Panic in the Disko Bay: Estimating Greenland Bowhead Whale Detection and Abundance with Distance Sampling

Misha Tseitlin, Louise Blackman, Dometa Tuomaite

15 February 2024

Project 1: Distance Sampling

School of Mathematics and Statistics

in partial fulfilment of the requirements for
MT5751: Estimating Animal Abundance and Biodiversity

Abstract

Bowhead whales (*Balaena mysticetus*) aggregate each spring to mate in west Greenland's Disko Bay (Qeqertarsuaq), providing a rare chance to measure this highly mobile species' abundance. Distance sampling (DS) uses likelihood estimation to generate estimates about abundance while accounting for imperfect detectability and has previously estimated bowhead whale numbers around Qeqertarsuaq. We analyse existing aerial survey data with simple DS methods to test the robustness of estimates without incorporating whale surfacing behaviour.[1] As data appears to violate DS assumptions, we compare different data processing approaches and their impact on the size effect and abundance estimates. Our simple DS, even after including all possible covariates, underestimates abundance relative to the original approach. Though DS remains a powerful tool to understand population sizes, we must approach cases of small sample sizes and simplified secondary analyses with caution.

Introduction

Annual springtime bowhead whale aggregations in Qeqertarsuaq allow analysis of both sub-population stock rebound after whaling moratoria and estimates of baseline population before wider climate change impacts in the Arctic. Using existing visual aerial surveys, we conduct a simplified replication of existing DS analysis using only covariates and adjustment terms (e.g., cosine and polynomial expansions) to test method robustness and data suitability for more accessible methods for understanding bowhead whale detection, abundance, and impacts of group size.

Methods

After splitting Qeqertarsuaq into 16 strata and 41 systematically-placed east-west transects (combined length–4445km), surveying used planes with two observation platforms to record declination angles and times of the first abeam whale sighting—generating perpendicular distances for DS—Beaufort sea state, and whale group size. Meaning resolved any differences in size or declination between observers. Authors left-truncated distances by 100 m due to obscured view close to the transect line, so we retain 58 observations of 74 whales[1].

Distance sampling estimates animal abundance using the perpendicular distance between the transect and observed animals. A detection function $g(x)$ is fitted to calculate the probability of detecting an animal at a distance x. We use this detection probability to account for unobserved animals. Covariates can be added to the scale parameter in half-normal and hazard-rate detection functions to model differences in detectability[2]. As $\sigma > 0$, we use a log link function:

$$\sigma(\mathbf{z}_i) = exp(\alpha_0 + \sum_{j=1}^{n} \alpha_j z_{ij})$$

where $\mathbf{z}_i = (z_{i1}, ..., z_{in})$ is the vector of covariates for animal i and $\alpha_0, , , \alpha_n$ are effect parameters to be estimated.

An initial plot of the detection distances shows possibly some line avoidance, as more whales were detected further away from the plane than on the transect. This could violate the assumption in distance sampling that animals are distributed independently with respect to the transect, or the assumption that all animals on the line are detected, depending on whether the whales are hiding or moving away[2]. Equally, this could be stochastic variation as there is only a difference of 4 observations between the 0-0.2 bin and the 0.4-0.6 bin. We considered 3 approaches to deal
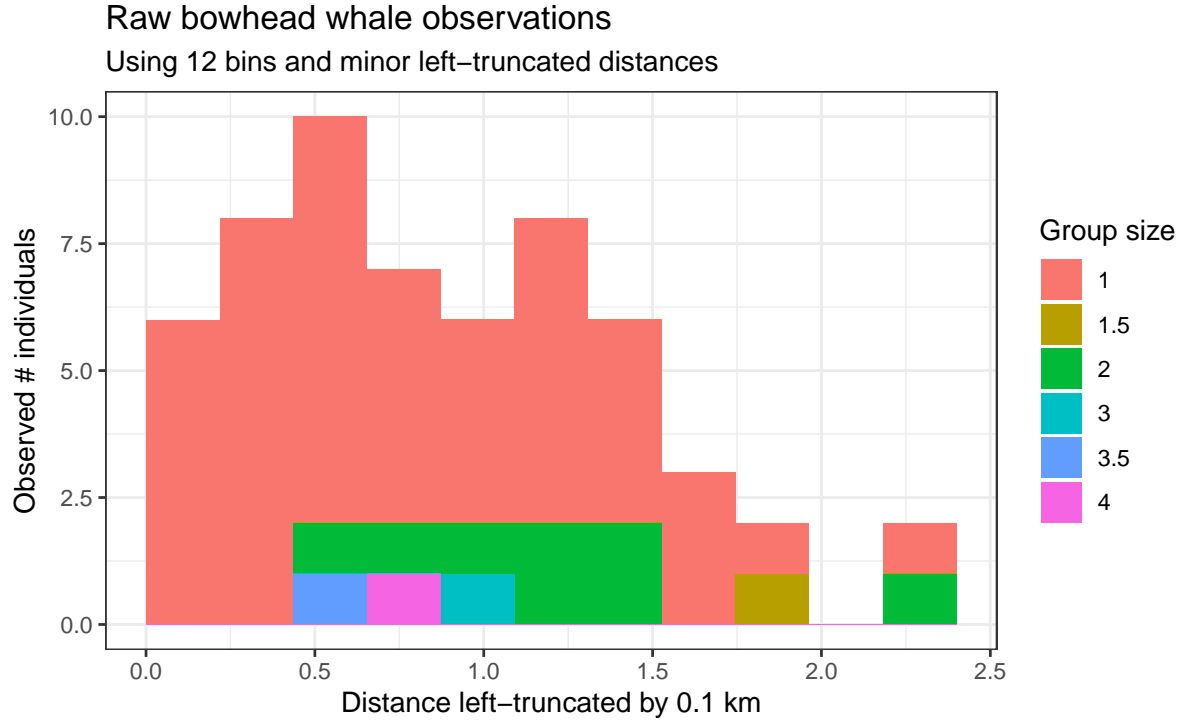
Figure 1: Original bowhead whale detection distances

