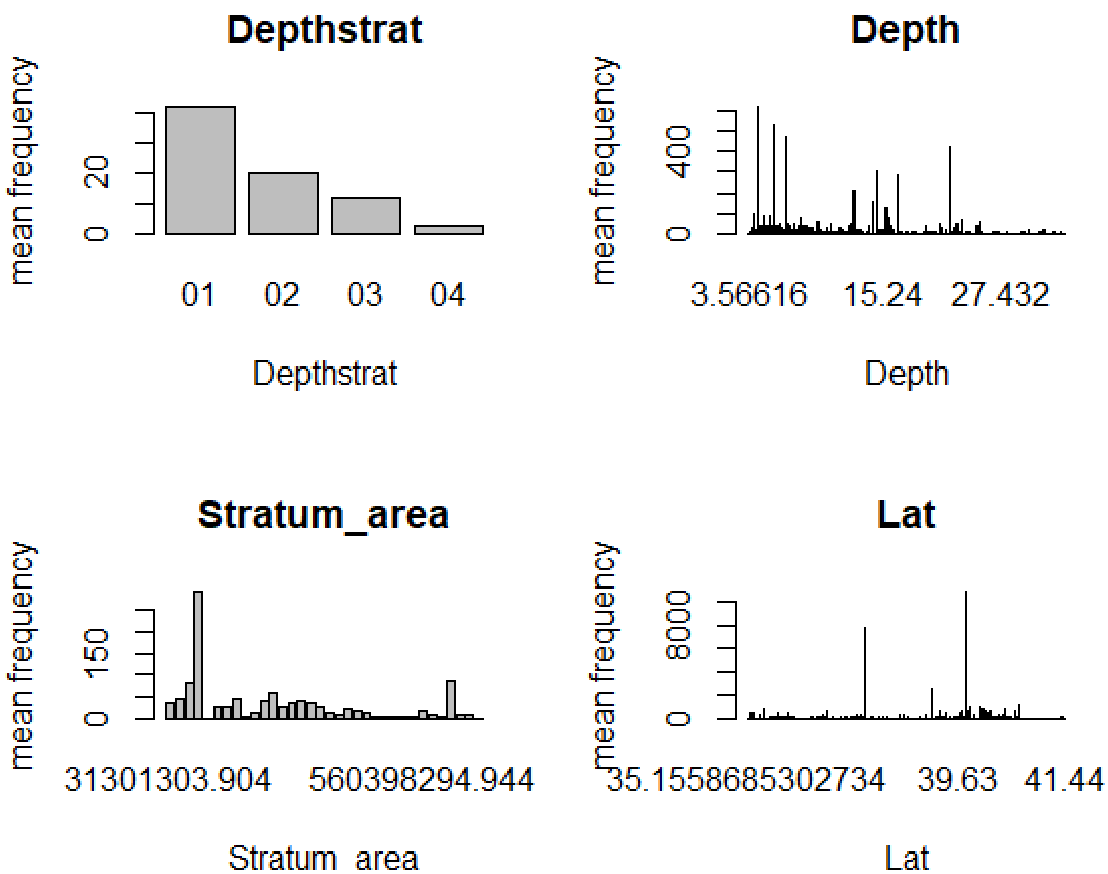
```
new_data <- new_data %>%
  dplyr::mutate(Month = as.Date(Time) %>%
                  format("%m")
                )
```

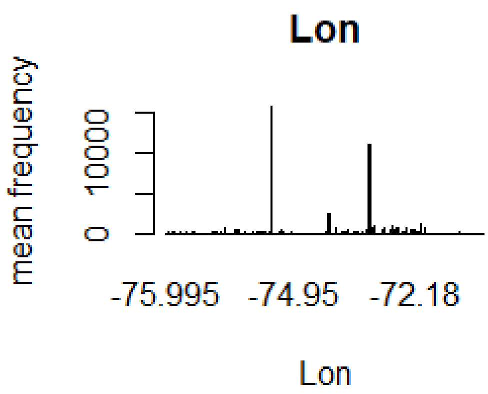
Plot data and decide if any data should be removed due to sparse sampling

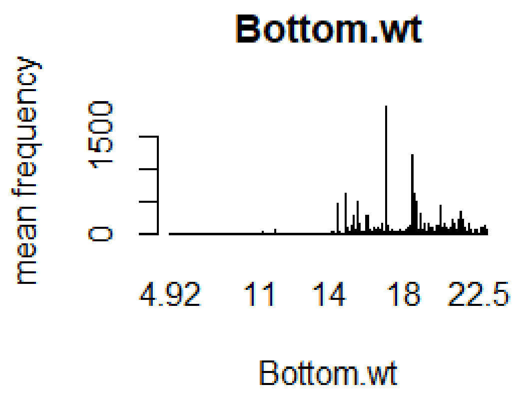
```
plot_obs(new_data,
         col = "FREQ")
```

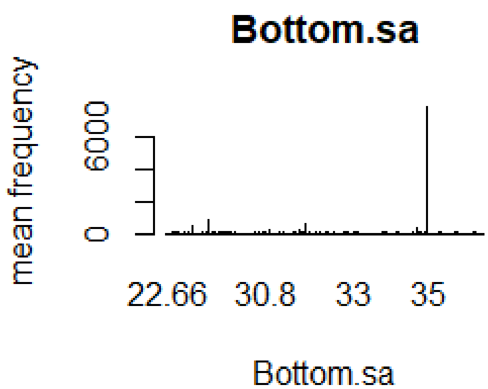


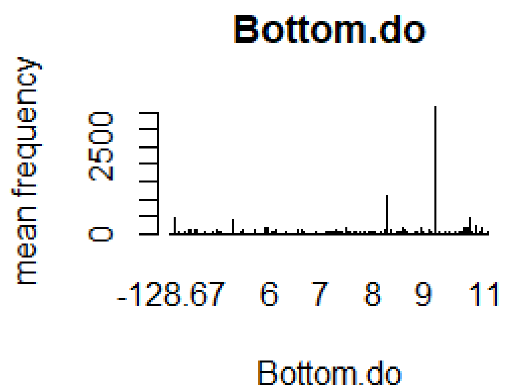

Date
mean frequency
0
1500
2007-09-27
2015-10-13
Date
Time
mean frequency
0
10000
2007-09-27 07:17:00
Time
Areasw
mean frequency
0
10000
11852.3519
25160
27540
Areasw
FREQ
mean frequency
0
10000
0
34
75
150
405
FREQ
Bf.totwght
mean frequency
0
10000
0
0.5
1.57
4.32
Bf.totwght
Region
mean frequency
0
60
01
05
09
13
RI
Region


Depthstrat
mean frequency
0
20
01
02
03
04
Depthstrat
Depth
mean frequency
0
400
3.56616
15.24
27.432
Depth
Stratum_area
mean frequency
0
150
31301303.904
560398294.944
Stratum_area
Lat
mean frequency
0
8000
35.1558685302734
39.63
41.44
Lat

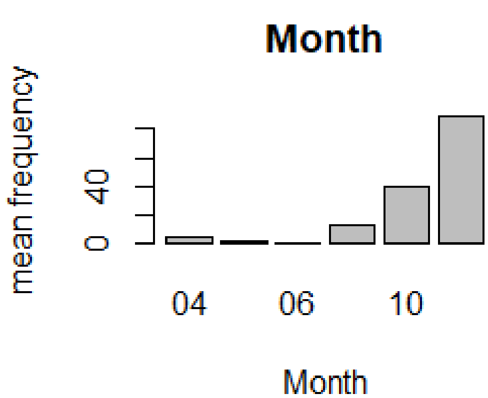









```
## Please manually remove sparsely sampled data!
```



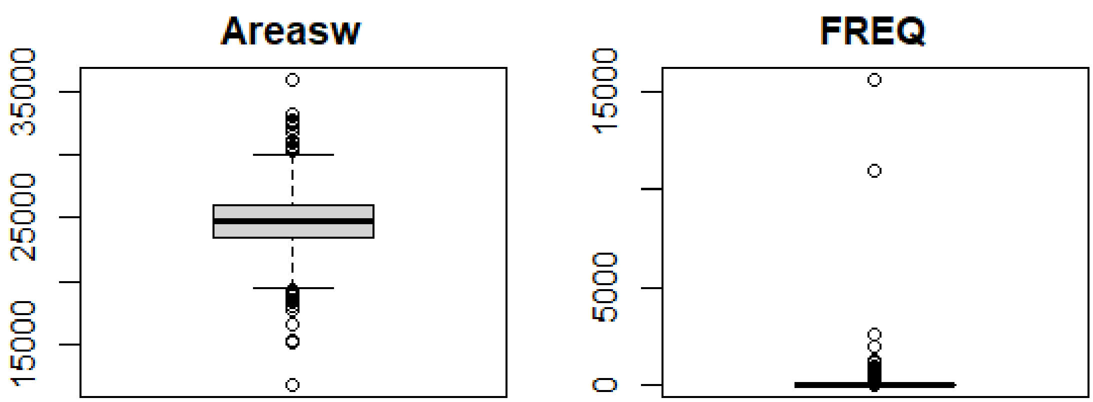
Remove sparse data

```
# remove Spring, September
length1 <- nrow(new_data)
```

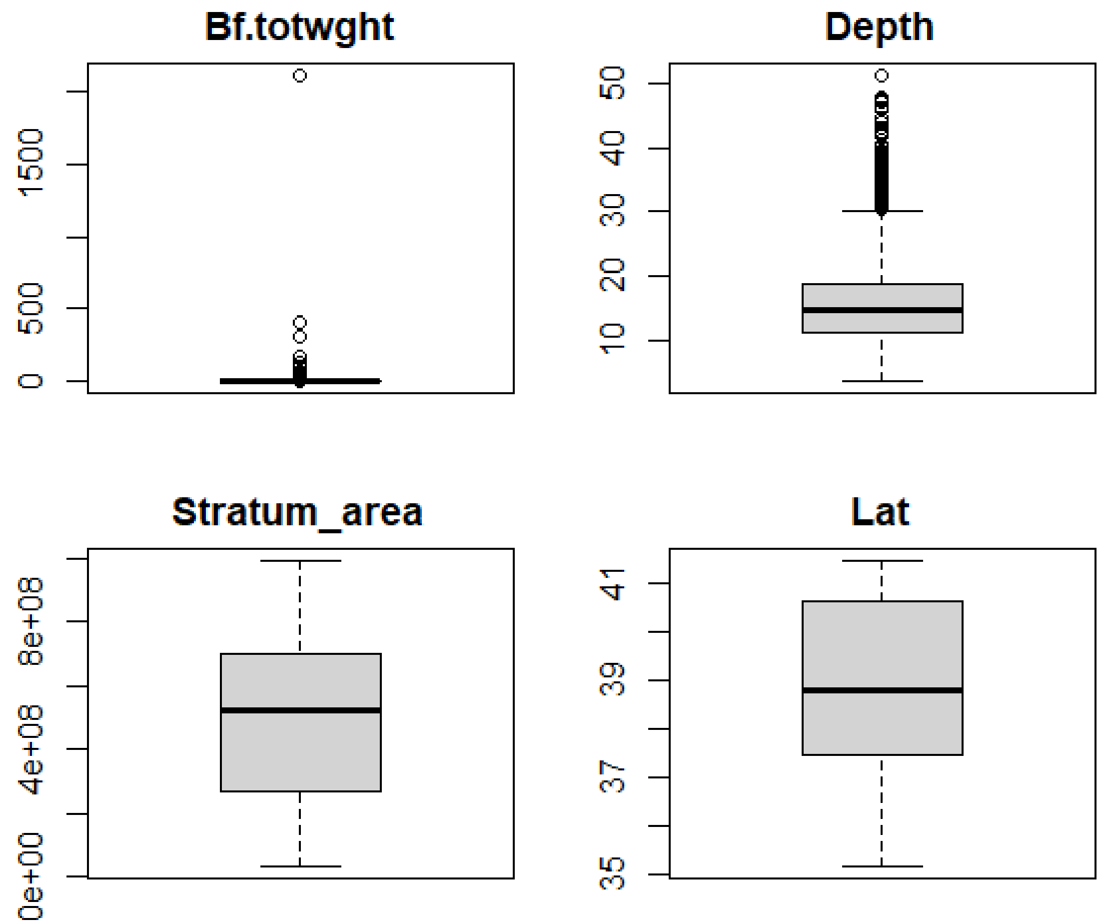
```
new_data <- new_data %>%
  dplyr::filter(Season != "SPRING")
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")

## 1806 rows removed
```

Check for outliers

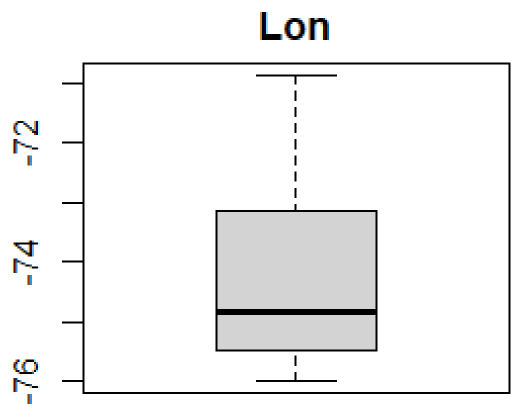
```
plot_outliers(new_data)

## SampleID is not numeric

## State is not numeric

## Year is not numeric

## Season is not numeric

## Date is not numeric

## Time is not numeric
```

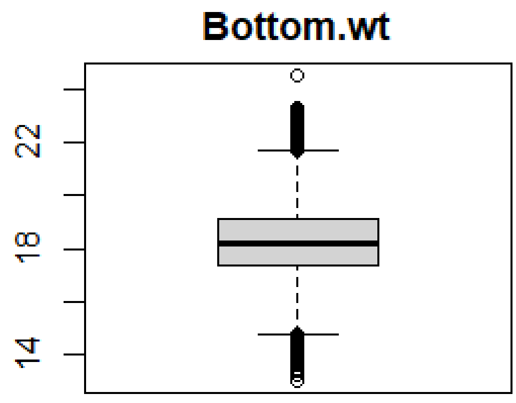


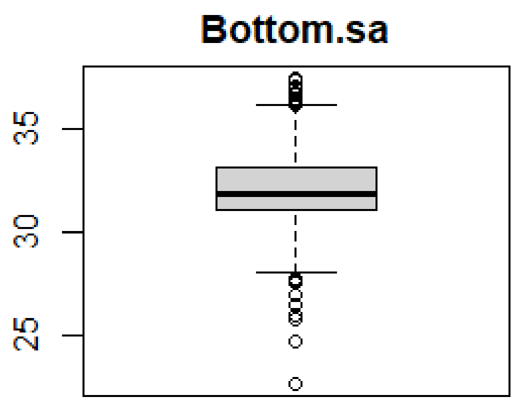

```
## Region is not numeric

## Depthstrat is not numeric
```



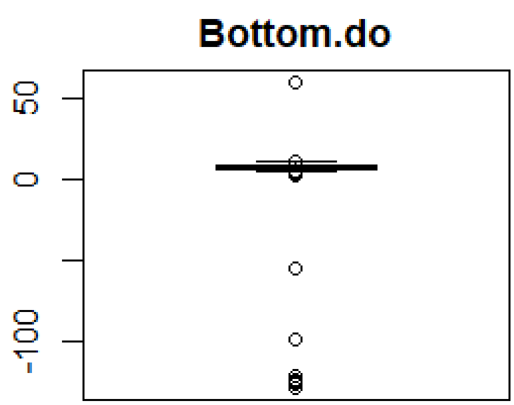
Bf.totwght
0
500
1500
Depth
10
20
30
40
50
Stratum_area
0e+00
4e+08
8e+08
Lat
35
37
39
41








```
## Month is not numeric

## Please manually remove any outliers!
```

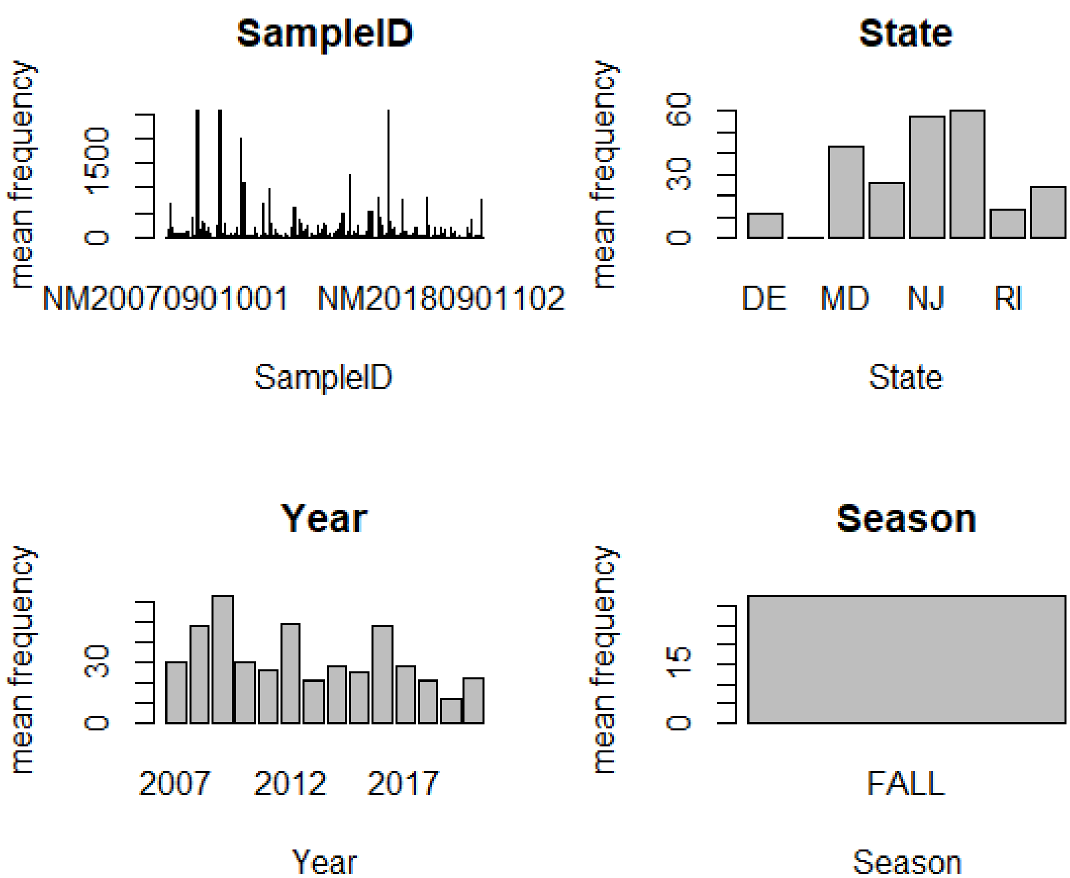


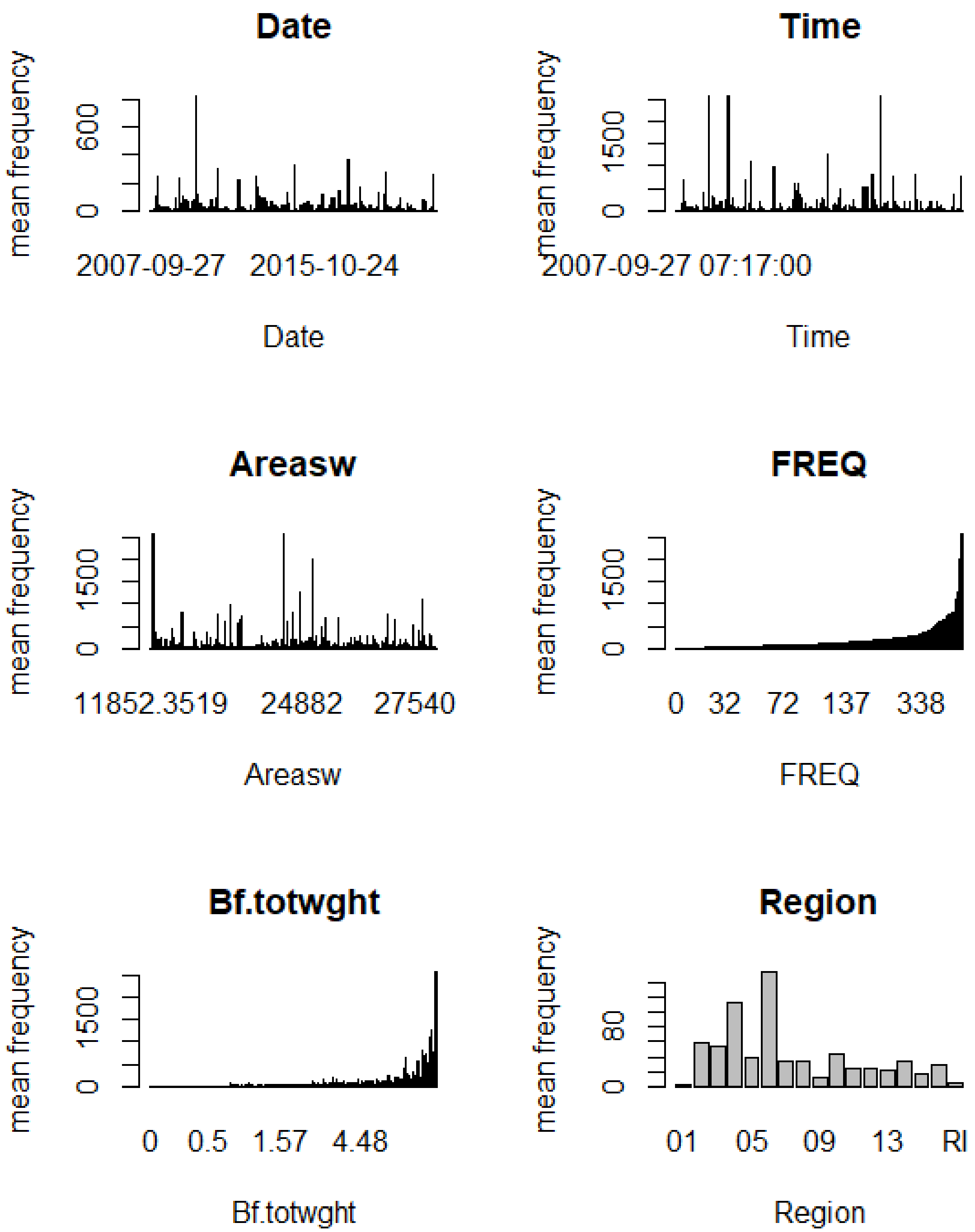
If needed, remove outliers

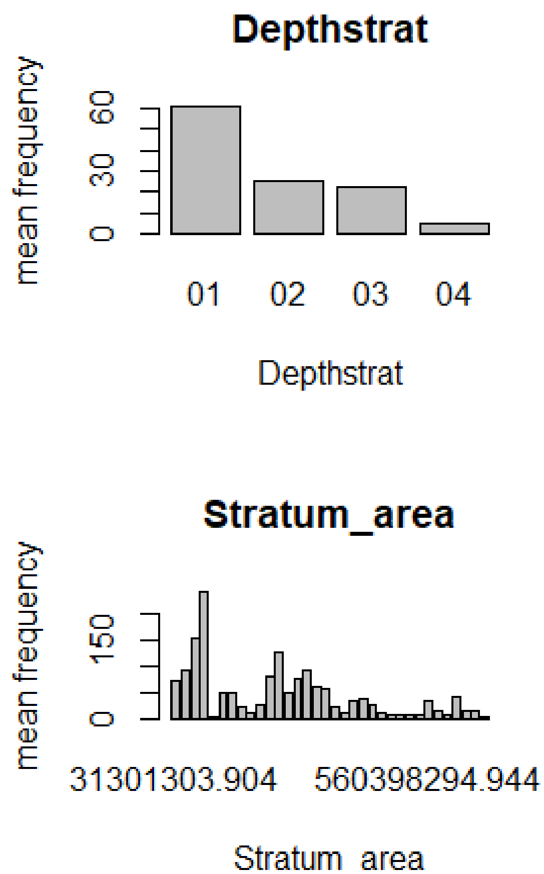
```
# remove incorrect DO values, lower FREQ > 5000
length1 <- nrow(new_data)
new_data <- new_data %>%
  dplyr::filter(Bottom.do < 20,
                Bottom.do > 0#,
               # FREQ < 5000
                ) %>%
  dplyr::mutate(FREQ = ifelse(FREQ > 2557, 2557, FREQ)) # 2557 is the third
largest count
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")

## 12 rows removed
```

Check for sparse sampling again

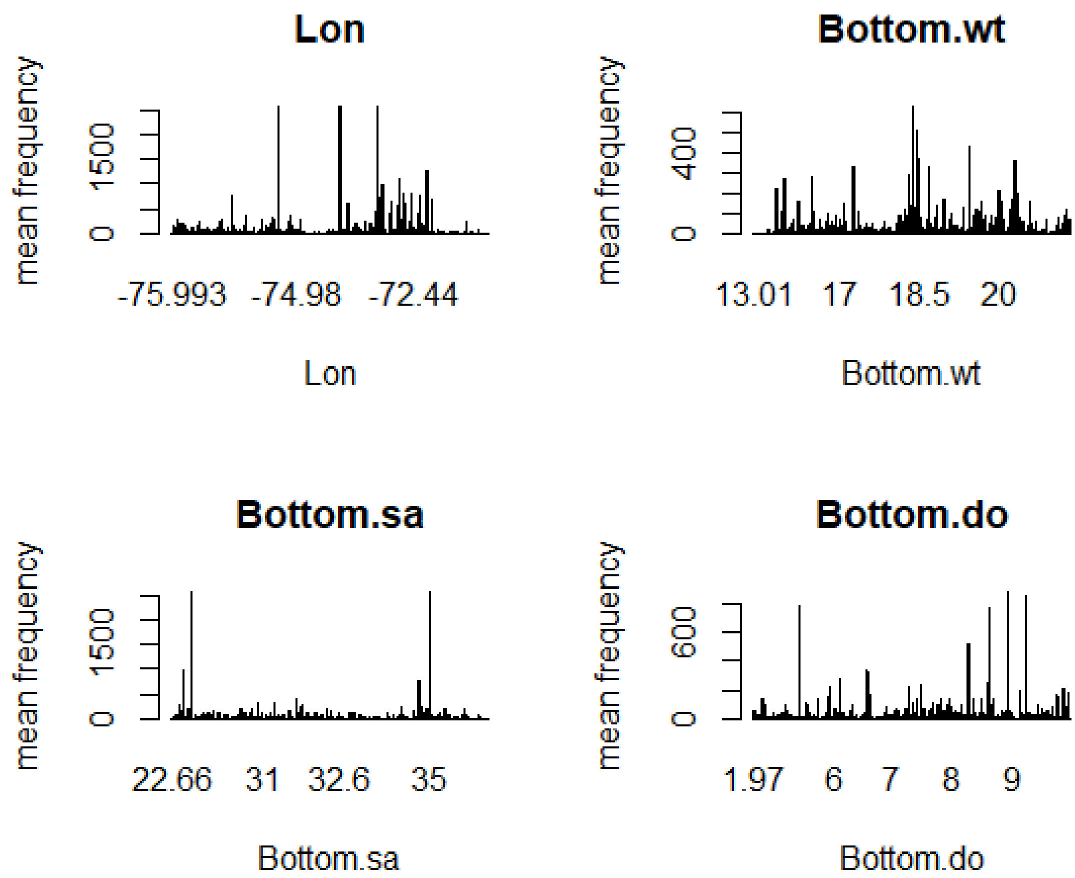
```
plot_obs(new_data,
         col = "FREQ")
```



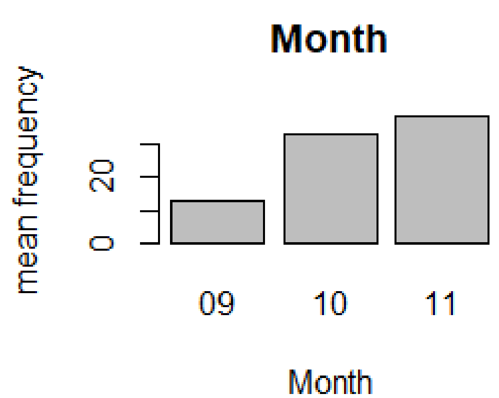

Date
mean frequency
0
600
2007-09-27
2015-10-24
Date
Time
mean frequency
0
1500
2007-09-27 07:17:00
Time
Areasw
mean frequency
0
1500
11852.3519
24882
27540
Areasw
FREQ
mean frequency
0
1500
0
32
72
137
338
FREQ
Bf.totwght
mean frequency
0
1500
0
0.5
1.57
4.48
Bf.totwght
Region
mean frequency
0
80
01
05
09
13
RI
Region



Depthstrat
mean frequency
0
30
60
01
02
03
04
Depthstrat
Depth
mean frequency
0
800
3.6576
15.24
27.432
Depth
Stratum_area
mean frequency
0
150
31301303.904
560398294.944
Stratum_area
Lat
mean frequency
0
1500
35.172
38.12
40.52
Lat




```
## Please manually remove sparsely sampled data!
```

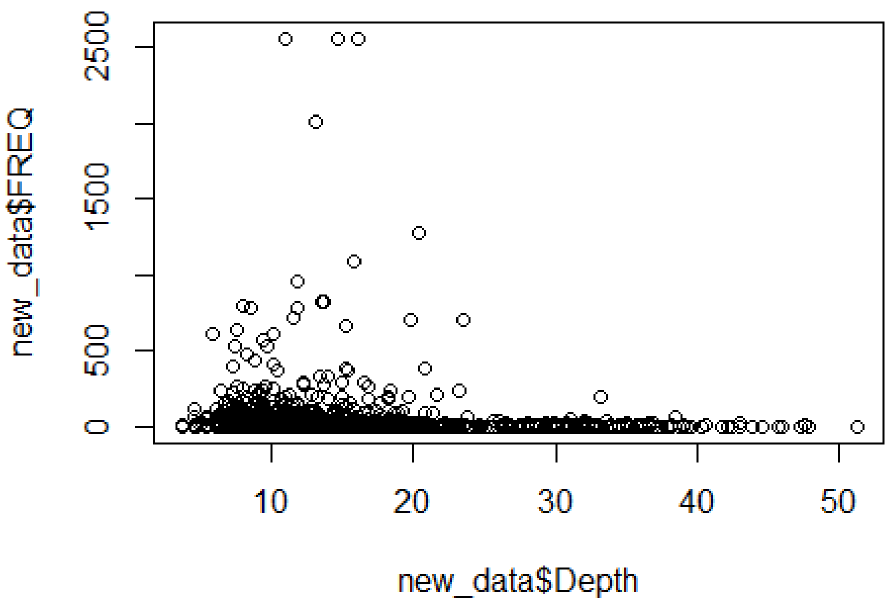


Check again for years with no catch

```
check_missing(new_data, f_col = "FREQ", t_col = "Year")

## Years with no catch: NONE
```

Depth is nonlinearly related to FREQ:

```
plot(new_data$Depth, new_data$FREQ)
```



Change Depth to category:

```
new_data <- new_data %>%
  dplyr::mutate(Depth = ifelse(Depth > 30, "Deep", "Shallow"))
new_data$Depth <- as.factor(new_data$Depth)
```

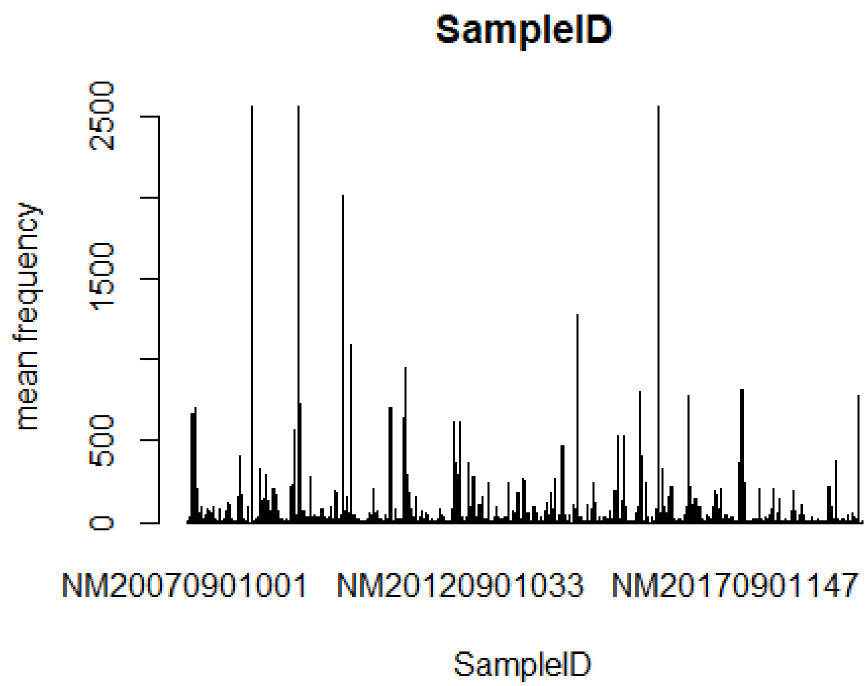
Change continuous variables to z-score

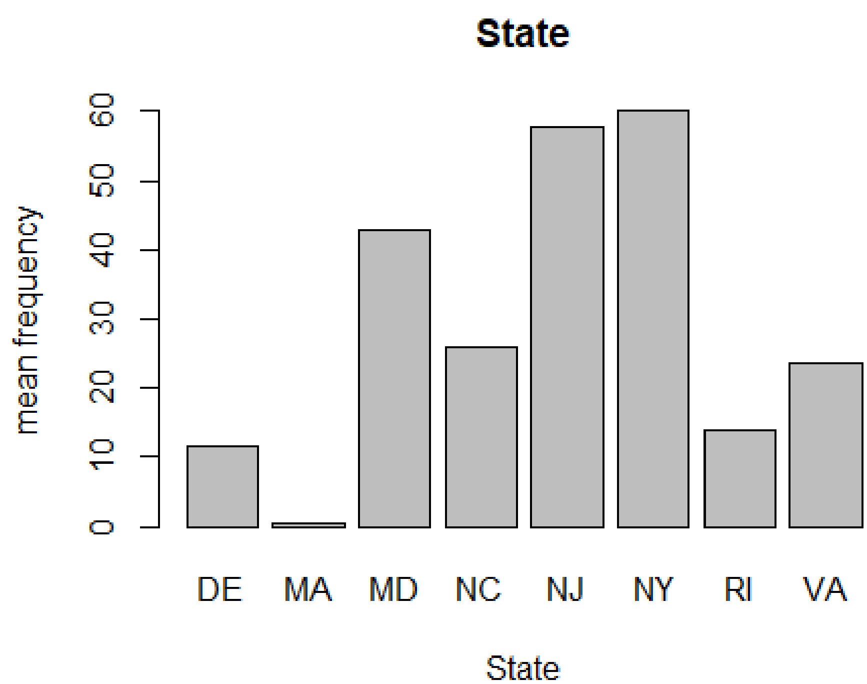
```
new_data2 <- standardize_data(new_data,
                              cols = c("Year", "Month", "Stratum", "Date",
                                       "Time", "Lat", "Lon", "Area",
                                       "Areasw", "Stratum_area",
                                       "SampleID", "State"),
                              f_col = "FREQ")

## SampleID has been changed to factor

## State has been changed to factor

## Year has been changed to factor

## Not sure what to do with Season, leaving as is

## Date has been changed to factor
```

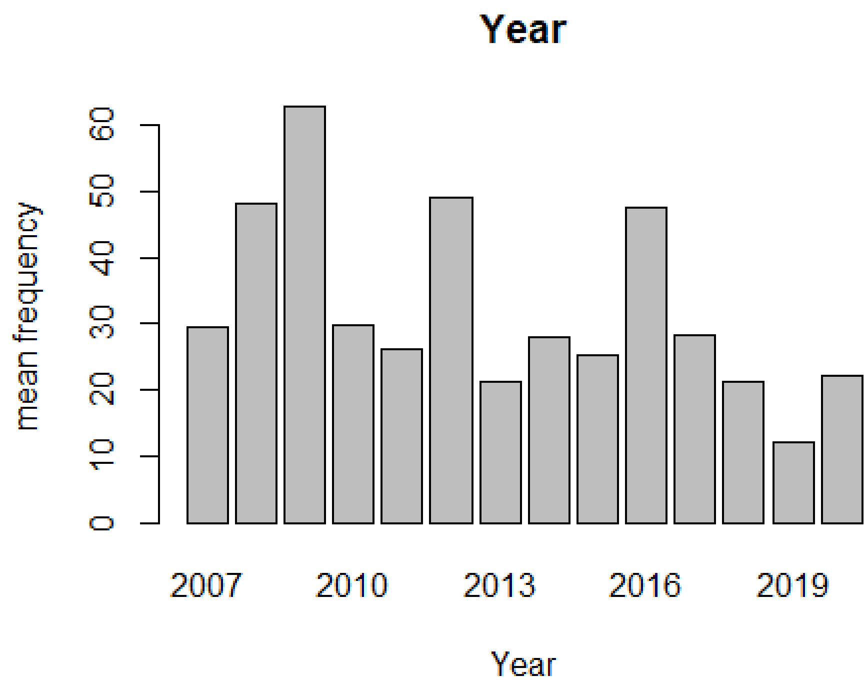
```
## Time has been changed to factor

## Areasw has been changed to factor

## Bf.totwght has been standardized

## Not sure what to do with Region, leaving as is

## Not sure what to do with Depthstrat, leaving as is

## Not sure what to do with Depth, leaving as is

## Stratum_area has been changed to factor

## Lat has been changed to factor

## Lon has been changed to factor

## Bottom.wt has been standardized

## Bottom.sa has been standardized

## Bottom.do has been standardized

## Month has been changed to factor

## Data standardized!

# change lat and lon, Areasw back to numeric
new_data2 <- new_data2 %>%
  dplyr::mutate(Lat = Lat %>%
                  as.character() %>%
                  as.numeric(),
                Lon = Lon %>%
                  as.character() %>%
                  as.numeric(),
                 Areasw = Areasw %>%
                  as.character() %>%
                  as.numeric())
head(new_data2)

## # A tibble: 6 x 19
##   SampleID   State Year  Season Date    Time       Areasw  FREQ Bf.totwght
Region
##   <fct>      <fct> <fct> <chr>  <fct>   <fct>       <dbl> <dbl>      <dbl>
<chr>
## 1 NM2007090~ NY    2007  FALL   2007-0~ 2007-09-~ 27800.     0    -0.0840
01
## 2 NM2007090~ NY    2007  FALL   2007-0~ 2007-09-~ 25146.     6    -0.0707
01
## 3 NM2007090~ NY    2007  FALL   2007-0~ 2007-09-~ 22701.     2     0.00896
02
## 4 NM2007090~ NY    2007  FALL   2007-0~ 2007-09-~ 23059.     6     0.258
```

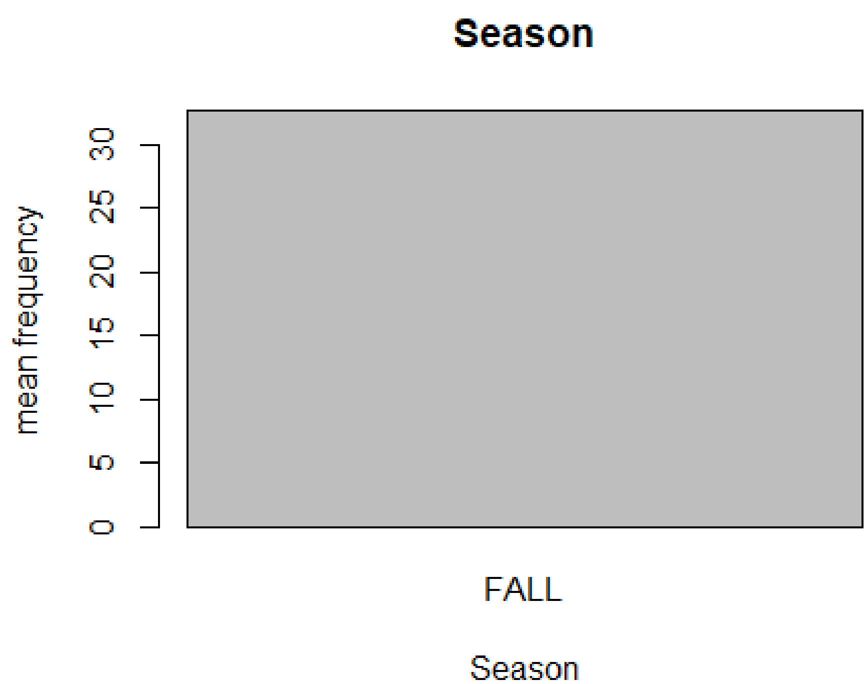
```
02
## 5 NM2007090~ NY    2007  FALL   2007-0~ 2007-09-~ 22826.     0   -0.0840
02
## 6 NM2007090~ NY    2007  FALL   2007-0~ 2007-09-~ 22531.     0   -0.0840
02
## # ... with 9 more variables: Depthstrat <chr>, Depth <fct>, Stratum_area
<fct>,
## #   Lat <dbl>, Lon <dbl>, Bottom.wt <dbl>, Bottom.sa <dbl>, Bottom.do
<dbl>,
## #   Month <fct>
```

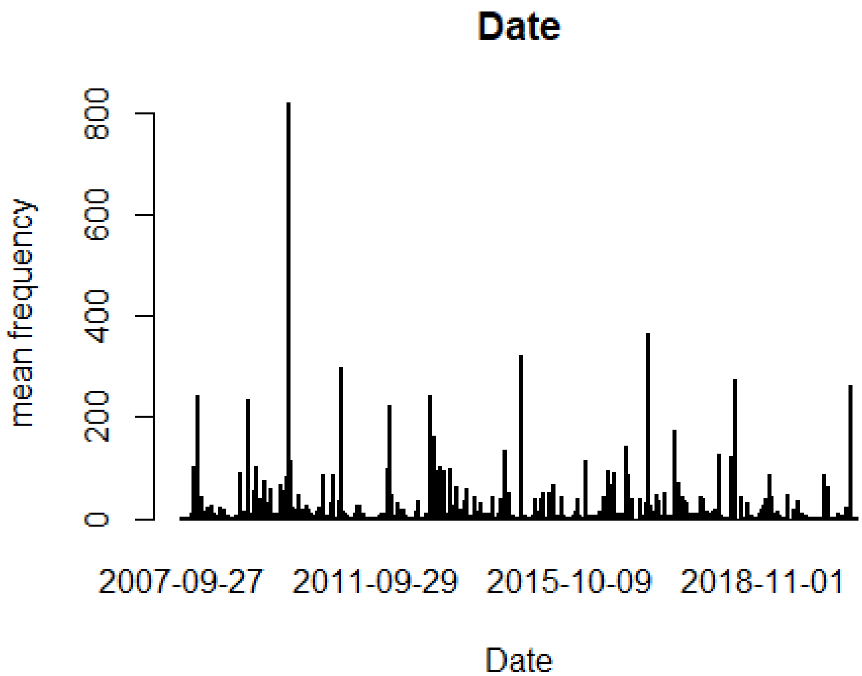
```
plot_obs(new_data2,
         col = "FREQ")
```

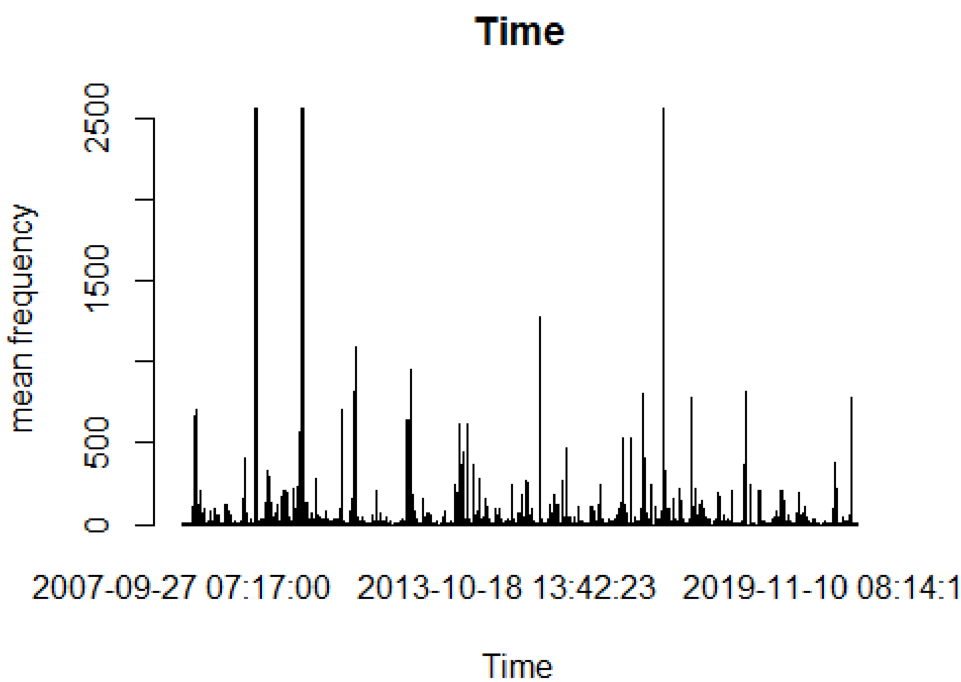

SampleID
mean frequency
0
500
1500
2500
NM20070901001
NM20120901033
NM20170901147
SampleID

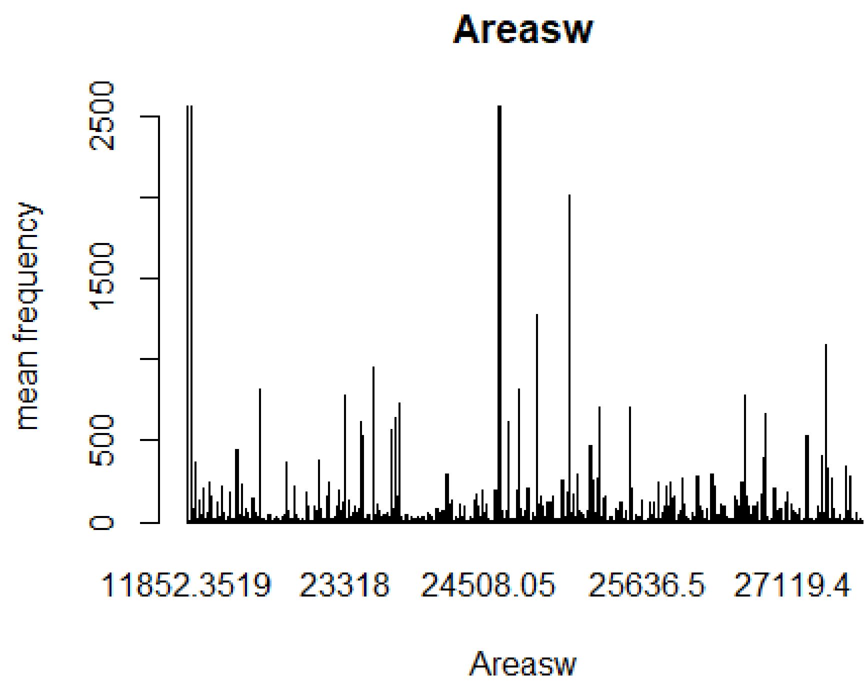

State
mean frequency
0
10
20
30
40
50
60
DE
MA
MD
NC
NJ
NY
RI
VA
State

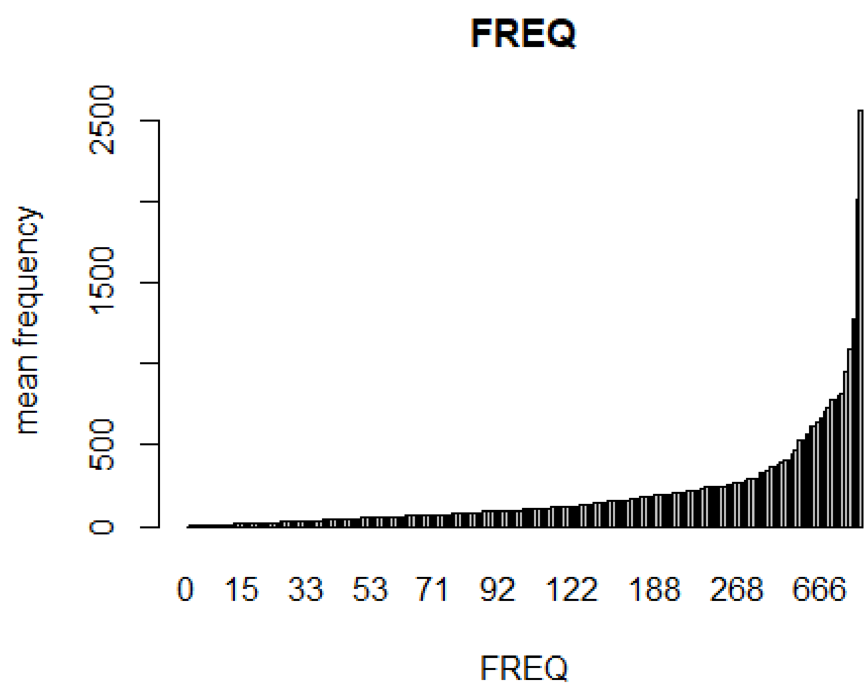

Year
mean frequency
0
10
20
30
40
50
60
2007
2010
2013
2016
2019
Year

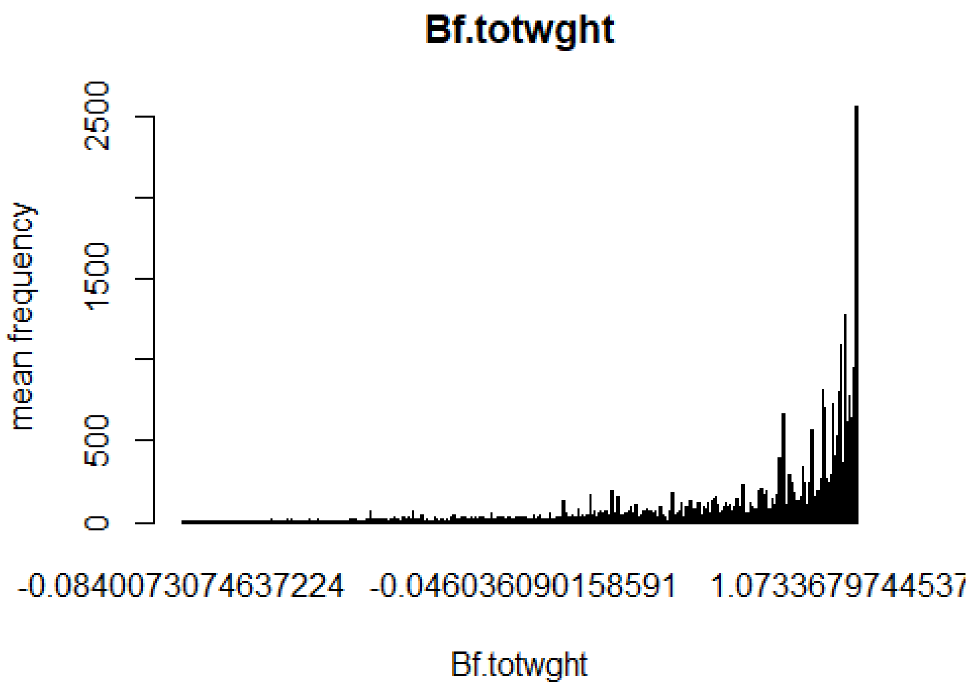

Season
mean frequency
0
5
10
15
20
25
30
FALL
Season

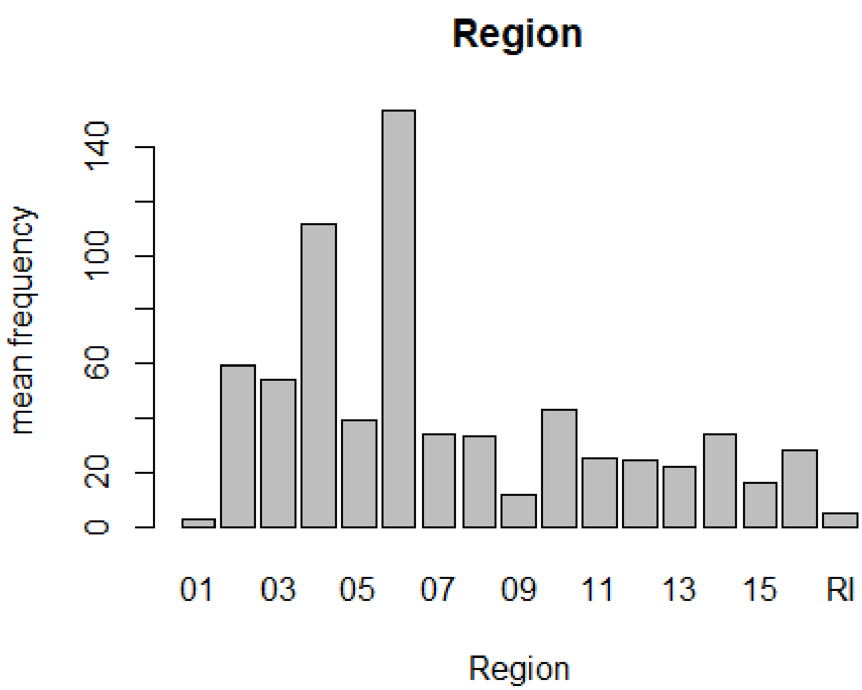

Date
mean frequency
0
200
400
600
800
2007-09-27
2011-09-29
2015-10-09
2018-11-01
Date

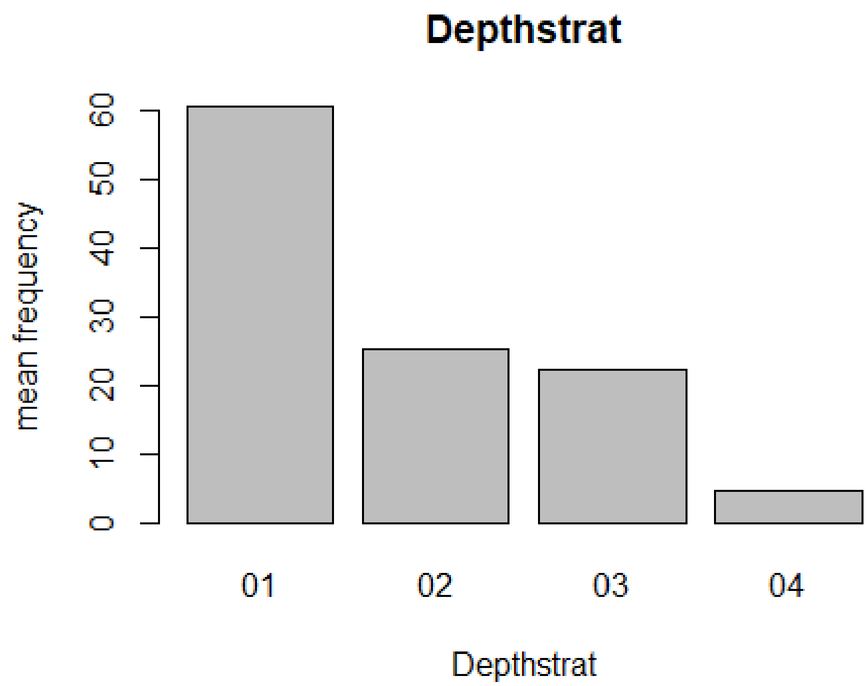

Time
mean frequency
0
500
1500
2500
2007-09-27 07:17:00
2013-10-18 13:42:23
2019-11-10 08:14:1
Time

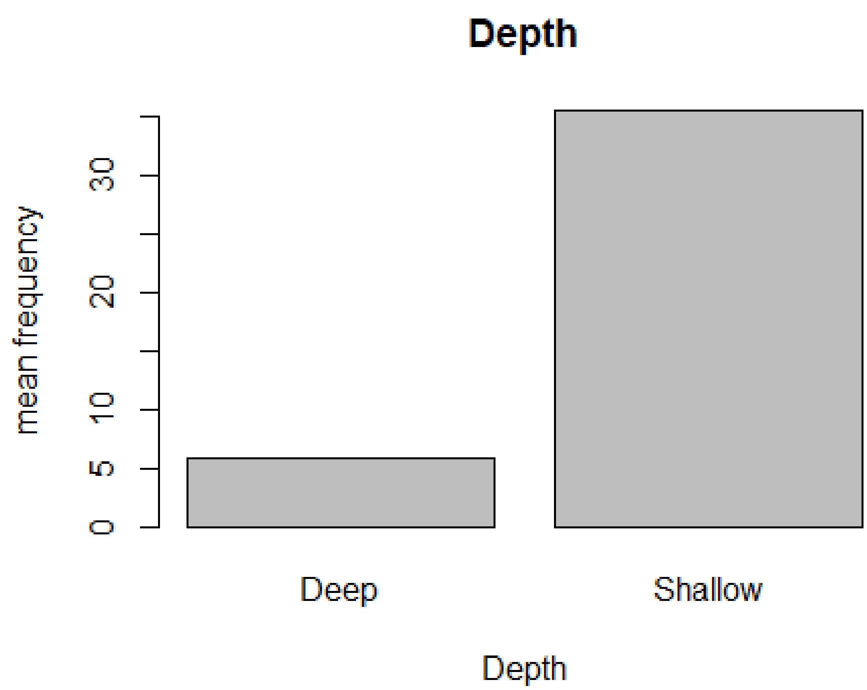

Areasw
mean frequency
0
500
1500
2500
11852.3519
23318
24508.05
25636.5
27119.4
Areasw

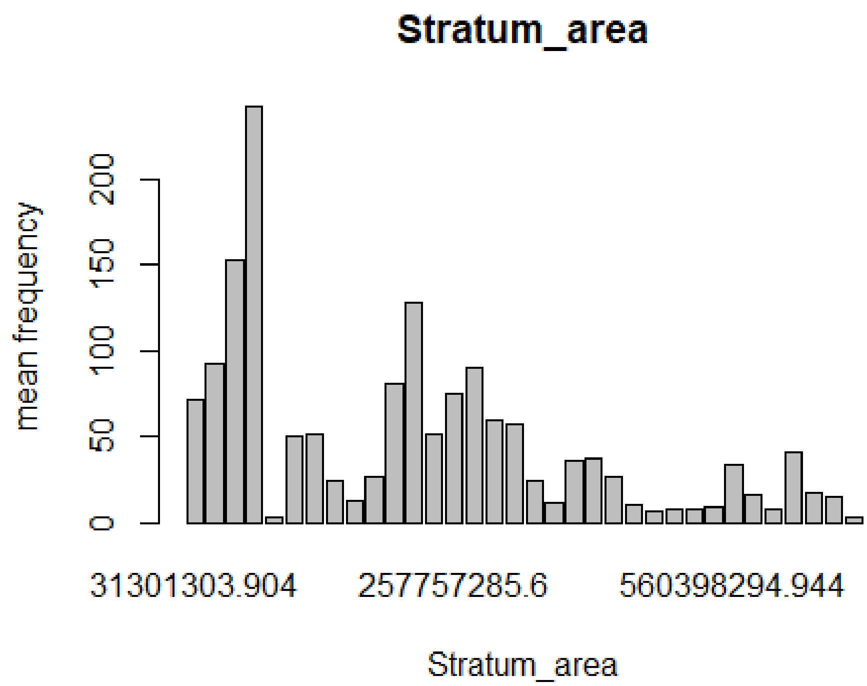

FREQ
mean frequency
0
500
1500
2500
0
15
33
53
71
92
122
188
268
666
FREQ


Bf.totwght
mean frequency
0
500
1500
2500
-0.0840073074637224
-0.046036090158591
1.0733679744537
Bf.totwght

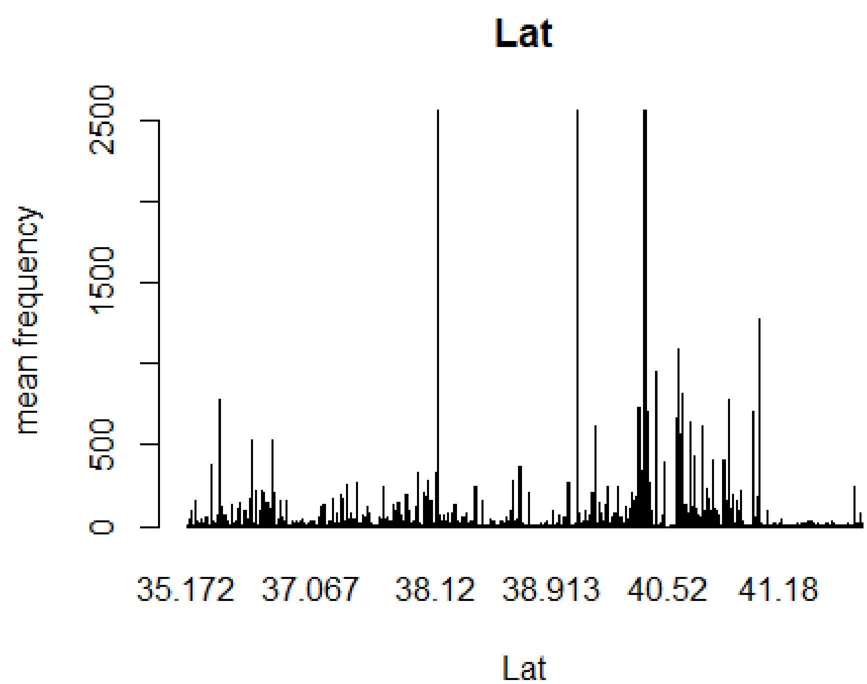

Region
mean frequency
0
20
60
100
140
01
03
05
07
09
11
13
15
RI
Region


Depthstrat
mean frequency
0
10
20
30
40
50
60
01
02
03
04
Depthstrat

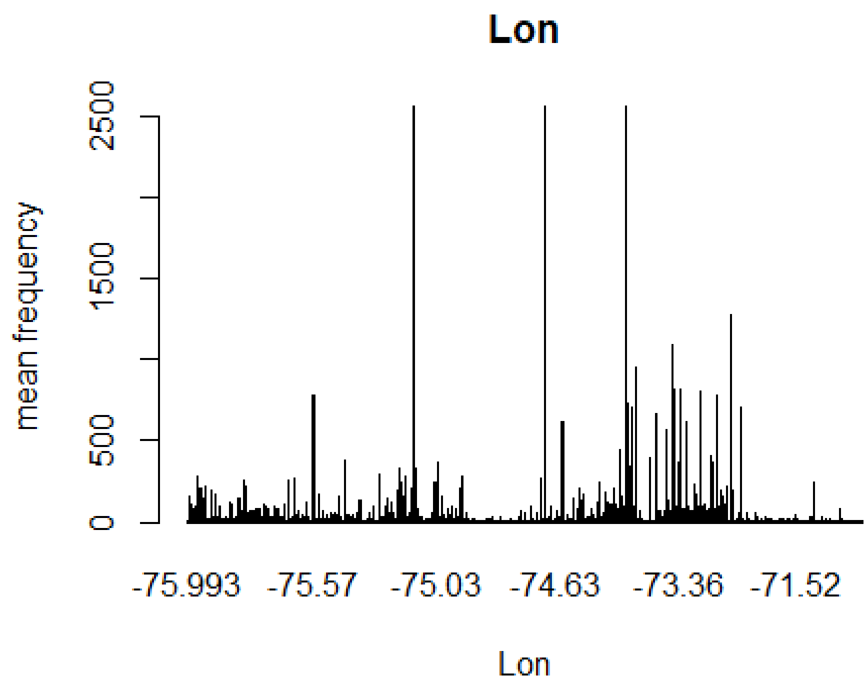

Depth
mean frequency
0
5
10
20
30
Deep
Shallow
Depth



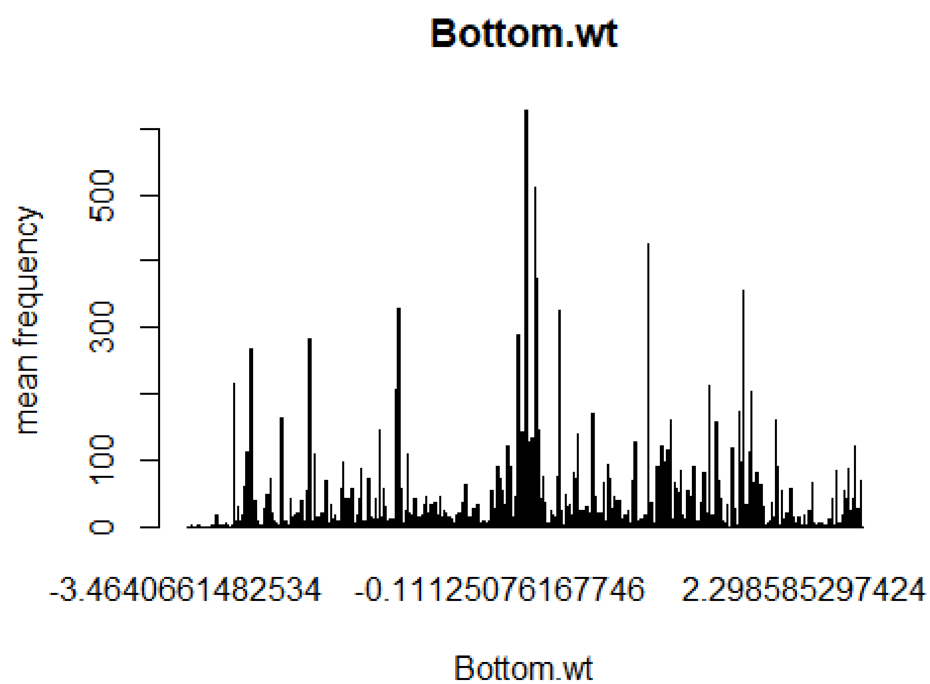
Stratum_area
mean frequency
0
50
100
150
200
31301303.904
257757285.6
560398294.944
Stratum_area



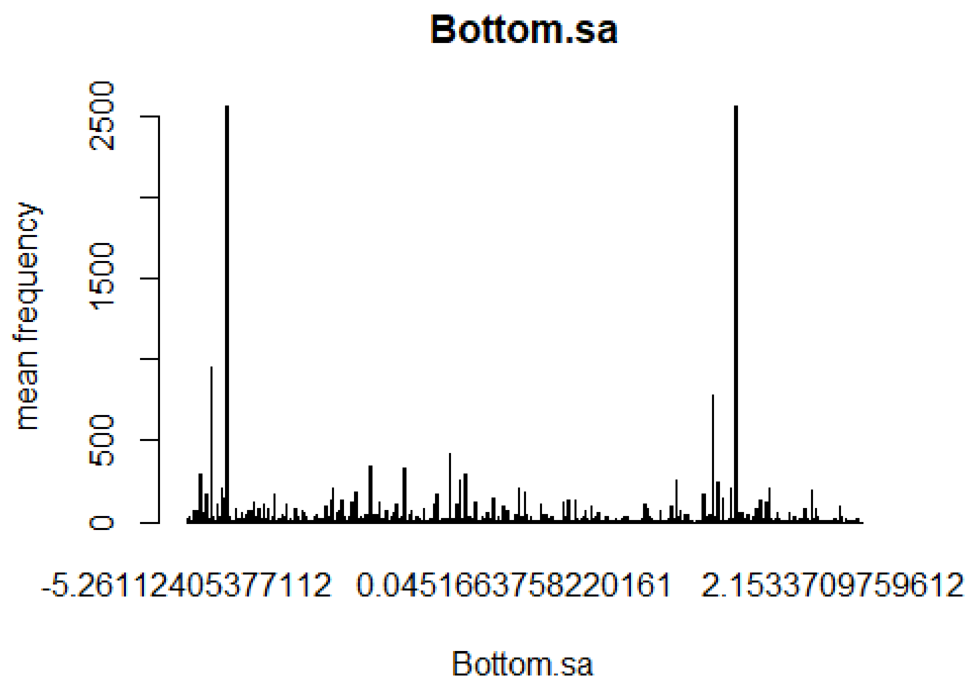
Lat
mean frequency
0
500
1500
2500
35.172
37.067
38.12
38.913
40.52
41.18
Lat

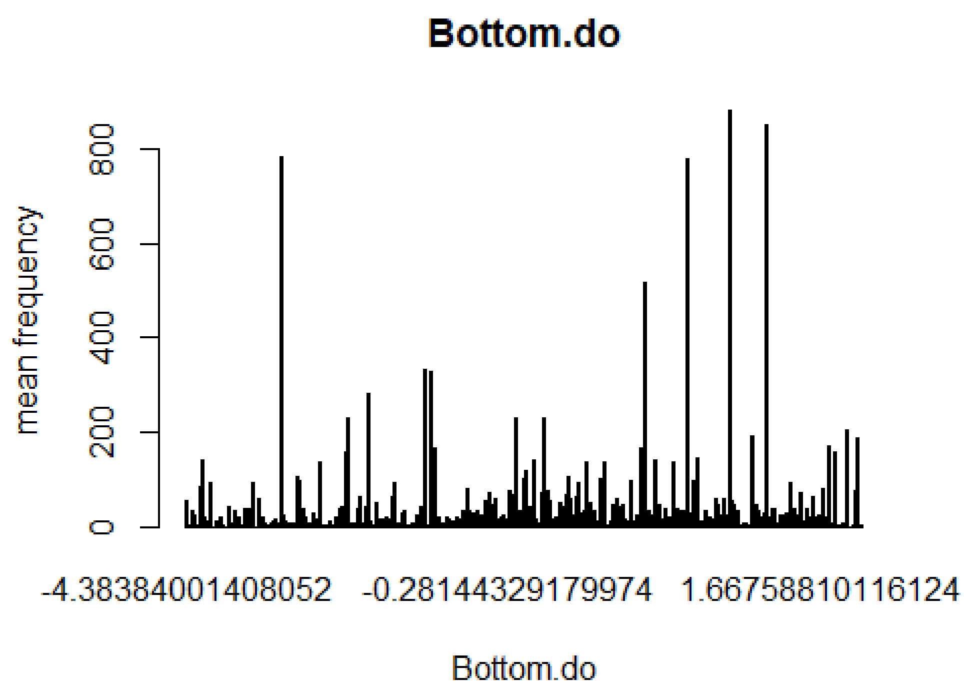


Lon
mean frequency
0
500
1500
2500
-75.993
-75.57
-75.03
-74.63
-73.36
-71.52
Lon

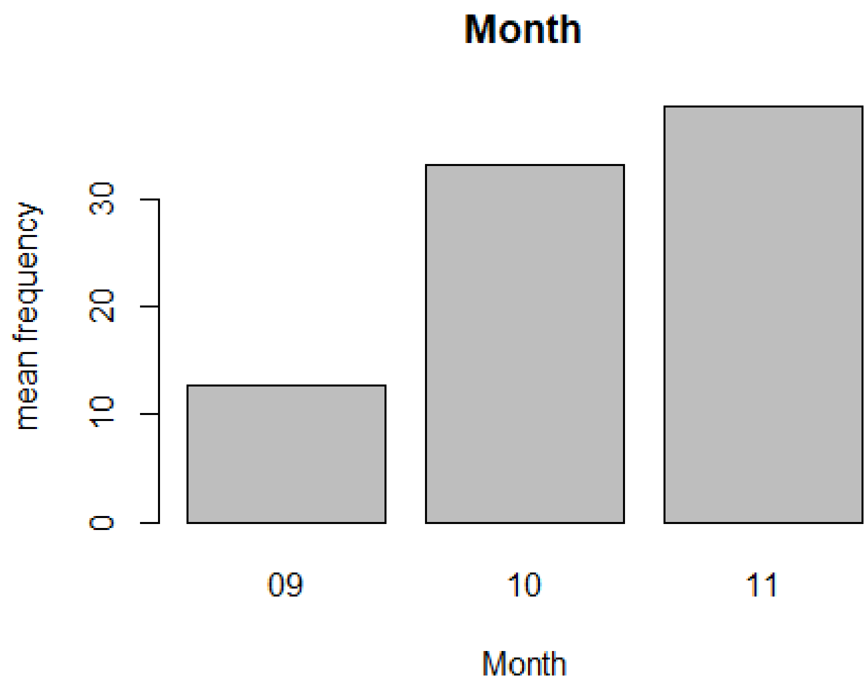


Bottom.wt
mean frequency
0
100
300
500
-3.4640661482534
-0.11125076167746
2.298585297424
Bottom.wt





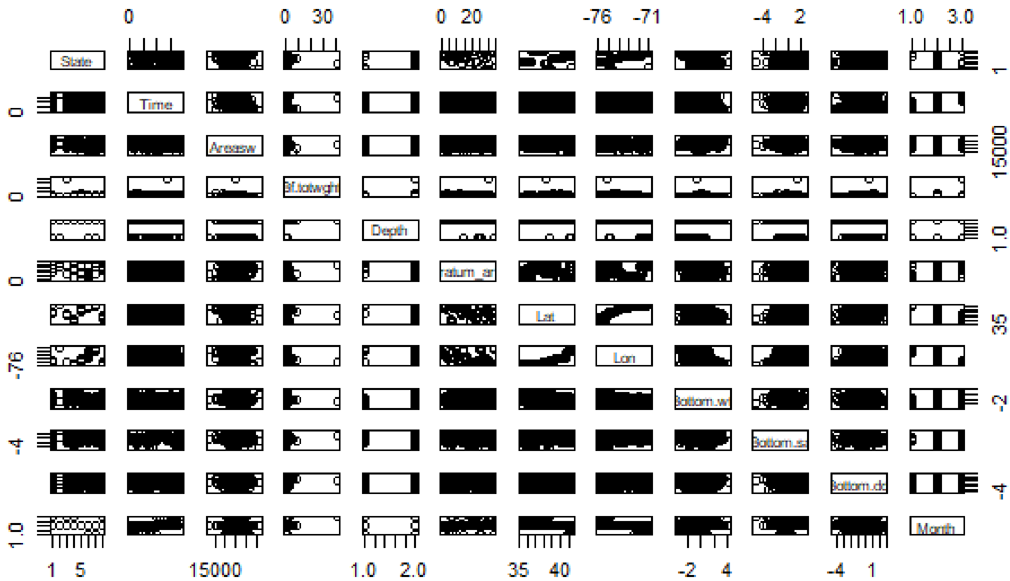

```
## Please manually remove sparsely sampled data!
```



Check colinearity

```
check_colin(new_data2,
            cols = c("Year", "Date", "Season", "Region", "Depthstrat",
                     "SampleID", "FREQ"))

##                State  Time Areasw Bf.totwght Depth Stratum_area   Lat   Lon
## State           1.00 -0.01   0.02      -0.03 -0.19         0.10  0.08  0.17
## Time           -0.01  1.00  -0.24       0.00 -0.01         0.04 -0.10 -0.09
## Areasw          0.02 -0.24   1.00      -0.03 -0.02         0.13 -0.11 -0.08
## Bf.totwght     -0.03  0.00  -0.03       1.00  0.01        -0.02  0.00 -0.01
## Depth          -0.19 -0.01  -0.02       0.01  1.00        -0.25 -0.43 -0.55
## Stratum_area    0.10  0.04   0.13      -0.02 -0.25         1.00 -0.18 -0.15
## Lat             0.08 -0.10  -0.11       0.00 -0.43        -0.18  1.00  0.92
## Lon             0.17 -0.09  -0.08      -0.01 -0.55        -0.15  0.92  1.00
## Bottom.wt      -0.12 -0.26   0.11      -0.01  0.24        -0.05 -0.34 -0.33
## Bottom.sa       0.06  0.31  -0.08       0.02 -0.07         0.03  0.02  0.07
## Bottom.do       0.02 -0.09   0.02       0.03  0.02        -0.05  0.01 -0.01
## Month           0.00  0.48  -0.03       0.05  0.28        -0.01 -0.57 -0.54
##                Bottom.wt Bottom.sa Bottom.do Month
## State              -0.12      0.06      0.02  0.00
## Time               -0.26      0.31     -0.09  0.48
## Areasw              0.11     -0.08      0.02 -0.03
## Bf.totwght         -0.01      0.02      0.03  0.05
## Depth               0.24     -0.07      0.02  0.28
## Stratum_area       -0.05      0.03     -0.05 -0.01
## Lat                -0.34      0.02      0.01 -0.57
```

```
## Lon                -0.33      0.07      -0.01  -0.54
## Bottom.wt           1.00     -0.25       0.04   0.06
## Bottom.sa          -0.25      1.00       0.00   0.19
## Bottom.do           0.04      0.00       1.00   0.00
## Month               0.06      0.19       0.00   1.00

## Lon is correlated with Lat
```



Lon and Lat are correlated - keep only Lat.

Check variance inflation factor - Remove covariates with GVIF > 3, starting with highest value, until all GVIF < 3

```
check_vif(model = "FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.sa +
Bottom.do + Month",
          data = new_data2)

## # A tibble: 7 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Bottom.do  1.06     1              1.03
## 2 Depth      1.32     1              1.15
## 3 Bottom.wt  2.01     1              1.42
## 4 Lat        2.11     1              1.45
## 5 Month      3.13     2              1.33
## 6 Bottom.sa  5.00     1              2.24
## 7 Year      12.8     13              1.10

## Variance inflation factor too high! Please remove a covariate.
```

```
# remove Bottom.sa

check_vif(model = "FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do +
Month",
          data = new_data2)

## # A tibble: 6 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Bottom.do  1.06     1              1.03
## 2 Depth      1.31     1              1.14
## 3 Bottom.wt  2.00     1              1.41
## 4 Lat        2.11     1              1.45
## 5 Year       2.89    13              1.04
## 6 Month      3.11     2              1.33

## Variance inflation factor too high! Please remove a covariate.

# remove Month

check_vif(model = "FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do",
          data = new_data2)

## # A tibble: 5 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Bottom.do  1.05     1              1.03
## 2 Depth      1.25     1              1.12
## 3 Lat        1.44     1              1.20
## 4 Year       1.66    13              1.02
## 5 Bottom.wt  1.74     1              1.32

## Passed!
```

Create effort for offset

```
new_data2$lnSweptArea <- log(new_data2$Areasw)
```

Step 2: Model selection without stratum

Compare negative binom, zero-infl neg binom, zero-altered neg binom

Comparison 1: Full models

```
compare_models(data = new_data2,
               nb_model = "FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do
+ offset(lnSweptArea)",

               zic_model = "FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do
+ offset(lnSweptArea)",
               zi_model = "~ Year + Depth + Lat + Bottom.wt + Bottom.do",
```

```
                 zac_model = "FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do
+ offset(lnSweptArea)",
                 za_model = "~ Year + Depth + Lat + Bottom.wt + Bottom.do",

                 r_dat = new_data2$FREQ) %>%
  try()
```

High percent positive tows, so theoretically the negative binomial model should be ok, but the dispersal is bad. Dispersal was bad on the zero models too.

Zero-inflated did not work.

Comparison 2: Try removing Lat from zero-inflated model

```
compare_models(data = new_data2,
                 models = "zi",

                 zic_model = "FREQ ~ Year + Depth + Bottom.wt + Bottom.do +
offset(lnSweptArea)",
                 zi_model = "~ Year + Depth + Bottom.wt + Bottom.do",

                 r_dat = new_data2$FREQ) %>%
  try()
```

Comparison 3: Null count for zero-inflated model

```
compare_models(data = new_data2,
                 models = "zi",

                 zic_model = "FREQ ~ offset(lnSweptArea)",
                 zi_model = "~ Year + Depth + Lat + Bottom.wt + Bottom.do",

                 r_dat = new_data2$FREQ) %>%
  try()
```

Comparison 4: Null binomial for zero-inflated model

```
compare_models(data = new_data2,
                 models = "zi",

                 zic_model = "FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do
+ offset(lnSweptArea)",
                 zi_model = "~ 1",

                 r_dat = new_data2$FREQ) %>%
  try()
```

Model selection for negative binomial model

```
model <- glmmTMB::glmmTMB(FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do +
offset(lnSweptArea),
            data = new_data2,
            family = nbinom2
        )
```

```
null <- glmmTMB::glmmTMB(FREQ ~ Year + offset(lnSweptArea),
          data = new_data2,
          family = nbinom2
        )

select_model(model = model, null_model = null)
```

Model selection for zero-inflated model

```
model <- glmmTMB::glmmTMB(FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do +
offset(lnSweptArea),
          ziformula = ~ 1,
          data = new_data2,
          family = nbinom2
        )

null <- glmmTMB::glmmTMB(FREQ ~ Year + offset(lnSweptArea),
          ziformula = ~ 1,
          data = new_data2,
          family = nbinom2
        )

select_model(model = model, null_model = null)
```

Model selection for zero-adjusted model

```
model <- glmmTMB::glmmTMB(FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do +
offset(lnSweptArea),
          ziformula = ~ Year + Depth + Lat + Bottom.wt + Bottom.do,
          data = new_data2,
          family = truncated_nbinom2(link = "log")
        )

null <- glmmTMB::glmmTMB(FREQ ~ offset(lnSweptArea),
          ziformula = ~ Year + Depth + Lat + Bottom.wt + Bottom.do,
          data = new_data2,
          family = truncated_nbinom2(link = "log")
        )

select_model(model = model, null_model = null) %>%
  try()
```

Final comparison

```
compare_models(data = new_data2,
               nb_model = "FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do
+ offset(lnSweptArea)",

               zic_model = "FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do
+ offset(lnSweptArea)",
               zi_model = "~ 1",
```

```
               zac_model = "FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do
+ offset(lnSweptArea)",
               za_model = "~ Year + Depth + Lat + Bottom.wt + Bottom.do",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 67.2

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do + offset(lnSweptArea)
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  13017.7  13122.9  -6489.8  12979.7     1860
##
##
## Dispersion parameter for nbinom2 family (): 0.21
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -15.53956    1.48548 -10.461  < 2e-16 ***
## Year2008      0.89628    0.26295   3.409 0.000653 ***
## Year2009      1.16969    0.27171   4.305 1.67e-05 ***
## Year2010      0.22521    0.27048   0.833 0.405056
## Year2011      0.14854    0.26284   0.565 0.571976
## Year2012      0.88254    0.28351   3.113 0.001852 **
## Year2013      0.50078    0.30476   1.643 0.100340
## Year2014      0.22372    0.27629   0.810 0.418100
## Year2015      0.64616    0.29386   2.199 0.027888 *
## Year2016      1.10104    0.29766   3.699 0.000216 ***
## Year2017      0.31959    0.26803   1.192 0.233124
## Year2018      0.49642    0.28029   1.771 0.076550 .
## Year2019     -0.04615    0.29465  -0.157 0.875530
## Year2020      0.76086    0.41286   1.843 0.065341 .
## DepthShallow  1.90913    0.19976   9.557  < 2e-16 ***
## Lat           0.16707    0.03507   4.764 1.90e-06 ***
## Bottom.wt     0.23451    0.07942   2.953 0.003149 **
## Bottom.do     0.18224    0.05878   3.100 0.001933 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.812

## Zero-inflated negative binomial
```

```
## model summary:

##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do + offset(lnSweptArea)
## Zero inflation:        ~1
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  13019.7  13130.4  -6489.8  12979.7     1859
##
##
## Dispersion parameter for nbinom2 family (): 0.21
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -15.53952    1.48548 -10.461  < 2e-16 ***
## Year2008       0.89628    0.26295   3.409 0.000653 ***
## Year2009       1.16967    0.27170   4.305 1.67e-05 ***
## Year2010       0.22521    0.27048   0.833 0.405056
## Year2011       0.14853    0.26284   0.565 0.571997
## Year2012       0.88253    0.28351   3.113 0.001853 **
## Year2013       0.50078    0.30476   1.643 0.100341
## Year2014       0.22371    0.27629   0.810 0.418125
## Year2015       0.64616    0.29386   2.199 0.027889 *
## Year2016       1.10103    0.29766   3.699 0.000216 ***
## Year2017       0.31958    0.26803   1.192 0.233137
## Year2018       0.49641    0.28029   1.771 0.076553 .
## Year2019      -0.04616    0.29465  -0.157 0.875519
## Year2020       0.76087    0.41286   1.843 0.065338 .
## DepthShallow   1.90912    0.19976   9.557  < 2e-16 ***
## Lat            0.16707    0.03507   4.764 1.90e-06 ***
## Bottom.wt      0.23451    0.07942   2.953 0.003149 **
## Bottom.do      0.18224    0.05878   3.100 0.001933 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -19.03    1070.35  -0.018    0.986

## Dispersal: 2.813

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:
## FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do + offset(lnSweptArea)
## Zero inflation:        ~Year + Depth + Lat + Bottom.wt + Bottom.do
```

```
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  12875.4  13080.3  -6400.7  12801.4     1842
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.00982
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -21.73490    3.02157  -7.193 6.33e-13 ***
## Year2008       1.07157    0.37079   2.890 0.003853 **
## Year2009       1.21031    0.36063   3.356 0.000790 ***
## Year2010       0.15099    0.35978   0.420 0.674728
## Year2011       0.07701    0.35305   0.218 0.827324
## Year2012       0.87017    0.37805   2.302 0.021348 *
## Year2013       1.09468    0.43407   2.522 0.011673 *
## Year2014       0.41585    0.38627   1.077 0.281666
## Year2015       0.88117    0.40606   2.170 0.030005 *
## Year2016       1.45849    0.41179   3.542 0.000397 ***
## Year2017       0.41752    0.36938   1.130 0.258336
## Year2018       1.33973    0.43129   3.106 0.001894 **
## Year2019       0.27224    0.40769   0.668 0.504284
## Year2020       1.19123    0.61164   1.948 0.051461 .
## DepthShallow   2.85382    0.26404  10.808  < 2e-16 ***
## Lat            0.23278    0.05237   4.445 8.80e-06 ***
## Bottom.wt      0.35225    0.11590   3.039 0.002373 **
## Bottom.do      0.18400    0.08547   2.153 0.031334 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.423106   1.429377  -0.996   0.3194
## Year2008     -0.088243   0.260864  -0.338   0.7352
## Year2009     -0.741396   0.296000  -2.505   0.0123 *
## Year2010     -0.298111   0.276418  -1.078   0.2808
## Year2011     -0.131782   0.263069  -0.501   0.6164
## Year2012     -0.675512   0.306054  -2.207   0.0273 *
## Year2013      0.205063   0.279842   0.733   0.4637
## Year2014      0.232315   0.268279   0.866   0.3865
## Year2015     -0.069943   0.282320  -0.248   0.8043
## Year2016     -0.051552   0.283952  -0.182   0.8559
## Year2017     -0.009516   0.259580  -0.037   0.9708
## Year2018      0.494496   0.258406   1.914   0.0557 .
## Year2019      0.439292   0.279183   1.573   0.1156
## Year2020      0.314070   0.381138   0.824   0.4099
## DepthShallow  0.301655   0.193627   1.558   0.1193
## Lat           0.011520   0.033858   0.340   0.7337
## Bottom.wt    -0.017851   0.066031  -0.270   0.7869
```

```
## Bottom.do    -0.091314   0.051113  -1.786   0.0740 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.259

##           dAIC  df
## zero_adj    0.0 37
## neg_binom 142.2 19
## zero_infl 144.2 20
```

Zero-adjusted is the winner! But all models have poor dispersal…

Step 3: Model selection with stratum

Add stratum factor

```
new_data2 <- new_data2 %>%
  dplyr::mutate(Stratum = as.factor(Stratum_area))
```

Check variance inflation factor - Remove covariates with GVIF > 3, starting with highest value, until all GVIF < 3

```
check_vif(model = "FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.sa +
Bottom.do + Month + Stratum",
          data = new_data2)

## # A tibble: 8 x 4
##   Covariate    GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>       <dbl> <dbl>             <dbl>
## 1 Bottom.do    1.09     1              1.05
## 2 Bottom.wt    2.37     1              1.54
## 3 Depth        3.09     1              1.76
## 4 Month        4.16     2              1.43
## 5 Bottom.sa    6.35     1              2.52
## 6 Year        18.0     13              1.12
## 7 Lat        209.       1             14.4
## 8 Stratum   1074.      33              1.11

## Variance inflation factor too high! Please remove a covariate.

check_vif(model = "FREQ ~ Year + Depth + Bottom.wt + Bottom.sa + Bottom.do +
Month + Stratum",
          data = new_data2)

## # A tibble: 7 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Bottom.do  1.09     1              1.05
## 2 Bottom.wt  2.34     1              1.53
## 3 Depth      3.08     1              1.75
## 4 Month      4.14     2              1.43
```

```
## 5 Bottom.sa  6.35     1               2.52
## 6 Stratum   10.9     33               1.04
## 7 Year      17.7     13               1.12

## Variance inflation factor too high! Please remove a covariate.

check_vif(model = "FREQ ~ Year + Depth + Bottom.wt + Bottom.do + Month +
Stratum",
          data = new_data2)

## # A tibble: 6 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Bottom.do  1.09     1              1.04
## 2 Bottom.wt  2.33     1              1.53
## 3 Depth      3.06     1              1.75
## 4 Year       3.32    13              1.05
## 5 Month      4.06     2              1.42
## 6 Stratum    8.56    33              1.03

## Variance inflation factor too high! Please remove a covariate.

check_vif(model = "FREQ ~ Year + Depth + Bottom.wt + Bottom.do + Stratum",
          data = new_data2)

## # A tibble: 5 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Bottom.do  1.09     1              1.04
## 2 Year       1.86    13              1.02
## 3 Bottom.wt  2.03     1              1.43
## 4 Depth      3.05     1              1.75
## 5 Stratum    4.47    33              1.02

## Variance inflation factor too high! Please remove a covariate.

check_vif(model = "FREQ ~ Year + Bottom.wt + Bottom.do + Stratum",
          data = new_data2)

## # A tibble: 4 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Bottom.do  1.09     1              1.04
## 2 Stratum    1.60    33              1.01
## 3 Year       1.84    13              1.02
## 4 Bottom.wt  2.03     1              1.43

## Passed!
```

Compare negative binom, zero-infl neg binom, zero-altered neg binom

Comparison 1: Full models

```
compare_models(data = new_data2,
               nb_model = "FREQ ~ Year + Bottom.wt + Bottom.do + Stratum + 
offset(lnSweptArea)",

               zic_model = "FREQ ~ Year + Bottom.wt + Bottom.do + Stratum + 
offset(lnSweptArea)",
               zi_model = "~ Year + Bottom.wt + Bottom.do + Stratum",

               zac_model = "FREQ ~ Year + Bottom.wt + Bottom.do + Stratum + 
offset(lnSweptArea)",
               za_model = "~ Year + Bottom.wt + Bottom.do + Stratum",

               r_dat = new_data2$FREQ) %>%
  try()
```

High percent positive tows, so theoretically the negative binomial model should be ok, but the dispersal is bad. Dispersal was bad on the zero models too.

Zero-inflated did not work.

Comparison 2: Try removing Bottom.do from zero-inflated model

```
compare_models(data = new_data2,
               models = "zi",

               zic_model = "FREQ ~ Year + Bottom.wt + Stratum + 
offset(lnSweptArea)",
               zi_model = "~ Year + Bottom.wt + Stratum",

               r_dat = new_data2$FREQ) %>%
  try()
```

Comparison 3: Try removing Bottom.wt from zero-inflated model

```
compare_models(data = new_data2,
               models = "zi",

               zic_model = "FREQ ~ Year + Stratum + offset(lnSweptArea)",
               zi_model = "~ Year + Stratum",

               r_dat = new_data2$FREQ) %>%
  try()
```

Comparison 4: Null binomial for zero-inflated model

```
compare_models(data = new_data2,
               models = "zi",

               zic_model = "FREQ ~ Year + Stratum + offset(lnSweptArea)",
               zi_model = "~ 1",
```

```
               r_dat = new_data2$FREQ) %>%
  try()
```

Comparison 5: Null count for zero-inflated model

```
compare_models(data = new_data2,
               models = "zi",

               zic_model = "FREQ ~ offset(lnSweptArea)",
               zi_model = "~ Year + Stratum ",

               r_dat = new_data2$FREQ) %>%
  try()
```

Model selection for negative binomial model

```
model <- glmmTMB::glmmTMB(FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea),
          data = new_data2,
          family = nbinom2
        )

null <- glmmTMB::glmmTMB(FREQ ~ Year + offset(lnSweptArea),
          data = new_data2,
          family = nbinom2
        )

select_model(model = model, null_model = null)
```

Model selection for zero-inflated model

```
model <- glmmTMB::glmmTMB(FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea),
          ziformula = ~ 1,
          data = new_data2,
          family = nbinom2
        )

null <- glmmTMB::glmmTMB(FREQ ~ Year + offset(lnSweptArea),
          ziformula = ~ 1,
          data = new_data2,
          family = nbinom2
        )

select_model(model = model, null_model = null)
```

Model selection for zero-adjusted model

```
model <- glmmTMB::glmmTMB(FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea),
          ziformula = ~ Year + Stratum + Bottom.wt + Bottom.do,
          data = new_data2,
          family = truncated_nbinom2(link = "log")
        )
```

```r
null <- glmmTMB::glmmTMB(FREQ ~ offset(lnSweptArea),
          ziformula = ~ Year + Stratum + Bottom.wt + Bottom.do,
          data = new_data2,
          family = truncated_nbinom2(link = "log")
        )

select_model(model = model, null_model = null) %>%
  try()
```

Final comparison

```r
compare_models(data = new_data2,
               nb_model = "FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea)",

               zic_model = "FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea)",
               zi_model = "~ 1",

               zac_model = "FREQ ~ Year + Stratum + Bottom.wt +
offset(lnSweptArea)",
               za_model = "~ Year + Stratum + Bottom.wt + Bottom.do",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 67.2

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  12759.6  13036.5  -6329.8  12659.6     1829
##
##
## Dispersion parameter for nbinom2 family (): 0.251
##
## Conditional model:
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)             -6.28424    0.45292 -13.875  < 2e-16 ***
## Year2008                 0.61433    0.26840   2.289 0.022087 *
## Year2009                 0.62293    0.28977   2.150 0.031575 *
## Year2010                 0.11468    0.28490   0.403 0.687289
## Year2011                -0.08273    0.25307  -0.327 0.743742
```

```
## Year2012                  0.94042    0.28845   3.260 0.001113 **
## Year2013                  0.45311    0.30477   1.487 0.137088
## Year2014                  0.16771    0.27587   0.608 0.543237
## Year2015                  0.20757    0.29454   0.705 0.480975
## Year2016                  1.11077    0.31155   3.565 0.000363 ***
## Year2017                  0.32419    0.26575   1.220 0.222499
## Year2018                  0.08668    0.28682   0.302 0.762498
## Year2019                  0.04569    0.30527   0.150 0.881033
## Year2020                  0.27886    0.42290   0.659 0.509645
## Stratum67541669.568       0.04487    0.56109   0.080 0.936262
## Stratum93842173.44        0.89996    0.56227   1.601 0.109467
## Stratum100324692          1.21814    0.59329   2.053 0.040054 *
## Stratum114833185.92      -3.42624    0.57352  -5.974 2.31e-09 ***
## Stratum115697521.728     -0.28871    0.55641  -0.519 0.603841
## Stratum115759260         -0.18114    0.52314  -0.346 0.729159
## Stratum118290529.152     -1.09932    0.57140  -1.924 0.054367 .
## Stratum163729897.344     -1.71337    0.59054  -2.901 0.003716 **
## Stratum172558470.24      -0.91319    0.52089  -1.753 0.079577 .
## Stratum179040988.8        0.31308    0.53473   0.585 0.558213
## Stratum183115714.752      0.58261    0.59356   0.982 0.326326
## Stratum185955675.264     -0.46370    0.56991  -0.814 0.415855
## Stratum257757285.6        0.18705    0.53484   0.350 0.726549
## Stratum259300742.4        0.34769    0.53832   0.646 0.518355
## Stratum267450194.304     -0.07034    0.50010  -0.141 0.888153
## Stratum276896149.92       0.20023    0.54680   0.366 0.714232
## Stratum320112940.32      -1.07207    0.49840  -2.151 0.031475 *
## Stratum336473582.4       -1.50732    0.51077  -2.951 0.003167 **
## Stratum355118540.544     -0.57200    0.48033  -1.191 0.233713
## Stratum366848812.224     -0.57423    0.48203  -1.191 0.233538
## Stratum374133928.32      -0.78805    0.49871  -1.580 0.114069
## Stratum413646422.399999 -1.92176    0.49826  -3.857 0.000115 ***
## Stratum419820249.599999 -1.95352    0.49715  -3.929 8.51e-05 ***
## Stratum523849237.919999 -2.27413    0.46779  -4.861 1.17e-06 ***
## Stratum543296793.599999 -2.18546    0.46107  -4.740 2.14e-06 ***
## Stratum555644447.999999 -1.92900    0.47285  -4.080 4.51e-05 ***
## Stratum560398294.944     -0.62227    0.47312  -1.315 0.188424
## Stratum598861238.4       -1.30684    0.47888  -2.729 0.006353 **
## Stratum697642473.600001 -2.28138    0.45472  -5.017 5.25e-07 ***
## Stratum703816300.800001 -0.96445    0.46303  -2.083 0.037260 *
## Stratum753206918.400001 -1.52114    0.45255  -3.361 0.000776 ***
## Stratum876683462.400002 -1.64004    0.44544  -3.682 0.000232 ***
## Stratum993986179.200003 -3.13471    0.45050  -6.958 3.44e-12 ***
## Bottom.wt                 0.11631    0.08294   1.402 0.160800
## Bottom.do                 0.10185    0.05690   1.790 0.073473 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.16

## Zero-inflated negative binomial
```

```
## model summary:

##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
## Zero inflation:        ~1
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  12761.6  13044.1  -6329.8  12659.6     1828
##
##
## Dispersion parameter for nbinom2 family (): 0.251
##
## Conditional model:
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)             -6.28424    0.45292 -13.875  < 2e-16 ***
## Year2008                 0.61432    0.26840   2.289 0.022089 *
## Year2009                 0.62293    0.28977   2.150 0.031574 *
## Year2010                 0.11468    0.28490   0.403 0.687305
## Year2011                -0.08272    0.25307  -0.327 0.743756
## Year2012                 0.94042    0.28845   3.260 0.001113 **
## Year2013                 0.45310    0.30477   1.487 0.137099
## Year2014                 0.16770    0.27587   0.608 0.543256
## Year2015                 0.20755    0.29454   0.705 0.481014
## Year2016                 1.11078    0.31155   3.565 0.000363 ***
## Year2017                 0.32418    0.26575   1.220 0.222507
## Year2018                 0.08668    0.28682   0.302 0.762504
## Year2019                 0.04569    0.30527   0.150 0.881018
## Year2020                 0.27886    0.42290   0.659 0.509648
## Stratum67541669.568      0.04490    0.56110   0.080 0.936219
## Stratum93842173.44       0.89996    0.56227   1.601 0.109467
## Stratum100324692         1.21810    0.59328   2.053 0.040058 *
## Stratum114833185.92     -3.42626    0.57352  -5.974 2.31e-09 ***
## Stratum115697521.728    -0.28873    0.55641  -0.519 0.603822
## Stratum115759260        -0.18115    0.52314  -0.346 0.729132
## Stratum118290529.152    -1.09936    0.57140  -1.924 0.054357 .
## Stratum163729897.344    -1.71342    0.59053  -2.901 0.003714 **
## Stratum172558470.24     -0.91318    0.52089  -1.753 0.079582 .
## Stratum179040988.8       0.31307    0.53472   0.585 0.558225
## Stratum183115714.752     0.58261    0.59356   0.982 0.326324
## Stratum185955675.264    -0.46374    0.56991  -0.814 0.415812
## Stratum257757285.6       0.18705    0.53484   0.350 0.726550
## Stratum259300742.4       0.34768    0.53832   0.646 0.518365
## Stratum267450194.304    -0.07032    0.50010  -0.141 0.888172
## Stratum276896149.92      0.20023    0.54680   0.366 0.714228
## Stratum320112940.32     -1.07205    0.49840  -2.151 0.031478 *
## Stratum336473582.4      -1.50732    0.51077  -2.951 0.003167 **
## Stratum355118540.544    -0.57198    0.48033  -1.191 0.233733
## Stratum366848812.224    -0.57422    0.48203  -1.191 0.233547
```

```
## Stratum374133928.32     -0.78805    0.49872  -1.580 0.114072
## Stratum413646422.399999 -1.92179    0.49826  -3.857 0.000115 ***
## Stratum419820249.599999 -1.95353    0.49715  -3.929 8.51e-05 ***
## Stratum523849237.919999 -2.27412    0.46779  -4.861 1.17e-06 ***
## Stratum543296793.599999 -2.18546    0.46107  -4.740 2.14e-06 ***
## Stratum555644447.999999 -1.92901    0.47285  -4.080 4.51e-05 ***
## Stratum560398294.944    -0.62223    0.47312  -1.315 0.188453
## Stratum598861238.4      -1.30687    0.47888  -2.729 0.006352 **
## Stratum697642473.600001 -2.28138    0.45472  -5.017 5.25e-07 ***
## Stratum703816300.800001 -0.96448    0.46303  -2.083 0.037254 *
## Stratum753206918.400001 -1.52115    0.45255  -3.361 0.000776 ***
## Stratum876683462.400002 -1.64005    0.44544  -3.682 0.000232 ***
## Stratum993986179.200003 -3.13470    0.45050  -6.958 3.45e-12 ***
## Bottom.wt                0.11631    0.08294   1.402 0.160803
## Bottom.do                0.10184    0.05690   1.790 0.073488 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -19.03    1304.94  -0.015    0.988

## Dispersal: 2.161

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ Year + Stratum + Bottom.wt + offset(lnSweptArea)
## Zero inflation:        ~Year + Stratum + Bottom.wt + Bottom.do
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  12564.6  13107.4  -6184.3  12368.6     1781
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.125
##
## Conditional model:
##                           Estimate Std. Error z value Pr(>|z|)
## (Intercept)             -7.1507677  0.5464308 -13.086  < 2e-16 ***
## Year2008                 0.9931600  0.3304466   3.006 0.002651 **
## Year2009                 0.7721901  0.3520052   2.194 0.028258 *
## Year2010                 0.4812895  0.3527669   1.364 0.172465
## Year2011                 0.0008386  0.3053459   0.003 0.997809
## Year2012                 1.2630739  0.3505167   3.603 0.000314 ***
## Year2013                 1.3290284  0.3951935   3.363 0.000771 ***
## Year2014                 0.5941722  0.3472019   1.711 0.087023 .
## Year2015                 0.6975395  0.3605489   1.935 0.053032 .
## Year2016                 1.8498680  0.3940987   4.694 2.68e-06 ***
```

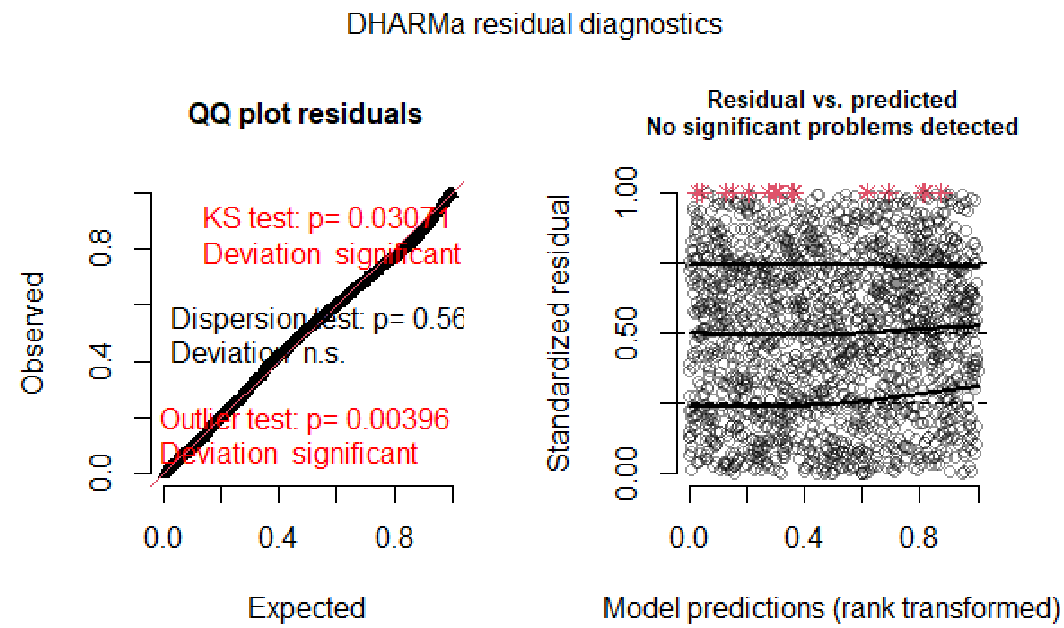
```
## Year2017                    0.7359523  0.3309377   2.224 0.026159 *
## Year2018                    0.7615671  0.3741439   2.035 0.041801 *
## Year2019                    0.6212155  0.3918366   1.585 0.112877
## Year2020                    0.7766195  0.5418709   1.433 0.151795
## Stratum67541669.568        -0.0919351  0.6450363  -0.143 0.886664
## Stratum93842173.44          0.9864960  0.6596710   1.495 0.134801
## Stratum100324692            1.7072395  0.7258771   2.352 0.018674 *
## Stratum114833185.92        -3.1898267  0.7146468  -4.464 8.06e-06 ***
## Stratum115697521.728       -0.6080159  0.6259713  -0.971 0.331391
## Stratum115759260           -0.2599018  0.6088749  -0.427 0.669484
## Stratum118290529.152       -1.3294743  0.6651426  -1.999 0.045632 *
## Stratum163729897.344       -1.3505558  0.7775303  -1.737 0.082390 .
## Stratum172558470.24        -0.9233853  0.6116342  -1.510 0.131119
## Stratum179040988.8          0.5603257  0.6596312   0.849 0.395629
## Stratum183115714.752        1.0228660  0.7611615   1.344 0.179006
## Stratum185955675.264       -0.6766365  0.6348950  -1.066 0.286539
## Stratum257757285.6          0.3280862  0.6553055   0.501 0.616609
## Stratum259300742.4          1.3608237  0.7336858   1.855 0.063628 .
## Stratum267450194.304       -0.2188650  0.5702628  -0.384 0.701129
## Stratum276896149.92         1.4710740  0.8275170   1.778 0.075454 .
## Stratum320112940.32        -1.0414871  0.5886449  -1.769 0.076845 .
## Stratum336473582.4         -1.3747874  0.6371342  -2.158 0.030946 *
## Stratum355118540.544       -0.8053318  0.5525523  -1.457 0.144985
## Stratum366848812.224       -0.7325105  0.5502357  -1.331 0.183101
## Stratum374133928.32        -0.9275175  0.5935441  -1.563 0.118129
## Stratum413646422.399999 -1.9056220  0.6131972  -3.108 0.001886 **
## Stratum419820249.599999 -1.9823669  0.6117460  -3.241 0.001193 **
## Stratum523849237.919999 -2.4646850  0.5418563  -4.549 5.40e-06 ***
## Stratum543296793.599999 -2.7386059  0.5179464  -5.287 1.24e-07 ***
## Stratum555644447.999999 -1.6184440  0.5825869  -2.778 0.005469 **
## Stratum560398294.944       -0.6840423  0.5629124  -1.215 0.224296
## Stratum598861238.4         -0.8652809  0.5802176  -1.491 0.135882
## Stratum697642473.600001 -2.0193881  0.5460310  -3.698 0.000217 ***
## Stratum703816300.800001 -1.0727578  0.5442506  -1.971 0.048715 *
## Stratum753206918.400001 -1.6676103  0.5244095  -3.180 0.001473 **
## Stratum876683462.400002 -1.3905350  0.5395867  -2.577 0.009965 **
## Stratum993986179.200003 -3.6870558  0.5188923  -7.106 1.20e-12 ***
## Bottom.wt                   0.2202827  0.1039576   2.119 0.034093 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                            Estimate Std. Error z value Pr(>|z|)
## (Intercept)                -1.96592    0.65538  -3.000 0.002703 **
## Year2008                   -0.06194    0.27669  -0.224 0.822868
## Year2009                   -0.80232    0.31484  -2.548 0.010823 *
## Year2010                   -0.34388    0.29304  -1.173 0.240599
## Year2011                   -0.14920    0.28041  -0.532 0.594674
## Year2012                   -0.79215    0.32457  -2.441 0.014663 *
## Year2013                    0.21472    0.29997   0.716 0.474107
```

```
## Year2014                  0.19892    0.28696   0.693 0.488181
## Year2015                 -0.02171    0.30057  -0.072 0.942415
## Year2016                 -0.07623    0.30399  -0.251 0.801988
## Year2017                 -0.00920    0.27603  -0.033 0.973412
## Year2018                  0.54757    0.27619   1.983 0.047409 *
## Year2019                  0.48773    0.30135   1.619 0.105553
## Year2020                  0.33443    0.40507   0.826 0.409019
## Stratum67541669.568      -0.09380    0.87553  -0.107 0.914685
## Stratum93842173.44       -0.05901    0.87475  -0.067 0.946217
## Stratum100324692          0.65720    0.80412   0.817 0.413766
## Stratum114833185.92       2.30053    0.73594   3.126 0.001772 **
## Stratum115697521.728     -1.42469    1.19578  -1.191 0.233485
## Stratum115759260          0.30340    0.76598   0.396 0.692032
## Stratum118290529.152      0.66248    0.78151   0.848 0.396607
## Stratum163729897.344      2.08236    0.75384   2.762 0.005739 **
## Stratum172558470.24       0.96688    0.72416   1.335 0.181818
## Stratum179040988.8        1.14524    0.72149   1.587 0.112440
## Stratum183115714.752      1.44864    0.76521   1.893 0.058341 .
## Stratum185955675.264     -1.20053    1.19776  -1.002 0.316195
## Stratum257757285.6        1.10707    0.71913   1.539 0.123692
## Stratum259300742.4        2.02859    0.71266   2.846 0.004420 **
## Stratum267450194.304     -0.44172    0.81363  -0.543 0.587197
## Stratum276896149.92       2.76719    0.71835   3.852 0.000117 ***
## Stratum320112940.32       1.00421    0.69724   1.440 0.149792
## Stratum336473582.4        1.80829    0.68762   2.630 0.008544 **
## Stratum355118540.544      0.13759    0.71917   0.191 0.848277
## Stratum366848812.224     -0.44930    0.77801  -0.577 0.563605
## Stratum374133928.32       0.98192    0.67953   1.445 0.148458
## Stratum413646422.399999   1.84926    0.68389   2.704 0.006850 **
## Stratum419820249.599999   1.91635    0.67603   2.835 0.004587 **
## Stratum523849237.919999   1.32763    0.66737   1.989 0.046663 *
## Stratum543296793.599999   0.03767    0.70059   0.054 0.957124
## Stratum555644447.999999   2.06153    0.65821   3.132 0.001736 **
## Stratum560398294.944      1.17800    0.67019   1.758 0.078797 .
## Stratum598861238.4        1.56926    0.67106   2.338 0.019363 *
## Stratum697642473.600001   1.97401    0.65288   3.024 0.002498 **
## Stratum703816300.800001   1.01683    0.66398   1.531 0.125668
## Stratum753206918.400001   1.03739    0.66045   1.571 0.116244
## Stratum876683462.400002   2.02050    0.64862   3.115 0.001839 **
## Stratum993986179.200003   1.55612    0.65377   2.380 0.017302 *
## Bottom.wt                -0.03989    0.07525  -0.530 0.596048
## Bottom.do                -0.08094    0.05489  -1.475 0.140302
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.913

##           dAIC df
## zero_adj    0  98
```

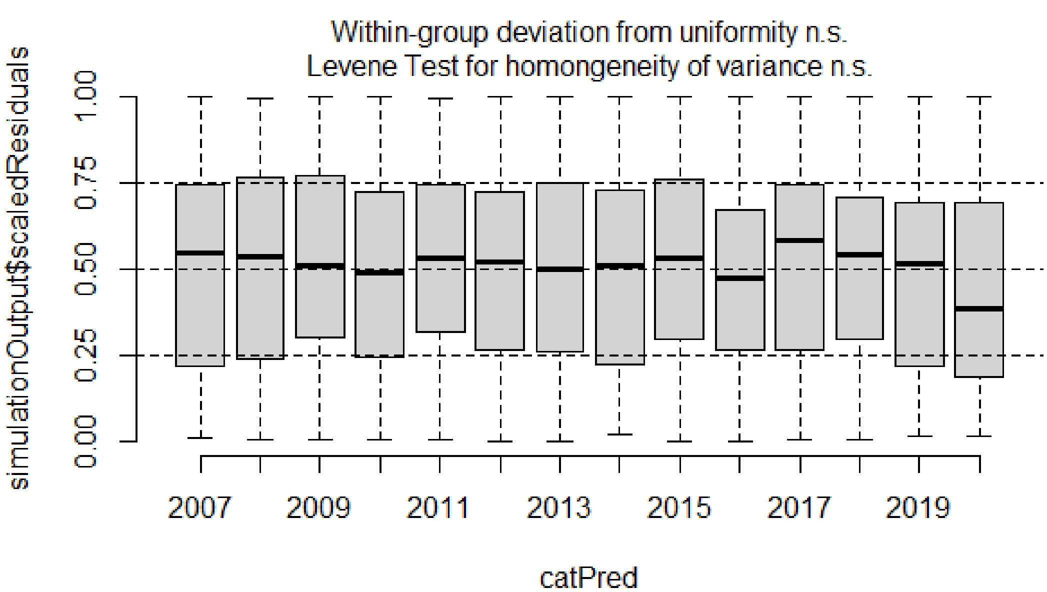
```
## neg_binom 195   50
## zero_infl 197   51
```

Step 3: DHARMA simulations

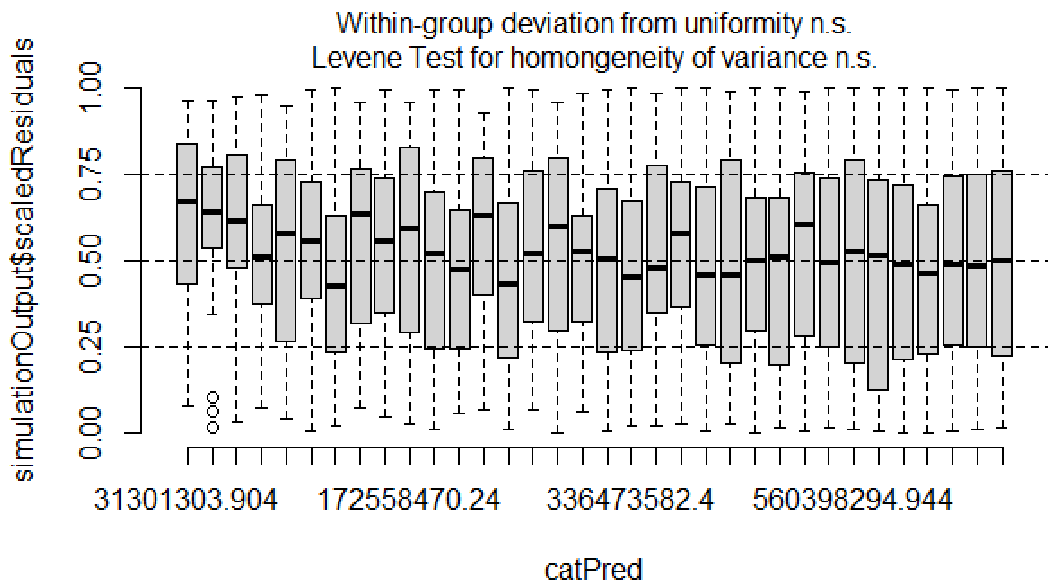
```
model1 <- glmmTMB::glmmTMB(FREQ ~ Year + Depth + Lat + Bottom.wt + Bottom.do
+ offset(lnSweptArea),
          ziformula = ~ Year + Depth + Lat + Bottom.wt + Bottom.do,
          data = new_data2,
          family = truncated_nbinom2(link = "log")
        )

model2 <- glmmTMB::glmmTMB(FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea),
          ziformula = ~ Year + Stratum + Bottom.wt + Bottom.do,
          data = new_data2,
          family = truncated_nbinom2(link = "log")
        )

AIC(model1, model2)

##        df      AIC
## model1 37 12875.41
## model2 99 12565.57

# model 2 is a lot better

model <- model2

sim <- simulateResiduals(model, quantreg = TRUE)
plot(sim) # problems

## DHARMa:testOutliers with type = binomial may have inflated Type I error
rates for integer-valued distributions. To get a more exact result, it is
recommended to re-run testOutliers with type = 'bootstrap'. See ?testOutliers
for details
```



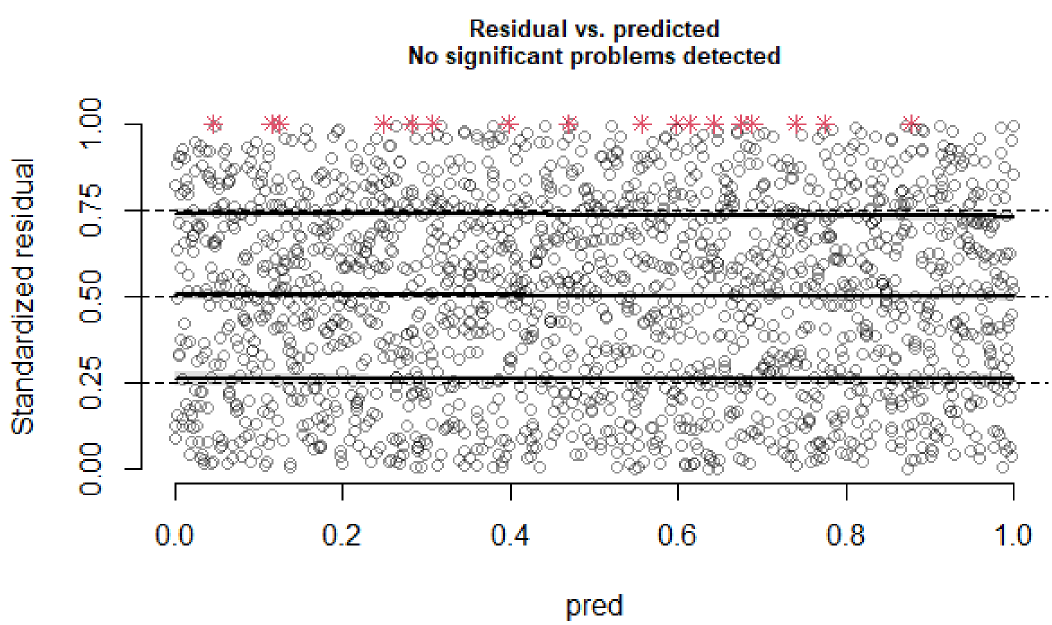

```
# Compare residuals vs. factors
plotResiduals(sim, quantreg = TRUE, form=new_data2$Year)
```



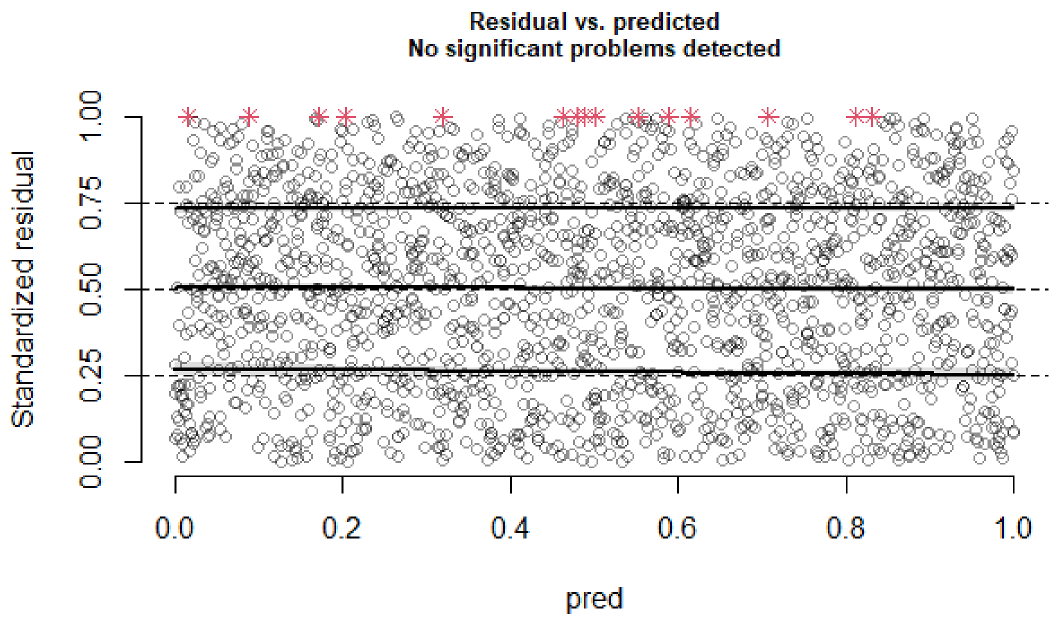

```
plotResiduals(sim, quantreg = TRUE, form=new_data2$Stratum)
```




```
plotResiduals(sim, quantreg = TRUE, form=new_data2$Bottom.do)
```

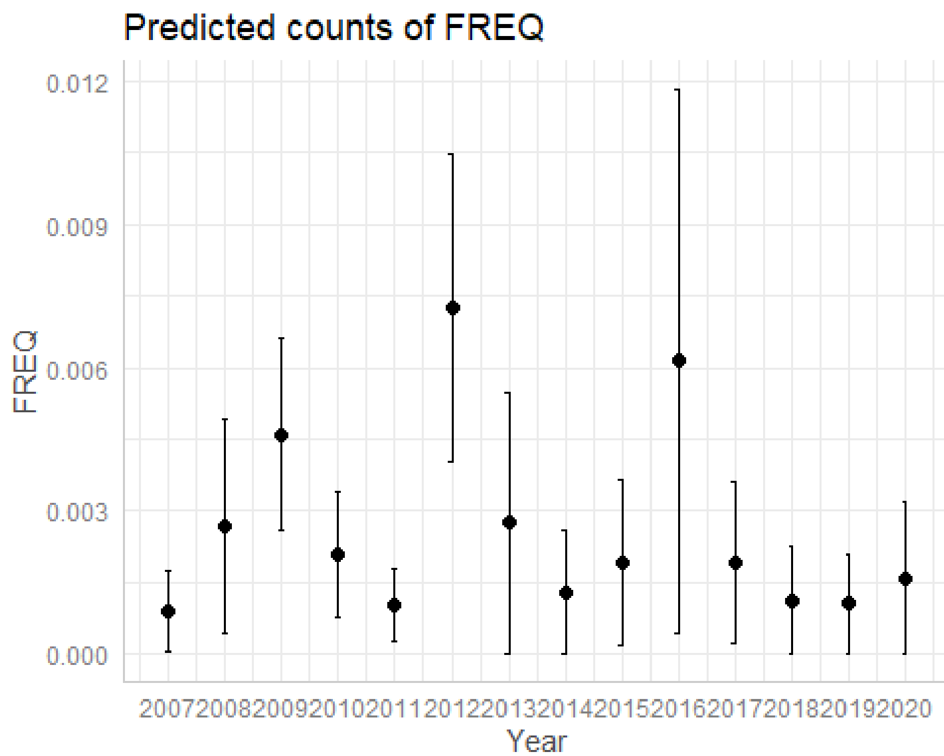



```
plotResiduals(sim, quantreg = TRUE, form=new_data2$Bottom.wt)
```



Step 4: Plot predictions

```
index.pred <- ggeffects::ggemmeans(model,
                                   terms="Year",
                                   type="fe.zi") %>%
  try()
plot(index.pred) %>%
  try()
```



Step 5: Save index

```
index.out <- as.data.frame(index.pred)
index.out$CV <- index.out$std.error/index.out$predicted

names(index.out)[c(1,2)] <- c("Year", "Frequency")

fname <- here::here("data/NEAMAP_index3.csv")
write.csv(index.out, fname, row.names=FALSE)
```

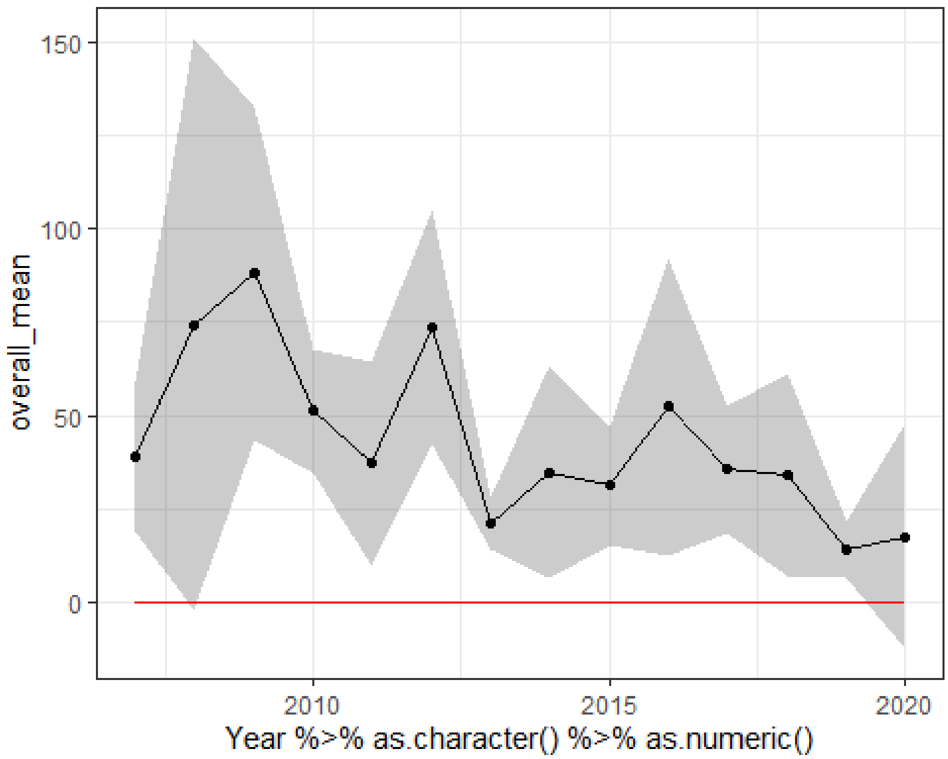
Develop nominal index

```
# using dataset `new_data`, which has sparse sampling removed
# not sure if/how to standardize for effort, so skipping that for now

stratum_weights <- new_data %>%
  dplyr::select(Stratum_area) %>%
  dplyr::distinct() %>%
  dplyr::mutate(weight = Stratum_area/sum(Stratum_area))

nominal_data <- new_data %>%
  dplyr::left_join(stratum_weights, by = "Stratum_area") %>%
  dplyr::rename(Stratum = Stratum_area) %>%
  dplyr::group_by(Year, Stratum) %>%
  dplyr::summarise(mean_freq = mean(FREQ) %>% round(digits = 3),
                   var_freq = var(FREQ) %>% round(digits = 3),
                   n_obs = length(SampleID)
                   ) %>%
```

```r
  dplyr::ungroup() %>%
  dplyr::group_by(Year)  %>%
  dplyr::mutate(mean_var = mean(var_freq, na.rm = TRUE),
                var_freq = ifelse(is.na(var_freq), mean_var, var_freq),
                n_strata = length(unique(Stratum)),
                weight = 1/n_strata) %>%
  dplyr::ungroup() %>%
  dplyr::group_by(Year)  %>%
  dplyr::summarise(overall_mean = sum(weight * mean_freq),

                   overall_var =  sum(weight^2 * var_freq / n_obs),

                   overall_sd = sqrt(overall_var),
                   .groups = "keep") %>%
  dplyr::ungroup()
```

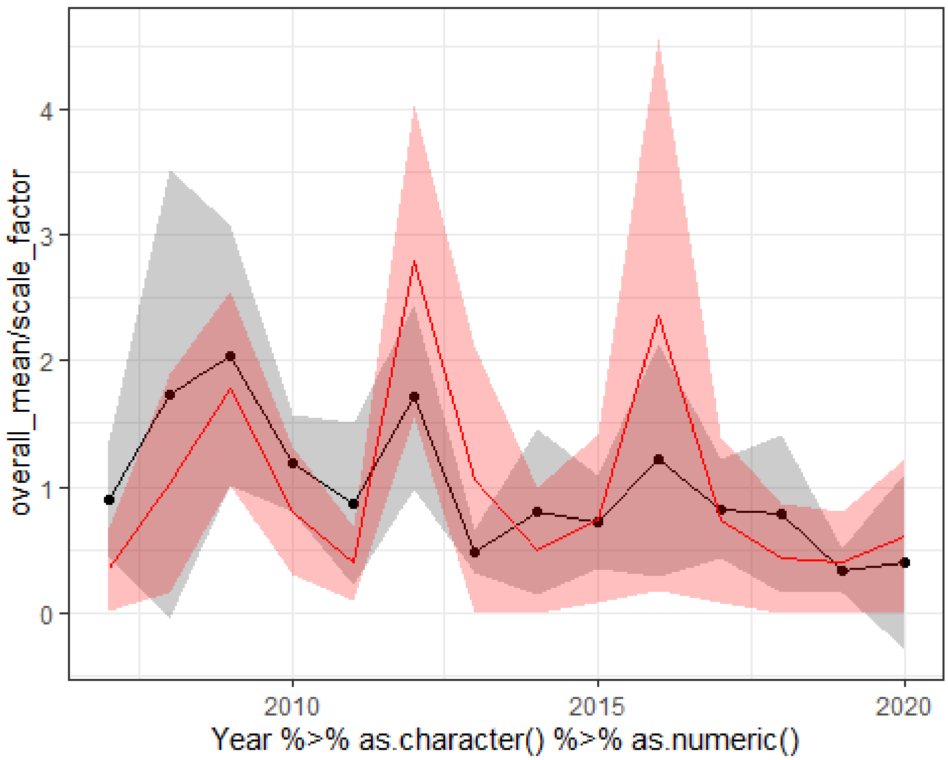
```
## `summarise()` has grouped output by 'Year'. You can override using the
`.groups` argument.
```

```r
write.csv(nominal_data, "VA_NEAMAP_index_nominal.csv", row.names=FALSE)

ggplot2::ggplot(nominal_data,
                ggplot2::aes(x = Year %>%
                                  as.character() %>%
                                  as.numeric(),
                             group = 1)) +
  ggplot2::geom_ribbon(ggplot2::aes(ymin = overall_mean - 1.96*overall_sd,
                                  ymax = overall_mean + 1.96*overall_sd),
                       alpha = 0.25) +
  ggplot2::geom_line(ggplot2::aes(y = overall_mean)) +
  ggplot2::geom_point(ggplot2::aes(y = overall_mean)) +
  ggplot2::theme_bw() +
  ggplot2::geom_line(data = index.out,
                     ggplot2::aes(x = Year %>%
                                  as.character() %>%
                                  as.numeric(),
                                y = Frequency),
               color = "red",
               inherit.aes = FALSE) +
  ggplot2::geom_ribbon(data = index.out,
                       ggplot2::aes(x = Year %>%
                                  as.character() %>%
                                  as.numeric(),
                                  ymin = conf.low,
                                  ymax = conf.high),
                       alpha = 0.25,
                       fill = "red",
                       inherit.aes = FALSE)
```




```
# standardize to mean
nominal_data <- nominal_data %>%
  dplyr::mutate(scale_factor = mean(overall_mean))
index.out <- index.out %>%
  dplyr::mutate(scale_factor = mean(Frequency))

ggplot2::ggplot(nominal_data,
                ggplot2::aes(x = Year %>%
                                    as.character() %>%
                                    as.numeric(),
                             group = 1)) +
  ggplot2::geom_ribbon(ggplot2::aes(ymin = (overall_mean -
1.96*overall_sd)/scale_factor,
                                    ymax = (overall_mean +
1.96*overall_sd)/scale_factor),
                       alpha = 0.25) +
  ggplot2::geom_line(ggplot2::aes(y = overall_mean/scale_factor)) +
  ggplot2::geom_point(ggplot2::aes(y = overall_mean/scale_factor)) +
  ggplot2::theme_bw() +
  ggplot2::geom_line(data = index.out,
                     ggplot2::aes(x = Year %>%
                                    as.character() %>%
                                    as.numeric(),
                                  y = Frequency/scale_factor),
               color = "red",
               inherit.aes = FALSE) +
```

```r
  ggplot2::geom_ribbon(data = index.out,
                       ggplot2::aes(x = Year %>%
                                      as.character() %>%
                                      as.numeric(),
                                    ymin = conf.low/scale_factor,
                                    ymax = conf.high/scale_factor),
                       alpha = 0.25,
                       fill = "red",
                       inherit.aes = FALSE)
```



The confidence intervals go below zero. Should we log-transform the data?

Bluefish index standardization - SEAMAP Fall Age-0

Katie Drew, ASMFC

2022-02-10

Bluefish index standardization

Step 1: Data processing

Read in data; SEAMAP data has already been subset to fall tows only for the age-0 index.

```
indata <- read.csv("SEAMAP_AGE0.csv")
head(indata)
```

```
##   SAMPLEID STATE       PROGRAM
GEAR
## 1  1989323    NC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
## 2  1989325    NC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
## 3  1989327    NC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
## 4  1989329    NC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
## 5  1989335    NC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
## 6  1989337    NC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
##   YEAR MONTH DAY  TIME DURATION EFFORT  SPECIES N WEIGHT AREA STRATUM
LATITUDE
## 1 1989    10  10 10:32       20      0          0  0.000   OB      57
33.72
## 2 1989    10  10 11:17       20      0          0  0.000   OB      57
33.75
## 3 1989    10  10 13:29       20      0          0  0.000   OB      57
33.91
## 4 1989    10  10 15:28       20      0          0  0.000   OB      59
34.03
## 5 1989    10  11  9:30       20      0          0  0.000   RB      67
35.08
## 6 1989    10  11 11:16       20      9 bluefish 3  0.331   RB      65
34.97
##   LONGITUDE DEPTH SURFACE.TEMP BOTTOM.TEMP SURFACE.SALINITY SALINITY
SEASON
## 1    -77.89     7         22.0        22.4             35.5     35.3
FALL
## 2    -77.90     7         21.9        22.3             35.4     35.3
```

```
FALL
## 3    -77.90    10          21.7         22.1              35.1     35.3
FALL
## 4    -77.87     8          22.0         22.3              35.2     34.9
FALL
## 5    -75.94     9          20.5         22.0              32.3     33.3
FALL
## 6    -76.12     7          21.0         21.9              32.4     33.0
FALL
##   POS_TOW N_A0
## 1       0    0
## 2       0    0
## 3       0    0
## 4       0    0
## 5       0    0
## 6       1    3

# modify data formatting
indata <- indata %>%
  dplyr::rename(FREQ = N_A0) %>%
  dplyr::mutate(YEAR = as.factor(YEAR),
                MONTH = as.factor(MONTH),
                AREA = as.factor(AREA),
                STRATUM = as.factor(STRATUM)) %>%
  dplyr::select(-c(PROGRAM, GEAR, SPECIES, DURATION, EFFORT)) # same data for
all entries

head(indata)

##   SAMPLEID STATE YEAR MONTH DAY  TIME N WEIGHT AREA STRATUM LATITUDE
LONGITUDE
## 1  1989323    NC 1989    10  10 10:32 0  0.000   OB      57    33.72    -
77.89
## 2  1989325    NC 1989    10  10 11:17 0  0.000   OB      57    33.75    -
77.90
## 3  1989327    NC 1989    10  10 13:29 0  0.000   OB      57    33.91    -
77.90
## 4  1989329    NC 1989    10  10 15:28 0  0.000   OB      59    34.03    -
77.87
## 5  1989335    NC 1989    10  11  9:30 0  0.000   RB      67    35.08    -
75.94
## 6  1989337    NC 1989    10  11 11:16 3  0.331   RB      65    34.97    -
76.12
##   DEPTH SURFACE.TEMP BOTTOM.TEMP SURFACE.SALINITY SALINITY SEASON POS_TOW
FREQ
## 1     7         22.0        22.4             35.5     35.3   FALL       0
0
## 2     7         21.9        22.3             35.4     35.3   FALL       0
0
## 3    10         21.7        22.1             35.1     35.3   FALL       0
```

```
0
## 4     8         22.0        22.3             35.2     34.9   FALL       0
0
## 5     9         20.5        22.0             32.3     33.3   FALL       0
0
## 6     7         21.0        21.9             32.4     33.0   FALL       1
3
```

```
summary(indata)
```

```
##     SAMPLEID          STATE                YEAR      MONTH         DAY
##  Min.   :1989323   Length:2803        2011   : 112   9 :  31   Min.   :
1.00
##  1st Qu.:1998456   Class :character   2012   : 112   10:2202   1st Qu.:
8.00
##  Median :2005601   Mode  :character   2009   : 111   11: 570   Median
:14.00
##  Mean   :2005365                      2010   : 111             Mean
:14.97
##  3rd Qu.:2012560                      2015   : 108             3rd
Qu.:22.00
##  Max.   :2019515                      2016   : 107             Max.
:31.00
##                                       (Other):2142
##      TIME                 N                WEIGHT         AREA
##  Length:2803        Min.   :   0.000   Min.   :  0.0000   FL:472
##  Class :character   1st Qu.:   0.000   1st Qu.:  0.0000   GA:683
##  Mode  :character   Median :   0.000   Median :  0.0000   LB:428
##                     Mean   :   7.669   Mean   :  0.8188   OB:419
##                     3rd Qu.:   1.000   3rd Qu.:  0.2390   RB:191
##                     Max.   :2771.000   Max.   :319.5680   SC:610
##
##     STRATUM        LATITUDE       LONGITUDE         DEPTH
##  51     : 174   Min.   :28.76   Min.   :-81.44   Min.   : 3.000
##  35     : 154   1st Qu.:31.00   1st Qu.:-81.08   1st Qu.: 8.000
##  49     : 150   Median :32.59   Median :-79.95   Median : 8.000
##  37     : 146   Mean   :32.33   Mean   :-79.59   Mean   : 8.412
##  33     : 144   3rd Qu.:33.80   3rd Qu.:-78.31   3rd Qu.: 9.000
##  41     : 142   Max.   :35.22   Max.   :-75.61   Max.   :14.000
##  (Other):1893
##   SURFACE.TEMP    BOTTOM.TEMP    SURFACE.SALINITY    SALINITY
##  Min.   :13.60   Min.   :14.40   Min.   :11.10    Min.   :28.60
##  1st Qu.:20.60   1st Qu.:20.60   1st Qu.:32.60    1st Qu.:33.30
##  Median :22.70   Median :22.80   Median :34.00    Median :34.40
##  Mean   :22.61   Mean   :22.58   Mean   :33.56    Mean   :34.17
##  3rd Qu.:24.60   3rd Qu.:24.40   3rd Qu.:35.10    3rd Qu.:35.30
##  Max.   :30.10   Max.   :29.10   Max.   :36.60    Max.   :36.80
##
##     SEASON             POS_TOW           FREQ
##  Length:2803        Min.   :0.000   Min.   :   0.000
```

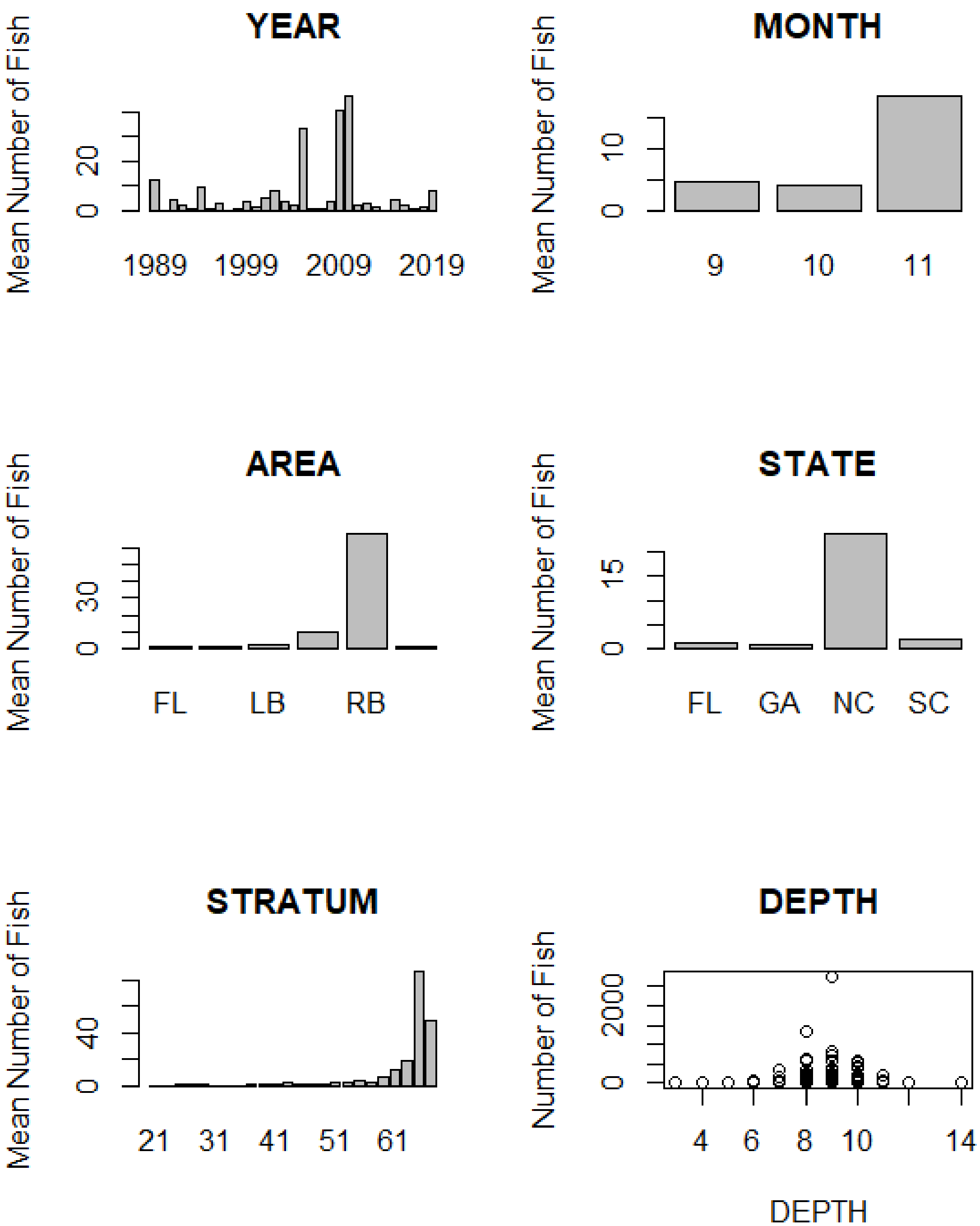
```
##  Class :character   1st Qu.:0.000   1st Qu.:   0.000
##  Mode  :character   Median :0.000   Median :   0.000
##                     Mean   :0.335   Mean   :   7.205
##                     3rd Qu.:1.000   3rd Qu.:   1.000
##                     Max.   :1.000   Max.   :2774.000
##
```

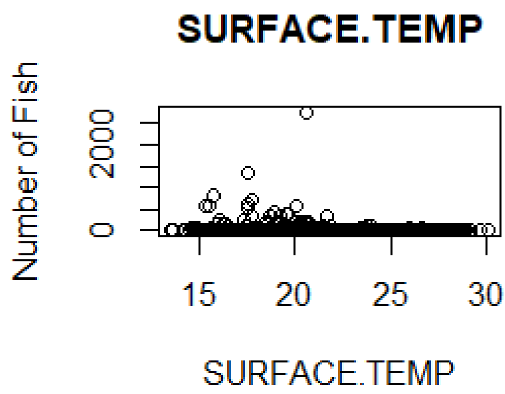
Remove tows that are missing covariates.

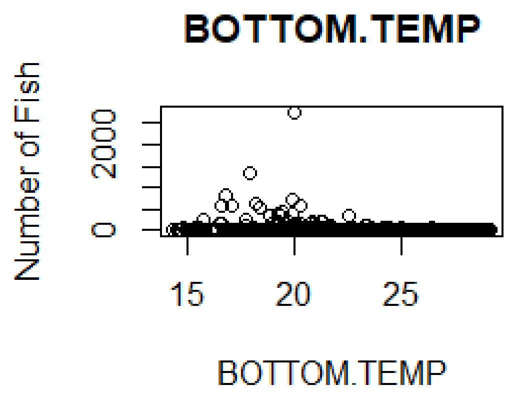
```
length(indata$YEAR)

## [1] 2803

new_data <- droplevels(indata[complete.cases(indata),])
length(new_data$YEAR)

## [1] 2803
```

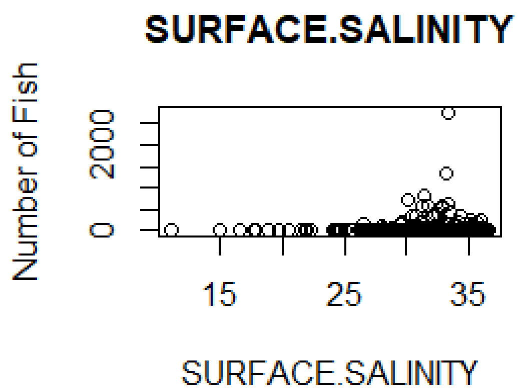
Plot data and decide if any data should be removed due to sparse sampling

```
plot_obs(new_data,
          re.v = "FREQ",p_vars=c("YEAR", "MONTH", "AREA", "STATE", "STRATUM",
"DEPTH", "SURFACE.TEMP", "BOTTOM.TEMP", "SURFACE.SALINITY", "SALINITY"))
```

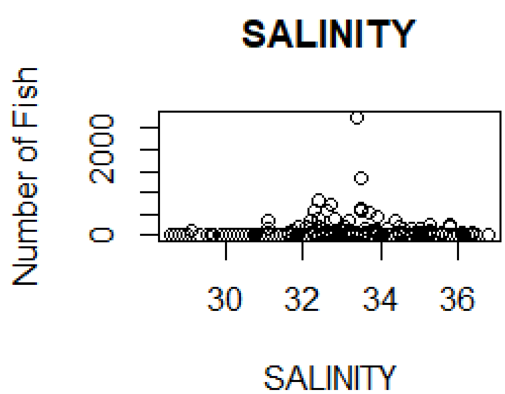

YEAR
Mean Number of Fish
0
20
1989
1999
2009
2019
MONTH
Mean Number of Fish
0
10
9
10
11
AREA
Mean Number of Fish
0
30
FL
LB
RB
STATE
Mean Number of Fish
0
15
FL
GA
NC
SC
STRATUM
Mean Number of Fish
0
40
21
31
41
51
61
DEPTH
Number of Fish
0
2000
4
6
8
10
14
DEPTH



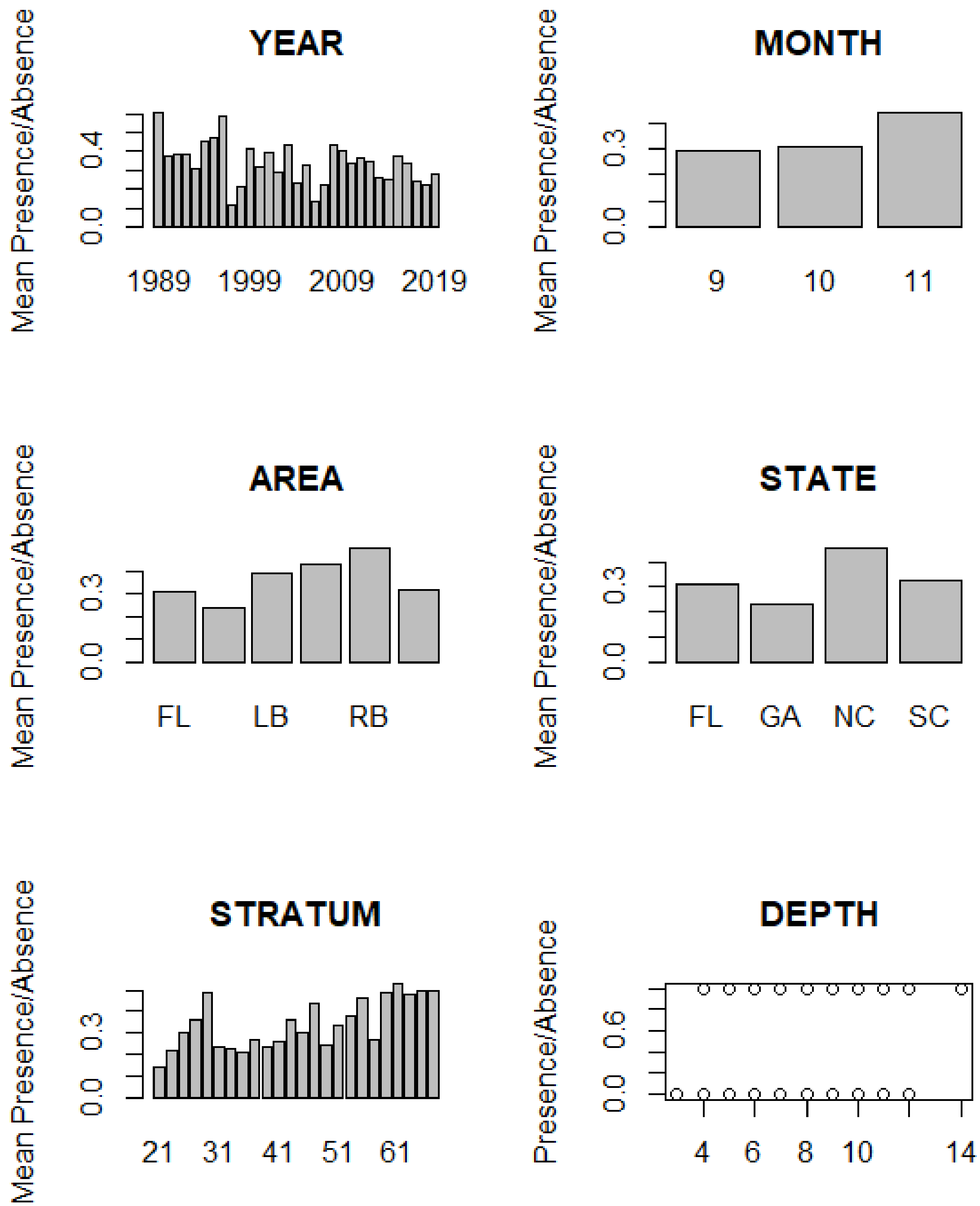




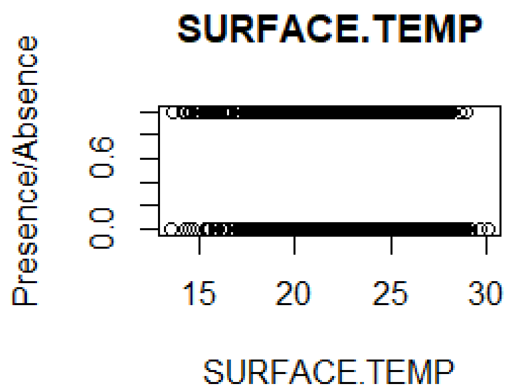

```
## Please manually remove sparsely sampled data!
```

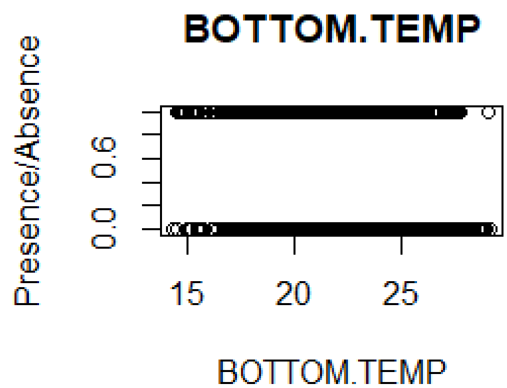


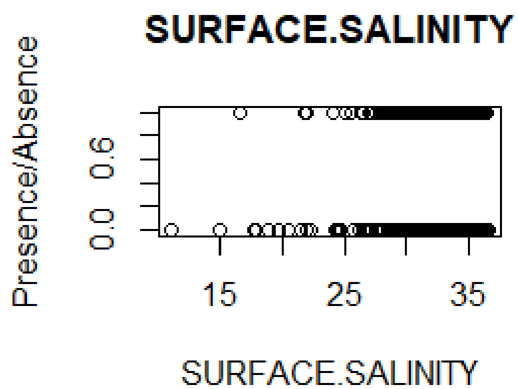

```
plot_obs(new_data,
          re.v = "POS_TOW",p_vars=c("YEAR", "MONTH", "AREA", "STATE",
"STRATUM", "DEPTH", "SURFACE.TEMP", "BOTTOM.TEMP", "SURFACE.SALINITY",
"SALINITY"))
```



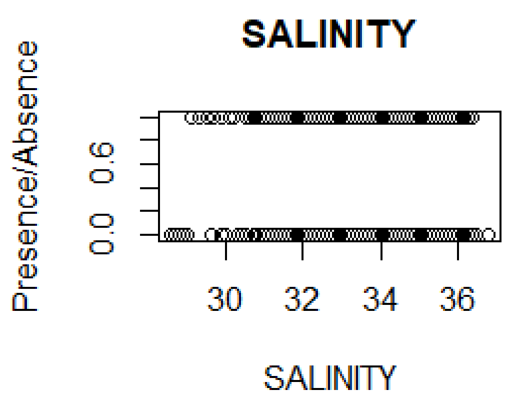
YEAR
Mean Presence/Absence
0.0
0.4
1989
1999
2009
2019
MONTH
Mean Presence/Absence
0.0
0.3
9
10
11
AREA
Mean Presence/Absence
0.0
0.3
FL
LB
RB
STATE
Mean Presence/Absence
0.0
0.3
FL
GA
NC
SC
STRATUM
Mean Presence/Absence
0.0
0.3
21
31
41
51
61
DEPTH
Presence/Absence
0.0
0.6
4
6
8
10
14
DEPTH





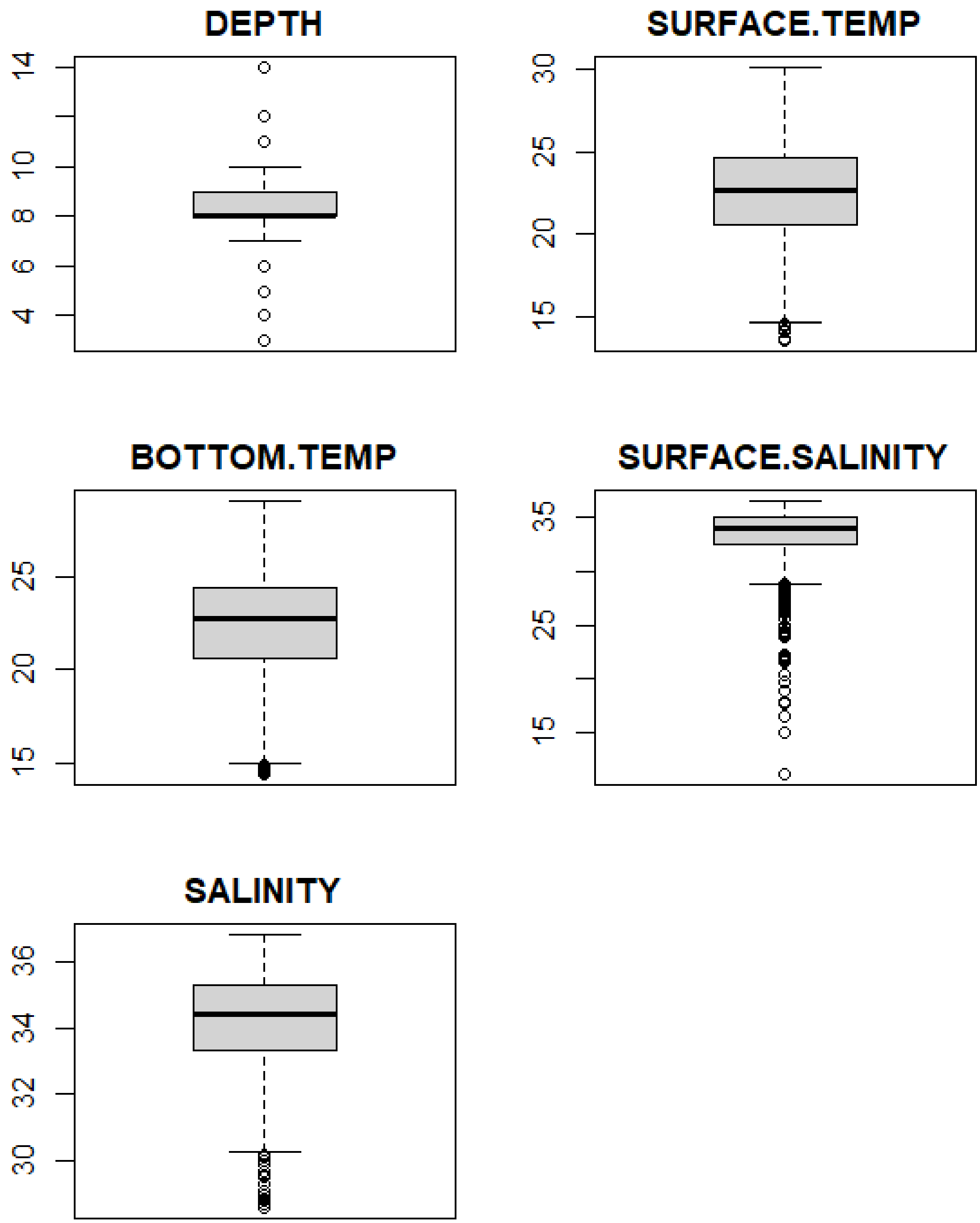



```
## Please manually remove sparsely sampled data!
```

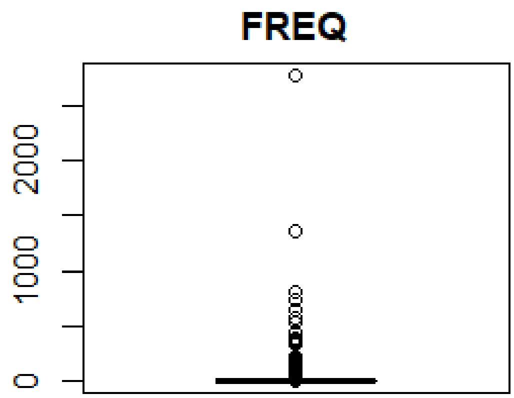


Check for outliers

```
plot_outliers(new_data, vars=c("DEPTH", "SURFACE.TEMP", "BOTTOM.TEMP",
"SURFACE.SALINITY", "SALINITY", "FREQ"))
```




```
## Please manually remove any outliers!
```



Change continuous variables to z-score and set factors.

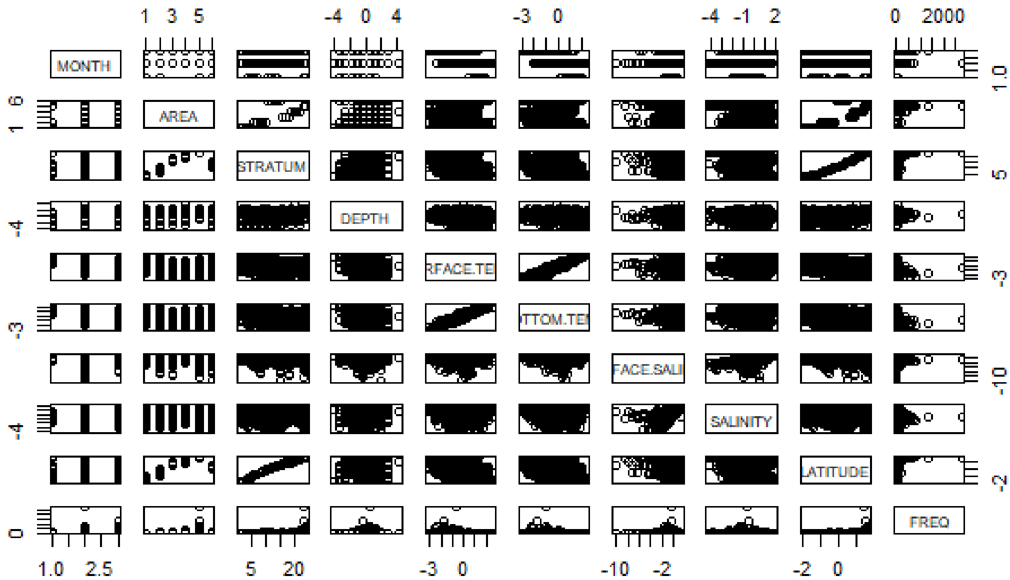
```
new_data2 <- standardize_data(new_data,
                              cont.cols = c("DEPTH", "SURFACE.TEMP",
"BOTTOM.TEMP", "SURFACE.SALINITY",
                                            "SALINITY", "LATITUDE"),
                              fac.cols = c("YEAR", "MONTH", "AREA",
"STRATUM", "STATE"),
                              N_col = "FREQ")

## YEAR is already a factor

## MONTH is already a factor

## AREA is already a factor

## STRATUM is already a factor

## Data standardized!

new_data2 <- as.data.frame(new_data2)
head(new_data2)

##   SAMPLEID STATE YEAR MONTH DAY  TIME N WEIGHT AREA STRATUM  LATITUDE
LONGITUDE
## 1  1989323    NC 1989    10  10 10:32 0  0.000   OB      57 0.7875065    -
77.89
## 2  1989325    NC 1989    10  10 11:17 0  0.000   OB      57 0.8045184    -
77.90
## 3  1989327    NC 1989    10  10 13:29 0  0.000   OB      57 0.8952486    -
77.90
## 4  1989329    NC 1989    10  10 15:28 0  0.000   OB      59 0.9632962    -
77.87
## 5  1989335    NC 1989    10  11  9:30 0  0.000   RB      67 1.5587131    -
75.94
## 6  1989337    NC 1989    10  11 11:16 3  0.331   RB      65 1.4963361    -
76.12
```

```
##        DEPTH SURFACE.TEMP BOTTOM.TEMP SURFACE.SALINITY   SALINITY SEASON
## 1 -1.0830105   -0.2163632 -0.06524466        0.8908869  0.8384552   FALL
## 2 -1.0830105   -0.2519874 -0.10231008        0.8450440  0.8384552   FALL
## 3  1.2184894   -0.3232357 -0.17644092        0.7075152  0.8384552   FALL
## 4 -0.3158438   -0.2163632 -0.10231008        0.7533581  0.5427594   FALL
## 5  0.4513228   -0.7507259 -0.21350634       -0.5760868 -0.6400236   FALL
## 6 -1.0830105   -0.5726050 -0.25057176       -0.5302439 -0.8617954   FALL
##   POS_TOW FREQ
## 1       0    0
## 2       0    0
## 3       0    0
## 4       0    0
## 5       0    0
## 6       1    3
```

Check colinearity

```
check_colin(new_data2,
            cols = c("MONTH", "AREA", "STRATUM", "DEPTH", "SURFACE.TEMP",
"BOTTOM.TEMP", "SURFACE.SALINITY", "SALINITY", "LATITUDE",  "FREQ"))
```

```
##                  MONTH  AREA STRATUM DEPTH SURFACE.TEMP BOTTOM.TEMP
## MONTH             1.00 -0.03   -0.05  0.03        -0.55       -0.56
## AREA             -0.03  1.00    0.63 -0.04        -0.31       -0.30
## STRATUM          -0.05  0.63    1.00  0.14        -0.38       -0.35
## DEPTH             0.03 -0.04    0.14  1.00        -0.07       -0.06
## SURFACE.TEMP     -0.55 -0.31   -0.38 -0.07         1.00        0.98
## BOTTOM.TEMP      -0.56 -0.30   -0.35 -0.06         0.98        1.00
## SURFACE.SALINITY  0.03 -0.12   -0.06  0.06         0.04       -0.01
## SALINITY         -0.01 -0.09    0.02  0.13         0.06        0.05
## LATITUDE         -0.03  0.66    0.99  0.12        -0.42       -0.39
## FREQ              0.08  0.07    0.14  0.03        -0.12       -0.12
##                  SURFACE.SALINITY SALINITY LATITUDE  FREQ
## MONTH                        0.03    -0.01    -0.03  0.08
## AREA                        -0.12    -0.09     0.66  0.07
## STRATUM                     -0.06     0.02     0.99  0.14
## DEPTH                        0.06     0.13     0.12  0.03
## SURFACE.TEMP                 0.04     0.06    -0.42 -0.12
## BOTTOM.TEMP                 -0.01     0.05    -0.39 -0.12
## SURFACE.SALINITY             1.00     0.75    -0.09 -0.04
## SALINITY                     0.75     1.00    -0.01 -0.05
## LATITUDE                    -0.09    -0.01     1.00  0.12
## FREQ                        -0.04    -0.05     0.12  1.00

## LATITUDE is correlated with STRATUM

## BOTTOM.TEMP is correlated with SURFACE.TEMP
```



Check variance inflation factor - Remove covariates with GVIF > 3, starting with highest value, until all GVIF < 3

```
lmmod <- lm("FREQ~YEAR+MONTH+STRATUM+DEPTH+BOTTOM.TEMP+SALINITY+LATITUDE",
data=new_data2)
vif(lmmod)

##                    GVIF Df GVIF^(1/(2*Df))
## YEAR           5.001167 30        1.027191
## MONTH          2.497439  2        1.257111
## STRATUM     1055.336454 23        1.163392
## DEPTH          1.446028  1        1.202509
## BOTTOM.TEMP    2.856435  1        1.690099
## SALINITY       2.591571  1        1.609836
## LATITUDE     472.358096  1       21.733801

# remove STRATUM
lmmod <- lm("FREQ~YEAR+MONTH+DEPTH+BOTTOM.TEMP+SALINITY+LATITUDE",
data=new_data2)
vif(lmmod)

##                 GVIF Df GVIF^(1/(2*Df))
## YEAR        3.633681 30        1.021737
## MONTH       2.367093  2        1.240377
## DEPTH       1.261159  1        1.123013
## BOTTOM.TEMP 2.505277  1        1.582807
## SALINITY    1.754444  1        1.324554
## LATITUDE    1.445174  1        1.202154
```

```r
# remove bottom temp
lmmod <- lm("FREQ~YEAR+MONTH+DEPTH+SALINITY+LATITUDE", data=new_data2)
vif(lmmod)
```

```
##              GVIF Df GVIF^(1/(2*Df))
## YEAR     2.841149 30        1.017556
## MONTH    1.423826  2        1.092356
## DEPTH    1.260275  1        1.122620
## SALINITY 1.722372  1        1.312392
## LATITUDE 1.032862  1        1.016298
```

```r
# remove month instead
lmmod <- lm("FREQ~YEAR+BOTTOM.TEMP+DEPTH+SALINITY+LATITUDE", data=new_data2)
vif(lmmod)
```

```
##                 GVIF Df GVIF^(1/(2*Df))
## YEAR        2.572145 30        1.015870
## BOTTOM.TEMP 1.506945  1        1.227577
## DEPTH       1.258781  1        1.121954
## SALINITY    1.745664  1        1.321236
## LATITUDE    1.248355  1        1.117298
```

```r
# STRATUM instead of LATITUDE
lmmod <- lm("FREQ~YEAR+BOTTOM.TEMP+DEPTH+SALINITY+STRATUM", data=new_data2)
vif(lmmod)
```

```
##                 GVIF Df GVIF^(1/(2*Df))
## YEAR        3.507102 30        1.021133
## BOTTOM.TEMP 1.708724  1        1.307182
## DEPTH       1.440777  1        1.200324
## SALINITY    2.583365  1        1.607285
## STRATUM     2.645848 23        1.021377
```

```r
# AREA instead of LATITUDE
lmmod <- lm("FREQ~YEAR+BOTTOM.TEMP+DEPTH+SALINITY+AREA", data=new_data2)
vif(lmmod)
```

```
##                 GVIF Df GVIF^(1/(2*Df))
## YEAR        3.103840 30        1.019057
## BOTTOM.TEMP 1.644901  1        1.282537
## DEPTH       1.289535  1        1.135577
## SALINITY    2.343139  1        1.530732
## AREA        1.977670  5        1.070571
```

```r
# Remove DEPTH
lmmod <- lm("FREQ~YEAR+BOTTOM.TEMP+SALINITY+AREA", data=new_data2)
vif(lmmod)
```

```
##                 GVIF Df GVIF^(1/(2*Df))
## YEAR        2.561857 30        1.015802
## BOTTOM.TEMP 1.641929  1        1.281378
```

```
## SALINITY    2.288918  1        1.512917
## AREA        1.898112  5        1.066184
```

Step 2: Model selection

Compare negative binomial, zero-inflated negative binomial, and zero-altered negative binomial.

Comparison 1: full models

Full model with latitude as a covariate.

```
full_mod <- compare_models(data = new_data2,
               nb_model = "FREQ~YEAR+DEPTH+BOTTOM.TEMP+SALINITY+LATITUDE",
               zic_model = "FREQ~YEAR+DEPTH+BOTTOM.TEMP+SALINITY+LATITUDE",
               zac_model = "FREQ~YEAR+DEPTH+BOTTOM.TEMP+SALINITY+LATITUDE",
               zi_model = "~ YEAR+DEPTH+BOTTOM.TEMP+SALINITY+LATITUDE",
               za_model = "~ YEAR+DEPTH+BOTTOM.TEMP+SALINITY+LATITUDE",
               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 26.7

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + LATITUDE
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   7465.1   7678.9  -3696.5   7393.1     2767
##
##
## Dispersion parameter for nbinom2 family (): 0.121
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.74653    0.40158   4.349 1.37e-05 ***
## YEAR1990    -3.32379    0.59869  -5.552 2.83e-08 ***
## YEAR1991    -1.46278    0.54913  -2.664 0.007726 **
## YEAR1992    -1.89250    0.52831  -3.582 0.000341 ***
## YEAR1993    -2.75056    0.56679  -4.853 1.22e-06 ***
## YEAR1994    -0.96366    0.51983  -1.854 0.063766 .
## YEAR1995    -1.51751    0.54956  -2.761 0.005757 **
## YEAR1996    -1.38509    0.53962  -2.567 0.010265 *
## YEAR1997    -3.02619    0.61980  -4.883 1.05e-06 ***
## YEAR1998    -2.46187    0.57102  -4.311 1.62e-05 ***
## YEAR1999    -1.57014    0.52217  -3.007 0.002639 **
## YEAR2000    -2.21513    0.55491  -3.992 6.56e-05 ***
## YEAR2001    -0.97878    0.50024  -1.957 0.050392 .
```

```
## YEAR2002    -1.04101    0.50193  -2.074 0.038077 *
## YEAR2003    -0.38800    0.50229  -0.772 0.439836
## YEAR2004    -1.74469    0.52262  -3.338 0.000843 ***
## YEAR2005     0.34249    0.49885   0.687 0.492364
## YEAR2006    -2.50922    0.54818  -4.577 4.71e-06 ***
## YEAR2007    -2.42405    0.54341  -4.461 8.16e-06 ***
## YEAR2008    -1.20247    0.49853  -2.412 0.015863 *
## YEAR2009     0.30238    0.48797   0.620 0.535479
## YEAR2010     0.47387    0.50184   0.944 0.345040
## YEAR2011    -1.13517    0.50469  -2.249 0.024497 *
## YEAR2012    -0.94497    0.51322  -1.841 0.065586 .
## YEAR2013    -1.22974    0.52016  -2.364 0.018072 *
## YEAR2014    -2.48207    0.57545  -4.313 1.61e-05 ***
## YEAR2015    -1.28854    0.50182  -2.568 0.010236 *
## YEAR2016    -0.94933    0.50283  -1.888 0.059031 .
## YEAR2017    -2.61715    0.53759  -4.868 1.13e-06 ***
## YEAR2018    -2.20511    0.57176  -3.857 0.000115 ***
## YEAR2019    -1.89734    0.57441  -3.303 0.000956 ***
## DEPTH        0.12617    0.07346   1.718 0.085867 .
## BOTTOM.TEMP -0.78512    0.08555  -9.177  < 2e-16 ***
## SALINITY    -0.16675    0.07890  -2.114 0.034551 *
## LATITUDE     0.63773    0.06705   9.511  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.59

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + LATITUDE
## Zero inflation:        ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + LATITUDE
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   7362.1   7783.7  -3610.0   7220.1     2732
##
##
## Dispersion parameter for nbinom2 family (): 0.214
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.07576    0.34571   6.004 1.92e-09 ***
## YEAR1990    -3.49681    0.48670  -7.185 6.74e-13 ***
## YEAR1991    -0.17803    0.51925  -0.343 0.731707
## YEAR1992    -0.95695    0.54648  -1.751 0.079927 .
## YEAR1993    -2.75309    0.45948  -5.992 2.08e-09 ***
## YEAR1994    -0.39205    0.47808  -0.820 0.412192
## YEAR1995    -2.25039    0.45522  -4.944 7.67e-07 ***
```

```
## YEAR1996    -1.85103    0.44651  -4.146 3.39e-05 ***
## YEAR1997    -2.91239    0.73570  -3.959 7.54e-05 ***
## YEAR1998    -2.36618    0.64518  -3.667 0.000245 ***
## YEAR1999    -1.18439    0.47644  -2.486 0.012921 *
## YEAR2000    -2.37996    0.45233  -5.262 1.43e-07 ***
## YEAR2001    -0.50979    0.44977  -1.133 0.257031
## YEAR2002    -0.33583    0.50764  -0.662 0.508268
## YEAR2003    -0.99372    0.43140  -2.304 0.021250 *
## YEAR2004    -1.55944    0.55084  -2.831 0.004640 **
## YEAR2005     1.20131    0.50643   2.372 0.017688 *
## YEAR2006    -0.76150    0.73672  -1.034 0.301311
## YEAR2007    -2.28899    0.55506  -4.124 3.73e-05 ***
## YEAR2008    -1.13312    0.45605  -2.485 0.012968 *
## YEAR2009     1.33271    0.50248   2.652 0.007995 **
## YEAR2010     1.54224    0.50318   3.065 0.002177 **
## YEAR2011    -0.85236    0.45250  -1.884 0.059609 .
## YEAR2012    -1.09213    0.45174  -2.418 0.015624 *
## YEAR2013    -1.33897    0.53614  -2.497 0.012509 *
## YEAR2014    -3.38364    0.48110  -7.033 2.02e-12 ***
## YEAR2015    -1.21074    0.46689  -2.593 0.009508 **
## YEAR2016    -0.95818    0.46008  -2.083 0.037284 *
## YEAR2017    -2.67662    0.49508  -5.406 6.43e-08 ***
## YEAR2018    -1.53444    0.60724  -2.527 0.011507 *
## YEAR2019     0.48141    0.66985   0.719 0.472334
## DEPTH        0.14246    0.07270   1.960 0.050047 .
## BOTTOM.TEMP -0.19339    0.08981  -2.153 0.031290 *
## SALINITY    -0.23302    0.08371  -2.784 0.005374 **
## LATITUDE     0.65854    0.07340   8.972  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.565e+00  1.489e+00  -1.051   0.2931
## YEAR1990    -1.853e+01  8.506e+03  -0.002   0.9983
## YEAR1991     2.303e+00  1.556e+00   1.480   0.1388
## YEAR1992     2.450e+00  1.537e+00   1.594   0.1109
## YEAR1993    -1.604e+01  4.742e+03  -0.003   0.9973
## YEAR1994     1.247e+00  1.543e+00   0.809   0.4187
## YEAR1995    -2.158e+01  2.188e+04  -0.001   0.9992
## YEAR1996    -2.072e+01  1.544e+04  -0.001   0.9989
## YEAR1997     9.678e-01  2.044e+00   0.474   0.6358
## YEAR1998     8.834e-01  1.693e+00   0.522   0.6018
## YEAR1999     4.404e-01  1.631e+00   0.270   0.7872
## YEAR2000    -1.155e+01  1.586e+03  -0.007   0.9942
## YEAR2001     1.302e+00  1.557e+00   0.836   0.4032
## YEAR2002     1.491e+00  1.558e+00   0.957   0.3385
## YEAR2003    -2.779e+00  4.639e+00  -0.599   0.5491
## YEAR2004     1.042e+00  1.574e+00   0.662   0.5078
## YEAR2005     1.853e+00  1.533e+00   1.208   0.2269
```

```
## YEAR2006     3.197e+00  1.568e+00   2.038   0.0415 *
## YEAR2007     1.103e+00  1.732e+00   0.637   0.5242
## YEAR2008     5.162e-01  1.621e+00   0.319   0.7501
## YEAR2009     2.179e+00  1.533e+00   1.421   0.1553
## YEAR2010     2.408e+00  1.521e+00   1.583   0.1133
## YEAR2011     6.645e-01  1.609e+00   0.413   0.6796
## YEAR2012    -5.303e-02  1.648e+00  -0.032   0.9743
## YEAR2013     1.026e+00  1.623e+00   0.632   0.5273
## YEAR2014    -2.261e+01  3.446e+04  -0.001   0.9995
## YEAR2015     3.811e-01  1.583e+00   0.241   0.8097
## YEAR2016     8.393e-01  1.568e+00   0.535   0.5925
## YEAR2017    -5.068e-01  2.777e+00  -0.182   0.8552
## YEAR2018     1.400e+00  1.629e+00   0.860   0.3899
## YEAR2019     3.692e+00  1.591e+00   2.321   0.0203 *
## DEPTH        2.011e-01  1.188e-01   1.693   0.0904 .
## BOTTOM.TEMP  1.160e+00  1.870e-01   6.203 5.53e-10 ***
## SALINITY    -1.631e-01  1.430e-01  -1.141   0.2540
## LATITUDE    -1.308e-01  1.326e-01  -0.986   0.3241
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.179

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + LATITUDE
## Zero inflation:        ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + LATITUDE
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   7326.7   7748.4  -3592.4   7184.7     2732
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.03
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.30705    1.26254   0.243 0.807853
## YEAR1990    -4.33816    0.73423  -5.908 3.45e-09 ***
## YEAR1991    -0.29166    0.68045  -0.429 0.668191
## YEAR1992    -1.05995    0.67250  -1.576 0.114997
## YEAR1993    -3.40434    0.61560  -5.530 3.20e-08 ***
## YEAR1994    -0.52745    0.59058  -0.893 0.371798
## YEAR1995    -2.30504    0.60044  -3.839 0.000124 ***
## YEAR1996    -2.34799    0.54813  -4.284 1.84e-05 ***
## YEAR1997    -2.62661    0.86482  -3.037 0.002388 **
## YEAR1998    -2.40739    0.74199  -3.244 0.001177 **
## YEAR1999    -1.22735    0.57796  -2.124 0.033703 *
```

```
## YEAR2000    -2.33988    0.64963  -3.602 0.000316 ***
## YEAR2001    -0.57858    0.55177  -1.049 0.294369
## YEAR2002    -0.26239    0.64051  -0.410 0.682057
## YEAR2003    -0.73345    0.53966  -1.359 0.174114
## YEAR2004    -1.49143    0.65507  -2.277 0.022801 *
## YEAR2005     1.28019    0.64583   1.982 0.047454 *
## YEAR2006    -0.74692    0.89995  -0.830 0.406561
## YEAR2007    -1.94412    0.72562  -2.679 0.007379 **
## YEAR2008    -1.09422    0.56238  -1.946 0.051693 .
## YEAR2009     1.36419    0.64253   2.123 0.033740 *
## YEAR2010     1.55146    0.64386   2.410 0.015969 *
## YEAR2011    -0.73488    0.57455  -1.279 0.200880
## YEAR2012    -1.14037    0.56611  -2.014 0.043968 *
## YEAR2013    -0.63894    0.67655  -0.944 0.344961
## YEAR2014    -3.17460    0.76165  -4.168 3.07e-05 ***
## YEAR2015    -1.32608    0.57161  -2.320 0.020347 *
## YEAR2016    -0.61579    0.57859  -1.064 0.287198
## YEAR2017    -2.71922    0.61161  -4.446 8.75e-06 ***
## YEAR2018    -1.45563    0.73836  -1.971 0.048674 *
## YEAR2019     0.06820    0.84860   0.080 0.935942
## DEPTH        0.26439    0.10422   2.537 0.011187 *
## BOTTOM.TEMP -0.36435    0.12771  -2.853 0.004332 **
## SALINITY    -0.19221    0.11202  -1.716 0.086196 .
## LATITUDE     0.75865    0.09701   7.820 5.27e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.18861    0.28693   0.657 0.510974
## YEAR1990     1.54668    0.41541   3.723 0.000197 ***
## YEAR1991     1.10552    0.38312   2.886 0.003907 **
## YEAR1992     1.42991    0.39901   3.584 0.000339 ***
## YEAR1993     0.98084    0.38881   2.523 0.011646 *
## YEAR1994     0.66445    0.37288   1.782 0.074758 .
## YEAR1995     0.40960    0.37752   1.085 0.277937
## YEAR1996    -0.14758    0.37640  -0.392 0.694993
## YEAR1997     2.15806    0.49941   4.321 1.55e-05 ***
## YEAR1998     1.41305    0.42217   3.347 0.000817 ***
## YEAR1999     0.59773    0.37917   1.576 0.114935
## YEAR2000     1.16818    0.39656   2.946 0.003222 **
## YEAR2001     0.58689    0.35819   1.638 0.101320
## YEAR2002     0.96472    0.38308   2.518 0.011791 *
## YEAR2003     0.03799    0.35878   0.106 0.915664
## YEAR2004     1.25682    0.39922   3.148 0.001643 **
## YEAR2005     0.90392    0.37565   2.406 0.016116 *
## YEAR2006     2.20215    0.47270   4.659 3.18e-06 ***
## YEAR2007     1.79763    0.42994   4.181 2.90e-05 ***
## YEAR2008     0.63956    0.36287   1.762 0.077986 .
## YEAR2009     0.99683    0.37392   2.666 0.007678 **
```

```
## YEAR2010     1.09678    0.37147   2.953 0.003152 **
## YEAR2011     0.80331    0.35815   2.243 0.024900 *
## YEAR2012     0.54721    0.36407   1.503 0.132831
## YEAR2013     1.22938    0.38442   3.198 0.001384 **
## YEAR2014     1.38521    0.40687   3.405 0.000663 ***
## YEAR2015     0.61602    0.36547   1.686 0.091884 .
## YEAR2016     0.92030    0.35909   2.563 0.010381 *
## YEAR2017     1.30895    0.38820   3.372 0.000747 ***
## YEAR2018     1.20664    0.44218   2.729 0.006356 **
## YEAR2019     1.57919    0.43805   3.605 0.000312 ***
## DEPTH        0.09821    0.05200   1.889 0.058938 .
## BOTTOM.TEMP  0.49921    0.05638   8.855  < 2e-16 ***
## SALINITY     0.04551    0.06010   0.757 0.448911
## LATITUDE    -0.36267    0.05409  -6.704 2.02e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.769

## neg_binom dispersal: 1.59

## zero_infl dispersal: 1.179

## zero_adj dispersal: 2.769

AICtab(full_mod)

##           dAIC  df
## zero_adj    0.0 71
## zero_infl  35.3 71
## neg_binom 138.4 36
```

Full model with stratum as a covariate instead of latitude.

```
full_mod_stratum <- compare_models(data = new_data2,
               nb_model = "FREQ~YEAR+DEPTH+BOTTOM.TEMP+SALINITY+STRATUM",
               zic_model = "FREQ~YEAR+DEPTH+BOTTOM.TEMP+SALINITY+STRATUM",
               zac_model = "FREQ~YEAR+DEPTH+BOTTOM.TEMP+SALINITY+STRATUM",
               zi_model = "~DEPTH+BOTTOM.TEMP+SALINITY+STRATUM",
               za_model = "~ YEAR+DEPTH+BOTTOM.TEMP+SALINITY+STRATUM",
               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 26.7

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## Data: data
```

```
##
##      AIC      BIC   logLik deviance df.resid
##   7240.2   7584.6  -3562.1   7124.2     2745
##
##
## Dispersion parameter for nbinom2 family (): 0.157
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.16598    0.69182  -1.685 0.091918 .
## YEAR1990    -2.67352    0.56379  -4.742 2.12e-06 ***
## YEAR1991    -0.82235    0.51420  -1.599 0.109759
## YEAR1992    -1.74683    0.48377  -3.611 0.000305 ***
## YEAR1993    -2.02862    0.52312  -3.878 0.000105 ***
## YEAR1994    -0.74635    0.48318  -1.545 0.122427
## YEAR1995    -0.98511    0.51048  -1.930 0.053636 .
## YEAR1996    -0.93410    0.51202  -1.824 0.068101 .
## YEAR1997    -2.92655    0.62322  -4.696 2.66e-06 ***
## YEAR1998    -2.39358    0.55234  -4.334 1.47e-05 ***
## YEAR1999    -1.41393    0.49331  -2.866 0.004154 **
## YEAR2000    -2.20778    0.51932  -4.251 2.13e-05 ***
## YEAR2001    -0.14331    0.47006  -0.305 0.760467
## YEAR2002    -1.10743    0.47287  -2.342 0.019184 *
## YEAR2003     0.54531    0.47449   1.149 0.250453
## YEAR2004    -1.64579    0.49738  -3.309 0.000936 ***
## YEAR2005    -0.16826    0.47208  -0.356 0.721532
## YEAR2006    -2.01629    0.49913  -4.040 5.35e-05 ***
## YEAR2007    -2.33097    0.51573  -4.520 6.19e-06 ***
## YEAR2008    -0.80144    0.45888  -1.747 0.080721 .
## YEAR2009     0.05666    0.48327   0.117 0.906663
## YEAR2010    -0.53962    0.47376  -1.139 0.254691
## YEAR2011    -0.69392    0.48056  -1.444 0.148744
## YEAR2012    -0.66664    0.51102  -1.305 0.192056
## YEAR2013    -1.72680    0.49452  -3.492 0.000480 ***
## YEAR2014    -2.44538    0.54241  -4.508 6.53e-06 ***
## YEAR2015    -0.82521    0.47359  -1.742 0.081426 .
## YEAR2016    -1.61388    0.47777  -3.378 0.000730 ***
## YEAR2017    -2.50039    0.50029  -4.998 5.80e-07 ***
## YEAR2018    -1.39467    0.54295  -2.569 0.010209 *
## YEAR2019    -1.26655    0.56030  -2.261 0.023790 *
## DEPTH       -0.10557    0.07246  -1.457 0.145121
## BOTTOM.TEMP -1.06900    0.08135 -13.140  < 2e-16 ***
## SALINITY    -0.15399    0.09850  -1.563 0.117988
## STRATUM23    3.01411    0.68975   4.370 1.24e-05 ***
## STRATUM25    2.91406    0.66579   4.377 1.20e-05 ***
## STRATUM27    3.77187    0.65890   5.724 1.04e-08 ***
## STRATUM29    3.48997    0.66782   5.226 1.73e-07 ***
## STRATUM31    1.99074    0.67902   2.932 0.003370 **
## STRATUM33    1.17088    0.64896   1.804 0.071191 .
## STRATUM35    0.48976    0.66002   0.742 0.458065
```

```
## STRATUM37    1.17389    0.65311   1.797 0.072273 .
## STRATUM39    1.12509    0.66507   1.692 0.090708 .
## STRATUM41    1.60267    0.64350   2.491 0.012754 *
## STRATUM43    2.67776    0.67525   3.966 7.32e-05 ***
## STRATUM45    1.77815    0.65478   2.716 0.006614 **
## STRATUM47    2.25985    0.64529   3.502 0.000462 ***
## STRATUM49    1.44082    0.64148   2.246 0.024699 *
## STRATUM51    1.93288    0.62670   3.084 0.002041 **
## STRATUM53    2.33229    0.64261   3.629 0.000284 ***
## STRATUM55    2.82518    0.63540   4.446 8.74e-06 ***
## STRATUM57    1.79238    0.65768   2.725 0.006424 **
## STRATUM59    3.63046    0.65019   5.584 2.35e-08 ***
## STRATUM61    4.28900    0.65898   6.509 7.59e-11 ***
## STRATUM63    4.01198    0.65197   6.154 7.57e-10 ***
## STRATUM65    5.04117    0.65637   7.680 1.59e-14 ***
## STRATUM67    4.33052    0.66904   6.473 9.62e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.673

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## Zero inflation:        ~DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   7206.7   7711.5  -3518.4   7036.7     2718
##
##
## Dispersion parameter for nbinom2 family (): 0.228
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.80628    0.81032  -0.995 0.319726
## YEAR1990    -2.60645    0.52618  -4.954 7.29e-07 ***
## YEAR1991    -0.90423    0.49878  -1.813 0.069848 .
## YEAR1992    -1.63816    0.47283  -3.465 0.000531 ***
## YEAR1993    -1.92618    0.50809  -3.791 0.000150 ***
## YEAR1994    -0.55041    0.45396  -1.212 0.225338
## YEAR1995    -0.71457    0.49633  -1.440 0.149952
## YEAR1996    -0.69674    0.47535  -1.466 0.142720
## YEAR1997    -2.69270    0.59595  -4.518 6.23e-06 ***
## YEAR1998    -2.00581    0.53992  -3.715 0.000203 ***
## YEAR1999    -1.07450    0.45830  -2.345 0.019052 *
## YEAR2000    -2.10506    0.50766  -4.147 3.38e-05 ***
## YEAR2001    -0.37579    0.44480  -0.845 0.398186
```

```
## YEAR2002    -0.64477    0.46249  -1.394 0.163282    
## YEAR2003     0.41479    0.45091   0.920 0.357626    
## YEAR2004    -1.39381    0.46954  -2.968 0.002993 ** 
## YEAR2005     0.25933    0.45154   0.574 0.565744    
## YEAR2006    -1.65069    0.51213  -3.223 0.001268 ** 
## YEAR2007    -2.14768    0.52343  -4.103 4.08e-05 ***
## YEAR2008    -0.86351    0.43685  -1.977 0.048079 *  
## YEAR2009     0.11579    0.47973   0.241 0.809266    
## YEAR2010    -0.22446    0.47540  -0.472 0.636820    
## YEAR2011    -0.77077    0.46762  -1.648 0.099293 .  
## YEAR2012    -0.91842    0.48471  -1.895 0.058120 .  
## YEAR2013    -1.47099    0.48042  -3.062 0.002200 ** 
## YEAR2014    -2.20181    0.53338  -4.128 3.66e-05 ***
## YEAR2015    -0.67169    0.44904  -1.496 0.134696    
## YEAR2016    -1.27761    0.45207  -2.826 0.004712 ** 
## YEAR2017    -2.30876    0.47393  -4.872 1.11e-06 ***
## YEAR2018    -0.79022    0.53475  -1.478 0.139474    
## YEAR2019    -0.83727    0.60241  -1.390 0.164570    
## DEPTH        0.03833    0.08524   0.450 0.652959    
## BOTTOM.TEMP -0.78906    0.10075  -7.832 4.80e-15 ***
## SALINITY     0.13843    0.10333   1.340 0.180348    
## STRATUM23    1.99288    0.83482   2.387 0.016977 *  
## STRATUM25    2.64998    0.82686   3.205 0.001351 ** 
## STRATUM27    2.94460    0.80740   3.647 0.000265 ***
## STRATUM29    2.61727    0.79953   3.273 0.001062 ** 
## STRATUM31    2.02526    0.87615   2.312 0.020802 *  
## STRATUM33    0.97628    0.78142   1.249 0.211533    
## STRATUM35    0.65362    0.82769   0.790 0.429710    
## STRATUM37    1.91084    0.82238   2.324 0.020150 *  
## STRATUM39    1.48984    0.82640   1.803 0.071417 .  
## STRATUM41    2.32578    0.82225   2.829 0.004676 ** 
## STRATUM43    2.96292    0.82852   3.576 0.000349 ***
## STRATUM45    1.75066    0.81187   2.156 0.031059 *  
## STRATUM47    1.96305    0.77968   2.518 0.011811 *  
## STRATUM49    1.71114    0.81462   2.101 0.035682 *  
## STRATUM51    2.11663    0.77942   2.716 0.006615 ** 
## STRATUM53    1.96299    0.77582   2.530 0.011400 *  
## STRATUM55    2.28237    0.76626   2.979 0.002896 ** 
## STRATUM57    1.85847    0.83095   2.237 0.025315 *  
## STRATUM59    3.37658    0.79174   4.265 2.00e-05 ***
## STRATUM61    3.72529    0.78427   4.750 2.03e-06 ***
## STRATUM63    3.62733    0.77997   4.651 3.31e-06 ***
## STRATUM65    5.22110    0.79521   6.566 5.18e-11 ***
## STRATUM67    4.71853    0.81077   5.820 5.89e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.2256     1.4339  -0.855  0.39269
```

```
## DEPTH         0.5107     0.1726   2.958  0.00309 **
## BOTTOM.TEMP   1.1005     0.2121   5.189 2.12e-07 ***
## SALINITY      0.9080     0.2257   4.022 5.76e-05 ***
## STRATUM23    -2.5559     1.8092  -1.413  0.15774
## STRATUM25    -0.7540     1.6038  -0.470  0.63824
## STRATUM27    -1.8140     1.6705  -1.086  0.27753
## STRATUM29    -3.6482     3.7356  -0.977  0.32877
## STRATUM31     0.8081     1.5896   0.508  0.61120
## STRATUM33    -1.3391     2.2061  -0.607  0.54385
## STRATUM35     0.7965     1.8200   0.438  0.66164
## STRATUM37     2.4657     1.5150   1.627  0.10364
## STRATUM39     1.5247     1.5256   0.999  0.31758
## STRATUM41     2.5553     1.5143   1.688  0.09151 .
## STRATUM43     1.5292     1.5224   1.004  0.31517
## STRATUM45     0.5510     1.6981   0.324  0.74559
## STRATUM47    -1.6665     2.7690  -0.602  0.54727
## STRATUM49     1.0414     1.5115   0.689  0.49081
## STRATUM51     0.8936     1.4727   0.607  0.54401
## STRATUM53    -0.8423     1.6927  -0.498  0.61876
## STRATUM55    -2.2213     2.5545  -0.870  0.38454
## STRATUM57     0.5904     1.5260   0.387  0.69885
## STRATUM59    -0.6500     1.5121  -0.430  0.66730
## STRATUM61    -1.2441     1.6545  -0.752  0.45208
## STRATUM63    -1.3627     1.7460  -0.781  0.43509
## STRATUM65     0.4526     1.4735   0.307  0.75870
## STRATUM67     0.4319     1.4961   0.289  0.77283
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.355

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## Zero inflation:        ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   7142.6   7825.5  -3456.3   6912.6     2688
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.139
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.41792    1.23169  -1.151 0.249652
## YEAR1990    -3.41082    0.74238  -4.594 4.34e-06 ***
## YEAR1991    -0.07055    0.67545  -0.104 0.916817
```

```
## YEAR1992    -1.20544    0.64300  -1.875 0.060834 .
## YEAR1993    -2.46661    0.60044  -4.108 3.99e-05 ***
## YEAR1994    -0.31195    0.55195  -0.565 0.571952
## YEAR1995    -1.37501    0.58897  -2.335 0.019565 *
## YEAR1996    -1.74912    0.54767  -3.194 0.001404 **
## YEAR1997    -2.44792    0.89614  -2.732 0.006302 **
## YEAR1998    -2.17487    0.74800  -2.908 0.003642 **
## YEAR1999    -1.02529    0.54378  -1.885 0.059367 .
## YEAR2000    -2.31701    0.60472  -3.832 0.000127 ***
## YEAR2001     0.17571    0.52920   0.332 0.739869
## YEAR2002    -0.12411    0.58142  -0.213 0.830974
## YEAR2003     0.08210    0.52998   0.155 0.876896
## YEAR2004    -1.11136    0.62062  -1.791 0.073336 .
## YEAR2005     0.95895    0.58884   1.629 0.103409
## YEAR2006     0.06785    0.80497   0.084 0.932826
## YEAR2007    -1.50108    0.68368  -2.196 0.028122 *
## YEAR2008    -0.50604    0.52119  -0.971 0.331584
## YEAR2009     0.76209    0.58146   1.311 0.189978
## YEAR2010     0.68952    0.59983   1.150 0.250339
## YEAR2011    -0.40278    0.55930  -0.720 0.471433
## YEAR2012    -1.09561    0.58128  -1.885 0.059456 .
## YEAR2013    -0.96339    0.64248  -1.499 0.133751
## YEAR2014    -3.05552    0.74258  -4.115 3.88e-05 ***
## YEAR2015    -0.69613    0.55365  -1.257 0.208627
## YEAR2016    -0.89567    0.58257  -1.537 0.124183
## YEAR2017    -2.25219    0.60335  -3.733 0.000189 ***
## YEAR2018    -0.38116    0.70765  -0.539 0.590140
## YEAR2019     0.54689    0.82771   0.661 0.508786
## DEPTH        0.14535    0.10375   1.401 0.161206
## BOTTOM.TEMP -0.67577    0.12008  -5.628 1.83e-08 ***
## SALINITY    -0.09271    0.13140  -0.706 0.480460
## STRATUM23    1.91831    1.24860   1.536 0.124447
## STRATUM25    2.52774    1.20580   2.096 0.036054 *
## STRATUM27    2.94297    1.21877   2.415 0.015748 *
## STRATUM29    2.38524    1.21927   1.956 0.050433 .
## STRATUM31    2.61324    1.32336   1.975 0.048302 *
## STRATUM33    1.48438    1.19646   1.241 0.214740
## STRATUM35    0.83977    1.23110   0.682 0.495157
## STRATUM37    1.56533    1.21597   1.287 0.197987
## STRATUM39    1.36096    1.22034   1.115 0.264751
## STRATUM41    2.58672    1.20606   2.145 0.031971 *
## STRATUM43    3.02150    1.24858   2.420 0.015522 *
## STRATUM45    1.70254    1.17133   1.454 0.146081
## STRATUM47    1.62828    1.16967   1.392 0.163897
## STRATUM49    1.46630    1.19660   1.225 0.220429
## STRATUM51    2.27949    1.15122   1.980 0.047696 *
## STRATUM53    2.35129    1.17655   1.998 0.045666 *
## STRATUM55    2.47188    1.16225   2.127 0.033437 *
## STRATUM57    1.86480    1.20540   1.547 0.121853
## STRATUM59    3.35780    1.18800   2.826 0.004707 **
```

```
## STRATUM61    3.89823    1.17870   3.307 0.000942 ***
## STRATUM63    4.03051    1.18303   3.407 0.000657 ***
## STRATUM65    5.00826    1.18809   4.215 2.49e-05 ***
## STRATUM67    4.22737    1.19232   3.546 0.000392 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.62768    0.60119   2.707 0.006781 **
## YEAR1990     1.55320    0.42551   3.650 0.000262 ***
## YEAR1991     1.24483    0.39910   3.119 0.001814 **
## YEAR1992     1.42278    0.40879   3.480 0.000501 ***
## YEAR1993     0.90143    0.40183   2.243 0.024877 *
## YEAR1994     0.63583    0.38513   1.651 0.098752 .
## YEAR1995     0.20172    0.38956   0.518 0.604592
## YEAR1996    -0.34358    0.38971  -0.882 0.377985
## YEAR1997     2.22289    0.51215   4.340 1.42e-05 ***
## YEAR1998     1.26928    0.43243   2.935 0.003333 **
## YEAR1999     0.50858    0.38869   1.308 0.190724
## YEAR2000     1.08597    0.41045   2.646 0.008150 **
## YEAR2001     0.71877    0.37051   1.940 0.052384 .
## YEAR2002     0.95316    0.40119   2.376 0.017509 *
## YEAR2003    -0.09803    0.37049  -0.265 0.791333
## YEAR2004     1.28240    0.41505   3.090 0.002003 **
## YEAR2005     0.91545    0.39552   2.315 0.020637 *
## YEAR2006     2.20951    0.48424   4.563 5.05e-06 ***
## YEAR2007     1.73642    0.44488   3.903 9.50e-05 ***
## YEAR2008     0.59990    0.37447   1.602 0.109162
## YEAR2009     1.05556    0.39589   2.666 0.007670 **
## YEAR2010     1.16497    0.39273   2.966 0.003014 **
## YEAR2011     0.97825    0.37348   2.619 0.008812 **
## YEAR2012     0.57842    0.37897   1.526 0.126939
## YEAR2013     1.31909    0.39610   3.330 0.000868 ***
## YEAR2014     1.32763    0.41820   3.175 0.001500 **
## YEAR2015     0.61393    0.37723   1.627 0.103635
## YEAR2016     1.09917    0.36720   2.993 0.002759 **
## YEAR2017     1.44687    0.39761   3.639 0.000274 ***
## YEAR2018     1.09686    0.45753   2.397 0.016514 *
## YEAR2019     1.75496    0.48536   3.616 0.000299 ***
## DEPTH        0.23909    0.05859   4.081 4.48e-05 ***
## BOTTOM.TEMP  0.75531    0.06582  11.476  < 2e-16 ***
## SALINITY     0.06659    0.07440   0.895 0.370810
## STRATUM23   -2.03208    0.60636  -3.351 0.000804 ***
## STRATUM25   -1.94255    0.58738  -3.307 0.000943 ***
## STRATUM27   -2.15385    0.58043  -3.711 0.000207 ***
## STRATUM29   -2.40334    0.58637  -4.099 4.15e-05 ***
## STRATUM31   -0.98326    0.59724  -1.646 0.099694 .
## STRATUM33   -0.72305    0.58599  -1.234 0.217242
## STRATUM35   -0.31878    0.58897  -0.541 0.588331
```

```
## STRATUM37   -0.50151    0.58724  -0.854 0.393097
## STRATUM39   -0.69765    0.59572  -1.171 0.241553
## STRATUM41   -0.41136    0.58580  -0.702 0.482544
## STRATUM43   -1.14701    0.59249  -1.936 0.052879 .
## STRATUM45   -1.07041    0.59250  -1.807 0.070827 .
## STRATUM47   -1.74465    0.56826  -3.070 0.002139 **
## STRATUM49   -0.73289    0.57959  -1.264 0.206055
## STRATUM51   -1.05749    0.56880  -1.859 0.063006 .
## STRATUM53   -1.68031    0.57532  -2.921 0.003493 **
## STRATUM55   -2.03210    0.56921  -3.570 0.000357 ***
## STRATUM57   -0.93904    0.58187  -1.614 0.106563
## STRATUM59   -2.16652    0.57548  -3.765 0.000167 ***
## STRATUM61   -2.38043    0.58391  -4.077 4.57e-05 ***
## STRATUM63   -2.28896    0.57597  -3.974 7.06e-05 ***
## STRATUM65   -2.52375    0.58209  -4.336 1.45e-05 ***
## STRATUM67   -2.62877    0.58798  -4.471 7.79e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.344

## neg_binom dispersal: 1.673

## zero_infl dispersal: 1.355

## zero_adj dispersal: 2.344

AICtab(full_mod_stratum)

##           dAIC  df
## zero_adj    0.0 115
## zero_infl  64.1 85
## neg_binom  97.6 58
```

Full model with state as a covariate instead of latitude.

```
full_mod_st <- compare_models(data = new_data2,
               nb_model = "FREQ~YEAR+DEPTH+BOTTOM.TEMP+SALINITY+STATE",
               zic_model = "FREQ~YEAR+DEPTH+BOTTOM.TEMP+SALINITY+STATE",
               zac_model = "FREQ~YEAR+DEPTH+BOTTOM.TEMP+SALINITY+STATE",
               zi_model = "~DEPTH+BOTTOM.TEMP+SALINITY+STATE",
               za_model = "~ YEAR+DEPTH+BOTTOM.TEMP+SALINITY+STATE",
               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 26.7

## Negative binomial

## model summary:
```

```
##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STATE
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   7293.7   7519.4  -3608.9   7217.7     2765
##
##
## Dispersion parameter for nbinom2 family (): 0.142
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.39078    0.41063   5.822 5.81e-09 ***
## YEAR1990    -2.99909    0.56161  -5.340 9.29e-08 ***
## YEAR1991    -1.07707    0.51101  -2.108 0.035055 *
## YEAR1992    -1.71840    0.48806  -3.521 0.000430 ***
## YEAR1993    -2.38561    0.53693  -4.443 8.87e-06 ***
## YEAR1994    -0.95798    0.48235  -1.986 0.047026 *
## YEAR1995    -1.57721    0.50653  -3.114 0.001847 **
## YEAR1996    -1.58014    0.50243  -3.145 0.001661 **
## YEAR1997    -2.97264    0.62934  -4.723 2.32e-06 ***
## YEAR1998    -2.73431    0.55420  -4.934 8.06e-07 ***
## YEAR1999    -1.85306    0.48732  -3.803 0.000143 ***
## YEAR2000    -2.34448    0.51921  -4.515 6.32e-06 ***
## YEAR2001    -0.55894    0.46965  -1.190 0.234003
## YEAR2002    -1.33301    0.47608  -2.800 0.005111 **
## YEAR2003     0.02464    0.46876   0.053 0.958071
## YEAR2004    -2.05628    0.49488  -4.155 3.25e-05 ***
## YEAR2005    -0.38695    0.46690  -0.829 0.407233
## YEAR2006    -2.41024    0.50132  -4.808 1.53e-06 ***
## YEAR2007    -2.23806    0.51077  -4.382 1.18e-05 ***
## YEAR2008    -0.89460    0.46485  -1.924 0.054294 .
## YEAR2009     0.20229    0.46990   0.431 0.666831
## YEAR2010    -0.34773    0.47156  -0.737 0.460879
## YEAR2011    -0.88691    0.47946  -1.850 0.064340 .
## YEAR2012    -1.23092    0.48117  -2.558 0.010523 *
## YEAR2013    -1.97542    0.48088  -4.108 3.99e-05 ***
## YEAR2014    -2.68666    0.53666  -5.006 5.55e-07 ***
## YEAR2015    -1.38753    0.47062  -2.948 0.003195 **
## YEAR2016    -1.99415    0.46974  -4.245 2.18e-05 ***
## YEAR2017    -2.57800    0.50552  -5.100 3.40e-07 ***
## YEAR2018    -1.89213    0.53861  -3.513 0.000443 ***
## YEAR2019    -2.82865    0.57989  -4.878 1.07e-06 ***
## DEPTH       -0.02433    0.07029  -0.346 0.729215
## BOTTOM.TEMP -1.12591    0.08378 -13.439  < 2e-16 ***
## SALINITY    -0.34680    0.07969  -4.352 1.35e-05 ***
## STATEGA     -2.40346    0.24589  -9.774  < 2e-16 ***
## STATENC      0.60565    0.21789   2.780 0.005443 **
## STATESC     -1.24448    0.22920  -5.430 5.64e-08 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.461

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STATE
## Zero inflation:        ~DEPTH + BOTTOM.TEMP + SALINITY + STATE
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   7277.3   7544.6  -3593.7   7187.3     2758
##
##
## Dispersion parameter for nbinom2 family (): 0.197
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.35390    0.40829   5.765 8.16e-09 ***
## YEAR1990    -3.00215    0.53157  -5.648 1.63e-08 ***
## YEAR1991    -0.82238    0.49674  -1.656 0.097810 .
## YEAR1992    -1.59526    0.46567  -3.426 0.000613 ***
## YEAR1993    -2.39730    0.51377  -4.666 3.07e-06 ***
## YEAR1994    -0.87950    0.44840  -1.961 0.049833 *
## YEAR1995    -1.70774    0.47640  -3.585 0.000338 ***
## YEAR1996    -1.65676    0.46123  -3.592 0.000328 ***
## YEAR1997    -3.02504    0.60374  -5.010 5.43e-07 ***
## YEAR1998    -2.62786    0.53579  -4.905 9.36e-07 ***
## YEAR1999    -1.65194    0.45137  -3.660 0.000252 ***
## YEAR2000    -2.35674    0.49922  -4.721 2.35e-06 ***
## YEAR2001    -0.54095    0.44154  -1.225 0.220515
## YEAR2002    -1.10330    0.46702  -2.362 0.018156 *
## YEAR2003    -0.06441    0.44784  -0.144 0.885637
## YEAR2004    -2.03374    0.46561  -4.368 1.25e-05 ***
## YEAR2005    -0.18731    0.45192  -0.414 0.678520
## YEAR2006    -2.16324    0.50089  -4.319 1.57e-05 ***
## YEAR2007    -2.21786    0.51098  -4.340 1.42e-05 ***
## YEAR2008    -0.91020    0.44201  -2.059 0.039472 *
## YEAR2009     0.49180    0.47155   1.043 0.296968
## YEAR2010    -0.05265    0.47326  -0.111 0.911420
## YEAR2011    -1.01511    0.45317  -2.240 0.025090 *
## YEAR2012    -1.29106    0.46426  -2.781 0.005421 **
## YEAR2013    -1.91434    0.46718  -4.098 4.17e-05 ***
## YEAR2014    -2.68120    0.53024  -5.057 4.27e-07 ***
## YEAR2015    -1.41410    0.44308  -3.192 0.001415 **
## YEAR2016    -1.94919    0.44111  -4.419 9.92e-06 ***
## YEAR2017    -2.58813    0.47835  -5.411 6.28e-08 ***
```

```
## YEAR2018    -1.62584    0.51775  -3.140 0.001688 **
## YEAR2019    -2.46336    0.57097  -4.314 1.60e-05 ***
## DEPTH        0.05311    0.08352   0.636 0.524834
## BOTTOM.TEMP -0.87369    0.10490  -8.328  < 2e-16 ***
## SALINITY    -0.21168    0.08354  -2.534 0.011286 *
## STATEGA     -1.93234    0.32984  -5.858 4.67e-09 ***
## STATENC      0.94104    0.25954   3.626 0.000288 ***
## STATESC     -0.82259    0.29531  -2.785 0.005345 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.1329     0.6405  -3.330 0.000868 ***
## DEPTH         0.3388     0.1735   1.952 0.050879 .
## BOTTOM.TEMP   0.9034     0.2020   4.473 7.72e-06 ***
## SALINITY      0.6602     0.2184   3.023 0.002502 **
## STATEGA       1.5570     0.6887   2.261 0.023771 *
## STATENC       0.7424     0.5237   1.418 0.156314
## STATESC       1.2730     0.6077   2.095 0.036178 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.336

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STATE
## Zero inflation:        ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STATE
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   7200.6   7646.0  -3525.3   7050.6     2728
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.0641
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.2388     0.7016   1.766  0.07745 .
## YEAR1990     -3.9021     0.7229  -5.398 6.74e-08 ***
## YEAR1991     -0.1019     0.6686  -0.152  0.87890
## YEAR1992     -0.9915     0.6465  -1.534  0.12513
## YEAR1993     -3.3125     0.6322  -5.240 1.61e-07 ***
## YEAR1994     -0.6248     0.5556  -1.125  0.26077
## YEAR1995     -2.3754     0.5636  -4.215 2.50e-05 ***
## YEAR1996     -2.5781     0.5233  -4.927 8.36e-07 ***
## YEAR1997     -2.7309     0.8855  -3.084  0.00204 **
```

```
## YEAR1998     -3.0571     0.7209  -4.241 2.23e-05 ***
## YEAR1999     -1.4851     0.5526  -2.687  0.00720 **
## YEAR2000     -2.6295     0.6155  -4.272 1.93e-05 ***
## YEAR2001     -0.2439     0.5346  -0.456  0.64825
## YEAR2002     -0.6555     0.6073  -1.079  0.28046
## YEAR2003     -0.2573     0.5286  -0.487  0.62641
## YEAR2004     -1.8506     0.6263  -2.955  0.00313 **
## YEAR2005      0.5374     0.6118   0.878  0.37973
## YEAR2006     -0.9994     0.8226  -1.215  0.22438
## YEAR2007     -1.9128     0.6923  -2.763  0.00572 **
## YEAR2008     -0.7669     0.5408  -1.418  0.15612
## YEAR2009      0.9974     0.6054   1.647  0.09947 .
## YEAR2010      0.6707     0.6251   1.073  0.28326
## YEAR2011     -0.7796     0.5522  -1.412  0.15796
## YEAR2012     -1.4584     0.5440  -2.681  0.00735 **
## YEAR2013     -1.5806     0.6193  -2.552  0.01071 *
## YEAR2014     -3.4828     0.7242  -4.809 1.52e-06 ***
## YEAR2015     -1.5302     0.5439  -2.813  0.00490 **
## YEAR2016     -1.7105     0.5539  -3.088  0.00201 **
## YEAR2017     -2.7578     0.6051  -4.557 5.18e-06 ***
## YEAR2018     -1.1683     0.7161  -1.631  0.10281
## YEAR2019     -1.4693     0.8218  -1.788  0.07380 .
## DEPTH         0.1626     0.1013   1.604  0.10870
## BOTTOM.TEMP  -0.7307     0.1335  -5.472 4.46e-08 ***
## SALINITY     -0.2934     0.1101  -2.665  0.00770 **
## STATEGA      -1.8592     0.4093  -4.543 5.55e-06 ***
## STATENC       1.0307     0.3224   3.196  0.00139 **
## STATESC      -0.6557     0.3675  -1.784  0.07436 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.23569    0.30879  -0.763 0.445294
## YEAR1990     1.41373    0.41791   3.383 0.000717 ***
## YEAR1991     1.19233    0.38953   3.061 0.002206 **
## YEAR1992     1.37911    0.40066   3.442 0.000577 ***
## YEAR1993     0.84182    0.39606   2.125 0.033546 *
## YEAR1994     0.72450    0.37706   1.921 0.054676 .
## YEAR1995     0.29432    0.38072   0.773 0.439486
## YEAR1996    -0.14982    0.38013  -0.394 0.693485
## YEAR1997     2.05009    0.50879   4.029 5.59e-05 ***
## YEAR1998     1.30239    0.42480   3.066 0.002170 **
## YEAR1999     0.65539    0.38089   1.721 0.085307 .
## YEAR2000     1.04544    0.40142   2.604 0.009205 **
## YEAR2001     0.53455    0.35825   1.492 0.135675
## YEAR2002     0.89719    0.39099   2.295 0.021752 *
## YEAR2003    -0.06476    0.36171  -0.179 0.857915
## YEAR2004     1.34750    0.40504   3.327 0.000878 ***
## YEAR2005     0.94581    0.38120   2.481 0.013096 *
```

```
## YEAR2006      1.97771    0.47432   4.170 3.05e-05 ***
## YEAR2007      1.53228    0.43185   3.548 0.000388 ***
## YEAR2008      0.52338    0.36513   1.433 0.151747
## YEAR2009      0.88157    0.38416   2.295 0.021744 *
## YEAR2010      1.04062    0.37909   2.745 0.006051 **
## YEAR2011      0.88341    0.36291   2.434 0.014924 *
## YEAR2012      0.49471    0.36798   1.344 0.178815
## YEAR2013      1.24373    0.38791   3.206 0.001345 **
## YEAR2014      1.24857    0.40981   3.047 0.002314 **
## YEAR2015      0.70091    0.36926   1.898 0.057676 .
## YEAR2016      1.11204    0.36051   3.085 0.002038 **
## YEAR2017      1.31797    0.39158   3.366 0.000763 ***
## YEAR2018      1.10198    0.44670   2.467 0.013627 *
## YEAR2019      1.72186    0.47588   3.618 0.000297 ***
## DEPTH         0.13869    0.05325   2.605 0.009198 **
## BOTTOM.TEMP   0.72893    0.06305  11.561  < 2e-16 ***
## SALINITY      0.18228    0.06209   2.935 0.003331 **
## STATEGA       1.34722    0.18297   7.363 1.80e-13 ***
## STATENC      -0.19400    0.15574  -1.246 0.212877
## STATESC       0.74019    0.16458   4.497 6.88e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.616

## neg_binom dispersal: 1.461

## zero_infl dispersal: 1.336

## zero_adj dispersal: 2.616

AICtab(full_mod_st)

##            dAIC df
## zero_adj    0.0 75
## zero_infl 76.7 45
## neg_binom 93.1 38
```

Full model with area as a covariate instead of latitude.

```
full_mod_area <- compare_models(data = new_data2,
               nb_model = "FREQ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY +
AREA",
               zic_model = "FREQ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY +
AREA",
               zac_model = "FREQ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY +
AREA",
               zi_model = "~DEPTH + BOTTOM.TEMP + SALINITY + AREA",
               za_model = "~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + AREA",
               r_dat = new_data2$FREQ) %>%
  try()
```

```
## Percent positive tows: 26.7

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + AREA
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   7304.8   7542.3  -3612.4   7224.8     2763
##
##
## Dispersion parameter for nbinom2 family (): 0.143
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.986440   0.410733   4.836 1.32e-06 ***
## YEAR1990    -2.873690   0.560501  -5.127 2.94e-07 ***
## YEAR1991    -1.014052   0.504953  -2.008 0.044621 *
## YEAR1992    -1.460832   0.486107  -3.005 0.002654 **
## YEAR1993    -2.359111   0.529360  -4.457 8.33e-06 ***
## YEAR1994    -0.865029   0.485341  -1.782 0.074699 .
## YEAR1995    -1.256843   0.508878  -2.470 0.013518 *
## YEAR1996    -1.187546   0.514554  -2.308 0.021004 *
## YEAR1997    -2.859142   0.620521  -4.608 4.07e-06 ***
## YEAR1998    -2.531209   0.545234  -4.642 3.44e-06 ***
## YEAR1999    -1.684916   0.498510  -3.380 0.000725 ***
## YEAR2000    -2.312428   0.517196  -4.471 7.78e-06 ***
## YEAR2001    -0.662192   0.469840  -1.409 0.158718
## YEAR2002    -1.106978   0.474116  -2.335 0.019553 *
## YEAR2003     0.193772   0.468374   0.414 0.679084
## YEAR2004    -1.782691   0.496122  -3.593 0.000327 ***
## YEAR2005    -0.106886   0.472019  -0.226 0.820857
## YEAR2006    -2.316812   0.494793  -4.682 2.84e-06 ***
## YEAR2007    -2.301868   0.511351  -4.502 6.75e-06 ***
## YEAR2008    -0.993821   0.463728  -2.143 0.032104 *
## YEAR2009    -0.168239   0.474914  -0.354 0.723150
## YEAR2010    -0.497508   0.476068  -1.045 0.296006
## YEAR2011    -1.084549   0.473749  -2.289 0.022063 *
## YEAR2012    -0.529965   0.506496  -1.046 0.295406
## YEAR2013    -1.723804   0.487193  -3.538 0.000403 ***
## YEAR2014    -2.443574   0.540351  -4.522 6.12e-06 ***
## YEAR2015    -1.014023   0.473742  -2.140 0.032318 *
## YEAR2016    -1.572581   0.475064  -3.310 0.000932 ***
## YEAR2017    -2.467732   0.502476  -4.911 9.05e-07 ***
## YEAR2018    -1.648663   0.546902  -3.015 0.002574 **
## YEAR2019    -2.181902   0.549712  -3.969 7.21e-05 ***
## DEPTH       -0.004862   0.068987  -0.070 0.943816
```

```
## BOTTOM.TEMP -1.012151   0.080633 -12.553  < 2e-16 ***
## SALINITY    -0.116661   0.094228  -1.238 0.215688    
## AREAGA      -1.762737   0.234907  -7.504 6.19e-14 ***
## AREALB      -0.565456   0.244592  -2.312 0.020787 *  
## AREAOB       0.579722   0.234041   2.477 0.013249 *  
## AREARB       1.754309   0.300720   5.834 5.42e-09 ***
## AREASC      -0.936922   0.234633  -3.993 6.52e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.762

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + AREA
## Zero inflation:        ~DEPTH + BOTTOM.TEMP + SALINITY + AREA
## Data: data
## 
##      AIC      BIC   logLik deviance df.resid 
##   7265.8   7556.8  -3583.9   7167.8     2754 
## 
## 
## Dispersion parameter for nbinom2 family (): 0.209 
## 
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)  1.79935    0.39958   4.503 6.70e-06 ***
## YEAR1990    -2.78423    0.51684  -5.387 7.16e-08 ***
## YEAR1991    -0.79542    0.47983  -1.658 0.097375 .  
## YEAR1992    -1.33864    0.45996  -2.910 0.003610 ** 
## YEAR1993    -2.27016    0.49850  -4.554 5.26e-06 ***
## YEAR1994    -0.59823    0.44741  -1.337 0.181188    
## YEAR1995    -1.12388    0.47998  -2.342 0.019205 *  
## YEAR1996    -1.05023    0.47359  -2.218 0.026584 *  
## YEAR1997    -2.70454    0.57778  -4.681 2.86e-06 ***
## YEAR1998    -2.28140    0.52432  -4.351 1.35e-05 ***
## YEAR1999    -1.23203    0.45664  -2.698 0.006975 ** 
## YEAR2000    -2.28014    0.49487  -4.608 4.07e-06 ***
## YEAR2001    -0.70153    0.43072  -1.629 0.103364    
## YEAR2002    -0.66451    0.46411  -1.432 0.152208    
## YEAR2003     0.13103    0.44133   0.297 0.766541    
## YEAR2004    -1.60137    0.46540  -3.441 0.000580 ***
## YEAR2005     0.23199    0.45142   0.514 0.607304    
## YEAR2006    -2.03994    0.49050  -4.159 3.20e-05 ***
## YEAR2007    -2.15654    0.51024  -4.227 2.37e-05 ***
## YEAR2008    -0.94018    0.43477  -2.162 0.030580 *  
## YEAR2009     0.05084    0.47281   0.108 0.914370    
## YEAR2010    -0.23852    0.46956  -0.508 0.611482    
```

```
## YEAR2011    -1.09946    0.44368  -2.478 0.013212 *
## YEAR2012    -0.71613    0.47847  -1.497 0.134470
## YEAR2013    -1.45803    0.47625  -3.061 0.002203 **
## YEAR2014    -2.28752    0.52700  -4.341 1.42e-05 ***
## YEAR2015    -0.89679    0.44586  -2.011 0.044286 *
## YEAR2016    -1.27569    0.44536  -2.864 0.004178 **
## YEAR2017    -2.33971    0.46618  -5.019 5.20e-07 ***
## YEAR2018    -1.03605    0.52493  -1.974 0.048415 *
## YEAR2019    -1.55751    0.54319  -2.867 0.004140 **
## DEPTH        0.14653    0.07999   1.832 0.066981 .
## BOTTOM.TEMP -0.77229    0.10380  -7.440 1.01e-13 ***
## SALINITY     0.08146    0.09850   0.827 0.408243
## AREAGA      -1.10005    0.29492  -3.730 0.000191 ***
## AREALB      -0.29734    0.28091  -1.059 0.289827
## AREAOB       0.81963    0.26791   3.059 0.002218 **
## AREARB       2.38204    0.33034   7.211 5.56e-13 ***
## AREASC      -0.23794    0.28530  -0.834 0.404267
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.3937     0.7356  -3.254  0.00114 **
## DEPTH         0.5607     0.1987   2.822  0.00477 **
## BOTTOM.TEMP   0.9078     0.1894   4.792 1.65e-06 ***
## SALINITY      0.6828     0.2090   3.267  0.00109 **
## AREAGA        2.1519     0.6563   3.279  0.00104 **
## AREALB        0.7436     0.6731   1.105  0.26928
## AREAOB        0.7053     0.5620   1.255  0.20949
## AREARB        1.4387     0.6946   2.071  0.03834 *
## AREASC        2.1949     0.6908   3.177  0.00149 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.583

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + AREA
## Zero inflation:        ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + AREA
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   7198.8   7667.9  -3520.4   7040.8     2724
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.0976
##
```

```
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)    
## (Intercept)  0.96839    0.57523   1.683 0.092285 .  
## YEAR1990    -3.76714    0.70456  -5.347 8.95e-08 ***
## YEAR1991    -0.16322    0.61740  -0.264 0.791502    
## YEAR1992    -0.64841    0.61433  -1.055 0.291209    
## YEAR1993    -3.10085    0.58795  -5.274 1.33e-07 ***
## YEAR1994    -0.47832    0.53363  -0.896 0.370064    
## YEAR1995    -1.78911    0.55985  -3.196 0.001395 ** 
## YEAR1996    -2.07857    0.52950  -3.926 8.65e-05 ***
## YEAR1997    -2.43772    0.85830  -2.840 0.004509 ** 
## YEAR1998    -2.52881    0.69687  -3.629 0.000285 ***
## YEAR1999    -1.34815    0.54492  -2.474 0.013360 *  
## YEAR2000    -2.49172    0.59738  -4.171 3.03e-05 ***
## YEAR2001    -0.35528    0.50896  -0.698 0.485141    
## YEAR2002    -0.29077    0.58775  -0.495 0.620799    
## YEAR2003    -0.12512    0.50211  -0.249 0.803206    
## YEAR2004    -1.45066    0.61090  -2.375 0.017567 *  
## YEAR2005     0.89145    0.59093   1.509 0.131414    
## YEAR2006    -0.72491    0.79099  -0.916 0.359431    
## YEAR2007    -1.87541    0.66969  -2.800 0.005104 ** 
## YEAR2008    -0.82534    0.51349  -1.607 0.107985    
## YEAR2009     0.52117    0.58503   0.891 0.373015    
## YEAR2010     0.56247    0.59528   0.945 0.344718    
## YEAR2011    -0.92293    0.51516  -1.792 0.073206 .  
## YEAR2012    -0.85365    0.54776  -1.558 0.119131    
## YEAR2013    -1.05437    0.63161  -1.669 0.095050 .  
## YEAR2014    -3.02985    0.72913  -4.155 3.25e-05 ***
## YEAR2015    -1.10066    0.53355  -2.063 0.039124 *  
## YEAR2016    -1.12629    0.55006  -2.048 0.040603 *  
## YEAR2017    -2.49440    0.58498  -4.264 2.01e-05 ***
## YEAR2018    -0.80576    0.70551  -1.142 0.253409    
## YEAR2019    -0.59903    0.75293  -0.796 0.426264    
## DEPTH        0.18558    0.09479   1.958 0.050254 .  
## BOTTOM.TEMP -0.59952    0.11799  -5.081 3.75e-07 ***
## SALINITY    -0.05218    0.12447  -0.419 0.675061    
## AREAGA      -0.86463    0.35420  -2.441 0.014645 *  
## AREALB       0.09437    0.33910   0.278 0.780785    
## AREAOB       1.18698    0.32303   3.675 0.000238 ***
## AREARB       2.36338    0.39869   5.928 3.07e-09 ***
## AREASC      -0.20897    0.34106  -0.613 0.540062    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)    
## (Intercept) -0.20166    0.31050  -0.649 0.516050    
## YEAR1990     1.47403    0.42060   3.505 0.000457 ***
## YEAR1991     1.19853    0.38987   3.074 0.002111 ** 
## YEAR1992     1.38102    0.40077   3.446 0.000569 ***
```

```
## YEAR1993     0.89206    0.39769   2.243 0.024892 *
## YEAR1994     0.68280    0.37781   1.807 0.070721 .
## YEAR1995     0.26130    0.38105   0.686 0.492873
## YEAR1996    -0.21869    0.38291  -0.571 0.567918
## YEAR1997     2.09240    0.50633   4.133 3.59e-05 ***
## YEAR1998     1.30328    0.42672   3.054 0.002257 **
## YEAR1999     0.57967    0.38388   1.510 0.131029
## YEAR2000     1.05121    0.40146   2.618 0.008833 **
## YEAR2001     0.56766    0.36129   1.571 0.116141
## YEAR2002     0.87549    0.39245   2.231 0.025693 *
## YEAR2003    -0.10561    0.36268  -0.291 0.770902
## YEAR2004     1.27418    0.40775   3.125 0.001778 **
## YEAR2005     0.88492    0.38368   2.306 0.021087 *
## YEAR2006     2.03770    0.47675   4.274 1.92e-05 ***
## YEAR2007     1.57702    0.43458   3.629 0.000285 ***
## YEAR2008     0.51885    0.36590   1.418 0.156185
## YEAR2009     0.90813    0.38582   2.354 0.018584 *
## YEAR2010     1.05030    0.38239   2.747 0.006021 **
## YEAR2011     0.89015    0.36577   2.434 0.014949 *
## YEAR2012     0.50241    0.37064   1.356 0.175257
## YEAR2013     1.20914    0.38849   3.112 0.001855 **
## YEAR2014     1.23193    0.41002   3.005 0.002659 **
## YEAR2015     0.60134    0.36952   1.627 0.103665
## YEAR2016     1.03783    0.36096   2.875 0.004038 **
## YEAR2017     1.28900    0.39025   3.303 0.000956 ***
## YEAR2018     1.03252    0.44905   2.299 0.021486 *
## YEAR2019     1.58045    0.46483   3.400 0.000674 ***
## DEPTH        0.16176    0.05393   2.999 0.002705 **
## BOTTOM.TEMP  0.70861    0.06296  11.254  < 2e-16 ***
## SALINITY     0.12372    0.07001   1.767 0.077185 .
## AREAGA       1.25272    0.18225   6.874 6.25e-12 ***
## AREALB       0.29993    0.18579   1.614 0.106443
## AREAOB      -0.04584    0.17642  -0.260 0.794968
## AREARB      -0.61072    0.21500  -2.841 0.004503 **
## AREASC       0.81035    0.17598   4.605 4.13e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.644

## neg_binom dispersal: 1.762

## zero_infl dispersal: 1.583

## zero_adj dispersal: 2.644
```

```
AICtab(full_mod_area)
```

```
##            dAIC df
## zero_adj    0   79
```

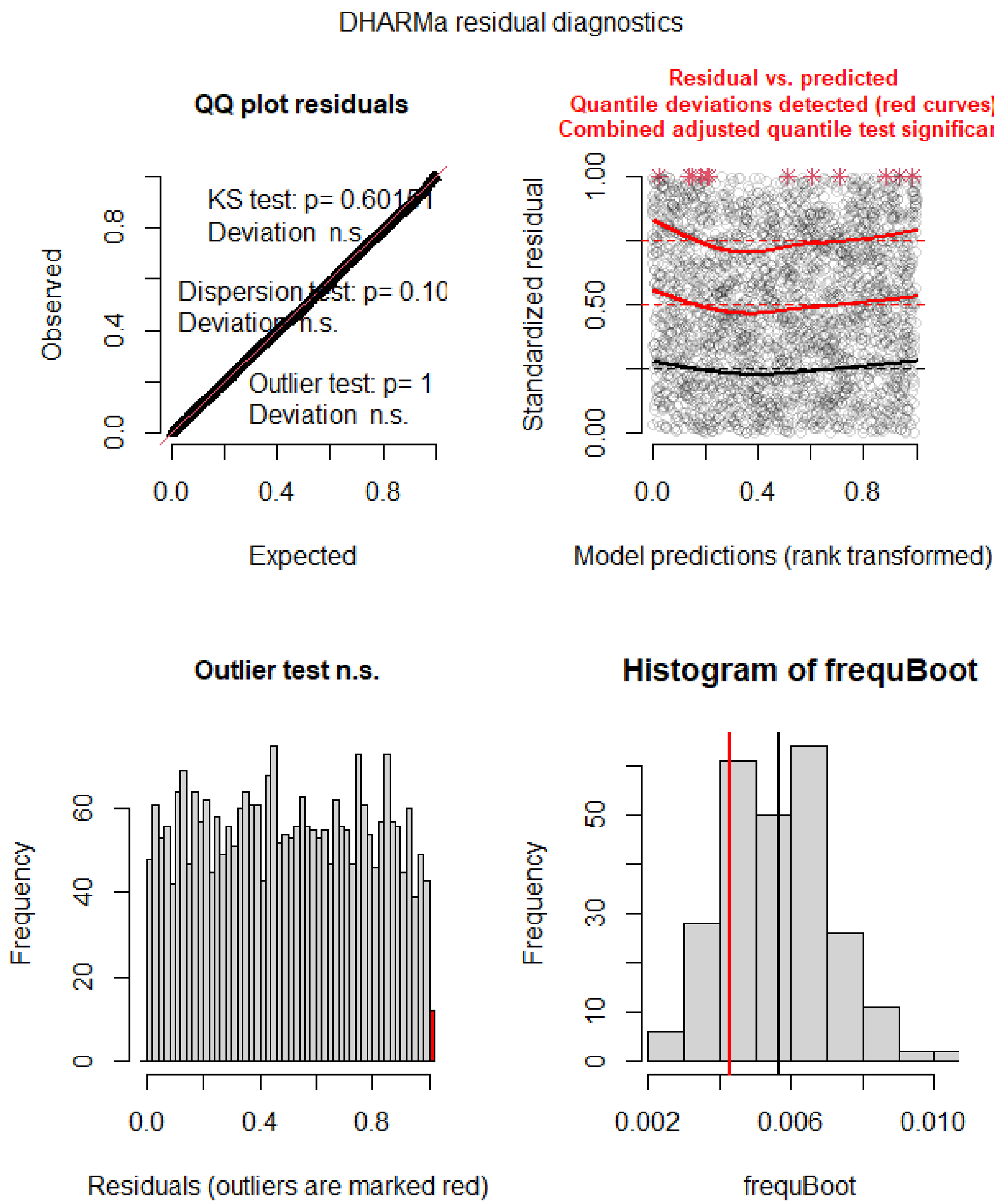
```
## zero_infl  67  49
## neg_binom 106  40
```

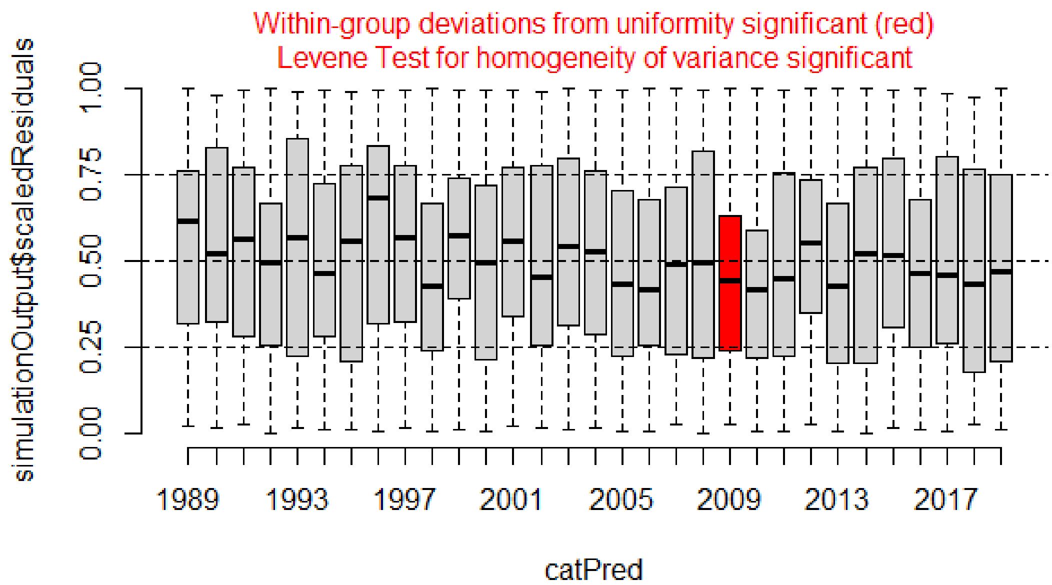
Compare alternate configurations with AIC. Zero-altered NB was the best performing model for all 4 configurations.

