# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## ABSTRACT

Summary of objectives, methods and results...

# RÉSUMÉ

Usually the french version of the abstract goes here, but it is included here to show how accents might be added to text and how non-numbered sections work.

# 1 INTRODUCTION

Hecate Strait and Queen Charlotte Sound contain some of the most productive fishing grounds in British Columbia (BC), providing key habitat for many commercially important groundfish species. This project will employ multivariate statistics and geostatistical approaches to analyse relationships between environmental factors and distribution and productivity of key groundfish populations in Hecate Strait and Queen Charlotte Sound. The project will improve advice for management of Pacific groundfish stocks by: (i) improving understanding of environmental drivers of groundfish distribution and productivity needed for ecosystem-based management; (ii) improved estimates of abundance for key species; (iii) identification of juvenile habitats; (iv) provision of a baseline for understanding impacts of environmental change on species distribution; and (v) identification of species that could indicate ecosystem change through shifts in distribution and productivity.

Distribution and abundance of groundfish species is associated with invariant (e.g., depth, bottom-type) and variable (e.g., temperature, salinity) environmental factors (Perry et al., 1994; Rooper et al., 2005; Rooper, 2008). Measuring relationships between these factors and distribution and abundance is the first step in understanding drivers of productivity (recruitment, growth, mortality), which is a critical component of ecosystem-based management. This project will employ statistical and hierarchical Bayesian geostatistical models (Rooper et al., 2005; Rooper, 2008; Lecomte et al., 2013*b*,*a*) to analyse relationships between environmental factors and distribution, abundance and size structure of a set of key, commercially harvested groundfish species in Hecate Strait and Queen Charlotte Sound. Models will utilize data from commercial trawl logbooks and fishery independent surveys. Temperature and salinity data from the Regional Oceanographic Model System (ROMS) (Masson and Fine, 2012) provided key model inputs.

Results for each species will include maps of predicted distribution and abundance of adults and juveniles; Bayesian predictive probability distributions of the relationships between habitat and environmental factors and abundance; and plots of the distribution of adults and juveniles along environmental gradients. Species most likely to be impacted by large environmental changes (e.g., ocean temperature) will be identified. Working with our external collaborator, results for a subset of species will be compared to results from Alaska to test generality of results and identify key differences. The project will provide updated estimates of abundance for the species of interest and will identify locations that may represent critical juvenile habitat.

In the long-term, the project will provide important baseline data for understanding potential future impacts of environmental change. Some species (e.g., Pacific Cod) are known to vary their habitat with depth to maintain a limited temperature range (Perry et al., 1994), indicating that their distribution could be strongly affected by long-term changes in ocean temperature. Published studies have shown that large-scale redistribution of north Pacific fish populations may occur under future climate scenarios, with the potential for large impacts on ecosystem structure and function (Jones and Cheung, 2014; Cheung et al., 2015). Groundfish indicator species most likely to be affected by environmental change will be identified in this project. Finally, ecosystem-based fishery management is based on principles of understanding the structure and function of the living components of marine ecosystems. In the US, NOAA is mandated to identify habitats essential for every managed fish species and identify those habitats that contribute most to survival, growth and productivity (Sigler et al., 2012). The analyses in this project will form an important component of this understanding for central and northern BC waters. Through partnership with our external collaborator, comparative analyses will test generality of results and identify differences between BC and Alaska.

# 2 METHODS

## 2.1 GROUNDFISH BOTTOM TRAWL SURVEY

Groundfish bottom trawl survey data for our analyses were collected during the DFO's biennial Hecate Strait Synoptic Trawl Survey and the Queen Charlotte Sound Synoptic Trawl Survey between 1984 and 2015. Tows began at pre-determined locations as part of a random, stratified sampling design with strata based on (?????). Fish were identified to species, lengths measured, sexed (0 = unknown, 1 = male, 2 = female, 3 = ?), and characterised by maturity (categories 1 to 7) according to a prioritized sampling protocol. A sub-sample of each fish species from every tow were grouped by species and weighed to the nearest kilogram (kg). Tow length and travel speed were also recorded. The data used these analyses are therefore the biomass of each species in each tow, and in some cases, standardised as catch per unit of trawl effort (catch per square kilometre).

## 2.2 ENVIRONMENTAL DATA

### 2.2.1 Survey Data

Highly spatially-resolved, commercial logbook data were available from BC's 100%-observed groundfish bottom trawl fishery. All these data are held in Oracle databases, co-managed by the Pacific Groundfish Statistics Program. During the groundfish bottom trawl survey average net depth (m) and average ocean temperature (°C) data were also collected from sensors attached to the net. Spatial bottom-type data at 100 m and 20 m resolution were available at the Pacific Biological Station. Scientists at NOAA's Alaska Fisheries Science Center are currently engaged in developing distribution maps for all commercially-fished species in Alaska. Some of these data were available, via our external collaborator Dr. Rooper at the Alaska Fisheries Science Center.

### 2.2.2 ROMS Data

ROMS is a terrain following equation model that has been used extensively globally (Haidvogel et al., 2008). The use of the model for the BC coast is forced by the North American Regional Reanalysis (NARR) (Mesinger et al., 2004) atmospheric data, and lateral boundary conditions are extracted from the Simple Ocean Data Assimilation project, or SODA (Carton and Giese, 2008). At the open boundaries tidal forcing is applied using the output from a North-East Pacific tidal model (Foreman et al., 2000). Furthermore, the model is forced by monthly fresh water discharge from major BC rivers, derived as in Morrison et al. (2012). Further details pertaining to the model, as well as an extensive model validation, can be found in Masson and Fine (2012).

The model was used to estimate the spatial and temporal variability of the salinity and temperature in the ocean off the British Columbia coast. ROMS was used to hindcast the period of 1979 to 2011 and the model domain extends from the Columbia River to the Alaska Panhandle. The horizontal grid resolution is 3 km, with 60 levels in the vertical available; however, only the deepest level was used for this study. The data used in this study are the average summer (June to August???) bottom temperature and salinity for each year.

ROMS bottom temperature data were highly related to bottom trawl survey average net temperature $R^2 = 0.84$. (can unhardwire once I'm sourcing our R code and value is in global environment)

#### 2.2.3 BCMCA Data

Substrate type (e.g., hard, sandy, muddy and undefined) and ocean depth for the study area was obtained from the BC Marine Conservation Analysis database. Substrate data was then gridded in 3 km cells with PBSMapping R script to convert the data to a usable format for our analyses.

### 2.3 SOFTWARE

A version of the hierarchical Bayesian model has already been written using the OpenBUGS programming language. This code was customized and refined for this project. All other statistical models and graphic outputs were developed using the R-programming language. Species abundance and data distribution maps were made using ArcGIS or a similar (e.g., QGIS, PBSMapping).

### 2.4 DATA MANAGEMENT

Spatially-gridded Canadian datasets and gridded predictions of distribution and abundance were transferred to a database managed by the Pacific Groundfish Statistics program. They were available internally to Pacific stock assessment scientists, and by request externally. Maps were incorporated into an intranet-based tool, making them available to Pacific scientists and managers.

### 2.5 CUMULATIVE DISTRIBUTION FUNCTIONS

The purpose of this component of the study was to identify significant associations between environmental parameters and the distributions of 20 species of northeast Pacific Ocean groundfish. It was also meant to be a continuation of the analysis presented by Perry et al. (1994). We used cumulative distribution functions (cdf's) of fish catch (CPUE) and the environmental factors described above (substrate, salinity, temperature and depth) (Perry and Smith, 1994; Perry et al., 1994). This technique calculates the empirical cdf's for the environmental parameters alone and the environmental parameters weighted by the CPUE of a particular species (Perry and Smith, 1994; Perry et al., 1994).

The probability associated with each observation in a cdf is 1/*n*, but the stratified random survey design results in a probability of 1/$n_h$ within each stratum (all symbols are provided in Table XX). Therefore the cdf for a given habitat variable ($x_i$) is of the form (Chambers and Dunstan, 1986):

$$f(t) = \sum_h \sum_i \frac{W_h}{n_h} \, I(x_{hi}) \tag{1}$$

with the indicator function

$$I\ (x_{hi}) = \begin{cases} 1, & \text{if } x_{hi} \leq t; \\ 0, & \text{otherwise.} \end{cases}$$

Where *t* represents an index, ranging from the lowest to the highest values of the habitat parameter at a step appropriate for the desired resolution. Equation 1 is calculated over all values of *t* for each habitat measurement ($x_{hi}$) available. The cdf's derived from equation 1 can be used to determine the proportion of the environmental-weighted catch within any range of the environmental variable during the survey. For example, the range of depths that occurred within the central 50% (between the 25th and 75th percentiles) or the central 95% (between the 2.5th and 97.5th percentiles) of the area surveyed can easily be calculated from each species' cdf figures.

Copied directly from Perry and Smith 1994. Including the survey stratification scheme via the $W_h$/$n_h$ terms ensures that we have an unbiased estimate of the frequency distribution for the habitat measurement. Ignoring the stratification by replacing $W_h$/$n_h$ with 1/*n* would result in either under- or overestimating the area associated with any particular value of the habitat measurement. However, the term $W_h$/$n_h$ does simplify to 1/*n* when the number of sets allocated to each stratum is proportional to the size of the stratum (i.e., $n_h$ = $nW_h$). That is, stratification can be ignored when the allocation of sets is strictly proportional to the stratum size. Next, we associate the catch of fish (in weight) of a particular species in each tow with the habitat parameters during that tow as weight in the form:

$$g(t) = \sum_h \sum_i \frac{W_h}{n_h} \frac{y_{hi}}{y_{st}} \, I(x_{hi}). \tag{2}$$

Scaling the number of fish caught ($Y_{hi}$) by the stratified mean number of fish caught ($\overline{y}_{st}$) in equation 2 results in *g(t)* summing to 1 over all values of *t*. If large values of ($Y_{hi}$)/$\overline{y}_{st}$ are consistently associated with particular habitat conditions, then this suggests a strong association between the fish species and those habitat conditions. The cumulative distribution functions calculated from equation 2 illustrate the range of conditions at which the species occurred and can be compared with the habitat conditions available in the sampled area as calculated with equation 1. For example, 50% of the depths surveyed in ..... (curve *f(t)*, Fig. XX) were less than

XX m, while 50% of the XX Arrowtooth Flounder (curve *g(t)*, Fig. XX) were caught at depths less than XX m. The curve *g(t)* can differ widely from the habitat curve *f(t)* depending on the range of conditions occupied by the fish. At one extreme, if the fish were associated with one depth only (e.g., 100 m), *g(t)* would be zero for $t < 100$ and equal to 1.0 for $t \geq 100$. If there was no particular association between fish distributions and the habitat variable within the area surveyed, for example if the fish were randomly distributed with respect to the habitat variable, then *g(t)* and *f(t)* would be almost identical.