with this problem. Firstly, treating the difference in detection as random variation and fitting detection functions as normal. This preserves all observations but will result in overestimating the detection probability (therefore underestimating abundance) if detection is not certain on the line. Another option is to bin the distances into 0-0.6km, 0.6-0.9km, . . . , 2.1-2.4km categories so that there are more observations in the first group than the others. This assumes the whales moved away from the observers, so those that would have been detected on the transect were instead seen further away. This method deals with the potentially violated assumptions, but in grouping the data we lose information.

Our final option was left-truncating the distances up to the peak in detections at 0.64km. This method ensures the monotonic decline in detections with increasing distance, but again loses information and assumes that detection is certain 0.64km from the plane. The monotonically decreasing detection allowed us to successfully fit complex models with multiple covariates and adjustment terms, unlike the other methods. However, given the small sample size these models likely suffer from over-fitting.

We used the `Distance` package in `R` to fit models using a maximum likelihood approach. For each data set (unedited, binned, truncated) we fitted half-normal (HN) and hazard rate (HR) detection functions with all combinations of adjustment terms (cosine, Hermite, polynomial) and covariates. Aikake Information Criterion (AIC) compares between models using the same manipulated data.

Results

Given raw data set the best model selected by AIC was half-normal detection function with size covariate and no adjustments. It estimates whale abundance as 212 with confidence interval (112, 401). However, size is not significant, therefore the model without covariates gives similar estimate of whale abundance (229) with confidence interval (112, 401).
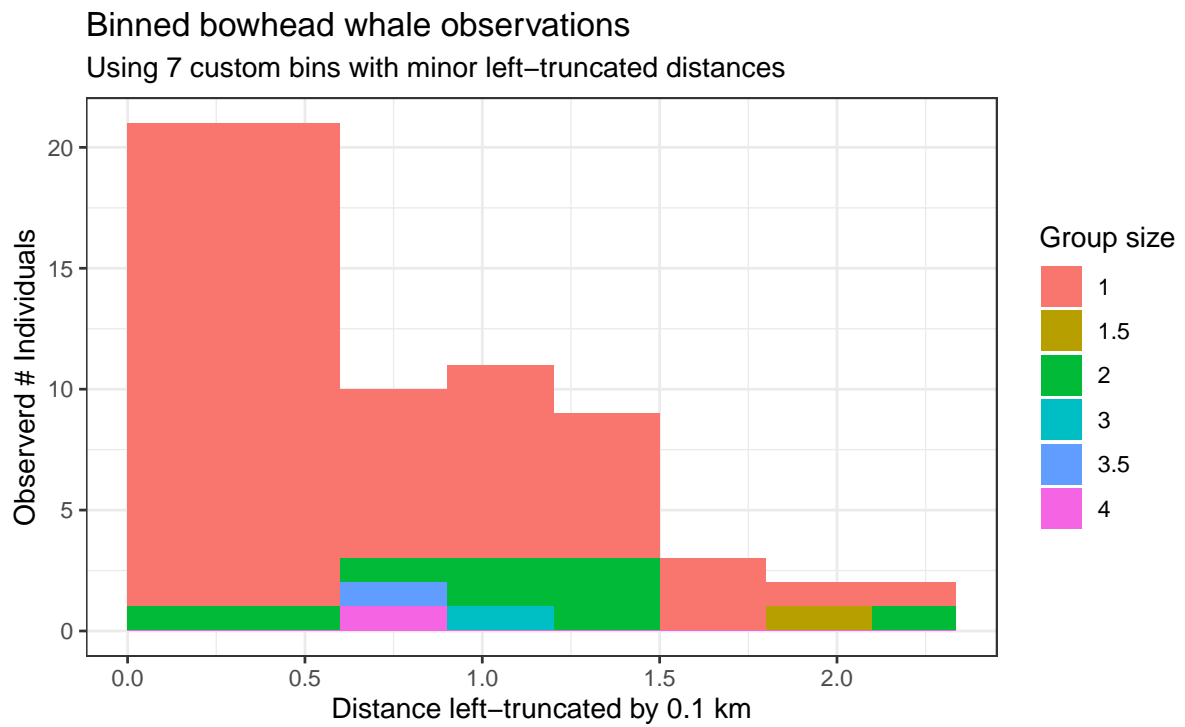
Figure 2: Original owhead whale detection distances with custom bins