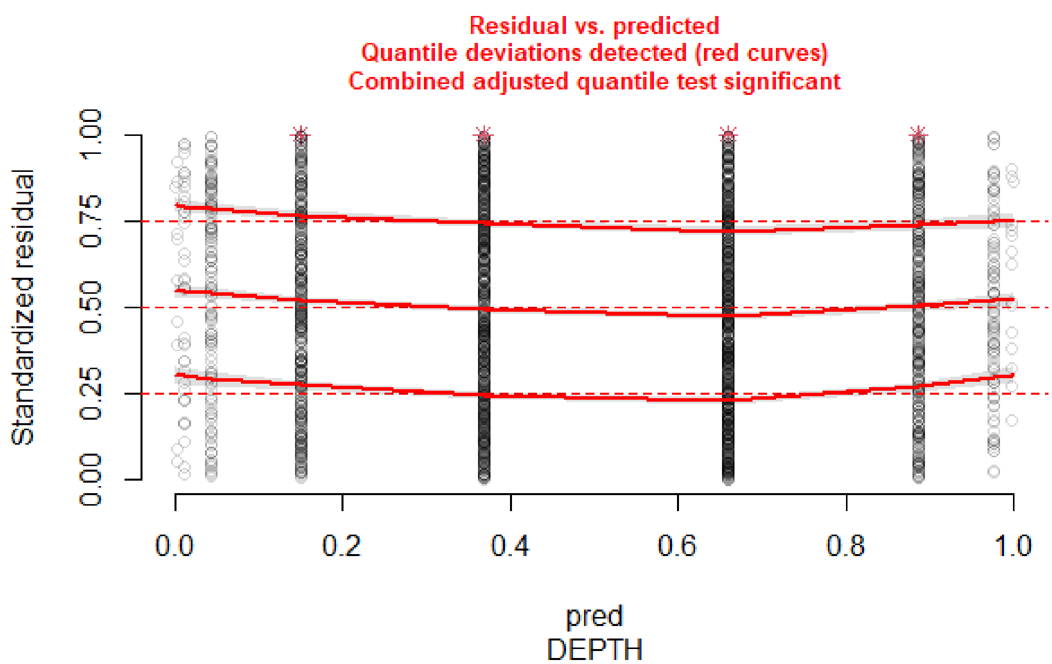
```
AICtab(full_mod)

##           dAIC  df
## zero_adj    0.0 71
## zero_infl  35.3 71
## neg_binom 138.4 36

AICtab(full_mod_stratum)

##           dAIC  df
## zero_adj    0.0 115
## zero_infl  64.1 85
## neg_binom  97.6 58

AICtab(full_mod_st)

##           dAIC df
## zero_adj   0.0 75
## zero_infl 76.7 45
## neg_binom 93.1 38

AICtab(full_mod_area)

##           dAIC df
## zero_adj    0   79
## zero_infl  67   49
## neg_binom 106   40

AICtab(full_mod$zero_adj,
full_mod_stratum$zero_adj,full_mod_st$zero_adj,full_mod_area$zero_adj)

##                           dAIC  df
## full_mod_stratum$zero_adj   0.0 115
## full_mod_area$zero_adj     56.2 79
## full_mod_st$zero_adj       58.0 75
## full_mod$zero_adj         184.1 71
```

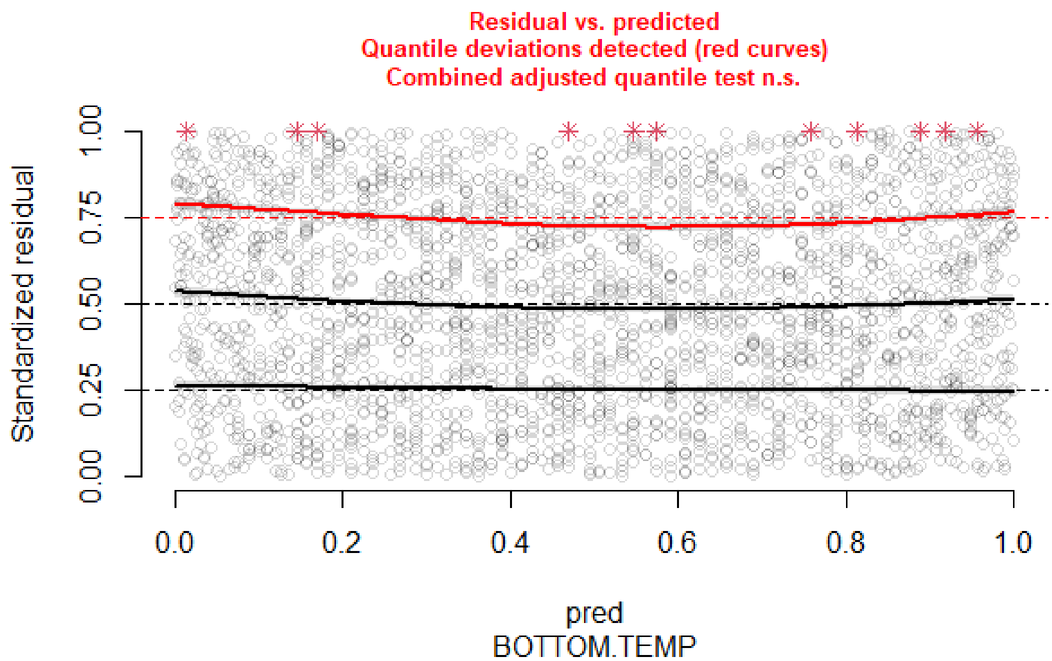
DHARMA simulations

Simulated residuals for the full model with latitude as a covariate.

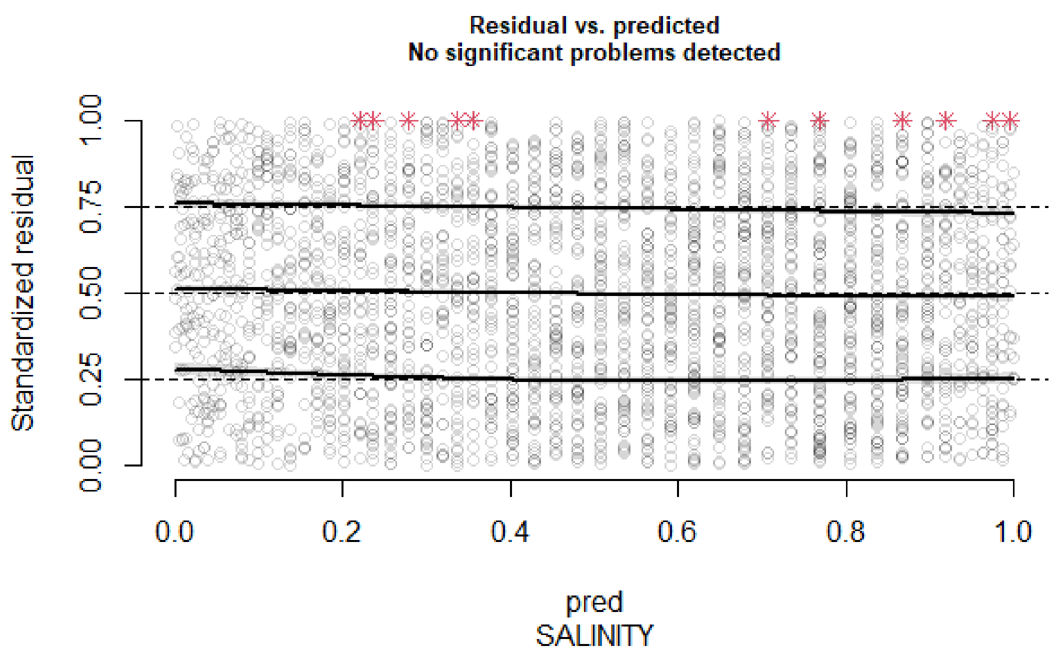



Within-group deviations from uniformity significant (red)
Levene Test for homogeneity of variance significant
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1989
1993
1997
2001
2005
2009
2013
2017
catPred

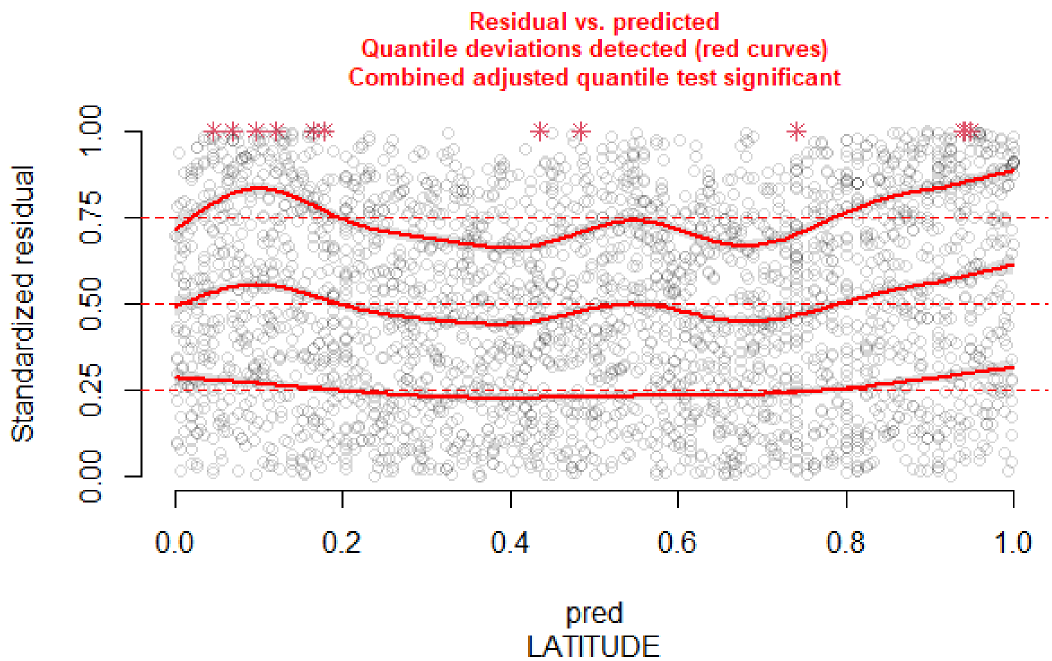

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

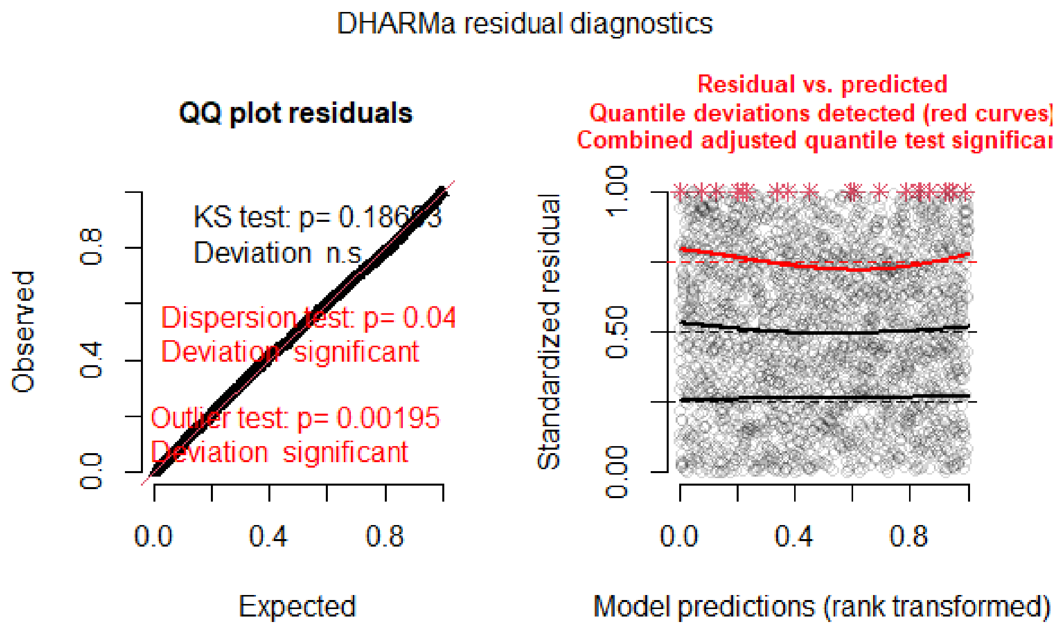


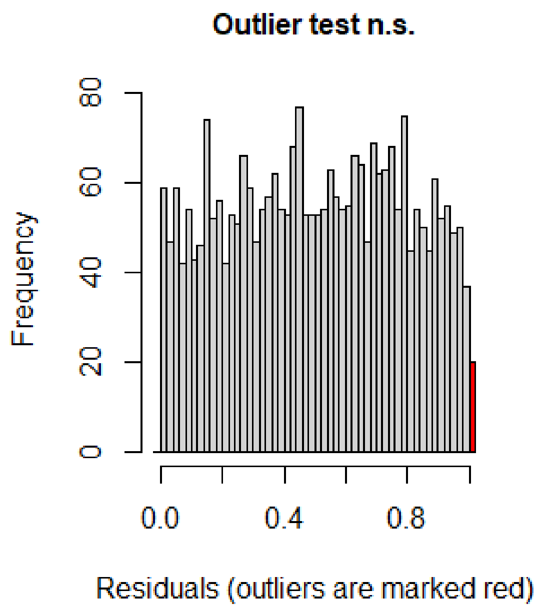
Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test n.s.
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

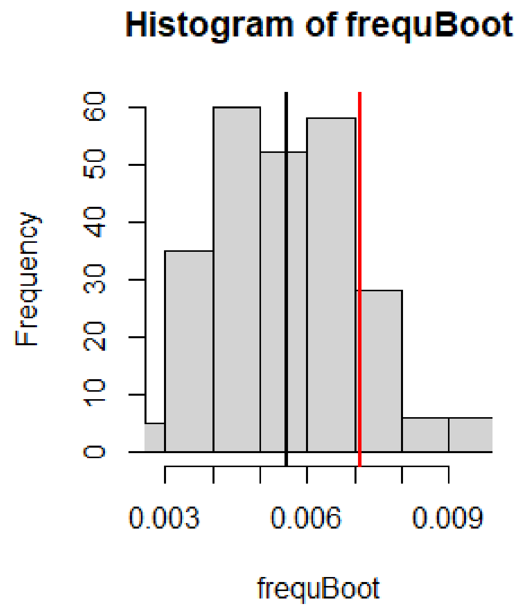


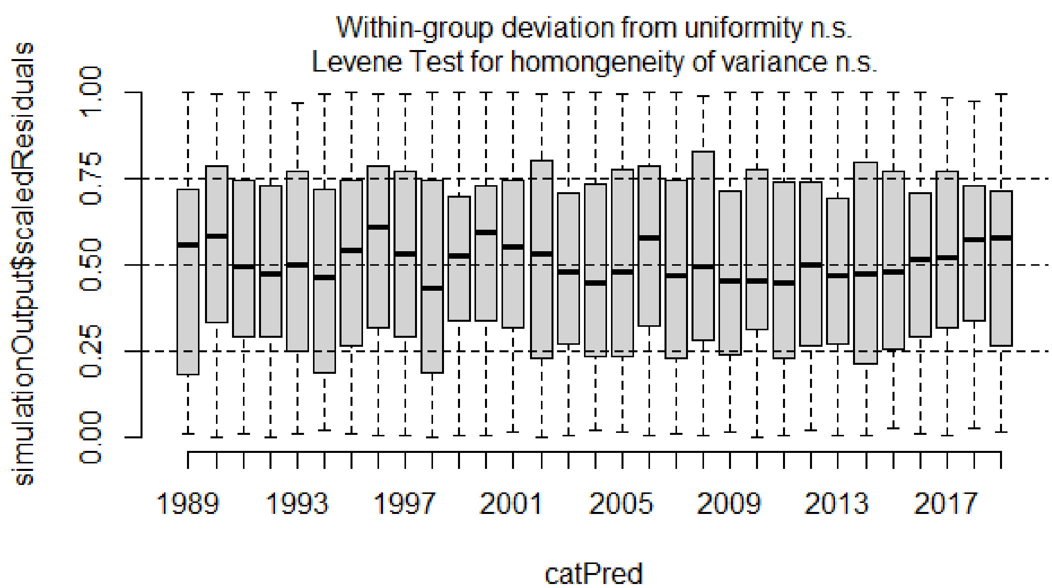
Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY

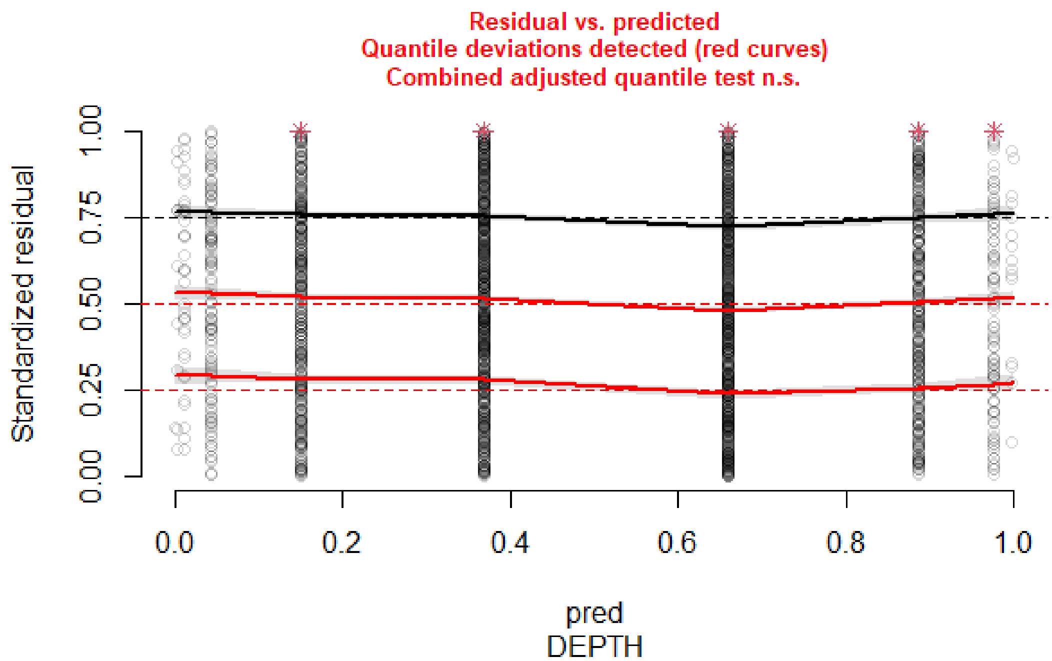

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
LATITUDE

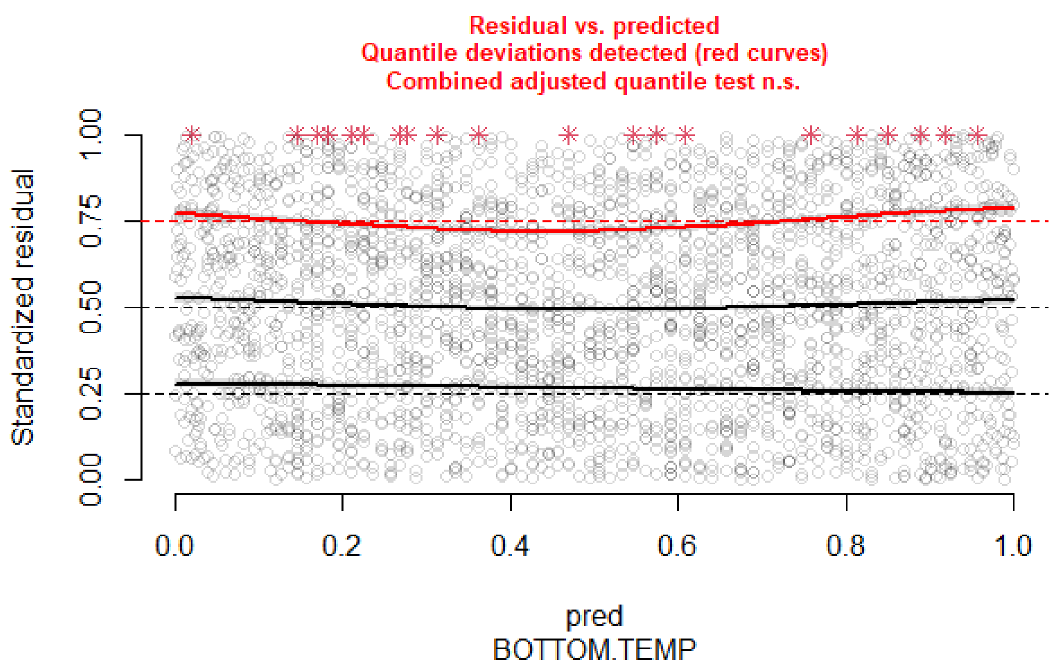

DHARMa residual diagnostics
QQ plot residuals
KS test: p= 0.18693
Deviation n.s.
Dispersion test: p= 0.04
Deviation significant
Outlier test: p= 0.00195
Deviation significant
Observed
Expected
0.0
0.4
0.8
Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.50
1.00
Model predictions (rank transformed)

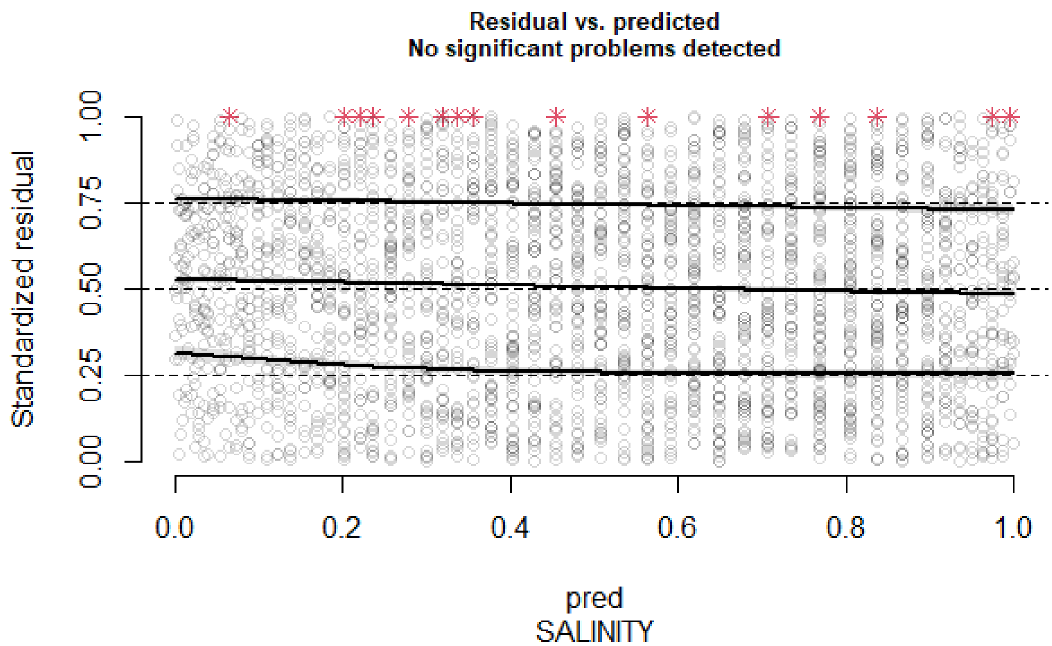

Outlier test n.s.
Frequency
0
20
40
60
80
0.0
0.4
0.8
Residuals (outliers are marked red)

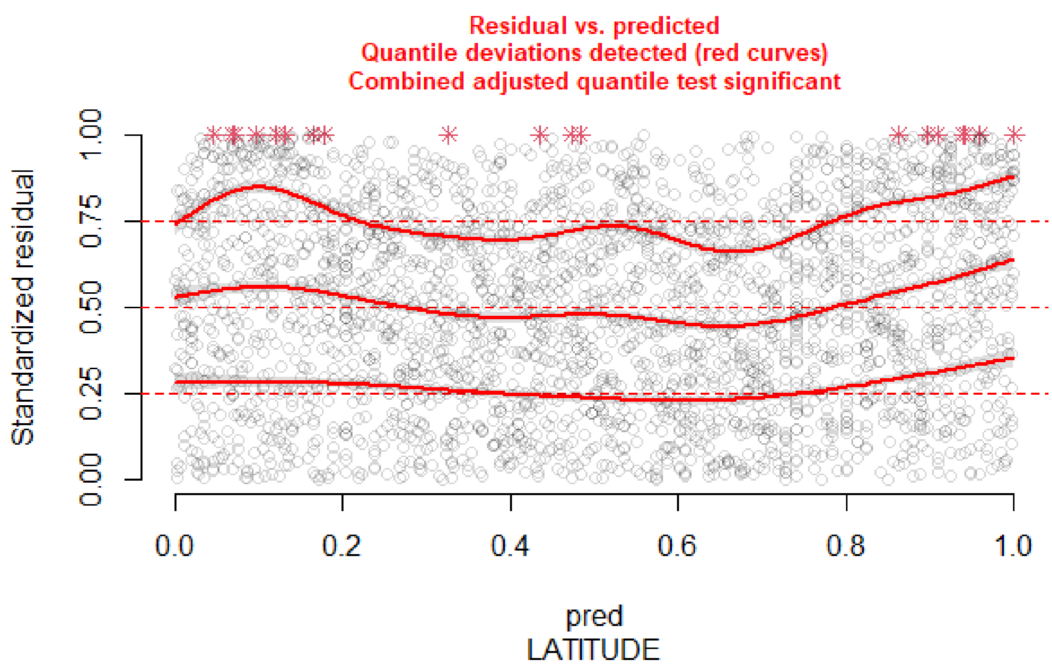

Histogram of frequBoot
Frequency
0
10
20
30
40
50
60
0.003
0.006
0.009
frequBoot

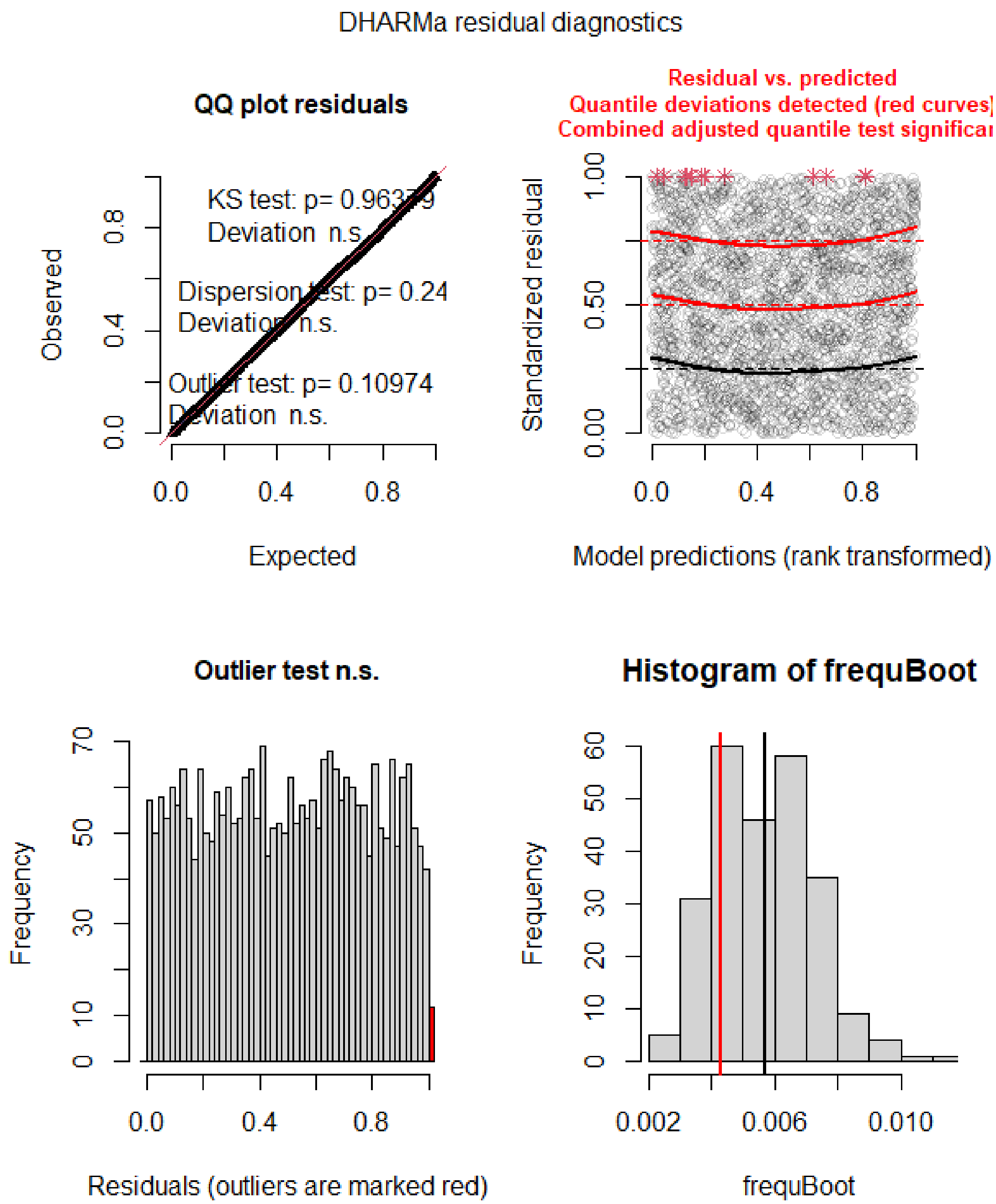

Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1989
1993
1997
2001
2005
2009
2013
2017
catPred

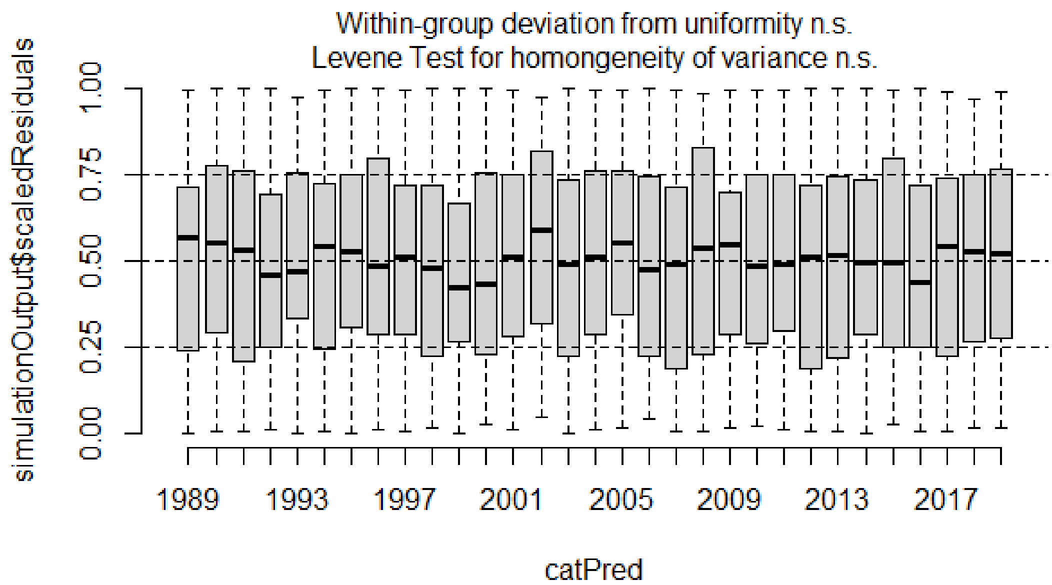

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test n.s.
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

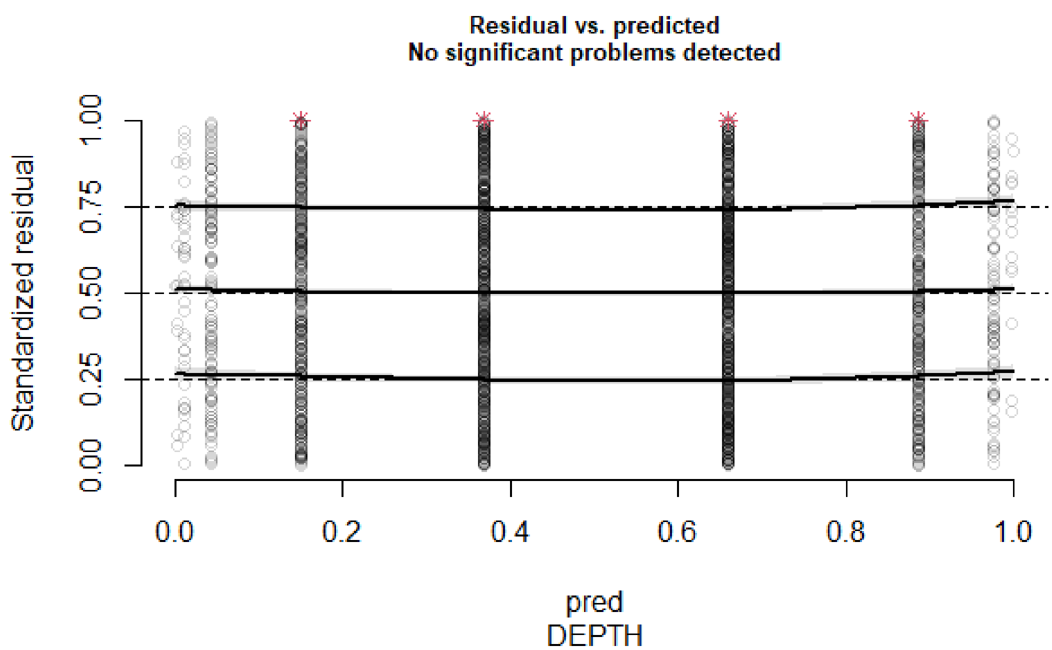

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test n.s.
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

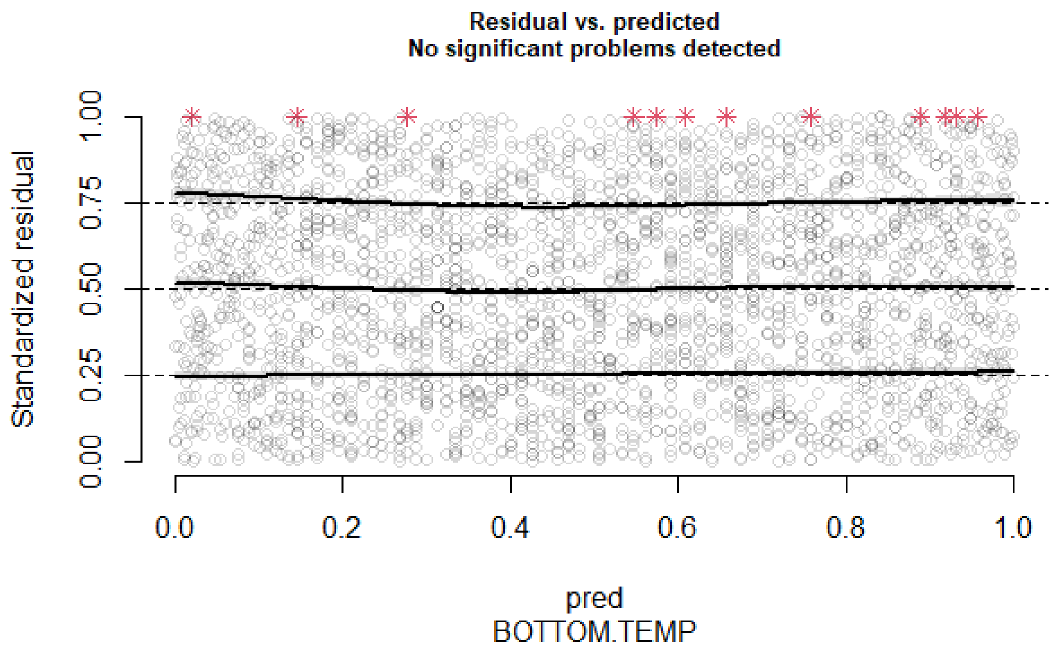

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY

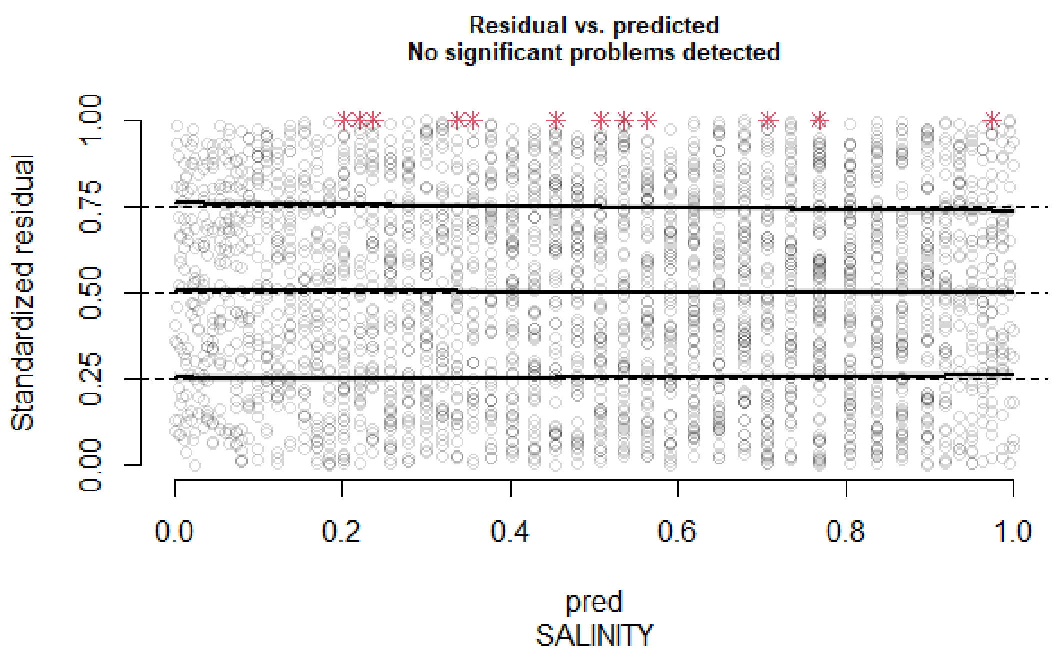

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
LATITUDE

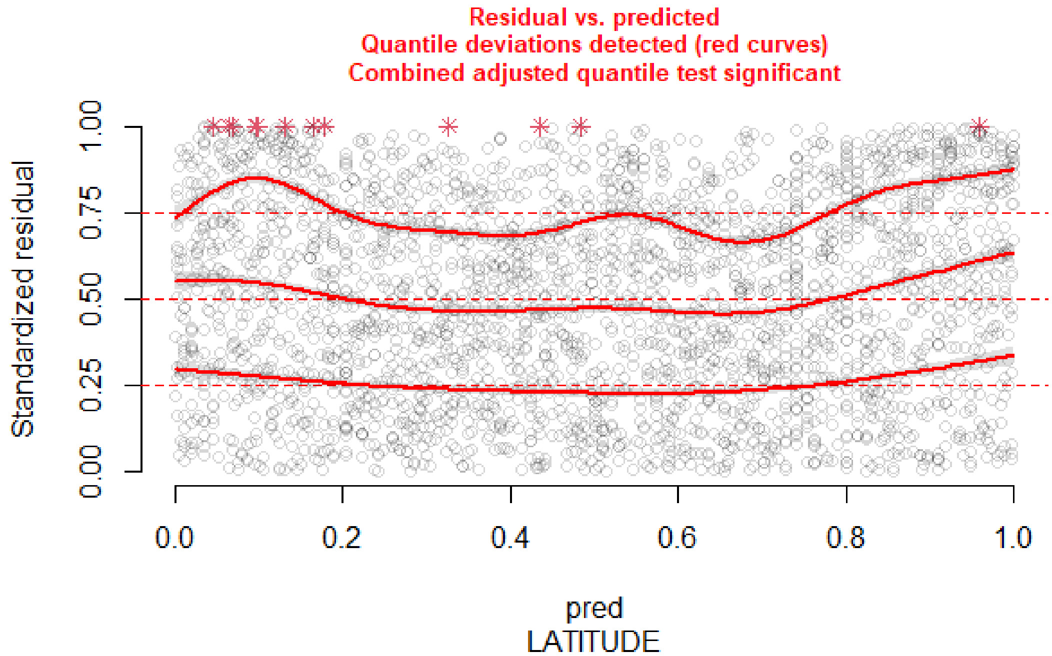

DHARMa residual diagnostics
QQ plot residuals
KS test: p= 0.96379
Deviation n.s.
Dispersion test: p= 0.24
Deviation n.s.
Outlier test: p= 0.10974
Deviation n.s.
Observed
Expected
Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significan
Standardized residual
Model predictions (rank transformed)
Outlier test n.s.
Frequency
Residuals (outliers are marked red)
Histogram of frequBoot
frequBoot


Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1989
1993
1997
2001
2005
2009
2013
2017
catPred

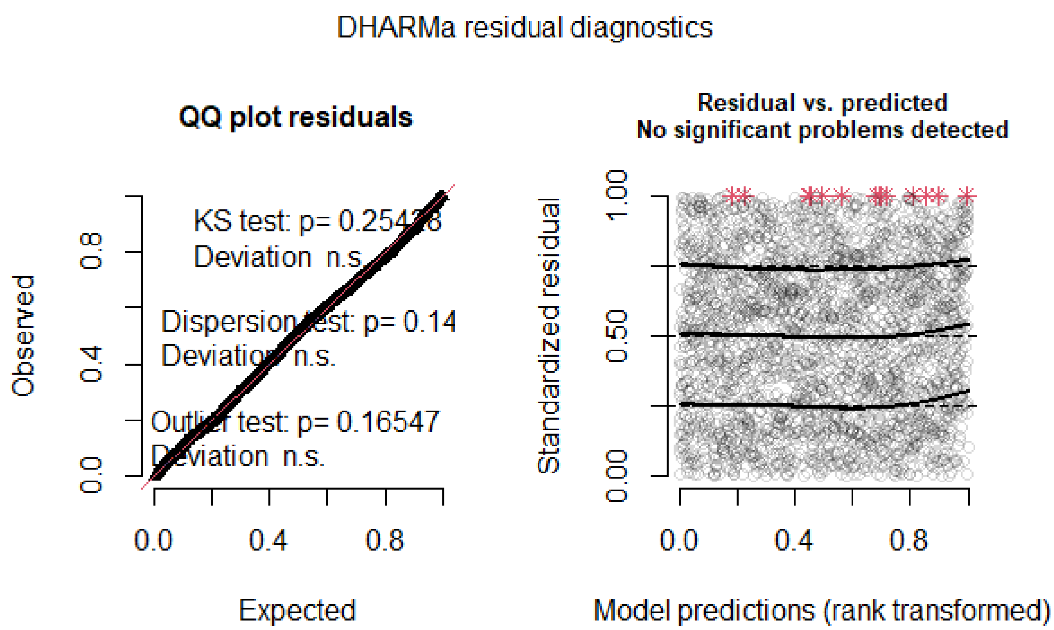

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

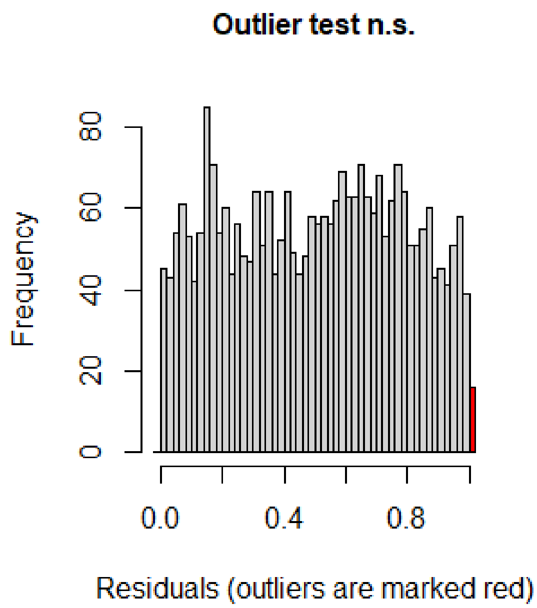

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

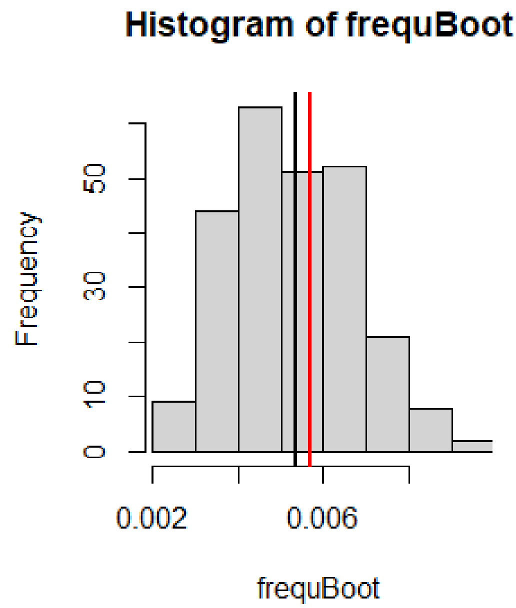

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY

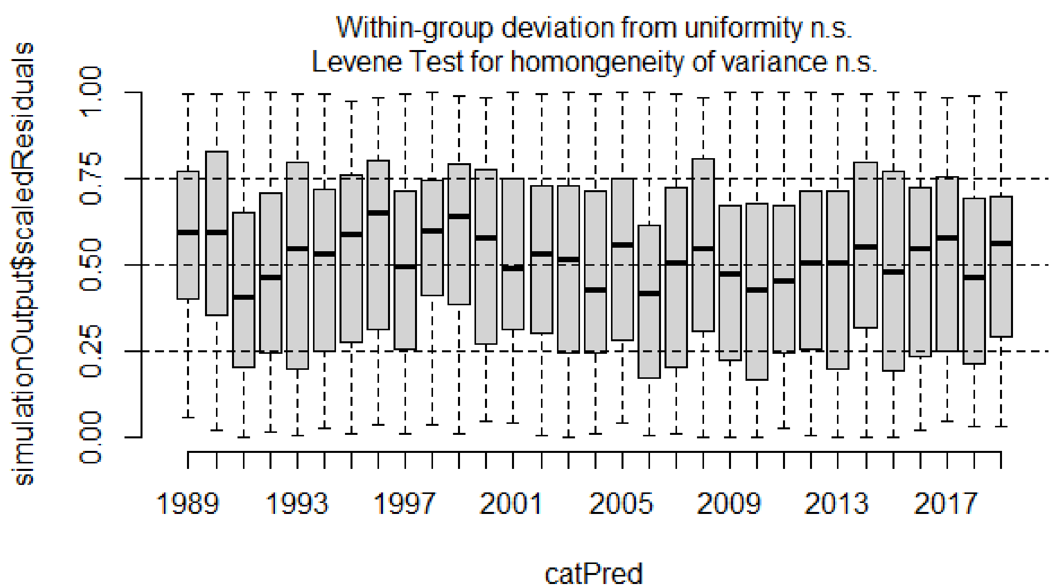


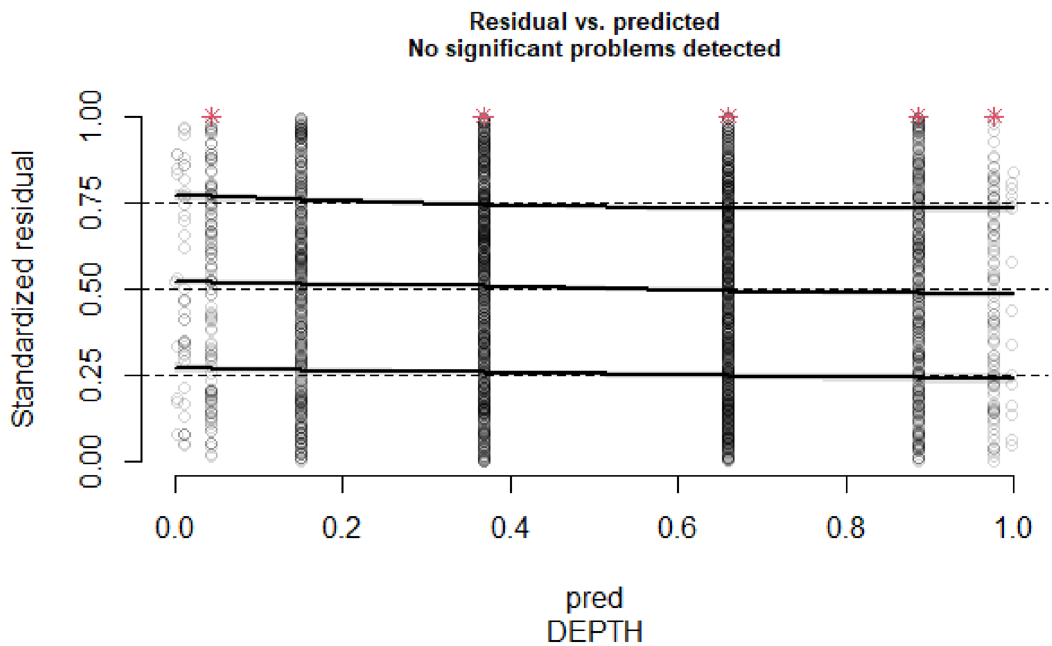
Simulated residuals for the full model with stratum as a covariate instead of latitude.

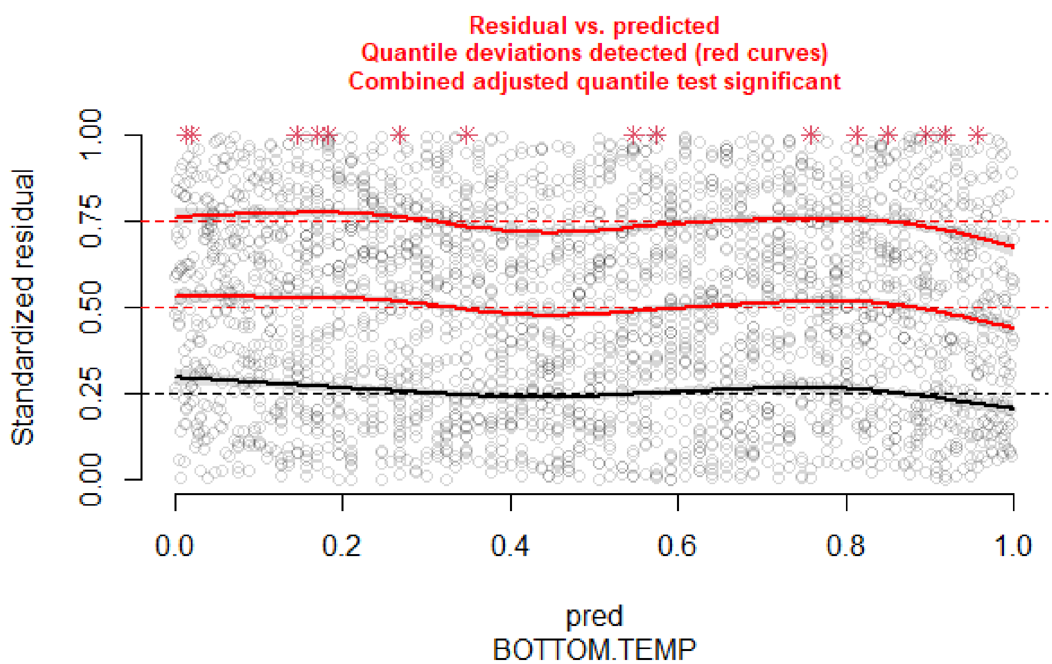


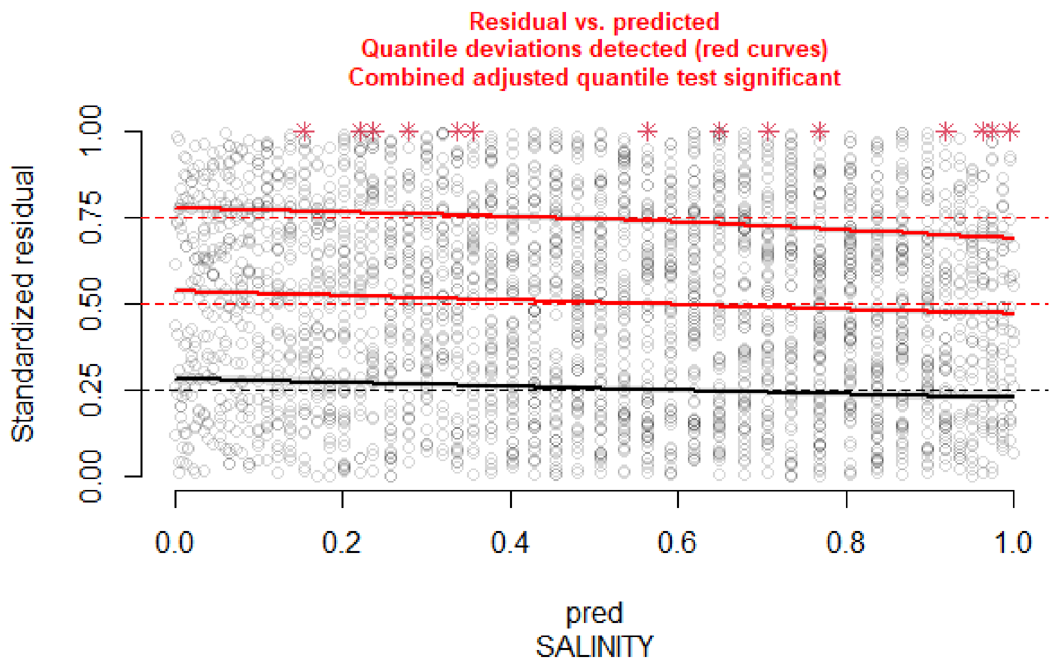

Outlier test n.s.
Frequency
0
20
40
60
80
0.0
0.4
0.8
Residuals (outliers are marked red)

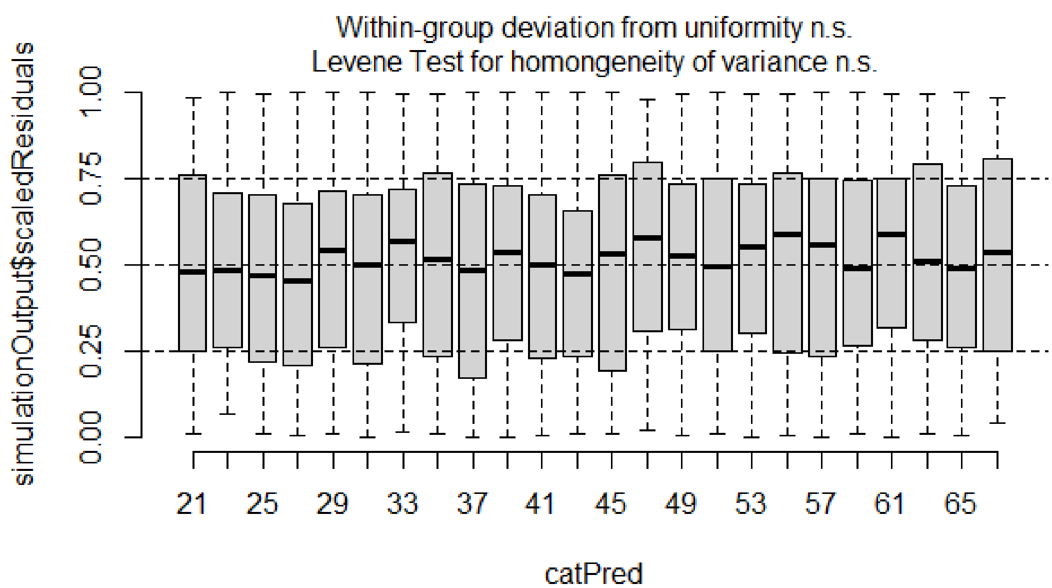

Histogram of frequBoot
Frequency
0
10
30
50
0.002
0.006
frequBoot

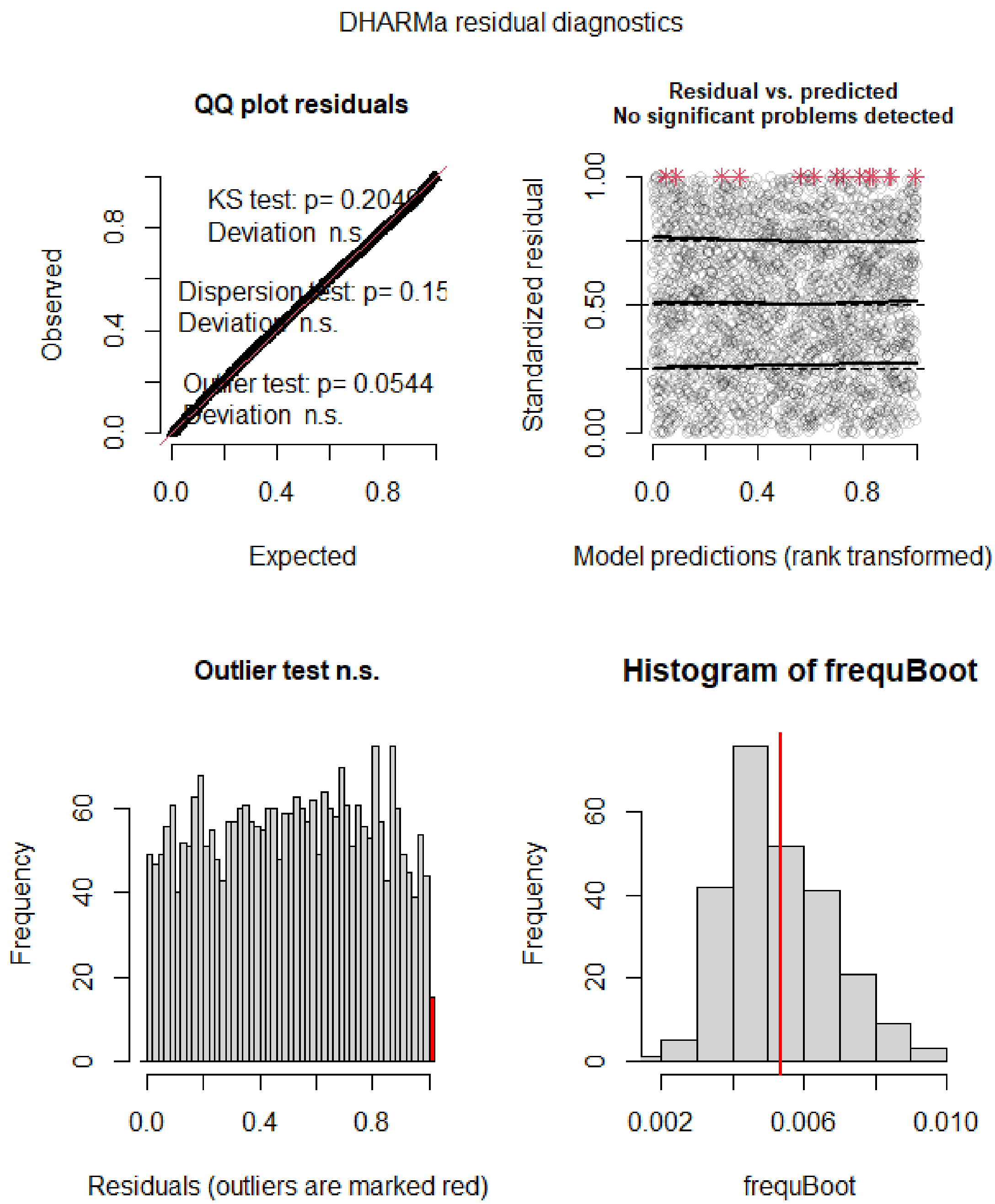

Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1989
1993
1997
2001
2005
2009
2013
2017
catPred

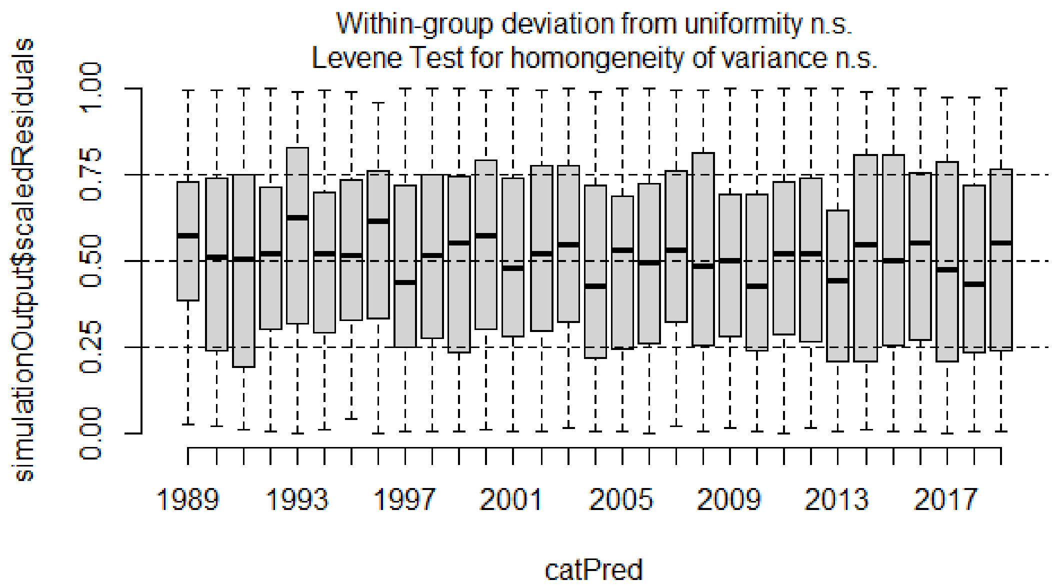

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

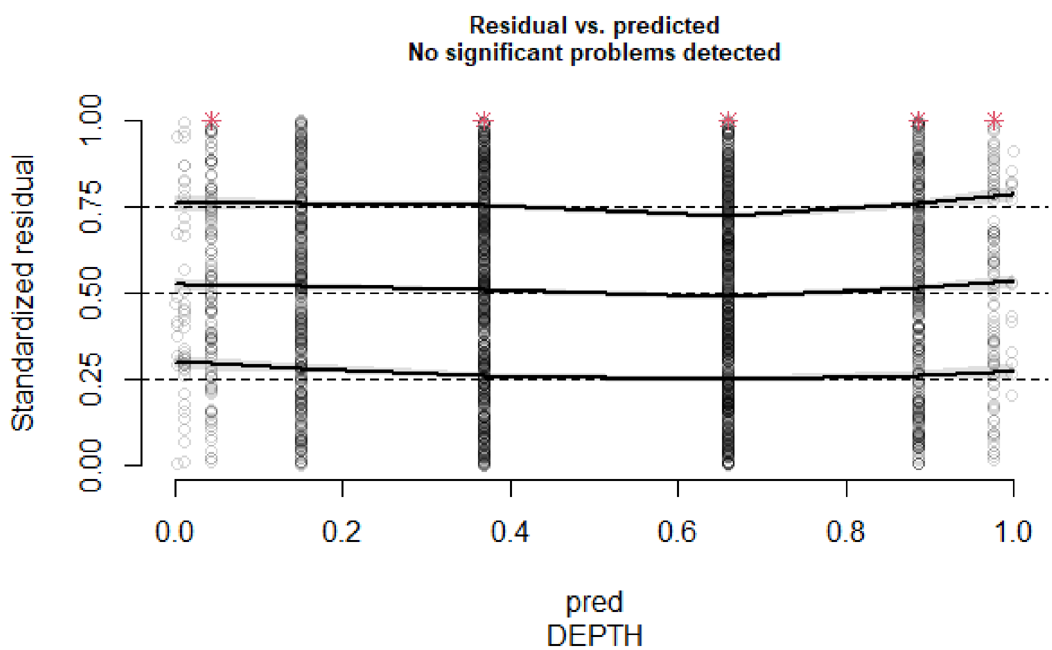

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

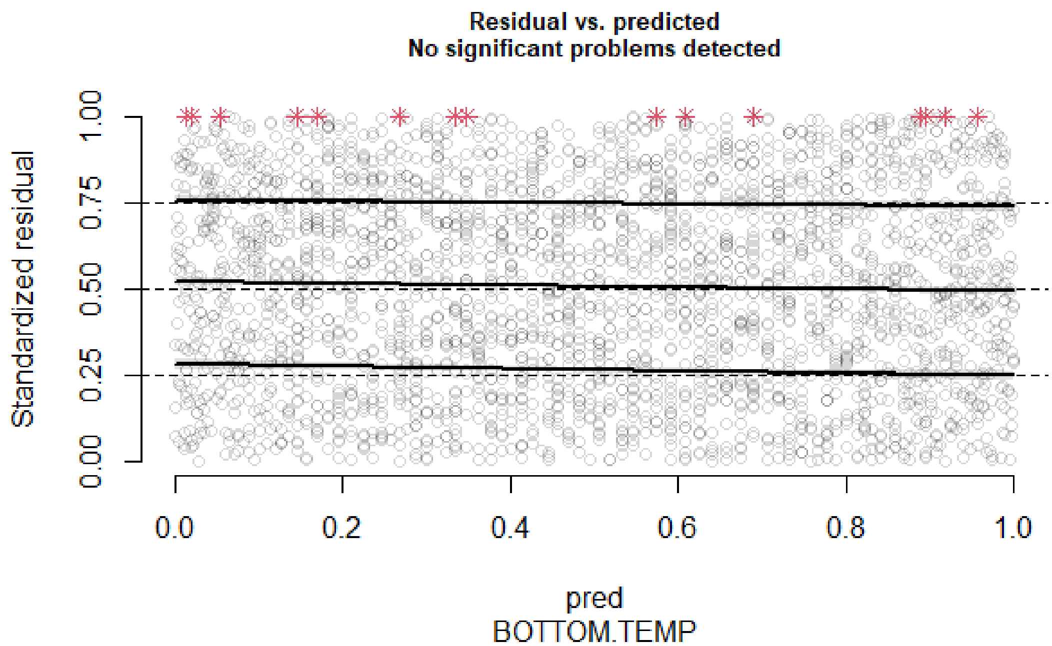

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY

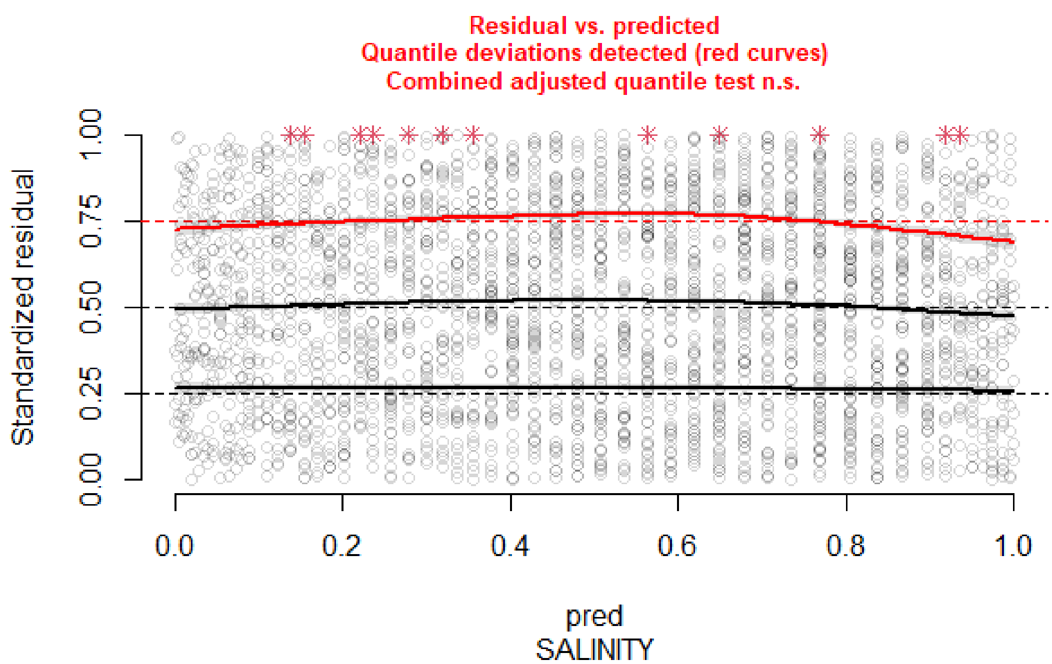

Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
21
25
29
33
37
41
45
49
53
57
61
65
catPred

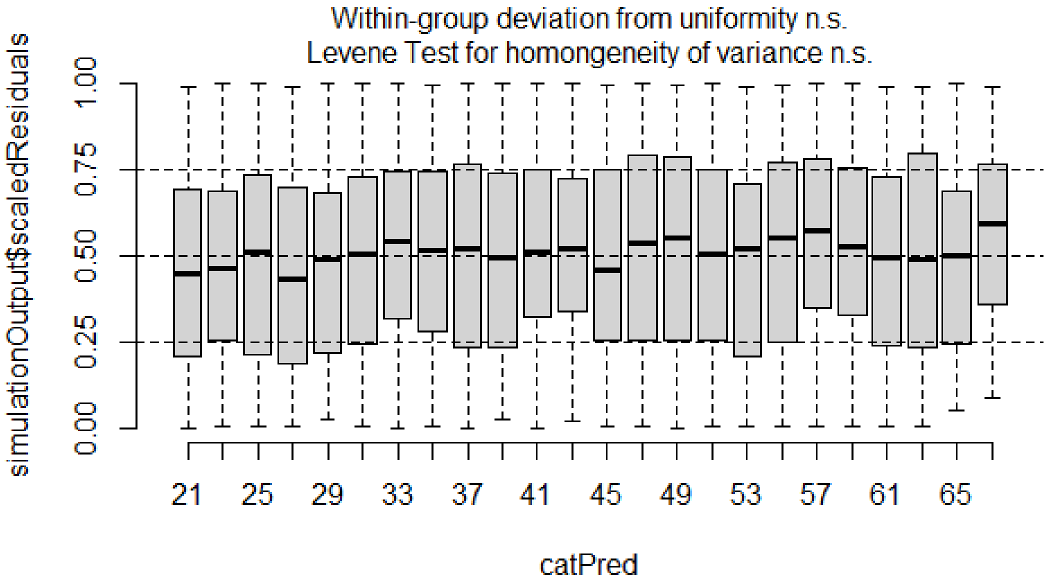

DHARMa residual diagnostics
QQ plot residuals
KS test: p= 0.2046
Deviation n.s.
Dispersion test: p= 0.15
Deviation n.s.
Outlier test: p= 0.0544
Deviation n.s.
Observed
Expected
0.0
0.4
0.8
Residual vs. predicted
No significant problems detected
Standardized residual
Model predictions (rank transformed)
0.00
0.50
1.00
Outlier test n.s.
Frequency
Residuals (outliers are marked red)
0
20
40
60
Histogram of frequBoot
frequBoot
0.002
0.006
0.010

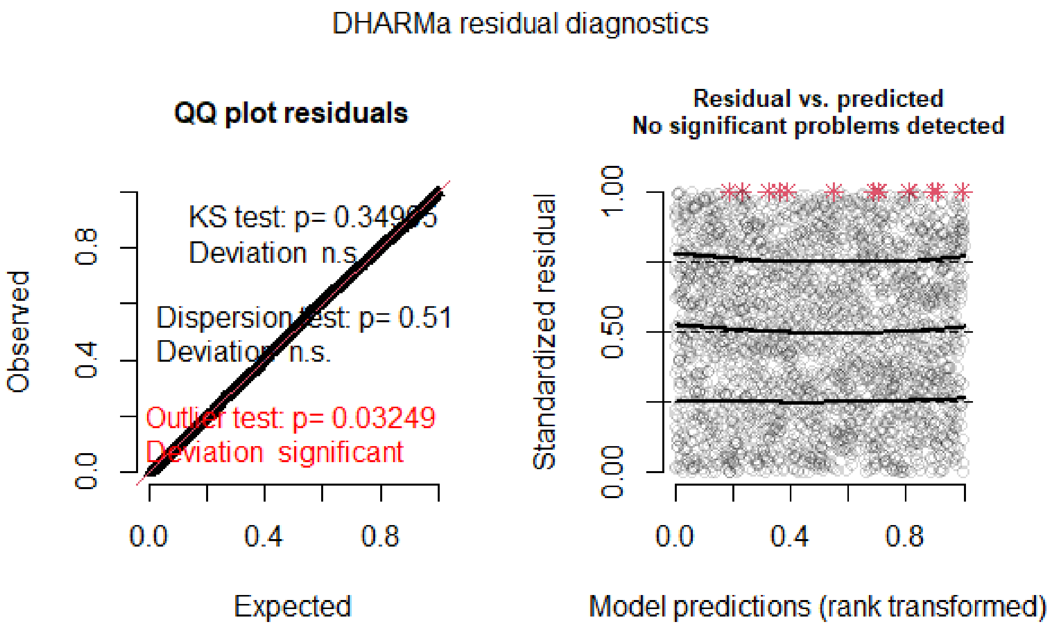

Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1989
1993
1997
2001
2005
2009
2013
2017
catPred

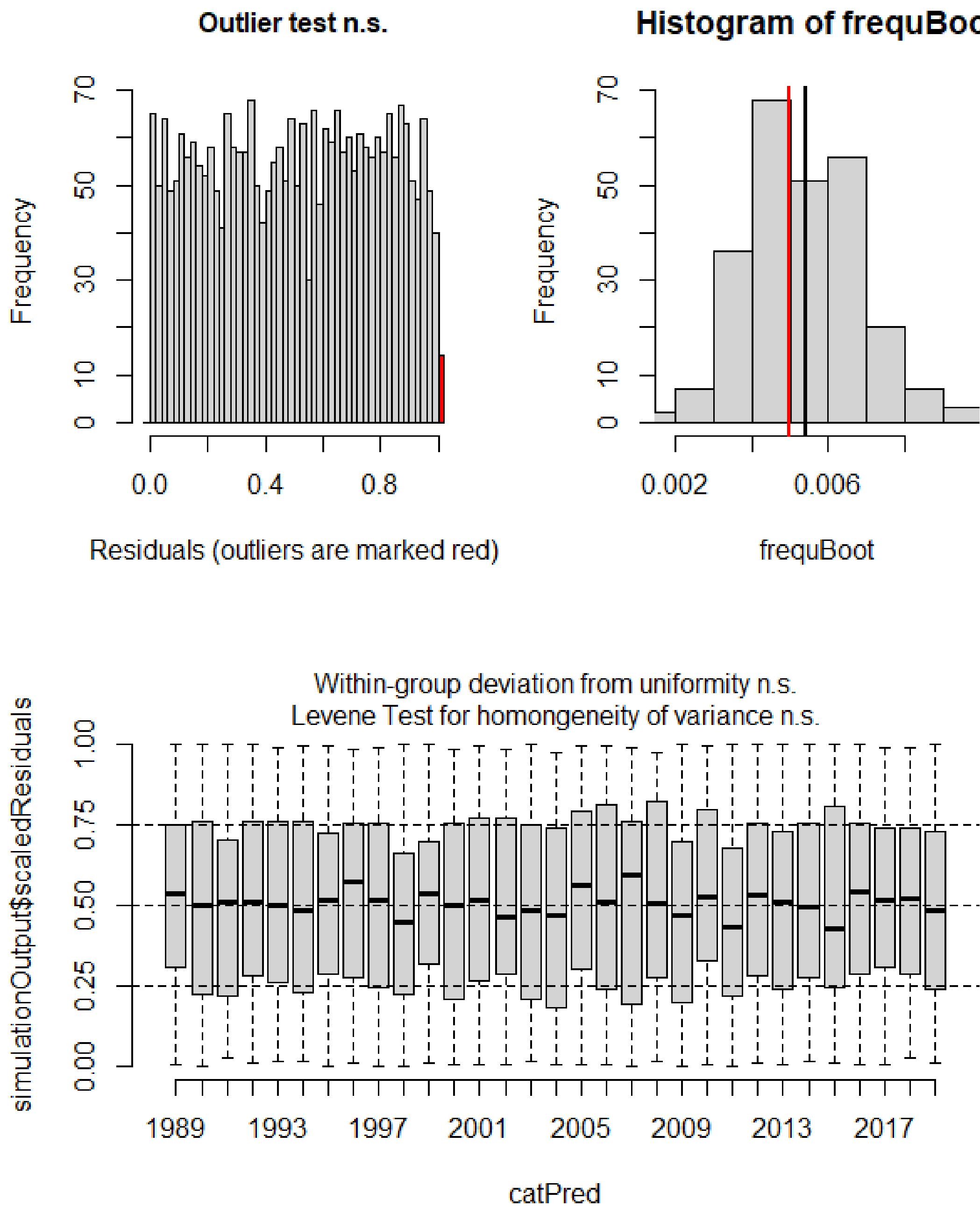

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

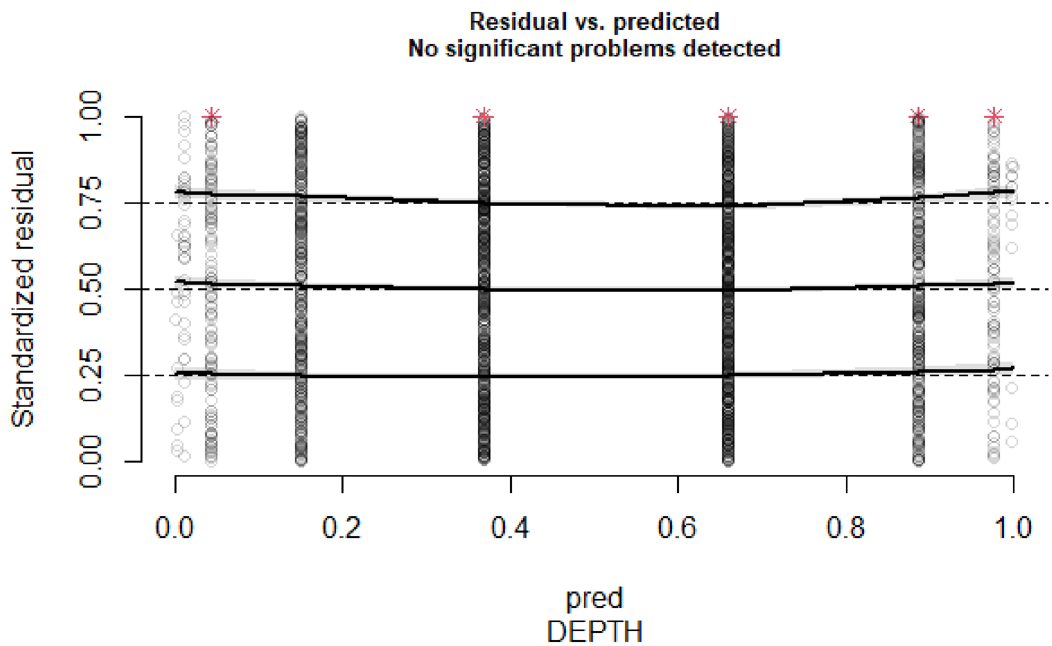

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

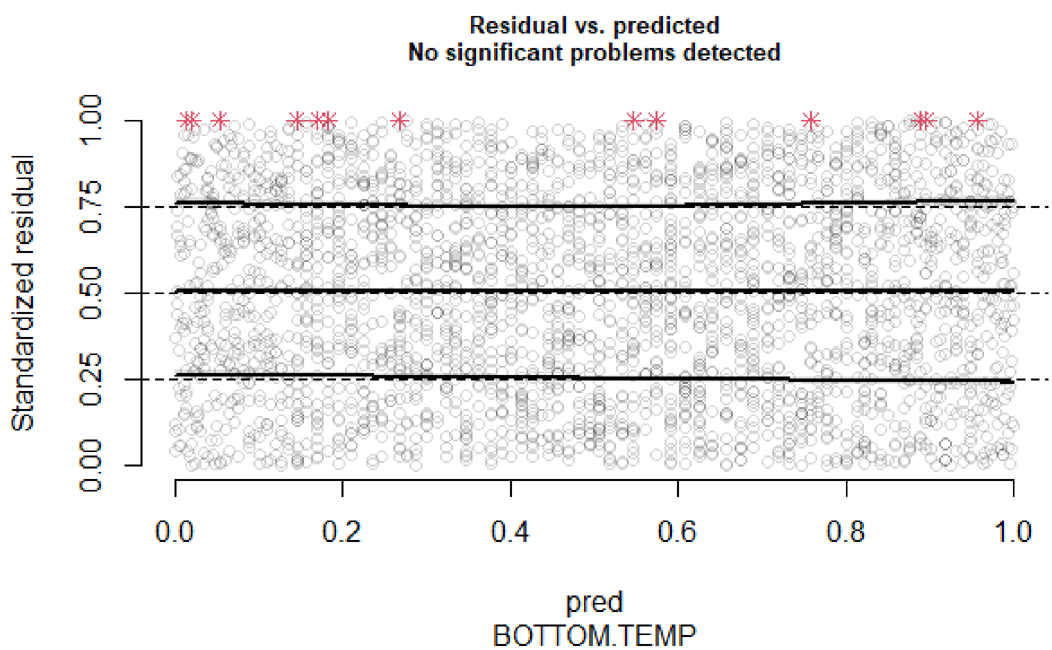

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test n.s.
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY

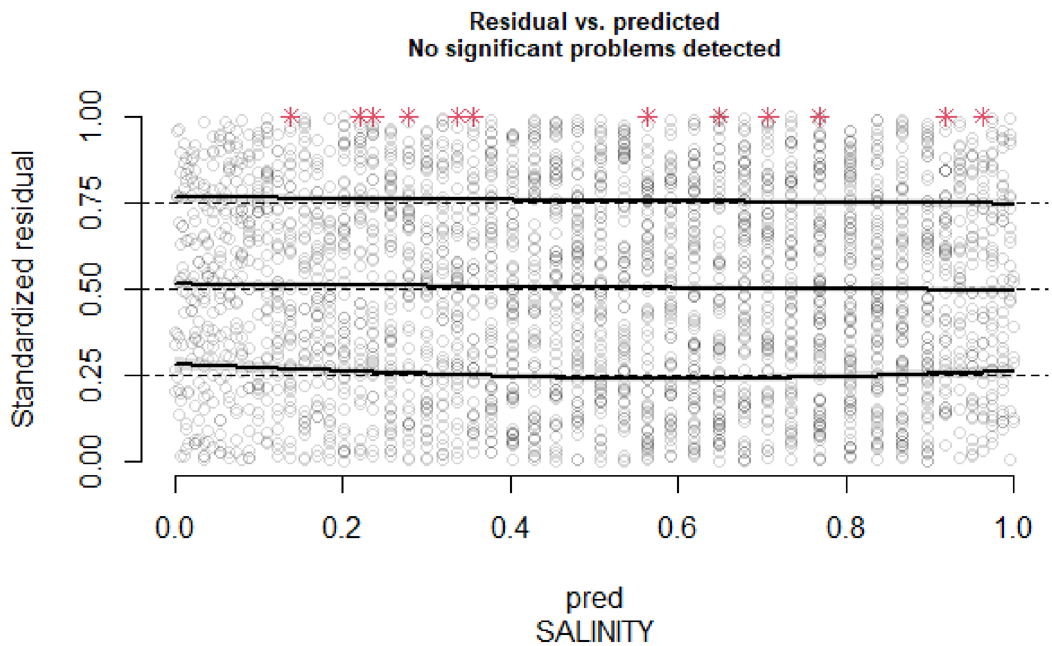

Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
21
25
29
33
37
41
45
49
53
57
61
65
catPred

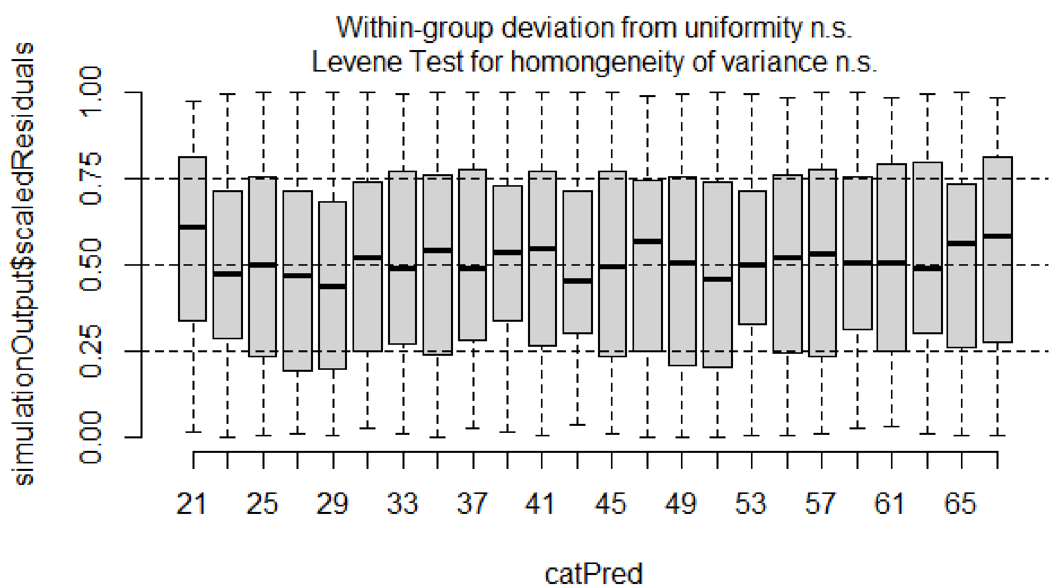

DHARMa residual diagnostics
QQ plot residuals
KS test: p= 0.34905
Deviation n.s.
Dispersion test: p= 0.51
Deviation n.s.
Outlier test: p= 0.03249
Deviation significant
Observed
Expected
0.0
0.4
0.8
Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.50
1.00
Model predictions (rank transformed)


Outlier test n.s.
Frequency
0
10
30
50
70
0.0
0.4
0.8
Residuals (outliers are marked red)
Histogram of frequBoot
Frequency
0
10
30
50
70
0.002
0.006
frequBoot
Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1989
1993
1997
2001
2005
2009
2013
2017
catPred

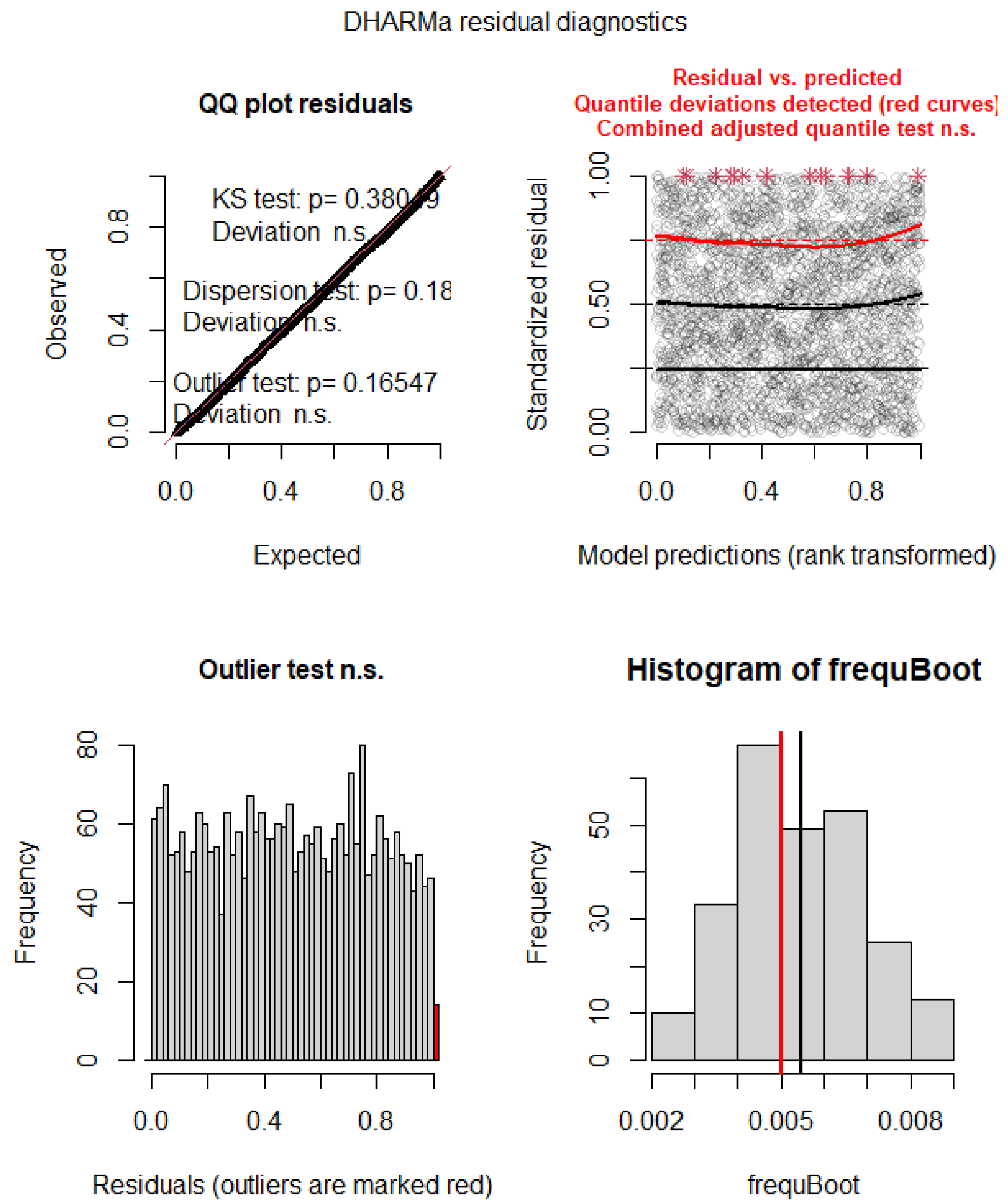

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

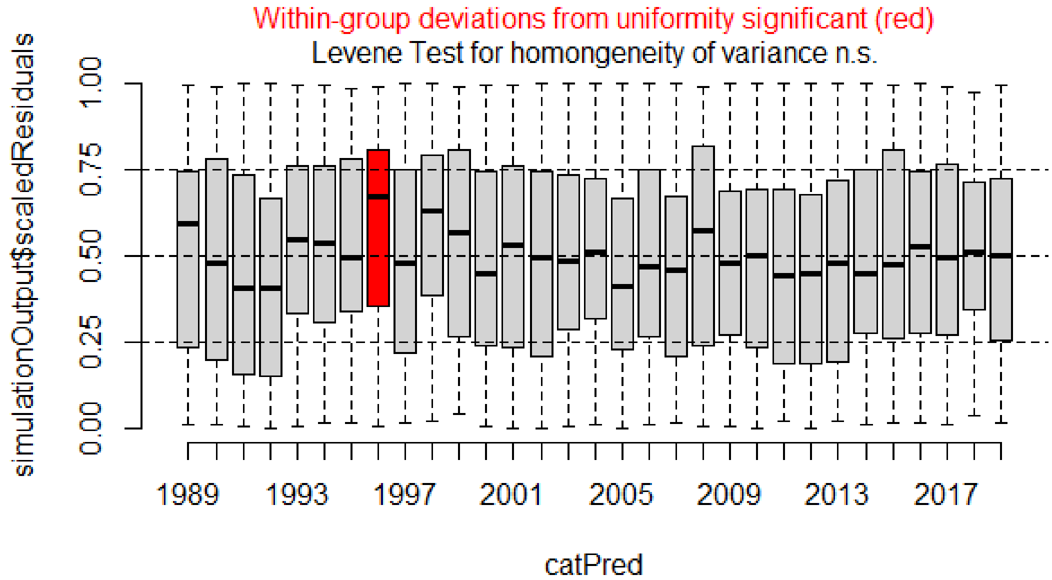

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

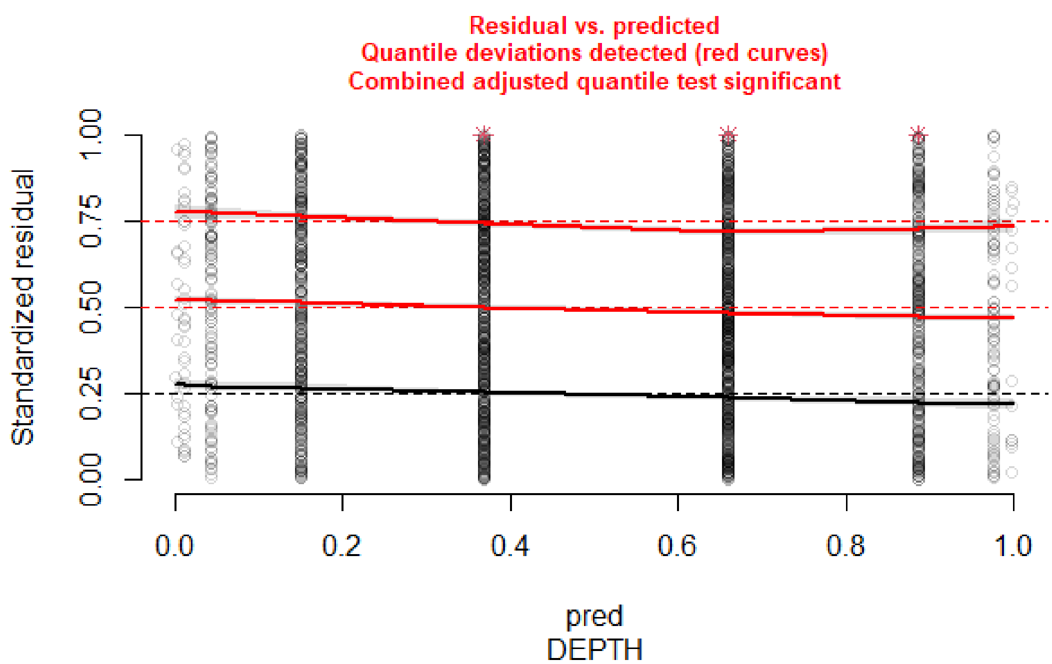

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY

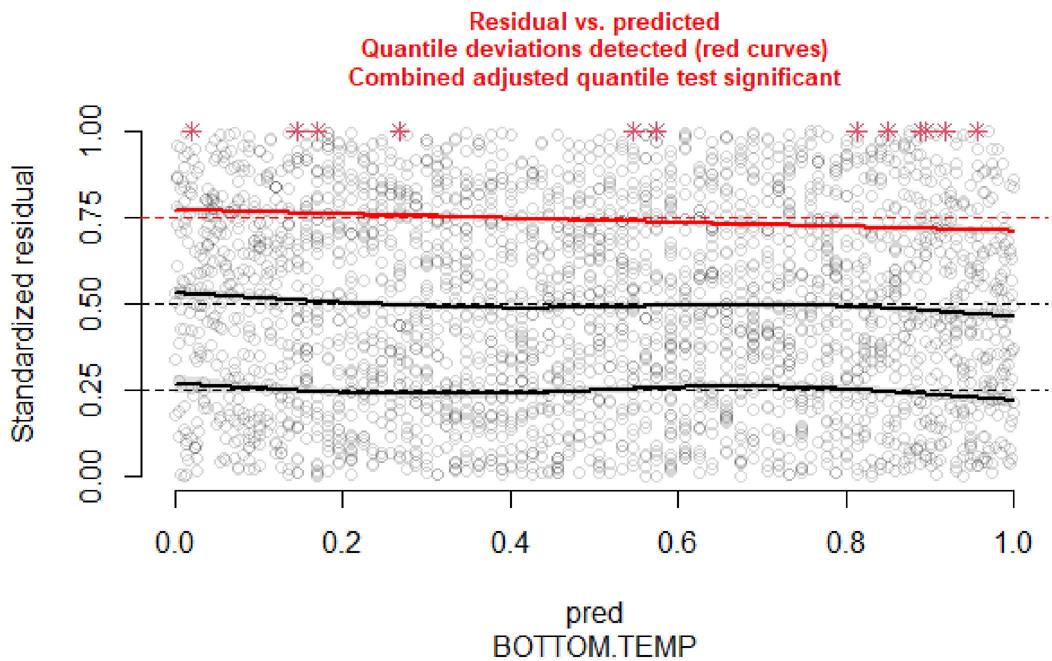

Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
21
25
29
33
37
41
45
49
53
57
61
65
catPred

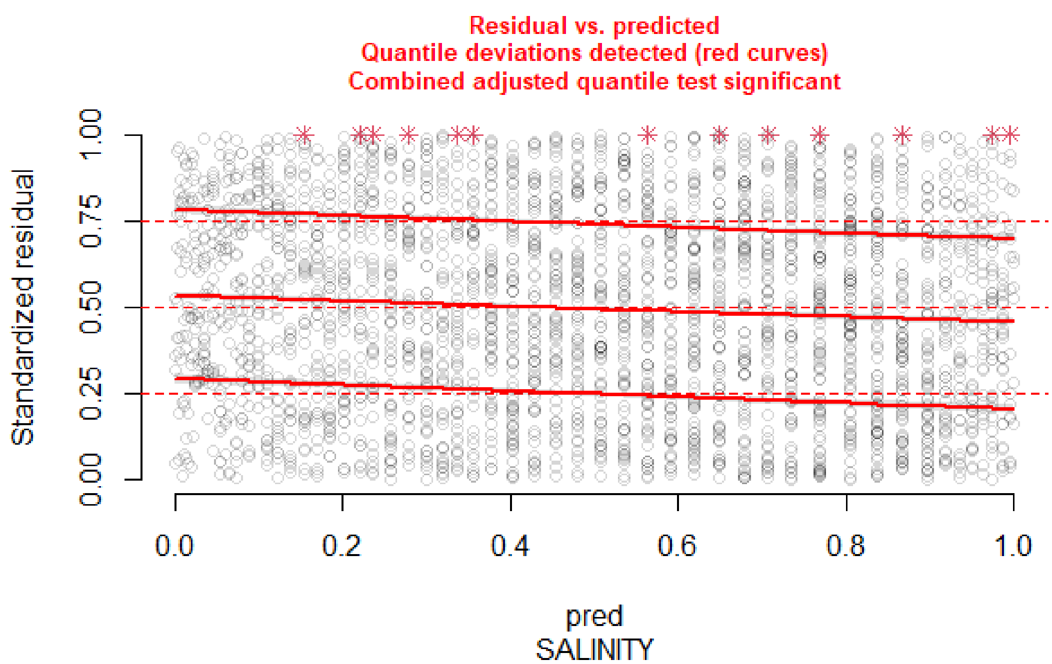
Simulated residuals for the full model with area as a covariate instead of latitude.

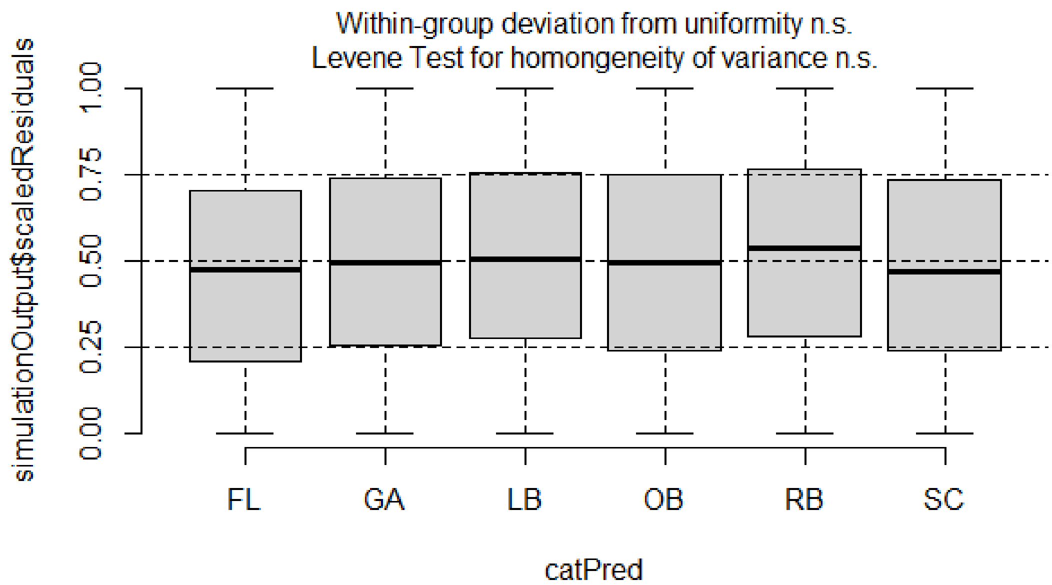


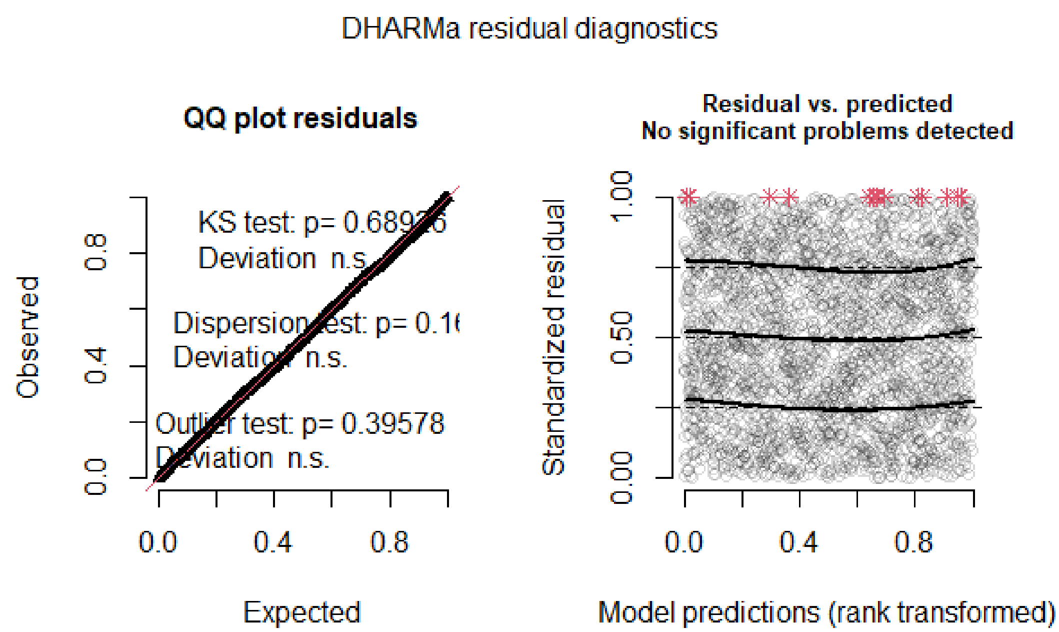

Within-group deviations from uniformity significant (red)
Levene Test for homogeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1989
1993
1997
2001
2005
2009
2013
2017
catPred

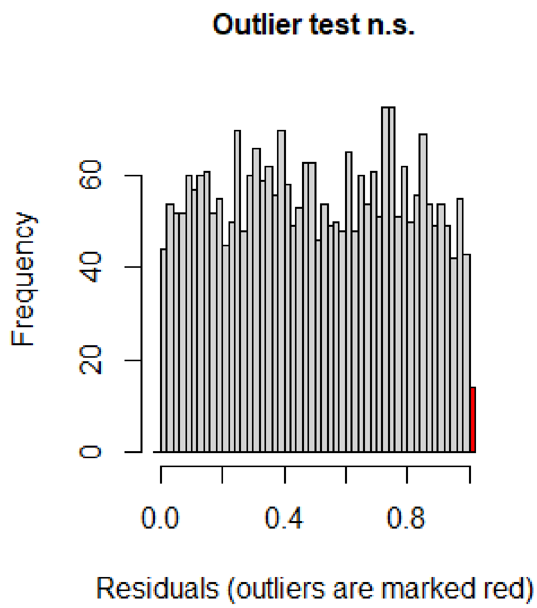

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH


Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

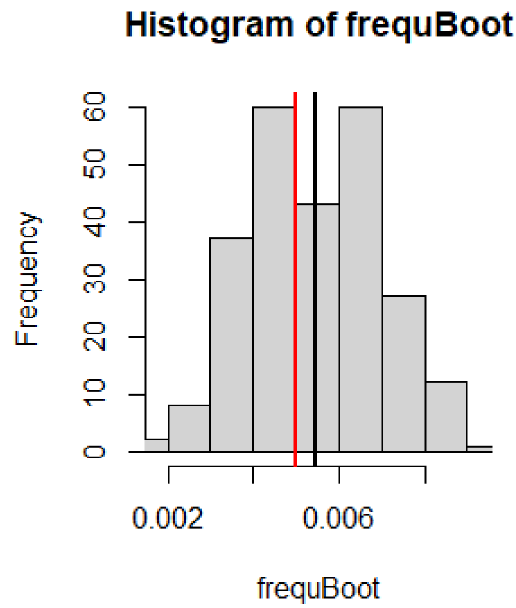

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY


Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
FL
GA
LB
OB
RB
SC
catPred

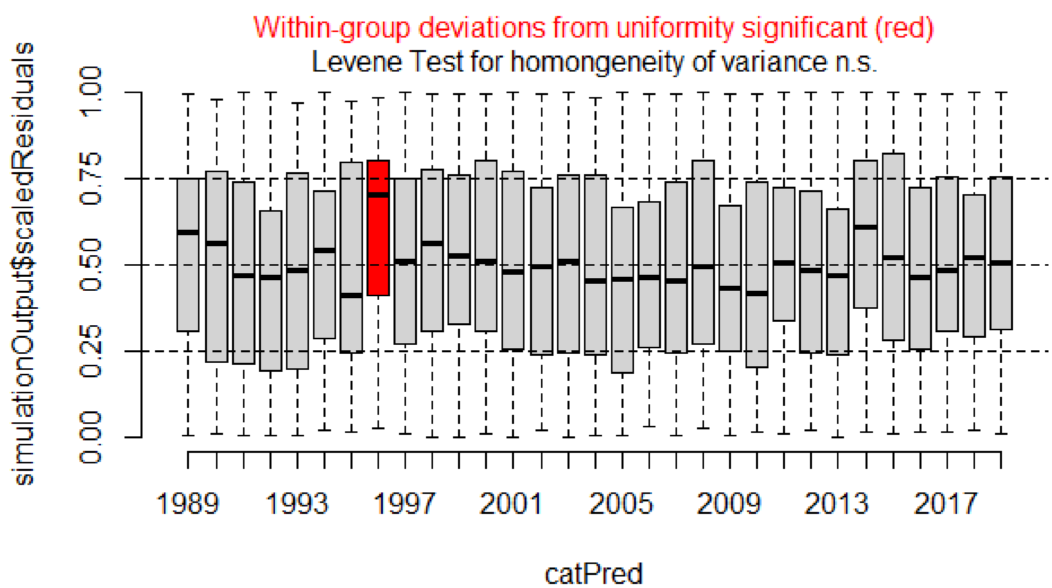

DHARMa residual diagnostics
QQ plot residuals
KS test: p= 0.68976
Deviation n.s.
Dispersion test: p= 0.16
Deviation n.s.
Outlier test: p= 0.39578
Deviation n.s.
Observed
Expected
0.0
0.4
0.8
Residual vs. predicted
No significant problems detected
Standardized residual
Model predictions (rank transformed)
0.00
0.50
1.00



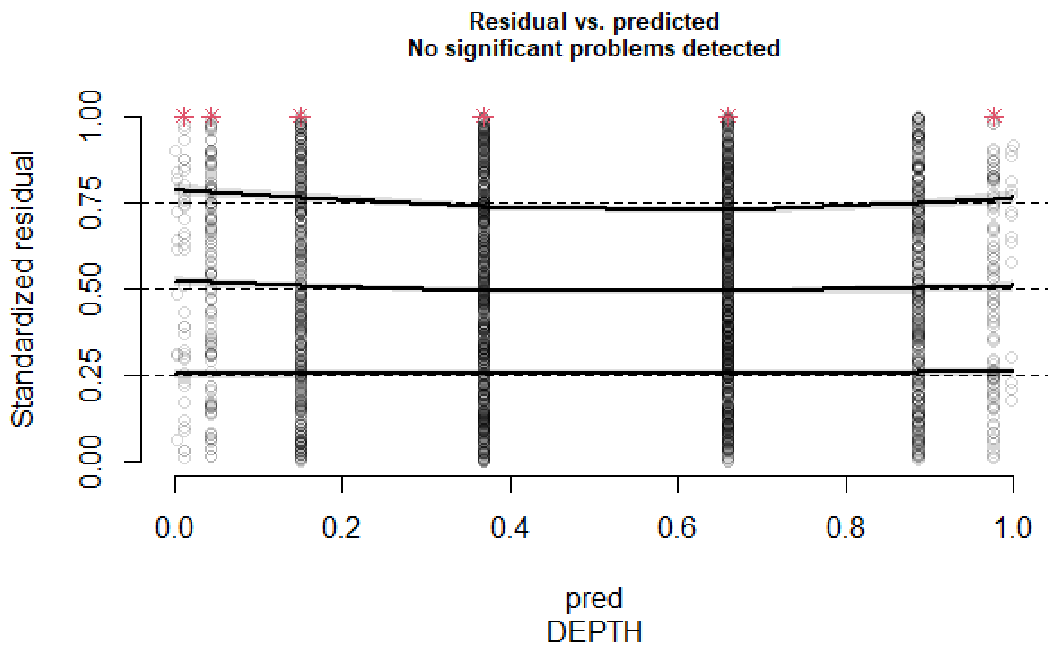
Outlier test n.s.
Frequency
0
20
40
60
0.0
0.4
0.8
Residuals (outliers are marked red)

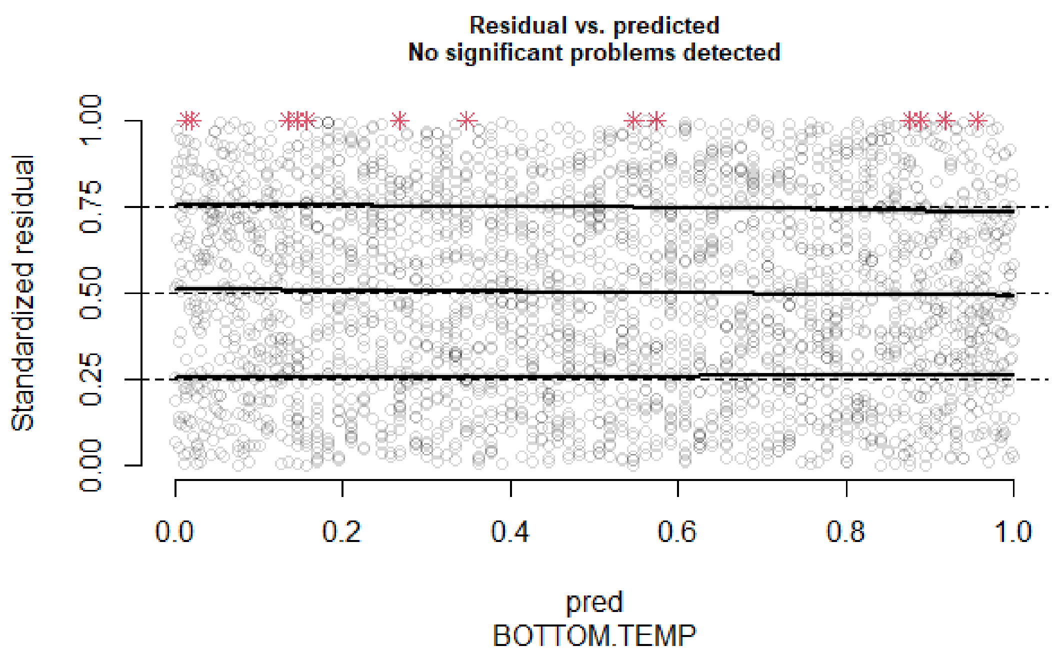


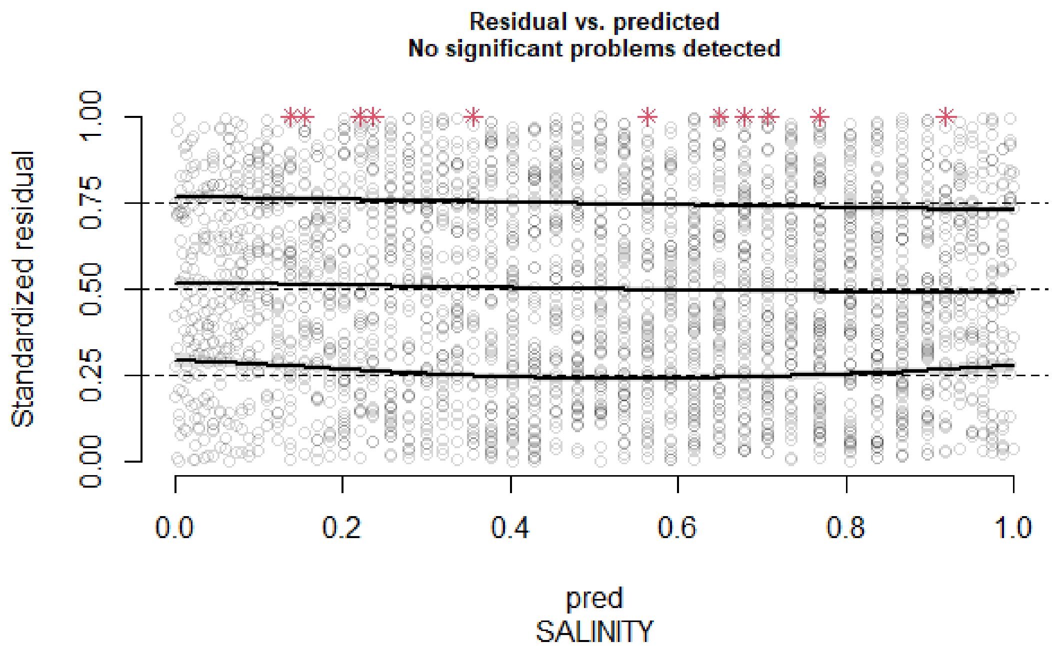
Histogram of frequBoot
Frequency
0
10
20
30
40
50
60
0.002
0.006
frequBoot

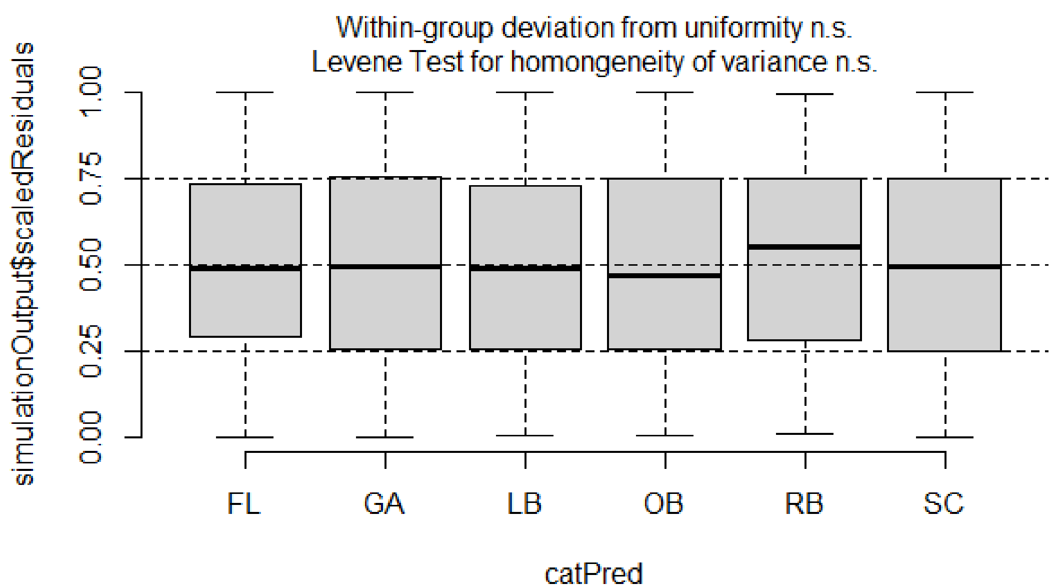


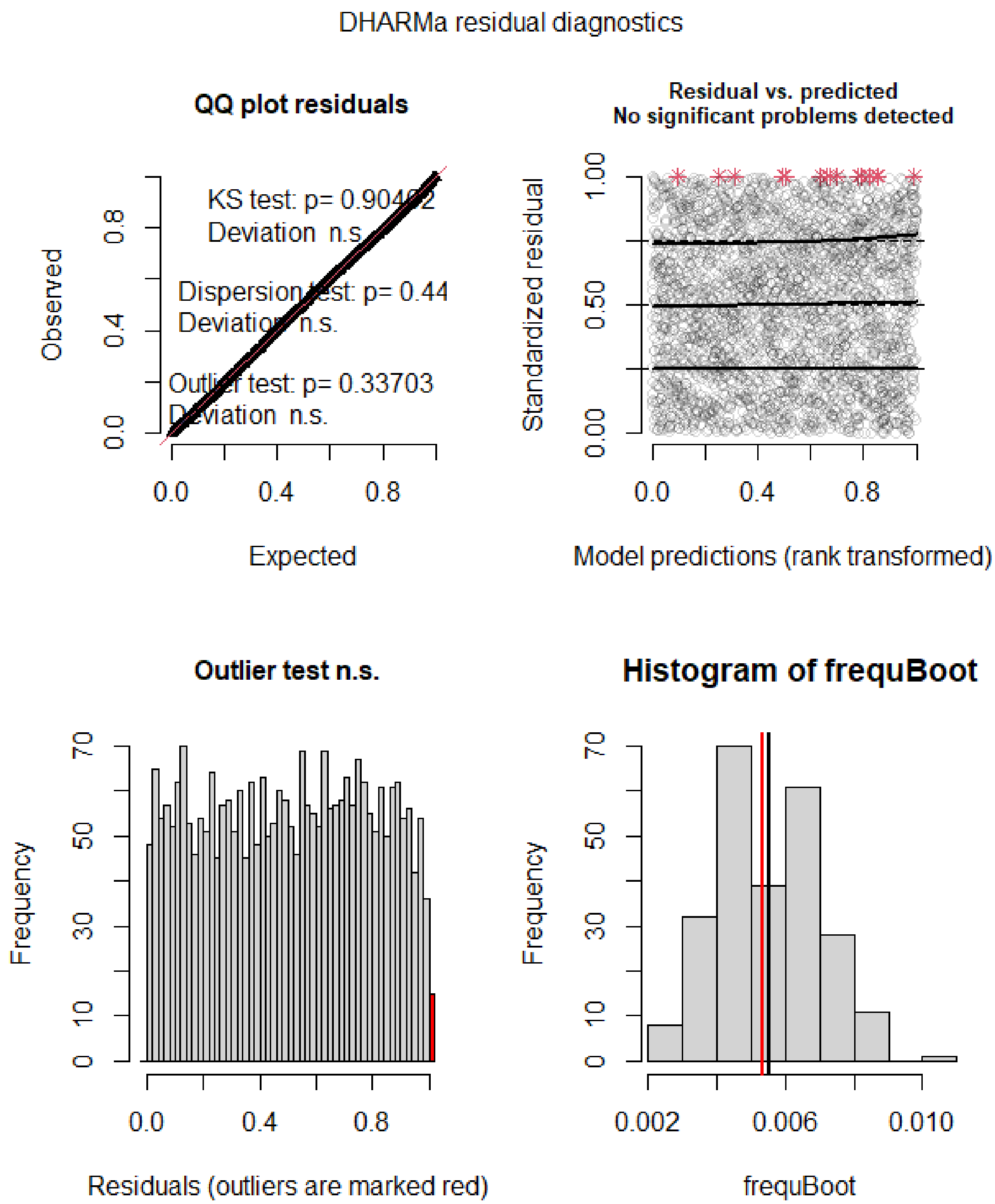
Within-group deviations from uniformity significant (red)
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1989
1993
1997
2001
2005
2009
2013
2017
catPred


Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

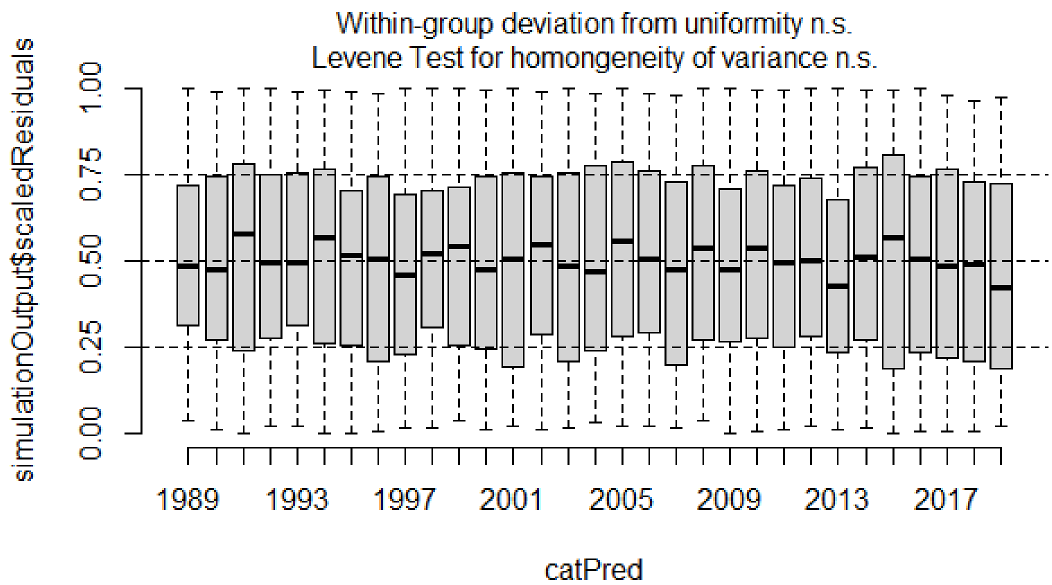

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

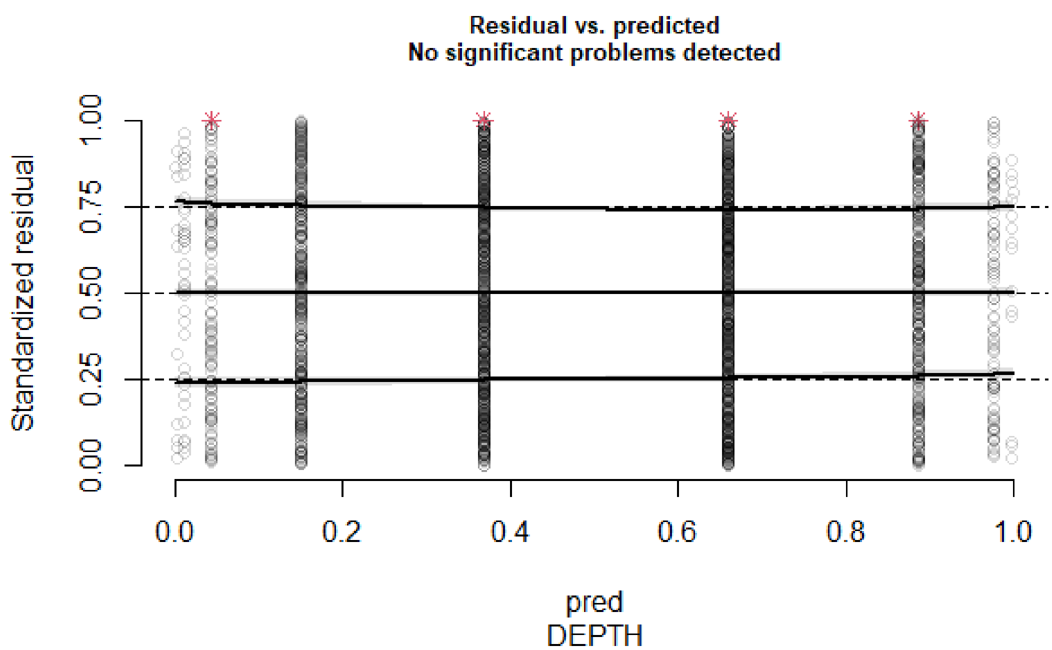

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY

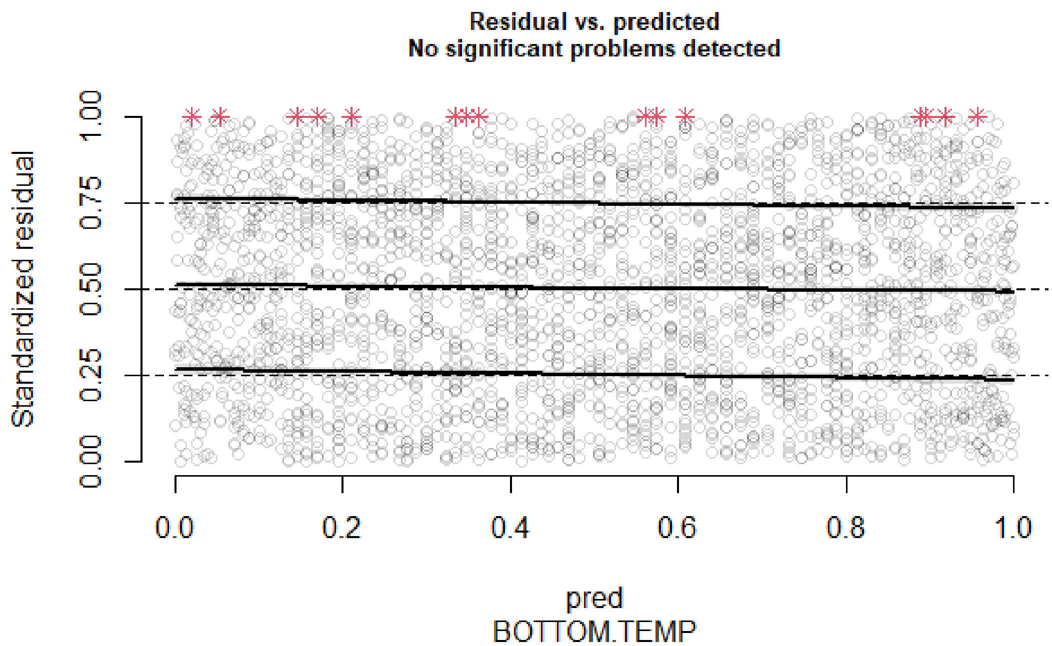

Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
FL
GA
LB
OB
RB
SC
catPred

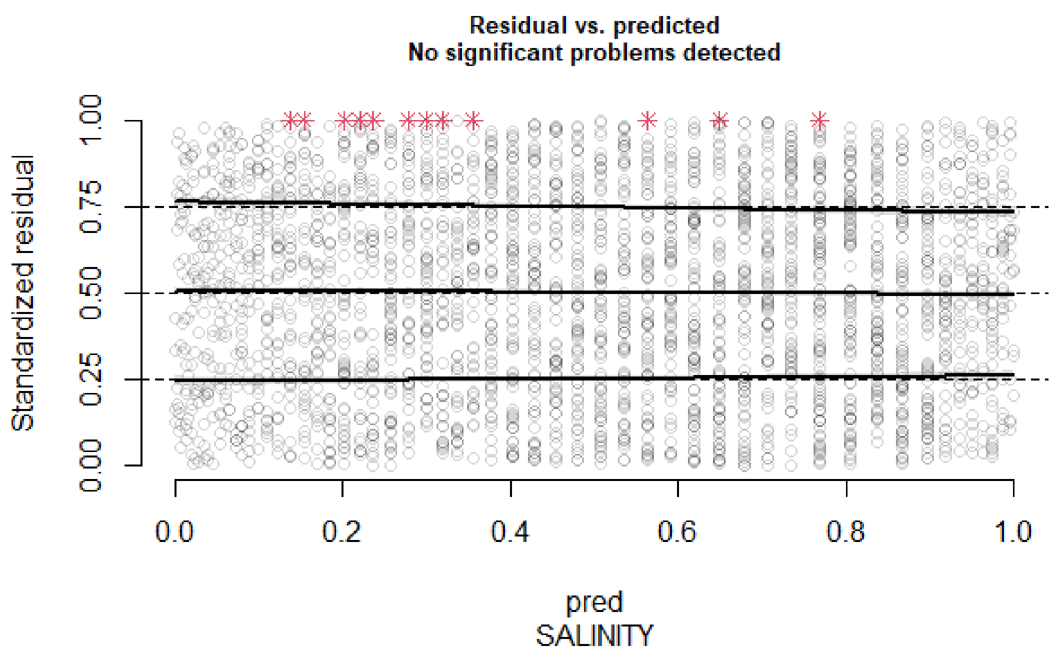


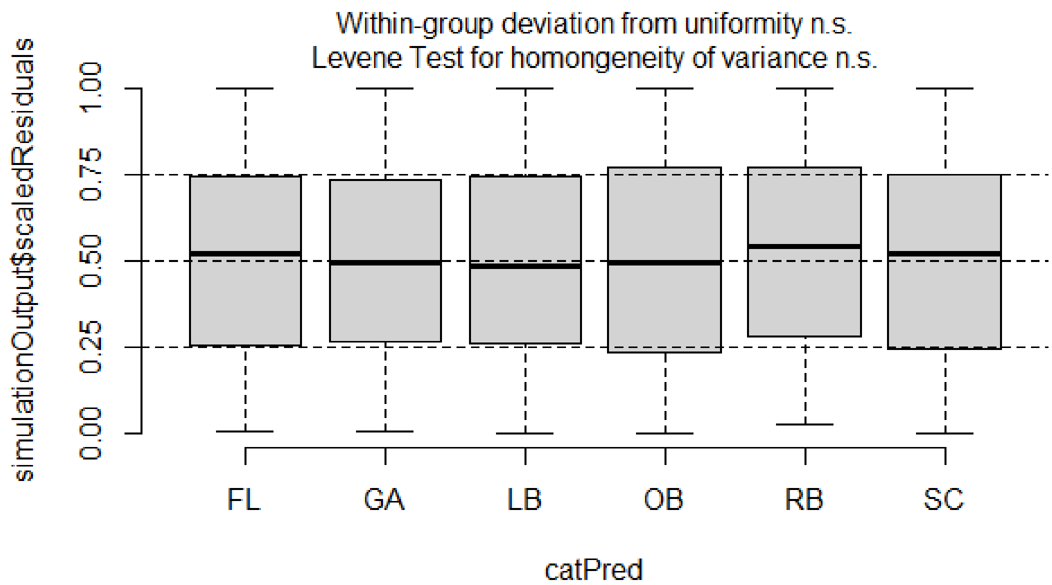
DHARMa residual diagnostics
QQ plot residuals
KS test: p= 0.90402
Deviation n.s.
Dispersion test: p= 0.44
Deviation n.s.
Outlier test: p= 0.33703
Deviation n.s.
Observed
Expected
0.0
0.4
0.8
Residual vs. predicted
No significant problems detected
Standardized residual
Model predictions (rank transformed)
0.00
0.50
1.00
Outlier test n.s.
Frequency
Residuals (outliers are marked red)
0
10
30
50
70
Histogram of frequBoot
frequBoot
0.002
0.006
0.010


Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1989
1993
1997
2001
2005
2009
2013
2017
catPred


Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH


Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP


Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY



Step 3: Winning model

Zero-altered negative binomial with stratum as factor has the best diagnostics as well as the lowest AIC.

```
zero_alt <- full_mod_stratum$zero_adj

null_mod <- glmmTMB(FREQ ~ 1,
                    ziformula =full_mod_stratum$zero_adj$call$ziformula,
                    data = new_data2,
                    family = truncated_nbinom2(link = "log"))
```

Test covariate importance

```
select_model(data = new_data2,
             model = zero_alt,
             null_model = null_mod) %>%
  try()

## Forward model selection:

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ 1
##
## Final Model:
## FREQ ~ STRATUM + YEAR + BOTTOM.TEMP
##
##
```

```
##            Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                                2744   7508.289 7626.289
## 2     + STRATUM 23 392.43803      2721   7115.851 7279.851
## 3        + YEAR 30 167.21925      2691   6948.632 7172.632
## 4 + BOTTOM.TEMP  1  33.72472      2690   6914.907 7140.907

## Bidirectional model selection:

## Stepwise Model Path 
## Analysis of Deviance Table
## 
## Initial Model:
## FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## 
## Final Model:
## FREQ ~ YEAR + BOTTOM.TEMP + STRATUM
## 
## 
##         Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                              2688   6912.616 7142.616
## 2 - SALINITY  1 0.4976379      2689   6913.113 7141.113
## 3    - DEPTH  1 1.7937454      2690   6914.907 7140.907

## Backward model selection:

## Stepwise Model Path 
## Analysis of Deviance Table
## 
## Initial Model:
## FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## 
## Final Model:
## FREQ ~ YEAR + BOTTOM.TEMP + STRATUM
## 
## 
##         Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                              2688   6912.616 7142.616
## 2 - SALINITY  1 0.4976379      2689   6913.113 7141.113
## 3    - DEPTH  1 1.7937454      2690   6914.907 7140.907
```

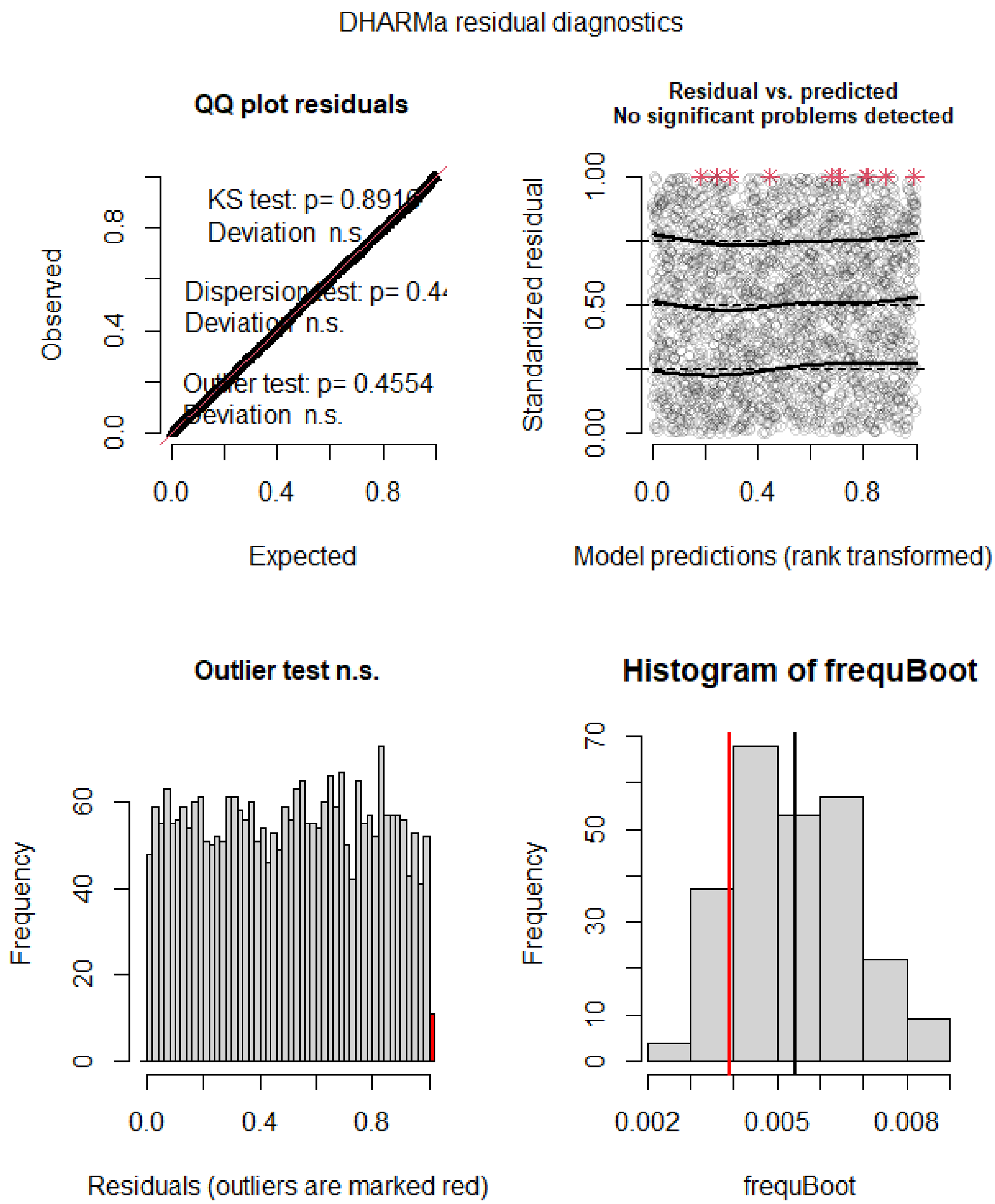
Model selection functions do not include the zero-altered component.

```
ZANB <- glmmTMB(FREQ ~ YEAR+BOTTOM.TEMP+STRATUM,
                    ziformula = ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY +
STRATUM,
                    data = new_data,
                    family = truncated_nbinom2(link = "log"))

ZANB.2 <- glmmTMB(FREQ ~YEAR+BOTTOM.TEMP+STRATUM,
                    ziformula = ~YEAR + DEPTH + BOTTOM.TEMP + STRATUM,
                    data = new_data2,
```

```
                    family = truncated_nbinom2(link = "log"))

ZANB.2.b <- glmmTMB(FREQ ~YEAR+DEPTH + BOTTOM.TEMP+STRATUM,
                    ziformula = ~YEAR + DEPTH + BOTTOM.TEMP + STRATUM,
                    data = new_data2,
                    family = truncated_nbinom2(link = "log"))

ZANB.3 <- glmmTMB(FREQ ~ YEAR+BOTTOM.TEMP+STRATUM,
                    ziformula = ~YEAR +BOTTOM.TEMP + STRATUM,
                    data = new_data2,
                    family = truncated_nbinom2(link = "log"))

ZANB.4 <- glmmTMB(FREQ ~ YEAR+BOTTOM.TEMP+STRATUM,
                    ziformula = ~YEAR +STRATUM,
                    data = new_data2,
                    family = truncated_nbinom2(link = "log"))

ZANB.5 <- glmmTMB(FREQ ~ YEAR+BOTTOM.TEMP+STRATUM,
                    ziformula = ~YEAR,
                    data = new_data2,
                    family = truncated_nbinom2(link = "log"))

ZANB.6 <- glmmTMB(FREQ ~ YEAR+BOTTOM.TEMP+STRATUM,
                    ziformula = ~STRATUM,
                    data = new_data2,
                    family = truncated_nbinom2(link = "log"))

ZANB.7 <-  glmmTMB(FREQ ~ YEAR+BOTTOM.TEMP+STRATUM,
                    ziformula = ~1,
                    data = new_data2,
                    family = truncated_nbinom2(link = "log"))

AICtab(ZANB, ZANB.2, ZANB.2.b, ZANB.3, ZANB.4, ZANB.5, ZANB.6, ZANB.7)

##          dAIC  df
## ZANB.2     0.0 112
## ZANB.2.b   0.2 113
## ZANB       1.2 113
## ZANB.3    16.3 111
## ZANB.4   163.8 110
## ZANB.6   279.1 80
## ZANB.5   330.6 87
## ZANB.7   431.7 57
```

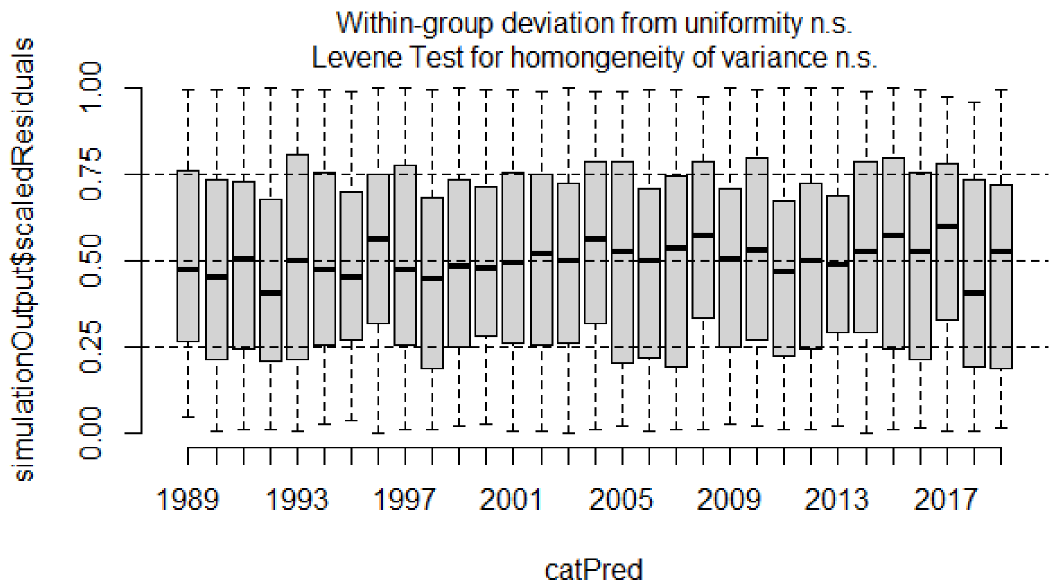
DHARMA simulations

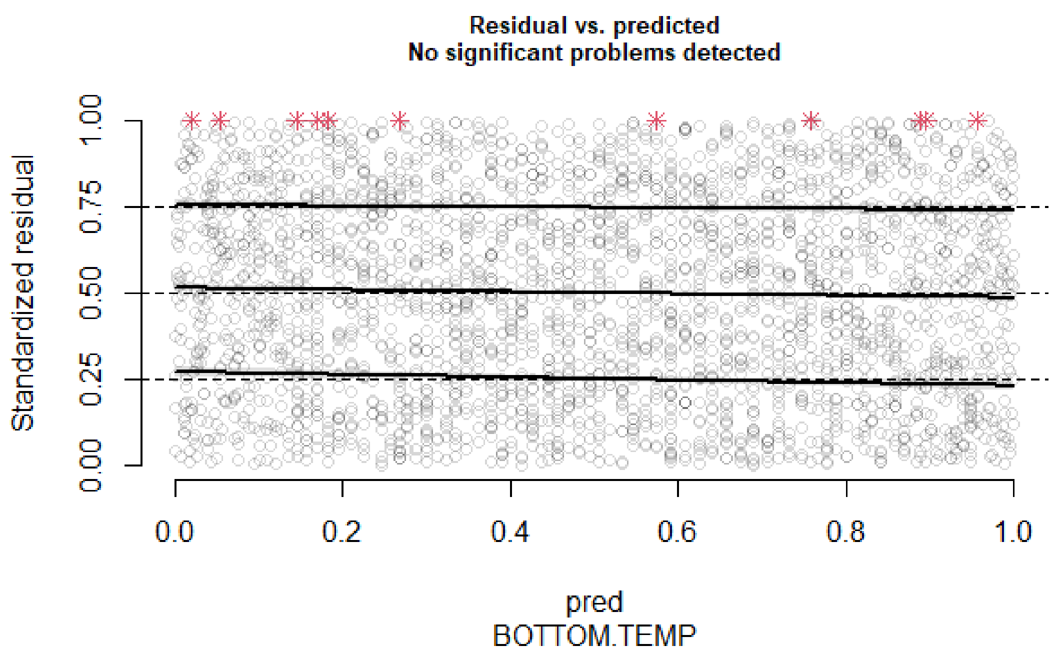
Compare residuals of the 3 models with the lowest AIC values: 1. Lowest AIC. FREQ ~YEAR+BOTTOM.TEMP+STRATUM, ziformula = ~YEAR + DEPTH + BOTTOM.TEMP + STRATUM

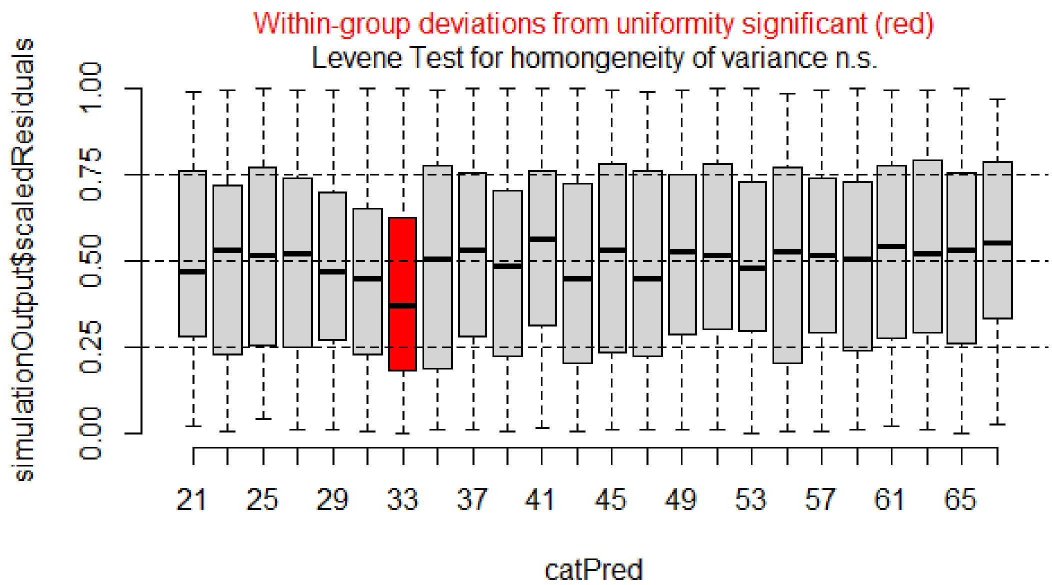


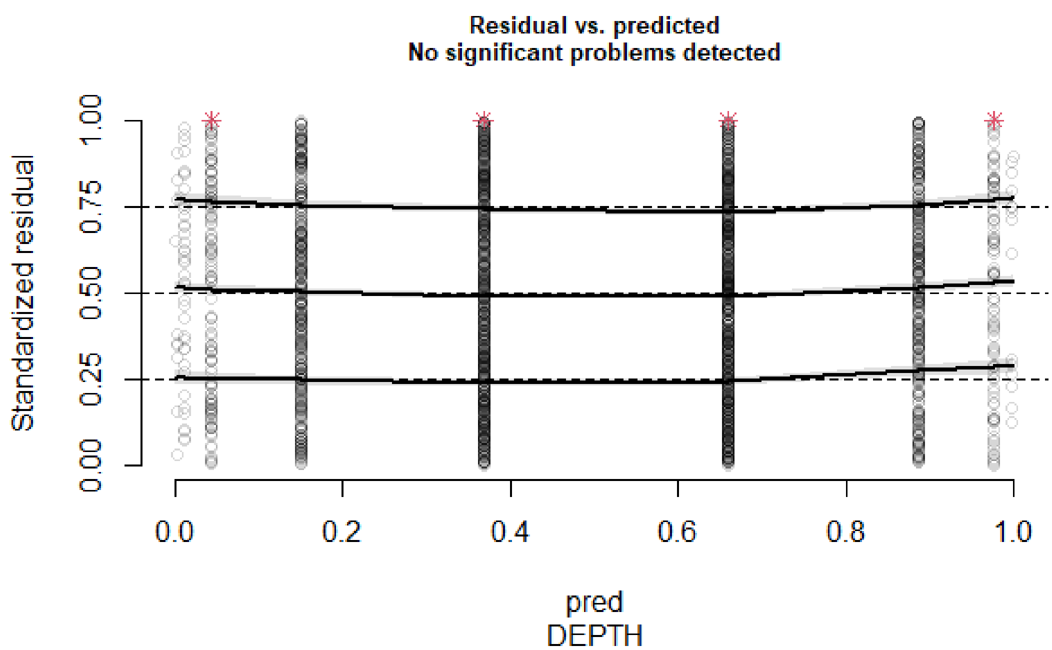

```
##
##  DHARMa bootstrapped outlier test
##
## data:  sim
## outliers at both margin(s) = 11, observations = 2803, p-value = 0.328
## alternative hypothesis: two.sided
##  percent confidence interval:
##  0.003210846 0.008205494
## sample estimates:
```

```
## outlier frequency (expected: 0.00541705315733143 )
##                                          0.003924367
```

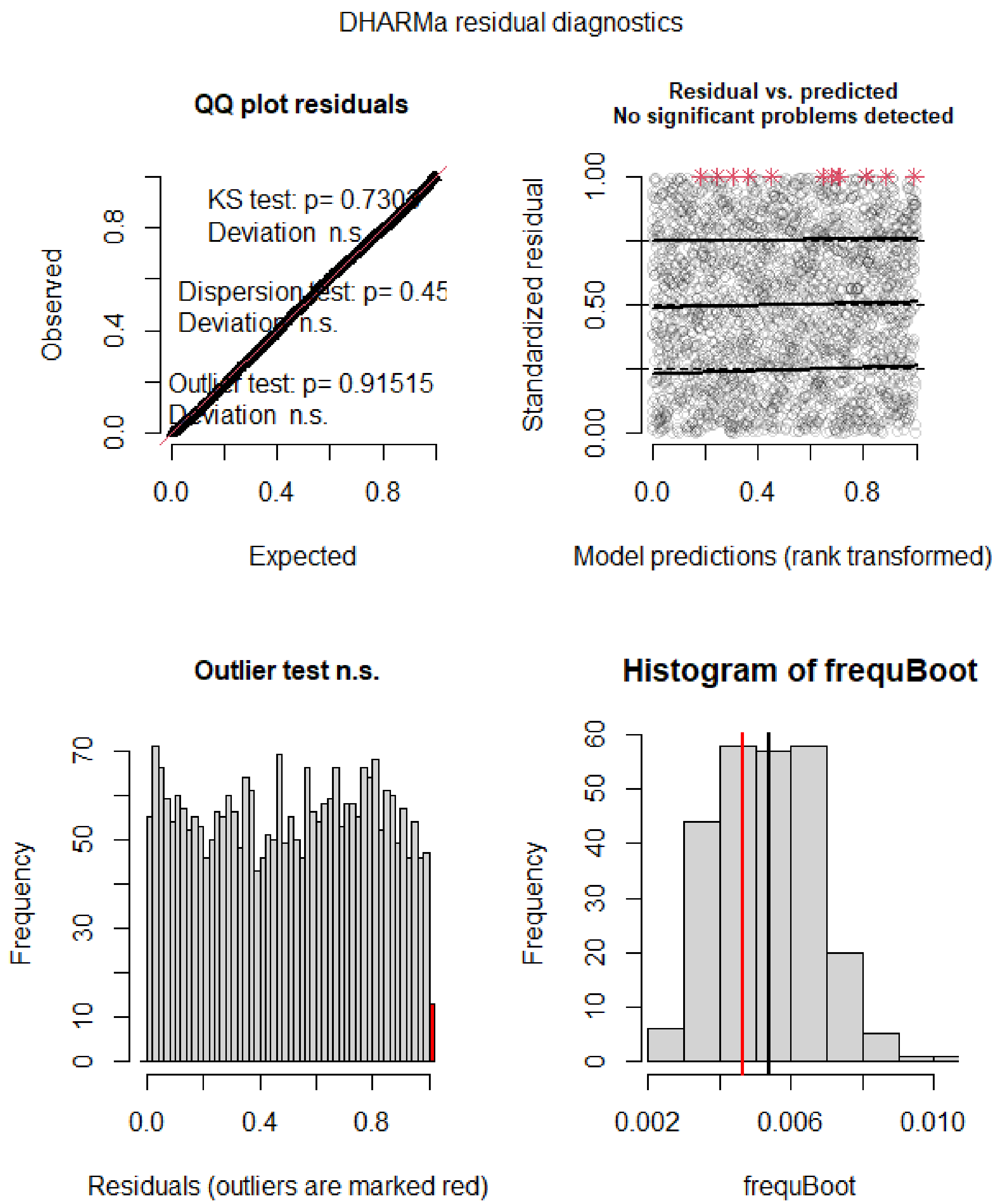






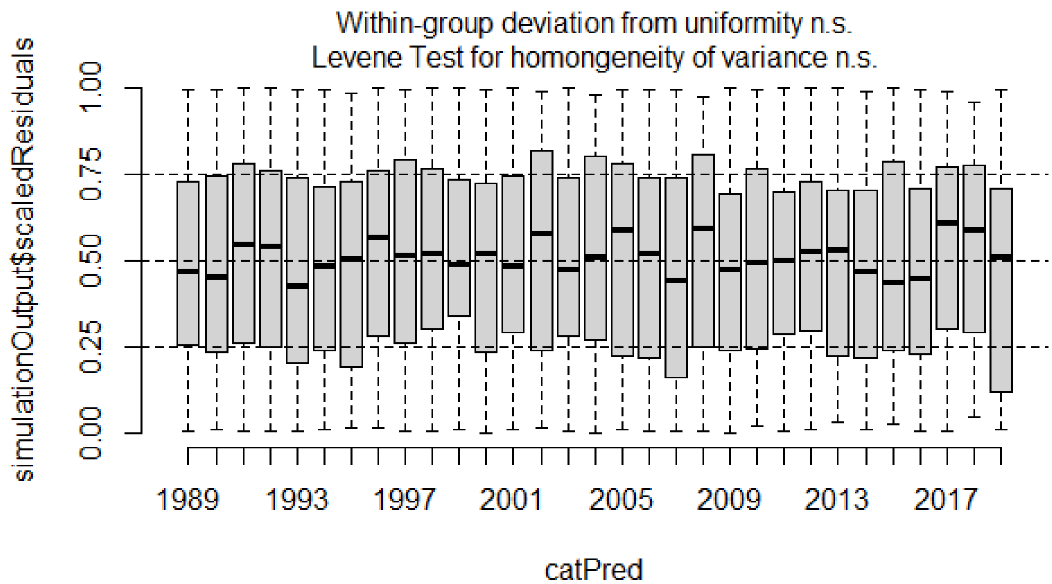


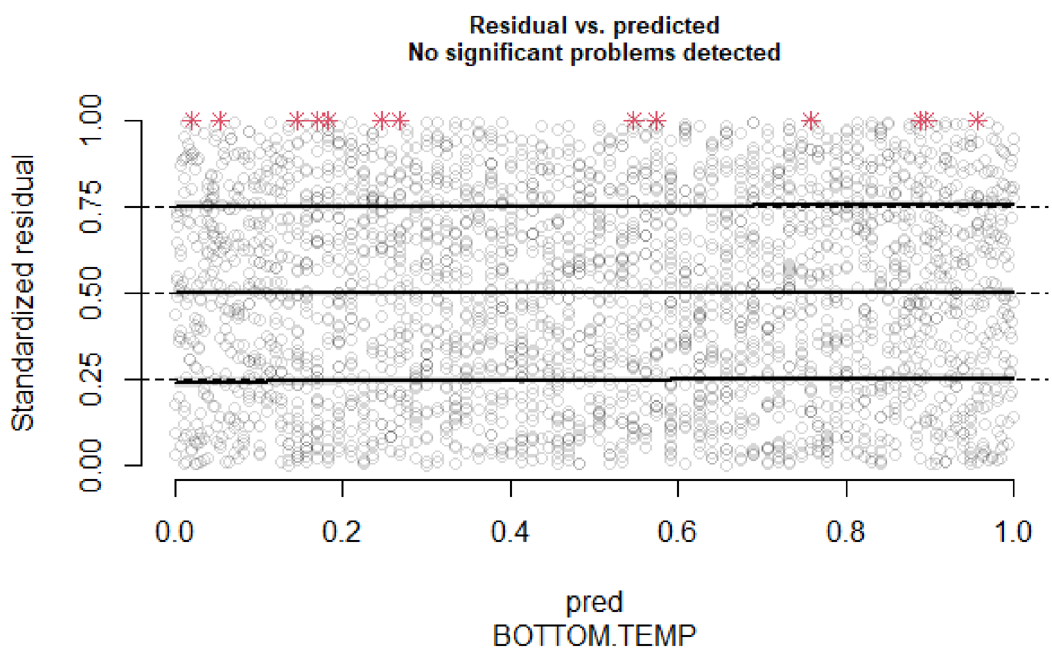
2. Full model for ZI-component. FREQ ~ YEAR+BOTTOM.TEMP+STRATUM, ziformula = ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM

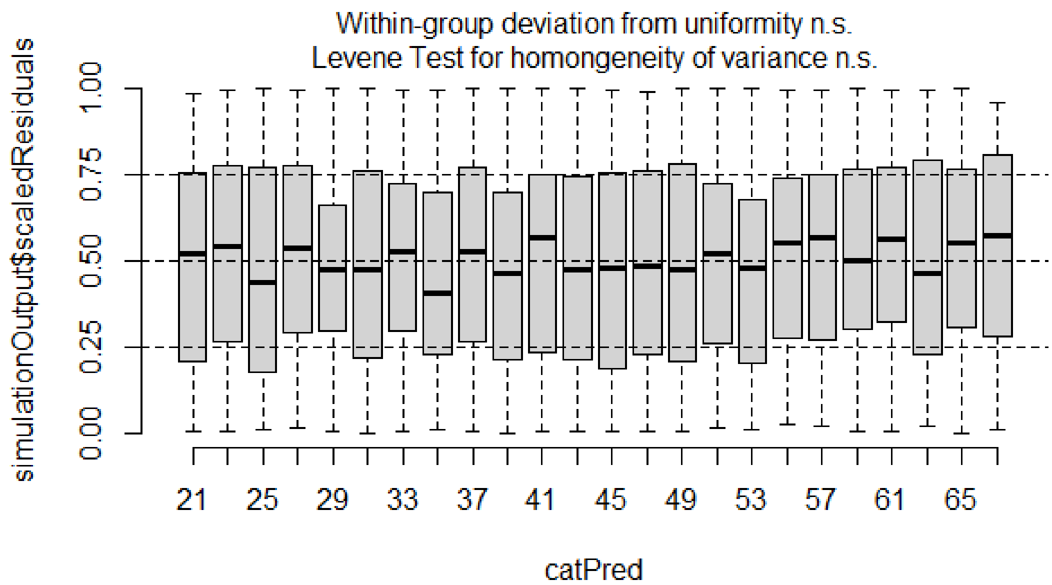


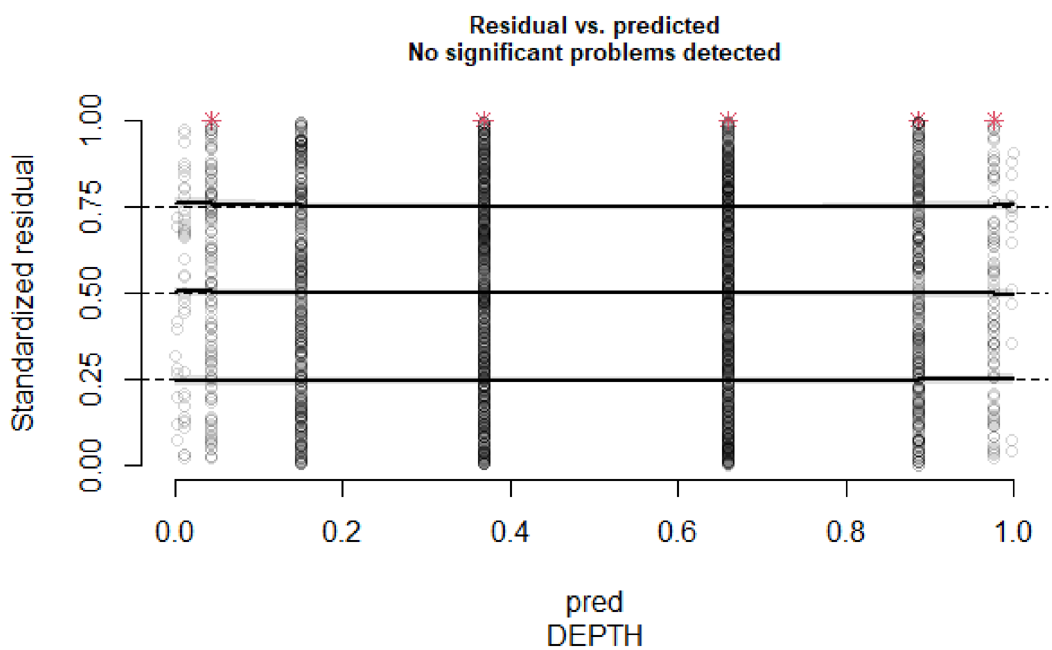

```
##
##  DHARMa bootstrapped outlier test
##
## data:  sim
## outliers at both margin(s) = 13, observations = 2803, p-value = 0.68
## alternative hypothesis: two.sided
##  percent confidence interval:
##  0.003210846 0.008125223
## sample estimates:
```

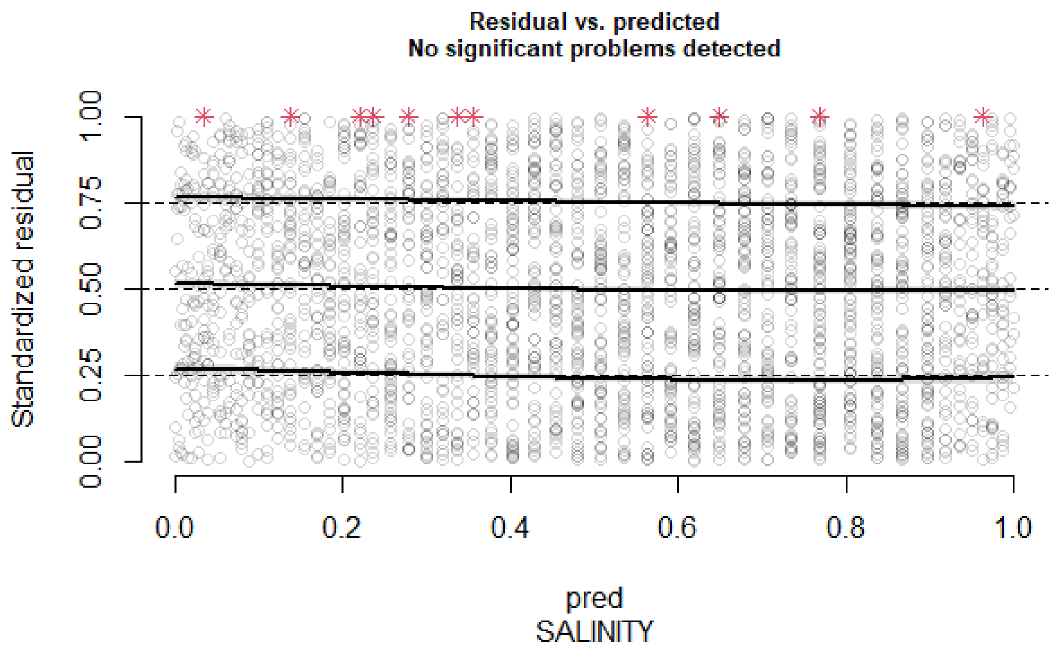
```
## outlier frequency (expected: 0.00536996075633325 )
##                                          0.004637888
```





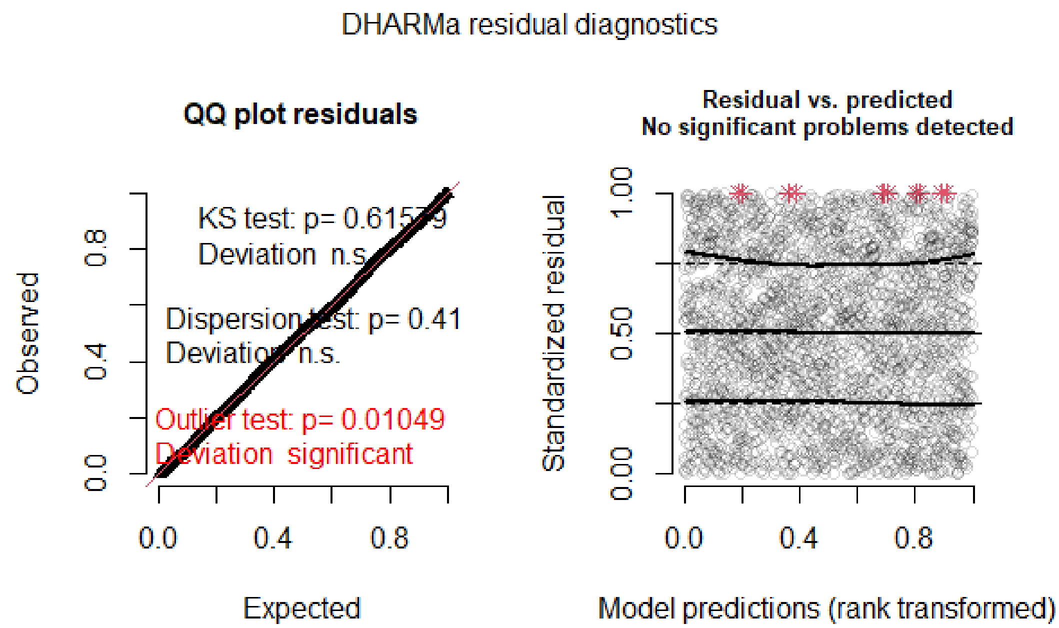

Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
21
25
29
33
37
41
45
49
53
57
61
65
catPred


Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

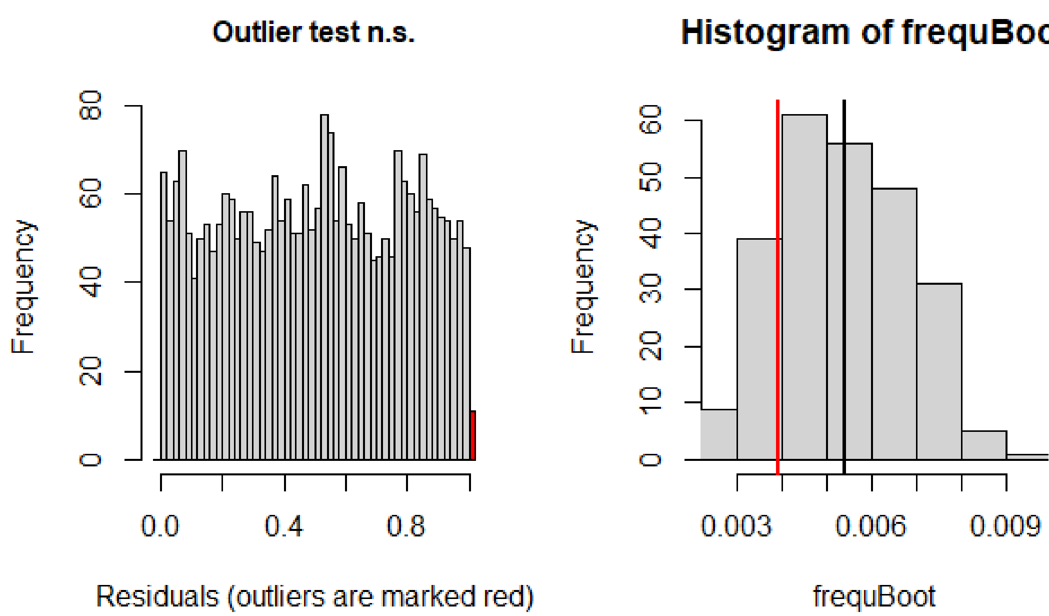


3. Depth included in count and ZI components. FREQ ~YEAR+DEPTH + BOTTOM.TEMP+STRATUM, ziformula = ~YEAR + DEPTH + BOTTOM.TEMP + STRATUM

```
sim <- simulateResiduals(ZANB.2.b)
plot(sim, quantreg=T)

## DHARMa:testOutliers with type = binomial may have inflated Type I error
rates for integer-valued distributions. To get a more exact result, it is
recommended to re-run testOutliers with type = 'bootstrap'. See ?testOutliers
for details
```

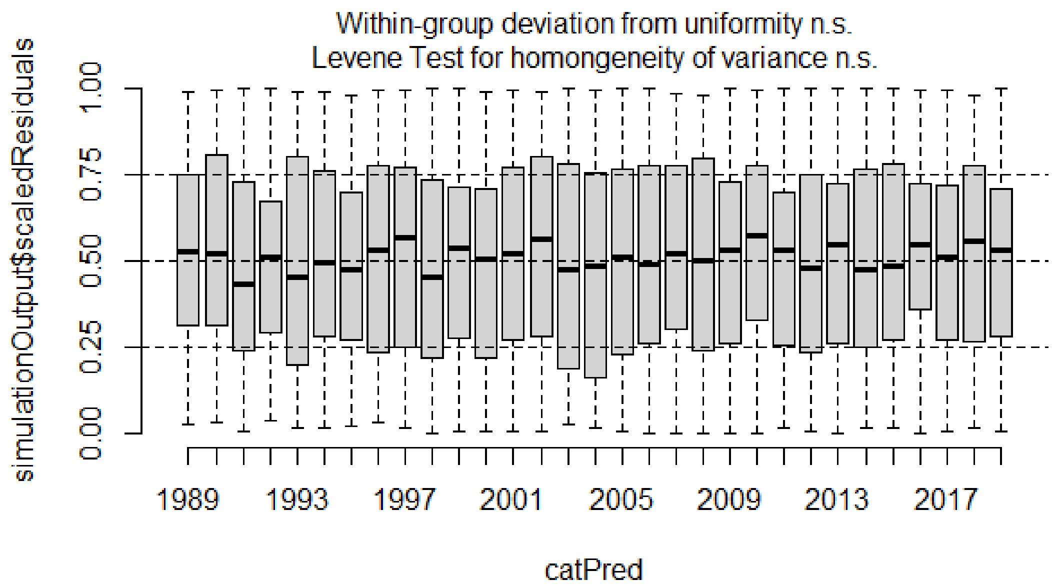


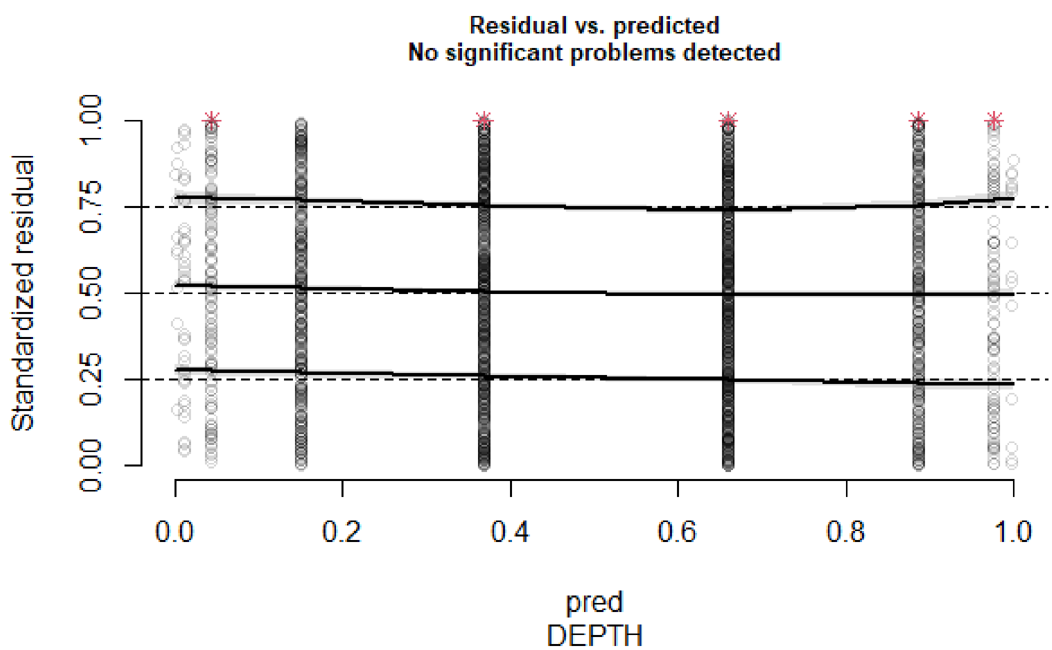

```r
testOutliers(sim, type="bootstrap", nBoot=250)
```

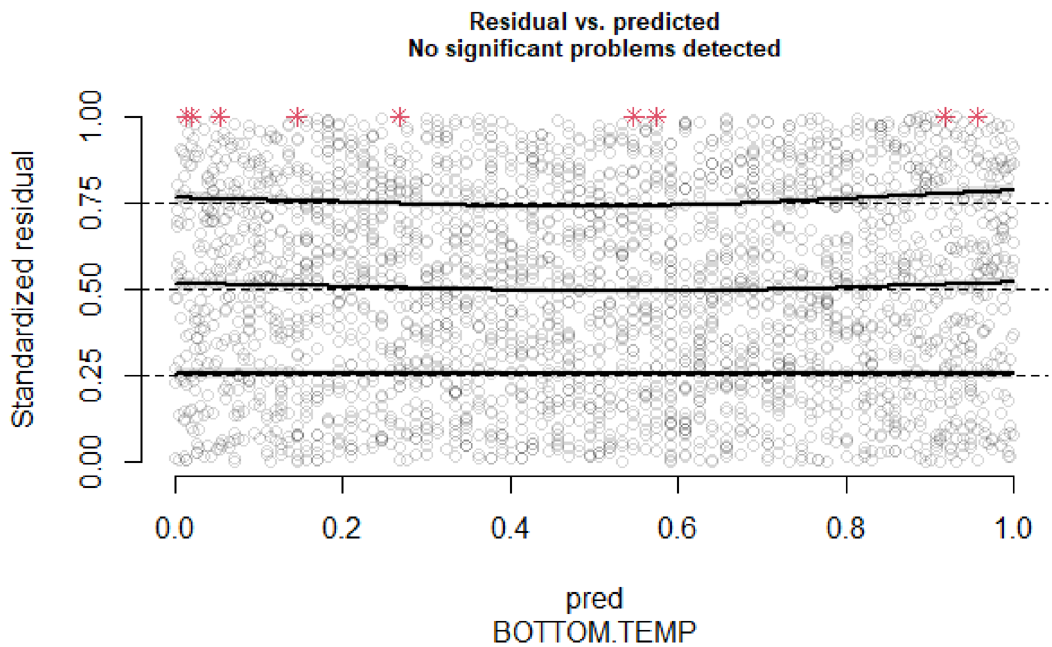


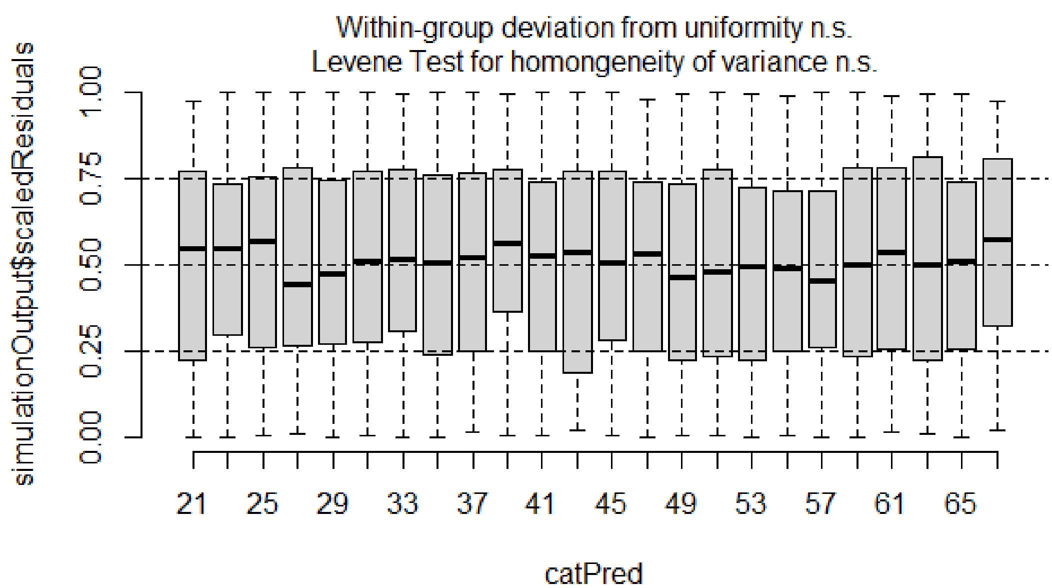

```
##
##  DHARMa bootstrapped outlier test
##
## data:  sim
## outliers at both margin(s) = 11, observations = 2803, p-value = 0.384
## alternative hypothesis: two.sided
##  percent confidence interval:
```

```
##  0.002854085 0.007848733
## sample estimates:
## outlier frequency (expected: 0.00537852301105958 )
##                                         0.003924367

# Compare residuals vs. factors
  vars <- names(ZANB.2.b$frame)
  vars <- vars[vars!="FREQ"]
  for(v in vars){
    plotResiduals(sim, form=new_data2[,v], sub=v, quantreg=T)
  }
```


Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1989
1993
1997
2001
2005
2009
2013
2017
catPred


Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH



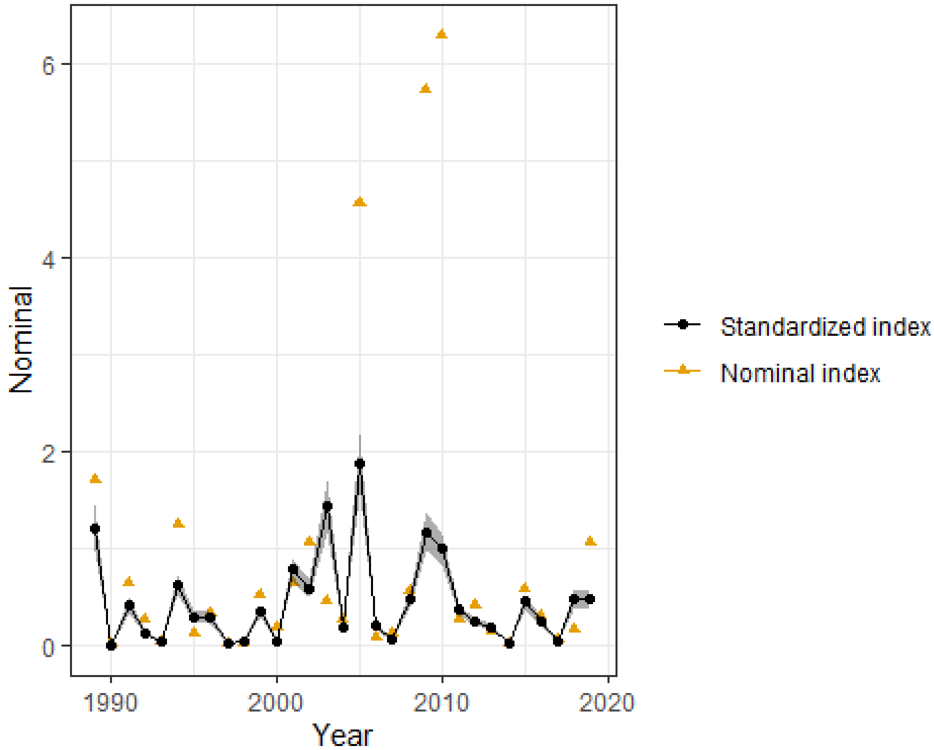


Step 4: Plot predictions

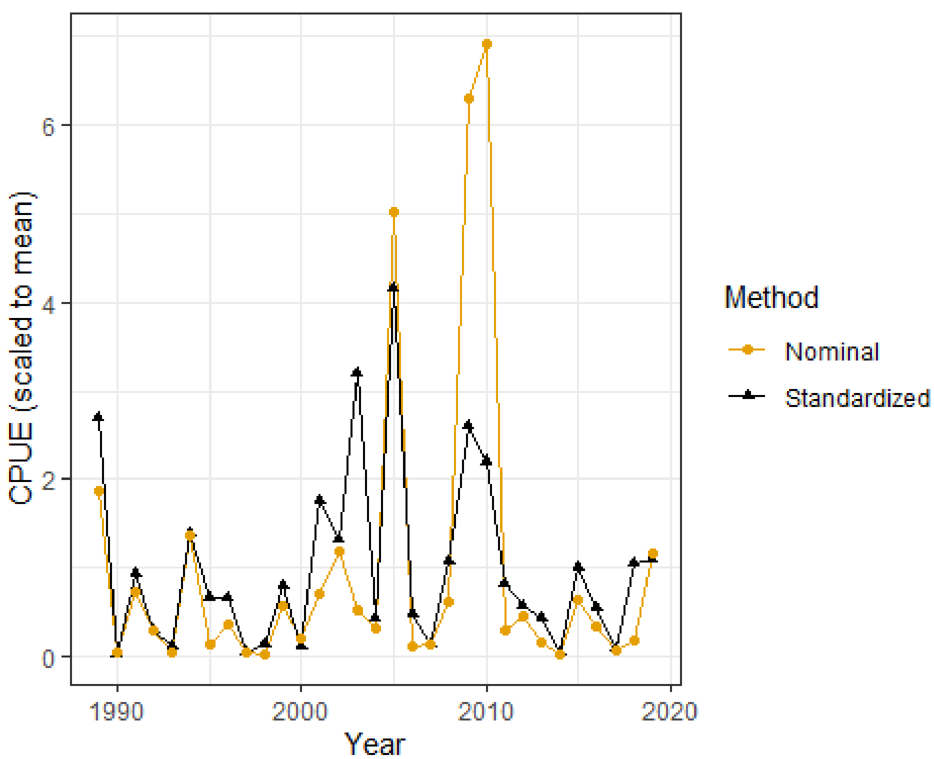
Including depth in the count model is marginally worse according to AIC (dAIC=0.2), but has better diagnostics. Final model:

FREQ ~YEAR+DEPTH + BOTTOM.TEMP+STRATUM, ziformula = ~YEAR + DEPTH + BOTTOM.TEMP + STRATUM

Plot standardized index from this work with index provided by the state.