The third step is to determine the strength of the association between catch and the habitat variable by assessing the degree of difference between the the two curves, *g(t)* and *f(t)*. Our test statistic's similar to that used for comparing empirical cdf's in Kolmogorov-Smirnov tests (see Conover 1980). We calculate the maximum absolute vertical distance between *g(t)* and *f(t)* as:

$$\max_{\forall t} |g(t) - f(t)| = \max_{\forall t} \left| \sum_h \sum_i \frac{W_h}{n_h} \left( \frac{y_{hi} - \overline{y}_{st}}{\overline{y}_{st}} \right) I(x_{hi}) \right| \tag{3}$$

where |*g(t)* - *f(t)*| indicates the absolute value of the difference between *g(t)* and *f(t)* at any point *t*. The stratified random survey design complicates the distributional assumptions for the test statistic in equation 3, and therefore the standard tables for the Kolmogorov-Smirnov test, or indeed any goodness-of-fit test, cannot be used (see Rao and Thomas 1989). Instead, we developed a randomization procedure (Noreen 1989) to evaluate the significance of the test statistics. We modelled the distribution of the test statistic under the null hypothesis of random association between fish catch (numbers) and habitat variable through Monte-Carlo sampling. This was done by randomizing the pairings of $(W_h/n_h)$[($y_{hi}$ - $\overline{y}_{st}$)/$\overline{y}_{st}$] and $x_{hi}$ over all *h* and *i* for the data within a survey and then calculating the test statistic in equation 3 for the new pairs. The $x_{hi}$ for the pairings were obtained by sampling with replacement the observed $x_{hi}$ with probability $W_h n_h$. This procedure was repeated a large number of times to give a pseudo-population of test statistics under the null hypothesis. The test statistic for XX species and depth in MONTH YEAR (Fig XX) was equal to XX (also indicated as "max" in Fig XX) which was greater than or equal to XX% (2000 out of 2001) of the test statistics in a randomized distribution.. There were 2001 test statistics when the original observed pairing of the data was included. Interpreting these results similar to a standard statistical hypothesis test, we note that the probability of obtaining a test statistic as large as XX by chance is close to zero (*p* = XX) and conclude that there was a very strong association by XX species with a specific range of depths available in the survey area during MONTH YEAR. This randomization test is a two-sided test, since it is the magnitudes of the absolute differences between *g(t)* and *f(t)* that are of interest. However, environmental conditions are often correlated, for example where temperature decreases with increasing depth, which suggests that an association between a species and a particular depth range may also be confounded by an associated with temperature. This problem can be explicitly considered by extending equations 2 and 3 to two (or more) habitat variables simultaneously. For *k* variables, equations 1 and 2 can be written as

$$f(t) = \sum_h \sum_i \frac{W_h}{n_h} I(x_{hi}) \tag{4}$$

$$g(\boldsymbol{t}) = \sum_h \sum_i \frac{W_h}{n_h} \frac{y_{hi}}{\overline{y}_{st}} \, I(\boldsymbol{x}_{hi}) \tag{5}$$

where

$$I\left(\boldsymbol{x}_{hi}\right) = \begin{cases} 1, & \text{if } (x_{hi} \leq t_1, x_{hi\ 2} \leq t_2, ..., x_{hi\ k} \leq t_k); \\ 0, & \text{otherwise.} \end{cases}$$

Boldface type text for $\boldsymbol{t}$ and $\boldsymbol{x}$ indicates vectors of habitat variables. In the two-variable case, *g(**t**)* can be represented as a three-dimensional surface in which the cumulative frequency forms the vertical axis. The test statistic (equation 3) is modified as

$$\max_{\forall t} |g(\boldsymbol{t}) - f(\boldsymbol{t})| = \max_{\forall t} \left| \sum_h \sum_i \frac{W_h}{n_h} \left( \frac{y_{hi} - \overline{y}_{st}}{\overline{y}_{st}} \right) I(\boldsymbol{x}_{hi}) \right|. \tag{6}$$

### 2.6 MODELS

The project will apply a mix of established and recently-published statistical approaches to achieve the deliverables of the project. An earlier study (Perry et al., 1994) applied a set of multivariate statistical models to classify groundfish species in Hecate Strait according to their relationships with invariant and variable environmental factors. Recent studies have further developed these types of approaches for Alaskan groundfish species (e.g., Rooper et al. (2005); Rooper (2008)). A problem with spatial datasets for many marine species is the high proportion of zero observations, which can bias results. One approach to solving this problem is to use a two-stage model to first predict presence and absence, then analyse relationships between environmental variables and abundance (Rooper and Martin, 2009). More recently, the problem has been addressed using new Bayesian hierarchical models which estimates both the probability of zero observations and abundance in a hierarchical framework (Lecomte et al., 2013*b*,*a*). Within this framework, a geostatistical approximation, consisting of a linear model with spatially-correlated errors, is used to efficiently predict spatial abundance as a function of environmental factors (Lecomte et al., 2013*a*). The model outputs spatial predictions of abundance, and predictive probability distributions of the effects of each environmental factor on abundance for each species. Models are calibrated with spatial abundance observations. Analyses can be done on different size-classes of the population to better understand differences in adult and juvenile distribution and improve understanding of productivity.

## 3 RESULTS

### 3.1 ENVIRONMENTAL COVARIATES

Environmental covariates considered during this analysis included average trawl net depth, and bottom temperature and salinity. Trawl net depth was recorded during each tow. Temperature and salinity data were collected during the groundfish trawl surveys and were also available from the Regional Oceanographic Model System (ROMS).

## 3.2 CUMULATIVE DISTRIBUTION FUNCTIONS

### 3.2.1 Survey Data

The relationships varied between environmental factors considered in this study (temperature, salinity and depth) and the twenty species of groundfish. For a better understanding of how to intepret the cumulative distribution results, we provide examples using Rock Sole and Longspine Thornyhead results. Rock Sole were most abundant at high temperatures with 50% of their biomass occurring at temperatures greater than 9.1°C and depths greater than 40.5 m (refer to figure). Fifty perent of Longspine Thornyhead occurred at temperatures cooler than 4.9°C and depths greater than 470 m. These examples are in opposite ends of the spectrum with respect to habitat use within the study area and therefore their catch weight cdf's were much different than the cdf's for both environmental variables (i.e., the distribution of available temperatures and depths within the survey area).

For comparative purposes between this study and Perry et al. (1994) we present the results of Arrowtooth Flounder and Spiny Dogfish for depth and survey temperature. With respect to depth, Perry et al. (1994) found that 25% of the Arrowtooth Flounder biomass occurred at depths shallower than 84 m, 50% occurred shallower than 100 m, and 75% occurred shallower than 120 m. Our study found a different distribution with 25% of the biomass shallower than 112 m, 50% shallower than 140.5 m and 75% shallower than 169.5 m (ref to figure). Twenty-five percent of Spiny Dogfish in Perry et al. (1994) were at depths less than 33 m, 50% at less than 50 m and 75% at less than 93 m. Similarly, our study found that 25% of the biomass was at depths less than 51 m, 50% less than 52.5 m and 75% less than 109 m (ref to figure).

For survey temperature, Perry et al. (1994) found that 25% of the Arrowtooth Flounder biomass occurred at temperatures less than 6.1°C, 50% at less than 6.5°C, and 75% at less than 6.7°C. Our results were similar with 25% of Arrowtooth Flounder in our study occurring at temperatures less than 5.9°C, 50% at less than 6.1°C and 75% at less than 6.5°C. Again, Spiny Dogfish cdf's were similar between studies. Twenty-five percent of Spiny Dogfish in the previous study were found at temperatures less than 6.6°C, 50% at less than 6.8°C and 75% at less than 9.6°C. Finally, in this study 25% of Spiny Dogfish occurred at temperatures less than 6.8°C, 50% at less than 8.8°C and 75% at less than 10.1°C. This is a considerably larger temperature range that is utilised by this species relative to Arrowtooth Flounder .

The randomization test (3)for determining significant differences from random distributions of the environmental variables indicated that Arrowtooth Flounder were (or were not) significantly different from the available depths or temperatures. Spiny Dogfish had (or did not have) a similar result. Refer to table XX for a summary of significant differences from random distributions.

The only species that did not show significant differences from random distributions of the environmental variables were.... (see Table Xx??)

In general, due to the confounding relationships among temperature, salinity and depth, species that were most abundant at low temperature were also most abundant at high salinity and deeper depths. For example, 50% of Shortspine Thornyhead were in temperatures less than 5.5 °C, salinities greater than 38.2 PSU and depths greater than 290 m.

### 3.3 ROMS DATA

Cumulative distribution functions of catch weight were also calculated based on ROMS depth, temperature and salinity data. Catch-weighted depth for Arrowtooth Flounder was at 59 m for 2.5% of the biomass, 25% at less than 108 m, 50% at less than 138 m, 75% at les than 164 m and 97.5% at less than 288 m. These catch-weighted depths were all similar (approximately $\pm$ 5 m) except for the 2.5th percentile which was 13 m deeper for the survey temperature. The catch-weighted depth percentiles for Spiny Dogfish were very similar between the two data types with all being witin $\pm$ 10 m of each other.

With respect to temperature-weighted catch, the 2.5th percentile for Arrowtooth Flounder was at 5.6°C, 25% of the biomass was colder than 6.5°C, 50% at less than 7.0°C, 75% at less than 7.8°C and 95% at less than 9.6°C. This represents a wider temperature range (4.0°C) for 95% of the biomass than the survey data (2.6°C); however, the the middle 50% of the ROMS biomass occurred within 1.3°C (6.5°C to 7.8°C), which was smaller than the 1.6°C temperature range for the survey data (5.9°C to 6.5°C). Spiny Dogfish had a similar temperature range between data types, 6.6°C for ROMS compared to 7.1°C for survey data. In contrast to Arrowtooth Flounder,the middle 50% of the biomass occurred in a narrower range of temperatures for ROMS data (7.7°C to 9.7°C) relative to survey data (6.8°C to 10.1°C). For both data types, 95% of the biomass was colder than 12.5°C.

The are no salinity-weighted catch percentiles to compare between data types or with Perry et al. (1994); however, the 2.5th percentile for Arrowtooth Flounder was for ROMS salinity was less than 34 PSU. The 25th percentile of biomass was for salinity less than 35.5 PSU, 50% at less than 36.4 PSU, 75% at less than 37.1 PSU and 97.5% at less tahn 38.2 PSU. Spiny Dogfish was similar with a 95th percentile range of from 33.9 to 37.7 PSU. The 25th percentile was for salinities less than 34.3 PSU, 50% at less than 34.3 PSU and 75% at salinities less than 35.6 PSU.