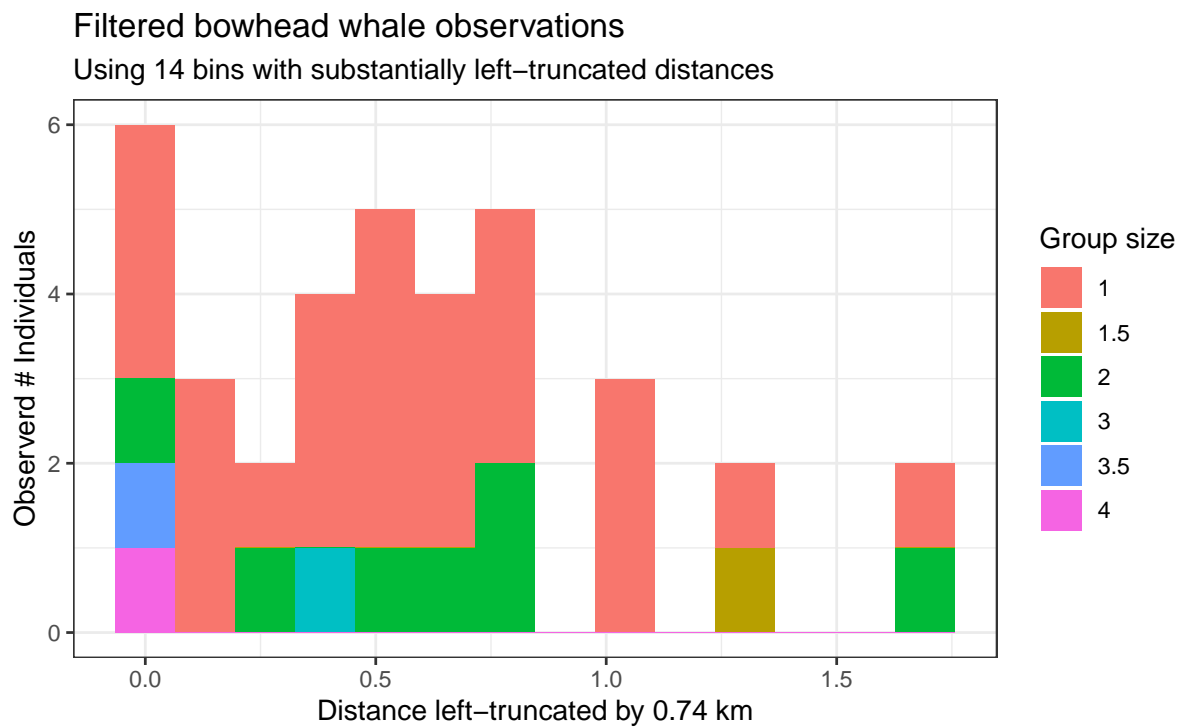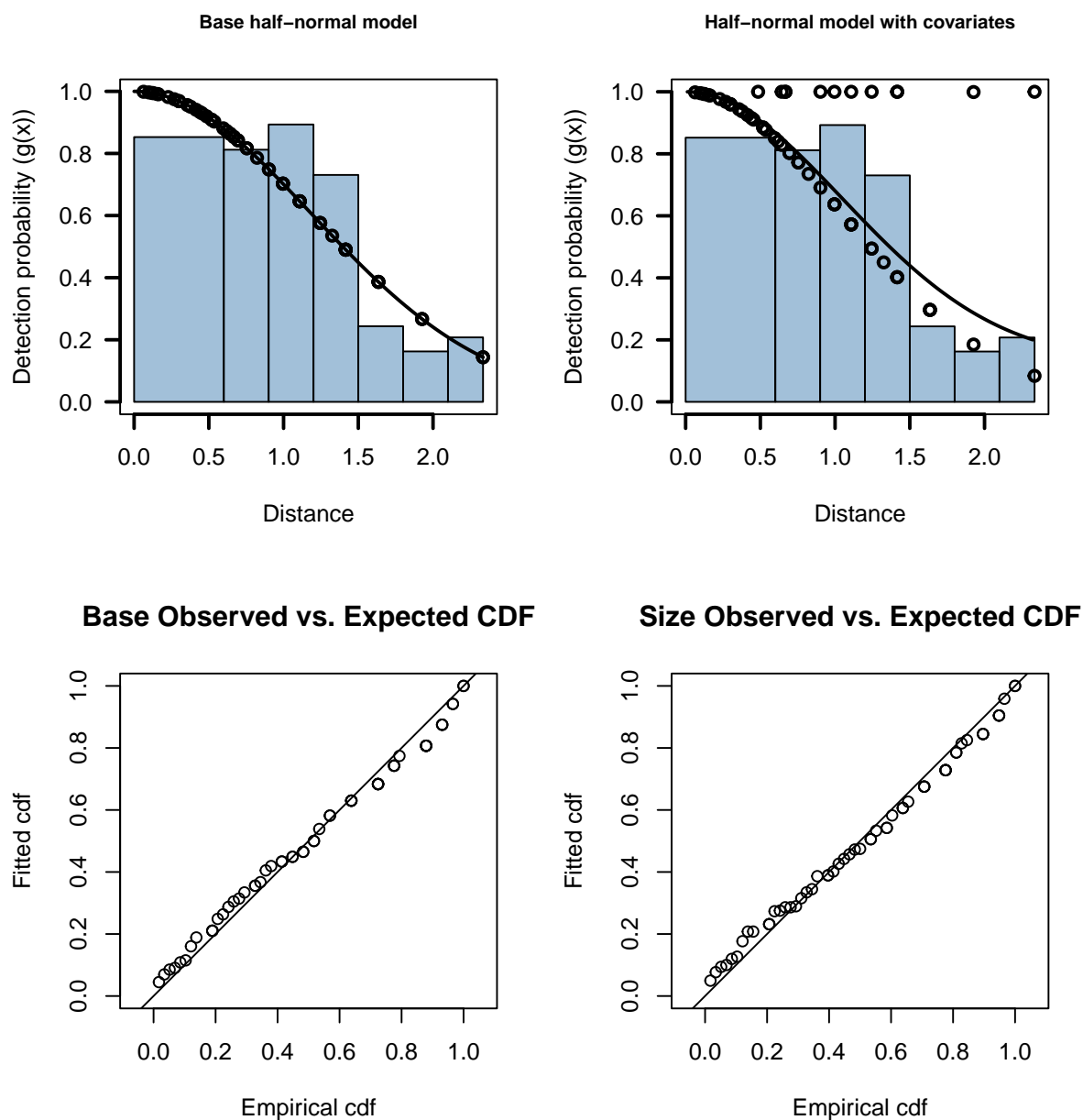Figure 3: More truncated bowhead whale detection distances