Comparison of standardized index and nominal index scaled to their time-series means.



Bluefish index standardization - NC PSIGNS

Katie Drew, ASMFC

2022-02-11

Bluefish index standardization

Step 1: Data processing

Load functions

Read in data

```
##            SAMPLEID YEAR MONTH DAY TIME EFFORT FREQ          AREA STRATUM
DEPTH
## 1 101013DARE21WANC. 2010    10  13  545  12.00    4 Pamlico Sound   WANC1
1.1
## 2 101013DARE22WANC. 2010    10  13  600  12.00    5 Pamlico Sound   WANC2
2.1
## 3 101013HYDE21WASH. 2010    10  13  745  12.42    7 Pamlico Sound   WASH1
1.1
## 4 101013HYDE22WASH. 2010    10  13  705  12.08    7 Pamlico Sound   WASH2
1.8
## 5 101014HYDE11WASH. 2010    10  14  715  12.17    0 Pamlico Sound   WASH1
0.6
## 6 101014HYDE12WASH. 2010    10  14  640  11.92    0 Pamlico Sound   WASH2
1.9
##   WGT_FACTOR SURFACE.TEMP BOTTOM.TEMP SURFACE.SALINITY BOTTOM.SALINITY
## 1      134.5         24.0        24.0             18.2            18.2
## 2       70.5         23.9        24.0             17.1            17.1
## 3       82.5         23.6        23.6             22.6            22.6
## 4       54.5         24.0        23.9             23.0            23.0
## 5       82.5         22.6        22.6             17.9            17.9
## 6       54.5         22.5        22.6             17.9            17.9
##   SURFACE.DO BOTTOM.DO BOTTOM.COMPOSITION WEIGHT
## 1        8.0       8.0                      1.25
## 2        8.1       8.2                      3.60
## 3        7.6       7.6              shell   2.00
## 4        7.9       8.0              shell   2.00
## 5        7.6       7.6           no grass   0.00
## 6        7.6       7.6           no grass   0.00
```

Remove missing data

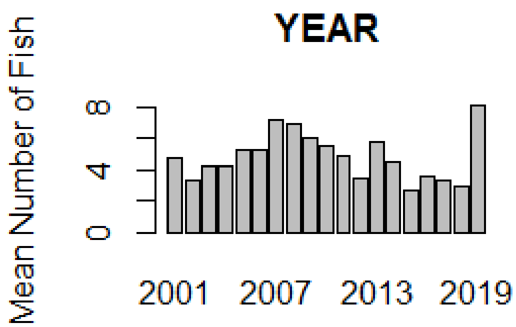
```
##    SAMPLEID              YEAR          MONTH           DAY
##  Length:5217         2002    : 280   5       : 532   Min.   : 1.00
##  Class :character    2003    : 280   6       : 532   1st Qu.: 9.00
```

```
##  Mode  :character   2004   : 280   11     : 532   Median :16.00
##                     2006   : 280   8      : 530   Mean   :16.27
##                     2007   : 280   10     : 529   3rd Qu.:23.00
##                     2008   : 280   7      : 520   Max.   :31.00
##                     (Other):3537   (Other):2042
##       TIME             EFFORT           FREQ                   AREA
##  Min.   : 310.0   Min.   : 8.67   Min.   :  0.000   Pamlico Sound:5217
##  1st Qu.: 545.0   1st Qu.:11.50   1st Qu.:  0.000
##  Median : 620.0   Median :11.92   Median :  1.000
##  Mean   : 628.9   Mean   :11.80   Mean   :  4.885
##  3rd Qu.: 700.0   3rd Qu.:12.17   3rd Qu.:  5.000
##  Max.   :2015.0   Max.   :17.75   Max.   :124.000
##  NA's   :1        NA's   :1
##   STRATUM         DEPTH        WGT_FACTOR      SURFACE.TEMP    BOTTOM.TEMP
##  WANC1:1113   Min.   :0.10   Min.   : 54.50   Min.   : 2.00   Min.   :
1.20
##  WANC2:1112   1st Qu.:0.80   1st Qu.: 54.50   1st Qu.:15.40   1st
Qu.:15.30
##  WASH1:1496   Median :1.60   Median : 82.50   Median :22.40   Median
:22.40
##  WASH2:1496   Mean   :1.42   Mean   : 83.01   Mean   :21.33   Mean
:21.27
##               3rd Qu.:1.90   3rd Qu.: 82.50   3rd Qu.:27.60   3rd
Qu.:27.50
##               Max.   :6.50   Max.   :134.50   Max.   :36.20   Max.
:36.20
##               NA's   :2                       NA's   :42      NA's   :45
##  SURFACE.SALINITY BOTTOM.SALINITY   SURFACE.DO       BOTTOM.DO
##  Min.   : 0.10    Min.   : 0.10   Min.   : 0.600   Min.   : 0.200
##  1st Qu.:15.30    1st Qu.:15.40   1st Qu.: 6.800   1st Qu.: 6.700
##  Median :19.10    Median :19.20   Median : 7.900   Median : 7.800
##  Mean   :18.83    Mean   :18.96   Mean   : 8.007   Mean   : 7.939
##  3rd Qu.:22.20    3rd Qu.:22.30   3rd Qu.: 9.200   3rd Qu.: 9.100
##  Max.   :35.70    Max.   :35.70   Max.   :15.600   Max.   :15.600
##  NA's   :46       NA's   :50      NA's   :89       NA's   :89
##  BOTTOM.COMPOSITION     WEIGHT
##  Length:5217        Min.   :  0.000
##  Class :character   1st Qu.:  0.000
##  Mode  :character   Median :  0.110
##                     Mean   :  3.073
##                     3rd Qu.:  2.350
##                     Max.   :152.400
##
```

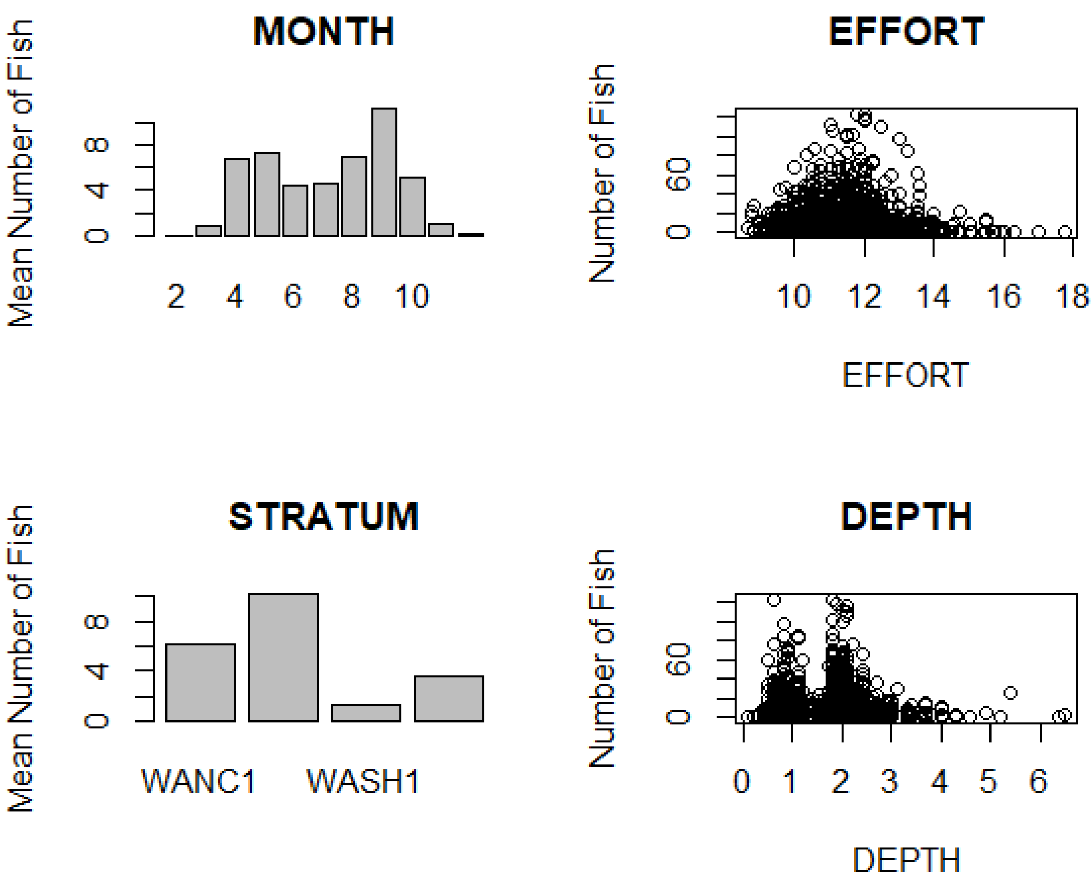
```
##    SAMPLEID              YEAR          MONTH           DAY
##  Length:5115        2016   : 280   8      : 526   Min.   : 1.00
##  Class :character   2017   : 280   6      : 521   1st Qu.: 9.00
##  Mode  :character   2010   : 278   7      : 517   Median :16.00
##                     2013   : 278   10     : 517   Mean   :16.25
##                     2004   : 277   5      : 514   3rd Qu.:23.00
```

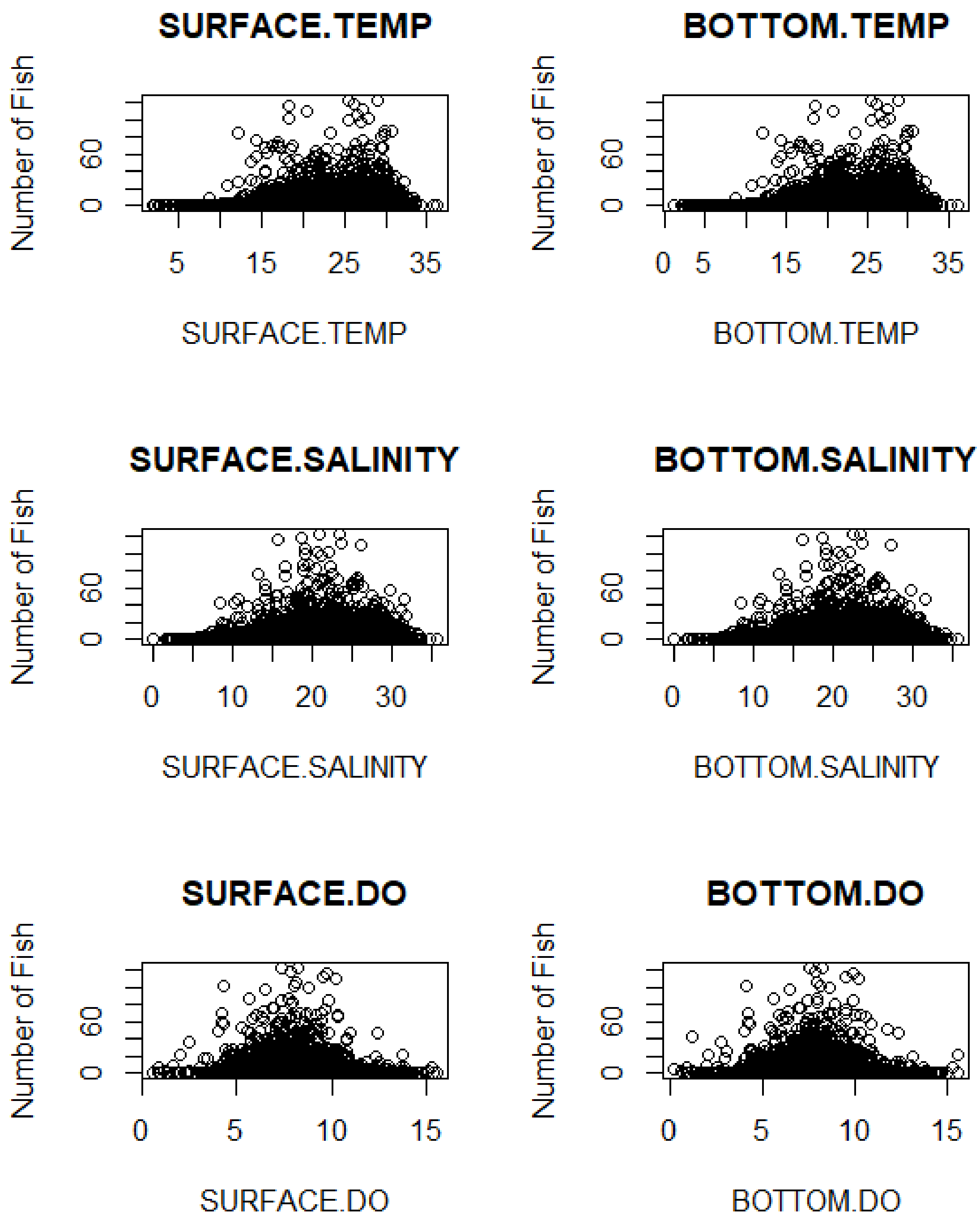
```
##                      2006   : 277   9      : 514   Max.   :31.00  
##                      (Other):3445   (Other):2006                  
##       TIME            EFFORT           FREQ                    AREA     
##  Min.   : 310.0   Min.   : 8.67   Min.   :  0.000   Pamlico Sound:5115  
##  1st Qu.: 545.0   1st Qu.:11.50   1st Qu.:  0.000                       
##  Median : 620.0   Median :11.92   Median :  1.000                       
##  Mean   : 628.5   Mean   :11.80   Mean   :  4.857                       
##  3rd Qu.: 700.0   3rd Qu.:12.17   3rd Qu.:  5.000                       
##  Max.   :2015.0   Max.   :17.75   Max.   :124.000                       
##                                                                         
##   STRATUM          DEPTH        WGT_FACTOR      SURFACE.TEMP    BOTTOM.TEMP 
##  WANC1:1093   Min.   :0.10   Min.   : 54.50   Min.   : 2.00   Min.   : 1.2  
##  WANC2:1095   1st Qu.:0.80   1st Qu.: 54.50   1st Qu.:15.40   1st Qu.:15.3  
##  WASH1:1464   Median :1.60   Median : 70.50   Median :22.50   Median :22.4  
##  WASH2:1463   Mean   :1.42   Mean   : 83.03   Mean   :21.36   Mean   :21.3  
##               3rd Qu.:1.90   3rd Qu.: 82.50   3rd Qu.:27.60   3rd Qu.:27.5  
##               Max.   :6.50   Max.   :134.50   Max.   :36.20   Max.   :36.2  
##                                                                             
##  SURFACE.SALINITY BOTTOM.SALINITY   SURFACE.DO       BOTTOM.DO     
##  Min.   : 0.10    Min.   : 0.10   Min.   : 0.600   Min.   : 0.200  
##  1st Qu.:15.30    1st Qu.:15.40   1st Qu.: 6.800   1st Qu.: 6.700  
##  Median :19.10    Median :19.20   Median : 7.900   Median : 7.800  
##  Mean   :18.82    Mean   :18.96   Mean   : 8.007   Mean   : 7.939  
##  3rd Qu.:22.20    3rd Qu.:22.30   3rd Qu.: 9.200   3rd Qu.: 9.100  
##  Max.   :35.70    Max.   :35.70   Max.   :15.600   Max.   :15.600  
##                                                                    
##  BOTTOM.COMPOSITION     WEIGHT           POS_TOW          REGION         
##  Length:5115        Min.   :  0.000   Min.   :0.0000   Length:5115       
##  Class :character   1st Qu.:  0.000   1st Qu.:0.0000   Class :character  
##  Mode  :character   Median :  0.110   Median :1.0000   Mode  :character  
##                     Mean   :  3.041   Mean   :0.5087                     
##                     3rd Qu.:  2.335   3rd Qu.:1.0000                     
##                     Max.   :152.400   Max.   :1.0000                     
##                                                                          
##    DEPTH.BIN   
##  Shallow:2557  
##  Deep   :2558  
##                
##                
##                
##                
## 
```

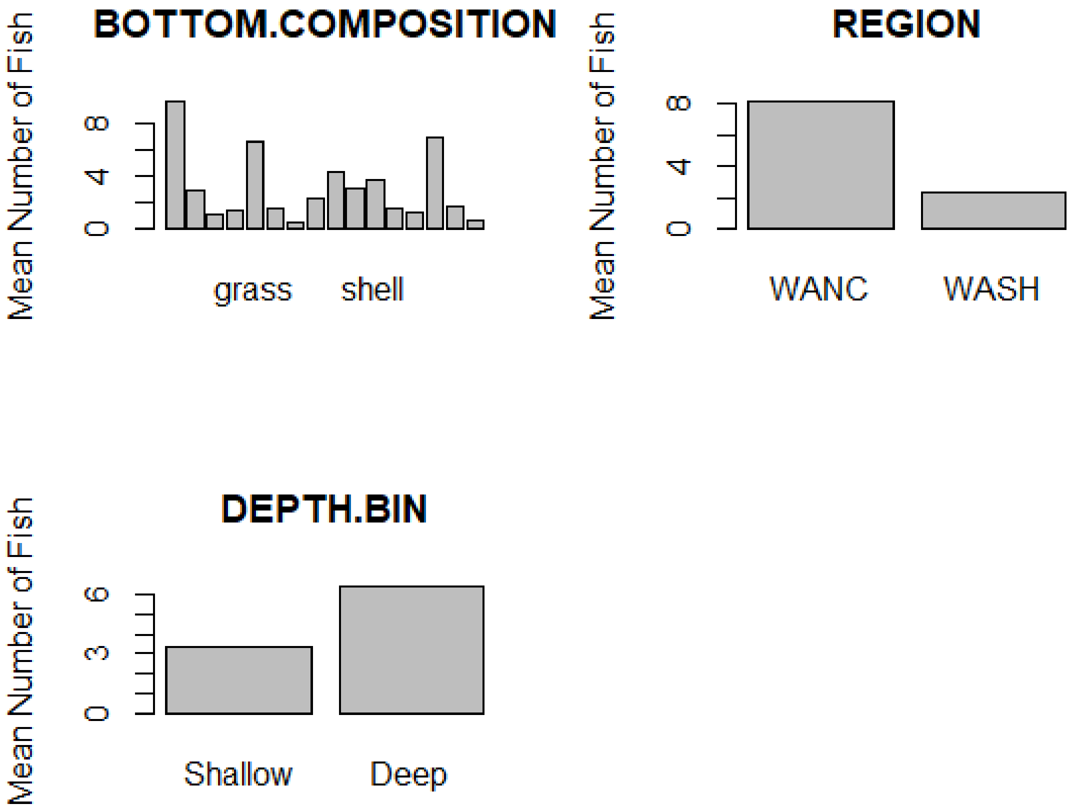
Plot data and decide if any data should be removed due to sparse sampling Mean CPUE by



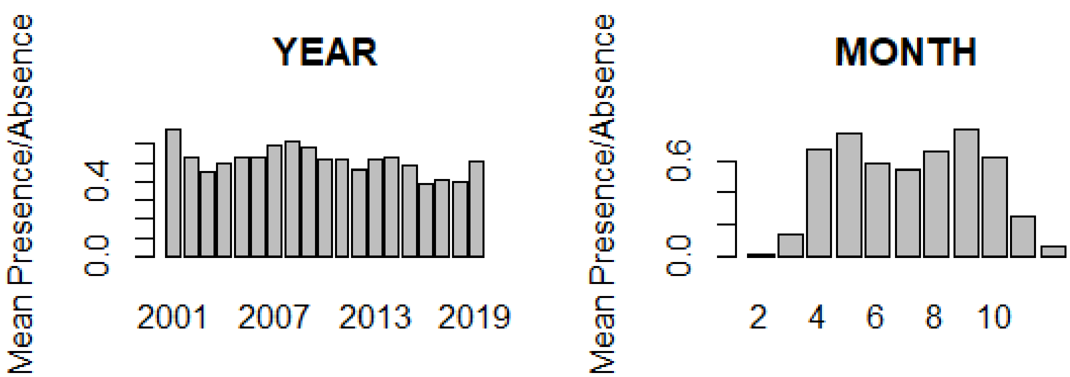
factor

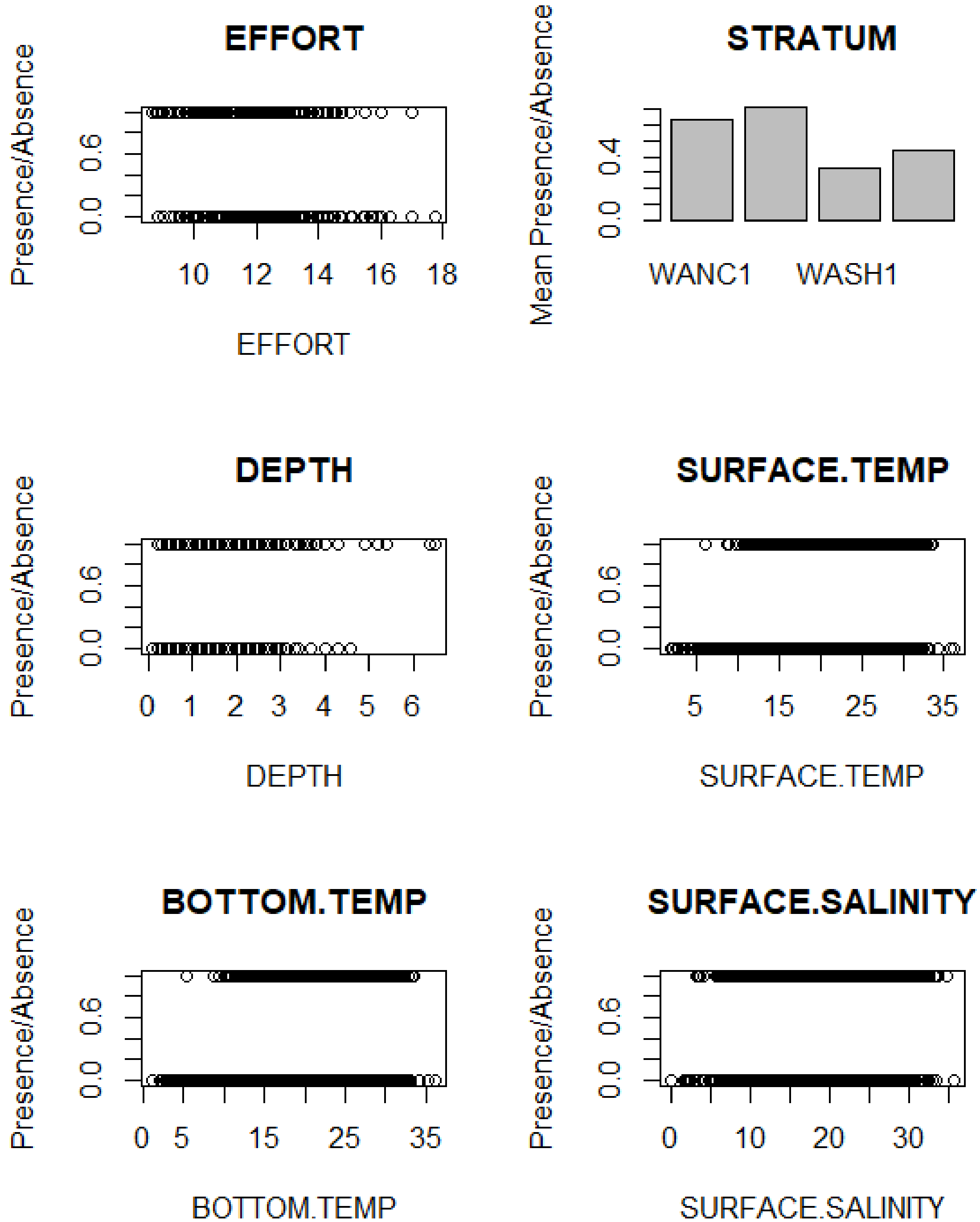



SURFACE.TEMP
BOTTOM.TEMP
SURFACE.SALINITY
BOTTOM.SALINITY
SURFACE.DO
BOTTOM.DO
Number of Fish
0
60
5
15
25
35
0
10
20
30
0
5
10
15

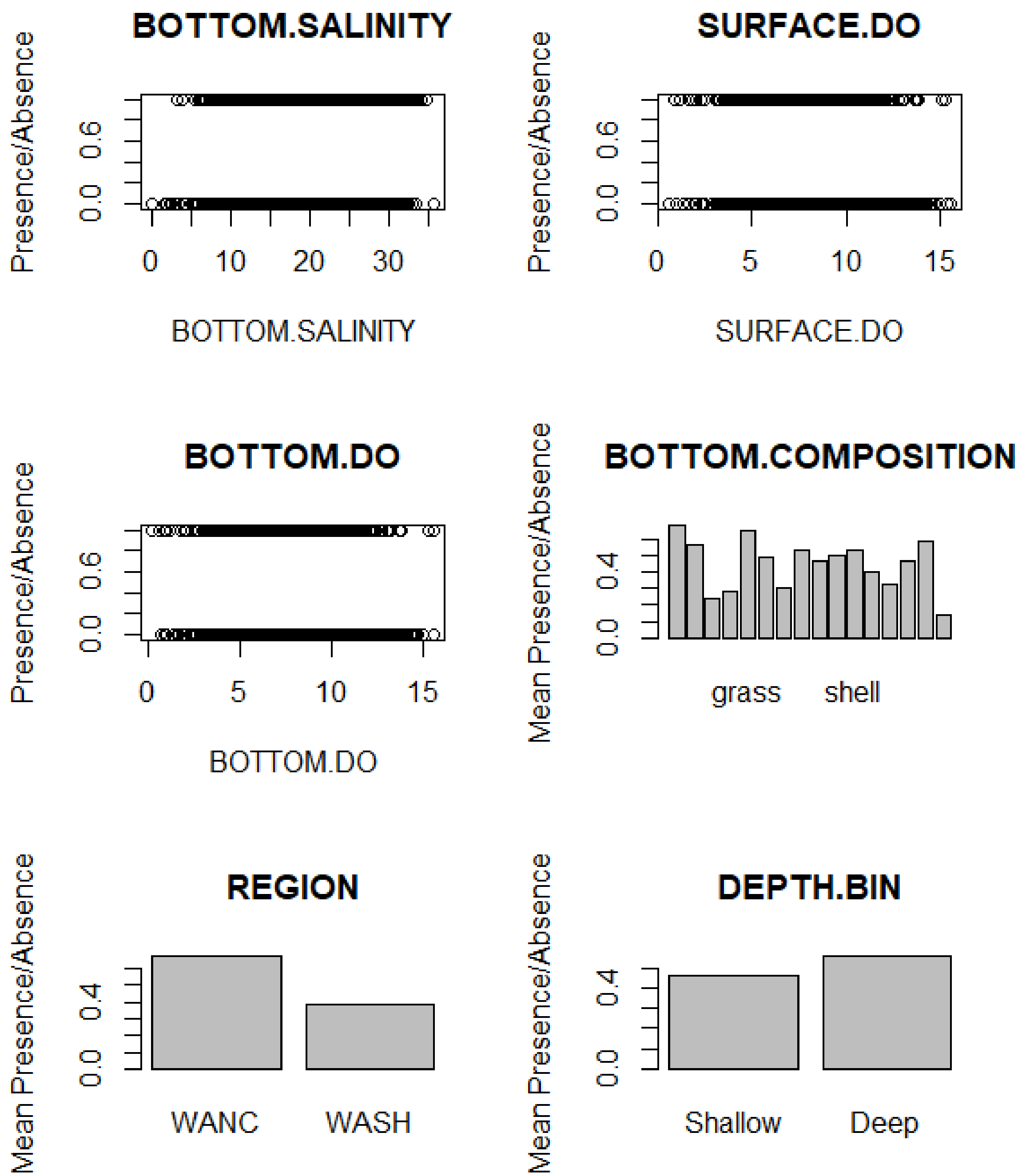


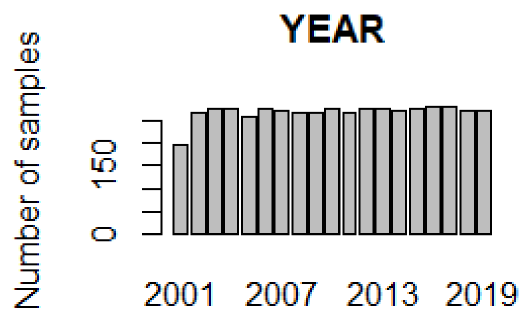
Proportion positive by factor



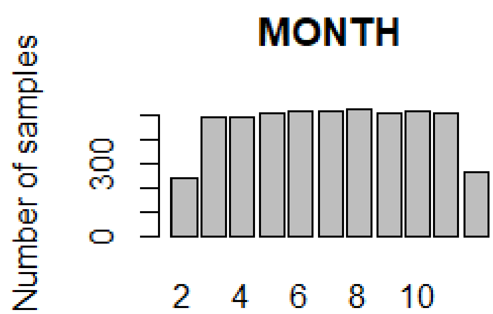


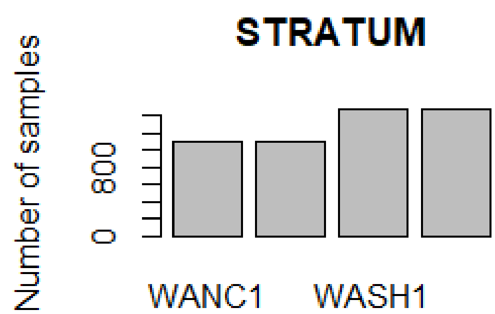
EFFORT
Presence/Absence
0.0 0.6
10 12 14 16 18
EFFORT
STRATUM
Mean Presence/Absence
0.0 0.4
WANC1 WASH1
DEPTH
Presence/Absence
0.0 0.6
0 1 2 3 4 5 6
DEPTH
SURFACE.TEMP
Presence/Absence
0.0 0.6
5 15 25 35
SURFACE.TEMP
BOTTOM.TEMP
Presence/Absence
0.0 0.6
0 5 15 25 35
BOTTOM.TEMP
SURFACE.SALINITY
Presence/Absence
0.0 0.6
0 10 20 30
SURFACE.SALINITY

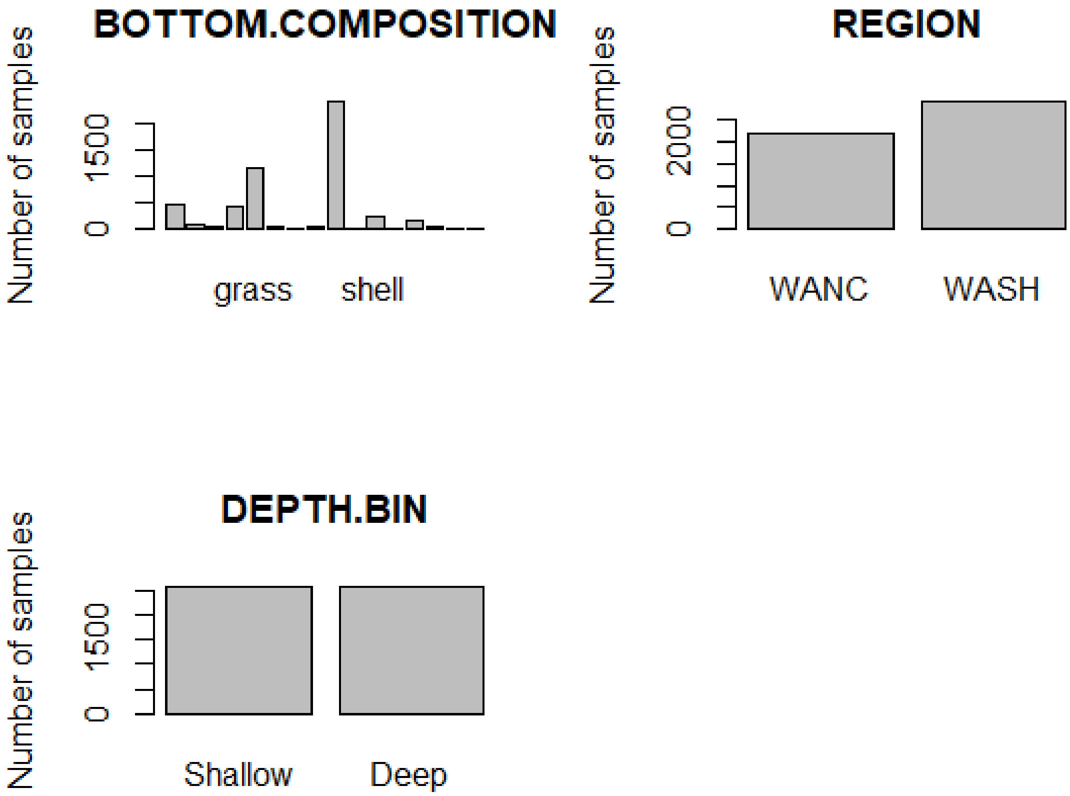

BOTTOM.SALINITY
Presence/Absence
0.0 0.6
0 10 20 30
BOTTOM.SALINITY
SURFACE.DO
Presence/Absence
0.0 0.6
0 5 10 15
SURFACE.DO
BOTTOM.DO
Presence/Absence
0.0 0.6
0 5 10 15
BOTTOM.DO
BOTTOM.COMPOSITION
Mean Presence/Absence
0.0 0.4
grass shell
REGION
Mean Presence/Absence
0.0 0.4
WANC WASH
DEPTH.BIN
Mean Presence/Absence
0.0 0.4
Shallow Deep

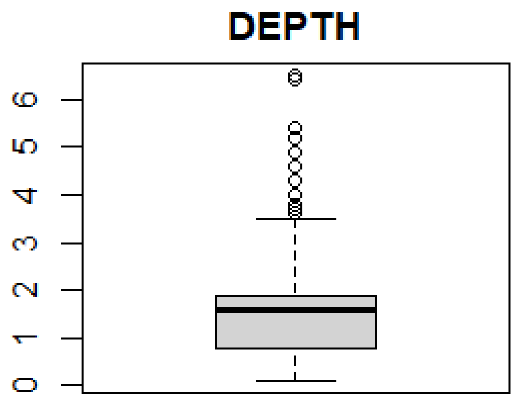


Number of samples by factor

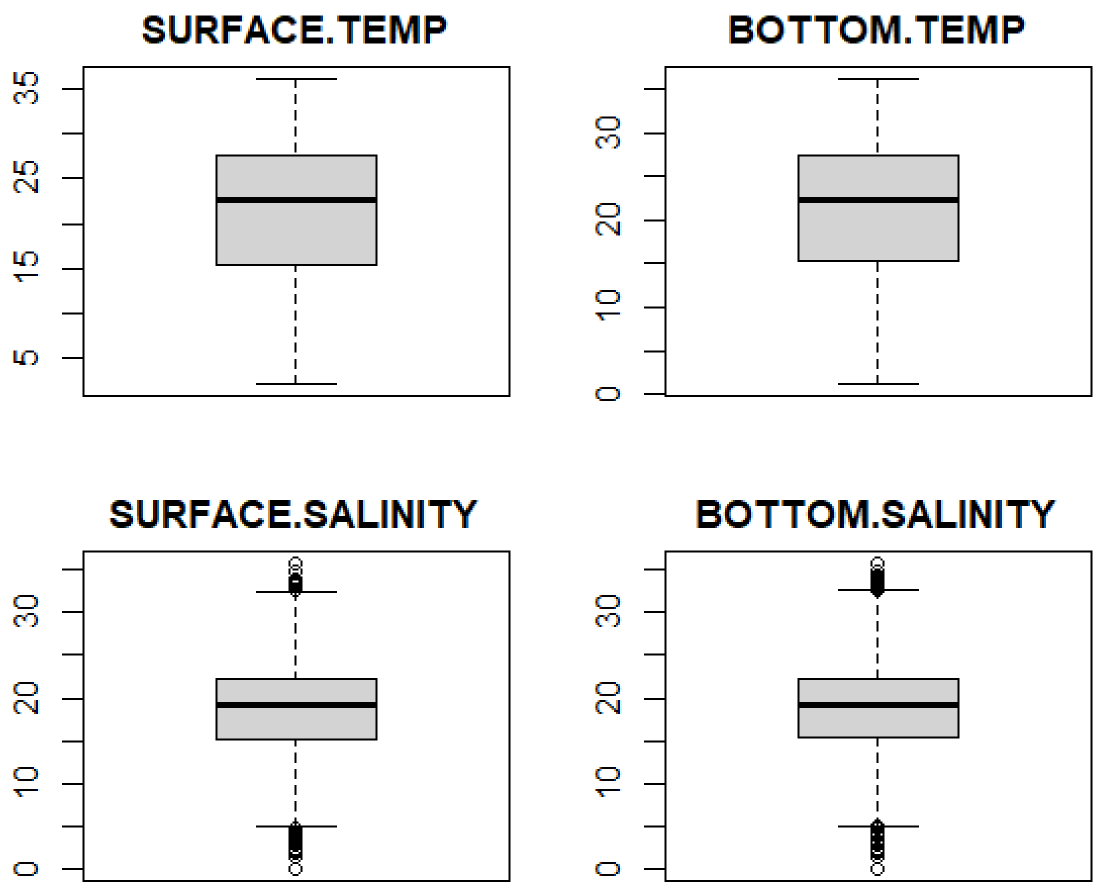


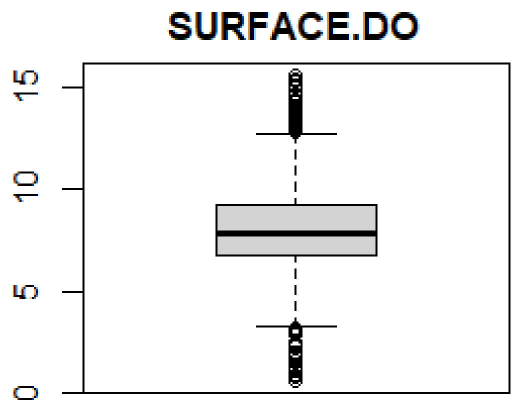


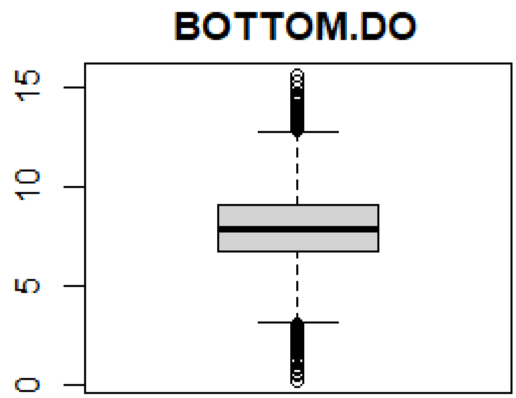

BOTTOM.COMPOSITION
Number of samples
0
1500
grass
shell
REGION
Number of samples
0
2000
WANC
WASH
DEPTH.BIN
Number of samples
0
1500
Shallow
Deep

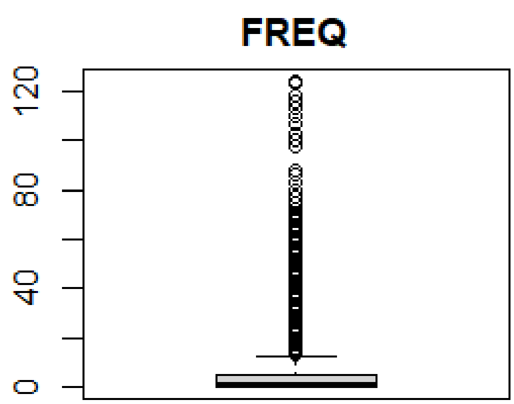


Check for outliers







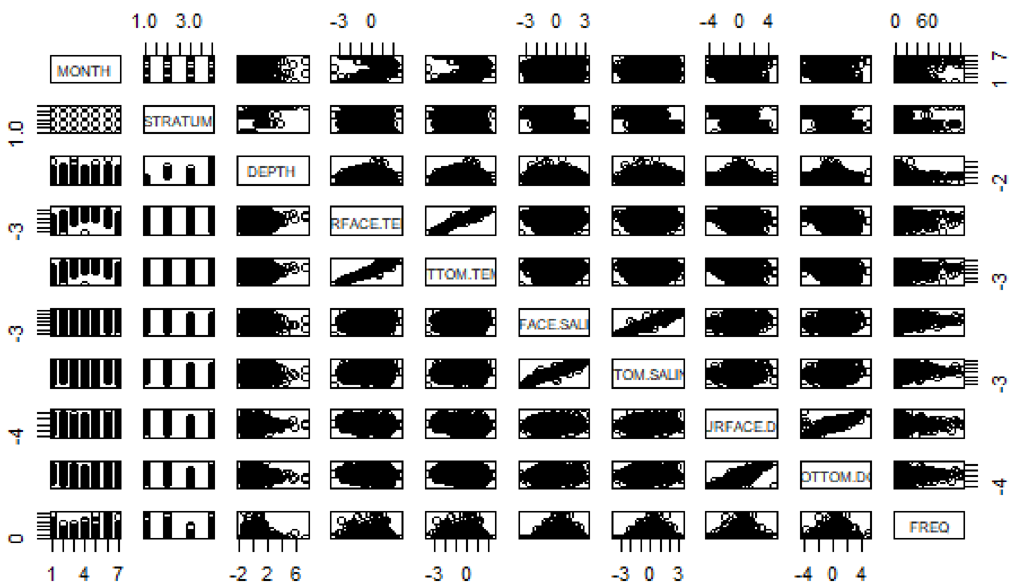


Change continuous variables to z-score

```
##            SAMPLEID YEAR MONTH DAY TIME EFFORT FREQ         AREA STRATUM
## 1 101013DARE21WANC. 2010    10  13  545  12.00    4 Pamlico Sound   WANC1
## 2 101013DARE22WANC. 2010    10  13  600  12.00    5 Pamlico Sound   WANC2
## 3 101013HYDE21WASH. 2010    10  13  745  12.42    7 Pamlico Sound   WASH1
## 4 101013HYDE22WASH. 2010    10  13  705  12.08    7 Pamlico Sound   WASH2
## 5 101014HYDE11WASH. 2010    10  14  715  12.17    0 Pamlico Sound   WASH1
## 6 101014HYDE12WASH. 2010    10  14  640  11.92    0 Pamlico Sound   WASH2
##        DEPTH WGT_FACTOR SURFACE.TEMP BOTTOM.TEMP SURFACE.SALINITY
## 1 -0.4675123      134.5   -0.2149786  -0.2022699       -0.1279402
## 2  0.9537219       70.5   -0.2371331  -0.2022699       -0.3346020
## 3 -0.4675123       82.5   -0.3035968  -0.2910966        0.6987069
## 4  0.5273517       54.5   -0.2149786  -0.2244766        0.7738566
## 5 -1.1781294       82.5   -0.5251425  -0.5131634       -0.1843025
## 6  0.6694751       54.5   -0.5472970  -0.5131634       -0.1843025
##   BOTTOM.SALINITY SURFACE.DO BOTTOM.DO BOTTOM.COMPOSITION WEIGHT POS_TOW
REGION
## 1      -0.1501408  0.3592611 0.3996155                      1.25       1
WANC
## 2      -0.3566919  0.4193269 0.5158796                      3.60       1
```

```
WANC
## 3        0.6760635  0.1189980 0.1670873              shell   2.00        1
WASH
## 4        0.7511730  0.2991954 0.3996155              shell   2.00        1
WASH
## 5       -0.2064729  0.1189980 0.1670873           no grass   0.00        0
WASH
## 6       -0.2064729  0.1189980 0.1670873           no grass   0.00        0
WASH
##    DEPTH.BIN lnEffort
## 1    Shallow 2.484907
## 2       Deep 2.484907
## 3    Shallow 2.519308
## 4       Deep 2.491551
## 5    Shallow 2.498974
## 6       Deep 2.478218
```

Check colinearity




```
##                   MONTH  STRATUM DEPTH SURFACE.TEMP BOTTOM.TEMP
SURFACE.SALINITY
## MONTH              1.00     0.00  0.01         0.23        0.23
0.16
## STRATUM            0.00     1.00  0.43         0.01        0.00            -
0.52
## DEPTH              0.01     0.43  1.00        -0.03       -0.04            -
0.03
## SURFACE.TEMP       0.23     0.01 -0.03         1.00        1.00
0.06
## BOTTOM.TEMP        0.23     0.00 -0.04         1.00        1.00
```

```
0.06
## SURFACE.SALINITY  0.16   -0.52 -0.03         0.06        0.06
1.00
## BOTTOM.SALINITY   0.16   -0.51 -0.01         0.06        0.06
0.99
## SURFACE.DO       -0.05   -0.31 -0.01        -0.30       -0.30
0.10
## BOTTOM.DO        -0.07   -0.36 -0.07        -0.31       -0.31
0.13
## FREQ              0.03   -0.21  0.13        -0.03       -0.03
0.20
##                  BOTTOM.SALINITY SURFACE.DO BOTTOM.DO  FREQ
## MONTH                       0.16      -0.05     -0.07  0.03
## STRATUM                    -0.51      -0.31     -0.36 -0.21
## DEPTH                      -0.01      -0.01     -0.07  0.13
## SURFACE.TEMP                0.06      -0.30     -0.31 -0.03
## BOTTOM.TEMP                 0.06      -0.30     -0.31 -0.03
## SURFACE.SALINITY            0.99       0.10      0.13  0.20
## BOTTOM.SALINITY             1.00       0.09      0.12  0.20
## SURFACE.DO                  0.09       1.00      0.94  0.10
## BOTTOM.DO                   0.12       0.94      1.00  0.12
## FREQ                        0.20       0.10      0.12  1.00
```

Check variance inflation factor - Remove covariates with GVIF > 3, starting with highest value, until all GVIF < 3

```
##                        GVIF Df GVIF^(1/(2*Df))
## YEAR               3.094222 18        1.031873
## MONTH              3.979892  6        1.121991
## STRATUM           11.063045  3        1.492723
## DEPTH              4.645561  1        2.155356
## SURFACE.TEMP     171.611514  1       13.100058
## BOTTOM.TEMP      172.856547  1       13.147492
## SURFACE.SALINITY  60.113913  1        7.753316
## BOTTOM.SALINITY   58.976267  1        7.679601
## SURFACE.DO        10.037739  1        3.168239
## BOTTOM.DO         10.092604  1        3.176886

##                        GVIF Df GVIF^(1/(2*Df))
## YEAR               2.675842 18        1.027718
## MONTH              3.954538  6        1.121393
## SURFACE.TEMP     171.181879  1       13.083649
## BOTTOM.TEMP      172.464500  1       13.132574
## SURFACE.SALINITY  59.961746  1        7.743497
## BOTTOM.SALINITY   58.854020  1        7.671637
## SURFACE.DO         9.753320  1        3.123031
## BOTTOM.DO          9.895053  1        3.145640
## REGION             2.436399  1        1.560897
## DEPTH.BIN          1.064029  1        1.031518
```

```
##                      GVIF Df GVIF^(1/(2*Df))
## YEAR             2.469037 18        1.025424
## MONTH            3.904840  6        1.120212
## REGION           2.414285  1        1.553797
## DEPTH.BIN        1.012202  1        1.006082
## BOTTOM.TEMP      3.775303  1        1.943014
## BOTTOM.SALINITY  3.172118  1        1.781044
## BOTTOM.DO        1.562779  1        1.250112

##                      GVIF Df GVIF^(1/(2*Df))
## YEAR             2.297494 18        1.023375
## REGION           2.342759  1        1.530607
## DEPTH.BIN        1.009297  1        1.004638
## BOTTOM.TEMP      1.159975  1        1.077021
## BOTTOM.SALINITY  2.982528  1        1.727000
## BOTTOM.DO        1.557756  1        1.248101

##                       GVIF Df GVIF^(1/(2*Df))
## YEAR              2.407363 18        1.024704
## REGION            2.295399  1        1.515057
## DEPTH.BIN         1.000844  1        1.000422
## SURFACE.TEMP      1.156068  1        1.075206
## SURFACE.SALINITY  3.031711  1        1.741181
## SURFACE.DO        1.541124  1        1.241420

##                      GVIF Df GVIF^(1/(2*Df))
## YEAR             2.353823 18        1.024064
## MONTH            1.199775  6        1.015294
## REGION           2.410226  1        1.552490
## DEPTH.BIN        1.005755  1        1.002873
## BOTTOM.SALINITY  3.172099  1        1.781039
## BOTTOM.DO        1.515231  1        1.230947

##                      GVIF Df GVIF^(1/(2*Df))
## YEAR             2.299084 18        1.023395
## STRATUM          2.373274  3        1.154936
## BOTTOM.TEMP      1.161086  1        1.077537
## BOTTOM.SALINITY  2.984017  1        1.727431
## BOTTOM.DO        1.562894  1        1.250158
```

Step 2: Model selection

Compare negative binom, zero-infl neg binom, zero-altered neg binom

Comparison 1: full models

```
##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO +
##     offset(lnEffort)
## Data: data
##
```

```
##       AIC       BIC   logLik deviance df.resid
##  18804.8  18965.7  -9376.4  18752.8     3577
##
##
## Dispersion parameter for nbinom2 family (): 0.507
##
## Conditional model:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -0.31579    0.13649  -2.314 0.020693 *
## YEAR2002        -0.53397    0.16927  -3.155 0.001607 **
## YEAR2003        -0.17302    0.17481  -0.990 0.322284
## YEAR2004         0.01127    0.16302   0.069 0.944900
## YEAR2005         0.10349    0.16656   0.621 0.534385
## YEAR2006        -0.09114    0.16141  -0.565 0.572291
## YEAR2007         0.20275    0.16086   1.260 0.207514
## YEAR2008         0.39250    0.16515   2.377 0.017471 *
## YEAR2009         0.10865    0.16244   0.669 0.503588
## YEAR2010         0.15696    0.16139   0.973 0.330762
## YEAR2011        -0.10999    0.16535  -0.665 0.505917
## YEAR2012        -0.38800    0.16371  -2.370 0.017788 *
## YEAR2013         0.42212    0.16336   2.584 0.009764 **
## YEAR2014         0.04031    0.16390   0.246 0.805712
## YEAR2015        -0.62578    0.16746  -3.737 0.000186 ***
## YEAR2016        -0.56259    0.16610  -3.387 0.000706 ***
## YEAR2017        -0.57610    0.16316  -3.531 0.000414 ***
## YEAR2018        -0.62727    0.16744  -3.746 0.000179 ***
## YEAR2019         0.44113    0.16018   2.754 0.005886 **
## STRATUMWANC2     0.56269    0.07456   7.546 4.48e-14 ***
## STRATUMWASH1    -1.59186    0.09459 -16.830  < 2e-16 ***
## STRATUMWASH2    -0.54100    0.09180  -5.893 3.78e-09 ***
## BOTTOM.TEMP     -0.14220    0.03005  -4.731 2.23e-06 ***
## BOTTOM.SALINITY  0.10266    0.04605   2.230 0.025780 *
## BOTTOM.DO        0.02577    0.03128   0.824 0.410072
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO +
##     offset(lnEffort)
## Zero inflation:
## ~YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO
## Data: data
##
##       AIC       BIC   logLik deviance df.resid
##  18531.0  18846.6  -9214.5  18429.0     3552
##
##
## Dispersion parameter for nbinom2 family (): 0.618
##
## Conditional model:
```

```
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -0.399581   0.124806  -3.202 0.001367 **
## YEAR2002        -0.289408   0.159117  -1.819 0.068937 .
## YEAR2003        -0.118416   0.165478  -0.716 0.474237
## YEAR2004         0.121797   0.154441   0.789 0.430325
## YEAR2005         0.239850   0.157249   1.525 0.127188
## YEAR2006         0.099271   0.150569   0.659 0.509697
## YEAR2007         0.464946   0.151352   3.072 0.002127 **
## YEAR2008         0.558474   0.152509   3.662 0.000250 ***
## YEAR2009         0.252874   0.150079   1.685 0.092000 .
## YEAR2010         0.332211   0.153058   2.170 0.029970 *
## YEAR2011         0.127423   0.156115   0.816 0.414377
## YEAR2012         0.004783   0.158818   0.030 0.975975
## YEAR2013         0.473624   0.150351   3.150 0.001632 **
## YEAR2014         0.103436   0.152659   0.678 0.498046
## YEAR2015        -0.557079   0.157172  -3.544 0.000394 ***
## YEAR2016        -0.077434   0.171058  -0.453 0.650782
## YEAR2017        -0.053553   0.167359  -0.320 0.748976
## YEAR2018        -0.502759   0.159981  -3.143 0.001674 **
## YEAR2019         0.627299   0.153657   4.082 4.46e-05 ***
## STRATUMWANC2     0.543335   0.068366   7.947 1.90e-15 ***
## STRATUMWASH1    -1.392683   0.094428 -14.749  < 2e-16 ***
## STRATUMWASH2    -0.521970   0.088638  -5.889 3.89e-09 ***
## BOTTOM.TEMP      0.029586   0.029434   1.005 0.314816
## BOTTOM.SALINITY -0.043248   0.045467  -0.951 0.341505
## BOTTOM.DO       -0.016148   0.030669  -0.527 0.598522
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                    Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -1.934e+01  1.217e+03  -0.016 0.987317
## YEAR2002         1.377e+01  1.217e+03   0.011 0.990972
## YEAR2003         1.107e+01  1.217e+03   0.009 0.992742
## YEAR2004         1.306e+01  1.217e+03   0.011 0.991438
## YEAR2005         1.226e+01  1.217e+03   0.010 0.991959
## YEAR2006         1.285e+01  1.217e+03   0.011 0.991571
## YEAR2007         1.336e+01  1.217e+03   0.011 0.991242
## YEAR2008         1.258e+01  1.217e+03   0.010 0.991749
## YEAR2009         1.258e+01  1.217e+03   0.010 0.991753
## YEAR2010         1.303e+01  1.217e+03   0.011 0.991457
## YEAR2011         1.384e+01  1.217e+03   0.011 0.990928
## YEAR2012         1.532e+01  1.217e+03   0.013 0.989955
## YEAR2013         1.203e+01  1.217e+03   0.010 0.992112
## YEAR2014         1.220e+01  1.217e+03   0.010 0.991999
## YEAR2015         1.204e+01  1.217e+03   0.010 0.992103
## YEAR2016         1.531e+01  1.217e+03   0.013 0.989958
## YEAR2017         1.667e+01  1.217e+03   0.014 0.989068
## YEAR2018         1.277e+01  1.217e+03   0.010 0.991626
## YEAR2019         1.339e+01  1.217e+03   0.011 0.991223
```

```
## STRATUMWANC2    -2.583e+01  3.811e+05   0.000 0.999946    
## STRATUMWASH1     3.154e+00  8.277e-01   3.810 0.000139 ***
## STRATUMWASH2     2.162e+00  8.277e-01   2.612 0.009010 ** 
## BOTTOM.TEMP      2.299e+00  3.716e-01   6.186 6.15e-10 ***
## BOTTOM.SALINITY -1.403e+00  2.554e-01  -5.491 4.00e-08 ***
## BOTTOM.DO       -3.834e-01  1.513e-01  -2.534 0.011277 *  
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: truncated_nbinom2  ( log )
## Formula:
## FREQ ~ YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO + 
##     offset(lnEffort)
## Zero inflation:
## ~YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO
## Data: data
## 
##      AIC      BIC   logLik deviance df.resid 
##  18601.5  18917.2  -9249.7  18499.5     3552 
## 
## 
## Dispersion parameter for truncated_nbinom2 family (): 0.521 
## 
## Conditional model:
##                  Estimate Std. Error z value Pr(>|z|)    
## (Intercept)     -0.532589   0.144519  -3.685 0.000228 ***
## YEAR2002        -0.259777   0.183707  -1.414 0.157339    
## YEAR2003         0.028237   0.191761   0.147 0.882934    
## YEAR2004         0.172229   0.174885   0.985 0.324717    
## YEAR2005         0.260576   0.177724   1.466 0.142599    
## YEAR2006         0.123075   0.171866   0.716 0.473922    
## YEAR2007         0.395374   0.169105   2.338 0.019385 *  
## YEAR2008         0.530831   0.173086   3.067 0.002163 ** 
## YEAR2009         0.360832   0.173797   2.076 0.037878 *  
## YEAR2010         0.352456   0.172772   2.040 0.041350 *  
## YEAR2011         0.143115   0.177436   0.807 0.419913    
## YEAR2012         0.007318   0.180853   0.040 0.967725    
## YEAR2013         0.583175   0.174555   3.341 0.000835 ***
## YEAR2014         0.163828   0.174391   0.939 0.347509    
## YEAR2015        -0.499778   0.181980  -2.746 0.006027 ** 
## YEAR2016         0.010812   0.194024   0.056 0.955560    
## YEAR2017        -0.033493   0.187037  -0.179 0.857880    
## YEAR2018        -0.263177   0.189254  -1.391 0.164346    
## YEAR2019         0.807580   0.175057   4.613 3.96e-06 ***
## STRATUMWANC2     0.474373   0.076123   6.232 4.61e-10 ***
## STRATUMWASH1    -1.400755   0.110221 -12.709  < 2e-16 ***
## STRATUMWASH2    -0.427372   0.102385  -4.174 2.99e-05 ***
## BOTTOM.TEMP     -0.097304   0.034446  -2.825 0.004731 ** 
## BOTTOM.SALINITY -0.020161   0.052307  -0.385 0.699919    
## BOTTOM.DO       -0.022316   0.034931  -0.639 0.522911    
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -1.91024    0.23268  -8.210  < 2e-16 ***
## YEAR2002         0.87704    0.26689   3.286 0.001016 **
## YEAR2003         0.52575    0.28188   1.865 0.062160 .
## YEAR2004         0.39946    0.26580   1.503 0.132866
## YEAR2005         0.21534    0.27371   0.787 0.431428
## YEAR2006         0.30629    0.26533   1.154 0.248354
## YEAR2007        -0.07835    0.27373  -0.286 0.774699
## YEAR2008        -0.15848    0.28703  -0.552 0.580851
## YEAR2009         0.50512    0.26532   1.904 0.056939 .
## YEAR2010         0.30931    0.26598   1.163 0.244863
## YEAR2011         0.50393    0.26515   1.901 0.057358 .
## YEAR2012         1.04192    0.25720   4.051 5.10e-05 ***
## YEAR2013         0.34361    0.26715   1.286 0.198374
## YEAR2014         0.31092    0.27047   1.150 0.250325
## YEAR2015         0.77052    0.26667   2.889 0.003860 **
## YEAR2016         1.43971    0.26436   5.446 5.15e-08 ***
## YEAR2017         1.43923    0.25670   5.607 2.06e-08 ***
## YEAR2018         1.19382    0.26554   4.496 6.93e-06 ***
## YEAR2019         0.55088    0.26429   2.084 0.037129 *
## STRATUMWANC2    -0.92990    0.16151  -5.758 8.53e-09 ***
## STRATUMWASH1     1.47366    0.14897   9.893  < 2e-16 ***
## STRATUMWASH2     0.81974    0.14887   5.506 3.66e-08 ***
## BOTTOM.TEMP      0.05332    0.04276   1.247 0.212449
## BOTTOM.SALINITY -0.25865    0.06985  -3.703 0.000213 ***
## BOTTOM.DO       -0.04916    0.05173  -0.950 0.341886
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##           dAIC  df
## zero_infl   0.0 51
## zero_adj   70.5 51
## neg_binom 273.8 26
```

Full model without an effort offset

```
##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  18806.6  18967.5  -9377.3  18754.6     3577
##
##
## Dispersion parameter for nbinom2 family (): 0.506
##
```

```
## Conditional model:
##                  Estimate Std. Error z value Pr(>|z|)    
## (Intercept)       2.09804    0.13590  15.438  < 2e-16 ***
## YEAR2002         -0.50903    0.16874  -3.017 0.002555 ** 
## YEAR2003         -0.15283    0.17446  -0.876 0.381011    
## YEAR2004          0.04172    0.16265   0.256 0.797569    
## YEAR2005          0.14121    0.16620   0.850 0.395550    
## YEAR2006         -0.04292    0.16100  -0.267 0.789779    
## YEAR2007          0.25076    0.16050   1.562 0.118203    
## YEAR2008          0.45677    0.16489   2.770 0.005604 ** 
## YEAR2009          0.15775    0.16215   0.973 0.330645    
## YEAR2010          0.18343    0.16111   1.139 0.254900    
## YEAR2011         -0.12113    0.16509  -0.734 0.463141    
## YEAR2012         -0.39353    0.16335  -2.409 0.015989 *  
## YEAR2013          0.44240    0.16310   2.712 0.006679 ** 
## YEAR2014          0.03528    0.16369   0.216 0.829362    
## YEAR2015         -0.63523    0.16718  -3.800 0.000145 ***
## YEAR2016         -0.55558    0.16573  -3.352 0.000801 ***
## YEAR2017         -0.55966    0.16273  -3.439 0.000583 ***
## YEAR2018         -0.57349    0.16695  -3.435 0.000592 ***
## YEAR2019          0.47530    0.15980   2.974 0.002937 ** 
## STRATUMWANC2      0.51727    0.07457   6.937 4.01e-12 ***
## STRATUMWASH1     -1.53250    0.09466 -16.189  < 2e-16 ***
## STRATUMWASH2     -0.51802    0.09192  -5.636 1.74e-08 ***
## BOTTOM.TEMP      -0.15563    0.03007  -5.175 2.27e-07 ***
## BOTTOM.SALINITY   0.10432    0.04618   2.259 0.023885 *  
## BOTTOM.DO         0.02711    0.03132   0.866 0.386727    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO
## Zero inflation:
## ~YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO
## Data: data
## 
##      AIC      BIC   logLik deviance df.resid 
##  18538.0  18853.7  -9218.0  18436.0     3552 
## 
## 
## Dispersion parameter for nbinom2 family (): 0.616 
## 
## Conditional model:
##                   Estimate Std. Error z value Pr(>|z|)    
## (Intercept)       2.021602   0.127559  15.848  < 2e-16 ***
## YEAR2002         -0.269406   0.161416  -1.669 0.095114 .  
## YEAR2003         -0.102806   0.167497  -0.614 0.539363    
## YEAR2004          0.146084   0.157044   0.930 0.352261    
## YEAR2005          0.268556   0.159606   1.683 0.092450 .  
## YEAR2006          0.139767   0.153253   0.912 0.361768    
```

```
## YEAR2007          0.499405   0.153772   3.248 0.001163 **
## YEAR2008          0.613652   0.155408   3.949 7.86e-05 ***
## YEAR2009          0.294692   0.152921   1.927 0.053969 .
## YEAR2010          0.351633   0.155866   2.256 0.024071 *
## YEAR2011          0.103729   0.158708   0.654 0.513381
## YEAR2012         -0.008593   0.161129  -0.053 0.957470
## YEAR2013          0.485444   0.153401   3.165 0.001553 **
## YEAR2014          0.087425   0.155433   0.562 0.573803
## YEAR2015         -0.577689   0.159705  -3.617 0.000298 ***
## YEAR2016         -0.076090   0.173177  -0.439 0.660388
## YEAR2017         -0.043768   0.169423  -0.258 0.796146
## YEAR2018         -0.453940   0.162462  -2.794 0.005204 **
## YEAR2019          0.655856   0.156744   4.184 2.86e-05 ***
## STRATUMWANC2      0.500941   0.068469   7.316 2.55e-13 ***
## STRATUMWASH1     -1.336412   0.094697 -14.112  < 2e-16 ***
## STRATUMWASH2     -0.501360   0.088956  -5.636 1.74e-08 ***
## BOTTOM.TEMP       0.013902   0.029584   0.470 0.638411
## BOTTOM.SALINITY -0.040288   0.045706  -0.881 0.378074
## BOTTOM.DO       -0.015069   0.030745  -0.490 0.624041
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -12.3025   113.9106  -0.108 0.913995
## YEAR2002          6.7574   113.9173   0.059 0.952699
## YEAR2003          4.0404   113.7993   0.036 0.971678
## YEAR2004          6.0267   113.8674   0.053 0.957790
## YEAR2005          5.2221   113.8609   0.046 0.963419
## YEAR2006          5.8278   113.8988   0.051 0.959193
## YEAR2007          6.3023   113.8975   0.055 0.955873
## YEAR2008          5.5311   113.9232   0.049 0.961277
## YEAR2009          5.5507   113.9085   0.049 0.961135
## YEAR2010          6.0000   113.8828   0.053 0.957982
## YEAR2011          6.7851   113.9347   0.060 0.952512
## YEAR2012          8.2876   113.9175   0.073 0.942004
## YEAR2013          5.0214   113.9033   0.044 0.964837
## YEAR2014          5.1480   113.8710   0.045 0.963940
## YEAR2015          4.9860   113.8652   0.044 0.965073
## YEAR2016          8.2970   113.8838   0.073 0.941922
## YEAR2017          9.6494   113.9075   0.085 0.932490
## YEAR2018          5.7889   113.8842   0.051 0.959460
## YEAR2019          6.3692   113.8994   0.056 0.955406
## STRATUMWANC2    -11.0711   265.8925  -0.042 0.966788
## STRATUMWASH1      3.1377     0.8191   3.830 0.000128 ***
## STRATUMWASH2      2.1434     0.8198   2.615 0.008934 **
## BOTTOM.TEMP       2.2937     0.3764   6.094  1.1e-09 ***
## BOTTOM.SALINITY  -1.4042     0.2610  -5.381  7.4e-08 ***
## BOTTOM.DO        -0.3778     0.1515  -2.494 0.012648 *
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: truncated_nbinom2  ( log )
## Formula:
## FREQ ~ YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO
## Zero inflation:
## ~YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  18602.8  18918.5  -9250.4  18500.8     3552
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.521
##
## Conditional model:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.892574   0.144059  13.137  < 2e-16 ***
## YEAR2002        -0.243779   0.183234  -1.330 0.183378
## YEAR2003         0.040333   0.191527   0.211 0.833209
## YEAR2004         0.190575   0.174557   1.092 0.274937
## YEAR2005         0.285612   0.177396   1.610 0.107393
## YEAR2006         0.156655   0.171434   0.914 0.360825
## YEAR2007         0.423482   0.168681   2.511 0.012054 *
## YEAR2008         0.577448   0.172845   3.341 0.000835 ***
## YEAR2009         0.395902   0.173554   2.281 0.022540 *
## YEAR2010         0.362559   0.172504   2.102 0.035576 *
## YEAR2011         0.109694   0.177110   0.619 0.535681
## YEAR2012        -0.013487   0.180511  -0.075 0.940441
## YEAR2013         0.590658   0.174402   3.387 0.000707 ***
## YEAR2014         0.142017   0.174181   0.815 0.414878
## YEAR2015        -0.525446   0.181612  -2.893 0.003813 **
## YEAR2016         0.004028   0.193715   0.021 0.983410
## YEAR2017        -0.030331   0.186711  -0.162 0.870952
## YEAR2018        -0.210476   0.188965  -1.114 0.265347
## YEAR2019         0.829717   0.174741   4.748 2.05e-06 ***
## STRATUMWANC2     0.430136   0.076141   5.649 1.61e-08 ***
## STRATUMWASH1    -1.339589   0.110339 -12.141  < 2e-16 ***
## STRATUMWASH2    -0.402296   0.102581  -3.922 8.79e-05 ***
## BOTTOM.TEMP     -0.114306   0.034479  -3.315 0.000916 ***
## BOTTOM.SALINITY -0.015284   0.052511  -0.291 0.770997
## BOTTOM.DO       -0.021045   0.034978  -0.602 0.547400
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -1.91032    0.23269  -8.210  < 2e-16 ***
## YEAR2002         0.87711    0.26689   3.286 0.001015 **
## YEAR2003         0.52584    0.28188   1.865 0.062115 .
## YEAR2004         0.39955    0.26580   1.503 0.132783
```

```
## YEAR2005         0.21541    0.27371   0.787 0.431281
## YEAR2006         0.30636    0.26534   1.155 0.248251
## YEAR2007        -0.07829    0.27373  -0.286 0.774866
## YEAR2008        -0.15838    0.28703  -0.552 0.581097
## YEAR2009         0.50518    0.26533   1.904 0.056910 .
## YEAR2010         0.30939    0.26598   1.163 0.244748
## YEAR2011         0.50399    0.26515   1.901 0.057330 .
## YEAR2012         1.04200    0.25720   4.051 5.09e-05 ***
## YEAR2013         0.34368    0.26715   1.286 0.198285
## YEAR2014         0.31099    0.27047   1.150 0.250218
## YEAR2015         0.77060    0.26667   2.890 0.003856 **
## YEAR2016         1.43982    0.26436   5.446 5.14e-08 ***
## YEAR2017         1.43929    0.25670   5.607 2.06e-08 ***
## YEAR2018         1.19390    0.26554   4.496 6.92e-06 ***
## YEAR2019         0.55095    0.26429   2.085 0.037107 *
## STRATUMWANC2    -0.92991    0.16151  -5.758 8.52e-09 ***
## STRATUMWASH1     1.47367    0.14897   9.893  < 2e-16 ***
## STRATUMWASH2     0.81974    0.14887   5.506 3.66e-08 ***
## BOTTOM.TEMP      0.05332    0.04276   1.247 0.212466
## BOTTOM.SALINITY -0.25864    0.06985  -3.703 0.000213 ***
## BOTTOM.DO       -0.04916    0.05173  -0.950 0.341907
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##           dAIC  df
## zero_infl   0.0 51
## zero_adj   64.8 51
## neg_binom 268.6 26
```

Full model with depth bin and region as separate factors

```
##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ YEAR + DEPTH.BIN + REGION + BOTTOM.TEMP + BOTTOM.SALINITY +
##     BOTTOM.DO + offset(lnEffort)
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  18825.8  18980.5  -9387.9  18775.8     3578
##
##
## Dispersion parameter for nbinom2 family (): 0.503
##
## Conditional model:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -0.46855    0.13181  -3.555 0.000378 ***
## YEAR2002        -0.50056    0.16884  -2.965 0.003030 **
## YEAR2003        -0.16634    0.17484  -0.951 0.341411
## YEAR2004         0.04062    0.16266   0.250 0.802811
## YEAR2005         0.13640    0.16635   0.820 0.412242
```

```
## YEAR2006        -0.06087    0.16124  -0.378 0.705768
## YEAR2007         0.23685    0.16069   1.474 0.140487
## YEAR2008         0.42859    0.16496   2.598 0.009372 **
## YEAR2009         0.12854    0.16229   0.792 0.428345
## YEAR2010         0.17809    0.16133   1.104 0.269632
## YEAR2011        -0.10111    0.16524  -0.612 0.540620
## YEAR2012        -0.36398    0.16343  -2.227 0.025941 *
## YEAR2013         0.43733    0.16323   2.679 0.007381 **
## YEAR2014         0.07079    0.16373   0.432 0.665464
## YEAR2015        -0.61242    0.16753  -3.656 0.000257 ***
## YEAR2016        -0.53999    0.16609  -3.251 0.001149 **
## YEAR2017        -0.57125    0.16315  -3.501 0.000463 ***
## YEAR2018        -0.61489    0.16762  -3.668 0.000244 ***
## YEAR2019         0.44577    0.16013   2.784 0.005374 **
## DEPTH.BINDeep    0.82583    0.05119  16.133  < 2e-16 ***
## REGIONWASH      -1.32868    0.07693 -17.271  < 2e-16 ***
## BOTTOM.TEMP     -0.13065    0.02999  -4.356 1.33e-05 ***
## BOTTOM.SALINITY  0.10983    0.04619   2.378 0.017423 *
## BOTTOM.DO        0.02911    0.03115   0.935 0.349980
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ YEAR + DEPTH.BIN + REGION + BOTTOM.TEMP + BOTTOM.SALINITY +
##     BOTTOM.DO + offset(lnEffort)
## Zero inflation:
## ~DEPTH.BIN + REGION + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  18621.4  18813.3  -9279.7  18559.4     3572
##
##
## Dispersion parameter for nbinom2 family (): 0.599
##
## Conditional model:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -0.437992   0.125395  -3.493 0.000478 ***
## YEAR2002        -0.347931   0.160702  -2.165 0.030383 *
## YEAR2003        -0.114494   0.169508  -0.675 0.499393
## YEAR2004         0.080507   0.157402   0.511 0.609019
## YEAR2005         0.219239   0.160934   1.362 0.173104
## YEAR2006         0.063098   0.154465   0.408 0.682912
## YEAR2007         0.410081   0.154579   2.653 0.007981 **
## YEAR2008         0.533190   0.156504   3.407 0.000657 ***
## YEAR2009         0.215431   0.154134   1.398 0.162207
## YEAR2010         0.282746   0.156647   1.805 0.071077 .
## YEAR2011         0.047932   0.157871   0.304 0.761422
## YEAR2012        -0.156812   0.157902  -0.993 0.320662
## YEAR2013         0.457217   0.154909   2.952 0.003162 **
```

```
## YEAR2014         0.091004   0.156549   0.581 0.561028    
## YEAR2015        -0.571714   0.161341  -3.544 0.000395 ***
## YEAR2016        -0.306934   0.164136  -1.870 0.061484 .  
## YEAR2017        -0.374217   0.158740  -2.357 0.018402 *  
## YEAR2018        -0.546729   0.162683  -3.361 0.000777 ***
## YEAR2019         0.552492   0.155938   3.543 0.000396 ***
## DEPTH.BINDeep    0.711830   0.051408  13.847  < 2e-16 ***
## REGIONWASH      -1.220174   0.074851 -16.301  < 2e-16 ***
## BOTTOM.TEMP      0.022045   0.029422   0.749 0.453691    
## BOTTOM.SALINITY -0.025272   0.045763  -0.552 0.580786    
## BOTTOM.DO       -0.001035   0.030855  -0.034 0.973239    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Zero-inflation model:
##                 Estimate Std. Error z value Pr(>|z|)    
## (Intercept)     -18.7231   944.4605  -0.020    0.984    
## DEPTH.BINDeep    -0.9496     0.2290  -4.146 3.38e-05 ***
## REGIONWASH       16.0487   944.4610   0.017    0.986    
## BOTTOM.TEMP       2.2384     0.3795   5.898 3.67e-09 ***
## BOTTOM.SALINITY  -0.9228     0.1561  -5.911 3.40e-09 ***
## BOTTOM.DO        -0.1709     0.1155  -1.479    0.139    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: truncated_nbinom2  ( log )
## Formula:          
## FREQ ~ YEAR + DEPTH.BIN + REGION + BOTTOM.TEMP + BOTTOM.SALINITY +  
##     BOTTOM.DO + offset(lnEffort)
## Zero inflation:        
## ~YEAR + DEPTH.BIN + REGION + BOTTOM.TEMP + BOTTOM.SALINITY +     BOTTOM.DO
## Data: data
## 
##      AIC      BIC   logLik deviance df.resid 
##  18618.0  18921.3  -9260.0  18520.0     3554 
## 
## 
## Dispersion parameter for truncated_nbinom2 family (): 0.512 
## 
## Conditional model:
##                 Estimate Std. Error z value Pr(>|z|)    
## (Intercept)     -0.68016    0.14057  -4.839 1.31e-06 ***
## YEAR2002        -0.21354    0.18369  -1.163 0.245028    
## YEAR2003         0.03515    0.19236   0.183 0.855016    
## YEAR2004         0.20808    0.17501   1.189 0.234444    
## YEAR2005         0.30058    0.17795   1.689 0.091190 .  
## YEAR2006         0.16366    0.17201   0.951 0.341362    
## YEAR2007         0.44149    0.16924   2.609 0.009089 ** 
## YEAR2008         0.58001    0.17321   3.349 0.000812 ***
## YEAR2009         0.38738    0.17406   2.226 0.026041 *  
## YEAR2010         0.38160    0.17315   2.204 0.027531 *  
```

```
## YEAR2011         0.15892    0.17777   0.894 0.371352
## YEAR2012         0.03950    0.18097   0.218 0.827221
## YEAR2013         0.60231    0.17489   3.444 0.000573 ***
## YEAR2014         0.20002    0.17464   1.145 0.252091
## YEAR2015        -0.48483    0.18250  -2.657 0.007893 **
## YEAR2016         0.04041    0.19459   0.208 0.835483
## YEAR2017        -0.01618    0.18760  -0.086 0.931250
## YEAR2018        -0.24574    0.18992  -1.294 0.195688
## YEAR2019         0.81524    0.17550   4.645 3.40e-06 ***
## DEPTH.BINDeep    0.68699    0.05873  11.699  < 2e-16 ***
## REGIONWASH      -1.10704    0.08697 -12.729  < 2e-16 ***
## BOTTOM.TEMP     -0.08070    0.03434  -2.350 0.018788 *
## BOTTOM.SALINITY -0.01487    0.05268  -0.282 0.777663
## BOTTOM.DO       -0.01972    0.03493  -0.565 0.572314
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -1.99137    0.22721  -8.764  < 2e-16 ***
## YEAR2002         0.88034    0.26732   3.293 0.000991 ***
## YEAR2003         0.53281    0.28217   1.888 0.058988 .
## YEAR2004         0.40343    0.26626   1.515 0.129722
## YEAR2005         0.21912    0.27420   0.799 0.424206
## YEAR2006         0.30909    0.26587   1.163 0.245017
## YEAR2007        -0.07729    0.27429  -0.282 0.778108
## YEAR2008        -0.15985    0.28761  -0.556 0.578352
## YEAR2009         0.50713    0.26584   1.908 0.056437 .
## YEAR2010         0.31201    0.26646   1.171 0.241624
## YEAR2011         0.50553    0.26568   1.903 0.057076 .
## YEAR2012         1.04498    0.25755   4.057 4.96e-05 ***
## YEAR2013         0.34836    0.26760   1.302 0.192997
## YEAR2014         0.31687    0.27086   1.170 0.242050
## YEAR2015         0.77812    0.26687   2.916 0.003548 **
## YEAR2016         1.44286    0.26431   5.459 4.79e-08 ***
## YEAR2017         1.44180    0.25682   5.614 1.98e-08 ***
## YEAR2018         1.20064    0.26553   4.522 6.14e-06 ***
## YEAR2019         0.55341    0.26472   2.091 0.036567 *
## DEPTH.BINDeep   -0.72278    0.07978  -9.059  < 2e-16 ***
## REGIONWASH       1.58725    0.12887  12.317  < 2e-16 ***
## BOTTOM.TEMP      0.05492    0.04280   1.283 0.199361
## BOTTOM.SALINITY -0.25694    0.06983  -3.680 0.000234 ***
## BOTTOM.DO       -0.04617    0.05169  -0.893 0.371765
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

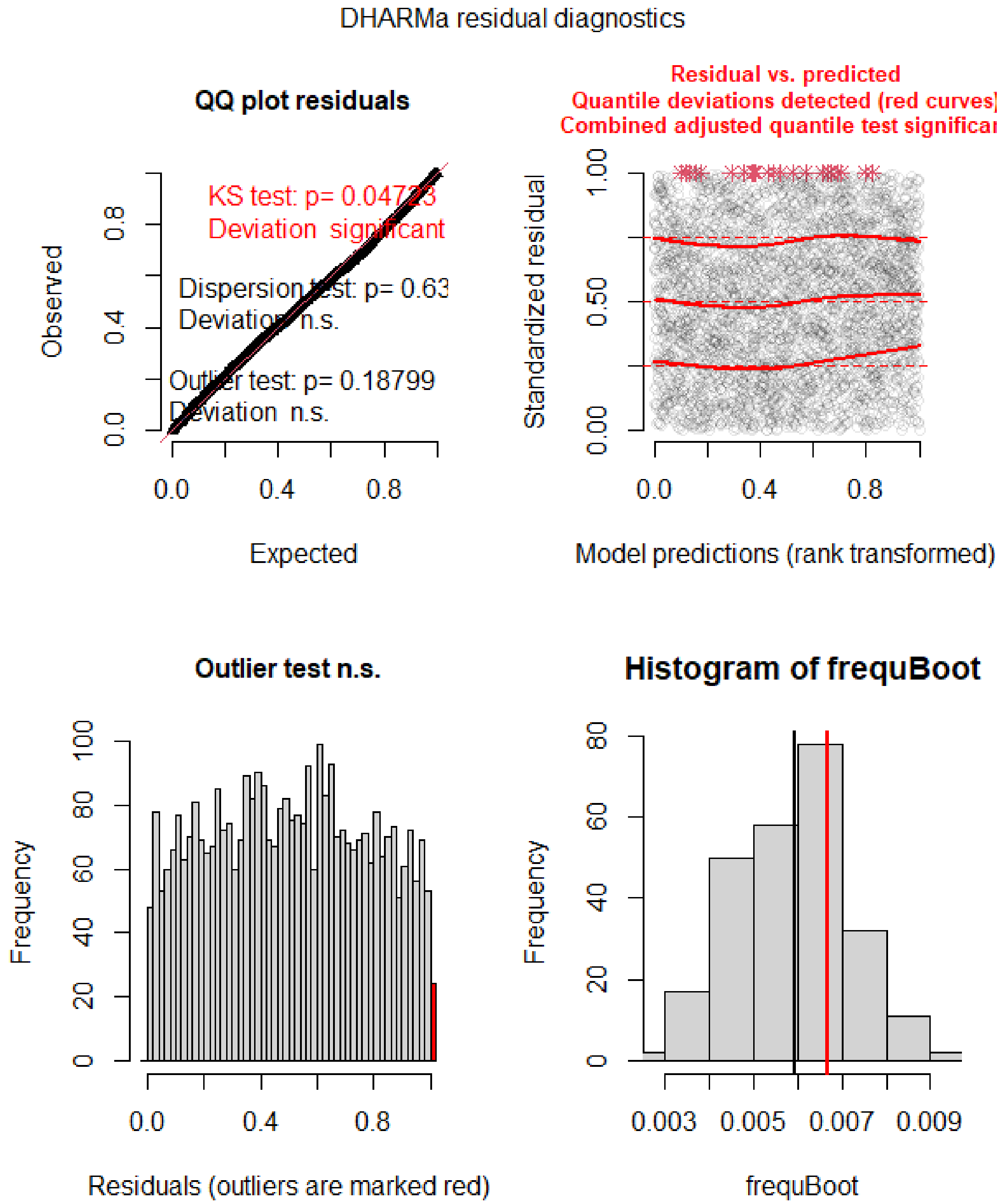
```
##           dAIC  df
## zero_adj    0.0 49
## zero_infl   3.4 31
## neg_binom 207.7 25
```

Comparison of models

```
##            dAIC  df
## zero_infl   0.0 51
## zero_adj   70.5 51
## neg_binom 273.8 26

##            dAIC  df
## zero_adj    0.0 49
## zero_infl   3.4 31
## neg_binom 207.7 25

##            dAIC  df
## zero_infl   0.0 51
## zero_adj   64.8 51
## neg_binom 268.6 26

##                         dAIC df
## full_mod$zero_infl       0.0 51
## full_mod_noE$zero_infl  7.0 51
## full_mod_DR$zero_adj   87.1 49
```

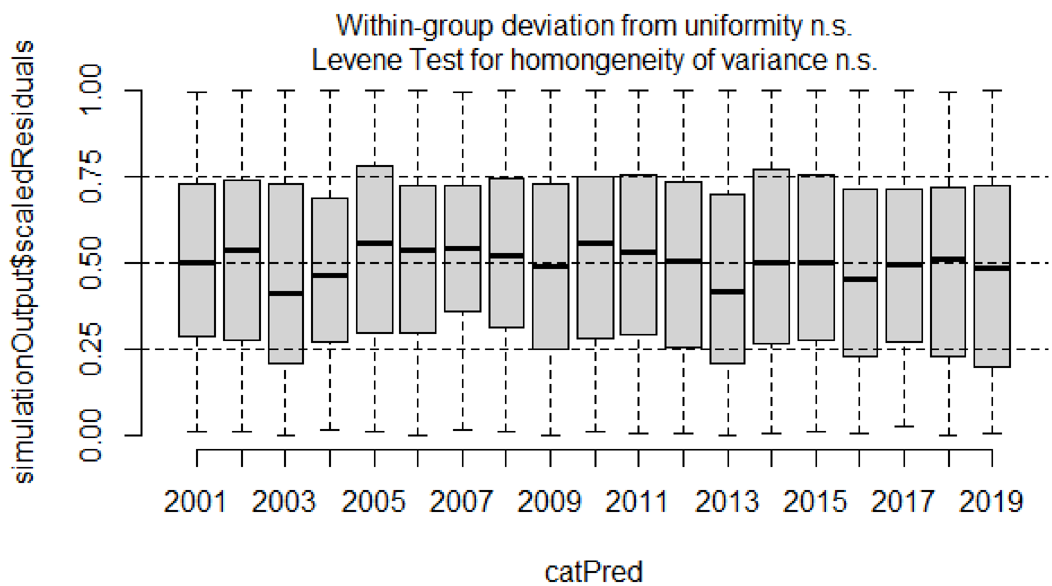
DHARMA simulations

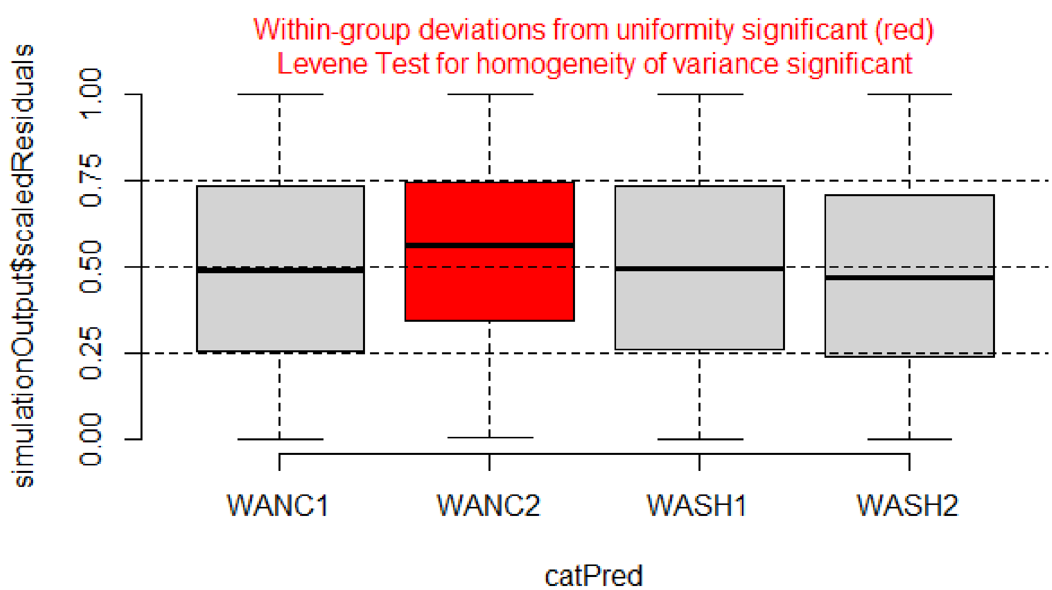
Full model, zero-inflated

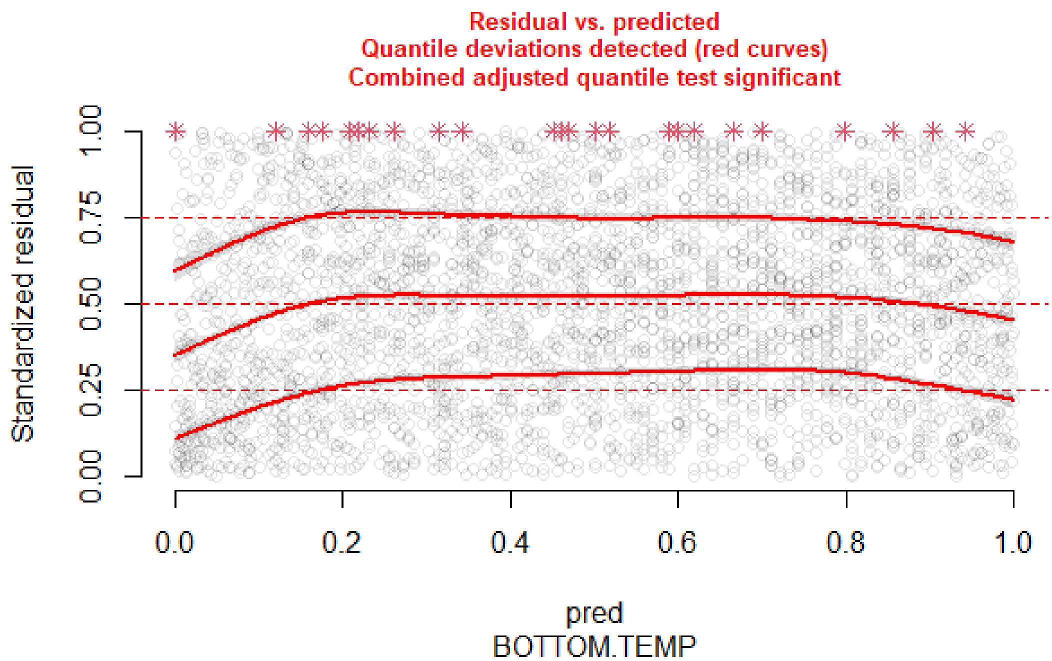


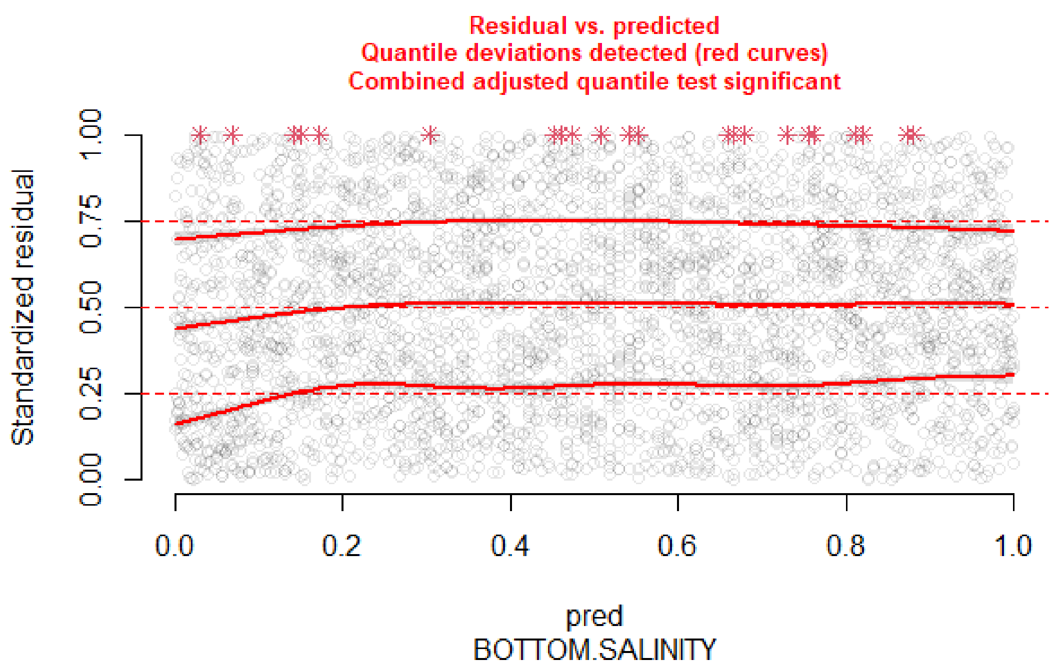

```
##
##  DHARMa bootstrapped outlier test
##
## data:  sim
## outliers at both margin(s) = 24, observations = 3603, p-value = 0.672
## alternative hypothesis: two.sided
```

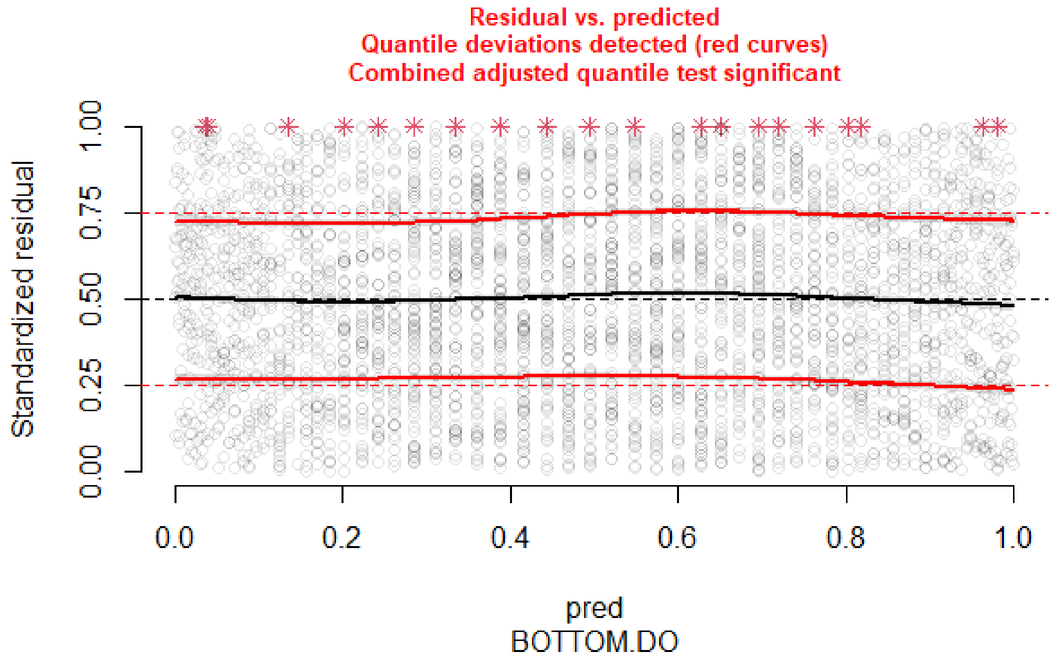
```
##  percent confidence interval:
##  0.003393006 0.008326395
## sample estimates:
## outlier frequency (expected: 0.00592950319178462 )
##                                         0.006661116
```






Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

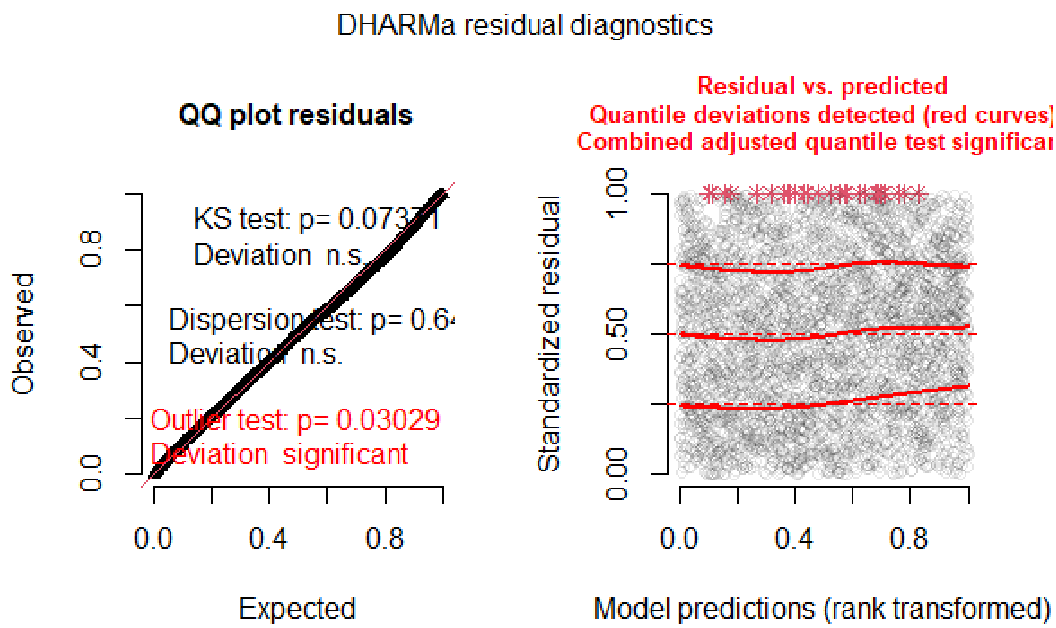

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.SALINITY

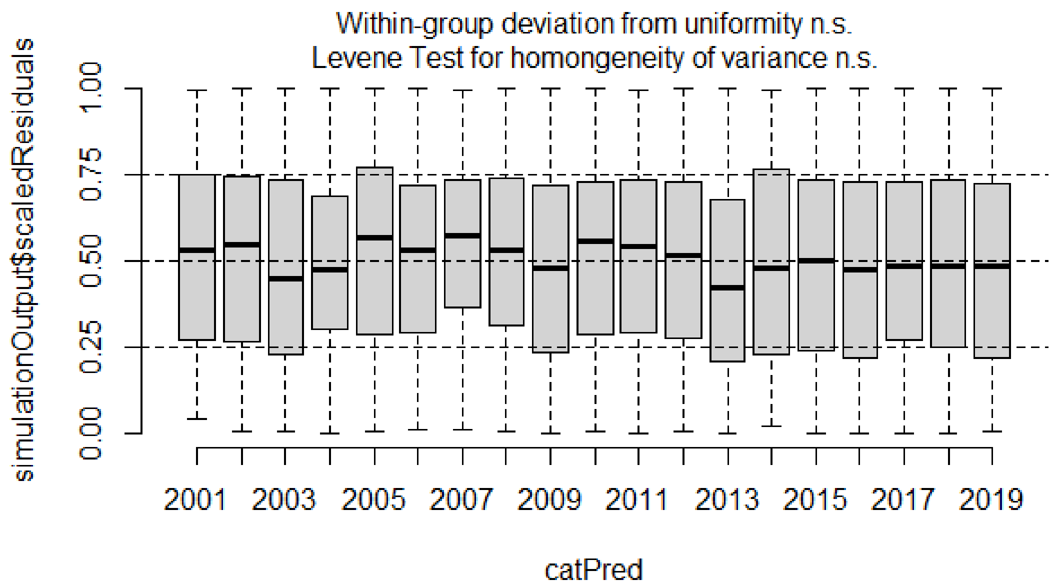


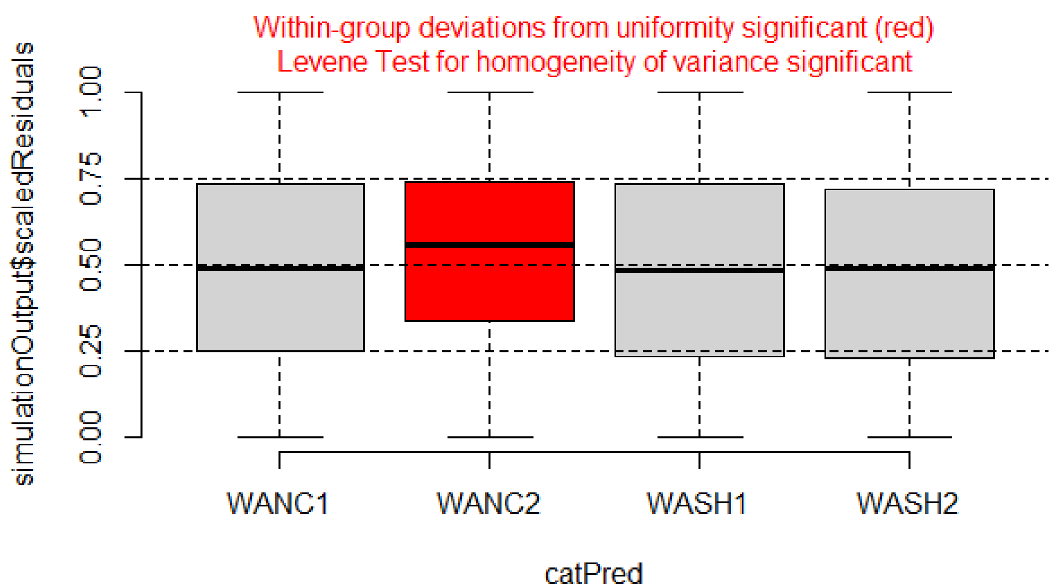

```
##     SAMPLEID              YEAR           MONTH           DAY
##  Length:5217        2002    : 280   5       : 532   Min.   : 1.00
##  Class :character   2003    : 280   6       : 532   1st Qu.: 9.00
##  Mode  :character   2004    : 280   11      : 532   Median :16.00
##                     2006    : 280   8       : 530   Mean   :16.27
##                     2007    : 280   10      : 529   3rd Qu.:23.00
##                     2008    : 280   7       : 520   Max.   :31.00
##                     (Other):3537   (Other):2042
##       TIME            EFFORT          FREQ                   AREA
##  Min.   : 310.0   Min.   : 8.67   Min.   :  0.000   Pamlico Sound:5217
##  1st Qu.: 545.0   1st Qu.:11.50   1st Qu.:  0.000
##  Median : 620.0   Median :11.92   Median :  1.000
##  Mean   : 628.9   Mean   :11.80   Mean   :  4.885
##  3rd Qu.: 700.0   3rd Qu.:12.17   3rd Qu.:  5.000
##  Max.   :2015.0   Max.   :17.75   Max.   :124.000
##  NA's   :1        NA's   :1
##   STRATUM         DEPTH         WGT_FACTOR       SURFACE.TEMP     BOTTOM.TEMP
##  WANC1:1113   Min.   :0.10   Min.   : 54.50   Min.   : 2.00   Min.   :
1.20
##  WANC2:1112   1st Qu.:0.80   1st Qu.: 54.50   1st Qu.:15.40   1st
Qu.:15.30
##  WASH1:1496   Median :1.60   Median : 82.50   Median :22.40   Median
:22.40
##  WASH2:1496   Mean   :1.42   Mean   : 83.01   Mean   :21.33   Mean
:21.27
##               3rd Qu.:1.90   3rd Qu.: 82.50   3rd Qu.:27.60   3rd
Qu.:27.50
##               Max.   :6.50   Max.   :134.50   Max.   :36.20   Max.
:36.20
```

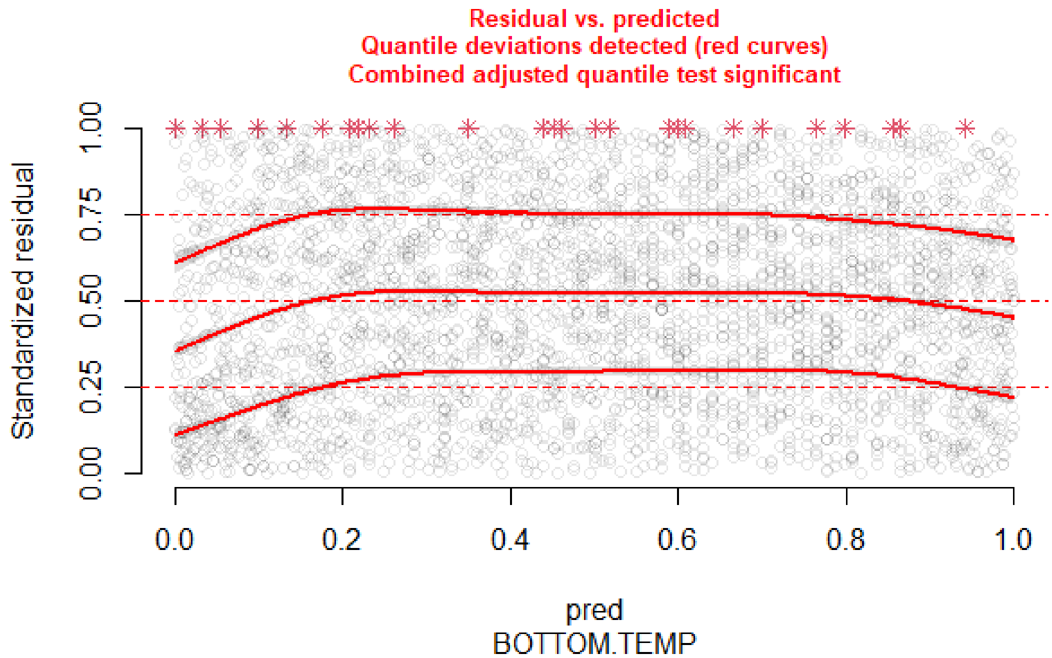
```
##                 NA's    :2                             NA's    :42       NA's    :45
##  SURFACE.SALINITY BOTTOM.SALINITY   SURFACE.DO       BOTTOM.DO
##  Min.   : 0.10    Min.   : 0.10    Min.    : 0.600   Min.    : 0.200
##  1st Qu.:15.30    1st Qu.:15.40    1st Qu.: 6.800    1st Qu.: 6.700
##  Median :19.10    Median :19.20    Median : 7.900    Median : 7.800
##  Mean   :18.83    Mean   :18.96    Mean    : 8.007   Mean    : 7.939
##  3rd Qu.:22.20    3rd Qu.:22.30    3rd Qu.: 9.200    3rd Qu.: 9.100
##  Max.   :35.70    Max.   :35.70    Max.    :15.600   Max.    :15.600
##  NA's   :46       NA's   :50       NA's    :89       NA's    :89
##  BOTTOM.COMPOSITION     WEIGHT
##  Length:5217         Min.   :  0.000
##  Class :character    1st Qu.:  0.000
##  Mode  :character    Median :  0.110
##                      Mean   :  3.073
##                      3rd Qu.:  2.350
##                      Max.   :152.400
##
```

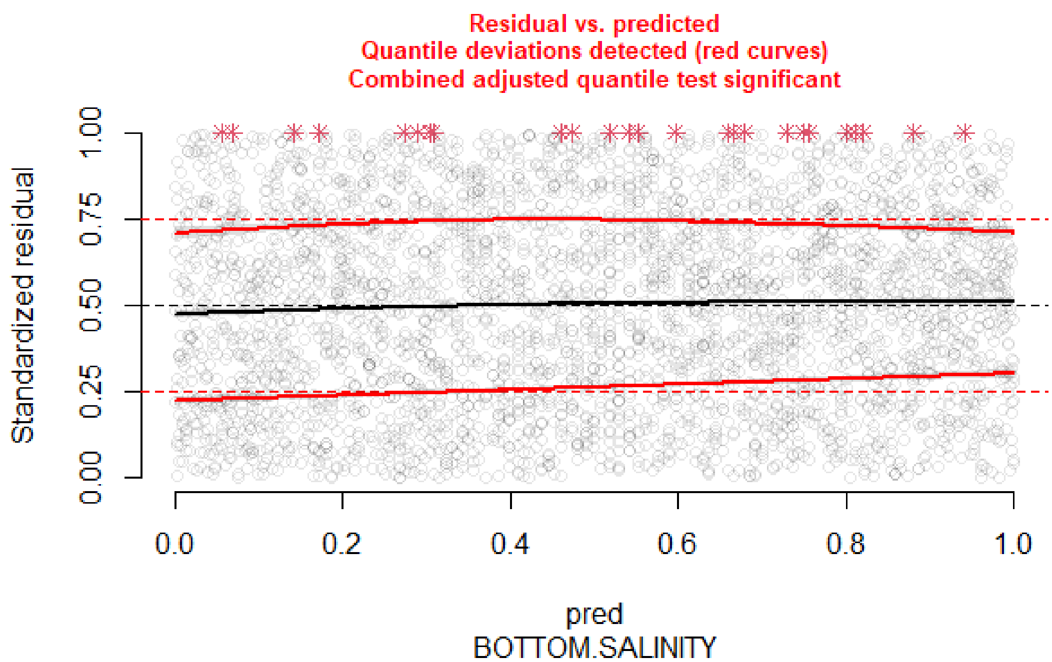
No effort offset, zero-inflated

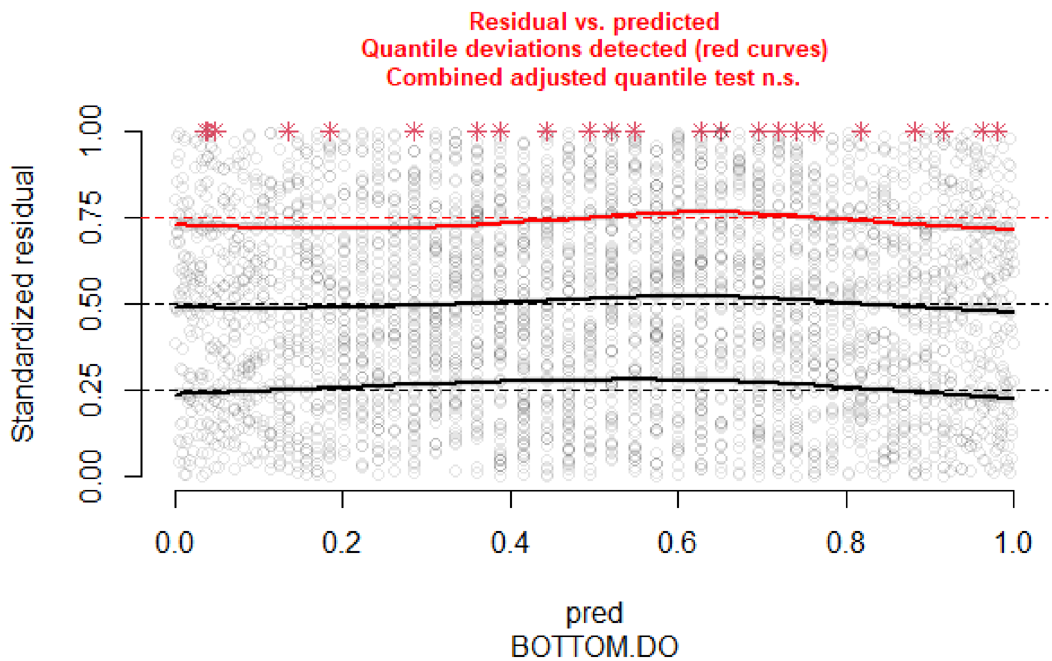



Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
2001
2003
2005
2007
2009
2011
2013
2015
2017
2019
catPred

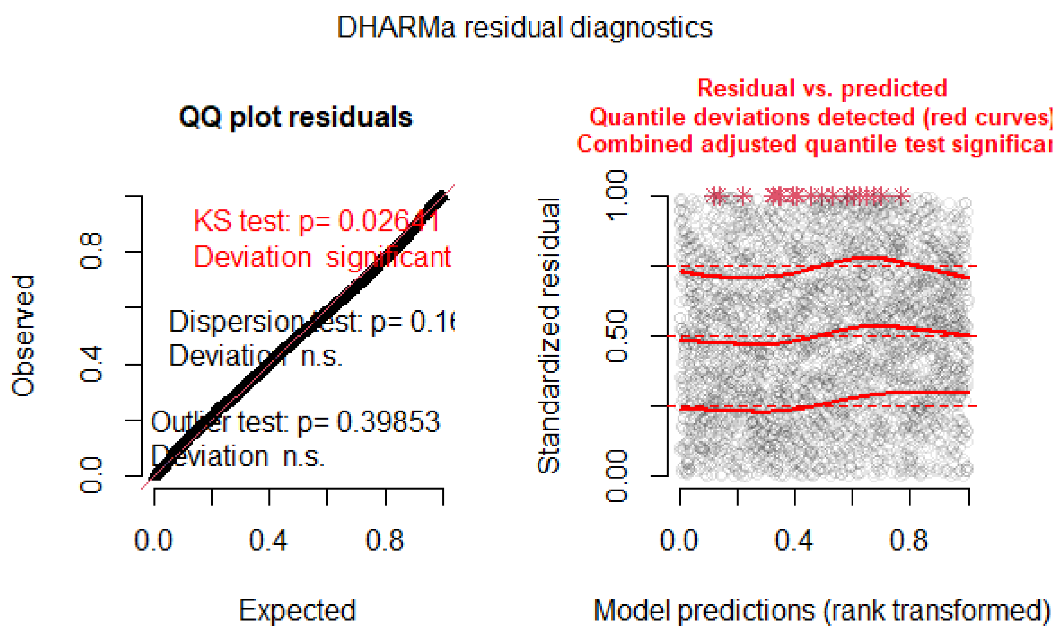

Within-group deviations from uniformity significant (red)
Levene Test for homogeneity of variance significant
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
WANC1
WANC2
WASH1
WASH2
catPred

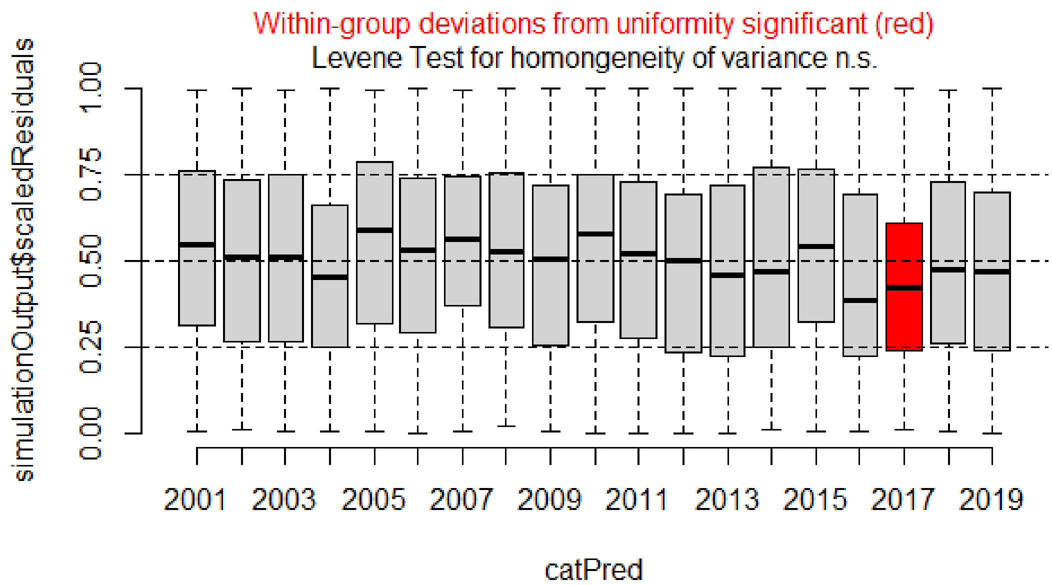

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP


Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.SALINITY

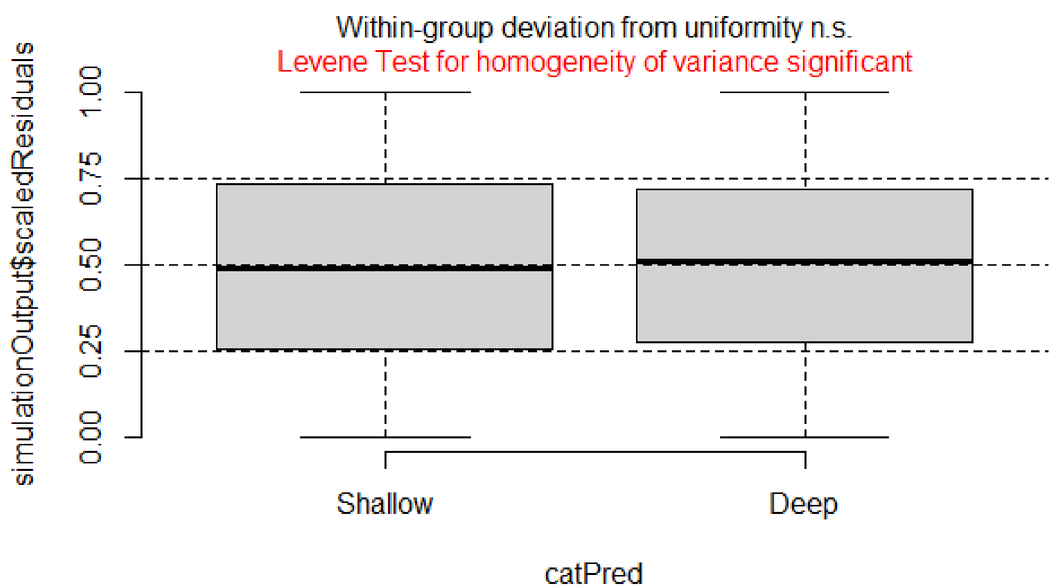


Model with depth and region separated, zero-inflated

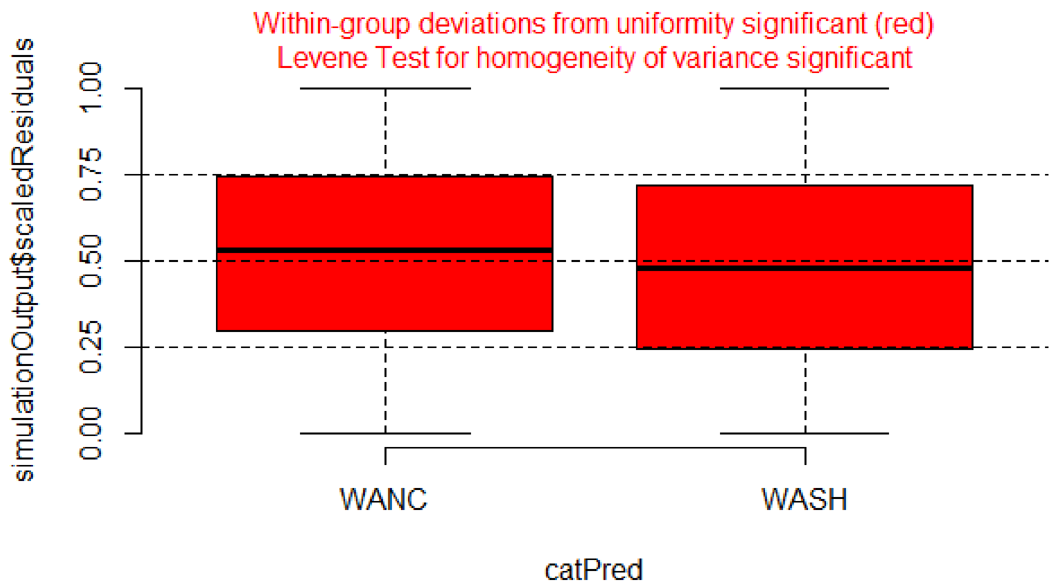




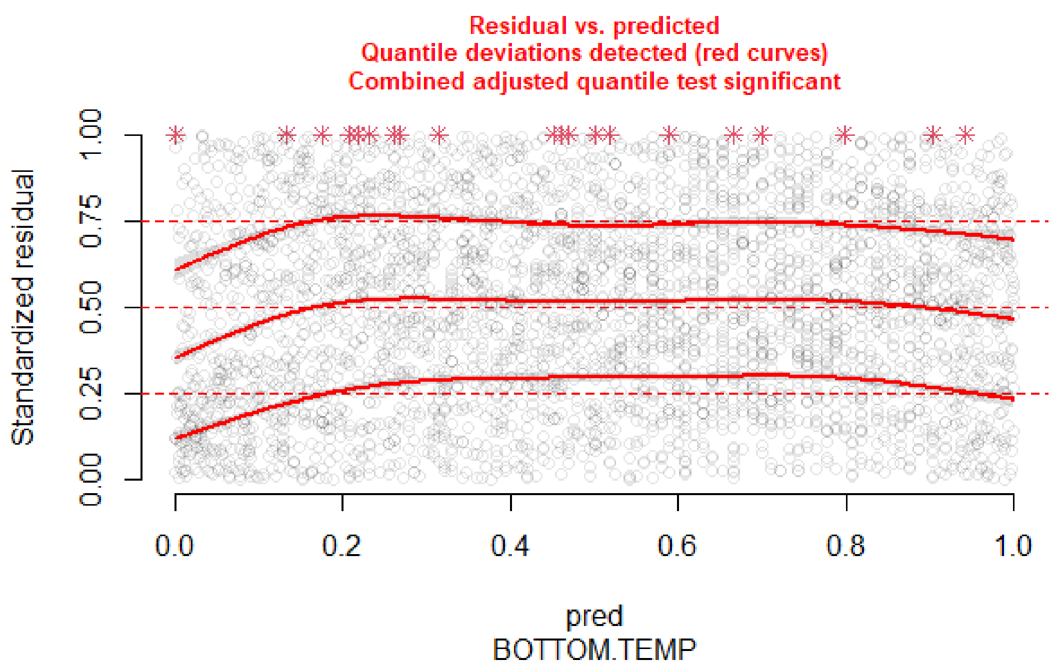
Within-group deviations from uniformity significant (red)
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
2001
2003
2005
2007
2009
2011
2013
2015
2017
2019
catPred



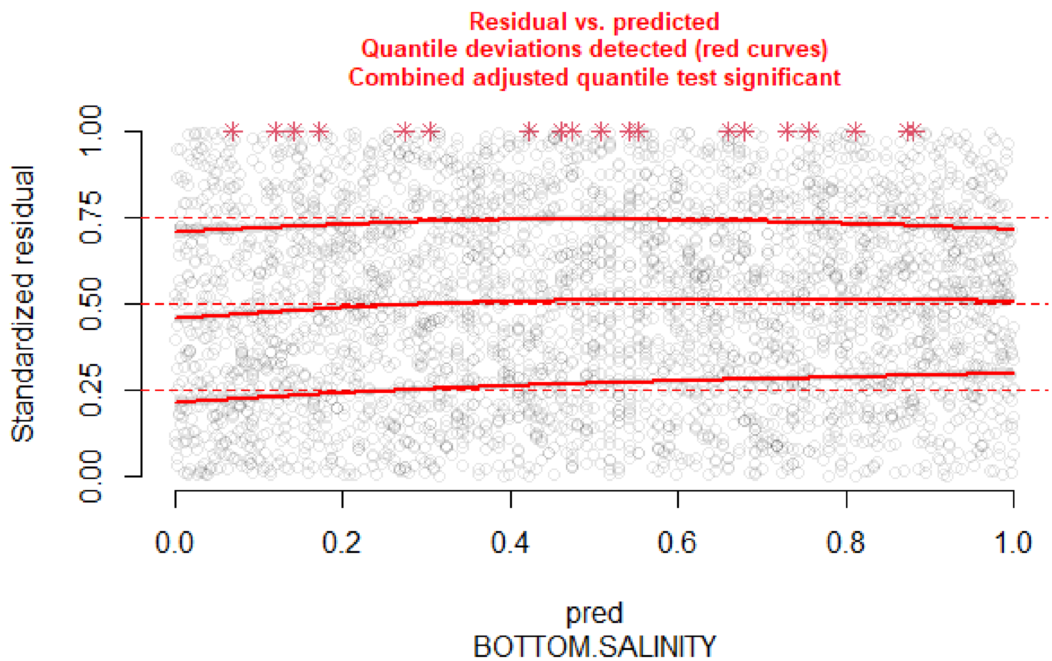
Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance significant
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
Shallow
Deep
catPred

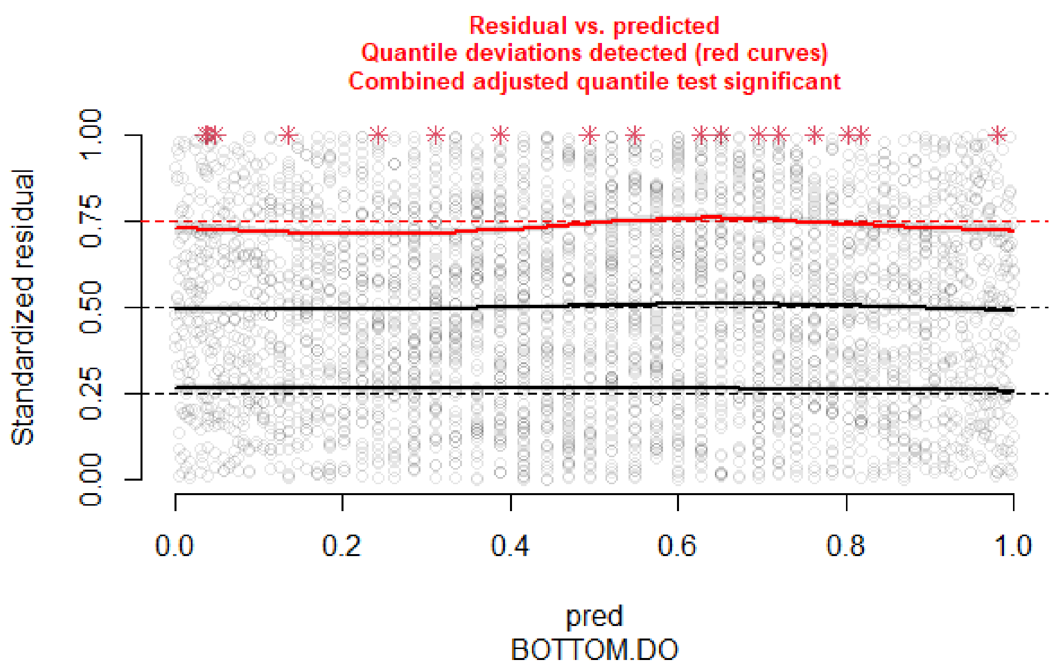


Within-group deviations from uniformity significant (red)
Levene Test for homogeneity of variance significant
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
WANC
WASH
catPred



Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP





###

Step 3: Winning model

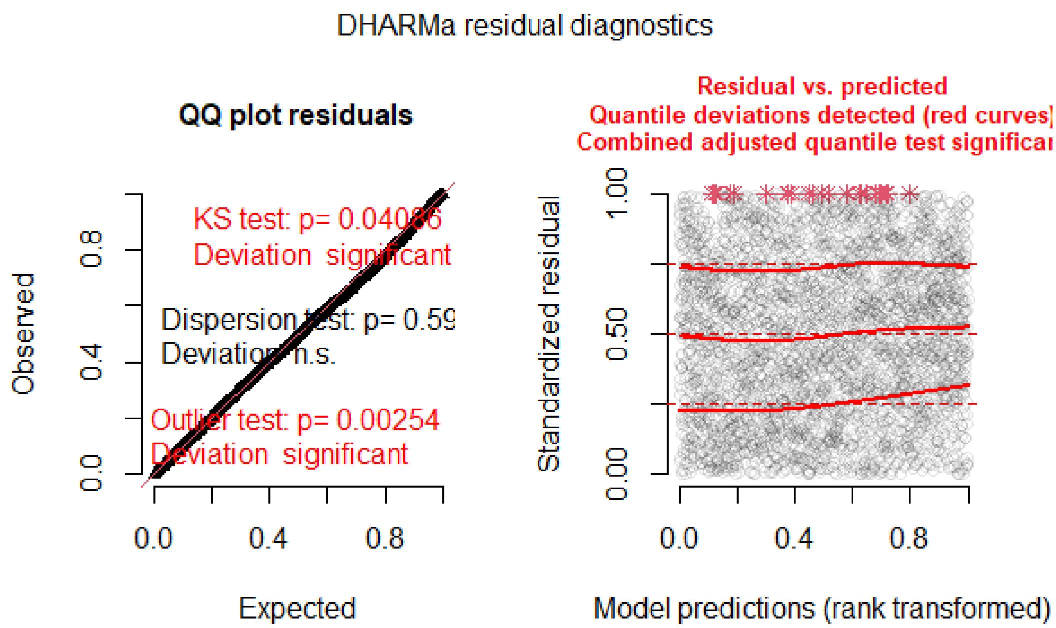
Select the type of model to use and update this chunk. Include all covariates.

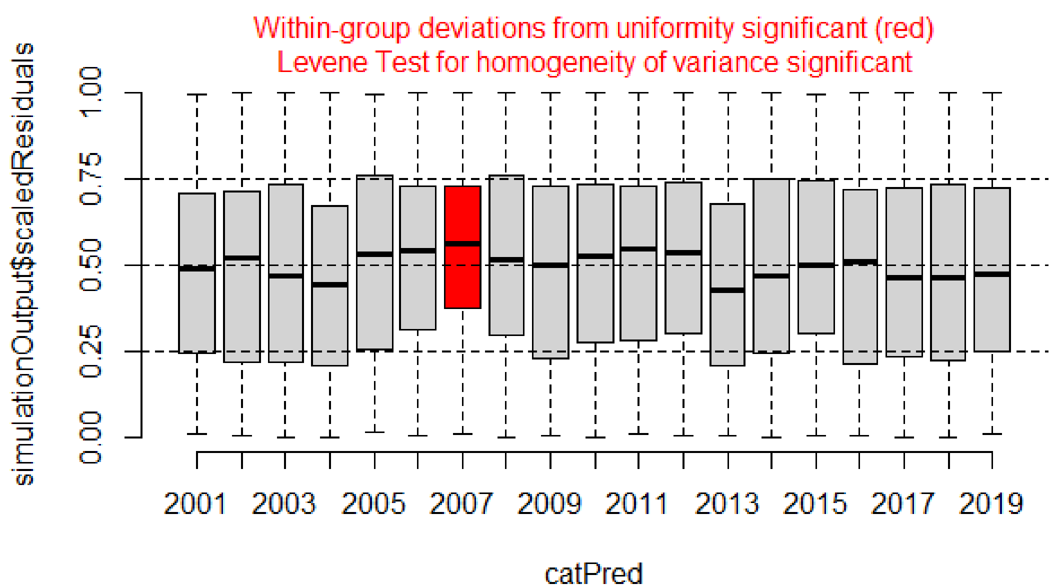
Test covariate importance

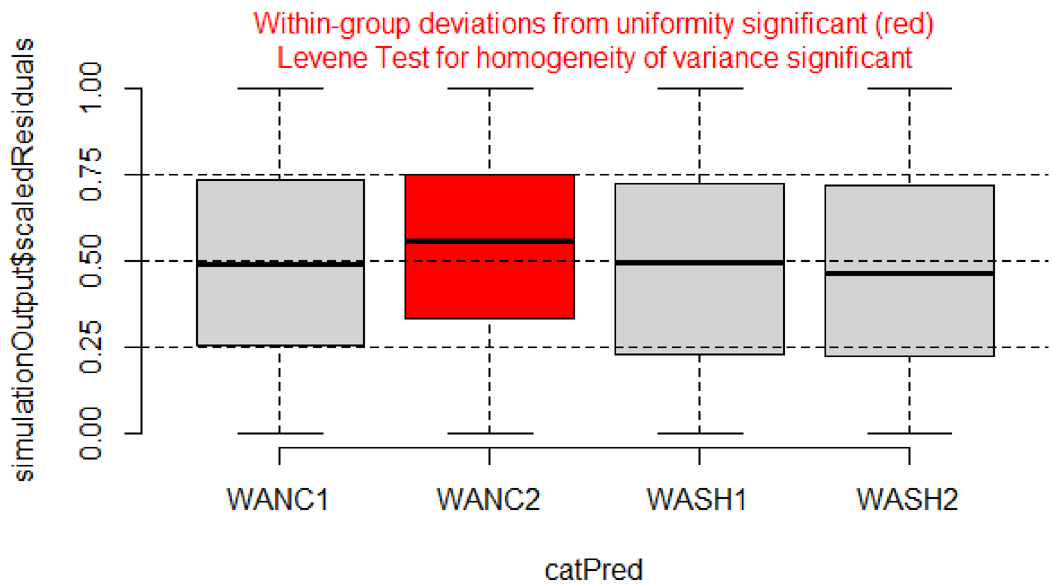
```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ 1
```

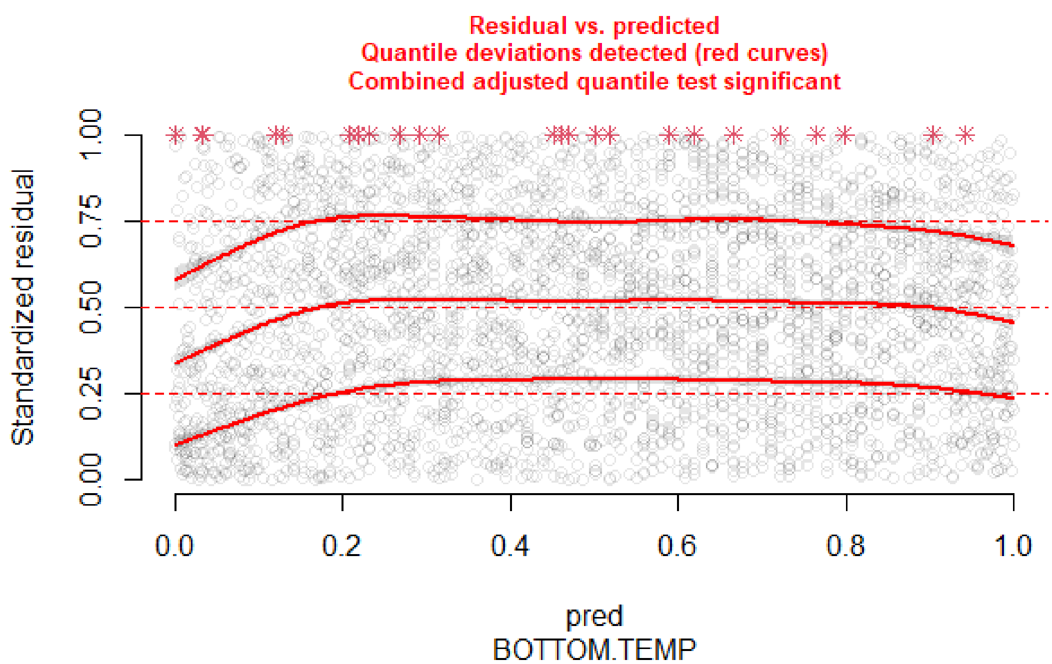
```
##
## Final Model:
## FREQ ~ STRATUM + YEAR
##
##
##        Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                            3576   19052.79 19106.79
## 2 + STRATUM  3 457.6197      3573   18595.17 18655.17
## 3    + YEAR 18 157.9971      3555   18437.18 18533.18
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO +
##     offset(lnEffort)
##
## Final Model:
## FREQ ~ YEAR + STRATUM + offset(lnEffort)
##
##
##                Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                                     3552   18428.98 18530.98
## 2       - BOTTOM.DO  1 0.2770063      3553   18429.26 18529.26
## 3 - BOTTOM.SALINITY  1 0.7924471      3554   18430.05 18528.05
## 4     - BOTTOM.TEMP  1 1.0845428      3555   18431.14 18527.14
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO +
##     offset(lnEffort)
##
## Final Model:
## FREQ ~ YEAR + STRATUM + offset(lnEffort)
##
##
##                Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                                     3552   18428.98 18530.98
## 2       - BOTTOM.DO  1 0.2770063      3553   18429.26 18529.26
## 3 - BOTTOM.SALINITY  1 0.7924471      3554   18430.05 18528.05
## 4     - BOTTOM.TEMP  1 1.0845428      3555   18431.14 18527.14

##        dAIC  df
## ZANB     0.0 48
## ZANB.2   4.4 47
## ZANB.3  32.2 46
## ZANB.4 113.4 45
## ZANB.5 194.3 27
## ZANB.7 303.7 24
## ZANB.6 310.4 42
```

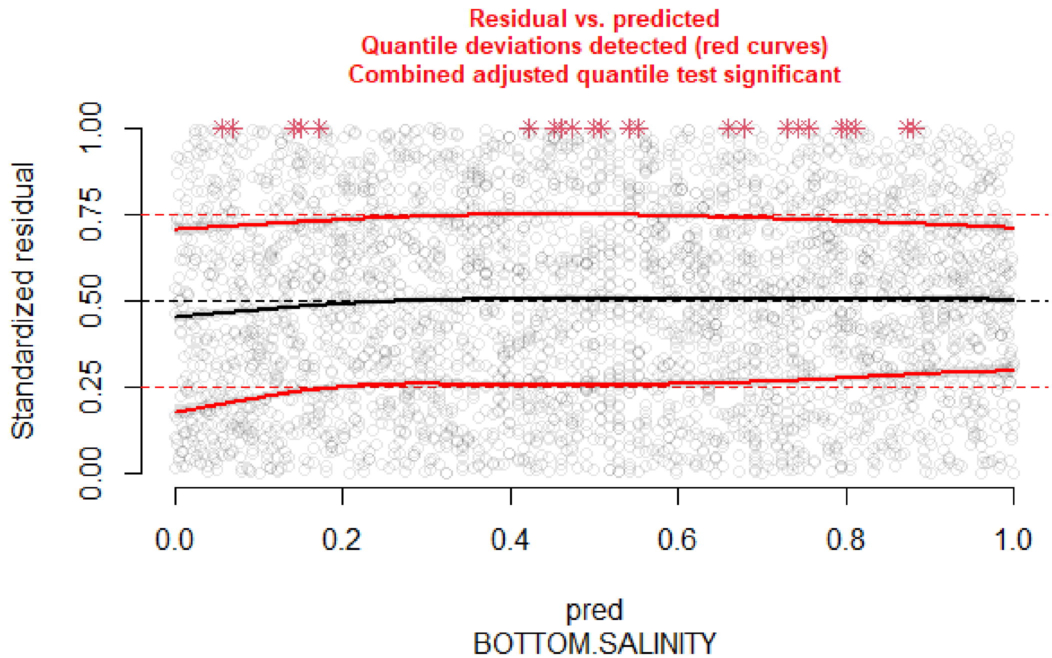
DHARMA simulations

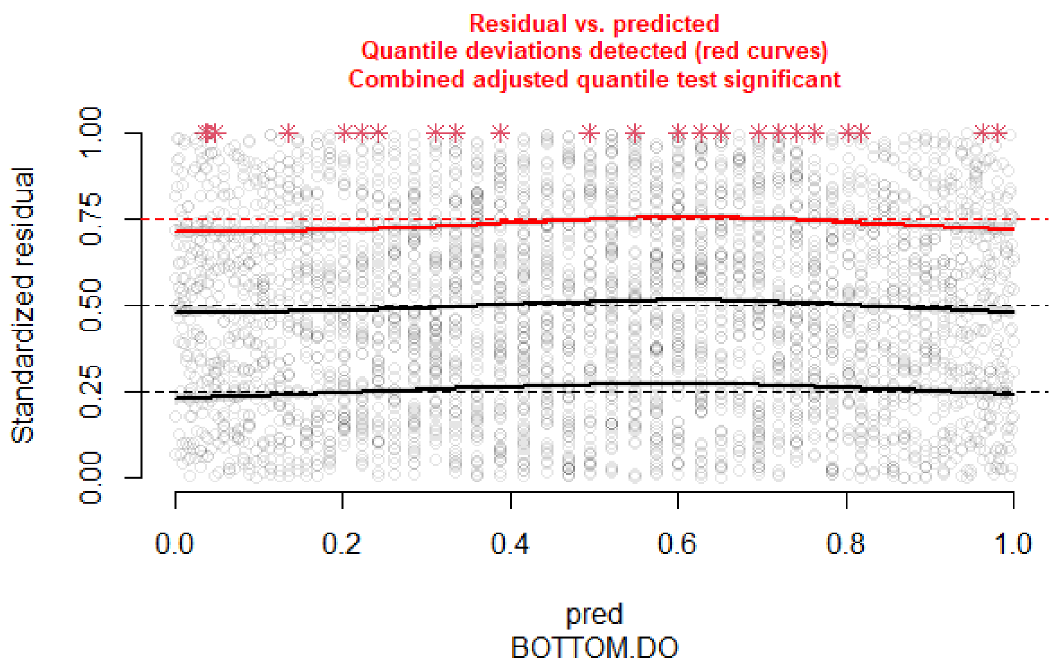





Within-group deviations from uniformity significant (red)
Levene Test for homogeneity of variance significant
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
WANC1
WANC2
WASH1
WASH2
catPred


Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP



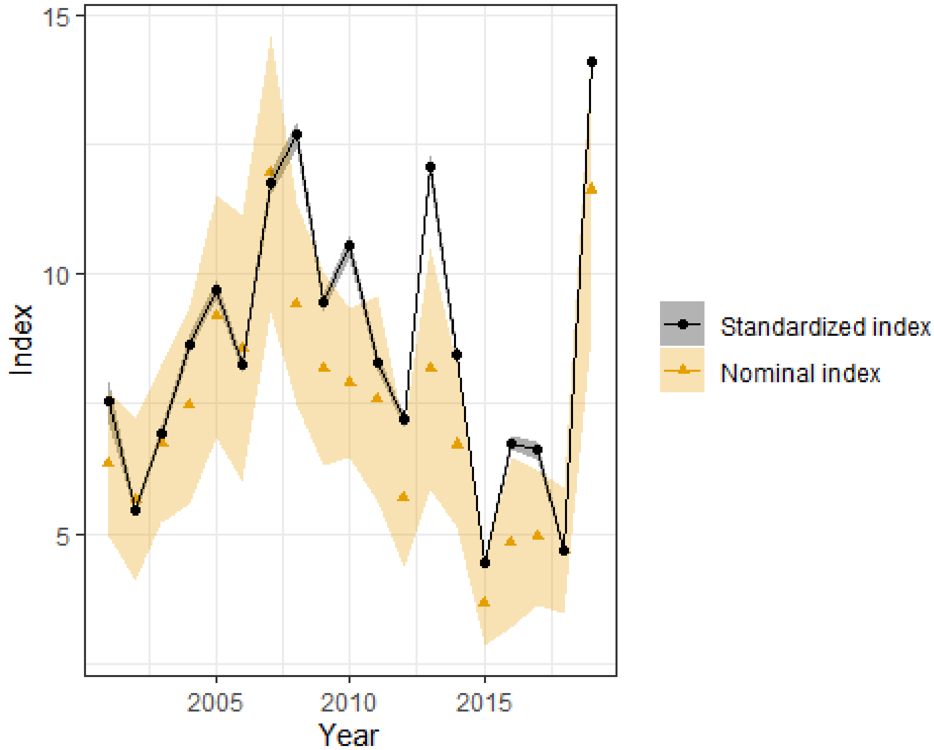


Step 4: Plot predictions

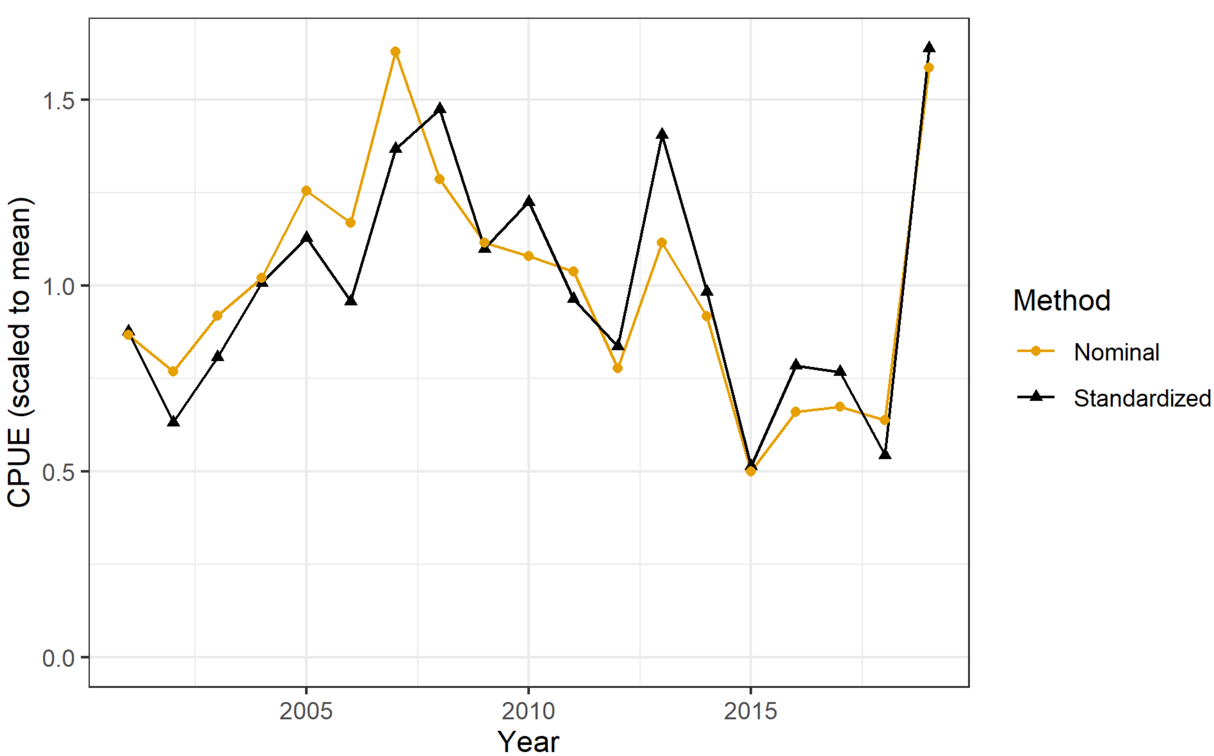
Final model:

FREQ ~YEAR+STRATUM, ziformula = ~ YEAR + STRATUM + BOTTOM.TEMP + BOTTOM.SALINITY + BOTTOM.DO

Plot standardized index from this work with nominal stratified mean index.



Comparison of standardized index and nominal index scaled to their time-series means.



Bluefish index standardization - ChesMMAP Survey

Abigail Tyrell

2022-05-18

Develop standardized index

Step 1: Data processing

Load functions

```
source(here::here("scripts/functions.R"))

## Loading required package: carData

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Warning: package 'glmmTMB' was built under R version 4.0.5

## Registered S3 methods overwritten by 'lme4':
##   method                          from
##   cooks.distance.influence.merMod car
##   influence.merMod                car
##   dfbeta.influence.merMod         car
##   dfbetas.influence.merMod        car

## Warning: package 'DHARMa' was built under R version 4.0.5

## This is DHARMa 0.4.4. For overview type '?DHARMa'. For recent changes,
type news(package = 'DHARMa')

## Warning: package 'bbmle' was built under R version 4.0.5

## Loading required package: stats4

## Warning: package 'ggeffects' was built under R version 4.0.5

## Warning: package 'ggplot2' was built under R version 4.0.5
```

Knit tables with `knitr::kable`

```
knit_print.data.frame <- function(x, ...) {
  res <- paste(c("", "", knitr::kable(x, digits = 2)), collapse = "\n")
```

```
  knitr::asis_output(res)
}

registerS3method(
  "knit_print", "data.frame", knit_print.data.frame,
  envir = asNamespace("knitr")
)
```

Read in data

```
indata <- readxl::read_excel(here::here("data/2021_bluefish_data_VA.xlsx"),
sheet = "Survey Data_ChesMMAP") %>%
  dplyr::rename(Stratum_area = Wgt.Factor)
head(indata, n = 2)
```

SampleID	State	Program	Gear	Year	Month	Date	Time	Areasw	Bf.count	Bf.totwght	Area	Stratum	Depth	Stratum_area	Lat	Lon	Bottom.wt	Bottom.sa	Bottom.do
CM20020301001	MD	ChesMMAP	BottomTrawl	2002	MAR	2002-04-04	2002-04-04 16:58:00	64712.99	0	0	ChesBay	101	2.44	450365801	39.14	-76.29	10.93	6.2	11.49
CM20020301002	MD	ChesMMAP	BottomTrawl	2002	MAR	2002-04-04	2002-04-04 15:55:00	52713.86	0	0	ChesBay	101	3.96	450365801	39.19	-76.35	9.86	4.4	9.25

```
# modify data formatting
indata <- indata %>%
  dplyr::rename(FREQ = Bf.count) %>%
  dplyr::select(-c(Program, Gear, Area)) # same value for all data
head(indata, n = 2)
```

Sampl	St	Y	M	D	Ti	Are	F	Bf.t	Str	D	Strat	L	L	Bot	Bot	Bot

eID	ate	ear	onth	ate	me	asw	REQ	otwght	atum	epth	um_area	at	on	tom.wt	tom.sa	tom.do
CM20020301001	MD	2002	MAR	2002-04-04	2002-04-04 16:58:00	64712.99	0	0	101	2.44	450365801	39.14	-76.29	10.93	6.2	11.49
CM20020301002	MD	2002	MAR	2002-04-04	2002-04-04 15:55:00	52713.86	0	0	101	3.96	450365801	39.19	-76.35	9.86	4.4	9.25

Remove missing data

```
new_data <- remove_missing(indata,
  f_col = "FREQ",
  t_col = "Year"
)

## Data before removing missing values:

##    SampleID            State              Year              Month
##  Length:6264        Length:6264        Length:6264        Length:6264
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##       Date                            Time                            Areasw
##  Min.   :2002-03-29 00:00:00   Min.   :2002-03-29 09:30:00   Min.   : 6306
##  1st Qu.:2006-03-30 00:00:00   1st Qu.:2006-03-30 12:32:15   1st Qu.:29848
##  Median :2010-09-06 00:00:00   Median :2010-09-06 09:04:00   Median :45017
##  Mean   :2010-07-28 05:25:58   Mean   :2010-07-28 18:01:05   Mean   :42959
##  3rd Qu.:2014-11-09 00:00:00   3rd Qu.:2014-11-09 10:08:35   3rd Qu.:52980
##  Max.   :2018-11-17 00:00:00   Max.   :2018-11-17 15:59:40   Max.   :98741
##
##       FREQ          Bf.totwght         Stratum             Depth
##  Min.   : 0.0000   Min.   : 0.00000   Length:6264        Min.   : 2.438
##  1st Qu.: 0.0000   1st Qu.: 0.00000   Class :character   1st Qu.: 7.407
```

```
##  Median : 0.0000   Median : 0.00000   Mode  :character   Median :10.363
##  Mean   : 0.1502   Mean   : 0.03152                      Mean   :11.199
##  3rd Qu.: 0.0000   3rd Qu.: 0.00000                      3rd Qu.:12.802
##  Max.   :85.0000   Max.   :13.72200                      Max.   :41.758
##                                                          NA's   :1
##   Stratum_area           Lat             Lon          Bottom.wt
##  Min.   : 20795800   Min.   :36.93   Min.   :-76.52   Min.   : 0.81
##  1st Qu.:405279801   1st Qu.:37.39   1st Qu.:-76.32   1st Qu.:13.30
##  Median :454035100   Median :37.76   Median :-76.20   Median :18.24
##  Mean   :507171662   Mean   :37.90   Mean   :-76.21   Mean   :18.00
##  3rd Qu.:776073103   3rd Qu.:38.34   3rd Qu.:-76.10   3rd Qu.:24.68
##  Max.   :897041604   Max.   :39.30   Max.   :-75.93   Max.   :29.44
##                                                       NA's   :213
##    Bottom.sa       Bottom.do
##  Min.   : 0.09   Min.   : 0.000
##  1st Qu.:15.34   1st Qu.: 4.870
##  Median :19.04   Median : 6.800
##  Mean   :19.05   Mean   : 6.662
##  3rd Qu.:22.87   3rd Qu.: 8.530
##  Max.   :34.98   Max.   :14.680
##  NA's   :239     NA's   :609
```

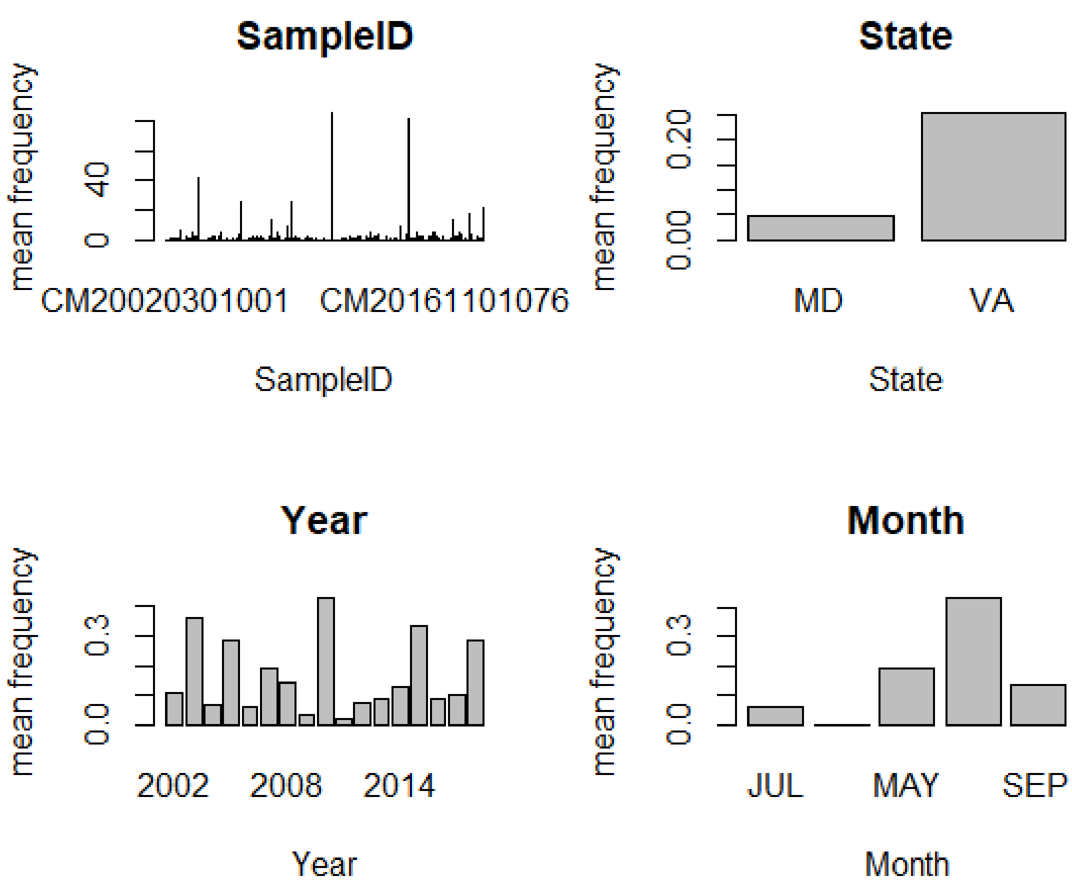
```
## Data after removing missing values:
```

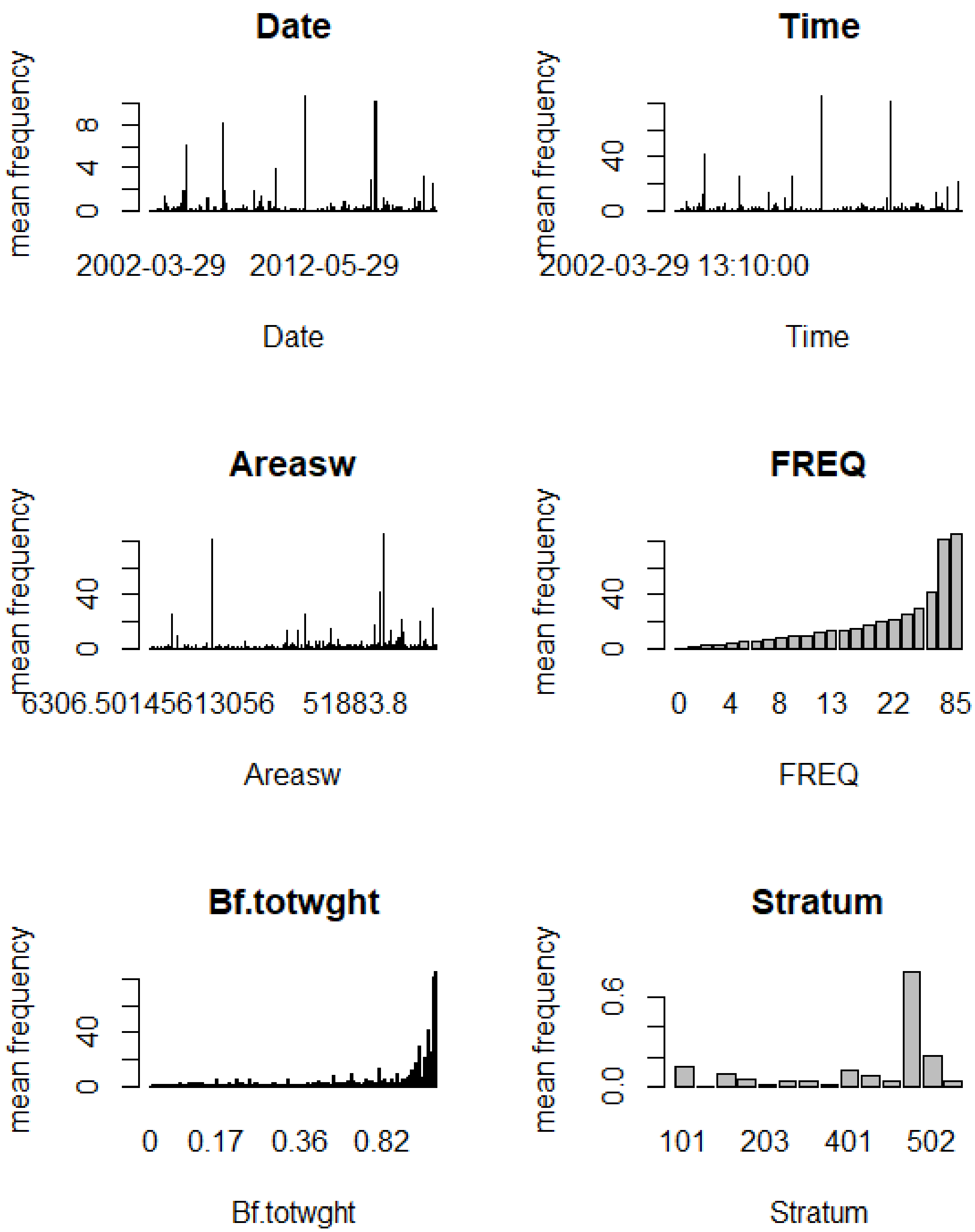
```
##    SampleID            State               Year              Month
##  Length:5620        Length:5620        Length:5620        Length:5620
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##       Date                            Time                          Areasw
##  Min.   :2002-03-29 00:00:00   Min.   :2002-03-29 13:10:00   Min.   : 6306
##  1st Qu.:2006-05-15 18:00:00   1st Qu.:2006-05-16 03:15:30   1st Qu.:29570
##  Median :2010-09-06 00:00:00   Median :2010-09-06 10:04:00   Median :44869
##  Mean   :2010-09-11 03:54:26   Mean   :2010-09-11 16:28:46   Mean   :42812
##  3rd Qu.:2015-05-28 00:00:00   3rd Qu.:2015-05-28 10:40:19   3rd Qu.:52932
##  Max.   :2018-11-17 00:00:00   Max.   :2018-11-17 15:59:40   Max.   :98741
##       FREQ           Bf.totwght          Stratum               Depth
##  Min.   : 0.0000   Min.   : 0.00000   Length:5620        Min.   : 2.438
##  1st Qu.: 0.0000   1st Qu.: 0.00000   Class :character   1st Qu.: 7.498
##  Median : 0.0000   Median : 0.00000   Mode  :character   Median :10.363
##  Mean   : 0.1637   Mean   : 0.03384                      Mean   :11.191
##  3rd Qu.: 0.0000   3rd Qu.: 0.00000                      3rd Qu.:12.802
##  Max.   :85.0000   Max.   :13.72200                      Max.   :41.758
##   Stratum_area           Lat             Lon          Bottom.wt
##  Min.   : 20795800   Min.   :36.93   Min.   :-76.52   Min.   : 2.30
##  1st Qu.:405279801   1st Qu.:37.41   1st Qu.:-76.32   1st Qu.:13.63
##  Median :454035100   Median :37.76   Median :-76.20   Median :19.04
##  Mean   :506759301   Mean   :37.91   Mean   :-76.21   Mean   :18.31
```

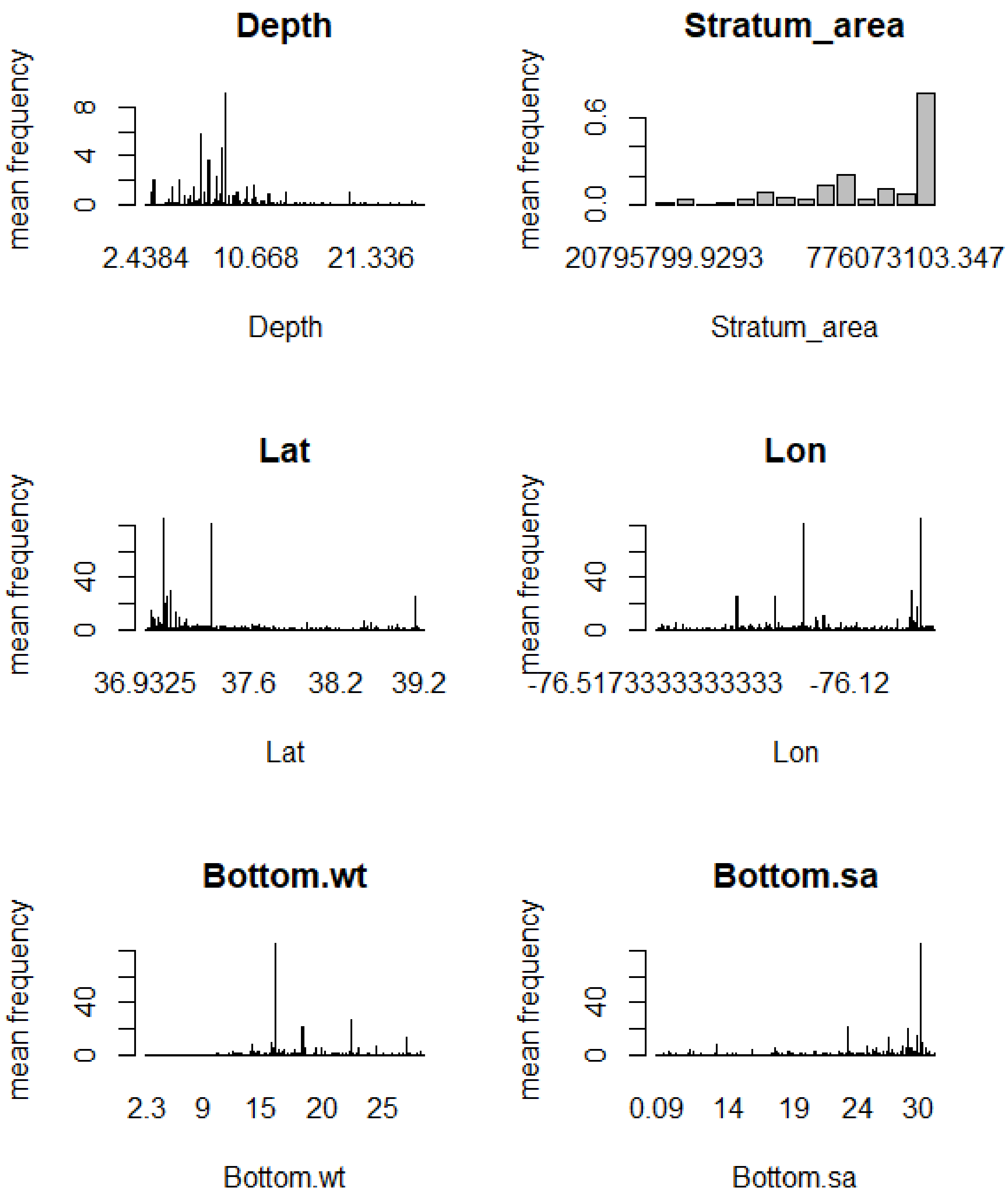
```
##   3rd Qu.:776073103   3rd Qu.:38.34   3rd Qu.:-76.10   3rd Qu.:24.83
##   Max.   :897041604   Max.   :39.30   Max.   :-75.93   Max.   :29.44
##     Bottom.sa        Bottom.do
##   Min.   : 0.09   Min.   : 0.000
##   1st Qu.:15.41   1st Qu.: 4.860
##   Median :19.05   Median : 6.800
##   Mean   :19.07   Mean   : 6.644
##   3rd Qu.:22.84   3rd Qu.: 8.510
##   Max.   :34.98   Max.   :14.680

## 644 datapoints removed (10.3% of data removed)

## All years had positive catch
```

Plot data and decide if any data should be removed due to sparse sampling

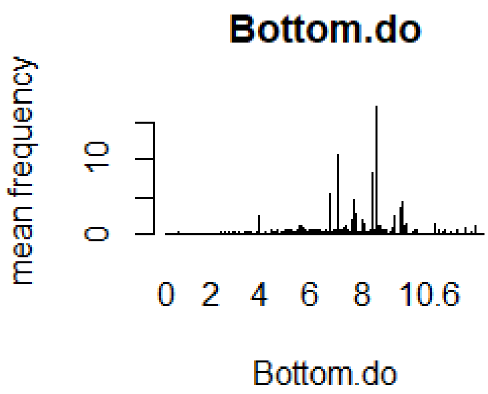
```
plot_obs(new_data,
  col = "FREQ"
)
```




Date
mean frequency
0
4
8
2002-03-29
2012-05-29
Date
Time
mean frequency
0
40
2002-03-29 13:10:00
Time
Areasw
mean frequency
0
40
6306.50145613056
51883.8
Areasw
FREQ
mean frequency
0
40
0
4
8
13
22
85
FREQ
Bf.totwght
mean frequency
0
40
0
0.17
0.36
0.82
Bf.totwght
Stratum
mean frequency
0.0
0.6
101
203
401
502
Stratum




```
## Please manually remove sparsely sampled data!
```

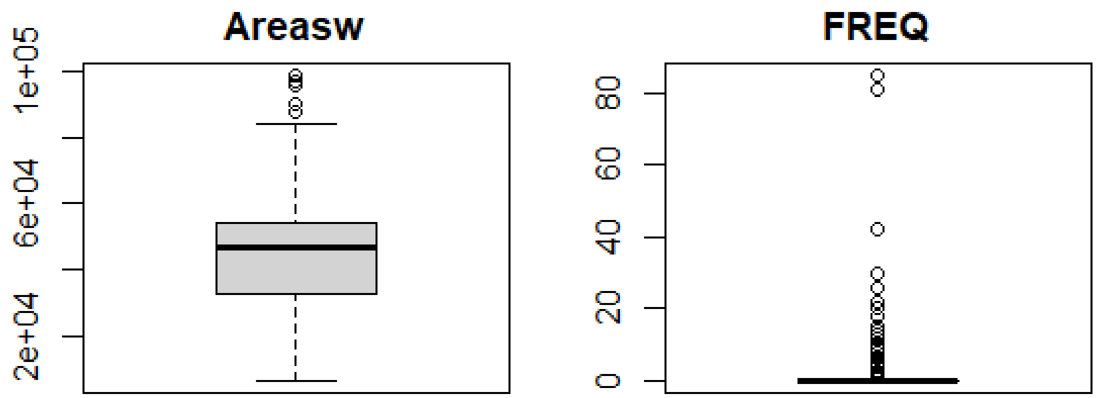


Remove sparse data

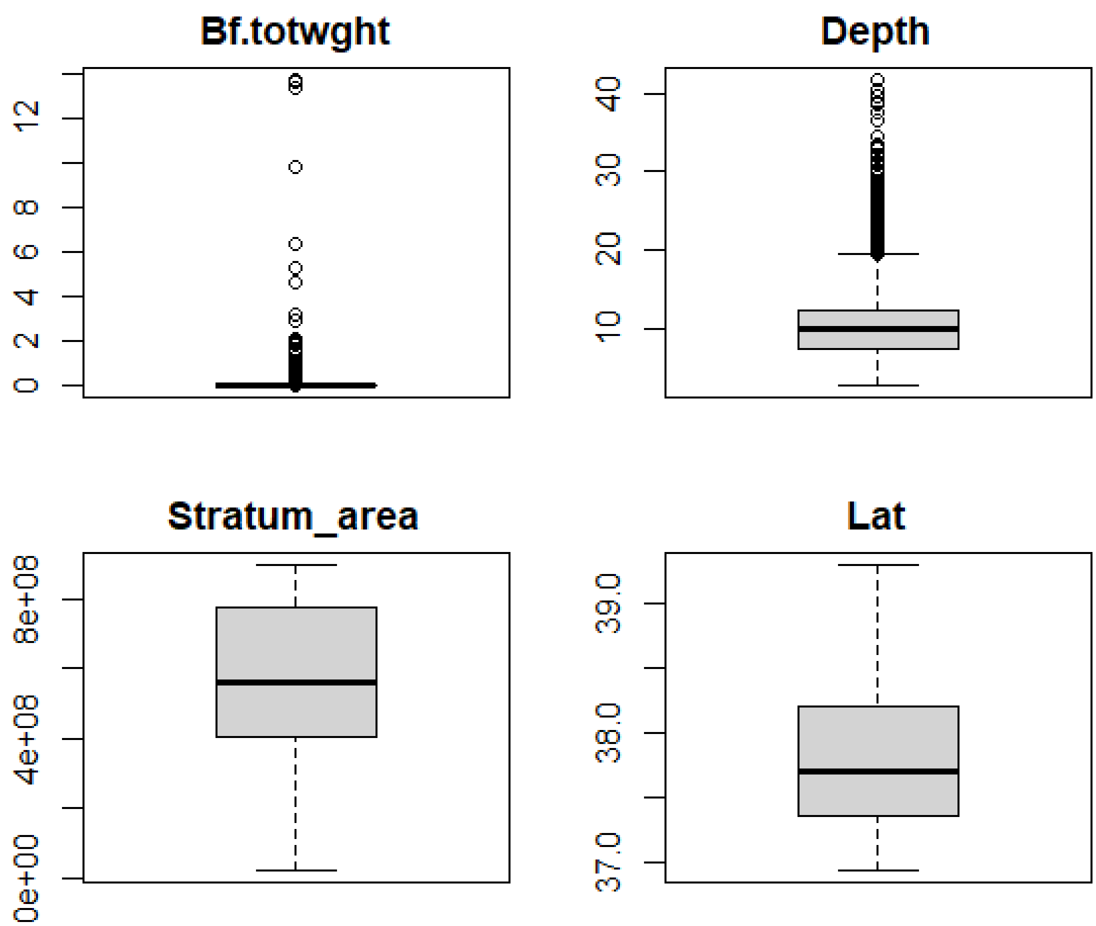
```
# remove March
length1 <- nrow(new_data)
new_data <- new_data %>%
  dplyr::filter(
    Month != "MAR",
    Bottom.wt > 12, # 12C is lower temp limit
    Bottom.do > 2
  ) # 2 mg/L is lower O2 limit
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")

## 1526 rows removed
```

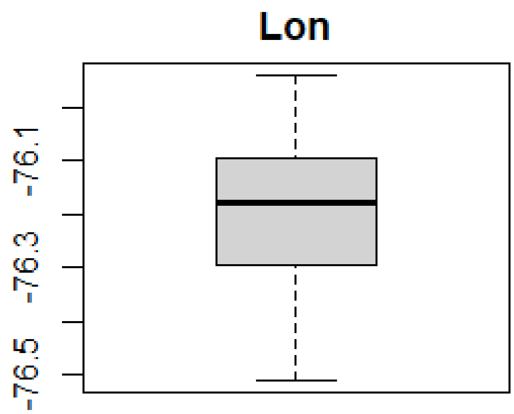
Check for outliers

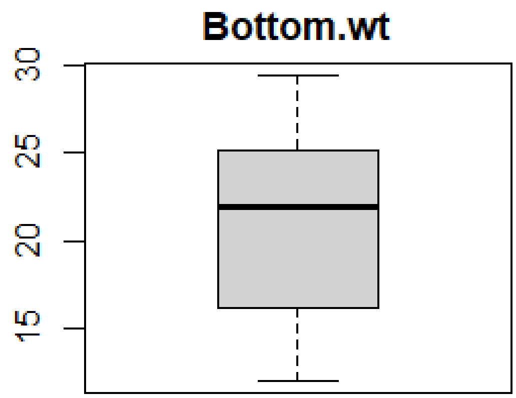
```
plot_outliers(new_data)

## SampleID is not numeric

## State is not numeric

## Year is not numeric

## Month is not numeric

## Date is not numeric

## Time is not numeric
```

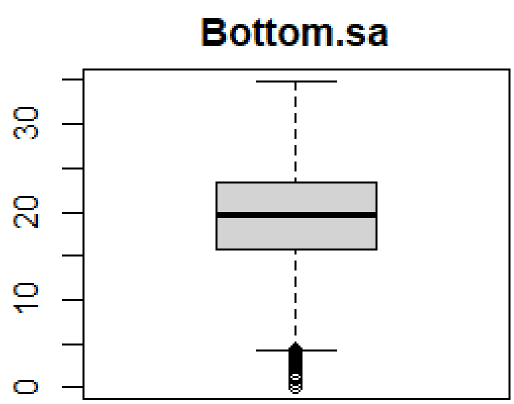



```
## Stratum is not numeric
```

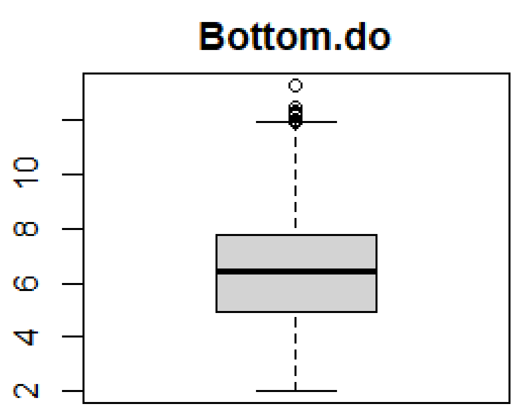









```
## Please manually remove any outliers!
```

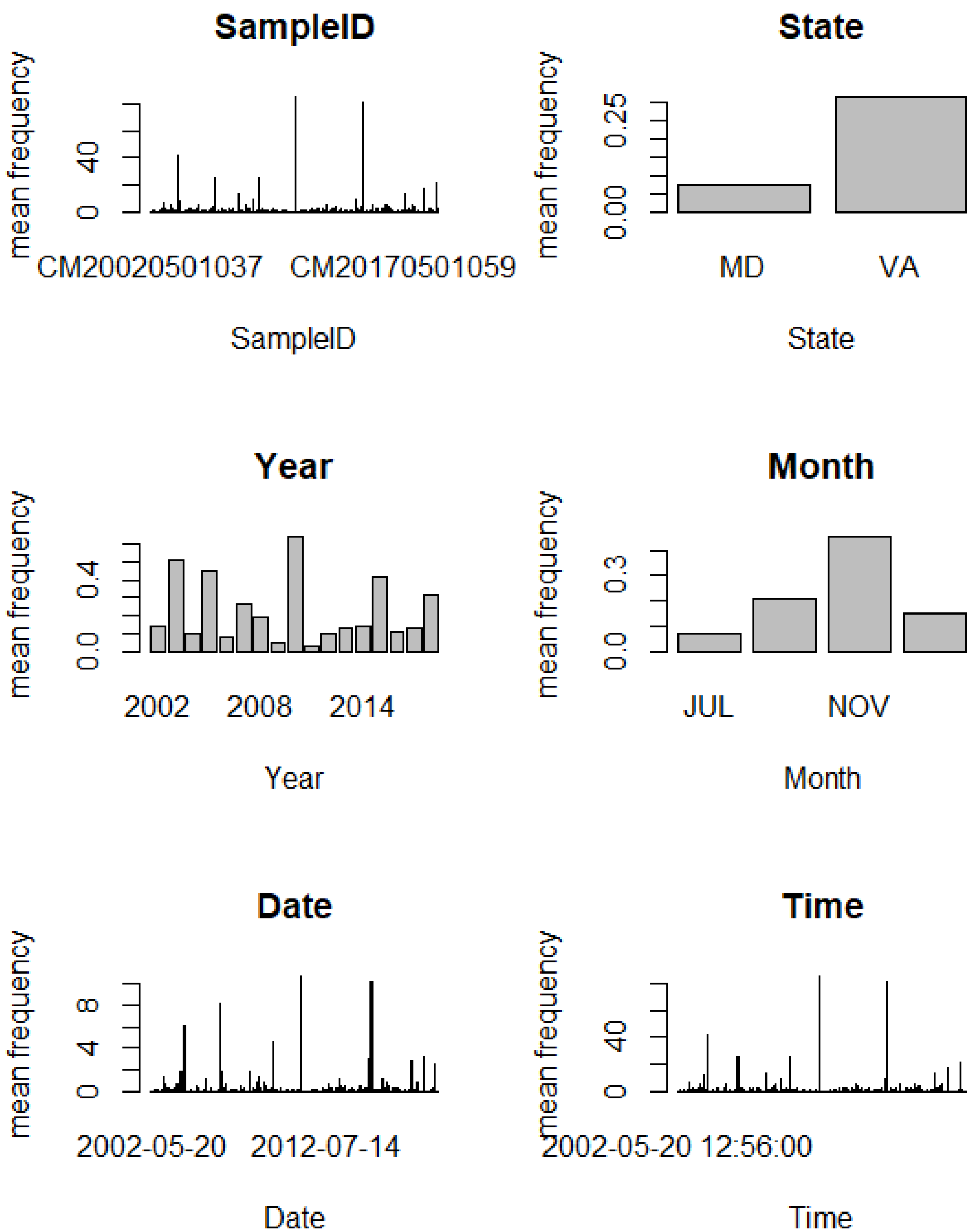


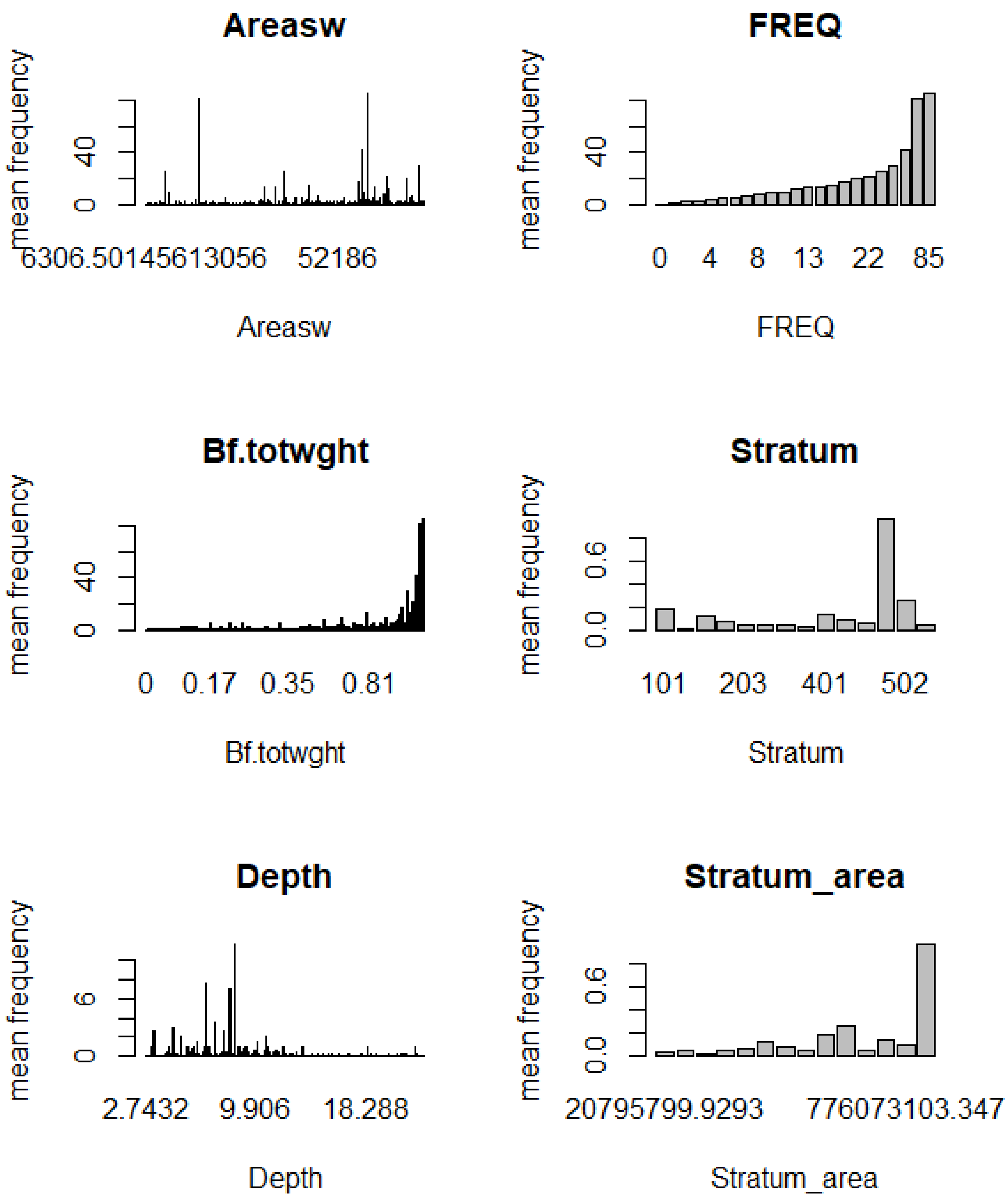
If needed, remove outliers

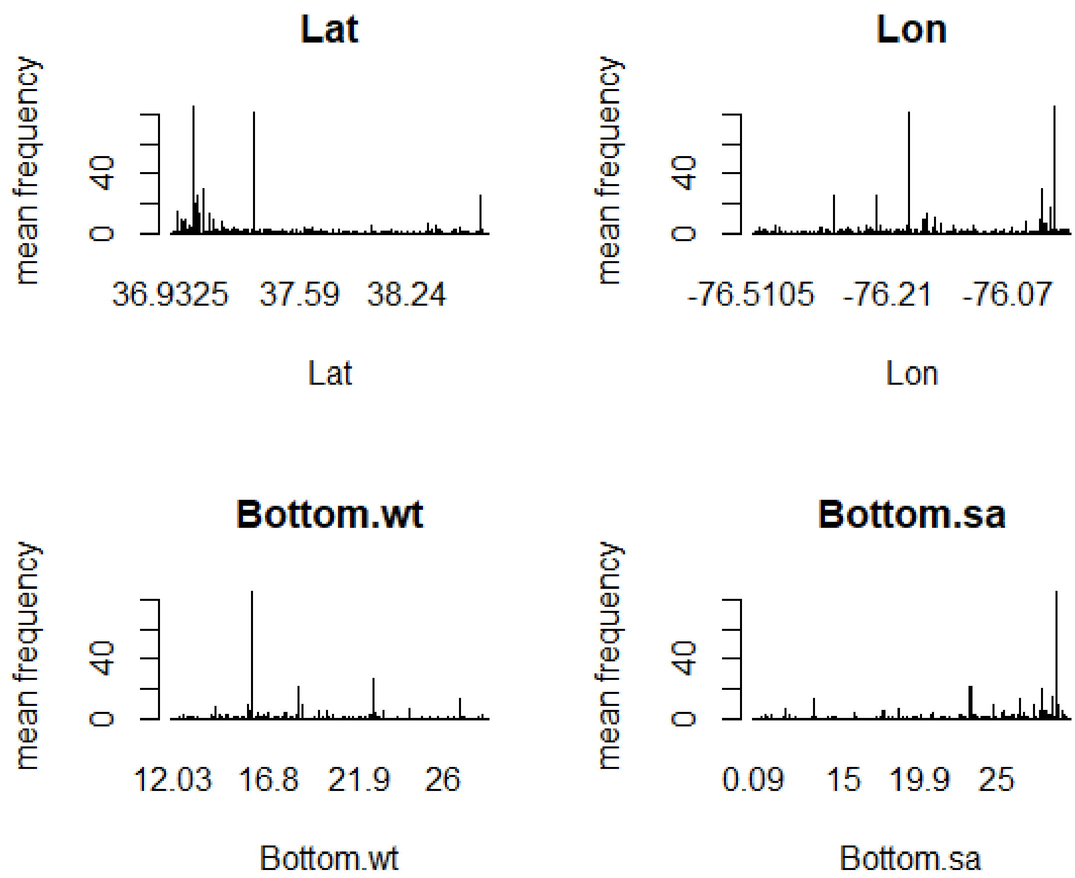
```
# None identified
```

Check for sparse sampling again

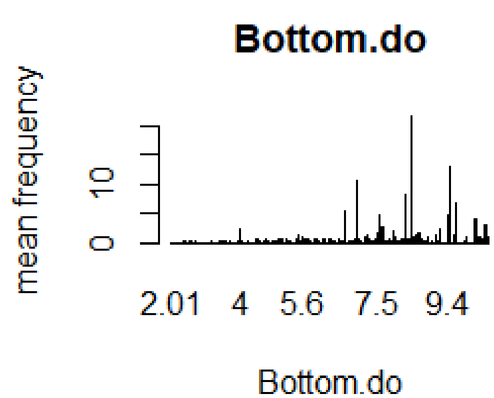
```
plot_obs(new_data,
  col = "FREQ"
)
```




Areasw
mean frequency
0
40
6306.50145613056
52186
Areasw
FREQ
mean frequency
0
40
0
4
8
13
22
85
FREQ
Bf.totwght
mean frequency
0
40
0
0.17
0.35
0.81
Bf.totwght
Stratum
mean frequency
0.0
0.6
101
203
401
502
Stratum
Depth
mean frequency
0
6
2.7432
9.906
18.288
Depth
Stratum_area
mean frequency
0.0
0.6
20795799.9293
776073103.347
Stratum_area




```
## Please manually remove sparsely sampled data!
```



Check again for years with no catch

```
check_missing(new_data, f_col = "FREQ", t_col = "Year")

## Years with no catch: NONE
```

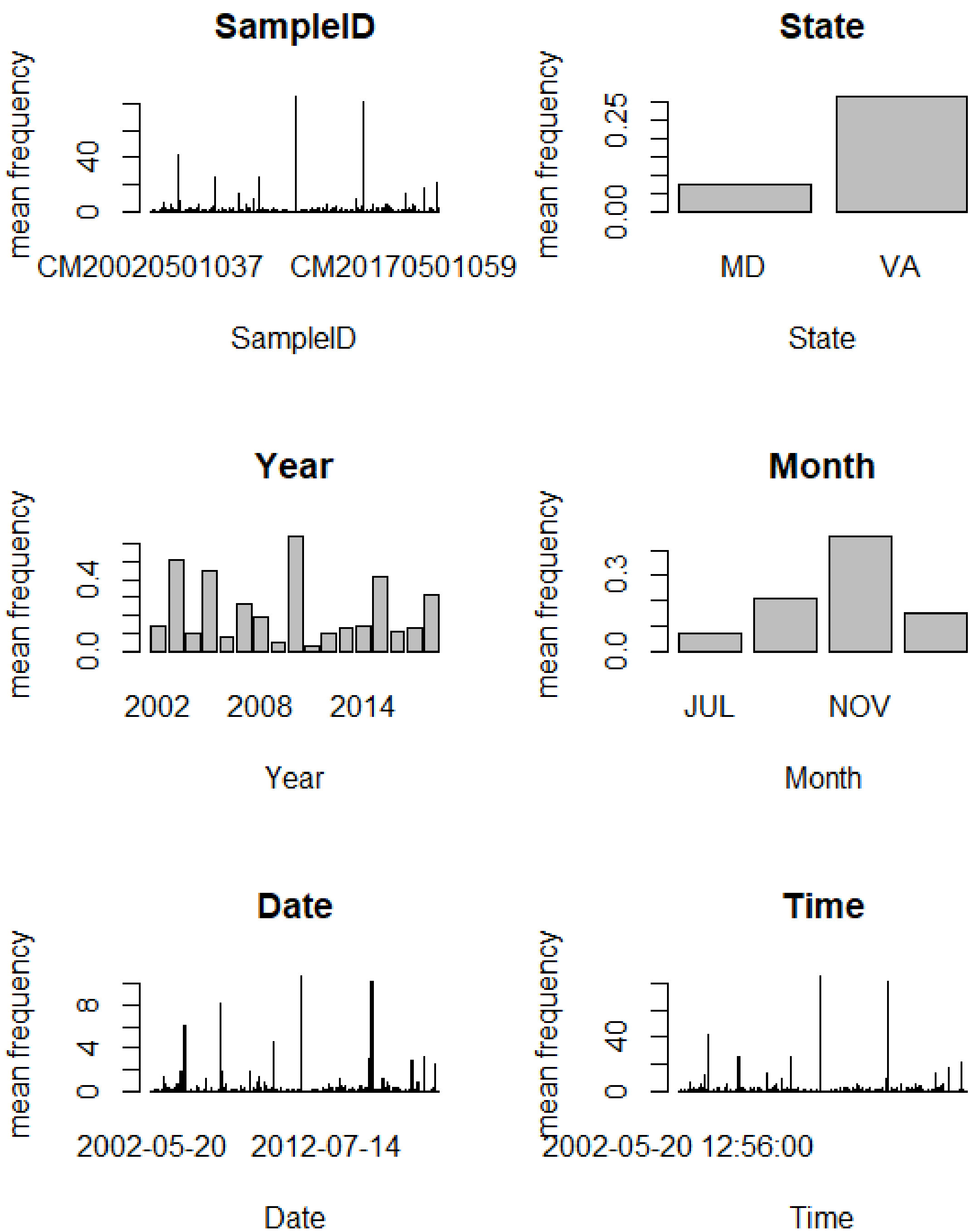
Change continuous variables to z-score

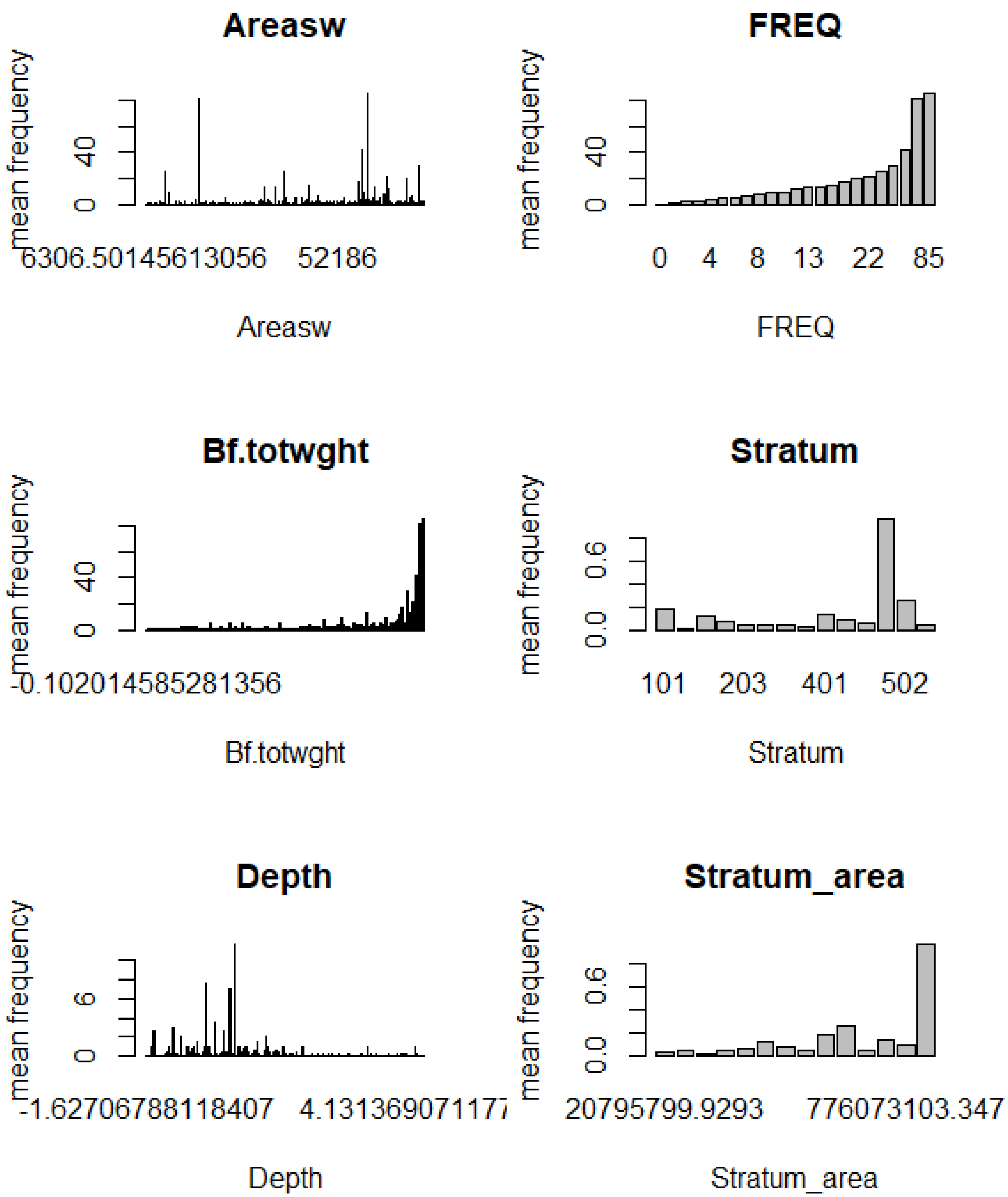
```
new_data2 <- standardize_data(new_data,
  cols = c(
    "Year", "Month", "Stratum", "Date",
    "Time", "Lat", "Lon", "Area",
    "Areasw", "Stratum_area",
    "SampleID", "State"
  ),
  f_col = "FREQ"
)

## SampleID has been changed to factor

## State has been changed to factor

## Year has been changed to factor

## Month has been changed to factor

## Date has been changed to factor

## Time has been changed to factor

## Areasw has been changed to factor

## Bf.totwght has been standardized

## Stratum has been changed to factor

## Depth has been standardized

## Stratum_area has been changed to factor

## Lat has been changed to factor

## Lon has been changed to factor

## Bottom.wt has been standardized

## Bottom.sa has been standardized

## Bottom.do has been standardized

## Data standardized!

# change lat, lon, Areasw back to numeric
new_data2 <- new_data2 %>%
  dplyr::mutate(
    Lat = Lat %>%
      as.character() %>%
      as.numeric(),
    Lon = Lon %>%
      as.character() %>%
```

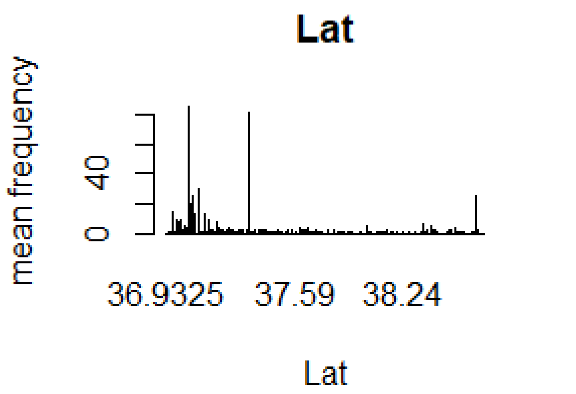
```
      as.numeric(),
    Areasw = Areasw %>%
      as.character() %>%
      as.numeric()
  )
head(new_data2, n = 2)
```

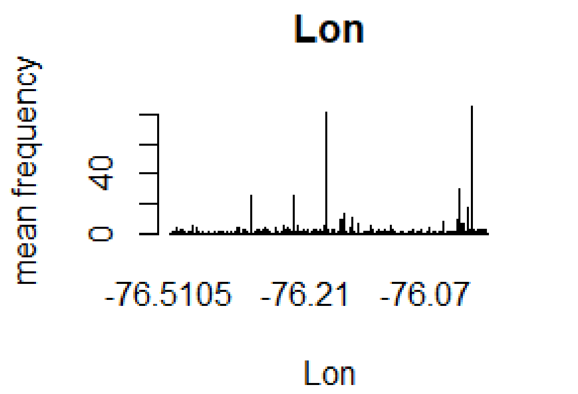
SampleID	State	Year	Month	Date	Time	Areasw	FREQ	Bf.totwght	Stratum	Depth	Stratum_area	Lat	Lon	Bottom.wt	Bottom.sa	Bottom.do
CM200205010 37	MD	2002	MAY	2002-05-20	2002-05-20 19:29:00	49427.75	0	-0.1	302	0.31	44704620 1.349	38.12	-76.26	-0.61	-1.81	0.28
CM200205010 65	VA	2002	MAY	2002-05-20	2002-05-20 15:37:00	47410.44	0	-0.1	403	2.13	22185289 9.225	37.71	-76.18	-0.51	-0.41	0.38

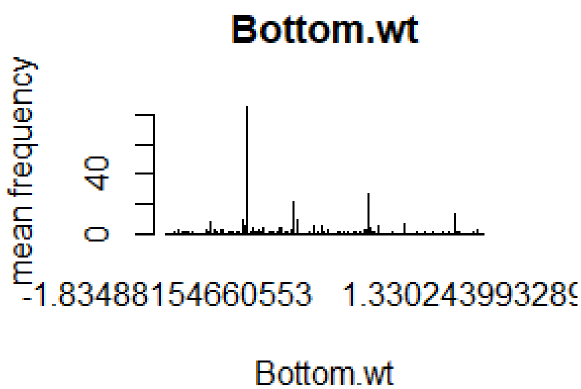
```
plot_obs(new_data2,
  col = "FREQ"
)
```

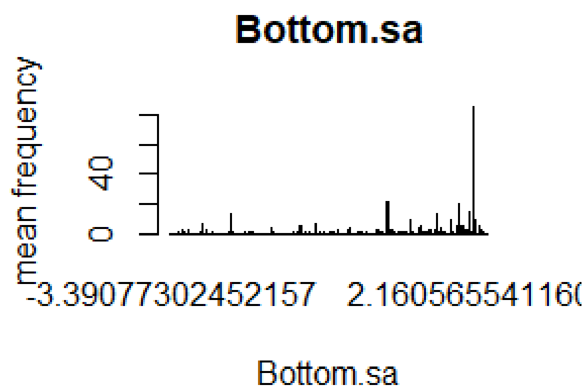

SampleID
mean frequency
0
40
CM20020501037
CM20170501059
SampleID
State
mean frequency
0.00
0.25
MD
VA
State
Year
mean frequency
0.0
0.4
2002
2008
2014
Year
Month
mean frequency
0.0
0.3
JUL
NOV
Month
Date
mean frequency
0
4
8
2002-05-20
2012-07-14
Date
Time
mean frequency
0
40
2002-05-20 12:56:00
Time


Areasw
mean frequency
0
40
6306.50145613056
52186
Areasw
FREQ
mean frequency
0
40
0
4
8
13
22
85
FREQ
Bf.totwght
mean frequency
0
40
-0.102014585281356
Bf.totwght
Stratum
mean frequency
0.0
0.6
101
203
401
502
Stratum
Depth
mean frequency
0
6
-1.62706788118407
Depth
Stratum_area
mean frequency
0.0
0.6
20795799.9293
776073103.347
Stratum_area

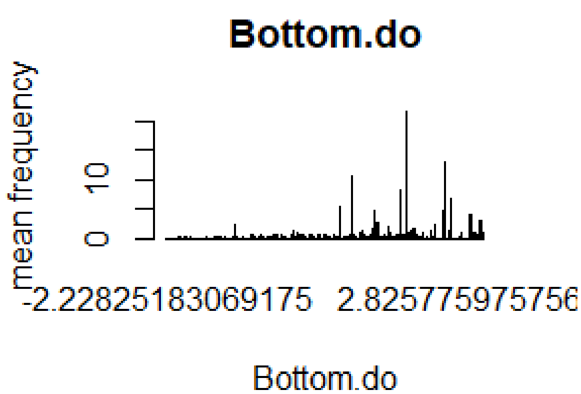









```
## Please manually remove sparsely sampled data!
```



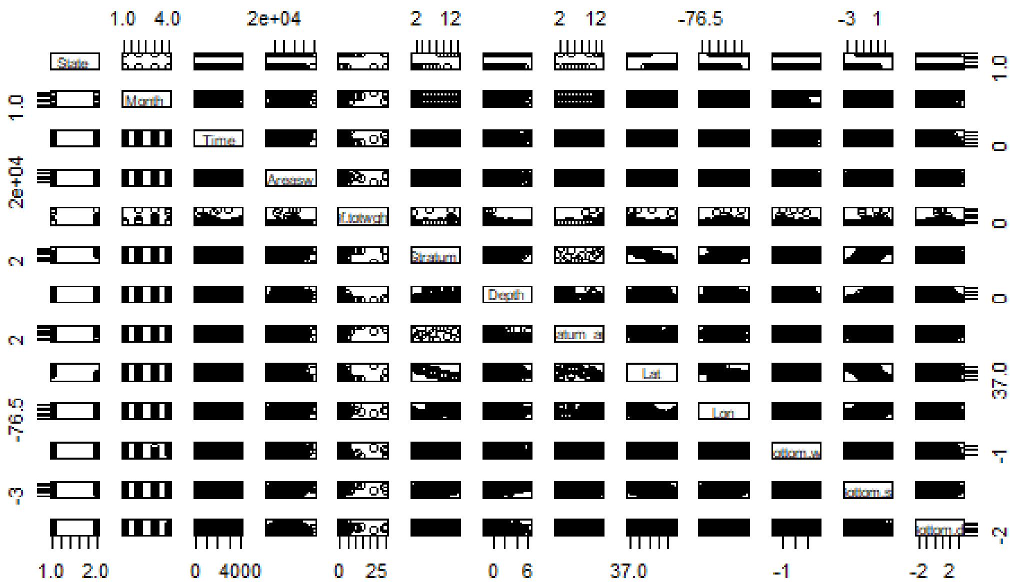
Check colinearity

```
check_colin(new_data2,
  cols = c(
```

```
    "Year", "Date",
    "SampleID", "FREQ"
  )
)
```

```
##              State Month  Time Areasw Bf.totwght Stratum Depth
Stratum_area
## State         1.00 -0.05 -0.01   0.29       0.03    0.86  0.06
0.51
## Month        -0.05  1.00 -0.02   0.09       0.04   -0.04  0.06        -
0.06
## Time         -0.01 -0.02  1.00  -0.12      -0.03   -0.01 -0.02
0.00
## Areasw        0.29  0.09 -0.12   1.00       0.03    0.37  0.15
0.10
## Bf.totwght    0.03  0.04 -0.03   0.03       1.00    0.03 -0.04
0.06
## Stratum       0.86 -0.04 -0.01   0.37       0.03    1.00  0.24
0.34
## Depth         0.06  0.06 -0.02   0.15      -0.04    0.24  1.00        -
0.59
## Stratum_area  0.51 -0.06  0.00   0.10       0.06    0.34 -0.59
1.00
## Lat          -0.85  0.05  0.00  -0.35      -0.04   -0.96 -0.06        -
0.49
## Lon           0.70 -0.03  0.01   0.23       0.02    0.70  0.06
0.42
## Bottom.wt     0.05 -0.04  0.09  -0.04      -0.02    0.00 -0.14
0.11
## Bottom.sa     0.66  0.10  0.12   0.32       0.04    0.79  0.26
0.19
## Bottom.do     0.05  0.14  0.01   0.05       0.06    0.05 -0.21
0.16
##                Lat   Lon Bottom.wt Bottom.sa Bottom.do
## State        -0.85  0.70      0.05      0.66      0.05
## Month         0.05 -0.03     -0.04      0.10      0.14
## Time          0.00  0.01      0.09      0.12      0.01
## Areasw       -0.35  0.23     -0.04      0.32      0.05
## Bf.totwght   -0.04  0.02     -0.02      0.04      0.06
## Stratum      -0.96  0.70      0.00      0.79      0.05
## Depth        -0.06  0.06     -0.14      0.26     -0.21
## Stratum_area -0.49  0.42      0.11      0.19      0.16
## Lat           1.00 -0.73     -0.01     -0.78     -0.10
## Lon          -0.73  1.00      0.04      0.59      0.09
## Bottom.wt    -0.01  0.04      1.00     -0.04     -0.49
## Bottom.sa    -0.78  0.59     -0.04      1.00      0.01
## Bottom.do    -0.10  0.09     -0.49      0.01      1.00

## Stratum is correlated with State
```

```
## Lat is correlated with State

## Lat is correlated with Stratum
```



Remove Stratum and State

```r
new_data2 <- new_data2 %>%
  dplyr::select(-c( # Stratum,
    State
  ))
```

Check variance inflation factor - Remove covariates with GVIF > 3, starting with highest value, until all GVIF < 3

```r
check_vif(
  model = "FREQ ~ Year + Month + Depth + Lat + Lon + Bottom.wt + Bottom.sa + 
Bottom.do + Stratum",
  data = new_data2
)

## # A tibble: 9 x 4
##   Covariate   GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>      <dbl> <dbl>             <dbl>
## 1 Bottom.do   1.94     1              1.39
## 2 Year        2.85    16              1.03
## 3 Lon         3.87     1              1.97
## 4 Depth       5.04     1              2.25
## 5 Bottom.sa   7.07     1              2.66
## 6 Bottom.wt   8.15     1              2.85
## 7 Month      11.0      3              1.49
```

```
## 8 Lat        31.9      1               5.65
## 9 Stratum   288.      13               1.24

## Variance inflation factor too high! Please remove a covariate.

# remove Month

check_vif(
  model = "FREQ ~ Year + Depth + Lat + Lon + Bottom.wt + Bottom.sa +
Bottom.do + Stratum",
  data = new_data2
)