The randomization test (3)for determining significant differences from random distributions of the ROMS environmental variables indicated that Arrowtooth Flounder were (or were not) significantly different from the available depths, temperatures and salinities. Spiny Dogfish had (or did not have) a similar result. Refer to table XX for a summary of significant differences from random distributions.

The only species that did not show significant differences from random distributions of the ROMS environmental variables were.... (see Table Xx??)

# 4 HOW REFERENCES WORK

## 4.1 HOW THE FIGURE AND TABLE REFERENCES WORK

A figure/table reference works by adding a reference name to a figure/table, then remembering what is was and using a `\ref` command to reference the figure/table. For example in Figure **??**, the figure reference code has a label tag like this `\label{fig:example-random-stuff}`. The figure can be referenced anywhere in the latex document by using this syntax: `\ref{fig:example-random-stuff}`. The numbering is taken care of for you and is separate for each type of reference. Here is a list of suggested prefixes to use for different reference types:

1. **sec**: - section
2. **subsec**: - subsection
3. **fig**: - figure
4. **tab**: - table
5. **eq**: - equation
6. **lst**: - code listing
7. **itm**: - enumerated list item (like this list)
8. **chap**: - appendix

## 4.2 HOW APPENDIX REFERENCES WORK

Appendix references are much like chapters of a book. They can be added or commented out easily at the bottom of *example.Rnw*. This helps with the incremental form of development where you make sure the main document is compiling and then when ready, uncomment the appendix inclusion code and the appendix will be included in the document. Once included, any appendix references will be resolved.

The code which adds an appendix is *knitr* code because you want the appendix added before the knitting process so that any figures or R expressions are resolved, just like in the main document. This is an example of how appendix code is added:

```
\rfoot{Appendix A -- Species Summaries}

<<appendix-A, child='appendix-A/appendix-A.Rnw'>>=
@
```

To reference this appendix, use this syntax: `\ref{chap:example.1}` which resolves to appendix **??**. This is also clickable and will take you directly to the appendix. The reference must be defined at the beginning of *appendix-A.Rnw* like this: `\label{chap:example.1}`. This method is repeated for all appendices. They will be lettered in the order in which they appear in example.Rnw, so it is very easy to change the order of appendices and rebuild the document.

# 5 SUMMARY

Here's a reference to an appendix:

## REFERENCES

Carton, J. and Giese, B. 2008. A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). Mon. Weather Rev. **136**: 2999–3017.

Chambers, R. and Dunstan, R. 1986. Estimating Distribution Fucntions from Survey Data. Biometrika **73**: 597–604.

Cheung, W., Brodeur, R., Okey, T. and Pauly, D. 2015. Projecting future changes in distributions of pelagic fish species of Northeast Pacific shelf seas. Prog. Oceanogr. **130**: 19–31.

Foreman, M., Crawford, W.R., Cherniawsky, J.Y., Henry, R.F. and Tarbotton, M.R. 2000. A high-resolution assimilating tidal model for the northeast Pacific Ocean. J. Geophys. Res. **105**: 28,629–28,651.

Haidvogel, D., Arango, H., Budgell, W., nd E. Curchitser, B.C., Lorenzo, E.D., Fennel, K., Geyer, W., Hermann, A., Lanerolle, L., Levin, J., McWilliams, J., Miller, A., Moore, A., T.M. Powell and, A.S., Sherwood, C., Signell, R., Warner, J. and Wilkin, J. 2008. Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the Regional Ocean Modeling System. J. Comput. Phys **227**: 3595–3624.

Jones, M. and Cheung, W. 2014. Multi-model ensemble projections of climate change effects on global marine biodiversity. ICES J. Mar. Sci. **72**: 741–752.

Lecomte, J., Benoît, H., Ancelet, S., Etienne, M., Bel, L. and Parent, E. 2013*a*. Compound Poisson-Gamma versus Delta-Gamma to handle zero-inflated continuous data under a variable sampling volume. Methods Ecol. Evol. **4**: 1159–1166.

Lecomte, J., Benoît, H., Etienne, M. and Parent, E. 2013*b*. Modeling the habitat associations and spatial distribution of benthic macroinvertebrates: A hierarchical Bayesian model for zero-inflated biomass data. Ecol. Model. **265**: 74–84.

Masson, D. and Fine, I. 2012. Modeling seasonal to interannual ocean variability of coastal British Columbia. J. Geophys. Res. **117**.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jovic, D., Woollen, J., Rogers, E., Berbery, E.H., Ek, M.B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D. and Shi, W. 2004. North American Regional Reanalysis(NARR): A long-term, consistent, high-resolution climate dataset for the North American domain, as a major improvement upon the earlier global reanalysis datasets in both resolution and accuracy. Bull. Amer. Meteor. Soc. **87**: 343–360.

Morrison, J., Foreman, M. and Masson, D. 2012. A method for estimating monthly freshwater discharge affecting British Columbia coastal waters. Atmos. Ocean **50**: 1–8.

Perry, R.I. and Smith, S. 1994. Identifying Habitat Associations of Marine Fishes Using Survey Data: An Application to the Northwest Atlantic. Can. J. Fish. Aquat. Sci. **51**: 589–602.

Perry, R., Stocker, M. and Fargo, J. 1994. Environmental effects on the distributions of groundfish in Hecate Strait, British Columbia. Can. J. Fish. Aquat. Sci. **51**: 1401–1409.

Rooper, C. 2008. An ecological analysis of rockfish (*Sebastes* spp.) assemblages in the North Pacific Ocean along broad-scale environmental gradients. Fish. Bull. **106**: 1–11.

Rooper, C. and Martin, M. 2009. Predicting presence and abundance of demersal fishes: a model application to shortspine thornyhead *Sebastolobus alascanus*. Mar. Ecol. Prog. Ser. **379**: 253–266.

Rooper, C., Zimmermann, M. and Spencer, P. 2005. Using ecologically based relationships to predict distribution of flathead sole *Hippoglossoides elassodon* in the eastern Bering Sea. Mar. Ecol. Prog. Ser. **290**: 251–262.

Sigler, M., Cameron, M., Eagleton, M., Faunce, C., Heifetz, J., Hesler, T., Laurel, B., Lindeberg, M., McConnaughey, R., Ryer, C. and Wilderbuer, T. 2012. Alaska Essential Fish Habitat Research Plan: A research plan for the National Marine Fisheries Service's Alaska Fisheries Science Center and Alaska Regional Office. AFSC Processed Rep. 2012-06, 21 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 17109 Pt. Lena Loop Road, Juneau, AK 99801 .

# 6 TABLES

*Table 1. Definitions of quantities associated with trawl survey calculations (Cochran 1977, p.89-92; Smith 1988) and equations in text.*

| | | |
|---|---|---|
| $n_h$ | $=$ | number of hauls or sets in stratum $h$ ($h$ = 1, ...,$L$) |
| $n$ | $=$ | $\sum_{h=1}^{L} n_h$ (in the stratified case), total number of hauls |
| $N_h$ | $=$ | total number of possible sets in stratum $h$ |
| $N$ | $=$ | $\sum_{h=1}^{L} N_h$, total number of possible sets overall |
| $W_h$ | $=$ | $N_h/N$, proportion of the survey area in stratum $h$ |
| $y_{hi}$ | $=$ | number of fish of a particular species caught in set $i$ ($i$ = 1, ...,$n_h$) and stratum $h$;$y_i$ is the same quantity but for the random (nonstratified) survey design |
| $\overline{y}_h$ | $=$ | estimated mean abundance of a particular species of fish in stratum $h$; $\overline{y}$ is the same quantity but for the random (nonstratified survey design) |
| $\overline{y}_{st}$ | $=$ | $\sum_{h=1}^{L} W_h \overline{y}_h$, estimated stratified mean abundance for a particular species of fish |
| $x_{hi}$ | $=$ | measurement for a hydrographic variable in set $i$ of stratum $h$; $x_{hij}$ indexes the measurement for hydrographic variable $j$ in set $i$ of stratum $h$ when more than one hydrographic variable is considered simultaneously |

Table 2. *Cumulative distribution function quantiles for survey temperature-weighted catch.*

| | Quantile | | | | |
|---|---|---|---|---|---|
| | **2.5%** | **25%** | **50%** | **75%** | **97.5%** |
| **Temperature** | 4.93 | 5.63 | 6.13 | 6.93 | 10.53 |
| **Arrowtooth Flounder** | 5.03 | 5.93 | 6.13 | 6.53 | 7.63 |
| **Bocaccio Rockfish** | 5.23 | 5.93 | 6.13 | 6.43 | 7.63 |
| **Canary Rockfish** | 5.63 | 6.03 | 6.33 | 6.73 | 7.53 |
| **Spiny Dogfish** | 5.43 | 6.83 | 8.83 | 10.13 | 12.53 |
| **Dover Sole** | 4.93 | 5.63 | 5.93 | 6.13 | 7.23 |
| **English Sole** | 5.83 | 6.53 | 7.13 | 8.23 | 11.43 |
| **Greenstripe Rockfish** | 5.43 | 5.83 | 6.03 | 6.23 | 6.83 |
| **Longspine Thornyhead** | 4.33 | 4.63 | 4.93 | 5.03 | 5.23 |
| **Pacific Cod** | 5.43 | 6.13 | 6.53 | 7.53 | 10.63 |
| **Petrale Sole** | 5.33 | 6.23 | 6.83 | 7.33 | 8.73 |
| **Pacific Ocean Perch** | 4.83 | 5.33 | 5.53 | 5.83 | 6.23 |
| **Ratfish** | 5.63 | 6.23 | 7.13 | 8.73 | 11.33 |
| **Rock Sole** | 6.53 | 7.93 | 9.03 | 10.23 | 12.93 |
| **Redstripe Rockfish** | 5.23 | 5.83 | 6.13 | 6.33 | 7.53 |
| **Rex Sole** | 5.13 | 5.83 | 6.23 | 6.73 | 7.93 |
| **Sablefish** | 4.63 | 5.23 | 5.83 | 6.43 | 8.33 |
| **Silvergrey Rockfish** | 5.13 | 5.53 | 5.93 | 6.23 | 6.93 |
| **Shortspine Thornyhead** | 4.53 | 5.03 | 5.33 | 5.63 | 6.03 |
| **Widow Rockfish** | 5.43 | 5.73 | 6.33 | 7.13 | 8.83 |
| **Yellowmouth Rockfish** | 5.13 | 5.53 | 5.73 | 6.03 | 6.33 |