Base half−normal model

Half−normal model with covariates

Base Observed vs. Expected CDF

Size Observed vs. Expected CDF

Figure 4: Detection functions using original data

Table 1: Estimated abundance from best-performing models (by AIC)

Detection.Function	Data	Size.Covariate	Estimate	SE	Lower.CI	Upper.CI	AIC
Half Normal	Base	Yes	212.07	65.62	112.13	401.06	86.03
Half Normal	Base	No	228.82	71.45	122.00	429.14	86.32
Half Normal	Binned	Yes	212.84	66.14	112.32	403.31	197.58
Half Normal	Binned	No	229.54	72.22	121.94	432.07	197.91
Hazard Rate	Truncated	Yes	1803.36	1346.22	461.56	7045.88	27.20
Half Normal	Truncated	No	519.51	90.74	363.25	743.00	28.43
Half Normal	Truncated	Yes	608.75	154.26	356.93	1038.20	28.73

Table 2: Estimates (and errors) of size on detectability

Model	Size.Coefficient	SE	Significant.
Base HN + size	6.16	5046.40	No
Binned HN + size	5.64	305.31	No
Truncated HN + size	-0.51	0.26	Yes

Using binning, AIC selects half-normal detection function with no adjustments and size. Like the model using raw data, it gives estimate of abundance of 213 and confidence interval (112, 403). Size is not significant.

The model chosen using truncated data is quite different. The best model is hazard rate with size covariate, with estimate of 1803 whales and confidence interval (461, 7045). It is worth noting that size is significant. Nonetheless, given huge standard error, there's a lot of uncertainty, therefore it is likely that the model largely overestimates.

Although the hazard rate model has lowest AIC, it overestimates whale abundance and the sharp decline is not realistic. This might be due to sample size and extrapolation to the area we have no data about. Therefore, it might be best to use half-normal detection function with size covariate to get the best estimate of whale abundance.

Discussion

While conducting the survey, several assumptions were violated. The area beneath the aircraft was not visible to surveyors, therefore, we cannot assume that certain detectability on the transect line. To account for this, passive monitoring such as cameras could be used. Moreover, since whales are never stationary, as they run or hide when noticing observers, the assumption of object detection at their initial location is violated. This results in less observations within close distance and we are likely underestimating the true whale abundance due to imperfect detection. Whales are underwater for periods of time, further creating availability bias. To address this, hidden Markov model could be used[3]. HMM copes with a small data by fitting a more restrictive model, which incorporates information on how the aircraft and whales move.