## # A tibble: 8 x 4
##   Covariate   GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>      <dbl> <dbl>             <dbl>
## 1 Bottom.wt   1.66     1              1.29
## 2 Bottom.do   1.71     1              1.31
## 3 Year        2.00    16              1.02
## 4 Lon         3.81     1              1.95
## 5 Depth       5.04     1              2.24
## 6 Bottom.sa   5.70     1              2.39
## 7 Lat        30.9      1              5.56
## 8 Stratum   272.      13              1.24

## Variance inflation factor too high! Please remove a covariate.

# remove Lat

check_vif(
  model = "FREQ ~ Year + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do +
Stratum",
  data = new_data2
)

## # A tibble: 7 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Bottom.wt  1.66     1              1.29
## 2 Bottom.do  1.71     1              1.31
## 3 Year       1.89    16              1.02
## 4 Lon        3.80     1              1.95
## 5 Bottom.sa  4.85     1              2.20
## 6 Depth      4.91     1              2.22
## 7 Stratum   40.6     13              1.15

## Variance inflation factor too high! Please remove a covariate.

### try keeping stratum, remove other covariates
check_vif(
  model = "FREQ ~ Year + Lon + Bottom.wt + Bottom.sa + Bottom.do + Stratum",
```

```
  data = new_data2
)

## # A tibble: 6 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Bottom.wt  1.64     1              1.28
## 2 Bottom.do  1.69     1              1.30
## 3 Year       1.84    16              1.02
## 4 Lon        3.78     1              1.94
## 5 Bottom.sa  4.80     1              2.19
## 6 Stratum   10.5     13              1.09

## Variance inflation factor too high! Please remove a covariate.

check_vif(
  model = "FREQ ~ Year + Lon + Bottom.wt + Bottom.do + Stratum",
  data = new_data2
)

## # A tibble: 5 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Year       1.20    16              1.01
## 2 Bottom.wt  1.60     1              1.27
## 3 Bottom.do  1.66     1              1.29
## 4 Lon        3.40     1              1.84
## 5 Stratum    4.14    13              1.06

## Variance inflation factor too high! Please remove a covariate.

check_vif(
  model = "FREQ ~ Year + Bottom.wt + Bottom.do + Stratum",
  data = new_data2
)

## # A tibble: 4 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Year       1.19    16              1.01
## 2 Stratum    1.26    13              1.01
## 3 Bottom.wt  1.60     1              1.27
## 4 Bottom.do  1.65     1              1.28

## Passed!
```

Step 2: Model selection

Compare negative binom, zero-infl neg binom, zero-altered neg binom

Include swept area offset as proxy for effort:

```r
new_data2$lnSweptArea <- log(new_data2$Areasw)
```

Comparison 1: Full models

```r
compare_models(
  data = new_data2,
  nb_model = "FREQ ~ Year + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do +
offset(lnSweptArea)",

  zic_model = "FREQ ~ Year + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do
+ offset(lnSweptArea)",
  zi_model = "~ Year + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do",

  zac_model = "FREQ ~ Year  + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do
+ offset(lnSweptArea)",
  za_model = "~ Year + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 7.5

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ Year + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do +
##     offset(lnSweptArea)
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2986.2   3131.5  -1470.1   2940.2     4071
##
##
## Dispersion parameter for nbinom2 family (): 0.0734
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -254.4770    60.8092  -4.185 2.85e-05 ***
## Year2003       1.6301     0.4848   3.363 0.000772 ***
## Year2004       0.3299     0.5106   0.646 0.518125
## Year2005       0.3692     0.4760   0.776 0.438031
## Year2006       0.1213     0.5175   0.234 0.814604
## Year2007       1.0296     0.4898   2.102 0.035561 *
## Year2008       1.3590     0.4815   2.823 0.004764 **
## Year2009      -0.1381     0.5799  -0.238 0.811811
## Year2010       0.2722     0.5069   0.537 0.591337
## Year2011      -0.4838     0.7095  -0.682 0.495296
## Year2012      -0.5739     0.5835  -0.984 0.325290
```

```
## Year2013       0.3743     0.5169   0.724 0.468930
## Year2014      -0.3882     0.4937  -0.786 0.431616
## Year2015       0.9525     0.4614   2.064 0.038991 *
## Year2016      -0.4609     0.4946  -0.932 0.351351
## Year2017       0.6906     0.4923   1.403 0.160733
## Year2018       0.9293     0.4685   1.984 0.047294 *
## Depth         -0.7006     0.1089  -6.434 1.24e-10 ***
## Lon           -3.1654     0.7979  -3.967 7.27e-05 ***
## Bottom.wt      0.3192     0.1228   2.600 0.009333 **
## Bottom.sa      0.8638     0.1214   7.117 1.10e-12 ***
## Bottom.do      0.6823     0.1284   5.314 1.07e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.031

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:
## FREQ ~ Year + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do +
##     offset(lnSweptArea)
## Zero inflation:        ~Year + Depth + Lon + Bottom.wt + Bottom.sa +
Bottom.do
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2938.2   3222.5  -1424.1   2848.2     4049
##
##
## Dispersion parameter for truncated_nbinom2 family (): 5.9e-07
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -151.2467   225.6666  -0.670 0.502716
## Year2003        1.1578     0.7333   1.579 0.114376
## Year2004        0.5991     0.8755   0.684 0.493827
## Year2005        1.0373     0.7687   1.349 0.177185
## Year2006       -0.4040     0.9635  -0.419 0.674968
## Year2007        1.0743     0.7615   1.411 0.158352
## Year2008        2.7911     0.8983   3.107 0.001890 **
## Year2009        0.2798     1.1999   0.233 0.815629
## Year2010        2.8281     1.1980   2.361 0.018246 *
## Year2011      -20.2813 33899.4190  -0.001 0.999523
## Year2012        0.2733     1.0771   0.254 0.799676
```

```
## Year2013        1.7902     0.9646   1.856 0.063459 .
## Year2014       -0.2423     0.8148  -0.297 0.766185
## Year2015        3.2694     0.8724   3.748 0.000178 ***
## Year2016        1.2827     0.9466   1.355 0.175387
## Year2017        2.2511     0.9897   2.275 0.022936 *
## Year2018        1.6382     0.7688   2.131 0.033092 *
## Depth          -1.2516     0.2337  -5.356 8.51e-08 ***
## Lon            -1.6456     1.5764  -1.044 0.296532
## Bottom.wt      -0.4714     0.2899  -1.626 0.103973
## Bottom.sa       0.2728     0.2272   1.201 0.229831
## Bottom.do       0.3775     0.3051   1.237 0.216060
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept) 147.98659   49.93822   2.963 0.003043 **
## Year2003     -1.54082    0.34173  -4.509 6.52e-06 ***
## Year2004     -0.26217    0.38279  -0.685 0.493414
## Year2005     -0.35101    0.35048  -1.002 0.316579
## Year2006     -0.28549    0.37835  -0.755 0.450516
## Year2007     -0.79853    0.34719  -2.300 0.021450 *
## Year2008     -0.13791    0.38066  -0.362 0.717124
## Year2009      0.10667    0.44384   0.240 0.810073
## Year2010      1.18054    0.53256   2.217 0.026642 *
## Year2011      0.47097    0.53667   0.878 0.380177
## Year2012      0.06287    0.41492   0.152 0.879570
## Year2013     -0.01106    0.37643  -0.029 0.976568
## Year2014     -0.02276    0.35454  -0.064 0.948818
## Year2015      0.08702    0.34778   0.250 0.802429
## Year2016      0.63822    0.37526   1.701 0.088993 .
## Year2017     -0.04691    0.37759  -0.124 0.901126
## Year2018     -0.49204    0.36013  -1.366 0.171851
## Depth         0.31301    0.08055   3.886 0.000102 ***
## Lon           1.90436    0.65476   2.909 0.003632 **
## Bottom.wt    -0.15493    0.07953  -1.948 0.051396 .
## Bottom.sa    -0.70319    0.09251  -7.601 2.93e-14 ***
## Bottom.do    -0.44164    0.08387  -5.266 1.40e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Warning in sqrt(pvar): NaNs produced

## Dispersal: NaN

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class
"c('simpleWarning', 'warning', 'condition')"

compare_models(
  data = new_data2,
```

```
  nb_model = "FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea)",

  zic_model = "FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea)",
  zi_model = "~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea)",

  zac_model = "FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea)",
  za_model = "~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea)",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 7.5

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2938.8   3147.3  -1436.4   2872.8     4061
##
##
## Dispersion parameter for nbinom2 family (): 0.0857
##
## Conditional model:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept) -12.47720    0.48696 -25.623  < 2e-16 ***
## Year2003      1.26602    0.46171   2.742  0.00611 **
## Year2004     -0.15461    0.49255  -0.314  0.75360
## Year2005      0.56026    0.45830   1.222  0.22153
## Year2006     -0.15741    0.51456  -0.306  0.75967
## Year2007      1.11898    0.47078   2.377  0.01746 *
## Year2008      0.89992    0.46964   1.916  0.05534 .
## Year2009     -0.26665    0.58322  -0.457  0.64752
## Year2010      0.38569    0.48681   0.792  0.42819
## Year2011     -0.78097    0.68976  -1.132  0.25754
## Year2012     -0.49272    0.57346  -0.859  0.39023
## Year2013      0.41594    0.50320   0.827  0.40847
## Year2014     -0.05131    0.47383  -0.108  0.91376
## Year2015      0.81676    0.44826   1.822  0.06845 .
```

```
## Year2016      0.27507    0.48298   0.570  0.56900
## Year2017      0.72237    0.48438   1.491  0.13587
## Year2018      0.11415    0.46319   0.246  0.80533
## Stratum102   -2.94235    1.12799  -2.608  0.00909 **
## Stratum201   -0.65559    0.47265  -1.387  0.16542
## Stratum202   -1.06698    0.50011  -2.134  0.03288 *
## Stratum203   -1.49414    0.77892  -1.918  0.05508 .
## Stratum301   -1.99857    0.46359  -4.311 1.62e-05 ***
## Stratum302   -2.03724    0.51603  -3.948 7.88e-05 ***
## Stratum303   -2.00855    0.82826  -2.425  0.01531 *
## Stratum401   -0.85335    0.38572  -2.212  0.02694 *
## Stratum402   -1.12504    0.38451  -2.926  0.00343 **
## Stratum403   -1.55985    0.57026  -2.735  0.00623 **
## Stratum501    0.95260    0.38244   2.491  0.01274 *
## Stratum502   -0.21168    0.38215  -0.554  0.57965
## Stratum503   -1.87744    0.58868  -3.189  0.00143 **
## Bottom.wt     0.27397    0.11549   2.372  0.01768 *
## Bottom.do     0.59632    0.12345   4.830 1.36e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.374

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
## Zero inflation:
## ~Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2964.9   3375.5  -1417.4   2834.9     4029
##
##
## Dispersion parameter for truncated_nbinom2 family (): 4.19e-09
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.855e+01  2.633e+03  -0.011 0.991349
## Year2003     9.994e-01  7.735e-01   1.292 0.196334
## Year2004     9.660e-02  9.690e-01   0.100 0.920585
## Year2005     8.984e-01  7.974e-01   1.127 0.259902
## Year2006    -7.327e-01  1.037e+00  -0.707 0.479873
```

```
## Year2007     1.301e+00  8.321e-01   1.563 0.118055
## Year2008     1.807e+00  1.015e+00   1.781 0.074929 .
## Year2009     1.811e-01  1.403e+00   0.129 0.897247
## Year2010     2.684e+00  1.211e+00   2.216 0.026708 *
## Year2011    -1.773e+01  6.288e+03  -0.003 0.997750
## Year2012    -7.196e-02  1.126e+00  -0.064 0.949028
## Year2013     1.086e+00  1.052e+00   1.032 0.302075
## Year2014    -2.604e-01  8.524e-01  -0.305 0.760007
## Year2015     2.517e+00  9.354e-01   2.691 0.007131 **
## Year2016     1.334e+00  1.035e+00   1.289 0.197405
## Year2017     1.988e+00  1.117e+00   1.779 0.075190 .
## Year2018     9.505e-01  8.418e-01   1.129 0.258877
## Stratum102  -2.766e+01  7.623e+05   0.000 0.999971
## Stratum201  -7.130e-01  1.201e+00  -0.594 0.552778
## Stratum202  -8.940e-01  1.166e+00  -0.767 0.443370
## Stratum203  -3.245e+01  1.000e+06   0.000 0.999974
## Stratum301  -1.865e+00  1.090e+00  -1.711 0.087139 .
## Stratum302  -3.834e+00  1.270e+00  -3.019 0.002540 **
## Stratum303  -3.312e+01  1.415e+06   0.000 0.999981
## Stratum401  -2.246e+00  8.962e-01  -2.506 0.012212 *
## Stratum402  -3.278e+00  8.953e-01  -3.661 0.000251 ***
## Stratum403  -4.682e+00  1.479e+00  -3.166 0.001545 **
## Stratum501  -8.684e-01  8.735e-01  -0.994 0.320103
## Stratum502  -1.732e+00  8.450e-01  -2.050 0.040384 *
## Stratum503  -3.704e+00  1.472e+00  -2.517 0.011848 *
## Bottom.wt   -2.569e-01  3.115e-01  -0.825 0.409655
## Bottom.do    7.150e-01  3.837e-01   1.863 0.062417 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -7.339485   0.417509 -17.579  < 2e-16 ***
## Year2003    -1.087229   0.334286  -3.252  0.00114 **
## Year2004     0.196090   0.383061   0.512  0.60872
## Year2005    -0.248657   0.355137  -0.700  0.48382
## Year2006     0.184364   0.379424   0.486  0.62703
## Year2007    -0.356038   0.349401  -1.019  0.30821
## Year2008     0.454263   0.383732   1.184  0.23649
## Year2009     0.730056   0.447320   1.632  0.10267
## Year2010     1.350140   0.535336   2.522  0.01167 *
## Year2011     1.006513   0.539153   1.867  0.06192 .
## Year2012     0.044861   0.417435   0.107  0.91442
## Year2013     0.031903   0.378773   0.084  0.93288
## Year2014     0.029556   0.357667   0.083  0.93414
## Year2015    -0.041760   0.350393  -0.119  0.90513
## Year2016     0.390035   0.375731   1.038  0.29924
## Year2017     0.188099   0.379177   0.496  0.61984
## Year2018     0.054268   0.358121   0.152  0.87955
## Stratum102   1.693244   1.055134   1.605  0.10855
```

```
## Stratum201  -0.008808   0.466213  -0.019  0.98493
## Stratum202  -0.195278   0.483681  -0.404  0.68641
## Stratum203  -0.416652   0.613171  -0.680  0.49682
## Stratum301   0.796115   0.463199   1.719  0.08566 .
## Stratum302   0.062036   0.465684   0.133  0.89402
## Stratum303  -0.298211   0.674356  -0.442  0.65833
## Stratum401  -0.715589   0.355567  -2.013  0.04416 *
## Stratum402  -0.862120   0.348881  -2.471  0.01347 *
## Stratum403  -0.638395   0.485545  -1.315  0.18858
## Stratum501  -1.745390   0.333524  -5.233 1.67e-07 ***
## Stratum502  -1.408687   0.346852  -4.061 4.88e-05 ***
## Stratum503   0.087000   0.555864   0.157  0.87563
## Bottom.wt   -0.107498   0.080628  -1.333  0.18244
## Bottom.do   -0.442557   0.086422  -5.121 3.04e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Warning in sqrt(pvar): NaNs produced

## Dispersal: NaN

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class
"c('simpleWarning', 'warning', 'condition')"
```

Zero models did not work - zero-altered did not have a dispersal. All covariates were statistically significant in the negative binomial model… try a null count model and a null binomial model.

Comparison 2: Null binomials

```
compare_models(
  data = new_data2,
  models = c("zi", "za"),

  zic_model = "FREQ ~ Year + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do
+ offset(lnSweptArea)",
  zi_model = "~ 1",

  zac_model = "FREQ ~ Year + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do
+ offset(lnSweptArea)",
  za_model = "~ 1",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 7.5

## Zero-inflated negative binomial

## Model did not work
```

```
## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"

# FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
compare_models(
  data = new_data2,
  models = c("zi", "za"),

  zic_model = "FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea)",
  zi_model = "~ 1",

  zac_model = "FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea)",
  za_model = "~ 1",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 7.5

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
## Zero inflation:        ~1
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2940.8   3155.6  -1436.4   2872.8     4060
##
##
## Dispersion parameter for nbinom2 family (): 0.0857
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -12.47723    0.48696 -25.623  < 2e-16 ***
## Year2003      1.26601    0.46171   2.742  0.00611 **
## Year2004     -0.15460    0.49255  -0.314  0.75362
## Year2005      0.56024    0.45830   1.222  0.22155
## Year2006     -0.15740    0.51456  -0.306  0.75968
## Year2007      1.11899    0.47078   2.377  0.01746 *
## Year2008      0.89993    0.46964   1.916  0.05534 .
```

```
## Year2009     -0.26664    0.58322  -0.457  0.64753
## Year2010      0.38572    0.48681   0.792  0.42816
## Year2011     -0.78096    0.68976  -1.132  0.25755
## Year2012     -0.49270    0.57346  -0.859  0.39025
## Year2013      0.41594    0.50320   0.827  0.40847
## Year2014     -0.05129    0.47383  -0.108  0.91380
## Year2015      0.81677    0.44826   1.822  0.06844 .
## Year2016      0.27508    0.48298   0.570  0.56898
## Year2017      0.72238    0.48438   1.491  0.13587
## Year2018      0.11417    0.46319   0.246  0.80531
## Stratum102   -2.94233    1.12799  -2.608  0.00909 **
## Stratum201   -0.65556    0.47265  -1.387  0.16544
## Stratum202   -1.06697    0.50011  -2.133  0.03288 *
## Stratum203   -1.49411    0.77892  -1.918  0.05509 .
## Stratum301   -1.99858    0.46359  -4.311 1.62e-05 ***
## Stratum302   -2.03725    0.51603  -3.948 7.88e-05 ***
## Stratum303   -2.00854    0.82826  -2.425  0.01531 *
## Stratum401   -0.85334    0.38572  -2.212  0.02695 *
## Stratum402   -1.12502    0.38451  -2.926  0.00344 **
## Stratum403   -1.55985    0.57026  -2.735  0.00623 **
## Stratum501    0.95261    0.38244   2.491  0.01274 *
## Stratum502   -0.21166    0.38215  -0.554  0.57966
## Stratum503   -1.87740    0.58868  -3.189  0.00143 **
## Bottom.wt     0.27396    0.11549   2.372  0.01769 *
## Bottom.do     0.59631    0.12345   4.830 1.36e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -17.74    2948.84  -0.006    0.995

## Dispersal: 1.375

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Still no good Zero inflated works with Stratum

Comparison 3: Null counts

```
compare_models(
  data = new_data2,
  models = c("zi", "za"),

  zic_model = "FREQ ~ offset(lnSweptArea)",
  zi_model = "~ Year + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do",
```

```
  zac_model = "FREQ ~ offset(lnSweptArea)",
  za_model = "~ Year + Depth + Lon + Bottom.wt + Bottom.sa + Bottom.do",

  r_dat = new_data2$FREQ
) %>%
  try()
```

```
## Percent positive tows: 7.5

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ offset(lnSweptArea)
## Zero inflation:        ~Year + Depth + Lon + Bottom.wt + Bottom.sa +
Bottom.do
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   3054.8   3206.4  -1503.4   3006.8     4070
##
##
## Dispersion parameter for nbinom2 family (): 0.103
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -11.3697     0.1592  -71.44   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) 359.53063  103.38768   3.477 0.000506 ***
## Year2003     -3.30323    1.19187  -2.771 0.005580 **
## Year2004     -0.29004    0.94298  -0.308 0.758401
## Year2005      0.31105    0.98442   0.316 0.752025
## Year2006     -0.26696    0.92753  -0.288 0.773483
## Year2007     -1.15753    1.00276  -1.154 0.248361
## Year2008     -0.33903    0.95069  -0.357 0.721380
## Year2009      0.60699    1.03571   0.586 0.557838
## Year2010      2.44170    1.16998   2.087 0.036892 *
## Year2011      0.83112    1.13028   0.735 0.462141
## Year2012      1.09466    1.14083   0.960 0.337294
## Year2013      0.08217    1.00950   0.081 0.935128
## Year2014      0.75397    1.02032   0.739 0.459938
## Year2015      0.93721    1.03452   0.906 0.364969
## Year2016      1.58974    1.05008   1.514 0.130046
## Year2017      0.14641    0.95521   0.153 0.878178
## Year2018     -0.42436    0.92903  -0.457 0.647829
## Depth         0.59194    0.15478   3.824 0.000131 ***
```

```
## Lon           4.71556    1.36135   3.464 0.000532 ***
## Bottom.wt    -0.60401    0.23671  -2.552 0.010719 *
## Bottom.sa    -1.45590    0.28662  -5.080 3.78e-07 ***
## Bottom.do    -0.88565    0.21675  -4.086 4.39e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 3.316

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ offset(lnSweptArea)
## Zero inflation:        ~Year + Depth + Lon + Bottom.wt + Bottom.sa +
Bottom.do
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   3038.9   3190.6  -1495.5   2990.9     4070
##
##
## Dispersion parameter for truncated_nbinom2 family (): 1.75e-21
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -56.77   29242.27  -0.002    0.998
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) 147.95473   47.50462   3.115  0.00184 **
## Year2003     -1.54077    0.33813  -4.557 5.20e-06 ***
## Year2004     -0.26214    0.38240  -0.686  0.49302
## Year2005     -0.35100    0.35024  -1.002  0.31626
## Year2006     -0.28546    0.37796  -0.755  0.45009
## Year2007     -0.79851    0.34636  -2.305  0.02114 *
## Year2008     -0.13789    0.38045  -0.362  0.71703
## Year2009      0.10668    0.44381   0.240  0.81003
## Year2010      1.18055    0.53269   2.216  0.02668 *
## Year2011      0.47099    0.53675   0.877  0.38023
## Year2012      0.06288    0.41492   0.152  0.87954
## Year2013     -0.01105    0.37640  -0.029  0.97659
## Year2014     -0.02275    0.35452  -0.064  0.94883
## Year2015      0.08701    0.34774   0.250  0.80243
## Year2016      0.63822    0.37500   1.702  0.08877 .
## Year2017     -0.04690    0.37751  -0.124  0.90114
## Year2018     -0.49201    0.35935  -1.369  0.17094
## Depth         0.31300    0.08036   3.895 9.82e-05 ***
## Lon           1.90395    0.62304   3.056  0.00224 **
## Bottom.wt    -0.15493    0.07928  -1.954  0.05067 .
```

```
## Bottom.sa    -0.70315    0.08741  -8.044 8.68e-16 ***
## Bottom.do    -0.44163    0.08324  -5.305 1.12e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: NaN

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

```
# ~ Year + Stratum + Bottom.wt + Bottom.do
compare_models(
  data = new_data2,
  models = c("zi", "za"),

  zic_model = "FREQ ~ offset(lnSweptArea)",
  zi_model = "~ Year + Stratum + Bottom.wt + Bottom.do",

  zac_model = "FREQ ~ offset(lnSweptArea)",
  za_model = "~ Year + Stratum + Bottom.wt + Bottom.do",

  r_dat = new_data2$FREQ
) %>%
  try()
```

```
## Percent positive tows: 7.5

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ offset(lnSweptArea)
## Zero inflation:        ~Year + Stratum + Bottom.wt + Bottom.do
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   3027.3   3242.0  -1479.6   2959.3     4060
##
##
## Dispersion parameter for nbinom2 family (): 0.0997
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -11.4009     0.1059  -107.6   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.13500    0.84904   0.159 0.873664
```

```
## Year2003      -2.40316    1.10691  -2.171 0.029927 *
## Year2004       0.48004    0.80436   0.597 0.550642
## Year2005       0.03160    0.79445   0.040 0.968274
## Year2006       0.24984    0.84263   0.296 0.766851
## Year2007      -1.70076    0.96935  -1.755 0.079339 .
## Year2008       0.40108    0.82589   0.486 0.627226
## Year2009       0.72153    0.90948   0.793 0.427575
## Year2010       3.37143    1.31741   2.559 0.010493 *
## Year2011       1.34273    0.98911   1.358 0.174615
## Year2012       1.48592    0.95795   1.551 0.120867
## Year2013       0.25088    0.82343   0.305 0.760612
## Year2014       0.62121    0.80761   0.769 0.441774
## Year2015       0.49645    0.79993   0.621 0.534851
## Year2016       0.55117    0.80802   0.682 0.495162
## Year2017       0.11448    0.80335   0.143 0.886685
## Year2018       1.37922    0.86764   1.590 0.111921
## Stratum102     2.43452    1.35521   1.796 0.072429 .
## Stratum201     0.50225    0.78991   0.636 0.524888
## Stratum202     0.71904    0.76688   0.938 0.348446
## Stratum203    -0.13131    1.03296  -0.127 0.898848
## Stratum301     1.99317    0.75687   2.633 0.008452 **
## Stratum302     1.13208    0.76997   1.470 0.141483
## Stratum303     0.73266    1.07258   0.683 0.494557
## Stratum401    -0.23901    0.66333  -0.360 0.718609
## Stratum402    -0.02568    0.62772  -0.041 0.967373
## Stratum403     0.23588    0.80476   0.293 0.769438
## Stratum501   -19.72544 2260.84788  -0.009 0.993039
## Stratum502    -1.53351    0.79375  -1.932 0.053363 .
## Stratum503     0.99499    0.86139   1.155 0.248048
## Bottom.wt     -0.75616    0.20900  -3.618 0.000297 ***
## Bottom.do     -0.80577    0.21200  -3.801 0.000144 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.42

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ offset(lnSweptArea)
## Zero inflation:        ~Year + Stratum + Bottom.wt + Bottom.do
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   3017.8   3232.6  -1474.9   2949.8     4060
##
##
## Dispersion parameter for truncated_nbinom2 family (): 1.1e-09
##
```

```
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -29.6     2825.9   -0.01    0.992
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.997640   0.413079   7.257 3.96e-13 ***
## Year2003    -0.981730   0.332631  -2.951  0.00316 **
## Year2004     0.213691   0.381558   0.560  0.57545
## Year2005    -0.273855   0.352408  -0.777  0.43710
## Year2006     0.089271   0.377637   0.236  0.81313
## Year2007    -0.580099   0.347245  -1.671  0.09481 .
## Year2008     0.172003   0.380717   0.452  0.65142
## Year2009     0.443069   0.445202   0.995  0.31963
## Year2010     1.167230   0.533525   2.188  0.02869 *
## Year2011     0.780584   0.537325   1.453  0.14630
## Year2012     0.055460   0.416383   0.133  0.89404
## Year2013    -0.019210   0.377711  -0.051  0.95944
## Year2014     0.014613   0.356512   0.041  0.96731
## Year2015    -0.130550   0.348854  -0.374  0.70824
## Year2016     0.294915   0.374484   0.788  0.43097
## Year2017     0.113398   0.378340   0.300  0.76439
## Year2018     0.006075   0.357268   0.017  0.98643
## Stratum102   1.690888   1.054788   1.603  0.10892
## Stratum201   0.196337   0.465084   0.422  0.67291
## Stratum202   0.188362   0.481408   0.391  0.69560
## Stratum203  -0.150884   0.609761  -0.247  0.80456
## Stratum301   0.957564   0.462960   2.068  0.03861 *
## Stratum302   0.430062   0.464793   0.925  0.35482
## Stratum303   0.228479   0.674034   0.339  0.73463
## Stratum401  -0.380063   0.354758  -1.071  0.28402
## Stratum402  -0.423195   0.348060  -1.216  0.22404
## Stratum403  -0.147766   0.484480  -0.305  0.76037
## Stratum501  -1.296452   0.333266  -3.890  0.00010 ***
## Stratum502  -0.887839   0.346085  -2.565  0.01031 *
## Stratum503   0.519324   0.553783   0.938  0.34836
## Bottom.wt   -0.094540   0.080180  -1.179  0.23836
## Bottom.do   -0.411748   0.086148  -4.780 1.76e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 5.028

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Zero-inflated worked Zero adjusted works with Stratum

Comparison 4: Year only

```
compare_models(
  data = new_data2,
  models = c("zi", "za"),

  zic_model = "FREQ ~ Year + offset(lnSweptArea)",
  zi_model = "~ Year",

  zac_model = "FREQ ~ Year + offset(lnSweptArea)",
  za_model = "~ Year",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 7.5

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ Year + offset(lnSweptArea)
## Zero inflation:        ~Year
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   3074.3   3295.4  -1502.2   3004.3     4059
##
##
## Dispersion parameter for nbinom2 family (): 0.108
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -12.56134    0.51322 -24.476  < 2e-16 ***
## Year2003      0.98215    0.56068   1.752 0.079821 .
## Year2004     -0.18359    0.70037  -0.262 0.793216
## Year2005      1.60140    0.61686   2.596 0.009430 **
## Year2006     -0.66020    0.58704  -1.125 0.260748
## Year2007      0.54149    0.56806   0.953 0.340471
## Year2008      2.07196    0.69062   3.000 0.002698 **
## Year2009     -0.91668    0.63336  -1.447 0.147804
## Year2010      3.71836    1.03077   3.607 0.000309 ***
## Year2011     -1.41743    0.72489  -1.955 0.050540 .
## Year2012     -0.56633    0.61485  -0.921 0.357002
## Year2013      0.22164    0.67863   0.327 0.743973
## Year2014      0.09226    0.63721   0.145 0.884876
## Year2015      2.38199    0.62704   3.799 0.000145 ***
## Year2016      0.41838    0.67923   0.616 0.537927
## Year2017      0.65049    0.66383   0.980 0.327131
```

```
## Year2018      1.33145    0.62526   2.129 0.033217 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.6699     2.6051  -0.641    0.522
## Year2003      -20.7646 61717.5203   0.000    1.000
## Year2004        0.7504     2.8615   0.262    0.793
## Year2005        1.6797     2.5790   0.651    0.515
## Year2006      -17.8001 12911.3933  -0.001    0.999
## Year2007      -16.2184  3844.4190  -0.004    0.997
## Year2008        2.7021     2.5875   1.044    0.296
## Year2009      -14.1191  4682.1834  -0.003    0.998
## Year2010        4.1599     2.6199   1.588    0.112
## Year2011      -18.4944 35384.6034   0.000    1.000
## Year2012      -15.2499  5901.8752  -0.003    0.998
## Year2013        0.9885     2.7265   0.363    0.717
## Year2014        0.4256     2.8542   0.149    0.881
## Year2015        2.3667     2.5751   0.919    0.358
## Year2016        1.5775     2.6369   0.598    0.550
## Year2017        1.8559     2.6107   0.711    0.477
## Year2018        1.8159     2.5854   0.702    0.482

## Dispersal: 1.425

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ Year + offset(lnSweptArea)
## Zero inflation:        ~Year
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   3082.5   3303.6  -1506.3   3012.5     4059
##
##
## Dispersion parameter for truncated_nbinom2 family (): 2.63e-09
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.992e+01  2.435e+03  -0.012 0.990198
## Year2003     8.968e-01  6.203e-01   1.446 0.148222
## Year2004    -2.111e-01  7.460e-01  -0.283 0.777183
## Year2005     1.637e+00  6.716e-01   2.438 0.014786 *
## Year2006    -1.083e+00  7.990e-01  -1.356 0.175127
## Year2007     3.648e-01  6.537e-01   0.558 0.576824
## Year2008     2.015e+00  7.451e-01   2.704 0.006857 **
## Year2009    -1.218e+00  1.036e+00  -1.176 0.239611
```

```
## Year2010      3.903e+00  1.215e+00   3.213 0.001315 **
## Year2011     -2.016e+01  1.370e+04  -0.001 0.998826
## Year2012     -7.881e-01  8.442e-01  -0.934 0.350503
## Year2013      2.088e-01  7.288e-01   0.286 0.774534
## Year2014      9.895e-02  6.851e-01   0.144 0.885164
## Year2015      2.492e+00  6.911e-01   3.606 0.000311 ***
## Year2016      4.167e-01  7.299e-01   0.571 0.568091
## Year2017      6.937e-01  7.201e-01   0.963 0.335409
## Year2018      1.442e+00  6.872e-01   2.099 0.035845 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.42775    0.26935   9.013  < 2e-16 ***
## Year2003    -0.87637    0.32191  -2.722  0.00648 **
## Year2004     0.29272    0.37305   0.785  0.43264
## Year2005    -0.25300    0.34219  -0.739  0.45970
## Year2006     0.26050    0.36793   0.708  0.47894
## Year2007    -0.51821    0.33698  -1.538  0.12410
## Year2008     0.34484    0.37277   0.925  0.35492
## Year2009     0.68823    0.43427   1.585  0.11301
## Year2010     1.16682    0.52730   2.213  0.02691 *
## Year2011     1.04421    0.52798   1.978  0.04796 *
## Year2012     0.18109    0.40251   0.450  0.65278
## Year2013     0.19008    0.36833   0.516  0.60582
## Year2014     0.06471    0.34895   0.185  0.85289
## Year2015    -0.04820    0.34096  -0.141  0.88758
## Year2016     0.34851    0.36747   0.948  0.34293
## Year2017     0.36666    0.36738   0.998  0.31826
## Year2018    -0.01346    0.34938  -0.039  0.96927
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: NaN

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Lowest dispersion for ZI model, so using for ZI and ZA.

Model selection for negative binomial model

```
model <- glmmTMB::glmmTMB(FREQ ~ Year + Lon + Bottom.wt + Bottom.sa +
Bottom.do + Depth + offset(lnSweptArea),
  data = new_data2,
  family = nbinom2
)

null <- glmmTMB::glmmTMB(FREQ ~ Year + offset(lnSweptArea),
  data = new_data2,
```

```
  family = nbinom2
)

select_model(model = model, null_model = null)

## Forward model selection:

## Stepwise Model Path 
## Analysis of Deviance Table
## 
## Initial Model:
## FREQ ~ Year + offset(lnSweptArea)
## 
## Final Model:
## FREQ ~ Year + Depth + Bottom.sa + Bottom.do + Lon + Bottom.wt + 
##     offset(lnSweptArea)
## 
## 
##          Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                               4076   3070.320 3106.320
## 2     + Depth  1 48.828360      4075   3021.491 3059.491
## 3 + Bottom.sa  1 42.688317      4074   2978.803 3018.803
## 4 + Bottom.do  1 18.828933      4073   2959.974 3001.974
## 5       + Lon  1 12.841193      4072   2947.133 2991.133
## 6 + Bottom.wt  1  6.962388      4071   2940.171 2986.171

## Bidirectional model selection:

## Stepwise Model Path 
## Analysis of Deviance Table
## 
## Initial Model:
## FREQ ~ Year + Lon + Bottom.wt + Bottom.sa + Bottom.do + Depth + 
##     offset(lnSweptArea)
## 
## Final Model:
## FREQ ~ Year + Lon + Bottom.wt + Bottom.sa + Bottom.do + Depth + 
##     offset(lnSweptArea)
## 
## 
##   Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                       4071   2940.171 2986.171

## Backward model selection:

## Stepwise Model Path 
## Analysis of Deviance Table
## 
## Initial Model:
## FREQ ~ Year + Lon + Bottom.wt + Bottom.sa + Bottom.do + Depth + 
##     offset(lnSweptArea)
```

```
##
## Final Model:
## FREQ ~ Year + Lon + Bottom.wt + Bottom.sa + Bottom.do + Depth +
##     offset(lnSweptArea)
##
##
##   Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                        4071   2940.171 2986.171
```

```
#
model <- glmmTMB::glmmTMB(FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea),
  data = new_data2,
  family = nbinom2
)

null <- glmmTMB::glmmTMB(FREQ ~ Year + Stratum + offset(lnSweptArea),
  data = new_data2,
  family = nbinom2
)

select_model(model = model, null_model = null)
```

```
## Forward model selection:
```

```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ Year + Stratum + offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Stratum + Bottom.do + Bottom.wt + offset(lnSweptArea)
##
##
##          Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                              4063   2897.772 2959.772
## 2 + Bottom.do  1 19.222573      4062   2878.549 2942.549
## 3 + Bottom.wt  1  5.757028      4061   2872.792 2938.792
```

```
## Bidirectional model selection:
```

```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
##
```

```
## 
##   Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                        4061   2872.792 2938.792

## Backward model selection:

## Stepwise Model Path 
## Analysis of Deviance Table
## 
## Initial Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
## 
## Final Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
## 
## 
##   Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                        4061   2872.792 2938.792
```

Model selection for zero-inflated model

```
model <- glmmTMB::glmmTMB(FREQ ~ Year + Lon + Bottom.wt + Bottom.sa + 
Bottom.do + Depth + offset(lnSweptArea),
  ziformula = ~Year,
  data = new_data2,
  family = nbinom2
)

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

null <- glmmTMB::glmmTMB(FREQ ~ offset(lnSweptArea),
  ziformula = ~Year,
  data = new_data2,
  family = nbinom2
)

select_model(model = model, null_model = null) %>%
  try()

## Forward model selection:

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Bottom.do + Depth + Bottom.sa + offset(lnSweptArea)
##
##
##          Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                               4075   3105.385 3143.385
## 2      + Year 16 101.06162      4059   3004.324 3074.324
## 3 + Bottom.do  1  48.52337      4058   2955.800 3027.800
## 4     + Depth  1  28.18870      4057   2927.611 3001.611
## 5 + Bottom.sa  1  34.54957      4056   2893.062 2969.062

## Bidirectional model selection:

## Error in MASS::stepAIC(model, direction = "both", scope = list(upper =
model$call,  :
##   AIC is not defined for this model, so 'stepAIC' cannot proceed

#

model <- glmmTMB::glmmTMB(FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea),
  ziformula = ~1,
  data = new_data2,
  family = nbinom2
)

null <- glmmTMB::glmmTMB(FREQ ~ Year + Stratum + offset(lnSweptArea),
  ziformula = ~1,
  data = new_data2,
```

```
  family = nbinom2
)

select_model(model = model, null_model = null) %>%
  try()

## Forward model selection:

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ Year + Stratum + offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Stratum + Bottom.do + Bottom.wt + offset(lnSweptArea)
##
##
##          Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                                4062   2897.772 2961.772
## 2 + Bottom.do  1 19.222573      4061   2878.549 2944.549
## 3 + Bottom.wt  1  5.757028      4060   2872.792 2940.792

## Bidirectional model selection:

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
##
##
##   Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                       4060   2872.792 2940.792

## Backward model selection:

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
##
##
```

```
##   Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                       4060   2872.792 2940.792
```

Model selection for zero-adjusted model

```
model <- glmmTMB::glmmTMB(FREQ ~ Year + Lon + Bottom.wt + Bottom.sa +
Bottom.do + Depth + offset(lnSweptArea),
  ziformula = ~Year,
  data = new_data2,
  family = truncated_nbinom2(link = "log")
)

null <- glmmTMB::glmmTMB(FREQ ~ offset(lnSweptArea),
  ziformula = ~Year,
  data = new_data2,
  family = truncated_nbinom2(link = "log")
)

select_model(model = model, null_model = null) %>%
  try()

## Forward model selection:

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Depth + Bottom.wt + offset(lnSweptArea)
##
##
##          Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                              4075   3109.338 3147.338
## 2      + Year 16 96.83566      4059   3012.502 3082.502
## 3     + Depth  1 33.03263      4058   2979.470 3051.470
## 4 + Bottom.wt  1 10.04593      4057   2969.424 3043.424

## Bidirectional model selection:

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ Year + Lon + Bottom.wt + Bottom.sa + Bottom.do + Depth +
##     offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Bottom.wt + Depth + offset(lnSweptArea)
##
##
##          Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                              4054   2966.640 3046.640
## 2       - Lon  1 1.0769118      4055   2967.716 3045.716
## 3 - Bottom.sa  1 0.4833284      4056   2968.200 3044.200
## 4 - Bottom.do  1 1.2240600      4057   2969.424 3043.424

## Backward model selection:

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
```

```
## FREQ ~ Year + Lon + Bottom.wt + Bottom.sa + Bottom.do + Depth +
##     offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Bottom.wt + Depth + offset(lnSweptArea)
##
##
##          Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                               4054   2966.640 3046.640
## 2       - Lon  1 1.0769118      4055   2967.716 3045.716
## 3 - Bottom.sa  1 0.4833284      4056   2968.200 3044.200
## 4 - Bottom.do  1 1.2240600      4057   2969.424 3043.424
```

```
#

model <- glmmTMB::glmmTMB(FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea),
  ziformula = ~ Year + Stratum + Bottom.wt + Bottom.do,
  data = new_data2,
  family = truncated_nbinom2(link = "log")
)
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')
```

```
null <- glmmTMB::glmmTMB(FREQ ~ offset(lnSweptArea),
  ziformula = ~ Year + Stratum + Bottom.wt + Bottom.do,
  data = new_data2,
  family = truncated_nbinom2(link = "log")
)

select_model(model = model, null_model = null) %>%
  try()
```

```
## Forward model selection:
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
```

```
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Stratum + offset(lnSweptArea)
##
##
##        Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                            4060   2949.837 3017.837
## 2    + Year 16 96.83566      4044   2853.001 2953.001
## 3 + Stratum 13 61.20161      4031   2791.800 2917.800

## Bidirectional model selection:

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')
```

```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
##
##
##   Step Df Deviance Resid. Df Resid. Dev     AIC
## 1                       4029    2782.57 2912.57

## Backward model selection:

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
##
## Final Model:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
##
##
##   Step Df Deviance Resid. Df Resid. Dev     AIC
## 1                       4029    2782.57 2912.57
```

Final comparison

```
compare_models(
  data = new_data2,
```

```
  nb_model = "FREQ ~ Year + Lon + Bottom.wt + Bottom.sa + Depth +
offset(lnSweptArea)",

  zic_model = "FREQ ~ Year + Bottom.sa + Bottom.do + Depth +
offset(lnSweptArea)",
  zi_model = "~ Year",

  zac_model = "FREQ ~ Year + Bottom.wt + Depth + offset(lnSweptArea)",
  za_model = "~ Year",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 7.5

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ Year + Lon + Bottom.wt + Bottom.sa + Depth + offset(lnSweptArea)
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   3014.3   3153.2  -1485.1   2970.3     4072
##
##
## Dispersion parameter for nbinom2 family (): 0.0696
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -188.21451   59.95742  -3.139  0.00169 **
## Year2003       1.38029    0.48968   2.819  0.00482 **
## Year2004       0.20775    0.50975   0.408  0.68361
## Year2005       0.36845    0.47776   0.771  0.44059
## Year2006      -0.19122    0.51680  -0.370  0.71138
## Year2007       0.92458    0.49178   1.880  0.06010 .
## Year2008       1.21324    0.48741   2.489  0.01281 *
## Year2009      -0.57834    0.57878  -0.999  0.31768
## Year2010       0.09782    0.51471   0.190  0.84928
## Year2011      -0.75537    0.70286  -1.075  0.28251
## Year2012      -0.73221    0.58766  -1.246  0.21278
## Year2013       0.23560    0.52195   0.451  0.65172
## Year2014      -0.45105    0.49783  -0.906  0.36492
## Year2015       0.97419    0.46890   2.078  0.03774 *
## Year2016      -0.52284    0.50086  -1.044  0.29655
## Year2017       0.24718    0.49104   0.503  0.61470
## Year2018       0.85313    0.47553   1.794  0.07280 .
```

```
## Lon          -2.29916    0.78676  -2.922  0.00347 **
## Bottom.wt    -0.05834    0.09831  -0.593  0.55289
## Bottom.sa     0.81385    0.12208   6.667 2.62e-11 ***
## Depth        -0.85046    0.10628  -8.002 1.23e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.181

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ Year + Bottom.sa + Bottom.do + Depth + offset(lnSweptArea)
## Zero inflation:        ~Year
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2969.1   3209.1  -1446.5   2893.1     4056
##
##
## Dispersion parameter for nbinom2 family (): 0.155
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -13.11222    0.55787 -23.504  < 2e-16 ***
## Year2003      1.13109    0.59468   1.902  0.05717 .
## Year2004      0.70342    0.75305   0.934  0.35025
## Year2005      0.53630    0.62589   0.857  0.39152
## Year2006     -0.21159    0.62522  -0.338  0.73505
## Year2007      0.74737    0.60086   1.244  0.21356
## Year2008      2.90418    0.73691   3.941 8.11e-05 ***
## Year2009     -0.28513    0.67136  -0.425  0.67105
## Year2010      2.07251    0.92559   2.239  0.02515 *
## Year2011      0.13267    1.84782   0.072  0.94276
## Year2012     -0.40523    0.65656  -0.617  0.53710
## Year2013      1.50222    0.76338   1.968  0.04908 *
## Year2014     -0.37037    0.60365  -0.614  0.53951
## Year2015      2.29581    0.64610   3.553  0.00038 ***
## Year2016      0.64200    0.73508   0.873  0.38246
## Year2017      2.18528    0.76461   2.858  0.00426 **
## Year2018      1.21515    0.63692   1.908  0.05641 .
## Bottom.sa     0.55604    0.09608   5.787 7.16e-09 ***
## Bottom.do     0.51419    0.09848   5.221 1.78e-07 ***
## Depth        -0.61951    0.10175  -6.088 1.14e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
```

```
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.354e+00  2.076e+00  -0.652    0.514
## Year2003    -2.116e+01  3.549e+04  -0.001    1.000
## Year2004     1.168e+00  2.187e+00   0.534    0.593
## Year2005    -3.466e-02  2.375e+00  -0.015    0.988
## Year2006    -1.644e+01  5.485e+03  -0.003    0.998
## Year2007    -1.844e+01  1.199e+04  -0.002    0.999
## Year2008     2.683e+00  2.077e+00   1.292    0.196
## Year2009    -1.450e+01  5.701e+03  -0.002    0.998
## Year2010     3.465e+00  2.127e+00   1.629    0.103
## Year2011     1.261e+00  3.453e+00   0.365    0.715
## Year2012    -1.430e+01  3.016e+03  -0.005    0.996
## Year2013     1.865e+00  2.103e+00   0.886    0.375
## Year2014    -1.664e+01  6.147e+03  -0.003    0.998
## Year2015     2.194e+00  2.063e+00   1.064    0.288
## Year2016     1.848e+00  2.111e+00   0.876    0.381
## Year2017     2.450e+00  2.086e+00   1.175    0.240
## Year2018     1.089e+00  2.111e+00   0.516    0.606

## Dispersal: 1.476

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ Year + Bottom.wt + Depth + offset(lnSweptArea)
## Zero inflation:        ~Year
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   3043.4   3277.2  -1484.7   2969.4     4057
##
##
## Dispersion parameter for truncated_nbinom2 family (): 3.05e-09
##
## Conditional model:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.096e+01  2.628e+03  -0.012  0.99060
## Year2003     9.814e-01  7.087e-01   1.385  0.16607
## Year2004     5.005e-01  8.444e-01   0.593  0.55335
## Year2005     1.333e+00  7.285e-01   1.830  0.06726 .
## Year2006    -5.648e-01  9.322e-01  -0.606  0.54460
## Year2007     1.044e+00  7.510e-01   1.390  0.16466
## Year2008     2.704e+00  8.844e-01   3.057  0.00223 **
## Year2009     1.692e-01  1.181e+00   0.143  0.88607
## Year2010     3.006e+00  1.176e+00   2.555  0.01061 *
## Year2011    -2.243e+01  9.192e+04   0.000  0.99981
## Year2012     4.722e-01  1.035e+00   0.456  0.64834
## Year2013     1.827e+00  9.506e-01   1.922  0.05465 .
## Year2014    -4.965e-03  7.682e-01  -0.006  0.99484
```

```
## Year2015     3.499e+00  8.339e-01   4.196 2.72e-05 ***
## Year2016     1.381e+00  9.184e-01   1.504  0.13267
## Year2017     2.023e+00  9.367e-01   2.160  0.03078 *
## Year2018     1.606e+00  7.588e-01   2.117  0.03428 *
## Bottom.wt   -7.130e-01  2.284e-01  -3.122  0.00180 **
## Depth       -1.295e+00  2.295e-01  -5.643 1.67e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.42775    0.26935   9.013  < 2e-16 ***
## Year2003    -0.87637    0.32191  -2.722  0.00648 **
## Year2004     0.29272    0.37305   0.785  0.43264
## Year2005    -0.25300    0.34219  -0.739  0.45970
## Year2006     0.26050    0.36793   0.708  0.47894
## Year2007    -0.51821    0.33698  -1.538  0.12410
## Year2008     0.34484    0.37277   0.925  0.35492
## Year2009     0.68823    0.43427   1.585  0.11301
## Year2010     1.16682    0.52730   2.213  0.02691 *
## Year2011     1.04422    0.52798   1.978  0.04795 *
## Year2012     0.18110    0.40251   0.450  0.65278
## Year2013     0.19008    0.36833   0.516  0.60582
## Year2014     0.06471    0.34895   0.185  0.85289
## Year2015    -0.04820    0.34096  -0.141  0.88758
## Year2016     0.34851    0.36747   0.948  0.34293
## Year2017     0.36666    0.36738   0.998  0.31826
## Year2018    -0.01346    0.34938  -0.039  0.96927
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: NaN

##           dAIC df
## zero_infl  0.0 38
## neg_binom 45.2 22
## zero_adj  74.4 37

#

compare_models(
  data = new_data2,
  nb_model = "FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea)",

  zic_model = "FREQ ~ Year + Stratum + Bottom.wt + Bottom.do +
offset(lnSweptArea)",
  zi_model = "~ 1",

  zac_model = "FREQ ~ Year + offset(lnSweptArea)",
```

```
  za_model = "~ Year + Stratum + Bottom.wt + Bottom.do",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 7.5

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2938.8   3147.3  -1436.4   2872.8     4061
##
##
## Dispersion parameter for nbinom2 family (): 0.0857
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -12.47720    0.48696 -25.623  < 2e-16 ***
## Year2003      1.26602    0.46171   2.742  0.00611 **
## Year2004     -0.15461    0.49255  -0.314  0.75360
## Year2005      0.56026    0.45830   1.222  0.22153
## Year2006     -0.15741    0.51456  -0.306  0.75967
## Year2007      1.11898    0.47078   2.377  0.01746 *
## Year2008      0.89992    0.46964   1.916  0.05534 .
## Year2009     -0.26665    0.58322  -0.457  0.64752
## Year2010      0.38569    0.48681   0.792  0.42819
## Year2011     -0.78097    0.68976  -1.132  0.25754
## Year2012     -0.49272    0.57346  -0.859  0.39023
## Year2013      0.41594    0.50320   0.827  0.40847
## Year2014     -0.05131    0.47383  -0.108  0.91376
## Year2015      0.81676    0.44826   1.822  0.06845 .
## Year2016      0.27507    0.48298   0.570  0.56900
## Year2017      0.72237    0.48438   1.491  0.13587
## Year2018      0.11415    0.46319   0.246  0.80533
## Stratum102   -2.94235    1.12799  -2.608  0.00909 **
## Stratum201   -0.65559    0.47265  -1.387  0.16542
## Stratum202   -1.06698    0.50011  -2.134  0.03288 *
## Stratum203   -1.49414    0.77892  -1.918  0.05508 .
## Stratum301   -1.99857    0.46359  -4.311 1.62e-05 ***
## Stratum302   -2.03724    0.51603  -3.948 7.88e-05 ***
## Stratum303   -2.00855    0.82826  -2.425  0.01531 *
## Stratum401   -0.85335    0.38572  -2.212  0.02694 *
## Stratum402   -1.12504    0.38451  -2.926  0.00343 **
```

```
## Stratum403   -1.55985    0.57026  -2.735  0.00623 **
## Stratum501    0.95260    0.38244   2.491  0.01274 *
## Stratum502   -0.21168    0.38215  -0.554  0.57965
## Stratum503   -1.87744    0.58868  -3.189  0.00143 **
## Bottom.wt     0.27397    0.11549   2.372  0.01768 *
## Bottom.do     0.59632    0.12345   4.830 1.36e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.374

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:
## FREQ ~ Year + Stratum + Bottom.wt + Bottom.do + offset(lnSweptArea)
## Zero inflation:        ~1
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2940.8   3155.6  -1436.4   2872.8     4060
##
##
## Dispersion parameter for nbinom2 family (): 0.0857
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -12.47723    0.48696 -25.623  < 2e-16 ***
## Year2003      1.26601    0.46171   2.742  0.00611 **
## Year2004     -0.15460    0.49255  -0.314  0.75362
## Year2005      0.56024    0.45830   1.222  0.22155
## Year2006     -0.15740    0.51456  -0.306  0.75968
## Year2007      1.11899    0.47078   2.377  0.01746 *
## Year2008      0.89993    0.46964   1.916  0.05534 .
## Year2009     -0.26664    0.58322  -0.457  0.64753
## Year2010      0.38572    0.48681   0.792  0.42816
## Year2011     -0.78096    0.68976  -1.132  0.25755
## Year2012     -0.49270    0.57346  -0.859  0.39025
## Year2013      0.41594    0.50320   0.827  0.40847
## Year2014     -0.05129    0.47383  -0.108  0.91380
## Year2015      0.81677    0.44826   1.822  0.06844 .
## Year2016      0.27508    0.48298   0.570  0.56898
## Year2017      0.72238    0.48438   1.491  0.13587
## Year2018      0.11417    0.46319   0.246  0.80531
## Stratum102   -2.94233    1.12799  -2.608  0.00909 **
## Stratum201   -0.65556    0.47265  -1.387  0.16544
## Stratum202   -1.06697    0.50011  -2.133  0.03288 *
## Stratum203   -1.49411    0.77892  -1.918  0.05509 .
## Stratum301   -1.99858    0.46359  -4.311 1.62e-05 ***
```

```
## Stratum302   -2.03725    0.51603  -3.948 7.88e-05 ***
## Stratum303   -2.00854    0.82826  -2.425  0.01531 *
## Stratum401   -0.85334    0.38572  -2.212  0.02695 *
## Stratum402   -1.12502    0.38451  -2.926  0.00344 **
## Stratum403   -1.55985    0.57026  -2.735  0.00623 **
## Stratum501    0.95261    0.38244   2.491  0.01274 *
## Stratum502   -0.21166    0.38215  -0.554  0.57966
## Stratum503   -1.87740    0.58868  -3.189  0.00143 **
## Bottom.wt     0.27396    0.11549   2.372  0.01769 *
## Bottom.do     0.59631    0.12345   4.830 1.36e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -17.74    2948.84  -0.006    0.995

## Dispersal: 1.375

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ Year + offset(lnSweptArea)
## Zero inflation:        ~Year + Stratum + Bottom.wt + Bottom.do
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2953.0   3268.9  -1426.5   2853.0     4044
##
##
## Dispersion parameter for truncated_nbinom2 family (): 1.59e-09
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -30.41955 3124.07963  -0.010 0.992231
## Year2003       0.89680    0.62026   1.446 0.148223
## Year2004      -0.21114    0.74605  -0.283 0.777171
## Year2005       1.63706    0.67159   2.438 0.014786 *
## Year2006      -1.08339    0.79902  -1.356 0.175133
## Year2007       0.36480    0.65374   0.558 0.576827
## Year2008       2.01462    0.74513   2.704 0.006857 **
## Year2009      -1.21800    1.03575  -1.176 0.239610
## Year2010       3.90265    1.21475   3.213 0.001315 **
## Year2011     -19.06529 7914.61496  -0.002 0.998078
## Year2012      -0.78814    0.84417  -0.934 0.350496
## Year2013       0.20877    0.72878   0.286 0.774524
## Year2014       0.09895    0.68510   0.144 0.885164
## Year2015       2.49205    0.69109   3.606 0.000311 ***
## Year2016       0.41669    0.72995   0.571 0.568102
```

```
## Year2017       0.69365    0.72010   0.963 0.335415
## Year2018       1.44212    0.68716   2.099 0.035846 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.997642   0.413079   7.257 3.96e-13 ***
## Year2003    -0.981732   0.332631  -2.951  0.00316 **
## Year2004     0.213692   0.381557   0.560  0.57544
## Year2005    -0.273855   0.352408  -0.777  0.43710
## Year2006     0.089271   0.377637   0.236  0.81313
## Year2007    -0.580099   0.347245  -1.671  0.09480 .
## Year2008     0.172003   0.380717   0.452  0.65142
## Year2009     0.443073   0.445202   0.995  0.31963
## Year2010     1.167229   0.533525   2.188  0.02869 *
## Year2011     0.780593   0.537326   1.453  0.14630
## Year2012     0.055460   0.416383   0.133  0.89404
## Year2013    -0.019209   0.377711  -0.051  0.95944
## Year2014     0.014613   0.356512   0.041  0.96730
## Year2015    -0.130550   0.348854  -0.374  0.70824
## Year2016     0.294916   0.374484   0.788  0.43097
## Year2017     0.113399   0.378340   0.300  0.76439
## Year2018     0.006075   0.357268   0.017  0.98643
## Stratum102   1.690895   1.054793   1.603  0.10892
## Stratum201   0.196330   0.465084   0.422  0.67292
## Stratum202   0.188358   0.481408   0.391  0.69560
## Stratum203  -0.150960   0.609748  -0.248  0.80446
## Stratum301   0.957557   0.462960   2.068  0.03861 *
## Stratum302   0.430062   0.464793   0.925  0.35482
## Stratum303   0.228604   0.674064   0.339  0.73450
## Stratum401  -0.380067   0.354758  -1.071  0.28402
## Stratum402  -0.423197   0.348060  -1.216  0.22403
## Stratum403  -0.147773   0.484480  -0.305  0.76036
## Stratum501  -1.296453   0.333267  -3.890  0.00010 ***
## Stratum502  -0.887844   0.346086  -2.565  0.01031 *
## Stratum503   0.519325   0.553784   0.938  0.34836
## Bottom.wt   -0.094541   0.080180  -1.179  0.23836
## Bottom.do   -0.411747   0.086148  -4.780 1.76e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: NaN

##           dAIC df
## neg_binom  0.0 33
## zero_infl  2.0 34
## zero_adj  14.2 50
```

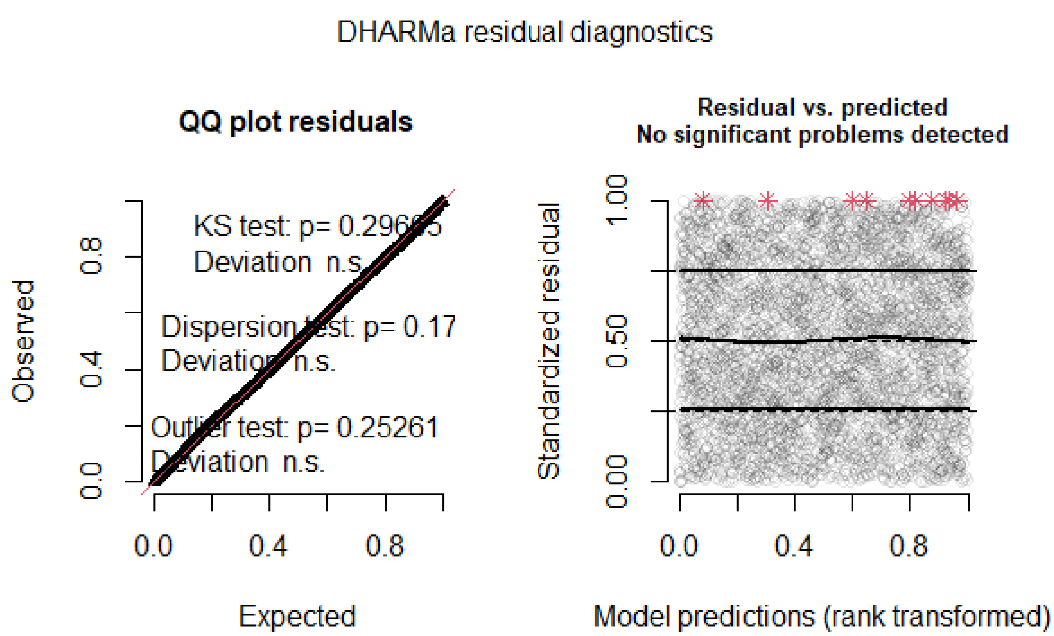
Zero-inflated is the winner! NB is best when including Stratum

Step 3: DHARMA simulations

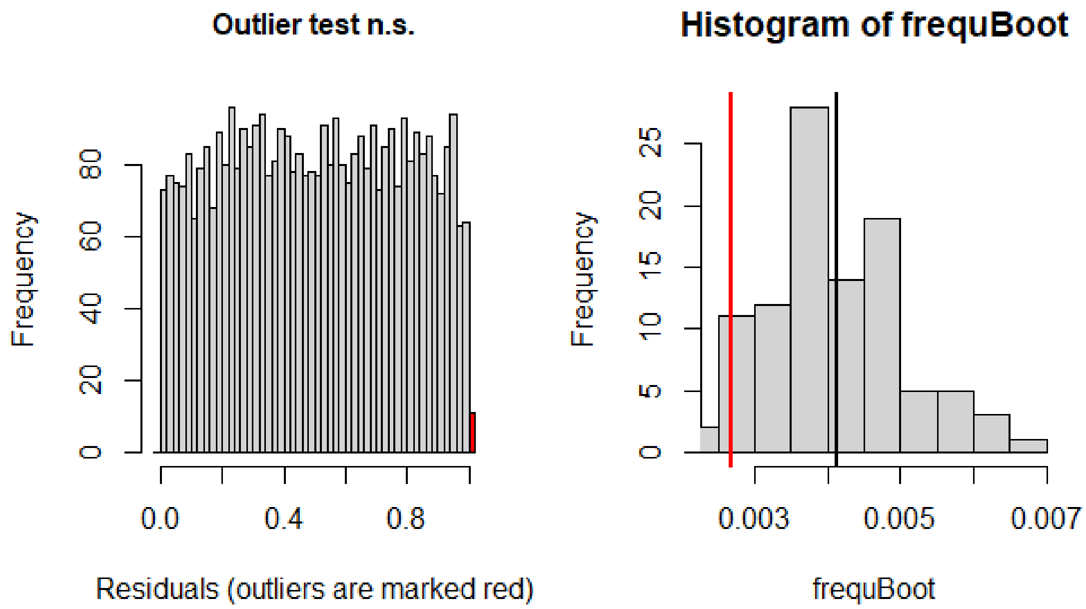
```
# model picked when excluding stratum
model1 <- glmmTMB(FREQ ~ Year + Bottom.sa + Bottom.do + Depth +
offset(lnSweptArea),
  ziformula = ~Year,
  data = new_data2,
  family = nbinom2
)

# model picked when including stratum
model <- glmmTMB(FREQ ~ Year + Bottom.wt + Bottom.do + Stratum +
offset(lnSweptArea),
  data = new_data2,
  family = nbinom2
)

AIC(model1, model)
```

	df	AIC
model1	38	2969.06
model	33	2938.79

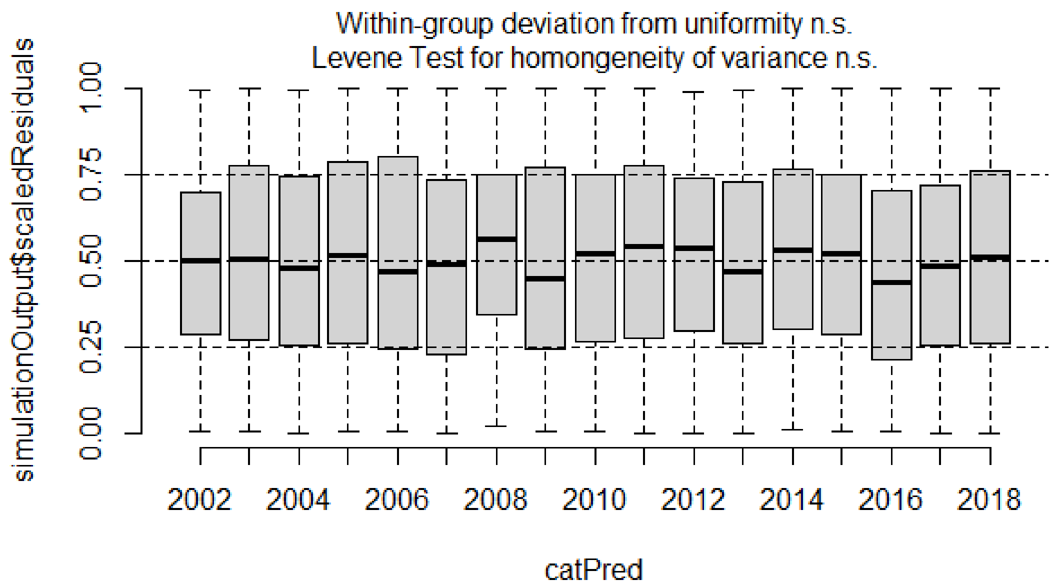
```
sim <- simulateResiduals(model)
plot(sim, quantreg = TRUE)
```



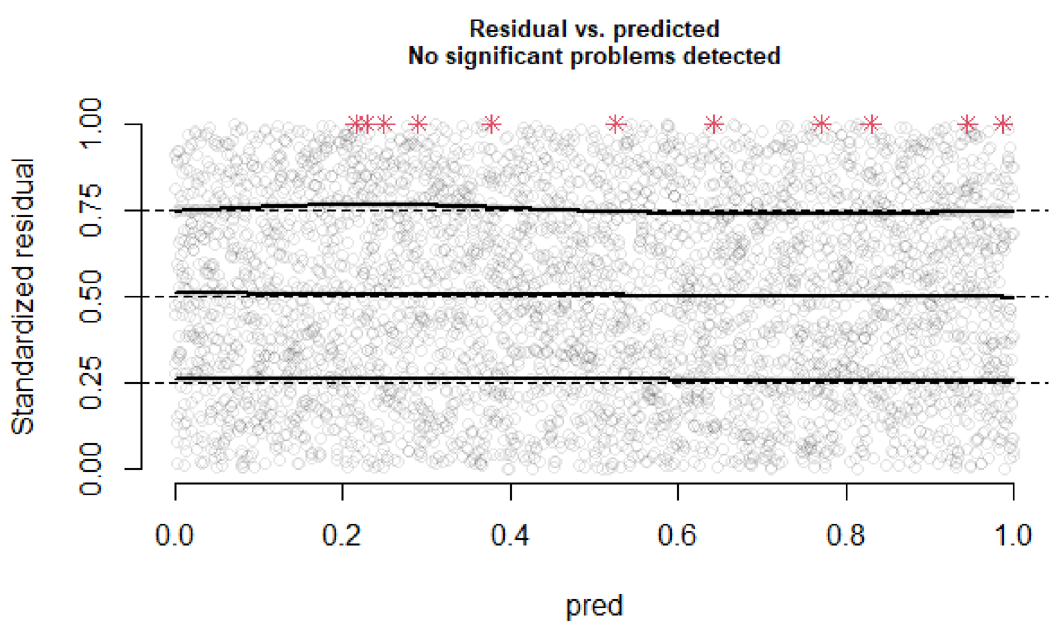

```
testOutliers(model, type = "bootstrap")
```



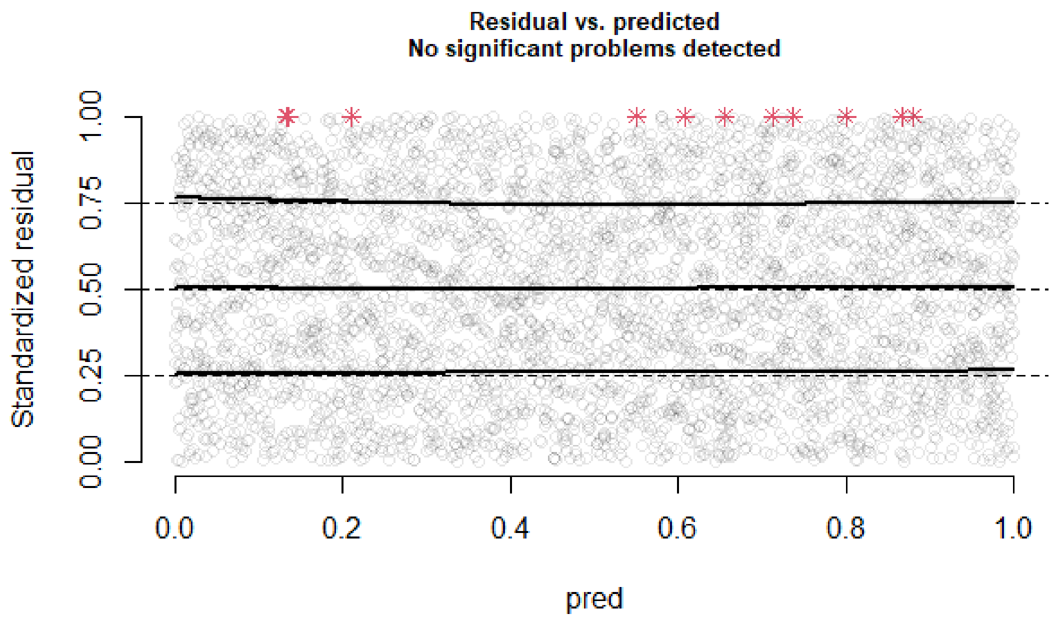

```
##
##  DHARMa bootstrapped outlier test
##
## data:  model
## outliers at both margin(s) = 11, observations = 4094, p-value = 0.16
## alternative hypothesis: two.sided
##  percent confidence interval:
##  0.002686859 0.006234734
## sample estimates:
## outlier frequency (expected: 0.00414264777723498 )
##                                           0.002686859

# Compare residuals vs. factors
plotResiduals(sim, quantreg = TRUE, form = new_data2$Year)
```



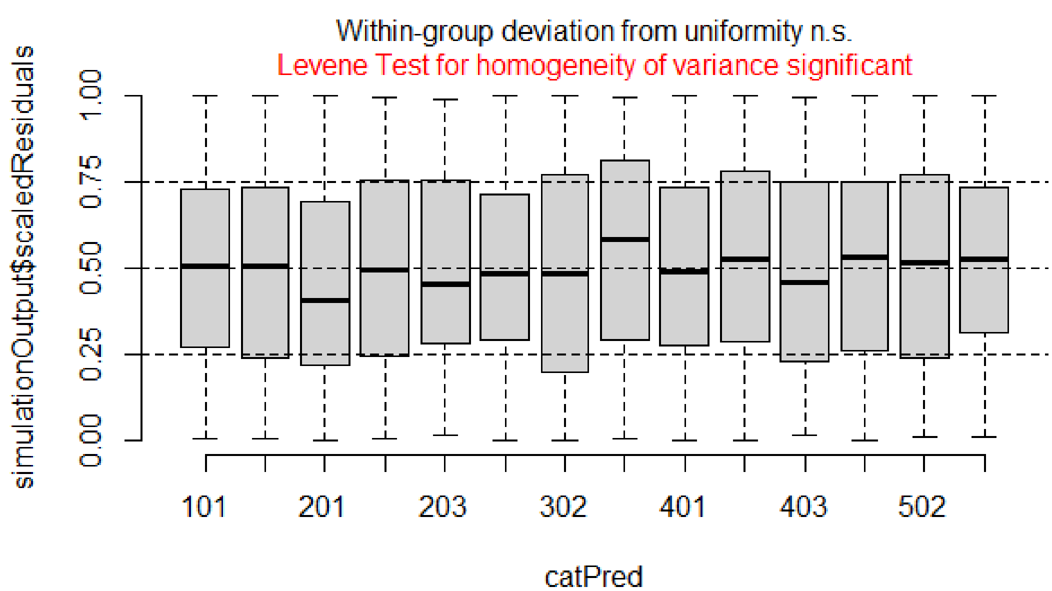

```
plotResiduals(sim, quantreg = TRUE, form = new_data2$Bottom.wt)
```




```
plotResiduals(sim, quantreg = TRUE, form = new_data2$Bottom.do)
```




```
plotResiduals(sim, quantreg = TRUE, form = new_data2$Stratum)
```



Step 4: Plot predictions

Add 2021 data

```
indata <- readxl::read_excel(here::here("data/2021
Data/2021_bluefish_data_NEAMAP_April_2022.xlsx"), sheet = "Survey
Data_ChesMMAP") %>%
```

```
  dplyr::rename(Stratum_area = Wgt.Factor)

indata <- indata %>%
  dplyr::rename(FREQ = Bf.count) %>%
  dplyr::select(-c(Program, Gear, Area)) # same value for all data

new_data <- remove_missing(indata,
  f_col = "FREQ",
  t_col = "Year"
)

## Data before removing missing values:

##    SampleID            STATE               Year               Month          
##  Length:6264        Length:6264        Length:6264        Length:6264       
##  Class :character   Class :character   Class :character   Class :character  
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character  
##                                                                             
##                                                                             
##                                                                             
##                                                                             
##       Date                            Time                            Areasw     
##  Min.   :2002-03-29 00:00:00   Min.   :2002-03-29 09:30:00   Min.   : 6306  
##  1st Qu.:2006-03-30 00:00:00   1st Qu.:2006-03-30 12:32:15   1st Qu.:29848  
##  Median :2010-09-06 00:00:00   Median :2010-09-06 09:04:00   Median :45017  
##  Mean   :2010-07-28 05:25:58   Mean   :2010-07-28 18:01:05   Mean   :42959  
##  3rd Qu.:2014-11-09 00:00:00   3rd Qu.:2014-11-09 10:08:35   3rd Qu.:52980  
##  Max.   :2018-11-17 00:00:00   Max.   :2018-11-17 15:59:40   Max.   :98741  
##                                                                             
##       FREQ            Bf.totwght          Stratum              Depth       
##  Min.   : 0.0000   Min.   : 0.00000   Length:6264        Min.   : 2.438  
##  1st Qu.: 0.0000   1st Qu.: 0.00000   Class :character   1st Qu.: 7.407  
##  Median : 0.0000   Median : 0.00000   Mode  :character   Median :10.363  
##  Mean   : 0.1502   Mean   : 0.03152                      Mean   :11.199  
##  3rd Qu.: 0.0000   3rd Qu.: 0.00000                      3rd Qu.:12.802  
##  Max.   :85.0000   Max.   :13.72200                      Max.   :41.758  
##                                                          NA's   :1       
##   Stratum_area           Lat             Lon           Bottom.wt    
##  Min.   : 20795800   Min.   :36.93   Min.   :-76.52   Min.   : 0.81  
##  1st Qu.:405279801   1st Qu.:37.39   1st Qu.:-76.32   1st Qu.:13.30  
##  Median :454035100   Median :37.76   Median :-76.20   Median :18.24  
##  Mean   :507171662   Mean   :37.90   Mean   :-76.21   Mean   :18.00  
##  3rd Qu.:776073103   3rd Qu.:38.34   3rd Qu.:-76.10   3rd Qu.:24.68  
##  Max.   :897041604   Max.   :39.30   Max.   :-75.93   Max.   :29.44  
##                                                       NA's   :213    
##    Bottom.sa       Bottom.do    
##  Min.   : 0.09   Min.   : 0.000  
##  1st Qu.:15.34   1st Qu.: 4.870  
##  Median :19.04   Median : 6.800  
##  Mean   :19.05   Mean   : 6.662  
```

```
##  3rd Qu.:22.87   3rd Qu.: 8.530  
##  Max.   :34.98   Max.   :14.680  
##  NA's   :239     NA's   :609

## Data after removing missing values:

##    SampleID            STATE               Year              Month          
##  Length:5620        Length:5620        Length:5620        Length:5620       
##  Class :character   Class :character   Class :character   Class :character  
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character  
##                                                                             
##                                                                             
##                                                                             
##       Date                             Time                           Areasw     
##  Min.   :2002-03-29 00:00:00   Min.   :2002-03-29 13:10:00   Min.   : 6306  
##  1st Qu.:2006-05-15 18:00:00   1st Qu.:2006-05-16 03:15:30   1st Qu.:29570  
##  Median :2010-09-06 00:00:00   Median :2010-09-06 10:04:00   Median :44869  
##  Mean   :2010-09-11 03:54:26   Mean   :2010-09-11 16:28:46   Mean   :42812  
##  3rd Qu.:2015-05-28 00:00:00   3rd Qu.:2015-05-28 10:40:19   3rd Qu.:52932  
##  Max.   :2018-11-17 00:00:00   Max.   :2018-11-17 15:59:40   Max.   :98741  
##       FREQ           Bf.totwght          Stratum              Depth       
##  Min.   : 0.0000   Min.   : 0.00000   Length:5620        Min.   : 2.438  
##  1st Qu.: 0.0000   1st Qu.: 0.00000   Class :character   1st Qu.: 7.498  
##  Median : 0.0000   Median : 0.00000   Mode  :character   Median :10.363  
##  Mean   : 0.1637   Mean   : 0.03384                      Mean   :11.191  
##  3rd Qu.: 0.0000   3rd Qu.: 0.00000                      3rd Qu.:12.802  
##  Max.   :85.0000   Max.   :13.72200                      Max.   :41.758  
##   Stratum_area            Lat             Lon           Bottom.wt    
##  Min.   : 20795800   Min.   :36.93   Min.   :-76.52   Min.   : 2.30  
##  1st Qu.:405279801   1st Qu.:37.41   1st Qu.:-76.32   1st Qu.:13.63  
##  Median :454035100   Median :37.76   Median :-76.20   Median :19.04  
##  Mean   :506759301   Mean   :37.91   Mean   :-76.21   Mean   :18.31  
##  3rd Qu.:776073103   3rd Qu.:38.34   3rd Qu.:-76.10   3rd Qu.:24.83  
##  Max.   :897041604   Max.   :39.30   Max.   :-75.93   Max.   :29.44  
##    Bottom.sa        Bottom.do     
##  Min.   : 0.09   Min.   : 0.000  
##  1st Qu.:15.41   1st Qu.: 4.860  
##  Median :19.05   Median : 6.800  
##  Mean   :19.07   Mean   : 6.644  
##  3rd Qu.:22.84   3rd Qu.: 8.510  
##  Max.   :34.98   Max.   :14.680

## 644 datapoints removed (10.3% of data removed)

## All years had positive catch

# remove March
length1 <- nrow(new_data)
new_data <- new_data %>%
  dplyr::filter(
    Month != "MAR",
```

```r
    Bottom.wt > 12, # 12C is lower temp limit
    Bottom.do > 2
  ) # 2 mg/L is lower O2 limit
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")

## 1526 rows removed

new_data2 <- standardize_data(new_data,
  cols = c(
    "Year", "Month", "Stratum", "Date",
    "Time", "Lat", "Lon", "Area",
    "Areasw", "Stratum_area",
    "SampleID", "State"
  ),
  f_col = "FREQ"
)

## SampleID has been changed to factor

## Not sure what to do with STATE, leaving as is

## Year has been changed to factor

## Month has been changed to factor

## Date has been changed to factor

## Time has been changed to factor

## Areasw has been changed to factor

## Bf.totwght has been standardized

## Stratum has been changed to factor

## Depth has been standardized

## Stratum_area has been changed to factor

## Lat has been changed to factor

## Lon has been changed to factor

## Bottom.wt has been standardized

## Bottom.sa has been standardized

## Bottom.do has been standardized

## Data standardized!

# change lat, lon, Areasw back to numeric
new_data2 <- new_data2 %>%
```

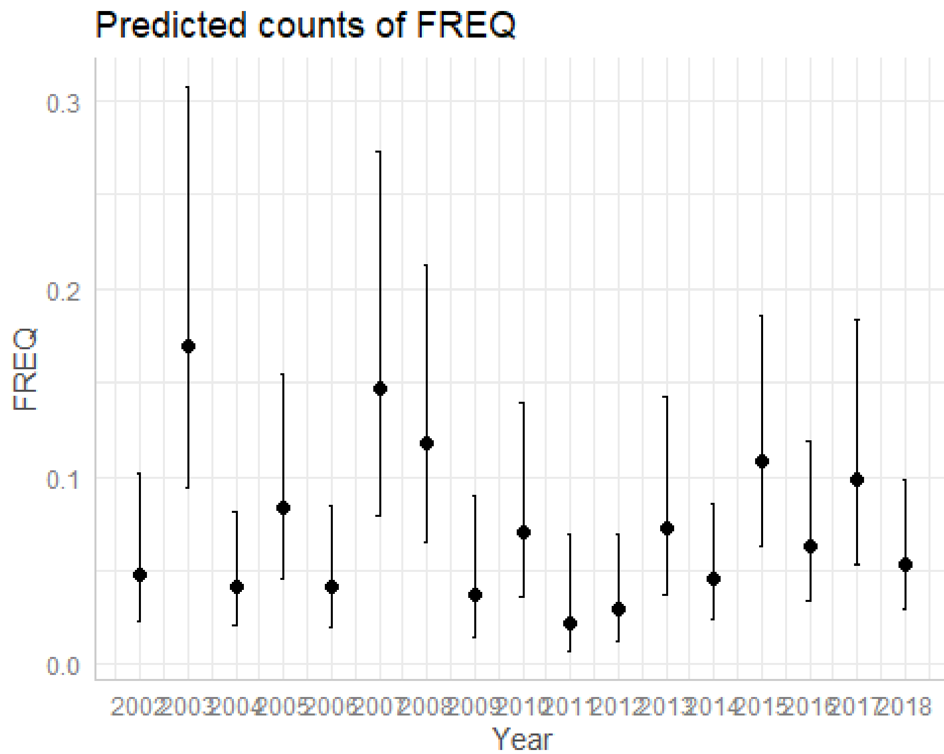
```
  dplyr::mutate(
    Lat = Lat %>%
      as.character() %>%
      as.numeric(),
    Lon = Lon %>%
      as.character() %>%
      as.numeric(),
    Areasw = Areasw %>%
      as.character() %>%
      as.numeric()
  )
head(new_data2, n = 2)
```

SampleID	STATE	Year	Month	Date	Time	Areasw	FREQ	Bf.totwght	Stratum	Depth	Stratum_area	Lat	Lon	Bottom.wt	Bottom.sa	Bottom.do
CM20020501037	MD	2002	MAY	2002-05-20	2002-05-20 19:29:00	49427.75	0	-0.1	302	0.31	447046201.349	38.12	-76.226	-0.61	-1.81	0.28
CM20020501065	VA	2002	MAY	2002-05-20	2002-05-20 15:37:00	47410.44	0	-0.1	403	2.13	221852899.225	37.71	-76.18	-0.51	-0.41	0.38

```
new_data2 <- new_data2 %>%
  dplyr::select(-c( # Stratum,
    STATE
  ))

new_data2$lnSweptArea <- log(new_data2$Areasw)

model <- glmmTMB(FREQ ~ Year + Bottom.wt + Bottom.do + Stratum +
offset(lnSweptArea),
  data = new_data2,
  family = nbinom2
)
```

ggemmeans method

```
index.pred <- ggeffects::ggemmeans(model,
  terms = "Year",
  type = "fe.zi"
)
plot(index.pred)
```

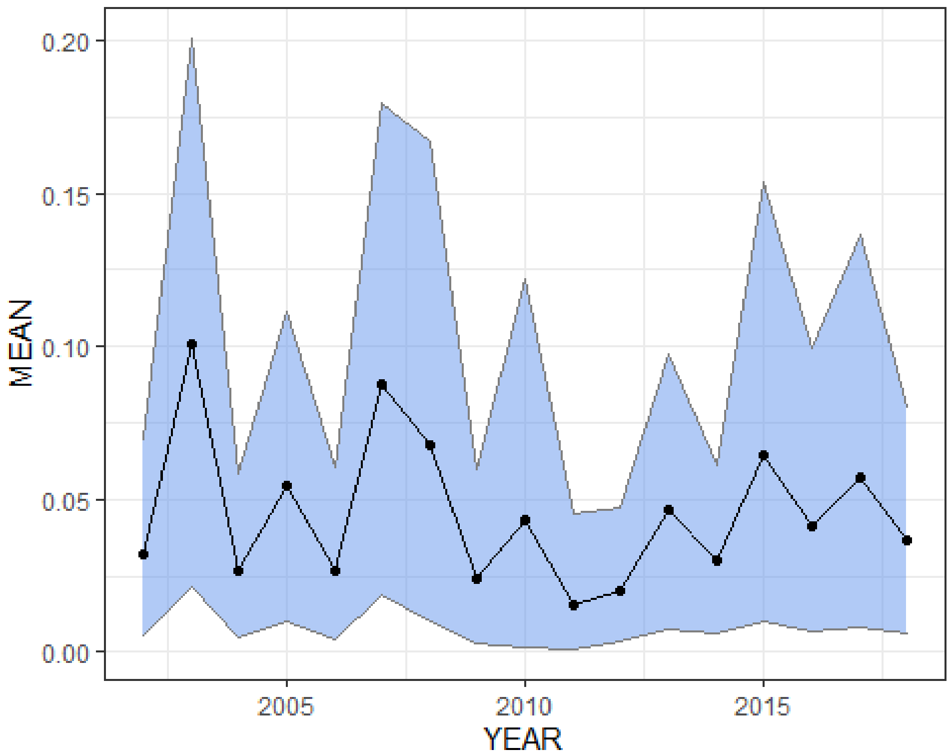


Bootstrap method

```
source(here::here("scripts/bootstrap_functions_AT_lnE2.R"))

index.pred <- boot.ZI_AT(
  best = model,
  type_z = "zero",
  nboots = 1000,
  ycol = "Year"
)

## starting fxn

## year name fixed

## starting boots

## session planned
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## finished boots

## Number of successful predictions: 1000
```

```
index.pred %>%
  plot_boot()
```



Step 5: Save index

```
index.out <- as.data.frame(index.pred)
index.out$CV <- abs(index.out$SE / index.out$MEAN)

names(index.out)[c(1, 2)] <- c("Year", "Frequency")

write.csv(index.out,
  here::here("reports/updated/ChesMMAP_index5.csv"),
  row.names = FALSE
)
```

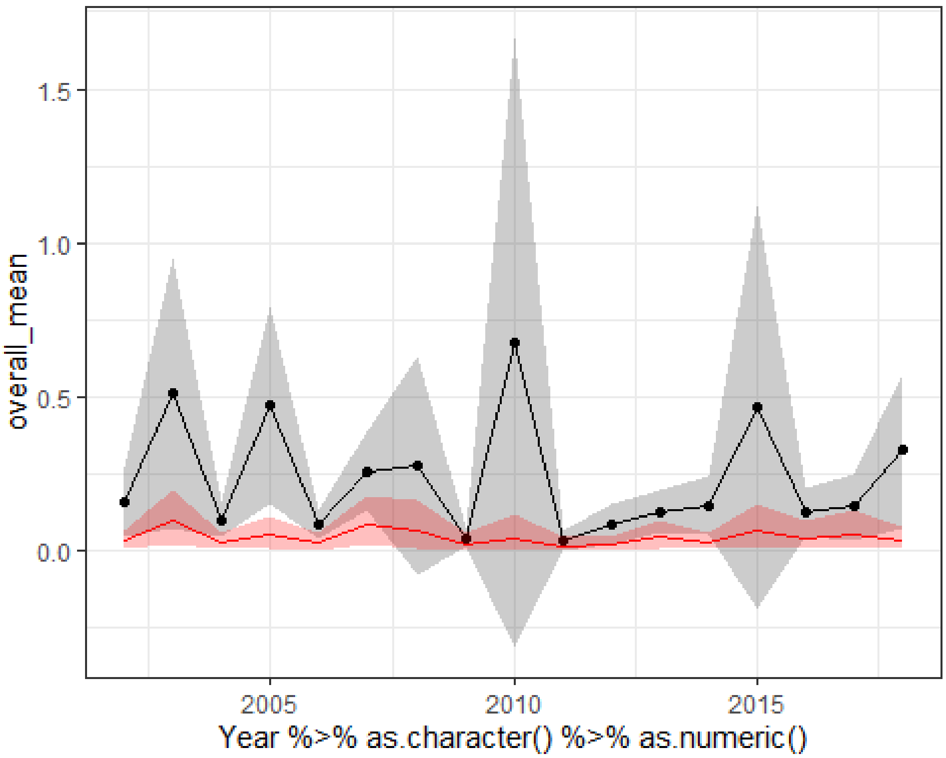
Develop nominal index

```
# weighting by stratum area

stratum_weights <- indata %>%
  dplyr::select(Stratum, Stratum_area) %>%
```

```r
  dplyr::distinct() %>%
  dplyr::mutate(weight = Stratum_area / sum(Stratum_area))

nominal_data <- indata %>%
  dplyr::left_join(stratum_weights, by = c("Stratum", "Stratum_area")) %>%
  #  dplyr::rename(FREQ = Bf.count) %>%
  dplyr::group_by(Year) %>%
  dplyr::mutate(year_total = sum(FREQ, na.rm = TRUE)) %>%
  dplyr::filter(
    year_total > 0,
    Month != "MAR",
    Bottom.wt > 12, # 12C is lower temp limit
    Bottom.do > 2
  ) %>% # 2 mg/L is lower O2 limit
  dplyr::ungroup() %>%
  dplyr::group_by(Year, Stratum, weight) %>%
  dplyr::summarise(
    mean_freq = mean(FREQ),
    var_freq = var(FREQ),

    mean_cpue = mean(FREQ / Areasw),
    var_cpue = var(FREQ / Areasw),

    n_obs = length(SampleID)
  ) %>%
  dplyr::ungroup() %>%
  dplyr::group_by(Year) %>%
  dplyr::mutate(
    mean_var = mean(var_freq, na.rm = TRUE),
    var_freq = ifelse(is.na(var_freq), mean_var, var_freq),

    mean_var_cpue = mean(var_cpue, na.rm = TRUE),
    var_cpue = ifelse(is.na(var_cpue), mean_var_cpue, var_cpue),

    n_strata = length(unique(Stratum))
  ) %>%
  dplyr::ungroup() %>%
  dplyr::group_by(Year) %>%
  dplyr::summarise(
    overall_mean = sum(weight * mean_freq),
    overall_var = sum(weight^2 * var_freq / n_obs),
    overall_sd = sqrt(overall_var),

    overall_mean_cpue = sum(weight * mean_cpue),
    overall_var_cpue = sum(weight^2 * var_cpue / n_obs),
    overall_sd_cpue = sqrt(overall_var_cpue),

    .groups = "keep"
```

```
  ) %>%
  dplyr::ungroup()

## `summarise()` has grouped output by 'Year', 'Stratum'. You can override
using
## the `.groups` argument.

write.csv(nominal_data,
  here::here("reports/updated/VA_ChesMMAP_index_nominal4.csv"),
  row.names = FALSE
)

ggplot2::ggplot(
  nominal_data,
  ggplot2::aes(
    x = Year %>%
      as.character() %>%
      as.numeric(),
    group = 1
  )
) +
  ggplot2::geom_ribbon(ggplot2::aes(
    ymin = overall_mean - 1.96 * overall_sd,
    ymax = overall_mean + 1.96 * overall_sd
  ),
  alpha = 0.25
  ) +
  ggplot2::geom_line(ggplot2::aes(y = overall_mean)) +
  ggplot2::geom_point(ggplot2::aes(y = overall_mean)) +
  ggplot2::theme_bw() +
  ggplot2::geom_line(
    data = index.out,
    ggplot2::aes(
      x = Year %>%
        as.character() %>%
        as.numeric(),
      y = Frequency
    ),
    color = "red",
    inherit.aes = FALSE
  ) +
  ggplot2::geom_ribbon(
    data = index.out,
    ggplot2::aes(
      x = Year %>%
        as.character() %>%
        as.numeric(),
      ymin = LCI,
      ymax = UCI
    ),
```

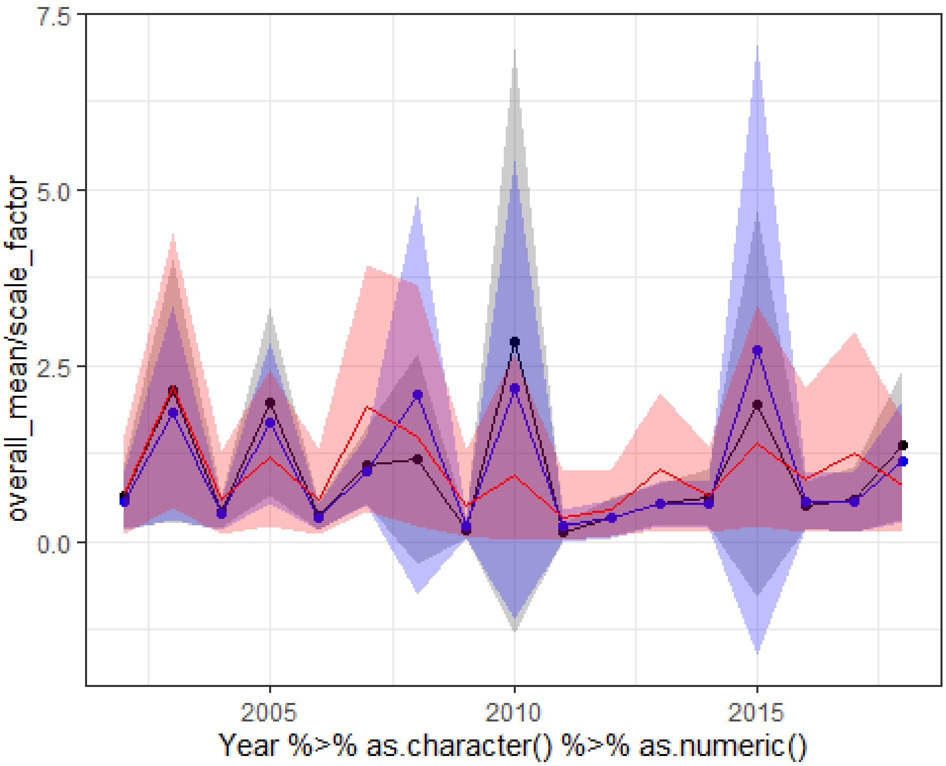
```r
    alpha = 0.25,
    fill = "red",
    inherit.aes = FALSE
  )
```




```r
# standardize to mean
nominal_data <- nominal_data %>%
  dplyr::mutate(scale_factor = mean(overall_mean))

nominal_cpue <- nominal_data %>%
  dplyr::mutate(scale_factor = mean(overall_mean_cpue))

index.out <- index.out %>%
  dplyr::mutate(scale_factor = mean(Frequency))

# plot
ggplot2::ggplot(
  nominal_data,
  ggplot2::aes(
    x = Year %>%
      as.character() %>%
      as.numeric(),
    group = 1
  )
) +
```

```
# nominal mean
ggplot2::geom_ribbon(ggplot2::aes(
  ymin = (overall_mean - 1.96 * overall_sd) / scale_factor,
  ymax = (overall_mean + 1.96 * overall_sd) / scale_factor
),
alpha = 0.25
) +
ggplot2::geom_line(ggplot2::aes(y = overall_mean / scale_factor)) +
ggplot2::geom_point(ggplot2::aes(y = overall_mean / scale_factor)) +

# nominal mean cpue
ggplot2::geom_ribbon(
  data = nominal_cpue,
  ggplot2::aes(
    ymin = (overall_mean_cpue - 1.96 * overall_sd_cpue) / scale_factor,
    ymax = (overall_mean_cpue + 1.96 * overall_sd_cpue) / scale_factor
  ),
  alpha = 0.25,
  fill = "blue"
) +
ggplot2::geom_line(
  data = nominal_cpue,
  ggplot2::aes(y = overall_mean_cpue / scale_factor),
  color = "blue"
) +
ggplot2::geom_point(
  data = nominal_cpue,
  ggplot2::aes(y = overall_mean_cpue / scale_factor),
  color = "blue"
) +

# model output
ggplot2::theme_bw() +
ggplot2::geom_line(
  data = index.out,
  ggplot2::aes(
    x = Year %>%
      as.character() %>%
      as.numeric(),
    y = Frequency / scale_factor
  ),
  color = "red",
  inherit.aes = FALSE
) +
ggplot2::geom_ribbon(
  data = index.out,
  ggplot2::aes(
    x = Year %>%
      as.character() %>%
      as.numeric(),
```

```
      ymin = LCI / scale_factor,
      ymax = UCI / scale_factor
    ),
    alpha = 0.25,
    fill = "red",
    inherit.aes = FALSE
  )
```



The confidence intervals go below zero. Should we log-transform the data?

Bluefish index standardization - SEAMAP Fall Age-0

Katie Drew, ASMFC

2022-02-10

Bluefish index standardization

Step 1: Data processing

Load functions

Read in data

```
##   SAMPLEID STATE       PROGRAM
GEAR
## 1  1989001    SC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
## 2  1989003    SC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
## 3  1989017    NC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
## 4  1989019    NC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
## 5  1989021    NC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
## 6  1989023    NC SEAMAP-SA CTS 75' Falcon trawl (double rigged) without
TED
##   YEAR MONTH DAY  TIME DURATION EFFORT  SPECIES  N WEIGHT AREA STRATUM
LATITUDE
## 1 1989     4  12 20:32       20      9 bluefish  3  0.274   SC      43
32.41
## 2 1989     4  12 23:29       20     12 bluefish  4  0.128   SC      43
32.50
## 3 1989     4  17 20:21       20     21 bluefish  7  0.673   OB      57
33.73
## 4 1989     4  17 21:47       20     24 bluefish  8  0.661   OB      57
33.75
## 5 1989     4  17 23:49       20     63 bluefish 21  1.602   OB      57
33.91
## 6 1989     4  18  2:25       20     39 bluefish 13  1.290   OB      59
34.06
##   LONGITUDE DEPTH SURFACE.TEMP BOTTOM.TEMP SURFACE.SALINITY SALINITY
SEASON
## 1    -80.28     7         14.4        14.4             34.1     33.8
SPRING
## 2    -80.16     9         14.7        14.6             33.0     33.0
SPRING
```

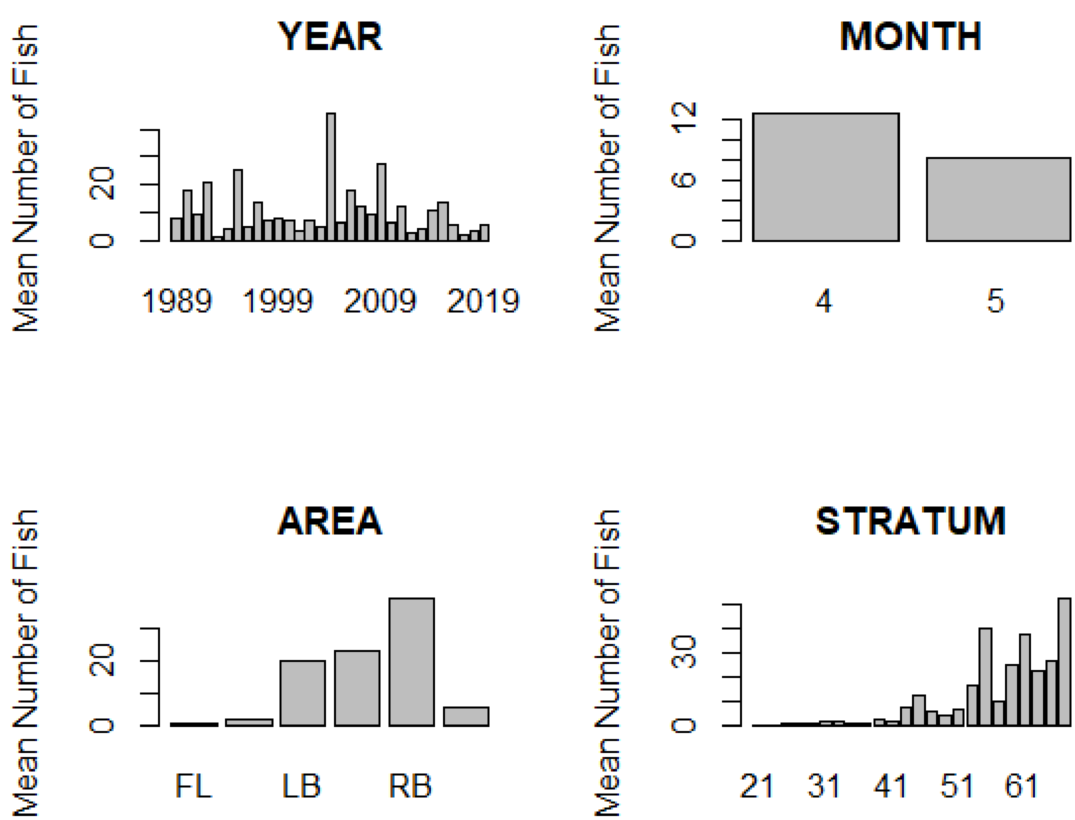
```
## 3    -77.89     8         15.7        15.7             34.6     34.7
SPRING
## 4    -77.90     7         16.0        15.8             34.0     34.1
SPRING
## 5    -77.90     8         16.4        14.7             26.7     34.9
SPRING
## 6    -77.87     9         15.0        13.9             21.0     34.9
SPRING
##   POS_TOW N_A1
## 1       1    2
## 2       1    1
## 3       1    7
## 4       1    6
## 5       1   18
## 6       1   11
```

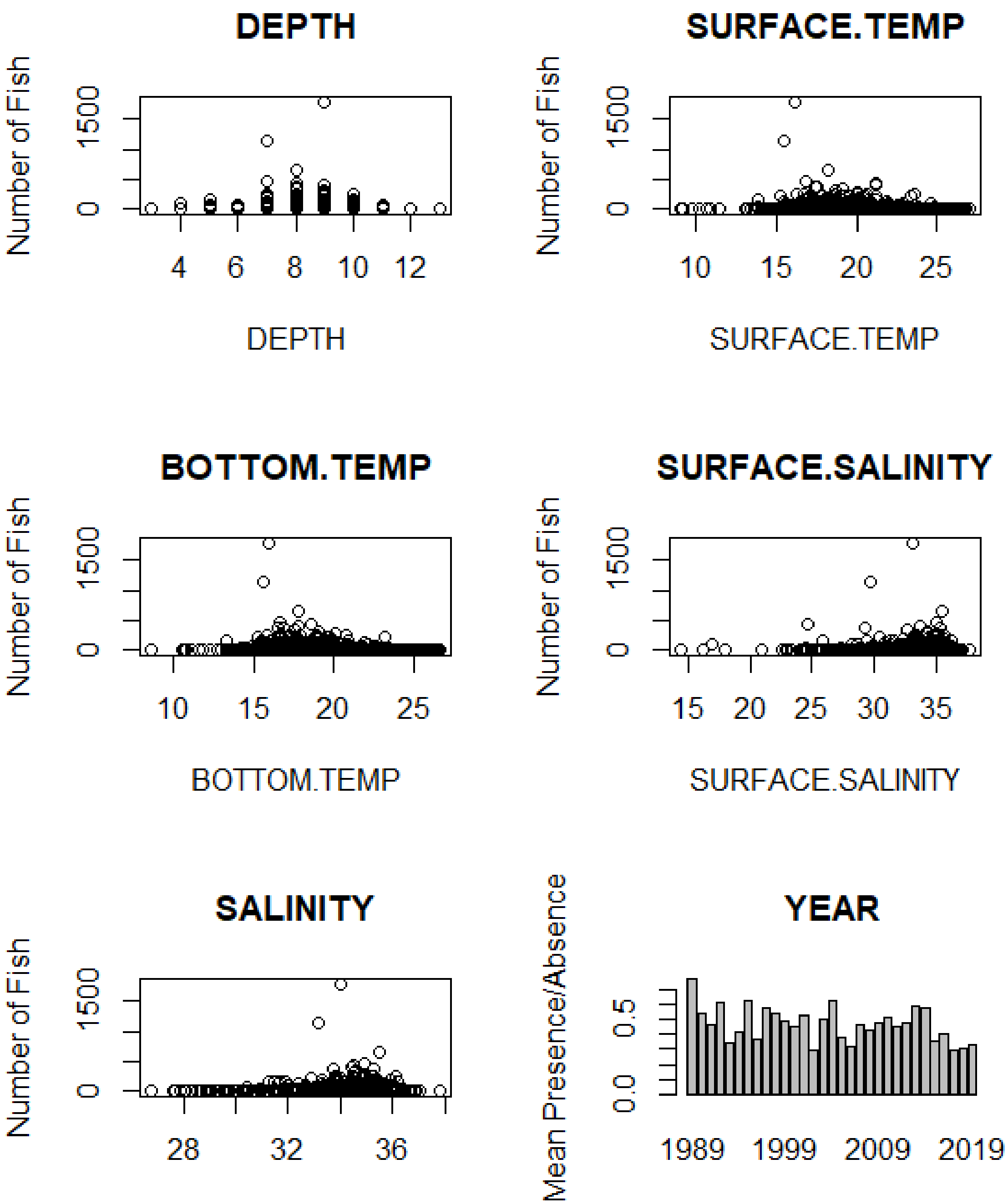
```
##     SAMPLEID           YEAR      MONTH         DAY            TIME
##  Min.   :1989001   2009   : 112   4:1770   Min.   : 1.00   Length:2816
##  1st Qu.:1998115   2010   : 112   5:1046   1st Qu.: 9.00   Class
:character
##  Median :2006008   2011   : 112            Median :17.00   Mode
:character
##  Mean   :2005102   2012   : 112            Mean   :16.46
##  3rd Qu.:2012138   2016   : 112            3rd Qu.:24.00
##  Max.   :2019175   2015   : 109            Max.   :31.00
##                    (Other):2147
##        N                WEIGHT         AREA       STRATUM        LATITUDE
##  Min.   :   0.00   Min.   :  0.0000   FL:491   51     : 171   Min.
:28.76
##  1st Qu.:   0.00   1st Qu.:  0.0000   GA:677   35     : 155   1st
Qu.:30.95
##  Median :   0.00   Median :  0.0000   LB:423   49     : 148   Median
:32.59
##  Mean   :  11.43   Mean   :  0.9536   OB:433   33     : 144   Mean
:32.31
##  3rd Qu.:   4.00   3rd Qu.:  0.4285   RB:187   41     : 141   3rd
Qu.:33.80
##  Max.   :1839.00   Max.   :100.5660   SC:605   37     : 138   Max.
:35.23
##                                                (Other):1919
##    LONGITUDE          DEPTH        SURFACE.TEMP    BOTTOM.TEMP
##  Min.   :-81.44   Min.   : 3.00   Min.   : 9.10   Min.   : 8.70
##  1st Qu.:-81.09   1st Qu.: 7.00   1st Qu.:18.90   1st Qu.:18.00
##  Median :-79.95   Median : 8.00   Median :20.60   Median :19.80
##  Mean   :-79.60   Mean   : 8.06   Mean   :20.47   Mean   :19.74
##  3rd Qu.:-78.29   3rd Qu.: 9.00   3rd Qu.:22.20   3rd Qu.:21.40
##  Max.   :-75.57   Max.   :13.00   Max.   :27.00   Max.   :26.60
##
##  SURFACE.SALINITY    SALINITY        SEASON            POS_TOW
##  Min.   :14.50    Min.   :26.80   Length:2816        Min.   :0.0000
```

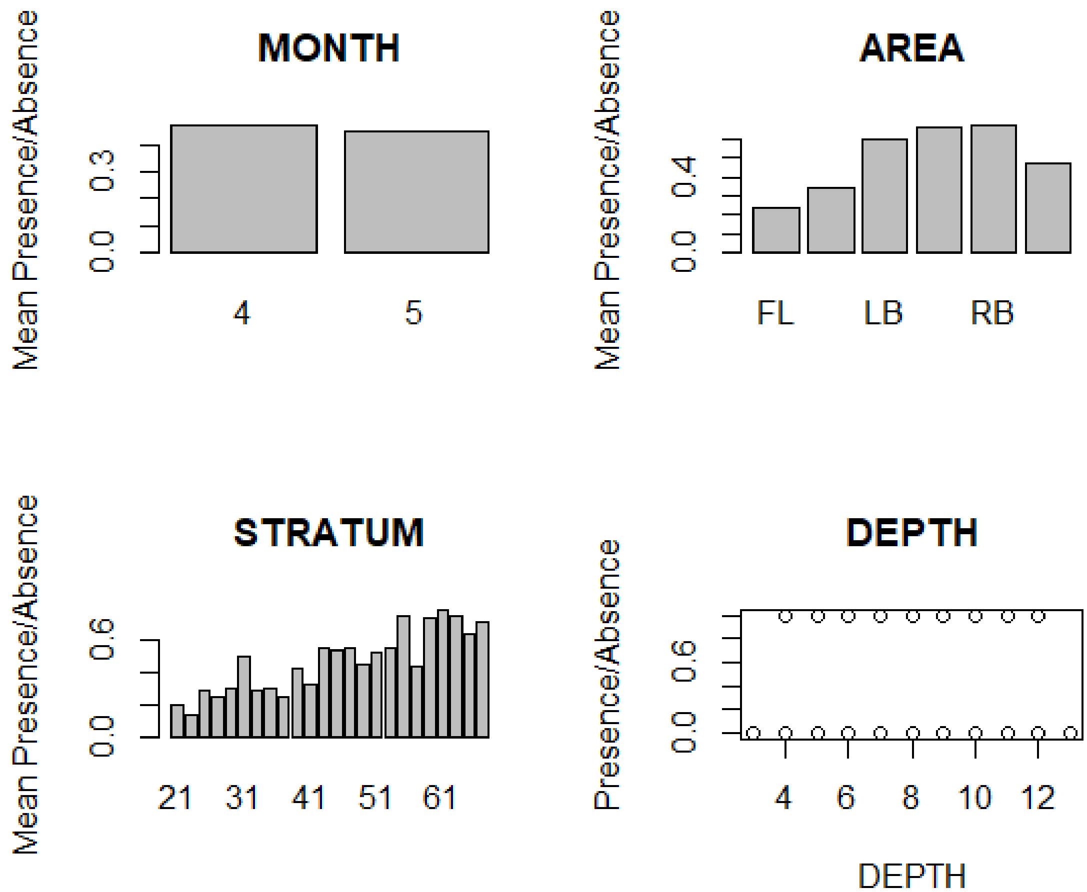
```
##  1st Qu.:32.40    1st Qu.:33.20   Class :character   1st Qu.:0.0000  
##  Median :33.80    Median :34.20   Mode  :character   Median :0.0000  
##  Mean   :33.41    Mean   :34.08                      Mean   :0.4613  
##  3rd Qu.:34.90    3rd Qu.:35.10                      3rd Qu.:1.0000  
##  Max.   :37.70    Max.   :37.80                      Max.   :1.0000  
##                                                                      
##       FREQ        
##  Min.   :   0.00  
##  1st Qu.:   0.00  
##  Median :   0.00  
##  Mean   :  10.85  
##  3rd Qu.:   4.00  
##  Max.   :1803.00  
## 
```

Remove missing data

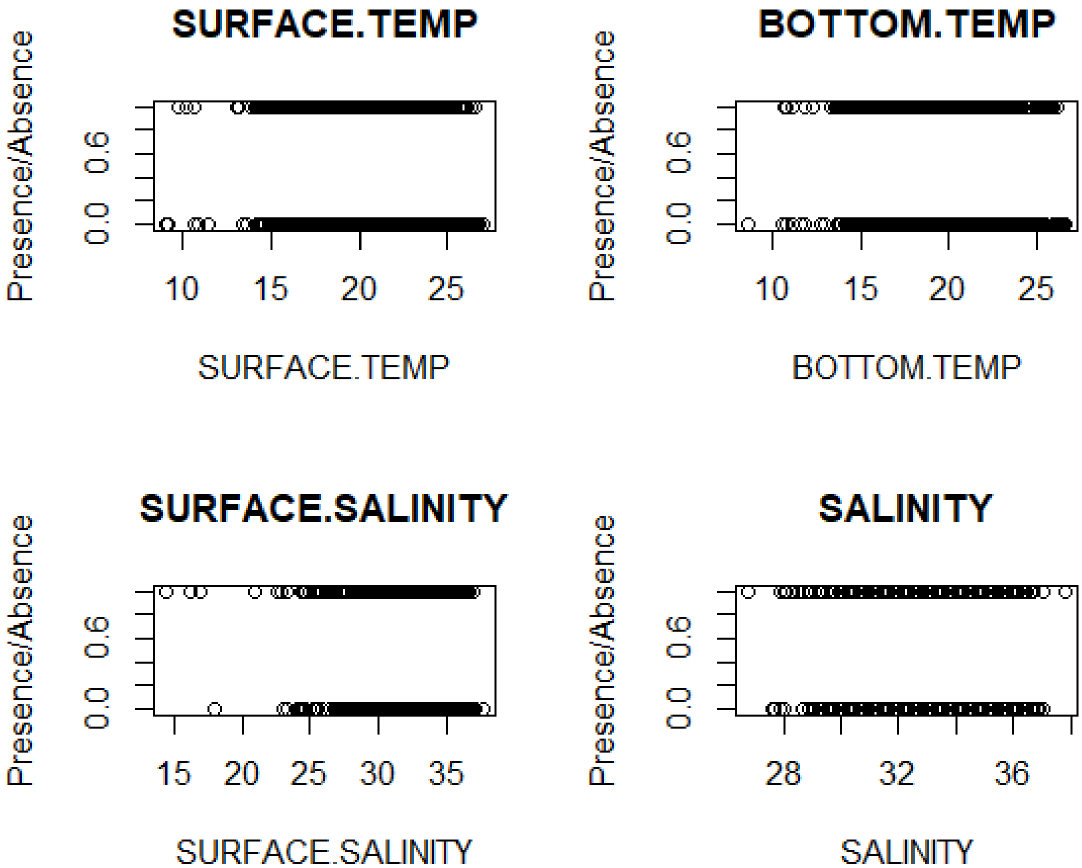
Plot data and decide if any data should be removed due to sparse sampling

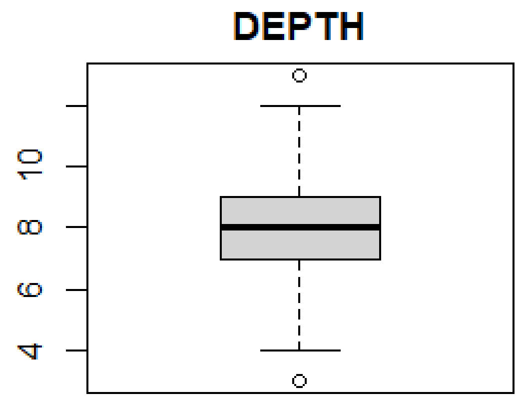



DEPTH
Number of Fish
0
1500
4
6
8
10
12
DEPTH
SURFACE.TEMP
Number of Fish
0
1500
10
15
20
25
SURFACE.TEMP
BOTTOM.TEMP
Number of Fish
0
1500
10
15
20
25
BOTTOM.TEMP
SURFACE.SALINITY
Number of Fish
0
1500
15
20
25
30
35
SURFACE.SALINITY
SALINITY
Number of Fish
0
1500
28
32
36
SALINITY
YEAR
Mean Presence/Absence
0.0
0.5
1989
1999
2009
2019

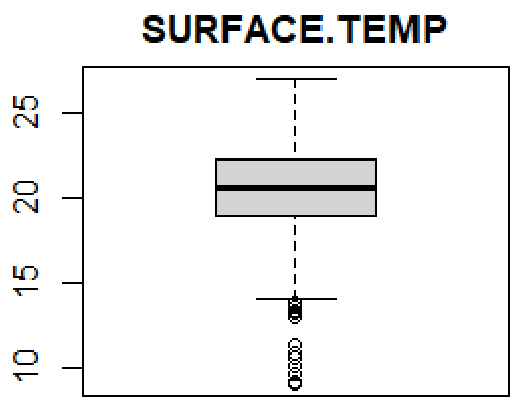


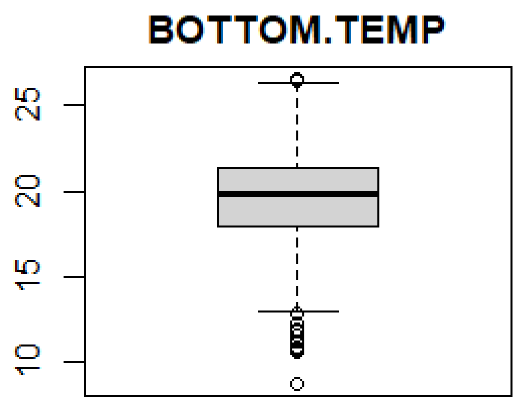
MONTH
Mean Presence/Absence
0.0
0.3
4
5
AREA
Mean Presence/Absence
0.0
0.4
FL
LB
RB
STRATUM
Mean Presence/Absence
0.0
0.6
21
31
41
51
61
DEPTH
Presence/Absence
0.0
0.6
4
6
8
10
12
DEPTH

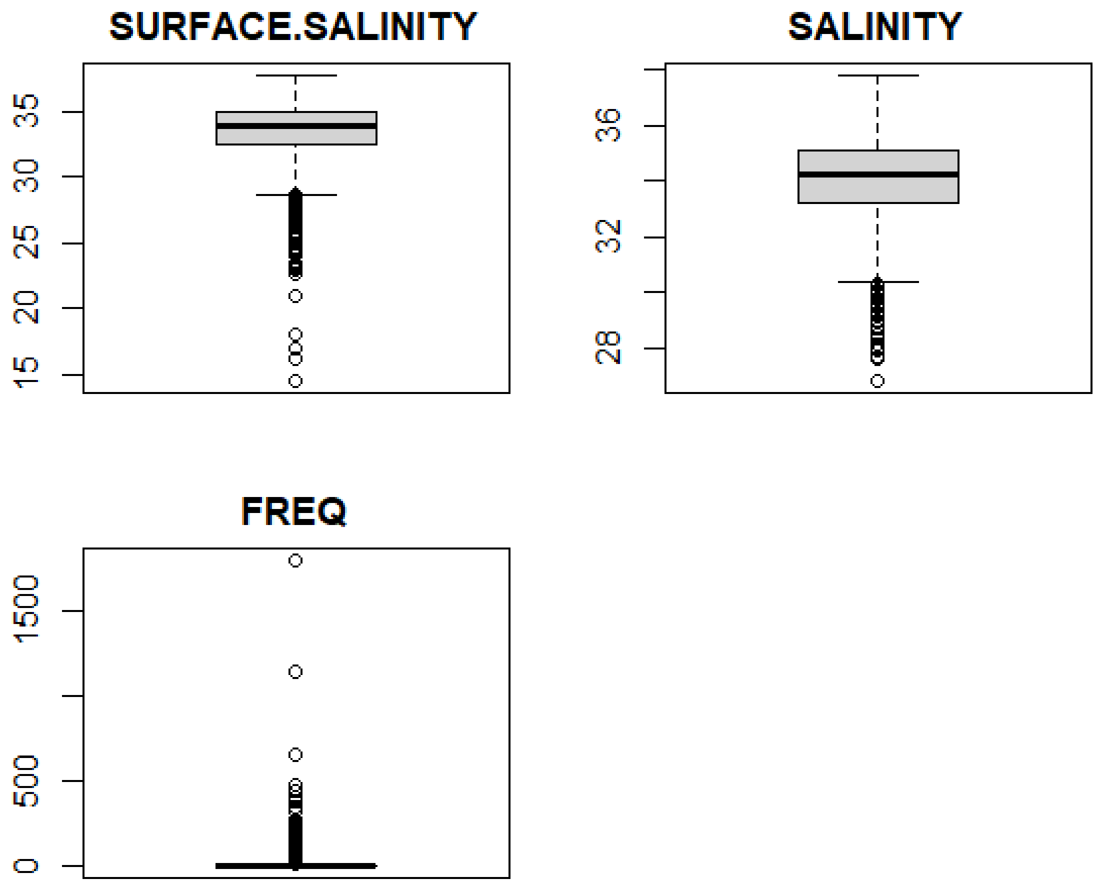

SURFACE.TEMP
Presence/Absence
0.0 0.6
10 15 20 25
SURFACE.TEMP
BOTTOM.TEMP
Presence/Absence
0.0 0.6
10 15 20 25
BOTTOM.TEMP
SURFACE.SALINITY
Presence/Absence
0.0 0.6
15 20 25 30 35
SURFACE.SALINITY
SALINITY
Presence/Absence
0.0 0.6
28 32 36
SALINITY



Check for outliers





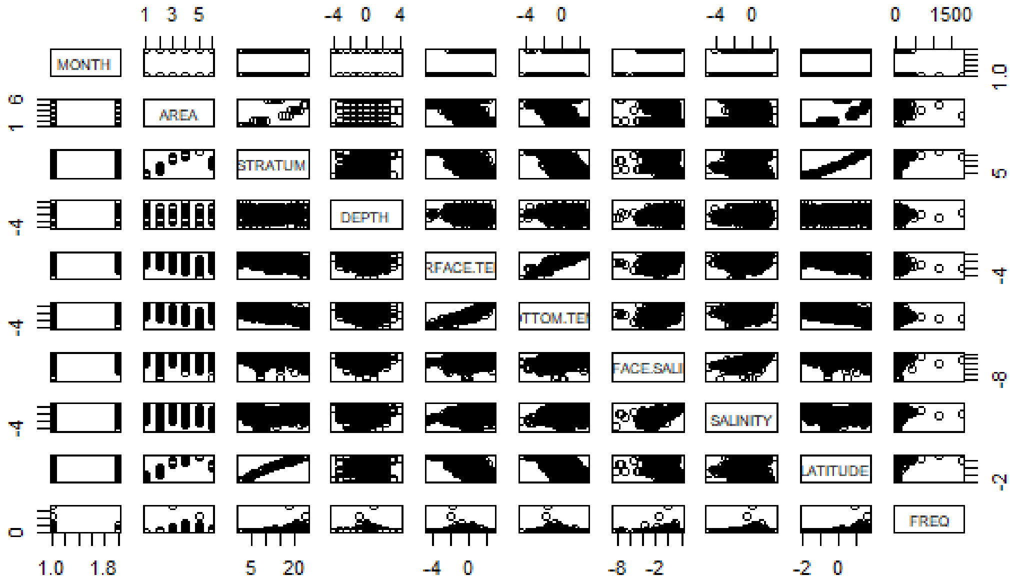


Change continuous variables to z-score

```
##   SAMPLEID YEAR MONTH DAY  TIME  N WEIGHT AREA STRATUM  LATITUDE LONGITUDE
## 1  1989001 1989     4  12 20:32  3  0.274   SC      43 0.0535692    -80.28
## 2  1989003 1989     4  12 23:29  4  0.128   SC      43 0.1041827    -80.16
## 3  1989017 1989     4  17 20:21  7  0.673   OB      57 0.7959012    -77.89
## 4  1989019 1989     4  17 21:47  8  0.661   OB      57 0.8071487    -77.90
## 5  1989021 1989     4  17 23:49 21  1.602   OB      57 0.8971283    -77.90
## 6  1989023 1989     4  18  2:25 13  1.290   OB      59 0.9814842    -77.87
##        DEPTH SURFACE.TEMP BOTTOM.TEMP SURFACE.SALINITY    SALINITY SEASON
## 1 -0.8538112    -2.479324   -2.188642        0.3029857 -0.18284933 SPRING
## 2  0.7576716    -2.356717   -2.106639       -0.1815412 -0.71007699 SPRING
## 3 -0.0480698    -1.948027   -1.655620        0.5232252  0.41028179 SPRING
## 4 -0.8538112    -1.825420   -1.614618        0.2589378  0.01486104 SPRING
## 5 -0.0480698    -1.661944   -2.065637       -2.9565587  0.54208870 SPRING
## 6  0.7576716    -2.234110   -2.393651       -5.4672888  0.54208870 SPRING
##   POS_TOW FREQ
## 1       1    2
## 2       1    1
## 3       1    7
## 4       1    6
```

```
## 5          1    18
## 6          1    11
```

Check colinearity




```
##                   MONTH  AREA STRATUM DEPTH SURFACE.TEMP BOTTOM.TEMP
## MONTH              1.00  0.08    0.11  0.02         0.42        0.40
## AREA               0.08  1.00    0.64 -0.03        -0.37       -0.35
## STRATUM            0.11  0.64    1.00  0.13        -0.55       -0.56
## DEPTH              0.02 -0.03    0.13  1.00        -0.09       -0.10
## SURFACE.TEMP       0.42 -0.37   -0.55 -0.09         1.00        0.93
## BOTTOM.TEMP        0.40 -0.35   -0.56 -0.10         0.93        1.00
## SURFACE.SALINITY   0.06 -0.13   -0.04  0.09         0.15        0.22
## SALINITY           0.08 -0.09    0.03  0.12         0.17        0.18
## LATITUDE           0.11  0.67    0.99  0.12        -0.54       -0.56
## FREQ              -0.04  0.09    0.20  0.02        -0.14       -0.14
##                   SURFACE.SALINITY SALINITY LATITUDE  FREQ
## MONTH                         0.06     0.08     0.11 -0.04
## AREA                         -0.13    -0.09     0.67  0.09
## STRATUM                      -0.04     0.03     0.99  0.20
## DEPTH                         0.09     0.12     0.12  0.02
## SURFACE.TEMP                  0.15     0.17    -0.54 -0.14
## BOTTOM.TEMP                   0.22     0.18    -0.56 -0.14
## SURFACE.SALINITY              1.00     0.81    -0.10 -0.01
## SALINITY                      0.81     1.00    -0.03  0.03
## LATITUDE                     -0.10    -0.03     1.00  0.19
## FREQ                         -0.01     0.03     0.19  1.00
```

Check variance inflation factor - Remove covariates with GVIF > 3, starting with highest value, until all GVIF < 3

```r
lmmod <- lm("FREQ~YEAR+MONTH+STRATUM+DEPTH+BOTTOM.TEMP+SALINITY+LATITUDE", data=new_data2)
vif(lmmod)
```

```
##                    GVIF Df GVIF^(1/(2*Df))
## YEAR           5.951125 30        1.030173
## MONTH          2.068978  1        1.438394
## STRATUM     1345.371703 23        1.169549
## DEPTH          1.353048  1        1.163206
## BOTTOM.TEMP    4.596526  1        2.143951
## SALINITY       3.454332  1        1.858583
## LATITUDE     486.811481  1       22.063805
```

```r
# remove STRATUM
lmmod <- lm("FREQ~YEAR+MONTH+DEPTH+BOTTOM.TEMP+SALINITY+LATITUDE", data=new_data2)
vif(lmmod)
```

```
##                 GVIF Df GVIF^(1/(2*Df))
## YEAR        3.980308 30        1.023290
## MONTH       2.028031  1        1.424089
## DEPTH       1.201662  1        1.096203
## BOTTOM.TEMP 4.311702  1        2.076464
## SALINITY    1.618598  1        1.272241
## LATITUDE    2.572823  1        1.604002
```

```r
# remove bottom temp
lmmod <- lm("FREQ~YEAR+MONTH+DEPTH+SALINITY+LATITUDE", data=new_data2)
vif(lmmod)
```

```
##              GVIF Df GVIF^(1/(2*Df))
## YEAR     2.046720 30        1.012009
## MONTH    1.138030  1        1.066785
## DEPTH    1.192438  1        1.091988
## SALINITY 1.618447  1        1.272182
## LATITUDE 1.044096  1        1.021810
```

```r
# remove month instead
lmmod <- lm("FREQ~YEAR+BOTTOM.TEMP+DEPTH+SALINITY+LATITUDE", data=new_data2)
vif(lmmod)
```

```
##                 GVIF Df GVIF^(1/(2*Df))
## YEAR        2.931298 30        1.018086
## BOTTOM.TEMP 2.419513  1        1.555478
## DEPTH       1.196395  1        1.093798
## SALINITY    1.609517  1        1.268667
## LATITUDE    1.748551  1        1.322328
```

```r
# STRATUM instead of LATITUDE
lmmod <- lm("FREQ~YEAR+BOTTOM.TEMP+DEPTH+SALINITY+STRATUM", data=new_data2)
vif(lmmod)
```

```
##                   GVIF Df GVIF^(1/(2*Df))
## YEAR        4.335238 30        1.024748
## BOTTOM.TEMP 2.601426  1        1.612894
## DEPTH       1.341137  1        1.158075
## SALINITY    3.442887  1        1.855502
## STRATUM     4.744603 23        1.034427
```

```
# AREA instead of LATITUDE
lmmod <- lm("FREQ~YEAR+BOTTOM.TEMP+DEPTH+SALINITY+AREA", data=new_data2)
vif(lmmod)
```

```
##                   GVIF Df GVIF^(1/(2*Df))
## YEAR        3.738173 30        1.022220
## BOTTOM.TEMP 2.481923  1        1.575412
## DEPTH       1.217386  1        1.103352
## SALINITY    2.941709  1        1.715141
## AREA        3.360743  5        1.128869
```

```
# Remove DEPTH
lmmod <- lm("FREQ~YEAR+BOTTOM.TEMP+SALINITY+AREA", data=new_data2)
vif(lmmod)
```

```
##                   GVIF Df GVIF^(1/(2*Df))
## YEAR        3.237557 30        1.019773
## BOTTOM.TEMP 2.471329  1        1.572046
## SALINITY    2.913608  1        1.706929
## AREA        3.286766  5        1.126359
```

Step 2: Model selection

Compare negative binom, zero-infl neg binom, zero-altered neg binom

Comparison 1: full models

```
##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + LATITUDE
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  12010.9  12224.8  -5969.4  11938.9     2780
##
##
## Dispersion parameter for nbinom2 family (): 0.21
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.409446   0.350118   4.026 5.68e-05 ***
## YEAR1990     1.220326   0.435915   2.799 0.005119 **
## YEAR1991     0.371250   0.438934   0.846 0.397665
## YEAR1992     0.662004   0.433609   1.527 0.126828
## YEAR1993    -1.208305   0.473067  -2.554 0.010643 *
## YEAR1994    -0.227804   0.445690  -0.511 0.609262
```

```
## YEAR1995     1.356135   0.434732   3.119 0.001812 **
## YEAR1996    -1.464740   0.449231  -3.261 0.001112 **
## YEAR1997    -0.255578   0.441738  -0.579 0.562876
## YEAR1998    -0.340975   0.458633  -0.743 0.457203
## YEAR1999    -0.265059   0.447671  -0.592 0.553794
## YEAR2000    -0.245669   0.433849  -0.566 0.571221
## YEAR2001    -0.766097   0.428884  -1.786 0.074057 .
## YEAR2002    -0.574479   0.426113  -1.348 0.177598
## YEAR2003     0.074970   0.444593   0.169 0.866090
## YEAR2004     0.868085   0.414840   2.093 0.036387 *
## YEAR2005    -0.965556   0.442755  -2.181 0.029199 *
## YEAR2006     0.392783   0.415977   0.944 0.345045
## YEAR2007     0.024788   0.413817   0.060 0.952234
## YEAR2008    -0.272308   0.421437  -0.646 0.518188
## YEAR2009     0.461192   0.414558   1.112 0.265927
## YEAR2010     0.239198   0.422642   0.566 0.571422
## YEAR2011     0.245503   0.410429   0.598 0.549732
## YEAR2012    -0.670271   0.420227  -1.595 0.110708
## YEAR2013    -0.114466   0.434202  -0.264 0.792070
## YEAR2014     0.148835   0.435089   0.342 0.732292
## YEAR2015    -0.176506   0.411597  -0.429 0.668045
## YEAR2016    -0.463942   0.419263  -1.107 0.268482
## YEAR2017    -1.537229   0.433058  -3.550 0.000386 ***
## YEAR2018    -1.009957   0.467890  -2.159 0.030886 *
## YEAR2019    -0.099292   0.427675  -0.232 0.816409
## DEPTH       -0.190315   0.052238  -3.643 0.000269 ***
## BOTTOM.TEMP -0.524478   0.073621  -7.124 1.05e-12 ***
## SALINITY    -0.002528   0.059693  -0.042 0.966225
## LATITUDE     1.215694   0.061290  19.835  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + LATITUDE
## Zero inflation:        ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + LATITUDE
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  11990.6  12412.6  -5924.3  11848.6     2745
##
##
## Dispersion parameter for nbinom2 family (): 0.25
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.48892    0.33422   4.455 8.39e-06 ***
## YEAR1990     1.13886    0.42537   2.677  0.00742 **
## YEAR1991     0.44284    0.44546   0.994  0.32017
## YEAR1992     0.69576    0.41496   1.677  0.09360 .
## YEAR1993    -1.27302    0.44300  -2.874  0.00406 **
## YEAR1994    -0.37834    0.41576  -0.910  0.36282
```

```
## YEAR1995     1.24623    0.41990   2.968  0.00300 **
## YEAR1996    -0.90507    0.47632  -1.900  0.05742 .
## YEAR1997     0.16314    0.43898   0.372  0.71016
## YEAR1998    -0.25016    0.45007  -0.556  0.57833
## YEAR1999    -0.23100    0.44432  -0.520  0.60314
## YEAR2000     0.00876    0.43981   0.020  0.98411
## YEAR2001    -0.87930    0.40201  -2.187  0.02872 *
## YEAR2002    -0.01455    0.47462  -0.031  0.97554
## YEAR2003    -0.07294    0.42135  -0.173  0.86256
## YEAR2004     1.00690    0.39956   2.520  0.01173 *
## YEAR2005    -0.71880    0.43519  -1.652  0.09860 .
## YEAR2006     1.06912    0.46451   2.302  0.02136 *
## YEAR2007     0.31736    0.41464   0.765  0.44404
## YEAR2008    -0.29412    0.39640  -0.742  0.45811
## YEAR2009     0.89140    0.42006   2.122  0.03383 *
## YEAR2010     0.04132    0.39715   0.104  0.91714
## YEAR2011     0.40552    0.41929   0.967  0.33347
## YEAR2012    -0.83775    0.40369  -2.075  0.03796 *
## YEAR2013    -0.14393    0.41844  -0.344  0.73087
## YEAR2014     0.10893    0.40811   0.267  0.78954
## YEAR2015     0.36850    0.43628   0.845  0.39831
## YEAR2016    -0.12033    0.42661  -0.282  0.77790
## YEAR2017    -1.38629    0.43136  -3.214  0.00131 **
## YEAR2018    -0.38181    0.52300  -0.730  0.46537
## YEAR2019     0.47449    0.47166   1.006  0.31442
## DEPTH       -0.09403    0.05711  -1.646  0.09970 .
## BOTTOM.TEMP -0.36287    0.08165  -4.444 8.81e-06 ***
## SALINITY     0.02310    0.06671   0.346  0.72919
## LATITUDE     1.06775    0.06845  15.600  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.815e+00  4.226e+00  -1.139  0.25457
## YEAR1990     8.891e-01  4.633e+00   0.192  0.84782
## YEAR1991     2.064e+00  4.399e+00   0.469  0.63887
## YEAR1992     8.203e-01  4.918e+00   0.167  0.86754
## YEAR1993    -2.384e+01  6.863e+05   0.000  0.99997
## YEAR1994    -2.678e+01  3.827e+05   0.000  0.99994
## YEAR1995     1.151e+00  4.456e+00   0.258  0.79614
## YEAR1996     4.939e+00  4.322e+00   1.143  0.25313
## YEAR1997     4.320e+00  4.272e+00   1.011  0.31194
## YEAR1998     2.542e+00  4.500e+00   0.565  0.57217
## YEAR1999     1.174e+00  5.199e+00   0.226  0.82140
## YEAR2000     3.554e+00  4.293e+00   0.828  0.40773
## YEAR2001    -1.738e+01  1.348e+04  -0.001  0.99897
## YEAR2002     3.710e+00  4.286e+00   0.866  0.38667
## YEAR2003    -1.023e+01  1.728e+03  -0.006  0.99528
## YEAR2004     1.273e+00  4.387e+00   0.290  0.77166
```

```
## YEAR2005     2.939e+00  4.629e+00   0.635  0.52550
## YEAR2006     4.848e+00  4.255e+00   1.139  0.25455
## YEAR2007     3.255e+00  4.273e+00   0.762  0.44619
## YEAR2008    -2.583e+01  1.222e+06   0.000  0.99998
## YEAR2009     3.363e+00  4.277e+00   0.786  0.43166
## YEAR2010    -2.261e+01  1.484e+05   0.000  0.99988
## YEAR2011     2.610e+00  4.334e+00   0.602  0.54703
## YEAR2012    -5.739e-01  5.159e+00  -0.111  0.91143
## YEAR2013     4.978e-01  5.537e+00   0.090  0.92835
## YEAR2014    -2.193e+01  7.298e+04   0.000  0.99976
## YEAR2015     4.122e+00  4.258e+00   0.968  0.33309
## YEAR2016     3.551e+00  4.282e+00   0.829  0.40698
## YEAR2017     2.131e+00  4.392e+00   0.485  0.62758
## YEAR2018     4.599e+00  4.265e+00   1.078  0.28089
## YEAR2019     4.065e+00  4.226e+00   0.962  0.33605
## DEPTH        5.250e-01  2.356e-01   2.228  0.02589 *
## BOTTOM.TEMP  9.566e-01  2.812e-01   3.402  0.00067 ***
## SALINITY     3.063e-01  2.511e-01   1.220  0.22250
## LATITUDE    -9.096e-01  2.058e-01  -4.419 9.93e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + LATITUDE
## Zero inflation:        ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + LATITUDE
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  11970.7  12392.7  -5914.4  11828.7     2745
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.163
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.87524    0.37597   2.328 0.019915 *
## YEAR1990     1.64648    0.47143   3.493 0.000478 ***
## YEAR1991     0.71615    0.47245   1.516 0.129565
## YEAR1992     0.90872    0.44973   2.021 0.043323 *
## YEAR1993    -0.65830    0.54728  -1.203 0.229030
## YEAR1994     0.09597    0.48692   0.197 0.843754
## YEAR1995     1.62271    0.45952   3.531 0.000414 ***
## YEAR1996    -0.76411    0.51120  -1.495 0.134984
## YEAR1997     0.22583    0.46721   0.483 0.628849
## YEAR1998    -0.10384    0.48930  -0.212 0.831941
## YEAR1999     0.08370    0.48554   0.172 0.863131
## YEAR2000     0.23685    0.47155   0.502 0.615472
## YEAR2001    -0.81337    0.44379  -1.833 0.066833 .
## YEAR2002     0.22378    0.50299   0.445 0.656388
## YEAR2003     0.30480    0.47636   0.640 0.522268
## YEAR2004     1.20952    0.42931   2.817 0.004842 **
```

```
## YEAR2005    -0.21600    0.50363  -0.429 0.668003
## YEAR2006     1.51824    0.50628   2.999 0.002710 **
## YEAR2007     0.50982    0.44310   1.151 0.249909
## YEAR2008     0.11451    0.45400   0.252 0.800868
## YEAR2009     1.09872    0.45410   2.420 0.015539 *
## YEAR2010     0.43635    0.44795   0.974 0.329998
## YEAR2011     0.69015    0.44542   1.549 0.121277
## YEAR2012    -0.81203    0.43453  -1.869 0.061657 .
## YEAR2013    -0.09979    0.44438  -0.225 0.822315
## YEAR2014     0.23860    0.44939   0.531 0.595466
## YEAR2015     0.56645    0.46599   1.216 0.224147
## YEAR2016     0.11012    0.46092   0.239 0.811179
## YEAR2017    -1.17242    0.48552  -2.415 0.015744 *
## YEAR2018    -0.18600    0.56824  -0.327 0.743423
## YEAR2019     0.78865    0.51333   1.536 0.124456
## DEPTH       -0.11726    0.06656  -1.762 0.078137 .
## BOTTOM.TEMP -0.43575    0.09688  -4.498 6.86e-06 ***
## SALINITY     0.01110    0.07881   0.141 0.887995
## LATITUDE     1.16990    0.08332  14.040  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.71127    0.35363  -2.011 0.044289 *
## YEAR1990     0.55288    0.43198   1.280 0.200590
## YEAR1991     0.77129    0.43284   1.782 0.074760 .
## YEAR1992     0.29560    0.43161   0.685 0.493419
## YEAR1993     1.90917    0.45063   4.237 2.27e-05 ***
## YEAR1994     1.12955    0.43033   2.625 0.008669 **
## YEAR1995     0.19925    0.43353   0.460 0.645800
## YEAR1996     1.72179    0.44047   3.909 9.27e-05 ***
## YEAR1997     0.86743    0.43543   1.992 0.046356 *
## YEAR1998     0.90027    0.44607   2.018 0.043569 *
## YEAR1999     0.95373    0.42943   2.221 0.026356 *
## YEAR2000     1.04251    0.42756   2.438 0.014758 *
## YEAR2001     0.72953    0.41280   1.767 0.077186 .
## YEAR2002     1.46706    0.42374   3.462 0.000536 ***
## YEAR2003     0.88556    0.42907   2.064 0.039028 *
## YEAR2004     0.24898    0.41748   0.596 0.550913
## YEAR2005     1.93618    0.43885   4.412 1.02e-05 ***
## YEAR2006     1.76075    0.42369   4.156 3.24e-05 ***
## YEAR2007     0.81782    0.41098   1.990 0.046600 *
## YEAR2008     1.13450    0.41286   2.748 0.005998 **
## YEAR2009     0.86748    0.40775   2.127 0.033382 *
## YEAR2010     0.82055    0.40696   2.016 0.043772 *
## YEAR2011     0.86872    0.40756   2.132 0.033048 *
## YEAR2012     0.56199    0.40892   1.374 0.169335
## YEAR2013     0.35300    0.42549   0.830 0.406751
## YEAR2014     0.47734    0.41899   1.139 0.254601
```

```
## YEAR2015     1.30721    0.41341   3.162 0.001567 **
## YEAR2016     1.23963    0.41021   3.022 0.002511 **
## YEAR2017     1.74701    0.42721   4.089 4.33e-05 ***
## YEAR2018     1.62830    0.46068   3.535 0.000408 ***
## YEAR2019     1.36860    0.43041   3.180 0.001474 **
## DEPTH        0.14963    0.04584   3.264 0.001097 **
## BOTTOM.TEMP  0.24426    0.06475   3.773 0.000162 ***
## SALINITY     0.09429    0.05460   1.727 0.084183 .
## LATITUDE    -0.64495    0.05712 -11.292  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##           dAIC df
## zero_adj   0.0 71
## zero_infl 19.9 71
## neg_binom 40.2 36
```

Try stratum instead of latitude.

```
##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  11925.9  12270.6  -5904.9  11809.9     2758
##
##
## Dispersion parameter for nbinom2 family (): 0.226
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.10337    0.50321   0.205 0.837243
## YEAR1990     0.63483    0.43427   1.462 0.143790
## YEAR1991    -0.26531    0.43307  -0.613 0.540122
## YEAR1992     0.09462    0.42486   0.223 0.823764
## YEAR1993    -2.10176    0.46940  -4.477 7.55e-06 ***
## YEAR1994    -0.67633    0.44006  -1.537 0.124320
## YEAR1995     0.53018    0.42938   1.235 0.216925
## YEAR1996    -1.82266    0.44606  -4.086 4.39e-05 ***
## YEAR1997    -0.69806    0.43893  -1.590 0.111750
## YEAR1998    -1.55304    0.47288  -3.284 0.001023 **
## YEAR1999    -0.53714    0.44019  -1.220 0.222379
## YEAR2000    -0.70167    0.42796  -1.640 0.101096
## YEAR2001    -1.00967    0.42467  -2.378 0.017428 *
## YEAR2002    -0.82116    0.42555  -1.930 0.053649 .
## YEAR2003    -1.16733    0.45289  -2.578 0.009952 **
## YEAR2004     0.57287    0.41326   1.386 0.165674
## YEAR2005    -1.27851    0.44536  -2.871 0.004095 **
## YEAR2006    -0.19554    0.41508  -0.471 0.637579
## YEAR2007    -0.22082    0.41120  -0.537 0.591265
```

```
## YEAR2008    -0.76048    0.41900  -1.815 0.069528 .
## YEAR2009    -0.12064    0.40967  -0.294 0.768396
## YEAR2010    -0.52245    0.42574  -1.227 0.219766
## YEAR2011    -0.30942    0.41201  -0.751 0.452660
## YEAR2012    -0.98724    0.41430  -2.383 0.017176 *
## YEAR2013    -0.48601    0.42676  -1.139 0.254767
## YEAR2014    -0.15440    0.42997  -0.359 0.719524
## YEAR2015    -0.78196    0.41453  -1.886 0.059243 .
## YEAR2016    -1.48647    0.42079  -3.533 0.000412 ***
## YEAR2017    -1.71657    0.43666  -3.931 8.46e-05 ***
## YEAR2018    -1.74745    0.47167  -3.705 0.000212 ***
## YEAR2019    -1.12232    0.43181  -2.599 0.009347 **
## DEPTH       -0.20030    0.05485  -3.652 0.000261 ***
## BOTTOM.TEMP -0.36428    0.07580  -4.806 1.54e-06 ***
## SALINITY    -0.43617    0.08939  -4.879 1.06e-06 ***
## STRATUM23   -0.14012    0.46548  -0.301 0.763396
## STRATUM25    0.53669    0.41886   1.281 0.200081
## STRATUM27    0.57273    0.40707   1.407 0.159441
## STRATUM29    0.52307    0.41785   1.252 0.210643
## STRATUM31    1.05134    0.39928   2.633 0.008461 **
## STRATUM33    0.38173    0.41463   0.921 0.357226
## STRATUM35   -0.10643    0.42019  -0.253 0.800052
## STRATUM37    0.15472    0.43452   0.356 0.721786
## STRATUM39    1.07303    0.42323   2.535 0.011234 *
## STRATUM41    0.49292    0.40753   1.210 0.226457
## STRATUM43    2.09732    0.42552   4.929 8.27e-07 ***
## STRATUM45    2.74544    0.43326   6.337 2.35e-10 ***
## STRATUM47    1.86263    0.40438   4.606 4.10e-06 ***
## STRATUM49    1.61681    0.40095   4.032 5.52e-05 ***
## STRATUM51    2.01561    0.39644   5.084 3.69e-07 ***
## STRATUM53    2.62365    0.40892   6.416 1.40e-10 ***
## STRATUM55    3.69799    0.41003   9.019  < 2e-16 ***
## STRATUM57    2.96102    0.40853   7.248 4.23e-13 ***
## STRATUM59    3.83232    0.41221   9.297  < 2e-16 ***
## STRATUM61    4.02964    0.42094   9.573  < 2e-16 ***
## STRATUM63    3.68222    0.41987   8.770  < 2e-16 ***
## STRATUM65    3.80056    0.41349   9.191  < 2e-16 ***
## STRATUM67    4.21655    0.41346  10.198  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## Zero inflation:        ~DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  11916.8  12421.9  -5873.4  11746.8     2731
##
##
## Dispersion parameter for nbinom2 family (): 0.275
```

```
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.03571    0.61606  -0.058 0.953774
## YEAR1990     0.76801    0.43799   1.753 0.079516 .
## YEAR1991    -0.23546    0.43059  -0.547 0.584499
## YEAR1992     0.28854    0.41844   0.690 0.490471
## YEAR1993    -1.84402    0.45443  -4.058 4.95e-05 ***
## YEAR1994    -0.65596    0.43892  -1.494 0.135052
## YEAR1995     0.69051    0.42480   1.625 0.104059
## YEAR1996    -1.57001    0.43761  -3.588 0.000334 ***
## YEAR1997    -0.38426    0.43326  -0.887 0.375134
## YEAR1998    -1.16580    0.46317  -2.517 0.011835 *
## YEAR1999    -0.42529    0.43561  -0.976 0.328912
## YEAR2000    -0.51140    0.42350  -1.208 0.227214
## YEAR2001    -1.03343    0.42164  -2.451 0.014247 *
## YEAR2002    -0.49897    0.44391  -1.124 0.260995
## YEAR2003    -0.96778    0.43989  -2.200 0.027802 *
## YEAR2004     0.62150    0.40974   1.517 0.129317
## YEAR2005    -1.10103    0.43586  -2.526 0.011534 *
## YEAR2006    -0.07897    0.41658  -0.190 0.849643
## YEAR2007    -0.14211    0.41338  -0.344 0.731021
## YEAR2008    -0.61008    0.41751  -1.461 0.143953
## YEAR2009     0.19874    0.40923   0.486 0.627217
## YEAR2010    -0.58319    0.42446  -1.374 0.169454
## YEAR2011    -0.12112    0.41362  -0.293 0.769646
## YEAR2012    -1.01186    0.41342  -2.448 0.014384 *
## YEAR2013    -0.43312    0.41840  -1.035 0.300579
## YEAR2014    -0.08680    0.42589  -0.204 0.838503
## YEAR2015    -0.58025    0.40980  -1.416 0.156793
## YEAR2016    -1.37464    0.41477  -3.314 0.000919 ***
## YEAR2017    -1.73295    0.44309  -3.911 9.19e-05 ***
## YEAR2018    -1.67081    0.47146  -3.544 0.000394 ***
## YEAR2019    -0.82322    0.43919  -1.874 0.060874 .
## DEPTH       -0.20992    0.06121  -3.429 0.000605 ***
## BOTTOM.TEMP -0.23376    0.08172  -2.861 0.004228 **
## SALINITY    -0.28391    0.09235  -3.074 0.002111 **
## STRATUM23    0.38713    0.68767   0.563 0.573456
## STRATUM25    0.64826    0.58322   1.112 0.266343
## STRATUM27    0.96513    0.56711   1.702 0.088786 .
## STRATUM29    0.65661    0.58397   1.124 0.260844
## STRATUM31    0.99446    0.52325   1.901 0.057360 .
## STRATUM33    0.95840    0.57936   1.654 0.098077 .
## STRATUM35    0.23988    0.57432   0.418 0.676185
## STRATUM37    0.98056    0.59178   1.657 0.097529 .
## STRATUM39    1.21633    0.55737   2.182 0.029090 *
## STRATUM41    0.48788    0.53490   0.912 0.361720
## STRATUM43    2.11493    0.54980   3.847 0.000120 ***
## STRATUM45    3.08496    0.58401   5.282 1.28e-07 ***
## STRATUM47    2.01370    0.55918   3.601 0.000317 ***
```

```
## STRATUM49    1.93319    0.54798   3.528 0.000419 ***
## STRATUM51    2.19902    0.54525   4.033 5.51e-05 ***
## STRATUM53    2.83758    0.53922   5.262 1.42e-07 ***
## STRATUM55    3.72659    0.53273   6.995 2.65e-12 ***
## STRATUM57    3.48694    0.55204   6.317 2.68e-10 ***
## STRATUM59    3.81927    0.53372   7.156 8.31e-13 ***
## STRATUM61    3.97721    0.54173   7.342 2.11e-13 ***
## STRATUM63    3.66969    0.54594   6.722 1.79e-11 ***
## STRATUM65    3.89346    0.54098   7.197 6.15e-13 ***
## STRATUM67    4.20902    0.53127   7.923 2.33e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.86969    1.66633  -1.722 0.085040 .
## DEPTH         -0.01733    0.18385  -0.094 0.924919
## BOTTOM.TEMP    0.65454    0.18995   3.446 0.000569 ***
## SALINITY       0.72838    0.24402   2.985 0.002836 **
## STRATUM23      1.62086    1.74097   0.931 0.351846
## STRATUM25      0.49623    1.80394   0.275 0.783255
## STRATUM27      1.34863    1.67590   0.805 0.420983
## STRATUM29      0.94982    1.90796   0.498 0.618610
## STRATUM31    -15.67983 2431.32164  -0.006 0.994854
## STRATUM33      2.30297    1.75260   1.314 0.188836
## STRATUM35      1.68794    1.94010   0.870 0.384287
## STRATUM37      3.14893    1.79847   1.751 0.079965 .
## STRATUM39      0.17584    2.66818   0.066 0.947454
## STRATUM41    -15.07710 2515.86981  -0.006 0.995218
## STRATUM43    -15.18627 2968.66735  -0.005 0.995918
## STRATUM45      1.44153    1.79042   0.805 0.420740
## STRATUM47      0.26282    2.08130   0.126 0.899511
## STRATUM49      1.49769    1.74485   0.858 0.390701
## STRATUM51      1.02255    1.77906   0.575 0.565445
## STRATUM53      0.79028    1.80614   0.438 0.661713
## STRATUM55    -17.32283 4134.39256  -0.004 0.996657
## STRATUM57      2.33881    1.64585   1.421 0.155307
## STRATUM59    -17.15622 4354.65154  -0.004 0.996857
## STRATUM61    -17.33668 4166.98431  -0.004 0.996680
## STRATUM63     -1.13020    3.91991  -0.288 0.773099
## STRATUM65      0.74112    1.71652   0.432 0.665918
## STRATUM67    -14.92348 2369.77084  -0.006 0.994975
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## Zero inflation:        ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
```

```
##  11876.1  12559.5  -5823.0  11646.1     2701
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.206
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.36105    0.68183  -0.530 0.596443
## YEAR1990     1.21089    0.47527   2.548 0.010842 *
## YEAR1991     0.20465    0.47601   0.430 0.667242
## YEAR1992     0.48746    0.44477   1.096 0.273084
## YEAR1993    -1.07403    0.55440  -1.937 0.052708 .
## YEAR1994    -0.18156    0.49265  -0.369 0.712481
## YEAR1995     0.96324    0.45625   2.111 0.034756 *
## YEAR1996    -1.06331    0.51331  -2.071 0.038314 *
## YEAR1997     0.01576    0.47302   0.033 0.973414
## YEAR1998    -0.67768    0.52249  -1.297 0.194624
## YEAR1999    -0.30607    0.48109  -0.636 0.524644
## YEAR2000    -0.07721    0.46811  -0.165 0.868987
## YEAR2001    -1.20780    0.44942  -2.687 0.007200 **
## YEAR2002    -0.01734    0.50343  -0.034 0.972523
## YEAR2003    -0.46582    0.50121  -0.929 0.352687
## YEAR2004     0.79766    0.43373   1.839 0.065906 .
## YEAR2005    -0.40724    0.51306  -0.794 0.427344
## YEAR2006     0.98205    0.50496   1.945 0.051795 .
## YEAR2007     0.11261    0.44545   0.253 0.800416
## YEAR2008    -0.32188    0.45791  -0.703 0.482091
## YEAR2009     0.57220    0.45148   1.267 0.205018
## YEAR2010    -0.53110    0.46834  -1.134 0.256798
## YEAR2011     0.23292    0.45717   0.509 0.610413
## YEAR2012    -1.29547    0.43382  -2.986 0.002824 **
## YEAR2013    -0.40017    0.44778  -0.894 0.371496
## YEAR2014    -0.03718    0.45675  -0.081 0.935121
## YEAR2015     0.12799    0.47326   0.270 0.786813
## YEAR2016    -0.85053    0.47247  -1.800 0.071829 .
## YEAR2017    -1.74824    0.51523  -3.393 0.000691 ***
## YEAR2018    -1.20529    0.57333  -2.102 0.035531 *
## YEAR2019     0.01347    0.52438   0.026 0.979501
## DEPTH       -0.14109    0.06954  -2.029 0.042455 *
## BOTTOM.TEMP -0.32859    0.09797  -3.354 0.000796 ***
## SALINITY    -0.17435    0.11178  -1.560 0.118807
## STRATUM23    0.17652    0.78421   0.225 0.821912
## STRATUM25    0.15178    0.66134   0.230 0.818480
## STRATUM27    1.11359    0.67025   1.661 0.096623 .
## STRATUM29    0.38565    0.64039   0.602 0.547032
## STRATUM31    0.60768    0.59467   1.022 0.306844
## STRATUM33    1.06281    0.65276   1.628 0.103490
## STRATUM35    0.29550    0.63740   0.464 0.642932
## STRATUM37    0.82163    0.66510   1.235 0.216699
## STRATUM39    1.13403    0.63020   1.799 0.071942 .
```

```
## STRATUM41    0.54098    0.62210   0.870 0.384517
## STRATUM43    1.99734    0.62648   3.188 0.001432 **
## STRATUM45    2.93460    0.64915   4.521 6.16e-06 ***
## STRATUM47    1.90286    0.61029   3.118 0.001821 **
## STRATUM49    1.58989    0.60729   2.618 0.008844 **
## STRATUM51    1.95723    0.60239   3.249 0.001158 **
## STRATUM53    2.68342    0.60788   4.414 1.01e-05 ***
## STRATUM55    3.36875    0.60526   5.566 2.61e-08 ***
## STRATUM57    3.41238    0.63451   5.378 7.53e-08 ***
## STRATUM59    3.27105    0.60925   5.369 7.92e-08 ***
## STRATUM61    3.42478    0.61293   5.588 2.30e-08 ***
## STRATUM63    3.29122    0.62055   5.304 1.13e-07 ***
## STRATUM65    3.68858    0.61942   5.955 2.60e-09 ***
## STRATUM67    4.14023    0.61271   6.757 1.41e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.20751    0.48391   0.429 0.668046
## YEAR1990     0.69973    0.43960   1.592 0.111440
## YEAR1991     1.06482    0.44253   2.406 0.016119 *
## YEAR1992     0.35743    0.43740   0.817 0.413839
## YEAR1993     2.32371    0.46560   4.991 6.01e-07 ***
## YEAR1994     1.25220    0.43841   2.856 0.004287 **
## YEAR1995     0.36189    0.43966   0.823 0.410441
## YEAR1996     1.77114    0.44729   3.960 7.51e-05 ***
## YEAR1997     1.06626    0.44427   2.400 0.016395 *
## YEAR1998     1.42305    0.46693   3.048 0.002306 **
## YEAR1999     0.94635    0.43577   2.172 0.029882 *
## YEAR2000     1.15256    0.43475   2.651 0.008023 **
## YEAR2001     0.64339    0.42041   1.530 0.125927
## YEAR2002     1.41551    0.43236   3.274 0.001061 **
## YEAR2003     1.34879    0.44399   3.038 0.002382 **
## YEAR2004     0.21108    0.42493   0.497 0.619369
## YEAR2005     2.11520    0.44941   4.707 2.52e-06 ***
## YEAR2006     1.91804    0.43155   4.445 8.81e-06 ***
## YEAR2007     0.80328    0.41792   1.922 0.054598 .
## YEAR2008     1.19909    0.41967   2.857 0.004274 **
## YEAR2009     0.91052    0.41488   2.195 0.028187 *
## YEAR2010     0.96944    0.41417   2.341 0.019250 *
## YEAR2011     0.98142    0.41651   2.356 0.018458 *
## YEAR2012     0.53537    0.41676   1.285 0.198930
## YEAR2013     0.53820    0.43215   1.245 0.212985
## YEAR2014     0.55345    0.42560   1.300 0.193468
## YEAR2015     1.56000    0.42214   3.695 0.000220 ***
## YEAR2016     1.56790    0.42096   3.725 0.000196 ***
## YEAR2017     1.87116    0.43626   4.289 1.79e-05 ***
## YEAR2018     1.82058    0.46936   3.879 0.000105 ***
## YEAR2019     1.72899    0.44280   3.905 9.44e-05 ***
```

```
## DEPTH        0.18265    0.04939   3.698 0.000217 ***
## BOTTOM.TEMP  0.14821    0.06804   2.178 0.029399 *
## SALINITY     0.37619    0.07989   4.709 2.49e-06 ***
## STRATUM23    0.12545    0.44810   0.280 0.779499
## STRATUM25   -0.53478    0.38285  -1.397 0.162470
## STRATUM27   -0.27603    0.38702  -0.713 0.475713
## STRATUM29   -0.44373    0.38985  -1.138 0.255036
## STRATUM31   -1.21083    0.36559  -3.312 0.000926 ***
## STRATUM33   -0.10963    0.38184  -0.287 0.774033
## STRATUM35    0.01757    0.38429   0.046 0.963528
## STRATUM37    0.22642    0.40224   0.563 0.573504
## STRATUM39   -0.81188    0.38786  -2.093 0.036331 *
## STRATUM41   -0.36134    0.37155  -0.973 0.330793
## STRATUM43   -1.34218    0.38827  -3.457 0.000547 ***
## STRATUM45   -1.34812    0.38490  -3.503 0.000461 ***
## STRATUM47   -1.38922    0.36839  -3.771 0.000163 ***
## STRATUM49   -1.02968    0.36251  -2.840 0.004505 **
## STRATUM51   -1.31421    0.36135  -3.637 0.000276 ***
## STRATUM53   -1.53917    0.37398  -4.116 3.86e-05 ***
## STRATUM55   -2.55595    0.38631  -6.616 3.68e-11 ***
## STRATUM57   -1.01717    0.36873  -2.759 0.005806 **
## STRATUM59   -2.58569    0.39082  -6.616 3.69e-11 ***
## STRATUM61   -2.67921    0.40906  -6.550 5.76e-11 ***
## STRATUM63   -2.19264    0.39600  -5.537 3.08e-08 ***
## STRATUM65   -1.88988    0.39091  -4.835 1.33e-06 ***
## STRATUM67   -2.11438    0.40417  -5.231 1.68e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##            dAIC  df
## zero_adj    0.0 115
## zero_infl  40.7 85
## neg_binom  49.8 58
```

Try AREA instead of STRATUM

```
##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + AREA
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##  12024.2  12261.9  -5972.1  11944.2     2776
##
##
## Dispersion parameter for nbinom2 family (): 0.208
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.13614    0.39688   0.343 0.731585
## YEAR1990     1.19332    0.44486   2.682 0.007308 **
```

```
## YEAR1991     0.14805    0.44202   0.335 0.737668    
## YEAR1992     0.46230    0.43527   1.062 0.288193    
## YEAR1993    -1.65529    0.47912  -3.455 0.000551 ***
## YEAR1994    -0.30048    0.44446  -0.676 0.499001    
## YEAR1995     1.19737    0.43837   2.731 0.006306 ** 
## YEAR1996    -1.58290    0.45248  -3.498 0.000468 ***
## YEAR1997    -0.61256    0.44527  -1.376 0.168919    
## YEAR1998    -0.92645    0.47454  -1.952 0.050904 .  
## YEAR1999    -0.41478    0.44632  -0.929 0.352723    
## YEAR2000    -0.48113    0.43746  -1.100 0.271406    
## YEAR2001    -0.79421    0.43100  -1.843 0.065372 .  
## YEAR2002    -0.42090    0.42916  -0.981 0.326716    
## YEAR2003    -0.52853    0.45790  -1.154 0.248401    
## YEAR2004     0.77494    0.41937   1.848 0.064619 .  
## YEAR2005    -1.17443    0.45007  -2.609 0.009069 ** 
## YEAR2006     0.25878    0.42227   0.613 0.539981    
## YEAR2007    -0.01006    0.41903  -0.024 0.980847    
## YEAR2008    -0.41710    0.42555  -0.980 0.327009    
## YEAR2009     0.37260    0.41581   0.896 0.370211    
## YEAR2010    -0.13795    0.43088  -0.320 0.748844    
## YEAR2011     0.18829    0.41820   0.450 0.652529    
## YEAR2012    -0.58886    0.42179  -1.396 0.162690    
## YEAR2013    -0.23399    0.43687  -0.536 0.592235    
## YEAR2014    -0.16063    0.43544  -0.369 0.712212    
## YEAR2015    -0.37044    0.42114  -0.880 0.379068    
## YEAR2016    -0.83034    0.42659  -1.946 0.051601 .  
## YEAR2017    -1.70908    0.44232  -3.864 0.000112 ***
## YEAR2018    -1.13178    0.46987  -2.409 0.016008 *  
## YEAR2019    -0.44712    0.43483  -1.028 0.303825    
## DEPTH       -0.19577    0.05332  -3.672 0.000241 ***
## BOTTOM.TEMP -0.52790    0.07521  -7.019 2.24e-12 ***
## SALINITY    -0.28498    0.08141  -3.501 0.000464 ***
## AREAGA       0.34433    0.19728   1.745 0.080913 .  
## AREALB       2.46625    0.20666  11.934  < 2e-16 ***
## AREAOB       3.06466    0.20044  15.290  < 2e-16 ***
## AREARB       3.46629    0.22600  15.338  < 2e-16 ***
## AREASC       1.49491    0.19194   7.788 6.79e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + AREA
## Zero inflation:        ~DEPTH + BOTTOM.TEMP + SALINITY + AREA
## Data: data
## 
##      AIC      BIC   logLik deviance df.resid 
##  12032.5  12323.8  -5967.3  11934.5     2767 
## 
## 
## Dispersion parameter for nbinom2 family (): 0.222 
## 
```

```
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.26202    0.42243   0.620 0.535079
## YEAR1990     1.18815    0.44970   2.642 0.008240 **
## YEAR1991     0.17996    0.44742   0.402 0.687523
## YEAR1992     0.48022    0.43634   1.101 0.271086
## YEAR1993    -1.66023    0.47653  -3.484 0.000494 ***
## YEAR1994    -0.31455    0.44720  -0.703 0.481817
## YEAR1995     1.14980    0.44080   2.608 0.009096 **
## YEAR1996    -1.49890    0.45343  -3.306 0.000947 ***
## YEAR1997    -0.50907    0.44508  -1.144 0.252721
## YEAR1998    -0.88995    0.47533  -1.872 0.061167 .
## YEAR1999    -0.40095    0.44889  -0.893 0.371745
## YEAR2000    -0.43530    0.43977  -0.990 0.322254
## YEAR2001    -0.83031    0.43480  -1.910 0.056177 .
## YEAR2002    -0.33447    0.44260  -0.756 0.449832
## YEAR2003    -0.53149    0.45869  -1.159 0.246574
## YEAR2004     0.82966    0.42494   1.952 0.050887 .
## YEAR2005    -1.13200    0.45257  -2.501 0.012375 *
## YEAR2006     0.33049    0.42952   0.769 0.441626
## YEAR2007     0.03628    0.42495   0.085 0.931954
## YEAR2008    -0.42737    0.42848  -0.997 0.318561
## YEAR2009     0.47551    0.42285   1.125 0.260789
## YEAR2010    -0.17632    0.43851  -0.402 0.687615
## YEAR2011     0.20935    0.42597   0.491 0.623097
## YEAR2012    -0.60333    0.42908  -1.406 0.159690
## YEAR2013    -0.26994    0.43904  -0.615 0.538660
## YEAR2014    -0.16620    0.43766  -0.380 0.704143
## YEAR2015    -0.29160    0.42567  -0.685 0.493324
## YEAR2016    -0.76943    0.42926  -1.792 0.073060 .
## YEAR2017    -1.66029    0.44536  -3.728 0.000193 ***
## YEAR2018    -1.07094    0.47234  -2.267 0.023370 *
## YEAR2019    -0.35981    0.44072  -0.816 0.414257
## DEPTH       -0.13942    0.06030  -2.312 0.020761 *
## BOTTOM.TEMP -0.47247    0.08129  -5.812 6.16e-09 ***
## SALINITY    -0.23679    0.08562  -2.766 0.005682 **
## AREAGA       0.37087    0.24419   1.519 0.128819
## AREALB       2.32185    0.24976   9.296  < 2e-16 ***
## AREAOB       2.87760    0.24146  11.917  < 2e-16 ***
## AREARB       3.28892    0.25660  12.817  < 2e-16 ***
## AREASC       1.47066    0.24287   6.055 1.40e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.61515    0.89455  -2.923  0.00346 **
## DEPTH          0.62858    0.31487   1.996  0.04590 *
## BOTTOM.TEMP    0.71177    0.30138   2.362  0.01819 *
## SALINITY       0.59392    0.36666   1.620  0.10528
```

```
## AREAGA         0.48588    0.93508   0.520  0.60333
## AREALB        -2.00768    2.77327  -0.724  0.46910
## AREAOB       -12.31059 1741.90943  -0.007  0.99436
## AREARB       -12.95452 1671.00420  -0.008  0.99381
## AREASC         0.02882    0.97077   0.030  0.97631
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + AREA
## Zero inflation:        ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + AREA
## Data: data
## 
##      AIC      BIC   logLik deviance df.resid 
##  11998.0  12467.5  -5920.0  11840.0     2737 
## 
## 
## Dispersion parameter for truncated_nbinom2 family (): 0.159 
## 
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept) -0.29883    0.46765  -0.639 0.522814    
## YEAR1990     1.66252    0.48861   3.403 0.000668 ***
## YEAR1991     0.60564    0.48071   1.260 0.207706    
## YEAR1992     0.74220    0.45478   1.632 0.102681    
## YEAR1993    -0.99735    0.55986  -1.781 0.074844 .  
## YEAR1994     0.09341    0.48909   0.191 0.848533    
## YEAR1995     1.53418    0.46752   3.282 0.001032 ** 
## YEAR1996    -0.87683    0.52107  -1.683 0.092425 .  
## YEAR1997    -0.14276    0.47445  -0.301 0.763495    
## YEAR1998    -0.55888    0.51709  -1.081 0.279778    
## YEAR1999    -0.09373    0.48486  -0.193 0.846714    
## YEAR2000     0.03020    0.47907   0.063 0.949728    
## YEAR2001    -0.86802    0.45372  -1.913 0.055730 .  
## YEAR2002     0.34214    0.50987   0.671 0.502190    
## YEAR2003    -0.18525    0.49877  -0.371 0.710331    
## YEAR2004     1.06273    0.44038   2.413 0.015813 *  
## YEAR2005    -0.39833    0.52007  -0.766 0.443723    
## YEAR2006     1.39944    0.51808   2.701 0.006909 ** 
## YEAR2007     0.46136    0.45591   1.012 0.311560    
## YEAR2008    -0.03770    0.46557  -0.081 0.935454    
## YEAR2009     0.98631    0.45751   2.156 0.031097 *  
## YEAR2010    -0.01220    0.46950  -0.026 0.979274    
## YEAR2011     0.64812    0.46138   1.405 0.160097    
## YEAR2012    -0.73468    0.44026  -1.669 0.095171 .  
## YEAR2013    -0.14927    0.45551  -0.328 0.743138    
## YEAR2014    -0.04914    0.45565  -0.108 0.914121    
## YEAR2015     0.44739    0.48380   0.925 0.355103    
## YEAR2016    -0.21135    0.47482  -0.445 0.656237    
## YEAR2017    -1.42856    0.49884  -2.864 0.004186 ** 
## YEAR2018    -0.34921    0.57276  -0.610 0.542056    
```

```
## YEAR2019     0.50450    0.52409   0.963 0.335737
## DEPTH       -0.13399    0.06907  -1.940 0.052400 .
## BOTTOM.TEMP -0.48289    0.09797  -4.929 8.27e-07 ***
## SALINITY    -0.19690    0.10304  -1.911 0.056023 .
## AREAGA       0.30754    0.28519   1.078 0.280867
## AREALB       2.25294    0.28867   7.805 5.97e-15 ***
## AREAOB       2.72778    0.28319   9.632  < 2e-16 ***
## AREARB       3.32016    0.30271  10.968  < 2e-16 ***
## AREASC       1.39308    0.27698   5.030 4.92e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.09090    0.38238   0.238 0.812095
## YEAR1990     0.56739    0.43389   1.308 0.190977
## YEAR1991     0.83302    0.43635   1.909 0.056255 .
## YEAR1992     0.34477    0.43219   0.798 0.425032
## YEAR1993     2.09309    0.45714   4.579 4.68e-06 ***
## YEAR1994     1.14583    0.43147   2.656 0.007916 **
## YEAR1995     0.25452    0.43461   0.586 0.558121
## YEAR1996     1.78383    0.44137   4.042 5.31e-05 ***
## YEAR1997     0.95371    0.43711   2.182 0.029123 *
## YEAR1998     1.10527    0.45663   2.421 0.015499 *
## YEAR1999     0.97311    0.43051   2.260 0.023797 *
## YEAR2000     1.08383    0.42869   2.528 0.011463 *
## YEAR2001     0.68820    0.41302   1.666 0.095658 .
## YEAR2002     1.35208    0.42425   3.187 0.001438 **
## YEAR2003     1.03953    0.43589   2.385 0.017085 *
## YEAR2004     0.23058    0.41842   0.551 0.581586
## YEAR2005     2.03201    0.44034   4.615 3.94e-06 ***
## YEAR2006     1.76018    0.42463   4.145 3.39e-05 ***
## YEAR2007     0.77637    0.41116   1.888 0.058995 .
## YEAR2008     1.13397    0.41304   2.745 0.006043 **
## YEAR2009     0.85683    0.40892   2.095 0.036138 *
## YEAR2010     0.84866    0.40773   2.081 0.037397 *
## YEAR2011     0.85191    0.40833   2.086 0.036949 *
## YEAR2012     0.50074    0.40918   1.224 0.221048
## YEAR2013     0.45284    0.42664   1.061 0.288497
## YEAR2014     0.53031    0.41930   1.265 0.205957
## YEAR2015     1.35943    0.41511   3.275 0.001057 **
## YEAR2016     1.30510    0.41311   3.159 0.001582 **
## YEAR2017     1.70767    0.42827   3.987 6.68e-05 ***
## YEAR2018     1.63639    0.46062   3.553 0.000381 ***
## YEAR2019     1.43815    0.43318   3.320 0.000900 ***
## DEPTH        0.14553    0.04618   3.152 0.001624 **
## BOTTOM.TEMP  0.26350    0.06524   4.039 5.36e-05 ***
## SALINITY     0.19240    0.07190   2.676 0.007450 **
## AREAGA      -0.32508    0.17595  -1.848 0.064663 .
## AREALB      -1.44549    0.17485  -8.267  < 2e-16 ***
```

```
## AREAOB      -1.61294    0.17741  -9.091  < 2e-16 ***
## AREARB      -1.60220    0.21683  -7.389 1.48e-13 ***
## AREASC      -0.88576    0.16214  -5.463 4.68e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##            dAIC df
## zero_adj    0.0 79
## neg_binom  26.2 40
## zero_infl  34.6 49
```