Table 4. *Arrowtooth Flounder mean lengths (mm), maturity, sex and total sample weight (kg) $\pm$ standard deviation (SD) from 1984 to 2015.*

| **Survey Year** | $\overline{L}$ | **SD** | $\overline{S}$ | **SD** | $\overline{M}$ | **SD** | $\overline{TotWt_s}$ | **SD** |
|---|---|---|---|---|---|---|---|---|
| 1984 | a | See Table 4 | | | | | | |
| 1987 | a | $\vartheta^2 = 1.538$; $\rho = 0.015$ | | | | | | |
| 1989 | a | $\vartheta^2$ estimated; $\rho = 0.059$ | | | | | | |
| 1991 | a | $\vartheta^2 = 0.962$; $\rho = 0.038$ | | | | | | |
| 1993 | a | $\vartheta^2 = 2.500$; $\rho = 0.100$ | | | | | | |
| 1998 | a | $h$ = Beta($\alpha = 12.7$, $\beta = 5.0$) | | | | | | |
| 2000 | a | $\ln(M)$ = Normal($\ln(0.2)$, $0.05$) | | | | | | |
| 2002 | a | $\ln(M)$ = Normal($\ln(0.2)$, $0.25$) | | | | | | |
| 2003 | a | $\ln(M)$ = Normal($\ln(0.3)$, $0.20$) | | | | | | |
| 2004 | a | $\ln(q_k)$ = Normal($\ln(1.0)$, $1.0$) | | | | | | |
| 2005 | a | $\ln(q_k)$ = Normal($\ln(0.5)$, $1.5$) | | | | | | |
| 2007 | a | $\hat{a}$ = $4.99$ yrs; $\hat{\gamma}$ = $1.27$ yrs | | | | | | |
| 2009 | a | $\hat{a}$ = $6.00$ yrs; $\hat{\gamma}$ = $1.00$ yrs | | | | | | |
| 2011 | Aa | $\hat{a}$ = $6.00$ yrs; $\hat{\gamma}$ = $1.00$ yrs | | | | | | |
| 2013 | a | $\hat{a}$ = $6.00$ yrs; $\hat{\gamma}$ = $1.00$ yrs | | | | | | |
| 2015 | a | $\hat{a}$ = $6.00$ yrs; $\hat{\gamma}$ = $1.00$ yrs | | | | | | |

*Table 3. Example using xtable with some pseudo-random seeded numbers. The function get.align makes the left column justified left and the rest justified right which is how most tables giving values are shown.*

| **ID** | $R_{s=1}$ | $R_{s=2}$ | $R_{s=3}$ | $R_{s=4}$ | $\overline{R}$ | $\sigma$ |
|:---|---:|---:|---:|---:|---:|---:|
| 1 | 19.56 | 15.11 | 18.43 | 1.09 | 13.55 | 8.52 |
| 2 | 2.63 | 14.49 | 12.90 | 16.82 | 11.71 | 6.26 |
| 3 | 19.54 | 12.88 | 8.27 | 1.65 | 10.59 | 7.54 |
| 4 | 2.00 | 7.83 | 15.02 | 12.43 | 9.32 | 5.72 |
| 5 | 14.37 | 2.22 | 4.06 | 14.94 | 8.90 | 6.70 |
| 6 | 4.97 | 4.50 | 11.41 | 17.99 | 9.72 | 6.35 |
| 7 | 10.67 | 16.40 | 2.16 | 17.77 | 11.75 | 7.10 |
| 8 | 3.97 | 15.28 | 3.83 | 6.93 | 7.50 | 5.38 |
| 9 | 14.11 | 2.63 | 1.99 | 2.66 | 5.35 | 5.85 |
| 10 | 7.24 | 16.63 | 11.03 | 6.77 | 10.42 | 4.56 |
| 11 | 9.65 | 11.02 | 14.15 | 4.09 | 9.73 | 4.20 |
| 12 | 5.60 | 14.69 | 18.49 | 12.08 | 12.72 | 5.43 |
| 13 | 19.37 | 14.96 | 3.11 | 8.09 | 11.38 | 7.21 |
| 14 | 3.09 | 3.06 | 14.50 | 8.60 | 7.31 | 5.45 |
| 15 | 6.30 | 12.32 | 18.37 | 9.38 | 11.59 | 5.14 |
| 16 | 8.81 | 16.03 | 8.28 | 14.20 | 11.83 | 3.87 |
| 17 | 9.17 | 11.69 | 4.64 | 16.03 | 10.38 | 4.76 |
| 18 | 4.83 | 18.61 | 7.07 | 2.08 | 8.15 | 7.27 |
| 19 | 12.64 | 1.01 | 6.02 | 1.14 | 5.20 | 5.48 |
| 20 | 13.61 | 10.84 | 1.68 | 19.64 | 11.44 | 7.47 |

*Table 5. Sensitivity cases for $q_k$; posterior quantiles.*

| **Index** | **Sensitivity 10** | | | **Sensitivity 11** | | |
|---|---|---|---|---|---|---|
| $\mathbf{q_k}$ | **2.5%** | **50%** | **97.5%** | **2.5%** | **50%** | **97.5%** |
| QCSSS | 0.081 | 0.158 | 0.508 | 0.029 | 0.083 | 0.226 |
| HSMAS | 0.079 | 0.121 | 0.155 | 0.035 | 0.081 | 0.136 |
| HSSS | 0.070 | 0.118 | 0.200 | 0.027 | 0.067 | 0.136 |
| WCVISS | 0.061 | 0.104 | 0.172 | 0.022 | 0.059 | 0.118 |

# 7 FIGURES

## 7.1 SURVEY DATA

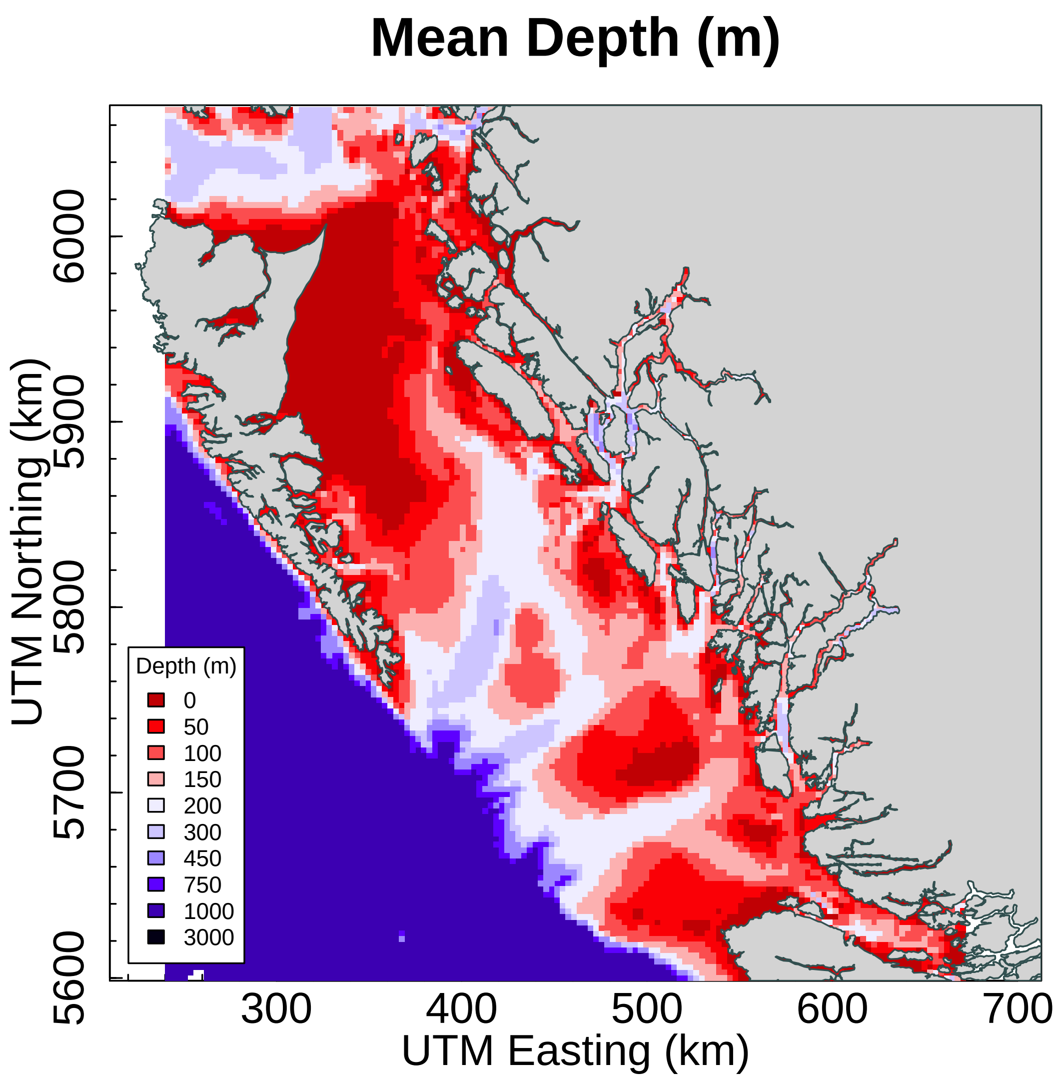



*Figure 1. Mean ocean depth (m) in grid cells of 3 km x 3 km for fishing events from all groundfish trawl surveys for all years.*