Violation of these assumptions bias estimates of whale abundance – models will underestimate the real number of whales; therefore, truncation of data is needed. However, truncation reduces survey area by 30% and since data is quite unbalanced and sample size is small, it is not a realistic approximation of how whale abundance and any variation seen in estimates is possibly due to stochasticity.
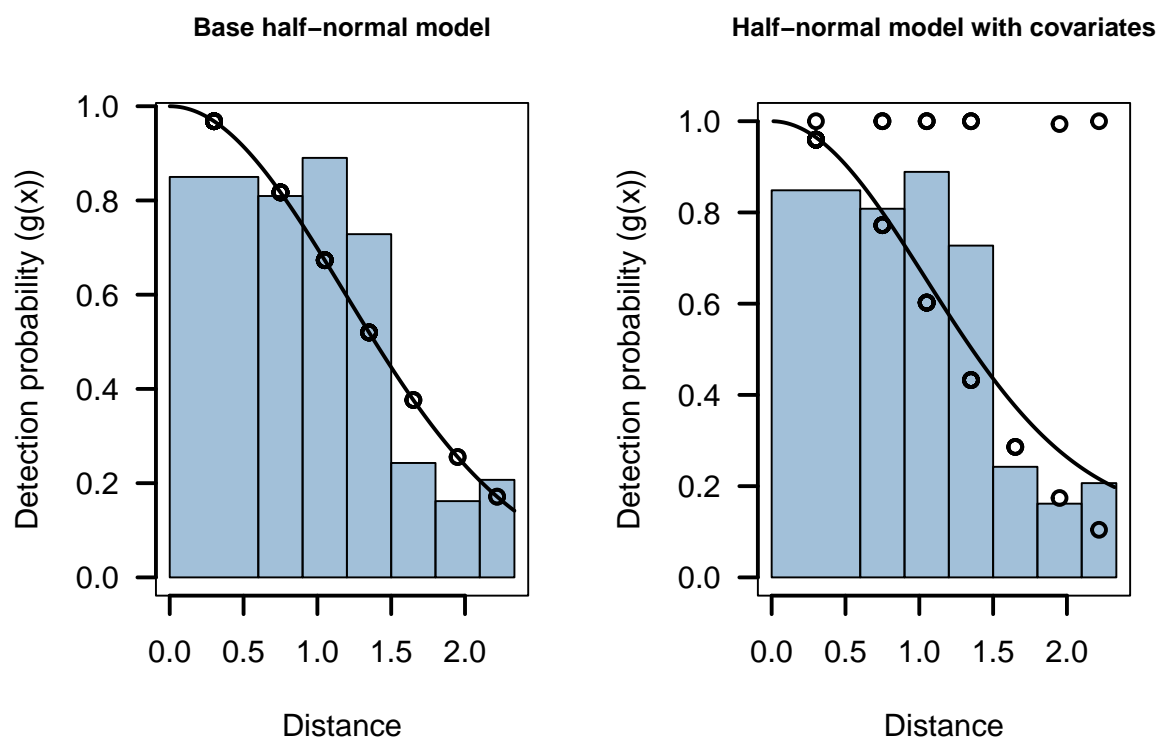
Figure 5: Detection functions using binned data