```
##            dAIC df
## zero_adj    0.0 71
## zero_infl  19.9 71
## neg_binom  40.2 36
```

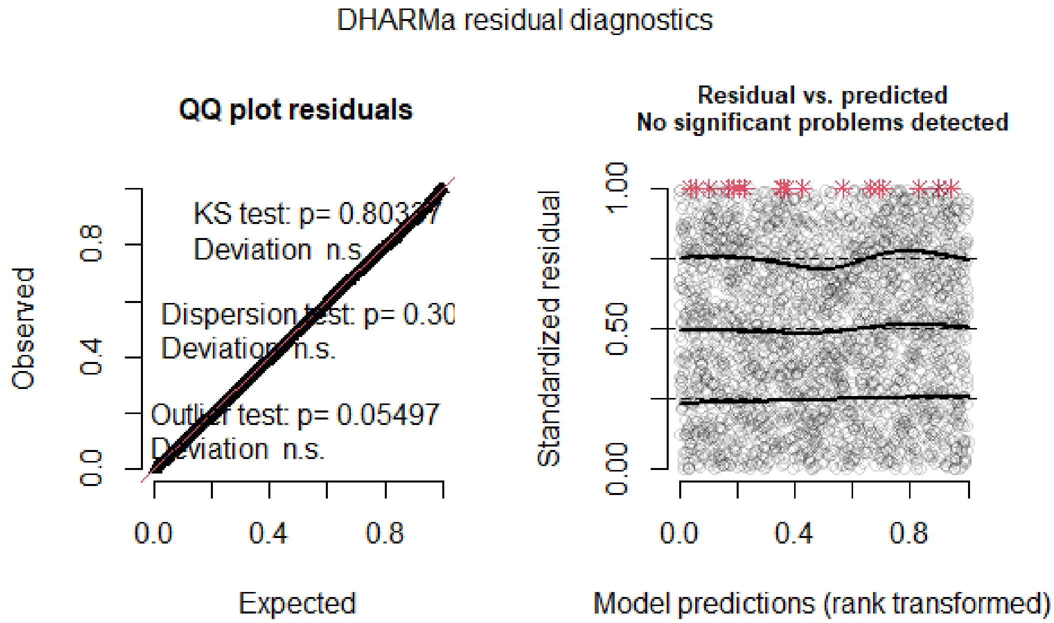
```
##            dAIC  df
## zero_adj    0.0 115
## zero_infl  40.7 85
## neg_binom  49.8 58
```

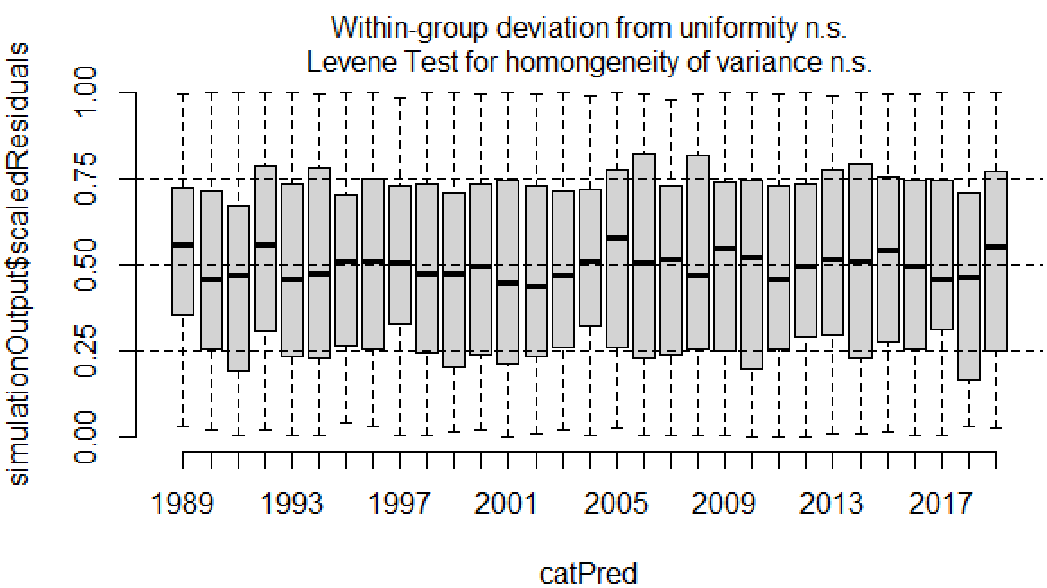
```
##            dAIC df
## zero_adj    0.0 79
## neg_binom  26.2 40
## zero_infl  34.6 49
```

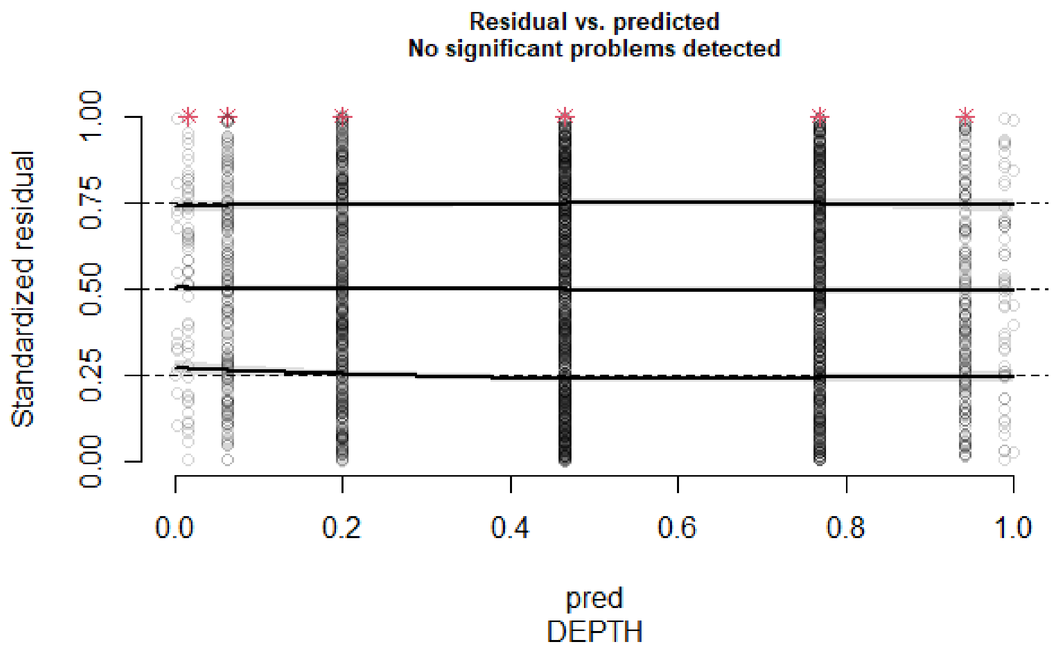
```
##                             dAIC  df
## full_mod_stratum$zero_adj    0.0 115
## full_mod$zero_adj           94.6 71
## full_mod_area$zero_adj     121.9 79
```

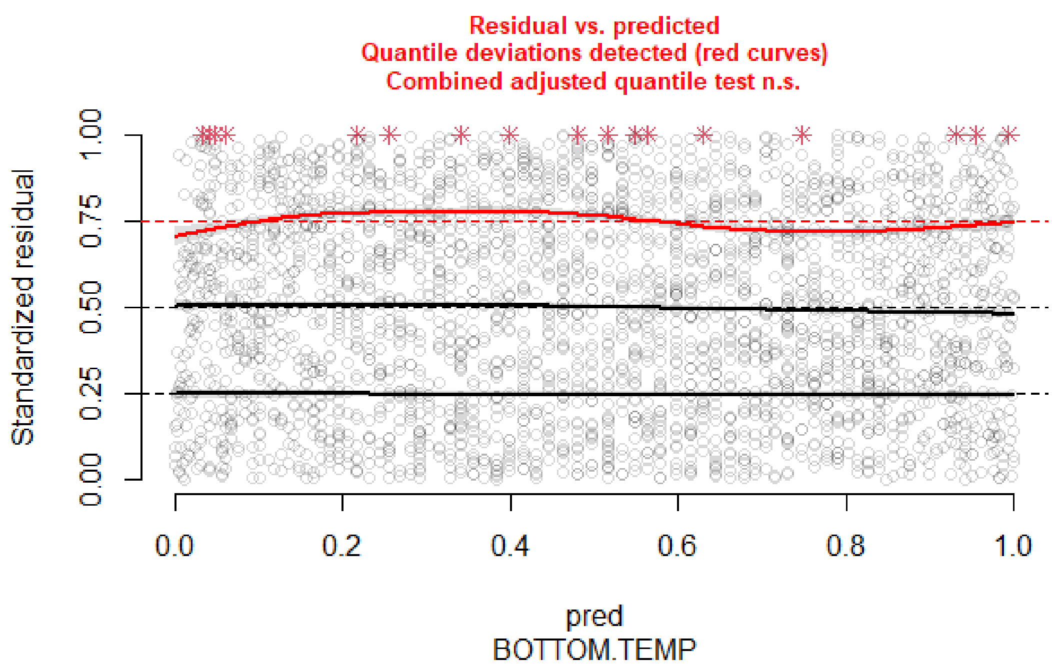
DHARMA simulations

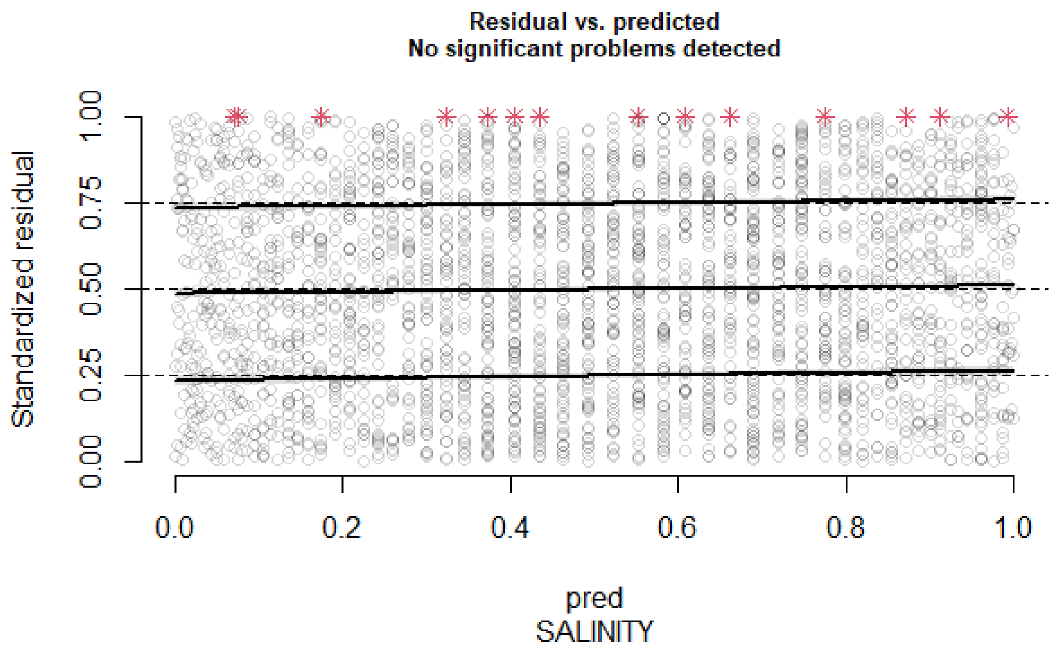
Full model with latitude

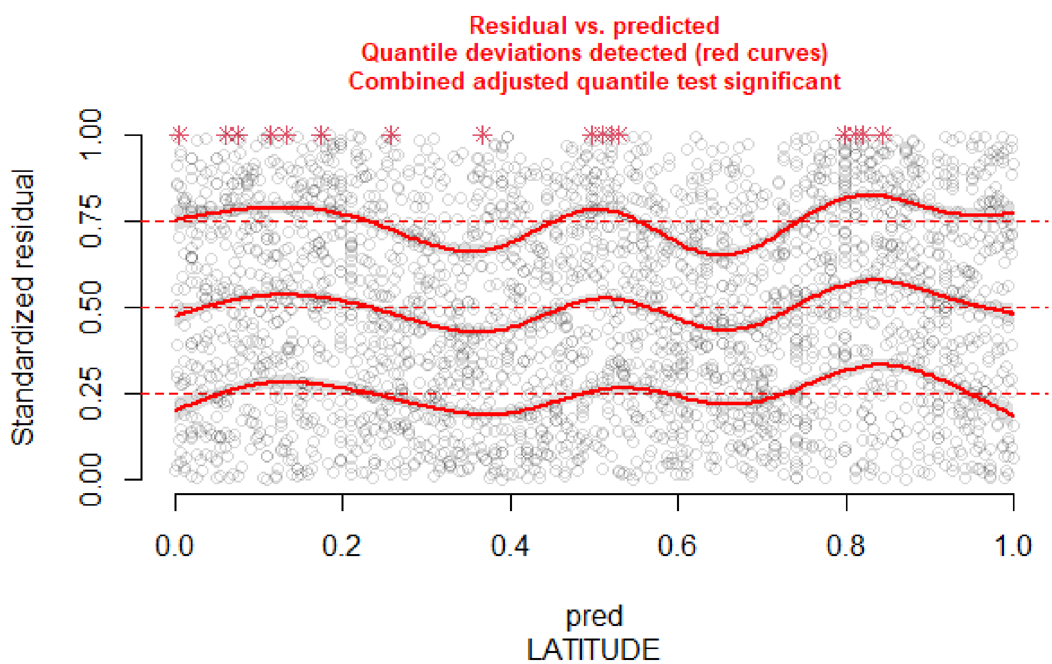




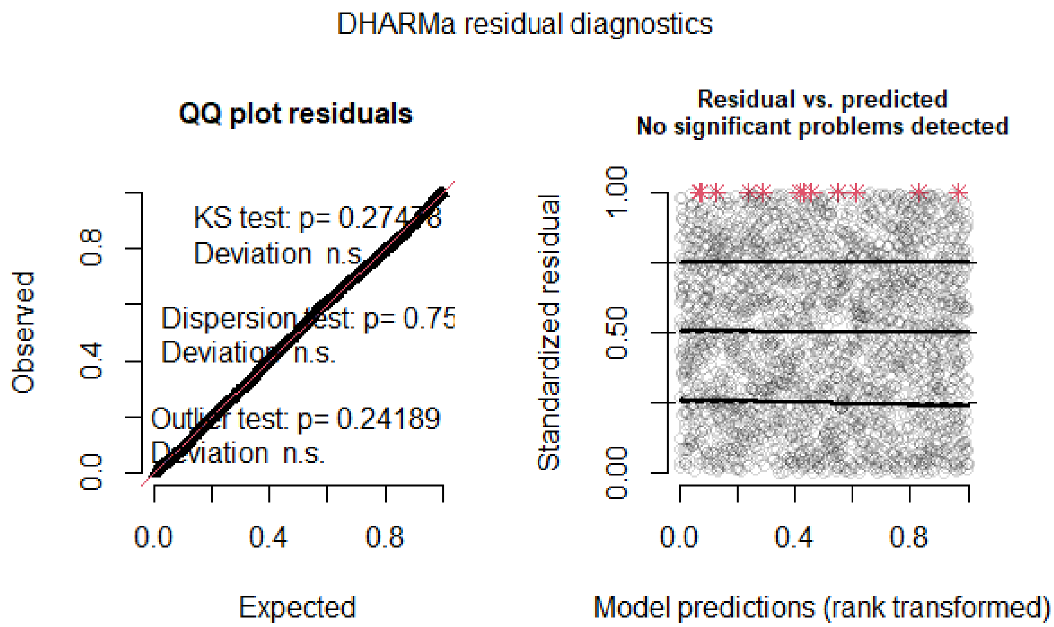

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

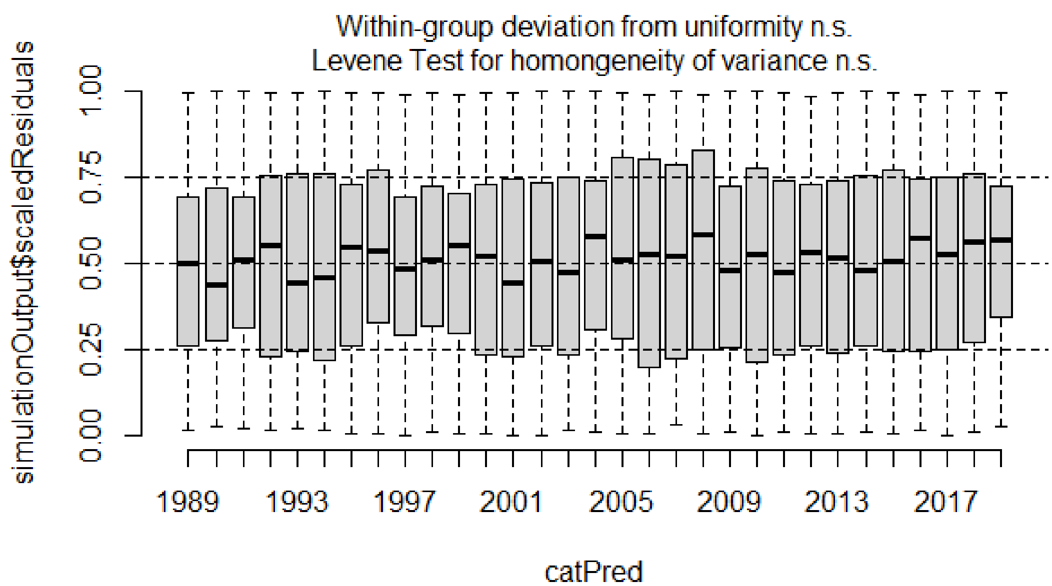

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test n.s.
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

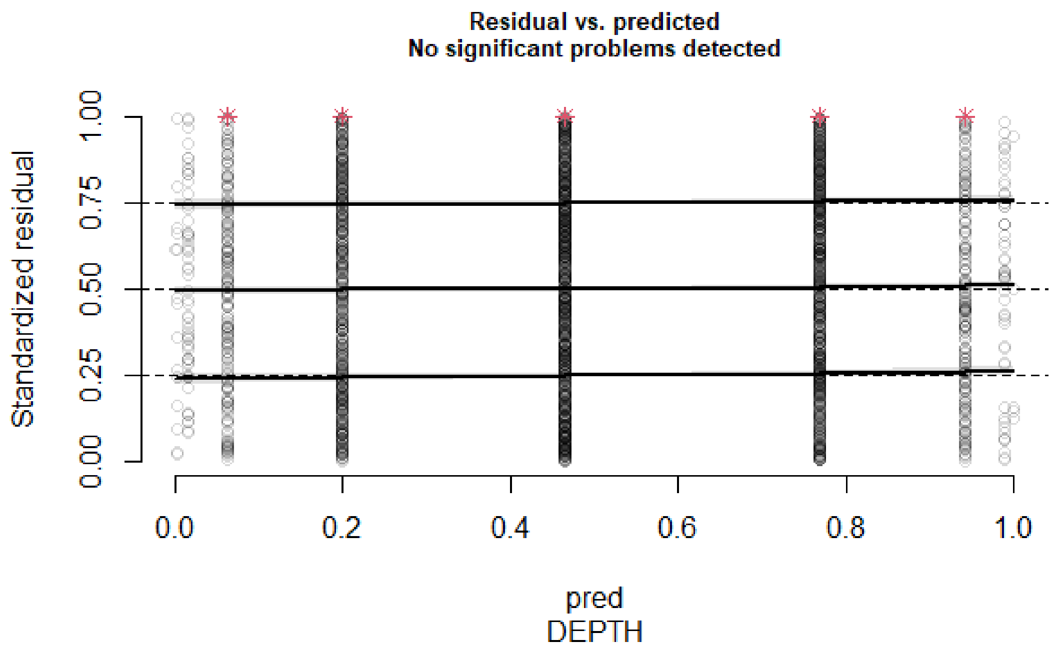

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY

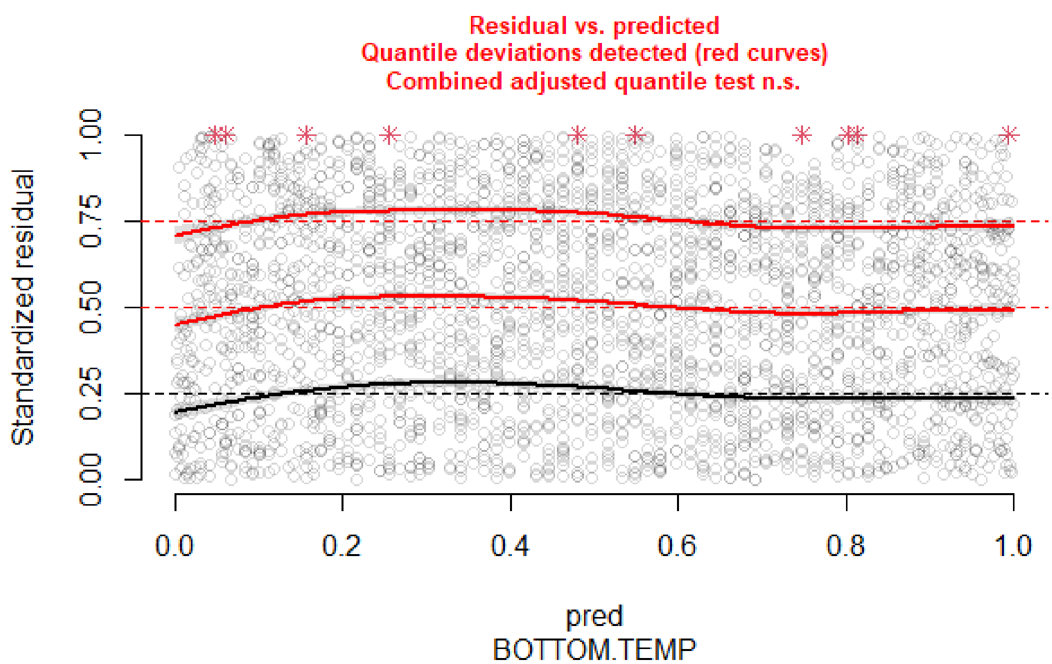

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
LATITUDE

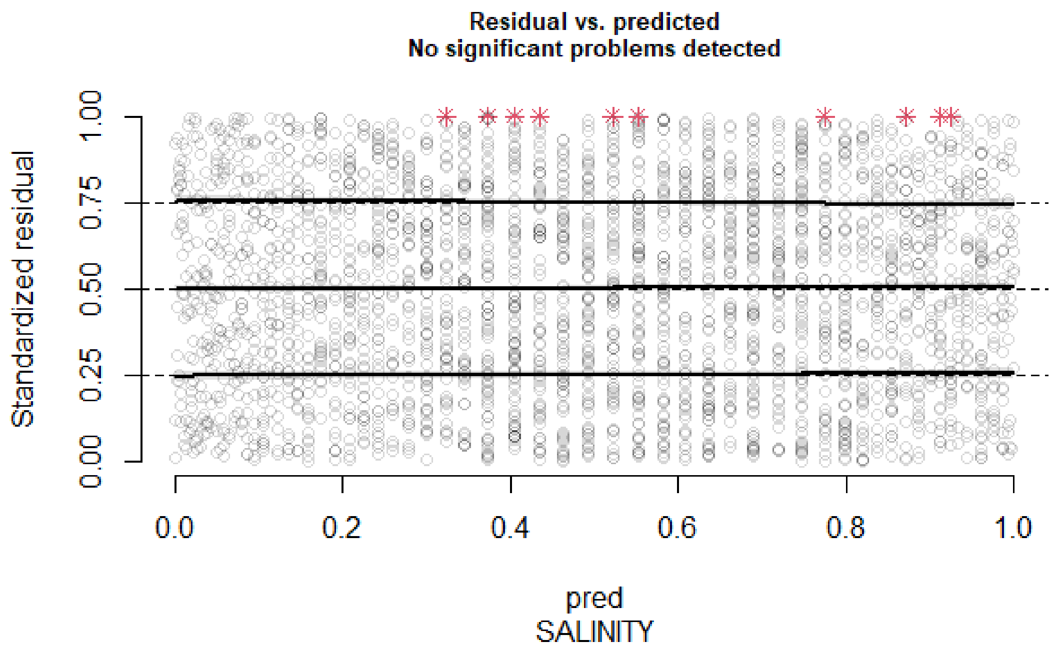
Full model with STRATUM instead of LATITUDE

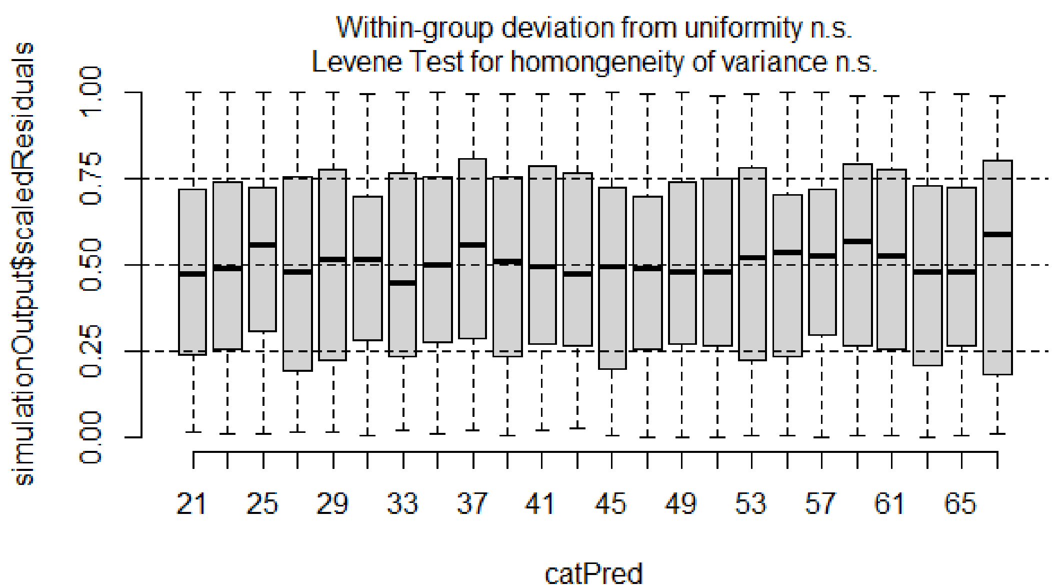




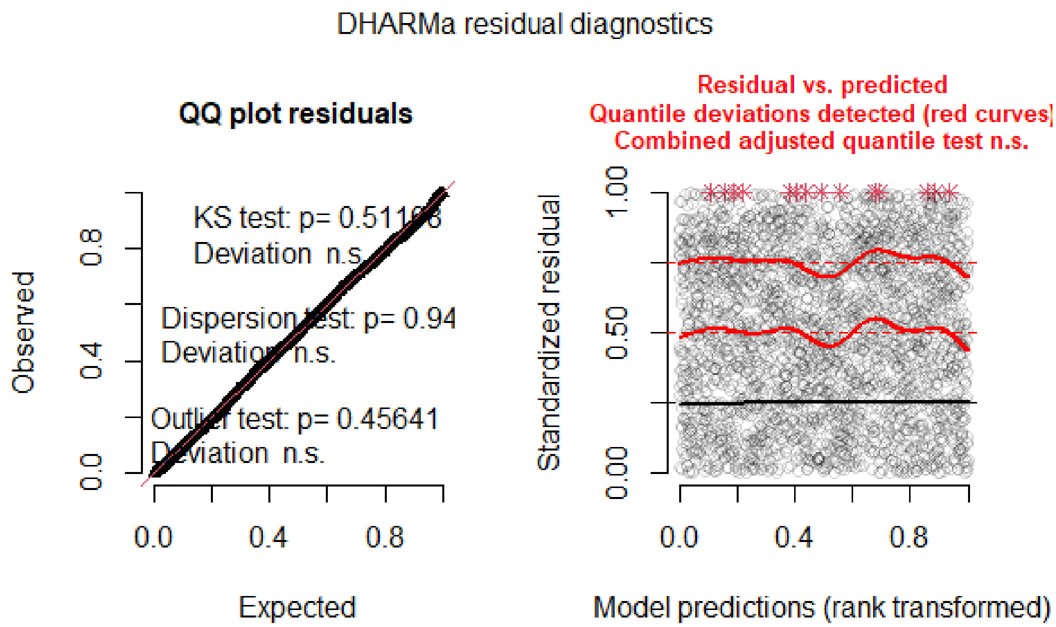

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

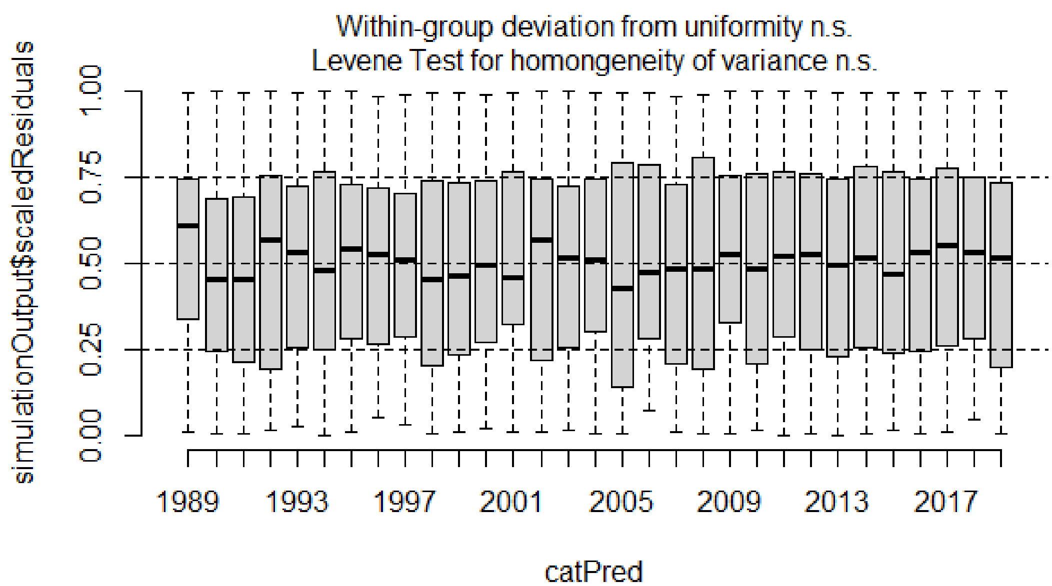

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test n.s.
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

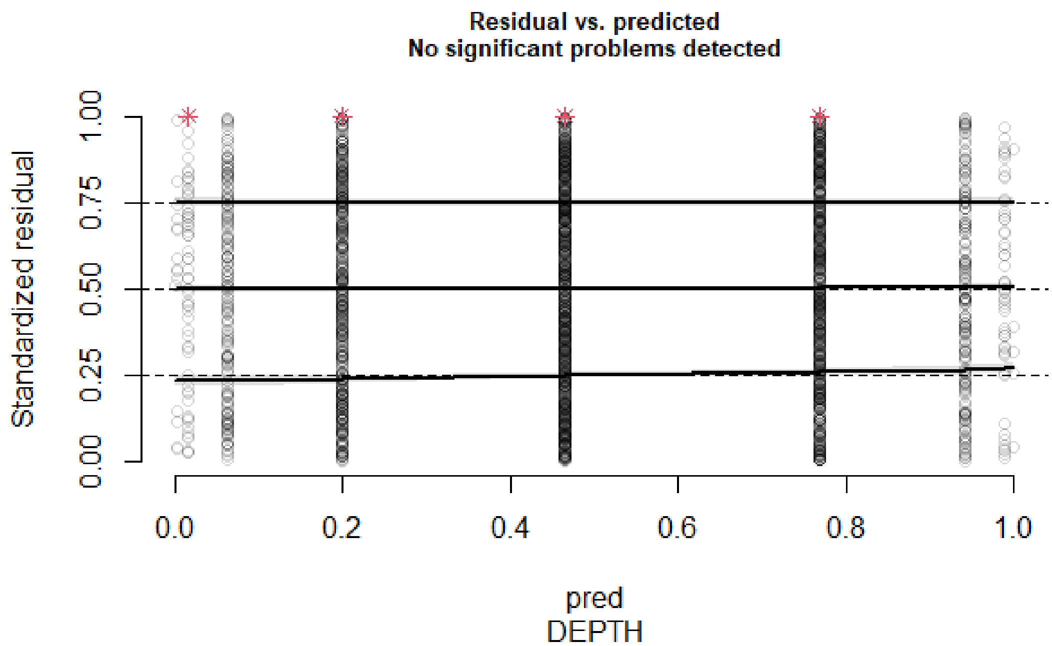

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY


Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
21
25
29
33
37
41
45
49
53
57
61
65
catPred

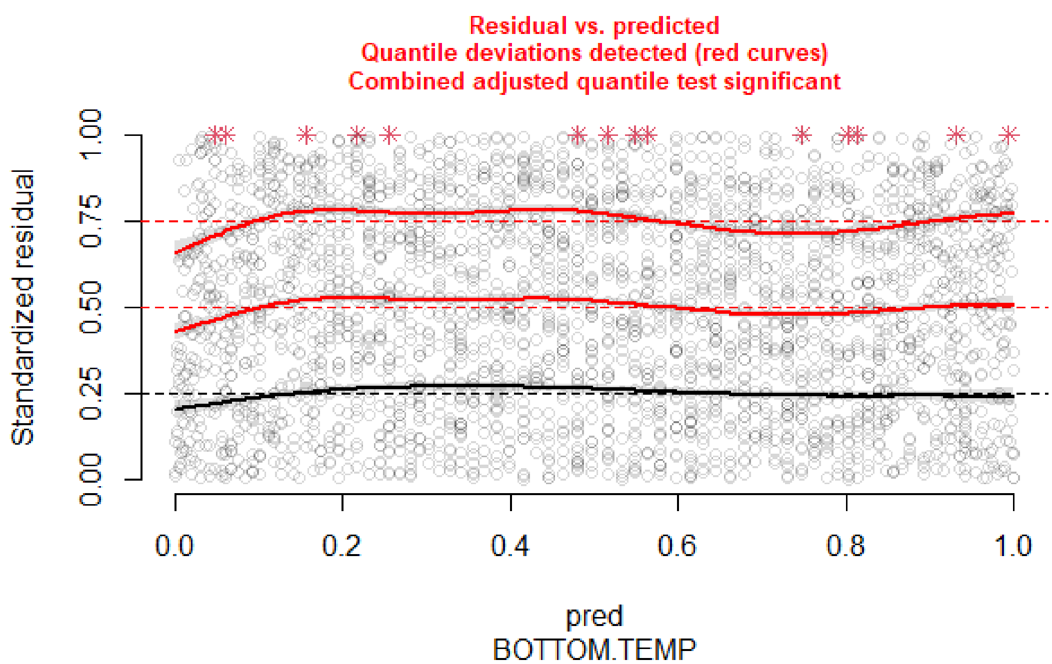
Full model with AREA instead of LATITUDE



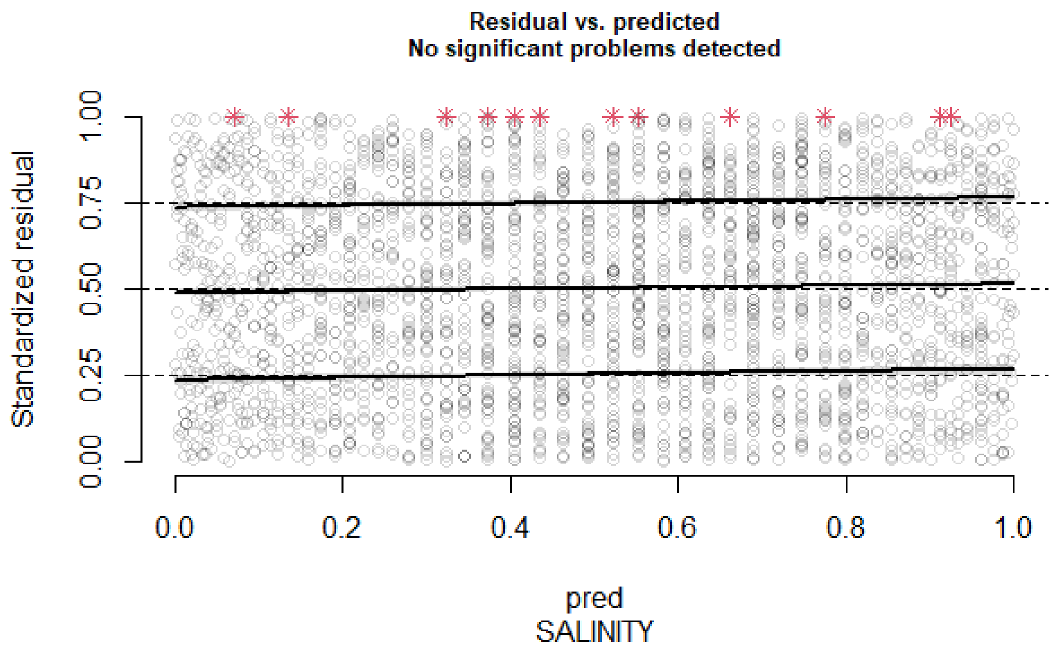


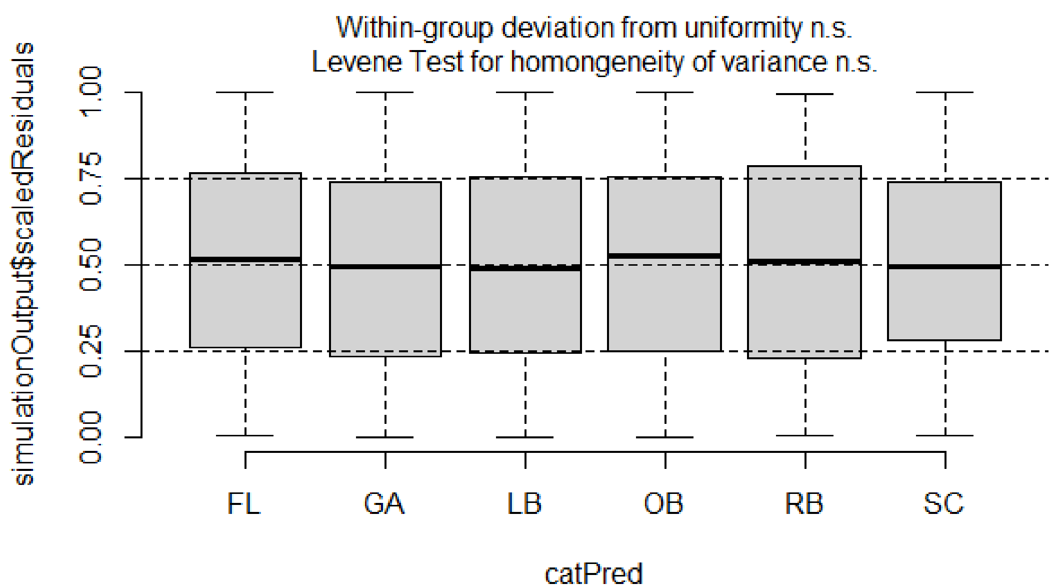


Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

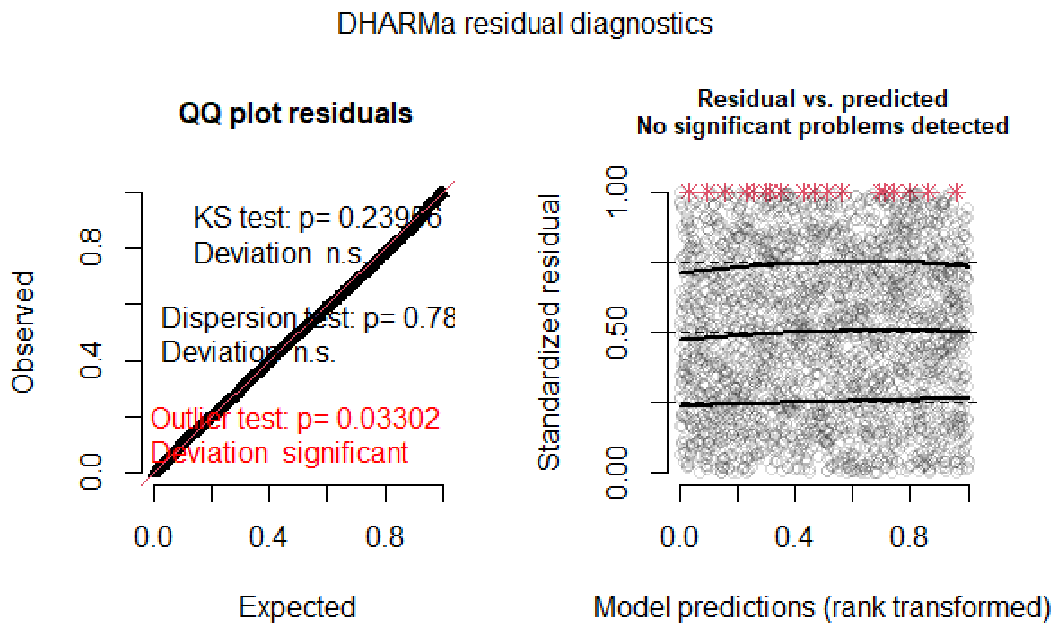


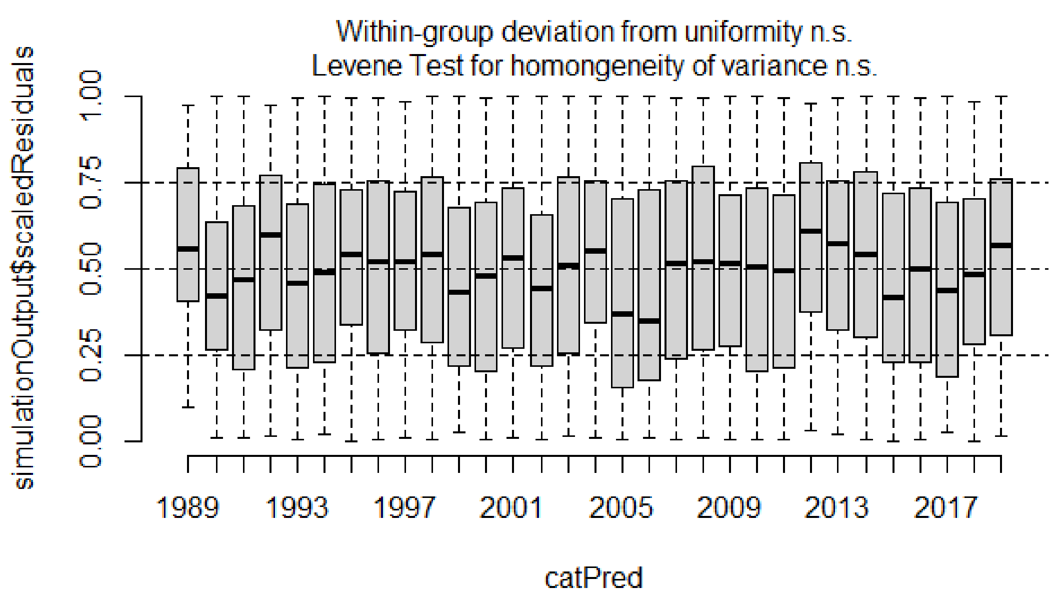
Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

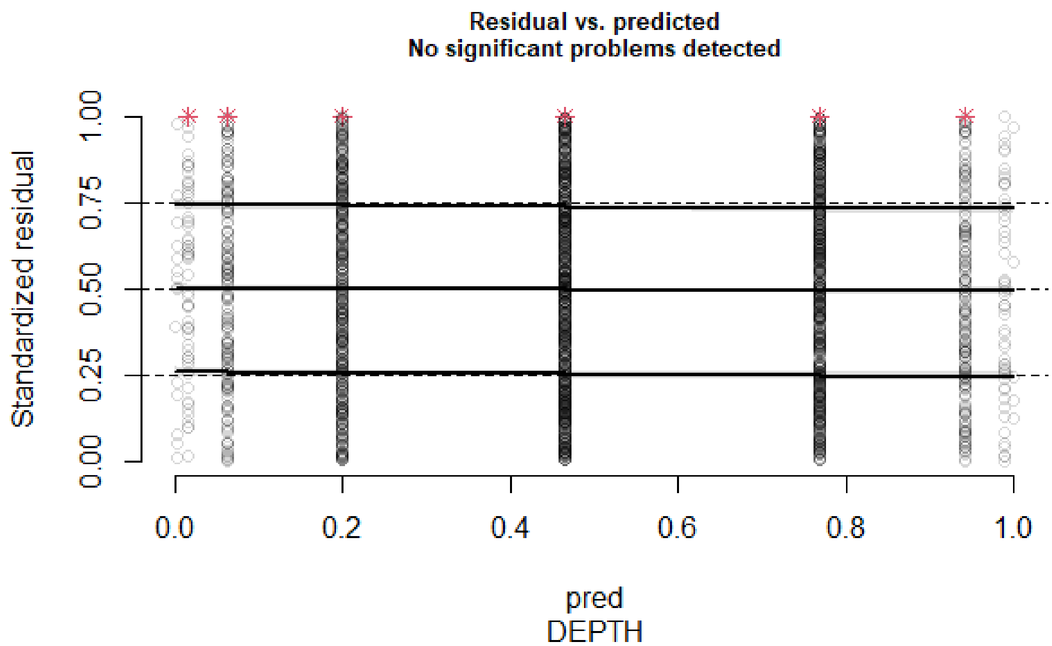

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY

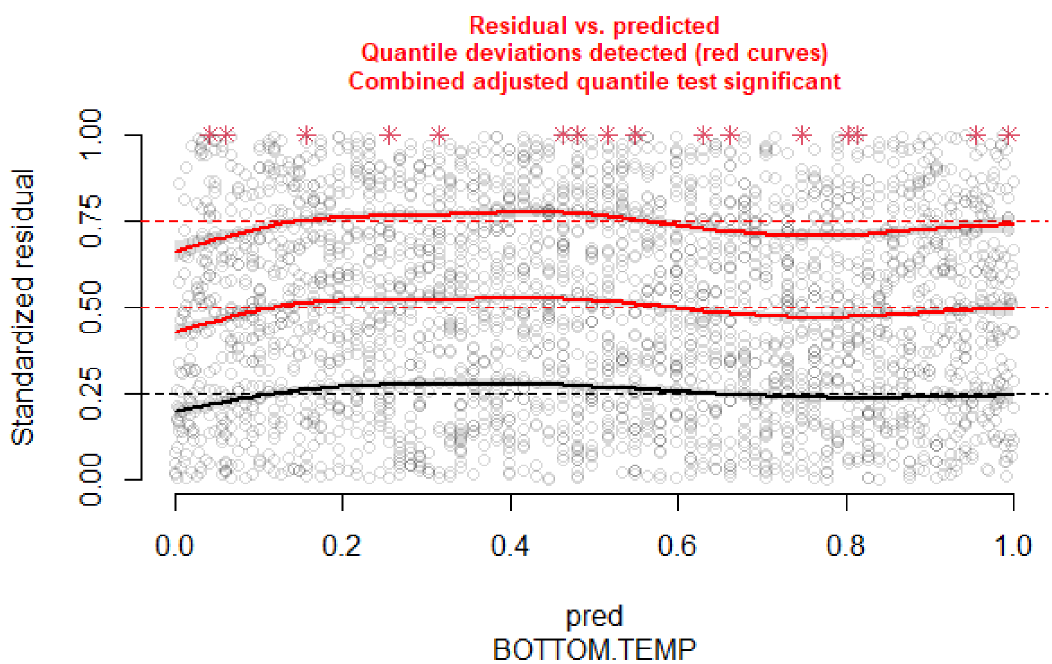

Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
FL
GA
LB
OB
RB
SC
catPred

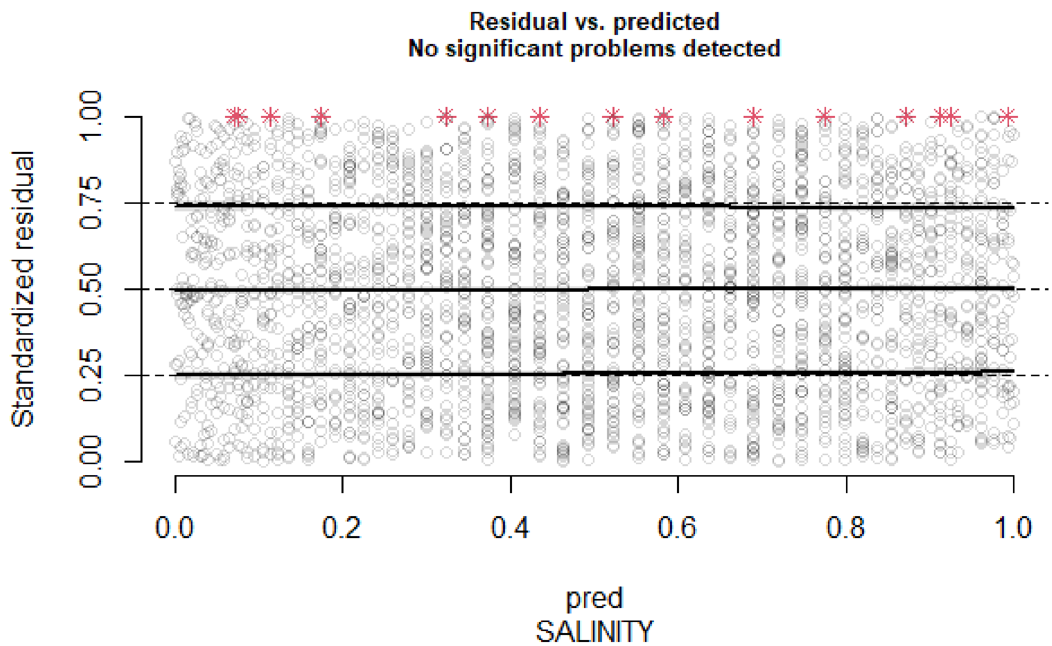
All distributions, full model with STRATUM



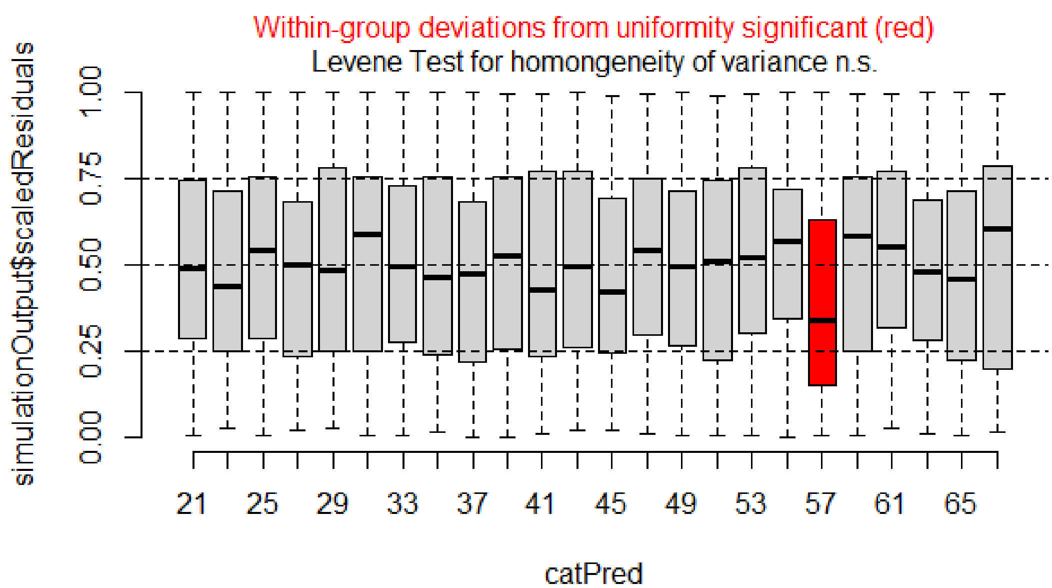



Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

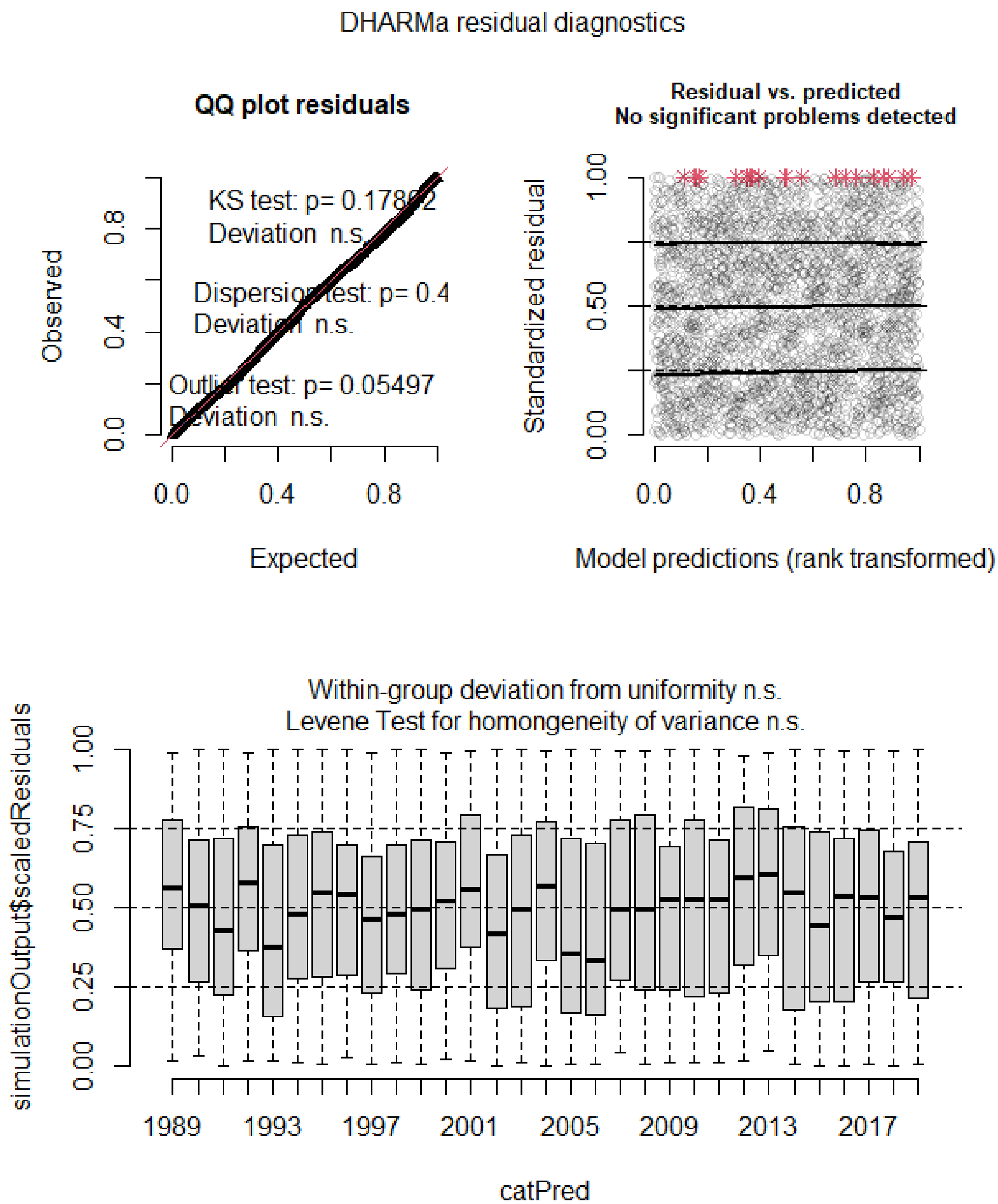

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

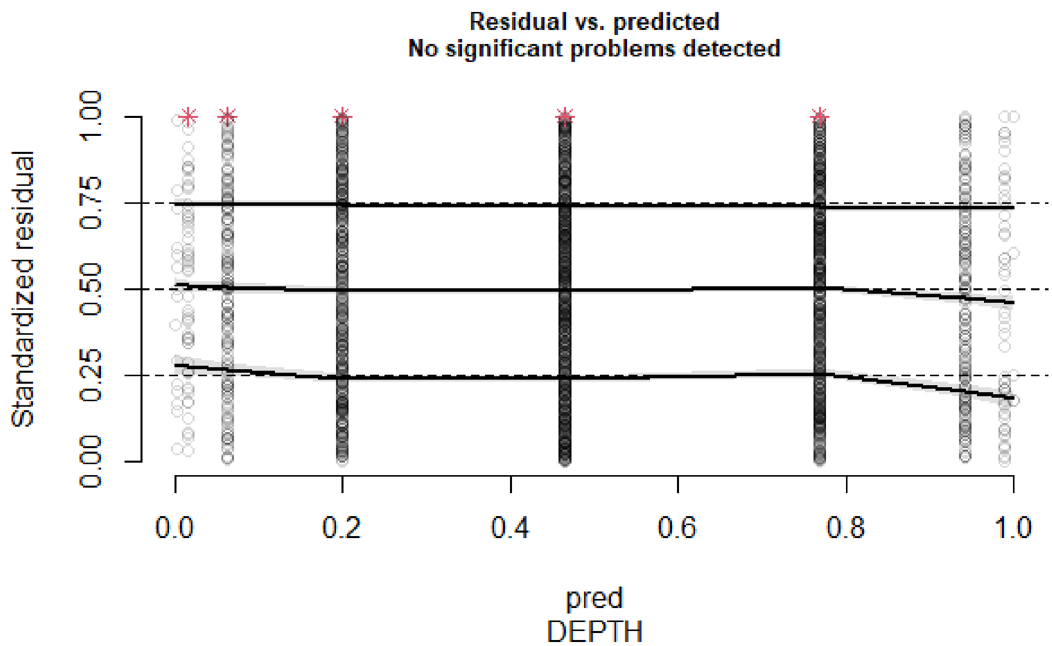


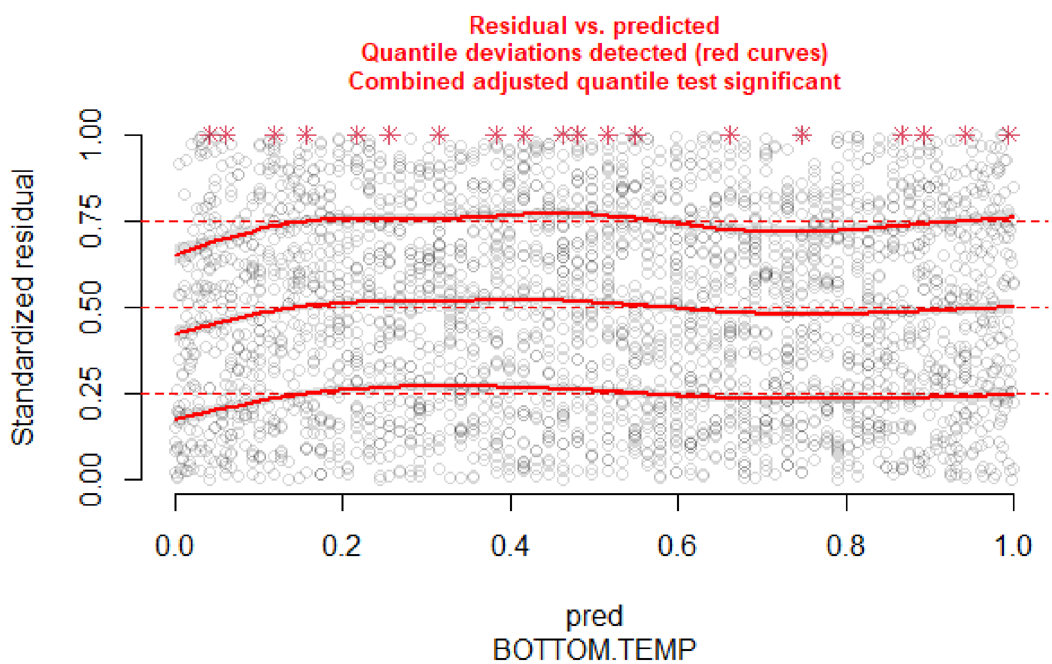
Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY

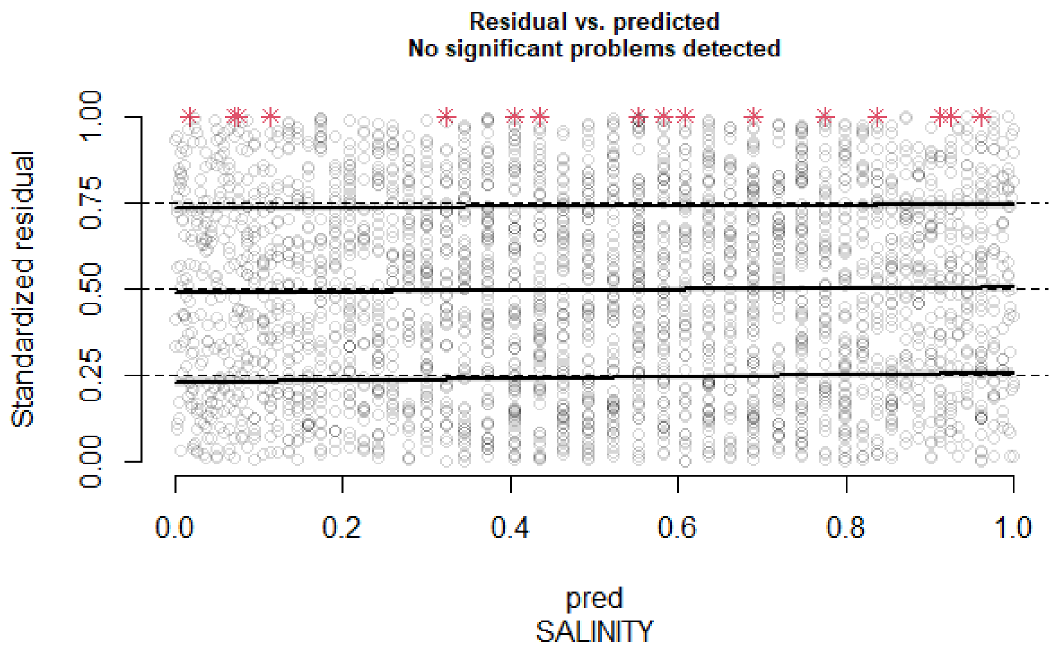


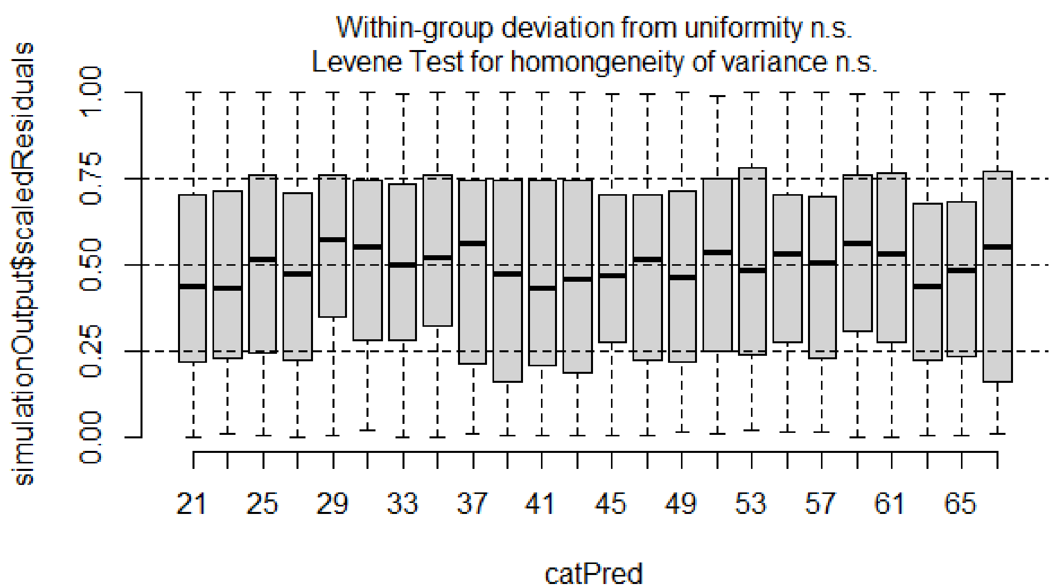
Within-group deviations from uniformity significant (red)
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
21
25
29
33
37
41
45
49
53
57
61
65
catPred

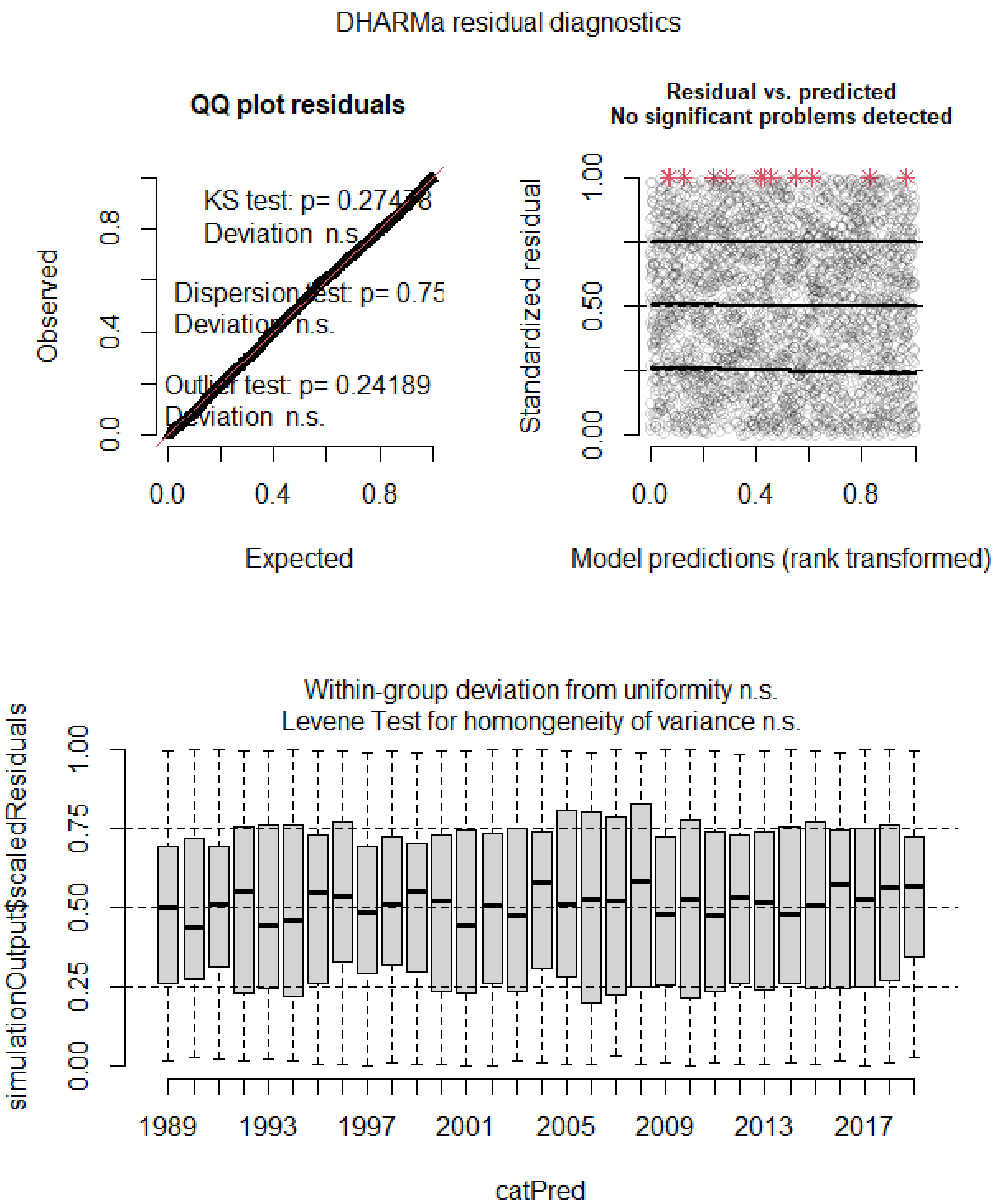

DHARMa residual diagnostics
QQ plot residuals
KS test: p= 0.17862
Deviation n.s.
Dispersion test: p= 0.4
Deviation n.s.
Outlier test: p= 0.05497
Deviation n.s.
Observed
Expected
Residual vs. predicted
No significant problems detected
Standardized residual
Model predictions (rank transformed)
Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
catPred
1989
1993
1997
2001
2005
2009
2013
2017

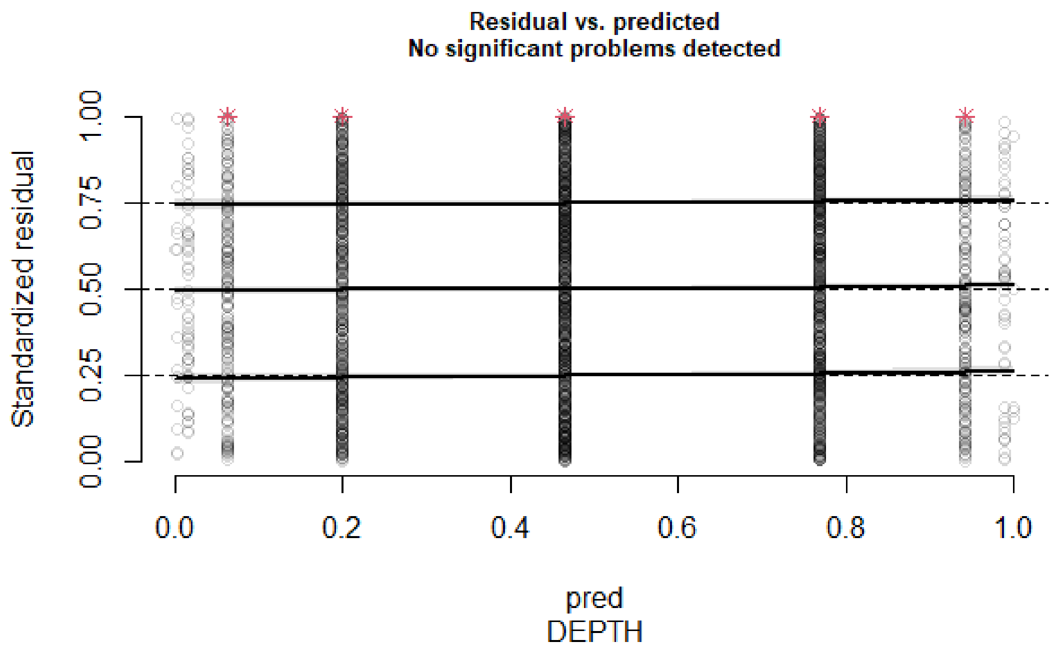

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH

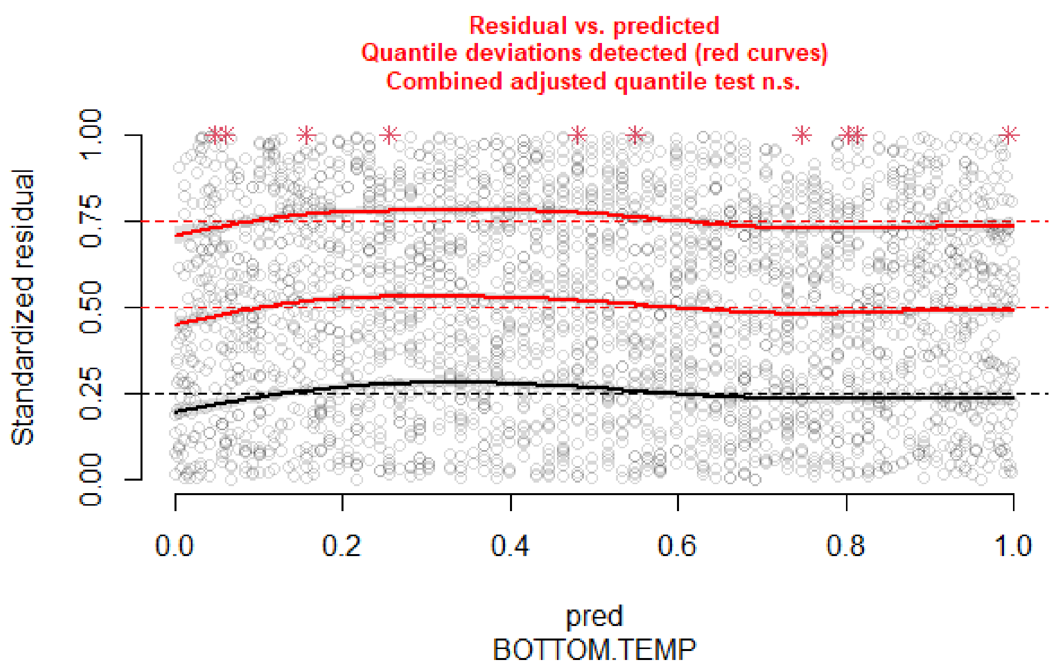

Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test significant
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP

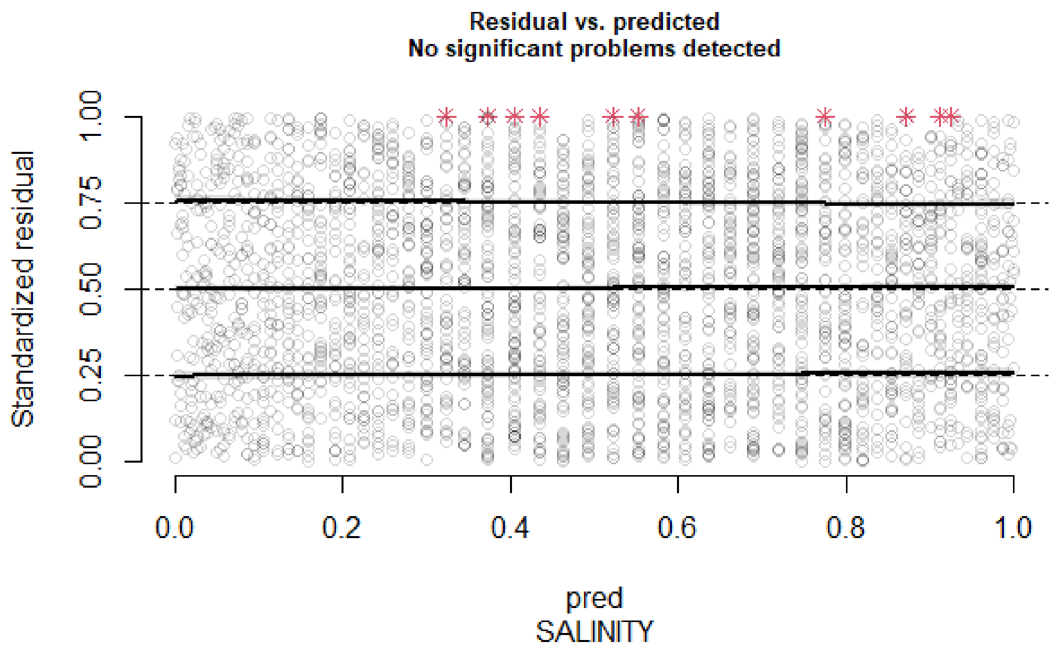

Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
SALINITY

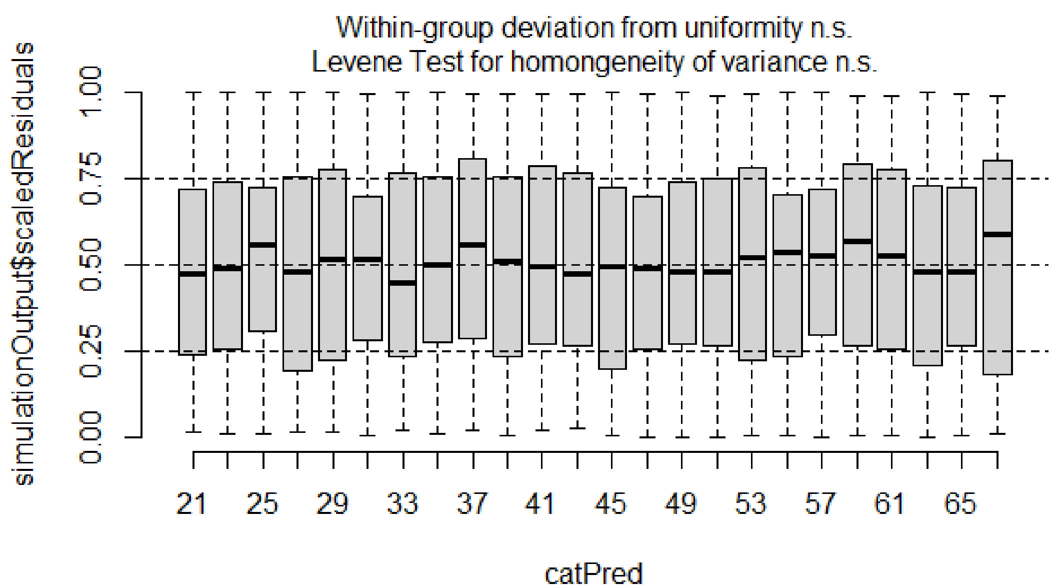

Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
21
25
29
33
37
41
45
49
53
57
61
65
catPred


DHARMa residual diagnostics
QQ plot residuals
KS test: p= 0.27478
Deviation n.s.
Dispersion test: p= 0.75
Deviation n.s.
Outlier test: p= 0.24189
Deviation n.s.
Observed
Expected
Residual vs. predicted
No significant problems detected
Standardized residual
Model predictions (rank transformed)
Within-group deviation from uniformity n.s.
Levene Test for homongeneity of variance n.s.
simulationOutput$scaledResiduals
catPred
1989
1993
1997
2001
2005
2009
2013
2017


Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH


Residual vs. predicted
Quantile deviations detected (red curves)
Combined adjusted quantile test n.s.
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP





###

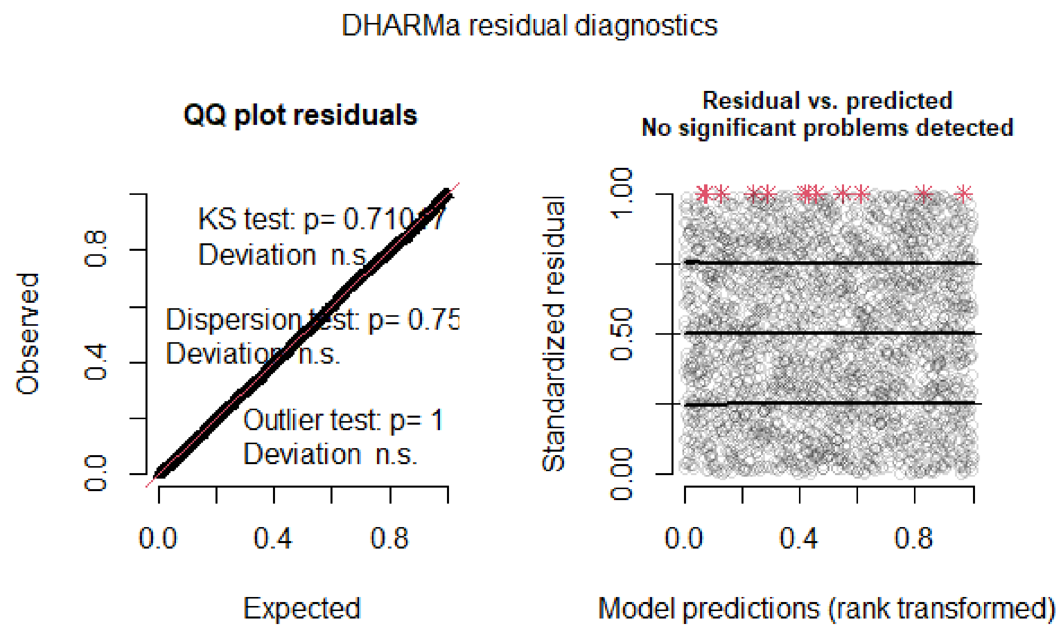
Step 3: Winning model Zero-altered negative binomial with STRATUM had the best AIC and diagnostics.

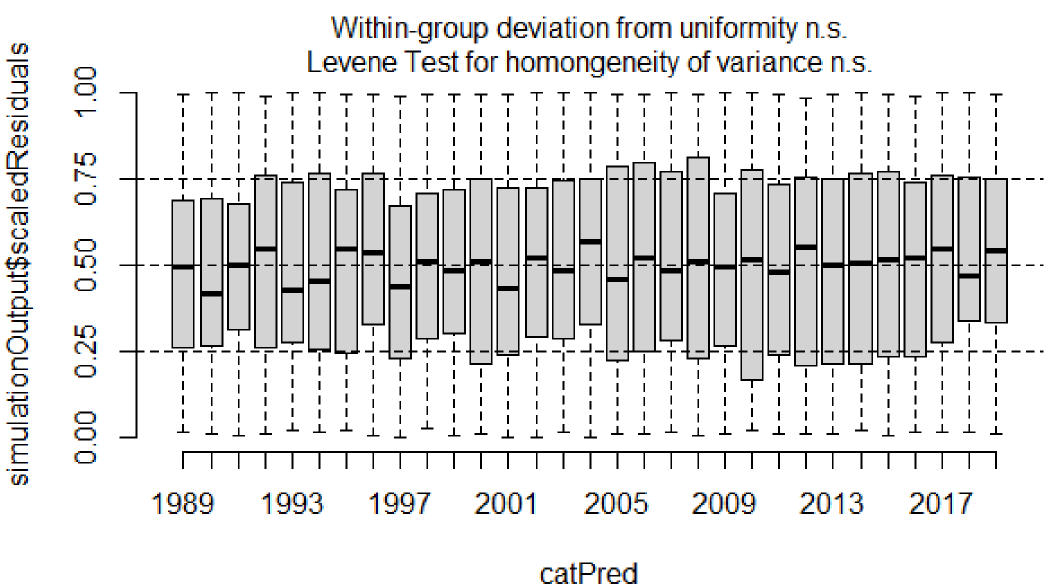
Test covariate importance

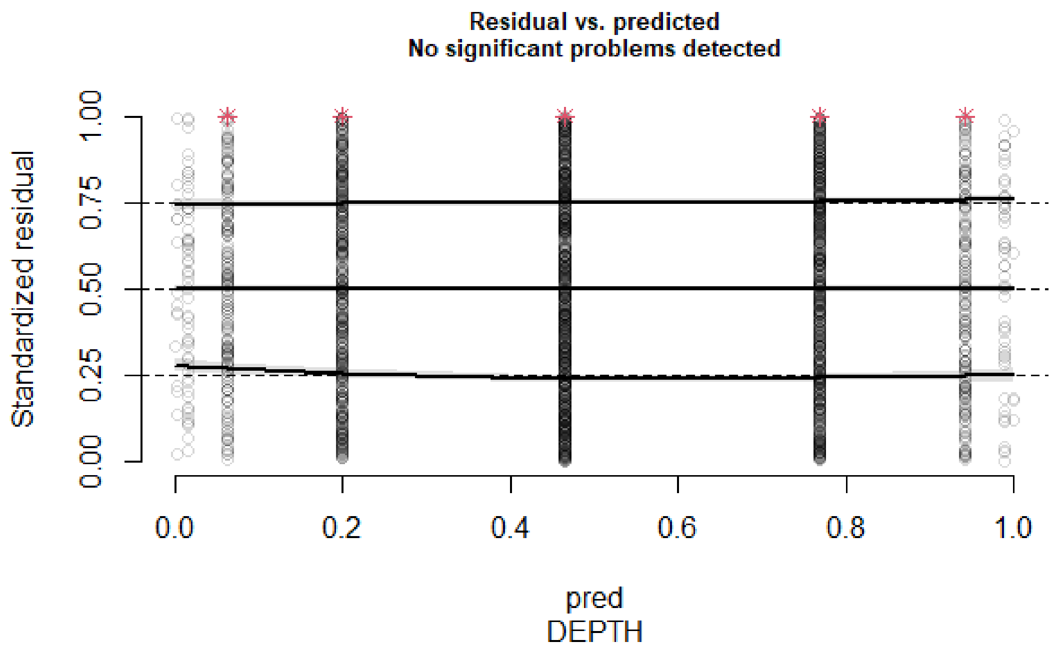
```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ 1
```

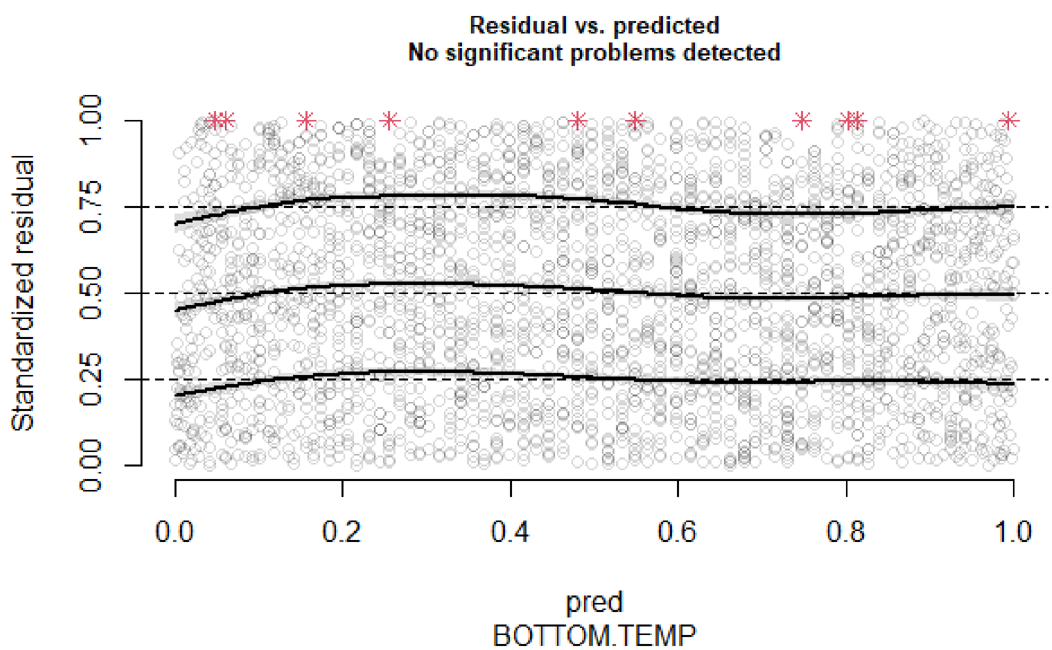
```
##
## Final Model:
## FREQ ~ STRATUM + YEAR + BOTTOM.TEMP + DEPTH + SALINITY
##
##
##            Step Df   Deviance Resid. Df Resid. Dev      AIC
## 1                                 2757   12191.51 12309.51
## 2     + STRATUM 23 385.096722      2734   11806.41 11970.41
## 3        + YEAR 30 140.179444      2704   11666.23 11890.23
## 4 + BOTTOM.TEMP  1  12.331711      2703   11653.90 11879.90
## 5       + DEPTH  1   5.358309      2702   11648.54 11876.54
## 6    + SALINITY  1   2.468314      2701   11646.08 11876.08
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
##
## Final Model:
## FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
##
##
##   Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                       2701   11646.08 11876.08
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
##
## Final Model:
## FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM
##
##
##   Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                       2701   11646.08 11876.08

##        dAIC  df
## ZANB     0.0 115
## ZANB.2  20.5 114
## ZANB.3  36.4 113
## ZANB.4  43.2 112
## ZANB.5 405.5 89
## ZANB.6 477.2 59
```

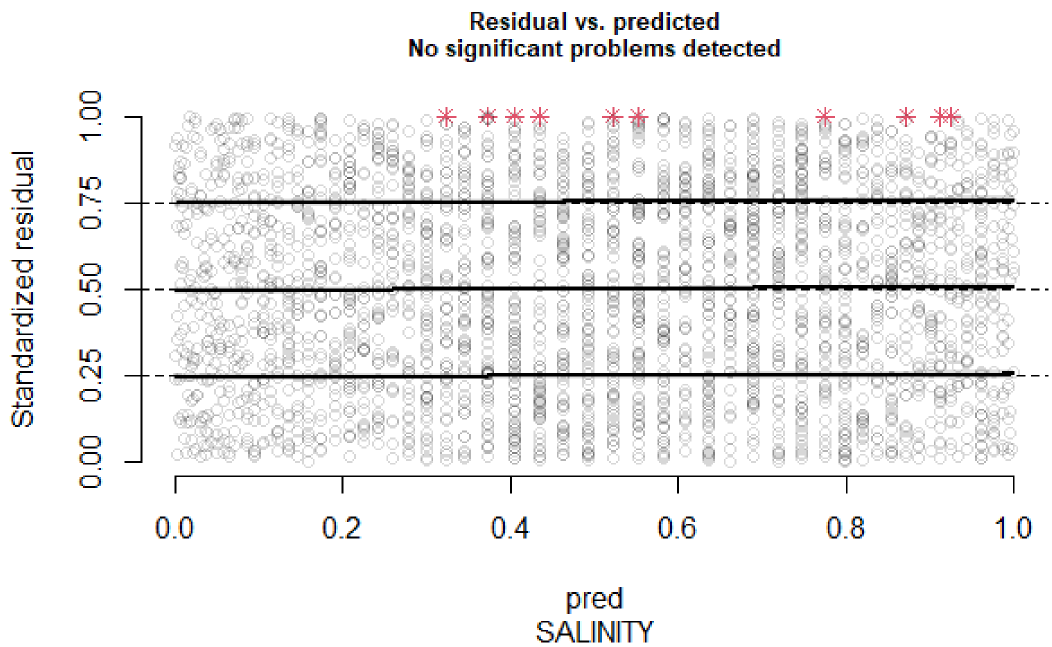
DHARMA simulations

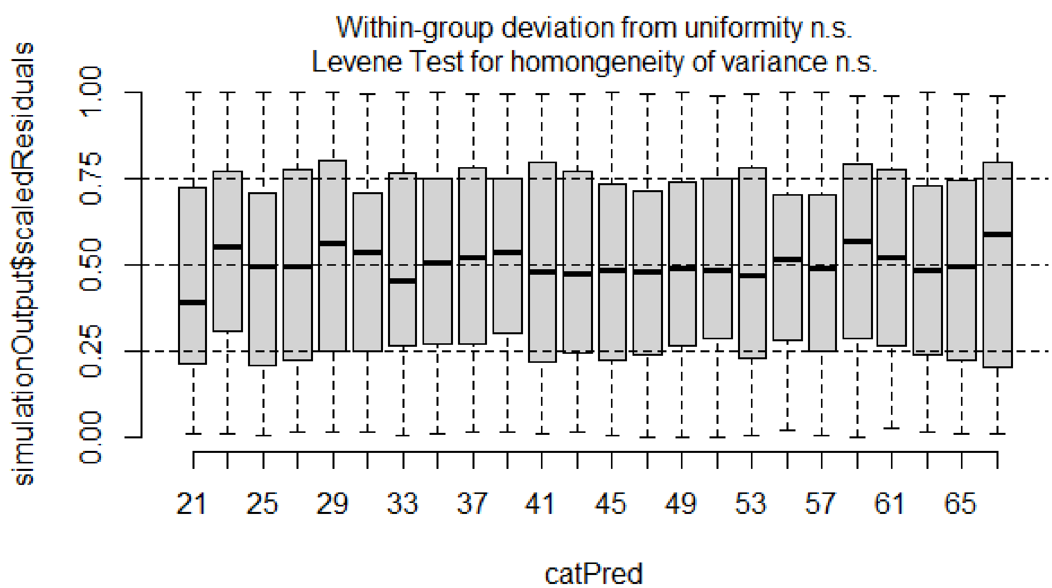





Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
DEPTH


Residual vs. predicted
No significant problems detected
Standardized residual
0.00
0.25
0.50
0.75
1.00
0.0
0.2
0.4
0.6
0.8
1.0
pred
BOTTOM.TEMP



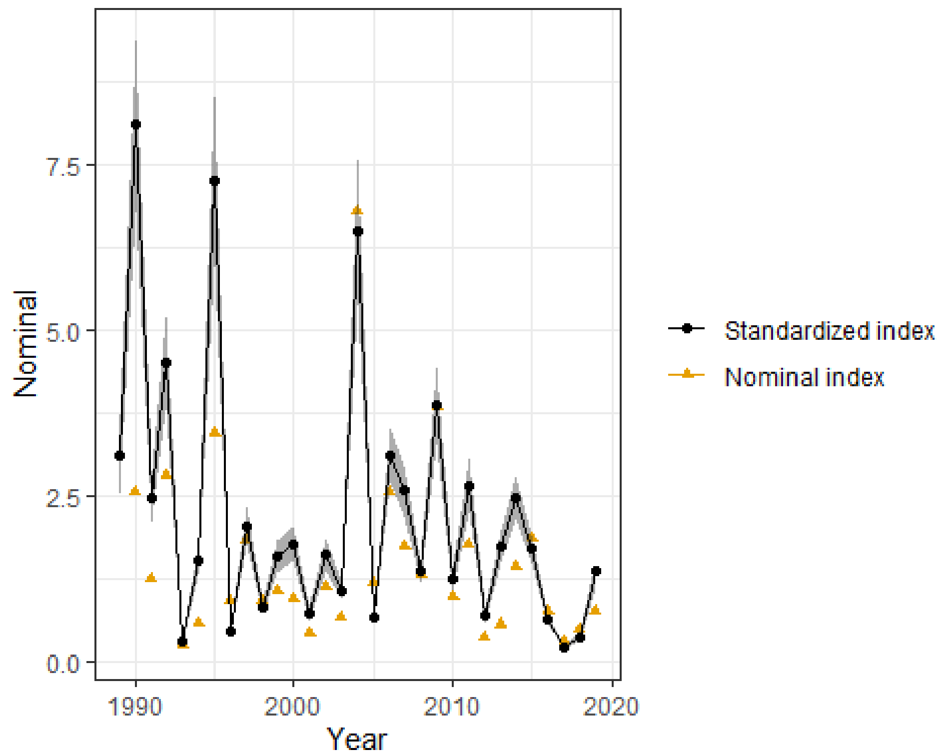


Step 4: Plot predictions

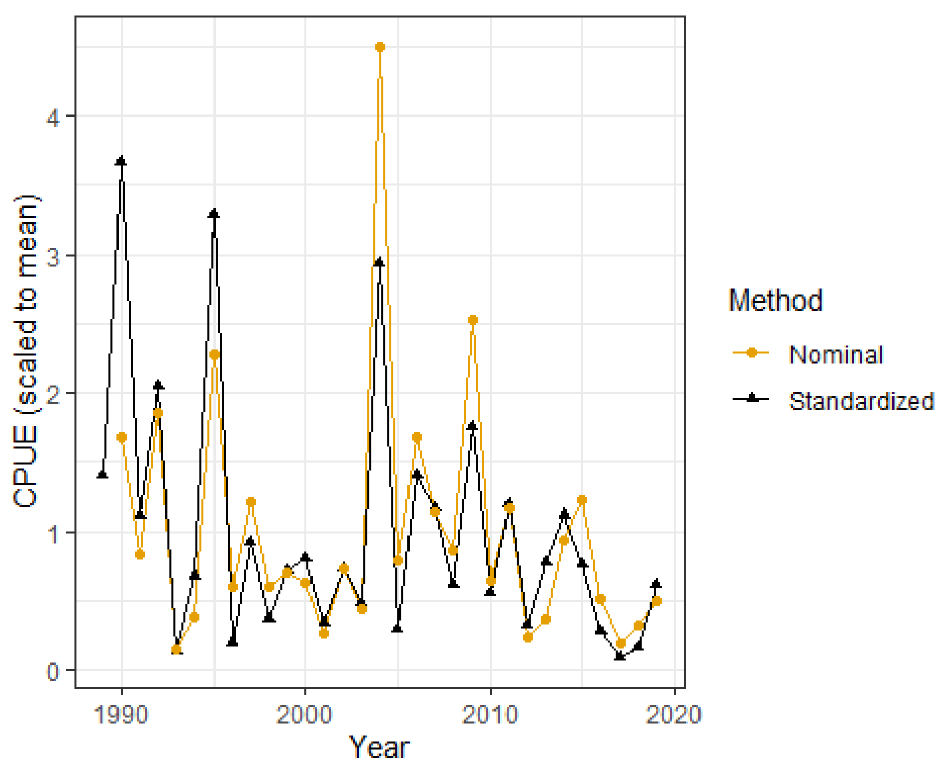
Final model:

FREQ ~ YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM ziformula = ~YEAR + DEPTH + BOTTOM.TEMP + SALINITY + STRATUM

Plot standardized index from this work with index provided by the state (also standardized, but with their preferred method).



Comparison of standardized index and nominal index (i.e., index provided by the state) scaled to their time-series means.



NH Seine

```
st <- 'NH'
source('singleIndices/get_packages.R')
source('functions/get_combos.R')
source('functions/cleanData.R')
source('functions/samplesByLevel.R')
source('functions/samplesByNumeric.R')
source('functions/Zscr.R')
source('functions/get_aicTab.R')
source('functions/get_combos.R')
source('functions/get_idxPlot.R')
source('functions/bootstrap_functions.R')
ctNH <- c(SampleID = 'text',
        State = 'text',
        Program = 'text',
        Gear = 'text',
        Year = 'text',
        Month = 'numeric',
        Day = 'numeric',
        Time = 'text',
        Effort = 'text',
        N = 'numeric',
        Weight = 'text',
        Area = 'text',
        Stratum = 'text',
        Depth = 'numeric',
        Wgt.Factor = 'numeric',
        Surface.Temp = 'numeric',
        Bottom.Temp = 'numeric',
        Salinith = 'numeric',
        DO = 'numeric',
        Bottom.Type = 'text')

nhSeine <- read_excel(path = 'data/2021_bluefish_data_NH.xlsx',
                    sheet = 'SurveyData',
                    col_types = ctNH,
                    na = c('NA', 'less than 6 ft')) %>%
  mutate(Depth = as.numeric(NA),
         Gear = 'Seine',
         Year = ifelse(Year == 992001, 2001, Year),
         Effort = 1, # effort is "Net Haul"
         Season = 'Summer-Fall') %>%
         cleanData(.) %>%

  # Include only the necessary data
  select(Year, Month, Surface.Temp, Bottom.Type, Salinity, N)
```

No samples during Nov / Dec.

```
nhSeine %>%
  group_by(Month) %>%
  summarize(N = sum(N))
## # A tibble: 7 x 2
```

```
##   Month     N
##   <fct> <dbl>
## 1 Jun      26
## 2 Jul     211
## 3 Aug     301
## 4 Sep     276
## 5 Oct      58
## 6 Nov       0
## 7 Dec       0
```

So remove these months.

```
nhSeine <- nhSeine %>%
  filter(!Month %in% c('Nov', 'Dec'))
```

Save these records for the nominal mean calculation

```
nhSeine_nomDat <- nhSeine
```

There are only a few records for bottom types sand/shell and gravel.

```
nhSeine %>%
  group_by(Bottom.Type) %>%
  count() %>%
  kbl() %>%
  kable_classic(full_width = FALSE)
```

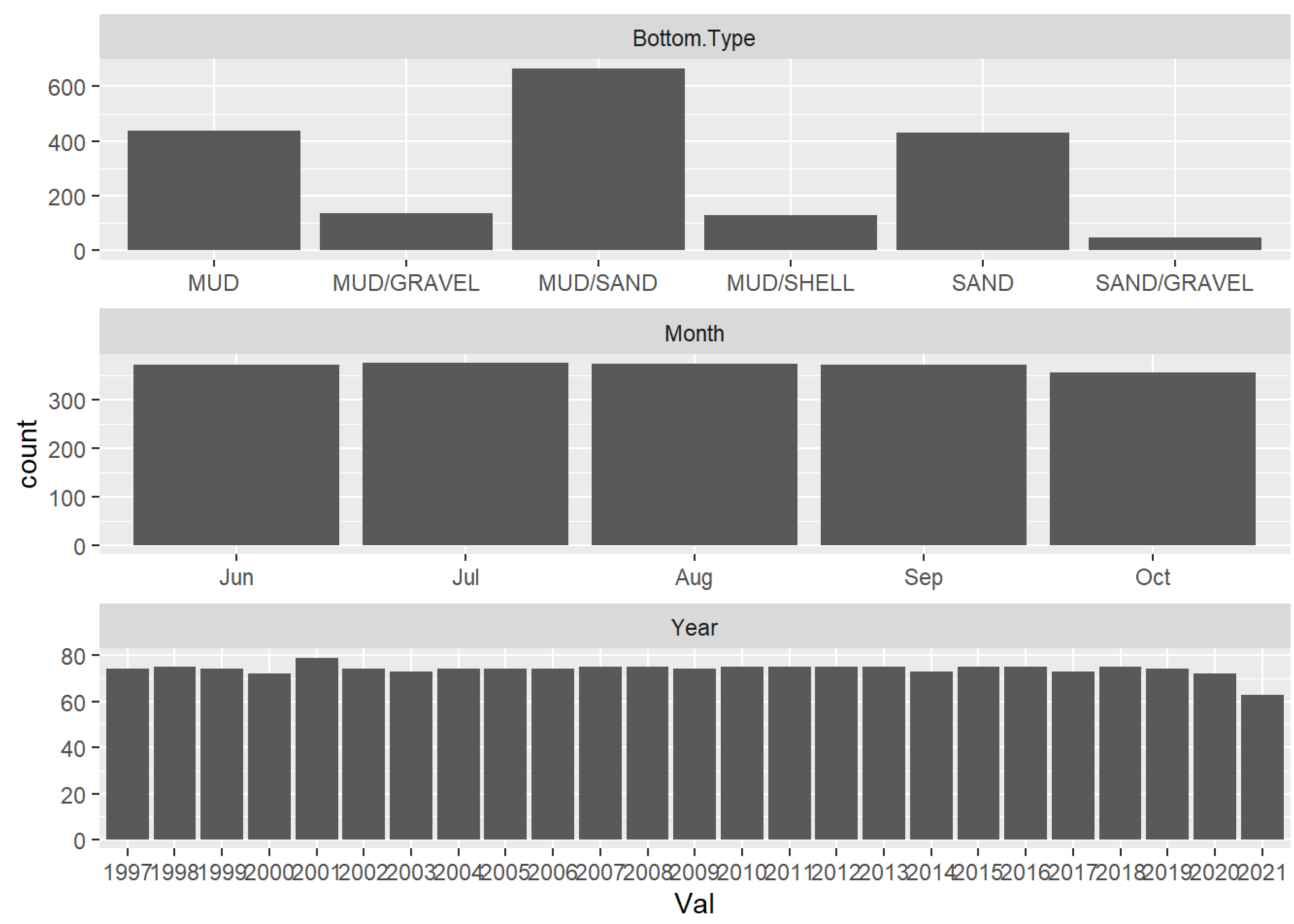
Bottom.Type	n
GRAVEL	17
MUD	439
MUD/GRAVEL	136
MUD/SAND	666
MUD/SHELL	128
SAND	430
SAND/GRAVEL	48
SAND/SHELL	3

Remove those records and make sure bottom type is a factor.

```
nhSeine <- nhSeine %>%
  filter(!Bottom.Type %in% c('GRAVEL', 'SAND/SHELL')) %>%
      mutate(Bottom.Type = factor(Bottom.Type))
```

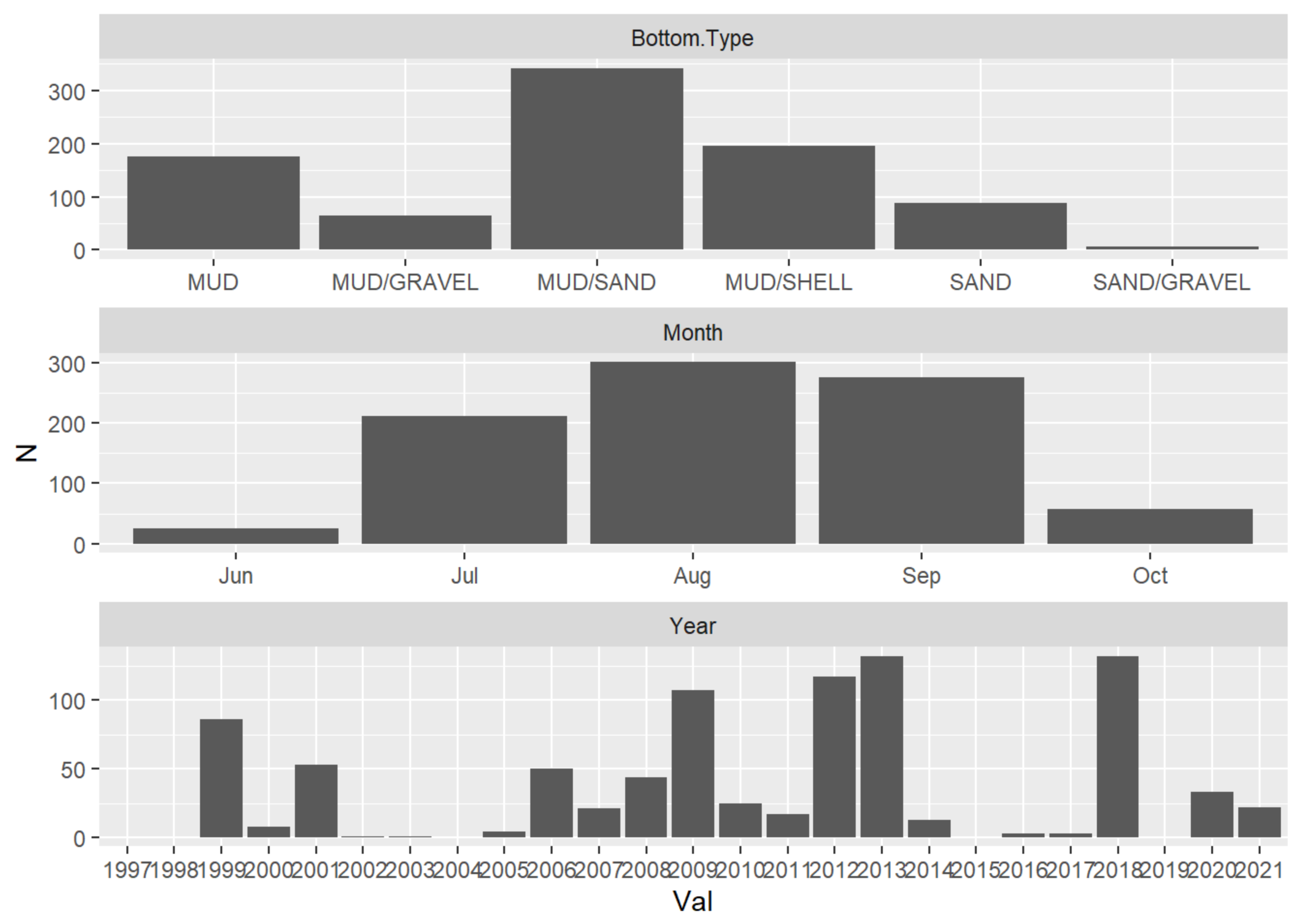
Number of records associated with various factors (y axis is not the count of bluefish)

```
samplesByLevel(nhSeine, useN = FALSE)
```



Counts of bluefish relative to the factor variables.

```
samplesByLevel(nhSeine, useN = TRUE)
```

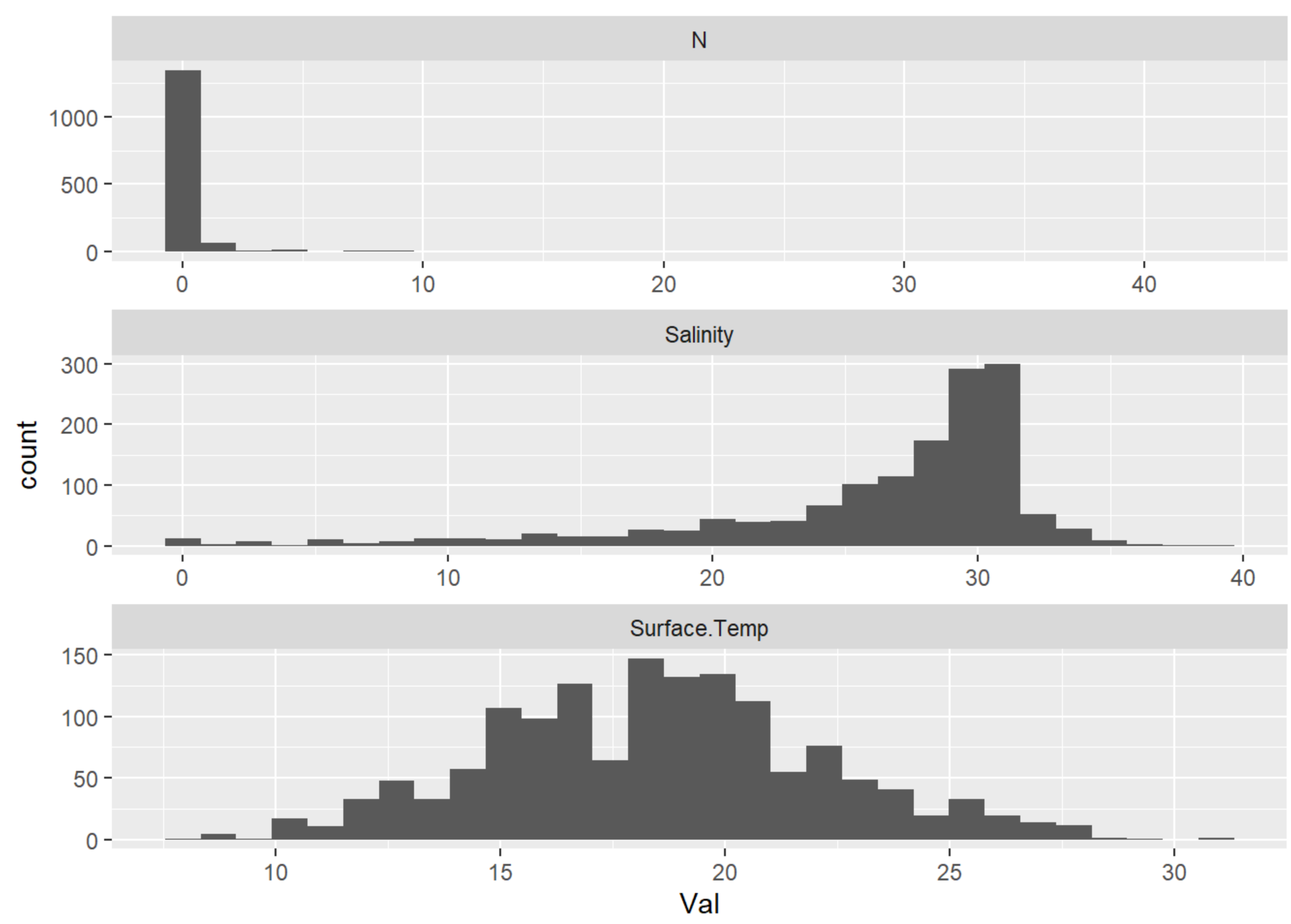


Remove years that had no bluefish records

```
nhSeine <- nhSeine %>%
  group_by(Year) %>%
  filter(!(sum(N) == 0)) %>%
  ungroup()
```

DIstribution of numeric variables

```
samplesByNumeric(nhSeine)
```

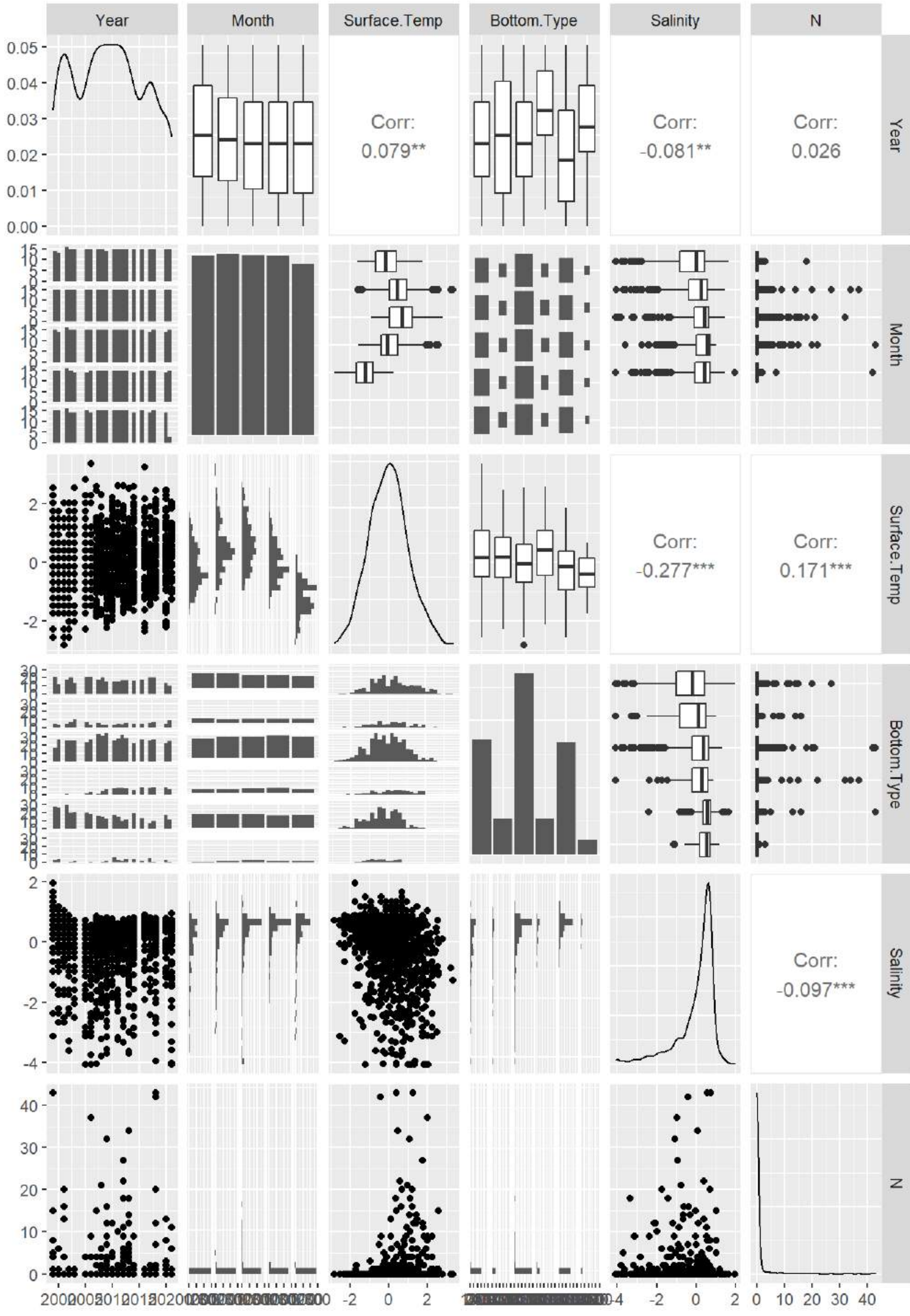


Standardize the data

```
stdData <- nhSeine %>%
  mutate(across(where(is.numeric) & -N, Z.scr))
```

Colinearity

```
stdData %>%
  mutate(Year = as.numeric(as.character(Year))) %>%
  ggpairs(bins = 30)
```



Year
Month
Surface.Temp
Bottom.Type
Salinity
N
Corr:
0.079**
Corr:
-0.081**
Corr:
0.026
Corr:
-0.277***
Corr:
0.171***
Corr:
-0.097***

```
stdData %>%
  select(where(is.numeric)) %>%
  cor(., use = 'complete.obs') %>%
  round(., digits = 3) %>%
  kbl() %>%
  kable_classic(full_width = FALSE)
```

	Surface.Temp	Salinity	N
Surface.Temp	1.000	-0.277	0.171
Salinity	-0.277	1.000	-0.097
N	0.171	-0.097	1.000

Run the models

```
var <- c('Year', 'Month', 'Bottom.Type', 'Surface.Temp', 'Salinity')
cbn <- get_combos(2:length(var))
modTxt <- lapply(cbn, function(x) c(var[1], var[x]))
mod <- modTxt %>%
  lapply(., function(x) paste(x, collapse = '+')) %>%
  lapply(., function(x) paste('N~', x, collapse = '')) %>%
  lapply(., as.formula)

AICList <- list()
modelList <- list()
predsList <- list()
for(m in 1:length(mod)){

  fullMod <- mod[[m]]
  modName <- paste(modTxt[[m]], collapse = '_')

  modElem <- str_split(fullMod, pattern = '[[:space:]]') %>%
  unlist()

  varIdx <- modElem %>%
    magrittr::is_in(c('+', '~', '1', '|')) %>%
    magrittr::not() %>%
    magrittr::extract()

  var2use <- modElem[varIdx]

  zim <- formula(~ .)
  zam <- formula(~ .)

  title <- paste(rmarkdown::metadata$title, modName) %>%
    str_replace_all(pattern = '\\.', replace = '-')
  options(knitr.duplicate.label = "allow")
  rmarkdown::render('singleIndices/diagnosticsTPL.Rmd',
                    output_file = file.path(paste0(st,
                                                   '/results/',
                                                   paste0(title, '.html'))))
  AICList[[m]] <- aic
  modelList[[m]] <- modSave
  predsList[[m]] <- preds %>%
    mutate(ModelTerms2 = modName)
```

```
}

predsAll <- bind_rows(predsList)
```

AIC Table

```
aicAll <- bind_rows(AICList) %>%
  mutate(DeltaAIC = AIC - min(AIC, na.rm = TRUE)) %>%
  arrange(DeltaAIC)

get_aicTab(aicAll, highlightRowN = 1)
```

Model	fullMod	zim	zam	AIC	DeltaAIC
neg_binom	~NYear + Month + Bottom.Type + Surface.Temp + Salinity	~.	~.	1399.504	0.000000
neg_binom	~NYear + Month + Surface.Temp + Salinity	~.	~.	1406.261	6.757837
neg_binom	~NYear + Month + Bottom.Type + Surface.Temp	~.	~.	1419.676	20.172085
zero_alt	~NYear + Month + Bottom.Type + Surface.Temp + Salinity	~.	~.	1420.069	20.565608
neg_binom	~NYear + Month + Surface.Temp	~.	~.	1420.901	21.397306
zero_alt	~NYear + Month + Surface.Temp + Salinity	~.	~.	1422.136	22.632656
neg_binom	~NYear + Month + Bottom.Type + Salinity	~.	~.	1425.540	26.036071
zero_infl	~NYear + Surface.Temp + Salinity	~.	~.	1428.127	28.623334
zero_alt	~NYear + Month + Surface.Temp	~.	~.	1431.901	32.397537
zero_alt	~NYear + Month + Bottom.Type + Surface.Temp	~.	~.	1433.432	33.928695
neg_binom	~NYear + Bottom.Type + Surface.Temp + Salinity	~.	~.	1434.456	34.952720
neg_binom	~NYear + Surface.Temp + Salinity	~.	~.	1438.562	39.058242
neg_binom	~NYear + Bottom.Type + Surface.Temp	~.	~.	1440.017	40.513442
neg_binom	~NYear + Surface.Temp	~.	~.	1441.938	42.434547
neg_binom	~NYear + Month + Salinity	~.	~.	1445.345	45.841661
zero_alt	~NYear + Bottom.Type + Surface.Temp + Salinity	~.	~.	1449.545	50.041200
zero_infl	~NYear + Bottom.Type + Surface.Temp	~.	~.	1449.570	50.066065
zero_alt	~NYear + Surface.Temp + Salinity	~.	~.	1452.173	52.669645
zero_alt	~NYear + Bottom.Type + Surface.Temp	~.	~.	1454.425	54.921021
zero_alt	~NYear + Surface.Temp	~.	~.	1454.609	55.105840
zero_alt	~NYear + Month + Bottom.Type + Salinity	~.	~.	1456.628	57.124687
zero_alt	~NYear + Month + Salinity	~.	~.	1460.767	61.263441
neg_binom	~NYear + Month + Bottom.Type	~.	~.	1480.384	80.880834
zero_infl	~NYear + Salinity	~.	~.	1495.288	95.784794
neg_binom	~NYear + Month	~.	~.	1496.695	97.191325
zero_alt	~NYear + Month + Bottom.Type	~.	~.	1497.759	98.255046
neg_binom	~NYear + Bottom.Type + Salinity	~.	~.	1505.582	106.078589
zero_alt	~NYear + Month	~.	~.	1506.392	106.888774

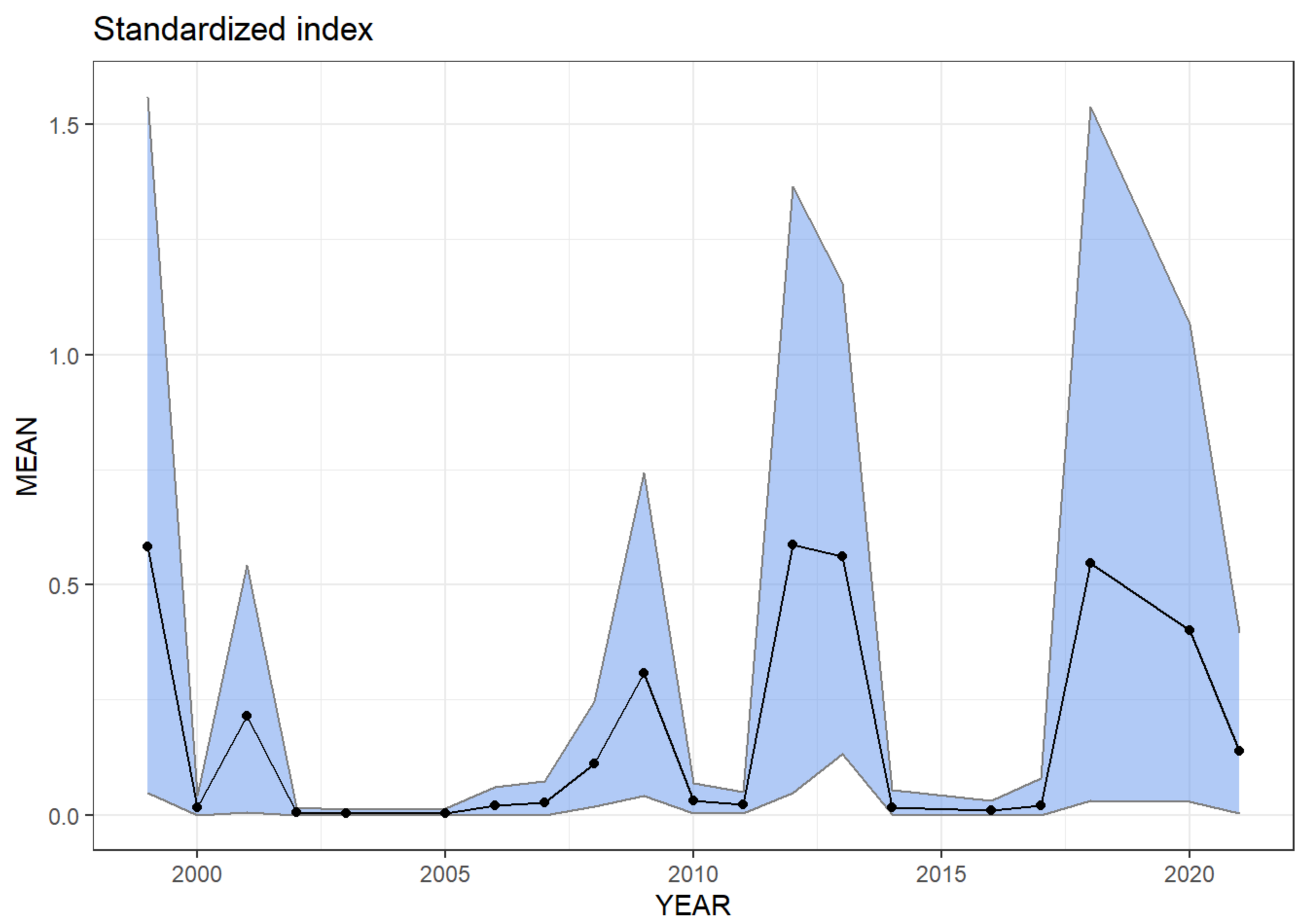
Model	fullMod	zim	zam	AIC	DeltaAIC
zero_alt	~NYear + Bottom.Type + Salinity	~.	~.	1525.002	125.498325
neg_binom	~NYear + Salinity	~.	~.	1528.696	129.191950
zero_alt	~NYear + Salinity	~.	~.	1530.862	131.357994
neg_binom	~NYear + Bottom.Type	~.	~.	1537.218	137.714182
zero_alt	~NYear + Bottom.Type	~.	~.	1551.458	151.954055
zero_infl	~NYear + Month	~.	~.	NA	NA
zero_infl	~NYear + Bottom.Type	~.	~.	NA	NA
zero_infl	~NYear + Surface.Temp	~.	~.	NA	NA
zero_infl	~NYear + Month + Bottom.Type	~.	~.	NA	NA
zero_infl	~NYear + Month + Surface.Temp	~.	~.	NA	NA
zero_infl	~NYear + Month + Salinity	~.	~.	NA	NA
zero_infl	~NYear + Bottom.Type + Salinity	~.	~.	NA	NA
zero_infl	~NYear + Month + Bottom.Type + Surface.Temp	~.	~.	NA	NA
zero_infl	~NYear + Month + Bottom.Type + Salinity	~.	~.	NA	NA
zero_infl	~NYear + Month + Surface.Temp + Salinity	~.	~.	NA	NA
zero_infl	~NYear + Bottom.Type + Surface.Temp + Salinity	~.	~.	NA	NA
zero_infl	~NYear + Month + Bottom.Type + Surface.Temp + Salinity	~.	~.	NA	NA

Chosen model is N = f(Year, Month, Bottom Type, SurfaceTemp, Salinity) with negative binomial error distribution

```
selectedModelIndex <- 15
print(mod[[selectedModelIndex]])
## N ~ Year + Month + Bottom.Type + Surface.Temp + Salinity
## <environment: 0x0000000027fef530>
```

Bootstrapped results

```
out <- boot.ZI_AT(best = modelList[[selectedModelIndex]][[1]], # a model
                  type_z = "nb", # "zero" if zero-inflated
                  nboots = 1000,
                  CI = 0.95,
                  print.prog = T)


stdPlot <- plot_boot(data = out) +
  ggtitle('Standardized index')
print(stdPlot)
```




```
options(knitr.kable.NA = "--")

out %>%
  as_tibble() %>%
  complete(YEAR = min(YEAR):max(YEAR)) %>%
  mutate(across(where(is.numeric), round, 3)) %>%
  kbl() %>%
  kable_classic(full_width = FALSE)
```

YEAR	MEAN	SE	LCI	UCI
1999	0.583	0.400	0.047	1.560
2000	0.017	0.011	0.000	0.042
2001	0.214	0.144	0.005	0.543
2002	0.004	0.005	0.000	0.017
2003	0.003	0.003	0.000	0.011
2004	–	–	–	–
2005	0.004	0.004	0.000	0.013
2006	0.020	0.018	0.000	0.059
2007	0.026	0.019	0.000	0.074

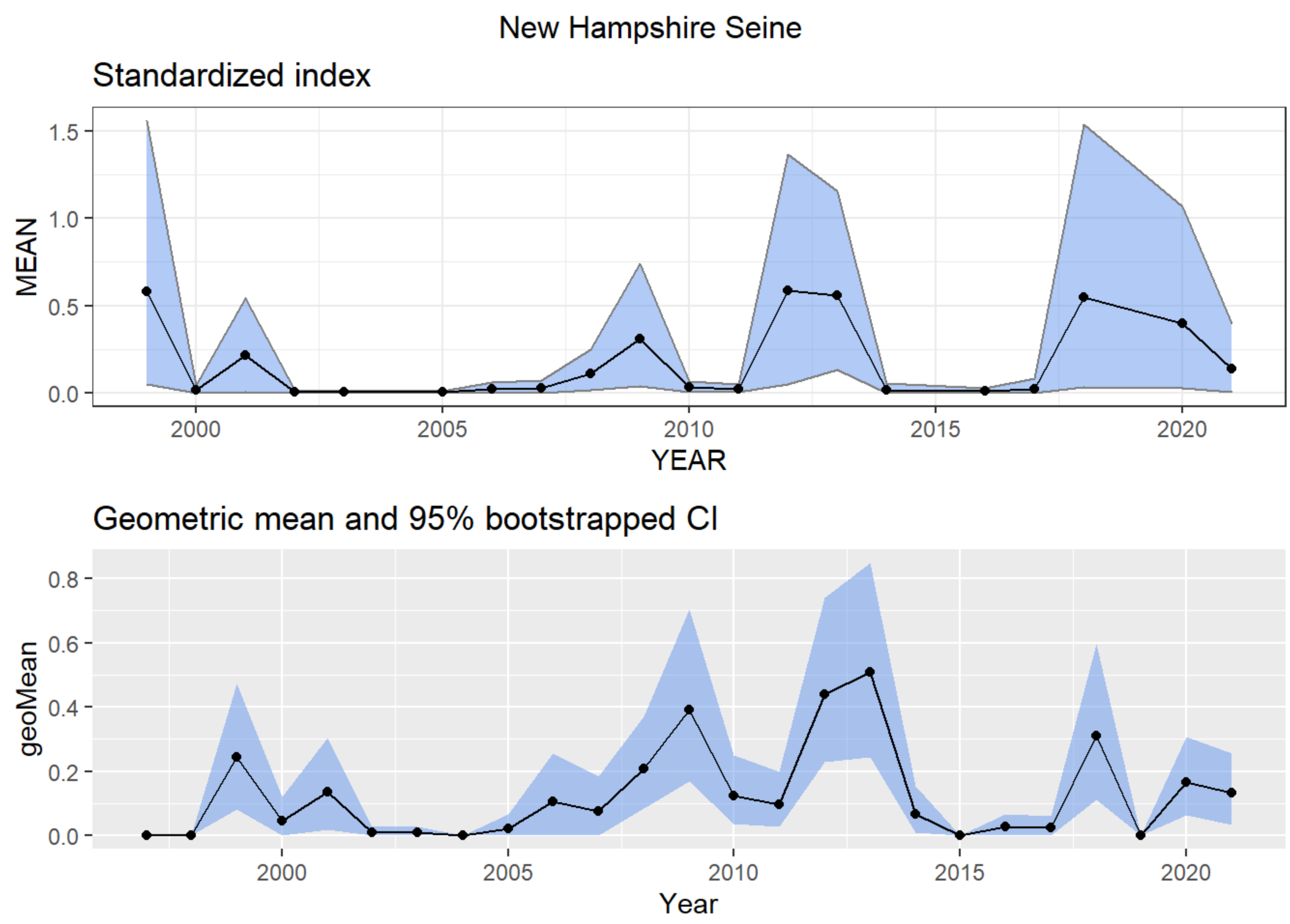
YEAR	MEAN	SE	LCI	UCI
2008	0.112	0.058	0.017	0.247
2009	0.308	0.181	0.041	0.744
2010	0.031	0.017	0.004	0.068
2011	0.022	0.012	0.003	0.049
2012	0.587	0.352	0.048	1.365
2013	0.561	0.271	0.133	1.155
2014	0.016	0.014	0.001	0.053
2015	–	–	–	–
2016	0.010	0.008	0.000	0.030
2017	0.021	0.022	0.000	0.080
2018	0.547	0.412	0.031	1.538
2019	–	–	–	–
2020	0.400	0.263	0.030	1.067
2021	0.139	0.108	0.003	0.396

Nominal index

```
bootFun <- function(x, n = 1000){
  bootMean_i <- numeric(n)
  for(i in 1:n){
    nd <- sample(x, size = length(x), replace = TRUE)
    bootMean_i[i] <- exp(mean(log(nd+1)))-1
  }
  bootPct <- quantile(bootMean_i, probs = c(0.025, 0.975))
  bootSD <- sd(bootMean_i)
  return(tibble(bootSD = bootSD, boot05 = bootPct[1], boot95 = bootPct[2]))
}

nominalGeoMean <- nhSeine_nomDat %>%
  group_by(Year) %>%
  summarize(mean = mean(N),
            sd = sd(N),
            geoMean = exp(mean(log(N + 1))) - 1,
            bootRes = bootFun(x = N),
            .groups = 'drop') %>%
  unpack(bootRes) %>%
  rename(geoMeanBootSD = bootSD) %>%
  mutate(geoCV = geoMeanBootSD / geoMean,
         across(where(is.numeric), round, 3),
         Year = as.numeric(as.character(Year)))

nominalGeoMean %>%
  complete(Year = min(Year):max(Year)) %>%
  kbl() %>%
  kable_classic(full_width = FALSE)
```

Year	mean	sd	geoMean	geoMeanBootSD	boot05	boot95	geoCV
1997	0.000	0.000	0.000	0.000	0.000	0.000	–
1998	0.000	0.000	0.000	0.000	0.000	0.000	–
1999	1.147	5.387	0.245	0.101	0.082	0.473	0.414
2000	0.107	0.709	0.045	0.031	0.000	0.122	0.688
2001	0.662	3.194	0.135	0.073	0.020	0.305	0.541
2002	0.013	0.115	0.009	0.009	0.000	0.028	0.989
2003	0.013	0.115	0.009	0.009	0.000	0.028	0.997
2004	0.000	0.000	0.000	0.000	0.000	0.000	–
2005	0.053	0.462	0.022	0.021	0.000	0.066	0.964
2006	0.667	4.400	0.106	0.068	0.000	0.257	0.638
2007	0.280	1.798	0.076	0.051	0.000	0.185	0.668
2008	0.587	2.574	0.209	0.075	0.086	0.372	0.362
2009	1.427	4.645	0.391	0.139	0.168	0.704	0.355
2010	0.333	1.563	0.124	0.055	0.038	0.250	0.449
2011	0.227	1.192	0.097	0.044	0.028	0.198	0.456
2012	1.560	4.751	0.439	0.134	0.230	0.740	0.305
2013	1.760	5.154	0.509	0.151	0.245	0.847	0.297
2014	0.173	1.070	0.066	0.037	0.009	0.155	0.569
2015	0.000	0.000	0.000	0.000	0.000	0.000	–
2016	0.040	0.197	0.028	0.017	0.000	0.067	0.587
2017	0.040	0.257	0.024	0.017	0.000	0.060	0.711
2018	1.760	7.339	0.309	0.125	0.111	0.595	0.406
2019	0.000	0.000	0.000	0.000	0.000	0.000	–
2020	0.440	1.810	0.167	0.061	0.063	0.308	0.367
2021	0.349	1.578	0.134	0.060	0.034	0.257	0.446

```
write_csv(nominalGeoMean, file = 'NH/nominalGeoMean_NH.csv')
nomPlot <- nominalGeoMean %>%
  ggplot(aes(x = Year)) +
    geom_ribbon(aes(ymin = boot05, ymax = boot95, group = 1),
                fill = 'cornflowerblue', alpha = 0.5) +
  geom_line(aes(y = geoMean), group = 1) +
  geom_point(aes(y = geoMean)) +
  ggtitle('Geometric mean and 95% bootstrapped CI')

grid.arrange(stdPlot, nomPlot, top = 'New Hampshire Seine')
```



Appendix: NB/Zero inflated/Hurdle models with selected terms

The following shows model diagnostics associated with the final selected terms. Note that all three error distributions are included below, but the **negative binomial** model was selected.

Year_Month_Bottom.Type_Surface.Temp_Salinity

Variance inflation factor

```
if(!any(modElem == '|') & length(all.vars(fullMod)) > 2){
  lm(fullMod, data = stdData) %>%
    car::vif(.)
```

```
}else{
  cat('Random effects within formula or model contains fewer than two
terms:',
      'VIF not computed', sep = ' ')
}
##                   GVIF Df GVIF^(1/(2*Df))
## Year          1.415886 19        1.009193
## Month         2.167725  4        1.101541
## Bottom.Type   1.358207  5        1.031090
## Surface.Temp 2.354849  1        1.534552
## Salinity      1.531308  1        1.237460
```

Model fits

```
coefTab <- function(x){
  if(class(x)[1] != 'try-error'){
    summary(x)$coefficients$cond %>%
    as_tibble(rownames = 'Parameter') %>%
    mutate(across(where(is.numeric), round, 3)) %>%
    kbl() %>%
    kable_classic(full_width = FALSE)
  }
}

# Negative binomial fit
neg_binom = glmmTMB(fullMod,
                    data = stdData,
                    family = nbinom2)
coefTab(neg_binom)
```

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.426	0.744	-4.608	0.000
Year2000	-3.623	0.906	-4.001	0.000
Year2001	-0.978	0.713	-1.371	0.170
Year2002	-4.929	1.314	-3.750	0.000
Year2003	-5.216	1.337	-3.901	0.000
Year2005	-5.021	1.095	-4.585	0.000
Year2006	-3.349	0.856	-3.912	0.000
Year2007	-3.200	0.941	-3.402	0.001
Year2008	-1.754	0.755	-2.324	0.020
Year2009	-0.741	0.723	-1.025	0.306
Year2010	-3.025	0.786	-3.849	0.000
Year2011	-3.384	0.832	-4.068	0.000
Year2012	0.022	0.758	0.029	0.977
Year2013	-0.158	0.690	-0.229	0.819
Year2014	-3.707	0.969	-3.827	0.000
Year2016	-4.219	1.031	-4.094	0.000

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
Year2017	-3.627	1.102	-3.290	0.001
Year2018	-0.024	0.721	-0.034	0.973
Year2020	-0.459	0.683	-0.672	0.502
Year2021	-1.407	0.701	-2.007	0.045
MonthJul	2.054	0.615	3.341	0.001
MonthAug	3.038	0.680	4.468	0.000
MonthSep	3.720	0.651	5.714	0.000
MonthOct	3.140	0.746	4.207	0.000
Bottom.TypeMUD/GRAVEL	-0.408	0.583	-0.699	0.484
Bottom.TypeMUD/SAND	0.708	0.359	1.970	0.049
Bottom.TypeMUD/SHELL	1.727	0.530	3.259	0.001
Bottom.TypeSAND	0.221	0.461	0.480	0.632
Bottom.TypeSAND/GRAVEL	1.874	0.960	1.952	0.051
Surface.Temp	1.379	0.264	5.227	0.000
Salinity	-0.799	0.181	-4.413	0.000

```
# Zero inflated fit
zero_infl = glmmTMB(fullMod,
                    ziformula = zim,
                    data = stdData,
                    family = nbinom2) %>%
  try()
coefTab(zero_infl)
```

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.856	NaN	NaN	NaN
Year2000	-1.524	NaN	NaN	NaN
Year2001	1.142	NaN	NaN	NaN
Year2002	-4.136	NaN	NaN	NaN
Year2003	-3.393	NaN	NaN	NaN
Year2005	-3.039	NaN	NaN	NaN
Year2006	0.194	NaN	NaN	NaN
Year2007	-2.610	NaN	NaN	NaN
Year2008	-1.250	NaN	NaN	NaN
Year2009	-0.304	NaN	NaN	NaN
Year2010	-1.908	NaN	NaN	NaN
Year2011	-2.187	NaN	NaN	NaN
Year2012	2.263	NaN	NaN	NaN
Year2013	0.800	NaN	NaN	NaN
Year2014	-2.437	NaN	NaN	NaN
Year2016	-3.291	NaN	NaN	NaN

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
Year2017	-3.089	NaN	NaN	NaN
Year2018	0.533	NaN	NaN	NaN
Year2020	-1.105	NaN	NaN	NaN
Year2021	-1.595	NaN	NaN	NaN
MonthJul	2.717	NaN	NaN	NaN
MonthAug	4.168	NaN	NaN	NaN
MonthSep	3.950	NaN	NaN	NaN
MonthOct	3.459	NaN	NaN	NaN
Bottom.TypeMUD/GRAVEL	-0.295	NaN	NaN	NaN
Bottom.TypeMUD/SAND	0.918	NaN	NaN	NaN
Bottom.TypeMUD/SHELL	0.826	NaN	NaN	NaN
Bottom.TypeSAND	0.699	NaN	NaN	NaN
Bottom.TypeSAND/GRAVEL	0.141	NaN	NaN	NaN
Surface.Temp	0.792	NaN	NaN	NaN
Salinity	-0.209	NaN	NaN	NaN

```
# Zero altered fit
zero_alt <- glmmTMB(fullMod,
                    ziformula = zam,
                    data = stdData,
                    family = truncated_nbinom2(link = "log")) %>%
  try()
coefTab(zero_alt)
```

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.465	1.362	0.342	0.733
Year2000	-2.663	1.342	-1.984	0.047
Year2001	0.565	1.129	0.501	0.617
Year2002	-22.202	15350.800	-0.001	0.999
Year2003	-22.394	16172.919	-0.001	0.999
Year2005	-3.163	1.883	-1.679	0.093
Year2006	-1.205	1.406	-0.857	0.391
Year2007	-2.378	1.315	-1.808	0.071
Year2008	-1.597	1.000	-1.597	0.110
Year2009	-0.623	0.957	-0.651	0.515
Year2010	-2.739	1.014	-2.701	0.007
Year2011	-2.858	1.119	-2.555	0.011
Year2012	-0.479	0.912	-0.525	0.599
Year2013	-0.548	0.896	-0.611	0.541
Year2014	-2.953	1.289	-2.291	0.022
Year2016	-21.304	4033.886	-0.005	0.996

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
Year2017	-3.749	1.815	-2.065	0.039
Year2018	-0.136	0.961	-0.141	0.888
Year2020	-1.161	0.824	-1.409	0.159
Year2021	-1.402	0.930	-1.508	0.132
MonthJul	-0.023	1.217	-0.019	0.985
MonthAug	0.257	1.353	0.190	0.850
MonthSep	0.228	1.304	0.175	0.861
MonthOct	0.962	1.429	0.673	0.501
Bottom.TypeMUD/GRAVEL	-0.259	0.649	-0.399	0.690
Bottom.TypeMUD/SAND	0.875	0.457	1.914	0.056
Bottom.TypeMUD/SHELL	1.032	0.571	1.808	0.071
Bottom.TypeSAND	0.959	0.717	1.338	0.181
Bottom.TypeSAND/GRAVEL	-0.584	1.381	-0.423	0.672
Surface.Temp	1.009	0.439	2.299	0.022
Salinity	-0.067	0.237	-0.283	0.777

```
aic <- tibble(
  Model = c('neg_binom', 'zero_infl', 'zero_alt'),
  fullMod = paste(as.character(fullMod), collapse = ''),
  zim = paste(as.character(zim), collapse = ''),
  zam = paste(as.character(zam), collapse = ''),
  AIC = c(AIC(neg_binom), AIC(zero_infl), AIC(zero_alt))
)

modSave <- list(neg_binom, zero_infl, zero_alt)

neg_binomPred <- ggemmeans(model = neg_binom, terms = 'Year', type = 'fe')
%>%
  try()
zero_inflPred <- ggemmeans(model = zero_infl, terms = 'Year', type = 'fe.zi')
%>%
  try()
## Error : Predicted values could not be computed. Try reducing number of
simulation, using argument `nsim` (e.g. `nsim = 100`)
# Help says fe for conditional model
zero_altPred <- ggemmeans(model = zero_alt, terms = 'Year', type = 'fe.zi')
%>%
  try()

makeNAPred <- function(x){
  if('try-error' %in% class(x)){
    out <- list(x = NA, predicted = NA, std.error = NA,
                conf.low = NA, conf.high = NA)
  }else{
    out <- x
  }
}

neg_binomPred <- makeNAPred(neg_binomPred)
```

```
zero_inflPred <- makeNAPred(zero_inflPred)
zero_altPred <- makeNAPred(zero_altPred)

neg_binomPredDF <- tibble(
  Model = 'neg_binom',
  paste(as.character(fullMod), collapse = ''),
  zim = paste(as.character(zim), collapse = ''),
  zam = paste(as.character(zam), collapse = ''),
  Year = neg_binomPred$x,
  Idx = neg_binomPred$predicted,
  se = neg_binomPred$std.error,
  conf.low = neg_binomPred$conf.low,
  conf.high = neg_binomPred$conf.high
)

zero_inflPredDF <- tibble(
  Model = 'zero_infl',
  paste(as.character(fullMod), collapse = ''),
  zim = paste(as.character(zim), collapse = ''),
  zam = paste(as.character(zam), collapse = ''),
  Year = zero_inflPred$x,
  Idx = zero_inflPred$predicted,
  se = zero_inflPred$std.error,
  conf.low = zero_inflPred$conf.low,
  conf.high = zero_inflPred$conf.high
)

zero_altPredDF <- tibble(
  Model = 'zero_alt',
  ModelTerms = paste(as.character(fullMod), collapse = ''),
  zim = paste(as.character(zim), collapse = ''),
  zam = paste(as.character(zam), collapse = ''),
  Year = zero_altPred$x,
  Idx = zero_altPred$predicted,
  se = ifelse(!is.null(zero_altPred$std.error), zero_altPred$std.error, NA),
  conf.low = ifelse(!is.null(zero_altPred$conf.low),
                    zero_altPred$conf.low, NA),
  conf.high = ifelse(!is.null(zero_altPred$conf.high),
                     zero_altPred$conf.high, NA)
)

preds <- bind_rows(neg_binomPredDF, zero_inflPredDF, zero_altPredDF)
```

Model diagnostics

Get the DHARMa residuals

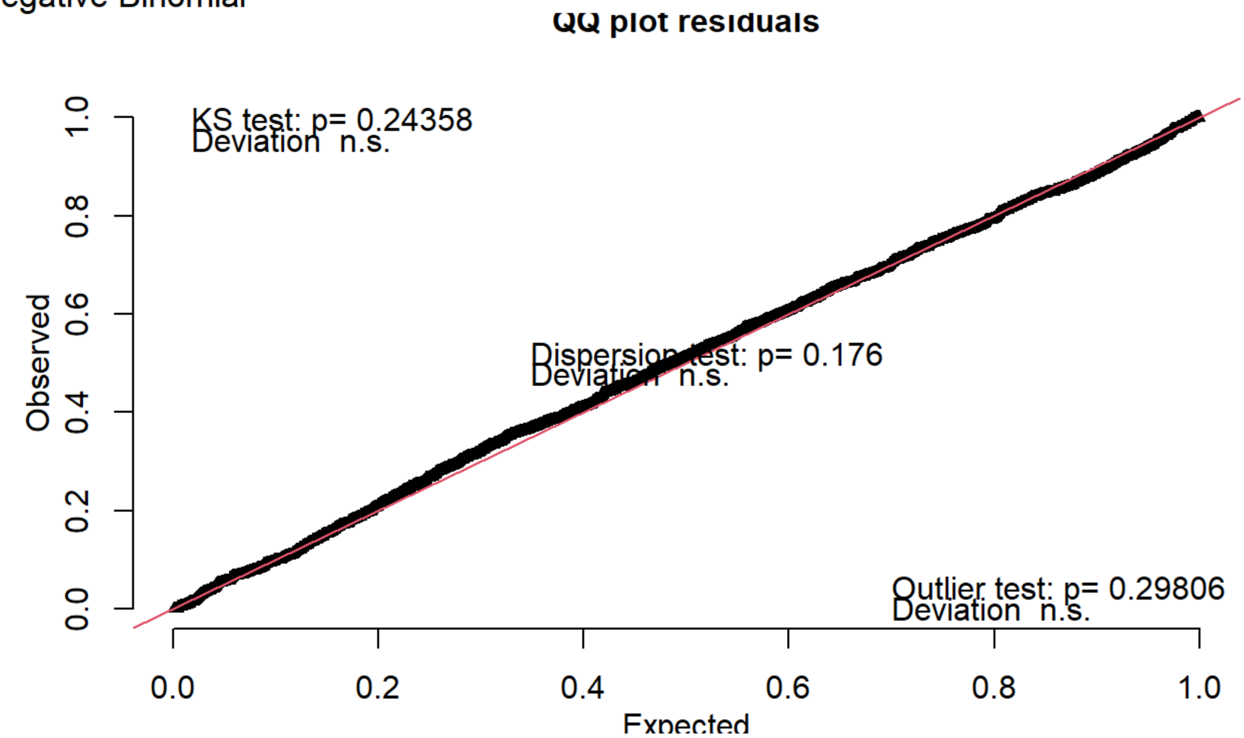
```
soNeg_binom <- simulateResiduals(neg_binom) %>%
  try()
soZero_infl <- simulateResiduals(zero_infl) %>%
  try()
soZero_alt <- simulateResiduals(zero_alt) %>%
  try()
```

```
if(class(soNeg_binom)[1] == 'try-error'){
  soNeg_binom <- NULL
}
if(class(soZero_infl)[1] == 'try-error'){
  soZero_infl <- NULL
}
if(class(soZero_alt)[1] == 'try-error'){
  soZero_alt <- NULL
}

# Be sure to maintain order bwt these two lists
soLst <- list(`Negative Binomial` = soNeg_binom,
              `Zero inflated` = soZero_infl,
              `Zero altered` = soZero_alt)
modLst <- list(`Negative Binomial` = neg_binom,
              `Zero inflated` = zero_infl,
              `Zero altered` = zero_alt)
anyNull <- which(sapply(soLst, is.null))
if(length(anyNull) > 0){
  soLst <- soLst[-anyNull]
  modLst <- modLst[-anyNull]
}

if(length(modLst) < 1){
  cat('---- No models converged from this run ----')
}
```

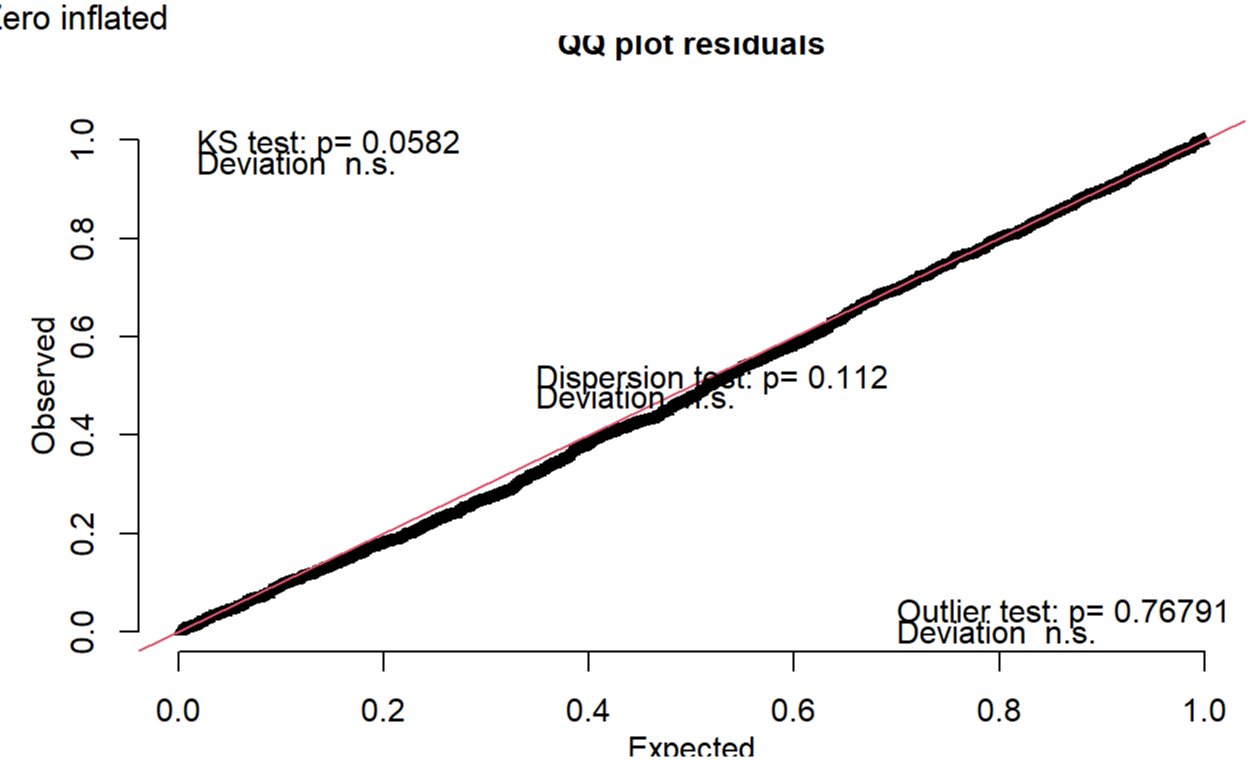
QQ plots

```
qq <- list()
for(i in 1:length(soLst)){
  qq[[i]] <- as.ggplot(~plotQQunif(soLst[[i]])) +
    ggtitle(names(soLst)[i])
}
grid.arrange(grobs = qq, ncol = 1)
```

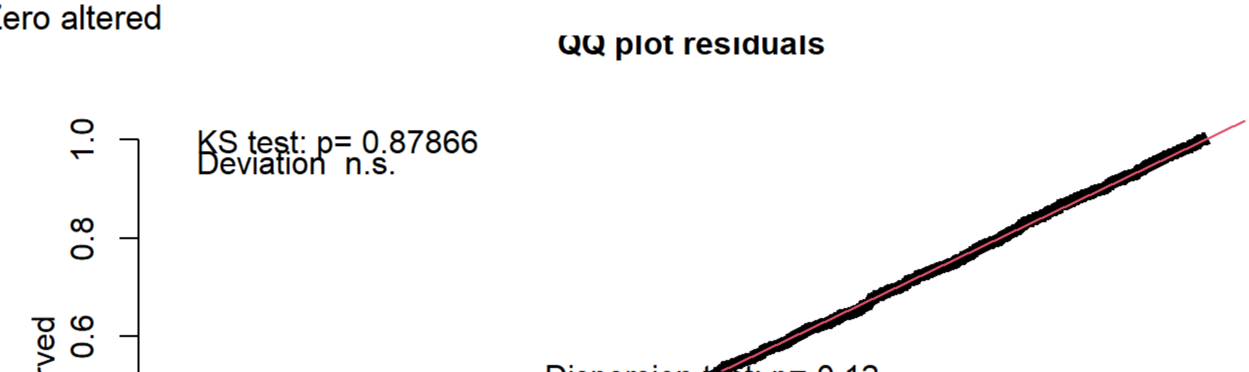
Negative Binomial



Zero inflated



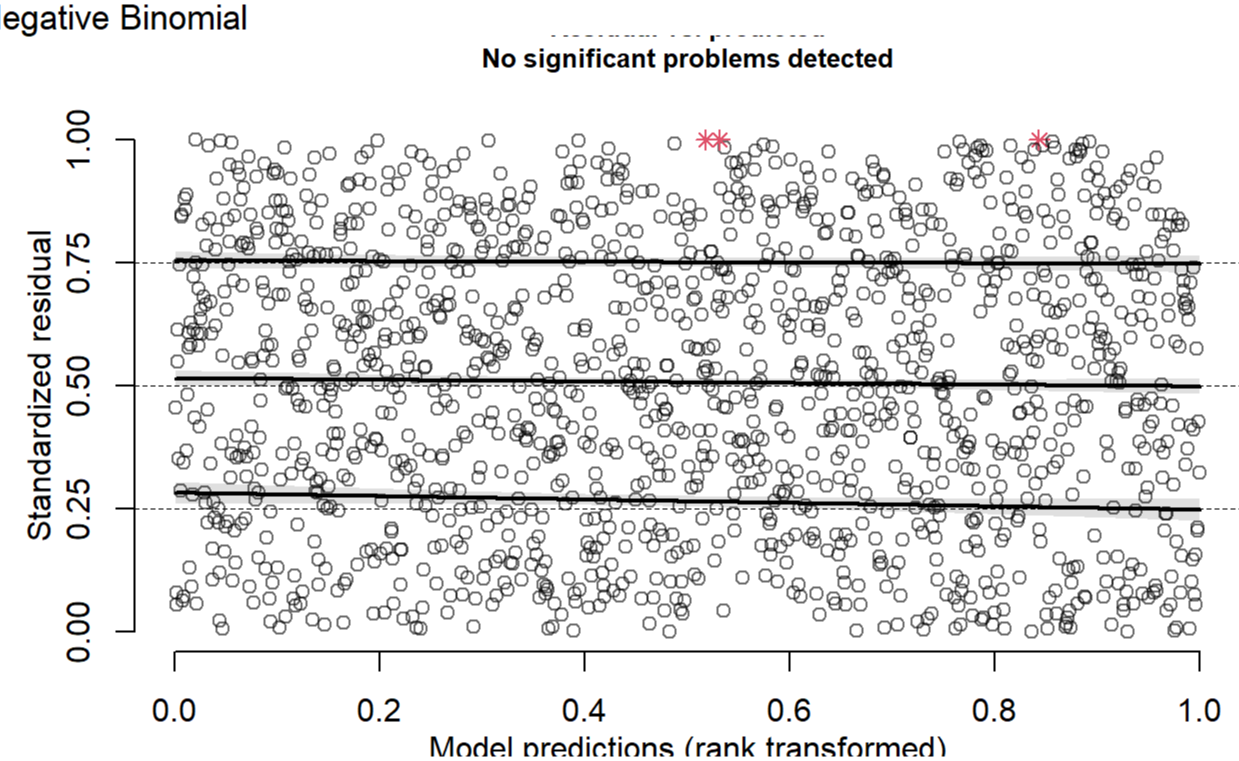
Zero altered



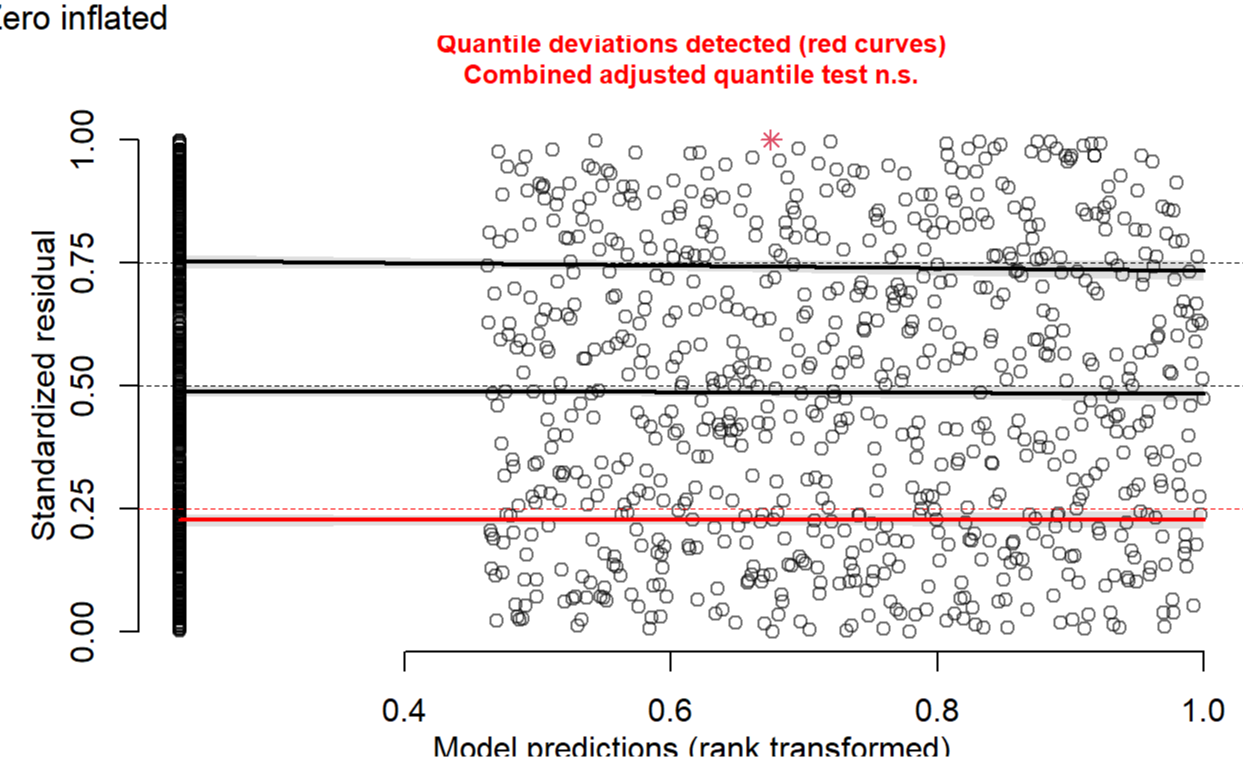
Residuals versus predicted values

```
residVsPred <- list()
for(i in 1:length(soLst)){
  residVsPred[[i]] <- as.ggplot(~plotResiduals(soLst[[i]], quantreg = TRUE))
+
    ggtitle(names(soLst)[i])
}
grid.arrange(grobs = residVsPred, ncol = 1)
```

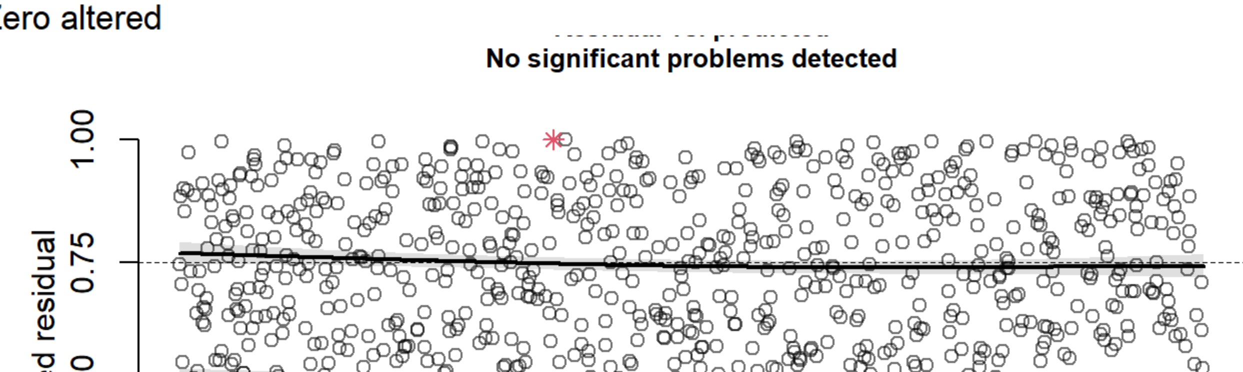
Negative Binomial



Zero inflated

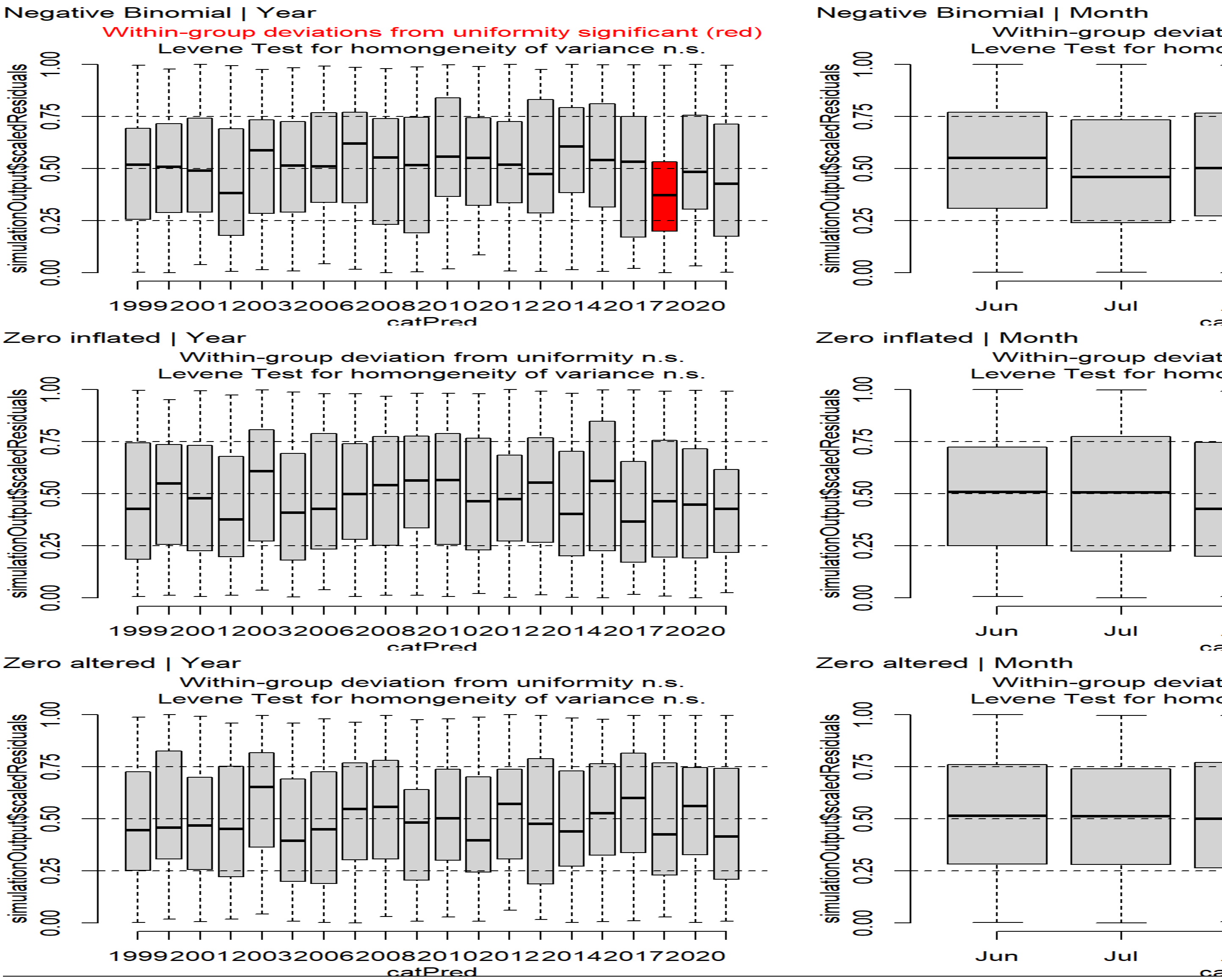


Zero altered



Residuals versus predictors

```
residVsVar <- list()
for(i in 1:length(soLst)){

  residVsVarM <- list()
  frameVar <- names(modLst[[i]]$frame)[-which(names(modLst[[i]]$frame) ==
'N')]
  for(j in 1:length(frameVar)){
    fv <- frameVar[j]
    predictor <- modLst[[i]]$frame[[fv]]
    tplot <- try(as.ggplot(~plotResiduals(soLst[[i]], form = predictor)))
    if(class(tplot)[1] != 'try-error'){
      residVsVarM[[j]] <- tplot
    }else{
      residVsVarM[[j]] <- ggplot() +
        annotate("text", x = 0, y = 0, size=8, label = 'Unsuccessful')
    }
    residVsVarM[[j]] <- residVsVarM[[j]] +
      ggtitle(paste(names(soLst)[i], '|', fv))
  }
  residVsVar[[i]] <- arrangeGrob(grobs = residVsVarM, nrow = 1)

}
grid.arrange(grobs = residVsVar, ncol = 1)
```

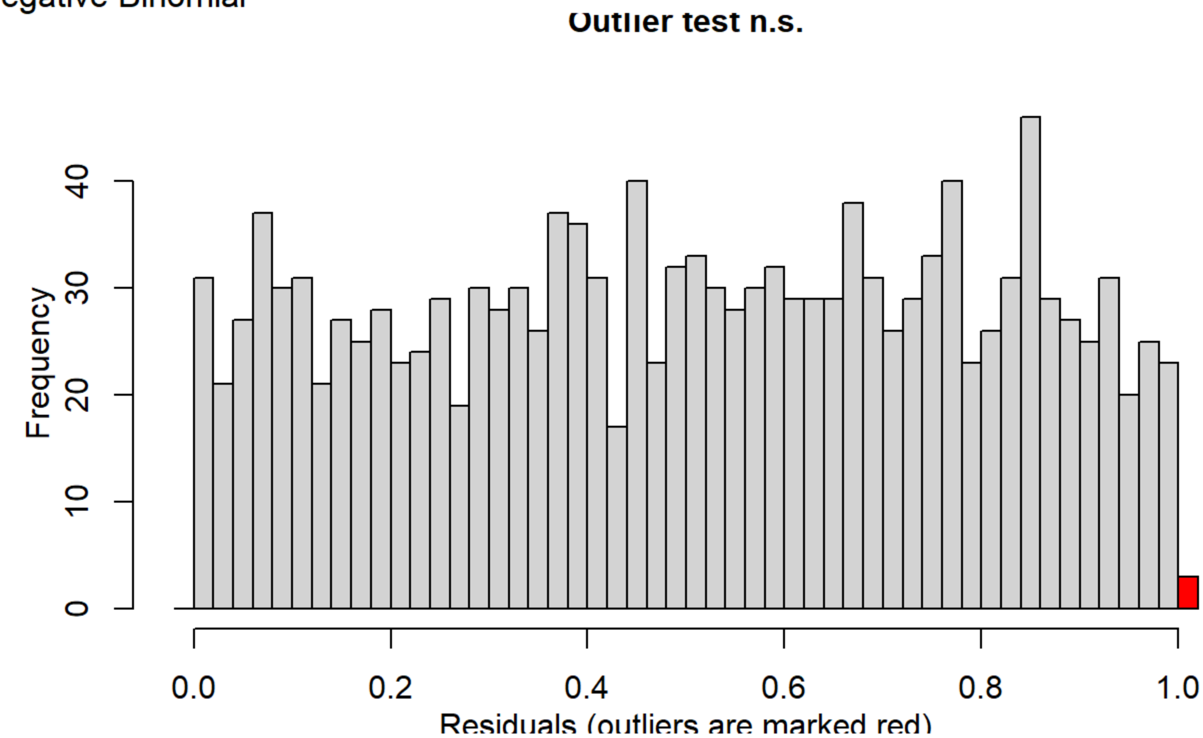



```
# unif <- list()
# for(i in 1:length(soLst)){
#   unif[[i]] <- as.ggplot(~testUniformity(soLst[[i]])) +
#     ggtitle(names(soLst)[i])
# }
# grid.arrange(grobs = unif, ncol = 1)
```

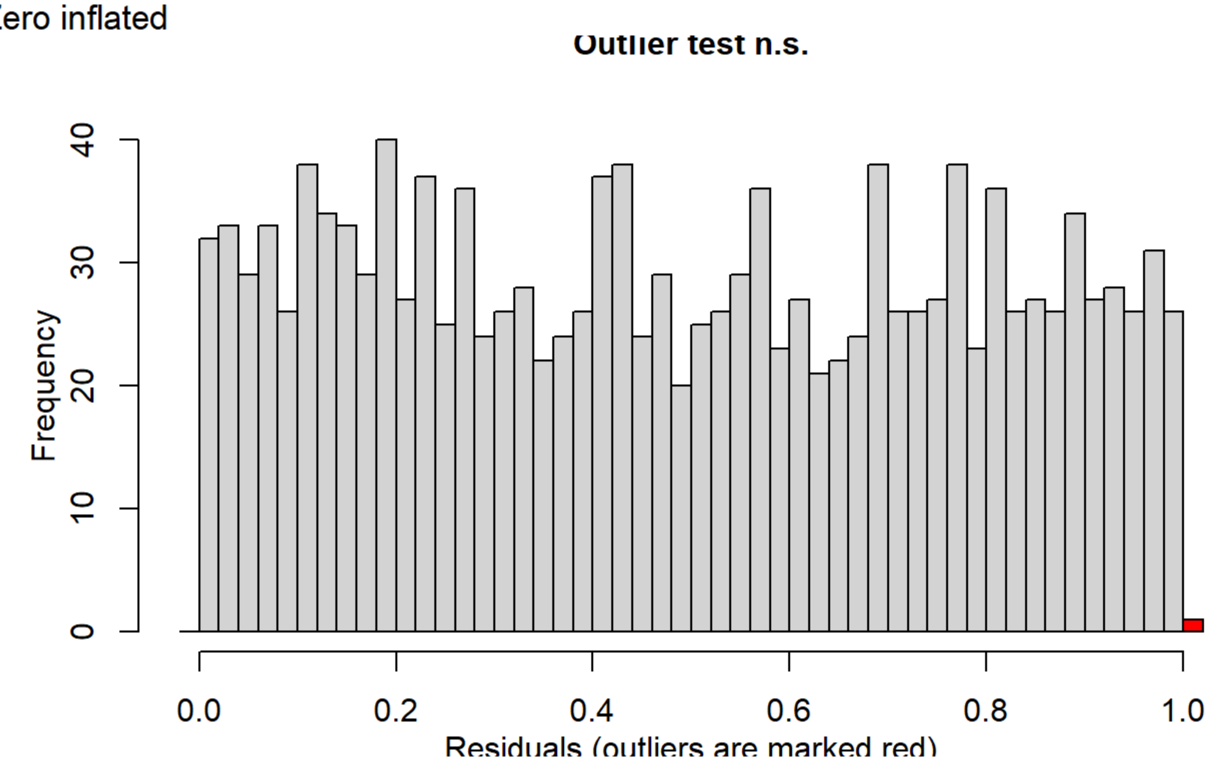
testOutliers() - tests if there are more simulation outliers than expected

```
out <- list()
for(i in 1:length(soLst)){
  out[[i]] <- as.ggplot(~testOutliers(soLst[[i]])) +
    ggtitle(names(soLst)[i])
}
grid.arrange(grobs = out, ncol = 1)
```

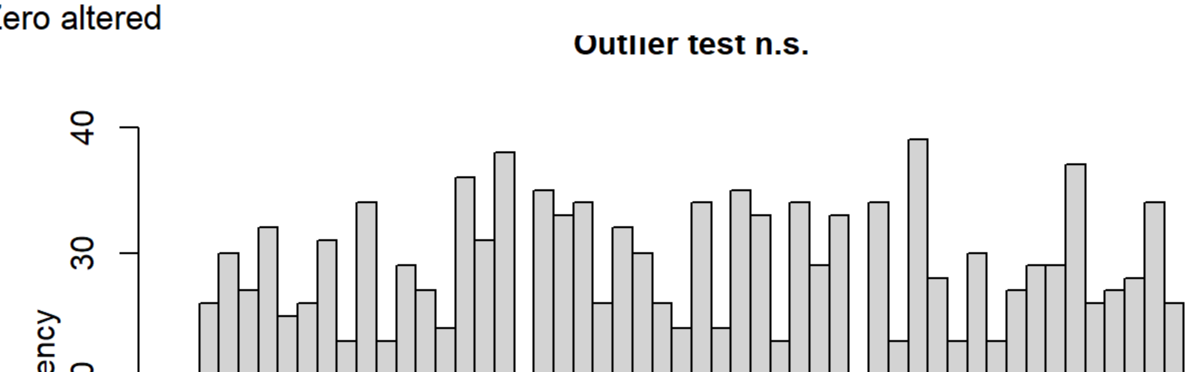
Negative Binomial



Zero inflated



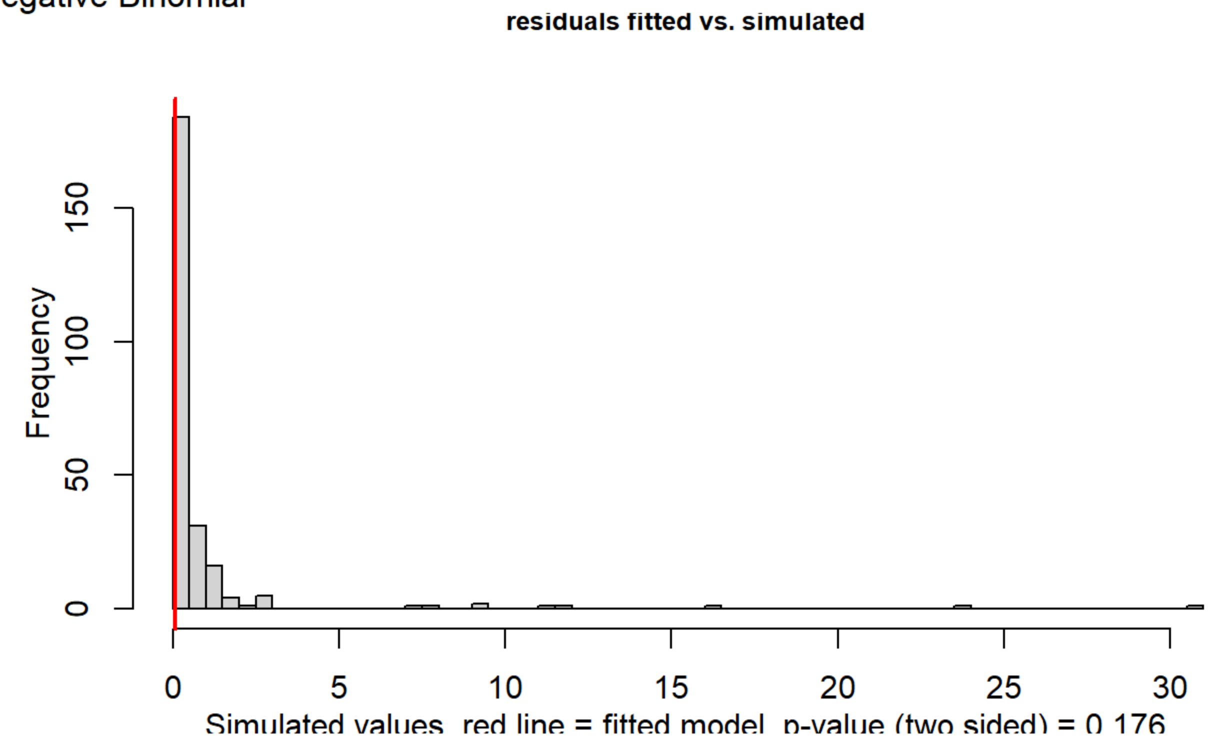
Zero altered



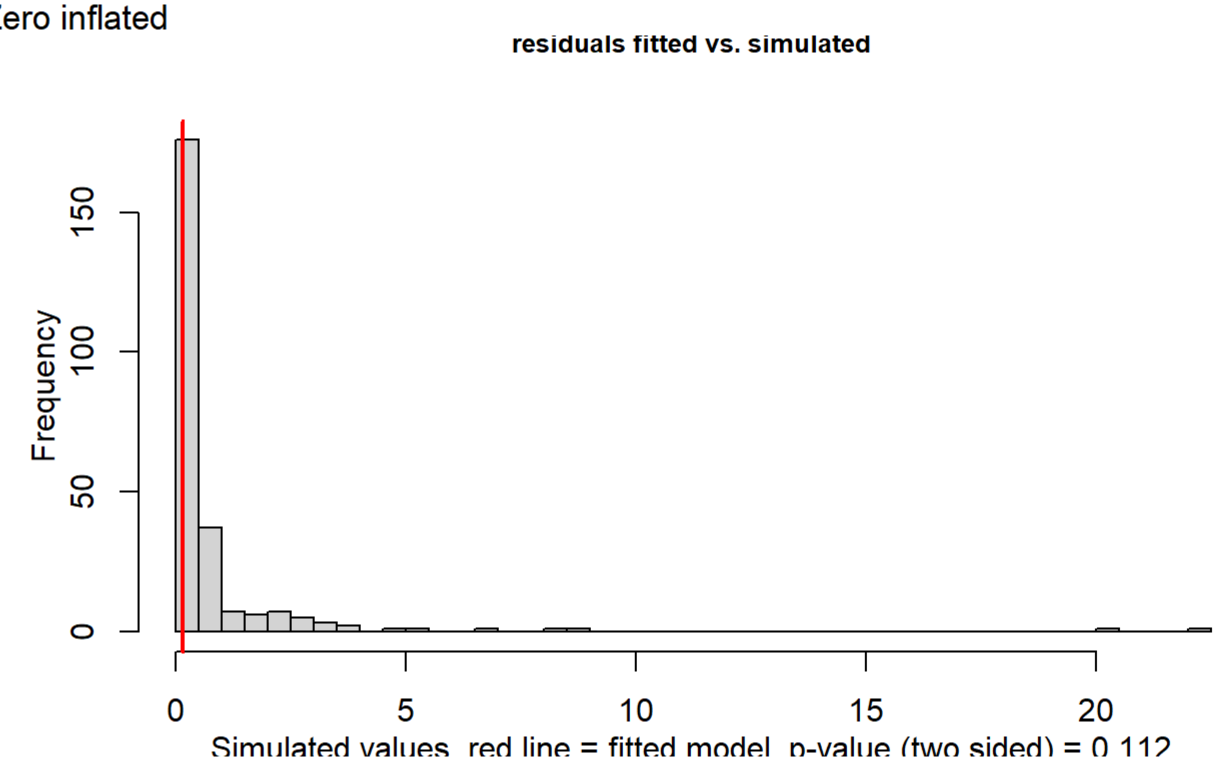
testDispersion() - tests if the simulated dispersion is equal to the observed dispersion

```
disp <- list()
for(i in 1:length(soLst)){
  disp[[i]] <- as.ggplot(~testDispersion(soLst[[i]])) +
    ggtitle(names(soLst)[i])
}
grid.arrange(grobs = disp, ncol = 1)
```

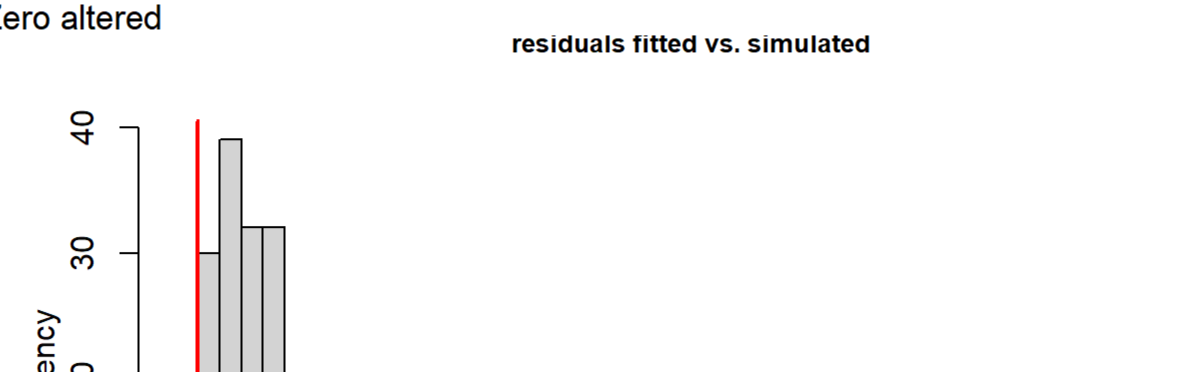
Negative Binomial



Zero inflated



Zero altered

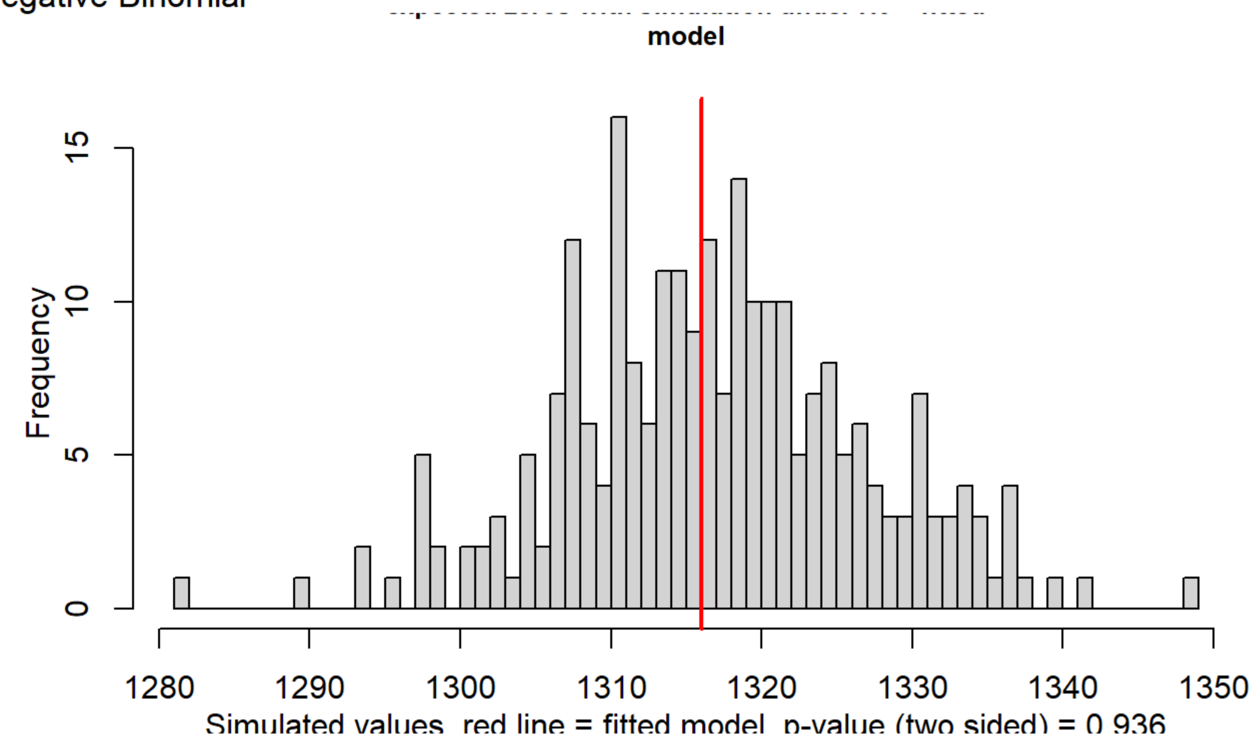


```
# quant <- list()
# for(i in 1:length(soLst)){
#   quant[[i]] <- as.ggplot(~testQuantiles(soLst[[i]])) +
#     ggtitle(names(soLst)[i])
# }
# grid.arrange(grobs = quant, ncol = 1)
```

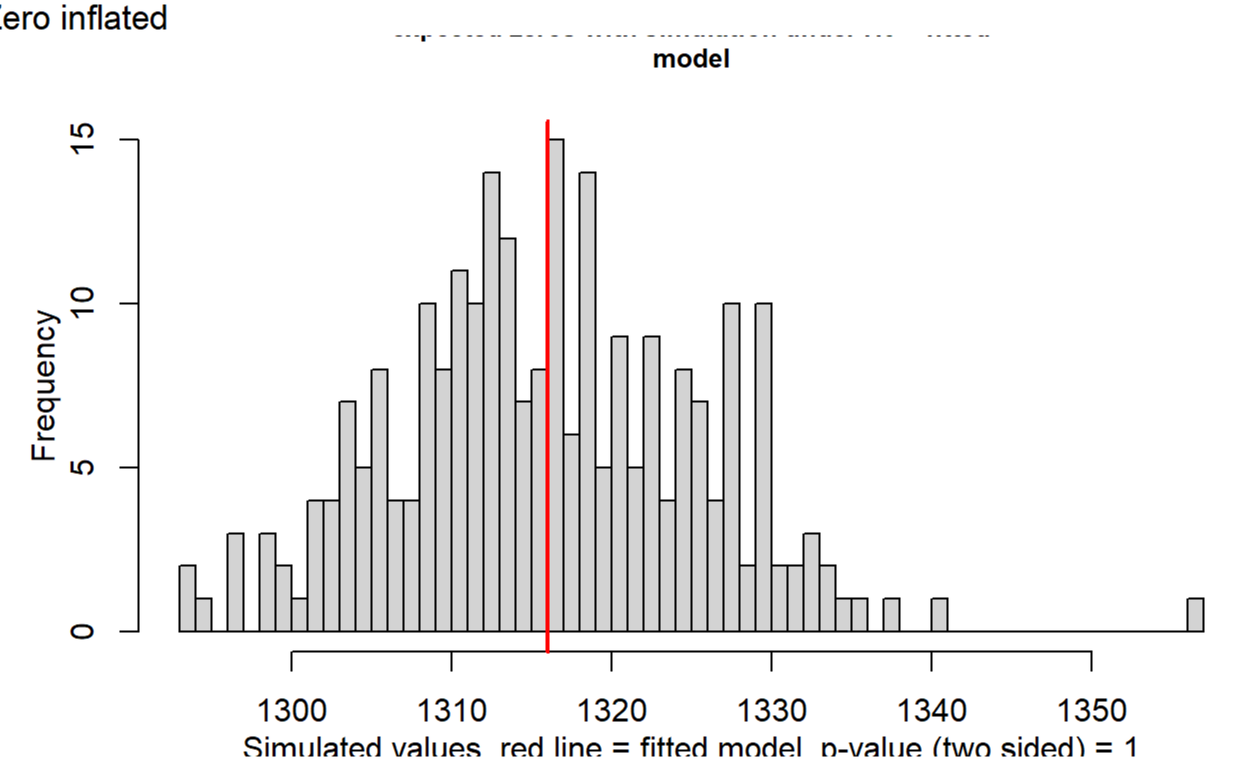
testZeroinflation() - tests if there are more zeros in the data than expected from the simulations

```
zi <- list()
for(i in 1:length(soLst)){
  zi[[i]] <- as.ggplot(~testZeroInflation(soLst[[i]])) +
    ggtitle(names(soLst)[i])
}
grid.arrange(grobs = zi, ncol = 1)
```

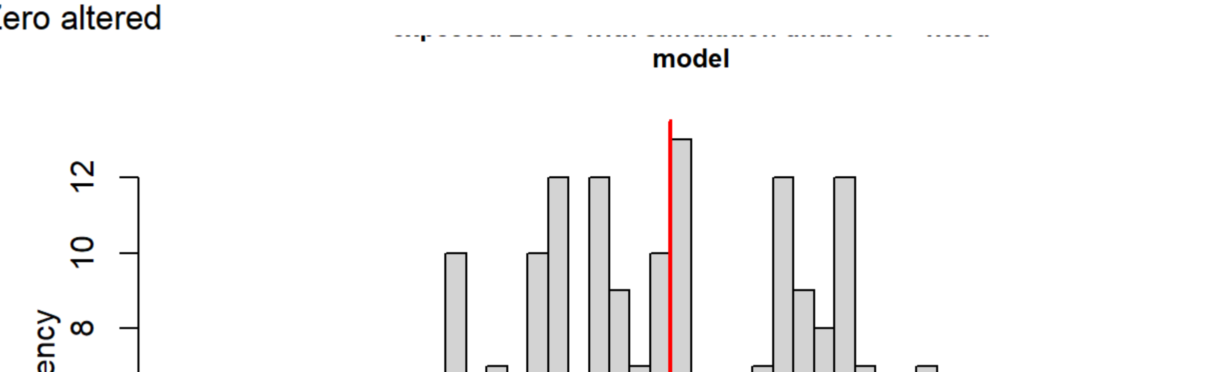
Negative Binomial



Zero inflated



Zero altered



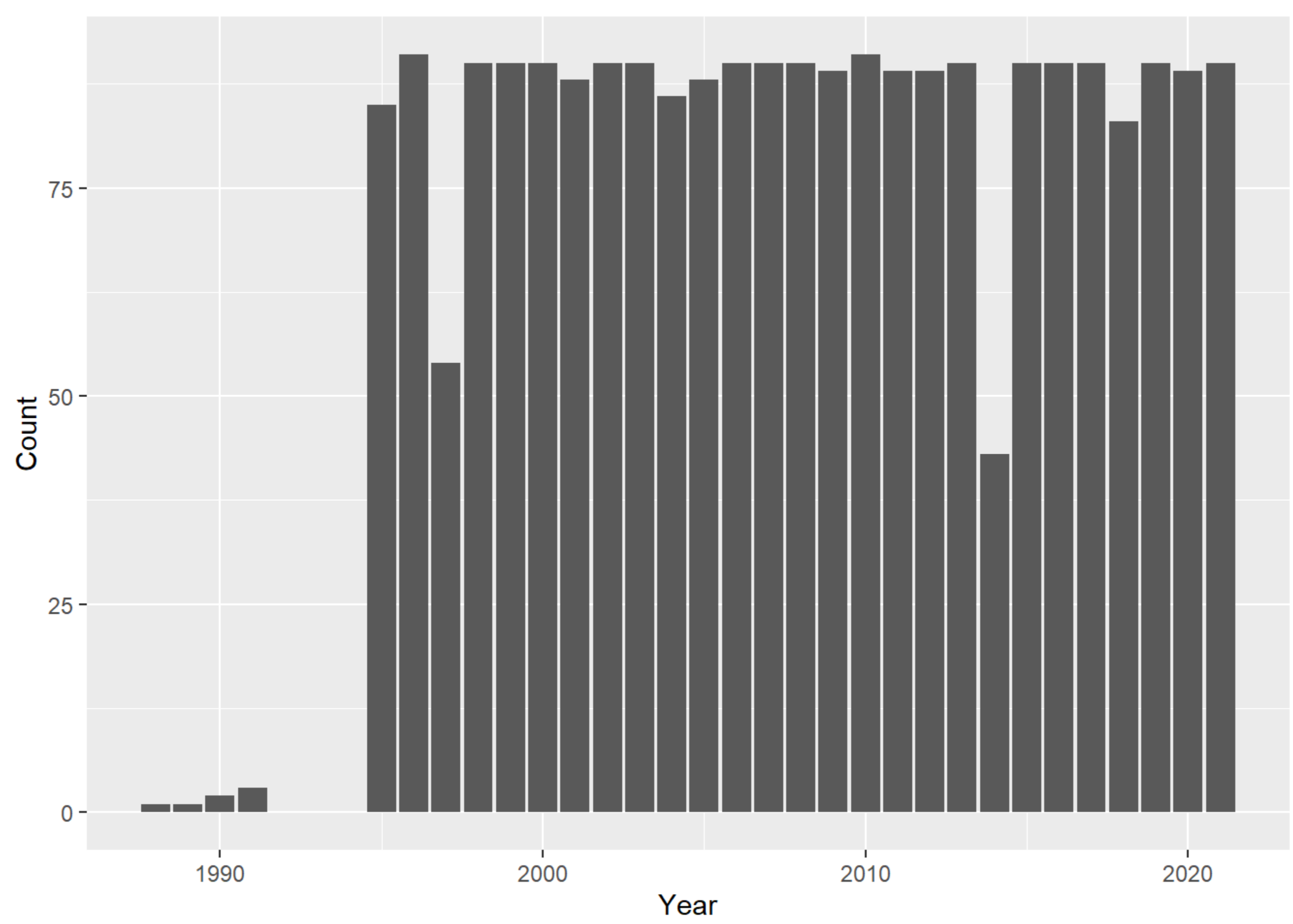
RI Seine

```
st <- 'RISeine'
source('singleIndices/get_packages.R')
source('functions/get_combos.R')
source('functions/cleanData.R')
source('functions/samplesByLevel.R')
source('functions/samplesByNumeric.R')
source('functions/Zscr.R')
source('functions/get_aicTab.R')
source('functions/get_combos.R')
source('functions/get_idxPlot.R')
source('functions/bootstrap_functions.R')
ctNBS <- c(SampleID = 'text',
           State = 'text',
           Program = 'text',
           Gear = 'text',
           Year = 'text',
           Month = 'text',
           Day = 'numeric',
           Time = 'text',
           Effort = 'text',
           N = 'numeric',
           Weight = 'text',
           Area = 'text',
           Stratum = 'text',
           Depth = 'numeric',
           Wgt.Factor = 'numeric',
           Surface.Temp = 'numeric',
           Bottom.Temp = 'numeric',
           Salinith = 'numeric',
           DO = 'numeric',
           Bottom.Type = 'text')

riSeine <- read_excel(path = 'data/2021
data/2021_bluefish_data_RI_April_2022.xlsx',
                       sheet = 'SurveyData_NBS',
                       col_types = ctNBS) %>%

  mutate(Month = month(ymd(paste('1900-', Month, '-01'))),
         Effort = 1, # All effort is "single haul"
         Season = 'Summer-Fall',

         # All Nov dates w/in first 4 days and very few; lump in with Oct
         Month = ifelse(Month == 11, 10, 11)) %>%
         cleanData(.) %>%

  # Include only the necessary data
  select(Year, Month, Station, Bottom.Temp, Salinity, DO, N)
```

A number of early time series years lack DO measurements, so drop this variable.

```
riSeine %>%
  group_by(Year) %>%
```

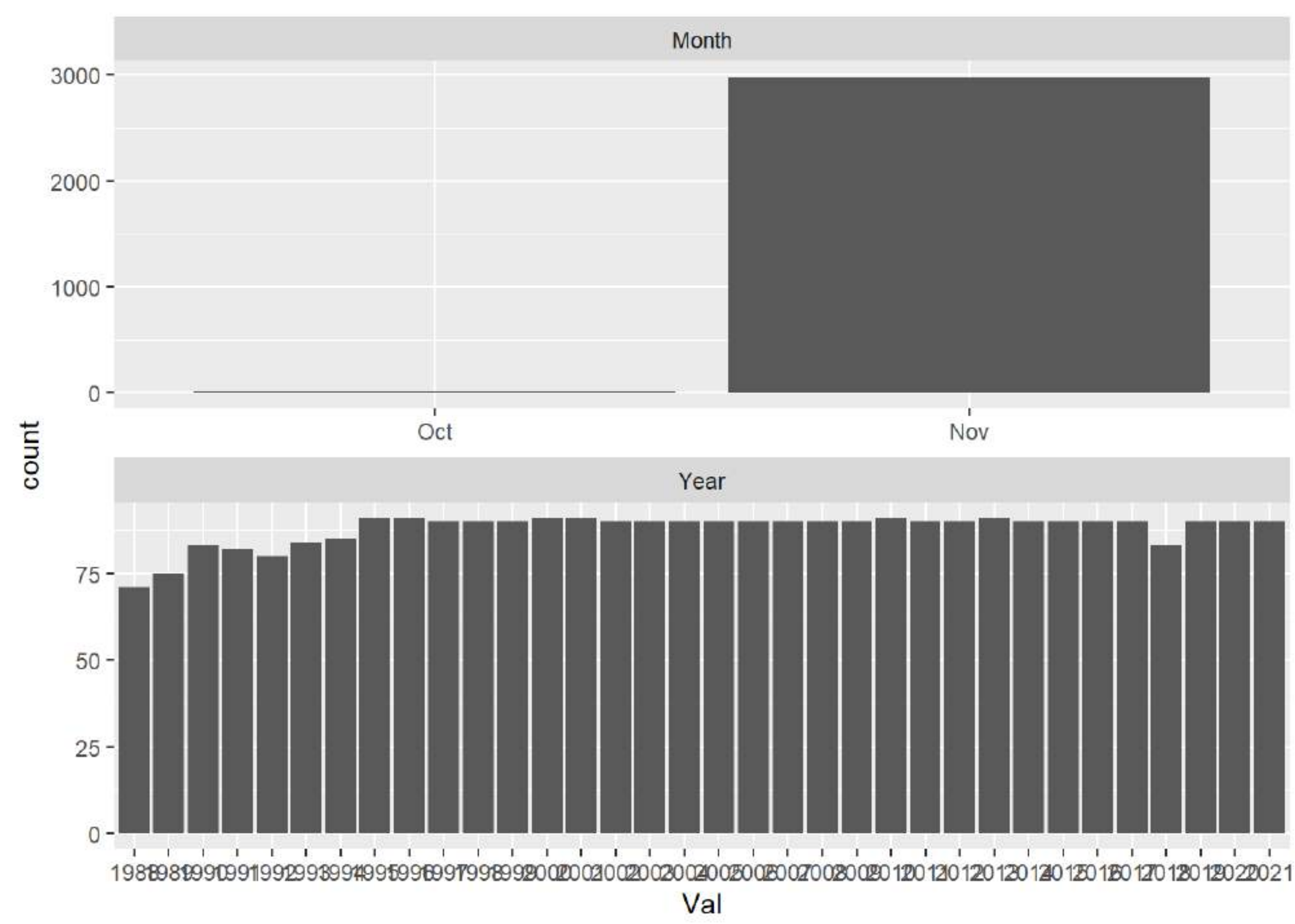
```
  summarize(Count = sum(!is.na(DO))) %>%
  mutate(Year = as.numeric(as.character(Year))) %>%
  ggplot(aes(x = Year, y = Count)) +
  geom_col()
```




```
riSeine <- riSeine %>%
  select(-DO)
```

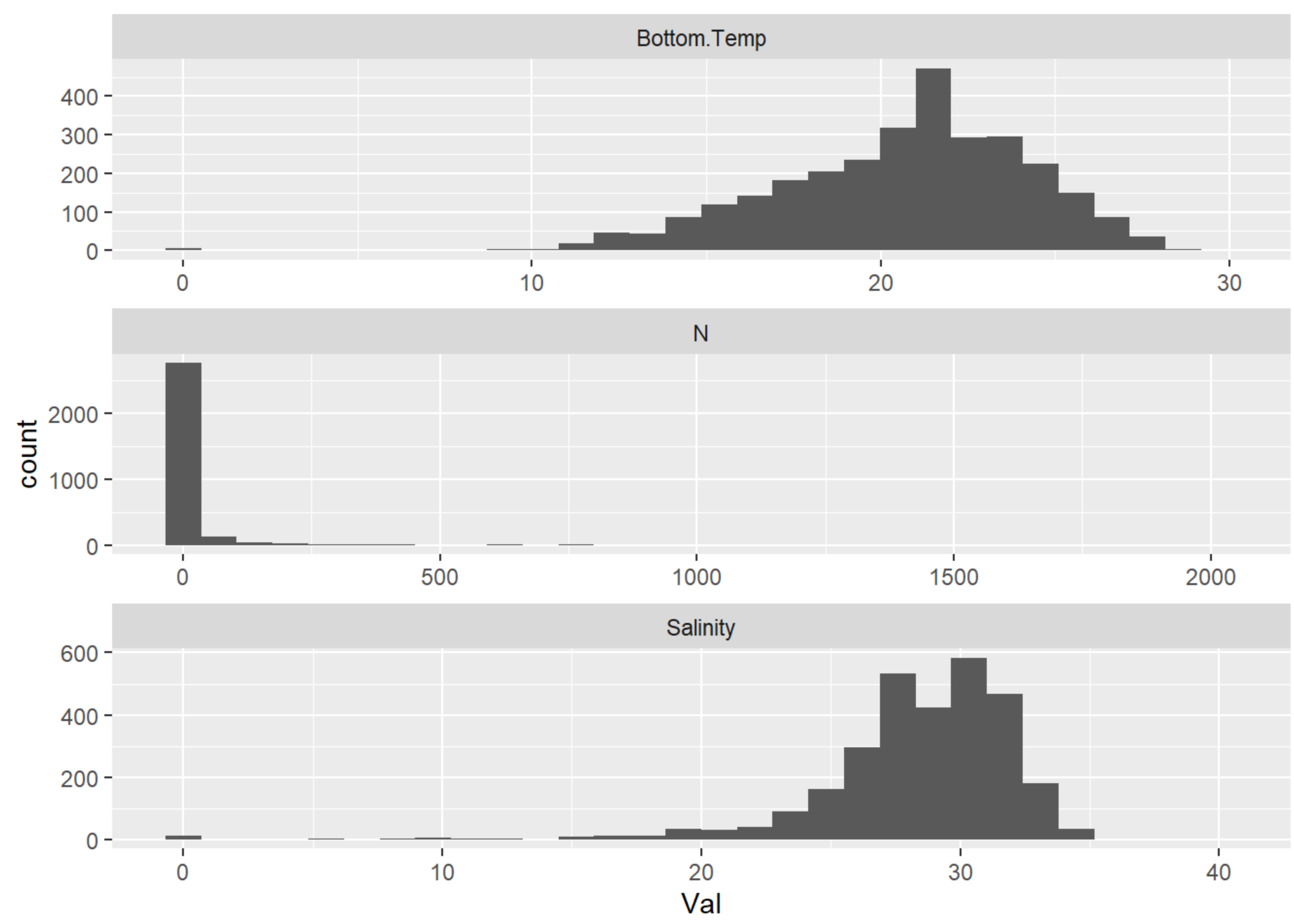
Number of records associated with various factors (y axis is not the count of bluefish). Only very few samples from October, so don't include this in the model.

```
samplesByLevel(riSeine, useN = FALSE)
```



Distribution of numeric variables

```
samplesByNumeric(riSeine)
```

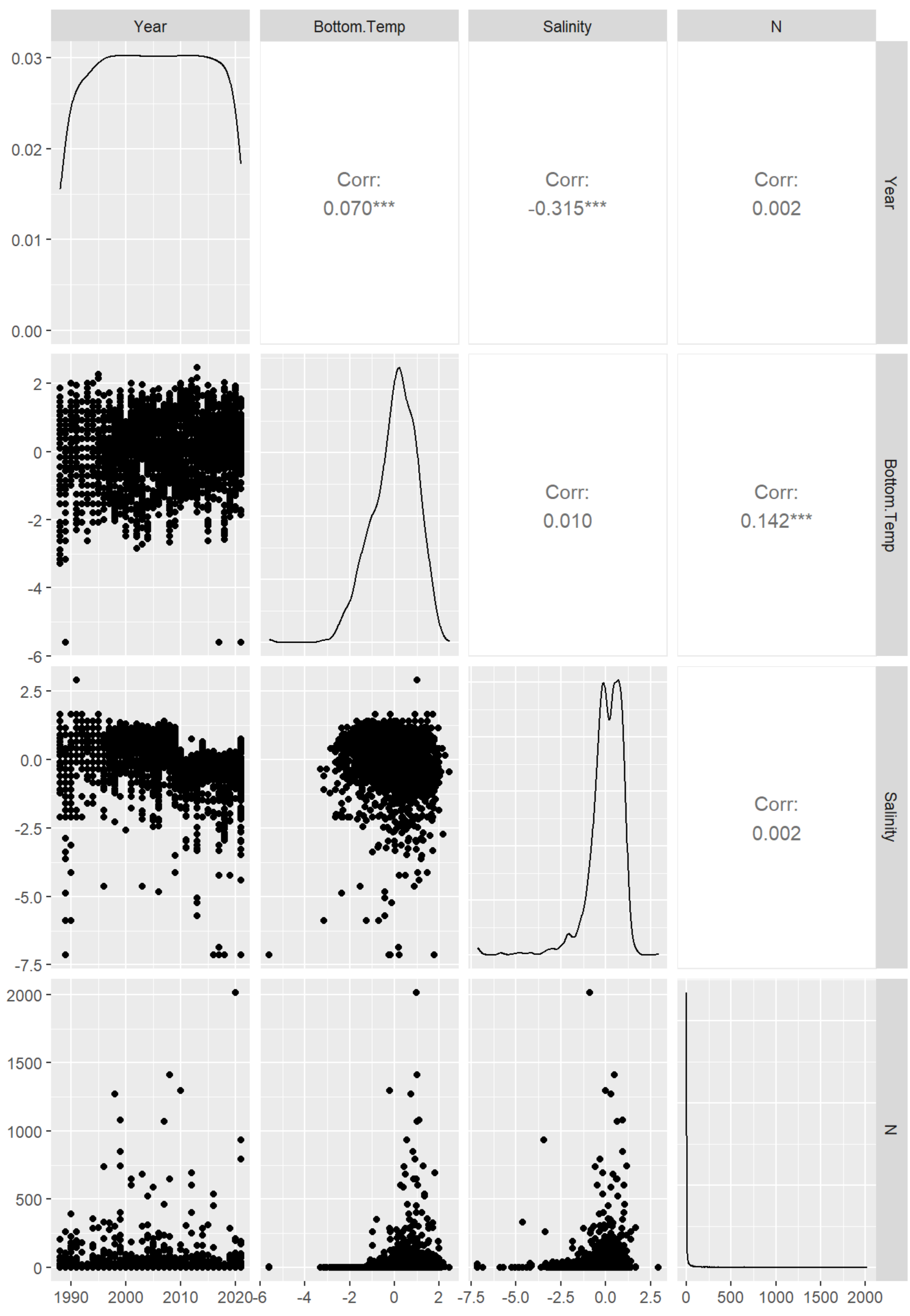


Standardize the data

```
stdData <- riSeine %>%
  mutate(across(where(is.numeric) & -N, Z.scr))
```

Colinearity

```
stdData %>%
  select(-Station, -Month) %>%
  mutate(Year = as.numeric(as.character(Year))) %>%
  ggpairs(bins = 30)
```


Year
Bottom.Temp
Salinity
N
Corr:
0.070***
Corr:
-0.315***
Corr:
0.002
Corr:
0.010
Corr:
0.142***
Corr:
0.002
0.03
0.02
0.01
0.00
1990
2000
2010
2020
-6
-4
-2
0
2
-7.5
-5.0
-2.5
0.0
2.5
500
1000
1500
2000

```
stdData %>%
  select(where(is.numeric)) %>%
  cor(., use = 'complete.obs') %>%
  round(., digits = 3) %>%
  kbl() %>%
  kable_classic(full_width = FALSE)
```

	Bottom.Temp	Salinity	N
Bottom.Temp	1.000	0.010	0.142
Salinity	0.010	1.000	0.002
N	0.142	0.002	1.000

Run the models

```
var <- c('Year', 'Bottom.Temp', 'Salinity')
cbn <- get_combos(2:length(var))
modTxt <- lapply(cbn, function(x) c(var[1], var[x]))
mod <- modTxt %>%
  lapply(., function(x) paste(x, collapse = '+')) %>%
  lapply(., function(x) paste('N~', x, collapse = '')) %>%
  lapply(., as.formula)

AICList <- list()
modelList <- list()
predsList <- list()
for(m in 1:length(mod)){

  fullMod <- mod[[m]]
  modName <- paste(modTxt[[m]], collapse = '_')

  modElem <- str_split(fullMod, pattern = '[[:space:]]') %>%
  unlist()

  varIdx <- modElem %>%
    magrittr::is_in(c('+', '~', '1', '|')) %>%
    magrittr::not() %>%
    magrittr::extract()

  var2use <- modElem[varIdx]

  zim <- formula(~ .)
  zam <- formula(~ .)

  title <- paste(rmarkdown::metadata$title, modName)
  options(knitr.duplicate.label = "allow")
  rmarkdown::render('singleIndices/diagnosticsTPL.Rmd',
                    output_file = file.path(paste0(st,
                                                   '/results/',
                                                   paste0(title, '.html'))))
  AICList[[m]] <- aic
  modelList[[m]] <- modSave
  predsList[[m]] <- preds %>%
    mutate(ModelTerms2 = modName)

}
```

```
predsAll <- bind_rows(predsList)
```

AIC Table

Some diagnostics issues (see below) but not terrible.

```
aicAll <- bind_rows(AICList) %>%
  mutate(DeltaAIC = AIC - min(AIC, na.rm = TRUE)) %>%
  arrange(DeltaAIC)
```

```
get_aicTab(aicAll, highlightRowN = 1)
```

Model	fullMod	zim	zam	AIC	DeltaAIC
zero_alt	~NYear + Bottom.Temp + Salinity	~.	~.	11229.04	0.00000
zero_infl	~NYear + Bottom.Temp + Salinity	~.	~.	11240.39	11.35428
zero_alt	~NYear + Bottom.Temp	~.	~.	11309.34	80.29695
zero_infl	~NYear + Bottom.Temp	~.	~.	11316.19	87.15189
neg_binom	~NYear + Bottom.Temp + Salinity	~.	~.	11500.58	271.53616
neg_binom	~NYear + Bottom.Temp	~.	~.	11582.26	353.22040
zero_alt	~NYear + Salinity	~.	~.	11884.91	655.87345
neg_binom	~NYear + Salinity	~.	~.	11936.70	707.66505
zero_infl	~NYear + Salinity	~.	~.	NA	NA

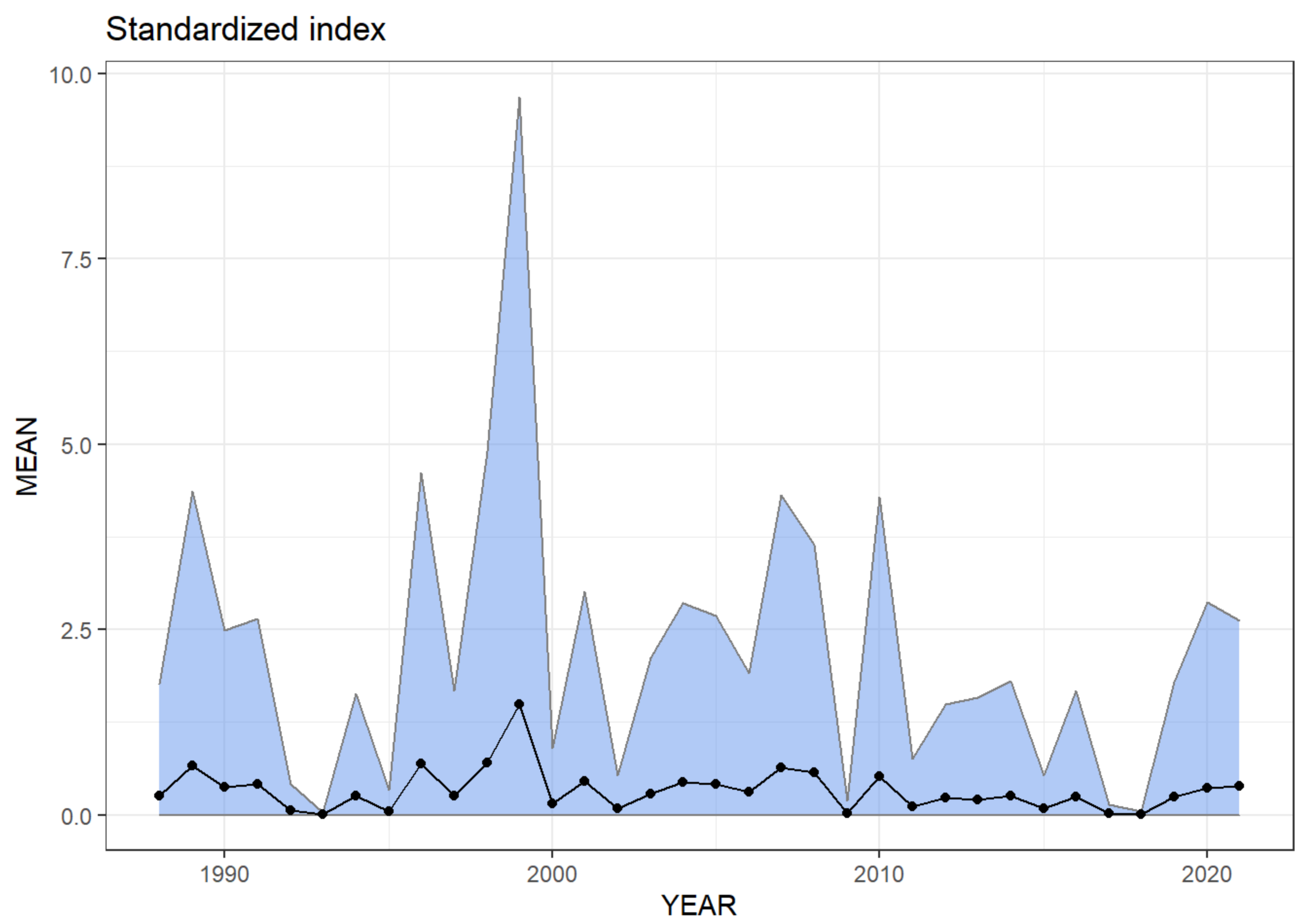
Chosen model is N = f(Year, Temperature, Salinity) hurdle model.