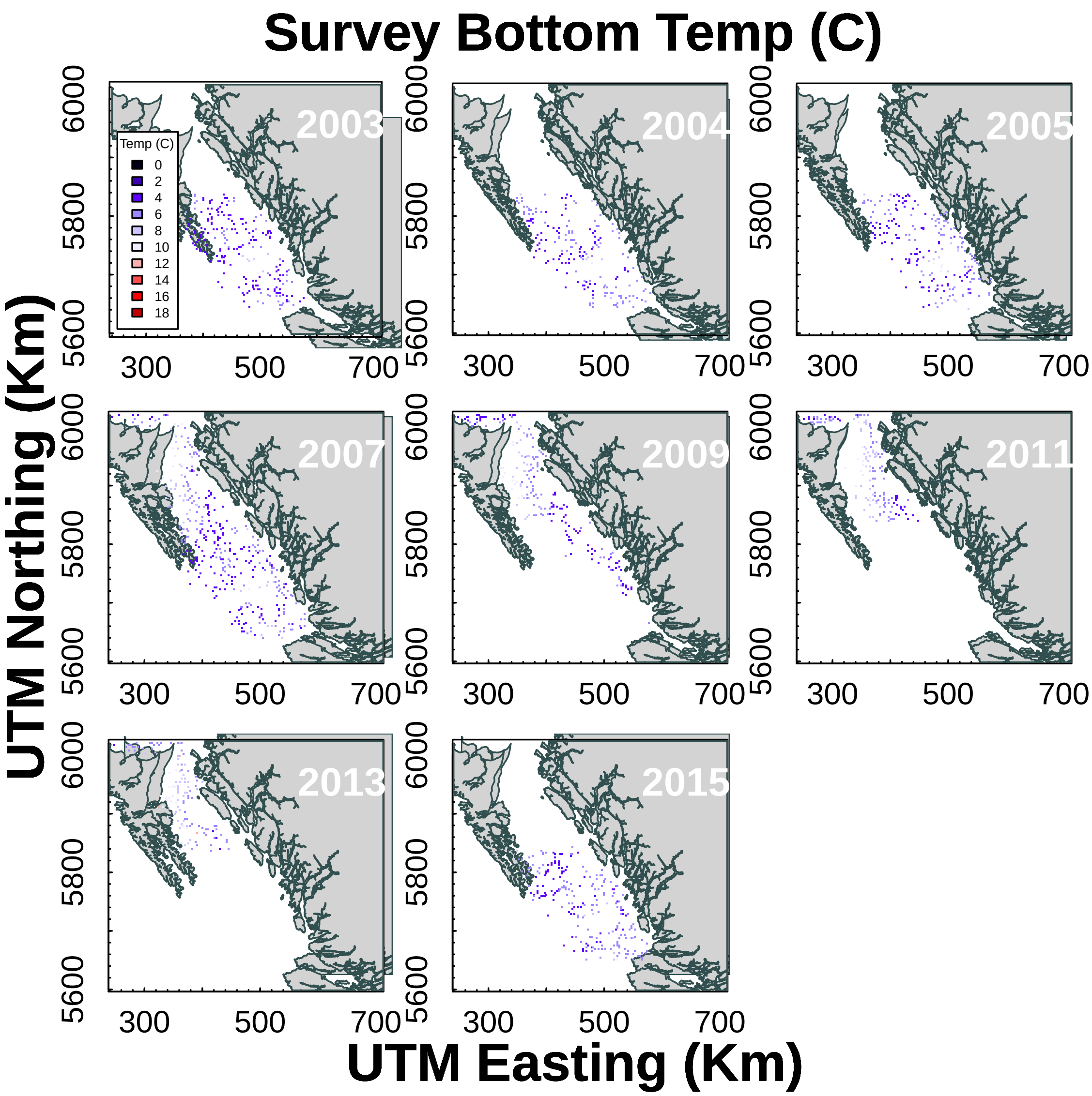



*Figure 2. Mean ocean temperature (°C) in grid cells of 3 km x 3 km for fishing events from groundfish trawl surveys from 2003 to 2011.*

## 7.2 BCMCA DATA

## 7.3 ROMS

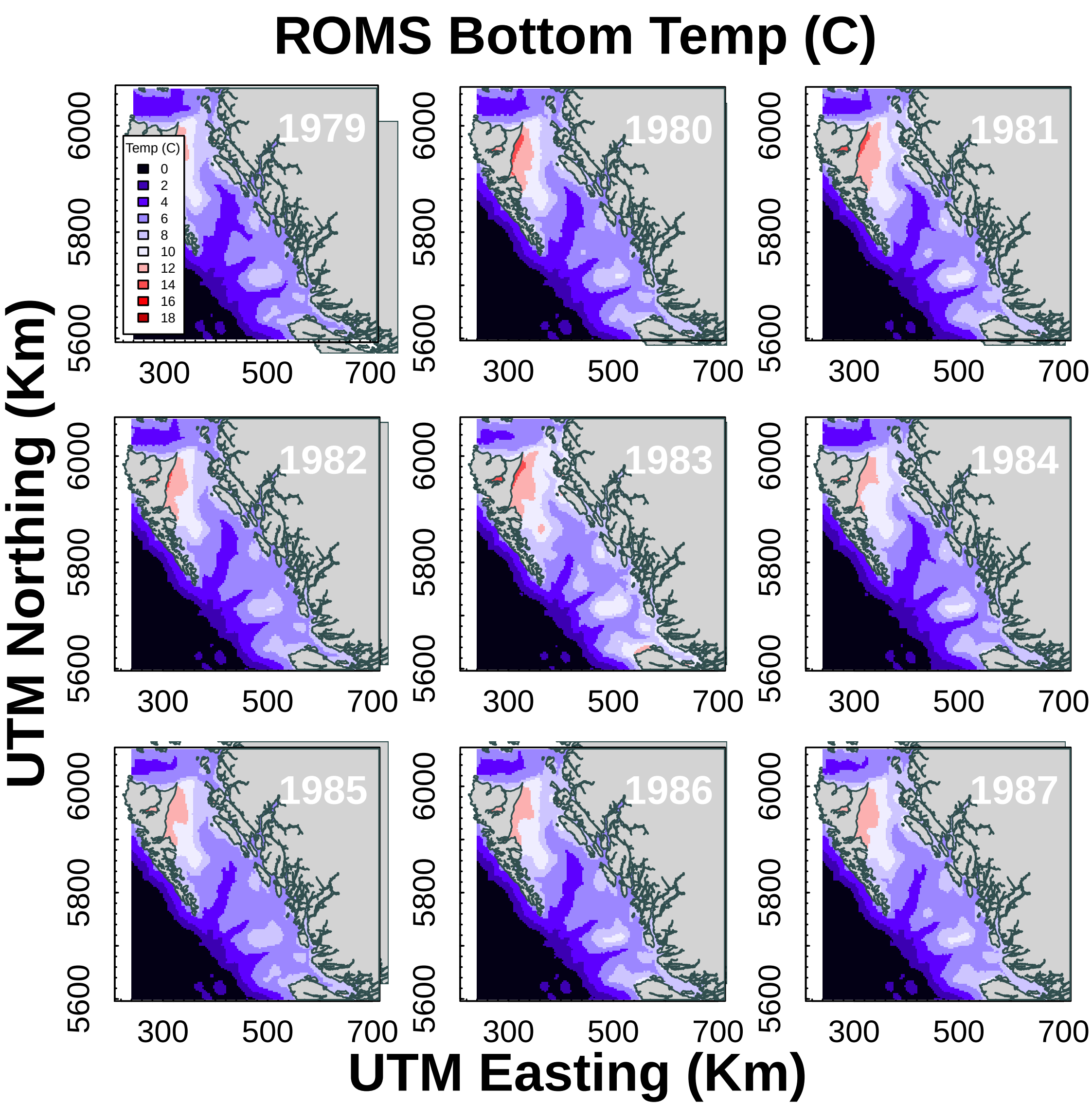



*Figure 3. Summer ROMS ocean bottom temperature (°C) from 1979 to 2011.*

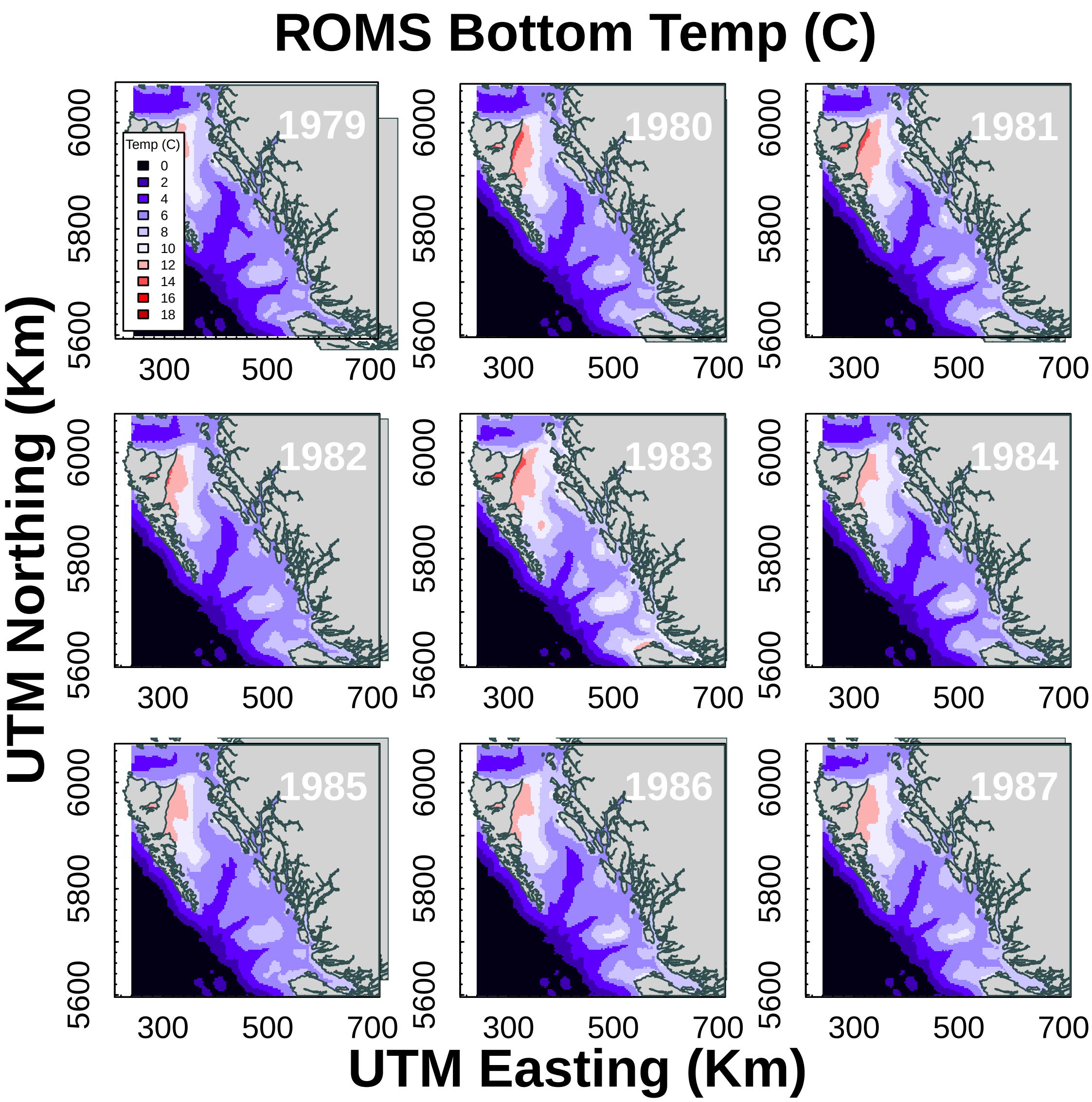



*Figure 4. ROMS mean bottom temperature (°C) in grid cells of 3 km x 3 km for fishing events from all groundfish trawl surveys 1979 to 1987.*