Code Appendix

```r
# import packages and data
library(statsecol)
library(tidyverse)
library(viridisLite)
library(ggridges)
library(Distance)
data(bowhead_LT)

# raw data plot
# looks like hazard rate would be the best fit here
ggplot(data = bowhead_LT) +
  geom_histogram(aes(x = distance, fill = as.factor(size)), bins= 21) +
  labs(x = "Distance truncated by 0.1 km",
       y = "Observerd # individuals",
       title = "Raw bowhead whale observations",
       subtitle = "Using 21 bins and minor left-truncated distances",
       fill = "Group size") +
  theme_bw()

# aggressive binning
# definitely half-normal is the better fit here
histBreaks2 <- c(0, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1, max(bowhead_LT$distance,
```
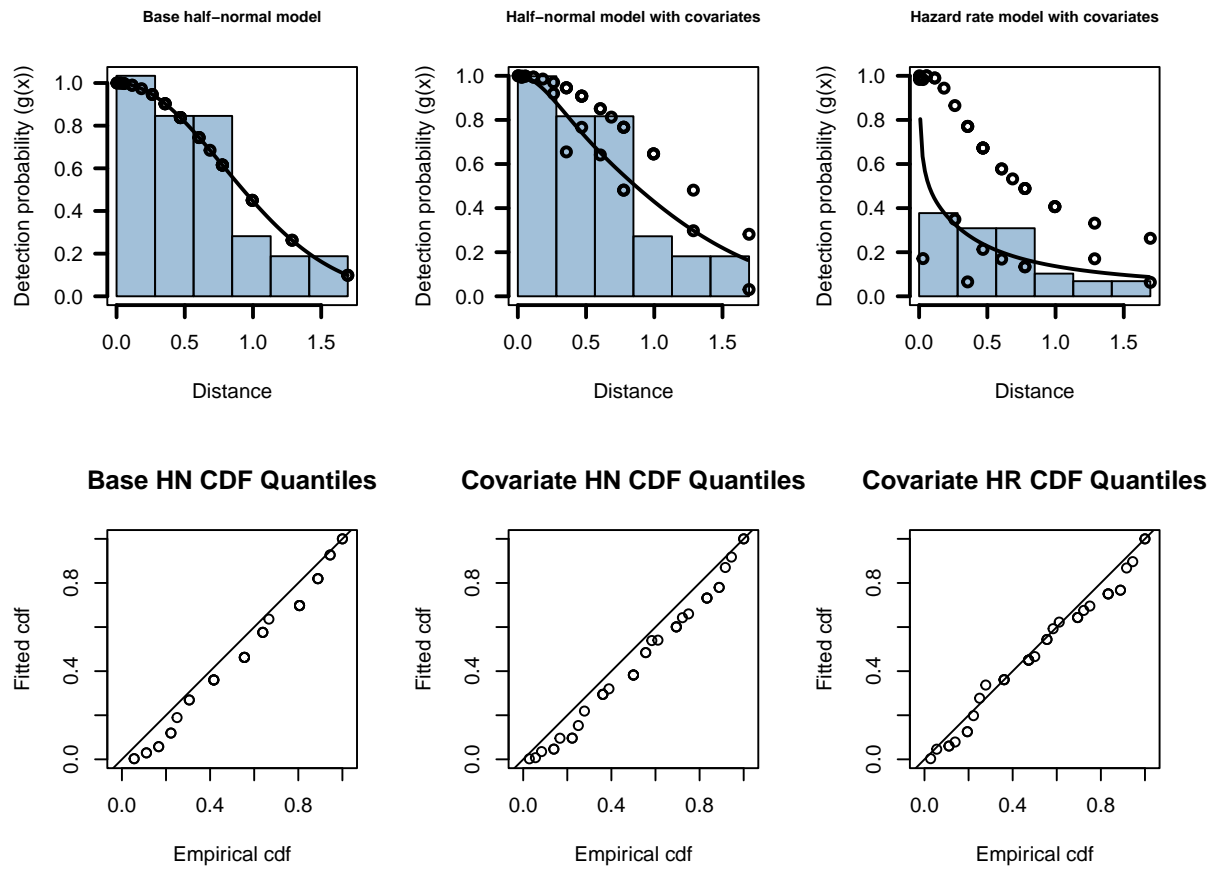
Figure 6: Detection functions using truncated data

```r
                                               na.rm = TRUE))
ggplot(data = bowhead_LT) +
  geom_histogram(aes(x = distance, fill = as.factor(size)),
                 breaks = histBreaks2) +
  labs(x = "Distance truncated by 0.1 km",
       y = "Observerd # Individuals",
       title = "Binned bowhead whale observations",
       subtitle = "Using 7 custom bins with minor left-truncated distances",
       fill = "Group size") +
  theme_bw()


#truncate to get decreasing observations with distance
bowhead_LT_Trunc = bowhead_LT %>% filter(distance >= 0.64) %>%
  mutate(distance = distance - 0.64)


# plot the new data
# both half-normal and hazard rate appear possible here
ggplot(data = bowhead_LT_Trunc) +
  geom_histogram(aes(x = distance, fill = as.factor(size)), bins = 14) +
  labs(x = "Distance left-truncated by 0.74 km",
       y = "Observerd # Individuals",
       title = "Filtered bowhead whale observations",
       subtitle = "Using 14 bins with substantially left-truncated distances",
       fill = "Group size") +
  theme_bw()


conversion.factor <- convert_units(distance_units = "kilometre",
                                   effort_units = "kilometre",
                                   area_units = "square kilometre")


# using model fits from separate alternatives,
# we'll only fit no-adjustment HN, HR, and size covariates
# ds() will automatically pick the low AIC version (adjustment = NULL)
baseRawHN <- ds(data = bowhead_LT, key = "hn",
                convert_units = conversion.factor)
baseRawHR <- ds(data = bowhead_LT, key = "hr",
                convert_units = conversion.factor)
# sometimes does not converge
sizeRawHN <- ds(data = bowhead_LT, key = "hn", adjustment = NULL,
                convert_units = conversion.factor,
                formula = ~size)
sizeRawHR <- ds(data = bowhead_LT, key = "hr", adjustment = NULL,
                convert_units = conversion.factor,
                formula = ~size)


# size + half-normal is the best fit on the raw data
summarize_ds_models(baseRawHN,
                    baseRawHR,
                    sizeRawHN,
                    sizeRawHR, output = "plain")
```

```r
par(mfrow = c(2,2))
plot(baseRawHN, which=2, pl.col = adjustcolor("steelblue",0.5), border=NULL,
     lwd = 2,
     breaks = histBreaks2,
     ylab = "Detection probability (g(x))", xlab = "Distance", las=1,
     main = "Base half-normal model")
# covariate model assumes perfect detection of large-sized groups
# this is a quirk of the limited data
# only Region 2 has any within-group variation in size
plot(sizeRawHN, which=2, pl.col = adjustcolor("steelblue",0.5),border=NULL,
     lwd = 2,
     breaks = histBreaks2,
     ylab = "Detection probability (g(x))", xlab = "Distance", las=1,
     main = "Half-normal model with covariates")
# both models have solid GOF
# covariate model is slightly better from visual QQ inspection
gof_ds(baseRawHN, main="Base Observed vs. Expected CDF", ks = TRUE)
gof_ds(sizeRawHN, main="Covariate Observed vs. Expected CDF", ks = TRUE)

# repeat for binned data
baseBinHN <- ds(data = bowhead_LT, key = "hn",
                cutpoints = histBreaks2,
                convert_units = conversion.factor)
baseBinHR <- ds(data = bowhead_LT, key = "hr",
                cutpoints = histBreaks2,
                convert_units = conversion.factor)
# often does not converge
sizeBinHN <- ds(data = bowhead_LT, key = "hn", adjustment = NULL,
                cutpoints = histBreaks2,
                convert_units = conversion.factor,
                formula = ~size)
sizeBinHR <- ds(data = bowhead_LT, key = "hr", adjustment = NULL,
                cutpoints = histBreaks2,
                convert_units = conversion.factor,
                formula = ~size)

#binned data also prefers models in the same order as unbinned
summarize_ds_models(baseBinHN,