```
selectedModelIndex <- 3
print(mod[[selectedModelIndex]])
## N ~ Year + Bottom.Temp + Salinity
## <environment: 0x0000000028bd4bc8>
```

Bootstrapped results

```
out <- boot.ZI_AT(best = modelList[[selectedModelIndex]][[3]], # a model
                  type_z = "zero", # "zero" if zero-inflated
                  nboots = 1200,
                  CI = 0.95,
                  print.prog = T)
```

```
stdPlot <- plot_boot(data = out) +
  ggtitle('Standardized index')
print(stdPlot)
```




```
options(knitr.kable.NA = "--")

out %>%
  as_tibble() %>%
  complete(YEAR = min(YEAR):max(YEAR)) %>%
  mutate(across(where(is.numeric), round, 3)) %>%
  kbl() %>%
  kable_classic(full_width = FALSE)
```

YEAR	MEAN	SE	LCI	UCI
1988	0.257	0.563	0	1.768
1989	0.662	1.294	0	4.369
1990	0.375	0.705	0	2.494
1991	0.415	0.751	0	2.641
1992	0.055	0.123	0	0.408
1993	0.004	0.009	0	0.030
1994	0.253	0.468	0	1.633
1995	0.051	0.097	0	0.333
1996	0.691	1.366	0	4.620

YEAR	MEAN	SE	LCI	UCI
1997	0.260	0.479	0	1.669
1998	0.702	1.385	0	4.901
1999	1.484	2.679	0	9.687
2000	0.145	0.265	0	0.896
2001	0.448	0.983	0	3.013
2002	0.081	0.153	0	0.535
2003	0.283	0.597	0	2.122
2004	0.446	0.845	0	2.862
2005	0.412	0.808	0	2.683
2006	0.304	0.562	0	1.905
2007	0.636	1.276	0	4.306
2008	0.573	1.160	0	3.642
2009	0.026	0.052	0	0.188
2010	0.519	1.421	0	4.287
2011	0.107	0.215	0	0.758
2012	0.229	0.459	0	1.485
2013	0.203	0.432	0	1.588
2014	0.256	0.518	0	1.799
2015	0.084	0.152	0	0.529
2016	0.240	0.462	0	1.671
2017	0.020	0.036	0	0.135
2018	0.007	0.015	0	0.050
2019	0.249	0.542	0	1.786
2020	0.363	0.842	0	2.868
2021	0.394	0.794	0	2.621

Nominal index

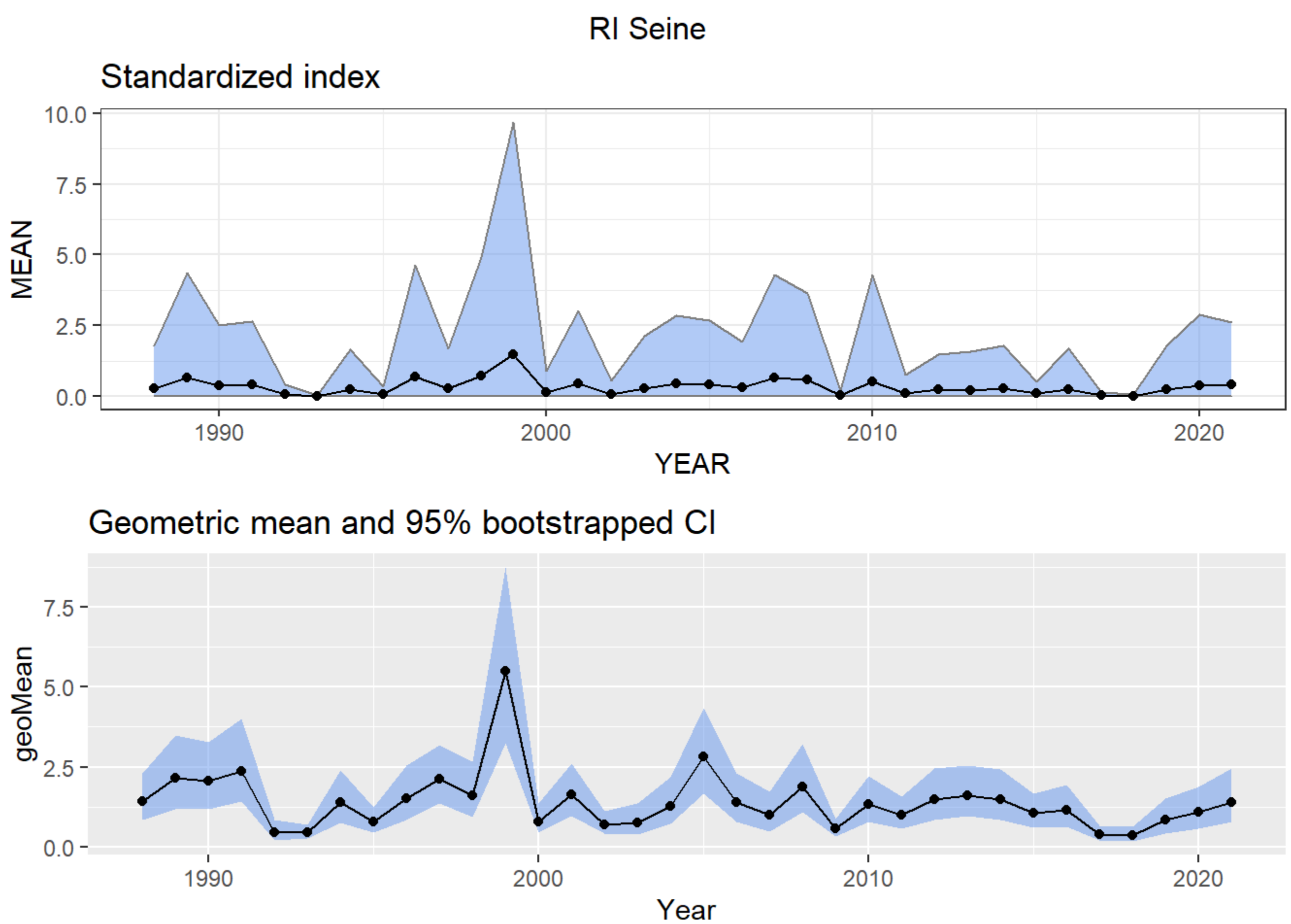
```
bootFun <- function(x, n = 1000){
  bootMean_i <- numeric(n)
  for(i in 1:n){
    nd <- sample(x, size = length(x), replace = TRUE)
    bootMean_i[i] <- exp(mean(log(nd+1)))-1
  }
  bootPct <- quantile(bootMean_i, probs = c(0.025, 0.975))
  bootSD <- sd(bootMean_i)
  return(tibble(bootSD = bootSD, boot05 = bootPct[1], boot95 = bootPct[2]))
}

nominalGeoMean <- riSeine %>%
  group_by(Year) %>%
```

```
  summarize(mean = mean(N),
            sd = sd(N),
            geoMean = exp(mean(log(N + 1))) - 1,
            bootRes = bootFun(x = N),
            .groups = 'drop') %>%
  unpack(bootRes) %>%
  rename(geoMeanBootSD = bootSD) %>%
  mutate(geoCV = geoMeanBootSD / geoMean,
         across(where(is.numeric), round, 3),
         Year = as.numeric(as.character(Year)))

nominalGeoMean %>%
  complete(Year = min(Year):max(Year)) %>%
  kbl() %>%
  kable_classic(full_width = FALSE)
```

Year	mean	sd	geoMean	geoMeanBootSD	boot05	boot95	geoCV
1988	8.831	29.372	1.435	0.382	0.828	2.316	0.266
1989	16.280	45.876	2.147	0.593	1.184	3.497	0.276
1990	19.771	59.049	2.046	0.547	1.165	3.272	0.267
1991	17.012	43.936	2.375	0.655	1.427	3.996	0.276
1992	4.325	23.293	0.460	0.161	0.203	0.831	0.351
1993	1.298	3.963	0.451	0.112	0.257	0.689	0.249
1994	16.047	54.713	1.393	0.425	0.742	2.406	0.305
1995	5.791	31.434	0.772	0.205	0.440	1.225	0.266
1996	21.429	88.120	1.524	0.428	0.838	2.549	0.281
1997	14.222	47.226	2.106	0.481	1.357	3.177	0.228
1998	27.500	140.084	1.600	0.454	0.920	2.655	0.284
1999	61.967	174.461	5.489	1.386	3.235	8.742	0.252
2000	5.725	18.121	0.797	0.239	0.437	1.369	0.300
2001	20.901	93.752	1.645	0.432	0.972	2.600	0.262
2002	3.144	10.430	0.695	0.186	0.403	1.111	0.268
2003	15.767	80.607	0.759	0.263	0.377	1.365	0.347
2004	19.600	72.727	1.272	0.372	0.728	2.190	0.292
2005	24.000	72.980	2.820	0.682	1.663	4.346	0.242
2006	13.467	40.202	1.391	0.405	0.774	2.311	0.291
2007	26.333	125.987	0.989	0.324	0.482	1.727	0.328
2008	33.633	164.396	1.878	0.549	1.081	3.214	0.292
2009	2.233	8.430	0.569	0.145	0.329	0.877	0.255
2010	22.769	137.971	1.325	0.378	0.795	2.221	0.285
2011	8.200	35.664	1.005	0.262	0.583	1.579	0.261
2012	25.989	107.134	1.475	0.417	0.845	2.443	0.283
2013	9.868	25.013	1.615	0.406	0.959	2.532	0.251
2014	13.900	44.139	1.470	0.424	0.833	2.436	0.288

Year	mean	sd	geoMean	geoMeanBootSD	boot05	boot95	geoCV
2015	7.456	34.081	1.044	0.279	0.591	1.675	0.268
2016	15.889	74.980	1.159	0.336	0.638	1.924	0.290
2017	1.833	7.001	0.393	0.124	0.196	0.673	0.315
2018	1.349	4.315	0.362	0.121	0.163	0.620	0.334
2019	11.022	44.187	0.856	0.277	0.415	1.497	0.323
2020	32.200	214.250	1.099	0.333	0.557	1.874	0.303
2021	27.944	130.450	1.393	0.425	0.783	2.460	0.305

```
write_csv(nominalGeoMean, file = 'RISeine/nominalGeoMean_RI.csv')
nomPlot <- nominalGeoMean %>%
  ggplot(aes(x = Year)) +
    geom_ribbon(aes(ymin = boot05, ymax = boot95, group = 1),
                fill = 'cornflowerblue', alpha = 0.5) +
  geom_line(aes(y = geoMean), group = 1) +
  geom_point(aes(y = geoMean)) +
  ggtitle('Geometric mean and 95% bootstrapped CI')

grid.arrange(stdPlot, nomPlot, top = 'RI Seine')
```



Appendix: NB/Zero inflated/Hurdle models with selected terms

The following shows model diagnostics associated with the final selected terms. Note that all three error distributions are included below, but the **zero altered** model was selected.

Year_Bottom.Temp_Salinity

Variance inflation factor

```
if(!any(modElem == '|') & length(all.vars(fullMod)) > 2){
  lm(fullMod, data = stdData) %>%
    car::vif(.)
}else{
  cat('Random effects within formula or model contains fewer than two
terms:',
      'VIF not computed', sep = ' ')
}
##                 GVIF Df GVIF^(1/(2*Df))
## Year        1.563268 33        1.006792
## Bottom.Temp 1.050964  1        1.025165
## Salinity    1.490664  1        1.220928
```

Model fits

```
coefTab <- function(x){
  if(class(x)[1] != 'try-error'){
    summary(x)$coefficients$cond %>%
    as_tibble(rownames = 'Parameter') %>%
    mutate(across(where(is.numeric), round, 3)) %>%
    kbl() %>%
    kable_classic(full_width = FALSE)
  }
}

# Negative binomial fit
neg_binom = glmmTMB(fullMod,
                    data = stdData,
                    family = nbinom2)
coefTab(neg_binom)
```

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.215	0.407	5.437	0.000
Year1989	1.425	0.631	2.259	0.024
Year1990	-0.040	0.540	-0.074	0.941
Year1991	-0.137	0.526	-0.260	0.795

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
Year1992	-2.060	0.552	-3.734	0.000
Year1993	-4.010	0.582	-6.888	0.000
Year1994	-0.369	0.523	-0.706	0.480
Year1995	-1.988	0.524	-3.795	0.000
Year1996	0.536	0.534	1.004	0.316
Year1997	-0.097	0.521	-0.187	0.852
Year1998	0.046	0.519	0.089	0.929
Year1999	1.411	0.514	2.743	0.006
Year2000	-0.996	0.521	-1.910	0.056
Year2001	-0.155	0.513	-0.302	0.763
Year2002	-1.050	0.525	-2.001	0.045
Year2003	-0.662	0.523	-1.266	0.205
Year2004	-0.141	0.519	-0.272	0.786
Year2005	-0.300	0.520	-0.578	0.563
Year2006	-0.627	0.522	-1.201	0.230
Year2007	-0.221	0.515	-0.429	0.668
Year2008	-0.062	0.519	-0.119	0.906
Year2009	-2.217	0.549	-4.038	0.000
Year2010	0.995	0.519	1.917	0.055
Year2011	-1.380	0.545	-2.531	0.011
Year2012	-0.944	0.538	-1.754	0.079
Year2013	0.518	0.527	0.981	0.327
Year2014	-0.747	0.570	-1.310	0.190
Year2015	-1.138	0.536	-2.123	0.034
Year2016	-0.177	0.529	-0.334	0.738
Year2017	-2.311	0.558	-4.142	0.000
Year2018	-3.489	0.591	-5.907	0.000
Year2019	-0.143	0.527	-0.272	0.786
Year2020	-0.575	0.536	-1.071	0.284
Year2021	-0.257	0.568	-0.453	0.651
Bottom.Temp	2.120	0.101	20.938	0.000
Salinity	-0.100	0.091	-1.100	0.271

```
# Zero inflated fit
zero_infl = glmmTMB(fullMod,
                    ziformula = zim,
                    data = stdData,
                    family = nbinom2) %>%
  try()
coefTab(zero_infl)
```

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	2.507	0.396	6.322	0.000
Year1989	0.548	0.564	0.972	0.331
Year1990	0.399	0.534	0.748	0.455
Year1991	0.318	0.514	0.618	0.537
Year1992	-1.057	0.569	-1.856	0.063
Year1993	-2.660	0.557	-4.777	0.000
Year1994	0.029	0.502	0.058	0.954
Year1995	-1.066	0.529	-2.014	0.044
Year1996	0.619	0.555	1.115	0.265
Year1997	0.012	0.507	0.023	0.981
Year1998	0.878	0.524	1.675	0.094
Year1999	1.313	0.481	2.731	0.006
Year2000	-0.284	0.557	-0.511	0.610
Year2001	0.692	0.525	1.316	0.188
Year2002	-0.935	0.522	-1.792	0.073
Year2003	0.613	0.583	1.052	0.293
Year2004	0.703	0.544	1.293	0.196
Year2005	0.418	0.499	0.838	0.402
Year2006	0.426	0.567	0.752	0.452
Year2007	1.356	0.613	2.211	0.027
Year2008	0.938	0.530	1.769	0.077
Year2009	-1.838	0.546	-3.368	0.001
Year2010	0.974	0.517	1.883	0.060
Year2011	-0.572	0.539	-1.062	0.288
Year2012	0.467	0.554	0.843	0.399
Year2013	-0.036	0.543	-0.067	0.947
Year2014	0.115	0.557	0.207	0.836
Year2015	-0.905	0.506	-1.787	0.074
Year2016	-0.051	0.510	-0.100	0.921
Year2017	-1.870	0.563	-3.323	0.001
Year2018	-2.233	0.608	-3.672	0.000
Year2019	0.259	0.567	0.457	0.648
Year2020	0.742	0.549	1.352	0.176
Year2021	0.674	0.577	1.168	0.243
Bottom.Temp	0.728	0.114	6.359	0.000
Salinity	-0.084	0.089	-0.943	0.346

```
# Zero altered fit
zero_alt <- glmmTMB(fullMod,
```

```
                    ziformula = zam,
                    data = stdData,
                    family = truncated_nbinom2(link = "log")) %>%
  try()
coefTab(zero_alt)
```

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-14.805	1879.965	-0.008	0.994
Year1989	0.821	0.738	1.113	0.266
Year1990	0.497	0.665	0.748	0.455
Year1991	0.308	0.638	0.483	0.629
Year1992	-0.788	0.756	-1.043	0.297
Year1993	-3.109	0.676	-4.598	0.000
Year1994	0.459	0.689	0.666	0.505
Year1995	-1.175	0.655	-1.795	0.073
Year1996	1.083	0.727	1.491	0.136
Year1997	-0.007	0.629	-0.012	0.991
Year1998	1.093	0.678	1.613	0.107
Year1999	1.436	0.609	2.356	0.018
Year2000	-0.251	0.689	-0.364	0.716
Year2001	0.746	0.656	1.138	0.255
Year2002	-0.811	0.672	-1.206	0.228
Year2003	0.949	0.780	1.217	0.224
Year2004	0.897	0.698	1.285	0.199
Year2005	0.465	0.625	0.744	0.457
Year2006	0.509	0.707	0.720	0.471
Year2007	1.617	0.805	2.007	0.045
Year2008	1.142	0.681	1.677	0.094
Year2009	-1.889	0.685	-2.760	0.006
Year2010	1.338	0.681	1.965	0.049
Year2011	-0.347	0.687	-0.506	0.613
Year2012	0.767	0.729	1.052	0.293
Year2013	0.047	0.674	0.070	0.945
Year2014	0.323	0.721	0.448	0.654
Year2015	-0.658	0.665	-0.989	0.323
Year2016	0.408	0.685	0.596	0.551
Year2017	-1.493	0.750	-1.991	0.046
Year2018	-2.021	0.800	-2.527	0.012
Year2019	0.761	0.772	0.986	0.324
Year2020	1.255	0.752	1.670	0.095
Year2021	1.125	0.769	1.461	0.144

Parameter	Estimate	Std. Error	z value	Pr(>|z|)
Bottom.Temp	0.767	0.154	4.979	0.000
Salinity	-0.031	0.117	-0.261	0.794

```
aic <- tibble(
  Model = c('neg_binom', 'zero_infl', 'zero_alt'),
  fullMod = paste(as.character(fullMod), collapse = ''),
  zim = paste(as.character(zim), collapse = ''),
  zam = paste(as.character(zam), collapse = ''),
  AIC = c(AIC(neg_binom), AIC(zero_infl), AIC(zero_alt))
)

modSave <- list(neg_binom, zero_infl, zero_alt)

neg_binomPred <- ggemmeans(model = neg_binom, terms = 'Year', type = 'fe')
%>%
  try()
zero_inflPred <- ggemmeans(model = zero_infl, terms = 'Year', type = 'fe.zi')
%>%
  try()
# Help says fe for conditional model
zero_altPred <- ggemmeans(model = zero_alt, terms = 'Year', type = 'fe.zi')
%>%
  try()

makeNAPred <- function(x){
  if('try-error' %in% class(x)){
    out <- list(x = NA, predicted = NA, std.error = NA,
                conf.low = NA, conf.high = NA)
  }else{
    out <- x
  }
}

neg_binomPred <- makeNAPred(neg_binomPred)
zero_inflPred <- makeNAPred(zero_inflPred)
zero_altPred <- makeNAPred(zero_altPred)

neg_binomPredDF <- tibble(
  Model = 'neg_binom',
  paste(as.character(fullMod), collapse = ''),
  zim = paste(as.character(zim), collapse = ''),
  zam = paste(as.character(zam), collapse = ''),
  Year = neg_binomPred$x,
  Idx = neg_binomPred$predicted,
  se = neg_binomPred$std.error,
  conf.low = neg_binomPred$conf.low,
  conf.high = neg_binomPred$conf.high
)

zero_inflPredDF <- tibble(
  Model = 'zero_infl',
  paste(as.character(fullMod), collapse = ''),
  zim = paste(as.character(zim), collapse = ''),
  zam = paste(as.character(zam), collapse = ''),
  Year = zero_inflPred$x,
```

```
  Idx = zero_inflPred$predicted,
  se = zero_inflPred$std.error,
  conf.low = zero_inflPred$conf.low,
  conf.high = zero_inflPred$conf.high
)

zero_altPredDF <- tibble(
  Model = 'zero_alt',
  ModelTerms = paste(as.character(fullMod), collapse = ''),
  zim = paste(as.character(zim), collapse = ''),
  zam = paste(as.character(zam), collapse = ''),
  Year = zero_altPred$x,
  Idx = zero_altPred$predicted,
  se = ifelse(!is.null(zero_altPred$std.error), zero_altPred$std.error, NA),
  conf.low = ifelse(!is.null(zero_altPred$conf.low),
                    zero_altPred$conf.low, NA),
  conf.high = ifelse(!is.null(zero_altPred$conf.high),
                     zero_altPred$conf.high, NA)
)

preds <- bind_rows(neg_binomPredDF, zero_inflPredDF, zero_altPredDF)
```

Model diagnostics

Get the DHARMa residuals

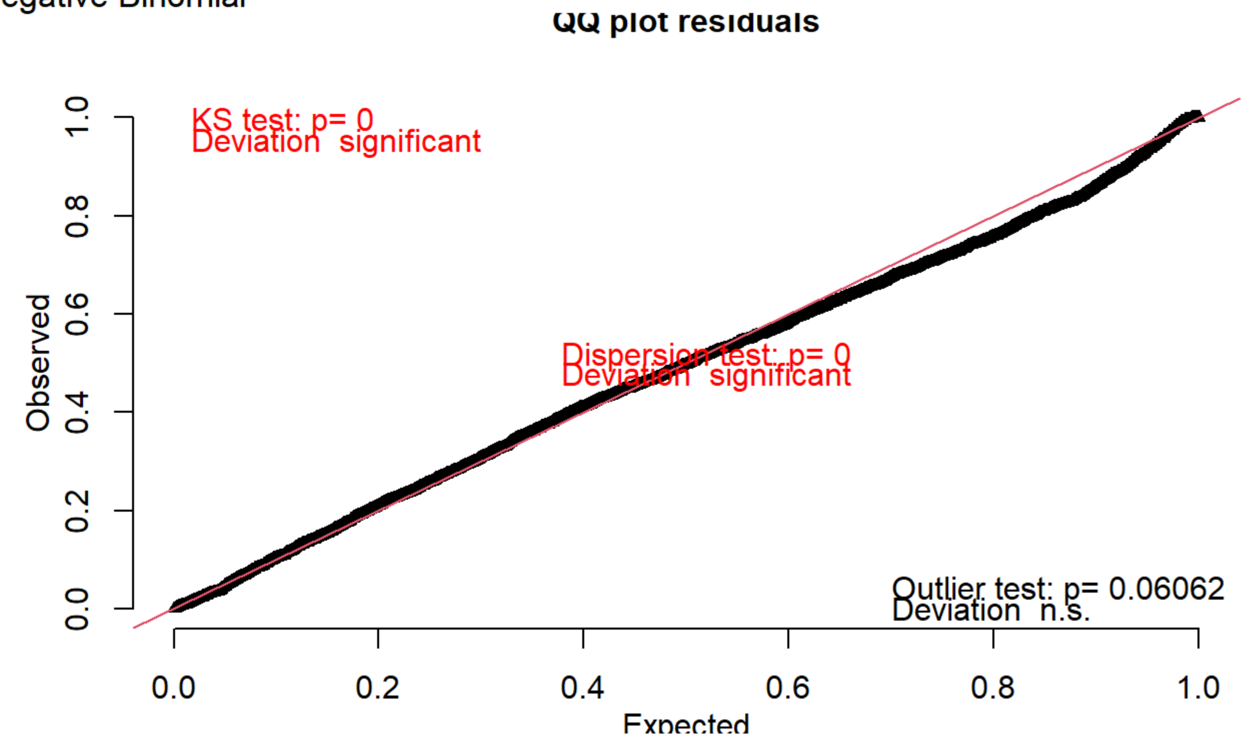
```
soNeg_binom <- simulateResiduals(neg_binom) %>%
  try()
soZero_infl <- simulateResiduals(zero_infl) %>%
  try()
soZero_alt <- simulateResiduals(zero_alt) %>%
  try()

if(class(soNeg_binom)[1] == 'try-error'){
  soNeg_binom <- NULL
}
if(class(soZero_infl)[1] == 'try-error'){
  soZero_infl <- NULL
}
if(class(soZero_alt)[1] == 'try-error'){
  soZero_alt <- NULL
}

# Be sure to maintain order bwt these two lists
soLst <- list(`Negative Binomial` = soNeg_binom,
              `Zero inflated` = soZero_infl,
              `Zero altered` = soZero_alt)
modLst <- list(`Negative Binomial` = neg_binom,
              `Zero inflated` = zero_infl,
              `Zero altered` = zero_alt)
anyNull <- which(sapply(soLst, is.null))
if(length(anyNull) > 0){
  soLst <- soLst[-anyNull]
  modLst <- modLst[-anyNull]
```

```
}

if(length(modLst) < 1){
  cat('---- No models converged from this run ----')
}
```

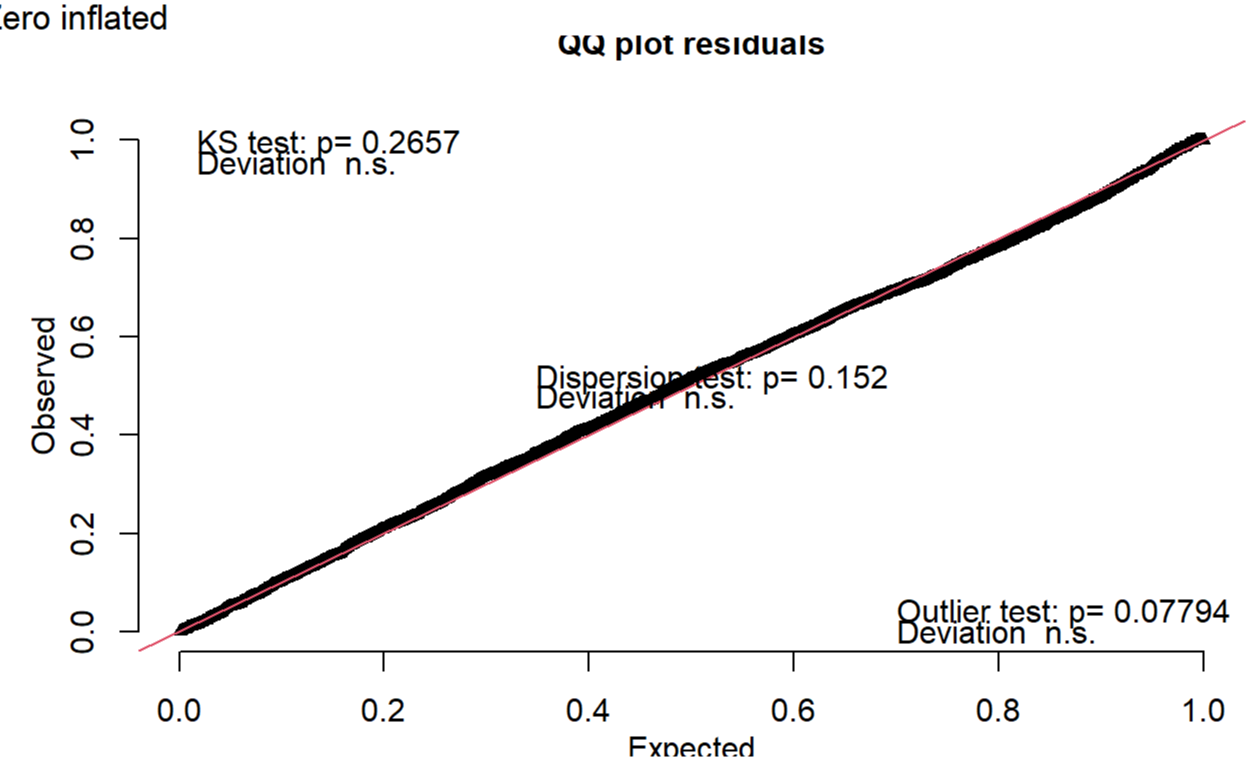
QQ plots

```
qq <- list()
for(i in 1:length(soLst)){
  qq[[i]] <- as.ggplot(~plotQQunif(soLst[[i]])) +
    ggtitle(names(soLst)[i])
}
grid.arrange(grobs = qq, ncol = 1)
```

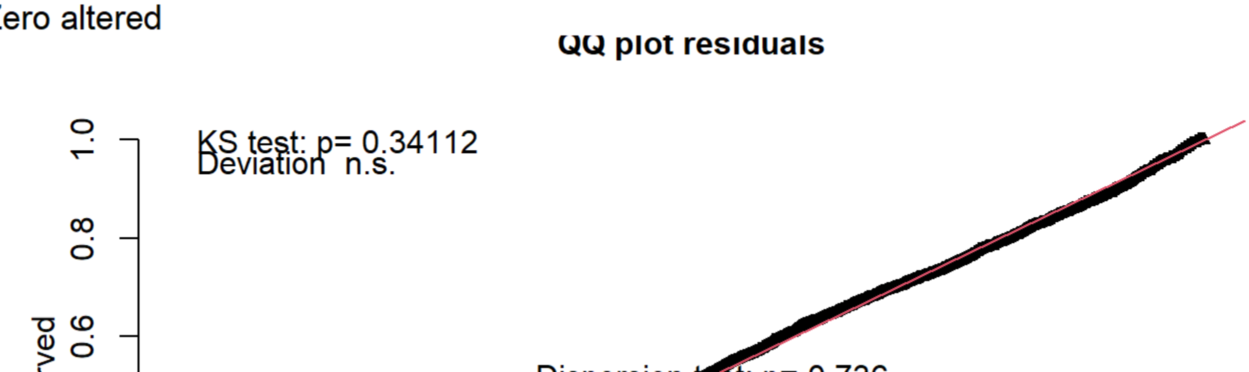
Negative Binomial



Zero inflated



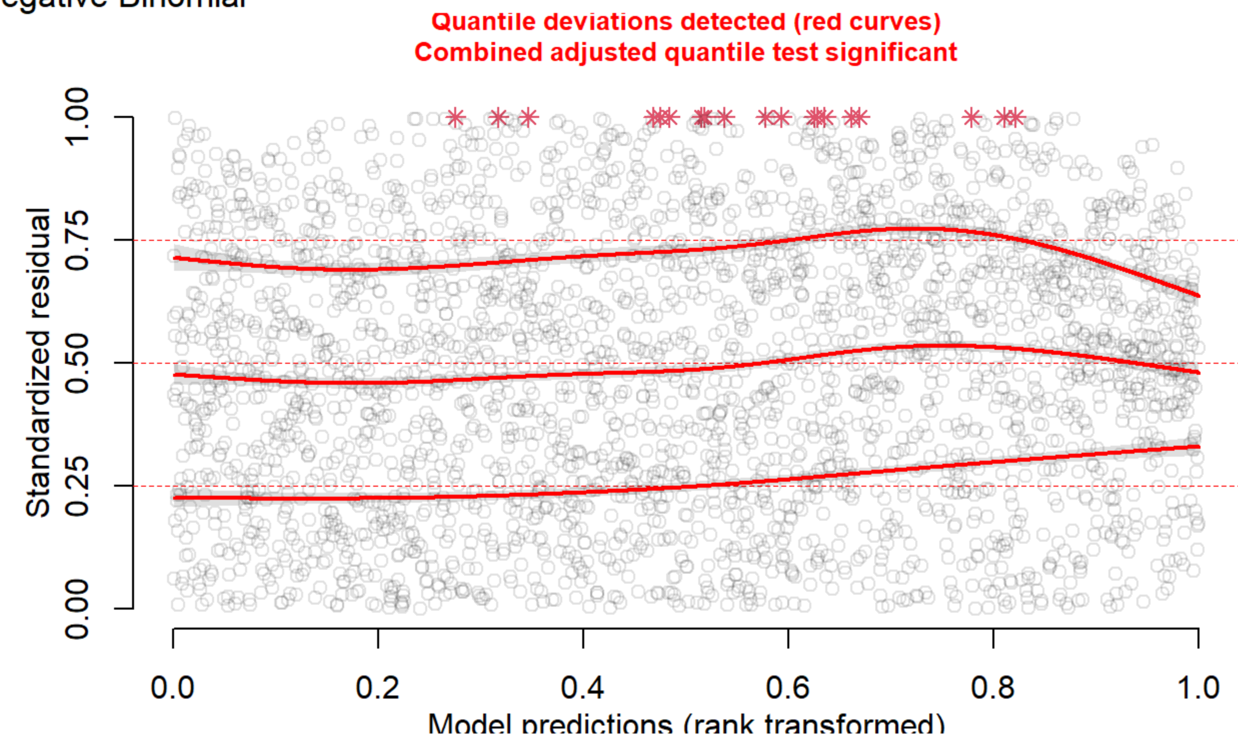
Zero altered



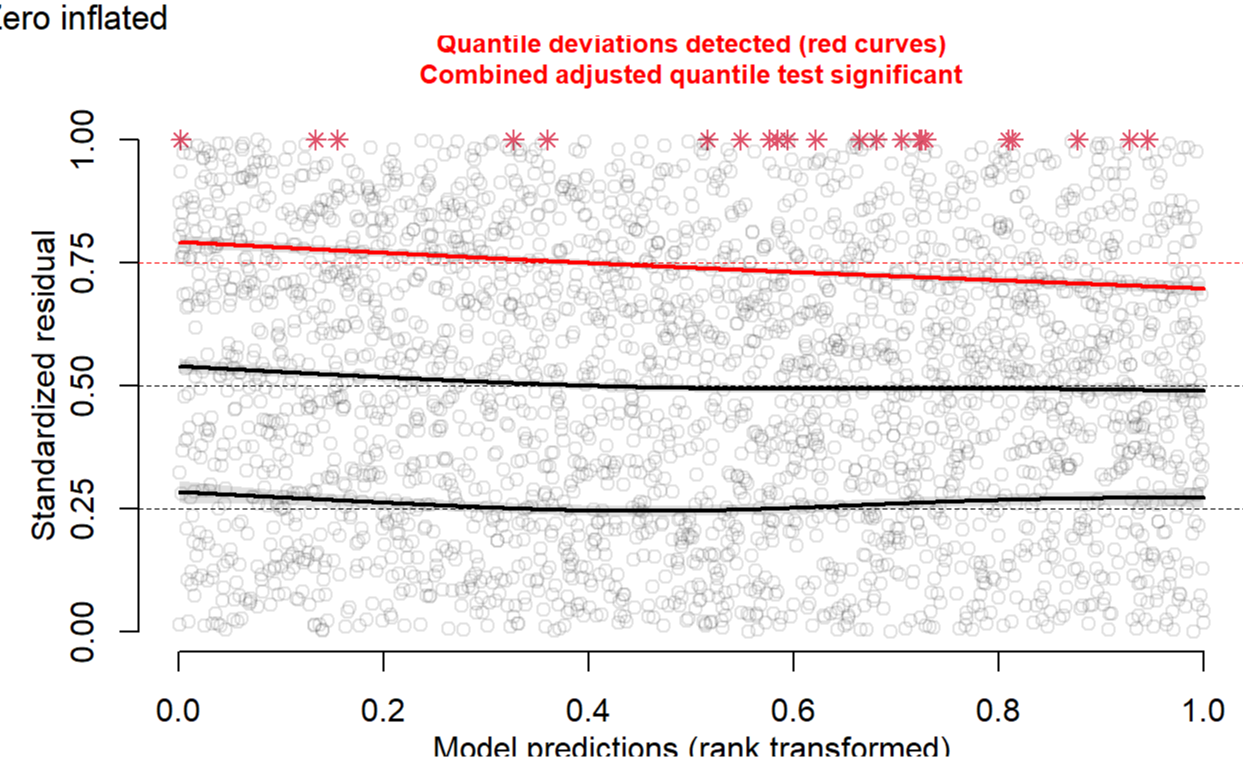
Residuals versus predicted values

```
residVsPred <- list()
for(i in 1:length(soLst)){
  residVsPred[[i]] <- as.ggplot(~plotResiduals(soLst[[i]], quantreg = TRUE))
+
    ggtitle(names(soLst)[i])
}
grid.arrange(grobs = residVsPred, ncol = 1)
```

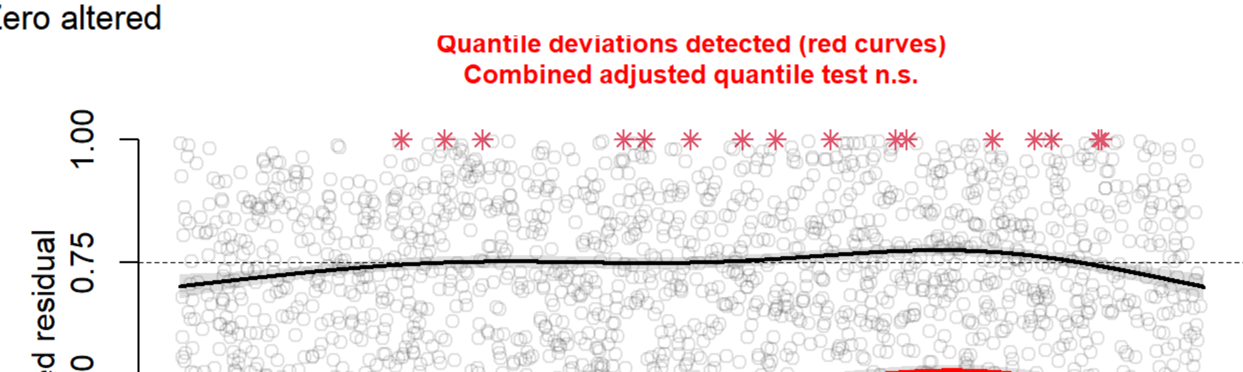
Negative Binomial



Zero inflated

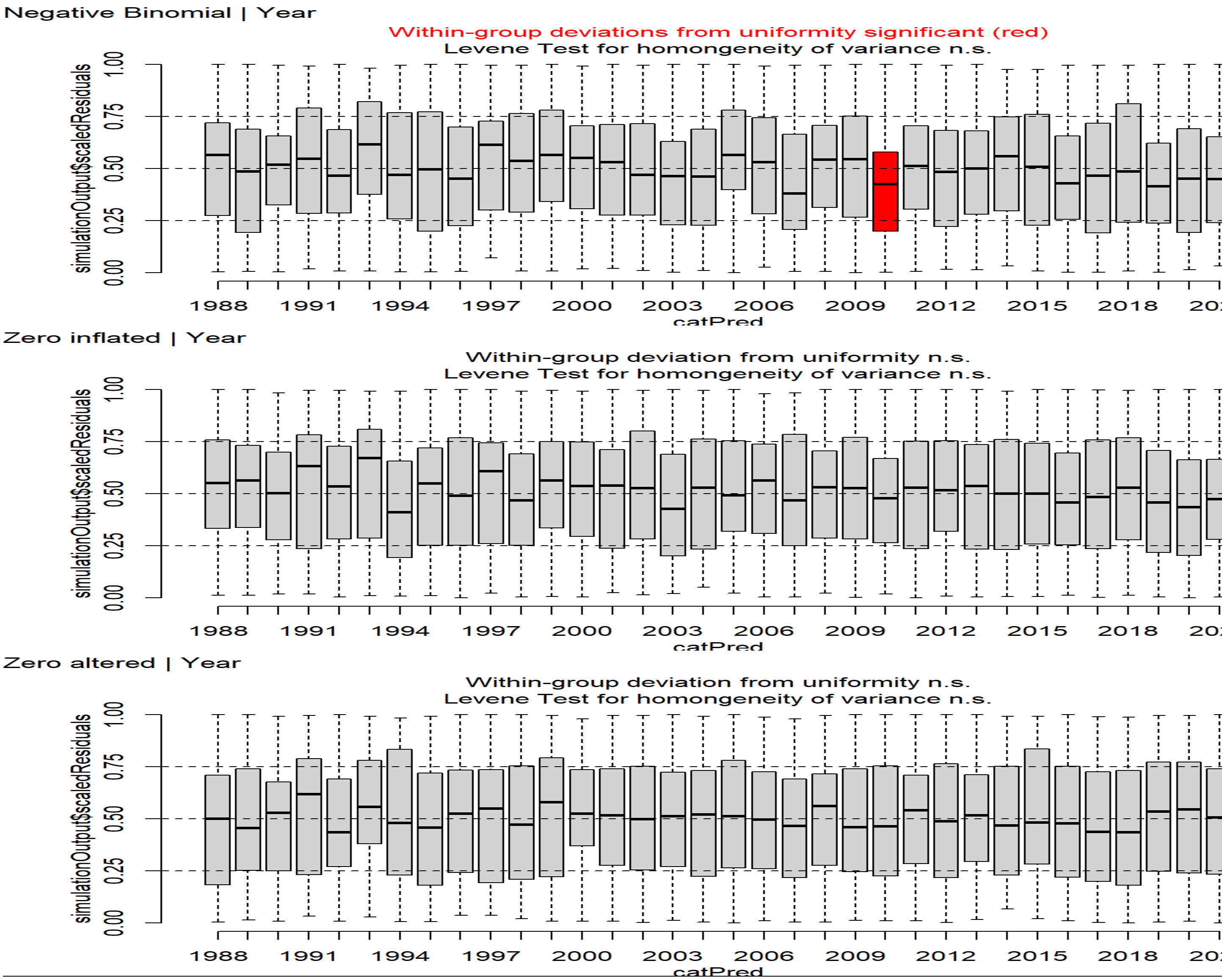


Zero altered



Residuals versus predictors

```
residVsVar <- list()
for(i in 1:length(soLst)){

  residVsVarM <- list()
  frameVar <- names(modLst[[i]]$frame)[-which(names(modLst[[i]]$frame) ==
'N')]
  for(j in 1:length(frameVar)){
    fv <- frameVar[j]
    predictor <- modLst[[i]]$frame[[fv]]
    tplot <- try(as.ggplot(~plotResiduals(soLst[[i]], form = predictor)))
    if(class(tplot)[1] != 'try-error'){
      residVsVarM[[j]] <- tplot
    }else{
      residVsVarM[[j]] <- ggplot() +
        annotate("text", x = 0, y = 0, size=8, label = 'Unsuccessful')
    }
    residVsVarM[[j]] <- residVsVarM[[j]] +
      ggtitle(paste(names(soLst)[i], '|', fv))
  }
  residVsVar[[i]] <- arrangeGrob(grobs = residVsVarM, nrow = 1)

}
grid.arrange(grobs = residVsVar, ncol = 1)
```

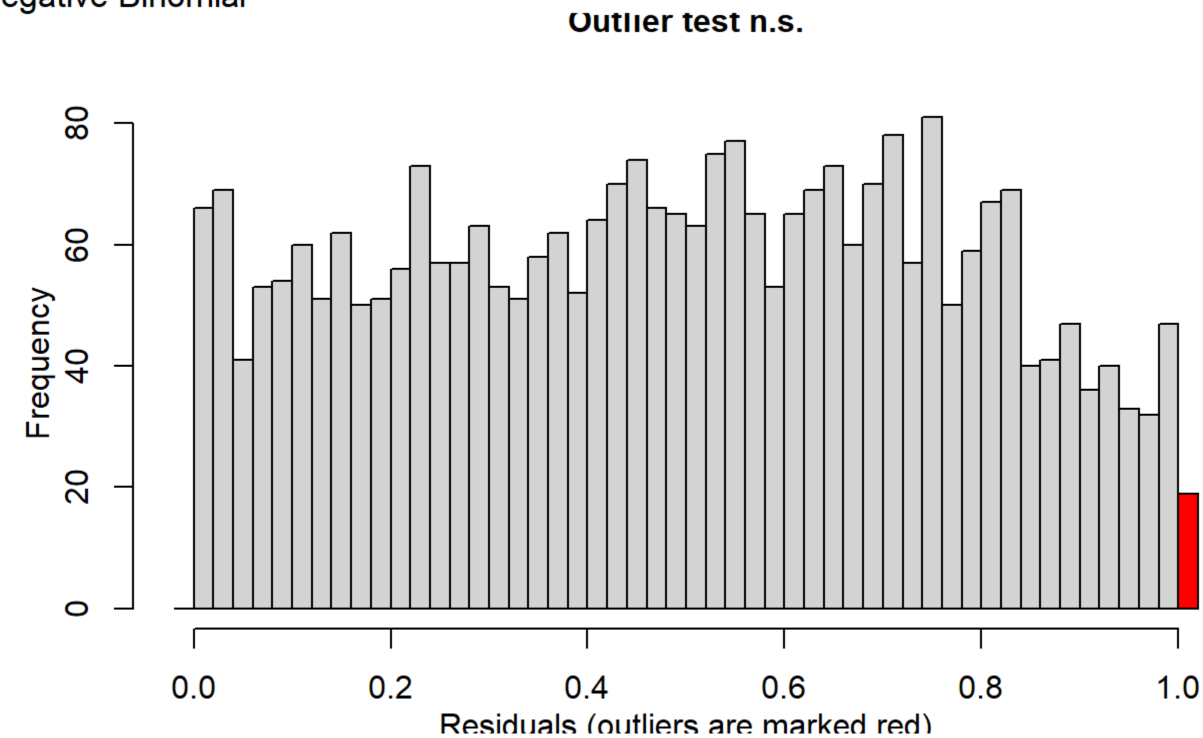



```
# unif <- list()
# for(i in 1:length(soLst)){
#   unif[[i]] <- as.ggplot(~testUniformity(soLst[[i]])) +
#     ggtitle(names(soLst)[i])
# }
# grid.arrange(grobs = unif, ncol = 1)
```

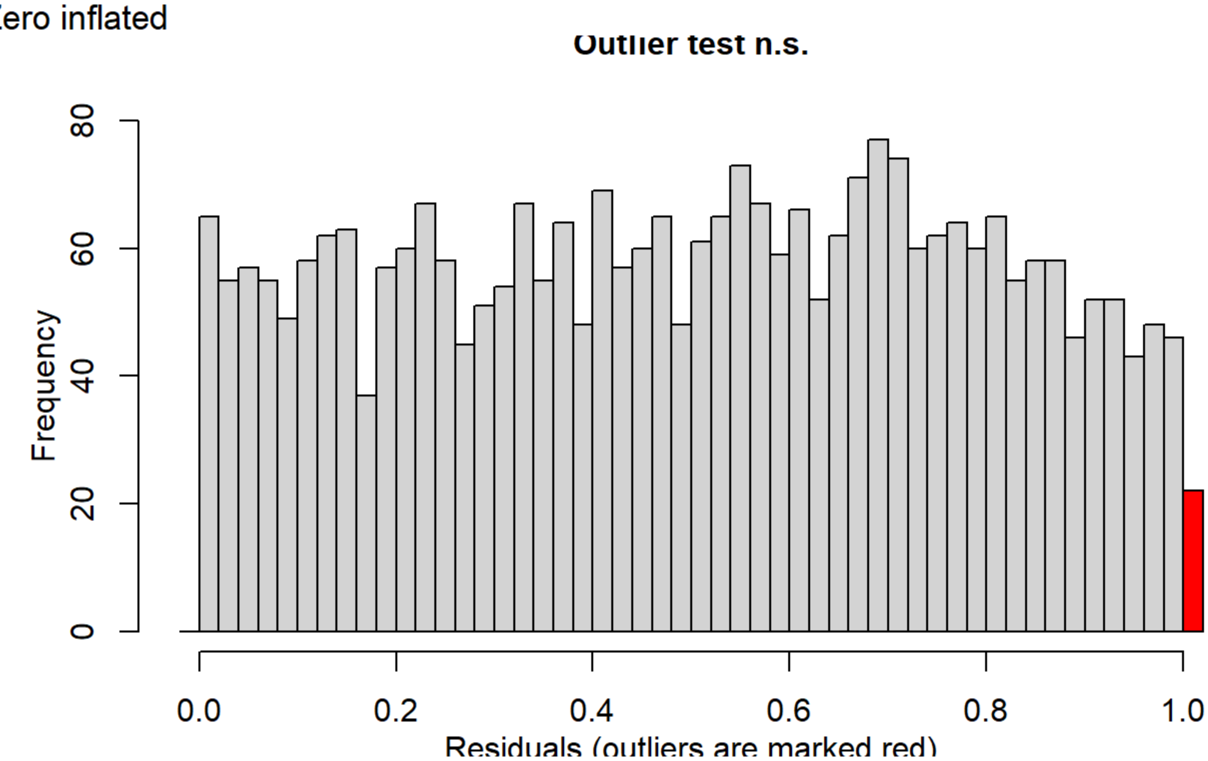
testOutliers() - tests if there are more simulation outliers than expected

```
out <- list()
for(i in 1:length(soLst)){
  out[[i]] <- as.ggplot(~testOutliers(soLst[[i]])) +
    ggtitle(names(soLst)[i])
}
grid.arrange(grobs = out, ncol = 1)
```

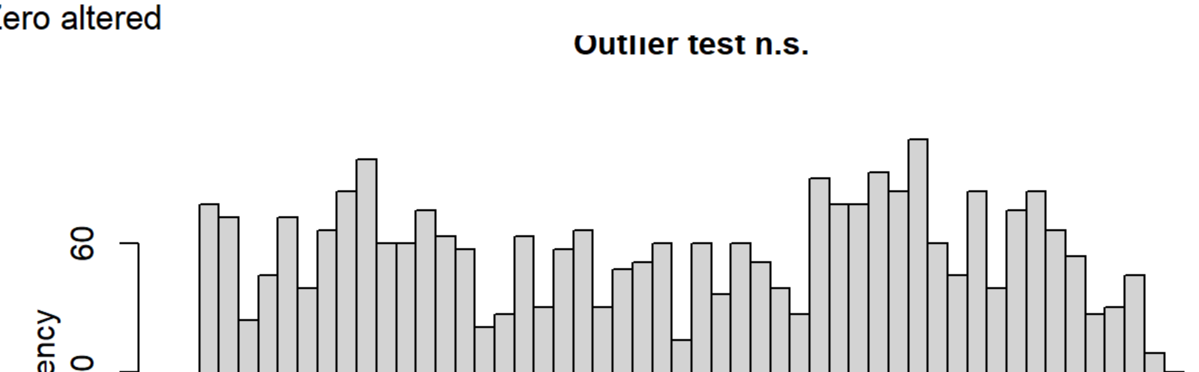
Negative Binomial



Zero inflated



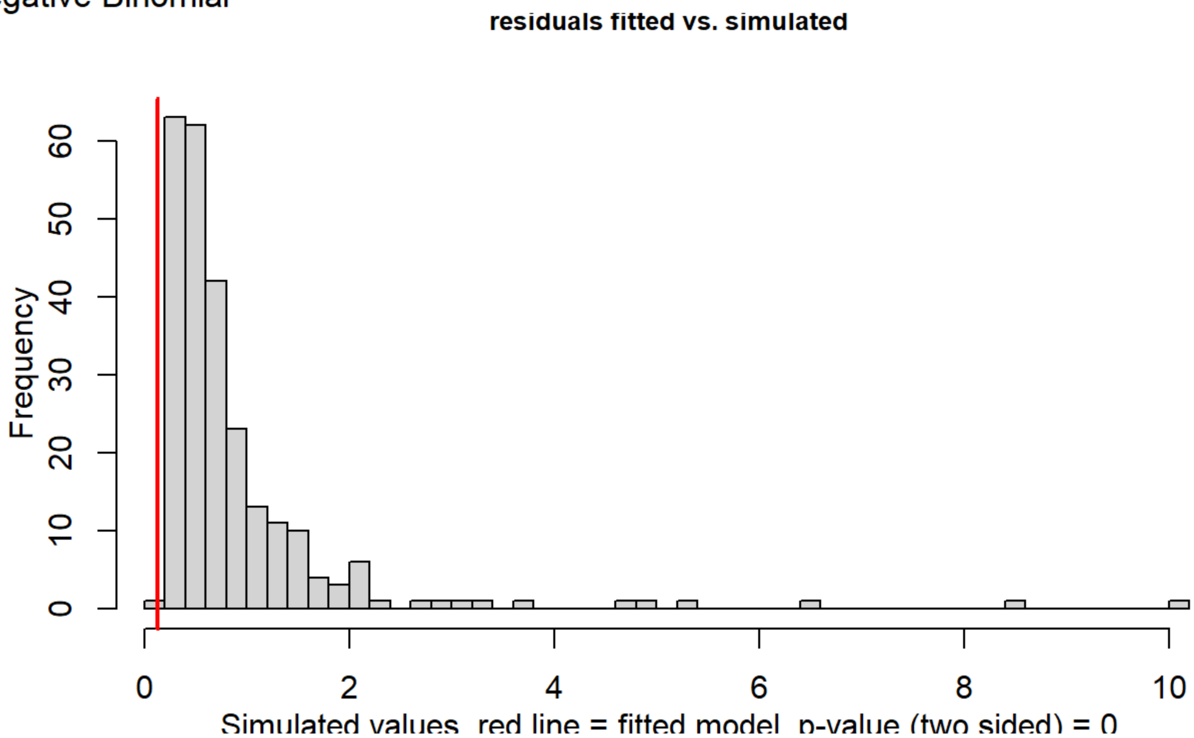
Zero altered



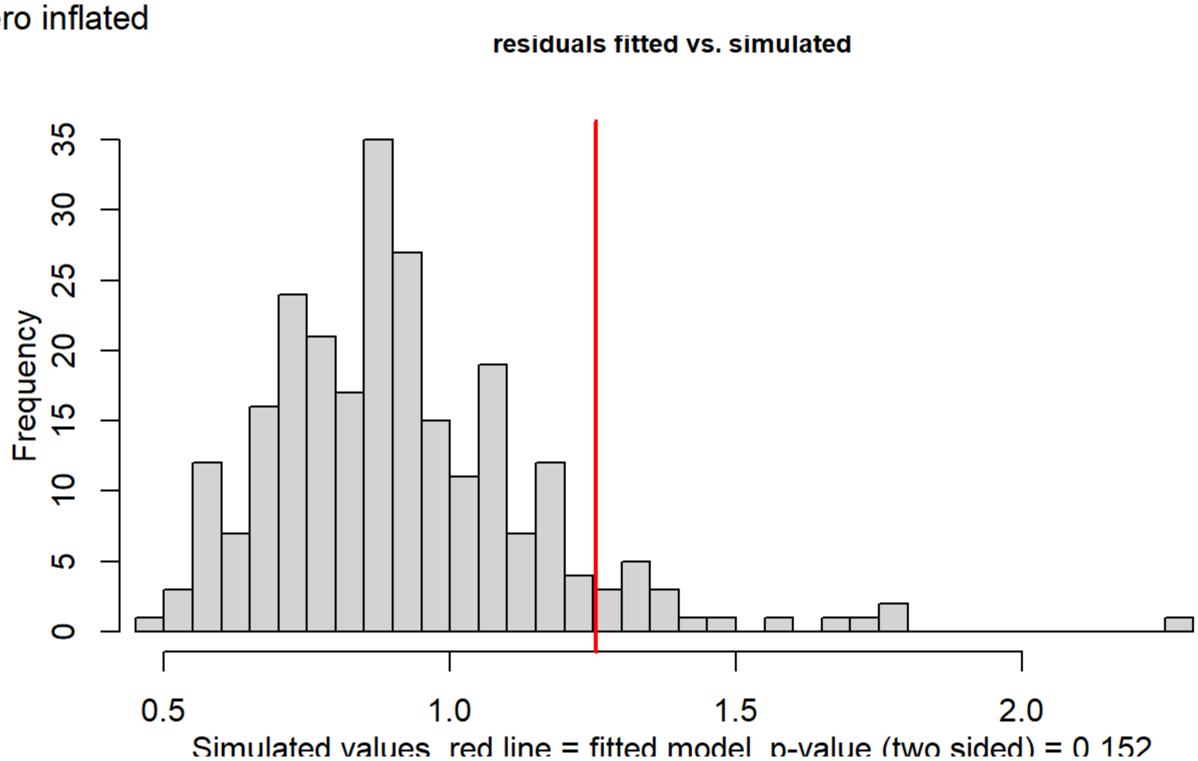
testDispersion() - tests if the simulated dispersion is equal to the observed dispersion

```
disp <- list()
for(i in 1:length(soLst)){
  disp[[i]] <- as.ggplot(~testDispersion(soLst[[i]])) +
    ggtitle(names(soLst)[i])
}
grid.arrange(grobs = disp, ncol = 1)
```

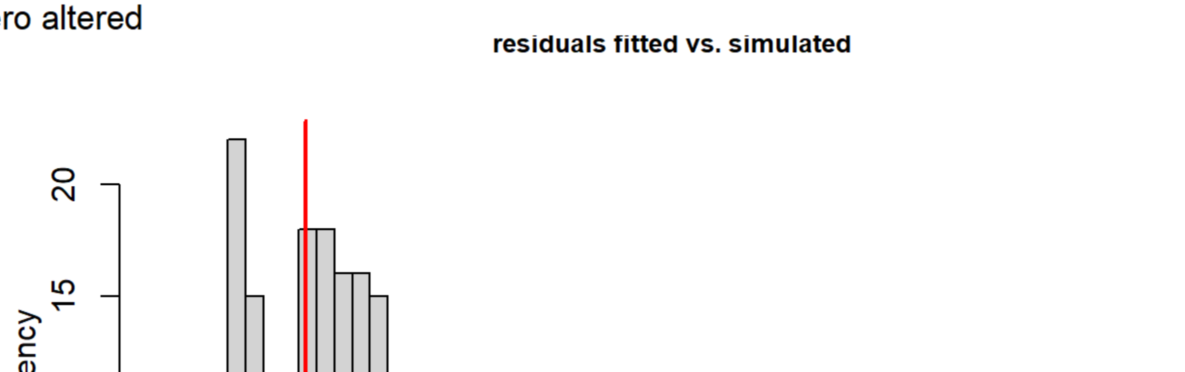
Negative Binomial



Zero inflated



Zero altered

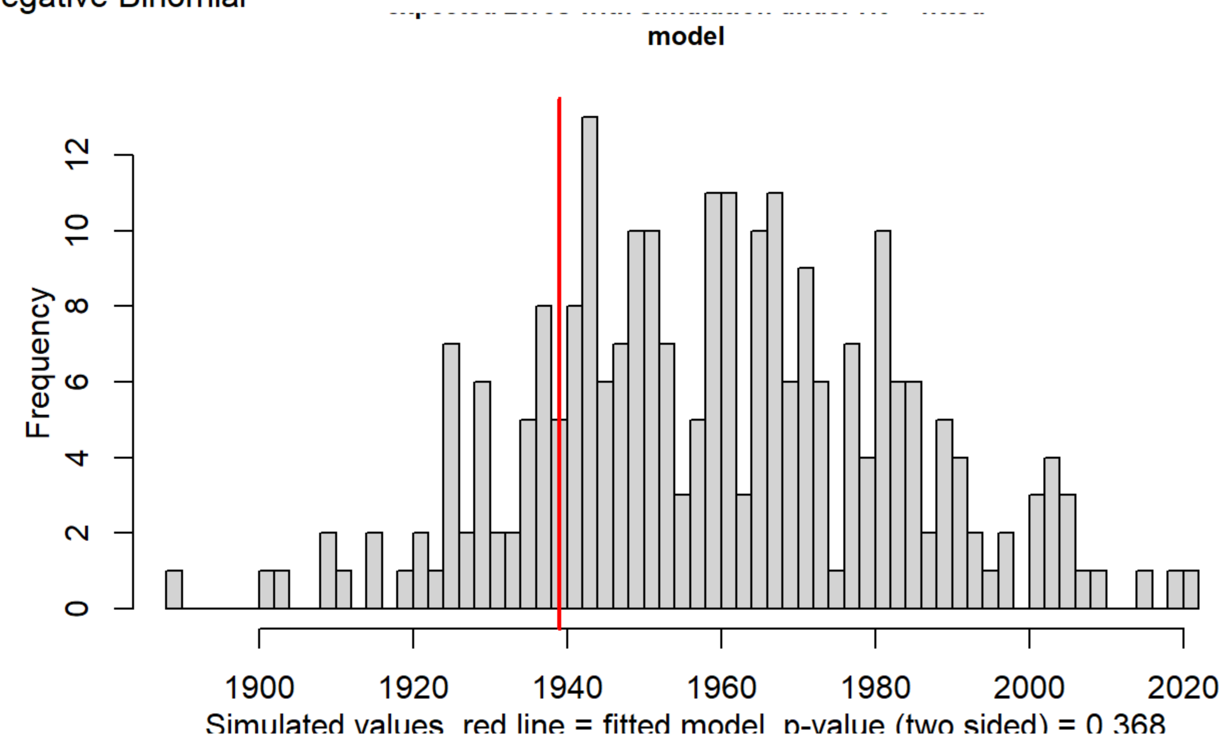


```
# quant <- list()
# for(i in 1:length(soLst)){
#   quant[[i]] <- as.ggplot(~testQuantiles(soLst[[i]])) +
#     ggtitle(names(soLst)[i])
# }
# grid.arrange(grobs = quant, ncol = 1)
```

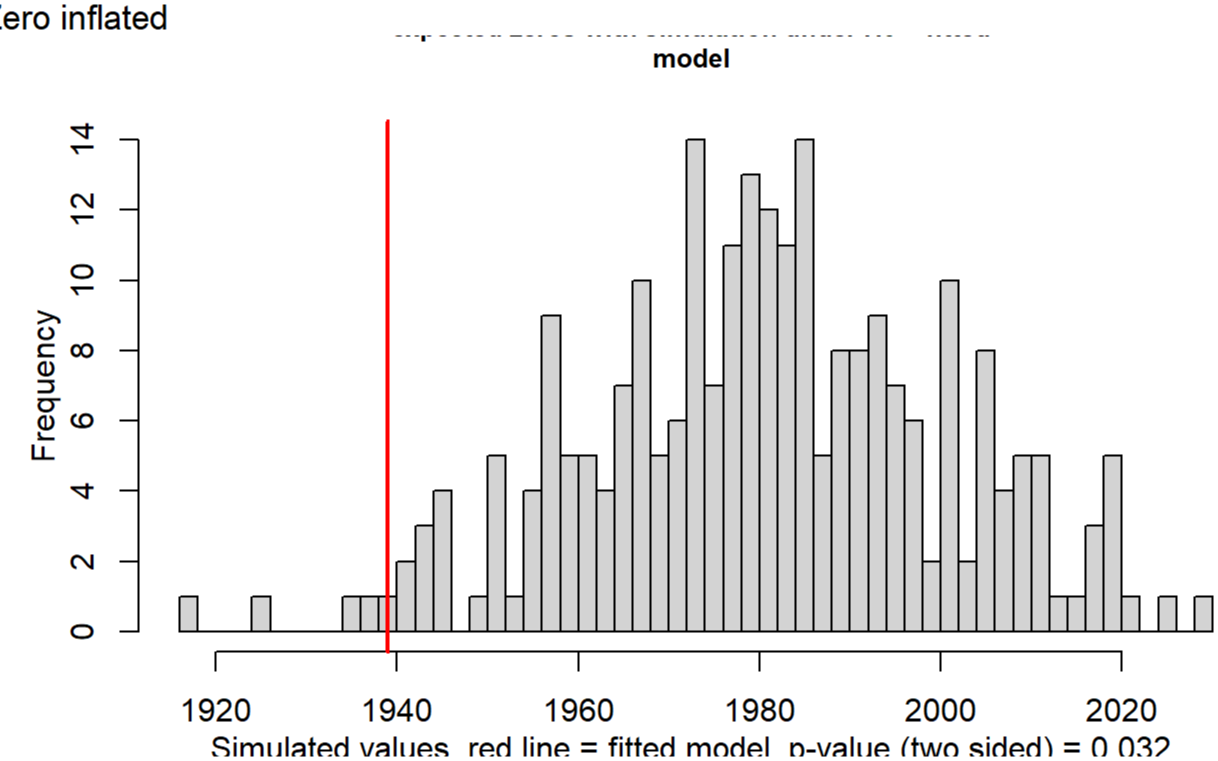
testZeroinflation() - tests if there are more zeros in the data than expected from the simulations

```
zi <- list()
for(i in 1:length(soLst)){
  zi[[i]] <- as.ggplot(~testZeroInflation(soLst[[i]])) +
    ggtitle(names(soLst)[i])
}
grid.arrange(grobs = zi, ncol = 1)
```

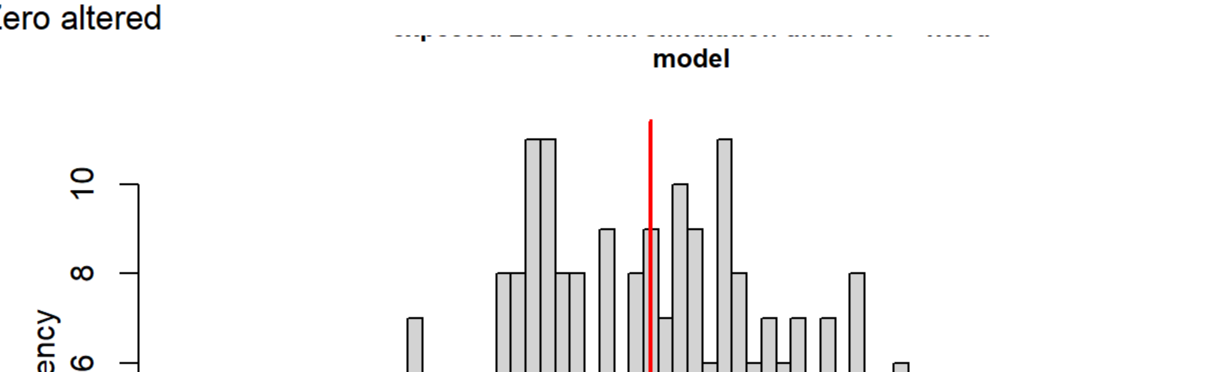
Negative Binomial



Zero inflated



Zero altered



Bluefish index standardization - NY WLIS (all yrs)

M Celestino

2022-06-13

Bluefish index standardization

Step 1: Data processing

Load functions & libraries

```
source("f:\\BF_WG\\FIG\\figFuns.R")
source("f:\\BF_WG\\FIG\\originalCode\\KDrew2\\bootstrap_functions_AT2.R")


library(car, quietly = TRUE, verbose=FALSE)

## Warning: package 'car' was built under R version 4.1.2

## Warning: package 'carData' was built under R version 4.1.2

library(lmtest, quietly = TRUE, verbose=FALSE)

## Warning: package 'lmtest' was built under R version 4.1.2

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

library(glmmTMB, quietly = TRUE, verbose=FALSE)
library(DHARMa, quietly = TRUE, verbose=FALSE)

## Warning: package 'DHARMa' was built under R version 4.1.2

## This is DHARMa 0.4.5. For overview type '?DHARMa'. For recent changes,
type news(package = 'DHARMa')

library(bbmle, quietly = TRUE, verbose=FALSE)
library(ggeffects, quietly = TRUE, verbose=FALSE)
library(ggplot2, quietly = TRUE, verbose=FALSE)

## Warning: package 'ggplot2' was built under R version 4.1.2

library(readxl, quietly = TRUE, verbose=FALSE)
library(lattice, quietly = TRUE, verbose=FALSE)

## Warning: package 'lattice' was built under R version 4.1.2
```

```
library(emmeans, quietly = TRUE, verbose=FALSE)

## Warning: package 'emmeans' was built under R version 4.1.2

library(rmarkdown, quietly = TRUE, verbose=FALSE)

## Warning: package 'rmarkdown' was built under R version 4.1.2

library(knitr, quietly = TRUE, verbose=FALSE)

## Warning: package 'knitr' was built under R version 4.1.2

library(reshape2, quietly=TRUE, verbose=FALSE)

## Warning: package 'reshape2' was built under R version 4.1.2

library(doBy, quietly=TRUE, verbose=FALSE)

## Warning: package 'doBy' was built under R version 4.1.2

library(dplyr, quietly=TRUE, verbose=FALSE,warn.conflicts=FALSE)
```

Read in data

```
indata <- read_excel("f:\\BF_WG\\FIG\\NY-
WLI\\2021_bluefish_data_NY_April_2022.xlsx", sheet="SurveyData",
range="A305:U4316",na="",col_names=FALSE)

## New names:
## * `` -> ...1
## * `` -> ...2
## * `` -> ...3
## * `` -> ...4
## * `` -> ...5
## * ...

cn <- read_excel("f:\\BF_WG\\FIG\\NY-
WLI\\2021_bluefish_data_NY_April_2022.xlsx", sheet="SurveyData",
range="A1:U2",na="")

colnames(indata) <- colnames(cn); rm(cn)

head(indata)

## # A tibble: 6 x 21
##   SampleID State Program     Gear   Year Month Day                 `Time
(mins)`
##      <dbl> <chr> <chr>       <chr> <dbl> <dbl> <dttm>              <lgl>
## 1 19871144 NY    Western Lo~ Beac~  1987     6 1987-06-15 00:00:00 NA
## 2 19871145 NY    Western Lo~ Beac~  1987     6 1987-06-15 00:00:00 NA
## 3 19871146 NY    Western Lo~ Beac~  1987     6 1987-06-15 00:00:00 NA
## 4 19871147 NY    Western Lo~ Beac~  1987     6 1987-06-18 00:00:00 NA
## 5 19871148 NY    Western Lo~ Beac~  1987     6 1987-06-18 00:00:00 NA
```

```
## 6 19871149 NY    Western Lo~ Beac~  1987     6 1987-06-18 00:00:00 NA
## # ... with 13 more variables: Effort <chr>, N <dbl>, Weight <lgl>, Area
<chr>,
## #   Stratum <lgl>, Depth <lgl>, Wgt.Factor <lgl>, TowValue <lgl>,
## #   Surface.Temp <dbl>, Bottom.Temp <lgl>, Salinity <dbl>, DO <dbl>,
## #   Bottom.Type <lgl>
```

```
str(indata)
```

```
## tibble [4,012 x 21] (S3: tbl_df/tbl/data.frame)
##  $ SampleID    : num [1:4012] 19871144 19871145 19871146 19871147 19871148
...
##  $ State       : chr [1:4012] "NY" "NY" "NY" "NY" ...
##  $ Program     : chr [1:4012] "Western Long Island Seine Survey" "Western
Long Island Seine Survey" "Western Long Island Seine Survey" "Western Long
Island Seine Survey" ...
##  $ Gear        : chr [1:4012] "Beach Seine" "Beach Seine" "Beach Seine"
"Beach Seine" ...
##  $ Year        : num [1:4012] 1987 1987 1987 1987 1987 ...
##  $ Month       : num [1:4012] 6 6 6 6 6 6 6 6 6 6 ...
##  $ Day         : POSIXct[1:4012], format: "1987-06-15" "1987-06-15" ...
##  $ Time (mins) : logi [1:4012] NA NA NA NA NA NA ...
##  $ Effort      : chr [1:4012] "Tow" "Tow" "Tow" "Tow" ...
##  $ N           : num [1:4012] 1 70 10 50 0 1 178 6 0 2 ...
##  $ Weight      : logi [1:4012] NA NA NA NA NA NA ...
##  $ Area        : chr [1:4012] "Jamaica Bay" "Jamaica Bay" "Jamaica Bay"
"Jamaica Bay" ...
##  $ Stratum     : logi [1:4012] NA NA NA NA NA NA ...
##  $ Depth       : logi [1:4012] NA NA NA NA NA NA ...
##  $ Wgt.Factor  : logi [1:4012] NA NA NA NA NA NA ...
##  $ TowValue    : logi [1:4012] NA NA NA NA NA NA ...
##  $ Surface.Temp: num [1:4012] 23 24 24 25 24 25 29 19 19 19 ...
##  $ Bottom.Temp : logi [1:4012] NA NA NA NA NA NA ...
##  $ Salinity    : num [1:4012] 29.6 29.2 29.2 28 28.4 29.2 30 30 NA 30 ...
##  $ DO          : num [1:4012] 7 6.8 7.2 8 7.8 7.6 8 8.6 8 8.2 ...
##  $ Bottom.Type : logi [1:4012] NA NA NA NA NA NA ...
```

```
summary(indata)
```

```
##     SampleID           State              Program              Gear          
##  Min.   :19871018   Length:4012        Length:4012        Length:4012       
##  1st Qu.:20001111   Class :character   Class :character   Class :character  
##  Median :20081066   Mode  :character   Mode  :character   Mode  :character  
##  Mean   :20067739                                                           
##  3rd Qu.:20151067                                                           
##  Max.   :20211210                                                           
##                                                                             
##       Year          Month              Day                         Time (mins)
##  Min.   :1987   Min.   : 6.000   Min.   :1987-06-15 00:00:00   
Mode:logical  
##  1st Qu.:2000   1st Qu.: 7.000   1st Qu.:2000-09-25 00:00:00   NA's:4012
```

```
##  Median :2008   Median : 8.000   Median :2008-07-01 00:00:00
##  Mean   :2007   Mean   : 7.846   Mean   :2007-04-10 05:46:21
##  3rd Qu.:2015   3rd Qu.: 9.000   3rd Qu.:2015-06-18 00:00:00
##  Max.   :2021   Max.   :10.000   Max.   :2021-10-21 00:00:00
##
##     Effort                N            Weight            Area
##  Length:4012        Min.   :   0.00   Mode:logical   Length:4012
##  Class :character   1st Qu.:   0.00   NA's:4012      Class :character
##  Mode  :character   Median :   1.00                  Mode  :character
##                     Mean   :  19.57
##                     3rd Qu.:   7.00
##                     Max.   :6676.00
##
##  Stratum          Depth          Wgt.Factor     TowValue        Surface.Temp
##  Mode:logical   Mode:logical   Mode:logical   Mode:logical   Min.   : 5.90
##  NA's:4012      NA's:4012      NA's:4012      NA's:4012      1st Qu.:19.10
##                                                              Median :22.00
##                                                              Mean   :21.36
##                                                              3rd Qu.:24.00
##                                                              Max.   :32.10
##                                                              NA's   :116
##  Bottom.Temp      Salinity           DO          Bottom.Type
##  Mode:logical   Min.   :10.00   Min.   : 0.000   Mode:logical
##  NA's:4012      1st Qu.:23.70   1st Qu.: 5.200   NA's:4012
##                 Median :25.10   Median : 6.590
##                 Mean   :24.87   Mean   : 6.847
##                 3rd Qu.:26.40   3rd Qu.: 8.130
##                 Max.   :36.60   Max.   :20.000
##                 NA's   :170     NA's   :403
```

Some data exploration to determine which years and covariates to keep; then remove missing values

```
# remove fields with too many (or all) NAs:
indata <- indata[,-c(8,11,13,14,15,16,18,21)]
summary(indata)
```

```
##     SampleID            State              Program              Gear
##  Min.   :19871018   Length:4012        Length:4012        Length:4012
##  1st Qu.:20001111   Class :character   Class :character   Class :character
##  Median :20081066   Mode  :character   Mode  :character   Mode  :character
##  Mean   :20067739
##  3rd Qu.:20151067
##  Max.   :20211210
##
##       Year          Month              Day
##  Min.   :1987   Min.   : 6.000   Min.   :1987-06-15 00:00:00
##  1st Qu.:2000   1st Qu.: 7.000   1st Qu.:2000-09-25 00:00:00
##  Median :2008   Median : 8.000   Median :2008-07-01 00:00:00
##  Mean   :2007   Mean   : 7.846   Mean   :2007-04-10 05:46:21
```

```
##  3rd Qu.:2015   3rd Qu.: 9.000   3rd Qu.:2015-06-18 00:00:00
##  Max.   :2021   Max.   :10.000   Max.   :2021-10-21 00:00:00
##
##     Effort                N                Area            Surface.Temp
##  Length:4012        Min.   :   0.00   Length:4012        Min.   : 5.90
##  Class :character   1st Qu.:   0.00   Class :character   1st Qu.:19.10
##  Mode  :character   Median :   1.00   Mode  :character   Median :22.00
##                     Mean   :  19.57                      Mean   :21.36
##                     3rd Qu.:   7.00                      3rd Qu.:24.00
##                     Max.   :6676.00                      Max.   :32.10
##                                                          NA's   :116
##     Salinity          DO
##  Min.   :10.00   Min.   : 0.000
##  1st Qu.:23.70   1st Qu.: 5.200
##  Median :25.10   Median : 6.590
##  Mean   :24.87   Mean   : 6.847
##  3rd Qu.:26.40   3rd Qu.: 8.130
##  Max.   :36.60   Max.   :20.000
##  NA's   :170     NA's   :403
```

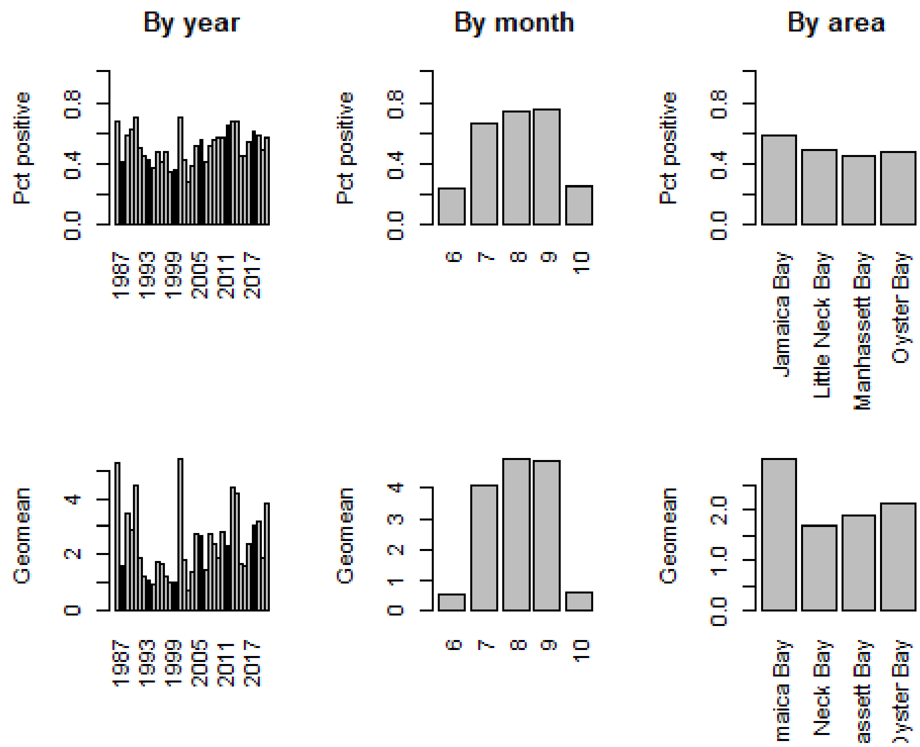
```
# Before dropping NAs, doing some EDA
with(indata, table(Year,Area))
```

```
##       Area
## Year   Jamaica Bay Little Neck Bay Manhassett Bay Oyster Bay
##   1987          41              21             21          0
##   1988          27              22             17          0
##   1989          29              18             24          0
##   1990          29              20             14          0
##   1991          44              10             16          0
##   1992          37              22             22          0
##   1993          39              24             20          0
##   1994          41              21             17          0
##   1995          26              15             14          0
##   1996          19              13             20          0
##   1997          37              17             20          0
##   1998          38              22             20          0
##   1999          34              17             13          0
##   2000          57              28             29          0
##   2001          47              25             30         22
##   2002          48              26             38         15
##   2003          53              34             35         22
##   2004          48              34             36         24
##   2005          55              27             28         24
##   2006          49              29             28         25
##   2007          50              30             30         30
##   2008          54              31             33         30
##   2009          51              30             30         30
##   2010          48              27             30         28
##   2011          54              33             33         30
```

```
##   2012          49              28              30          29
##   2013          57              29              30          29
##   2014          62              29              33          30
##   2015          64              32              33          30
##   2016          60              33              34          30
##   2017          62              33              33          30
##   2018          56              33              33          28
##   2019          61              31              33          26
##   2020          32              25              30          22
##   2021          57              29              29          29
```

```
indata$posTow <- ifelse(indata$N>0,1,0)

with(indata, {
   print(mean(posTow))

   op <- par(mfrow=c(2,3))

   print("Pct positive:")
   barplot(tapply(posTow,Year,mean),las=3,ylim=c(0,1),ylab="Pct
positive",main="By year")
   barplot(tapply(posTow,Month,mean),las=3,ylim=c(0,1),ylab="Pct
positive",main="By month")
   barplot(tapply(posTow,Area,mean),las=3,ylim=c(0,1),ylab="Pct
positive",main="By area")
   print(round(tapply(posTow,list(Month,Area),mean,na.rm=TRUE),2))

   print("Geomean:")
   barplot(tapply(N,Year,geoMean,alpha=NULL,warn=FALSE),las=3,ylab="Geomean")

barplot(tapply(N,Month,geoMean,alpha=NULL,warn=FALSE),las=3,ylab="Geomean")
   barplot(tapply(N,Area,geoMean,alpha=NULL,warn=FALSE),las=3,ylab="Geomean")

print(round(tapply(N,list(Month,Area),geoMean,alpha=NULL,warn=FALSE,na.rm=TRU
E),2))

   par(op)
   }
)
```

```
## [1] 0.5171984
## [1] "Pct positive:"

##    Jamaica Bay Little Neck Bay Manhassett Bay Oyster Bay
## 6         0.31            0.22           0.18       0.11
## 7         0.79            0.56           0.58       0.59
## 8         0.78            0.76           0.69       0.66
## 9         0.77            0.77           0.73       0.73
## 10        0.28            0.25           0.19       0.28
## [1] "Geomean:"
```




```
##    Jamaica Bay Little Neck Bay Manhassett Bay Oyster Bay
## 6         0.93            0.37           0.26       0.21
## 7         5.74            2.46           3.89       3.87
## 8         6.04            3.53           5.20       3.88
## 9         5.72            4.70           3.66       4.81
## 10        0.78            0.50           0.39       0.52
```

```
# Now dropping NAs:
apply(indata, 2, function(x) sum(is.na(x))) # how many records would we lose
due to NAs?
```

```
##     SampleID        State      Program         Gear         Year
Month
##            0            0            0            0            0
0
##          Day       Effort            N         Area Surface.Temp
Salinity
##            0            0            0            0          116
170
##           DO       posTow
##          403            0
```

```
with(indata, tapply(DO,Year,function(x) {sum(is.na(x))})) # in which years do
we lose DOs?
```

```
## 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001
2002
```

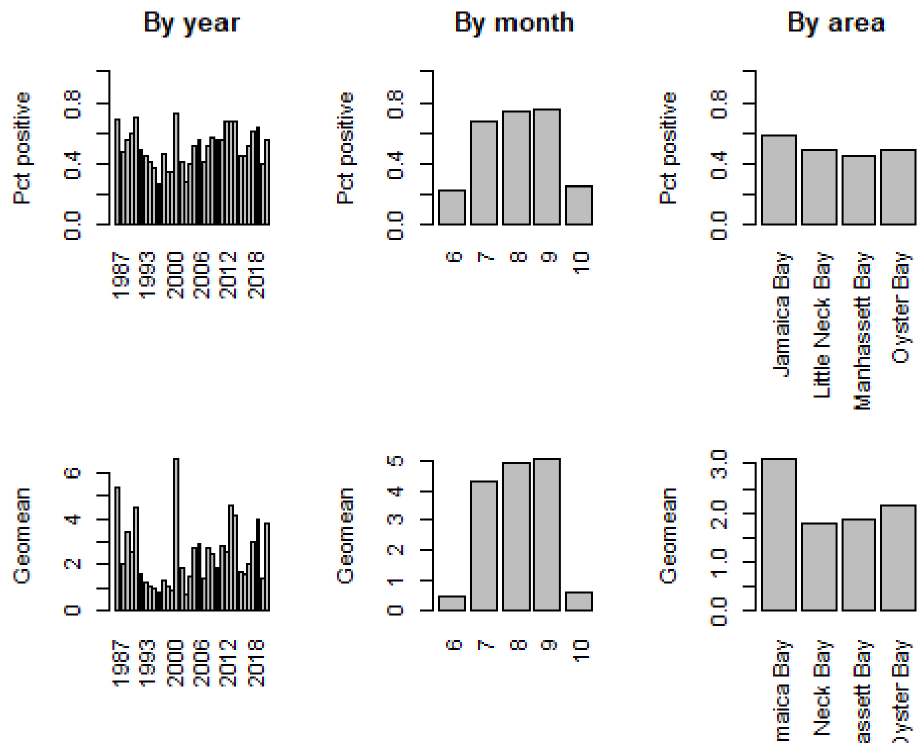
```
##    0   12    4    0    0    7    1   15    1   37   74   26    0    7   18
11
## 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
2018
##    1    5    0   19    0    2    5   26   21   11    4    0    0    0    9
0
## 2019 2020 2021
##   22   64    1
```

```
indata2 <- indata[complete.cases(indata),]

summary(indata2)
```

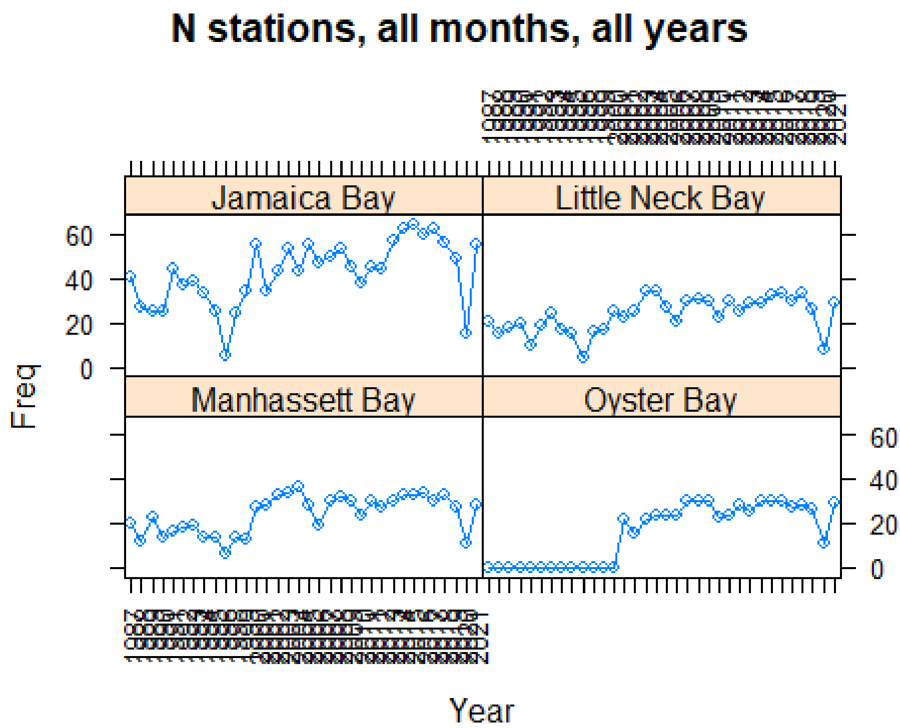
```
##     SampleID            State              Program              Gear
##  Min.   :19871018   Length:3597        Length:3597        Length:3597
##  1st Qu.:20011102   Class :character   Class :character   Class :character
##  Median :20081111   Mode  :character   Mode  :character   Mode  :character
##  Mean   :20069754
##  3rd Qu.:20151072
##  Max.   :20211210
##       Year          Month             Day
##  Min.   :1987   Min.   : 6.000   Min.   :1987-06-15 00:00:00
##  1st Qu.:2001   1st Qu.: 7.000   1st Qu.:2001-08-28 00:00:00
##  Median :2008   Median : 8.000   Median :2008-08-12 00:00:00
##  Mean   :2007   Mean   : 7.837   Mean   :2007-06-22 13:04:15
##  3rd Qu.:2015   3rd Qu.: 9.000   3rd Qu.:2015-06-18 00:00:00
##  Max.   :2021   Max.   :10.000   Max.   :2021-10-21 00:00:00
##     Effort                N                Area            Surface.Temp
##  Length:3597        Min.   :   0.00   Length:3597        Min.   : 5.90
##  Class :character   1st Qu.:   0.00   Class :character   1st Qu.:19.20
##  Mode  :character   Median :   1.00   Mode  :character   Median :21.90
##                     Mean   :  19.67                      Mean   :21.37
##                     3rd Qu.:   7.00                      3rd Qu.:24.00
##                     Max.   :6676.00                      Max.   :32.10
##     Salinity           DO             posTow
##  Min.   :10.00   Min.   : 0.000   Min.   :0.000
##  1st Qu.:23.60   1st Qu.: 5.200   1st Qu.:0.000
##  Median :25.10   Median : 6.580   Median :1.000
##  Mean   :24.86   Mean   : 6.847   Mean   :0.519
##  3rd Qu.:26.50   3rd Qu.: 8.140   3rd Qu.:1.000
##  Max.   :36.60   Max.   :20.000   Max.   :1.000
```

```
with(indata2, {
   print(mean(posTow))

   op <- par(mfrow=c(2,3))

   print("Pct positive:")
   barplot(tapply(posTow,Year,mean),las=3,ylim=c(0,1),ylab="Pct
positive",main="By year")
```

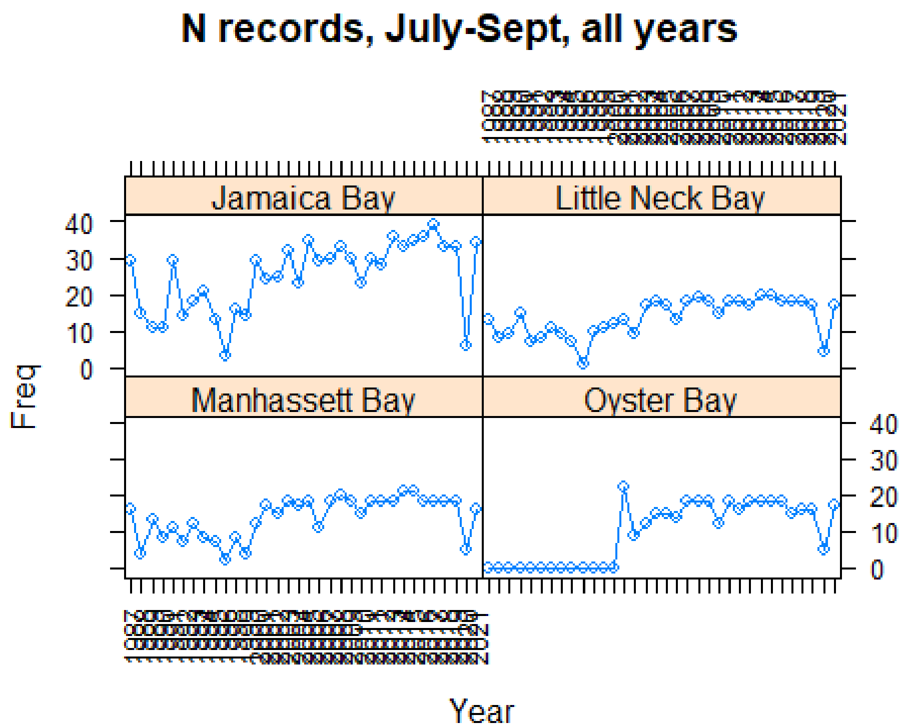
```
   barplot(tapply(posTow,Month,mean),las=3,ylim=c(0,1),ylab="Pct
positive",main="By month")
   barplot(tapply(posTow,Area,mean),las=3,ylim=c(0,1),ylab="Pct
positive",main="By area")
   print(round(tapply(posTow,list(Month,Area),mean,na.rm=TRUE),2))

   print("Geomean:")
   barplot(tapply(N,Year,geoMean,alpha=NULL,warn=FALSE),las=3,ylab="Geomean")

barplot(tapply(N,Month,geoMean,alpha=NULL,warn=FALSE),las=3,ylab="Geomean")
   barplot(tapply(N,Area,geoMean,alpha=NULL,warn=FALSE),las=3,ylab="Geomean")

print(round(tapply(N,list(Month,Area),geoMean,alpha=NULL,warn=FALSE,na.rm=TRU
E),2))
   par(op)
   }
)

## [1] 0.5190436
## [1] "Pct positive:"

##    Jamaica Bay Little Neck Bay Manhassett Bay Oyster Bay
## 6         0.30            0.22           0.19       0.12
## 7         0.80            0.55           0.59       0.60
## 8         0.78            0.78           0.69       0.66
## 9         0.77            0.77           0.70       0.73
## 10        0.28            0.24           0.17       0.28
## [1] "Geomean:"
```




```
##    Jamaica Bay Little Neck Bay Manhassett Bay Oyster Bay
## 6         0.88            0.37           0.26       0.22
## 7         6.13            2.49           3.95       4.10
## 8         6.01            3.71           5.09       3.93
## 9         6.05            4.99           3.63       4.87
## 10        0.84            0.48           0.39       0.50

# NY doesn't seem to sample the same number of stations every year &
something happened in 1996 with sampling effort
xyplot(Freq~Year | Area, as.table=TRUE,
   data=as.data.frame(with(indata2, table(Year,Area))),type="o",
   main="N stations, all months, all years",scales=list(x=list(rot=90)))
```




```
# Something happened in 1996 with sampling effort
xyplot(Freq~Year | Area, as.table=TRUE,

data=as.data.frame(with(indata2[is.element(indata2$Month,7:9),],table(Year,Ar
ea))),
   type="o",main="N records, July-Sept, all
years",scales=list(x=list(rot=90)))
```




```
# Nevertheless, I'm proposing to subset to: omit Oyster bay (sampled 1x per
mo, vs 2x per mo & started in 2001), include July-September, and for now, use
all years


with(indata2[is.element(indata2$Month,7:9),],
tapply(N,list(Year,Area),length))

##      Jamaica Bay Little Neck Bay Manhassett Bay Oyster Bay
## 1987          29              13             16         NA
## 1988          15               8              4         NA
## 1989          11               9             13         NA
## 1990          11              15              8         NA
## 1991          29               7             11         NA
## 1992          14               8              7         NA
## 1993          18              11             12         NA
## 1994          21               9              8         NA
## 1995          13               7              7         NA
## 1996           3               1              2         NA
## 1998          16              10              8         NA
## 1999          14              11              4         NA
## 2000          29              12             12         NA
## 2001          24              13             17         22
## 2002          25               9             15          9
## 2003          32              17             18         12
## 2004          23              18             17         15
## 2005          35              17             18         15
```

```
## 2006           29               13             11          14
## 2007           30               18             18          18
## 2008           33               19             20          18
## 2009           30               18             18          18
## 2010           23               15             15          12
## 2011           30               18             18          18
## 2012           28               18             18          16
## 2013           36               17             18          18
## 2014           33               20             21          18
## 2015           35               20             21          18
## 2016           36               18             18          18
## 2017           39               18             18          15
## 2018           33               18             18          16
## 2019           33               17             18          16
## 2020            6                4              5           5
## 2021           34               17             16          17
```

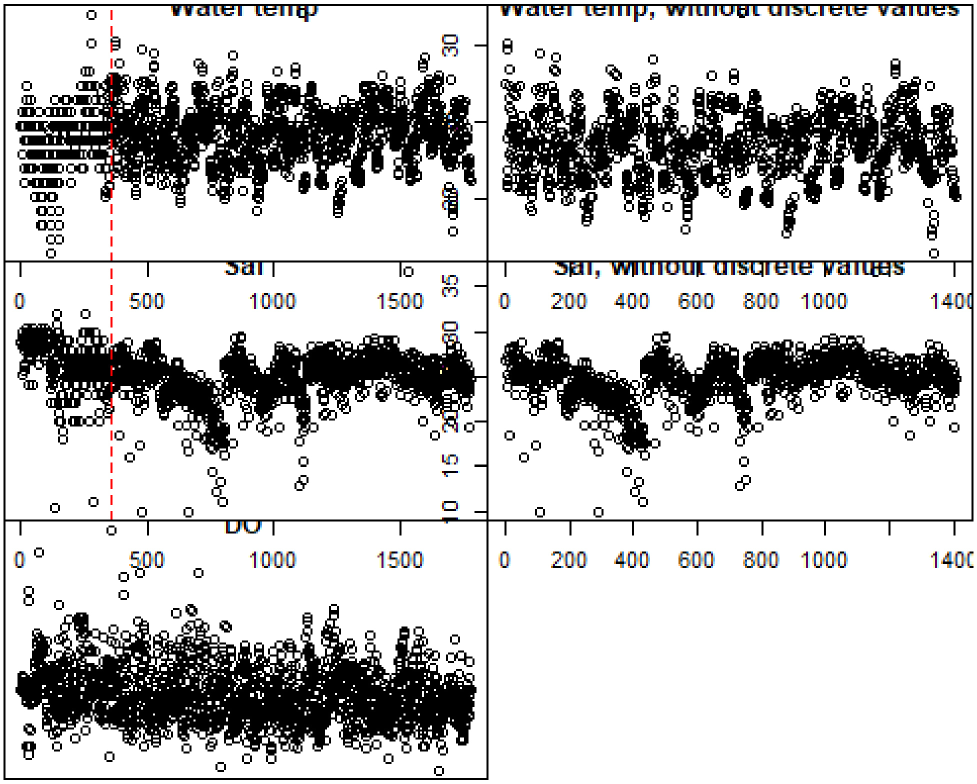
```
# sampling is less consistent prior to 2003, but keeping all years for now


# subsetting to include just July - Sept, inclusive (and omit Oyster Bay)
indata2 <- indata2[is.element(indata2$Month,7:9) &
(!is.element(indata2$Area,"Oyster Bay")), ]
with(indata2[is.element(indata2$Month,7:9),],
tapply(N,list(Year,Area),length))
```

```
##      Jamaica Bay Little Neck Bay Manhassett Bay
## 1987          29              13             16
## 1988          15               8              4
## 1989          11               9             13
## 1990          11              15              8
## 1991          29               7             11
## 1992          14               8              7
## 1993          18              11             12
## 1994          21               9              8
## 1995          13               7              7
## 1996           3               1              2
## 1998          16              10              8
## 1999          14              11              4
## 2000          29              12             12
## 2001          24              13             17
## 2002          25               9             15
## 2003          32              17             18
## 2004          23              18             17
## 2005          35              17             18
## 2006          29              13             11
## 2007          30              18             18
## 2008          33              19             20
## 2009          30              18             18
## 2010          23              15             15
```

```
## 2011          30              18              18
## 2012          28              18              18
## 2013          36              17              18
## 2014          33              20              21
## 2015          35              20              21
## 2016          36              18              18
## 2017          39              18              18
## 2018          33              18              18
## 2019          33              17              18
## 2020           6               4               5
## 2021          34              17              16
```

Check for outliers

```
# Check for outliers or weirdo values
op <- par(mfrow=c(3,2),mar=c(0,0,0,0))
plot(indata2$Surface.Temp,main="Water temp") # remove the first ~360 records?
readings prior to 1999? No.
   abline(v=360,col="red",lty=2)
   plot(indata2$Surface.Temp[indata2$Year>1998],main="Water temp, without
discrete values")
plot(indata2$Salinity, main="Sal") # remove the first ~360 records; readings
prior to 1999? No.
   abline(v=360,col="red",lty=2)
   plot(indata2$Salinity[indata2$Year>1998],main="Sal, without discrete
values")
plot(indata2$DO,main="DO")
par(op)
```



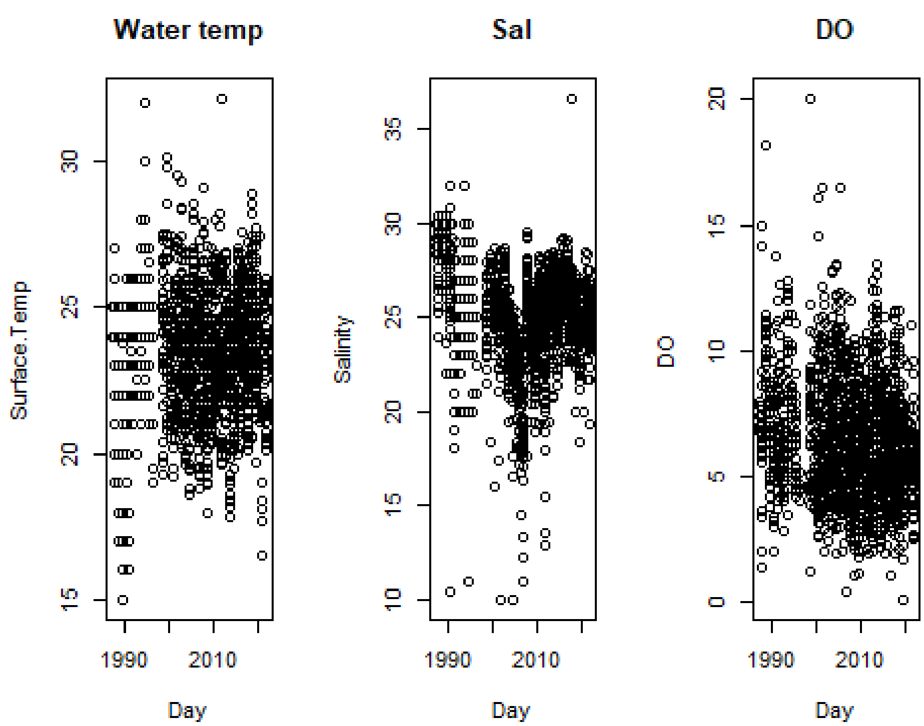

```
# Could remove one weird salinity value, but not going to.
indata2[indata2$Salinity>=30,]

## # A tibble: 28 x 14
##    SampleID State Program     Gear   Year Month Day                 Effort
N
##       <dbl> <chr> <chr>       <chr> <dbl> <dbl> <dttm>              <chr>
<dbl>
##  1 19871153 NY    Western Lo~ Beac~  1987     7 1987-07-02 00:00:00 Tow
0
##  2 19871025 NY    Western Lo~ Beac~  1987     7 1987-07-14 00:00:00 Tow
41
##  3 19871155 NY    Western Lo~ Beac~  1987     7 1987-07-14 00:00:00 Tow
133
##  4 19871159 NY    Western Lo~ Beac~  1987     7 1987-07-16 00:00:00 Tow
2
##  5 19871167 NY    Western Lo~ Beac~  1987     9 1987-09-08 00:00:00 Tow
0
##  6 19871172 NY    Western Lo~ Beac~  1987     9 1987-09-10 00:00:00 Tow
0
##  7 19881030 NY    Western Lo~ Beac~  1988     7 1988-07-06 00:00:00 Tow
30
##  8 19881045 NY    Western Lo~ Beac~  1988     9 1988-09-21 00:00:00 Tow
0
##  9 19881046 NY    Western Lo~ Beac~  1988     9 1988-09-26 00:00:00 Tow
0
## 10 19881134 NY    Western Lo~ Beac~  1988     9 1988-09-26 00:00:00 Tow
```

```
9
## # ... with 18 more rows, and 5 more variables: Area <chr>, Surface.Temp
<dbl>,
## #   Salinity <dbl>, DO <dbl>, posTow <dbl>
```

```
indata2[indata2$Salinity>=30 & indata2$Year>1998,]
```

```
## # A tibble: 1 x 14
##   SampleID State Program      Gear   Year Month Day                    Effort
N
##      <dbl> <chr> <chr>        <chr> <dbl> <dbl> <dttm>                 <chr>
<dbl>
## 1 20171149 NY    Western Lon~ Beac~  2017     8 2017-08-31 00:00:00 Tow
3
## # ... with 5 more variables: Area <chr>, Surface.Temp <dbl>, Salinity
<dbl>,
## #   DO <dbl>, posTow <dbl>
```

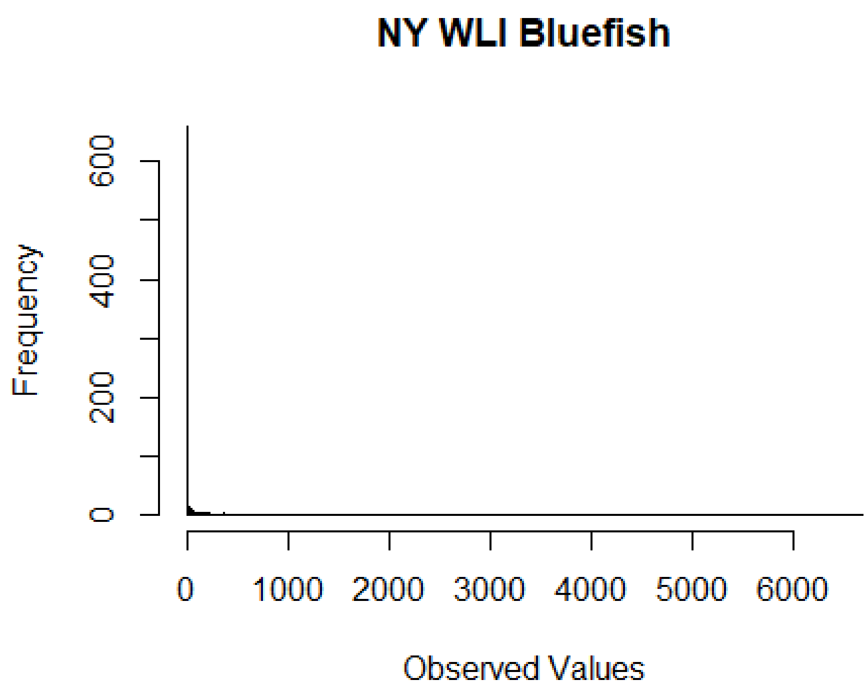
```
#indata2 <- indata2[indata2$Year>1998 & indata2$Salinity<30,]
op <- par(mfrow=c(1,3))
with(indata2, plot(Day,Surface.Temp,main="Water temp"))
with(indata2, plot(Day,Salinity, main="Sal"))
with(indata2, plot(Day,DO,main="DO"))
```




```
op <- par(mfrow=c(1,1))
```

Plot distribution of catches:

```
hist(indata2$N,main="NY WLI Bluefish",xlab="Observed Values",
      ylab="Frequency", breaks=seq(0,max(indata2$N),1))
```

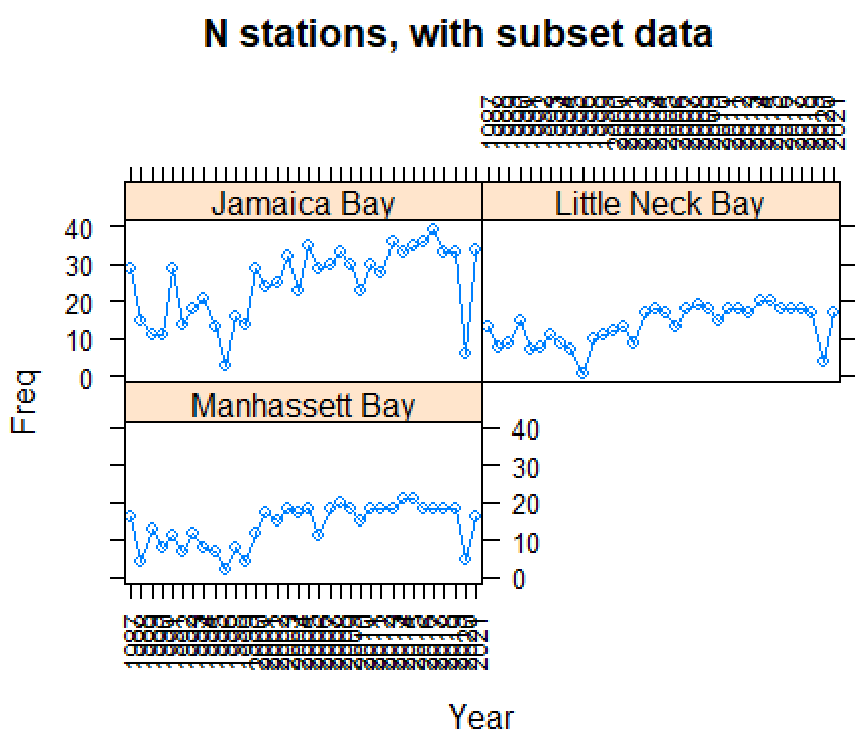



```
# look at values of large catches:
sort(indata2$N,decreasing=TRUE)[1:50]

##  [1] 6676 2058 1602 1440  945  910  850  842  781  741  735  727  702  660
589
## [16]  538  487  473  426  425  395  382  373  373  366  355  345  325  324
320
## [31]  288  284  273  252  248  247  240  239  235  232  228  228  217  208
204
## [46]  201  199  199  198  198
```

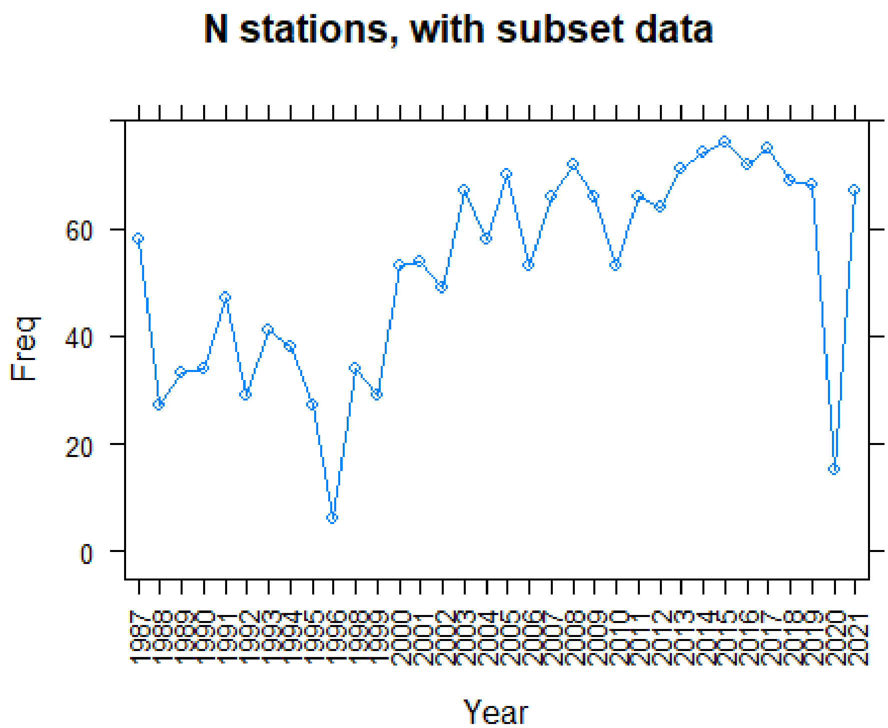
Finally, check to see if there are any years with zero catch that need to be removed:

```
# You also want to check to see if you have any years where you have zero
# catch (no positive tows/hauls/sets)
diff(as.numeric(names(xtabs(N~Year, indata2))))

##  [1] 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

# There are no records from 1997 (lost due to no DOs)
xtabs(N~Year, indata2)

## Year
##  1987  1988  1989  1990  1991  1992  1993  1994  1995  1996  1998  1999
2000
```

```
## 10909   605  2550   513  1108   743   289   184   261    50   231   195 
390 
##  2001  2002  2003  2004  2005  2006  2007  2008  2009  2010  2011  2012 
2013 
##  3129   613   602   912  3830   851  1204  3445  1067  1072  1850   667 
1415 
##  2014  2015  2016  2017  2018  2019  2020  2021 
##  2501  1820  2314  1953  2270  3373   727  2357

table(indata2$Year)

## 
## 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1998 1999 2000 2001 2002 
2003 
##   58   27   33   34   47   29   41   38   27    6   34   29   53   54   49 
67 
## 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 
2019 
##   58   70   53   66   72   66   53   66   64   71   74   76   72   75   69 
68 
## 2020 2021 
##   15   67

# All years have catch, variable stations per year, but reasonable for now:
xyplot(Freq~Year | Area, as.table=TRUE,
   data=as.data.frame(with(indata2, table(Year,Area))),type="o",
   main="N stations, with subset data",scales=list(x=list(rot=90)))
```



```
xyplot(Freq~Year, as.table=TRUE,ylim=c(-5,80),
   data=as.data.frame(with(indata2, table(Year))),type="o",
   main="N stations, with subset data",scales=list(x=list(rot=90)))
```




```
# Not dropping yet, but keep an eye on whether I may want to drop out 1996 &
2020
```

Create a dataframe with all the factors we need (z-score all continuous variables; all others as.factors)

```
dat = data.frame(CPUE = indata2$N,
                 YEAR = as.factor(indata2$Year),
                 MONTH = as.factor(indata2$Month),
                 AREA = as.factor(indata2$Area),
                 STEMP = Z.scr(indata2$Surface.Temp),
                 SSAL = Z.scr(indata2$Salinity),
                 SDO = Z.scr(indata2$DO))

# Calculate the proportion of positive tows/sets/hauls
dat$PosTow <- ifelse(dat$CPUE > 0, 1, 0)
mean(dat$PosTow)

## [1] 0.7276811

head(dat)

##   CPUE YEAR MONTH        AREA     STEMP     SSAL       SDO PosTow
## 1  425 1987     7 Jamaica Bay 0.2931763 1.452187 0.1287865      1
```

```
## 2    1 1987     7 Jamaica Bay 0.2931763 1.299106 0.3832448      1
## 3 2058 1987     7 Jamaica Bay 0.2931763 1.605268 0.2136059      1
## 4    5 1987     7 Jamaica Bay 0.7684646 1.299106 0.3832448      1
## 5    0 1987     7 Jamaica Bay 0.2931763 1.605268 0.1287865      0
## 6    7 1987     7 Jamaica Bay 0.2931763 1.299106 0.6377031      1
```

Check collinearity

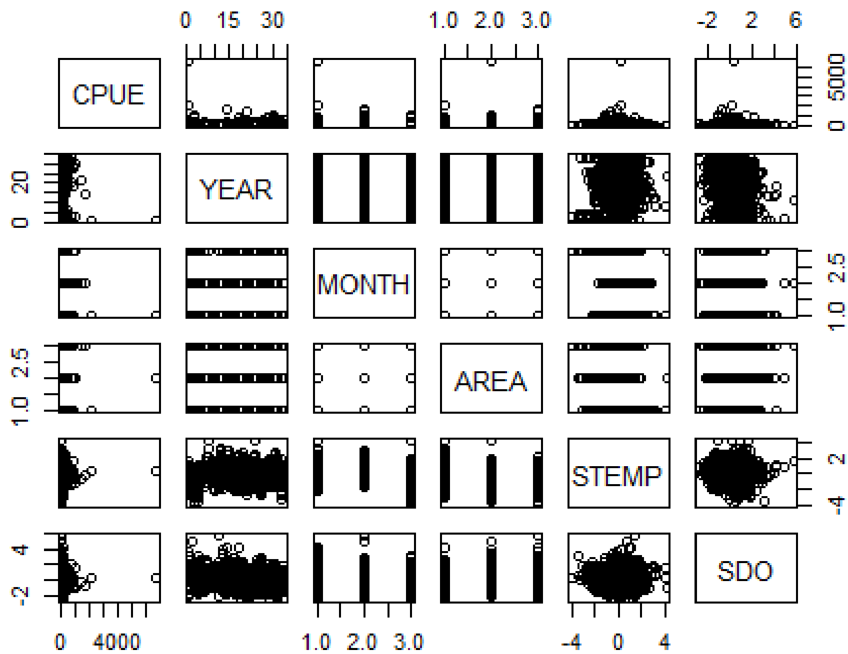
```
mod = lm(CPUE ~ YEAR + MONTH + AREA + STEMP + SSAL + SDO, data = dat)
vif(mod)
```

```
##           GVIF Df GVIF^(1/(2*Df))
## YEAR  3.634306 33        1.019744
## MONTH 1.453091  2        1.097926
## AREA  1.435064  2        1.094505
## STEMP 1.884181  1        1.372655
## SSAL  2.110505  1        1.452758
## SDO   1.299648  1        1.140021
```

```
# You want GVIF to be less than ~3.
# Year is > 3, but can't drop it, so dropping next highest, SSAL
mod = lm(CPUE ~ YEAR + MONTH + AREA + STEMP + SDO, data = dat)
vif(mod)
```

```
##           GVIF Df GVIF^(1/(2*Df))
## YEAR  1.776916 33        1.008748
## MONTH 1.434552  2        1.094408
## AREA  1.371777  2        1.082233
## STEMP 1.883775  1        1.372507
## SDO   1.295581  1        1.138236
```

```
pairs(~ CPUE + YEAR + MONTH + AREA + STEMP + SDO, data = dat)
```



Step 2: Model selection

Compare negative binom, zero-infl neg binom, zero-altered neg binom

Model 1: Negative Binomial

```
tmb1.NB <- glmmTMB(CPUE ~ YEAR + MONTH + AREA + STEMP + SDO, data = dat,
family = nbinom2)

# Check the Std. Error of the estimates; high SEs indicate problems with the
fit.
summary(tmb1.NB)

##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12681.8  12906.7  -6299.9  12599.8     1740
##
##
## Dispersion parameter for nbinom2 family (): 0.263
##
## Conditional model:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)            5.220198   0.272042  19.189  < 2e-16 ***
## YEAR1988              -2.035414   0.469556  -4.335 1.46e-05 ***
## YEAR1989              -0.818436   0.447472  -1.829 0.067396 .
```

```
## YEAR1990            -2.482592   0.439094  -5.654 1.57e-08 ***
## YEAR1991            -2.127448   0.388832  -5.471 4.47e-08 ***
## YEAR1992            -1.934689   0.456271  -4.240 2.23e-05 ***
## YEAR1993            -3.316592   0.407549  -8.138 4.02e-16 ***
## YEAR1994            -3.709637   0.427389  -8.680  < 2e-16 ***
## YEAR1995            -2.997083   0.469498  -6.384 1.73e-10 ***
## YEAR1996            -3.143413   0.853545  -3.683 0.000231 ***
## YEAR1998            -3.414891   0.455802  -7.492 6.78e-14 ***
## YEAR1999            -3.400056   0.456238  -7.452 9.17e-14 ***
## YEAR2000            -3.199663   0.391894  -8.165 3.22e-16 ***
## YEAR2001            -1.296715   0.386453  -3.355 0.000792 ***
## YEAR2002            -2.735037   0.399241  -6.851 7.35e-12 ***
## YEAR2003            -3.009040   0.359867  -8.362  < 2e-16 ***
## YEAR2004            -2.517191   0.379049  -6.641 3.12e-11 ***
## YEAR2005            -1.290976   0.367312  -3.515 0.000440 ***
## YEAR2006            -2.510606   0.385817  -6.507 7.65e-11 ***
## YEAR2007            -2.292544   0.358689  -6.391 1.64e-10 ***
## YEAR2008            -1.465851   0.365633  -4.009 6.10e-05 ***
## YEAR2009            -2.495966   0.364725  -6.843 7.73e-12 ***
## YEAR2010            -2.326355   0.384959  -6.043 1.51e-09 ***
## YEAR2011            -1.950963   0.358649  -5.440 5.34e-08 ***
## YEAR2012            -2.968313   0.363299  -8.170 3.07e-16 ***
## YEAR2013            -2.245035   0.358826  -6.257 3.93e-10 ***
## YEAR2014            -1.771625   0.356732  -4.966 6.83e-07 ***
## YEAR2015            -2.181165   0.352387  -6.190 6.03e-10 ***
## YEAR2016            -1.794673   0.371543  -4.830 1.36e-06 ***
## YEAR2017            -1.971540   0.354843  -5.556 2.76e-08 ***
## YEAR2018            -1.894480   0.364989  -5.191 2.10e-07 ***
## YEAR2019            -1.385919   0.363913  -3.808 0.000140 ***
## YEAR2020            -1.388115   0.575498  -2.412 0.015864 *
## YEAR2021            -1.719852   0.362824  -4.740 2.14e-06 ***
## MONTH8              -0.031502   0.123501  -0.255 0.798666
## MONTH9               0.024094   0.158909   0.152 0.879487
## AREALittle Neck Bay -0.009538   0.131843  -0.072 0.942327
## AREAManhassett Bay   0.214005   0.132745   1.612 0.106930
## STEMP                0.083051   0.073997   1.122 0.261712
## SDO                 -0.018421   0.058764  -0.313 0.753920
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# dispersion of a glmmTMB model:
disp3(tmb1.NB)
```

```
## df.resid
## 2.253395
```

Model 2: ZINB

```
ZINB = glmmTMB(CPUE ~ YEAR + MONTH + AREA + STEMP + SDO, ziformula = ~ YEAR +
MONTH + AREA + STEMP + SDO, data = dat, family=nbinom2)
summary(ZINB)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
## Zero inflation:        ~YEAR + MONTH + AREA + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12562.9  13007.2  -6200.4  12400.9     1700
##
##
## Dispersion parameter for nbinom2 family (): 0.304
##
## Conditional model:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)          5.33732    0.26165  20.398  < 2e-16 ***
## YEAR1988            -2.08836    0.44644  -4.678 2.90e-06 ***
## YEAR1989            -0.85557    0.43573  -1.964 0.049587 *
## YEAR1990            -2.73788    0.41252  -6.637 3.20e-11 ***
## YEAR1991            -2.15344    0.36879  -5.839 5.25e-09 ***
## YEAR1992            -1.96801    0.43846  -4.488 7.17e-06 ***
## YEAR1993            -3.26620    0.38606  -8.460  < 2e-16 ***
## YEAR1994            -3.49149    0.40418  -8.638  < 2e-16 ***
## YEAR1995            -2.85024    0.46994  -6.065 1.32e-09 ***
## YEAR1996            -2.57762    1.08756  -2.370 0.017783 *
## YEAR1998            -2.70120    0.46144  -5.854 4.80e-09 ***
## YEAR1999            -3.39172    0.44819  -7.568 3.80e-14 ***
## YEAR2000            -3.13814    0.37793  -8.304  < 2e-16 ***
## YEAR2001            -1.32527    0.36474  -3.634 0.000280 ***
## YEAR2002            -2.26728    0.40236  -5.635 1.75e-08 ***
## YEAR2003            -2.86185    0.35604  -8.038 9.13e-16 ***
## YEAR2004            -1.99066    0.39685  -5.016 5.27e-07 ***
## YEAR2005            -0.81888    0.35813  -2.287 0.022223 *
## YEAR2006            -2.33333    0.37531  -6.217 5.06e-10 ***
## YEAR2007            -2.39022    0.34879  -6.853 7.24e-12 ***
## YEAR2008            -1.55649    0.35225  -4.419 9.93e-06 ***
## YEAR2009            -2.45763    0.34980  -7.026 2.13e-12 ***
## YEAR2010            -2.24936    0.36457  -6.170 6.83e-10 ***
## YEAR2011            -2.08191    0.34167  -6.093 1.11e-09 ***
## YEAR2012            -2.92151    0.34498  -8.469  < 2e-16 ***
## YEAR2013            -2.19403    0.34138  -6.427 1.30e-10 ***
## YEAR2014            -1.81133    0.33890  -5.345 9.05e-08 ***
## YEAR2015            -2.09022    0.34282  -6.097 1.08e-09 ***
## YEAR2016            -1.32998    0.37077  -3.587 0.000334 ***
## YEAR2017            -1.89884    0.33755  -5.625 1.85e-08 ***
## YEAR2018            -1.69503    0.35224  -4.812 1.49e-06 ***
## YEAR2019            -1.29638    0.34717  -3.734 0.000188 ***
## YEAR2020            -1.12656    0.56430  -1.996 0.045891 *
## YEAR2021            -1.47164    0.35516  -4.144 3.42e-05 ***
## MONTH8              -0.21678    0.12061  -1.797 0.072275 .
## MONTH9              -0.47032    0.15426  -3.049 0.002297 **
## AREALittle Neck Bay  0.09514    0.12745   0.746 0.455389
```

```
## AREAManhassett Bay   0.37197    0.12893   2.885 0.003914 **
## STEMP               -0.11061    0.06799  -1.627 0.103773
## SDO                 -0.10476    0.05598  -1.872 0.061266 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)         -7.479e+00  2.032e+00  -3.680 0.000233 ***
## YEAR1988             3.084e+00  2.560e+00   1.205 0.228274
## YEAR1989            -1.569e+00  1.621e+00  -0.968 0.332930
## YEAR1990            -2.974e+01  5.470e+05   0.000 0.999957
## YEAR1991            -1.381e+01  1.396e+03  -0.010 0.992105
## YEAR1992             6.097e-02  2.155e+00   0.028 0.977429
## YEAR1993            -2.656e+01  1.182e+06   0.000 0.999982
## YEAR1994            -1.188e+01  2.240e+03  -0.005 0.995767
## YEAR1995             2.931e+00  1.686e+00   1.738 0.082126 .
## YEAR1996             1.687e+00  4.857e+00   0.347 0.728352
## YEAR1998             4.742e+00  1.945e+00   2.438 0.014762 *
## YEAR1999             1.738e+00  2.163e+00   0.804 0.421667
## YEAR2000            -1.817e+00  2.080e+00  -0.874 0.382269
## YEAR2001            -2.062e+01  1.420e+04  -0.001 0.998841
## YEAR2002             4.846e+00  1.675e+00   2.893 0.003821 **
## YEAR2003             1.733e+00  1.410e+00   1.229 0.218993
## YEAR2004             2.650e+00  1.401e+00   1.891 0.058578 .
## YEAR2005             3.364e+00  1.463e+00   2.300 0.021433 *
## YEAR2006             9.700e-01  1.599e+00   0.607 0.544144
## YEAR2007             2.172e-01  1.843e+00   0.118 0.906185
## YEAR2008             4.036e-01  1.379e+00   0.293 0.769735
## YEAR2009            -1.920e+00  1.533e+00  -1.252 0.210550
## YEAR2010            -1.632e+01  3.250e+03  -0.005 0.995993
## YEAR2011            -2.184e+00  2.075e+00  -1.053 0.292541
## YEAR2012            -1.878e+01  8.626e+03  -0.002 0.998263
## YEAR2013            -2.994e+01  2.480e+05   0.000 0.999904
## YEAR2014            -2.554e+01  8.438e+04   0.000 0.999759
## YEAR2015             1.408e+00  1.364e+00   1.032 0.302206
## YEAR2016             4.610e+00  1.512e+00   3.048 0.002302 **
## YEAR2017            -1.919e+01  5.464e+03  -0.004 0.997197
## YEAR2018             1.432e+00  1.743e+00   0.822 0.411203
## YEAR2019             1.317e-01  1.656e+00   0.080 0.936615
## YEAR2020             6.071e-01  2.397e+00   0.253 0.800099
## YEAR2021             2.245e+00  1.301e+00   1.725 0.084437 .
## MONTH8              -1.747e+00  7.823e-01  -2.233 0.025546 *
## MONTH9              -7.228e+00  1.640e+00  -4.406 1.05e-05 ***
## AREALittle Neck Bay  2.559e+00  1.261e+00   2.030 0.042399 *
## AREAManhassett Bay   4.366e+00  1.339e+00   3.260 0.001115 **
## STEMP               -3.229e+00  7.544e-01  -4.280 1.87e-05 ***
## SDO                 -2.536e+00  6.992e-01  -3.627 0.000287 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# High SEs for YEAR in zero-infl model so removing:
ZINB = glmmTMB(CPUE ~ YEAR + MONTH + AREA + STEMP + SDO, ziformula = ~ MONTH
+ AREA + STEMP + SDO, data = dat, family=nbinom2)
summary(ZINB)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
## Zero inflation:        ~MONTH + AREA + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12604.4  12867.7  -6254.2  12508.4     1733
##
##
## Dispersion parameter for nbinom2 family (): 0.298
##
## Conditional model:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)           5.33034    0.26535  20.088  < 2e-16 ***
## YEAR1988             -2.12071    0.44881  -4.725 2.30e-06 ***
## YEAR1989             -0.82749    0.44376  -1.865 0.062220 .
## YEAR1990             -2.64593    0.42830  -6.178 6.50e-10 ***
## YEAR1991             -2.17570    0.37304  -5.832 5.47e-09 ***
## YEAR1992             -1.97814    0.44335  -4.462 8.13e-06 ***
## YEAR1993             -3.27920    0.39323  -8.339  < 2e-16 ***
## YEAR1994             -3.56023    0.41054  -8.672  < 2e-16 ***
## YEAR1995             -2.94545    0.46305  -6.361 2.01e-10 ***
## YEAR1996             -2.62179    1.08049  -2.426 0.015246 *
## YEAR1998             -2.99437    0.45621  -6.564 5.25e-11 ***
## YEAR1999             -3.42802    0.44441  -7.714 1.22e-14 ***
## YEAR2000             -3.14052    0.38459  -8.166 3.19e-16 ***
## YEAR2001             -1.34595    0.37084  -3.629 0.000284 ***
## YEAR2002             -2.50015    0.39399  -6.346 2.21e-10 ***
## YEAR2003             -2.92915    0.35696  -8.206 2.29e-16 ***
## YEAR2004             -2.21841    0.38806  -5.717 1.09e-08 ***
## YEAR2005             -0.97530    0.36103  -2.701 0.006904 **
## YEAR2006             -2.40345    0.37819  -6.355 2.08e-10 ***
## YEAR2007             -2.36848    0.35357  -6.699 2.10e-11 ***
## YEAR2008             -1.55853    0.35679  -4.368 1.25e-05 ***
## YEAR2009             -2.48357    0.35619  -6.973 3.11e-12 ***
## YEAR2010             -2.30082    0.37029  -6.214 5.18e-10 ***
## YEAR2011             -2.05593    0.34851  -5.899 3.65e-09 ***
## YEAR2012             -2.95770    0.35019  -8.446  < 2e-16 ***
## YEAR2013             -2.22826    0.34671  -6.427 1.30e-10 ***
## YEAR2014             -1.83646    0.34449  -5.331 9.77e-08 ***
## YEAR2015             -2.14870    0.34637  -6.203 5.52e-10 ***
## YEAR2016             -1.54749    0.36516  -4.238 2.26e-05 ***
## YEAR2017             -1.95468    0.34286  -5.701 1.19e-08 ***
## YEAR2018             -1.77255    0.35630  -4.975 6.53e-07 ***
## YEAR2019             -1.34449    0.35243  -3.815 0.000136 ***
```

```
## YEAR2020             -1.16483    0.57832  -2.014 0.043991 *
## YEAR2021             -1.58111    0.35705  -4.428 9.50e-06 ***
## MONTH8               -0.17571    0.12141  -1.447 0.147821
## MONTH9               -0.32289    0.15858  -2.036 0.041734 *
## AREALittle Neck Bay   0.08231    0.12868   0.640 0.522416
## AREAManhassett Bay    0.35587    0.13210   2.694 0.007061 **
## STEMP                -0.07070    0.07083  -0.998 0.318249
## SDO                  -0.08991    0.05654  -1.590 0.111785
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -17.3120  1260.5059  -0.014  0.98904
## MONTH8                 -1.4954     0.4989  -2.998  0.00272 **
## MONTH9                 -3.3436     0.8324  -4.017 5.90e-05 ***
## AREALittle Neck Bay    14.6866  1260.5059   0.012  0.99070
## AREAManhassett Bay     15.4173  1260.5058   0.012  0.99024
## STEMP                  -1.6382     0.3187  -5.140 2.75e-07 ***
## SDO                    -1.5056     0.2985  -5.043 4.58e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# High SEs in area, so removing AREA in zi model:
ZINB = glmmTMB(CPUE ~ YEAR + MONTH + AREA + STEMP + SDO, ziformula = ~ MONTH
+ STEMP + SDO, data = dat, family=nbinom2)
summary(ZINB)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
## Zero inflation:        ~MONTH + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12615.3  12867.6  -6261.7  12523.3     1735
##
##
## Dispersion parameter for nbinom2 family (): 0.294
##
## Conditional model:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)           5.34922    0.26861  19.914  < 2e-16 ***
## YEAR1988             -2.13011    0.45218  -4.711 2.47e-06 ***
## YEAR1989             -0.83439    0.44563  -1.872 0.061152 .
## YEAR1990             -2.56728    0.43964  -5.840 5.24e-09 ***
## YEAR1991             -2.16555    0.37664  -5.750 8.95e-09 ***
## YEAR1992             -1.98619    0.44530  -4.460 8.18e-06 ***
## YEAR1993             -3.28922    0.39623  -8.301  < 2e-16 ***
## YEAR1994             -3.56767    0.41345  -8.629  < 2e-16 ***
## YEAR1995             -2.94157    0.46683  -6.301 2.95e-10 ***
```

```
## YEAR1996            -2.65168    1.08146  -2.452 0.014208 *
## YEAR1998            -2.97649    0.46016  -6.468 9.91e-11 ***
## YEAR1999            -3.41597    0.44715  -7.639 2.18e-14 ***
## YEAR2000            -3.12468    0.39050  -8.002 1.23e-15 ***
## YEAR2001            -1.31116    0.37380  -3.508 0.000452 ***
## YEAR2002            -2.51121    0.39608  -6.340 2.30e-10 ***
## YEAR2003            -2.94118    0.35959  -8.179 2.85e-16 ***
## YEAR2004            -2.19958    0.39172  -5.615 1.96e-08 ***
## YEAR2005            -0.99712    0.36309  -2.746 0.006030 **
## YEAR2006            -2.38361    0.38238  -6.234 4.56e-10 ***
## YEAR2007            -2.37348    0.35600  -6.667 2.61e-11 ***
## YEAR2008            -1.51180    0.36013  -4.198 2.69e-05 ***
## YEAR2009            -2.45809    0.35965  -6.835 8.22e-12 ***
## YEAR2010            -2.27217    0.37385  -6.078 1.22e-09 ***
## YEAR2011            -2.04755    0.35067  -5.839 5.25e-09 ***
## YEAR2012            -2.94528    0.35305  -8.342  < 2e-16 ***
## YEAR2013            -2.23211    0.34949  -6.387 1.69e-10 ***
## YEAR2014            -1.80692    0.34771  -5.197 2.03e-07 ***
## YEAR2015            -2.12486    0.34861  -6.095 1.09e-09 ***
## YEAR2016            -1.57068    0.36545  -4.298 1.72e-05 ***
## YEAR2017            -1.94544    0.34580  -5.626 1.85e-08 ***
## YEAR2018            -1.74782    0.35894  -4.869 1.12e-06 ***
## YEAR2019            -1.33620    0.35495  -3.764 0.000167 ***
## YEAR2020            -1.22293    0.56843  -2.151 0.031442 *
## YEAR2021            -1.59739    0.35878  -4.452 8.49e-06 ***
## MONTH8              -0.18163    0.12270  -1.480 0.138777
## MONTH9              -0.33787    0.15891  -2.126 0.033491 *
## AREALittle Neck Bay  0.06039    0.12907   0.468 0.639845
## AREAManhassett Bay   0.26684    0.12886   2.071 0.038376 *
## STEMP               -0.07112    0.07125  -0.998 0.318163
## SDO                 -0.08115    0.05722  -1.418 0.156096
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -3.2453     0.4820  -6.733 1.67e-11 ***
## MONTH8       -1.5474     0.5943  -2.604 0.009219 **
## MONTH9       -5.4638     1.5996  -3.416 0.000636 ***
## STEMP        -2.3600     0.3638  -6.486 8.79e-11 ***
## SDO          -1.4287     0.3195  -4.472 7.75e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tmbZI.resids(ZINB,y.dat=dat$CPUE)
```

```
## df.resid
## 2.121708
```

Model 3: ZANB (aka hurdle model)

```
ZANB <- glmmTMB(CPUE ~  YEAR + MONTH + AREA + STEMP + SDO, ziformula = ~
YEAR + MONTH + AREA + STEMP + SDO, data = dat,family = truncated_nbinom2(link
= "log"))
summary(ZANB)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
## Zero inflation:        ~YEAR + MONTH + AREA + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12413.5  12857.7  -6125.7  12251.5     1700
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.047
##
## Conditional model:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)           4.38701    0.55700   7.876 3.38e-15 ***
## YEAR1988             -2.54041    0.65682  -3.868 0.000110 ***
## YEAR1989             -1.23706    0.65963  -1.875 0.060740 .
## YEAR1990             -3.22338    0.62650  -5.145 2.67e-07 ***
## YEAR1991             -2.81219    0.53356  -5.271 1.36e-07 ***
## YEAR1992             -2.53645    0.62960  -4.029 5.61e-05 ***
## YEAR1993             -3.93262    0.55761  -7.053 1.76e-12 ***
## YEAR1994             -4.04493    0.59672  -6.779 1.21e-11 ***
## YEAR1995             -3.25620    0.70335  -4.630 3.66e-06 ***
## YEAR1996             -3.33171    1.34192  -2.483 0.013036 *
## YEAR1998             -3.12798    0.68182  -4.588 4.48e-06 ***
## YEAR1999             -3.74749    0.68112  -5.502 3.76e-08 ***
## YEAR2000             -3.50215    0.58911  -5.945 2.77e-09 ***
## YEAR2001             -1.75175    0.55706  -3.145 0.001663 **
## YEAR2002             -2.74159    0.59133  -4.636 3.55e-06 ***
## YEAR2003             -3.11854    0.56036  -5.565 2.62e-08 ***
## YEAR2004             -2.41858    0.59596  -4.058 4.94e-05 ***
## YEAR2005             -0.97267    0.55681  -1.747 0.080663 .
## YEAR2006             -2.58296    0.59159  -4.366 1.26e-05 ***
## YEAR2007             -2.64605    0.54974  -4.813 1.48e-06 ***
## YEAR2008             -1.96904    0.55080  -3.575 0.000350 ***
## YEAR2009             -2.96249    0.53095  -5.580 2.41e-08 ***
## YEAR2010             -2.74716    0.55247  -4.973 6.61e-07 ***
## YEAR2011             -2.50546    0.52224  -4.798 1.61e-06 ***
## YEAR2012             -3.59595    0.50991  -7.052 1.76e-12 ***
## YEAR2013             -2.73697    0.51801  -5.284 1.27e-07 ***
## YEAR2014             -2.26436    0.52282  -4.331 1.48e-05 ***
## YEAR2015             -2.35454    0.54179  -4.346 1.39e-05 ***
## YEAR2016             -1.28924    0.59802  -2.156 0.031094 *
## YEAR2017             -2.15899    0.53255  -4.054 5.03e-05 ***
## YEAR2018             -2.01920    0.54936  -3.676 0.000237 ***
```

```
## YEAR2019            -1.68346    0.53254  -3.161 0.001571 **
## YEAR2020            -1.16255    0.90177  -1.289 0.197336
## YEAR2021            -1.78221    0.54624  -3.263 0.001104 **
## MONTH8              -0.21020    0.17376  -1.210 0.226406
## MONTH9              -0.55484    0.23038  -2.408 0.016025 *
## AREALittle Neck Bay  0.03850    0.18907   0.204 0.838633
## AREAManhassett Bay   0.49405    0.18887   2.616 0.008903 **
## STEMP               -0.21019    0.09619  -2.185 0.028885 *
## SDO                 -0.11134    0.08051  -1.383 0.166663
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)         -1.20284    0.32824  -3.664 0.000248 ***
## YEAR1988            -0.13162    0.58390  -0.225 0.821656
## YEAR1989            -0.88406    0.56758  -1.558 0.119334
## YEAR1990            -0.11598    0.50370  -0.230 0.817893
## YEAR1991            -2.58870    1.06075  -2.440 0.014669 *
## YEAR1992            -0.74306    0.63293  -1.174 0.240397
## YEAR1993             0.17986    0.50391   0.357 0.721136
## YEAR1994             1.08980    0.48048   2.268 0.023321 *
## YEAR1995             0.85394    0.51226   1.667 0.095514 .
## YEAR1996             0.52966    0.93374   0.567 0.570547
## YEAR1998             1.12028    0.49559   2.261 0.023790 *
## YEAR1999             1.28019    0.50895   2.515 0.011892 *
## YEAR2000             0.81047    0.42847   1.892 0.058551 .
## YEAR2001            -0.42223    0.52288  -0.808 0.419375
## YEAR2002             0.68758    0.45150   1.523 0.127791
## YEAR2003             1.13015    0.40526   2.789 0.005292 **
## YEAR2004             0.69143    0.42272   1.636 0.101906
## YEAR2005             0.36939    0.43594   0.847 0.396805
## YEAR2006             0.87267    0.43061   2.027 0.042704 *
## YEAR2007             0.96362    0.41221   2.338 0.019404 *
## YEAR2008             0.04752    0.41870   0.113 0.909645
## YEAR2009            -0.08120    0.43802  -0.185 0.852936
## YEAR2010            -0.24299    0.49088  -0.495 0.620600
## YEAR2011            -0.09450    0.44369  -0.213 0.831346
## YEAR2012            -0.34913    0.47469  -0.735 0.462043
## YEAR2013            -0.59871    0.47096  -1.271 0.203633
## YEAR2014            -0.41626    0.45187  -0.921 0.356950
## YEAR2015             0.91435    0.40295   2.269 0.023259 *
## YEAR2016             1.31968    0.40868   3.229 0.001242 **
## YEAR2017             0.25801    0.41902   0.616 0.538061
## YEAR2018             0.36034    0.43687   0.825 0.409472
## YEAR2019            -0.72482    0.49616  -1.461 0.144054
## YEAR2020             0.17755    0.65577   0.271 0.786582
## YEAR2021             0.21713    0.42917   0.506 0.612906
## MONTH8              -0.27876    0.14105  -1.976 0.048125 *
## MONTH9              -0.74444    0.15756  -4.725 2.30e-06 ***
```

```
## AREALittle Neck Bay  0.16676    0.15936   1.046 0.295378
## AREAManhassett Bay   0.47372    0.14902   3.179 0.001478 **
## STEMP               -0.51293    0.08127  -6.311 2.77e-10 ***
## SDO                 -0.21073    0.06753  -3.120 0.001806 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Acceptable standard errors

tmbZI.resids(ZANB,dat$CPUE)

## df.resid
## 1.621377

# Dispersion looks ok-ish, or at least the best of the three model forms
```

Preliminary model comparisons

```
AICtab(tmb1.NB, ZANB, ZINB) # ZANB wins by a lot.

##          dAIC  df
## ZANB       0.0 81
## ZINB     201.9 46
## tmb1.NB  268.4 41
```

Step 3: Factor selection (by model)

Negative Binomial

```
NB1 <- glmmTMB(CPUE ~ YEAR + MONTH + AREA + STEMP + SDO, data = dat, family =
nbinom2)

summary(NB1)

##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12681.8  12906.7  -6299.9  12599.8     1740
##
##
## Dispersion parameter for nbinom2 family (): 0.263
##
## Conditional model:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          5.220198   0.272042  19.189  < 2e-16 ***
## YEAR1988            -2.035414   0.469556  -4.335 1.46e-05 ***
## YEAR1989            -0.818436   0.447472  -1.829 0.067396 .
## YEAR1990            -2.482592   0.439094  -5.654 1.57e-08 ***
## YEAR1991            -2.127448   0.388832  -5.471 4.47e-08 ***
## YEAR1992            -1.934689   0.456271  -4.240 2.23e-05 ***
## YEAR1993            -3.316592   0.407549  -8.138 4.02e-16 ***
```

```
## YEAR1994             -3.709637   0.427389  -8.680  < 2e-16 ***
## YEAR1995             -2.997083   0.469498  -6.384 1.73e-10 ***
## YEAR1996             -3.143413   0.853545  -3.683 0.000231 ***
## YEAR1998             -3.414891   0.455802  -7.492 6.78e-14 ***
## YEAR1999             -3.400056   0.456238  -7.452 9.17e-14 ***
## YEAR2000             -3.199663   0.391894  -8.165 3.22e-16 ***
## YEAR2001             -1.296715   0.386453  -3.355 0.000792 ***
## YEAR2002             -2.735037   0.399241  -6.851 7.35e-12 ***
## YEAR2003             -3.009040   0.359867  -8.362  < 2e-16 ***
## YEAR2004             -2.517191   0.379049  -6.641 3.12e-11 ***
## YEAR2005             -1.290976   0.367312  -3.515 0.000440 ***
## YEAR2006             -2.510606   0.385817  -6.507 7.65e-11 ***
## YEAR2007             -2.292544   0.358689  -6.391 1.64e-10 ***
## YEAR2008             -1.465851   0.365633  -4.009 6.10e-05 ***
## YEAR2009             -2.495966   0.364725  -6.843 7.73e-12 ***
## YEAR2010             -2.326355   0.384959  -6.043 1.51e-09 ***
## YEAR2011             -1.950963   0.358649  -5.440 5.34e-08 ***
## YEAR2012             -2.968313   0.363299  -8.170 3.07e-16 ***
## YEAR2013             -2.245035   0.358826  -6.257 3.93e-10 ***
## YEAR2014             -1.771625   0.356732  -4.966 6.83e-07 ***
## YEAR2015             -2.181165   0.352387  -6.190 6.03e-10 ***
## YEAR2016             -1.794673   0.371543  -4.830 1.36e-06 ***
## YEAR2017             -1.971540   0.354843  -5.556 2.76e-08 ***
## YEAR2018             -1.894480   0.364989  -5.191 2.10e-07 ***
## YEAR2019             -1.385919   0.363913  -3.808 0.000140 ***
## YEAR2020             -1.388115   0.575498  -2.412 0.015864 *
## YEAR2021             -1.719852   0.362824  -4.740 2.14e-06 ***
## MONTH8               -0.031502   0.123501  -0.255 0.798666
## MONTH9                0.024094   0.158909   0.152 0.879487
## AREALittle Neck Bay -0.009538   0.131843  -0.072 0.942327
## AREAManhassett Bay    0.214005   0.132745   1.612 0.106930
## STEMP                 0.083051   0.073997   1.122 0.261712
## SDO                  -0.018421   0.058764  -0.313 0.753920
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(NB1,test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
##        Df   AIC     LRT Pr(>Chi)
## <none>    12682
## YEAR   33 12874 257.643   <2e-16 ***
## MONTH   2 12678   0.142   0.9314
## AREA    2 12681   3.239   0.1980
## STEMP   1 12681   1.270   0.2597
## SDO     1 12680   0.098   0.7543
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# drop MONTH first ('least' significant)
NB2 <- glmmTMB(CPUE ~ YEAR + AREA + STEMP + SDO, data = dat, family =
nbinom2)
summary(NB2)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + AREA + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12678.0  12891.9  -6300.0  12600.0     1742
##
##
## Dispersion parameter for nbinom2 family (): 0.263
##
## Conditional model:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)             5.22286    0.27090  19.280  < 2e-16 ***
## YEAR1988               -2.04481    0.46838  -4.366 1.27e-05 ***
## YEAR1989               -0.85526    0.43432  -1.969 0.048932 *
## YEAR1990               -2.50187    0.43566  -5.743 9.31e-09 ***
## YEAR1991               -2.13546    0.38757  -5.510 3.59e-08 ***
## YEAR1992               -1.96346    0.44735  -4.389 1.14e-05 ***
## YEAR1993               -3.32497    0.40607  -8.188 2.65e-16 ***
## YEAR1994               -3.71893    0.42025  -8.849  < 2e-16 ***
## YEAR1995               -3.01585    0.46527  -6.482 9.05e-11 ***
## YEAR1996               -3.14061    0.85318  -3.681 0.000232 ***
## YEAR1998               -3.39276    0.43170  -7.859 3.87e-15 ***
## YEAR1999               -3.41097    0.45428  -7.508 5.98e-14 ***
## YEAR2000               -3.21618    0.38650  -8.321  < 2e-16 ***
## YEAR2001               -1.31156    0.37946  -3.456 0.000548 ***
## YEAR2002               -2.72705    0.38705  -7.046 1.84e-12 ***
## YEAR2003               -3.02460    0.35527  -8.514  < 2e-16 ***
## YEAR2004               -2.51744    0.36859  -6.830 8.50e-12 ***
## YEAR2005               -1.28452    0.35041  -3.666 0.000247 ***
## YEAR2006               -2.52398    0.37720  -6.691 2.21e-11 ***
## YEAR2007               -2.30415    0.35703  -6.454 1.09e-10 ***
## YEAR2008               -1.48404    0.35976  -4.125 3.71e-05 ***
## YEAR2009               -2.49810    0.35984  -6.942 3.86e-12 ***
## YEAR2010               -2.32323    0.37938  -6.124 9.14e-10 ***
## YEAR2011               -1.96076    0.35718  -5.490 4.03e-08 ***
## YEAR2012               -2.97134    0.35933  -8.269  < 2e-16 ***
## YEAR2013               -2.24792    0.35563  -6.321 2.60e-10 ***
## YEAR2014               -1.77446    0.35260  -5.033 4.84e-07 ***
## YEAR2015               -2.17870    0.34992  -6.226 4.78e-10 ***
## YEAR2016               -1.77335    0.34725  -5.107 3.28e-07 ***
## YEAR2017               -1.96743    0.34600  -5.686 1.30e-08 ***
```

```
## YEAR2018            -1.89121    0.35980  -5.256 1.47e-07 ***
## YEAR2019            -1.39447    0.35592  -3.918 8.93e-05 ***
## YEAR2020            -1.37511    0.57108  -2.408 0.016043 *
## YEAR2021            -1.71214    0.35700  -4.796 1.62e-06 ***
## AREALittle Neck Bay -0.01057    0.13169  -0.080 0.936022
## AREAManhassett Bay   0.20772    0.13037   1.593 0.111095
## STEMP                0.07176    0.06252   1.148 0.251094
## SDO                 -0.01619    0.05752  -0.281 0.778347
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(NB2, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + AREA + STEMP + SDO
##        Df   AIC     LRT Pr(>Chi)
## <none>    12678
## YEAR   33 12874 261.572   <2e-16 ***
## AREA    2 12677   3.171   0.2049
## STEMP   1 12677   1.314   0.2516
## SDO     1 12676   0.079   0.7786
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# next drop SDO ('least' significant)
NB3 <- glmmTMB(CPUE ~ YEAR + AREA + STEMP, data = dat, family = nbinom2)
summary(NB3)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + AREA + STEMP
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12676.1  12884.5  -6300.0  12600.1     1743
##
##
## Dispersion parameter for nbinom2 family (): 0.263
##
## Conditional model:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)          5.22374    0.27095  19.279  < 2e-16 ***
## YEAR1988            -2.06630    0.46215  -4.471 7.78e-06 ***
## YEAR1989            -0.86590    0.43269  -2.001 0.045372 *
## YEAR1990            -2.50382    0.43561  -5.748 9.04e-09 ***
## YEAR1991            -2.13911    0.38737  -5.522 3.35e-08 ***
## YEAR1992            -1.96250    0.44735  -4.387 1.15e-05 ***
## YEAR1993            -3.32997    0.40574  -8.207 2.26e-16 ***
## YEAR1994            -3.72447    0.41983  -8.871  < 2e-16 ***
## YEAR1995            -3.00700    0.46427  -6.477 9.37e-11 ***
```

```
## YEAR1996            -3.12711    0.85185  -3.671 0.000242 ***
## YEAR1998            -3.39715    0.43141  -7.874 3.42e-15 ***
## YEAR1999            -3.39776    0.45188  -7.519 5.52e-14 ***
## YEAR2000            -3.20930    0.38593  -8.316  < 2e-16 ***
## YEAR2001            -1.30607    0.37917  -3.445 0.000572 ***
## YEAR2002            -2.71967    0.38622  -7.042 1.90e-12 ***
## YEAR2003            -3.02025    0.35493  -8.509  < 2e-16 ***
## YEAR2004            -2.52368    0.36789  -6.860 6.90e-12 ***
## YEAR2005            -1.28291    0.35040  -3.661 0.000251 ***
## YEAR2006            -2.52085    0.37708  -6.685 2.31e-11 ***
## YEAR2007            -2.29662    0.35611  -6.449 1.12e-10 ***
## YEAR2008            -1.47647    0.35883  -4.115 3.88e-05 ***
## YEAR2009            -2.48700    0.35772  -6.952 3.59e-12 ***
## YEAR2010            -2.31892    0.37908  -6.117 9.52e-10 ***
## YEAR2011            -1.94856    0.35461  -5.495 3.91e-08 ***
## YEAR2012            -2.97156    0.35945  -8.267  < 2e-16 ***
## YEAR2013            -2.25476    0.35474  -6.356 2.07e-10 ***
## YEAR2014            -1.76631    0.35148  -5.025 5.03e-07 ***
## YEAR2015            -2.17200    0.34916  -6.221 4.95e-10 ***
## YEAR2016            -1.76580    0.34621  -5.100 3.39e-07 ***
## YEAR2017            -1.95806    0.34446  -5.684 1.31e-08 ***
## YEAR2018            -1.89951    0.35849  -5.299 1.17e-07 ***
## YEAR2019            -1.38474    0.35428  -3.909 9.28e-05 ***
## YEAR2020            -1.37431    0.57114  -2.406 0.016116 *
## YEAR2021            -1.70439    0.35598  -4.788 1.69e-06 ***
## AREALittle Neck Bay -0.01742    0.12953  -0.134 0.893013
## AREAManhassett Bay   0.19963    0.12718   1.570 0.116480
## STEMP                0.07006    0.06233   1.124 0.261014
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(NB3, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + AREA + STEMP
##        Df   AIC     LRT Pr(>Chi)
## <none>    12676
## YEAR   33 12872 261.638   <2e-16 ***
## AREA    2 12675   3.092   0.2131
## STEMP   1 12675   1.262   0.2613
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# next drop STEMP ('least' significant)
NB4 <- glmmTMB(CPUE ~ YEAR + AREA, data = dat, family = nbinom2)
summary(NB4)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + AREA
```

```
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12675.3  12878.3  -6300.7  12601.3     1744
##
##
## Dispersion parameter for nbinom2 family (): 0.263
##
## Conditional model:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)          5.27002    0.26835  19.639  < 2e-16 ***
## YEAR1988            -2.14391    0.45645  -4.697 2.64e-06 ***
## YEAR1989            -0.94541    0.42694  -2.214 0.026800 *
## YEAR1990            -2.60317    0.42614  -6.109 1.00e-09 ***
## YEAR1991            -2.15434    0.38741  -5.561 2.68e-08 ***
## YEAR1992            -2.01208    0.44534  -4.518 6.24e-06 ***
## YEAR1993            -3.33328    0.40602  -8.210  < 2e-16 ***
## YEAR1994            -3.72081    0.42013  -8.856  < 2e-16 ***
## YEAR1995            -3.03357    0.46400  -6.538 6.24e-11 ***
## YEAR1996            -3.15112    0.85201  -3.698 0.000217 ***
## YEAR1998            -3.38156    0.43154  -7.836 4.65e-15 ***
## YEAR1999            -3.40236    0.45205  -7.527 5.21e-14 ***
## YEAR2000            -3.28200    0.38057  -8.624  < 2e-16 ***
## YEAR2001            -1.32360    0.37894  -3.493 0.000478 ***
## YEAR2002            -2.74580    0.38573  -7.118 1.09e-12 ***
## YEAR2003            -3.05647    0.35366  -8.642  < 2e-16 ***
## YEAR2004            -2.55641    0.36695  -6.967 3.25e-12 ***
## YEAR2005            -1.25642    0.35000  -3.590 0.000331 ***
## YEAR2006            -2.54140    0.37702  -6.741 1.58e-11 ***
## YEAR2007            -2.35223    0.35254  -6.672 2.52e-11 ***
## YEAR2008            -1.52644    0.35628  -4.284 1.83e-05 ***
## YEAR2009            -2.53188    0.35563  -7.119 1.08e-12 ***
## YEAR2010            -2.33930    0.37900  -6.172 6.73e-10 ***
## YEAR2011            -1.97318    0.35396  -5.575 2.48e-08 ***
## YEAR2012            -2.98081    0.35945  -8.293  < 2e-16 ***
## YEAR2013            -2.30940    0.35125  -6.575 4.87e-11 ***
## YEAR2014            -1.81953    0.34841  -5.222 1.77e-07 ***
## YEAR2015            -2.18866    0.34916  -6.268 3.65e-10 ***
## YEAR2016            -1.79035    0.34558  -5.181 2.21e-07 ***
## YEAR2017            -2.00323    0.34223  -5.854 4.81e-09 ***
## YEAR2018            -1.88558    0.35855  -5.259 1.45e-07 ***
## YEAR2019            -1.40855    0.35391  -3.980 6.89e-05 ***
## YEAR2020            -1.40355    0.57078  -2.459 0.013932 *
## YEAR2021            -1.72737    0.35577  -4.855 1.20e-06 ***
## AREALittle Neck Bay -0.05781    0.12477  -0.463 0.643131
## AREAManhassett Bay   0.17720    0.12574   1.409 0.158758
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(NB4, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + AREA
##        Df   AIC     LRT Pr(>Chi)
## <none>    12675
## YEAR   33 12878 268.495   <2e-16 ***
## AREA    2 12674   3.026   0.2202
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# next drop AREA ('least' significant)
NB5 <- glmmTMB(CPUE ~ YEAR , data = dat, family = nbinom2)
summary(NB5)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12674.4  12866.3  -6302.2  12604.4     1746
##
##
## Dispersion parameter for nbinom2 family (): 0.262
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   5.2368     0.2565  20.417  < 2e-16 ***
## YEAR1988     -2.1275     0.4567  -4.658 3.19e-06 ***
## YEAR1989     -0.8895     0.4262  -2.087 0.036892 *
## YEAR1990     -2.5229     0.4241  -5.949 2.69e-09 ***
## YEAR1991     -2.0767     0.3844  -5.402 6.58e-08 ***
## YEAR1992     -1.9934     0.4456  -4.474 7.68e-06 ***
## YEAR1993     -3.2840     0.4027  -8.154 3.51e-16 ***
## YEAR1994     -3.6594     0.4141  -8.836  < 2e-16 ***
## YEAR1995     -2.9681     0.4591  -6.465 1.01e-10 ***
## YEAR1996     -3.1166     0.8490  -3.671 0.000242 ***
## YEAR1998     -3.3208     0.4268  -7.780 7.25e-15 ***
## YEAR1999     -3.3311     0.4498  -7.406 1.30e-13 ***
## YEAR2000     -3.2410     0.3745  -8.654  < 2e-16 ***
## YEAR2001     -1.1773     0.3697  -3.185 0.001449 **
## YEAR2002     -2.7103     0.3810  -7.113 1.14e-12 ***
## YEAR2003     -3.0413     0.3526  -8.625  < 2e-16 ***
## YEAR2004     -2.4816     0.3641  -6.815 9.40e-12 ***
## YEAR2005     -1.2347     0.3471  -3.557 0.000375 ***
## YEAR2006     -2.4607     0.3726  -6.604 4.01e-11 ***
## YEAR2007     -2.3331     0.3526  -6.616 3.69e-11 ***
## YEAR2008     -1.3688     0.3450  -3.968 7.25e-05 ***
## YEAR2009     -2.4539     0.3528  -6.956 3.51e-12 ***
## YEAR2010     -2.2298     0.3723  -5.989 2.11e-09 ***
```

```
## YEAR2011     -1.9035     0.3522  -5.404 6.51e-08 ***
## YEAR2012     -2.8929     0.3561  -8.123 4.54e-16 ***
## YEAR2013     -2.2446     0.3466  -6.475 9.46e-11 ***
## YEAR2014     -1.7164     0.3430  -5.004 5.63e-07 ***
## YEAR2015     -2.0610     0.3413  -6.039 1.55e-09 ***
## YEAR2016     -1.7668     0.3452  -5.118 3.08e-07 ***
## YEAR2017     -1.9772     0.3422  -5.778 7.57e-09 ***
## YEAR2018     -1.7434     0.3485  -5.003 5.66e-07 ***
## YEAR2019     -1.3328     0.3495  -3.814 0.000137 ***
## YEAR2020     -1.3559     0.5668  -2.392 0.016735 *
## YEAR2021     -1.6764     0.3508  -4.778 1.77e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(NB5, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR
##        Df   AIC    LRT  Pr(>Chi)
## <none>    12674
## YEAR   33 12880 271.79 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
disp3(NB5)
```

```
## df.resid
## 2.232711
```

```
# Let's compare the models using AIC to see if adding more factors
# improves how much deviance is explained.

AICtab(NB1, NB2, NB3, NB4, NB5)
```

```
##     dAIC df
## NB5  0.0 35
## NB4  1.0 37
## NB3  1.7 38
## NB2  3.6 39
## NB1  7.5 41
```

```
lrtest(NB5, NB4, NB3, NB2, NB1)
```

```
## Likelihood ratio test
##
## Model 1: CPUE ~ YEAR
## Model 2: CPUE ~ YEAR + AREA
## Model 3: CPUE ~ YEAR + AREA + STEMP
## Model 4: CPUE ~ YEAR + AREA + STEMP + SDO
## Model 5: CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
```

```
##   #Df  LogLik Df  Chisq Pr(>Chisq)
## 1  35 -6302.2
## 2  37 -6300.7  2 3.0262     0.2202
## 3  38 -6300.0  1 1.2619     0.2613
## 4  39 -6300.0  1 0.0791     0.7786
## 5  41 -6299.9  2 0.1421     0.9314
```

```
sapply(list(NB5,NB4,NB3,NB2,NB1),disp3)
```

```
## df.resid df.resid df.resid df.resid df.resid
## 2.232711 2.218173 2.255218 2.247060 2.253395
```

ZINB

```
ZINB1 = glmmTMB(CPUE ~ YEAR + MONTH + AREA + STEMP + SDO, ziformula = ~ MONTH
+ STEMP + SDO, data = dat, family=nbinom2)
summary(ZINB1)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
## Zero inflation:        ~MONTH + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12615.3  12867.6  -6261.7  12523.3     1735
##
##
## Dispersion parameter for nbinom2 family (): 0.294
##
## Conditional model:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)           5.34922    0.26861  19.914  < 2e-16 ***
## YEAR1988             -2.13011    0.45218  -4.711 2.47e-06 ***
## YEAR1989             -0.83439    0.44563  -1.872 0.061152 .
## YEAR1990             -2.56728    0.43964  -5.840 5.24e-09 ***
## YEAR1991             -2.16555    0.37664  -5.750 8.95e-09 ***
## YEAR1992             -1.98619    0.44530  -4.460 8.18e-06 ***
## YEAR1993             -3.28922    0.39623  -8.301  < 2e-16 ***
## YEAR1994             -3.56767    0.41345  -8.629  < 2e-16 ***
## YEAR1995             -2.94157    0.46683  -6.301 2.95e-10 ***
## YEAR1996             -2.65168    1.08146  -2.452 0.014208 *
## YEAR1998             -2.97649    0.46016  -6.468 9.91e-11 ***
## YEAR1999             -3.41597    0.44715  -7.639 2.18e-14 ***
## YEAR2000             -3.12468    0.39050  -8.002 1.23e-15 ***
## YEAR2001             -1.31116    0.37380  -3.508 0.000452 ***
## YEAR2002             -2.51121    0.39608  -6.340 2.30e-10 ***
## YEAR2003             -2.94118    0.35959  -8.179 2.85e-16 ***
## YEAR2004             -2.19958    0.39172  -5.615 1.96e-08 ***
## YEAR2005             -0.99712    0.36309  -2.746 0.006030 **
## YEAR2006             -2.38361    0.38238  -6.234 4.56e-10 ***
## YEAR2007             -2.37348    0.35600  -6.667 2.61e-11 ***
## YEAR2008             -1.51180    0.36013  -4.198 2.69e-05 ***
```

```
## YEAR2009           -2.45809    0.35965  -6.835 8.22e-12 ***
## YEAR2010           -2.27217    0.37385  -6.078 1.22e-09 ***
## YEAR2011           -2.04755    0.35067  -5.839 5.25e-09 ***
## YEAR2012           -2.94528    0.35305  -8.342  < 2e-16 ***
## YEAR2013           -2.23211    0.34949  -6.387 1.69e-10 ***
## YEAR2014           -1.80692    0.34771  -5.197 2.03e-07 ***
## YEAR2015           -2.12486    0.34861  -6.095 1.09e-09 ***
## YEAR2016           -1.57068    0.36545  -4.298 1.72e-05 ***
## YEAR2017           -1.94544    0.34580  -5.626 1.85e-08 ***
## YEAR2018           -1.74782    0.35894  -4.869 1.12e-06 ***
## YEAR2019           -1.33620    0.35495  -3.764 0.000167 ***
## YEAR2020           -1.22293    0.56843  -2.151 0.031442 *
## YEAR2021           -1.59739    0.35878  -4.452 8.49e-06 ***
## MONTH8             -0.18163    0.12270  -1.480 0.138777
## MONTH9             -0.33787    0.15891  -2.126 0.033491 *
## AREALittle Neck Bay  0.06039    0.12907   0.468 0.639845
## AREAManhassett Bay   0.26684    0.12886   2.071 0.038376 *
## STEMP              -0.07112    0.07125  -0.998 0.318163
## SDO                -0.08115    0.05722  -1.418 0.156096
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -3.2453     0.4820  -6.733 1.67e-11 ***
## MONTH8       -1.5474     0.5943  -2.604 0.009219 **
## MONTH9       -5.4638     1.5996  -3.416 0.000636 ***
## STEMP        -2.3600     0.3638  -6.486 8.79e-11 ***
## SDO          -1.4287     0.3195  -4.472 7.75e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZINB1, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
##        Df   AIC     LRT Pr(>Chi)
## <none>    12615
## YEAR   33 12809 259.423  < 2e-16 ***
## MONTH   2 12616   4.868  0.08766 .
## AREA    2 12616   4.475  0.10671
## STEMP   1 12614   0.986  0.32069
## SDO     1 12615   1.988  0.15859
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tmbZI.resids(ZINB1,y.dat=dat$CPUE)
```

```
## df.resid
## 2.121708

# remove STEMP from conditional model (least sig):
ZINB2 = glmmTMB(CPUE ~ YEAR + MONTH + AREA + SDO, ziformula = ~ MONTH + STEMP
+ SDO, data = dat, family=nbinom2)
summary(ZINB2)

##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + AREA + SDO
## Zero inflation:        ~MONTH + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12614.3  12861.1  -6262.2  12524.3     1736
##
##
## Dispersion parameter for nbinom2 family (): 0.293
##
## Conditional model:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)           5.30546    0.26504  20.017  < 2e-16 ***
## YEAR1988             -2.06373    0.44826  -4.604 4.15e-06 ***
## YEAR1989             -0.75679    0.43931  -1.723 0.084945 .
## YEAR1990             -2.48839    0.43315  -5.745 9.20e-09 ***
## YEAR1991             -2.15558    0.37682  -5.721 1.06e-08 ***
## YEAR1992             -1.93898    0.44352  -4.372 1.23e-05 ***
## YEAR1993             -3.29067    0.39643  -8.301  < 2e-16 ***
## YEAR1994             -3.59411    0.41314  -8.700  < 2e-16 ***
## YEAR1995             -2.92664    0.46667  -6.271 3.58e-10 ***
## YEAR1996             -2.63806    1.08174  -2.439 0.014739 *
## YEAR1998             -3.06824    0.45039  -6.812 9.59e-12 ***
## YEAR1999             -3.41993    0.44740  -7.644 2.11e-14 ***
## YEAR2000             -3.08300    0.38874  -7.931 2.18e-15 ***
## YEAR2001             -1.30954    0.37432  -3.498 0.000468 ***
## YEAR2002             -2.54310    0.39506  -6.437 1.22e-10 ***
## YEAR2003             -2.92109    0.35909  -8.135 4.13e-16 ***
## YEAR2004             -2.21496    0.39102  -5.665 1.47e-08 ***
## YEAR2005             -1.06626    0.35668  -2.989 0.002795 **
## YEAR2006             -2.38524    0.38260  -6.234 4.54e-10 ***
## YEAR2007             -2.33445    0.35435  -6.588 4.46e-11 ***
## YEAR2008             -1.47600    0.35893  -4.112 3.92e-05 ***
## YEAR2009             -2.45404    0.35984  -6.820 9.11e-12 ***
## YEAR2010             -2.28305    0.37436  -6.099 1.07e-09 ***
## YEAR2011             -2.03354    0.35089  -5.795 6.82e-09 ***
## YEAR2012             -2.95731    0.35317  -8.374  < 2e-16 ***
## YEAR2013             -2.20658    0.34923  -6.318 2.64e-10 ***
## YEAR2014             -1.79203    0.34766  -5.155 2.54e-07 ***
## YEAR2015             -2.13689    0.34867  -6.129 8.86e-10 ***
## YEAR2016             -1.62077    0.36239  -4.472 7.73e-06 ***
```

```
## YEAR2017            -1.95083    0.34595  -5.639 1.71e-08 ***
## YEAR2018            -1.78803    0.35677  -5.012 5.39e-07 ***
## YEAR2019            -1.34054    0.35537  -3.772 0.000162 ***
## YEAR2020            -1.24183    0.56847  -2.185 0.028924 *
## YEAR2021            -1.61760    0.35822  -4.516 6.31e-06 ***
## MONTH8              -0.18013    0.12301  -1.464 0.143089
## MONTH9              -0.25919    0.13848  -1.872 0.061249 .
## AREALittle Neck Bay  0.09088    0.12529   0.725 0.468242
## AREAManhassett Bay   0.29507    0.12612   2.340 0.019301 *
## SDO                 -0.08416    0.05728  -1.469 0.141747
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -3.3026     0.4922  -6.710 1.95e-11 ***
## MONTH8       -1.5500     0.6057  -2.559 0.010492 *
## MONTH9       -5.5317     1.6248  -3.405 0.000663 ***
## STEMP        -2.3724     0.3696  -6.419 1.37e-10 ***
## SDO          -1.4552     0.3270  -4.450 8.59e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZINB2, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + AREA + SDO
##        Df   AIC     LRT Pr(>Chi)
## <none>    12614
## YEAR   33 12807 258.496  < 2e-16 ***
## MONTH   2 12614   3.891  0.14295
## AREA    2 12616   5.564  0.06193 .
## SDO     1 12614   2.132  0.14424
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tmbZI.resids(ZINB2,y.dat=dat$CPUE)
```

```
## df.resid
## 2.165834
```

```
# remove MONTH from conditional model (least sig):
ZINB3 = glmmTMB(CPUE ~ YEAR + AREA + SDO, ziformula = ~ MONTH + STEMP + SDO,
data = dat, family=nbinom2)
summary(ZINB3)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + AREA + SDO
## Zero inflation:        ~MONTH + STEMP + SDO
```

```
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12614.2  12850.1  -6264.1  12528.2     1738
##
##
## Dispersion parameter for nbinom2 family (): 0.291
##
## Conditional model:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)          5.28671    0.26571  19.897  < 2e-16 ***
## YEAR1988            -2.15990    0.44654  -4.837 1.32e-06 ***
## YEAR1989            -0.90039    0.42958  -2.096 0.036086 *
## YEAR1990            -2.59221    0.42996  -6.029 1.65e-09 ***
## YEAR1991            -2.21302    0.37622  -5.882 4.05e-09 ***
## YEAR1992            -2.06070    0.43679  -4.718 2.38e-06 ***
## YEAR1993            -3.35003    0.39568  -8.466  < 2e-16 ***
## YEAR1994            -3.72816    0.40806  -9.136  < 2e-16 ***
## YEAR1995            -3.02425    0.46273  -6.536 6.34e-11 ***
## YEAR1996            -2.60092    1.08324  -2.401 0.016348 *
## YEAR1998            -3.30920    0.43229  -7.655 1.93e-14 ***
## YEAR1999            -3.48139    0.44574  -7.810 5.70e-15 ***
## YEAR2000            -3.21717    0.38309  -8.398  < 2e-16 ***
## YEAR2001            -1.43022    0.36840  -3.882 0.000103 ***
## YEAR2002            -2.73334    0.38356  -7.126 1.03e-12 ***
## YEAR2003            -3.02648    0.35457  -8.536  < 2e-16 ***
## YEAR2004            -2.41327    0.37745  -6.394 1.62e-10 ***
## YEAR2005            -1.24724    0.34575  -3.607 0.000309 ***
## YEAR2006            -2.53398    0.37488  -6.759 1.39e-11 ***
## YEAR2007            -2.35371    0.35474  -6.635 3.24e-11 ***
## YEAR2008            -1.58552    0.35424  -4.476 7.61e-06 ***
## YEAR2009            -2.59095    0.35397  -7.320 2.49e-13 ***
## YEAR2010            -2.41106    0.36986  -6.519 7.09e-11 ***
## YEAR2011            -2.08327    0.35011  -5.950 2.68e-09 ***
## YEAR2012            -3.05812    0.35027  -8.731  < 2e-16 ***
## YEAR2013            -2.33145    0.34364  -6.784 1.17e-11 ***
## YEAR2014            -1.91873    0.34224  -5.606 2.07e-08 ***
## YEAR2015            -2.22397    0.34718  -6.406 1.50e-10 ***
## YEAR2016            -1.84581    0.34168  -5.402 6.58e-08 ***
## YEAR2017            -2.11477    0.33616  -6.291 3.15e-10 ***
## YEAR2018            -1.89719    0.35387  -5.361 8.26e-08 ***
## YEAR2019            -1.48534    0.34824  -4.265 2.00e-05 ***
## YEAR2020            -1.36339    0.56549  -2.411 0.015910 *
## YEAR2021            -1.74118    0.35276  -4.936 7.98e-07 ***
## AREALittle Neck Bay  0.07086    0.12497   0.567 0.570706
## AREAManhassett Bay   0.30614    0.12666   2.417 0.015651 *
## SDO                 -0.05833    0.05563  -1.049 0.294372
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept)  -3.4303     0.5142  -6.671 2.54e-11 ***
## MONTH8       -1.4749     0.6083  -2.424  0.01533 *  
## MONTH9       -5.4320     1.6376  -3.317  0.00091 ***
## STEMP        -2.4109     0.3815  -6.320 2.62e-10 ***
## SDO          -1.4904     0.3407  -4.375 1.22e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZINB3, test="Chi")
```

```
## Single term deletions
## 
## Model:
## CPUE ~ YEAR + AREA + SDO
##        Df   AIC     LRT Pr(>Chi)    
## <none>    12614                     
## YEAR   33 12820 271.715  < 2e-16 ***
## AREA    2 12616   6.035  0.04893 *  
## SDO     1 12613   1.089  0.29679    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tmbZI.resids(ZINB3,y.dat=dat$CPUE)
```

```
## df.resid 
##  2.20197
```

```
# remove SDO from conditional model (least sig):
ZINB4 = glmmTMB(CPUE ~ YEAR + AREA, ziformula = ~ MONTH + STEMP + SDO, data =
dat, family=nbinom2)
summary(ZINB4)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + AREA
## Zero inflation:        ~MONTH + STEMP + SDO
## Data: dat
## 
##      AIC      BIC   logLik deviance df.resid 
##  12613.3  12843.7  -6264.7  12529.3     1739 
## 
## 
## Dispersion parameter for nbinom2 family (): 0.29 
## 
## Conditional model:
##                       Estimate Std. Error z value Pr(>|z|)    
## (Intercept)            5.28474    0.26602  19.866  < 2e-16 ***
## YEAR1988              -2.22984    0.44188  -5.046 4.51e-07 ***
## YEAR1989              -0.93143    0.42884  -2.172  0.02986 *  
## YEAR1990              -2.58963    0.43029  -6.018 1.76e-09 ***
```

```
## YEAR1991             -2.22387    0.37642  -5.908 3.46e-09 ***
## YEAR1992             -2.05180    0.43704  -4.695 2.67e-06 ***
## YEAR1993             -3.36617    0.39593  -8.502  < 2e-16 ***
## YEAR1994             -3.74910    0.40797  -9.190  < 2e-16 ***
## YEAR1995             -2.98996    0.46198  -6.472 9.67e-11 ***
## YEAR1996             -2.55038    1.08285  -2.355  0.01851 *
## YEAR1998             -3.32606    0.43244  -7.691 1.45e-14 ***
## YEAR1999             -3.43217    0.44369  -7.735 1.03e-14 ***
## YEAR2000             -3.18994    0.38258  -8.338  < 2e-16 ***
## YEAR2001             -1.41199    0.36893  -3.827  0.00013 ***
## YEAR2002             -2.70447    0.38295  -7.062 1.64e-12 ***
## YEAR2003             -3.00916    0.35436  -8.492  < 2e-16 ***
## YEAR2004             -2.43402    0.37699  -6.457 1.07e-10 ***
## YEAR2005             -1.24548    0.34607  -3.599  0.00032 ***
## YEAR2006             -2.52118    0.37494  -6.724 1.77e-11 ***
## YEAR2007             -2.32547    0.35400  -6.569 5.06e-11 ***
## YEAR2008             -1.55488    0.35337  -4.400 1.08e-05 ***
## YEAR2009             -2.54768    0.35193  -7.239 4.51e-13 ***
## YEAR2010             -2.39238    0.36974  -6.470 9.77e-11 ***
## YEAR2011             -2.03871    0.34788  -5.860 4.62e-09 ***
## YEAR2012             -3.06055    0.35081  -8.724  < 2e-16 ***
## YEAR2013             -2.34996    0.34349  -6.841 7.84e-12 ***
## YEAR2014             -1.88454    0.34099  -5.527 3.26e-08 ***
## YEAR2015             -2.19735    0.34662  -6.339 2.31e-10 ***
## YEAR2016             -1.81794    0.34082  -5.334 9.60e-08 ***
## YEAR2017             -2.08055    0.33492  -6.212 5.23e-10 ***
## YEAR2018             -1.92881    0.35245  -5.473 4.44e-08 ***
## YEAR2019             -1.44827    0.34670  -4.177 2.95e-05 ***
## YEAR2020             -1.35996    0.56571  -2.404  0.01622 *
## YEAR2021             -1.71112    0.35194  -4.862 1.16e-06 ***
## AREALittle Neck Bay  0.05013    0.12382   0.405  0.68558
## AREAManhassett Bay   0.27890    0.12408   2.248  0.02459 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -3.4485     0.5183  -6.653 2.86e-11 ***
## MONTH8       -1.5066     0.6264  -2.405 0.016160 *
## MONTH9       -5.5231     1.6596  -3.328 0.000874 ***
## STEMP        -2.4278     0.3839  -6.325 2.53e-10 ***
## SDO          -1.4680     0.3433  -4.276 1.90e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZINB4, test="Chi")
```

```
## Single term deletions
##
## Model:
```

```
## CPUE ~ YEAR + AREA
##        Df   AIC     LRT Pr(>Chi)
## <none>    12613
## YEAR   33 12818 270.638  < 2e-16 ***
## AREA    2 12615   5.274  0.07156 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tmbZI.resids(ZINB4,y.dat=dat$CPUE)
```

```
## df.resid
## 2.238909
```

```
# remove AREA from conditional model (least sig):
ZINB5 = glmmTMB(CPUE ~ YEAR, ziformula = ~ MONTH + STEMP + SDO, data = dat,
family=nbinom2)
summary(ZINB5)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR
## Zero inflation:        ~MONTH + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12614.6  12834.0  -6267.3  12534.6     1741
##
##
## Dispersion parameter for nbinom2 family (): 0.289
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   5.3187     0.2557  20.801  < 2e-16 ***
## YEAR1988     -2.2050     0.4428  -4.980 6.36e-07 ***
## YEAR1989     -0.8651     0.4286  -2.018 0.043549 *
## YEAR1990     -2.4826     0.4289  -5.789 7.10e-09 ***
## YEAR1991     -2.1585     0.3742  -5.769 7.99e-09 ***
## YEAR1992     -2.0322     0.4379  -4.641 3.47e-06 ***
## YEAR1993     -3.3535     0.3932  -8.528  < 2e-16 ***
## YEAR1994     -3.7381     0.4029  -9.277  < 2e-16 ***
## YEAR1995     -2.9716     0.4577  -6.492 8.46e-11 ***
## YEAR1996     -2.5905     1.0802  -2.398 0.016480 *
## YEAR1998     -3.3115     0.4280  -7.738 1.01e-14 ***
## YEAR1999     -3.3763     0.4420  -7.638 2.20e-14 ***
## YEAR2000     -3.2080     0.3772  -8.505  < 2e-16 ***
## YEAR2001     -1.2573     0.3606  -3.487 0.000489 ***
## YEAR2002     -2.7208     0.3789  -7.181 6.91e-13 ***
## YEAR2003     -3.0223     0.3535  -8.549  < 2e-16 ***
## YEAR2004     -2.3744     0.3745  -6.340 2.30e-10 ***
## YEAR2005     -1.2655     0.3436  -3.683 0.000231 ***
## YEAR2006     -2.4655     0.3713  -6.640 3.13e-11 ***
## YEAR2007     -2.2997     0.3545  -6.487 8.75e-11 ***
```

```
## YEAR2008     -1.3818     0.3431  -4.028 5.63e-05 ***
## YEAR2009     -2.4749     0.3498  -7.076 1.49e-12 ***
## YEAR2010     -2.3013     0.3640  -6.322 2.59e-10 ***
## YEAR2011     -1.9560     0.3467  -5.642 1.69e-08 ***
## YEAR2012     -2.9717     0.3482  -8.535  < 2e-16 ***
## YEAR2013     -2.3186     0.3396  -6.827 8.67e-12 ***
## YEAR2014     -1.7878     0.3364  -5.314 1.07e-07 ***
## YEAR2015     -2.0699     0.3396  -6.094 1.10e-09 ***
## YEAR2016     -1.8084     0.3408  -5.306 1.12e-07 ***
## YEAR2017     -2.0517     0.3354  -6.117 9.53e-10 ***
## YEAR2018     -1.7898     0.3435  -5.211 1.88e-07 ***
## YEAR2019     -1.3981     0.3430  -4.076 4.58e-05 ***
## YEAR2020     -1.3697     0.5622  -2.436 0.014837 *
## YEAR2021     -1.7051     0.3476  -4.906 9.30e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -3.5147     0.5318  -6.609 3.86e-11 ***
## MONTH8       -1.5212     0.6445  -2.360 0.018263 *
## MONTH9       -5.6015     1.7016  -3.292 0.000995 ***
## STEMP        -2.4508     0.3913  -6.263 3.77e-10 ***
## SDO          -1.4921     0.3521  -4.238 2.25e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

drop1(ZINB5, test="Chi")

## Single term deletions
##
## Model:
## CPUE ~ YEAR
##        Df   AIC    LRT  Pr(>Chi)
## <none>    12615
## YEAR   33 12826 276.91 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

tmbZI.resids(ZINB5,y.dat=dat$CPUE)

## df.resid
## 2.322094

# all other terms in conditional and zero-infl model sig, so retaining

# but also try a year-only model:
ZINB0 <-  glmmTMB(CPUE ~ YEAR, data = dat, family=nbinom2)#ziformula = ~
YEAR,
summary(ZINB0)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12674.4  12866.3  -6302.2  12604.4     1746
##
##
## Dispersion parameter for nbinom2 family (): 0.262
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   5.2368     0.2565  20.417  < 2e-16 ***
## YEAR1988     -2.1275     0.4567  -4.658 3.19e-06 ***
## YEAR1989     -0.8895     0.4262  -2.087 0.036892 *
## YEAR1990     -2.5229     0.4241  -5.949 2.69e-09 ***
## YEAR1991     -2.0767     0.3844  -5.402 6.58e-08 ***
## YEAR1992     -1.9934     0.4456  -4.474 7.68e-06 ***
## YEAR1993     -3.2840     0.4027  -8.154 3.51e-16 ***
## YEAR1994     -3.6594     0.4141  -8.836  < 2e-16 ***
## YEAR1995     -2.9681     0.4591  -6.465 1.01e-10 ***
## YEAR1996     -3.1166     0.8490  -3.671 0.000242 ***
## YEAR1998     -3.3208     0.4268  -7.780 7.25e-15 ***
## YEAR1999     -3.3311     0.4498  -7.406 1.30e-13 ***
## YEAR2000     -3.2410     0.3745  -8.654  < 2e-16 ***
## YEAR2001     -1.1773     0.3697  -3.185 0.001449 **
## YEAR2002     -2.7103     0.3810  -7.113 1.14e-12 ***
## YEAR2003     -3.0413     0.3526  -8.625  < 2e-16 ***
## YEAR2004     -2.4816     0.3641  -6.815 9.40e-12 ***
## YEAR2005     -1.2347     0.3471  -3.557 0.000375 ***
## YEAR2006     -2.4607     0.3726  -6.604 4.01e-11 ***
## YEAR2007     -2.3331     0.3526  -6.616 3.69e-11 ***
## YEAR2008     -1.3688     0.3450  -3.968 7.25e-05 ***
## YEAR2009     -2.4539     0.3528  -6.956 3.51e-12 ***
## YEAR2010     -2.2298     0.3723  -5.989 2.11e-09 ***
## YEAR2011     -1.9035     0.3522  -5.404 6.51e-08 ***
## YEAR2012     -2.8929     0.3561  -8.123 4.54e-16 ***
## YEAR2013     -2.2446     0.3466  -6.475 9.46e-11 ***
## YEAR2014     -1.7164     0.3430  -5.004 5.63e-07 ***
## YEAR2015     -2.0610     0.3413  -6.039 1.55e-09 ***
## YEAR2016     -1.7668     0.3452  -5.118 3.08e-07 ***
## YEAR2017     -1.9772     0.3422  -5.778 7.57e-09 ***
## YEAR2018     -1.7434     0.3485  -5.003 5.66e-07 ***
## YEAR2019     -1.3328     0.3495  -3.814 0.000137 ***
## YEAR2020     -1.3559     0.5668  -2.392 0.016735 *
## YEAR2021     -1.6764     0.3508  -4.778 1.77e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tmbZI.resids(ZINB0,y.dat=dat$CPUE)
```

```
##   df.resid
## 0.03931795
```

```
AICtab(ZINB1, ZINB2, ZINB3, ZINB4, ZINB5, ZINB0)
```

```
##       dAIC df
## ZINB4  0.0 42
## ZINB3  0.9 43
## ZINB2  1.0 45
## ZINB5  1.3 40
## ZINB1  2.0 46
## ZINB0 61.1 35
```

```
lrtest(ZINB0, ZINB5, ZINB4, ZINB3, ZINB2, ZINB1)
```

```
## Likelihood ratio test
##
## Model 1: CPUE ~ YEAR
## Model 2: CPUE ~ YEAR
## Model 3: CPUE ~ YEAR + AREA
## Model 4: CPUE ~ YEAR + AREA + SDO
## Model 5: CPUE ~ YEAR + MONTH + AREA + SDO
## Model 6: CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
##   #Df  LogLik Df   Chisq Pr(>Chisq)
## 1  35 -6302.2
## 2  40 -6267.3  5 69.7798  1.139e-13 ***
## 3  42 -6264.7  2  5.2744    0.07156 .
## 4  43 -6264.1  1  1.0885    0.29679
## 5  45 -6262.2  2  3.8905    0.14295
## 6  46 -6261.7  1  0.9861    0.32069
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#
sapply(list(ZINB4,ZINB3,ZINB2,ZINB5,ZINB1,ZINB0),tmbZI.resids,dat$CPUE)
```

```
##   df.resid   df.resid   df.resid   df.resid   df.resid   df.resid
## 2.23890942 2.20197005 2.16583361 2.32209364 2.12170838 0.03931795
```

```
tmbZI.resids(ZINB4,dat$CPUE)
```

```
## df.resid
## 2.238909
```

ZANB

```
ZANB1 <- glmmTMB(CPUE ~ YEAR + MONTH + AREA + STEMP + SDO, ziformula = ~ YEAR
+ MONTH + AREA + STEMP + SDO, data = dat,family = truncated_nbinom2(link =
"log"))
summary(ZANB1)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
```

```
## Zero inflation:        ~YEAR + MONTH + AREA + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12413.5  12857.7  -6125.7  12251.5     1700
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.047
##
## Conditional model:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)           4.38701    0.55700   7.876 3.38e-15 ***
## YEAR1988             -2.54041    0.65682  -3.868 0.000110 ***
## YEAR1989             -1.23706    0.65963  -1.875 0.060740 .
## YEAR1990             -3.22338    0.62650  -5.145 2.67e-07 ***
## YEAR1991             -2.81219    0.53356  -5.271 1.36e-07 ***
## YEAR1992             -2.53645    0.62960  -4.029 5.61e-05 ***
## YEAR1993             -3.93262    0.55761  -7.053 1.76e-12 ***
## YEAR1994             -4.04493    0.59672  -6.779 1.21e-11 ***
## YEAR1995             -3.25620    0.70335  -4.630 3.66e-06 ***
## YEAR1996             -3.33171    1.34192  -2.483 0.013036 *
## YEAR1998             -3.12798    0.68182  -4.588 4.48e-06 ***
## YEAR1999             -3.74749    0.68112  -5.502 3.76e-08 ***
## YEAR2000             -3.50215    0.58911  -5.945 2.77e-09 ***
## YEAR2001             -1.75175    0.55706  -3.145 0.001663 **
## YEAR2002             -2.74159    0.59133  -4.636 3.55e-06 ***
## YEAR2003             -3.11854    0.56036  -5.565 2.62e-08 ***
## YEAR2004             -2.41858    0.59596  -4.058 4.94e-05 ***
## YEAR2005             -0.97267    0.55681  -1.747 0.080663 .
## YEAR2006             -2.58296    0.59159  -4.366 1.26e-05 ***
## YEAR2007             -2.64605    0.54974  -4.813 1.48e-06 ***
## YEAR2008             -1.96904    0.55080  -3.575 0.000350 ***
## YEAR2009             -2.96249    0.53095  -5.580 2.41e-08 ***
## YEAR2010             -2.74716    0.55247  -4.973 6.61e-07 ***
## YEAR2011             -2.50546    0.52224  -4.798 1.61e-06 ***
## YEAR2012             -3.59595    0.50991  -7.052 1.76e-12 ***
## YEAR2013             -2.73697    0.51801  -5.284 1.27e-07 ***
## YEAR2014             -2.26436    0.52282  -4.331 1.48e-05 ***
## YEAR2015             -2.35454    0.54179  -4.346 1.39e-05 ***
## YEAR2016             -1.28924    0.59802  -2.156 0.031094 *
## YEAR2017             -2.15899    0.53255  -4.054 5.03e-05 ***
## YEAR2018             -2.01920    0.54936  -3.676 0.000237 ***
## YEAR2019             -1.68346    0.53254  -3.161 0.001571 **
## YEAR2020             -1.16255    0.90177  -1.289 0.197336
## YEAR2021             -1.78221    0.54624  -3.263 0.001104 **
## MONTH8               -0.21020    0.17376  -1.210 0.226406
## MONTH9               -0.55484    0.23038  -2.408 0.016025 *
## AREALittle Neck Bay   0.03850    0.18907   0.204 0.838633
## AREAManhassett Bay    0.49405    0.18887   2.616 0.008903 **
## STEMP                -0.21019    0.09619  -2.185 0.028885 *
```

```
## SDO                   -0.11134    0.08051  -1.383 0.166663
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -1.20284    0.32824  -3.664 0.000248 ***
## YEAR1988              -0.13162    0.58390  -0.225 0.821656
## YEAR1989              -0.88406    0.56758  -1.558 0.119334
## YEAR1990              -0.11598    0.50370  -0.230 0.817893
## YEAR1991              -2.58870    1.06075  -2.440 0.014669 *
## YEAR1992              -0.74306    0.63293  -1.174 0.240397
## YEAR1993               0.17986    0.50391   0.357 0.721136
## YEAR1994               1.08980    0.48048   2.268 0.023321 *
## YEAR1995               0.85394    0.51226   1.667 0.095514 .
## YEAR1996               0.52966    0.93374   0.567 0.570547
## YEAR1998               1.12028    0.49559   2.261 0.023790 *
## YEAR1999               1.28019    0.50895   2.515 0.011892 *
## YEAR2000               0.81047    0.42847   1.892 0.058551 .
## YEAR2001              -0.42223    0.52288  -0.808 0.419375
## YEAR2002               0.68758    0.45150   1.523 0.127791
## YEAR2003               1.13015    0.40526   2.789 0.005292 **
## YEAR2004               0.69143    0.42272   1.636 0.101906
## YEAR2005               0.36939    0.43594   0.847 0.396805
## YEAR2006               0.87267    0.43061   2.027 0.042704 *
## YEAR2007               0.96362    0.41221   2.338 0.019404 *
## YEAR2008               0.04752    0.41870   0.113 0.909645
## YEAR2009              -0.08120    0.43802  -0.185 0.852936
## YEAR2010              -0.24299    0.49088  -0.495 0.620600
## YEAR2011              -0.09450    0.44369  -0.213 0.831346
## YEAR2012              -0.34913    0.47469  -0.735 0.462043
## YEAR2013              -0.59871    0.47096  -1.271 0.203633
## YEAR2014              -0.41626    0.45187  -0.921 0.356950
## YEAR2015               0.91435    0.40295   2.269 0.023259 *
## YEAR2016               1.31968    0.40868   3.229 0.001242 **
## YEAR2017               0.25801    0.41902   0.616 0.538061
## YEAR2018               0.36034    0.43687   0.825 0.409472
## YEAR2019              -0.72482    0.49616  -1.461 0.144054
## YEAR2020               0.17755    0.65577   0.271 0.786582
## YEAR2021               0.21713    0.42917   0.506 0.612906
## MONTH8                -0.27876    0.14105  -1.976 0.048125 *
## MONTH9                -0.74444    0.15756  -4.725 2.30e-06 ***
## AREALittle Neck Bay    0.16676    0.15936   1.046 0.295378
## AREAManhassett Bay     0.47372    0.14902   3.179 0.001478 **
## STEMP                 -0.51293    0.08127  -6.311 2.77e-10 ***
## SDO                   -0.21073    0.06753  -3.120 0.001806 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Acceptable standard errors
drop1(ZANB1, test="Chi")

## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
##        Df   AIC     LRT Pr(>Chi)
## <none>    12414
## YEAR   33 12512 164.330  < 2e-16 ***
## MONTH   2 12415   5.719  0.05729 .
## AREA    2 12417   7.770  0.02055 *
## STEMP   1 12416   4.607  0.03185 *
## SDO     1 12413   1.883  0.16997
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# dispersion:
tmbZI.resids(ZANB1, y.dat=dat$CPUE)

## df.resid
## 1.621377

# remove SDO from conditional model
ZANB2 <- glmmTMB(CPUE ~ YEAR + MONTH + AREA + STEMP, ziformula = ~ YEAR +
MONTH + AREA + STEMP + SDO, data = dat,family = truncated_nbinom2(link =
"log"))
summary(ZANB2)

##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + AREA + STEMP
## Zero inflation:        ~YEAR + MONTH + AREA + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12413.3  12852.1  -6126.7  12253.3     1701
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.0459
##
## Conditional model:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)            4.37732    0.56470   7.752 9.08e-15 ***
## YEAR1988              -2.69359    0.65037  -4.142 3.45e-05 ***
## YEAR1989              -1.36022    0.65367  -2.081 0.037445 *
## YEAR1990              -3.27038    0.62644  -5.221 1.78e-07 ***
## YEAR1991              -2.86718    0.53346  -5.375 7.67e-08 ***
## YEAR1992              -2.57490    0.62911  -4.093 4.26e-05 ***
## YEAR1993              -3.97234    0.56061  -7.086 1.38e-12 ***
## YEAR1994              -4.12730    0.59609  -6.924 4.39e-12 ***
## YEAR1995              -3.23291    0.70627  -4.577 4.71e-06 ***
```

```
## YEAR1996            -3.24346    1.34317  -2.415 0.015744 *
## YEAR1998            -3.23571    0.67725  -4.778 1.77e-06 ***
## YEAR1999            -3.68032    0.68194  -5.397 6.78e-08 ***
## YEAR2000            -3.50829    0.59368  -5.909 3.43e-09 ***
## YEAR2001            -1.76674    0.56253  -3.141 0.001686 **
## YEAR2002            -2.74268    0.59361  -4.620 3.83e-06 ***
## YEAR2003            -3.12649    0.56266  -5.557 2.75e-08 ***
## YEAR2004            -2.52003    0.59089  -4.265 2.00e-05 ***
## YEAR2005            -1.00625    0.55932  -1.799 0.072009 .
## YEAR2006            -2.59944    0.59476  -4.371 1.24e-05 ***
## YEAR2007            -2.62015    0.55243  -4.743 2.11e-06 ***
## YEAR2008            -1.96034    0.55333  -3.543 0.000396 ***
## YEAR2009            -2.92793    0.53306  -5.493 3.96e-08 ***
## YEAR2010            -2.76001    0.55390  -4.983 6.27e-07 ***
## YEAR2011            -2.44819    0.52301  -4.681 2.86e-06 ***
## YEAR2012            -3.65317    0.51108  -7.148 8.81e-13 ***
## YEAR2013            -2.81549    0.51582  -5.458 4.81e-08 ***
## YEAR2014            -2.23704    0.52539  -4.258 2.06e-05 ***
## YEAR2015            -2.34196    0.54386  -4.306 1.66e-05 ***
## YEAR2016            -1.29431    0.59932  -2.160 0.030801 *
## YEAR2017            -2.12165    0.53641  -3.955 7.64e-05 ***
## YEAR2018            -2.12661    0.54378  -3.911 9.20e-05 ***
## YEAR2019            -1.66475    0.53453  -3.114 0.001843 **
## YEAR2020            -1.18438    0.90490  -1.309 0.190584
## YEAR2021            -1.75956    0.54866  -3.207 0.001341 **
## MONTH8              -0.15830    0.16983  -0.932 0.351290
## MONTH9              -0.50722    0.22748  -2.230 0.025766 *
## AREALittle Neck Bay -0.01627    0.18705  -0.087 0.930698
## AREAManhassett Bay   0.45002    0.18692   2.408 0.016058 *
## STEMP               -0.22000    0.09674  -2.274 0.022961 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)         -1.20283    0.32824  -3.664 0.000248 ***
## YEAR1988            -0.13172    0.58391  -0.226 0.821527
## YEAR1989            -0.88396    0.56758  -1.557 0.119366
## YEAR1990            -0.11597    0.50370  -0.230 0.817916
## YEAR1991            -2.58871    1.06076  -2.440 0.014670 *
## YEAR1992            -0.74301    0.63293  -1.174 0.240424
## YEAR1993             0.17987    0.50391   0.357 0.721125
## YEAR1994             1.08982    0.48048   2.268 0.023318 *
## YEAR1995             0.85394    0.51226   1.667 0.095515 .
## YEAR1996             0.52968    0.93374   0.567 0.570538
## YEAR1998             1.12028    0.49559   2.260 0.023791 *
## YEAR1999             1.28020    0.50896   2.515 0.011892 *
## YEAR2000             0.81044    0.42847   1.891 0.058559 .
## YEAR2001            -0.42217    0.52288  -0.807 0.419440
## YEAR2002             0.68759    0.45150   1.523 0.127784
```

```
## YEAR2003             1.13018    0.40526   2.789 0.005291 **
## YEAR2004             0.69146    0.42272   1.636 0.101889
## YEAR2005             0.36943    0.43594   0.847 0.396759
## YEAR2006             0.87265    0.43061   2.027 0.042710 *
## YEAR2007             0.96367    0.41221   2.338 0.019399 *
## YEAR2008             0.04754    0.41870   0.114 0.909601
## YEAR2009            -0.08119    0.43802  -0.185 0.852947
## YEAR2010            -0.24297    0.49088  -0.495 0.620619
## YEAR2011            -0.09448    0.44369  -0.213 0.831378
## YEAR2012            -0.34913    0.47469  -0.735 0.462046
## YEAR2013            -0.59871    0.47096  -1.271 0.203638
## YEAR2014            -0.41624    0.45187  -0.921 0.356966
## YEAR2015             0.91437    0.40295   2.269 0.023255 *
## YEAR2016             1.31970    0.40868   3.229 0.001241 **
## YEAR2017             0.25801    0.41902   0.616 0.538059
## YEAR2018             0.36032    0.43687   0.825 0.409494
## YEAR2019            -0.72480    0.49616  -1.461 0.144064
## YEAR2020             0.17759    0.65576   0.271 0.786536
## YEAR2021             0.21713    0.42917   0.506 0.612907
## MONTH8              -0.27877    0.14105  -1.976 0.048112 *
## MONTH9              -0.74444    0.15756  -4.725 2.30e-06 ***
## AREALittle Neck Bay  0.16675    0.15936   1.046 0.295406
## AREAManhassett Bay   0.47370    0.14902   3.179 0.001479 **
## STEMP               -0.51294    0.08127  -6.311 2.77e-10 ***
## SDO                 -0.21072    0.06753  -3.120 0.001806 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Acceptable standard errors
drop1(ZANB2, test="Chi")

## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + AREA + STEMP
##        Df   AIC     LRT Pr(>Chi)
## <none>    12413
## YEAR   33 12512 164.505  < 2e-16 ***
## MONTH   2 12414   4.838  0.08901 .
## AREA    2 12416   6.950  0.03097 *
## STEMP   1 12416   4.963  0.02590 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# dispersion:
tmbZI.resids(ZANB2, y.dat=dat$CPUE)

## df.resid
## 1.671789
```

```r
# remove MONTH from conditional model
ZANB3 <- glmmTMB(CPUE ~ YEAR + AREA + STEMP, ziformula = ~ YEAR + MONTH +
AREA + STEMP + SDO, data = dat,family = truncated_nbinom2(link = "log"))
summary(ZANB3)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + AREA + STEMP
## Zero inflation:        ~YEAR + MONTH + AREA + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12414.2  12842.0  -6129.1  12258.2     1703
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.0435
##
## Conditional model:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           4.271800   0.583699   7.318 2.51e-13 ***
## YEAR1988             -2.682231   0.657175  -4.081 4.48e-05 ***
## YEAR1989             -1.381976   0.642555  -2.151 0.031496 *
## YEAR1990             -3.298625   0.628734  -5.246 1.55e-07 ***
## YEAR1991             -2.926962   0.535211  -5.469 4.53e-08 ***
## YEAR1992             -2.623344   0.622468  -4.214 2.50e-05 ***
## YEAR1993             -4.045676   0.562681  -7.190 6.48e-13 ***
## YEAR1994             -4.313829   0.595235  -7.247 4.25e-13 ***
## YEAR1995             -3.329737   0.702842  -4.738 2.16e-06 ***
## YEAR1996             -3.221004   1.350399  -2.385 0.017069 *
## YEAR1998             -3.772328   0.633644  -5.953 2.63e-09 ***
## YEAR1999             -3.781316   0.679412  -5.566 2.61e-08 ***
## YEAR2000             -3.637900   0.589907  -6.167 6.96e-10 ***
## YEAR2001             -1.911728   0.556817  -3.433 0.000596 ***
## YEAR2002             -3.087370   0.577214  -5.349 8.86e-08 ***
## YEAR2003             -3.213629   0.560074  -5.738 9.59e-09 ***
## YEAR2004             -2.823208   0.574234  -4.916 8.81e-07 ***
## YEAR2005             -1.373275   0.537839  -2.553 0.010670 *
## YEAR2006             -2.793403   0.586432  -4.763 1.90e-06 ***
## YEAR2007             -2.567211   0.561035  -4.576 4.74e-06 ***
## YEAR2008             -2.054857   0.549784  -3.738 0.000186 ***
## YEAR2009             -3.172396   0.523906  -6.055 1.40e-09 ***
## YEAR2010             -2.998817   0.548164  -5.471 4.48e-08 ***
## YEAR2011             -2.493062   0.525322  -4.746 2.08e-06 ***
## YEAR2012             -3.810740   0.508992  -7.487 7.06e-14 ***
## YEAR2013             -2.962199   0.515199  -5.750 8.94e-09 ***
## YEAR2014             -2.432151   0.520455  -4.673 2.97e-06 ***
## YEAR2015             -2.507818   0.545926  -4.594 4.35e-06 ***
## YEAR2016             -1.769296   0.563705  -3.139 0.001697 **
## YEAR2017             -2.435203   0.518966  -4.692 2.70e-06 ***
## YEAR2018             -2.362261   0.539187  -4.381 1.18e-05 ***
## YEAR2019             -1.882621   0.524448  -3.590 0.000331 ***
```

```
## YEAR2020            -1.417775   0.905058  -1.567 0.117231
## YEAR2021            -1.993150   0.543195  -3.669 0.000243 ***
## AREALittle Neck Bay -0.002769   0.188138  -0.015 0.988257
## AREAManhassett Bay   0.536020   0.187090   2.865 0.004170 **
## STEMP               -0.112476   0.085332  -1.318 0.187470
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)         -1.20288    0.32825  -3.665 0.000248 ***
## YEAR1988            -0.13156    0.58390  -0.225 0.821732
## YEAR1989            -0.88395    0.56758  -1.557 0.119376
## YEAR1990            -0.11592    0.50371  -0.230 0.817991
## YEAR1991            -2.58863    1.06074  -2.440 0.014671 *
## YEAR1992            -0.74299    0.63293  -1.174 0.240442
## YEAR1993             0.17994    0.50391   0.357 0.721027
## YEAR1994             1.08987    0.48049   2.268 0.023313 *
## YEAR1995             0.85400    0.51227   1.667 0.095493 .
## YEAR1996             0.52973    0.93374   0.567 0.570496
## YEAR1998             1.12036    0.49559   2.261 0.023780 *
## YEAR1999             1.28024    0.50896   2.515 0.011889 *
## YEAR2000             0.81051    0.42847   1.892 0.058540 .
## YEAR2001            -0.42213    0.52288  -0.807 0.419487
## YEAR2002             0.68760    0.45151   1.523 0.127784
## YEAR2003             1.13024    0.40526   2.789 0.005288 **
## YEAR2004             0.69150    0.42272   1.636 0.101871
## YEAR2005             0.36944    0.43594   0.847 0.396742
## YEAR2006             0.87271    0.43061   2.027 0.042697 *
## YEAR2007             0.96370    0.41222   2.338 0.019395 *
## YEAR2008             0.04763    0.41870   0.114 0.909424
## YEAR2009            -0.08117    0.43802  -0.185 0.852983
## YEAR2010            -0.24294    0.49088  -0.495 0.620676
## YEAR2011            -0.09440    0.44369  -0.213 0.831514
## YEAR2012            -0.34908    0.47469  -0.735 0.462111
## YEAR2013            -0.59863    0.47096  -1.271 0.203693
## YEAR2014            -0.41621    0.45187  -0.921 0.357007
## YEAR2015             0.91441    0.40295   2.269 0.023251 *
## YEAR2016             1.31974    0.40868   3.229 0.001241 **
## YEAR2017             0.25810    0.41902   0.616 0.537929
## YEAR2018             0.36041    0.43687   0.825 0.409385
## YEAR2019            -0.72478    0.49617  -1.461 0.144081
## YEAR2020             0.17762    0.65577   0.271 0.786498
## YEAR2021             0.21716    0.42917   0.506 0.612859
## MONTH8              -0.27876    0.14105  -1.976 0.048122 *
## MONTH9              -0.74444    0.15756  -4.725 2.30e-06 ***
## AREALittle Neck Bay  0.16676    0.15936   1.046 0.295375
## AREAManhassett Bay   0.47369    0.14902   3.179 0.001479 **
## STEMP               -0.51292    0.08127  -6.311 2.77e-10 ***
## SDO                 -0.21072    0.06753  -3.120 0.001806 **
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Acceptable standard errors
drop1(ZANB3, test="Chi")

## Single term deletions
##
## Model:
## CPUE ~ YEAR + AREA + STEMP
##        Df   AIC     LRT  Pr(>Chi)
## <none>    12414
## YEAR   33 12516 168.128 < 2.2e-16 ***
## AREA    2 12420   9.726  0.007728 **
## STEMP   1 12414   1.744  0.186679
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# dispersion:
tmbZI.resids(ZANB3, y.dat=dat$CPUE)

## df.resid
## 1.732446

# remove STEMP from conditional model
ZANB4 <- glmmTMB(CPUE ~ YEAR + AREA, ziformula = ~ YEAR + MONTH + AREA +
STEMP + SDO, data = dat,family = truncated_nbinom2(link = "log"))
summary(ZANB4)

##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + AREA
## Zero inflation:        ~YEAR + MONTH + AREA + STEMP + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12413.9  12836.3  -6130.0  12259.9     1704
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.0431
##
## Conditional model:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)          4.18452    0.58348   7.172 7.41e-13 ***
## YEAR1988            -2.58062    0.65042  -3.968 7.26e-05 ***
## YEAR1989            -1.25473    0.63468  -1.977 0.048049 *
## YEAR1990            -3.13268    0.61819  -5.067 4.03e-07 ***
## YEAR1991            -2.90071    0.53345  -5.438 5.40e-08 ***
## YEAR1992            -2.53617    0.61857  -4.100 4.13e-05 ***
## YEAR1993            -4.03494    0.56182  -7.182 6.88e-13 ***
## YEAR1994            -4.33511    0.59356  -7.304 2.80e-13 ***
## YEAR1995            -3.27944    0.70140  -4.676 2.93e-06 ***
```

```
## YEAR1996            -3.17618    1.35143  -2.350 0.018761 *  
## YEAR1998            -3.79299    0.63262  -5.996 2.03e-09 ***
## YEAR1999            -3.75735    0.68058  -5.521 3.37e-08 ***
## YEAR2000            -3.51387    0.58235  -6.034 1.60e-09 ***
## YEAR2001            -1.87782    0.55610  -3.377 0.000733 ***
## YEAR2002            -3.04562    0.57556  -5.292 1.21e-07 ***
## YEAR2003            -3.15532    0.55729  -5.662 1.50e-08 ***
## YEAR2004            -2.76922    0.57158  -4.845 1.27e-06 ***
## YEAR2005            -1.42015    0.53497  -2.655 0.007939 ** 
## YEAR2006            -2.76360    0.58405  -4.732 2.23e-06 ***
## YEAR2007            -2.46644    0.55698  -4.428 9.50e-06 ***
## YEAR2008            -1.96667    0.54500  -3.609 0.000308 ***
## YEAR2009            -3.09531    0.51990  -5.954 2.62e-09 ***
## YEAR2010            -2.94977    0.54618  -5.401 6.64e-08 ***
## YEAR2011            -2.44862    0.52405  -4.673 2.97e-06 ***
## YEAR2012            -3.80104    0.50799  -7.482 7.29e-14 ***
## YEAR2013            -2.87237    0.51053  -5.626 1.84e-08 ***
## YEAR2014            -2.33947    0.51513  -4.542 5.58e-06 ***
## YEAR2015            -2.47635    0.54440  -4.549 5.40e-06 ***
## YEAR2016            -1.71576    0.56254  -3.050 0.002288 ** 
## YEAR2017            -2.35434    0.51490  -4.572 4.82e-06 ***
## YEAR2018            -2.38208    0.53816  -4.426 9.58e-06 ***
## YEAR2019            -1.83889    0.52247  -3.520 0.000432 ***
## YEAR2020            -1.41033    0.89757  -1.571 0.116118    
## YEAR2021            -1.95952    0.54094  -3.622 0.000292 ***
## AREALittle Neck Bay  0.05987    0.18153   0.330 0.741542    
## AREAManhassett Bay   0.56591    0.18561   3.049 0.002297 ** 
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Zero-inflation model:
##                     Estimate Std. Error z value Pr(>|z|)    
## (Intercept)         -1.20282    0.32824  -3.664 0.000248 ***
## YEAR1988            -0.13171    0.58391  -0.226 0.821539    
## YEAR1989            -0.88405    0.56758  -1.558 0.119335    
## YEAR1990            -0.11599    0.50370  -0.230 0.817882    
## YEAR1991            -2.58872    1.06075  -2.440 0.014669 *  
## YEAR1992            -0.74303    0.63293  -1.174 0.240408    
## YEAR1993             0.17986    0.50391   0.357 0.721147    
## YEAR1994             1.08979    0.48048   2.268 0.023323 *  
## YEAR1995             0.85393    0.51226   1.667 0.095517 .  
## YEAR1996             0.52965    0.93374   0.567 0.570555    
## YEAR1998             1.12028    0.49559   2.260 0.023790 *  
## YEAR1999             1.28019    0.50895   2.515 0.011892 *  
## YEAR2000             0.81047    0.42847   1.892 0.058549 .  
## YEAR2001            -0.42222    0.52288  -0.807 0.419389    
## YEAR2002             0.68757    0.45150   1.523 0.127796    
## YEAR2003             1.13015    0.40526   2.789 0.005292 ** 
## YEAR2004             0.69143    0.42272   1.636 0.101905    
## YEAR2005             0.36941    0.43594   0.847 0.396781    
```

```
## YEAR2006             0.87267    0.43061   2.027 0.042706 *
## YEAR2007             0.96360    0.41221   2.338 0.019407 *
## YEAR2008             0.04751    0.41870   0.113 0.909659
## YEAR2009            -0.08121    0.43802  -0.185 0.852913
## YEAR2010            -0.24298    0.49088  -0.495 0.620611
## YEAR2011            -0.09449    0.44369  -0.213 0.831361
## YEAR2012            -0.34912    0.47469  -0.735 0.462047
## YEAR2013            -0.59871    0.47096  -1.271 0.203631
## YEAR2014            -0.41628    0.45187  -0.921 0.356920
## YEAR2015             0.91434    0.40295   2.269 0.023260 *
## YEAR2016             1.31967    0.40868   3.229 0.001242 **
## YEAR2017             0.25798    0.41902   0.616 0.538111
## YEAR2018             0.36032    0.43687   0.825 0.409490
## YEAR2019            -0.72481    0.49616  -1.461 0.144059
## YEAR2020             0.17757    0.65576   0.271 0.786556
## YEAR2021             0.21713    0.42917   0.506 0.612899
## MONTH8              -0.27877    0.14105  -1.976 0.048112 *
## MONTH9              -0.74445    0.15756  -4.725 2.30e-06 ***
## AREALittle Neck Bay  0.16676    0.15936   1.046 0.295356
## AREAManhassett Bay   0.47371    0.14902   3.179 0.001479 **
## STEMP               -0.51293    0.08127  -6.311 2.77e-10 ***
## SDO                 -0.21073    0.06753  -3.120 0.001806 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Acceptable standard errors
drop1(ZANB4, test="Chi")

## Single term deletions
##
## Model:
## CPUE ~ YEAR + AREA
##        Df   AIC     LRT  Pr(>Chi)
## <none>    12414
## YEAR   33 12515 166.893 < 2.2e-16 ***
## AREA    2 12420  10.237  0.005985 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# dispersion:
tmbZI.resids(ZANB4, y.dat=dat$CPUE)

## df.resid
## 1.782778

# all terms in zero-infl model significant, so keeping all


# Let's compare to a model with year only.
ZANB0 <- glmmTMB(CPUE ~ YEAR, ziformula = ~ YEAR, data = dat,family =
```

```
truncated_nbinom2(link = "log"))
summary(ZANB0)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR
## Zero inflation:         ~YEAR
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12508.6  12887.1  -6185.3  12370.6     1712
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.0375
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.0932     0.6249   6.550 5.75e-11 ***
## YEAR1988     -2.5244     0.6547  -3.856 0.000115 ***
## YEAR1989     -1.1079     0.6384  -1.735 0.082684 .
## YEAR1990     -2.8947     0.6206  -4.665 3.09e-06 ***
## YEAR1991     -2.7457     0.5319  -5.162 2.44e-07 ***
## YEAR1992     -2.4863     0.6227  -3.993 6.53e-05 ***
## YEAR1993     -4.0167     0.5549  -7.238 4.56e-13 ***
## YEAR1994     -4.3000     0.5838  -7.365 1.77e-13 ***
## YEAR1995     -3.2370     0.6926  -4.674 2.96e-06 ***
## YEAR1996     -3.2095     1.3570  -2.365 0.018018 *
## YEAR1998     -3.7510     0.6240  -6.011 1.84e-09 ***
## YEAR1999     -3.6153     0.6813  -5.307 1.12e-07 ***
## YEAR2000     -3.5306     0.5721  -6.171 6.79e-10 ***
## YEAR2001     -1.5175     0.5465  -2.777 0.005492 **
## YEAR2002     -3.0716     0.5674  -5.413 6.18e-08 ***
## YEAR2003     -3.1693     0.5579  -5.681 1.34e-08 ***
## YEAR2004     -2.6124     0.5715  -4.571 4.85e-06 ***
## YEAR2005     -1.4430     0.5315  -2.715 0.006631 **
## YEAR2006     -2.6231     0.5800  -4.523 6.10e-06 ***
## YEAR2007     -2.4079     0.5599  -4.300 1.70e-05 ***
## YEAR2008     -1.5458     0.5323  -2.904 0.003684 **
## YEAR2009     -2.9216     0.5188  -5.631 1.79e-08 ***
## YEAR2010     -2.7315     0.5370  -5.086 3.65e-07 ***
## YEAR2011     -2.2668     0.5252  -4.316 1.59e-05 ***
## YEAR2012     -3.5977     0.5053  -7.120 1.08e-12 ***
## YEAR2013     -2.7926     0.5038  -5.543 2.97e-08 ***
## YEAR2014     -2.1148     0.5098  -4.148 3.35e-05 ***
## YEAR2015     -2.1642     0.5376  -4.026 5.68e-05 ***
## YEAR2016     -1.6837     0.5647  -2.981 0.002870 **
## YEAR2017     -2.2910     0.5174  -4.428 9.52e-06 ***
## YEAR2018     -2.0427     0.5260  -3.883 0.000103 ***
## YEAR2019     -1.7177     0.5174  -3.320 0.000900 ***
## YEAR2020     -1.4139     0.8922  -1.585 0.113035
## YEAR2021     -1.9276     0.5338  -3.611 0.000305 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.14505    0.30684  -3.732  0.00019 ***
## YEAR1988    -0.10776    0.55538  -0.194  0.84615
## YEAR1989    -0.35900    0.54576  -0.658  0.51067
## YEAR1990     0.26956    0.48561   0.555  0.57883
## YEAR1991    -2.68368    1.05640  -2.540  0.01107 *
## YEAR1992    -0.68763    0.61982  -1.109  0.26725
## YEAR1993    -0.12347    0.48633  -0.254  0.79958
## YEAR1994     0.49111    0.45943   1.069  0.28509
## YEAR1995     0.77036    0.49755   1.548  0.12155
## YEAR1996     1.14503    0.87225   1.313  0.18927
## YEAR1998     0.66546    0.46765   1.423  0.15474
## YEAR1999     0.93740    0.48330   1.940  0.05243 .
## YEAR2000     0.87935    0.41348   2.127  0.03344 *
## YEAR2001    -0.75918    0.50822  -1.494  0.13522
## YEAR2002     0.42110    0.43238   0.974  0.33011
## YEAR2003     1.05544    0.39239   2.690  0.00715 **
## YEAR2004     0.79675    0.40649   1.960  0.04999 *
## YEAR2005    -0.07131    0.41853  -0.170  0.86472
## YEAR2006     0.72384    0.41596   1.740  0.08183 .
## YEAR2007     0.83968    0.39520   2.125  0.03361 *
## YEAR2008     0.11920    0.40700   0.293  0.76962
## YEAR2009    -0.07871    0.42477  -0.185  0.85299
## YEAR2010    -0.44193    0.47748  -0.926  0.35468
## YEAR2011    -0.16713    0.42990  -0.389  0.69745
## YEAR2012    -0.54134    0.46116  -1.174  0.24045
## YEAR2013    -0.66323    0.45885  -1.445  0.14834
## YEAR2014    -0.49721    0.44002  -1.130  0.25848
## YEAR2015     0.66220    0.38718   1.710  0.08721 .
## YEAR2016     0.92191    0.38782   2.377  0.01745 *
## YEAR2017     0.06414    0.40576   0.158  0.87440
## YEAR2018    -0.05268    0.41895  -0.126  0.89993
## YEAR2019    -0.86983    0.48561  -1.791  0.07326 .
## YEAR2020     0.45190    0.62781   0.720  0.47165
## YEAR2021     0.06624    0.41590   0.159  0.87346
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Let's compare the models using AIC to see if adding more factors
# improves how much deviance is explained.