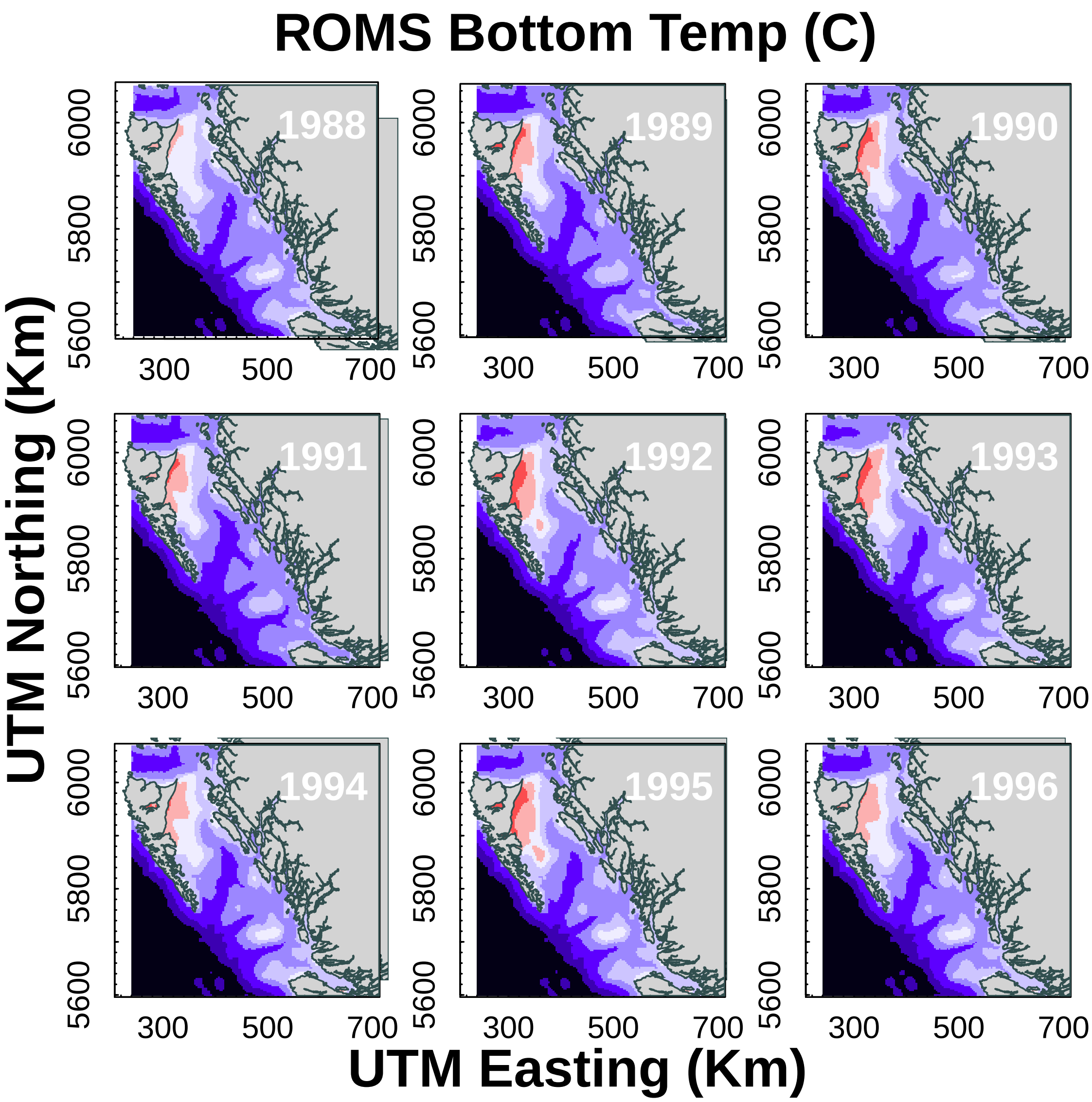



*Figure 4 (Continued). ROMS mean bottom temperature (°C) in grid cells of 3 km x 3 km for fishing events from all groundfish trawl surveys 1988 to 1996.*

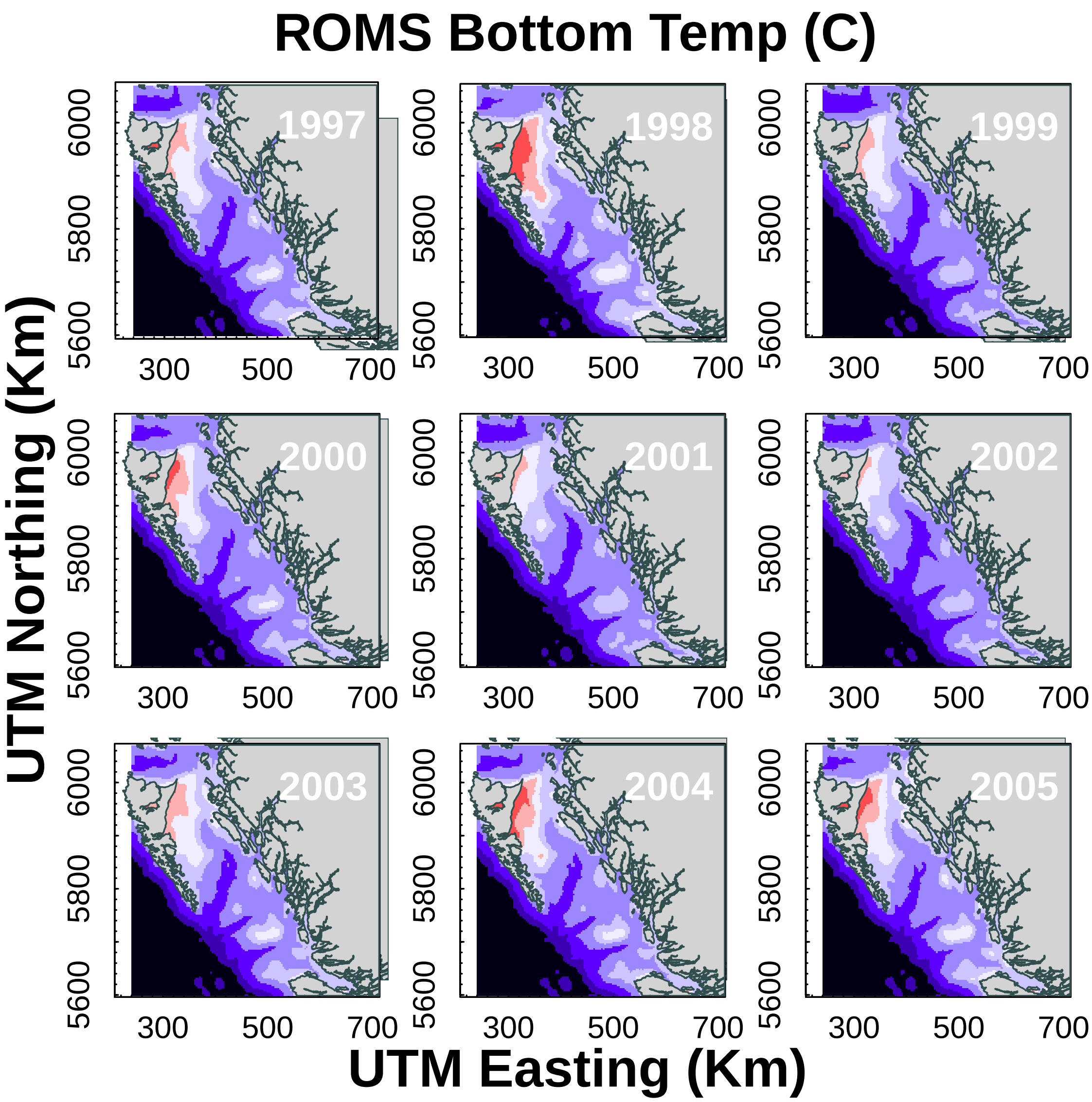



*Figure 4 (Continued). ROMS mean bottom temperature (°C) in grid cells of 3 km x 3 km for fishing events from all groundfish trawl surveys 1997 to 2005. The grey dashed line is the 1:1 line.*

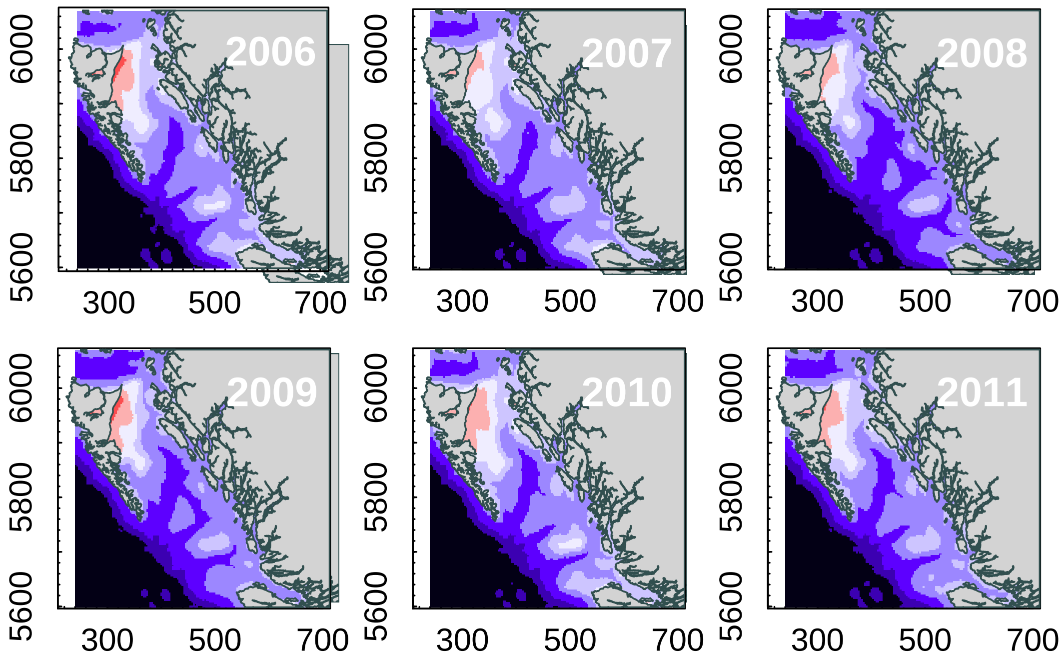



*Figure 4 (Continued). ROMS mean bottom temperature (°C) in grid cells of 3 km x 3 km for fishing events from all groundfish trawl surveys 2006 to 2011.*