                    baseBinHR,
                    sizeBinHN,
                    sizeBinHR, output = "plain")

# binned models have a very similar fit (if not identical) to the raw data
par(mfrow = c(1,2))
plot(baseBinHN, which=2, pl.col = adjustcolor("steelblue",0.5),border=NULL,
     lwd = 2,
     ylab = "Detection probability (g(x))", xlab = "Distance", las=1,
     main = "Base half-normal model")
# covariate model still assumes perfect detection of large-sized groups
# still a quirk of the data due to between-strata variation
plot(sizeBinHN, which=2, pl.col = adjustcolor("steelblue",0.5),border=NULL,
```

```
      lwd = 2,
      ylab = "Detection probability (g(x))", xlab = "Distance", las=1,
      main = "Half-normal model with covariates")
# GOF produce output but cannot generate plots
# bootstrapping impossible so KS test doesn't happen
gof_ds(baseBinHN, main="Base Observed vs. Expected CDF", ks = TRUE)
gof_ds(sizeBinHN, main="Covariate Observed vs. Expected CDF", ks = TRUE)

# finally repeat for truncated data
# truncated data prefers complex fits w/ more covariates and adjustments
baseTruncHN <- ds(data = bowhead_LT_Trunc, key = "hn",
                convert_units = conversion.factor)
baseTruncHR <- ds(data = bowhead_LT_Trunc, key = "hr",
                convert_units = conversion.factor)
sizeTruncHN <- ds(data = bowhead_LT_Trunc, key = "hn", adjustment = NULL,
                convert_units = conversion.factor,
                formula = ~size)
sizeTruncHR <- ds(data = bowhead_LT_Trunc, key = "hr", adjustment = NULL,
                convert_units = conversion.factor,
                formula = ~size)

# truncated data shows very different rankings
# hazard rate is preferred w/ size covariate
# then basic hanf-normal, then size + HN, finally base HR
summarize_ds_models(baseTruncHN,
                    baseTruncHR,
                    sizeTruncHN,
                    sizeTruncHR, output = "plain")

# looks like the binned models have exactly the same fit as the raw data
par(mfrow = c(2,3))
plot(baseTruncHN, which=2, pl.col = adjustcolor("steelblue",0.5),border=NULL,
      lwd = 2,
      ylab = "Detection probability (g(x))", xlab = "Distance", las=1,
      main = "Base half-normal model")
# covariate model no longer assumes perfect detection of large-sized groups
plot(sizeTruncHN, which=2, pl.col = adjustcolor("steelblue",0.5),border=NULL,
      lwd = 2,
      ylab = "Detection probability (g(x))", xlab = "Distance", las=1,
      main = "Half-normal model with covariates")
plot(sizeTruncHR, which=2, pl.col = adjustcolor("steelblue",0.5),border=NULL,
      lwd = 2,
      ylab = "Detection probability (g(x))", xlab = "Distance", las=1,
      main = "Hazard rate model with covariates")
# covariate model still slightly better from visual QQ plot
gof_ds(baseTruncHN, main="Base Half-Normal Observed vs. Expected CDF",
        ks = TRUE)
gof_ds(sizeTruncHN, main="Covariate Half-Normal Observed vs. Expected CDF",
        ks = TRUE)
gof_ds(sizeTruncHR, main="Covariate Hazard Rate Observed vs. Expected CDF",
        ks = TRUE)
```

```r
# output the models for consistency and processing speed
save(baseRawHN, baseRawHR, sizeRawHN, sizeRawHR,
     baseBinHN, baseBinHR, sizeBinHN, sizeBinHR,
     baseTruncHN, baseTruncHR, sizeTruncHN, sizeTruncHR,
     file="df-models.RData")

# break out abundance under both models by region
# generate abundance estimates for the best raw data models
region_table <- unique(bowhead_LT[,c("Region.Label", "Area")])
sample_table <- unique(bowhead_LT[,c("Region.Label",
                                     "Sample.Label",
                                     "Effort")])
observation_table <- unique(bowhead_LT[,c("object",
                                          "Region.Label",
                                          "Sample.Label")])
# raw model
baseRawHN_N <- dht(model = baseRawHN$ddf,
                   region.table = region_table,
                   sample.table = sample_table,
                   obs.table = observation_table)
# covariate model
# lower overall SE driven by Regions 2 and 9
# Regions 3, 11, 12, and 15 have increased SE
# abundance decreases in regions 2 and 9 while increasing elsewhere
# Regions 2 and 9 are the only sub-regions with size > 1 observations
sizeRawHN_N <- dht(model = sizeRawHN$ddf,
                   region.table = region_table,
                   sample.table = sample_table,
                   obs.table = observation_table)
# repeat abundance estimation process for remaining models of interest
# base binned model
baseBinHN_N <- dht(model = baseBinHN$ddf,
                   region.table = region_table,
                   sample.table = sample_table,