AICtab(ZANB1, ZANB2, ZANB3, ZANB4, ZANB0)
```

```
##       dAIC df
## ZANB2  0.0 80
## ZANB1  0.1 81
```

```
## ZANB4  0.6 77
## ZANB3  0.8 78
## ZANB0 95.3 69
```

```
lrtest(ZANB0, ZANB4, ZANB3, ZANB2, ZANB1)
```

```
## Likelihood ratio test
##
## Model 1: CPUE ~ YEAR
## Model 2: CPUE ~ YEAR + AREA
## Model 3: CPUE ~ YEAR + AREA + STEMP
## Model 4: CPUE ~ YEAR + MONTH + AREA + STEMP
## Model 5: CPUE ~ YEAR + MONTH + AREA + STEMP + SDO
##   #Df  LogLik Df    Chisq Pr(>Chisq)
## 1  69 -6185.3
## 2  77 -6130.0  8 110.7072    < 2e-16 ***
## 3  78 -6129.1  1   1.7436    0.18668
## 4  80 -6126.7  2   4.8380    0.08901 .
## 5  81 -6125.7  1   1.8832    0.16997
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sapply(list(ZANB2,ZANB1,ZANB4,ZANB3,ZANB0),tmbZI.resids,y.dat=dat$CPUE)
```

```
## df.resid df.resid df.resid df.resid df.resid
## 1.671789 1.621377 1.782778 1.732446 1.729076
```

Model comparison:

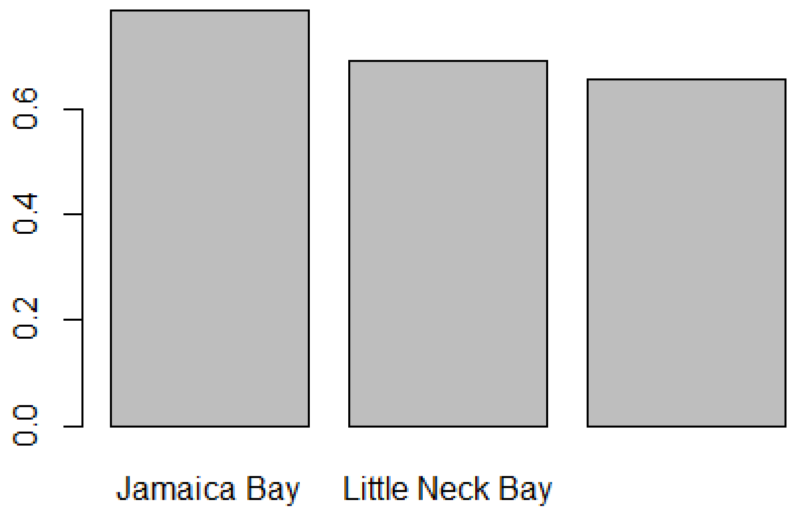
```
AICtab(NB5,ZINB4,ZANB4)
```

```
##       dAIC  df
## ZANB4   0.0 77
## ZINB4 199.4 42
## NB5   260.4 35
```

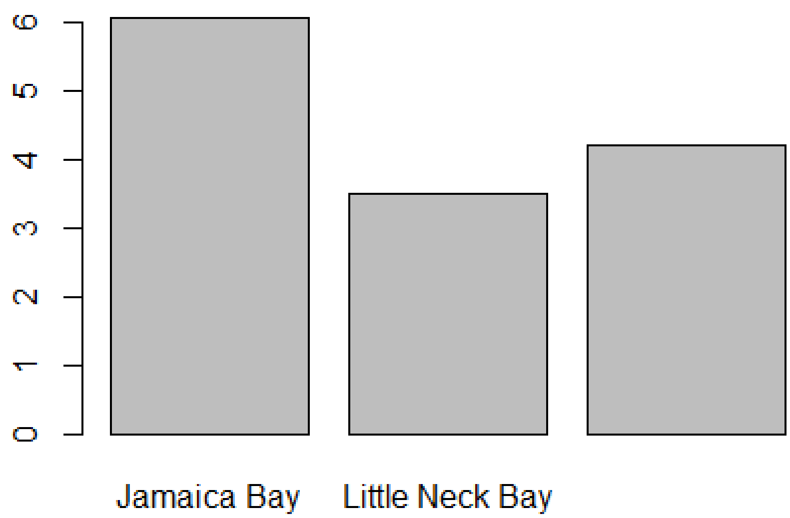
Model diagnostics

ZANB

```
# examine factors with problematic diagnostics:
barplot(with(dat, tapply(CPUE,AREA,pctPos)))
```



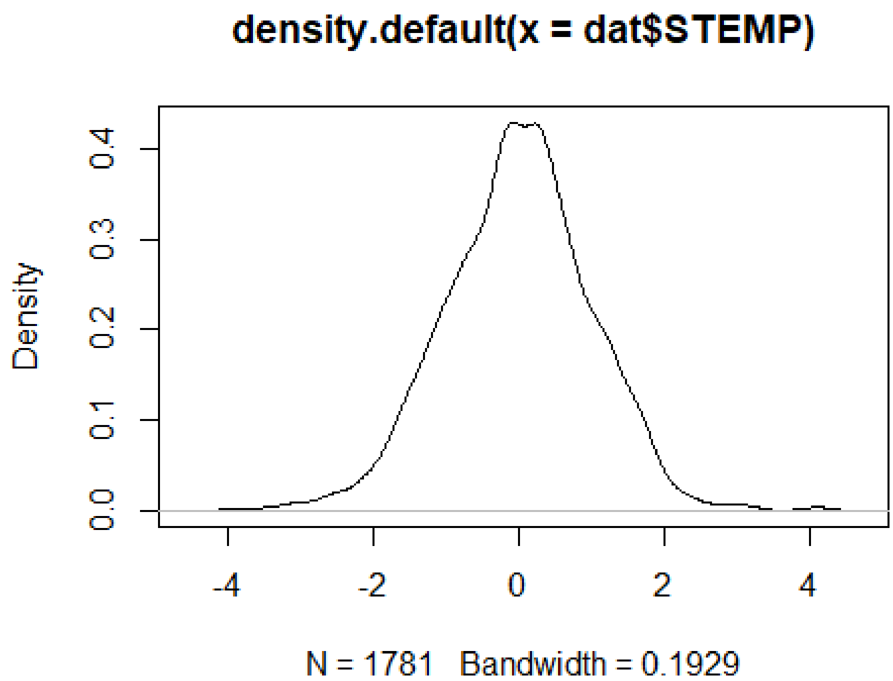

```
barplot(with(dat, tapply(CPUE,AREA,geoMean,alpha=NULL,warn=FALSE)))
```



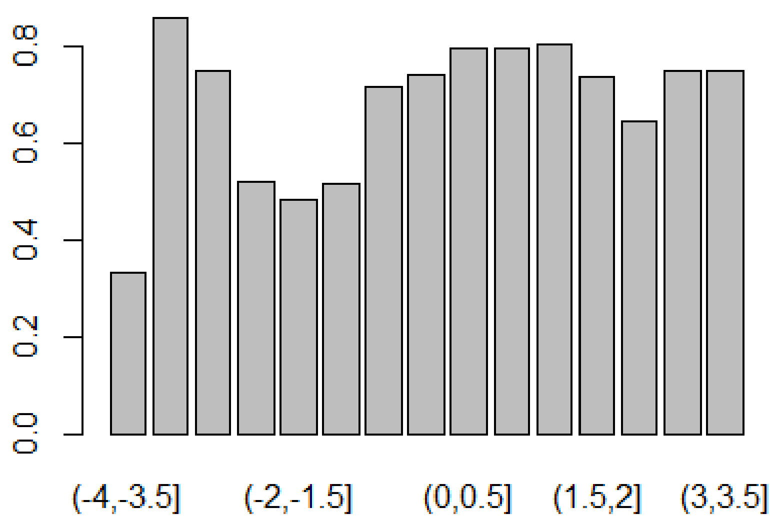

```
with(dat, tapply(CPUE,list(AREA,YEAR),pctPos))
```

```
##                      1987      1988      1989      1990      1991
1992
## Jamaica Bay     0.7586207 0.7333333 0.9090909 0.6363636 0.9655172
1.0000000
## Little Neck Bay 0.5384615 0.8750000 1.0000000 0.6666667 1.0000000
1.0000000
## Manhassett Bay  0.9375000 0.7500000 0.6153846 0.8750000 1.0000000
0.4285714
##                      1993      1994      1995 1996   1998      1999
2000
## Jamaica Bay     0.8888889 0.9047619 0.6153846    1 0.6875 0.5000000
0.6551724
## Little Neck Bay 0.6363636 0.1111111 0.4285714    0 0.5000 0.5454545
0.5833333
## Manhassett Bay  0.7500000 0.6250000 0.7142857    0 0.6250 0.7500000
0.3333333
##                      2001      2002      2003      2004      2005
2006
## Jamaica Bay     0.8333333 0.7600000 0.7812500 0.7826087 0.8571429
0.5172414
## Little Neck Bay 0.7692308 0.7777778 0.3529412 0.5555556 0.8235294
0.7692308
## Manhassett Bay  1.0000000 0.4666667 0.2222222 0.3529412 0.5555556
0.6363636
##                      2007      2008      2009      2010      2011
2012
## Jamaica Bay     0.7000000 0.7878788 0.9000000 0.7826087 0.8000000
0.7857143
## Little Neck Bay 0.5555556 0.7368421 0.7222222 0.9333333 0.7222222
0.8888889
## Manhassett Bay  0.3888889 0.6500000 0.6111111 0.8000000 0.8333333
0.8888889
##                      2013      2014      2015      2016      2017
2018
## Jamaica Bay     0.9166667 0.8181818 0.7428571 0.6388889 0.8205128
0.8181818
## Little Neck Bay 0.8235294 0.8500000 0.5000000 0.5555556 0.7222222
0.7222222
## Manhassett Bay  0.7777778 0.8571429 0.5238095 0.3888889 0.6111111
0.7222222
##                      2019      2020      2021
## Jamaica Bay     0.9090909 0.8333333 0.7647059
## Little Neck Bay 0.8235294 0.5000000 0.7647059
## Manhassett Bay  0.8888889 0.6000000 0.6875000
```

```
plot(density(dat$STEMP))
```




```
barplot(with(dat, tapply(CPUE,cut(STEMP,breaks=seq(-4,4,by=0.5)),pctPos)))
```



```
if(FALSE) {
### 1. remove 1996 (this didn't help):
# removing 1996 to see what that does, so save original dataset
.all <- dat
# dat <- .all
dat <- dat[dat$YEAR!=1996,]
# levels(dat$YEAR)
dat <- droplevels(dat)
#levels(dat$YEAR)
ZANB4.1 <- glmmTMB(CPUE ~ YEAR + AREA, ziformula = ~ YEAR + MONTH + AREA +
STEMP + SDO,
                  data = dat,family = truncated_nbinom2(link = "log"))
summary(ZANB4.1)
# Acceptable standard errors
drop1(ZANB4.1, test="Chi")
# dispersion:
tmbZI.resids(ZANB4.1, y.dat=dat$CPUE)

### 2. remove observations at STEMP <=-3.5 (this didn't help either)
dat <- .all
dat <- dat[dat$STEMP> -3.5,]
ZANB4.2 <- glmmTMB(CPUE ~ YEAR + AREA, ziformula = ~ YEAR + MONTH + AREA +
STEMP + SDO,
                  data = dat,family = truncated_nbinom2(link = "log"))
summary(ZANB4.2)
# Acceptable standard errors
drop1(ZANB4.2, test="Chi")
# dispersion:
tmbZI.resids(ZANB4.2, y.dat=dat$CPUE)

### 3. replace large catches with 95th percentile of time series catch (this
didn't help)
dat <- .all
which(dat$CPUE>quantile(dat$CPUE,0.95))
plot(sort(dat$CPUE,d=TRUE)[1:50]) # so just replace the largest value
dat[which.max(dat$CPUE),"CPUE"] <- quantile(dat$CPUE,0.95)
ZANB4.3 <- glmmTMB(CPUE ~ YEAR + AREA, ziformula = ~ YEAR + MONTH + AREA +
STEMP + SDO,
                  data = dat,family = truncated_nbinom2(link = "log"))
summary(ZANB4.3)
# Acceptable standard errors
drop1(ZANB4.3, test="Chi")
# dispersion:
tmbZI.resids(ZANB4.3, y.dat=dat$CPUE)
}

### 4. remove STEMP; this mostly worked.
#.all <- dat # since commenting out the steps above, don't need this step
#dat <- .all # since commenting out the steps above, don't need this step
ZANB4.4 <- glmmTMB(CPUE ~ YEAR + AREA, ziformula = ~ YEAR + MONTH + AREA +
```

```
SDO,
                  data = dat,family = truncated_nbinom2(link = "log"))
summary(ZANB4.4)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + AREA
## Zero inflation:        ~YEAR + MONTH + AREA + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12453.8  12870.7  -6150.9  12301.8     1705
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.0431
##
## Conditional model:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)             4.18452    0.58349   7.172 7.42e-13 ***
## YEAR1988               -2.58062    0.65042  -3.968 7.26e-05 ***
## YEAR1989               -1.25474    0.63469  -1.977 0.048048 *
## YEAR1990               -3.13268    0.61820  -5.067 4.03e-07 ***
## YEAR1991               -2.90072    0.53345  -5.438 5.40e-08 ***
## YEAR1992               -2.53619    0.61857  -4.100 4.13e-05 ***
## YEAR1993               -4.03497    0.56182  -7.182 6.87e-13 ***
## YEAR1994               -4.33514    0.59356  -7.304 2.80e-13 ***
## YEAR1995               -3.27945    0.70141  -4.676 2.93e-06 ***
## YEAR1996               -3.17614    1.35147  -2.350 0.018766 *
## YEAR1998               -3.79301    0.63262  -5.996 2.03e-09 ***
## YEAR1999               -3.75740    0.68057  -5.521 3.37e-08 ***
## YEAR2000               -3.51390    0.58235  -6.034 1.60e-09 ***
## YEAR2001               -1.87786    0.55610  -3.377 0.000733 ***
## YEAR2002               -3.04564    0.57556  -5.292 1.21e-07 ***
## YEAR2003               -3.15536    0.55729  -5.662 1.50e-08 ***
## YEAR2004               -2.76925    0.57158  -4.845 1.27e-06 ***
## YEAR2005               -1.42016    0.53497  -2.655 0.007939 **
## YEAR2006               -2.76363    0.58405  -4.732 2.23e-06 ***
## YEAR2007               -2.46648    0.55698  -4.428 9.50e-06 ***
## YEAR2008               -1.96669    0.54500  -3.609 0.000308 ***
## YEAR2009               -3.09533    0.51990  -5.954 2.62e-09 ***
## YEAR2010               -2.94980    0.54618  -5.401 6.64e-08 ***
## YEAR2011               -2.44863    0.52405  -4.673 2.98e-06 ***
## YEAR2012               -3.80110    0.50799  -7.483 7.29e-14 ***
## YEAR2013               -2.87240    0.51053  -5.626 1.84e-08 ***
## YEAR2014               -2.33956    0.51512  -4.542 5.58e-06 ***
## YEAR2015               -2.47637    0.54441  -4.549 5.40e-06 ***
## YEAR2016               -1.71579    0.56255  -3.050 0.002288 **
## YEAR2017               -2.35434    0.51490  -4.572 4.82e-06 ***
## YEAR2018               -2.38210    0.53816  -4.426 9.58e-06 ***
## YEAR2019               -1.83890    0.52247  -3.520 0.000432 ***
## YEAR2020               -1.41039    0.89755  -1.571 0.116097
```

```
## YEAR2021            -1.95952    0.54094  -3.622 0.000292 ***
## AREALittle Neck Bay  0.05986    0.18153   0.330 0.741582
## AREAManhassett Bay   0.56593    0.18561   3.049 0.002296 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)         -1.34383    0.32338  -4.156 3.25e-05 ***
## YEAR1988             0.37362    0.57742   0.647  0.51760
## YEAR1989            -0.37587    0.55350  -0.679  0.49709
## YEAR1990             0.21207    0.49441   0.429  0.66797
## YEAR1991            -2.63427    1.05989  -2.485  0.01294 *
## YEAR1992            -0.67897    0.62813  -1.081  0.27972
## YEAR1993             0.04574    0.49676   0.092  0.92663
## YEAR1994             0.77925    0.47023   1.657  0.09748 .
## YEAR1995             0.77687    0.50781   1.530  0.12606
## YEAR1996             0.74877    0.89193   0.839  0.40119
## YEAR1998             0.82046    0.47909   1.713  0.08680 .
## YEAR1999             0.89347    0.49392   1.809  0.07046 .
## YEAR2000             0.95127    0.42240   2.252  0.02432 *
## YEAR2001            -0.62850    0.51712  -1.215  0.22421
## YEAR2002             0.44033    0.44396   0.992  0.32129
## YEAR2003             1.13897    0.40140   2.837  0.00455 **
## YEAR2004             0.83992    0.41610   2.019  0.04353 *
## YEAR2005             0.04959    0.42812   0.116  0.90779
## YEAR2006             0.92488    0.42670   2.168  0.03019 *
## YEAR2007             0.82575    0.40584   2.035  0.04188 *
## YEAR2008             0.06228    0.41497   0.150  0.88070
## YEAR2009            -0.13529    0.43373  -0.312  0.75511
## YEAR2010            -0.46576    0.48675  -0.957  0.33863
## YEAR2011            -0.25372    0.44000  -0.577  0.56419
## YEAR2012            -0.54548    0.46991  -1.161  0.24572
## YEAR2013            -0.51440    0.46722  -1.101  0.27091
## YEAR2014            -0.48766    0.44781  -1.089  0.27616
## YEAR2015             0.64438    0.39535   1.630  0.10312
## YEAR2016             0.86781    0.39899   2.175  0.02963 *
## YEAR2017             0.24541    0.41586   0.590  0.55510
## YEAR2018            -0.06073    0.42787  -0.142  0.88712
## YEAR2019            -0.95024    0.49365  -1.925  0.05424 .
## YEAR2020             0.47243    0.64323   0.734  0.46267
## YEAR2021             0.05096    0.42589   0.120  0.90475
## MONTH8              -0.42143    0.13668  -3.083  0.00205 **
## MONTH9              -0.40406    0.14435  -2.799  0.00512 **
## AREALittle Neck Bay  0.62717    0.14070   4.457 8.30e-06 ***
## AREAManhassett Bay   0.83016    0.13783   6.023 1.71e-09 ***
## SDO                 -0.25926    0.06554  -3.956 7.63e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Acceptable standard errors
drop1(ZANB4.4, test="Chi")

## Single term deletions
##
## Model:
## CPUE ~ YEAR + AREA
##        Df   AIC     LRT  Pr(>Chi)
## <none>    12454
## YEAR   33 12555 166.893 < 2.2e-16 ***
## AREA    2 12460  10.237  0.005985 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# dispersion:
tmbZI.resids(ZANB4.4, y.dat=dat$CPUE)

## df.resid
## 1.742674

### 5. Lastly,  remove AREA too. This made things a little better yet.
dat <- .all

## Error in eval(expr, envir, enclos): object '.all' not found

ZANB4.5 <- glmmTMB(CPUE ~ YEAR , ziformula = ~ YEAR + MONTH + SDO,
                  data = dat,family = truncated_nbinom2(link = "log"))
summary(ZANB4.5)

##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR
## Zero inflation:        ~YEAR + MONTH + SDO
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##  12497.6  12892.5  -6176.8  12353.6     1709
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.0375
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   4.0932     0.6249   6.550 5.75e-11 ***
## YEAR1988     -2.5243     0.6546  -3.856 0.000115 ***
## YEAR1989     -1.1077     0.6384  -1.735 0.082719 .
## YEAR1990     -2.8946     0.6206  -4.665 3.09e-06 ***
## YEAR1991     -2.7456     0.5319  -5.162 2.44e-07 ***
## YEAR1992     -2.4861     0.6227  -3.993 6.53e-05 ***
## YEAR1993     -4.0166     0.5549  -7.238 4.55e-13 ***
## YEAR1994     -4.2998     0.5838  -7.365 1.77e-13 ***
## YEAR1995     -3.2369     0.6926  -4.674 2.96e-06 ***
```

```
## YEAR1996     -3.2096     1.3568  -2.366 0.018005 *
## YEAR1998     -3.7509     0.6240  -6.011 1.84e-09 ***
## YEAR1999     -3.6152     0.6813  -5.307 1.12e-07 ***
## YEAR2000     -3.5305     0.5721  -6.171 6.79e-10 ***
## YEAR2001     -1.5174     0.5465  -2.777 0.005493 **
## YEAR2002     -3.0715     0.5674  -5.414 6.18e-08 ***
## YEAR2003     -3.1692     0.5578  -5.681 1.34e-08 ***
## YEAR2004     -2.6123     0.5715  -4.571 4.85e-06 ***
## YEAR2005     -1.4429     0.5315  -2.715 0.006631 **
## YEAR2006     -2.6230     0.5800  -4.523 6.11e-06 ***
## YEAR2007     -2.4078     0.5599  -4.300 1.71e-05 ***
## YEAR2008     -1.5457     0.5323  -2.904 0.003684 **
## YEAR2009     -2.9216     0.5188  -5.631 1.79e-08 ***
## YEAR2010     -2.7314     0.5370  -5.086 3.65e-07 ***
## YEAR2011     -2.2667     0.5252  -4.316 1.59e-05 ***
## YEAR2012     -3.5976     0.5053  -7.120 1.08e-12 ***
## YEAR2013     -2.7925     0.5038  -5.543 2.97e-08 ***
## YEAR2014     -2.1146     0.5098  -4.148 3.35e-05 ***
## YEAR2015     -2.1640     0.5376  -4.026 5.68e-05 ***
## YEAR2016     -1.6835     0.5647  -2.981 0.002871 **
## YEAR2017     -2.2910     0.5174  -4.428 9.51e-06 ***
## YEAR2018     -2.0425     0.5260  -3.883 0.000103 ***
## YEAR2019     -1.7175     0.5173  -3.320 0.000900 ***
## YEAR2020     -1.4138     0.8922  -1.585 0.113053
## YEAR2021     -1.9275     0.5338  -3.611 0.000305 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.959840   0.312323  -3.073  0.00212 **
## YEAR1988     0.208539   0.567524   0.368  0.71328
## YEAR1989    -0.264628   0.548533  -0.482  0.62950
## YEAR1990     0.303284   0.488854   0.620  0.53500
## YEAR1991    -2.700505   1.057575  -2.554  0.01066 *
## YEAR1992    -0.644318   0.623209  -1.034  0.30120
## YEAR1993     0.003991   0.492058   0.008  0.99353
## YEAR1994     0.679969   0.464174   1.465  0.14295
## YEAR1995     0.779390   0.502183   1.552  0.12066
## YEAR1996     0.817668   0.876378   0.933  0.35082
## YEAR1998     0.792972   0.472073   1.680  0.09300 .
## YEAR1999     0.880296   0.485831   1.812  0.07000 .
## YEAR2000     0.895968   0.416117   2.153  0.03131 *
## YEAR2001    -0.596468   0.511766  -1.166  0.24381
## YEAR2002     0.465353   0.436745   1.065  0.28665
## YEAR2003     1.118739   0.395684   2.827  0.00469 **
## YEAR2004     0.871743   0.410386   2.124  0.03365 *
## YEAR2005     0.047754   0.422422   0.113  0.90999
## YEAR2006     0.858038   0.420402   2.041  0.04125 *
## YEAR2007     0.867432   0.399013   2.174  0.02971 *
```

```
## YEAR2008     0.111629   0.410113   0.272  0.78548
## YEAR2009    -0.067909   0.428408  -0.158  0.87405
## YEAR2010    -0.377693   0.482027  -0.784  0.43330
## YEAR2011    -0.168823   0.433785  -0.389  0.69714
## YEAR2012    -0.471738   0.464028  -1.017  0.30934
## YEAR2013    -0.541091   0.462259  -1.171  0.24179
## YEAR2014    -0.437519   0.443394  -0.987  0.32376
## YEAR2015     0.674597   0.390471   1.728  0.08405 .
## YEAR2016     0.898666   0.392849   2.288  0.02216 *
## YEAR2017     0.216314   0.410245   0.527  0.59800
## YEAR2018    -0.016907   0.422232  -0.040  0.96806
## YEAR2019    -0.880218   0.488795  -1.801  0.07174 .
## YEAR2020     0.541597   0.634279   0.854  0.39317
## YEAR2021     0.081663   0.420437   0.194  0.84599
## MONTH8      -0.413386   0.134849  -3.066  0.00217 **
## MONTH9      -0.385924   0.142482  -2.709  0.00676 **
## SDO         -0.172270   0.062543  -2.754  0.00588 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Acceptable standard errors
drop1(ZANB4.5, test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR
##        Df   AIC    LRT  Pr(>Chi)
## <none>    12498
## YEAR   33 12600 168.09 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# dispersion:
tmbZI.resids(ZANB4.5, y.dat=dat$CPUE)
```

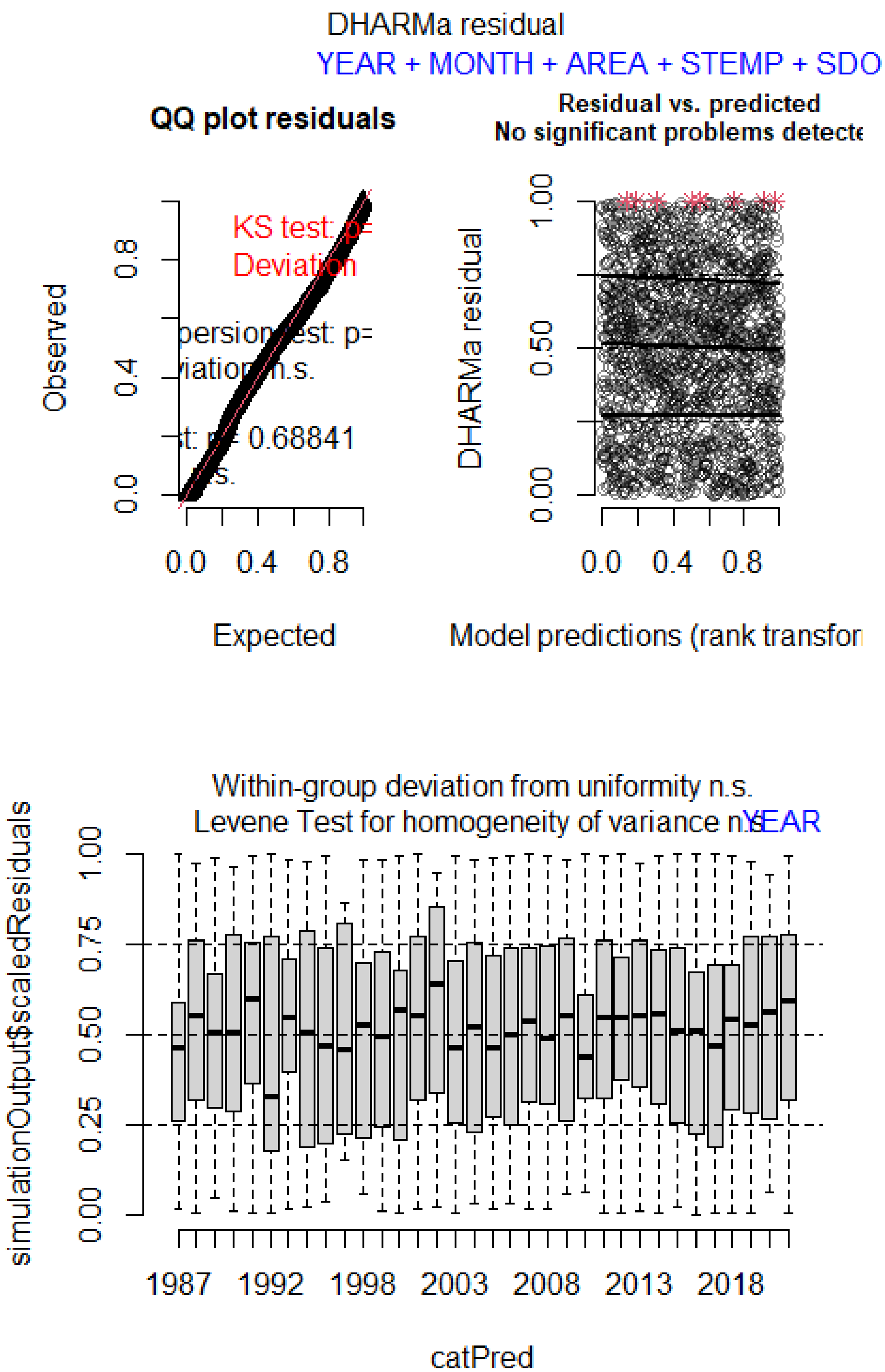
```
## df.resid
## 1.758761
```

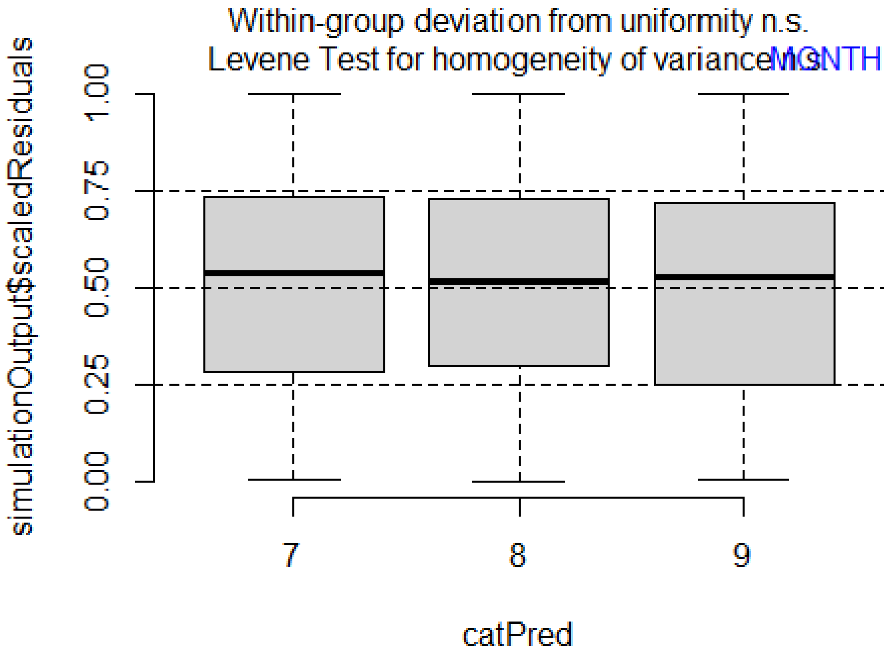
```
AICtab(ZANB4,ZANB4.4,ZANB4.5) # can't compare the other ZANBs becaus they use
different data
```

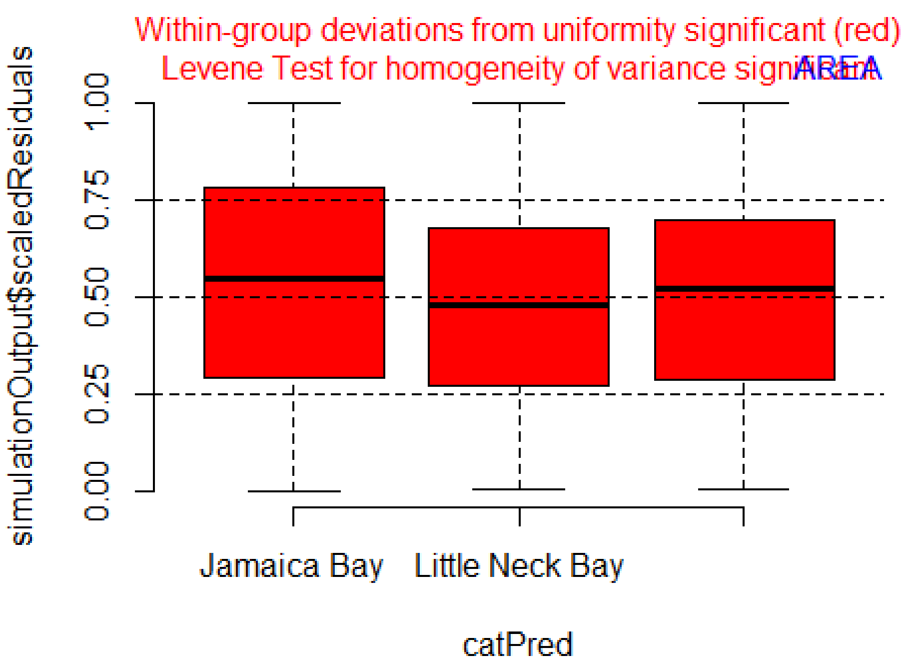
```
##         dAIC df
## ZANB4    0.0 77
## ZANB4.4 39.9 76
## ZANB4.5 83.7 72
```

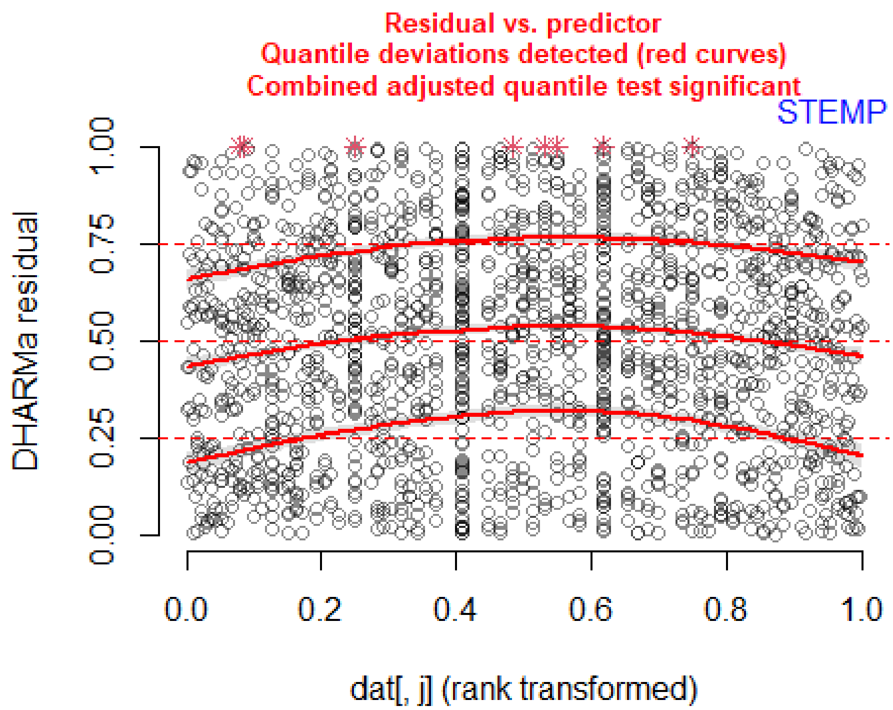
```
models <- list(ZANB4) # good to 'good enough' diagnostics
for(i in 1:length(models)) {
   ZANB.sim <- simulateResiduals(models[[i]])
   this <- names(dat)[names(dat) %in% names(models[[i]]$frame) & names(dat)
```

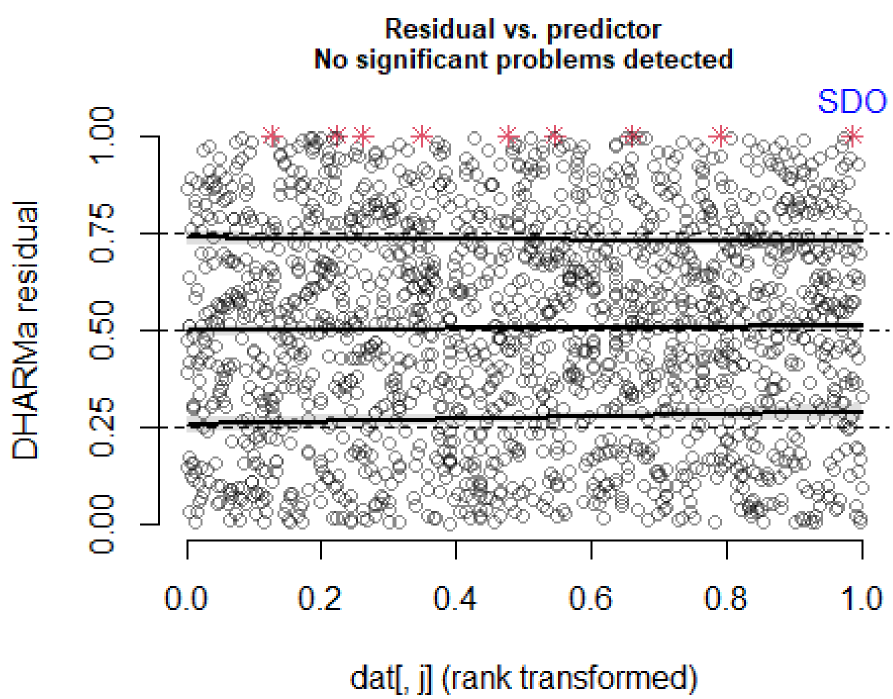
```
!= "CPUE"]
   print(this)
   plot(ZANB.sim); mtext(paste(this,collapse=" +
"),side=3,adj=0,outer=FALSE,line=1,col="blue")

   for(j in this) {
      plotResiduals(ZANB.sim, form=dat[,j]); mtext(j,3,col="blue",adj=1)
   }
}

## [1] "YEAR"  "MONTH" "AREA"  "STEMP" "SDO"
```


DHARMa residual
YEAR + MONTH + AREA + STEMP + SDO
QQ plot residuals
KS test: p=
Deviation
persion test: p=
viation n.s.
t: p= 0.68841
n.s.
Observed
Expected
Residual vs. predicted
No significant problems detecte
DHARMa residual
Model predictions (rank transfor
Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
YEAR
simulationOutput$scaledResiduals
catPred
1987
1992
1998
2003
2008
2013
2018


Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
MONTH
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
7
8
9
catPred


Within-group deviations from uniformity significant (red)
Levene Test for homogeneity of variance significant
AREA
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
Jamaica Bay
Little Neck Bay
catPred





All diagnostics generally seemed 'good enough' to bad; I stepped through a number of remedial measures for ZANB; removing STEMP and then AREA had the best diagnostics, but the model with the lowest AIC (ZANB4) seemed to have 'good enough' diagnostics, so going with ZANB4.

Model comparisons

```
AICtab(ZANB4,ZANB4.4,ZANB4.5) # ZANB4.1-3 not comparable with other models
because using different data set

##         dAIC df
## ZANB4    0.0 77
## ZANB4.4 39.9 76
## ZANB4.5 83.7 72
```

Generate and compare indices

```
# by-hand with bootstrapping :


# requires final dataset object called 'dat'
out <- byHand2(theModel=ZANB4,nBoot=1000,Method="ZANB4")

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
```

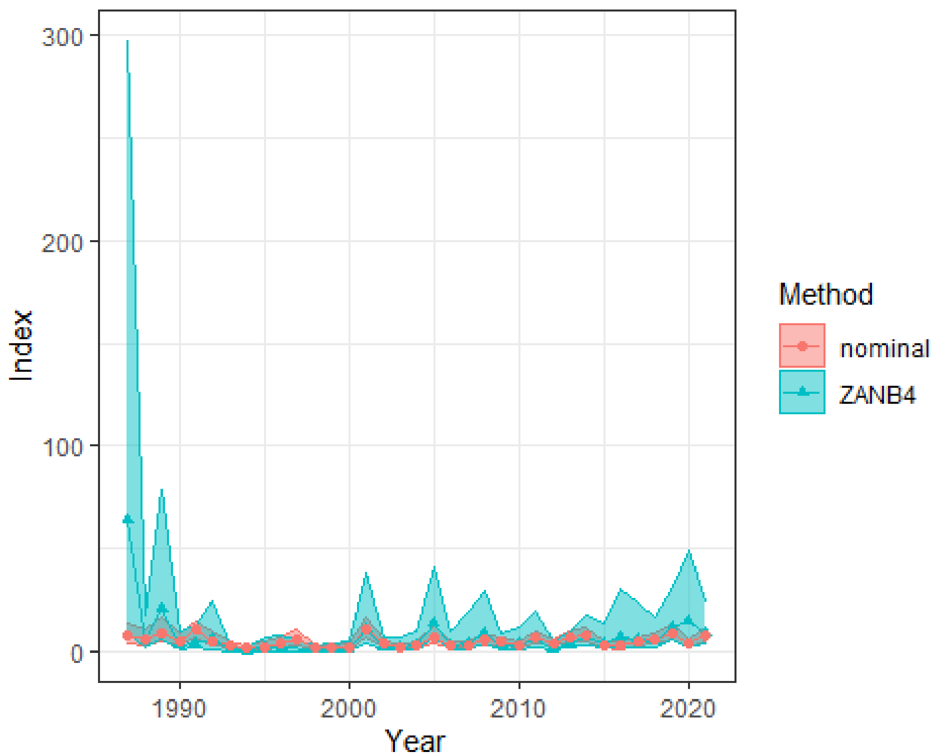
```
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

# calc nominal index using all data (i.e., not removing records with missing
env vars, but still subsetting to just July-Sept and dropping Oyster Bay)
dat <- as.data.frame(indata[is.element(indata$Month,7:9) &
(!is.element(indata$Area,"Oyster Bay")), ])

# generate nominal index with SEs:
nom <- matrix(NA,nrow=max(dat$Year),ncol=7)# year, sd, median, LCL, UCL
   colnames(nom) <- c("Year","Index","SE","LCI","UCI","CV","Method")

nom <- as.data.frame(nom)


for(i in unique(dat$Year)) {
   xx <- dat[dat$Year==i,"N"]

   storeg <- NULL
   ldata <- log(xx+1)
   for(t in 1:500){
     dodo <- sample(ldata,length(ldata),replace=TRUE)
     gmean <- exp(mean(dodo))-1
     storeg <- c(storeg,gmean)
   }
   SEB <- sd(storeg)

   #Percentile Method
   CIPB <- as.numeric(quantile(storeg,probs=c(0.025,0.50,0.975)))

   .gm <- geoMean(xx,crxn=1,alpha=0.05,warn=FALSE)
   .cv <- SEB/.gm[2]
```

```
   nom[i,1:6] <- c(i,.gm[2],SEB,CIPB[-2],.cv)
}
nom <- nom[!is.na(nom[,1]),]
nom[,"Method"] <- "nominal"

# combine nominal and standardized:
tmp <- rbind(out,nom)
tmp
```

```
##       Year      Index         SE       LCI        UCI        CV  Method
## 1987  1987 63.9782604 75.0608498 8.4699625 297.250567 1.1732243   ZANB4
## 1988  1988  4.9732127  4.2560685 1.9770681  17.608136 0.8557986   ZANB4
## 1989  1989 20.7955752 19.3569192 6.3510229  79.009701 0.9308191   ZANB4
## 1990  1990  2.8548794  2.2895273 1.3071341   9.945454 0.8019699   ZANB4
## 1991  1991  4.3611998  2.8976383 2.1419001  13.309157 0.6644131   ZANB4
## 1992  1992  5.6888088  6.5087271 1.5303250  24.917279 1.1441283   ZANB4
## 1993  1993  1.0867904  0.7049588 0.4865646   3.193067 0.6486612   ZANB4
## 1994  1994  0.5933040  0.3828384 0.2440816   1.774151 0.6452651   ZANB4
## 1995  1995  1.8788384  1.7425760 0.3627120   6.858525 0.9274752   ZANB4
## 1996  1996  2.3304683  1.9509066 0.2884644   7.630132 0.8371307   ZANB4
## 1998  1998  1.0065596  0.8006863 0.2871121   3.368262 0.7954683   ZANB4
## 1999  1999  0.9679305  0.9333829 0.2090800   3.817145 0.9643078   ZANB4
## 2000  2000  1.5108161  1.2998177 0.4616276   5.094886 0.8603415   ZANB4
## 2001  2001 10.5429453  8.7304773 4.3530840  38.115379 0.8280871   ZANB4
## 2002  2002  2.5218185  1.3527054 1.2912583   6.559251 0.5364008   ZANB4
## 2003  2003  1.8960544  1.6012205 0.7281807   6.619694 0.8445013   ZANB4
## 2004  2004  3.3202638  2.2921349 1.5015841   9.898730 0.6903472   ZANB4
## 2005  2005 14.1407915  9.0378821 6.8785199  41.688818 0.6391355   ZANB4
## 2006  2006  3.1245054  2.2269535 1.1144504   9.449716 0.7127379   ZANB4
## 2007  2007  4.0560429  5.0965622 1.0406653  19.869309 1.2565356   ZANB4
## 2008  2008  8.8618468  6.9036366 3.8298559  30.003147 0.7790291   ZANB4
## 2009  2009  2.9437288  1.8407432 1.5197945   8.468835 0.6253101   ZANB4
## 2010  2010  3.5075931  3.1119612 0.9183280  11.994163 0.8872070   ZANB4
## 2011  2011  5.6347388  4.3805278 2.5060665  19.465947 0.7774145   ZANB4
## 2012  2012  1.5236365  1.0745673 0.7212997   4.964302 0.7052649   ZANB4
## 2013  2013  3.9952366  2.0953934 1.9503601  10.246314 0.5244729   ZANB4
## 2014  2014  6.6387902  3.8085245 2.9191759  18.095155 0.5736775   ZANB4
## 2015  2015  4.0966518  3.0781639 1.5795703  13.404500 0.7513853   ZANB4
## 2016  2016  7.3123070  7.2693022 1.6866695  30.187021 0.9941189   ZANB4
## 2017  2017  5.7229576  6.0404958 1.9526777  23.622803 1.0554850   ZANB4
## 2018  2018  5.4175088  3.9274955 2.2156810  17.172143 0.7249634   ZANB4
## 2019  2019 11.4012690  6.5475691 5.6573970  31.115376 0.5742842   ZANB4
## 2020  2020 15.0051259 12.3365022 2.2786777  48.938640 0.8221525   ZANB4
## 2021  2021  8.5810661  5.3235919 3.7308419  24.714389 0.6203882   ZANB4
## 19871 1987  7.6172431  2.4337701 4.2632915  13.383109 0.3195080 nominal
## 19881 1988  5.9326923  2.1009913 3.1269356  10.952922 0.3541379 nominal
## 19891 1989  9.2143008  3.3726290 4.5707128  17.199956 0.3660212 nominal
## 19901 1990  5.1067940  1.5162224 2.9503867   9.053101 0.2969030 nominal
## 19911 1991 10.4412613  2.0025557 7.1520546  14.890882 0.1917925 nominal
## 19921 1992  5.4271972  1.8964944 2.8512258   9.898468 0.3494427 nominal
```

```
## 19931 1993  3.4334556  0.7531264 2.3078929   5.162610 0.2193494 nominal
## 19941 1994  1.8774332  0.5225342 1.1505533   3.178766 0.2783237 nominal
## 19951 1995  2.2096637  0.9072889 1.0264290   4.533934 0.4106004 nominal
## 19961 1996  4.2733913  1.1477617 2.4652299   6.992844 0.2685833 nominal
## 1997  1997  5.9131411  2.0055882 3.3380375  11.161340 0.3391748 nominal
## 19981 1998  2.0055070  0.5599937 1.1132019   3.249192 0.2792280 nominal
## 19991 1999  1.7135656  0.6597864 0.8496928   3.271598 0.3850371 nominal
## 20001 2000  2.2328797  0.7234597 1.1365324   3.920393 0.3240030 nominal
## 20011 2001 10.9218314  2.5198539 6.8918184  16.831323 0.2307172 nominal
## 20021 2002  3.8240406  0.9725425 2.3640156   6.065029 0.2543233 nominal
## 20031 2003  1.7949660  0.4515031 1.0610331   2.952335 0.2515385 nominal
## 20041 2004  2.9819912  0.9149456 1.6472515   5.219134 0.3068237 nominal
## 20051 2005  7.4269511  2.1562930 4.4940676  12.847822 0.2903335 nominal
## 20061 2006  2.8236664  0.7909153 1.6314781   4.766222 0.2801022 nominal
## 20071 2007  2.7173495  0.7247170 1.7215551   4.547209 0.2667000 nominal
## 20081 2008  6.0093921  1.4315289 3.5212399   8.958721 0.2382153 nominal
## 20091 2009  4.9846550  1.1086200 3.2233916   7.420836 0.2224066 nominal
## 20101 2010  3.4894276  0.6904951 2.3904174   4.971006 0.1978821 nominal
## 20111 2011  6.6763465  1.3727467 4.2007861   9.631277 0.2056135 nominal
## 20121 2012  4.1379212  0.7133568 2.9111621   5.692216 0.1723950 nominal
## 20131 2013  7.2705344  1.3559419 5.1880067  10.369116 0.1864982 nominal
## 20141 2014  7.6755588  1.6789857 5.1220554  11.706200 0.2187444 nominal
## 20151 2015  3.2869597  0.8757269 2.0347654   5.368609 0.2664246 nominal
## 20161 2016  2.6316422  0.7761426 1.5447420   4.605915 0.2949271 nominal
## 20171 2017  4.8192585  0.9760656 3.2792787   7.108166 0.2025344 nominal
## 20181 2018  5.7538812  1.2731570 3.7942133   8.661626 0.2212693 nominal
## 20191 2019  8.7360530  2.1663776 5.7785771  14.192439 0.2479813 nominal
## 20201 2020  3.8266146  1.0336339 2.2897335   6.459929 0.2701171 nominal
## 20211 2021  7.6024120  1.7843406 4.7077402  11.643868 0.2347072 nominal
```

```r
ggplot(tmp, aes(x=as.numeric(Year), y=Index, color=Method, shape=Method)) +
  geom_ribbon(aes(x=as.numeric(Year), ymin=LCI, ymax=UCI, fill=Method),
alpha=0.5) +
  geom_line() + geom_point() +
  xlab("Year") + theme_bw()
```




```
matplot(unique(unfactor(dat$YEAR)),sapply(split(tmp,tmp$Method), function(x)
scaleToMean(x$Index)),
        ylab="Index scaled to mean",type="o",xlab="Year")

## Error in unfactor(dat$YEAR): You are trying to convert something that is
not a factor

round(cor(sapply(split(tmp,tmp$Method), function(x) x$Index)),2)

## Error in cor(sapply(split(tmp, tmp$Method), function(x) x$Index)): supply
both 'x' and 'y' or a matrix-like 'x'

round(cor(sapply(split(tmp,tmp$Method), function(x)
x$Index),method="spearman"),2)

## Error in cor(sapply(split(tmp, tmp$Method), function(x) x$Index), method =
"spearman"): supply both 'x' and 'y' or a matrix-like 'x'

#AICtab(ZANB4,ZANB4.4)
Sys.time()

## [1] "2022-06-13 17:45:54 EDT"
```

Summary:

- Recall that the first iteration of model fitting was done with all fish collected (not just YOYs); the final dataset from NY pared down to just YOYs (and that is why this script needed to be updated). Final model glm index done with just YOYs

- No DOs collected in 1997, so that year dropped out early.
- Excluded Oyster Bay (sampling not started until 2001).
- Model diagnostics weren't perfect so I stepped through a number of remedial measures, some of which, got the diagnostics to good or good enough. The final model (ZANB4) didn't have the best diagnostics, but they appeared 'good enough' and ZANB4 had the lowest AIC.
- Exploration of a truncated time series that included Oyster Bay (1991+) performed similar to the entire time series models.

Bluefish index standardization - NJ Delaware River Seine

M Celestino

2022-06-07 10:52:46

Bluefish index standardization

Step 1: Data processing

Load functions & libraries

```
source("f:\\BF_WG\\FIG\\figFuns.R")
source("f:\\BF_WG\\FIG\\originalCode\\KDrew2\\bootstrap_functions_AT2.R")

library(car, quietly = FALSE)

## Warning: package 'car' was built under R version 4.1.2

## Loading required package: carData

## Warning: package 'carData' was built under R version 4.1.2

library(lmtest, quietly = FALSE)

## Warning: package 'lmtest' was built under R version 4.1.2

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

library(glmmTMB, quietly = FALSE)
library(DHARMa, quietly = FALSE)

## Warning: package 'DHARMa' was built under R version 4.1.2

## This is DHARMa 0.4.5. For overview type '?DHARMa'. For recent changes,
type news(package = 'DHARMa')

library(bbmle, quietly = FALSE)

## Loading required package: stats4

library(ggeffects, quietly = FALSE)
library(ggplot2, quietly = FALSE)
```

```
## Warning: package 'ggplot2' was built under R version 4.1.2

require(readxl, quietly = FALSE)

## Loading required package: readxl

require(lattice, quietly = FALSE)

## Loading required package: lattice

## Warning: package 'lattice' was built under R version 4.1.2

require(emmeans, quietly = FALSE)

## Loading required package: emmeans

## Warning: package 'emmeans' was built under R version 4.1.2

require(rmarkdown, quietly = FALSE)

## Loading required package: rmarkdown

## Warning: package 'rmarkdown' was built under R version 4.1.2

require(knitr, quietly = FALSE)

## Loading required package: knitr

## Warning: package 'knitr' was built under R version 4.1.2

library(reshape2, quietly=TRUE, verbose=FALSE)

## Warning: package 'reshape2' was built under R version 4.1.2

library(doBy, quietly=TRUE, verbose=FALSE)

## Warning: package 'doBy' was built under R version 4.1.2

library(dplyr, quietly=TRUE, verbose=FALSE,warn.conflicts=FALSE)
```

Read in data

```
indata <- read.csv("f:\\BF_WG\\FIG\\njSeine\\blueSeineData2021.csv", header =
TRUE)
head(indata)

##   Year id STRATUM TIDE   TOW TEMPBOT DOBOT SALBOT  dataMonth .order MONTH
## 1 1980  1       1   NA 449NJ    26.5   8.5   13.2 septemberA      7     9
## 2 1980  2       1    7 449NJ    17.5   7.7   15.2   octoberA      9    10
## 3 1980  3       1    7   505    25.0   7.1    9.0 septemberA      7     9
## 4 1980  4       1    7   515    18.2   7.7   14.0   octoberA      9    10
## 5 1980  5       1    7     1      NA    NA     NA septemberA      7     9
## 6 1980  6       1    7     1    19.0   7.3   10.4   octoberA      9    10
##   SEASON CPUE
## 1      3    1
```

```
## 2      4    0
## 3      3    0
## 4      4    0
## 5      3    0
## 6      4    3
```

```
names(indata)[1] <- "YEAR"
head(indata)
```

```
##   YEAR id STRATUM TIDE   TOW TEMPBOT DOBOT SALBOT  dataMonth .order MONTH
## 1 1980  1       1   NA 449NJ    26.5   8.5   13.2 septemberA      7     9
## 2 1980  2       1    7 449NJ    17.5   7.7   15.2   octoberA      9    10
## 3 1980  3       1    7   505    25.0   7.1    9.0 septemberA      7     9
## 4 1980  4       1    7   515    18.2   7.7   14.0   octoberA      9    10
## 5 1980  5       1    7     1      NA    NA     NA septemberA      7     9
## 6 1980  6       1    7     1    19.0   7.3   10.4   octoberA      9    10
##   SEASON CPUE
## 1      3    1
## 2      4    0
## 3      3    0
## 4      4    0
## 5      3    0
## 6      4    3
```

```
summary(indata)
```

```
##       YEAR            id           STRATUM           TIDE      
##  Min.   :1980   Min.   :  1.0   Min.   :1.000   Min.   :1.000  
##  1st Qu.:1997   1st Qu.: 59.0   1st Qu.:1.000   1st Qu.:2.000  
##  Median :2005   Median :123.0   Median :2.000   Median :3.000  
##  Mean   :2005   Mean   :138.6   Mean   :1.989   Mean   :2.803  
##  3rd Qu.:2013   3rd Qu.:198.0   3rd Qu.:3.000   3rd Qu.:4.000  
##  Max.   :2021   Max.   :732.0   Max.   :3.000   Max.   :8.000  
##                                                 NA's   :220    
##      TOW               TEMPBOT          DOBOT            SALBOT       
##  Length:8830        Min.   : 5.00   Min.   : 0.000   Min.   : 0.0000  
##  Class :character   1st Qu.:18.70   1st Qu.: 6.100   1st Qu.: 0.1000  
##  Mode  :character   Median :24.00   Median : 6.900   Median : 0.1000  
##                     Mean   :22.26   Mean   : 6.989   Mean   : 0.9882  
##                     3rd Qu.:26.50   3rd Qu.: 7.800   3rd Qu.: 0.9000  
##                     Max.   :34.00   Max.   :16.500   Max.   :16.0000  
##                     NA's   :90      NA's   :406      NA's   :223      
##   dataMonth             .order           MONTH            SEASON     
##  Length:8830        Min.   : 2.000   Min.   : 6.000   Min.   :2.000  
##  Class :character   1st Qu.: 5.000   1st Qu.: 8.000   1st Qu.:3.000  
##  Mode  :character   Median : 7.000   Median : 9.000   Median :3.000  
##                     Mean   : 6.753   Mean   : 8.657   Mean   :3.246  
##                     3rd Qu.: 9.000   3rd Qu.:10.000   3rd Qu.:4.000  
##                     Max.   :12.000   Max.   :12.000   Max.   :4.000  
##                                                                      
##       CPUE
```

```
##  Min.   : 0.0000
##  1st Qu.: 0.0000
##  Median : 0.0000
##  Mean   : 0.3952
##  3rd Qu.: 0.0000
##  Max.   :27.0000
##
```

Some data exploration to determine which years and covariates to keep; then remove missing values

```
# limit to just yrs >2001 (length data not recorded prior to 2001, and June
sampling started 2002) & only regions 1 & 2
indata <- indata[indata$YEAR>2001,]
indata$dataMonth <- reorder(indata$dataMonth,indata$.order,median)

keepVars <- c("STRATUM","TEMPBOT","DOBOT","SALBOT","TIDE","dataMonth")

# how many records would we lose to NAs:
apply(indata[,keepVars], 2, function(x) sum(is.na(x)))
```

```
##   STRATUM   TEMPBOT     DOBOT    SALBOT      TIDE dataMonth
##         0        71       347       155         0         0
```

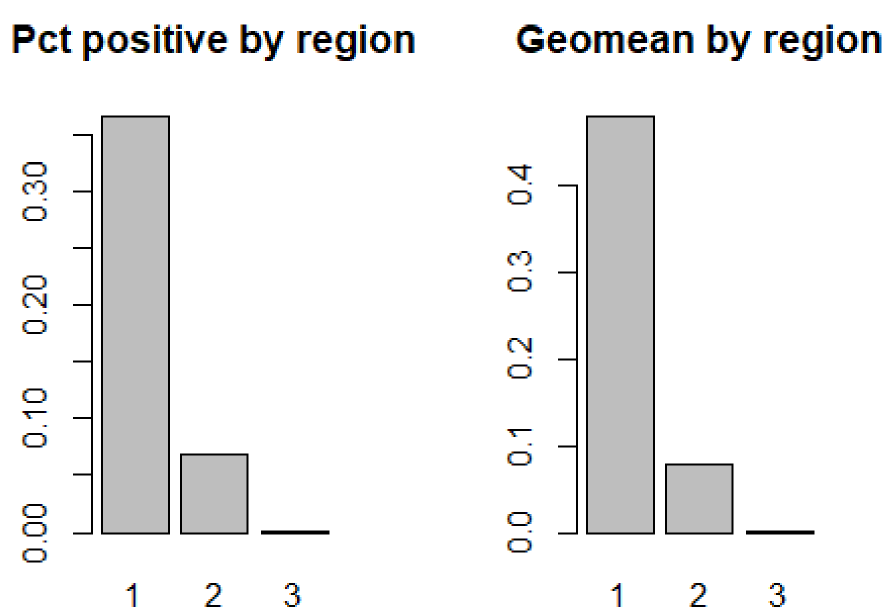
```
with(indata, tapply(CPUE,list(YEAR,STRATUM),geoMean,alpha=NULL,warn=FALSE))
```

```
##              1          2           3
## 2002 0.4107364 0.16323137 0.000000000
## 2003 0.4296663 0.00000000 0.000000000
## 2004 0.3463416 0.01585158 0.000000000
## 2005 0.6858573 0.23089223 0.000000000
## 2006 0.6827602 0.07659534 0.000000000
## 2007 0.7088958 0.15313884 0.000000000
## 2008 0.5868975 0.18674405 0.000000000
## 2009 0.5165609 0.02031792 0.000000000
## 2010 0.5902168 0.18352290 0.000000000
## 2011 0.3187649 0.01047778 0.000000000
## 2012 0.3134526 0.04793263 0.000000000
## 2013 0.5677036 0.01781666 0.000000000
## 2014 0.8499345 0.04251849 0.000000000
## 2015 0.3738383 0.05335500 0.000000000
## 2016 0.2508176 0.18842251 0.000000000
## 2017 0.6873056 0.04218962 0.009540389
## 2018 0.1810787 0.00000000 0.000000000
## 2019 0.3260934 0.02930224 0.000000000
## 2021 0.3936103 0.02213018 0.000000000
```

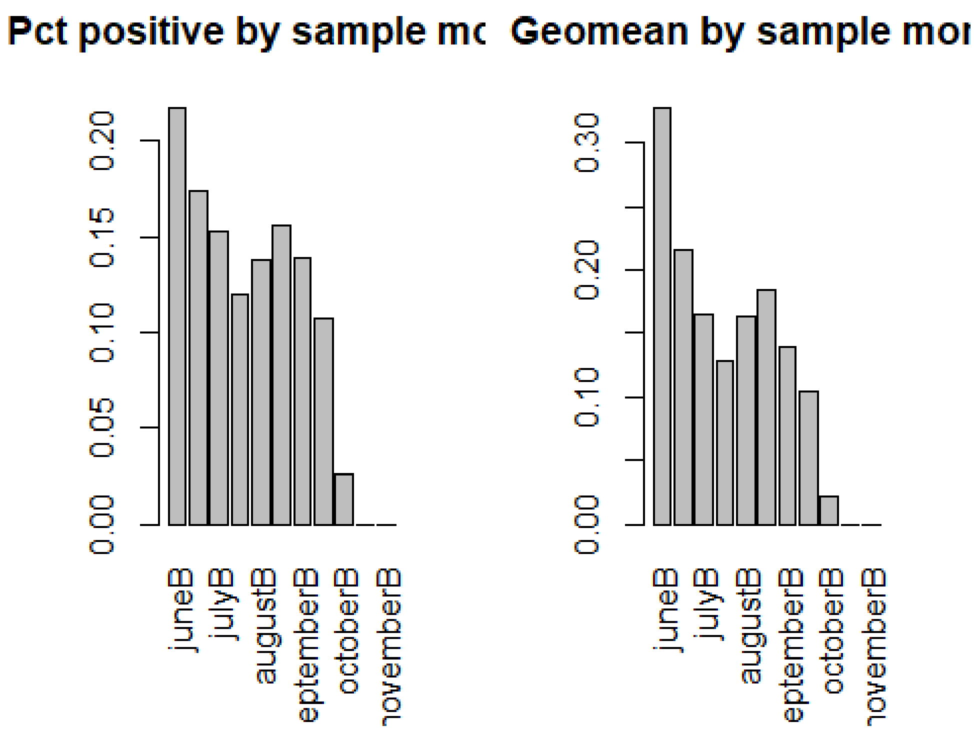
```
with(indata, tapply(CPUE,list(YEAR,STRATUM),function(x)
{sum(x>0)/length(x)}))
```

```
##              1          2          3
## 2002 0.3250000 0.16875000 0.00000000
## 2003 0.3164557 0.00000000 0.00000000
## 2004 0.2875000 0.01898734 0.00000000
## 2005 0.4625000 0.16455696 0.00000000
## 2006 0.5000000 0.06250000 0.00000000
## 2007 0.4625000 0.13750000 0.00000000
## 2008 0.4788732 0.14583333 0.00000000
## 2009 0.3875000 0.02531646 0.00000000
## 2010 0.3750000 0.13125000 0.00000000
## 2011 0.2250000 0.01503759 0.00000000
## 2012 0.2750000 0.04375000 0.00000000
## 2013 0.4250000 0.02547771 0.00000000
## 2014 0.5789474 0.04411765 0.00000000
## 2015 0.2875000 0.04605263 0.00000000
## 2016 0.2500000 0.11888112 0.00000000
## 2017 0.4927536 0.05555556 0.01369863
## 2018 0.2000000 0.00000000 0.00000000
## 2019 0.2638889 0.04166667 0.00000000
## 2021 0.3333333 0.03157895 0.00000000
```

```
op <- par(mfrow=c(1,2))
barplot(with(indata, tapply(CPUE,STRATUM,function(x)
{sum(x>0)/length(x)})),main="Pct positive by region")
barplot(with(indata,
tapply(CPUE,STRATUM,geoMean,alpha=NULL,warn=FALSE)),main="Geomean by region")
```



```
barplot(with(indata, tapply(CPUE,dataMonth,function(x)
{sum(x>0)/length(x)})),main="Pct positive by sample month",las=3)
barplot(with(indata,
tapply(CPUE,dataMonth,geoMean,alpha=NULL,warn=FALSE)),main="Geomean by sample
month",las=3)
```

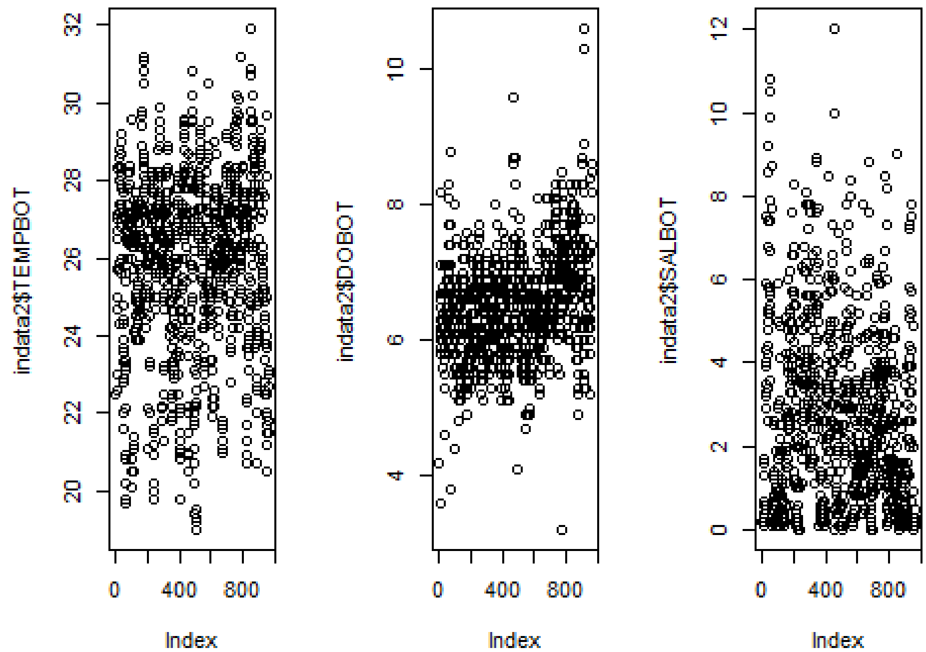



```
# So subsetting to just regio 1 and just juneB to septB
indata2 <- indata[indata$STRATUM<2, ]
# just region 1
indata2 <- indata2[is.element(indata2$.order,2:8),]; table(indata2$dataMonth)
# juneB to septB

##
##      juneB      julyA      julyB    augustA    augustB septemberA
septemberB
##        144        142        144        152        152        154
150
##   octoberA   octoberB  novemberA  novemberB
##          0          0          0          0

indata2 <- indata2[complete.cases(indata2[,keepVars]),] # keep only complete
cases
summary(indata2)

##       YEAR            id             STRATUM       TIDE            TOW
##  Min.   :2002   Min.   :  1.0   Min.   :1   Min.   :1.000   Length:963
##  1st Qu.:2006   1st Qu.: 49.5   1st Qu.:1   1st Qu.:1.000   Class
```

```
:character
##  Median :2011   Median :103.0   Median :1   Median :2.000   Mode
:character
##  Mean   :2011   Mean   :105.8   Mean   :1   Mean   :2.594
##  3rd Qu.:2015   3rd Qu.:161.0   3rd Qu.:1   3rd Qu.:4.000
##  Max.   :2021   Max.   :224.0   Max.   :1   Max.   :6.000
##
##     TEMPBOT          DOBOT            SALBOT           dataMonth
##  Min.   :19.00   Min.   : 3.200   Min.   : 0.00   septemberB:148
##  1st Qu.:24.50   1st Qu.: 6.000   1st Qu.: 0.70   augustB   :144
##  Median :26.30   Median : 6.500   Median : 2.20   septemberA:138
##  Mean   :25.89   Mean   : 6.531   Mean   : 2.62   julyA     :135
##  3rd Qu.:27.50   3rd Qu.: 7.000   3rd Qu.: 3.90   julyB     :135
##  Max.   :31.90   Max.   :10.600   Max.   :12.00   juneB     :132
##                                                    (Other)   :131
##      .order          MONTH           SEASON           CPUE
##  Min.   :2.000   Min.   :6.000   Min.   :2.000   Min.   : 0.000
##  1st Qu.:3.000   1st Qu.:7.000   1st Qu.:3.000   1st Qu.: 0.000
##  Median :5.000   Median :8.000   Median :3.000   Median : 0.000
##  Mean   :5.065   Mean   :7.751   Mean   :2.871   Mean   : 1.072
##  3rd Qu.:7.000   3rd Qu.:9.000   3rd Qu.:3.000   3rd Qu.: 1.000
##  Max.   :8.000   Max.   :9.000   Max.   :3.000   Max.   :27.000
##
```

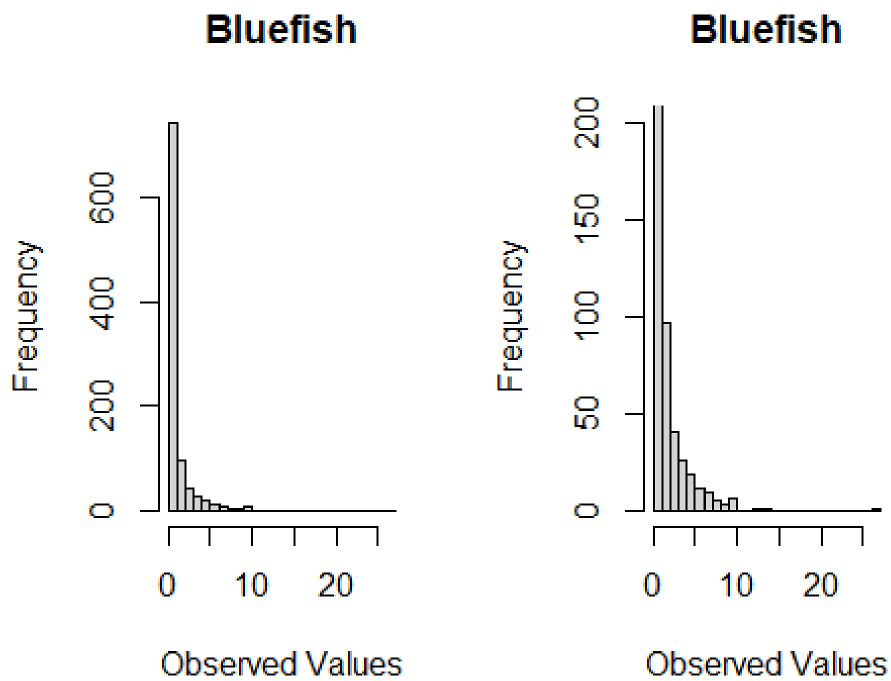
Check for outliers

```
# all values generally seem ok
op <- par(mfrow=c(1,3))
plot(indata2$TEMPBOT)
plot(indata2$DOBOT)
plot(indata2$SALBOT)
```



Plot distribution of catches:

```
op <- par(mfrow=c(1,2))
hist(indata2$CPUE,main="Bluefish",xlab="Observed Values",ylab="Frequency",
     breaks=seq(0,max(indata2$CPUE),1))
hist(indata2$CPUE,main="Bluefish",xlab="Observed Values",ylab="Frequency",
     breaks=seq(0,max(indata2$CPUE),1),ylim=c(0,200))
```




```
par(op)
```

Finally, check to see if there are any years with zero catch that need to be removed:

```
diff(as.numeric(names(xtabs(CPUE~YEAR, indata2))))

##  [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2
```

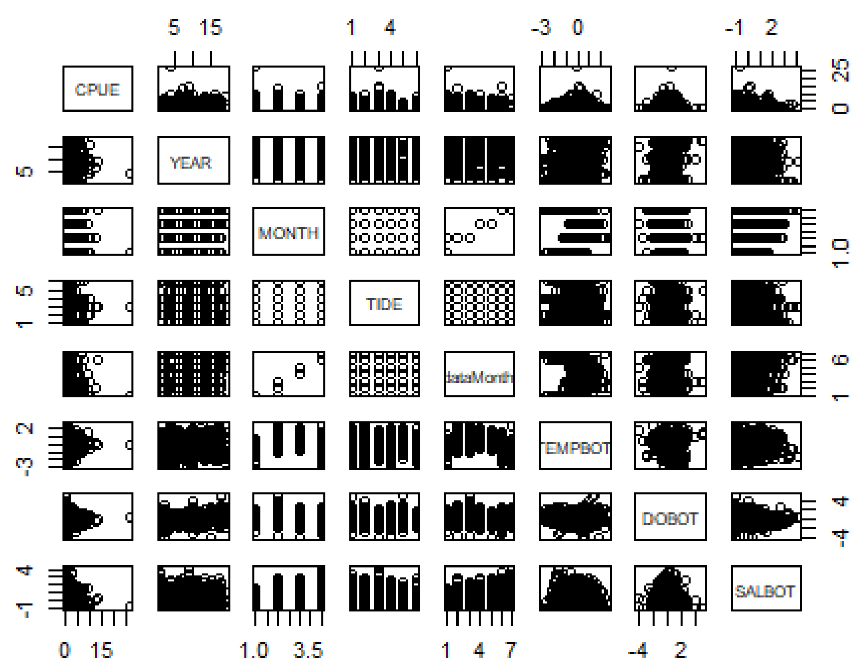
Create a dataframe with all the factors we need (z-score all continuous variables; all others as.factors)

```
dat = data.frame(CPUE = indata2$CPUE,
                 YEAR = as.factor(indata2$YEAR),
                 MONTH = as.factor(indata2$MONTH),
                 STRATUM = as.factor(indata2$STRATUM),
                 TIDE = as.factor(indata2$TIDE),
                 dataMonth = as.factor(indata2$dataMonth),
                 TEMPBOT = Z.scr(indata2$TEMPBOT),
                 DOBOT = Z.scr(indata2$DOBOT),
                 SALBOT = Z.scr(indata2$SALBOT))

# Calculate the proportion of positive tows/sets/hauls
dat$PosTow <- ifelse(dat$CPUE > 0, 1, 0)
mean(dat$PosTow) # 45%

## [1] 0.4423676

head(dat)
```

```
##   CPUE YEAR MONTH STRATUM TIDE dataMonth    TEMPBOT       DOBOT     SALBOT
## 1    4 2002     6       1    5     juneB -1.4586272 -2.94184538 -1.0921188
## 2    0 2002     6       1    5     juneB -1.4155620 -0.67019767 -1.1372411
## 3    5 2002     6       1    3     juneB -1.2433011  0.21322089 -1.0921188
## 4    6 2002     6       1    3     juneB -1.2433011 -0.29158972 -1.0469965
## 5    2 2002     6       1    3     juneB -0.3819969  0.71803149 -0.4152844
## 6    1 2002     6       1    4     juneB -0.5111925 -0.03918442 -0.5957736
##   PosTow
## 1      1
## 2      0
## 3      1
## 4      1
## 5      1
## 6      1
```

Check collinearity

```
pairs(~CPUE+YEAR+MONTH+TIDE+dataMonth+TEMPBOT+DOBOT+SALBOT,data=dat)
```



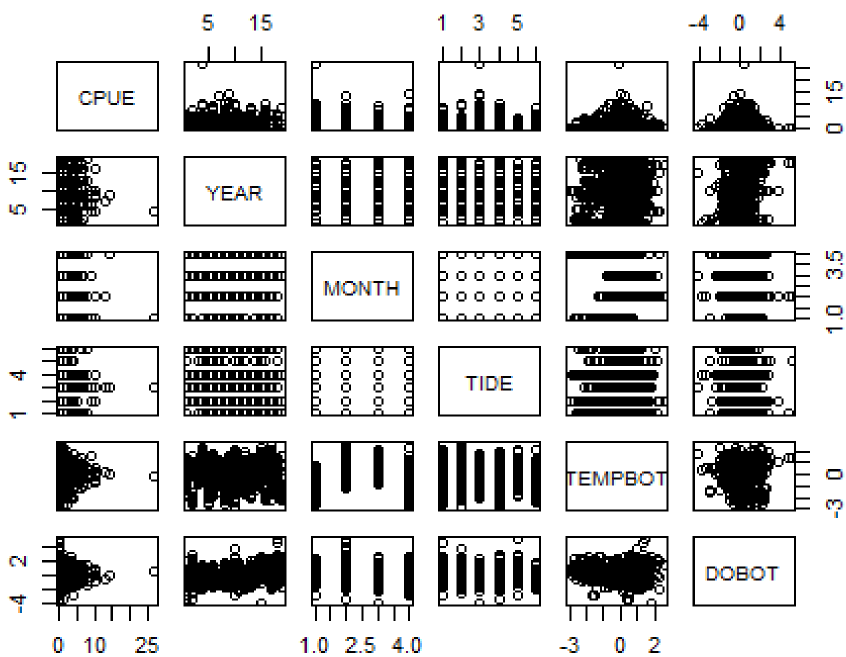

```
# Check the variance inflation factor for a more statistical check.
mod = lm(CPUE ~ YEAR+MONTH+TIDE+TEMPBOT+DOBOT+SALBOT, data = dat)     # won't
run with both month and dataMonth (so I removed dataMonth)
vif(mod)

##              GVIF Df GVIF^(1/(2*Df))
## YEAR     3.491343 18        1.035340
## MONTH    2.419357  3        1.158644
```

```
## TIDE    1.341003  5        1.029776
## TEMPBOT 2.127597  1        1.458628
## DOBOT   1.380355  1        1.174885
## SALBOT  1.864598  1        1.365503
```

```
# Drop SALBOT based on pairs plot; removing MONTH resulted in YEAR=2.1
mod = lm(CPUE ~ YEAR+MONTH+TIDE+TEMPBOT+DOBOT, data = dat)
vif(mod)
```

```
##             GVIF Df GVIF^(1/(2*Df))
## YEAR    2.159315 18        1.021613
## MONTH   2.067990  3        1.128734
## TIDE    1.331807  5        1.029068
## TEMPBOT 2.115109  1        1.454342
## DOBOT   1.380275  1        1.174851
```

```
pairs(~ CPUE+YEAR+MONTH+TIDE+TEMPBOT+DOBOT, data=dat)
```




```
round(with(dat, cor(cbind(CPUE,as.numeric(as.character(YEAR)),

as.numeric(as.character(MONTH)),as.numeric(as.character(TIDE)),TEMPBOT,DOBOT)
)),2)
```

```
##          CPUE                   TEMPBOT DOBOT
## CPUE     1.00 -0.09 -0.26  0.06    0.01  0.00
##         -0.09  1.00  0.05 -0.07    0.04  0.32
##         -0.26  0.05  1.00 -0.01   -0.27 -0.11
##          0.06 -0.07 -0.01  1.00    0.04 -0.06
```

```
## TEMPBOT  0.01  0.04 -0.27  0.04    1.00 -0.16
## DOBOT    0.00  0.32 -0.11 -0.06   -0.16  1.00
```

```
summary(lm(DOBOT~as.numeric(as.character(YEAR)),data=dat))
```

```
##
## Call:
## lm(formula = DOBOT ~ as.numeric(as.character(YEAR)), data = dat)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.5141 -0.5674 -0.0313  0.5684  4.6434
##
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    -1.216e+02  1.151e+01  -10.57   <2e-16 ***
## as.numeric(as.character(YEAR))  6.049e-02  5.725e-03   10.57   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.947 on 961 degrees of freedom
## Multiple R-squared:  0.1041, Adjusted R-squared:  0.1032
## F-statistic: 111.6 on 1 and 961 DF,  p-value: < 2.2e-16
```

Step 2: Model selection

Compare negative binom, zero-infl neg binom, zero-altered neg binom

Model 1: Negative Binomial

```
tmb1.NB = glmmTMB(CPUE ~ YEAR+MONTH+TIDE+TEMPBOT+DOBOT, data = dat, family =
nbinom2)

summary(tmb1.NB)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + TEMPBOT + DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2562.8   2708.9  -1251.4   2502.8      933
##
##
## Dispersion parameter for nbinom2 family (): 0.891
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) 0.771826   0.254532   3.032 0.002427 **
## YEAR2003    0.375196   0.301853   1.243 0.213877
## YEAR2004    0.009939   0.346475   0.029 0.977114
## YEAR2005    0.609835   0.282852   2.156 0.031082 *
## YEAR2006    0.172731   0.323609   0.534 0.593504
```

```
## YEAR2007      0.751340   0.282680   2.658 0.007862 **
## YEAR2008      0.297093   0.292535   1.016 0.309829
## YEAR2009      0.285803   0.292973   0.976 0.329299
## YEAR2010      0.621865   0.284927   2.183 0.029069 *
## YEAR2011      0.049164   0.321410   0.153 0.878427
## YEAR2012     -0.220089   0.312244  -0.705 0.480896
## YEAR2013      0.381073   0.286145   1.332 0.182943
## YEAR2014      0.744675   0.280288   2.657 0.007888 **
## YEAR2015     -0.533291   0.321409  -1.659 0.097070 .
## YEAR2016     -0.795012   0.350050  -2.271 0.023139 *
## YEAR2017      0.198670   0.295138   0.673 0.500856
## YEAR2018     -0.715246   0.360876  -1.982 0.047482 *
## YEAR2019     -0.168026   0.311133  -0.540 0.589166
## YEAR2021      0.032188   0.477817   0.067 0.946291
## MONTH7       -0.792705   0.169023  -4.690 2.73e-06 ***
## MONTH8       -1.350921   0.176870  -7.638 2.21e-14 ***
## MONTH9       -1.241808   0.158619  -7.829 4.92e-15 ***
## TIDE2        -0.528911   0.144888  -3.650 0.000262 ***
## TIDE3         0.501300   0.153301   3.270 0.001075 **
## TIDE4        -0.105492   0.158015  -0.668 0.504381
## TIDE5         0.106635   0.254830   0.418 0.675615
## TIDE6        -0.276543   0.270735  -1.021 0.307040
## TEMPBOT       0.102389   0.078160   1.310 0.190202
## DOBOT        -0.025874   0.062258  -0.416 0.677706
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
disp3(tmb1.NB)
```

```
## df.resid
##  1.16382
```

Model 2: ZINB

```
ZINB = glmmTMB(CPUE ~ YEAR+MONTH+TIDE+TEMPBOT+DOBOT,
               ziformula = ~ YEAR+MONTH+TIDE+TEMPBOT+DOBOT, data = dat,
family=nbinom2,

control=glmmTMBControl(optCtrl=list(iter.max=5e6,eval.max=5e6)))
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')
```

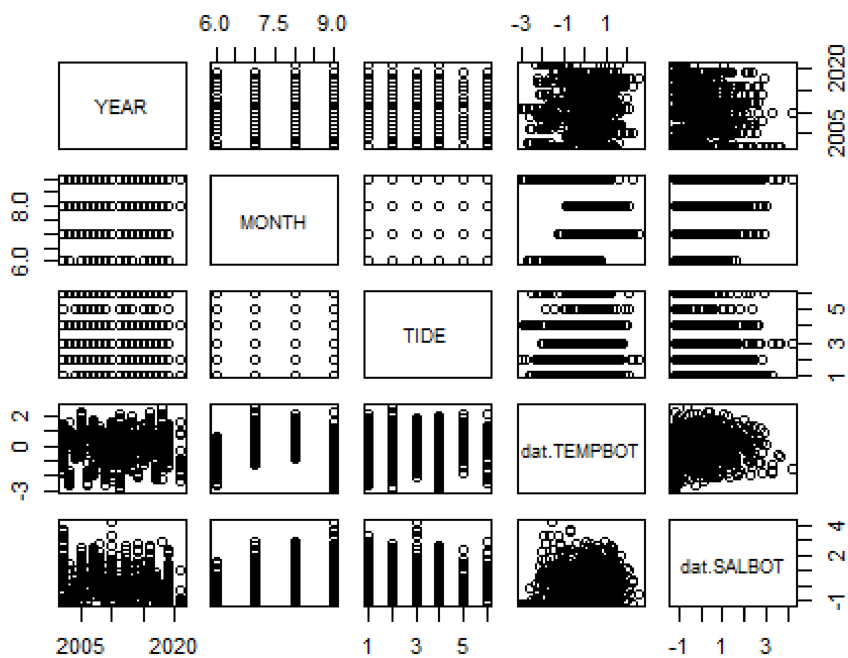
```
summary(ZINB)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + TEMPBOT + DOBOT
## Zero inflation:        ~YEAR + MONTH + TIDE + TEMPBOT + DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
```

```
##        NA        NA        NA        NA       904
##
##
## Dispersion parameter for nbinom2 family (): 1.14
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.51033        NaN     NaN      NaN
## YEAR2003     0.61138        NaN     NaN      NaN
## YEAR2004     0.16304        NaN     NaN      NaN
## YEAR2005     0.77783        NaN     NaN      NaN
## YEAR2006     0.10747        NaN     NaN      NaN
## YEAR2007     0.80376        NaN     NaN      NaN
## YEAR2008     0.44683        NaN     NaN      NaN
## YEAR2009     0.22898        NaN     NaN      NaN
## YEAR2010     0.75770        NaN     NaN      NaN
## YEAR2011     0.65416        NaN     NaN      NaN
## YEAR2012     0.23898        NaN     NaN      NaN
## YEAR2013     0.41221        NaN     NaN      NaN
## YEAR2014     0.79814        NaN     NaN      NaN
## YEAR2015    -0.12870        NaN     NaN      NaN
## YEAR2016    -0.53940        NaN     NaN      NaN
## YEAR2017     0.54606        NaN     NaN      NaN
## YEAR2018    -0.20802        NaN     NaN      NaN
## YEAR2019     0.05881        NaN     NaN      NaN
## YEAR2021     0.91731        NaN     NaN      NaN
## MONTH7      -0.64575        NaN     NaN      NaN
## MONTH8      -1.01331        NaN     NaN      NaN
## MONTH9      -1.20514        NaN     NaN      NaN
## TIDE2       -0.54965        NaN     NaN      NaN
## TIDE3        0.58218        NaN     NaN      NaN
## TIDE4        0.07342        NaN     NaN      NaN
## TIDE5        0.08960        NaN     NaN      NaN
## TIDE6        0.14425        NaN     NaN      NaN
## TEMPBOT     -0.15697        NaN     NaN      NaN
## DOBOT       -0.07447        NaN     NaN      NaN
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3071.96        NaN     NaN      NaN
## YEAR2003      957.62        NaN     NaN      NaN
## YEAR2004      826.83        NaN     NaN      NaN
## YEAR2005       75.10        NaN     NaN      NaN
## YEAR2006    -1601.54        NaN     NaN      NaN
## YEAR2007    -1632.00        NaN     NaN      NaN
## YEAR2008     1495.75        NaN     NaN      NaN
## YEAR2009      -10.99        NaN     NaN      NaN
## YEAR2010      589.63        NaN     NaN      NaN
## YEAR2011     1320.07        NaN     NaN      NaN
## YEAR2012     2094.09        NaN     NaN      NaN
```

```
## YEAR2013      -50.11        NaN       NaN       NaN
## YEAR2014     -136.39        NaN       NaN       NaN
## YEAR2015     1298.53        NaN       NaN       NaN
## YEAR2016     -193.71        NaN       NaN       NaN
## YEAR2017     1564.66        NaN       NaN       NaN
## YEAR2018     2282.00        NaN       NaN       NaN
## YEAR2019      927.63        NaN       NaN       NaN
## YEAR2021     2290.02        NaN       NaN       NaN
## MONTH7        633.84        NaN       NaN       NaN
## MONTH8       1806.34        NaN       NaN       NaN
## MONTH9       -134.01        NaN       NaN       NaN
## TIDE2        -678.85        NaN       NaN       NaN
## TIDE3         525.40        NaN       NaN       NaN
## TIDE4         785.57        NaN       NaN       NaN
## TIDE5        -112.29        NaN       NaN       NaN
## TIDE6        1282.12        NaN       NaN       NaN
## TEMPBOT     -1130.93        NaN       NaN       NaN
## DOBOT        -128.81        NaN       NaN       NaN
```

```
pairs(data.frame("YEAR"=as.numeric(as.character(dat$YEAR)),
   "MONTH"=as.numeric(as.character(dat$MONTH)),
   "TIDE"=as.numeric(as.character(dat$TIDE)),
   dat$TEMPBOT,dat$SALBOT))
```




```
round(cor(data.frame("YEAR"=as.numeric(as.character(dat$YEAR)),
   "MONTH"=as.numeric(as.character(dat$MONTH)),
```

```
   "TIDE"=as.numeric(as.character(dat$TIDE)),
   dat$TEMPBOT,dat$SALBOT)),2)

##               YEAR MONTH  TIDE dat.TEMPBOT dat.SALBOT
## YEAR          1.00  0.05 -0.07        0.04      -0.14
## MONTH         0.05  1.00 -0.01       -0.27       0.27
## TIDE         -0.07 -0.01  1.00        0.04      -0.02
## dat.TEMPBOT   0.04 -0.27  0.04        1.00       0.07
## dat.SALBOT   -0.14  0.27 -0.02        0.07       1.00

# Month and temperature related, dropping temperature; on second look, I
could have dropped salinity

#fixef(ZINB)

# dropping TEMP in both parts due to corr with MONTH:
ZINB1 = glmmTMB(CPUE ~ YEAR+MONTH+TIDE+DOBOT, ziformula = ~
YEAR+MONTH+TIDE+DOBOT,
                data = dat, family=nbinom2,

control=glmmTMBControl(optCtrl=list(iter.max=5e6,eval.max=5e6)))

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

summary(ZINB1)

## Warning in sqrt(diag(vcovs)): NaNs produced

##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + DOBOT
## Zero inflation:        ~YEAR + MONTH + TIDE + DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##       NA       NA       NA       NA      906
##
##
## Dispersion parameter for nbinom2 family (): 1.31
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.58774    0.26074   2.254   0.0242 *
## YEAR2003     0.26509    0.33703   0.787   0.4315
## YEAR2004     0.18263    0.40835   0.447   0.6547
## YEAR2005     0.41096    0.28669   1.433   0.1517
## YEAR2006     0.10495    0.33184   0.316   0.7518
## YEAR2007     0.49292    0.29153   1.691   0.0909 .
## YEAR2008     0.22118    0.30702   0.720   0.4713
## YEAR2009     0.09360    0.30354   0.308   0.7578
```

```
## YEAR2010     0.60369    0.30110   2.005   0.0450 *
## YEAR2011     0.66901    0.39745   1.683   0.0923 .
## YEAR2012     0.14459    0.35819   0.404   0.6865
## YEAR2013     0.25895    0.30249   0.856   0.3920
## YEAR2014     0.49538    0.30062   1.648   0.0994 .
## YEAR2015    -0.41359    0.32933  -1.256   0.2092
## YEAR2016    -0.71284    0.37977  -1.877   0.0605 .
## YEAR2017     0.28211    0.31119   0.907   0.3646
## YEAR2018    -0.81900    0.40984  -1.998   0.0457 *
## YEAR2019    -0.28506    0.31731  -0.898   0.3690
## YEAR2021     1.74402    0.78176   2.231   0.0257 *
## MONTH7      -0.62322    0.14315  -4.354 1.34e-05 ***
## MONTH8      -0.88819    0.16490  -5.386 7.19e-08 ***
## MONTH9      -1.01547    0.15659  -6.485 8.88e-11 ***
## TIDE2       -0.46008    0.14095  -3.264   0.0011 **
## TIDE3        0.74352    0.15529   4.788 1.68e-06 ***
## TIDE4        0.27538    0.16618   1.657   0.0975 .
## TIDE5        0.11666    0.23694   0.492   0.6225
## TIDE6        0.45938    0.31637   1.452   0.1465
## DOBOT       -0.06903    0.06488  -1.064   0.2874
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -23.81193 1277.64768  -0.019 0.985130
## YEAR2003       -1.16857    1.16120  -1.006 0.314248
## YEAR2004        2.16631    1.60862   1.347 0.178079
## YEAR2005       -9.81736        NaN     NaN      NaN
## YEAR2006        1.66653    1.80835   0.922 0.356752
## YEAR2007      -21.22949 6393.75862  -0.003 0.997351
## YEAR2008        0.30839    1.50379   0.205 0.837512
## YEAR2009       -1.68415    1.92269  -0.876 0.381066
## YEAR2010       -0.80905    1.17385  -0.689 0.490683
## YEAR2011        4.18934    1.51936   2.757 0.005828 **
## YEAR2012        2.73400    1.53171   1.785 0.074271 .
## YEAR2013       -1.58631    1.27648  -1.243 0.213970
## YEAR2014       -2.58754    1.59015  -1.627 0.103687
## YEAR2015       -0.03447    1.55282  -0.022 0.982290
## YEAR2016       -0.63423    2.30623  -0.275 0.783313
## YEAR2017        1.08586    1.32358   0.820 0.411991
## YEAR2018       -1.96106    2.40468  -0.816 0.414776
## YEAR2019      -20.75081 4815.29600  -0.004 0.996562
## YEAR2021        5.48649    1.70327   3.221 0.001277 **
## MONTH7         17.39814 1277.64616   0.014 0.989135
## MONTH8         20.61748 1277.64632   0.016 0.987125
## MONTH9         19.68625 1277.64631   0.015 0.987707
## TIDE2           0.40695    0.88005   0.462 0.643784
## TIDE3           3.34220    1.14891   2.909 0.003626 **
## TIDE4           4.49985    1.19905   3.753 0.000175 ***
```

```
## TIDE5          -0.77474    2.31180  -0.335 0.737530
## TIDE6           6.46878    1.62886   3.971 7.15e-05 ***
## DOBOT          -0.47729    0.32064  -1.489 0.136611
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# dropping year from zero-infl
ZINB2 = glmmTMB(CPUE ~ YEAR+MONTH+TIDE+DOBOT, ziformula = ~ MONTH+TIDE+DOBOT,
data = dat, family=nbinom2)
summary(ZINB2)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + DOBOT
## Zero inflation:        ~MONTH + TIDE + DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2560.8   2750.7  -1241.4   2482.8      924
##
##
## Dispersion parameter for nbinom2 family (): 1.29
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.651140   0.253664   2.567 0.010260 *
## YEAR2003     0.357055   0.311822   1.145 0.252184
## YEAR2004    -0.006976   0.345480  -0.020 0.983890
## YEAR2005     0.632223   0.278269   2.272 0.023087 *
## YEAR2006     0.093203   0.311775   0.299 0.764983
## YEAR2007     0.686047   0.278483   2.464 0.013758 *
## YEAR2008     0.255286   0.291475   0.876 0.381115
## YEAR2009     0.180468   0.286378   0.630 0.528579
## YEAR2010     0.667435   0.286537   2.329 0.019842 *
## YEAR2011     0.102498   0.321585   0.319 0.749932
## YEAR2012    -0.111153   0.310973  -0.357 0.720766
## YEAR2013     0.367849   0.284996   1.291 0.196802
## YEAR2014     0.658677   0.280608   2.347 0.018909 *
## YEAR2015    -0.419099   0.314931  -1.331 0.183267
## YEAR2016    -0.642851   0.343808  -1.870 0.061512 .
## YEAR2017     0.258500   0.291684   0.886 0.375492
## YEAR2018    -0.619366   0.362769  -1.707 0.087761 .
## YEAR2019    -0.126716   0.306476  -0.413 0.679268
## YEAR2021     0.055314   0.478445   0.116 0.907960
## MONTH7      -0.682014   0.162898  -4.187 2.83e-05 ***
## MONTH8      -1.032898   0.169152  -6.106 1.02e-09 ***
## MONTH9      -1.101435   0.166074  -6.632 3.31e-11 ***
## TIDE2       -0.559581   0.167316  -3.344 0.000824 ***
## TIDE3        0.714020   0.175613   4.066 4.79e-05 ***
## TIDE4        0.229282   0.218786   1.048 0.294650
## TIDE5        0.073094   0.250032   0.292 0.770028
```

```
## TIDE6        0.102351   0.348910   0.293 0.769259
## DOBOT       -0.078989   0.079067  -0.999 0.317789
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -4.6560     2.3742  -1.961   0.0499 *
## MONTH7           0.6276     1.9227   0.326   0.7441
## MONTH8           2.3056     1.6446   1.402   0.1609
## MONTH9           2.0115     1.6779   1.199   0.2306
## TIDE2          -13.8742  2541.5795  -0.005   0.9956
## TIDE3            1.8132     1.7791   1.019   0.3081
## TIDE4            2.2304     2.0402   1.093   0.2743
## TIDE5          -18.9442 19397.4331  -0.001   0.9992
## TIDE6            2.4718     1.9926   1.240   0.2148
## DOBOT           -0.2681     0.4223  -0.635   0.5255
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# dropping tide from zero-infl (large SEs):
ZINB3 = glmmTMB(CPUE ~ YEAR+MONTH+TIDE+DOBOT, ziformula = ~ MONTH+DOBOT, data
= dat, family=nbinom2)
summary(ZINB3)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + DOBOT
## Zero inflation:        ~MONTH + DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2567.6   2733.2  -1249.8   2499.6      929
##
##
## Dispersion parameter for nbinom2 family (): 1.09
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.70244    0.24784   2.834 0.004593 **
## YEAR2003     0.40796    0.30506   1.337 0.181115
## YEAR2004     0.01218    0.35188   0.035 0.972379
## YEAR2005     0.70495    0.28098   2.509 0.012112 *
## YEAR2006     0.14508    0.31765   0.457 0.647857
## YEAR2007     0.75716    0.28015   2.703 0.006879 **
## YEAR2008     0.28897    0.29035   0.995 0.319624
## YEAR2009     0.23114    0.28799   0.803 0.422205
## YEAR2010     0.72153    0.28795   2.506 0.012218 *
## YEAR2011     0.12646    0.32602   0.388 0.698097
## YEAR2012    -0.15348    0.31475  -0.488 0.625811
## YEAR2013     0.37657    0.28369   1.327 0.184373
```

```
## YEAR2014     0.70260    0.27924   2.516 0.011866 *
## YEAR2015    -0.45119    0.31606  -1.428 0.153432
## YEAR2016    -0.63825    0.34682  -1.840 0.065722 .
## YEAR2017     0.26418    0.29043   0.910 0.363027
## YEAR2018    -0.62062    0.35595  -1.744 0.081241 .
## YEAR2019    -0.09820    0.30798  -0.319 0.749826
## YEAR2021     0.10240    0.48207   0.212 0.831779
## MONTH7      -0.65339    0.15707  -4.160 3.18e-05 ***
## MONTH8      -1.07647    0.21262  -5.063 4.13e-07 ***
## MONTH9      -1.19855    0.17867  -6.708 1.97e-11 ***
## TIDE2       -0.53584    0.14278  -3.753 0.000175 ***
## TIDE3        0.52859    0.15243   3.468 0.000525 ***
## TIDE4       -0.06246    0.15780  -0.396 0.692215
## TIDE5        0.08648    0.24917   0.347 0.728543
## TIDE6       -0.26712    0.27016  -0.989 0.322786
## DOBOT       -0.10201    0.07232  -1.410 0.158406
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -18.7948  2851.2762  -0.007    0.995
## MONTH7        15.8309  2851.2762   0.006    0.996
## MONTH8        17.0788  2851.2761   0.006    0.995
## MONTH9        16.3939  2851.2762   0.006    0.995
## DOBOT         -0.6626     0.4823  -1.374    0.169
```

```
# dropping month from zero-infl (large SEs):
ZINB4 = glmmTMB(CPUE ~ YEAR+MONTH+TIDE+DOBOT, ziformula = ~ DOBOT, data =
dat, family=nbinom2)
summary(ZINB4)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + DOBOT
## Zero inflation:          ~DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2565.6   2716.5  -1251.8   2503.6      932
##
##
## Dispersion parameter for nbinom2 family (): 0.952
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.72949    0.25233   2.891 0.003839 **
## YEAR2003     0.38745    0.30329   1.277 0.201436
## YEAR2004    -0.02564    0.34738  -0.074 0.941158
## YEAR2005     0.68557    0.28241   2.428 0.015201 *
## YEAR2006     0.13254    0.32045   0.414 0.679163
```

```
## YEAR2007     0.76873    0.28174   2.729 0.006362 **
## YEAR2008     0.30277    0.29235   1.036 0.300371
## YEAR2009     0.22254    0.28855   0.771 0.440572
## YEAR2010     0.67901    0.28477   2.384 0.017105 *
## YEAR2011     0.07784    0.32242   0.241 0.809220
## YEAR2012    -0.19619    0.31417  -0.624 0.532308
## YEAR2013     0.38050    0.28547   1.333 0.182575
## YEAR2014     0.72501    0.28022   2.587 0.009674 **
## YEAR2015    -0.47130    0.31767  -1.484 0.137911
## YEAR2016    -0.68601    0.34875  -1.967 0.049176 *
## YEAR2017     0.24282    0.29320   0.828 0.407571
## YEAR2018    -0.63167    0.35782  -1.765 0.077511 .
## YEAR2019    -0.11951    0.31001  -0.385 0.699869
## YEAR2021     0.05145    0.47825   0.108 0.914326
## MONTH7      -0.68860    0.15157  -4.543 5.54e-06 ***
## MONTH8      -1.25631    0.16221  -7.745 9.57e-15 ***
## MONTH9      -1.27216    0.15523  -8.195 2.50e-16 ***
## TIDE2       -0.53411    0.14414  -3.705 0.000211 ***
## TIDE3        0.50906    0.15317   3.323 0.000889 ***
## TIDE4       -0.09584    0.15759  -0.608 0.543063
## TIDE5        0.10213    0.25335   0.403 0.686867
## TIDE6       -0.29155    0.26966  -1.081 0.279602
## DOBOT       -0.07194    0.07300  -0.986 0.324372
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -4.040      2.235  -1.808   0.0706 .
## DOBOT         -1.036      0.710  -1.460   0.1443
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Standard errors are much better.

# We need a new function to get residuals for ZI-models.
tmbZI.resids(ZINB4,dat$CPUE)
```

```
## df.resid
## 1.182707
```

Model 3: ZANB (aka hurdle model)

```r
ZANB <- glmmTMB(CPUE ~ YEAR+MONTH+TIDE+TEMPBOT+DOBOT, ziformula = ~
YEAR+MONTH+TIDE+TEMPBOT+DOBOT, data = dat,family = truncated_nbinom2(link =
"log"))


summary(ZANB)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + TEMPBOT + DOBOT
## Zero inflation:        ~YEAR + MONTH + TIDE + TEMPBOT + DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2557.0   2844.3  -1219.5   2439.0      904
##
##
## Dispersion parameter for truncated_nbinom2 family (): 1.19
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.576748   0.345380   1.670 0.094941 .
## YEAR2003     0.249967   0.419017   0.597 0.550804
## YEAR2004    -0.112863   0.519567  -0.217 0.828032
## YEAR2005     0.564234   0.375176   1.504 0.132603
## YEAR2006    -0.216111   0.447050  -0.483 0.628801
## YEAR2007     0.515834   0.372784   1.384 0.166440
## YEAR2008    -0.009458   0.399611  -0.024 0.981118
## YEAR2009    -0.151569   0.392735  -0.386 0.699548
## YEAR2010     0.712405   0.380808   1.871 0.061376 .
## YEAR2011     0.724831   0.469808   1.543 0.122873
## YEAR2012     0.252705   0.453931   0.557 0.577729
## YEAR2013     0.100941   0.380435   0.265 0.790756
## YEAR2014     0.291814   0.369362   0.790 0.429498
## YEAR2015    -0.169967   0.466617  -0.364 0.715669
## YEAR2016    -0.443683   0.551671  -0.804 0.421251
## YEAR2017    -0.029094   0.400823  -0.073 0.942136
## YEAR2018    -0.707190   0.576102  -1.228 0.219619
## YEAR2019     0.171866   0.442122   0.389 0.697476
## YEAR2021     1.918630   0.829289   2.314 0.020691 *
## MONTH7      -0.560131   0.212739  -2.633 0.008465 **
## MONTH8      -0.823759   0.233903  -3.522 0.000429 ***
## MONTH9      -1.186555   0.211064  -5.622 1.89e-08 ***
## TIDE2       -0.602163   0.210153  -2.865 0.004166 **
## TIDE3        0.878537   0.197813   4.441 8.94e-06 ***
## TIDE4        0.329941   0.211219   1.562 0.118270
## TIDE5       -0.261118   0.343344  -0.761 0.446947
## TIDE6        0.238528   0.374150   0.638 0.523786
## TEMPBOT     -0.065676   0.116466  -0.564 0.572817
## DOBOT       -0.047038   0.090078  -0.522 0.601533
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.00093    0.37119  -2.697  0.00701 **
## YEAR2003    -0.39596    0.41490  -0.954  0.33990
## YEAR2004    -0.06410    0.44740  -0.143  0.88608
```

```
## YEAR2005    -0.64933    0.40731  -1.594  0.11089
## YEAR2006    -0.44571    0.44275  -1.007  0.31409
## YEAR2007    -0.90134    0.40662  -2.217  0.02664 *
## YEAR2008    -0.54418    0.40350  -1.349  0.17745
## YEAR2009    -0.57484    0.40996  -1.402  0.16085
## YEAR2010    -0.48173    0.40896  -1.178  0.23882
## YEAR2011     0.50819    0.46092   1.103  0.27022
## YEAR2012     0.44295    0.41634   1.064  0.28736
## YEAR2013    -0.73751    0.40358  -1.827  0.06764 .
## YEAR2014    -1.19909    0.40701  -2.946  0.00322 **
## YEAR2015     0.65015    0.42846   1.517  0.12917
## YEAR2016     0.88776    0.44550   1.993  0.04629 *
## YEAR2017    -0.48907    0.41012  -1.192  0.23307
## YEAR2018     0.61370    0.45400   1.352  0.17645
## YEAR2019     0.46492    0.42476   1.095  0.27371
## YEAR2021     1.27013    0.81465   1.559  0.11897
## MONTH7       0.99591    0.26953   3.695  0.00022 ***
## MONTH8       1.72770    0.27336   6.320 2.61e-10 ***
## MONTH9       1.33974    0.24490   5.471 4.48e-08 ***
## TIDE2        0.47191    0.19460   2.425  0.01531 *
## TIDE3       -0.04247    0.23061  -0.184  0.85388
## TIDE4        0.45239    0.22630   1.999  0.04560 *
## TIDE5       -0.44707    0.35932  -1.244  0.21342
## TIDE6        0.70554    0.38234   1.845  0.06500 .
## TEMPBOT     -0.15473    0.10375  -1.491  0.13585
## DOBOT        0.04008    0.08368   0.479  0.63193
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Acceptable standard errors

tmbZI.resids(ZANB, y.dat=dat$CPUE)

## df.resid
## 1.339399
```

Preliminary model comparisons

```
AICtab(tmb1.NB, ZINB4, ZANB)

##         dAIC df
## ZANB     0.0 59
## tmb1.NB  5.8 30
## ZINB4    8.6 31
```

Step 3: Factor selection (by model)

Negative Binomial

```
NB1 <- glmmTMB(CPUE ~ YEAR+MONTH+TIDE+TEMPBOT+DOBOT, data = dat, family =
nbinom2)
summary(NB1)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + TEMPBOT + DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2562.8   2708.9  -1251.4   2502.8      933
##
##
## Dispersion parameter for nbinom2 family (): 0.891
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.771826   0.254532   3.032 0.002427 **
## YEAR2003     0.375196   0.301853   1.243 0.213877
## YEAR2004     0.009939   0.346475   0.029 0.977114
## YEAR2005     0.609835   0.282852   2.156 0.031082 *
## YEAR2006     0.172731   0.323609   0.534 0.593504
## YEAR2007     0.751340   0.282680   2.658 0.007862 **
## YEAR2008     0.297093   0.292535   1.016 0.309829
## YEAR2009     0.285803   0.292973   0.976 0.329299
## YEAR2010     0.621865   0.284927   2.183 0.029069 *
## YEAR2011     0.049164   0.321410   0.153 0.878427
## YEAR2012    -0.220089   0.312244  -0.705 0.480896
## YEAR2013     0.381073   0.286145   1.332 0.182943
## YEAR2014     0.744675   0.280288   2.657 0.007888 **
## YEAR2015    -0.533291   0.321409  -1.659 0.097070 .
## YEAR2016    -0.795012   0.350050  -2.271 0.023139 *
## YEAR2017     0.198670   0.295138   0.673 0.500856
## YEAR2018    -0.715246   0.360876  -1.982 0.047482 *
## YEAR2019    -0.168026   0.311133  -0.540 0.589166
## YEAR2021     0.032188   0.477817   0.067 0.946291
## MONTH7      -0.792705   0.169023  -4.690 2.73e-06 ***
## MONTH8      -1.350921   0.176870  -7.638 2.21e-14 ***
## MONTH9      -1.241808   0.158619  -7.829 4.92e-15 ***
## TIDE2       -0.528911   0.144888  -3.650 0.000262 ***
## TIDE3        0.501300   0.153301   3.270 0.001075 **
## TIDE4       -0.105492   0.158015  -0.668 0.504381
## TIDE5        0.106635   0.254830   0.418 0.675615
## TIDE6       -0.276543   0.270735  -1.021 0.307040
## TEMPBOT      0.102389   0.078160   1.310 0.190202
## DOBOT       -0.025874   0.062258  -0.416 0.677706
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(NB1,test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + TIDE + TEMPBOT + DOBOT
```

```
##         Df    AIC    LRT  Pr(>Chi)
## <none>     2562.8
## YEAR    18 2592.2 65.409 2.658e-07 ***
## MONTH    3 2641.3 84.551 < 2.2e-16 ***
## TIDE     5 2598.5 45.713 1.039e-08 ***
## TEMPBOT  1 2562.5  1.719    0.1898
## DOBOT    1 2560.9  0.173    0.6776
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# drop DOBOT
NB2 <- glmmTMB(CPUE ~ YEAR+MONTH+TIDE+TEMPBOT, data = dat, family = nbinom2)
summary(NB2)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + TEMPBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2560.9   2702.2  -1251.5   2502.9      934
##
##
## Dispersion parameter for nbinom2 family (): 0.891
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.774639   0.254255   3.047 0.002314 **
## YEAR2003     0.355058   0.297889   1.192 0.233295
## YEAR2004     0.015744   0.346081   0.045 0.963715
## YEAR2005     0.601877   0.282224   2.133 0.032956 *
## YEAR2006     0.161177   0.322281   0.500 0.616996
## YEAR2007     0.744024   0.282081   2.638 0.008349 **
## YEAR2008     0.296203   0.292338   1.013 0.310954
## YEAR2009     0.277323   0.292089   0.949 0.342393
## YEAR2010     0.612123   0.283773   2.157 0.030999 *
## YEAR2011     0.029269   0.317590   0.092 0.926571
## YEAR2012    -0.227651   0.311686  -0.730 0.465153
## YEAR2013     0.377088   0.285777   1.320 0.186995
## YEAR2014     0.735365   0.279183   2.634 0.008439 **
## YEAR2015    -0.548648   0.319071  -1.720 0.085520 .
## YEAR2016    -0.829905   0.339874  -2.442 0.014614 *
## YEAR2017     0.168803   0.286180   0.590 0.555291
## YEAR2018    -0.743949   0.354345  -2.100 0.035773 *
## YEAR2019    -0.198024   0.302786  -0.654 0.513108
## YEAR2021     0.007821   0.474041   0.016 0.986836
## MONTH7      -0.786674   0.168392  -4.672 2.99e-06 ***
## MONTH8      -1.337033   0.173668  -7.699 1.37e-14 ***
## MONTH9      -1.233259   0.157271  -7.842 4.45e-15 ***
## TIDE2       -0.525235   0.144591  -3.633 0.000281 ***
## TIDE3        0.505158   0.153000   3.302 0.000961 ***
```

```
## TIDE4       -0.100159   0.157490  -0.636 0.524797
## TIDE5        0.103570   0.254551   0.407 0.684100
## TIDE6       -0.272008   0.270406  -1.006 0.314453
## TEMPBOT      0.106634   0.077485   1.376 0.168760
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(NB2,test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + TIDE + TEMPBOT
##         Df    AIC    LRT  Pr(>Chi)
## <none>      2560.9
## YEAR    18 2595.0 70.065 4.406e-08 ***
## MONTH    3 2641.3 86.401 < 2.2e-16 ***
## TIDE     5 2596.6 45.641 1.075e-08 ***
## TEMPBOT  1 2560.8  1.898    0.1683
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# drop TEMPBOT:
NB3 <- glmmTMB(CPUE ~ YEAR+MONTH+TIDE, data = dat, family = nbinom2)
summary(NB3)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2560.8   2697.2  -1252.4   2504.8      935
##
##
## Dispersion parameter for nbinom2 family (): 0.89
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.71958    0.25143   2.862 0.004210 **
## YEAR2003     0.33206    0.29796   1.114 0.265087
## YEAR2004    -0.03178    0.34483  -0.092 0.926571
## YEAR2005     0.65586    0.27975   2.344 0.019056 *
## YEAR2006     0.09873    0.31908   0.309 0.757007
## YEAR2007     0.76146    0.28193   2.701 0.006915 **
## YEAR2008     0.31186    0.29234   1.067 0.286072
## YEAR2009     0.20658    0.28795   0.717 0.473117
## YEAR2010     0.65221    0.28280   2.306 0.021099 *
## YEAR2011     0.02259    0.31776   0.071 0.943325
## YEAR2012    -0.22581    0.31216  -0.723 0.469442
## YEAR2013     0.37844    0.28551   1.325 0.185009
## YEAR2014     0.71737    0.27915   2.570 0.010174 *
```

```
## YEAR2015    -0.49919    0.31663  -1.577 0.114892    
## YEAR2016    -0.77696    0.33822  -2.297 0.021609 *  
## YEAR2017     0.18392    0.28587   0.643 0.519989    
## YEAR2018    -0.69062    0.35289  -1.957 0.050344 .  
## YEAR2019    -0.18240    0.30301  -0.602 0.547198    
## YEAR2021    -0.02068    0.47307  -0.044 0.965136    
## MONTH7      -0.68240    0.15008  -4.547 5.45e-06 ***
## MONTH8      -1.23498    0.15664  -7.884 3.16e-15 ***
## MONTH9      -1.26553    0.15546  -8.140 3.94e-16 ***
## TIDE2       -0.52111    0.14435  -3.610 0.000306 ***
## TIDE3        0.50473    0.15290   3.301 0.000963 ***
## TIDE4       -0.09418    0.15734  -0.599 0.549451    
## TIDE5        0.10212    0.25480   0.401 0.688572    
## TIDE6       -0.29098    0.27027  -1.077 0.281651    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(NB3,test="Chi")
```

```
## Single term deletions
## 
## Model:
## CPUE ~ YEAR + MONTH + TIDE
##        Df    AIC    LRT  Pr(>Chi)    
## <none>    2560.8                     
## YEAR   18 2593.6 68.744 7.362e-08 ***
## MONTH   3 2639.8 85.012 < 2.2e-16 ***
## TIDE    5 2596.2 45.330 1.243e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Let's compare to a model with year only.
NB0 <- glmmTMB(CPUE ~ YEAR, data = dat, family = nbinom2)

# Let's compare the models using AIC to see if adding more factors
# improves how much deviance is explained.

AICtab(NB0, NB1, NB2, NB3)
```

```
##     dAIC  df
## NB3   0.0 28
## NB2   0.1 29
## NB1   1.9 30
## NB0 117.6 20
```

```
sapply(list(NB0, NB1, NB2, NB3),formula)
```

```
## [[1]]
## CPUE ~ YEAR
## 
## [[2]]
```

```
## CPUE ~ YEAR + MONTH + TIDE + TEMPBOT + DOBOT
##
## [[3]]
## CPUE ~ YEAR + MONTH + TIDE + TEMPBOT
##
## [[4]]
## CPUE ~ YEAR + MONTH + TIDE
```

```
lrtest(NB0, NB3, NB2, NB1)
```

```
## Likelihood ratio test
##
## Model 1: CPUE ~ YEAR
## Model 2: CPUE ~ YEAR + MONTH + TIDE
## Model 3: CPUE ~ YEAR + MONTH + TIDE + TEMPBOT
## Model 4: CPUE ~ YEAR + MONTH + TIDE + TEMPBOT + DOBOT
##   #Df  LogLik Df    Chisq Pr(>Chisq)
## 1  20 -1319.2
## 2  28 -1252.4  8 133.6465     <2e-16 ***
## 3  29 -1251.5  1   1.8981     0.1683
## 4  30 -1251.4  1   0.1729     0.6776
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# So NB3 wins
sapply(list(NB3, NB2, NB1, NB0),disp3)
```

```
## df.resid df.resid df.resid df.resid
## 1.191702 1.164203 1.163820 1.154704
```

ZINB

```
ZINB1 = glmmTMB(CPUE ~ YEAR+MONTH+TIDE+DOBOT, ziformula = ~ DOBOT, data =
dat, family=nbinom2)
summary(ZINB1)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + DOBOT
## Zero inflation:        ~DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2565.6   2716.5  -1251.8   2503.6      932
##
##
## Dispersion parameter for nbinom2 family (): 0.952
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.72949    0.25233   2.891 0.003839 **
## YEAR2003     0.38745    0.30329   1.277 0.201436
## YEAR2004    -0.02564    0.34738  -0.074 0.941158
```

```
## YEAR2005     0.68557    0.28241   2.428 0.015201 *
## YEAR2006     0.13254    0.32045   0.414 0.679163
## YEAR2007     0.76873    0.28174   2.729 0.006362 **
## YEAR2008     0.30277    0.29235   1.036 0.300371
## YEAR2009     0.22254    0.28855   0.771 0.440572
## YEAR2010     0.67901    0.28477   2.384 0.017105 *
## YEAR2011     0.07784    0.32242   0.241 0.809220
## YEAR2012    -0.19619    0.31417  -0.624 0.532308
## YEAR2013     0.38050    0.28547   1.333 0.182575
## YEAR2014     0.72501    0.28022   2.587 0.009674 **
## YEAR2015    -0.47130    0.31767  -1.484 0.137911
## YEAR2016    -0.68601    0.34875  -1.967 0.049176 *
## YEAR2017     0.24282    0.29320   0.828 0.407571
## YEAR2018    -0.63167    0.35782  -1.765 0.077511 .
## YEAR2019    -0.11951    0.31001  -0.385 0.699869
## YEAR2021     0.05145    0.47825   0.108 0.914326
## MONTH7      -0.68860    0.15157  -4.543 5.54e-06 ***
## MONTH8      -1.25631    0.16221  -7.745 9.57e-15 ***
## MONTH9      -1.27216    0.15523  -8.195 2.50e-16 ***
## TIDE2       -0.53411    0.14414  -3.705 0.000211 ***
## TIDE3        0.50906    0.15317   3.323 0.000889 ***
## TIDE4       -0.09584    0.15759  -0.608 0.543063
## TIDE5        0.10213    0.25335   0.403 0.686867
## TIDE6       -0.29155    0.26966  -1.081 0.279602
## DOBOT       -0.07194    0.07300  -0.986 0.324372
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -4.040      2.235  -1.808   0.0706 .
## DOBOT         -1.036      0.710  -1.460   0.1443
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZINB1,test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + TIDE + DOBOT
##        Df    AIC    LRT  Pr(>Chi)
## <none>    2565.6
## YEAR   18 2591.9 62.344 8.507e-07 ***
## MONTH   3 2638.4 78.846 < 2.2e-16 ***
## TIDE    5 2595.4 39.830 1.616e-07 ***
## DOBOT   1 2559.1 -4.466         1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# drop DOBOT from conditional:
ZINB2 <- glmmTMB(CPUE ~ YEAR+MONTH+TIDE, ziformula = ~ DOBOT, data = dat,
family=nbinom2)
summary(ZINB2)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE
## Zero inflation:        ~DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2559.1   2705.2  -1249.6   2499.1      933
##
##
## Dispersion parameter for nbinom2 family (): 0.919
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.73180    0.25126   2.913 0.003585 **
## YEAR2003     0.36090    0.29958   1.205 0.228333
## YEAR2004    -0.04340    0.34299  -0.127 0.899315
## YEAR2005     0.64885    0.27792   2.335 0.019560 *
## YEAR2006     0.08623    0.31704   0.272 0.785647
## YEAR2007     0.75502    0.28020   2.695 0.007048 **
## YEAR2008     0.30326    0.29057   1.044 0.296637
## YEAR2009     0.20704    0.28645   0.723 0.469822
## YEAR2010     0.64870    0.28088   2.309 0.020916 *
## YEAR2011     0.08055    0.32014   0.252 0.801344
## YEAR2012    -0.23958    0.31064  -0.771 0.440555
## YEAR2013     0.37131    0.28369   1.309 0.190575
## YEAR2014     0.71638    0.27757   2.581 0.009854 **
## YEAR2015    -0.50216    0.31537  -1.592 0.111326
## YEAR2016    -0.76119    0.33933  -2.243 0.024881 *
## YEAR2017     0.18219    0.28445   0.640 0.521866
## YEAR2018    -0.68858    0.35203  -1.956 0.050466 .
## YEAR2019    -0.09576    0.30561  -0.313 0.754024
## YEAR2021     0.07712    0.48551   0.159 0.873786
## MONTH7      -0.67336    0.15096  -4.460 8.18e-06 ***
## MONTH8      -1.24885    0.15711  -7.949 1.88e-15 ***
## MONTH9      -1.27361    0.15537  -8.197 2.46e-16 ***
## TIDE2       -0.51621    0.14424  -3.579 0.000345 ***
## TIDE3        0.49430    0.15225   3.247 0.001167 **
## TIDE4       -0.10143    0.15698  -0.646 0.518184
## TIDE5        0.12724    0.25572   0.498 0.618782
## TIDE6       -0.29359    0.26881  -1.092 0.274751
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)   -7.715      4.574  -1.687   0.0916 .
## DOBOT           2.827      1.824   1.550   0.1212
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZINB2,test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + TIDE
##        Df    AIC    LRT  Pr(>Chi)
## <none>    2559.1
## YEAR   18 2589.4 66.344 1.859e-07 ***
## MONTH   3 2642.9 89.757 < 2.2e-16 ***
## TIDE    5 2593.7 44.579 1.767e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# drop DOBOT from ZI:
ZINB3 <- glmmTMB(CPUE ~ YEAR+MONTH+TIDE, data = dat, family=nbinom2)
summary(ZINB3)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2560.8   2697.2  -1252.4   2504.8      935
##
##
## Dispersion parameter for nbinom2 family (): 0.89
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.71958    0.25143   2.862 0.004210 **
## YEAR2003     0.33206    0.29796   1.114 0.265087
## YEAR2004    -0.03178    0.34483  -0.092 0.926571
## YEAR2005     0.65586    0.27975   2.344 0.019056 *
## YEAR2006     0.09873    0.31908   0.309 0.757007
## YEAR2007     0.76146    0.28193   2.701 0.006915 **
## YEAR2008     0.31186    0.29234   1.067 0.286072
## YEAR2009     0.20658    0.28795   0.717 0.473117
## YEAR2010     0.65221    0.28280   2.306 0.021099 *
## YEAR2011     0.02259    0.31776   0.071 0.943325
## YEAR2012    -0.22581    0.31216  -0.723 0.469442
## YEAR2013     0.37844    0.28551   1.325 0.185009
## YEAR2014     0.71737    0.27915   2.570 0.010174 *
## YEAR2015    -0.49919    0.31663  -1.577 0.114892
## YEAR2016    -0.77696    0.33822  -2.297 0.021609 *
## YEAR2017     0.18392    0.28587   0.643 0.519989
```

```
## YEAR2018    -0.69062    0.35289  -1.957 0.050344 .
## YEAR2019    -0.18240    0.30301  -0.602 0.547198
## YEAR2021    -0.02068    0.47307  -0.044 0.965136
## MONTH7      -0.68240    0.15008  -4.547 5.45e-06 ***
## MONTH8      -1.23498    0.15664  -7.884 3.16e-15 ***
## MONTH9      -1.26553    0.15546  -8.140 3.94e-16 ***
## TIDE2       -0.52111    0.14435  -3.610 0.000306 ***
## TIDE3        0.50473    0.15290   3.301 0.000963 ***
## TIDE4       -0.09418    0.15734  -0.599 0.549451
## TIDE5        0.10212    0.25480   0.401 0.688572
## TIDE6       -0.29098    0.27027  -1.077 0.281651
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZINB3,test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + TIDE
##        Df    AIC    LRT  Pr(>Chi)
## <none>    2560.8
## YEAR   18 2593.6 68.744 7.362e-08 ***
## MONTH   3 2639.8 85.012 < 2.2e-16 ***
## TIDE    5 2596.2 45.330 1.243e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# year only model:
ZINB0 <- glmmTMB(CPUE ~ YEAR, ziformula = ~ YEAR, data = dat, family=nbinom2,

control=glmmTMBControl(optCtrl=list(iter.max=5e6,eval.max=5e6)))
summary(ZINB0) # bombs: large SEs in ZI YEAR
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR
## Zero inflation:        ~YEAR
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2687.6   2877.5  -1304.8   2609.6      924
##
##
## Dispersion parameter for nbinom2 family (): 0.819
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.18775    0.29515   0.636   0.5247
## YEAR2003     0.24071    0.40127   0.600   0.5486
## YEAR2004    -0.69857    0.39915  -1.750   0.0801 .
## YEAR2005     0.62410    0.36564   1.707   0.0878 .
```

```
## YEAR2006    -0.13896    0.37612  -0.370   0.7118
## YEAR2007     0.27554    0.34665   0.795   0.4267
## YEAR2008    -0.11754    0.35535  -0.331   0.7408
## YEAR2009     0.05738    0.35266   0.163   0.8708
## YEAR2010     0.57638    0.37893   1.521   0.1282
## YEAR2011     0.66310    0.46402   1.429   0.1530
## YEAR2012    -0.27016    0.45150  -0.598   0.5496
## YEAR2013     0.02443    0.35218   0.069   0.9447
## YEAR2014     0.28671    0.34647   0.828   0.4079
## YEAR2015    -0.27016    0.45925  -0.588   0.5564
## YEAR2016    -0.83754    0.53958  -1.552   0.1206
## YEAR2017     0.11939    0.38144   0.313   0.7543
## YEAR2018    -0.89765    0.56149  -1.599   0.1099
## YEAR2019    -0.07594    0.43423  -0.175   0.8612
## YEAR2021     1.00656    0.90451   1.113   0.2658
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.378e+00  1.054e+00  -1.308   0.1910
## YEAR2003     2.018e-01  1.296e+00   0.156   0.8763
## YEAR2004    -1.553e+01  5.272e+03  -0.003   0.9977
## YEAR2005    -4.733e-01  1.421e+00  -0.333   0.7392
## YEAR2006    -2.070e+01  2.893e+04  -0.001   0.9994
## YEAR2007    -1.824e+01  1.087e+04  -0.002   0.9987
## YEAR2008    -1.929e+01  1.668e+04  -0.001   0.9991
## YEAR2009    -1.126e+01  4.122e+03  -0.003   0.9978
## YEAR2010    -8.999e-02  1.304e+00  -0.069   0.9450
## YEAR2011     1.826e+00  1.127e+00   1.620   0.1052
## YEAR2012     5.399e-01  1.339e+00   0.403   0.6867
## YEAR2013    -1.994e+01  2.128e+04  -0.001   0.9993
## YEAR2014    -2.148e+01  3.394e+04  -0.001   0.9995
## YEAR2015     6.641e-01  1.312e+00   0.506   0.6127
## YEAR2016     2.668e-01  1.815e+00   0.147   0.8831
## YEAR2017    -2.168e+00  6.095e+00  -0.356   0.7221
## YEAR2018    -2.084e-01  2.535e+00  -0.082   0.9345
## YEAR2019     6.669e-01  1.253e+00   0.532   0.5945
## YEAR2021     3.165e+00  1.323e+00   2.391   0.0168 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
ZINB0 <- glmmTMB(CPUE ~ YEAR, data = dat, family=nbinom2,

control=glmmTMBControl(optCtrl=list(iter.max=5e6,eval.max=5e6)))
summary(ZINB0)
```

```
##  Family: nbinom2  ( log )
## Formula:          CPUE ~ YEAR
## Data: dat
```

```
##
##      AIC      BIC   logLik deviance df.resid
##   2678.5   2775.9  -1319.2   2638.5      943
##
##
## Dispersion parameter for nbinom2 family (): 0.589
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.03703    0.22307  -0.166   0.8682
## YEAR2003     0.19665    0.31437   0.625   0.5316
## YEAR2004    -0.47379    0.36593  -1.295   0.1954
## YEAR2005     0.70304    0.29879   2.353   0.0186 *
## YEAR2006     0.08581    0.34066   0.252   0.8011
## YEAR2007     0.50031    0.30222   1.655   0.0978 .
## YEAR2008     0.10721    0.31241   0.343   0.7315
## YEAR2009     0.28215    0.30938   0.912   0.3618
## YEAR2010     0.59383    0.30710   1.934   0.0532 .
## YEAR2011    -0.05394    0.33242  -0.162   0.8711
## YEAR2012    -0.40480    0.32845  -1.232   0.2178
## YEAR2013     0.24920    0.30880   0.807   0.4197
## YEAR2014     0.51145    0.30201   1.694   0.0904 .
## YEAR2015    -0.44395    0.33176  -1.338   0.1808
## YEAR2016    -0.89727    0.35434  -2.532   0.0113 *
## YEAR2017     0.31576    0.30596   1.032   0.3021
## YEAR2018    -0.85907    0.36665  -2.343   0.0191 *
## YEAR2019    -0.25066    0.32235  -0.778   0.4368
## YEAR2021    -0.71021    0.50030  -1.420   0.1557
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
AICtab(ZINB1,ZINB2,ZINB3,ZINB0)
```

```
##       dAIC  df
## ZINB2   0.0 30
## ZINB3   1.7 28
## ZINB1   6.5 31
## ZINB0 119.4 20
```

```
lrtest(ZINB0,ZINB3,ZINB2,ZINB1)
```

```
## Likelihood ratio test
##
## Model 1: CPUE ~ YEAR
## Model 2: CPUE ~ YEAR + MONTH + TIDE
## Model 3: CPUE ~ YEAR + MONTH + TIDE
## Model 4: CPUE ~ YEAR + MONTH + TIDE + DOBOT
##   #Df  LogLik Df    Chisq Pr(>Chisq)
## 1  20 -1319.2
## 2  28 -1252.4  8 133.6465    < 2e-16 ***
## 3  30 -1249.5  2   5.7322    0.05692 .
```

```
## 4  31 -1251.8  1   4.4660    0.03458 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# equal support for all models; lrtest suggests ZINB3 not improved with
additional terms, so ZINB3 wins
sapply(list(ZINB3,ZINB2,ZINB1,ZINB0),tmbZI.resids,y.dat=dat$CPUE)
```

```
## df.resid df.resid df.resid df.resid
## 1.302556 1.171509 1.182707 1.379993
```

ZANB (aka hurdle model)

```
ZANB1 <- glmmTMB(CPUE ~ YEAR+MONTH+TIDE+TEMPBOT+DOBOT, ziformula = ~
YEAR+MONTH+TIDE+TEMPBOT+DOBOT, data = dat,family = truncated_nbinom2(link =
"log"))
summary(ZANB1)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + TEMPBOT + DOBOT
## Zero inflation:        ~YEAR + MONTH + TIDE + TEMPBOT + DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2557.0   2844.3  -1219.5   2439.0      904
##
##
## Dispersion parameter for truncated_nbinom2 family (): 1.19
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.576748   0.345380   1.670 0.094941 .
## YEAR2003     0.249967   0.419017   0.597 0.550804
## YEAR2004    -0.112863   0.519567  -0.217 0.828032
## YEAR2005     0.564234   0.375176   1.504 0.132603
## YEAR2006    -0.216111   0.447050  -0.483 0.628801
## YEAR2007     0.515834   0.372784   1.384 0.166440
## YEAR2008    -0.009458   0.399611  -0.024 0.981118
## YEAR2009    -0.151569   0.392735  -0.386 0.699548
## YEAR2010     0.712405   0.380808   1.871 0.061376 .
## YEAR2011     0.724831   0.469808   1.543 0.122873
## YEAR2012     0.252705   0.453931   0.557 0.577729
## YEAR2013     0.100941   0.380435   0.265 0.790756
## YEAR2014     0.291814   0.369362   0.790 0.429498
## YEAR2015    -0.169967   0.466617  -0.364 0.715669
## YEAR2016    -0.443683   0.551671  -0.804 0.421251
## YEAR2017    -0.029094   0.400823  -0.073 0.942136
## YEAR2018    -0.707190   0.576102  -1.228 0.219619
## YEAR2019     0.171866   0.442122   0.389 0.697476
## YEAR2021     1.918630   0.829289   2.314 0.020691 *
## MONTH7      -0.560131   0.212739  -2.633 0.008465 **
## MONTH8      -0.823759   0.233903  -3.522 0.000429 ***
```

```
## MONTH9      -1.186555   0.211064  -5.622 1.89e-08 ***
## TIDE2       -0.602163   0.210153  -2.865 0.004166 **
## TIDE3        0.878537   0.197813   4.441 8.94e-06 ***
## TIDE4        0.329941   0.211219   1.562 0.118270
## TIDE5       -0.261118   0.343344  -0.761 0.446947
## TIDE6        0.238528   0.374150   0.638 0.523786
## TEMPBOT     -0.065676   0.116466  -0.564 0.572817
## DOBOT       -0.047038   0.090078  -0.522 0.601533
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.00093    0.37119  -2.697  0.00701 **
## YEAR2003    -0.39596    0.41490  -0.954  0.33990
## YEAR2004    -0.06410    0.44740  -0.143  0.88608
## YEAR2005    -0.64933    0.40731  -1.594  0.11089
## YEAR2006    -0.44571    0.44275  -1.007  0.31409
## YEAR2007    -0.90134    0.40662  -2.217  0.02664 *
## YEAR2008    -0.54418    0.40350  -1.349  0.17745
## YEAR2009    -0.57484    0.40996  -1.402  0.16085
## YEAR2010    -0.48173    0.40896  -1.178  0.23882
## YEAR2011     0.50819    0.46092   1.103  0.27022
## YEAR2012     0.44295    0.41634   1.064  0.28736
## YEAR2013    -0.73751    0.40358  -1.827  0.06764 .
## YEAR2014    -1.19909    0.40701  -2.946  0.00322 **
## YEAR2015     0.65015    0.42846   1.517  0.12917
## YEAR2016     0.88776    0.44550   1.993  0.04629 *
## YEAR2017    -0.48907    0.41012  -1.192  0.23307
## YEAR2018     0.61370    0.45400   1.352  0.17645
## YEAR2019     0.46492    0.42476   1.095  0.27371
## YEAR2021     1.27013    0.81465   1.559  0.11897
## MONTH7       0.99591    0.26953   3.695  0.00022 ***
## MONTH8       1.72770    0.27336   6.320 2.61e-10 ***
## MONTH9       1.33974    0.24490   5.471 4.48e-08 ***
## TIDE2        0.47191    0.19460   2.425  0.01531 *
## TIDE3       -0.04247    0.23061  -0.184  0.85388
## TIDE4        0.45239    0.22630   1.999  0.04560 *
## TIDE5       -0.44707    0.35932  -1.244  0.21342
## TIDE6        0.70554    0.38234   1.845  0.06500 .
## TEMPBOT     -0.15473    0.10375  -1.491  0.13585
## DOBOT        0.04008    0.08368   0.479  0.63193
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZANB1,test="Chi")
```

```
## Single term deletions
##
## Model:
```

```
## CPUE ~ YEAR + MONTH + TIDE + TEMPBOT + DOBOT
##         Df    AIC    LRT  Pr(>Chi)
## <none>     2557.0
## YEAR    18 2550.3 29.364   0.04411 *
## MONTH    3 2586.2 35.242 1.083e-07 ***
## TIDE     5 2599.5 52.538 4.184e-10 ***
## TEMPBOT  1 2555.3  0.318   0.57285
## DOBOT    1 2555.3  0.273   0.60103
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# remove TEMPBOT in conditional:
ZANB2 <- glmmTMB(CPUE ~ YEAR+MONTH+TIDE+DOBOT, ziformula = ~
YEAR+MONTH+TIDE+TEMPBOT+DOBOT, data = dat,family = truncated_nbinom2(link =
"log"))
summary(ZANB2)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE + DOBOT
## Zero inflation:        ~YEAR + MONTH + TIDE + TEMPBOT + DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2555.3   2837.8  -1219.7   2439.3      905
##
##
## Dispersion parameter for truncated_nbinom2 family (): 1.18
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.61305    0.33866   1.810  0.07026 .
## YEAR2003     0.24688    0.41891   0.589  0.55563
## YEAR2004    -0.08409    0.51651  -0.163  0.87067
## YEAR2005     0.52219    0.36775   1.420  0.15561
## YEAR2006    -0.17858    0.44230  -0.404  0.68640
## YEAR2007     0.49981    0.37163   1.345  0.17865
## YEAR2008    -0.02510    0.39844  -0.063  0.94977
## YEAR2009    -0.11788    0.38779  -0.304  0.76113
## YEAR2010     0.67791    0.37549   1.805  0.07101 .
## YEAR2011     0.69355    0.46483   1.492  0.13569
## YEAR2012     0.24366    0.45363   0.537  0.59118
## YEAR2013     0.08611    0.37954   0.227  0.82052
## YEAR2014     0.29130    0.36896   0.790  0.42980
## YEAR2015    -0.21978    0.45881  -0.479  0.63192
## YEAR2016    -0.48690    0.54645  -0.891  0.37291
## YEAR2017    -0.06005    0.39737  -0.151  0.87988
## YEAR2018    -0.74788    0.57144  -1.309  0.19061
## YEAR2019     0.14037    0.43883   0.320  0.74907
## YEAR2021     1.91103    0.83017   2.302  0.02134 *
## MONTH7      -0.61483    0.18991  -3.238  0.00121 **
```

```
## MONTH8      -0.87923    0.21301  -4.128 3.67e-05 ***
## MONTH9      -1.16968    0.20905  -5.595 2.20e-08 ***
## TIDE2       -0.60370    0.21049  -2.868  0.00413 **
## TIDE3        0.88375    0.19796   4.464 8.03e-06 ***
## TIDE4        0.32579    0.21131   1.542  0.12312
## TIDE5       -0.25254    0.34280  -0.737  0.46131
## TIDE6        0.24936    0.37367   0.667  0.50456
## DOBOT       -0.04120    0.08947  -0.461  0.64513
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.00094    0.37119  -2.697  0.00701 **
## YEAR2003    -0.39596    0.41490  -0.954  0.33990
## YEAR2004    -0.06410    0.44740  -0.143  0.88607
## YEAR2005    -0.64933    0.40731  -1.594  0.11089
## YEAR2006    -0.44571    0.44275  -1.007  0.31409
## YEAR2007    -0.90134    0.40662  -2.217  0.02664 *
## YEAR2008    -0.54418    0.40351  -1.349  0.17745
## YEAR2009    -0.57485    0.40996  -1.402  0.16085
## YEAR2010    -0.48172    0.40896  -1.178  0.23883
## YEAR2011     0.50819    0.46092   1.103  0.27022
## YEAR2012     0.44296    0.41634   1.064  0.28736
## YEAR2013    -0.73752    0.40358  -1.827  0.06763 .
## YEAR2014    -1.19911    0.40701  -2.946  0.00322 **
## YEAR2015     0.65015    0.42846   1.517  0.12917
## YEAR2016     0.88776    0.44550   1.993  0.04629 *
## YEAR2017    -0.48908    0.41012  -1.193  0.23306
## YEAR2018     0.61370    0.45400   1.352  0.17645
## YEAR2019     0.46493    0.42476   1.095  0.27370
## YEAR2021     1.27012    0.81465   1.559  0.11897
## MONTH7       0.99591    0.26953   3.695  0.00022 ***
## MONTH8       1.72771    0.27336   6.320 2.61e-10 ***
## MONTH9       1.33975    0.24490   5.471 4.48e-08 ***
## TIDE2        0.47192    0.19460   2.425  0.01531 *
## TIDE3       -0.04246    0.23061  -0.184  0.85391
## TIDE4        0.45240    0.22630   1.999  0.04560 *
## TIDE5       -0.44707    0.35932  -1.244  0.21342
## TIDE6        0.70555    0.38234   1.845  0.06499 .
## TEMPBOT     -0.15473    0.10375  -1.491  0.13586
## DOBOT        0.04009    0.08368   0.479  0.63189
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZANB2,test="Chi")
```

```
## Single term deletions
##
## Model:
```

```
## CPUE ~ YEAR + MONTH + TIDE + DOBOT
##        Df    AIC    LRT  Pr(>Chi)
## <none>    2555.3
## YEAR   18 2548.4 29.133   0.04678 *
## MONTH   3 2584.4 35.122 1.148e-07 ***
## TIDE    5 2598.3 53.002 3.360e-10 ***
## DOBOT   1 2553.5  0.213   0.64475
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# drop DOBOT:
ZANB3 <- glmmTMB(CPUE ~ YEAR+MONTH+TIDE, ziformula = ~
YEAR+MONTH+TIDE+TEMPBOT+DOBOT, data = dat,family = truncated_nbinom2(link =
"log"))
summary(ZANB3)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE
## Zero inflation:        ~YEAR + MONTH + TIDE + TEMPBOT + DOBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2553.5   2831.1  -1219.8   2439.5      906
##
##
## Dispersion parameter for truncated_nbinom2 family (): 1.18
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.61983    0.33803   1.834  0.06670 .
## YEAR2003     0.20170    0.40717   0.495  0.62035
## YEAR2004    -0.08555    0.51631  -0.166  0.86840
## YEAR2005     0.49967    0.36460   1.370  0.17054
## YEAR2006    -0.21086    0.43667  -0.483  0.62919
## YEAR2007     0.48074    0.36915   1.302  0.19282
## YEAR2008    -0.02903    0.39800  -0.073  0.94185
## YEAR2009    -0.13967    0.38476  -0.363  0.71661
## YEAR2010     0.65635    0.37200   1.764  0.07767 .
## YEAR2011     0.66147    0.45895   1.441  0.14951
## YEAR2012     0.21747    0.45014   0.483  0.62902
## YEAR2013     0.07660    0.37874   0.202  0.83971
## YEAR2014     0.27294    0.36647   0.745  0.45640
## YEAR2015    -0.24090    0.45642  -0.528  0.59764
## YEAR2016    -0.54812    0.53000  -1.034  0.30105
## YEAR2017    -0.11157    0.38162  -0.292  0.77002
## YEAR2018    -0.79930    0.56125  -1.424  0.15441
## YEAR2019     0.08945    0.42507   0.210  0.83332
## YEAR2021     1.87342    0.82560   2.269  0.02326 *
## MONTH7      -0.59547    0.18503  -3.218  0.00129 **
## MONTH8      -0.85283    0.20488  -4.162 3.15e-05 ***
```

```
## MONTH9      -1.16033    0.20796  -5.580 2.41e-08 ***
## TIDE2       -0.59219    0.20876  -2.837  0.00456 **
## TIDE3        0.88911    0.19761   4.499 6.82e-06 ***
## TIDE4        0.33420    0.21057   1.587  0.11248
## TIDE5       -0.25235    0.34247  -0.737  0.46121
## TIDE6        0.25911    0.37282   0.695  0.48706
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.00094    0.37119  -2.697  0.00701 **
## YEAR2003    -0.39596    0.41490  -0.954  0.33990
## YEAR2004    -0.06410    0.44740  -0.143  0.88608
## YEAR2005    -0.64934    0.40731  -1.594  0.11089
## YEAR2006    -0.44571    0.44275  -1.007  0.31409
## YEAR2007    -0.90133    0.40662  -2.217  0.02665 *
## YEAR2008    -0.54419    0.40350  -1.349  0.17745
## YEAR2009    -0.57485    0.40996  -1.402  0.16085
## YEAR2010    -0.48173    0.40896  -1.178  0.23882
## YEAR2011     0.50818    0.46092   1.103  0.27022
## YEAR2012     0.44295    0.41634   1.064  0.28736
## YEAR2013    -0.73751    0.40358  -1.827  0.06764 .
## YEAR2014    -1.19909    0.40701  -2.946  0.00322 **
## YEAR2015     0.65014    0.42846   1.517  0.12917
## YEAR2016     0.88775    0.44550   1.993  0.04629 *
## YEAR2017    -0.48906    0.41012  -1.192  0.23308
## YEAR2018     0.61370    0.45400   1.352  0.17645
## YEAR2019     0.46491    0.42475   1.095  0.27372
## YEAR2021     1.27013    0.81465   1.559  0.11897
## MONTH7       0.99591    0.26953   3.695  0.00022 ***
## MONTH8       1.72770    0.27336   6.320 2.61e-10 ***
## MONTH9       1.33975    0.24490   5.471 4.48e-08 ***
## TIDE2        0.47192    0.19460   2.425  0.01531 *
## TIDE3       -0.04246    0.23061  -0.184  0.85392
## TIDE4        0.45241    0.22630   1.999  0.04560 *
## TIDE5       -0.44707    0.35932  -1.244  0.21342
## TIDE6        0.70554    0.38234   1.845  0.06499 .
## TEMPBOT     -0.15473    0.10375  -1.491  0.13586
## DOBOT        0.04008    0.08368   0.479  0.63193
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZANB3,test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + TIDE
##        Df    AIC    LRT  Pr(>Chi)
```

```
## <none>    2553.5
## YEAR   18 2547.2 29.627   0.04123 *
## MONTH   3 2582.7 35.200 1.105e-07 ***
## TIDE    5 2596.4 52.873 3.571e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# remove DOBOT from ZI:
ZANB4 <- glmmTMB(CPUE ~ YEAR+MONTH+TIDE, ziformula = ~
YEAR+MONTH+TIDE+TEMPBOT, data = dat,family = truncated_nbinom2(link = "log"))
summary(ZANB4)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE
## Zero inflation:        ~YEAR + MONTH + TIDE + TEMPBOT
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2551.7   2824.5  -1219.9   2439.7      907
##
##
## Dispersion parameter for truncated_nbinom2 family (): 1.18
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.61984    0.33803   1.834  0.06670 .
## YEAR2003     0.20167    0.40717   0.495  0.62038
## YEAR2004    -0.08559    0.51631  -0.166  0.86833
## YEAR2005     0.49964    0.36460   1.370  0.17057
## YEAR2006    -0.21087    0.43667  -0.483  0.62917
## YEAR2007     0.48071    0.36915   1.302  0.19285
## YEAR2008    -0.02905    0.39800  -0.073  0.94181
## YEAR2009    -0.13968    0.38476  -0.363  0.71659
## YEAR2010     0.65633    0.37200   1.764  0.07768 .
## YEAR2011     0.66146    0.45895   1.441  0.14951
## YEAR2012     0.21746    0.45015   0.483  0.62903
## YEAR2013     0.07658    0.37874   0.202  0.83977
## YEAR2014     0.27293    0.36647   0.745  0.45642
## YEAR2015    -0.24091    0.45642  -0.528  0.59762
## YEAR2016    -0.54815    0.53001  -1.034  0.30103
## YEAR2017    -0.11157    0.38162  -0.292  0.77002
## YEAR2018    -0.79932    0.56125  -1.424  0.15440
## YEAR2019     0.08942    0.42508   0.210  0.83339
## YEAR2021     1.87341    0.82560   2.269  0.02326 *
## MONTH7      -0.59547    0.18503  -3.218  0.00129 **
## MONTH8      -0.85283    0.20488  -4.162 3.15e-05 ***
## MONTH9      -1.16033    0.20796  -5.580 2.41e-08 ***
## TIDE2       -0.59219    0.20876  -2.837  0.00456 **
## TIDE3        0.88911    0.19762   4.499 6.82e-06 ***
## TIDE4        0.33420    0.21057   1.587  0.11248
```

```
## TIDE5       -0.25235    0.34247  -0.737  0.46120
## TIDE6        0.25911    0.37283   0.695  0.48706
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.99967    0.37091  -2.695 0.007035 **
## YEAR2003    -0.37492    0.41248  -0.909 0.363378
## YEAR2004    -0.07845    0.44632  -0.176 0.860474
## YEAR2005    -0.65206    0.40730  -1.601 0.109389
## YEAR2006    -0.43727    0.44205  -0.989 0.322571
## YEAR2007    -0.89446    0.40635  -2.201 0.027721 *
## YEAR2008    -0.54469    0.40324  -1.351 0.176767
## YEAR2009    -0.56526    0.40924  -1.381 0.167201
## YEAR2010    -0.47927    0.40854  -1.173 0.240736
## YEAR2011     0.53869    0.45615   1.181 0.237616
## YEAR2012     0.44455    0.41628   1.068 0.285566
## YEAR2013    -0.73405    0.40303  -1.821 0.068557 .
## YEAR2014    -1.18441    0.40547  -2.921 0.003488 **
## YEAR2015     0.67150    0.42558   1.578 0.114602
## YEAR2016     0.93578    0.43421   2.155 0.031151 *
## YEAR2017    -0.44705    0.40024  -1.117 0.264012
## YEAR2018     0.65210    0.44683   1.459 0.144455
## YEAR2019     0.50815    0.41549   1.223 0.221320
## YEAR2021     1.30249    0.81138   1.605 0.108432
## MONTH7       0.98884    0.26914   3.674 0.000239 ***
## MONTH8       1.70479    0.26913   6.334 2.38e-10 ***
## MONTH9       1.32270    0.24232   5.459 4.80e-08 ***
## TIDE2        0.46956    0.19449   2.414 0.015767 *
## TIDE3       -0.04897    0.23023  -0.213 0.831545
## TIDE4        0.44285    0.22539   1.965 0.049440 *
## TIDE5       -0.43820    0.35854  -1.222 0.221649
## TIDE6        0.70021    0.38203   1.833 0.066821 .
## TEMPBOT     -0.16168    0.10265  -1.575 0.115248
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZANB4,test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + TIDE
##        Df    AIC    LRT   Pr(>Chi)
## <none>    2551.8
## YEAR   18 2545.4 29.627   0.04123 *
## MONTH   3 2580.9 35.200 1.105e-07 ***
## TIDE    5 2594.6 52.873 3.571e-10 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# remove TEMPBOT from ZI:
ZANB5 <- glmmTMB(CPUE ~ YEAR+MONTH+TIDE, ziformula = ~ YEAR+MONTH+TIDE, data
= dat,family = truncated_nbinom2(link = "log"))
summary(ZANB5)
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR + MONTH + TIDE
## Zero inflation:        ~YEAR + MONTH + TIDE
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2552.2   2820.1  -1221.1   2442.2      908
##
##
## Dispersion parameter for truncated_nbinom2 family (): 1.18
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.61984    0.33803   1.834  0.06670 .
## YEAR2003     0.20169    0.40717   0.495  0.62036
## YEAR2004    -0.08556    0.51631  -0.166  0.86838
## YEAR2005     0.49965    0.36460   1.370  0.17056
## YEAR2006    -0.21087    0.43667  -0.483  0.62916
## YEAR2007     0.48071    0.36915   1.302  0.19285
## YEAR2008    -0.02903    0.39800  -0.073  0.94185
## YEAR2009    -0.13967    0.38476  -0.363  0.71660
## YEAR2010     0.65633    0.37200   1.764  0.07768 .
## YEAR2011     0.66144    0.45895   1.441  0.14952
## YEAR2012     0.21746    0.45014   0.483  0.62904
## YEAR2013     0.07657    0.37873   0.202  0.83978
## YEAR2014     0.27293    0.36647   0.745  0.45642
## YEAR2015    -0.24091    0.45642  -0.528  0.59762
## YEAR2016    -0.54815    0.53000  -1.034  0.30102
## YEAR2017    -0.11159    0.38162  -0.292  0.76998
## YEAR2018    -0.79932    0.56125  -1.424  0.15440
## YEAR2019     0.08944    0.42507   0.210  0.83335
## YEAR2021     1.87342    0.82560   2.269  0.02326 *
## MONTH7      -0.59547    0.18503  -3.218  0.00129 **
## MONTH8      -0.85283    0.20488  -4.163 3.15e-05 ***
## MONTH9      -1.16033    0.20796  -5.580 2.41e-08 ***
## TIDE2       -0.59219    0.20876  -2.837  0.00456 **
## TIDE3        0.88910    0.19761   4.499 6.82e-06 ***
## TIDE4        0.33419    0.21056   1.587  0.11249
## TIDE5       -0.25236    0.34247  -0.737  0.46119
## TIDE6        0.25910    0.37282   0.695  0.48707
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Zero-inflation model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.90497    0.36566  -2.475 0.013328 *
## YEAR2003    -0.31921    0.41219  -0.774 0.438691
## YEAR2004    -0.00510    0.44394  -0.011 0.990834
## YEAR2005    -0.72419    0.40406  -1.792 0.073088 .
## YEAR2006    -0.35643    0.43866  -0.813 0.416479
## YEAR2007    -0.92588    0.40522  -2.285 0.022319 *
## YEAR2008    -0.56195    0.40247  -1.396 0.162637
## YEAR2009    -0.45240    0.40309  -1.122 0.261727
## YEAR2010    -0.50688    0.40803  -1.242 0.214139
## YEAR2011     0.59700    0.45344   1.317 0.187972
## YEAR2012     0.44728    0.41708   1.072 0.283539
## YEAR2013    -0.70574    0.40087  -1.761 0.078320 .
## YEAR2014    -1.13773    0.40366  -2.819 0.004824 **
## YEAR2015     0.62294    0.42303   1.473 0.140870
## YEAR2016     0.86518    0.43282   1.999 0.045614 *
## YEAR2017    -0.44101    0.39935  -1.104 0.269453
## YEAR2018     0.57683    0.44502   1.296 0.194907
## YEAR2019     0.51200    0.41588   1.231 0.218279
## YEAR2021     1.36694    0.80906   1.690 0.091114 .
## MONTH7       0.80036    0.23936   3.344 0.000827 ***
## MONTH8       1.53111    0.24367   6.284 3.31e-10 ***
## MONTH9       1.35656    0.24078   5.634 1.76e-08 ***
## TIDE2        0.46320    0.19410   2.386 0.017017 *
## TIDE3       -0.06228    0.22982  -0.271 0.786400
## TIDE4        0.43127    0.22495   1.917 0.055214 .
## TIDE5       -0.45578    0.35806  -1.273 0.203046
## TIDE6        0.72779    0.38149   1.908 0.056427 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
drop1(ZANB5,test="Chi")
```

```
## Single term deletions
##
## Model:
## CPUE ~ YEAR + MONTH + TIDE
##        Df    AIC    LRT  Pr(>Chi)
## <none>    2552.2
## YEAR   18 2545.9 29.627   0.04123 *
## MONTH   3 2581.4 35.200 1.105e-07 ***
## TIDE    5 2595.1 52.873 3.571e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
tmbZI.resids(ZANB5,dat$CPUE)
```

```
## df.resid
## 1.333353
```

```r
# try a model with year only:
ZANB0 <- glmmTMB(CPUE ~ YEAR, ziformula = ~ YEAR, data = dat,family =
truncated_nbinom2(link = "log"))
summary(ZANB0)                 # bombed if I left the second part empty
```

```
##  Family: truncated_nbinom2  ( log )
## Formula:          CPUE ~ YEAR
## Zero inflation:        ~YEAR
## Data: dat
##
##      AIC      BIC   logLik deviance df.resid
##   2676.5   2866.4  -1299.2   2598.5      924
##
##
## Dispersion parameter for truncated_nbinom2 family (): 0.266
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.44263    0.54089  -0.818    0.413
## YEAR2003     0.26177    0.49654   0.527    0.598
## YEAR2004    -0.84446    0.60703  -1.391    0.164
## YEAR2005     0.68168    0.45842   1.487    0.137
## YEAR2006    -0.39574    0.51821  -0.764    0.445
## YEAR2007     0.23830    0.45567   0.523    0.601
## YEAR2008    -0.25473    0.47987  -0.531    0.596
## YEAR2009     0.06169    0.47338   0.130    0.896
## YEAR2010     0.62925    0.47529   1.324    0.186
## YEAR2011     0.72456    0.59153   1.225    0.221
## YEAR2012    -0.29181    0.54382  -0.537    0.592
## YEAR2013    -0.09781    0.46937  -0.208    0.835
## YEAR2014     0.15453    0.45153   0.342    0.732
## YEAR2015    -0.29178    0.55294  -0.528    0.598
## YEAR2016    -0.89668    0.62552  -1.434    0.152
## YEAR2017     0.12964    0.46889   0.276    0.782
## YEAR2018    -0.96013    0.64753  -1.483    0.138
## YEAR2019    -0.08229    0.52888  -0.156    0.876
## YEAR2021     1.10305    1.20795   0.913    0.361
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.33024    0.27336   1.208   0.2270
## YEAR2003    -0.09844    0.39075  -0.252   0.8011
## YEAR2004     0.18058    0.42590   0.424   0.6716
## YEAR2005    -0.61792    0.38425  -1.608   0.1078
## YEAR2006    -0.53091    0.41921  -1.266   0.2054
## YEAR2007    -0.76557    0.38677  -1.979   0.0478 *
## YEAR2008    -0.43943    0.38428  -1.144   0.2528
## YEAR2009    -0.47867    0.38628  -1.239   0.2153
## YEAR2010    -0.44803    0.39170  -1.144   0.2527
## YEAR2011     0.71121    0.43298   1.643   0.1005
```

```
## YEAR2012     0.41697    0.39573   1.054   0.2920
## YEAR2013    -0.58617    0.38556  -1.520   0.1284
## YEAR2014    -0.91803    0.39052  -2.351   0.0187 *
## YEAR2015     0.47414    0.39983   1.186   0.2357
## YEAR2016     0.76838    0.41227   1.864   0.0624 .
## YEAR2017    -0.47332    0.38278  -1.236   0.2163
## YEAR2018     0.68833    0.42359   1.625   0.1042
## YEAR2019     0.33624    0.39292   0.856   0.3921
## YEAR2021     1.80983    0.79596   2.274   0.0230 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
AICtab(ZANB1, ZANB2, ZANB3, ZANB4, ZANB5, ZANB0)
```

```
##       dAIC  df
## ZANB4   0.0 56
## ZANB5   0.5 55
## ZANB3   1.8 57
## ZANB2   3.6 58
## ZANB1   5.2 59
## ZANB0 124.7 39
```

```
sapply(list(ZANB1, ZANB2, ZANB3, ZANB4, ZANB5, ZANB0),
function(x){x$modelInfo$allForm$combForm} )
```

```
## [[1]]
## CPUE ~ YEAR + MONTH + TIDE + TEMPBOT + DOBOT + (YEAR + MONTH +
##     TIDE + TEMPBOT + DOBOT) + 1
##
## [[2]]
## CPUE ~ YEAR + MONTH + TIDE + DOBOT + (YEAR + MONTH + TIDE + TEMPBOT +
##     DOBOT) + 1
##
## [[3]]
## CPUE ~ YEAR + MONTH + TIDE + (YEAR + MONTH + TIDE + TEMPBOT +
##     DOBOT) + 1
##
## [[4]]
## CPUE ~ YEAR + MONTH + TIDE + (YEAR + MONTH + TIDE + TEMPBOT) +
##     1
##
## [[5]]
## CPUE ~ YEAR + MONTH + TIDE + (YEAR + MONTH + TIDE) + 1
##
## [[6]]
## CPUE ~ YEAR + YEAR + 1
```

```
lrtest(ZANB0, ZANB5, ZANB4, ZANB3, ZANB2, ZANB1)
```

```
## Likelihood ratio test
##
```

```
## Model 1: CPUE ~ YEAR
## Model 2: CPUE ~ YEAR + MONTH + TIDE
## Model 3: CPUE ~ YEAR + MONTH + TIDE
## Model 4: CPUE ~ YEAR + MONTH + TIDE
## Model 5: CPUE ~ YEAR + MONTH + TIDE + DOBOT
## Model 6: CPUE ~ YEAR + MONTH + TIDE + TEMPBOT + DOBOT
##   #Df  LogLik Df    Chisq Pr(>Chisq)
## 1  39 -1299.2
## 2  55 -1221.1 16 156.2416     <2e-16 ***
## 3  56 -1219.9  1   2.4837     0.1150
## 4  57 -1219.8  1   0.2295     0.6319
## 5  58 -1219.7  1   0.2126     0.6448
## 6  59 -1219.5  1   0.3179     0.5728
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
sapply(list(ZANB5, ZANB3, ZANB4, ZANB2, ZANB1,
ZANB0),tmbZI.resids,y.dat=dat$CPUE)
```

```
## df.resid df.resid df.resid df.resid df.resid df.resid
## 1.333353 1.339003 1.338048 1.340044 1.339399 1.792334
```

```
# So ZANB4, though equal AIC support for ZANB 5 and 3
```

Compare models

```
AICtab(ZANB4,ZINB3,NB3) # all about the same
```

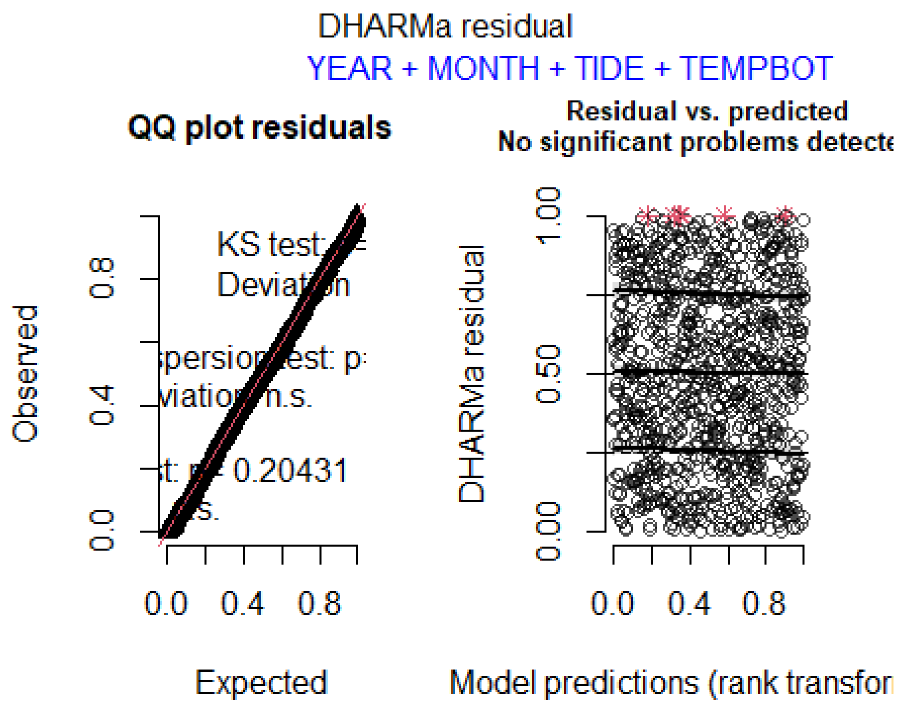
```
##       dAIC df
## ZANB4  0.0 56
## ZINB3  9.1 28
## NB3    9.1 28
```

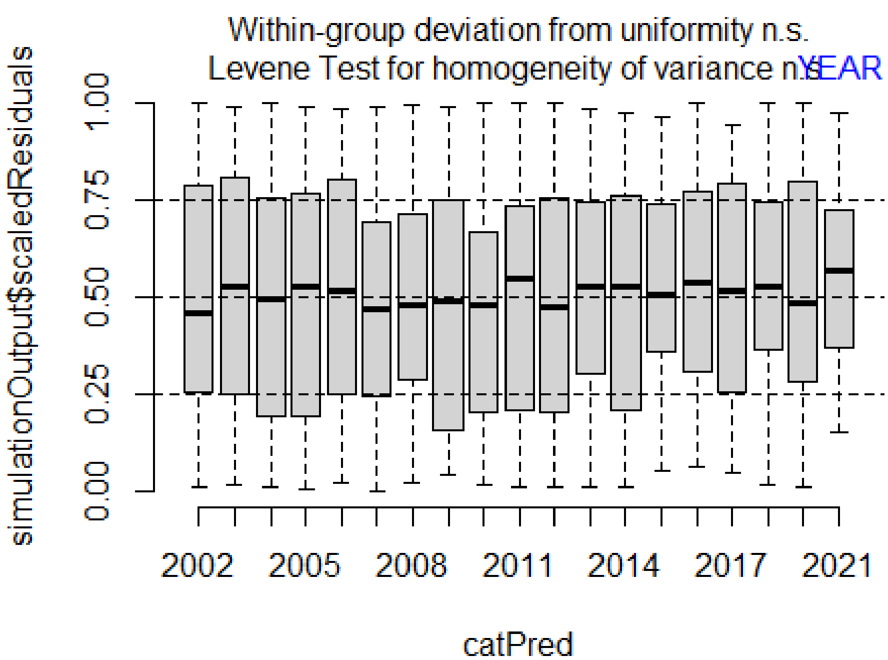
Model diagnostics

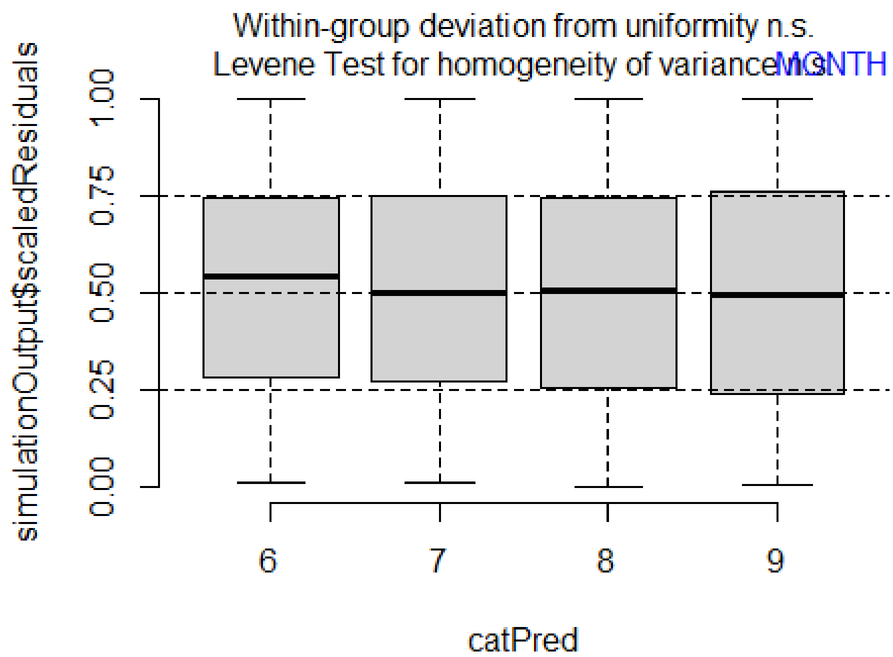
ZINB

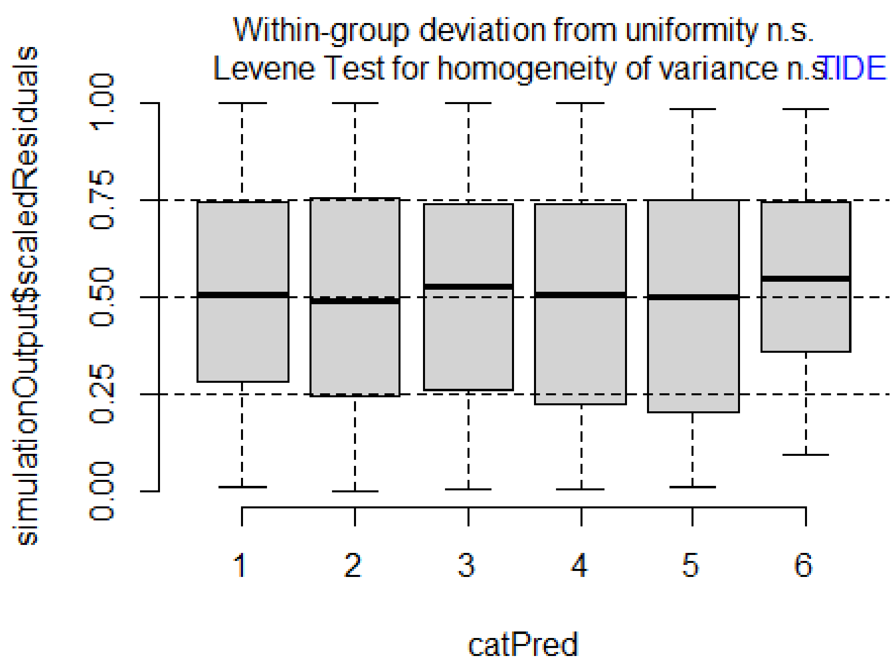
```
models <- list(ZANB4,ZINB3,NB3) # ZINB3 and NB3 had similarly good
diagnostics, therefore going with the simpler model (NB3)
for(i in 1:length(models)) {
   NB.sim <- simulateResiduals(models[[i]])
   this <- names(dat)[names(dat) %in% names(models[[i]]$frame) & names(dat)
!= "CPUE"]
   print(this)
   plot(NB.sim); mtext(paste(this,collapse=" +
"),side=3,adj=0,outer=FALSE,line=1,col="blue")

   for(j in this) {
      plotResiduals(NB.sim, form=dat[,j]); mtext(j,3,col="blue",adj=1)
   }
}
```

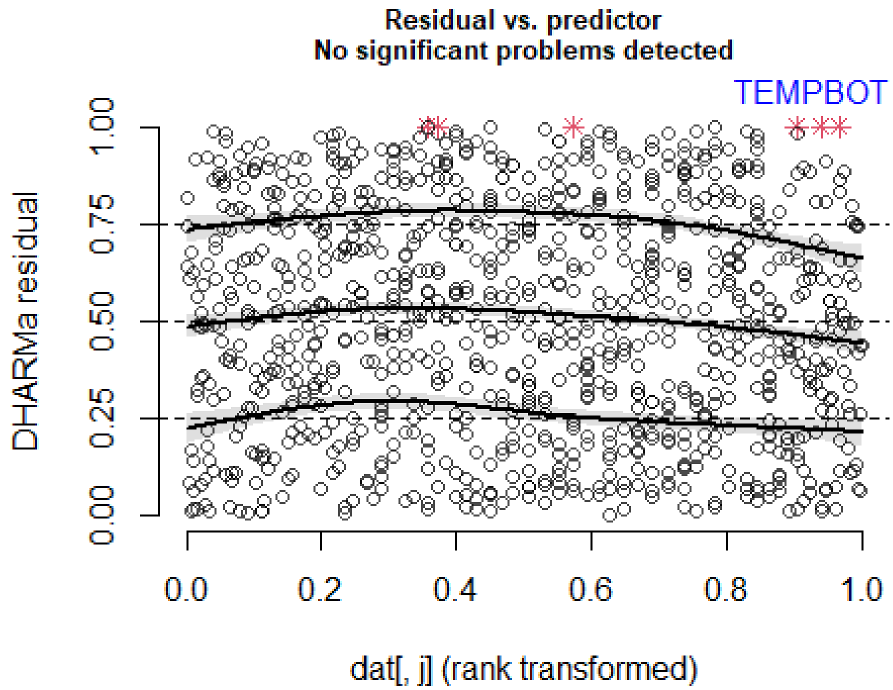
```
## [1] "YEAR"    "MONTH"   "TIDE"    "TEMPBOT"
```


DHARMa residual
YEAR + MONTH + TIDE + TEMPBOT
QQ plot residuals
Residual vs. predicted
No significant problems detecte
KS test:
Deviation
spersion test: p
viation n.s.
t: p = 0.20431
s.
Observed
Expected
DHARMa residual
Model predictions (rank transfor
0.0
0.4
0.8
0.00
0.50
1.00

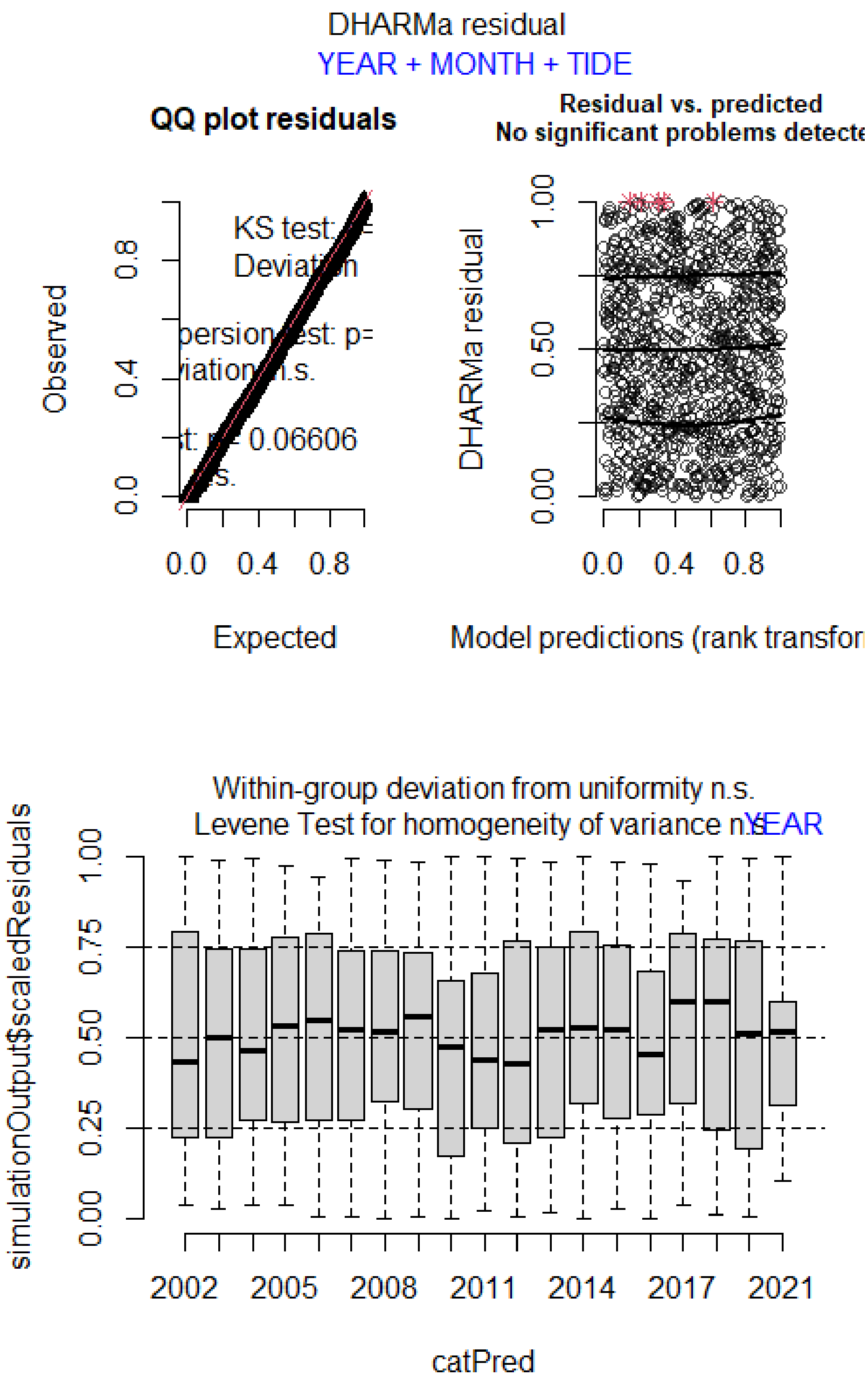

Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
YEAR
simulationOutput$scaledResiduals
catPred
0.00
0.25
0.50
0.75
1.00
2002
2005
2008
2011
2014
2017
2021


Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
MONTH
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
6
7
8
9
catPred

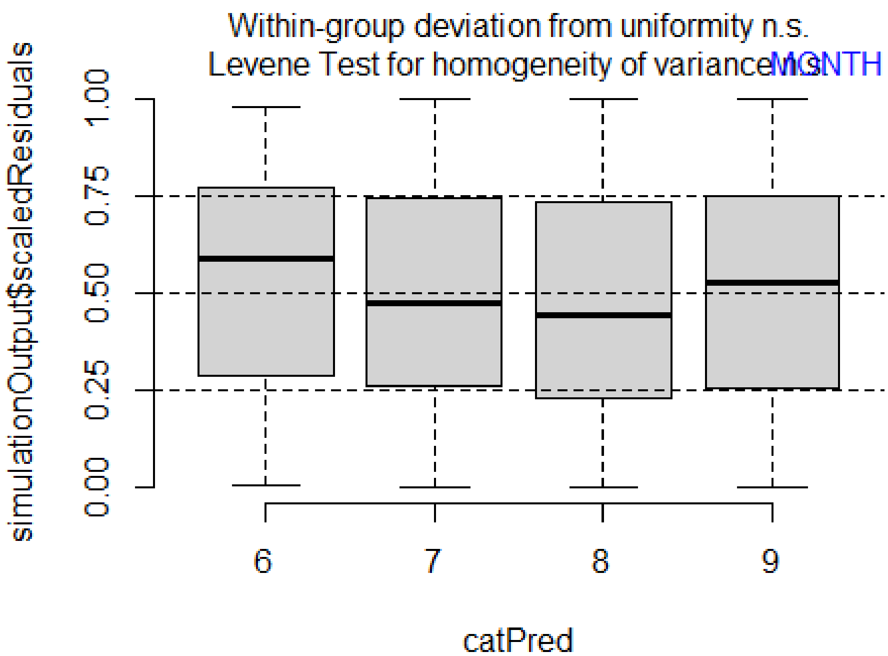

Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
TIDE
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
1
2
3
4
5
6
catPred

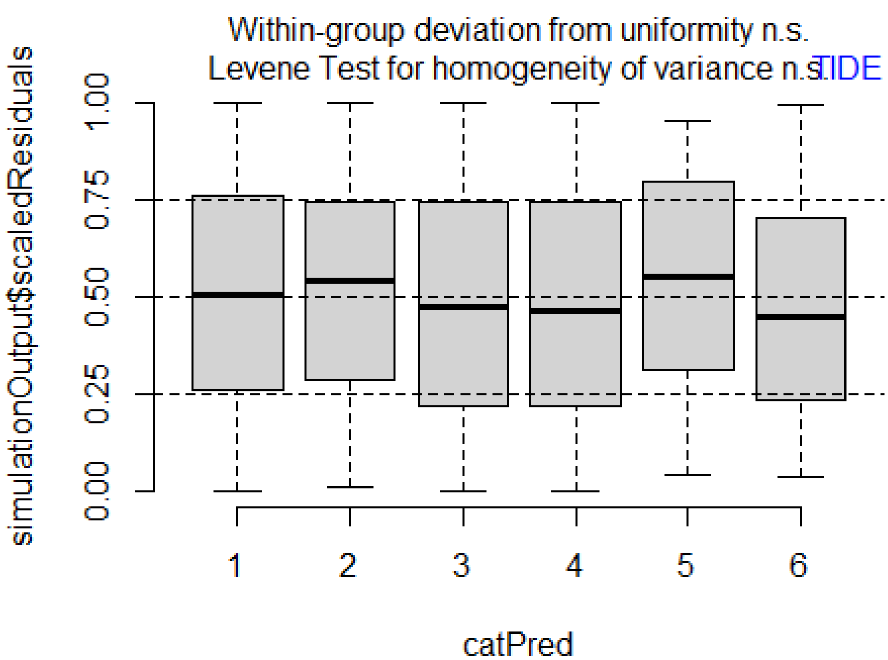



```
## [1] "YEAR"  "MONTH" "TIDE"
```

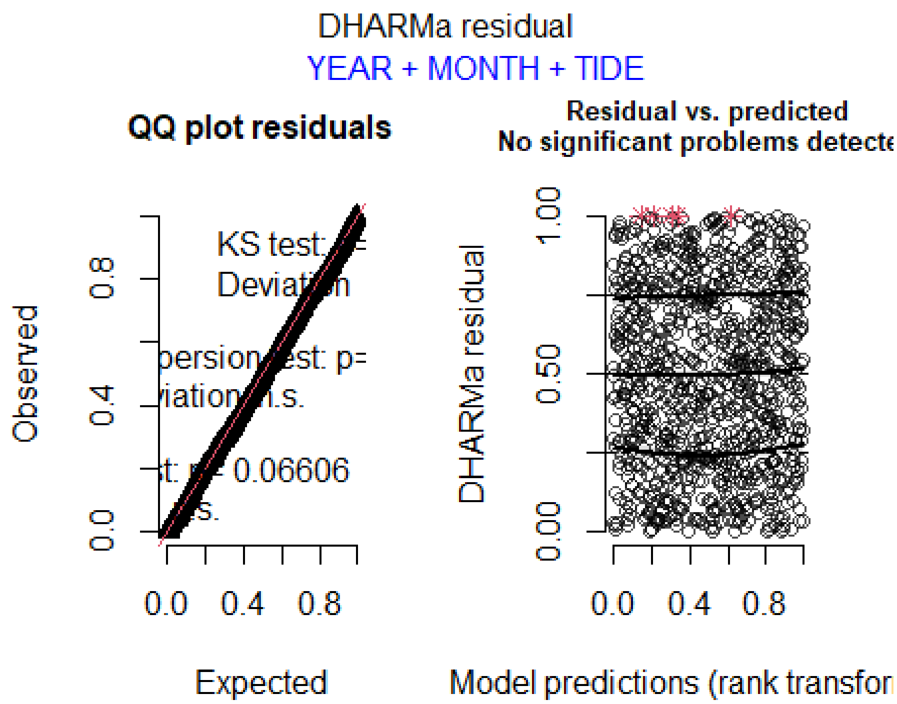


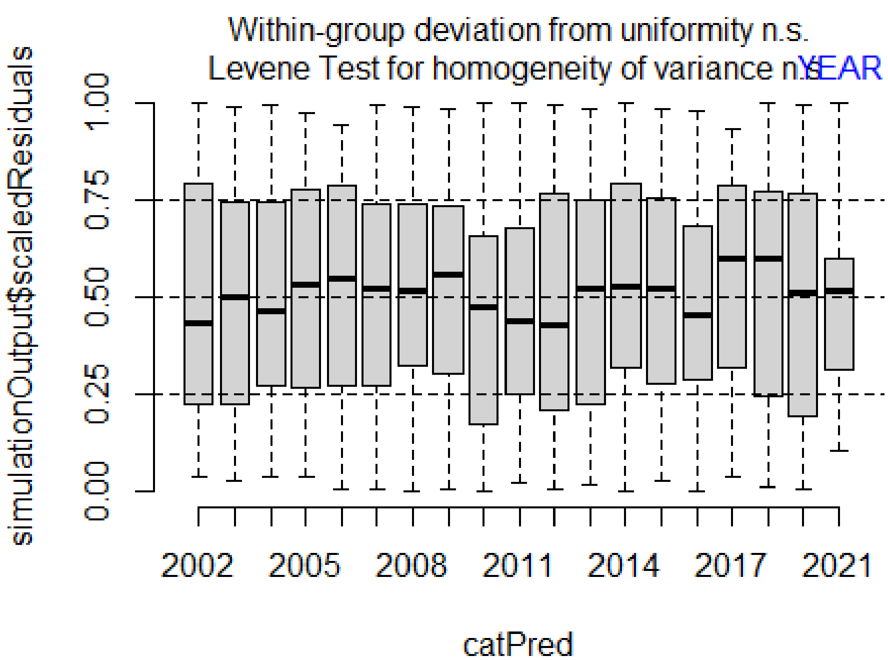
DHARMa residual
YEAR + MONTH + TIDE
QQ plot residuals
KS test:
Deviation
persion test: p=
iation n.s.
t: p= 0.06606
n.s.
Observed
Expected
Residual vs. predicted
No significant problems detecte
DHARMa residual
Model predictions (rank transfor
Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
YEAR
simulationOutput$scaledResiduals
catPred
2002
2005
2008
2011
2014
2017
2021

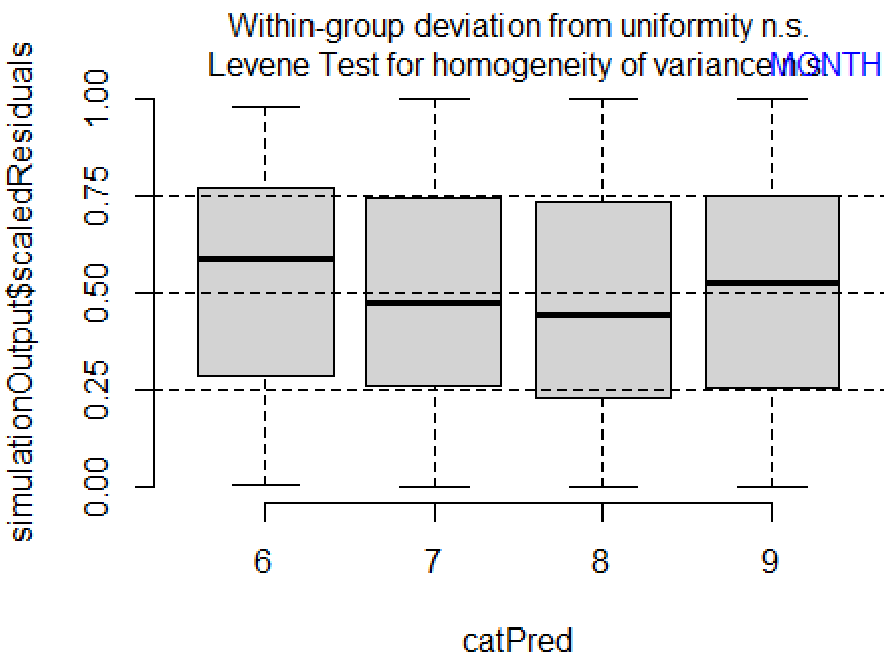


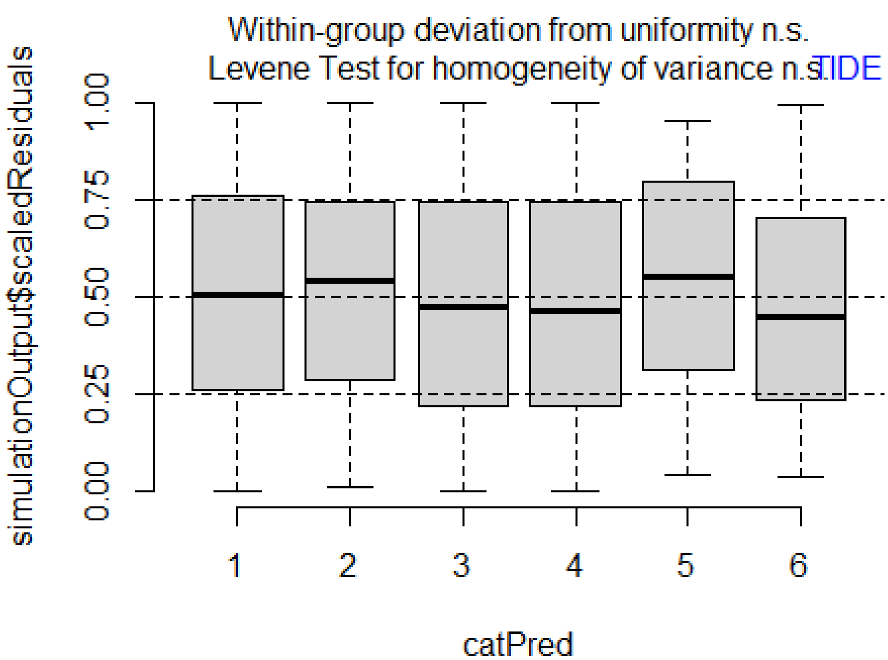



```
## [1] "YEAR"  "MONTH" "TIDE"
```


DHARMa residual
YEAR + MONTH + TIDE
QQ plot residuals
Observed
Expected
0.0
0.4
0.8
KS test:
Deviation
persion test: p=
viation n.s.
t: p= 0.06606
n.s.
Residual vs. predicted
No significant problems detecte
DHARMa residual
0.00
0.50
1.00
Model predictions (rank transfor


Within-group deviation from uniformity n.s.
Levene Test for homogeneity of variance n.s.
YEAR
simulationOutput$scaledResiduals
0.00
0.25
0.50
0.75
1.00
2002
2005
2008
2011
2014
2017
2021
catPred





Generate and compare indices

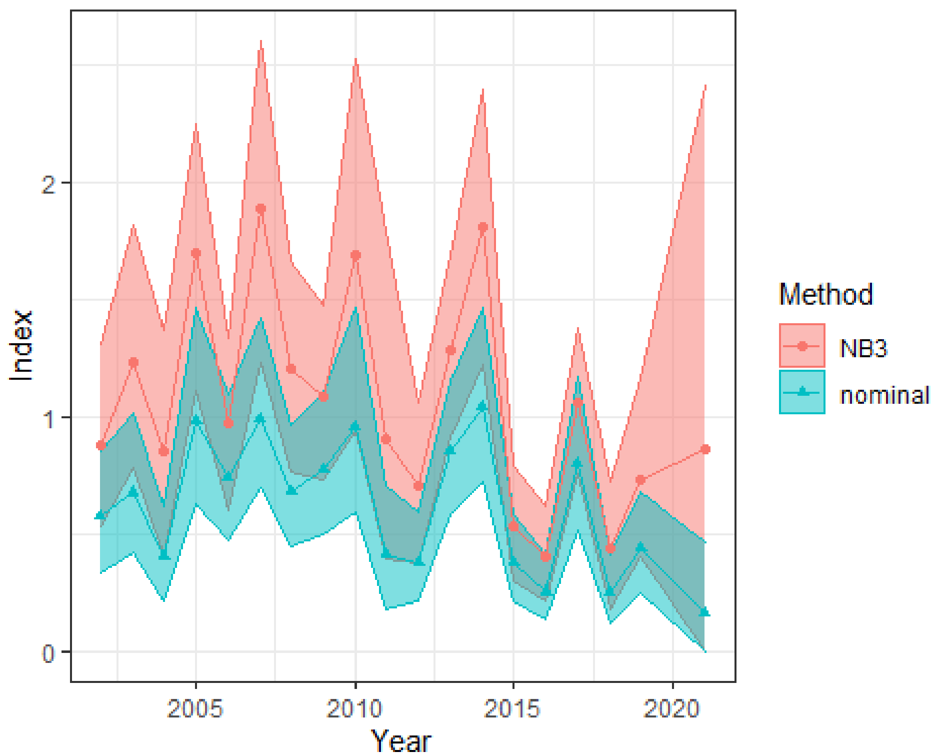
```
# by-hand with bootstrapping :
# requires final dataset object called 'dat'
out <- byHand1(theModel=NB3,nBoot=1000,Method="NB3")
```

```r
# nominal index:
nom <- matrix(NA,nrow=max(as.numeric(as.character(dat$YEAR))),ncol=7)# year,
sd, median, LCL, UCL
   colnames(nom) <- c("Year","Index","SE","LCI","UCI","CV","Method")

nom <- as.data.frame(nom)

for(i in unique(unfactor(dat$YEAR))) {
   xx <- dat[as.numeric(as.character(dat$YEAR))==i,"CPUE"]

   storeg <- NULL
   ldata <- log(xx+1)
   for(t in 1:1000){
     dodo <- sample(ldata,length(ldata),replace=TRUE)
     gmean <- exp(mean(dodo))-1
     storeg <- c(storeg,gmean)
   }
   SEB <- sd(storeg)

   #Percentile Method
   CIPB <- as.numeric(quantile(storeg,probs=c(0.025,0.50,0.975)))

   .gm <- geoMean(xx,crxn=1,alpha=0.05,warn=FALSE)
   .cv <- SEB/.gm[2]


   nom[i,1:6] <- c(i,.gm[2],SEB,CIPB[-2],.cv)
}
nom <- nom[!is.na(nom[,1]),]
nom[,"Method"] <- "nominal"

# combine nominal and standardized:
tmp <- rbind(out,nom)
tmp
```

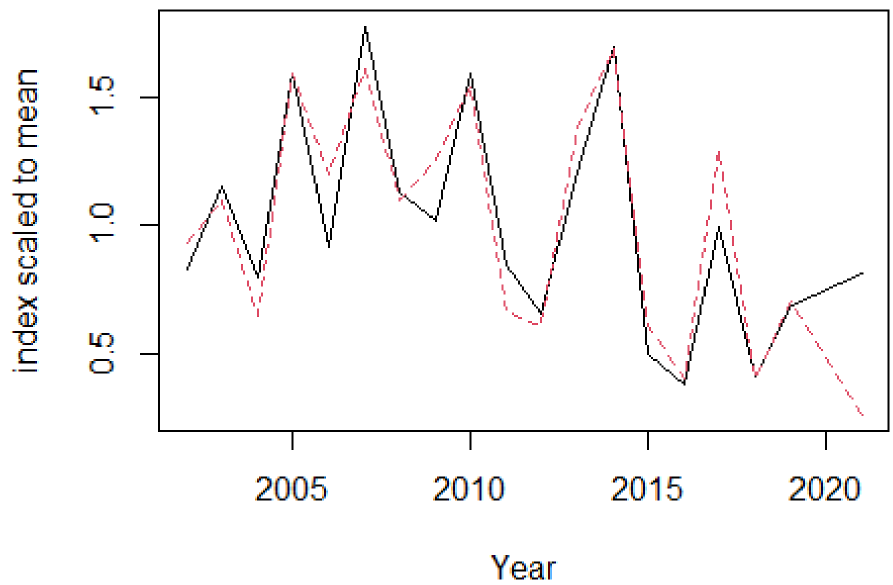
```
##        Year     Index         SE          LCI       UCI        CV  Method
## 2002   2002 0.8815738 0.20477799 5.352974e-01 1.3113629 0.2322869     NB3
## 2003   2003 1.2287746 0.26964690 7.843287e-01 1.8204463 0.2194438     NB3
## 2004   2004 0.8539983 0.24763174 4.165961e-01 1.3729316 0.2899675     NB3
## 2005   2005 1.6986143 0.30295086 1.113813e+00 2.2486926 0.1783518     NB3
## 2006   2006 0.9730515 0.18456255 6.060284e-01 1.3383793 0.1896740     NB3
## 2007   2007 1.8878019 0.35403258 1.233009e+00 2.6037382 0.1875369     NB3
## 2008   2008 1.2042003 0.23663186 7.690686e-01 1.6601846 0.1965054     NB3
## 2009   2009 1.0838689 0.18974318 7.328033e-01 1.4739398 0.1750610     NB3
## 2010   2010 1.6924213 0.40711696 9.400741e-01 2.5248011 0.2405530     NB3
## 2011   2011 0.9017150 0.35235737 3.950958e-01 1.7955159 0.3907636     NB3
## 2012   2012 0.7033801 0.17461525 3.802887e-01 1.0578527 0.2482516     NB3
```

```
## 2013  2013 1.2871007 0.19544462 9.159786e-01 1.6903995 0.1518487     NB3
## 2014  2014 1.8063854 0.30109381 1.225393e+00 2.3987252 0.1666830     NB3
## 2015  2015 0.5351332 0.12426402 3.015545e-01 0.7946824 0.2322114     NB3
## 2016  2016 0.4053505 0.10597295 2.139956e-01 0.6204458 0.2614354     NB3
## 2017  2017 1.0595768 0.16431240 7.564357e-01 1.3783026 0.1550736     NB3
## 2018  2018 0.4419028 0.14621899 1.805434e-01 0.7270971 0.3308849     NB3
## 2019  2019 0.7345876 0.19814136 4.026398e-01 1.1616077 0.2697314     NB3
## 2021  2021 0.8635320 0.69219740 1.015000e-08 2.4174455 0.8015886     NB3
## 20021 2002 0.5767191 0.13033872 3.352372e-01 0.8535770 0.2260004 nominal
## 20031 2003 0.6759429 0.15630239 4.249464e-01 1.0204357 0.2312361 nominal
## 20041 2004 0.4022508 0.10210334 2.171410e-01 0.6155132 0.2538301 nominal
## 20051 2005 0.9822209 0.21767096 6.247493e-01 1.4660122 0.2216110 nominal
## 20061 2006 0.7411502 0.16006529 4.700956e-01 1.0910933 0.2159687 nominal
## 20071 2007 0.9929787 0.18986769 6.966105e-01 1.4200557 0.1912102 nominal
## 20081 2008 0.6792706 0.13157915 4.487125e-01 0.9673487 0.1937065 nominal
## 20091 2009 0.7759143 0.15752617 5.029000e-01 1.1035944 0.2030201 nominal
## 20101 2010 0.9561256 0.21312262 5.980458e-01 1.4640739 0.2229023 nominal
## 20111 2011 0.4092160 0.13943718 1.840791e-01 0.7070874 0.3407423 nominal
## 20121 2012 0.3776576 0.09559435 2.186709e-01 0.5965552 0.2531244 nominal
## 20131 2013 0.8518273 0.14950201 5.851162e-01 1.1537544 0.1755074 nominal
## 20141 2014 1.0423567 0.18596554 7.214479e-01 1.4635641 0.1784087 nominal
## 20151 2015 0.3758388 0.09859721 2.166150e-01 0.5855239 0.2623391 nominal
## 20161 2016 0.2533894 0.07202607 1.341129e-01 0.4126131 0.2842506 nominal
## 20171 2017 0.8000664 0.16377715 5.163782e-01 1.1750784 0.2047044 nominal
## 20181 2018 0.2505815 0.07545499 1.198188e-01 0.4127517 0.3011196 nominal
## 20191 2019 0.4374584 0.11070604 2.457819e-01 0.6841019 0.2530665 nominal
## 20211 2021 0.1643056 0.13626479 0.000000e+00 0.4672819 0.8293376 nominal
```

```
ggplot(tmp, aes(x=as.numeric(Year), y=Index, color=Method, shape=Method)) +
  geom_ribbon(aes(x=as.numeric(Year), ymin=LCI, ymax=UCI, fill=Method), 
alpha=0.5) +
  geom_line() + geom_point() +
  xlab("Year") + theme_bw()
```




```r
# scale indices to respective means:
matplot(unique(tmp$Year),sapply(split(tmp,tmp$Method), function(x)
scaleToMean(x$Index)),
        type="l",xlab="Year",ylab="index scaled to mean")
```




```
round(cor(sapply(split(tmp,tmp$Method), function(x) x$Index)),2)

##           NB3 nominal
## NB3     1.00    0.91
## nominal 0.91    1.00

round(cor(sapply(split(tmp,tmp$Method), function(x)
x$Index),method="spearman"),2)

##           NB3 nominal
## NB3     1.00    0.92
## nominal 0.92    1.00

Sys.time()

## [1] "2022-06-07 11:03:41 EDT"
```

Summary:

- ZANB4, ZINB3, NB3 all were all within 0.3 delta AIC units of each other. Diagnostics for all three models were at least 'good enough,' though ZINB3 and NB3 were good. Given the AIC results and diagnostics, I went with the simplest model (NB3).

- ZINB factor selection was problematic; there were convergence problems or large SE issues until I removed all zero-inflated model terms. This is another line of support for not choosing ZINB as the final model (and see previous bullet). I suspect that ZINB without a zero-inflated model is equivalent to a NB model (which is why the diagnostics and delta AIC look identical).

Bluefish index standardization - Maryland

Abigail Tyrell

2022-05-18

Develop standardized index

Step 1: Data processing

Load functions

```
source(here::here("scripts/functions.R"))

## Loading required package: carData

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Warning: package 'glmmTMB' was built under R version 4.0.5

## Registered S3 methods overwritten by 'lme4':
##   method                          from
##   cooks.distance.influence.merMod car
##   influence.merMod                car
##   dfbeta.influence.merMod         car
##   dfbetas.influence.merMod        car

## Warning: package 'DHARMa' was built under R version 4.0.5

## This is DHARMa 0.4.4. For overview type '?DHARMa'. For recent changes,
type news(package = 'DHARMa')

## Warning: package 'bbmle' was built under R version 4.0.5

## Loading required package: stats4

## Warning: package 'ggeffects' was built under R version 4.0.5

## Warning: package 'ggplot2' was built under R version 4.0.5
```

Knit tables with `knitr::kable`

```
knit_print.data.frame <- function(x, ...) {
  res <- paste(c("", "", knitr::kable(x, digits = 2)), collapse = "\n")
```

```r
  knitr::asis_output(res)
}

registerS3method(
  "knit_print", "data.frame", knit_print.data.frame,
  envir = asNamespace("knitr")
)
```

Read in data

```r
indata <- readxl::read_excel(here::here("data/2021_bluefish_data_MD.xlsx"),
  sheet = "SurveyData",
  na = "."
)
head(indata, n = 2)
```

STATE	PROGRAM	GEAR	YEAR	MONTH	DAY	ROUND	SYSTEM	RIVER	SITETYPE	SITENO	SITENAME	HAULNO	ITISCODE	SPECNAME	SPECCNT	SIZEMIN	SIZEMAX	BOTTYPE1	BOTTYPE2	DEPMAX	TEMPWATR	SAL	DISOFFS	BF_SUBSET_
MD	CHESBAYSEINESURVEY	SEINE	1981	7	13	I	CHOPTANK	CHOPTANKRIVER	PERMANENT	67	WARWICKCR	1	168559	BLUEFISH	0	NA	NA	SAND	TYPE1ONLY	2	29.9	5.1	33	SUBSET
MD	CHES	SEIN	1981	7	13	I	CHOP	CHOP	PERM	67	WARW	2	1685	BLUE	0	NA	NA	SAND	TYPE	2	29.9	5.1	33	SUBS

BAYSEINESURVEY	E	TANK	TANKRIVER	ANENT	ICKCR	59	FISH	1ONLY	ET

```r
# modify data formatting
indata <- indata %>%
  dplyr::rename(FREQ = SPECCNT) %>%
  dplyr::mutate(
    SampleID = as.factor(dplyr::row_number()),
    HAULNO = as.factor(HAULNO),
    DEPMAX = as.factor(DEPMAX)
  ) %>%
  dplyr::select(-SIZEMIN, -SIZEMAX) %>% # too much missing data
  dplyr::select(-c(STATE, PROGRAM, GEAR, ITISCODE, SPECNAME, BF_SUBSET_)) #
same data for all entries

head(indata, n = 2)
```

YEAR	MONTH	DAY	ROUND	SYSTEM	RIVER	SITETYPE	SITENO	SITENAME	HAULNO	FREQ	BOTTYPE1	BOTTYPE2	DEPMAX	TEMPWATR	SAL	DISOFFS	SampleID
1981	7	13	I	CHOPTANK	CHOPTANKRIVER	PERMANENT	67	WARWICKCR	1	0	SAND	TYPE1ONLY	2	29.9	5.1	33	1
1981	7	13	I	CHOPTANK	CHOPTANKRIVER	PERMANENT	67	WARWICKCR	2	0	SAND	TYPE1ONLY	2	29.9	5.1	33	2

Remove missing data

```
# remove UNKNOWN bottom types
length1 <- nrow(indata)
new_data <- indata %>%
  dplyr::filter(
    BOTTYPE1 != "UNKNOWN",
    BOTTYPE2 != "UNKNOWN"
  )
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")

## 10 rows removed

# remove HAULNO2
length1 <- nrow(new_data)
new_data <- new_data %>%
  dplyr::filter(HAULNO != 2) %>%
  dplyr::select(-HAULNO) # remove because there is only 1 value now
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")

## 1500 rows removed

new_data <- remove_missing(new_data,
  f_col = "FREQ",
  t_col = "YEAR"
)

## Data before removing missing values:

##       YEAR          MONTH             DAY           ROUND          
##  Min.   :1981   Min.   :7.000   Min.   : 2.00   Length:1591       
##  1st Qu.:1991   1st Qu.:7.000   1st Qu.: 8.00   Class :character  
##  Median :2001   Median :8.000   Median :11.00   Mode  :character  
##  Mean   :2001   Mean   :7.992   Mean   :11.65                     
##  3rd Qu.:2011   3rd Qu.:9.000   3rd Qu.:15.00                     
##  Max.   :2021   Max.   :9.000   Max.   :31.00                     
##                                                                   
##     SYSTEM             RIVER             SITETYPE             SITENO      
##  Length:1591        Length:1591        Length:1591        Min.   :  4.00  
##  Class :character   Class :character   Class :character   1st Qu.: 11.00  
##  Mode  :character   Mode  :character   Mode  :character   Median : 52.00  
##                                                           Mean   : 58.04  
##                                                           3rd Qu.: 88.00  
##                                                           Max.   :172.00  
##                                                                         
##    SITENAME              FREQ          BOTTYPE1           BOTTYPE2        
##  Length:1591        Min.   : 0.0000   Length:1591        Length:1591       
##  Class :character   1st Qu.: 0.0000   Class :character   Class :character  
##  Mode  :character   Median : 0.0000   Mode  :character   Mode  :character  
```

```
##                     Mean   : 0.5035
##                     3rd Qu.: 0.0000
##                     Max.   :54.0000
##
##   DEPMAX       TEMPWATR          SAL            DISOFFS         SampleID
##  1   : 11   Min.   : 1.90   Min.   : 0.000   Min.   : 6.0   1      :   1
##  2   :122   1st Qu.:22.18   1st Qu.: 2.400   1st Qu.:18.0   3      :   1
##  2.5 :  1   Median :26.10   Median : 6.100   Median :30.0   5      :   1
##  3   :405   Mean   :21.15   Mean   : 6.246   Mean   :24.9   7      :   1
##  4   :978   3rd Qu.:27.73   3rd Qu.: 9.600   3rd Qu.:30.0   9      :   1
##  5   : 31   Max.   :33.50   Max.   :20.000   Max.   :33.0   11     :   1
##  NA's: 43   NA's   :23      NA's   :62       NA's   :2      (Other):1585
```

```
## Data after removing missing values:
```

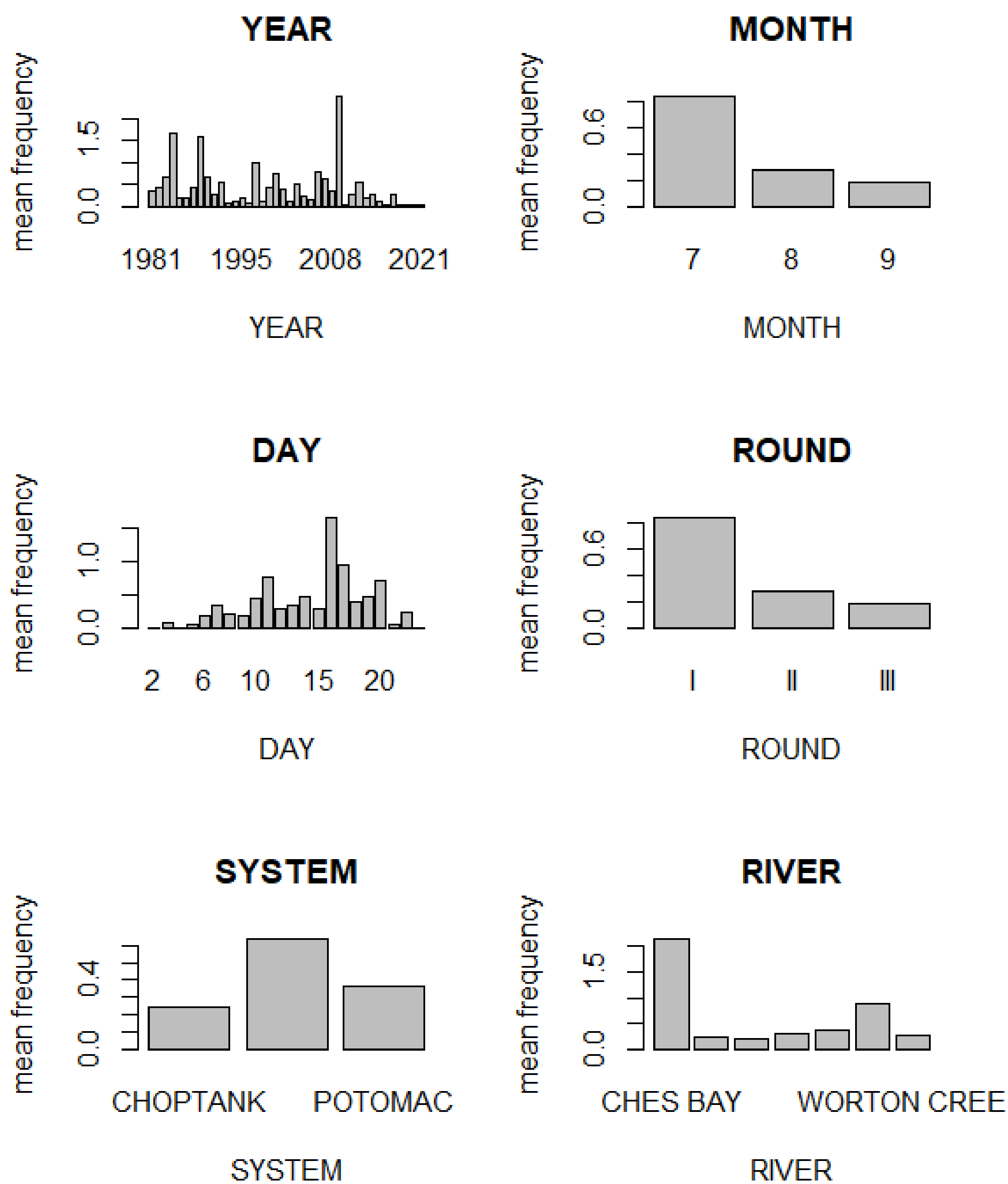
```
##       YEAR          MONTH            DAY          ROUND
##  Min.   :1981   Min.   :7.000   Min.   : 2.00   Length:1482
##  1st Qu.:1992   1st Qu.:7.000   1st Qu.: 8.00   Class :character
##  Median :2002   Median :8.000   Median :11.00   Mode  :character
##  Mean   :2002   Mean   :8.003   Mean   :11.46
##  3rd Qu.:2012   3rd Qu.:9.000   3rd Qu.:15.00
##  Max.   :2021   Max.   :9.000   Max.   :31.00
##
##     SYSTEM             RIVER             SITETYPE            SITENO
##  Length:1482        Length:1482        Length:1482        Min.   :  4.00
##  Class :character   Class :character   Class :character   1st Qu.: 11.00
##  Mode  :character   Mode  :character   Mode  :character   Median : 52.00
##                                                           Mean   : 59.43
##                                                           3rd Qu.: 88.00
##                                                           Max.   :172.00
##
##    SITENAME              FREQ          BOTTYPE1          BOTTYPE2
##  Length:1482        Min.   : 0.0000   Length:1482        Length:1482
##  Class :character   1st Qu.: 0.0000   Class :character   Class :character
##  Mode  :character   Median : 0.0000   Mode  :character   Mode  :character
##                     Mean   : 0.4366
##                     3rd Qu.: 0.0000
##                     Max.   :54.0000
##
##  DEPMAX       TEMPWATR          SAL            DISOFFS          SampleID
##  1  : 10   Min.   : 1.90   Min.   : 0.000   Min.   : 6.00   1      :   1
##  2  :116   1st Qu.:22.32   1st Qu.: 2.500   1st Qu.:18.00   3      :   1
##  2.5:  1   Median :26.10   Median : 6.100   Median :30.00   5      :   1
##  3  :386   Mean   :21.35   Mean   : 6.275   Mean   :24.74   7      :   1
##  4  :940   3rd Qu.:27.70   3rd Qu.: 9.690   3rd Qu.:30.00   9      :   1
##  5  : 29   Max.   :33.50   Max.   :20.000   Max.   :33.00   11     :   1
##                                                             (Other):1476
```

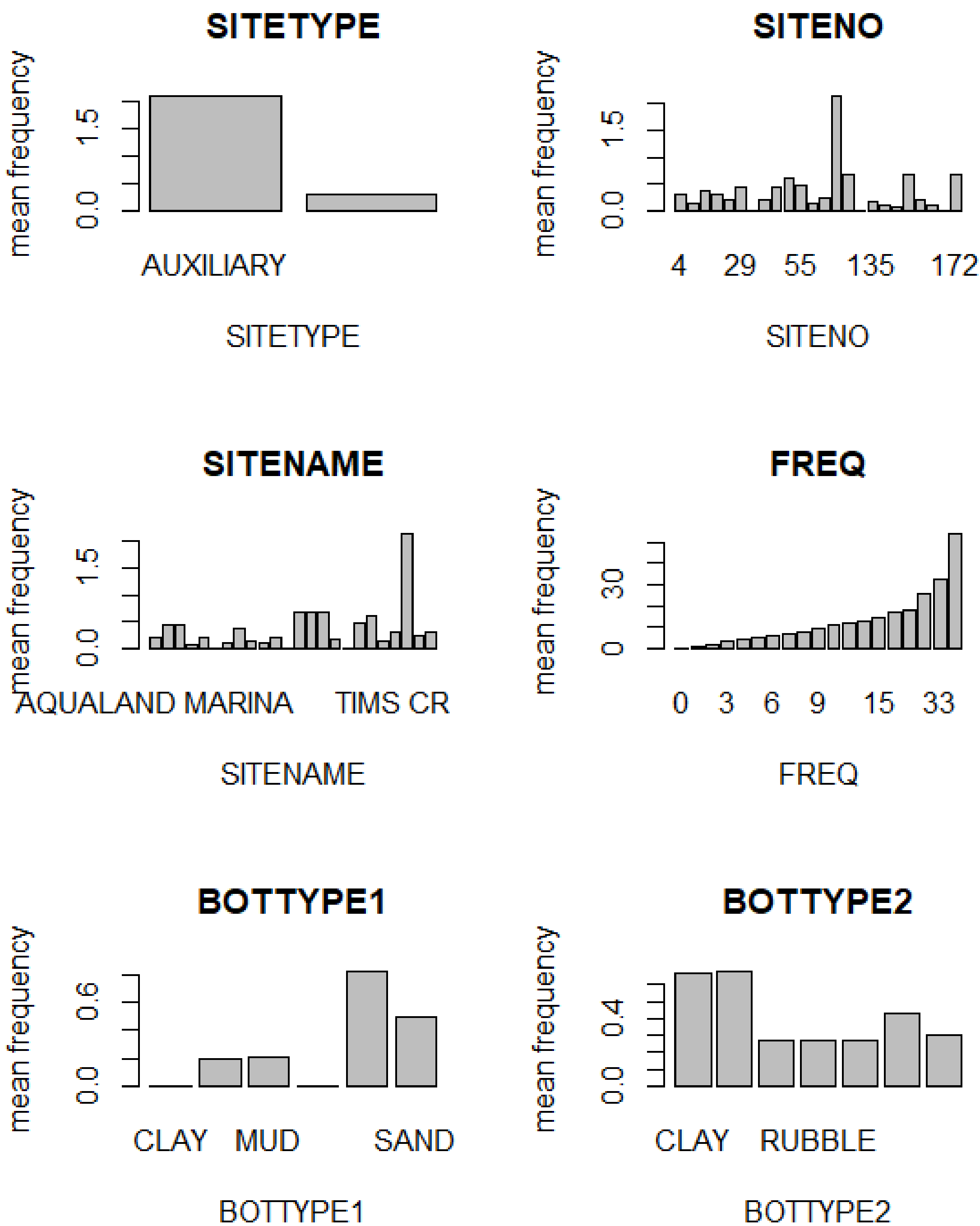
```
## 109 datapoints removed (6.9% of data removed)
```

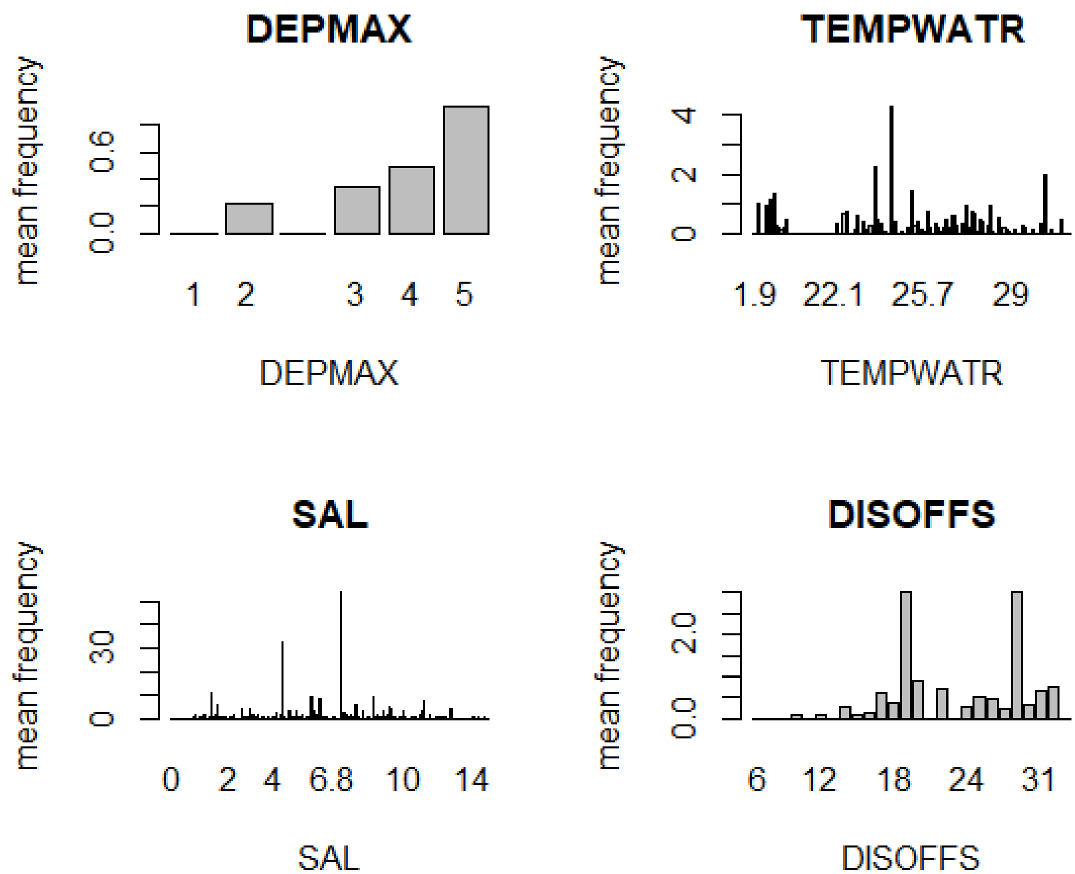
```
## All years had positive catch
```

Plot data and decide if any data should be removed due to sparse sampling

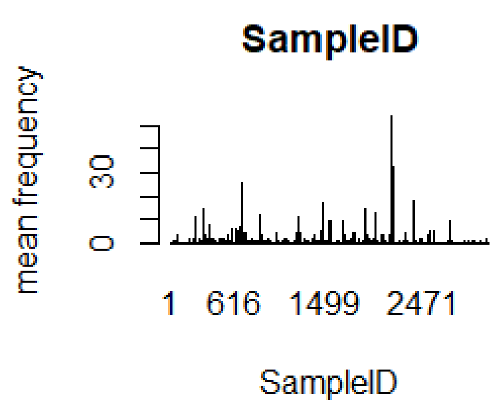
```
plot_obs(new_data,
  col = "FREQ"
)
```




SITETYPE
mean frequency
0.0
1.5
AUXILIARY
SITETYPE
SITENO
mean frequency
0.0
1.5
4
29
55
135
172
SITENO
SITENAME
mean frequency
0.0
1.5
AQUALAND MARINA
TIMS CR
SITENAME
FREQ
mean frequency
0
30
0
3
6
9
15
33
FREQ
BOTTYPE1
mean frequency
0.0
0.6
CLAY
MUD
SAND
BOTTYPE1
BOTTYPE2
mean frequency
0.0
0.4
CLAY
RUBBLE
BOTTYPE2