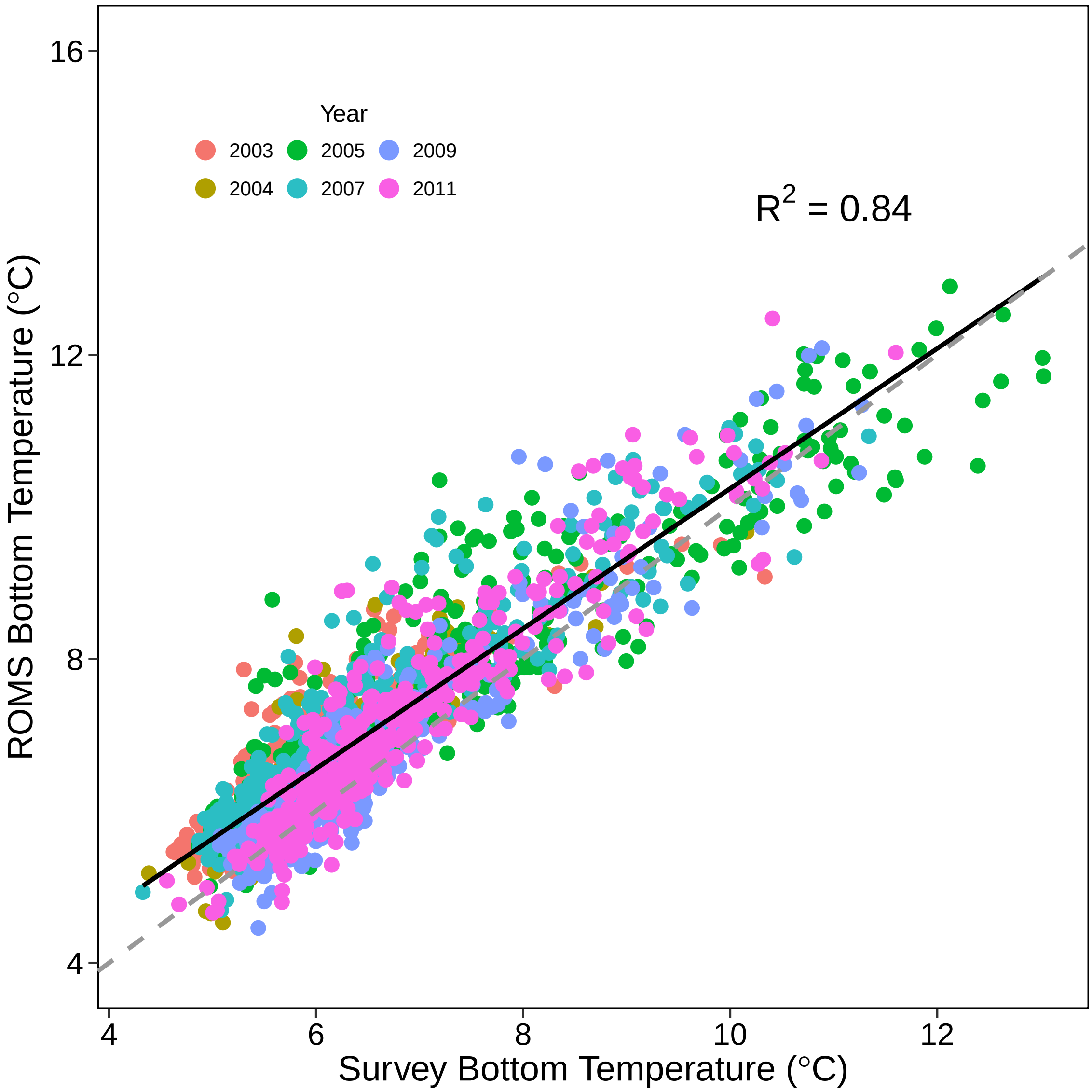



*Figure 5. Relationship between ROMS and survey ocean bottom temperature (°C) for 2003 to 2011.*

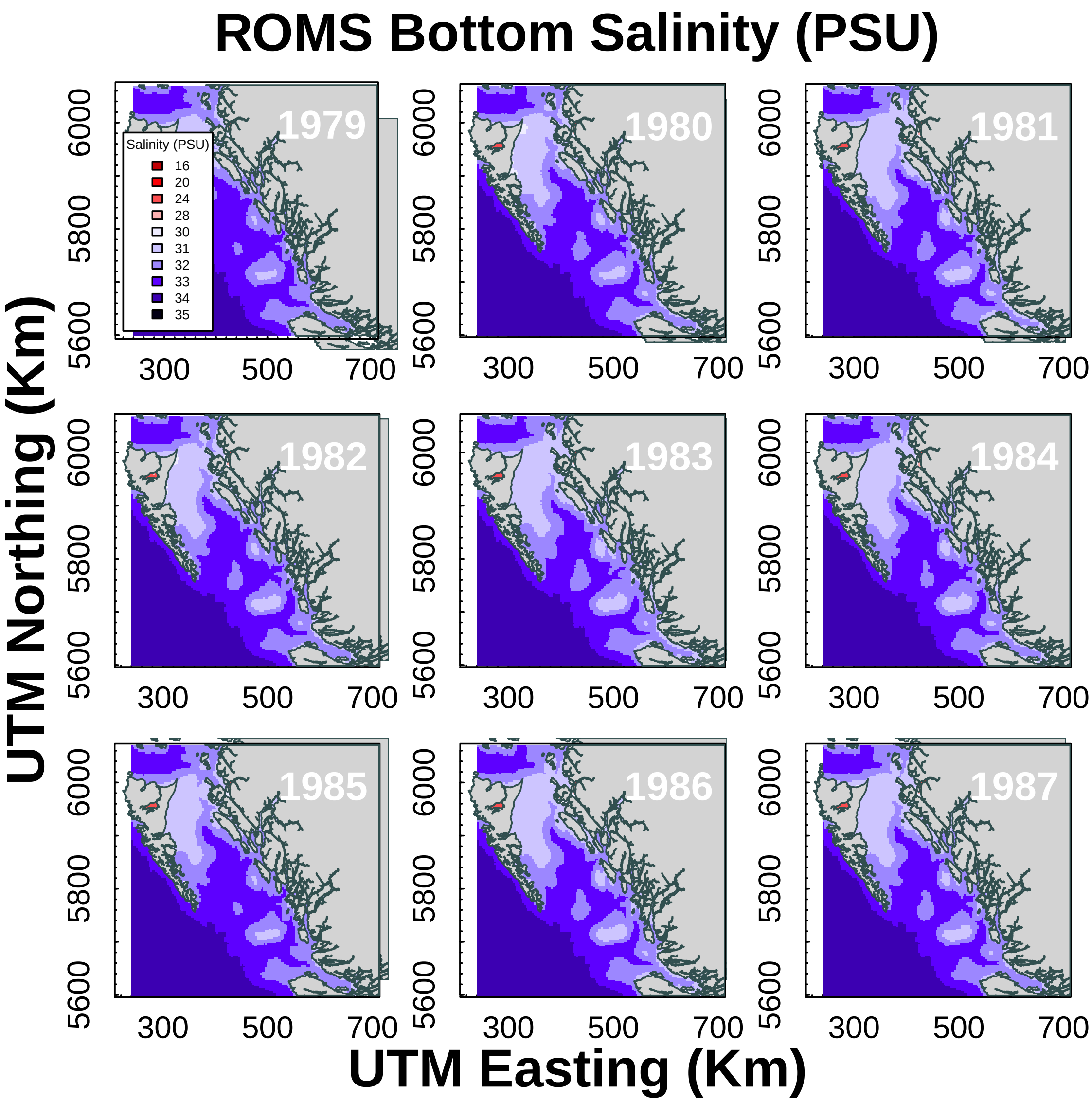



*Figure 6. ROMS mean bottom salinity (PSU) in grid cells of 3 km x 3 km for fishing events from all groundfish trawl surveys 1979 to 1987.*

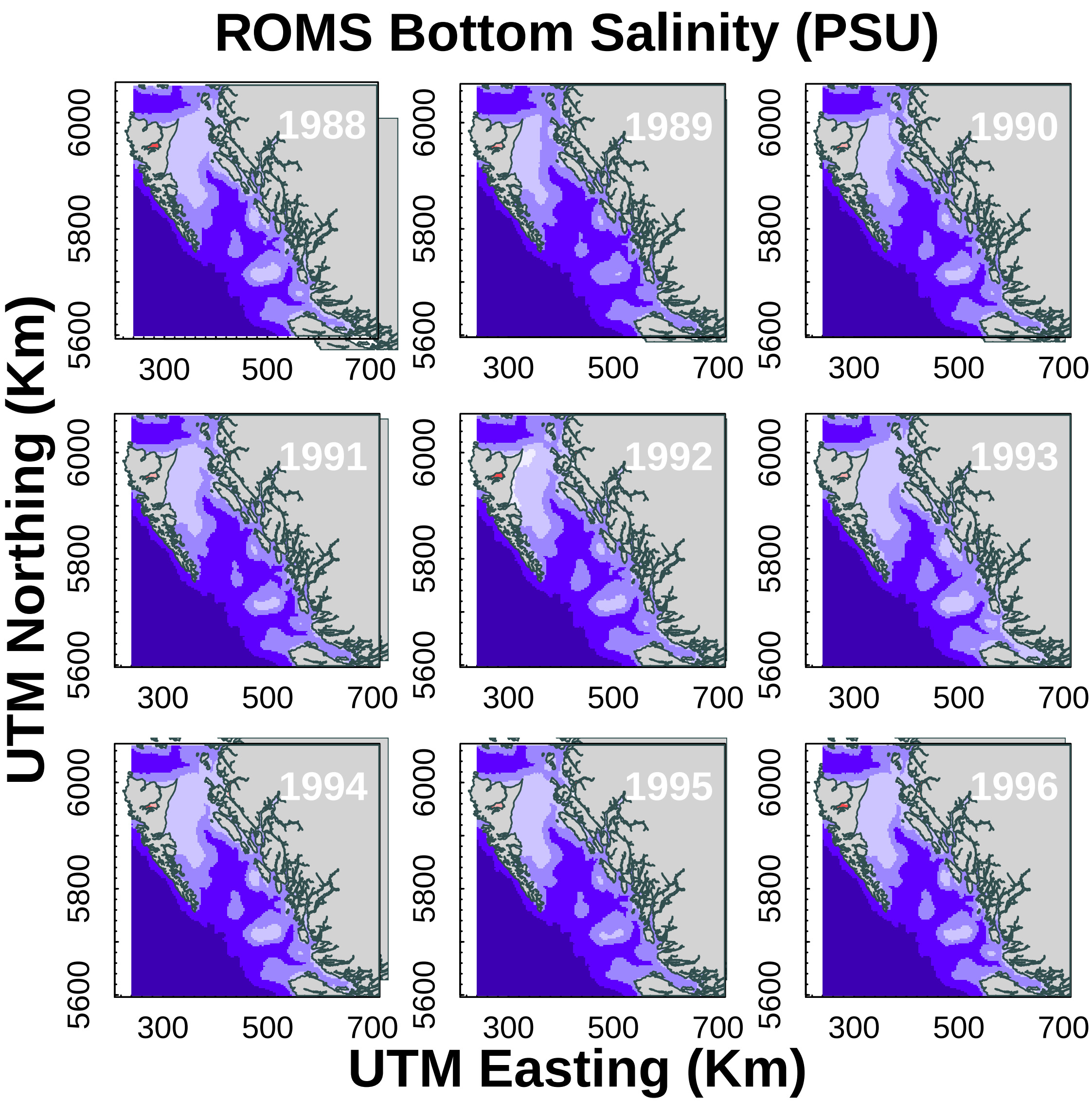



*Figure 6 (Continued). ROMS mean bottom salinity (PSU) in grid cells of 3 km x 3 km for fishing events from all groundfish trawl surveys 1988 to 1996.*