                   obs.table = observation_table)
# size covariate binned data model
sizeBinHN_N <- dht(model = sizeBinHN$ddf,
                   region.table = region_table,
                   sample.table = sample_table,
                   obs.table = observation_table)
# base truncated data model
baseTruncHN_N <- dht(model = baseTruncHN$ddf,
                     region.table = region_table,
                     sample.table = sample_table,
                     obs.table = observation_table)
# size covariate truncated data model
sizeTruncHN_N <- dht(model = sizeTruncHN$ddf,
                     region.table = region_table,
                     sample.table = sample_table,
                     obs.table = observation_table)
```

```r
# check variance using bootstraps rather than delta method
# delta-method approximation assumes independence between
# uncertainty in the detection function and variability in encounter rate

# takes ~3 minutes on my computer, likely longer for most
# check or delete the "cores = " if you don't have 10 cores on your computer
est.boot <- bootdht(model=sizeRawHN, flatfile=bowhead_LT,
                    summary_fun=bootdht_Nhat_summarize,
                    convert_units=conversion.factor, nboot=999, cores=10)
alpha <- 0.05
# the lower bound of our estimate is very poorly constrained
# 95% CI between 0 and 276
bootci <- quantile(est.boot$Nhat, probs = c(alpha/2, 1-alpha/2),
                   na.rm=TRUE)

# plot generated bootstraps
par(mfrow = c(1,1))
hist(est.boot$Nhat, nc=30,
     main="Distribution of bootstrap estimates\nwithout model uncertainty",
     xlab="Estimated abundance")
abline(v=bootci, lwd=2, lty=2)

# repeat boostraps w/ resample by region rather than by transect
# our upper CI increases slightly but the lower CI stays at 0
#est.bootStrata <- bootdht(model=bowhead.hn.null.size, flatfile=bowhead_LT,
#                summary_fun=bootdht_Nhat_summarize,
#                resample_strata = TRUE,
#                convert_units=conversion.factor, nboot=999, cores=10)
#bootciStrata <- quantile(est.bootStrata$Nhat, probs = c(alpha/2, 1-alpha/2),
#                na.rm=TRUE)


# this following section is for transparency on full model fitting
# attempted with the original, minorly truncated data
# play around with some adjustment terms with the size model
# adjustment terms generally aren't included in covariates for whatever reason
# Cosine(2) is preferred but never monotonic so that's not it
#bowhead.hn.cos.size <- ds(data = bowhead_LT, key = "hn", adjustment = "cos",
#                      convert_units = conversion.factor,
#                      formula = ~size,
#                      nadj = c(1,2,3,4))
# Hermite(4) is preferred but is non-monotonic; Hermite(4,6) is not as bad
#bowhead.hn.herm.size <- ds(data = bowhead_LT, key = "hn", adjustment = "herm",
#                      convert_units = conversion.factor,
#                      formula = ~size,
#                      nadj = c(2,3,4))
# Poly(4,6) ends up being non-monotonic, so leaving it at Polynomial(4)
#bowhead.hn.poly.size <- ds(data = bowhead_LT, key = "hn", adjustment = "poly",
#                      convert_units = conversion.factor,
#                      formula = ~size,
```

```
#                         nadj = c(1,2,3,4))
# now repeat for half-normal
# non-monotonic and g(x) > 1 with a cosine adjustment
#bowhead.hr.cos.size <- ds(data = bowhead_LT, key = "hr", adjustment = "cos",
#                         convert_units = conversion.factor,
#                         formula = ~size,
#                         nadj = c(1,2,3,4))
#Hermite(4,6) surprisingly matches everything
#bowhead.hr.herm.size <- ds(data = bowhead_LT, key = "hr", adjustment = "herm",
#                         convert_units = conversion.factor,
#                         formula = ~size,
#                         nadj = c(2,3,4))
#Polynomial adjustments don't match monotonicity
#bowhead.hr.poly.size <- ds(data = bowhead_LT, key = "hr", adjustment = "poly",
#                         convert_units = conversion.factor,
#                         formula = ~size,
#                         nadj = c(1,2,3,4))
```

References

1. Rekdal, S. L. *et al.* Trends in bowhead whales in west greenland: Aerial surveysvs. Genetic capture-recapture analyses. *Marine Mammal Science* **31**, 133–154 (2014).

2. Buckland, S. T., Rexstad, E. A., Marques, T. A. & Oedekoven, C. S. *Distance Sampling: Methods and Applications.* (Springer International Publish, 2015).

3. Borchers, D. L., Zucchini, W., Heide-Jørgensen, M. P., Cañadas, A. & Langrock, R. Using hidden markov models to deal with availability bias on line transect surveys. *Biometrics* **69**, 703–713 (2013).