```
## Please manually remove sparsely sampled data!
```

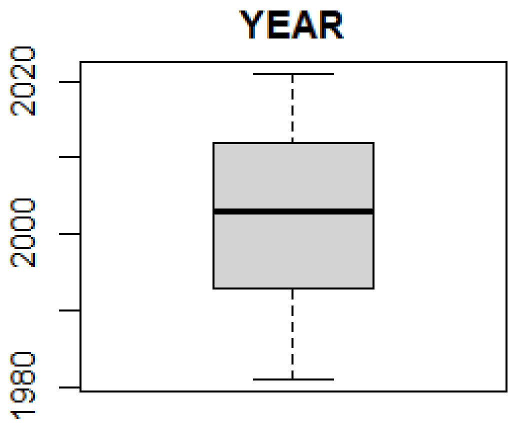


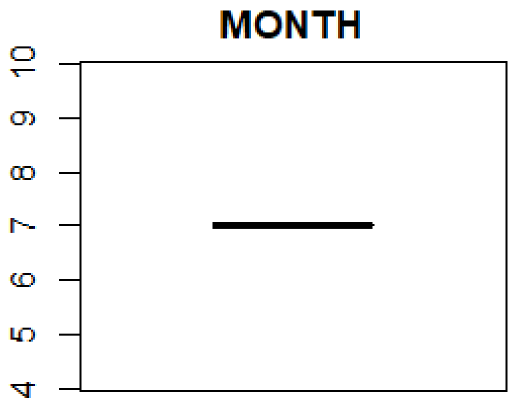
If data has to be removed, edit this chunk to remove it

```r
# dropping August and September
new_data <- new_data %>%
  dplyr::filter(MONTH == 7)
```

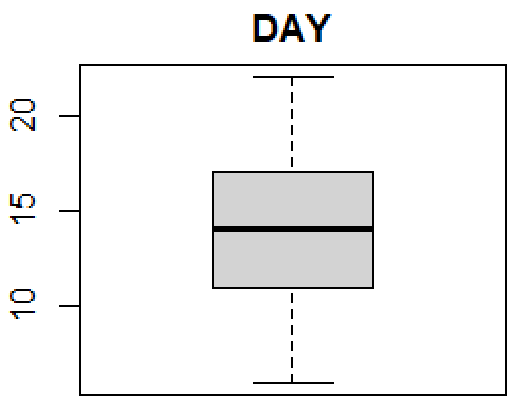
Check for outliers

```r
plot_outliers(new_data)
```

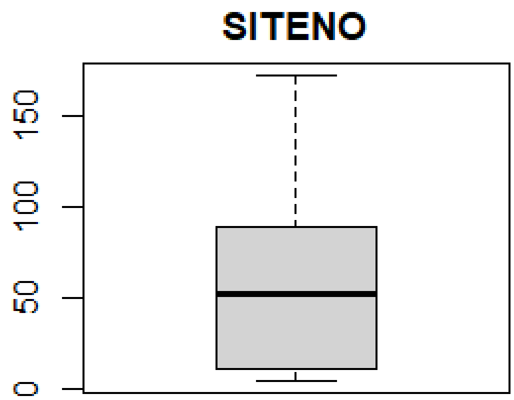




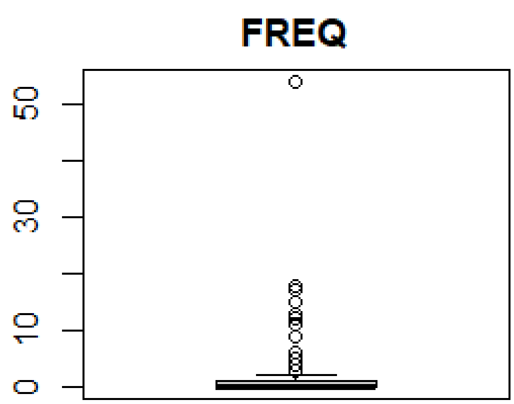

```
## ROUND is not numeric

## SYSTEM is not numeric

## RIVER is not numeric

## SITETYPE is not numeric
```

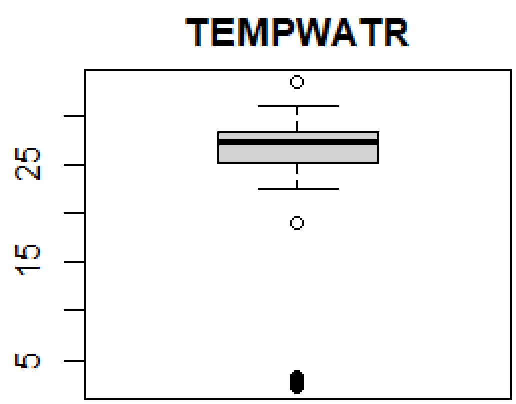


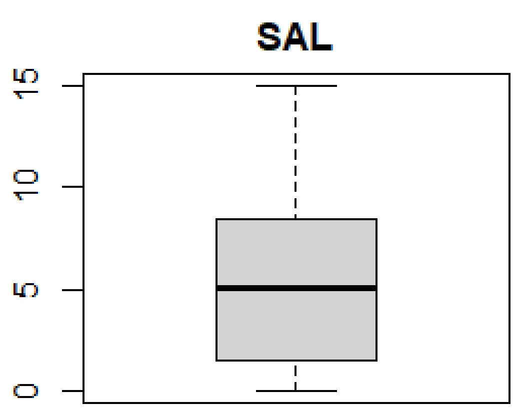

```
## SITENAME is not numeric
```




```
## BOTTYPE1 is not numeric

## BOTTYPE2 is not numeric

## DEPMAX is not numeric
```

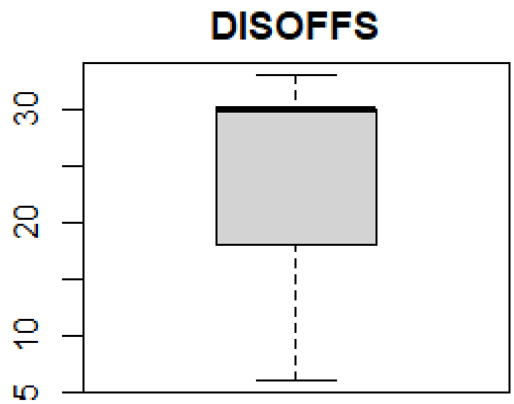







```
## SampleID is not numeric
```

```
## Please manually remove any outliers!
```

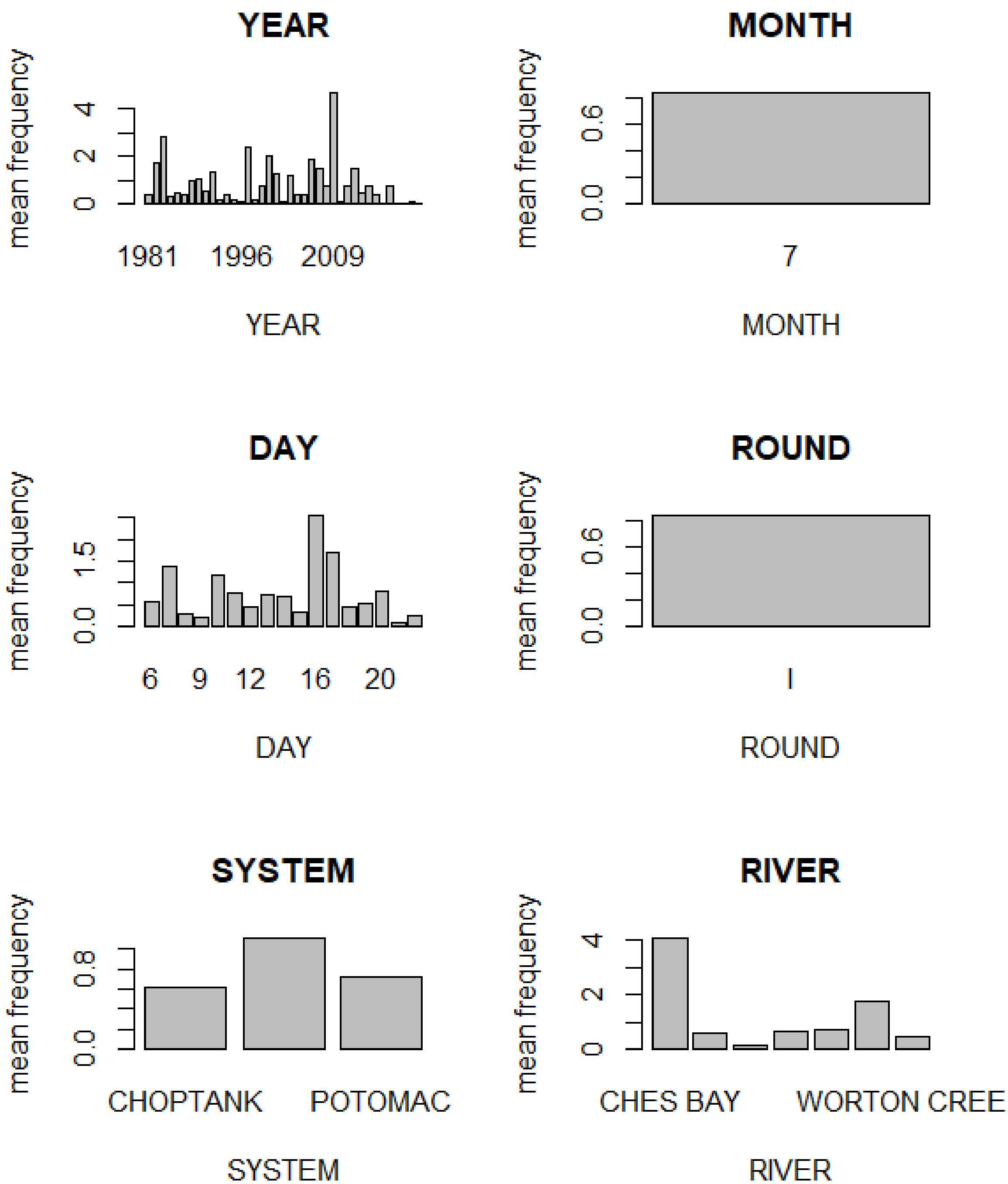


If needed, remove outliers

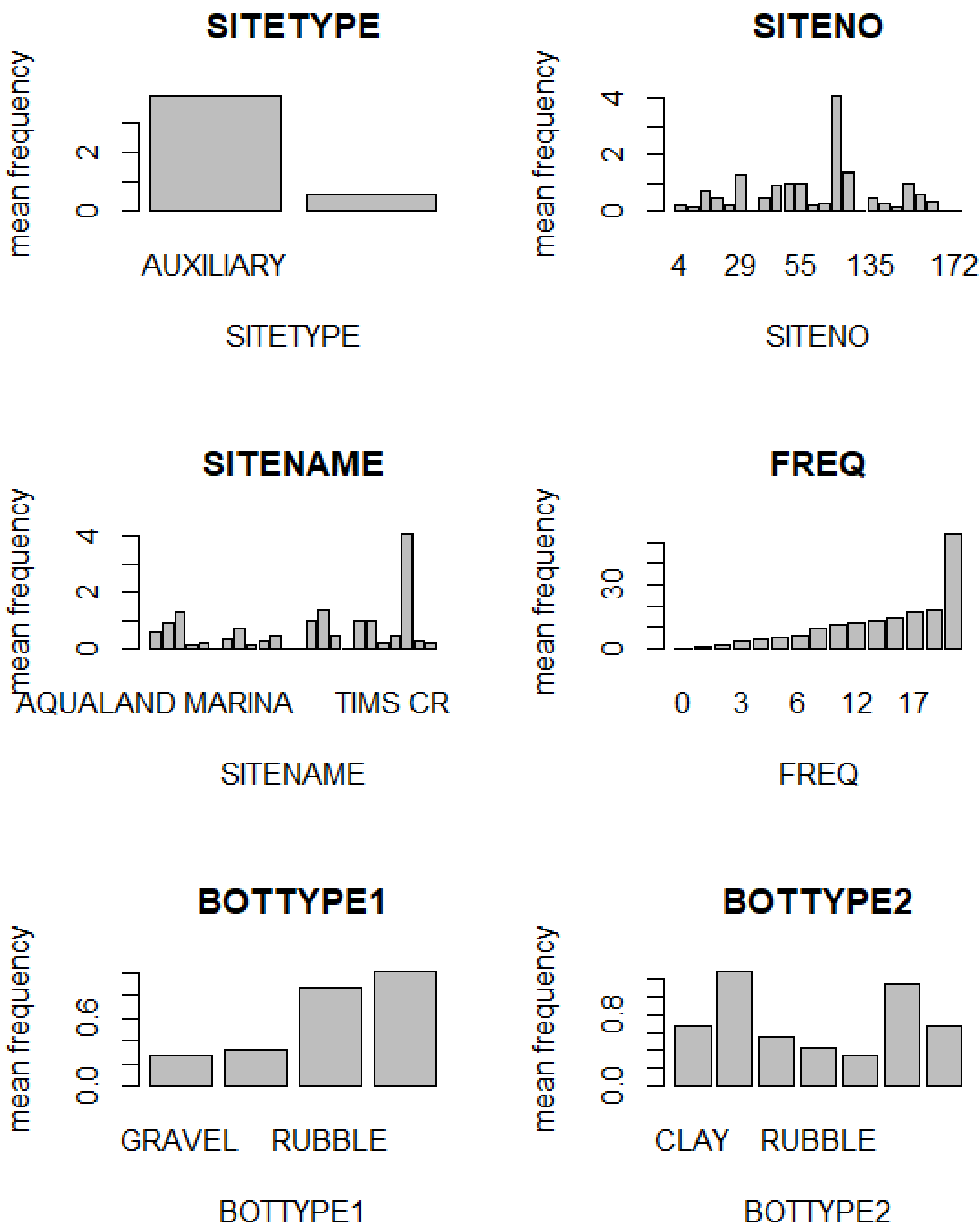
```
# Fix temperatures <5
new_data <- new_data %>%
  dplyr::mutate(TEMPWATR = ifelse(TEMPWATR < 5, TEMPWATR * 10, TEMPWATR))
```

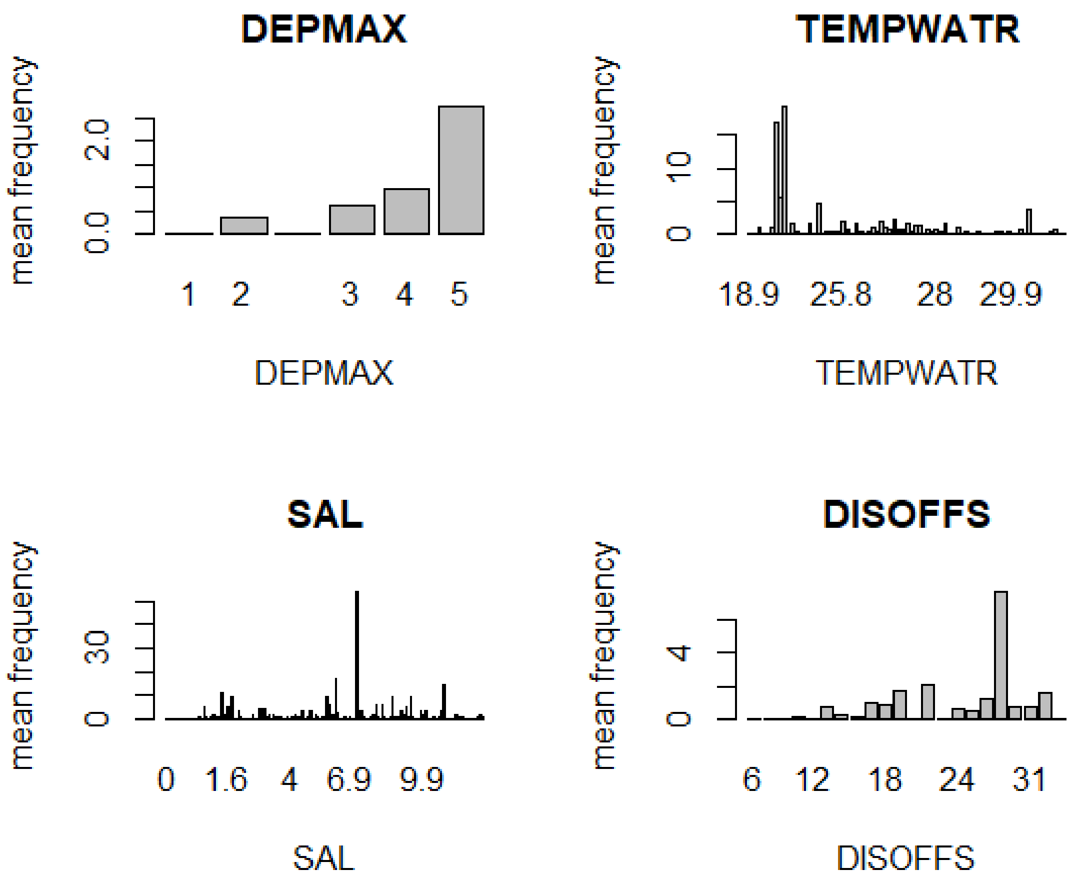
Check for sparse sampling again

```
plot_obs(new_data,
  col = "FREQ"
)
```

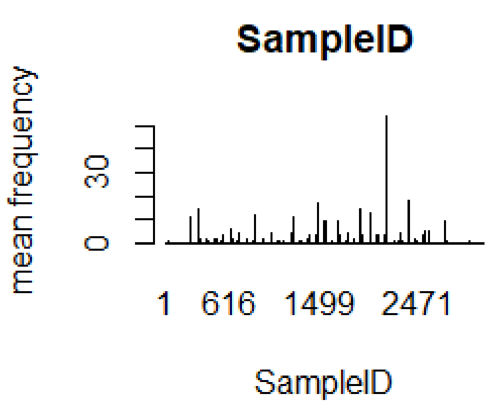


YEAR
mean frequency
0
2
4
1981
1996
2009
YEAR
MONTH
mean frequency
0.0
0.6
7
MONTH
DAY
mean frequency
0.0
1.5
6
9
12
16
20
DAY
ROUND
mean frequency
0.0
0.6
I
ROUND
SYSTEM
mean frequency
0.0
0.8
CHOPTANK
POTOMAC
SYSTEM
RIVER
mean frequency
0
2
4
CHES BAY
WORTON CREEI
RIVER


SITETYPE
mean frequency
0
2
AUXILIARY
SITETYPE
SITENO
mean frequency
0
2
4
4
29
55
135
172
SITENO
SITENAME
mean frequency
0
2
4
AQUALAND MARINA
TIMS CR
SITENAME
FREQ
mean frequency
0
30
0
3
6
12
17
FREQ
BOTTYPE1
mean frequency
0.0
0.6
GRAVEL
RUBBLE
BOTTYPE1
BOTTYPE2
mean frequency
0.0
0.8
CLAY
RUBBLE
BOTTYPE2




```
## Please manually remove sparsely sampled data!
```



Remove ROUND column

```
new_data <- new_data %>%
  dplyr::select(-ROUND)
```

Check again for years with no catch

```
check_missing(new_data, f_col = "FREQ", t_col = "YEAR")

## Years with no catch: c(2016, 2018, 2019, 2021)
```

Remove years with no catch (?)

```
new_data <- new_data %>%
  dplyr::filter(!YEAR %in% c(2016, 2018, 2019, 2021))
```

Change continuous variables to z-score

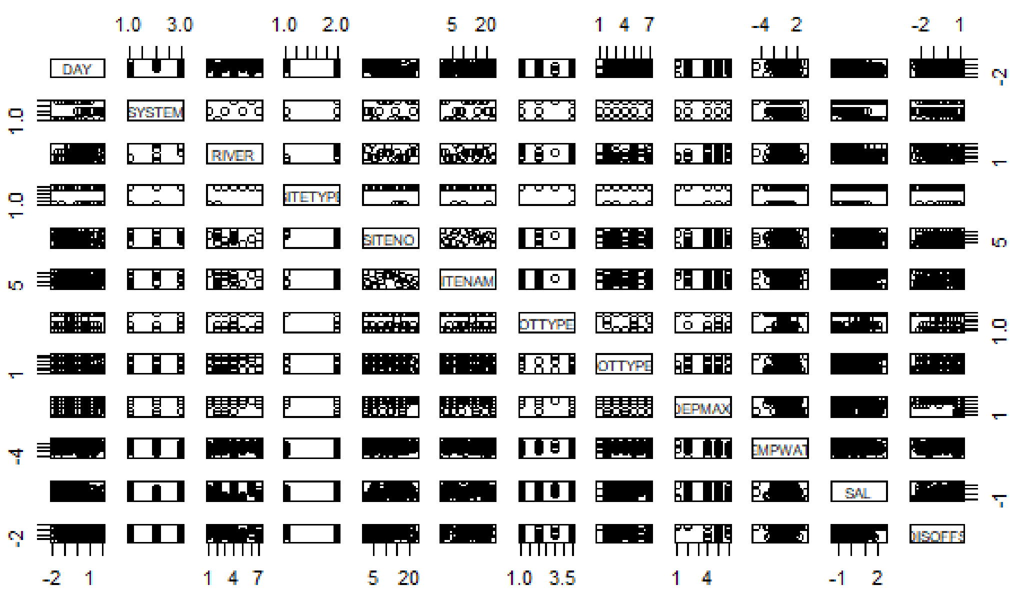
```
new_data2 <- standardize_data(new_data,
  cols = c("YEAR", "MONTH", "SYSTEM", "RIVER", "SITETYPE", "SITENO",
"SITENAME", "BOTTYPE1", "BOTTYPE2", "SampleID"),
  f_col = "FREQ"
)

## YEAR has been changed to factor

## MONTH has been changed to factor

## DAY has been standardized

## SYSTEM has been changed to factor

## RIVER has been changed to factor

## SITETYPE has been changed to factor

## SITENO has been changed to factor

## SITENAME has been changed to factor

## BOTTYPE1 has been changed to factor

## BOTTYPE2 has been changed to factor

## Not sure what to do with DEPMAX, leaving as is

## TEMPWATR has been standardized

## SAL has been standardized

## DISOFFS has been standardized

## SampleID is already a factor

## Data standardized!

head(new_data2, n = 2)
```

YE	MON	DA	SYSTEM	RIVER	SITETYPE	SITE	SITENA	FR	BOTTYP	BOTTYP	DEPM	TEMPWA	SA	DISOF	Sampl

AR	TH	Y				NO	ME	EQ	E1	E2	AX	TR	L	FS	eID
1981	7	-0.23	CHOPTANK	CHOPTANK RIVER	PERMANENT	67	WARWICK CR	0	SAND	TYPE 1 ONLY	2	1.34	-0.06	1.15	1
1981	7	0.28	CHOPTANK	CHOPTANK RIVER	PERMANENT	29	CASTLE HAVEN	0	SAND	TYPE 1 ONLY	3	-0.40	0.39	1.00	3

Check colinearity

```
check_colin(new_data2,
  cols = c("SampleID", "YEAR", "FREQ", "MONTH")
)
```




```
##              DAY SYSTEM RIVER SITETYPE SITENO SITENAME BOTTYPE1 BOTTYPE2
DEPMAX
## DAY         1.00   0.10  0.22    -0.20  -0.48     0.35     0.04    -0.26
0.04
## SYSTEM      0.10   1.00  0.49     0.08   0.01     0.12    -0.26     0.05
0.27
## RIVER       0.22   0.49  1.00     0.47  -0.27     0.05    -0.35    -0.04
0.25
## SITETYPE -0.20   0.08  0.47     1.00  -0.25    -0.32    -0.17     0.18   -
```

```
0.12
## SITENO   -0.48   0.01 -0.27    -0.25   1.00    -0.02    -0.05     0.13    
0.10
## SITENAME  0.35   0.12  0.05    -0.32  -0.02     1.00    -0.05    -0.17   -
0.13
## BOTTYPE1  0.04  -0.26 -0.35    -0.17  -0.05    -0.05     1.00    -0.21   -
0.22
## BOTTYPE2 -0.26   0.05 -0.04     0.18   0.13    -0.17    -0.21     1.00   -
0.11
## DEPMAX    0.04   0.27  0.25    -0.12   0.10    -0.13    -0.22    -0.11    
1.00
## TEMPWATR  0.10  -0.09  0.05     0.23  -0.10    -0.01     0.01     0.07   -
0.11
## SAL      -0.43  -0.10 -0.23    -0.01   0.50     0.00     0.17     0.29   -
0.17
## DISOFFS   0.02  -0.26 -0.34    -0.07   0.03     0.04     0.32     0.12   -
0.43
##          TEMPWATR   SAL DISOFFS
## DAY          0.10 -0.43    0.02
## SYSTEM      -0.09 -0.10   -0.26
## RIVER        0.05 -0.23   -0.34
## SITETYPE     0.23 -0.01   -0.07
## SITENO      -0.10  0.50    0.03
## SITENAME    -0.01  0.00    0.04
## BOTTYPE1     0.01  0.17    0.32
## BOTTYPE2     0.07  0.29    0.12
## DEPMAX      -0.11 -0.17   -0.43
## TEMPWATR     1.00 -0.07    0.08
## SAL         -0.07  1.00    0.31
## DISOFFS      0.08  0.31    1.00
```

Check variance inflation factor - Remove covariates with GVIF > 3, starting with highest value, until all GVIF < 3

```
check_vif(
  model = "FREQ ~ YEAR + SITENAME + BOTTYPE1 + BOTTYPE2 + DEPMAX + TEMPWATR + 
SAL + DISOFFS",
  data = new_data2
)

## # A tibble: 8 x 4
##   Covariate    GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>       <dbl> <dbl>             <dbl>
## 1 TEMPWATR     1.93     1              1.39
## 2 DISOFFS      3.03     1              1.74
## 3 DEPMAX       7.59     4              1.29
## 4 SAL          9.38     1              3.06
## 5 BOTTYPE1    14.3      3              1.56
## 6 BOTTYPE2    38.6      6              1.36
```

```
## 7 YEAR         57.1     34               1.06
## 8 SITENAME   2084.      22               1.19
```

```
## Variance inflation factor too high! Please remove a covariate.
```

```
# remove SITENAME

check_vif(
  model = "FREQ ~ YEAR + BOTTYPE1 + BOTTYPE2 + DEPMAX + TEMPWATR + SAL +
DISOFFS",
  data = new_data2
)
```

```
## # A tibble: 7 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 SAL        1.65     1              1.28
## 2 TEMPWATR   1.65     1              1.28
## 3 DISOFFS    1.76     1              1.32
## 4 DEPMAX     3.58     4              1.17
## 5 BOTTYPE1   4.36     3              1.28
## 6 BOTTYPE2   6.42     6              1.17
## 7 YEAR       9.25    34              1.03
```

```
## Variance inflation factor too high! Please remove a covariate.
```

```
# remove BOTTYPE2

check_vif(
  model = "FREQ ~ YEAR + BOTTYPE1 + DEPMAX + TEMPWATR + SAL + DISOFFS",
  data = new_data2
)
```

```
## # A tibble: 6 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 SAL        1.35     1              1.16
## 2 BOTTYPE1   1.56     3              1.08
## 3 TEMPWATR   1.62     1              1.27
## 4 DISOFFS    1.70     1              1.30
## 5 DEPMAX     3.16     4              1.15
## 6 YEAR       5.36    34              1.03
```

```
## Variance inflation factor too high! Please remove a covariate.
```

```
# remove DEPMAX

check_vif(
  model = "FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS",
  data = new_data2
)
```

```
## # A tibble: 5 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 SAL        1.32     1              1.15
## 2 DISOFFS    1.40     1              1.19
## 3 BOTTYPE1   1.49     3              1.07
## 4 TEMPWATR   1.57     1              1.25
## 5 YEAR       2.47    34              1.01

## Passed!
```

Step 2: Model selection

Compare negative binom, zero-infl neg binom, zero-altered neg binom

Comparison 1: full models

```
compare_models(
  data = new_data2,
  nb_model = "FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS",

  zic_model = "FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS",
  zi_model = "~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS",

  zac_model = "FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS",
  za_model = "~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 29.4

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##    977.3   1148.5   -446.6    893.3      394
##
##
## Dispersion parameter for nbinom2 family (): 0.367
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.90030    0.75908  -2.503 0.012300 *
## YEAR1984       1.70808    0.92570   1.845 0.065012 .
## YEAR1985       1.66811    0.88732   1.880 0.060116 .
```

```
## YEAR1986       -0.05221    1.01805  -0.051 0.959095
## YEAR1987        0.66927    1.18745   0.564 0.573014
## YEAR1988        0.13997    1.03047   0.136 0.891959
## YEAR1989        0.79381    0.91965   0.863 0.388048
## YEAR1990        1.07806    0.91581   1.177 0.239127
## YEAR1991        0.28383    1.04449   0.272 0.785818
## YEAR1992        1.05487    0.91301   1.155 0.247934
## YEAR1993       -0.59534    1.13437  -0.525 0.599710
## YEAR1994        0.06467    1.03190   0.063 0.950026
## YEAR1995       -1.05994    1.15821  -0.915 0.360111
## YEAR1996       -1.72939    1.36985  -1.262 0.206780
## YEAR1997        2.26991    0.92290   2.460 0.013911 *
## YEAR1998       -1.00664    1.13646  -0.886 0.375742
## YEAR1999        0.32444    0.94642   0.343 0.731743
## YEAR2000        0.73307    0.90142   0.813 0.416083
## YEAR2001        0.80835    0.91217   0.886 0.375518
## YEAR2002       -2.36058    1.38277  -1.707 0.087796 .
## YEAR2003        1.03151    0.93172   1.107 0.268248
## YEAR2004       -0.55425    0.99467  -0.557 0.577379
## YEAR2005        0.44693    1.00483   0.445 0.656479
## YEAR2006        1.17978    0.92312   1.278 0.201236
## YEAR2007        1.47182    0.93262   1.578 0.114530
## YEAR2008        0.71169    0.93013   0.765 0.444180
## YEAR2009        1.51296    0.89417   1.692 0.090640 .
## YEAR2010       -1.48144    1.35354  -1.094 0.273737
## YEAR2011        1.00943    0.93885   1.075 0.282295
## YEAR2012        1.31946    0.92112   1.432 0.152015
## YEAR2013        0.71222    0.97600   0.730 0.465555
## YEAR2014        1.06537    0.92978   1.146 0.251865
## YEAR2015       -0.32669    0.98912  -0.330 0.741189
## YEAR2017        0.68768    0.94577   0.727 0.467157
## YEAR2020       -1.30306    1.35010  -0.965 0.334466
## BOTTYPE1MUD    -0.05814    0.64987  -0.089 0.928718
## BOTTYPE1RUBBLE  0.82279    0.93892   0.876 0.380860
## BOTTYPE1SAND    1.03888    0.37299   2.785 0.005348 **
## TEMPWATR       -0.55803    0.15328  -3.641 0.000272 ***
## SAL             0.21414    0.14728   1.454 0.145951
## DISOFFS         0.07169    0.15713   0.456 0.648198
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 0.99

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work
```

```
## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class
"c('simpleWarning', 'warning', 'condition')"
```

Zero-inflated and zero-adjusted did not work. Negative binomial actually seems ok. SAL and DISOFFS were both not statistically significant in the negative binomial. Try dropping each one from the zero models.

Comparison 2: Removing SAL from zero models

```
compare_models(
  data = new_data2,
  models = c("zi", "za"),

  zic_model = "FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + DISOFFS",
  zi_model = "~ YEAR + BOTTYPE1 + TEMPWATR + DISOFFS",

  zac_model = "FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + DISOFFS",
  za_model = "~ YEAR + BOTTYPE1 + TEMPWATR + DISOFFS",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 29.4

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Comparison 3: Removing DISOFFS from zero models

```
compare_models(
  data = new_data2,
  models = c("zi", "za"),

  zic_model = "FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL",
  zi_model = "~ YEAR + BOTTYPE1 + TEMPWATR + SAL",

  zac_model = "FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL",
  za_model = "~ YEAR + BOTTYPE1 + TEMPWATR + SAL",

  r_dat = new_data2$FREQ
) %>%
  try()
```

```
## Percent positive tows: 29.4

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Still not working. Try removing BOTTYPE1

Comparison 4: Removing BOTTYPE1 from zero models

```
compare_models(
  data = new_data2,
  models = c("zi", "za"),

  zic_model = "FREQ ~ YEAR + TEMPWATR",
  zi_model = "~ YEAR + TEMPWATR",

  zac_model = "FREQ ~ YEAR + TEMPWATR",
  za_model = "~ YEAR + TEMPWATR",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 29.4

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Still not working. Try removing TEMPWATR

Comparison 5: Removing TEMPWATR from zero models

```
compare_models(
  data = new_data2,
  models = c("zi", "za"),

  zic_model = "FREQ ~ YEAR",
  zi_model = "~ YEAR",
```

```
  zac_model = "FREQ ~ YEAR",
  za_model = "~ YEAR",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 29.4

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Still not working. Try null binomial model and count model

Comparison 6: Removing YEAR from binomial models

```
compare_models(
  data = new_data2,
  models = c("zi", "za"),

  zic_model = "FREQ ~ YEAR",
  zi_model = "~ 1",

  zac_model = "FREQ ~ YEAR",
  za_model = "~ 1",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 29.4

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR
## Zero inflation:        ~1
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   1000.6   1151.4   -463.3    926.6      399
##
```

```
##
## Dispersion parameter for nbinom2 family (): 0.299
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.099e+00  7.270e-01  -1.511  0.13073
## YEAR1984     1.645e+00  9.407e-01   1.749  0.08031 .
## YEAR1985     2.145e+00  9.014e-01   2.379  0.01736 *
## YEAR1986    -8.001e-02  1.018e+00  -0.079  0.93733
## YEAR1987     2.513e-01  1.157e+00   0.217  0.82808
## YEAR1988     5.311e-06  1.028e+00   0.000  1.00000
## YEAR1989     1.099e+00  9.287e-01   1.183  0.23682
## YEAR1990     1.173e+00  9.257e-01   1.267  0.20523
## YEAR1991     4.055e-01  1.031e+00   0.393  0.69408
## YEAR1992     1.386e+00  9.325e-01   1.487  0.13709
## YEAR1993    -7.732e-01  1.134e+00  -0.682  0.49529
## YEAR1994     5.154e-06  1.028e+00   0.000  1.00000
## YEAR1995    -7.732e-01  1.134e+00  -0.682  0.49528
## YEAR1996    -1.466e+00  1.336e+00  -1.097  0.27249
## YEAR1997     1.981e+00  9.173e-01   2.160  0.03081 *
## YEAR1998    -6.931e-01  1.143e+00  -0.606  0.54431
## YEAR1999     8.363e-01  9.410e-01   0.889  0.37419
## YEAR2000     1.792e+00  9.078e-01   1.974  0.04840 *
## YEAR2001     1.306e+00  9.209e-01   1.419  0.15605
## YEAR2002    -1.466e+00  1.336e+00  -1.097  0.27249
## YEAR2003     1.266e+00  9.535e-01   1.327  0.18439
## YEAR2004     1.431e-01  9.928e-01   0.144  0.88539
## YEAR2005     1.431e-01  9.928e-01   0.144  0.88538
## YEAR2006     1.712e+00  9.095e-01   1.882  0.05983 .
## YEAR2007     1.478e+00  9.155e-01   1.614  0.10642
## YEAR2008     8.363e-01  9.410e-01   0.889  0.37419
## YEAR2009     2.645e+00  8.955e-01   2.953  0.00315 **
## YEAR2010    -1.466e+00  1.336e+00  -1.097  0.27249
## YEAR2011     8.363e-01  9.410e-01   0.889  0.37419
## YEAR2012     1.478e+00  9.155e-01   1.614  0.10642
## YEAR2013     3.254e-01  9.758e-01   0.334  0.73876
## YEAR2014     8.363e-01  9.410e-01   0.889  0.37419
## YEAR2015     1.431e-01  9.928e-01   0.144  0.88539
## YEAR2017     8.363e-01  9.410e-01   0.889  0.37419
## YEAR2020    -1.466e+00  1.336e+00  -1.097  0.27249
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -17.63    3265.01  -0.005    0.996

## Dispersal: 1.072

## Zero-altered negative binomial
```

```
## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Zero-inflated model worked, zero-adjusted did not.

Comparison 7: Removing YEAR from count models

```
compare_models(
  data = new_data2,
  models = c("zi", "za"),

  zic_model = "FREQ ~ 1",
  zi_model = "~ YEAR",

  zac_model = "FREQ ~ 1",
  za_model = "~ YEAR",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 29.4

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ 1
## Zero inflation:           ~YEAR
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##    997.4   1148.3   -461.7    923.4      399
##
##
## Dispersion parameter for truncated_nbinom2 family (): 1.9e-09
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -18.22    5214.94  -0.003    0.997
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  6.931e-01  6.124e-01   1.132   0.2577
## YEAR1984    -5.108e-01  8.612e-01  -0.593   0.5531
```

```
## YEAR1985    -1.163e+00  8.367e-01  -1.390   0.1645
## YEAR1986     5.108e-01  8.991e-01   0.568   0.5699
## YEAR1987     2.231e-01  1.037e+00   0.215   0.8296
## YEAR1988     9.163e-01  9.874e-01   0.928   0.3534
## YEAR1989    -5.390e-01  8.274e-01  -0.651   0.5147
## YEAR1990    -8.473e-01  8.274e-01  -1.024   0.3058
## YEAR1991    -2.877e-01  8.898e-01  -0.323   0.7464
## YEAR1992     4.611e-07  8.660e-01   0.000   1.0000
## YEAR1993     1.792e+00  1.208e+00   1.484   0.1379
## YEAR1994     1.705e+00  1.211e+00   1.408   0.1591
## YEAR1995     1.012e+00  9.828e-01   1.029   0.3033
## YEAR1996     1.792e+00  1.208e+00   1.484   0.1379
## YEAR1997    -1.792e+00  9.052e-01  -1.979   0.0478 *
## YEAR1998     9.163e-01  9.874e-01   0.928   0.3534
## YEAR1999    -5.390e-01  8.274e-01  -0.651   0.5147
## YEAR2000    -8.473e-01  8.274e-01  -1.024   0.3058
## YEAR2001    -2.231e-01  8.367e-01  -0.267   0.7897
## YEAR2002     1.792e+00  1.208e+00   1.484   0.1379
## YEAR2003     2.877e-01  9.129e-01   0.315   0.7527
## YEAR2004     1.012e+00  9.828e-01   1.029   0.3033
## YEAR2005     5.108e-01  8.991e-01   0.568   0.5699
## YEAR2006    -5.390e-01  8.274e-01  -0.651   0.5147
## YEAR2007     5.108e-01  8.991e-01   0.568   0.5699
## YEAR2008    -5.390e-01  8.274e-01  -0.651   0.5147
## YEAR2009    -2.231e-01  8.367e-01  -0.267   0.7897
## YEAR2010     1.792e+00  1.208e+00   1.484   0.1379
## YEAR2011    -5.390e-01  8.274e-01  -0.651   0.5147
## YEAR2012     1.012e+00  9.828e-01   1.029   0.3033
## YEAR2013     1.178e-01  8.580e-01   0.137   0.8908
## YEAR2014     5.108e-01  8.991e-01   0.568   0.5699
## YEAR2015     1.792e+00  1.208e+00   1.484   0.1379
## YEAR2017     1.012e+00  9.828e-01   1.029   0.3033
## YEAR2020     1.792e+00  1.208e+00   1.484   0.1379
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.913

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Zero-altered model worked, zero-inflated did not.

Model selection for negative binomial model

```
model <- glmmTMB::glmmTMB(FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS,
  data = new_data2,
  family = nbinom2
)

null <- glmmTMB::glmmTMB(FREQ ~ YEAR,
```

```
  data = new_data2,
  family = nbinom2
)

select_model(model = model, null_model = null)

## Forward model selection:

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ YEAR
##
## Final Model:
## FREQ ~ YEAR + TEMPWATR + BOTTYPE1 + SAL
##
##
##          Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                                400   926.5728 998.5728
## 2 + TEMPWATR  1 17.663477       399   908.9094 982.9094
## 3 + BOTTYPE1  3 12.365551       396   896.5438 976.5438
## 4      + SAL  1  3.085092       395   893.4587 975.4587

## Bidirectional model selection:

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS
##
## Final Model:
## FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL
##
##
##        Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                              394   893.2504 977.2504
## 2 - DISOFFS  1 0.2082725       395   893.4587 975.4587

## Backward model selection:

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS
##
## Final Model:
## FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL
##
```

```
##
##        Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                             394   893.2504 977.2504
## 2 - DISOFFS  1 0.2082725      395   893.4587 975.4587
```

Model selection for zero-inflated model

```
model <- glmmTMB::glmmTMB(FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS,
  ziformula = ~1,
  data = new_data2,
  family = nbinom2
)

null <- glmmTMB::glmmTMB(FREQ ~ YEAR,
  ziformula = ~1,
  data = new_data2,
  family = nbinom2
)

select_model(model = model, null_model = null)
```

```
## Forward model selection:
```

```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ YEAR
##
## Final Model:
## FREQ ~ YEAR + TEMPWATR + BOTTYPE1 + SAL
##
##
##         Step Df  Deviance Resid. Df Resid. Dev       AIC
## 1                              399   926.5728 1000.5728
## 2 + TEMPWATR  1 17.663477       398   908.9094  984.9094
## 3 + BOTTYPE1  3 12.365551       395   896.5438  978.5438
## 4      + SAL  1  3.085092       394   893.4587  977.4587
```

```
## Bidirectional model selection:
```

```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS
##
## Final Model:
## FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL
##
##
##        Step Df  Deviance Resid. Df Resid. Dev      AIC
```

```
## 1                               393   893.2504 979.2504
## 2 - DISOFFS  1 0.2082725        394   893.4587 977.4587

## Backward model selection:

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS
##
## Final Model:
## FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL
##
##
##        Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                               393   893.2504 979.2504
## 2 - DISOFFS  1 0.2082725        394   893.4587 977.4587

# forward selection only
forward <- MASS::stepAIC(model,
  direction = "forward",
  scope = list(upper = "FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS"),
  trace = FALSE
)

print(forward$anova)

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS
##
## Final Model:
## FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS
##
##
##   Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                         393   893.2504 979.2504
```

Model selection for zero-adjusted model

```
model <- glmmTMB::glmmTMB(FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS,
  ziformula = ~YEAR,
  data = new_data2,
  family = truncated_nbinom2(link = "log")
)

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

# full model does not converge

null <- glmmTMB::glmmTMB(FREQ ~ 1,
  ziformula = ~YEAR,
  data = new_data2,
  family = truncated_nbinom2(link = "log")
)
# forward selection only
forward <- MASS::stepAIC(null,
  direction = "forward",
  scope = list(upper = "FREQ ~ YEAR + BOTTYPE1 + TEMPWATR + SAL + DISOFFS"),
  trace = FALSE
)

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

print(forward$anova)

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
```

```
## FREQ ~ 1
##
## Final Model:
## FREQ ~ TEMPWATR + BOTTYPE1 + DISOFFS
##
##
##          Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                                399   923.4378 997.4378
## 2 + TEMPWATR  1 13.412085        398   910.0257 986.0257
## 3 + BOTTYPE1  3 10.361144        395   899.6646 981.6646
## 4   + DISOFFS  1  2.909223        394   896.7553 980.7553
```

Final comparison

```
compare_models(
  data = new_data2,
  nb_model = "FREQ ~ YEAR + TEMPWATR + BOTTYPE1 + SAL",

  zic_model = "FREQ ~ YEAR + TEMPWATR + BOTTYPE1 + SAL",
  zi_model = "~ 1",

  zac_model = "FREQ ~ TEMPWATR + BOTTYPE1 + DISOFFS",
  za_model = "~ YEAR",

  r_dat = new_data2$FREQ
) %>%
  try()

## Percent positive tows: 29.4

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + TEMPWATR + BOTTYPE1 + SAL
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##    975.5   1142.6   -446.7    893.5      395
##
##
## Dispersion parameter for nbinom2 family (): 0.367
##
## Conditional model:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.92162    0.75826  -2.534 0.011268 *
## YEAR1984        1.75107    0.92143   1.900 0.057384 .
## YEAR1985        1.64019    0.88607   1.851 0.064158 .
## YEAR1986       -0.06209    1.01799  -0.061 0.951363
## YEAR1987        0.67250    1.18851   0.566 0.571508
## YEAR1988        0.15008    1.03000   0.146 0.884154
```

```
## YEAR1989        0.79617    0.91931   0.866 0.386466
## YEAR1990        1.07805    0.91678   1.176 0.239631
## YEAR1991        0.28015    1.04474   0.268 0.788584
## YEAR1992        1.05893    0.91378   1.159 0.246520
## YEAR1993       -0.59667    1.13462  -0.526 0.598978
## YEAR1994        0.08813    1.03088   0.085 0.931873
## YEAR1995       -1.08168    1.15773  -0.934 0.350142
## YEAR1996       -1.75003    1.36799  -1.279 0.200800
## YEAR1997        2.21704    0.91424   2.425 0.015308 *
## YEAR1998       -1.03118    1.13555  -0.908 0.363832
## YEAR1999        0.29468    0.94462   0.312 0.755077
## YEAR2000        0.71846    0.90068   0.798 0.425055
## YEAR2001        0.77893    0.90897   0.857 0.391478
## YEAR2002       -2.37254    1.38180  -1.717 0.085980 .
## YEAR2003        1.00798    0.93130   1.082 0.279103
## YEAR2004       -0.54508    0.99418  -0.548 0.583506
## YEAR2005        0.43231    1.00386   0.431 0.666725
## YEAR2006        1.14729    0.92047   1.246 0.212615
## YEAR2007        1.37193    0.90666   1.513 0.130240
## YEAR2008        0.66906    0.92563   0.723 0.469790
## YEAR2009        1.49972    0.89372   1.678 0.093335 .
## YEAR2010       -1.53950    1.34954  -1.141 0.253970
## YEAR2011        1.01578    0.93990   1.081 0.279817
## YEAR2012        1.25229    0.90995   1.376 0.168752
## YEAR2013        0.66542    0.96928   0.687 0.492393
## YEAR2014        1.03401    0.92736   1.115 0.264849
## YEAR2015       -0.35558    0.98602  -0.361 0.718386
## YEAR2017        0.65913    0.94283   0.699 0.484494
## YEAR2020       -1.37966    1.33996  -1.030 0.303184
## TEMPWATR       -0.55573    0.15300  -3.632 0.000281 ***
## BOTTYPE1MUD    -0.02293    0.64570  -0.036 0.971670
## BOTTYPE1RUBBLE  0.84881    0.93884   0.904 0.365942
## BOTTYPE1SAND    1.09251    0.35341   3.091 0.001993 **
## SAL             0.23886    0.13720   1.741 0.081702 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 0.992

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ YEAR + TEMPWATR + BOTTYPE1 + SAL
## Zero inflation:        ~1
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##    977.5   1148.7   -446.7    893.5      394
##
```

```
##
## Dispersion parameter for nbinom2 family (): 0.367
##
## Conditional model:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.92161    0.75825  -2.534 0.011269 *
## YEAR1984        1.75109    0.92144   1.900 0.057382 .
## YEAR1985        1.64021    0.88608   1.851 0.064155 .
## YEAR1986       -0.06208    1.01799  -0.061 0.951373
## YEAR1987        0.67250    1.18851   0.566 0.571506
## YEAR1988        0.15004    1.03000   0.146 0.884182
## YEAR1989        0.79613    0.91931   0.866 0.386483
## YEAR1990        1.07804    0.91678   1.176 0.239635
## YEAR1991        0.28014    1.04474   0.268 0.788588
## YEAR1992        1.05896    0.91378   1.159 0.246509
## YEAR1993       -0.59666    1.13462  -0.526 0.598981
## YEAR1994        0.08815    1.03088   0.086 0.931855
## YEAR1995       -1.08168    1.15773  -0.934 0.350142
## YEAR1996       -1.75004    1.36799  -1.279 0.200799
## YEAR1997        2.21705    0.91425   2.425 0.015308 *
## YEAR1998       -1.03119    1.13555  -0.908 0.363828
## YEAR1999        0.29471    0.94462   0.312 0.755049
## YEAR2000        0.71848    0.90068   0.798 0.425043
## YEAR2001        0.77894    0.90897   0.857 0.391472
## YEAR2002       -2.37256    1.38180  -1.717 0.085979 .
## YEAR2003        1.00797    0.93130   1.082 0.279106
## YEAR2004       -0.54504    0.99418  -0.548 0.583530
## YEAR2005        0.43233    1.00386   0.431 0.666709
## YEAR2006        1.14731    0.92048   1.246 0.212605
## YEAR2007        1.37190    0.90666   1.513 0.130247
## YEAR2008        0.66907    0.92563   0.723 0.469785
## YEAR2009        1.49973    0.89372   1.678 0.093333 .
## YEAR2010       -1.53949    1.34953  -1.141 0.253971
## YEAR2011        1.01576    0.93990   1.081 0.279825
## YEAR2012        1.25229    0.90995   1.376 0.168752
## YEAR2013        0.66537    0.96928   0.686 0.492423
## YEAR2014        1.03402    0.92736   1.115 0.264846
## YEAR2015       -0.35560    0.98602  -0.361 0.718369
## YEAR2017        0.65912    0.94283   0.699 0.484501
## YEAR2020       -1.37965    1.33996  -1.030 0.303188
## TEMPWATR       -0.55572    0.15300  -3.632 0.000281 ***
## BOTTYPE1MUD    -0.02296    0.64570  -0.036 0.971638
## BOTTYPE1RUBBLE  0.84881    0.93884   0.904 0.365942
## BOTTYPE1SAND    1.09249    0.35341   3.091 0.001993 **
## SAL             0.23886    0.13720   1.741 0.081702 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
```

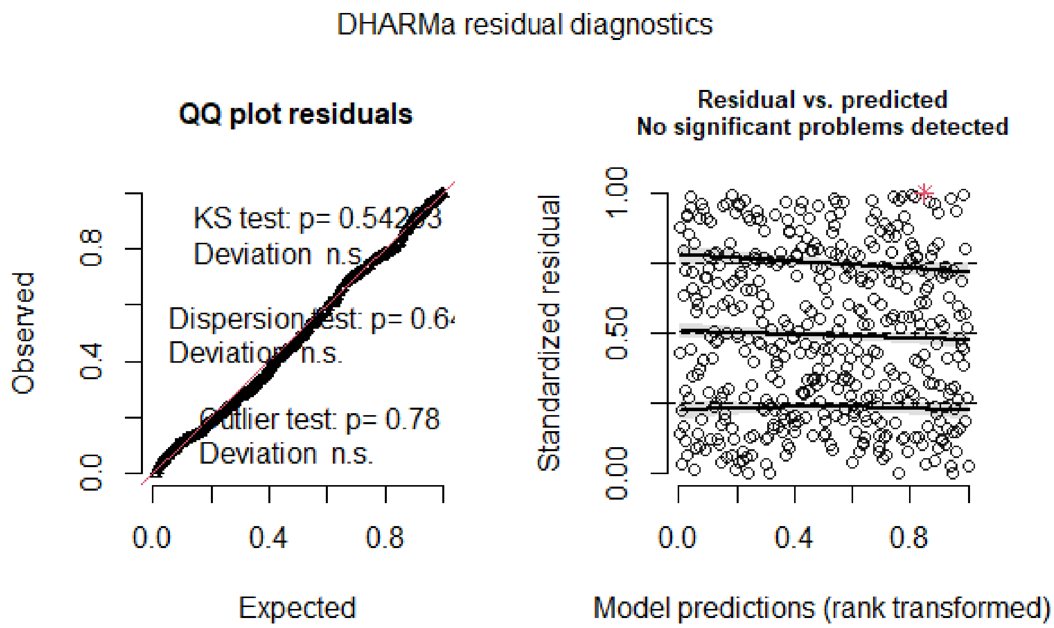
```
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -16.35    4784.38  -0.003    0.997

## Dispersal: 0.994

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ TEMPWATR + BOTTYPE1 + DISOFFS
## Zero inflation:        ~YEAR
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##    980.8   1152.0   -448.4    896.8      394
##
##
## Dispersion parameter for truncated_nbinom2 family (): 8.61e-09
##
## Conditional model:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -18.6464  4355.6116  -0.004 0.996584
## TEMPWATR         -0.6674     0.1941  -3.439 0.000585 ***
## BOTTYPE1MUD      -0.7363     1.1168  -0.659 0.509723
## BOTTYPE1RUBBLE    0.9564     1.3376   0.715 0.474570
## BOTTYPE1SAND      1.8897     0.6447   2.931 0.003378 **
## DISOFFS          -0.4275     0.2624  -1.629 0.103270
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  6.931e-01  6.124e-01   1.132   0.2577
## YEAR1984    -5.108e-01  8.612e-01  -0.593   0.5531
## YEAR1985    -1.163e+00  8.367e-01  -1.390   0.1645
## YEAR1986     5.108e-01  8.991e-01   0.568   0.5699
## YEAR1987     2.231e-01  1.037e+00   0.215   0.8296
## YEAR1988     9.163e-01  9.874e-01   0.928   0.3534
## YEAR1989    -5.390e-01  8.274e-01  -0.651   0.5147
## YEAR1990    -8.473e-01  8.274e-01  -1.024   0.3058
## YEAR1991    -2.877e-01  8.898e-01  -0.323   0.7464
## YEAR1992    -1.334e-06  8.660e-01   0.000   1.0000
## YEAR1993     1.792e+00  1.208e+00   1.484   0.1379
## YEAR1994     1.705e+00  1.211e+00   1.408   0.1591
## YEAR1995     1.012e+00  9.828e-01   1.029   0.3033
## YEAR1996     1.792e+00  1.208e+00   1.484   0.1379
## YEAR1997    -1.792e+00  9.052e-01  -1.979   0.0478 *
## YEAR1998     9.163e-01  9.874e-01   0.928   0.3534
## YEAR1999    -5.390e-01  8.274e-01  -0.651   0.5147
## YEAR2000    -8.473e-01  8.274e-01  -1.024   0.3058
```

```
## YEAR2001    -2.231e-01  8.367e-01  -0.267   0.7897
## YEAR2002     1.792e+00  1.208e+00   1.484   0.1379
## YEAR2003     2.877e-01  9.129e-01   0.315   0.7527
## YEAR2004     1.012e+00  9.828e-01   1.029   0.3033
## YEAR2005     5.108e-01  8.991e-01   0.568   0.5699
## YEAR2006    -5.390e-01  8.274e-01  -0.651   0.5147
## YEAR2007     5.108e-01  8.991e-01   0.568   0.5699
## YEAR2008    -5.390e-01  8.274e-01  -0.651   0.5147
## YEAR2009    -2.231e-01  8.367e-01  -0.267   0.7897
## YEAR2010     1.792e+00  1.208e+00   1.484   0.1379
## YEAR2011    -5.390e-01  8.274e-01  -0.651   0.5147
## YEAR2012     1.012e+00  9.828e-01   1.029   0.3033
## YEAR2013     1.178e-01  8.580e-01   0.137   0.8908
## YEAR2014     5.108e-01  8.991e-01   0.568   0.5699
## YEAR2015     1.792e+00  1.208e+00   1.484   0.1379
## YEAR2017     1.012e+00  9.828e-01   1.029   0.3033
## YEAR2020     1.792e+00  1.208e+00   1.484   0.1379
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.495

##            dAIC df
## neg_binom  0.0 41
## zero_infl  2.0 42
## zero_adj   5.3 42
```

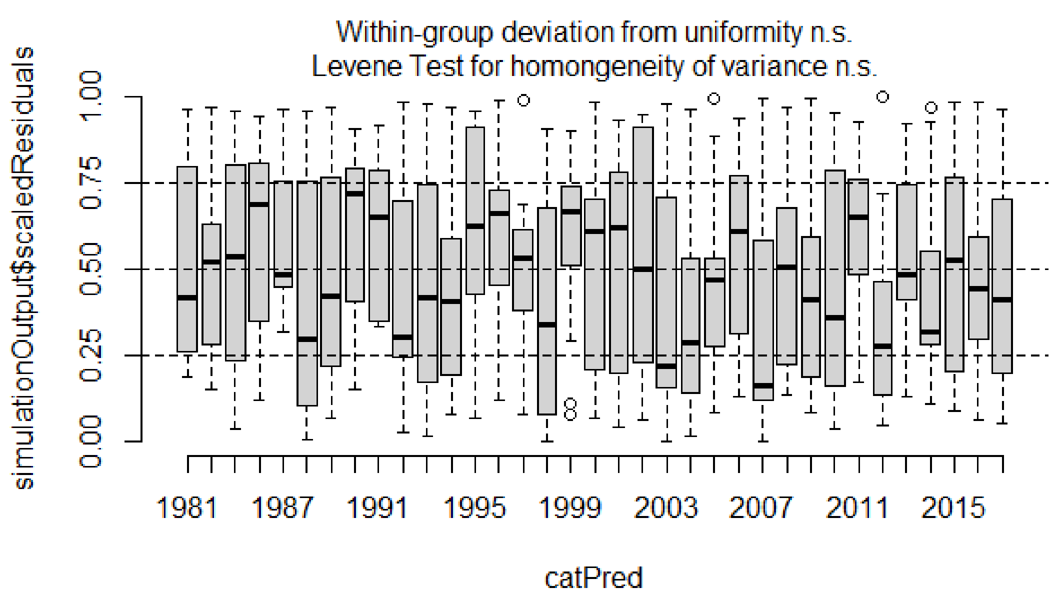
Negative binomial is the winner!

Step 3: DHARMA simulations

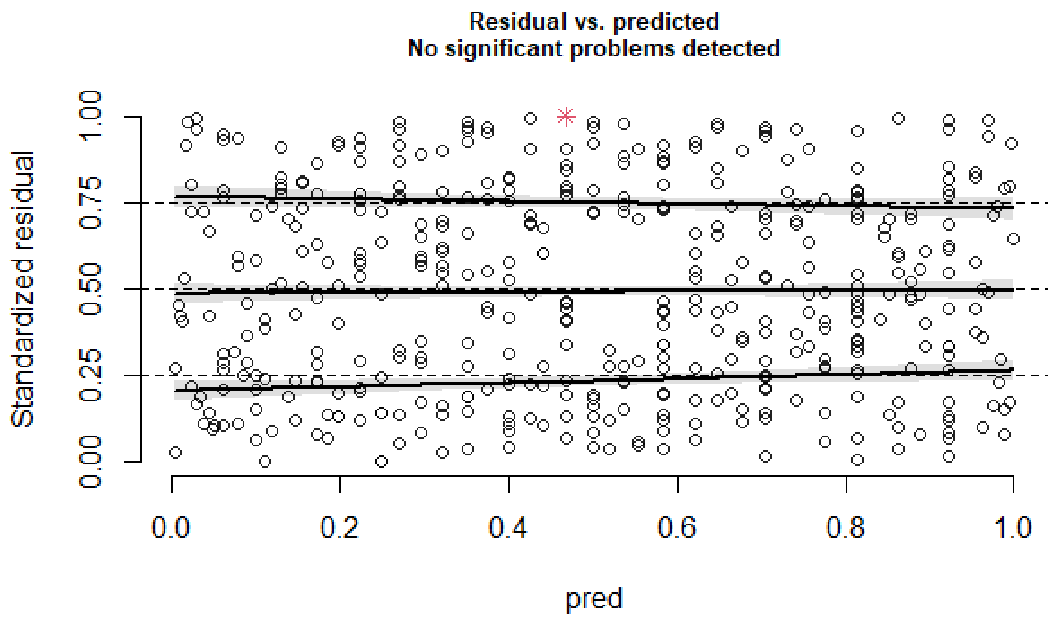
```
model <- glmmTMB(FREQ ~ YEAR + TEMPWATR + BOTTYPE1 + SAL,
  data = new_data2,
  family = nbinom2
)

sim <- simulateResiduals(model)
plot(sim)
```



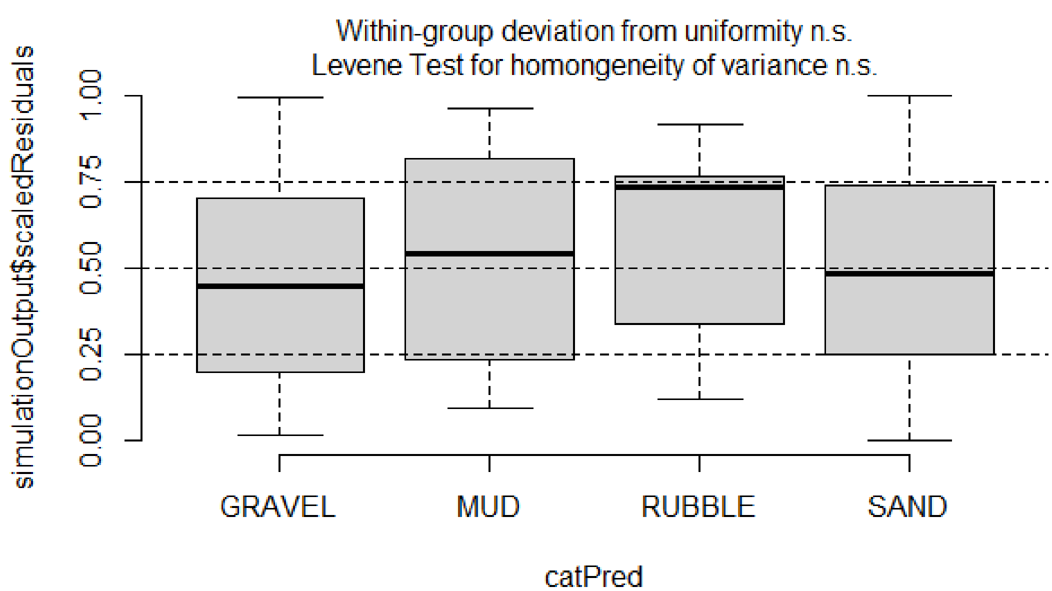

```r
# Compare residuals vs. factors
plotResiduals(sim, form = new_data2$YEAR)
```



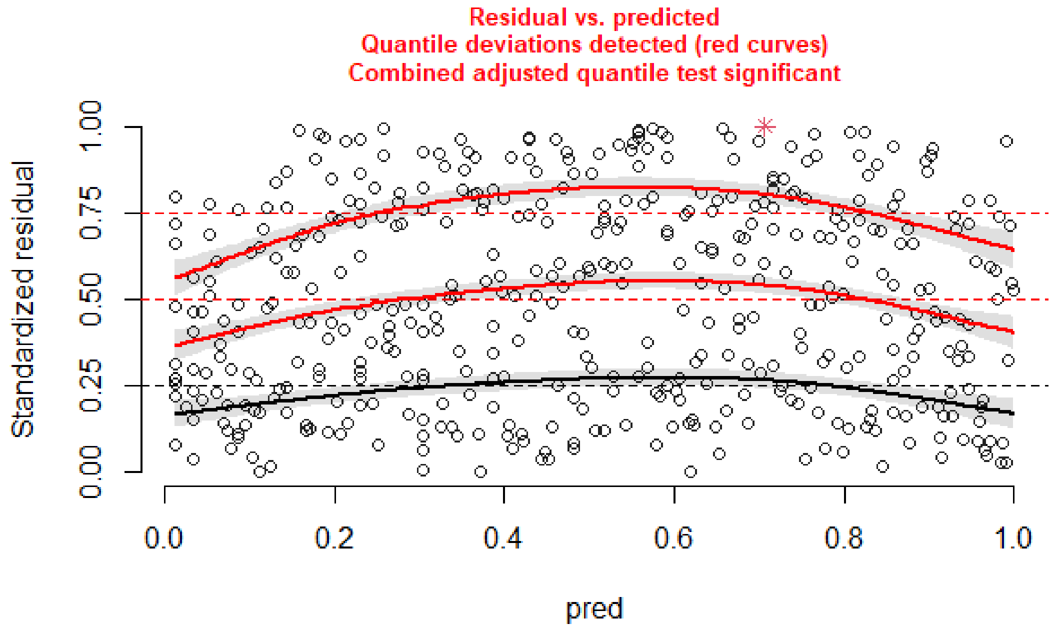

```r
plotResiduals(sim, form = new_data2$TEMPWATR)
```




```
plotResiduals(sim, form = new_data2$BOTTYPE1)
```

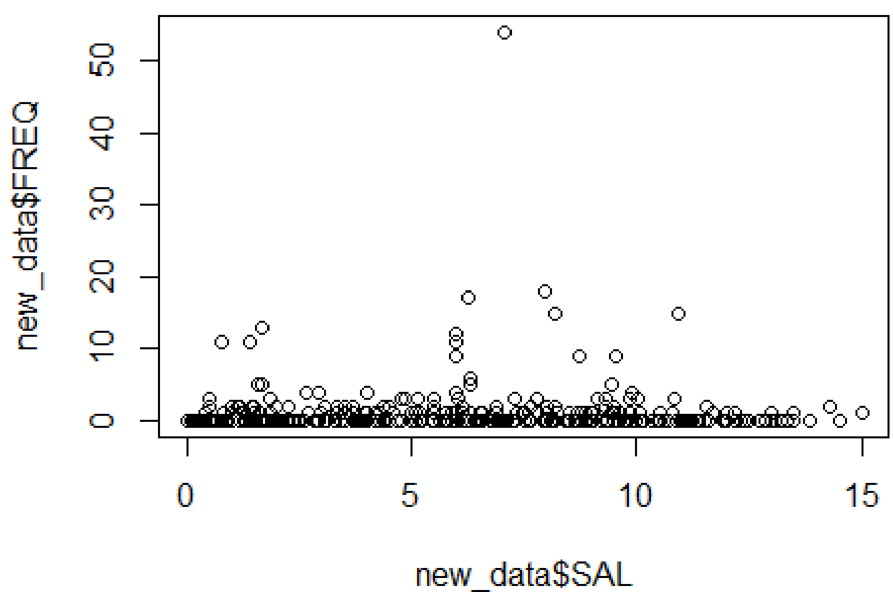



```
plotResiduals(sim, form = new_data2$SAL)
```



It seems like the problem could be because there are few observations at the highest salinities

```
plot(new_data$SAL, new_data$FREQ)
```

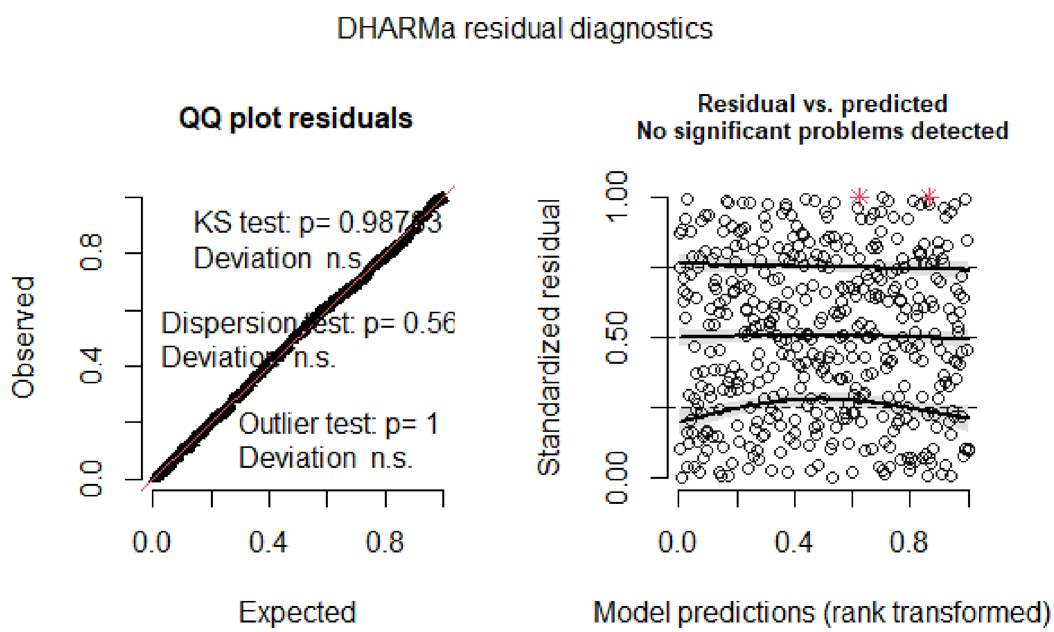


There’s a problem with the salinity residuals, so try removing SAL. [Should probably use a GAM??]

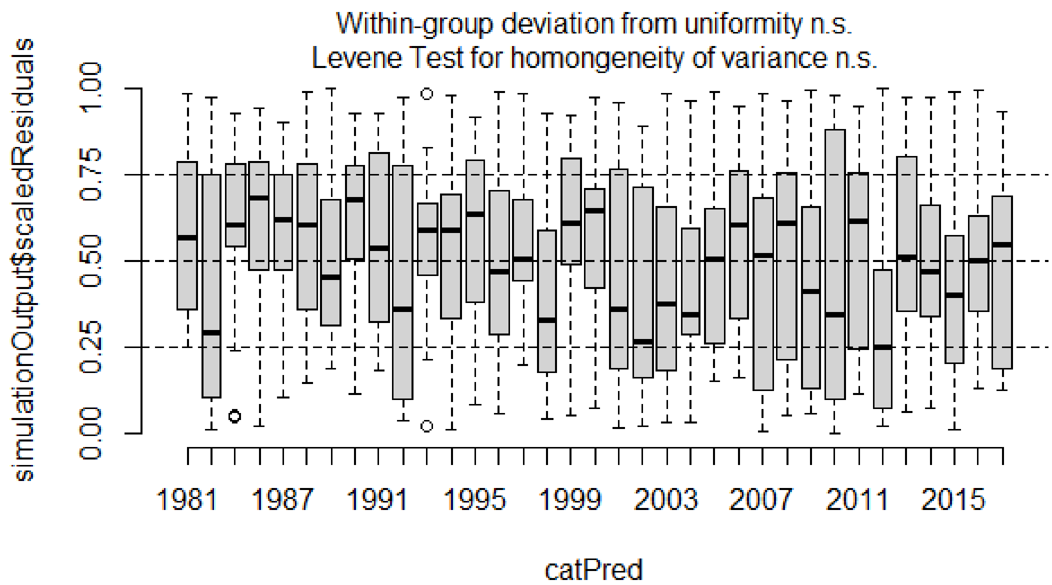
```
model2 <- glmmTMB(FREQ ~ YEAR + TEMPWATR + BOTTYPE1,
  data = new_data2,
  family = nbinom2
)

AIC(model, model2)
```

	df	AIC
model	41	975.46
model2	40	976.54

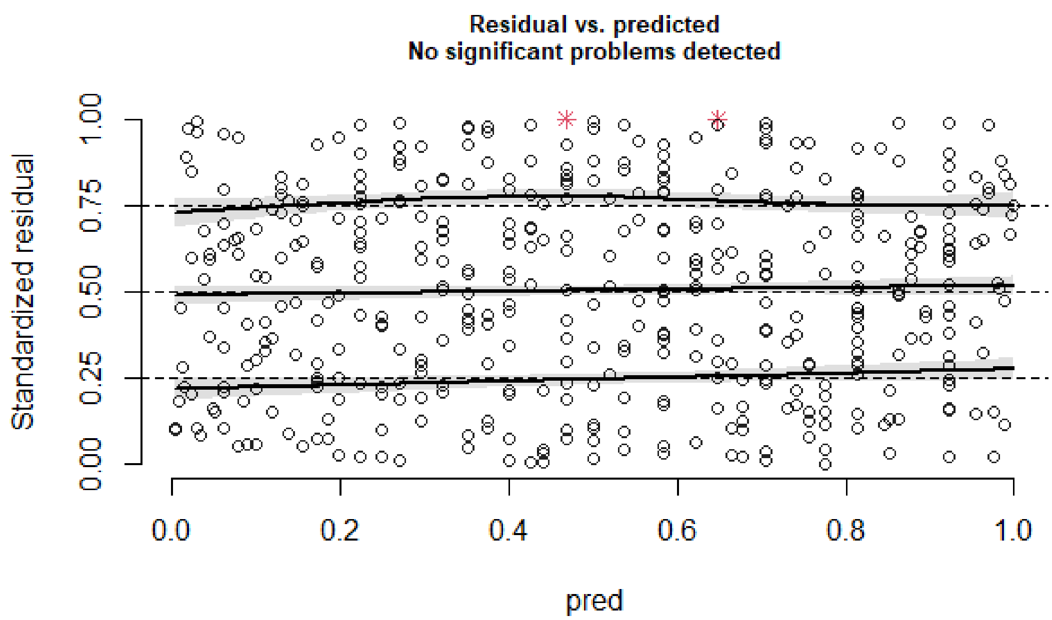
```
# dropping salinity has a negligible effect on AIC

sim <- simulateResiduals(model2)
plot(sim)
```



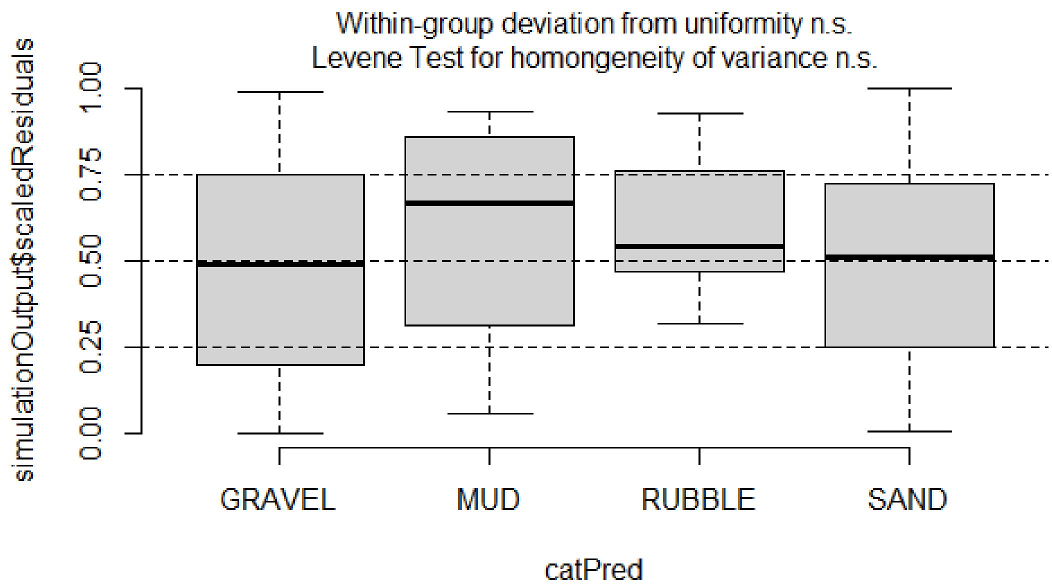

```
# Compare residuals vs. factors
plotResiduals(sim, form = new_data2$YEAR)
```




```
plotResiduals(sim, form = new_data2$TEMPWATR)
```




```
plotResiduals(sim, form = new_data2$BOTTYPE1)
```



Step 4: Plot predictions

Add 2021 data

```
indata <- readxl::read_excel(here::here("data/2021 Data/2021_bluefish_data_MD_April_2022.xlsx"),
  sheet = "MDChesBay seine survey",
  na = "."
)
head(indata, n = 2)
```

STATE	PROGRAM	GEAR	YEAR	MONTH	DAY	ROUND	SYSTEM	RIVER	SITETYPE	SITENO	SITENAME	HAULNO	ITISCODE	SPECNAME	SPECCNT	SIZEMIN	SIZEMAX	BOTTYPE1	BOTTYPE2	DEPMAX	TEMPWATR	SAL	DISOFFS	BF_SUBSET_
MD	CHESBAYS	SEINE	1981	7	13	I	CHOPTANK	CHOPTANK	PERMANENN	67	WARWICKC	1	168559	BLUEFISH	0	NA	NA	SAND	TYPE1ONL	2	29.9	5.1	33	SUBSET

	EINESURVEY							RIVER	T		R								Y					
MD	CHESBAYSEINESURVEY	SEINE	1981	7	13	I	CHOPTANK	CHOPTANKRIVER	PERMANENT	67	WARWICKCR	2	168559	BLUEFISH	0	NA	NA	SAND	TYPE1ONLY	2	29.9	5.1	33	SUBSET

```
# modify data formatting
indata <- indata %>%
  dplyr::rename(FREQ = SPECCNT) %>%
  dplyr::mutate(
    SampleID = as.factor(dplyr::row_number()),
    HAULNO = as.factor(HAULNO),
    DEPMAX = as.factor(DEPMAX)
  ) %>%
  dplyr::select(-SIZEMIN, -SIZEMAX) %>% # too much missing data
  dplyr::select(-c(STATE, PROGRAM, GEAR, ITISCODE, SPECNAME, BF_SUBSET_)) #
same data for all entries

# remove UNKNOWN bottom types
length1 <- nrow(indata)
new_data <- indata %>%
  dplyr::filter(
    BOTTYPE1 != "UNKNOWN",
    BOTTYPE2 != "UNKNOWN"
  )
```

```
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")

## 10 rows removed

# remove HAULNO2
length1 <- nrow(new_data)
new_data <- new_data %>%
  dplyr::filter(HAULNO != 2) %>%
  dplyr::select(-HAULNO) # remove because there is only 1 value now
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")

## 1500 rows removed

new_data <- remove_missing(new_data,
  f_col = "FREQ",
  t_col = "YEAR"
)

## Data before removing missing values:

##       YEAR          MONTH            DAY            ROUND
##  Min.   :1981   Min.   :7.000   Min.   : 2.00   Length:1591
##  1st Qu.:1991   1st Qu.:7.000   1st Qu.: 8.00   Class :character
##  Median :2001   Median :8.000   Median :11.00   Mode  :character
##  Mean   :2001   Mean   :7.992   Mean   :11.65
##  3rd Qu.:2011   3rd Qu.:9.000   3rd Qu.:15.00
##  Max.   :2021   Max.   :9.000   Max.   :31.00
##
##     SYSTEM             RIVER             SITETYPE             SITENO
##  Length:1591        Length:1591        Length:1591        Min.   :  4.00
##  Class :character   Class :character   Class :character   1st Qu.: 11.00
##  Mode  :character   Mode  :character   Mode  :character   Median : 52.00
##                                                           Mean   : 58.04
##                                                           3rd Qu.: 88.00
##                                                           Max.   :172.00
##
##    SITENAME              FREQ           BOTTYPE1           BOTTYPE2
##  Length:1591        Min.   : 0.0000   Length:1591        Length:1591
##  Class :character   1st Qu.: 0.0000   Class :character   Class :character
##  Mode  :character   Median : 0.0000   Mode  :character   Mode  :character
##                     Mean   : 0.5035
##                     3rd Qu.: 0.0000
##                     Max.   :54.0000
##
##   DEPMAX       TEMPWATR          SAL            DISOFFS          SampleID
##  1    : 11   Min.   :17.50   Min.   : 0.000   Min.   : 6.0   1       :   1
##  2    :122   1st Qu.:25.30   1st Qu.: 2.400   1st Qu.:18.0   3       :   1
##  2.5  :  1   Median :27.00   Median : 6.100   Median :30.0   5       :   1
##  3    :405   Mean   :26.67   Mean   : 6.246   Mean   :24.9   7       :   1
```

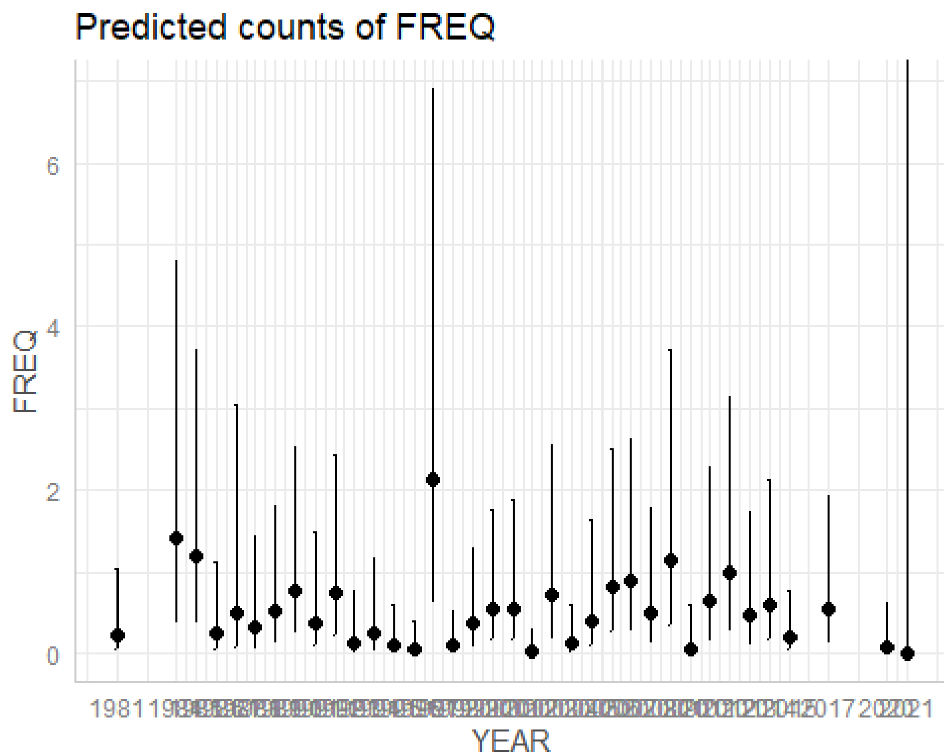
```
##  4   :978   3rd Qu.:28.20   3rd Qu.: 9.600   3rd Qu.:30.0   9      :   1
##  5   : 31   Max.   :33.50   Max.   :20.000   Max.   :33.0   11     :   1
##  NA's: 43   NA's   :23      NA's   :62       NA's   :2      (Other):1585

## Data after removing missing values:

##       YEAR          MONTH            DAY           ROUND
##  Min.   :1981   Min.   :7.000   Min.   : 2.00   Length:1482
##  1st Qu.:1992   1st Qu.:7.000   1st Qu.: 8.00   Class :character
##  Median :2002   Median :8.000   Median :11.00   Mode  :character
##  Mean   :2002   Mean   :8.003   Mean   :11.46
##  3rd Qu.:2012   3rd Qu.:9.000   3rd Qu.:15.00
##  Max.   :2021   Max.   :9.000   Max.   :31.00
##
##     SYSTEM             RIVER             SITETYPE             SITENO
##  Length:1482        Length:1482        Length:1482        Min.   :  4.00
##  Class :character   Class :character   Class :character   1st Qu.: 11.00
##  Mode  :character   Mode  :character   Mode  :character   Median : 52.00
##                                                           Mean   : 59.43
##                                                           3rd Qu.: 88.00
##                                                           Max.   :172.00
##
##    SITENAME              FREQ           BOTTYPE1           BOTTYPE2
##  Length:1482        Min.   : 0.0000   Length:1482        Length:1482
##  Class :character   1st Qu.: 0.0000   Class :character   Class :character
##  Mode  :character   Median : 0.0000   Mode  :character   Mode  :character
##                     Mean   : 0.4366
##                     3rd Qu.: 0.0000
##                     Max.   :54.0000
##
##  DEPMAX       TEMPWATR          SAL            DISOFFS          SampleID
##  1  : 10   Min.   :17.50   Min.   : 0.000   Min.   : 6.00   1      :   1
##  2  :116   1st Qu.:25.20   1st Qu.: 2.500   1st Qu.:18.00   3      :   1
##  2.5:  1   Median :26.90   Median : 6.100   Median :30.00   5      :   1
##  3  :386   Mean   :26.62   Mean   : 6.275   Mean   :24.74   7      :   1
##  4  :940   3rd Qu.:28.20   3rd Qu.: 9.690   3rd Qu.:30.00   9      :   1
##  5  : 29   Max.   :33.50   Max.   :20.000   Max.   :33.00   11     :   1
##                                                             (Other):1476

## 109 datapoints removed (6.9% of data removed)

## All years had positive catch

# dropping August and September
new_data <- new_data %>%
  dplyr::filter(MONTH == 7)

# Fix temperatures <5
new_data <- new_data %>%
  dplyr::mutate(TEMPWATR = ifelse(TEMPWATR < 5, TEMPWATR * 10, TEMPWATR))
```

```r
new_data <- new_data %>%
  dplyr::select(-ROUND)

new_data <- new_data %>%
  dplyr::filter(!YEAR %in% c(2016, 2018, 2019))

new_data2 <- standardize_data(new_data,
  cols = c("YEAR", "MONTH", "SYSTEM", "RIVER", "SITETYPE", "SITENO",
"SITENAME", "BOTTYPE1", "BOTTYPE2", "SampleID"),
  f_col = "FREQ"
)

## YEAR has been changed to factor

## MONTH has been changed to factor

## DAY has been standardized

## SYSTEM has been changed to factor

## RIVER has been changed to factor

## SITETYPE has been changed to factor

## SITENO has been changed to factor

## SITENAME has been changed to factor

## BOTTYPE1 has been changed to factor

## BOTTYPE2 has been changed to factor

## Not sure what to do with DEPMAX, leaving as is

## TEMPWATR has been standardized

## SAL has been standardized

## DISOFFS has been standardized

## SampleID is already a factor

## Data standardized!

model2 <- glmmTMB(FREQ ~ YEAR + TEMPWATR + BOTTYPE1,
  data = new_data2,
  family = nbinom2
)
```

ggemmeans method

```r
index.pred <- ggeffects::ggemmeans(model2,
  terms = "YEAR",
  type = "fixed"
```

```
)
plot(index.pred)
```



Bootstrap method

```
source(here::here("scripts/bootstrap_functions_AT_lnE2.R"))
message("sourced")

## sourced

index.pred <- boot.ZI_AT(
  best = model2,
  type_z = "nb",
  nboots = 1000
)

## starting fxn

## year name fixed

## starting boots

## session planned

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')
```

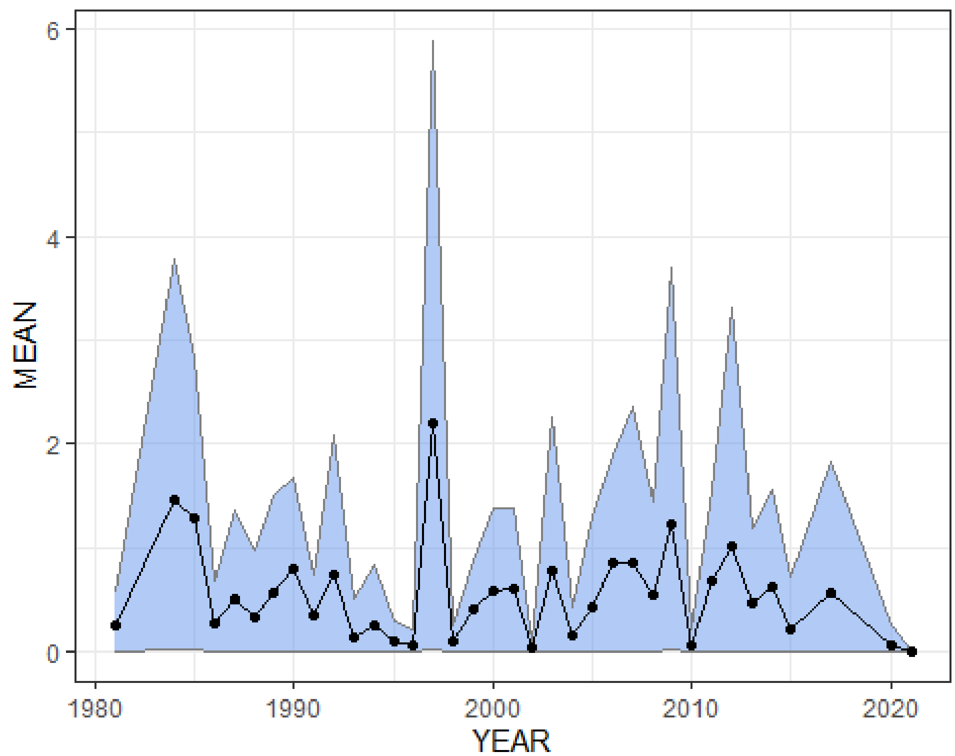
```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
```

```
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
```

```
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## finished boots

## Number of successful predictions: 949

message("predicted")

## predicted

index.pred %>%
  plot_boot()
```




```
message("plotted")

## plotted
```

Step 5: Save index

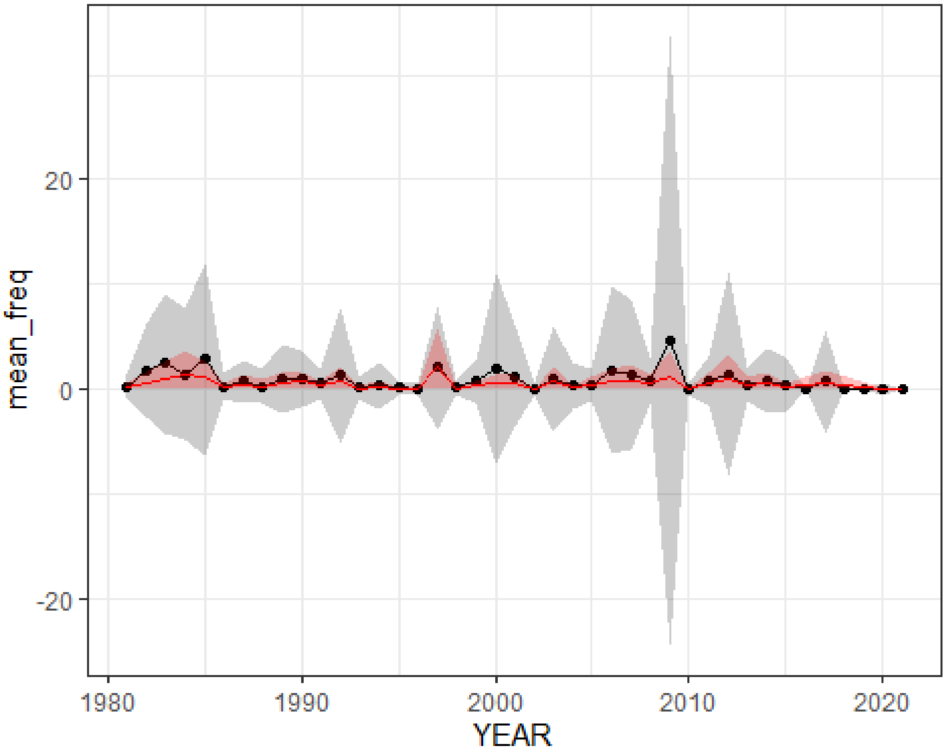
```
index.out <- as.data.frame(index.pred)
index.out$CV <- abs(index.out$SE / index.out$MEAN)

names(index.out)[c(1, 2)] <- c("Year", "Frequency")

write.csv(index.out, here::here("reports/updated/MD_index5.csv"), row.names =
FALSE)
```

Develop nominal index

```
# not sure if/how to standardize for effort, so skipping that for now
# also not sure what stratum weights are, so just weighting everything
equally for now
# are system/river columns equivalent to strata?

nominal_data <- indata %>%
  dplyr::group_by(YEAR) %>%
  dplyr::mutate(year_total = sum(FREQ, na.rm = TRUE)) %>%
  dplyr::filter(
    year_total > 0,
    MONTH %in% unique(new_data2$MONTH),
    HAULNO == 1
```

```r
  ) %>%
  dplyr::summarise(
    mean_freq = mean(FREQ),
    var_freq = var(FREQ),
    n_obs = length(FREQ),
    sd_freq = sd(FREQ), # treating all as one strata
    cv_freq = (sd_freq / mean_freq)
  )

write.csv(nominal_data,
  here::here("reports/updatedMD_index_nominal3.csv"),
  row.names = FALSE
)

ggplot2::ggplot(
  nominal_data,
  ggplot2::aes(x = YEAR)
) +
  ggplot2::geom_ribbon(ggplot2::aes(
    ymin = mean_freq - 1.96 * sd_freq,
    ymax = mean_freq + 1.96 * sd_freq
  ),
  alpha = 0.25
  ) +
  ggplot2::geom_line(ggplot2::aes(y = mean_freq)) +
  ggplot2::geom_point(ggplot2::aes(y = mean_freq)) +
  ggplot2::theme_bw() +
  ggplot2::geom_line(
    data = index.out,
    ggplot2::aes(
      x = Year %>%
        as.character() %>%
        as.numeric(),
      y = Frequency
    ),
    color = "red",
    inherit.aes = FALSE
  ) +
  ggplot2::geom_ribbon(
    data = index.out,
    ggplot2::aes(
      x = Year %>%
        as.character() %>%
        as.numeric(),
      ymin = LCI,
      ymax = UCI
    ),
    alpha = 0.25,
    fill = "red",
```

```
    inherit.aes = FALSE
  )
```



The confidence intervals go below zero. Should we log-transform the data?

Bluefish index standardization - Virginia Striped Bass Seine Survey

Abigail Tyrell

2022-05-18

Develop standardized index

Step 1: Data processing

Load functions

```
source(here::here("scripts/functions.R"))

## Loading required package: carData

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

## Warning: package 'glmmTMB' was built under R version 4.0.5

## Registered S3 methods overwritten by 'lme4':
##   method                          from
##   cooks.distance.influence.merMod car
##   influence.merMod                car
##   dfbeta.influence.merMod         car
##   dfbetas.influence.merMod        car

## Warning: package 'DHARMa' was built under R version 4.0.5

## This is DHARMa 0.4.4. For overview type '?DHARMa'. For recent changes,
type news(package = 'DHARMa')

## Warning: package 'bbmle' was built under R version 4.0.5

## Loading required package: stats4

## Warning: package 'ggeffects' was built under R version 4.0.5

## Warning: package 'ggplot2' was built under R version 4.0.5
```

Knit tables with `knitr::kable`

```r
knit_print.data.frame = function(x, ...) {
  res = paste(c("", "", knitr::kable(x, digits = 2)), collapse = "\n")
  knitr::asis_output(res)
}

registerS3method(
  "knit_print", "data.frame", knit_print.data.frame,
  envir = asNamespace("knitr")
)
```

Read in data

```r
indata <- readxl::read_excel(here::here("data/2021_bluefish_data_VA.xlsx"),
                             sheet = "SurveyData_SBSS",
                             col_types = c(rep("guess", 17), "numeric"))
head(indata, n = 2)
```

SampleID	Cruiseno	Stanum	State	Program	Gear	Haul	Year	Month	Day	Round	Bf.count	Area	Lat	Lon	Bottom.wt	Bottom.sa	Bottom.do
SB670707RA44	SB670707RA	44	VA	Virginia Striped Bass Seine Survey	BeachSeine	1	1967	7	7	1	1	RA	38.02	-76.92	25.6	1.6	NA
SB670811RA8	SB670811RA	8	VA	Virginia Striped Bass Seine Survey	BeachSeine	1	1967	8	11	3	1	RA	37.61	-76.44	24.5	15.0	NA

```r
# modify data formatting
indata <- indata %>%
```

```r
  dplyr::rename(FREQ = Bf.count,
                Stratum = Stanum) %>%
  dplyr::mutate(Haul = as.factor(Haul),
                Round = as.factor(Round),
                Stratum = as.factor(Stratum)) %>%
  dplyr::select(-c(State, Program, Gear)) # same value for all data

head(indata, n = 2)
```

SampleID	Cruiseno	Stratum	Haul	Year	Month	Day	Round	FREQ	Area	Lat	Lon	Bottom.wt	Bottom.sa	Bottom.do
SB670707RA44	SB670707RA	44	1	1967	7	7	1	1	RA	38.02	-76.92	25.6	1.6	NA
SB670811RA8	SB670811RA	8	1	1967	8	11	3	1	RA	37.61	-76.44	24.5	15.0	NA

Remove missing data

```r
# remove anything not Haul 1
length1 <- nrow(indata)
new_data <- indata %>%
  dplyr::filter(Haul == 1) %>%
  dplyr::select(-Haul) # remove because there is only 1 value now
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")
```

```
## 3348 rows removed
```

```r
# Bottom.do has a lot of missing data, so removing
new_data <- new_data %>%
  dplyr::select(-Bottom.do)

new_data <- remove_missing(new_data,
                           f_col = "FREQ",
                           t_col = "Year")
```

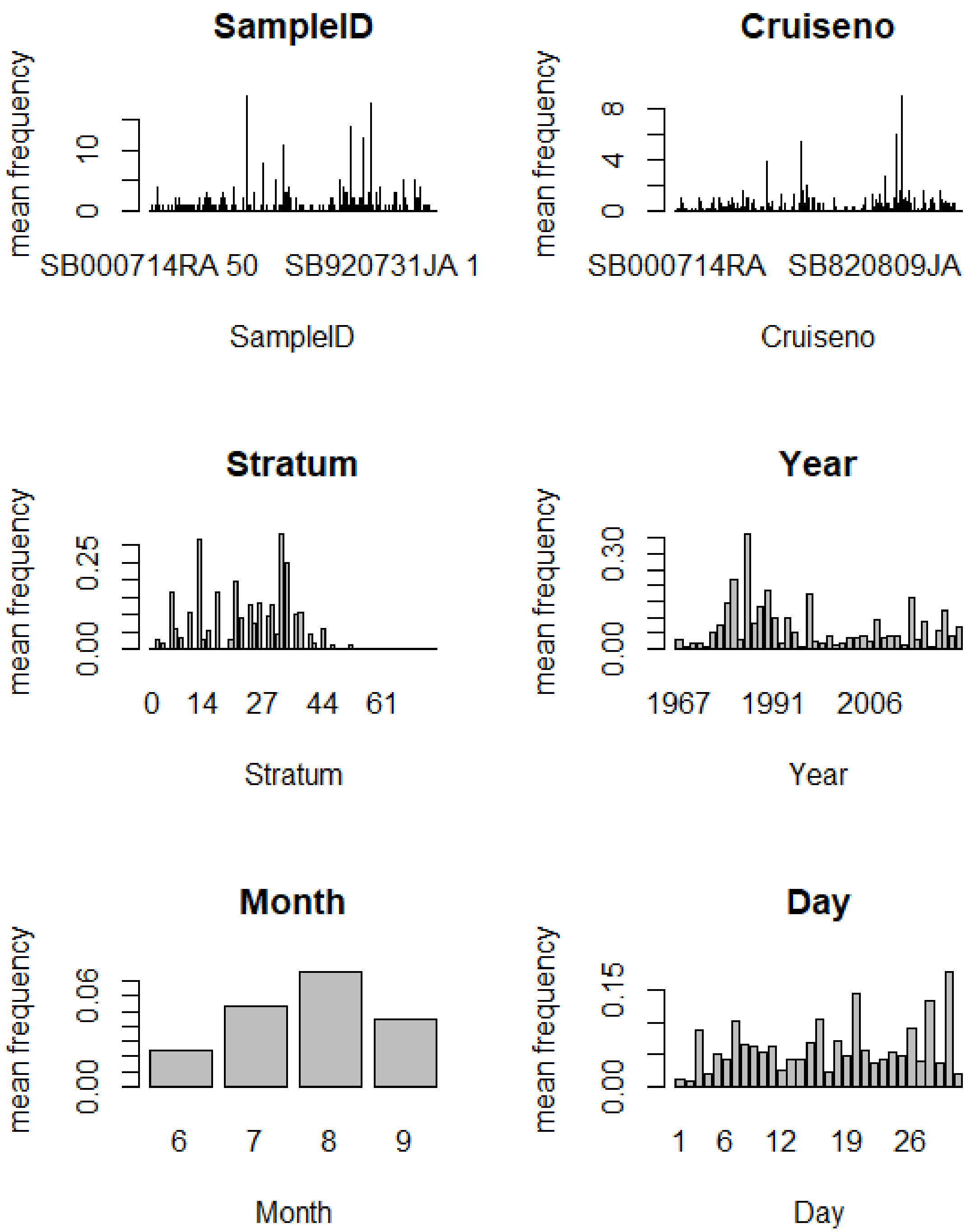
```
## Data before removing missing values:

##    SampleID           Cruiseno            Stratum          Year
##  Length:7551        Length:7551        44     : 465   Min.   :1967
##  Class :character   Class :character   50     : 413   1st Qu.:1990
##  Mode  :character   Mode  :character   28     : 394   Median :2000
##                                        37     : 386   Mean   :1998
##                                        41     : 384   3rd Qu.:2010
```

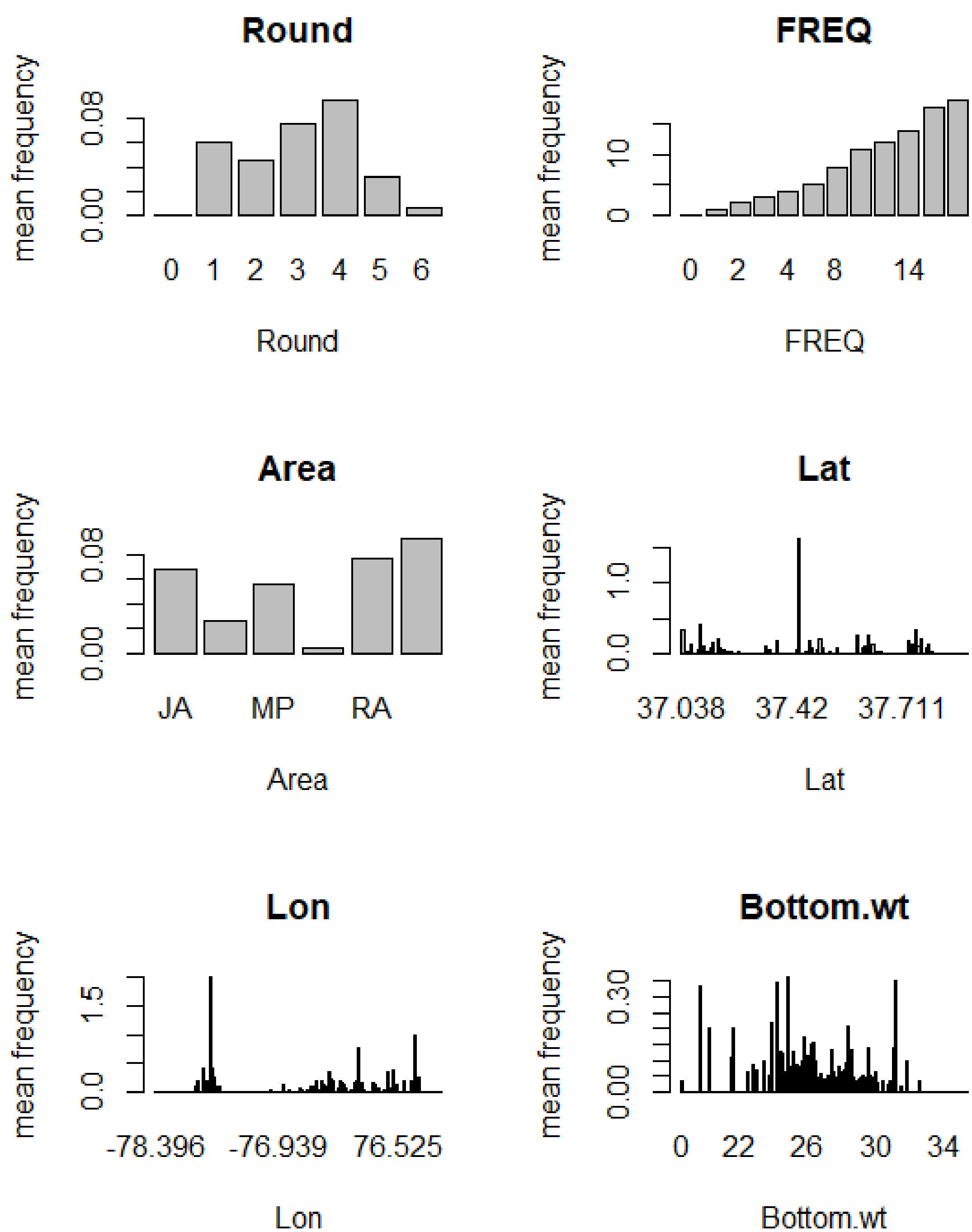
```
##                                         36     : 330   Max.   :2020  
##                                         (Other):5179                 
##      Month            Day          Round         FREQ              Area          
##  Min.   :6.000   Min.   : 1.00   0:   1   Min.   : 0.00000   Length:7551       
##  1st Qu.:7.000   1st Qu.: 8.00   1:1486   1st Qu.: 0.00000   Class 
:character  
##  Median :8.000   Median :15.00   2:1562   Median : 0.00000   Mode  
:character  
##  Mean   :7.801   Mean   :15.24   3:1535   Mean   : 0.05456                     
##  3rd Qu.:8.000   3rd Qu.:23.00   4:1437   3rd Qu.: 0.00000                     
##  Max.   :9.000   Max.   :31.00   5:1356   Max.   :19.00000                     
##                                  6: 174                                        
##       Lat             Lon           Bottom.wt       Bottom.sa     
##  Min.   :37.04   Min.   :-78.40   Min.   : 0.00   Min.   : 0.000  
##  1st Qu.:37.30   1st Qu.:-77.01   1st Qu.:26.50   1st Qu.: 0.100  
##  Median :37.56   Median :-76.89   Median :28.00   Median : 0.700  
##  Mean   :37.58   Mean   :-70.13   Mean   :27.72   Mean   : 3.892  
##  3rd Qu.:37.74   3rd Qu.:-76.74   3rd Qu.:29.50   3rd Qu.: 6.400  
##  Max.   :38.26   Max.   : 77.40   Max.   :39.90   Max.   :25.000  
##  NA's   :74      NA's   :74       NA's   :25      NA's   :42      

## Data after removing missing values:

##    SampleID           Cruiseno            Stratum          Year     
##  Length:7432        Length:7432        44     : 461   Min.   :1967  
##  Class :character   Class :character   50     : 411   1st Qu.:1990  
##  Mode  :character   Mode  :character   28     : 394   Median :2000  
##                                         37     : 382   Mean   :1998  
##                                         41     : 381   3rd Qu.:2010  
##                                         36     : 330   Max.   :2020  
##                                         (Other):5073                 
##      Month          Day          Round         FREQ              Area          
##  Min.   :6.0   Min.   : 1.00   0:   1   Min.   : 0.00000   Length:7432       
##  1st Qu.:7.0   1st Qu.: 8.00   1:1473   1st Qu.: 0.00000   Class 
:character  
##  Median :8.0   Median :15.00   2:1540   Median : 0.00000   Mode  
:character  
##  Mean   :7.8   Mean   :15.26   3:1504   Mean   : 0.05517                     
##  3rd Qu.:8.0   3rd Qu.:23.00   4:1411   3rd Qu.: 0.00000                     
##  Max.   :9.0   Max.   :31.00   5:1329   Max.   :19.00000                     
##                                6: 174                                        
##       Lat             Lon           Bottom.wt      Bottom.sa     
##  Min.   :37.04   Min.   :-78.40   Min.   : 0.0   Min.   : 0.000  
##  1st Qu.:37.30   1st Qu.:-77.01   1st Qu.:26.5   1st Qu.: 0.100  
##  Median :37.56   Median :-76.89   Median :28.0   Median : 0.700  
##  Mean   :37.58   Mean   :-70.11   Mean   :27.7   Mean   : 3.929  
##  3rd Qu.:37.74   3rd Qu.:-76.74   3rd Qu.:29.5   3rd Qu.: 6.500  
##  Max.   :38.26   Max.   : 77.40   Max.   :39.9   Max.   :25.000  
##
```

```
## 119 datapoints removed (1.6% of data removed)

## Data after removing years with no catch:

##    SampleID           Cruiseno            Stratum          Year     
##  Length:6759        Length:6759        44     : 412   Min.   :1967  
##  Class :character   Class :character   50     : 369   1st Qu.:1990  
##  Mode  :character   Mode  :character   28     : 360   Median :2001  
##                                        37     : 347   Mean   :1999  
##                                        41     : 340   3rd Qu.:2010  
##                                        36     : 302   Max.   :2020  
##                                        (Other):4629                 
##      Month            Day         Round         FREQ              Area          
##  Min.   :6.000   Min.   : 1.00   0:   1   Min.   : 0.00000   Length:6759       
##  1st Qu.:7.000   1st Qu.: 8.00   1:1310   1st Qu.: 0.00000   Class 
:character  
##  Median :8.000   Median :14.00   2:1382   Median : 0.00000   Mode  
:character  
##  Mean   :7.808   Mean   :15.06   3:1364   Mean   : 0.06066                     
##  3rd Qu.:8.000   3rd Qu.:22.00   4:1315   3rd Qu.: 0.00000                     
##  Max.   :9.000   Max.   :31.00   5:1232   Max.   :19.00000                     
##                                  6: 155                                        
##       Lat             Lon           Bottom.wt       Bottom.sa     
##  Min.   :37.04   Min.   :-78.40   Min.   : 0.00   Min.   : 0.000  
##  1st Qu.:37.30   1st Qu.:-77.04   1st Qu.:26.60   1st Qu.: 0.100  
##  Median :37.56   Median :-76.90   Median :28.10   Median : 0.800  
##  Mean   :37.58   Mean   :-73.87   Mean   :27.78   Mean   : 3.964  
##  3rd Qu.:37.74   3rd Qu.:-76.78   3rd Qu.:29.50   3rd Qu.: 6.600  
##  Max.   :38.26   Max.   : 77.40   Max.   :39.90   Max.   :25.000  
## 

## 673 datapoints removed (9.1% of data removed)
```

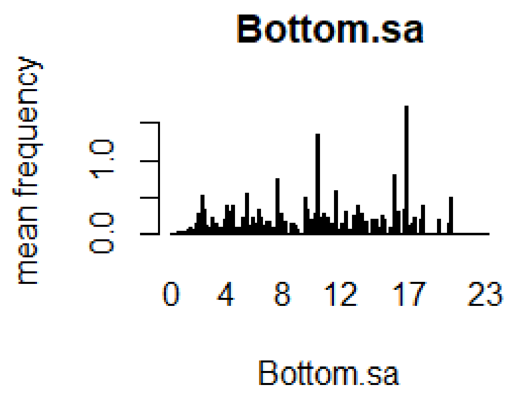
Plot data and decide if any data should be removed due to sparse sampling

```
plot_obs(new_data,
         col = "FREQ")
```


SampleID
mean frequency
0
10
SB000714RA 50
SB920731JA 1
SampleID
Cruiseno
mean frequency
0
4
8
SB000714RA
SB820809JA
Cruiseno
Stratum
mean frequency
0.00
0.25
0
14
27
44
61
Stratum
Year
mean frequency
0.00
0.30
1967
1991
2006
Year
Month
mean frequency
0.00
0.06
6
7
8
9
Month
Day
mean frequency
0.00
0.15
1
6
12
19
26
Day




```
## Please manually remove sparsely sampled data!
```

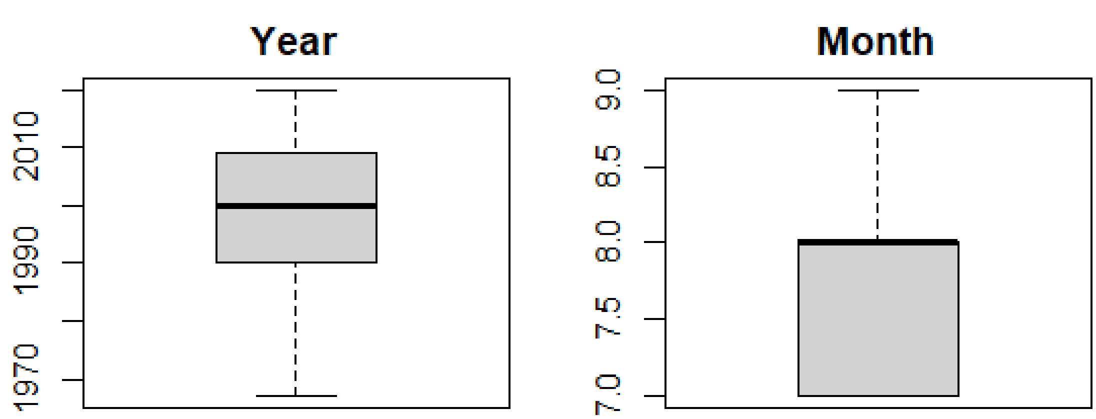


Remove sparse data

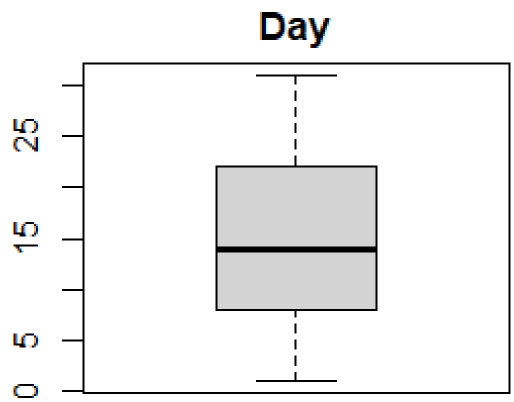
```
# remove June
length1 <- nrow(new_data)
new_data <- new_data %>%
  dplyr::filter(Month != 6)
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")

## 82 rows removed
```

Check for outliers

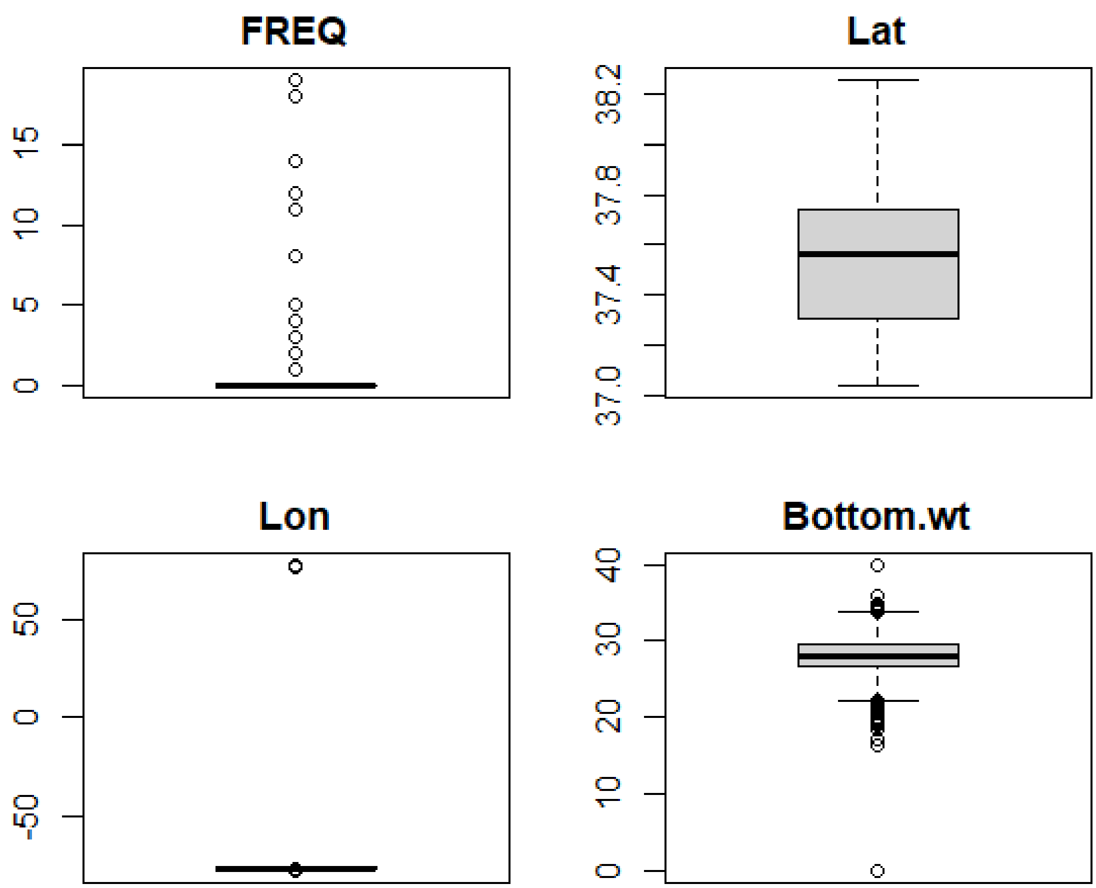
```
plot_outliers(new_data)

## SampleID is not numeric

## Cruiseno is not numeric

## Stratum is not numeric
```



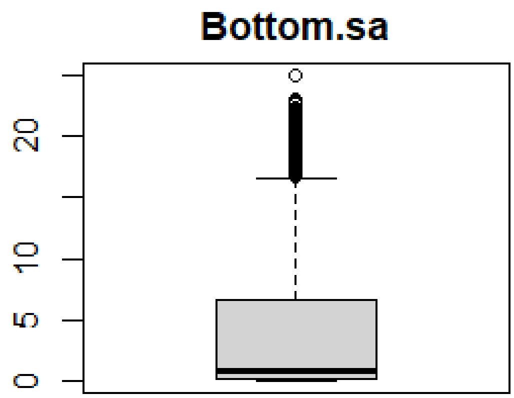

```
## Round is not numeric
```




```
## Area is not numeric
```




```
## Please manually remove any outliers!
```

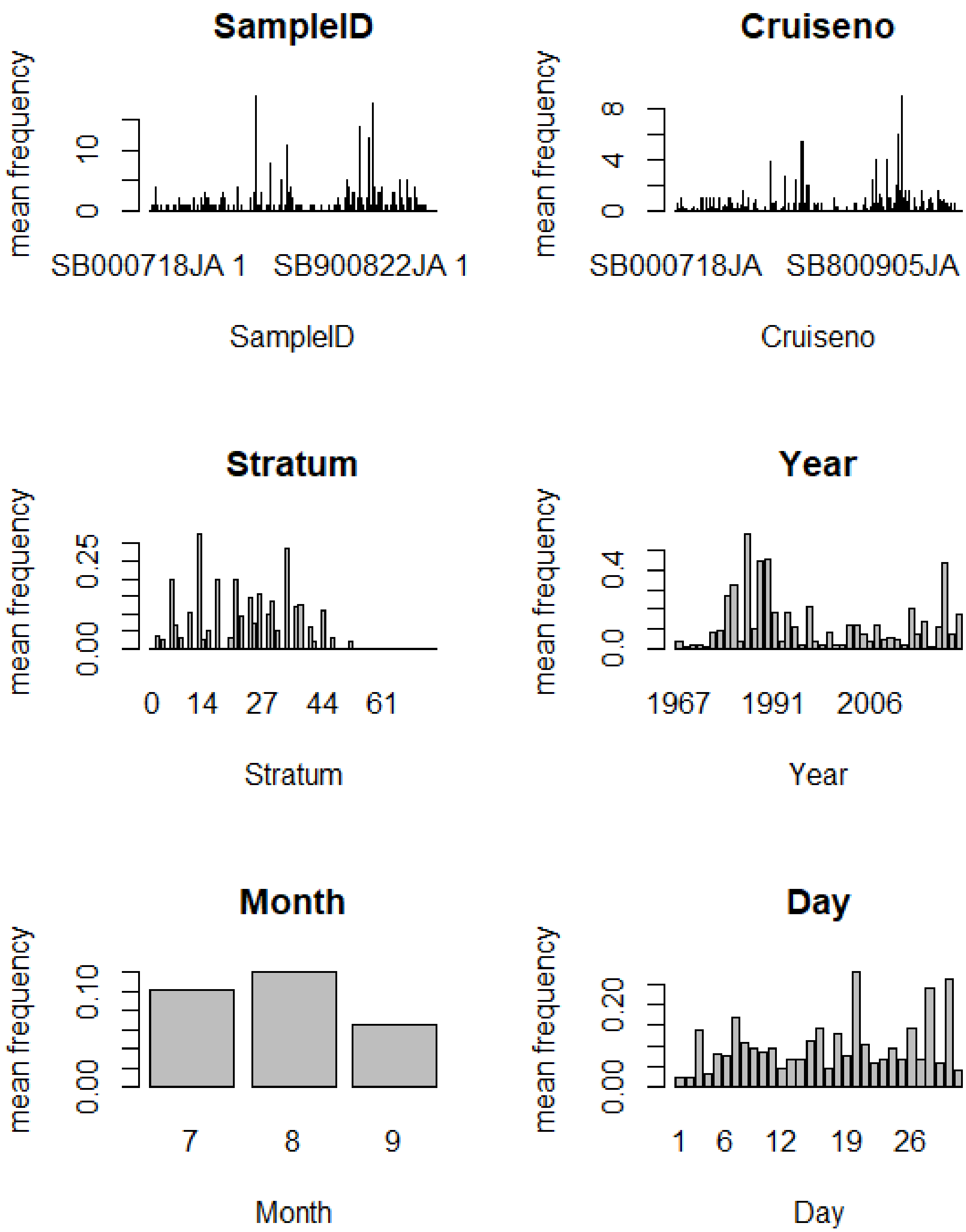


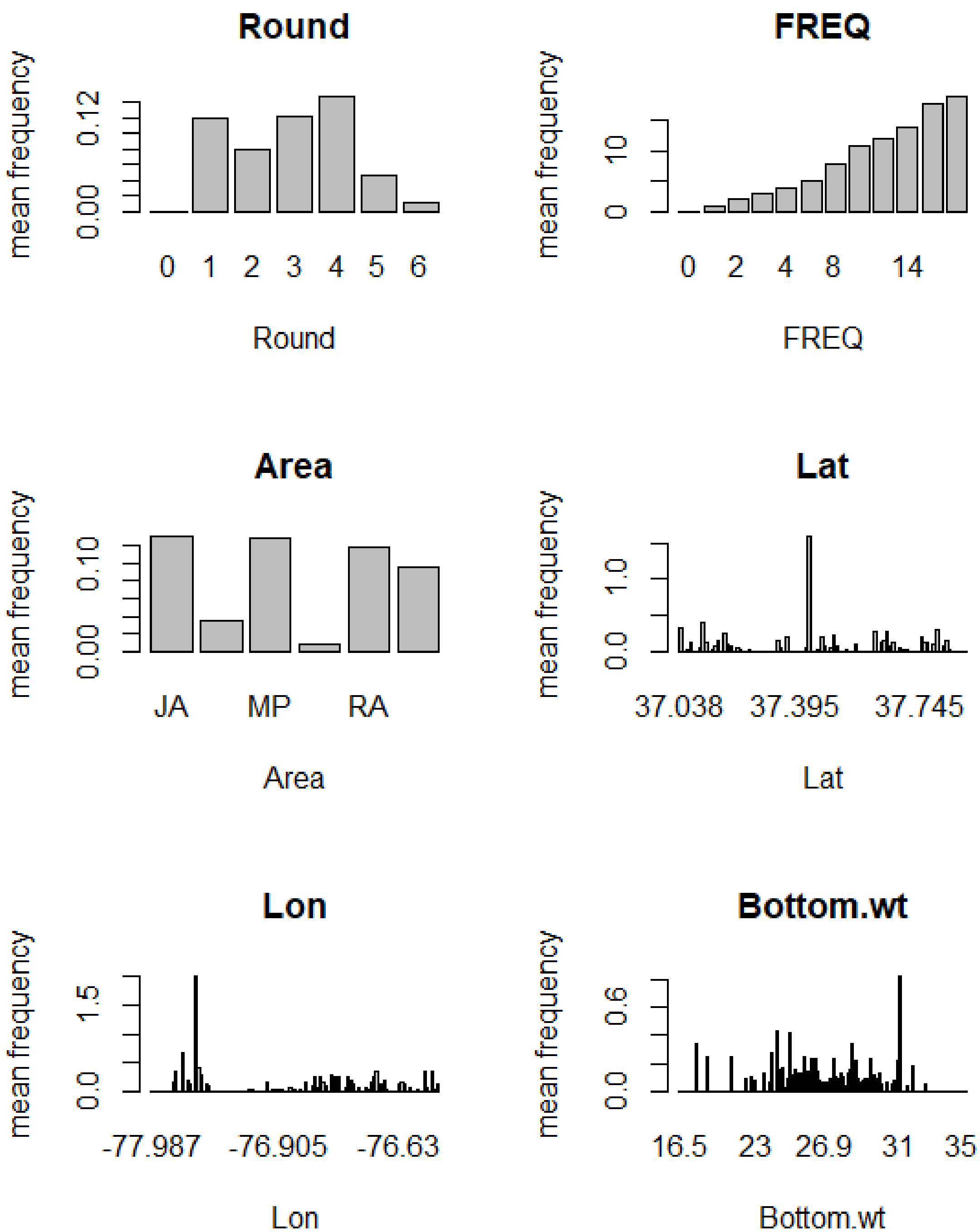
If needed, remove outliers

```
new_data <- new_data %>%
  dplyr::mutate(Lon = -abs(Lon)) %>% # Fix positive Lon
  dplyr::filter(Bottom.wt > 0, # remove temp = 0
                Bottom.sa > 0.2) # several stations with 0 salinity
```

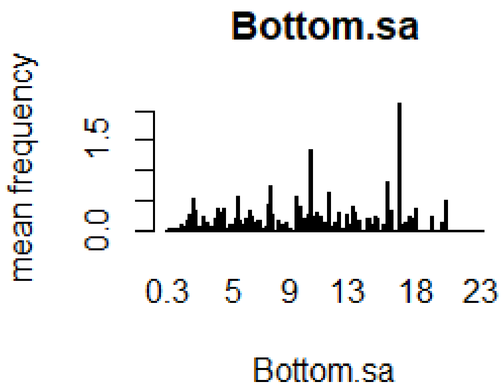
Check for sparse sampling again

```
plot_obs(new_data,
         col = "FREQ")
```


SampleID
mean frequency
0
10
SB000718JA 1
SB900822JA 1
SampleID
Cruiseno
mean frequency
0
4
8
SB000718JA
SB800905JA
Cruiseno
Stratum
mean frequency
0.00
0.25
0
14
27
44
61
Stratum
Year
mean frequency
0.0
0.4
1967
1991
2006
Year
Month
mean frequency
0.00
0.10
7
8
9
Month
Day
mean frequency
0.00
0.20
1
6
12
19
26
Day




```
## Please manually remove sparsely sampled data!
```



Check again for years with no catch

```
check_missing(new_data, f_col = "FREQ", t_col = "Year")

## Years with no catch: NONE
```

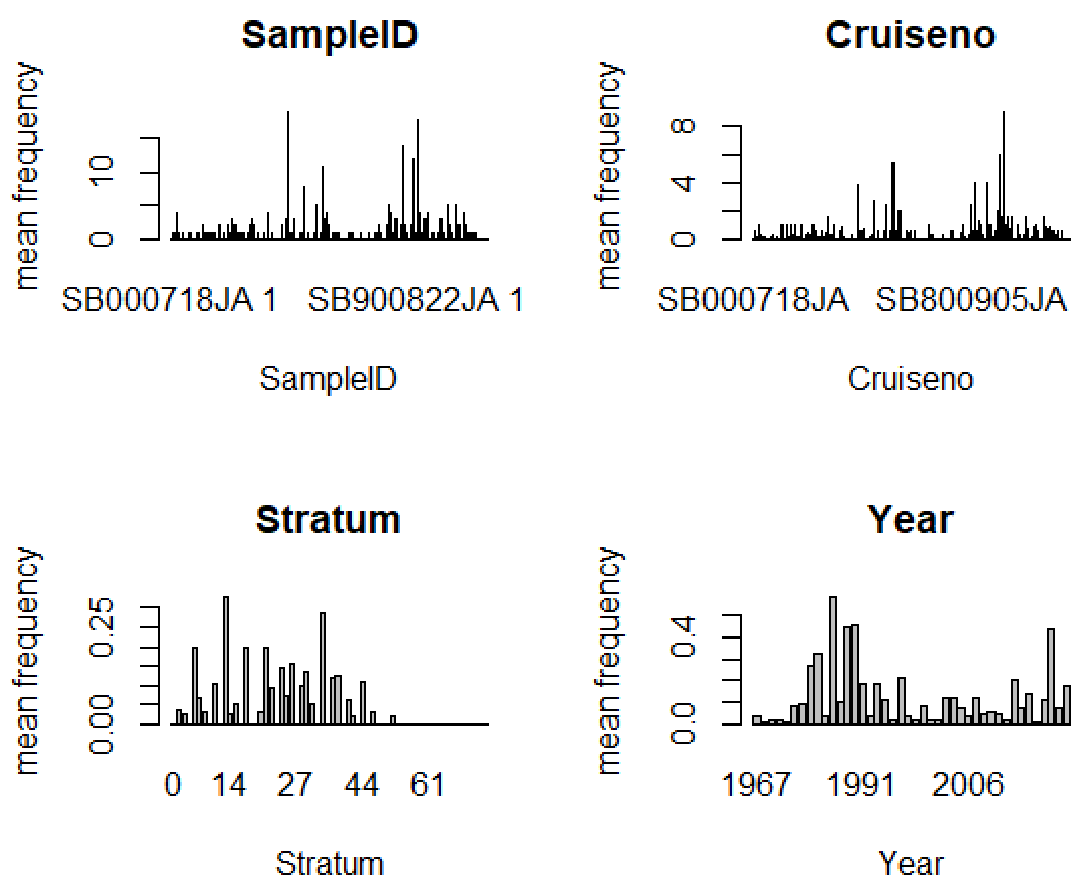
Change continuous variables to z-score

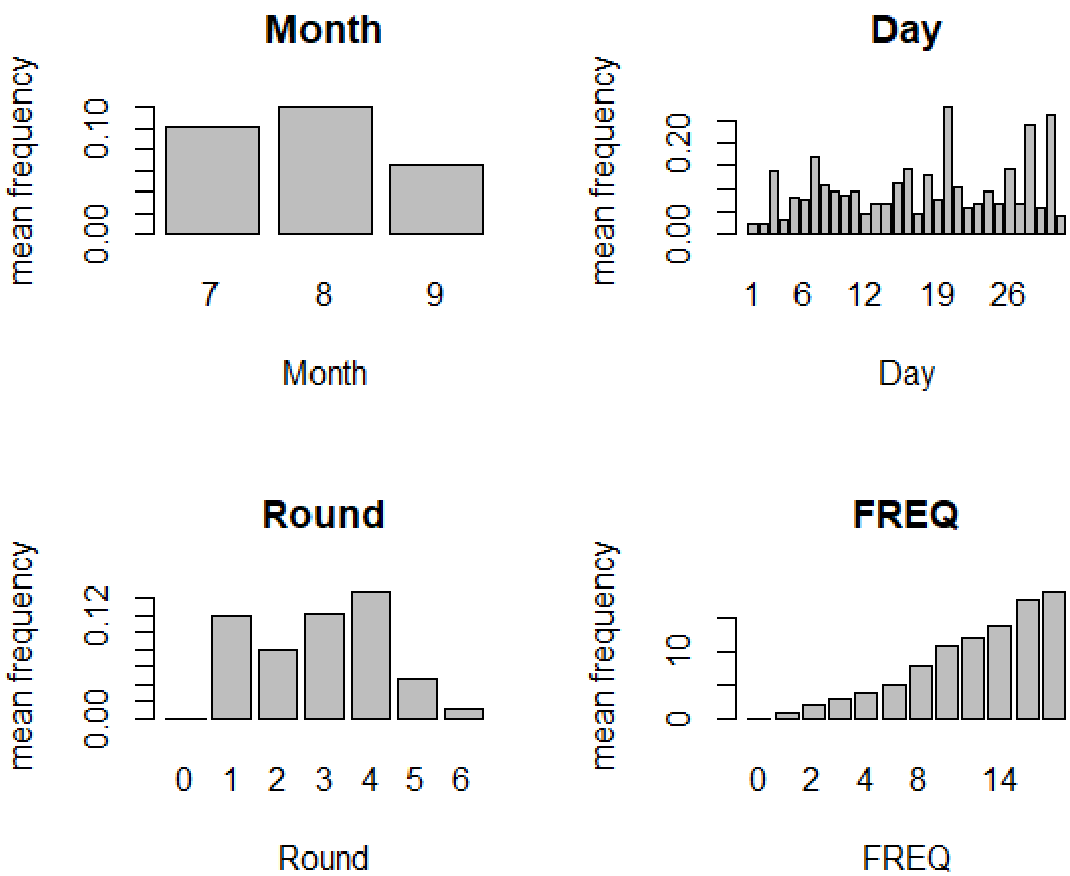
```
new_data2 <- standardize_data(new_data,
                              cols = c("Year", "Month", "Stratum", "Day",
                                       "Round", "Lat", "Lon", "Area",
                                       "SampleID", "Cruiseno"),
                              f_col = "FREQ")

## SampleID has been changed to factor

## Cruiseno has been changed to factor

## Stratum is already a factor

## Year has been changed to factor

## Month has been changed to factor

## Day has been changed to factor

## Round is already a factor

## Area has been changed to factor

## Lat has been changed to factor

## Lon has been changed to factor

## Bottom.wt has been standardized

## Bottom.sa has been standardized

## Data standardized!
```

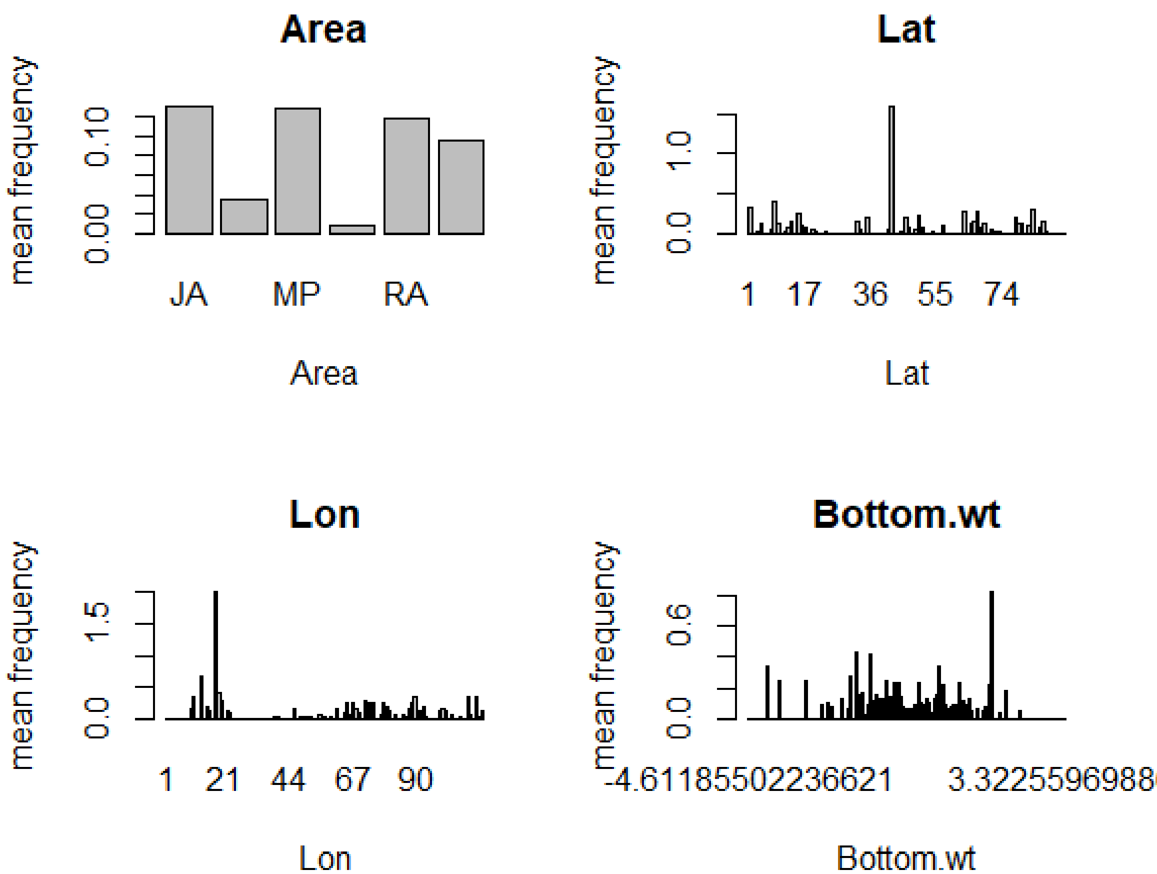
```r
# change Lat and Lon back to numeric
new_data2 <- new_data2 %>%
  dplyr::mutate(Lat = as.numeric(Lat),
                Lon = as.numeric(Lon))
head(new_data2, n = 2)
```

SampleID	Cruiseno	Stratum	Year	Month	Day	Round	FREQ	Area	Lat	Lon	Bottom.wt	Bottom.sa
SB670707RA 44	SB670707RA	44	1967	7	7	1	1	RA	86	46	-0.87	-0.85
SB670811RA 8	SB670811RA	8	1967	8	11	3	1	RA	67	112	-1.32	1.46

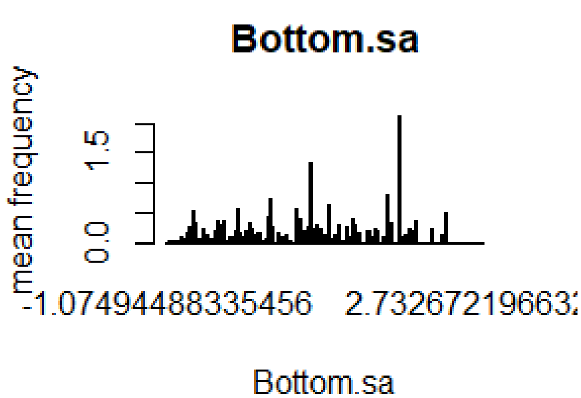
```r
plot_obs(new_data2,
         col = "FREQ")
```




Month
mean frequency
0.00
0.10
7
8
9
Month
Day
mean frequency
0.00
0.20
1
6
12
19
26
Day
Round
mean frequency
0.00
0.12
0
1
2
3
4
5
6
Round
FREQ
mean frequency
0
10
0
2
4
8
14
FREQ




```
## Please manually remove sparsely sampled data!
```

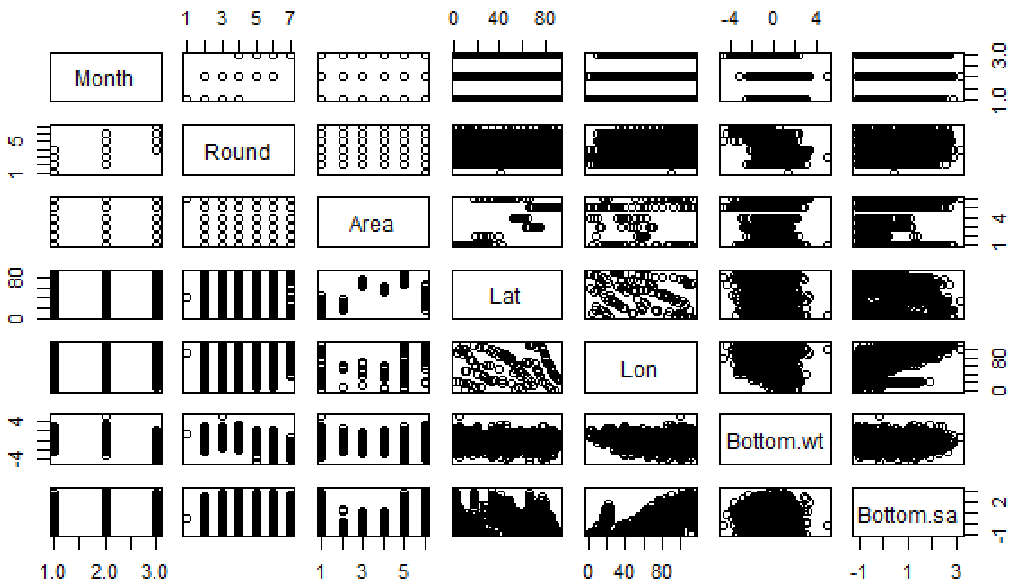


Check colinearity

```
check_colin(new_data2,
            cols = c("Year", "Day",
```

```
                     "SampleID", "Cruiseno", "FREQ",
                     "Stratum"))
```

```
##             Month Round  Area   Lat   Lon Bottom.wt Bottom.sa
## Month        1.00  0.89 -0.03 -0.01 -0.03     -0.48      0.05
## Round        0.89  1.00 -0.02  0.00 -0.05     -0.43      0.04
## Area        -0.03 -0.02  1.00  0.69  0.05     -0.14      0.36
## Lat         -0.01  0.00  0.69  1.00 -0.30     -0.08     -0.12
## Lon         -0.03 -0.05  0.05 -0.30  1.00     -0.11      0.73
## Bottom.wt -0.48 -0.43 -0.14 -0.08 -0.11      1.00     -0.11
## Bottom.sa  0.05  0.04  0.36 -0.12  0.73     -0.11      1.00

## Round is correlated with Month
```



Remove Round

```
new_data2 <- new_data2 %>%
  dplyr::select(-Round)
```

Check variance inflation factor - Remove covariates with GVIF > 3, starting with highest value, until all GVIF < 3

```
check_vif(model = "FREQ ~ Year + Month + Lat + Lon + Bottom.wt + Bottom.sa",
          data = new_data2)
```

```
## # A tibble: 6 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Lat        1.28     1              1.13
## 2 Month      1.53     2              1.11
```

```
## 3 Bottom.wt  1.69     1              1.30
## 4 Year       3.09    41              1.01
## 5 Bottom.sa  3.77     1              1.94
## 6 Lon        5.36     1              2.31

## Variance inflation factor too high! Please remove a covariate.

# remove Lon

check_vif(model = "FREQ ~ Year + Month + Lat + Bottom.wt + Bottom.sa",
          data = new_data2)

## # A tibble: 5 x 4
##   Covariate  GVIF    Df `GVIF^(1/(2*Df))`
##   <chr>     <dbl> <dbl>             <dbl>
## 1 Lat        1.05     1              1.02
## 2 Bottom.sa  1.11     1              1.05
## 3 Year       1.40    41              1.00
## 4 Month      1.46     2              1.10
## 5 Bottom.wt  1.69     1              1.30

## Passed!
```

Step 2: Model selection

Compare negative binom, zero-infl neg binom, zero-altered neg binom

Comparison 1: Full models

```
compare_models(data = new_data2,
               nb_model = "FREQ ~ Year + Month + Lat + Bottom.wt +
Bottom.sa",

               zic_model = "FREQ ~ Year + Month + Lat + Bottom.wt +
Bottom.sa",
               zi_model = "~ Year + Month + Lat + Bottom.wt + Bottom.sa",

               zac_model = "FREQ ~ Year + Month + Lat + Bottom.wt +
Bottom.sa",
               za_model = "~ Year + Month + Lat + Bottom.wt + Bottom.sa",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 5.6

## Negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ Year + Month + Lat + Bottom.wt + Bottom.sa
## Data: data
```

```
##
##      AIC      BIC   logLik deviance df.resid
##   2124.7   2426.7  -1014.4   2028.7     3937
##
##
## Dispersion parameter for nbinom2 family (): 0.119
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.750857   0.711515  -5.272 1.35e-07 ***
## Year1968    -0.929276   1.266124  -0.734 0.462977
## Year1969    -0.468944   1.043367  -0.449 0.653105
## Year1970    -0.543023   1.031988  -0.526 0.598755
## Year1971    -0.992458   1.255766  -0.790 0.429341
## Year1972     1.236977   0.828190   1.494 0.135283
## Year1980     1.603626   0.830614   1.931 0.053526 .
## Year1982     2.905001   0.989803   2.935 0.003336 **
## Year1985     2.864566   0.875863   3.271 0.001073 **
## Year1986     0.589952   1.062354   0.555 0.578673
## Year1987     3.589679   0.854454   4.201 2.66e-05 ***
## Year1988     1.905142   0.881684   2.161 0.030711 *
## Year1989     3.066972   0.815381   3.761 0.000169 ***
## Year1990     3.217257   0.788470   4.080 4.50e-05 ***
## Year1991     1.966717   0.788014   2.496 0.012568 *
## Year1992     0.561337   0.972440   0.577 0.563773
## Year1993     2.422830   0.804678   3.011 0.002604 **
## Year1995     1.613860   0.824749   1.957 0.050372 .
## Year1996    -0.416678   1.272807  -0.327 0.743389
## Year1997     2.258537   0.750391   3.010 0.002614 **
## Year1998     0.658644   0.904225   0.728 0.466364
## Year1999    -0.493954   0.945891  -0.522 0.601524
## Year2000     1.459554   0.836246   1.745 0.080921 .
## Year2001    -0.290463   1.033538  -0.281 0.778681
## Year2002    -0.441147   0.932663  -0.473 0.636215
## Year2003     1.731134   0.909595   1.903 0.057015 .
## Year2004     1.808494   0.882218   2.050 0.040370 *
## Year2005     1.410321   0.831621   1.696 0.089910 .
## Year2006     0.468407   0.901018   0.520 0.603158
## Year2007     1.783720   0.784364   2.274 0.022960 *
## Year2008     0.801279   0.854376   0.938 0.348320
## Year2009     0.818200   0.837652   0.977 0.328679
## Year2010     0.635729   0.830203   0.766 0.443824
## Year2011     0.087698   1.038672   0.084 0.932712
## Year2012     1.801366   0.778897   2.313 0.020738 *
## Year2013     1.055577   0.871957   1.211 0.226055
## Year2014     1.699067   0.794366   2.139 0.032444 *
## Year2015    -0.741196   1.261153  -0.588 0.556725
## Year2016     1.888254   0.873556   2.162 0.030651 *
## Year2018     3.122193   0.834319   3.742 0.000182 ***
## Year2019     1.422583   0.884611   1.608 0.107804
```

```
## Year2020     2.194226   0.833425   2.633 0.008469 **
## Month8      -0.035217   0.189945  -0.185 0.852910
## Month9      -1.165626   0.287911  -4.049 5.15e-05 ***
## Lat          0.001092   0.002946   0.371 0.710841
## Bottom.wt   -0.376712   0.111331  -3.384 0.000715 ***
## Bottom.sa    0.495826   0.090721   5.465 4.62e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.11

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class
"c('simpleWarning', 'warning', 'condition')"
```

Zero-inflated and zero-adjusted did not work. Negative binomial actually seems ok. Lat is not statistically significant in the negative binomial. Try from the zero models.

Comparison 2: Removing Lat from zero models

```
compare_models(data = new_data2,
               models = c("zi", "za"),

               zic_model = "FREQ ~ Year + Month + Bottom.wt + Bottom.sa",
               zi_model = "~ Year + Month + Bottom.wt + Bottom.sa",

               zac_model = "FREQ ~ Year + Month + Bottom.wt + Bottom.sa",
               za_model = "~ Year + Month + Bottom.wt + Bottom.sa",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 5.6

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Still no good… try dropping each of Month, Bottom.wt, Bottom.sa

Comparison 3: Removing Month from zero models

```
compare_models(data = new_data2,
               models = c("zi", "za"),

               zic_model = "FREQ ~ Year + Bottom.wt + Bottom.sa",
               zi_model = "~ Year + Bottom.wt + Bottom.sa",

               zac_model = "FREQ ~ Year + Bottom.wt + Bottom.sa",
               za_model = "~ Year + Bottom.wt + Bottom.sa",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 5.6

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Comparison 4: Removing Bottom.wt from zero models

```
compare_models(data = new_data2,
               models = c("zi", "za"),

               zic_model = "FREQ ~ Year + Month + Bottom.sa",
               zi_model = "~ Year + Month + Bottom.sa",

               zac_model = "FREQ ~ Year + Month + Bottom.sa",
               za_model = "~ Year + Month + Bottom.sa",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 5.6

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work
```

```
## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Comparison 5: Removing Bottom.sa from zero models

```
compare_models(data = new_data2,
               models = c("zi", "za"),

               zic_model = "FREQ ~ Year + Month + Bottom.wt",
               zi_model = "~ Year + Month + Bottom.wt",

               zac_model = "FREQ ~ Year + Month + Bottom.wt",
               za_model = "~ Year + Month + Bottom.wt",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 5.6

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Zero-inflated worked! (doesn't work after fixing salinity)

Try zero-adjusted model with just one of Month, Bottom.wt, Bottom.sa

Comparison 6: Zero with Month

```
compare_models(data = new_data2,
               models = c("zi", "za"),

               zic_model = "FREQ ~ Year + Month",
               zi_model = "~ Year + Month",

               zac_model = "FREQ ~ Year + Month",
               za_model = "~ Year + Month",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 5.6

## Zero-inflated negative binomial

## Model did not work
```

```
## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Comparison 7: Zero with Bottom.wt

```
compare_models(data = new_data2,
              models = c("zi", "za"),

               zic_model = "FREQ ~ Year + Bottom.wt",
               zi_model = "~ Year + Bottom.wt",

               zac_model = "FREQ ~ Year + Bottom.wt",
               za_model = "~ Year + Bottom.wt",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 5.6

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Comparison 8: Zero with Bottom.sa

```
compare_models(data = new_data2,
               models = c("zi", "za"),

               zic_model = "FREQ ~ Year + Bottom.sa",
               zi_model = "~ Year + Bottom.sa",

               zac_model = "FREQ ~ Year + Bottom.sa",
               za_model = "~ Year + Bottom.sa",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 5.6

## Zero-inflated negative binomial

## Model did not work
```

```
## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Still not working. Try with Year only

Comparison 9: Year only

```
compare_models(data = new_data2,
               models = c("zi", "za"),

               zic_model = "FREQ ~ Year",
               zi_model = "~ Year",

               zac_model = "FREQ ~ Year",
               za_model = "~ Year",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 5.6

## Zero-inflated negative binomial

## Model did not work

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Still not working. Trying intercept-only models.

Comparison 10: Intercept binomial model

```
compare_models(data = new_data2,
               models = c("zi", "za"),

               zic_model = "FREQ ~ Year",
               zi_model = "~ 1",

               zac_model = "FREQ ~ Year",
               za_model = "~ 1",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 5.6
```

```
## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ Year
## Zero inflation:        ~1
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2171.1   2447.9  -1041.5   2083.1     3941
##
##
## Dispersion parameter for nbinom2 family (): 0.102
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.35574    0.66887  -5.017 5.25e-07 ***
## Year1968    -1.16605    1.24660  -0.935 0.349589
## Year1969    -0.48374    1.02610  -0.471 0.637330
## Year1970    -0.74689    1.01412  -0.736 0.461436
## Year1971    -1.31709    1.24059  -1.062 0.288387
## Year1972     0.83001    0.81883   1.014 0.310749
## Year1980     0.97825    0.81217   1.204 0.228403
## Year1982     2.04355    0.98362   2.078 0.037749 *
## Year1985     2.23359    0.86425   2.584 0.009754 **
## Year1986    -0.06199    1.05271  -0.059 0.953041
## Year1987     2.81341    0.84588   3.326 0.000881 ***
## Year1988     1.09695    0.85828   1.278 0.201222
## Year1989     2.55339    0.80930   3.155 0.001605 **
## Year1990     2.55723    0.77316   3.307 0.000941 ***
## Year1991     1.68176    0.77790   2.162 0.030623 *
## Year1992     0.15028    0.95566   0.157 0.875042
## Year1993     1.69223    0.77859   2.173 0.029746 *
## Year1995     1.12572    0.80800   1.393 0.163553
## Year1996    -0.93472    1.25768  -0.743 0.457353
## Year1997     1.80419    0.73729   2.447 0.014403 *
## Year1998     0.12691    0.89135   0.142 0.886780
## Year1999    -0.59551    0.91848  -0.648 0.516751
## Year2000     0.89188    0.82264   1.084 0.278291
## Year2001    -0.73022    1.01480  -0.720 0.471789
## Year2002    -0.71613    0.91458  -0.783 0.433617
## Year2003     1.21567    0.89799   1.354 0.175811
## Year2004     1.24120    0.87142   1.424 0.154346
## Year2005     0.71668    0.81239   0.882 0.377673
## Year2006    -0.02865    0.88347  -0.032 0.974127
## Year2007     1.23049    0.76360   1.611 0.107089
## Year2008     0.27996    0.83037   0.337 0.735999
## Year2009     0.50586    0.81934   0.617 0.536973
## Year2010     0.25179    0.80816   0.312 0.755374
```

```
## Year2011    -0.78744    1.01255  -0.778 0.436760
## Year2012     1.74630    0.75551   2.311 0.020810 *
## Year2013     0.72865    0.85571   0.852 0.394480
## Year2014     1.39967    0.78607   1.781 0.074977 .
## Year2015    -1.22923    1.24398  -0.988 0.323081
## Year2016     1.18668    0.84411   1.406 0.159771
## Year2018     2.52094    0.82216   3.066 0.002168 **
## Year2019     0.79079    0.86012   0.919 0.357894
## Year2020     1.58995    0.80832   1.967 0.049185 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -16.68    2965.99  -0.006    0.996

## Dispersal: 1.155

## Zero-altered negative binomial

## Model did not work

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Comparison 11: Intercept count model

```
compare_models(data = new_data2,
               models = c("zi", "za"),

               zic_model = "FREQ ~ 1",
               zi_model = "~ Year",

               zac_model = "FREQ ~ 1",
               za_model = "~ Year",

               r_dat = new_data2$FREQ) %>%
  try()

## Percent positive tows: 5.6

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ 1
## Zero inflation:        ~Year
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2189.5   2466.2  -1050.7   2101.5     3941
```

```
##
##
## Dispersion parameter for nbinom2 family (): 0.202
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.1438     0.1556  -7.351 1.97e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    1.38455    0.73156   1.893   0.0584 .
## Year1968       1.32508    1.27628   1.038   0.2992
## Year1969       0.57538    1.06773   0.539   0.5900
## Year1970       0.87068    1.05058   0.829   0.4072
## Year1971       1.48541    1.26885   1.171   0.2417
## Year1972      -0.98700    0.93706  -1.053   0.2922
## Year1980      -0.21401    0.95633  -0.224   0.8229
## Year1982      -3.40853    4.13938  -0.823   0.4103
## Year1985     -19.98986 8108.92912  -0.002   0.9980
## Year1986       0.07693    1.11218   0.069   0.9449
## Year1987     -19.98986 8108.92912  -0.002   0.9980
## Year1988      -1.09647    1.03670  -1.058   0.2902
## Year1989     -18.89767 6091.59144  -0.003   0.9975
## Year1990      -4.01808    4.32390  -0.929   0.3527
## Year1991      -1.78198    1.04358  -1.708   0.0877 .
## Year1992      -0.19177    1.02293  -0.188   0.8513
## Year1993      -2.06930    1.13146  -1.829   0.0674 .
## Year1995      -0.57824    0.94857  -0.610   0.5421
## Year1996       1.07628    1.29053   0.834   0.4043
## Year1997      -2.15867    1.03442  -2.087   0.0369 *
## Year1998      -0.16140    0.95146  -0.170   0.8653
## Year1999       0.70199    0.95732   0.733   0.4634
## Year2000      -0.56285    0.94670  -0.595   0.5522
## Year2001       0.85225    1.05150   0.810   0.4177
## Year2002       0.83662    0.95186   0.879   0.3794
## Year2003      -1.18514    1.12790  -1.051   0.2934
## Year2004      -2.19844    1.38495  -1.587   0.1124
## Year2005      -0.80415    0.91147  -0.882   0.3776
## Year2006       0.38202    0.97393   0.392   0.6949
## Year2007      -1.26968    0.90001  -1.411   0.1583
## Year2008      -0.12369    0.90543  -0.137   0.8913
## Year2009       0.06676    0.93343   0.072   0.9430
## Year2010       0.37206    0.91548   0.406   0.6844
## Year2011       1.65737    1.26211   1.313   0.1891
## Year2012       0.05651    0.93414   0.060   0.9518
## Year2013      -0.42363    0.97888  -0.433   0.6652
## Year2014      -1.00386    0.93973  -1.068   0.2854
## Year2015       1.39233    1.27302   1.094   0.2741
```

```
## Year2016      -0.66768    1.01306  -0.659   0.5098
## Year2018     -18.54199 5510.72791  -0.003   0.9973
## Year2019      -1.15009    1.00126  -1.149   0.2507
## Year2020      -0.88533    0.99406  -0.891   0.3731
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 0.908

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ 1
## Zero inflation:        ~Year
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2179.4   2456.2  -1045.7   2091.4     3941
##
##
## Dispersion parameter for truncated_nbinom2 family (): 9.33e-09
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -17.84    2180.96  -0.008    0.993
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.32023    0.58769   5.650 1.61e-08 ***
## Year1968     1.19063    1.16463   1.022  0.30663
## Year1969     0.49754    0.92542   0.538  0.59082
## Year1970     0.76574    0.92400   0.829  0.40726
## Year1971     1.34321    1.16396   1.154  0.24850
## Year1972    -0.73354    0.70639  -1.038  0.29907
## Year1980    -0.17392    0.77855  -0.223  0.82323
## Year1982    -1.61548    0.80052  -2.018  0.04359 *
## Year1985    -1.99110    0.69707  -2.856  0.00428 **
## Year1986     0.06413    0.92861   0.069  0.94495
## Year1987    -1.99108    0.69707  -2.856  0.00429 **
## Year1988    -0.80254    0.74934  -1.071  0.28417
## Year1989    -1.86798    0.67644  -2.761  0.00575 **
## Year1990    -1.68048    0.66123  -2.541  0.01104 *
## Year1991    -1.16847    0.67602  -1.728  0.08391 .
## Year1992    -0.15613    0.83235  -0.188  0.85121
## Year1993    -1.28731    0.66947  -1.923  0.05450 .
## Year1995    -0.45231    0.74616  -0.606  0.54439
## Year1996     0.95644    1.16588   0.820  0.41201
## Year1997    -1.32025    0.63911  -2.066  0.03885 *
## Year1998    -0.13181    0.77826  -0.169  0.86551
```

```
## Year1999     0.61165    0.82781   0.739  0.45998
## Year2000    -0.44103    0.74607  -0.591  0.55442
## Year2001     0.74880    0.92408   0.810  0.41776
## Year2002     0.73441    0.82733   0.888  0.37471
## Year2003    -0.85636    0.78528  -1.091  0.27549
## Year2004    -1.33431    0.71263  -1.872  0.06116 .
## Year2005    -0.61218    0.70552  -0.868  0.38556
## Year2006     0.32607    0.82910   0.393  0.69411
## Year2007    -0.90594    0.66665  -1.359  0.17416
## Year2008    -0.10135    0.74390  -0.136  0.89163
## Year2009     0.05564    0.77713   0.072  0.94292
## Year2010     0.31736    0.77586   0.409  0.68251
## Year2011     1.50809    1.16335   1.296  0.19486
## Year2012     0.04706    0.77717   0.061  0.95171
## Year2013    -0.33709    0.77977  -0.432  0.66552
## Year2014    -0.74435    0.70648  -1.054  0.29206
## Year2015     1.25448    1.16434   1.077  0.28129
## Year2016    -0.51688    0.78136  -0.662  0.50829
## Year2018    -1.86161    0.68457  -2.719  0.00654 **
## Year2019    -0.83532    0.72521  -1.152  0.24939
## Year2020    -0.66699    0.74797  -0.892  0.37254
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 3.3

## Error in UseMethod("logLik") :
##   no applicable method for 'logLik' applied to an object of class "NULL"
```

Finally!

Model selection for negative binomial model

```
model <- glmmTMB::glmmTMB(FREQ ~ Year + Month + Lat + Bottom.wt + Bottom.sa,
          data = new_data2,
          family = nbinom2
        )

null <- glmmTMB::glmmTMB(FREQ ~ Year,
          data = new_data2,
          family = nbinom2
        )

select_model(model = model, null_model = null)

## Forward model selection:

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ Year
```

```
##
## Final Model:
## FREQ ~ Year + Bottom.sa + Month + Bottom.wt
##
##
##           Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                              3942   2083.081 2169.081
## 2 + Bottom.sa  1 30.43117      3941   2052.650 2140.650
## 3     + Month  2 11.15548      3939   2041.495 2133.495
## 4 + Bottom.wt  1 12.62542      3938   2028.869 2122.869

## Bidirectional model selection:

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ Year + Month + Lat + Bottom.wt + Bottom.sa
##
## Final Model:
## FREQ ~ Year + Month + Bottom.wt + Bottom.sa
##
##
##    Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                         3937   2028.731 2124.731
## 2 - Lat  1 0.1377321      3938   2028.869 2122.869

## Backward model selection:

## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ Year + Month + Lat + Bottom.wt + Bottom.sa
##
## Final Model:
## FREQ ~ Year + Month + Bottom.wt + Bottom.sa
##
##
##    Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                         3937   2028.731 2124.731
## 2 - Lat  1 0.1377321      3938   2028.869 2122.869
```

Model selection for zero-inflated model

```
null <- glmmTMB::glmmTMB(FREQ ~ Year,
          ziformula = ~ 1,
          data = new_data2,
          family = nbinom2
        )

model <- glmmTMB::glmmTMB(FREQ ~ Year + Month + Lat + Bottom.wt + Bottom.sa,
```

```
          ziformula = ~ 1,
          data = new_data2,
          family = nbinom2
        )

select_model(model = model, null_model = null)

## Forward model selection:

## Stepwise Model Path 
## Analysis of Deviance Table
## 
## Initial Model:
## FREQ ~ Year
## 
## Final Model:
## FREQ ~ Year + Bottom.sa + Month + Bottom.wt
## 
## 
##          Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                              3941   2083.081 2171.081
## 2 + Bottom.sa  1 30.43117      3940   2052.650 2142.650
## 3     + Month  2 11.15548      3938   2041.495 2135.495
## 4 + Bottom.wt  1 12.62542      3937   2028.869 2124.869

## Bidirectional model selection:

## Stepwise Model Path 
## Analysis of Deviance Table
## 
## Initial Model:
## FREQ ~ Year + Month + Lat + Bottom.wt + Bottom.sa
## 
## Final Model:
## FREQ ~ Year + Month + Bottom.wt + Bottom.sa
## 
## 
##    Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                         3936   2028.731 2126.731
## 2 - Lat  1 0.1377319      3937   2028.869 2124.869

## Backward model selection:

## Stepwise Model Path 
## Analysis of Deviance Table
## 
## Initial Model:
## FREQ ~ Year + Month + Lat + Bottom.wt + Bottom.sa
## 
## Final Model:
## FREQ ~ Year + Month + Bottom.wt + Bottom.sa
```

```
##
##
##    Step Df  Deviance Resid. Df Resid. Dev      AIC
## 1                         3936   2028.731 2126.731
## 2 - Lat  1 0.1377319      3937   2028.869 2124.869
```

All the alternate models are 20+ AIC units away from the original model. Keeping the original – test with more covariates in count model?

Model selection for zero-adjusted model

```
model <- glmmTMB::glmmTMB(FREQ ~ Year + Month + Lat + Bottom.wt + Bottom.sa,
          ziformula = ~ Year,
          data = new_data2,
          family = truncated_nbinom2(link = "log")
        )

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

# full model does not converge

null <- glmmTMB::glmmTMB(FREQ ~ 1,
          ziformula = ~ Year,
          data = new_data2,
          family = truncated_nbinom2(link = "log")
        )

# forward selection only
  forward <- MASS::stepAIC(null,
    direction = "forward",
    scope = list(
      upper = FREQ ~ Year + Month + Lat + Bottom.wt + Bottom.sa,
      lower = null
    ),
    trace = FALSE
  )

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; non-positive-
definite
## Hessian matrix. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
  print(forward$anova)
```

```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## FREQ ~ 1
##
## Final Model:
## FREQ ~ Bottom.sa + Bottom.wt
##
##
##          Step Df Deviance Resid. Df Resid. Dev      AIC
## 1                              3941   2091.418 2179.418
## 2 + Bottom.sa  1 5.450093      3940   2085.968 2175.968
## 3 + Bottom.wt  1 2.391612      3939   2083.576 2175.576
```

Final comparison

```
compare_models(data = new_data2,
               nb_model = "FREQ ~ Year + Bottom.sa + Month + Bottom.wt",

               zic_model = "FREQ ~ Year + Bottom.sa + Month + Bottom.wt",
               zi_model = "~ 1",

               zac_model = "FREQ ~ Bottom.sa + Bottom.wt",
               za_model = "~ Year",

               r_dat = new_data2$FREQ) %>%
  try()
```

```
## Percent positive tows: 5.6

## Negative binomial

## model summary:
```

```
##  Family: nbinom2  ( log )
## Formula:          FREQ ~ Year + Bottom.sa + Month + Bottom.wt
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2122.9   2418.5  -1014.4   2028.9     3938
##
##
## Dispersion parameter for nbinom2 family (): 0.119
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.69676    0.69583  -5.313 1.08e-07 ***
## Year1968    -0.92755    1.26611  -0.733 0.463800
## Year1969    -0.46309    1.04264  -0.444 0.656930
## Year1970    -0.54484    1.03177  -0.528 0.597456
## Year1971    -0.99921    1.25544  -0.796 0.426088
## Year1972     1.23021    0.82761   1.486 0.137159
## Year1980     1.60846    0.82992   1.938 0.052614 .
## Year1982     2.89643    0.98905   2.928 0.003406 **
## Year1985     2.86859    0.87500   3.278 0.001044 **
## Year1986     0.59048    1.06200   0.556 0.578203
## Year1987     3.56455    0.85109   4.188 2.81e-05 ***
## Year1988     1.90001    0.88154   2.155 0.031137 *
## Year1989     3.04000    0.81133   3.747 0.000179 ***
## Year1990     3.19887    0.78617   4.069 4.72e-05 ***
## Year1991     1.97184    0.78752   2.504 0.012284 *
## Year1992     0.55259    0.97141   0.569 0.569455
## Year1993     2.43469    0.80417   3.028 0.002465 **
## Year1995     1.62621    0.82358   1.975 0.048317 *
## Year1996    -0.42159    1.27232  -0.331 0.740373
## Year1997     2.27177    0.74902   3.033 0.002422 **
## Year1998     0.65402    0.90384   0.724 0.469305
## Year1999    -0.48153    0.94470  -0.510 0.610247
## Year2000     1.45958    0.83585   1.746 0.080771 .
## Year2001    -0.28572    1.03331  -0.277 0.782155
## Year2002    -0.42790    0.93160  -0.459 0.646003
## Year2003     1.73222    0.90893   1.906 0.056679 .
## Year2004     1.81224    0.88175   2.055 0.039852 *
## Year2005     1.41438    0.83103   1.702 0.088764 .
## Year2006     0.47063    0.90046   0.523 0.601217
## Year2007     1.78805    0.78384   2.281 0.022540 *
## Year2008     0.80756    0.85422   0.945 0.344467
## Year2009     0.81685    0.83754   0.975 0.329416
## Year2010     0.64939    0.82881   0.784 0.433323
## Year2011     0.09783    1.03812   0.094 0.924919
## Year2012     1.82049    0.77669   2.344 0.019082 *
## Year2013     1.07033    0.87056   1.229 0.218896
## Year2014     1.70230    0.79389   2.144 0.032013 *
## Year2015    -0.73879    1.26091  -0.586 0.557933
```

```
## Year2016     1.90463    0.87199   2.184 0.028945 *
## Year2018     3.12572    0.83378   3.749 0.000178 ***
## Year2019     1.42403    0.88418   1.611 0.107275
## Year2020     2.20636    0.83227   2.651 0.008025 **
## Bottom.sa    0.49133    0.08983   5.469 4.51e-08 ***
## Month8      -0.03753    0.18974  -0.198 0.843192
## Month9      -1.17404    0.28688  -4.092 4.27e-05 ***
## Bottom.wt   -0.38419    0.10952  -3.508 0.000451 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 1.119

## Zero-inflated negative binomial

## model summary:

##  Family: nbinom2  ( log )
## Formula:          FREQ ~ Year + Bottom.sa + Month + Bottom.wt
## Zero inflation:        ~1
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2124.9   2426.8  -1014.4   2028.9     3937
##
##
## Dispersion parameter for nbinom2 family (): 0.119
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.69676    0.69583  -5.313 1.08e-07 ***
## Year1968    -0.92757    1.26611  -0.733 0.463794
## Year1969    -0.46312    1.04264  -0.444 0.656911
## Year1970    -0.54486    1.03177  -0.528 0.597446
## Year1971    -0.99923    1.25545  -0.796 0.426081
## Year1972     1.23017    0.82761   1.486 0.137171
## Year1980     1.60845    0.82992   1.938 0.052614 .
## Year1982     2.89645    0.98906   2.929 0.003406 **
## Year1985     2.86858    0.87500   3.278 0.001044 **
## Year1986     0.59046    1.06200   0.556 0.578219
## Year1987     3.56453    0.85108   4.188 2.81e-05 ***
## Year1988     1.89997    0.88154   2.155 0.031140 *
## Year1989     3.03997    0.81133   3.747 0.000179 ***
## Year1990     3.19885    0.78617   4.069 4.72e-05 ***
## Year1991     1.97180    0.78751   2.504 0.012286 *
## Year1992     0.55262    0.97140   0.569 0.569429
## Year1993     2.43467    0.80417   3.028 0.002466 **
## Year1995     1.62621    0.82358   1.975 0.048316 *
## Year1996    -0.42161    1.27232  -0.331 0.740364
## Year1997     2.27175    0.74902   3.033 0.002422 **
## Year1998     0.65397    0.90384   0.724 0.469343
```

```
## Year1999    -0.48145    0.94469  -0.510 0.610305
## Year2000     1.45957    0.83585   1.746 0.080773 .
## Year2001    -0.28577    1.03331  -0.277 0.782121
## Year2002    -0.42798    0.93161  -0.459 0.645947
## Year2003     1.73221    0.90893   1.906 0.056679 .
## Year2004     1.81223    0.88175   2.055 0.039853 *
## Year2005     1.41435    0.83103   1.702 0.088772 .
## Year2006     0.47061    0.90046   0.523 0.601234
## Year2007     1.78805    0.78384   2.281 0.022540 *
## Year2008     0.80754    0.85422   0.945 0.344481
## Year2009     0.81684    0.83754   0.975 0.329416
## Year2010     0.64940    0.82881   0.784 0.433317
## Year2011     0.09784    1.03811   0.094 0.924915
## Year2012     1.82048    0.77669   2.344 0.019083 *
## Year2013     1.07030    0.87056   1.229 0.218907
## Year2014     1.70231    0.79389   2.144 0.032012 *
## Year2015    -0.73880    1.26092  -0.586 0.557925
## Year2016     1.90461    0.87199   2.184 0.028946 *
## Year2018     3.12571    0.83378   3.749 0.000178 ***
## Year2019     1.42404    0.88418   1.611 0.107271
## Year2020     2.20638    0.83227   2.651 0.008025 **
## Bottom.sa    0.49134    0.08983   5.470 4.51e-08 ***
## Month8      -0.03753    0.18974  -0.198 0.843215
## Month9      -1.17404    0.28688  -4.092 4.27e-05 ***
## Bottom.wt   -0.38418    0.10952  -3.508 0.000452 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -17.05    3428.03  -0.005    0.996

## Dispersal: 1.119

## Zero-altered negative binomial

## model summary:

##  Family: truncated_nbinom2  ( log )
## Formula:          FREQ ~ Bottom.sa + Bottom.wt
## Zero inflation:        ~Year
## Data: data
##
##      AIC      BIC   logLik deviance df.resid
##   2175.6   2464.9  -1041.8   2083.6     3939
##
##
## Dispersion parameter for truncated_nbinom2 family (): 1.88e-09
##
## Conditional model:
##              Estimate Std. Error z value Pr(>|z|)
```

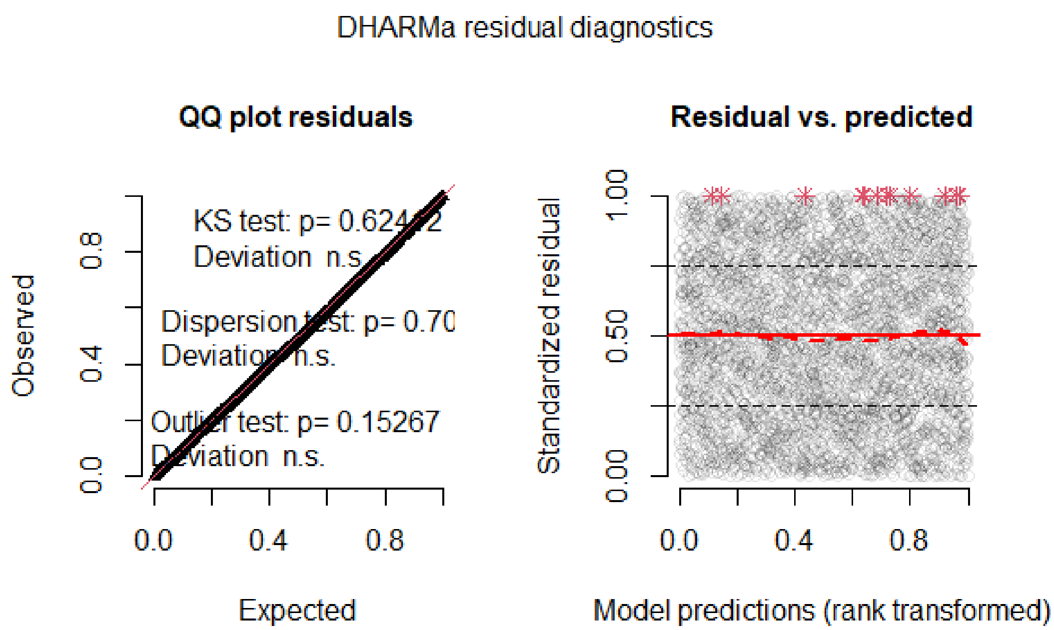
```
## (Intercept)  -19.5269  5150.3102  -0.004   0.9970
## Bottom.sa      0.3483     0.1546   2.254   0.0242 *
## Bottom.wt      0.2552     0.1658   1.540   0.1237
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.32023    0.58769   5.650 1.61e-08 ***
## Year1968     1.19063    1.16463   1.022  0.30663
## Year1969     0.49747    0.92540   0.538  0.59087
## Year1970     0.76575    0.92401   0.829  0.40726
## Year1971     1.34321    1.16397   1.154  0.24850
## Year1972    -0.73354    0.70639  -1.038  0.29907
## Year1980    -0.17393    0.77855  -0.223  0.82322
## Year1982    -1.61548    0.80052  -2.018  0.04359 *
## Year1985    -1.99109    0.69707  -2.856  0.00429 **
## Year1986     0.06417    0.92862   0.069  0.94491
## Year1987    -1.99109    0.69707  -2.856  0.00429 **
## Year1988    -0.80253    0.74934  -1.071  0.28418
## Year1989    -1.86798    0.67644  -2.761  0.00575 **
## Year1990    -1.68049    0.66123  -2.541  0.01104 *
## Year1991    -1.16847    0.67602  -1.728  0.08391 .
## Year1992    -0.15616    0.83235  -0.188  0.85118
## Year1993    -1.28731    0.66947  -1.923  0.05450 .
## Year1995    -0.45233    0.74615  -0.606  0.54437
## Year1996     0.95644    1.16588   0.820  0.41201
## Year1997    -1.32025    0.63911  -2.066  0.03885 *
## Year1998    -0.13182    0.77826  -0.169  0.86550
## Year1999     0.61155    0.82779   0.739  0.46004
## Year2000    -0.44103    0.74607  -0.591  0.55443
## Year2001     0.74880    0.92408   0.810  0.41776
## Year2002     0.73448    0.82735   0.888  0.37468
## Year2003    -0.85638    0.78528  -1.091  0.27547
## Year2004    -1.33431    0.71263  -1.872  0.06116 .
## Year2005    -0.61218    0.70552  -0.868  0.38556
## Year2006     0.32611    0.82911   0.393  0.69408
## Year2007    -0.90594    0.66665  -1.359  0.17416
## Year2008    -0.10135    0.74390  -0.136  0.89163
## Year2009     0.05566    0.77713   0.072  0.94291
## Year2010     0.31736    0.77586   0.409  0.68251
## Year2011     1.50808    1.16335   1.296  0.19486
## Year2012     0.04707    0.77718   0.061  0.95170
## Year2013    -0.33707    0.77977  -0.432  0.66555
## Year2014    -0.74435    0.70648  -1.054  0.29206
## Year2015     1.25448    1.16434   1.077  0.28129
## Year2016    -0.51687    0.78137  -0.661  0.50830
## Year2018    -1.86161    0.68457  -2.719  0.00654 **
## Year2019    -0.83532    0.72522  -1.152  0.24939
## Year2020    -0.66699    0.74797  -0.892  0.37254
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Dispersal: 2.599

##             dAIC df
## neg_binom   0.0 47
## zero_infl   2.0 48
## zero_adj   52.7 46
```

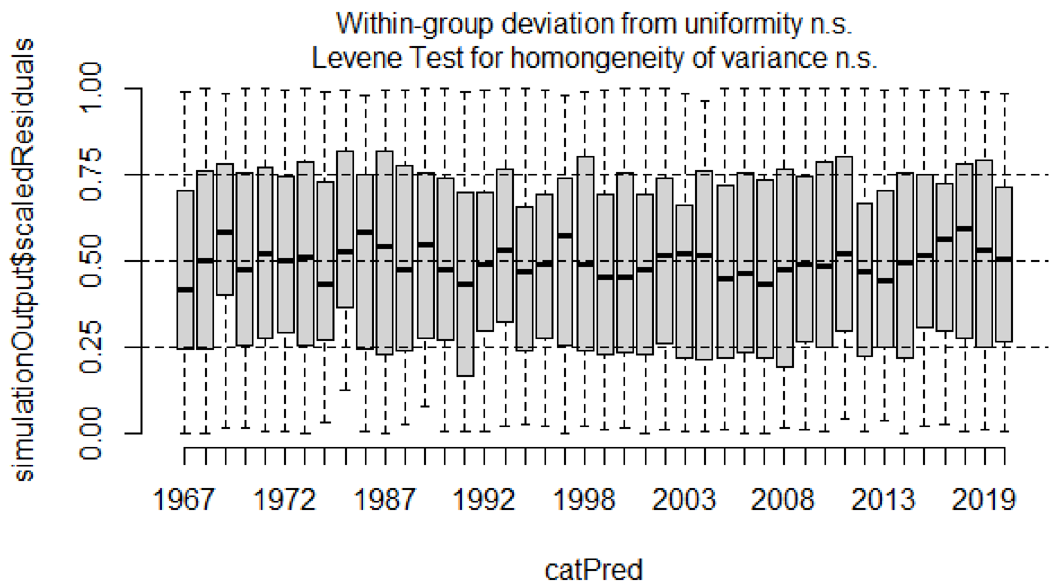
Negative binomial is the winner!

Step 3: DHARMA simulations

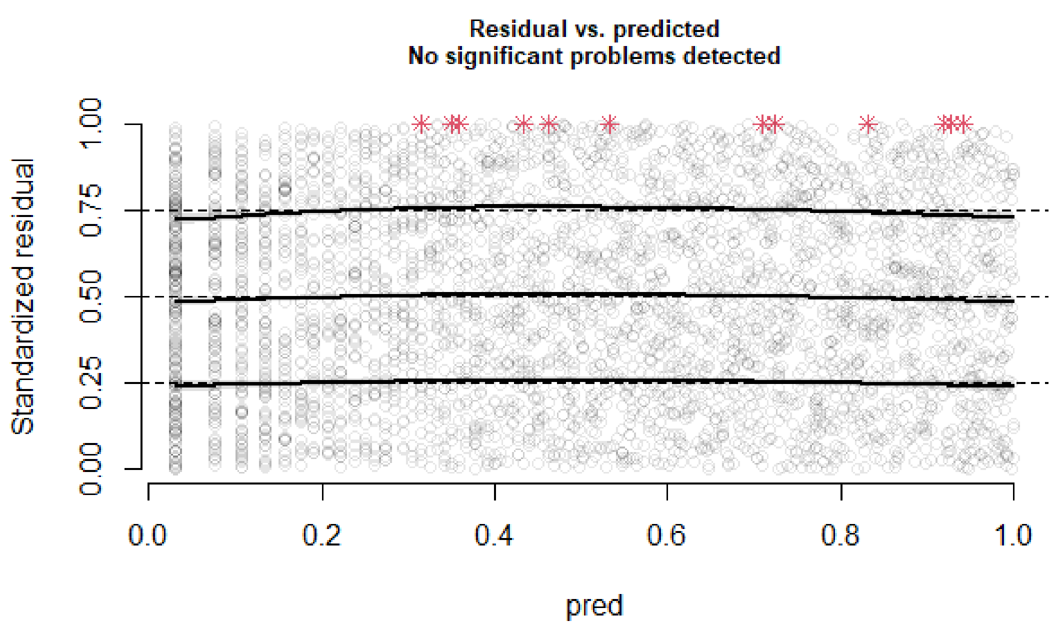
```
model <- glmmTMB(FREQ ~ Year + Bottom.sa + Month + Bottom.wt,
                    data = new_data2,
                    family = nbinom2)

sim <- simulateResiduals(model, quantreg = TRUE)
plot(sim)
```



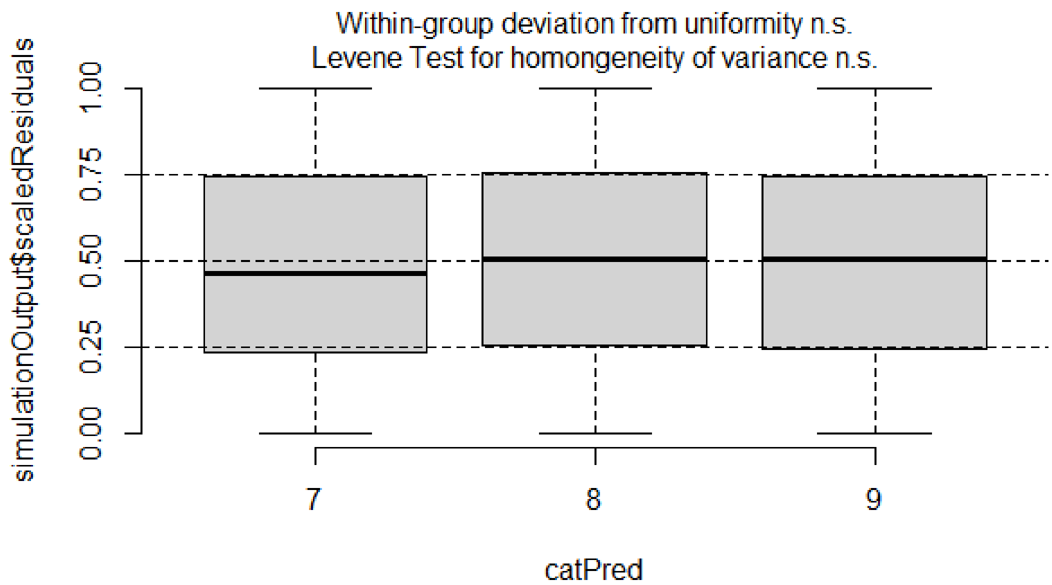

```
# Compare residuals vs. factors
plotResiduals(sim, quantreg = TRUE, form=new_data2$Year)
```



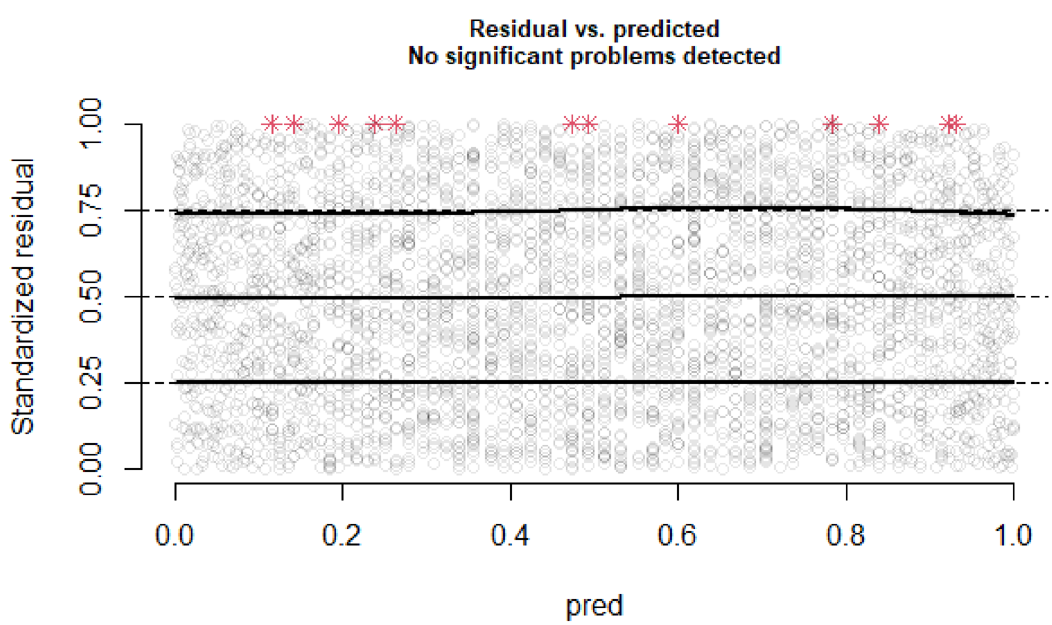

```
plotResiduals(sim, quantreg = TRUE, form=new_data2$Bottom.sa)
```




```
plotResiduals(sim, quantreg = TRUE, form=new_data2$Month)
```




```
plotResiduals(sim, quantreg = TRUE, form=new_data2$Bottom.wt)
```



Step 4: Plot predictions

Add 2021 data

```
indata <- readxl::read_excel(here::here("data/2021 
Data/2021_bluefish_data_NEAMAP_April_2022.xlsx"),
                                    sheet = "SurveyData_SBSS",
```

```
                              col_types = c(rep("guess", 17), "numeric"))
head(indata, n = 2)
```

SampleID	Cruiseno	Stanum	State	Program	Gear	Haul	Year	Month	Day	Round	Bf.count	Area	Lat	Lon	Bottom.wt	Bottom.sa	Bottom.do
SB670707RA44	SB670707RA	44	VA	Virginia Striped Bass Seine Survey	Beach Seine	1	1967	7	7	1	1	RA	38.02	-76.92	25.6	1.6	NA
SB670811RA 8	SB670811RA	8	VA	Virginia Striped Bass Seine Survey	Beach Seine	1	1967	8	11	3	1	RA	37.61	-76.44	24.5	15.0	NA

```
# modify data formatting
indata <- indata %>%
  dplyr::rename(FREQ = Bf.count,
                Stratum = Stanum) %>%
  dplyr::mutate(Haul = as.factor(Haul),
                Round = as.factor(Round),
                Stratum = as.factor(Stratum)) %>%
  dplyr::select(-c(State, Program, Gear)) # same value for all data
head(indata, n = 2)
```

SampleID	Cruiseno	Stratum	Haul	Year	Month	Day	Round	FREQ	Area	Lat	Lon	Bottom.wt	Bottom.sa	Bottom.do
SB670707RA	SB670707RA	44	1	1967	7	7	1	1	RA	38.0	-76	25.6	1.6	NA

44										2	.92			
SB670811RA8	SB670811RA	8	1	1967	8	11	3	1	RA	37.61	-76.44	24.5	15.0	NA

```r
# remove anything not Haul 1
length1 <- nrow(indata)
new_data <- indata %>%
  dplyr::filter(Haul == 1) %>%
  dplyr::select(-Haul) # remove because there is only 1 value now
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")

## 3423 rows removed

# remove June
length1 <- nrow(new_data)
new_data <- new_data %>%
  dplyr::filter(Month != 6)
rows_removed <- length1 - nrow(new_data)
message(rows_removed, " rows removed")

## 135 rows removed

# not sure what Bottom.do is, but it has a lot of missing data, so removing
new_data <- new_data %>%
  dplyr::select(-Bottom.do)

new_data <- remove_missing(new_data,
                           f_col = "FREQ",
                           t_col = "Year")

## Data before removing missing values:

##    SampleID           Cruiseno            Stratum           Year     
##  Length:7550        Length:7550        44     : 467   Min.   :1967  
##  Class :character   Class :character   50     : 411   1st Qu.:1990  
##  Mode  :character   Mode  :character   28     : 396   Median :2000  
##                                        37     : 388   Mean   :1998  
##                                        41     : 386   3rd Qu.:2010  
##                                        36     : 334   Max.   :2021  
##                                        (Other):5168                 
##      Month            Day          Round        FREQ              Area          
##  Min.   :7.000   Min.   : 1.00   0:   1   Min.   : 0.00000   Length:7550       
##  1st Qu.:7.000   1st Qu.: 8.00   1:1377   1st Qu.: 0.00000   Class 
:character  
##  Median :8.000   Median :14.00   2:1589   Median : 0.00000   Mode  
:character  
##  Mean   :7.825   Mean   :15.21   3:1562   Mean   : 0.05483                     
```

```
##  3rd Qu.:8.000   3rd Qu.:23.00   4:1464   3rd Qu.: 0.00000
##  Max.   :9.000   Max.   :31.00   5:1383   Max.   :19.00000
##                                  6: 174
##       Lat             Lon          Bottom.wt       Bottom.sa
##  Min.   :37.04   Min.   :-78.40   Min.   : 0.00   Min.   : 0.000
##  1st Qu.:37.30   1st Qu.:-77.04   1st Qu.:26.60   1st Qu.: 0.100
##  Median :37.56   Median :-76.90   Median :28.10   Median : 0.700
##  Mean   :37.58   Mean   :-71.50   Mean   :27.75   Mean   : 3.934
##  3rd Qu.:37.74   3rd Qu.:-76.74   3rd Qu.:29.60   3rd Qu.: 6.500
##  Max.   :38.26   Max.   : 77.40   Max.   :39.90   Max.   :25.000
##  NA's   :74      NA's   :74       NA's   :24      NA's   :42

## Data after removing missing values:

##    SampleID           Cruiseno            Stratum          Year
##  Length:7432        Length:7432        44     : 463   Min.   :1967
##  Class :character   Class :character   50     : 409   1st Qu.:1990
##  Mode  :character   Mode  :character   28     : 396   Median :2000
##                                        37     : 385   Mean   :1998
##                                        41     : 383   3rd Qu.:2010
##                                        36     : 334   Max.   :2021
##                                        (Other):5062
##      Month            Day         Round         FREQ              Area
##  Min.   :7.000   Min.   : 1.00   0:   1   Min.   : 0.00000   Length:7432
##  1st Qu.:7.000   1st Qu.: 8.00   1:1365   1st Qu.: 0.00000   Class
:character
##  Median :8.000   Median :15.00   2:1567   Median : 0.00000   Mode
:character
##  Mean   :7.824   Mean   :15.23   3:1531   Mean   : 0.05544
##  3rd Qu.:8.000   3rd Qu.:23.00   4:1438   3rd Qu.: 0.00000
##  Max.   :9.000   Max.   :31.00   5:1356   Max.   :19.00000
##                                  6: 174
##       Lat             Lon          Bottom.wt       Bottom.sa
##  Min.   :37.04   Min.   :-78.40   Min.   : 0.00   Min.   : 0.000
##  1st Qu.:37.30   1st Qu.:-77.04   1st Qu.:26.60   1st Qu.: 0.100
##  Median :37.56   Median :-76.90   Median :28.10   Median : 0.800
##  Mean   :37.58   Mean   :-71.47   Mean   :27.74   Mean   : 3.972
##  3rd Qu.:37.74   3rd Qu.:-76.74   3rd Qu.:29.60   3rd Qu.: 6.600
##  Max.   :38.26   Max.   : 77.40   Max.   :39.90   Max.   :25.000
##

## 118 datapoints removed (1.6% of data removed)

## Data after removing years with no catch:

##    SampleID           Cruiseno            Stratum          Year
##  Length:6798        Length:6798        44     : 416   Min.   :1967
##  Class :character   Class :character   50     : 369   1st Qu.:1991
##  Mode  :character   Mode  :character   28     : 364   Median :2001
##                                        37     : 352   Mean   :1999
##                                        41     : 344   3rd Qu.:2010
```

```
##                                         36      : 308   Max.   :2021
##                                         (Other):4645
##      Month            Day          Round          FREQ              Area
##  Min.   :7.000   Min.   : 1.00   0:   1   Min.   : 0.00000   Length:6798
##  1st Qu.:7.000   1st Qu.: 8.00   1:1241   1st Qu.: 0.00000   Class
:character
##  Median :8.000   Median :14.00   2:1409   Median : 0.00000   Mode
:character
##  Mean   :7.824   Mean   :15.09   3:1391   Mean   : 0.06061
##  3rd Qu.:8.000   3rd Qu.:22.00   4:1342   3rd Qu.: 0.00000
##  Max.   :9.000   Max.   :31.00   5:1259   Max.   :19.00000
##                                  6: 155
##       Lat             Lon            Bottom.wt       Bottom.sa
##  Min.   :37.04   Min.   :-78.40   Min.   : 0.00   Min.   : 0.000
##  1st Qu.:37.30   1st Qu.:-77.04   1st Qu.:26.60   1st Qu.: 0.100
##  Median :37.56   Median :-76.90   Median :28.10   Median : 0.800
##  Mean   :37.58   Mean   :-74.49   Mean   :27.82   Mean   : 4.005
##  3rd Qu.:37.74   3rd Qu.:-76.78   3rd Qu.:29.60   3rd Qu.: 6.700
##  Max.   :38.26   Max.   : 77.40   Max.   :39.90   Max.   :25.000
##

## 634 datapoints removed (8.5% of data removed)

new_data <- new_data %>%
  dplyr::mutate(Lon = -abs(Lon)) %>%
  dplyr::filter(Bottom.wt > 0,
                Bottom.sa > 0.2) # several stations with 0 salinity

new_data2 <- standardize_data(new_data,
                             cols = c("Year", "Month", "Stratum", "Day",
                                      "Round", "Lat", "Lon", "Area",
                                      "SampleID", "Cruiseno"),
                             f_col = "FREQ")

## SampleID has been changed to factor

## Cruiseno has been changed to factor

## Stratum is already a factor

## Year has been changed to factor

## Month has been changed to factor

## Day has been changed to factor

## Round is already a factor

## Area has been changed to factor

## Lat has been changed to factor
```

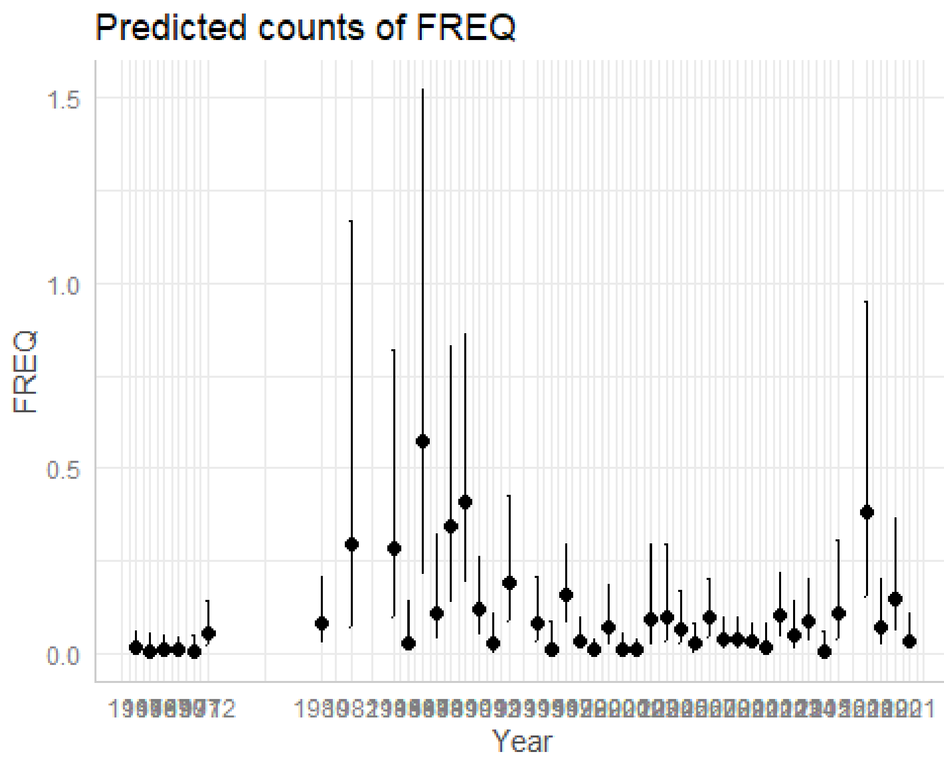
```
## Lon has been changed to factor

## Bottom.wt has been standardized

## Bottom.sa has been standardized

## Data standardized!

# change Lat and Lon back to numeric
new_data2 <- new_data2 %>%
  dplyr::mutate(Lat = as.numeric(Lat),
                Lon = as.numeric(Lon))
head(new_data2, n = 2)
```

SampleID	Cruiseno	Stratum	Year	Month	Day	Round	FREQ	Area	Lat	Lon	Bottom.wt	Bottom.sa
SB670707RA 44	SB670707RA	44	1967	7	7	1	1	RA	86	47	-0.89	-0.85
SB670811RA 8	SB670811RA	8	1967	8	11	3	1	RA	67	113	-1.34	1.46

```
new_data2 <- new_data2 %>%
  dplyr::select(-Round)

model <- glmmTMB(FREQ ~ Year + Bottom.sa + Month + Bottom.wt,
                    data = new_data2,
                    family = nbinom2)
```

ggemmeans method

```
index.pred <- ggeffects::ggemmeans(model,
                                   terms="Year",
                                   type="fixed")
plot(index.pred)
```



Bootstrap method

```
source(here::here("scripts/bootstrap_functions_AT_lnE2.R"))

index.pred <- boot.ZI_AT(best = model,
                         type_z = "nb",
                         nboots = 1000,
                         ycol = "Year"
                         )

## starting fxn

## year name fixed

## starting boots

## session planned

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
```

```
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
```

```
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')
```

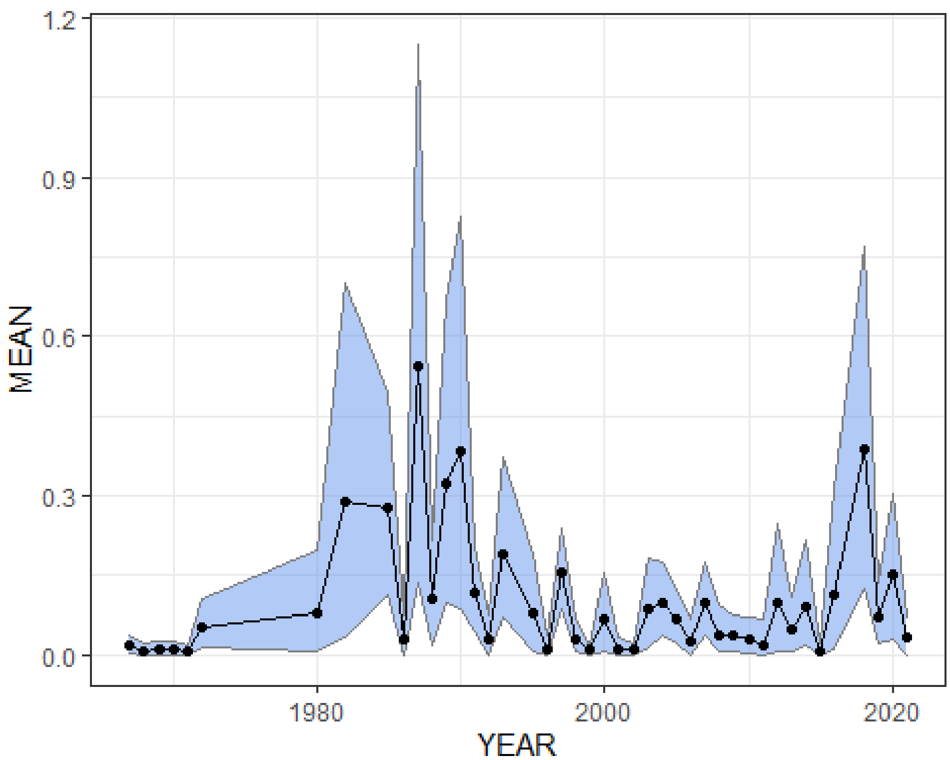
```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; extreme or very
small
## eigenvalues detected. See vignette('troubleshooting')
```

```
## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## Warning in fitTMB(TMBStruc): Model convergence problem; singular
convergence
## (7). See vignette('troubleshooting')

## finished boots

## Number of successful predictions: 933
```

```
index.pred %>%
  plot_boot()
```



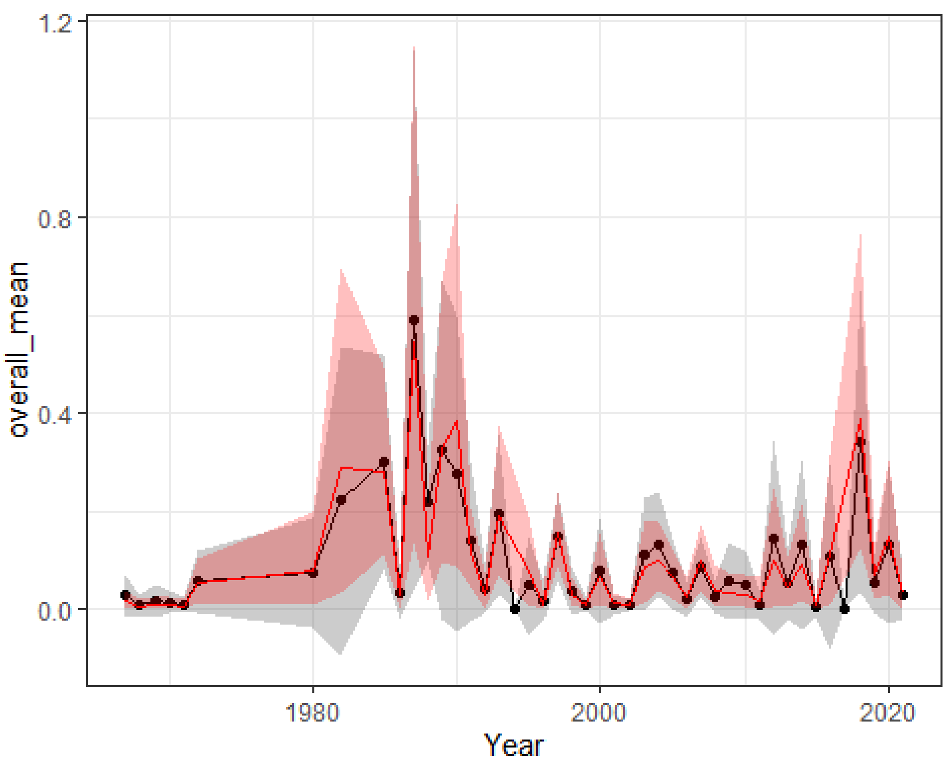
Step 5: Save index

```
index.out <- as.data.frame(index.pred)
index.out$CV <- abs(index.out$SE/index.out$MEAN)

names(index.out)[c(1,2)] <- c("Year", "Frequency")

write.csv(index.out,
          here::here("reports/updated/VA_SBSS_index5.csv"),
          row.names=FALSE)
```

Develop nominal index

```
# not sure if/how to standardize for effort, so skipping that for now
# also not sure what stratum weights are, so just weighting everything
equally for now
# are system/river columns equivalent to strata?

# limit to just years when bluefish were observed
# and the same months used for model data
nominal_data <- indata %>%
  dplyr::group_by(Year) %>%
  dplyr::mutate(year_total = sum(FREQ, na.rm = TRUE)) %>%
  dplyr::filter(year_total > 0,
                Month %in% unique(new_data2$Month),
                Bottom.wt > 0,
                Bottom.sa > 0.2,
                Haul == 1
                ) %>%
  dplyr::ungroup() %>%
  dplyr::group_by(Year, Stratum) %>%
  dplyr::summarise(mean_freq = mean(FREQ),
                   var_freq = var(FREQ),
                   n_obs = length(SampleID)
                   ) %>%
  dplyr::ungroup() %>%
  dplyr::group_by(Year)  %>%
  dplyr::mutate(mean_var = mean(var_freq, na.rm = TRUE),
                var_freq = ifelse(is.na(var_freq), mean_var, var_freq),
                n_strata = length(unique(Stratum)),
                weight = 1/n_strata) %>%
  dplyr::ungroup() %>%
  dplyr::group_by(Year)  %>%
  dplyr::summarise(overall_mean = sum(weight * mean_freq),

                   overall_var =  sum(weight^2 * var_freq / n_obs),

                   overall_sd = sqrt(overall_var),
                   .groups = "keep") %>%
  dplyr::ungroup()

## `summarise()` has grouped output by 'Year'. You can override using the
`.groups`
## argument.

write.csv(nominal_data,
          here::here("reports/updated/VA_SBSS_index_nominal4.csv"),
          row.names=FALSE)

ggplot2::ggplot(nominal_data,
                ggplot2::aes(x = Year)) +
  ggplot2::geom_ribbon(ggplot2::aes(ymin = overall_mean - 1.96*overall_sd,
```

```
                                ymax = overall_mean + 1.96*overall_sd),
                   alpha = 0.25) +
ggplot2::geom_line(ggplot2::aes(y = overall_mean)) +
ggplot2::geom_point(ggplot2::aes(y = overall_mean)) +
ggplot2::theme_bw() +
ggplot2::geom_line(data = index.out,
                 ggplot2::aes(x = Year %>%
                                as.character() %>%
                                as.numeric(),
                              y = Frequency),
           color = "red",
           inherit.aes = FALSE) +
ggplot2::geom_ribbon(data = index.out,
                   ggplot2::aes(x = Year %>%
                                as.character() %>%
                                as.numeric(),
                                ymin = LCI,
                                ymax = UCI),
                   alpha = 0.25,
                   fill = "red",
                   inherit.aes = FALSE)
```



The confidence intervals go below zero. Should we log-transform the data?