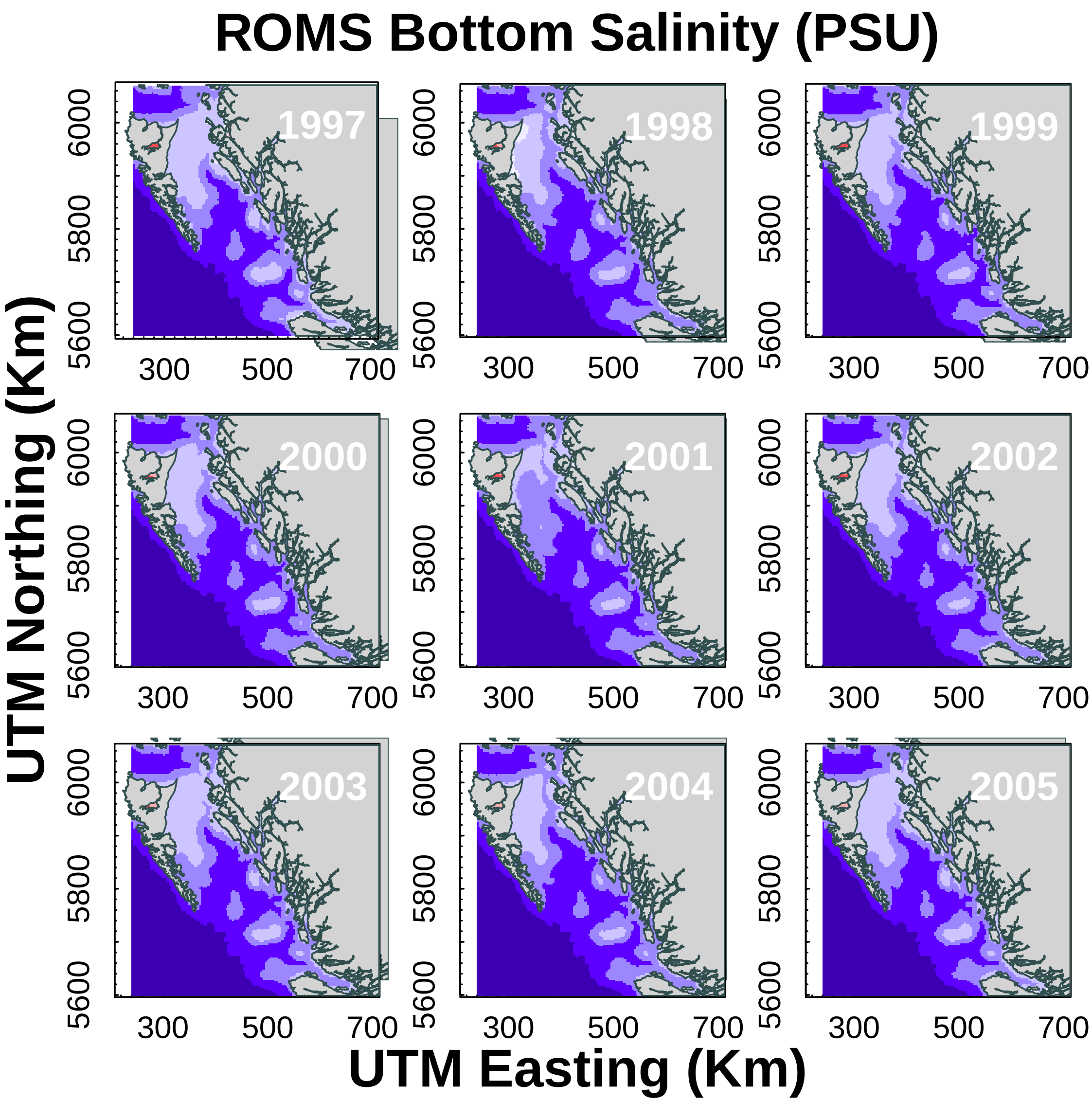



*Figure 6 (Continued). ROMS mean bottom salinity (PSU) in grid cells of 3 km x 3 km for fishing events from all groundfish trawl surveys 1997 to 2005.*

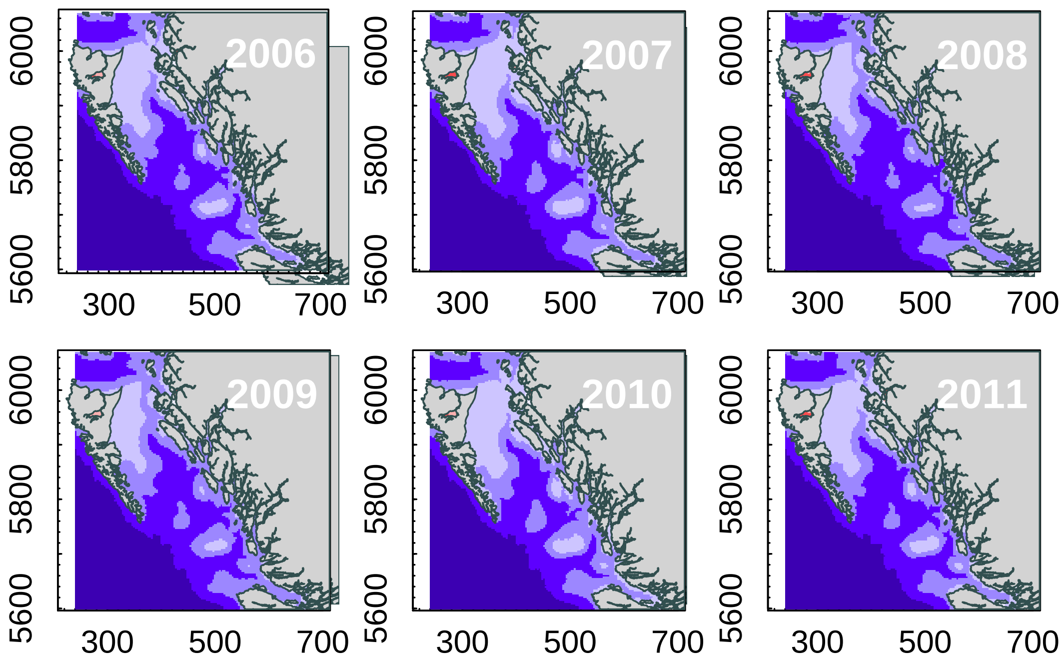
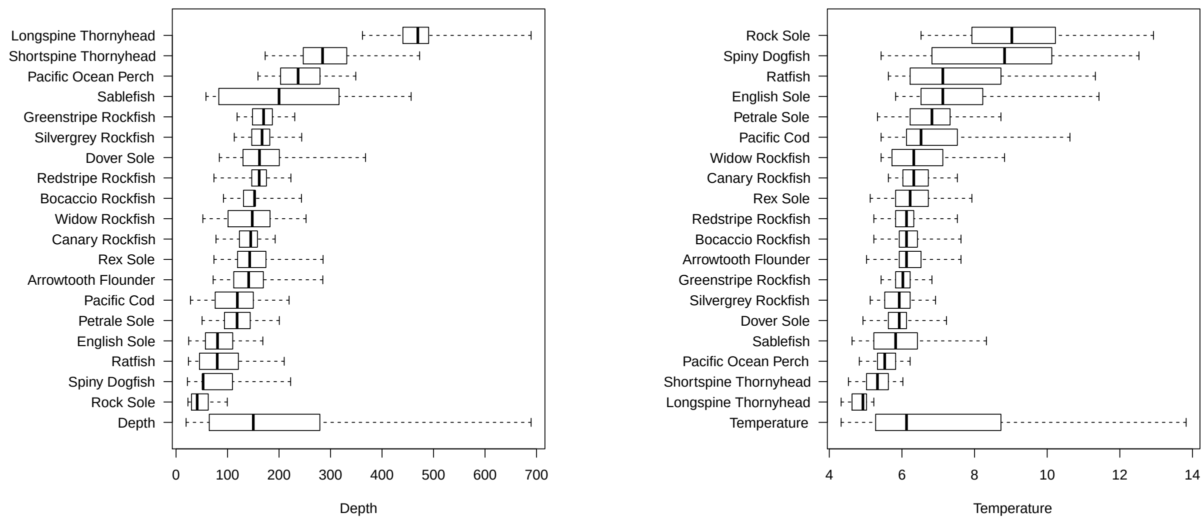



*Figure 6 (Continued). ROMS mean bottom salinity (PSU) in grid cells of 3 km x 3 km for fishing events from all groundfish trawl surveys 2006 to 2011.*




*(a) Plots of quartiles of available depth and groundfish species based on average survey depth (m).*

*(b) Plots of quartiles of available temperature and groundfish species based on average survey temperature ($^{\circ}C$).*

*Figure 7. Quartiles of available habitat and groundfish species based on ROMS 3 km gridded data. Whiskers extend to the 2.5th and 97.5th percentiles, the end of the boxes extend to the 25th and 75th percentiles, and the vertical black bar represents the median (50th percentile). Species are arranged in increasing order based on the value of their median.*

## 7.4 CUMULATIVE DISTRIBUTION FUNCTIONS