Spatiotemporal Models to Improve Stock Size Abundance Indices Derived from Stratified Bottom-Trawl Surveys, with an Application to Witch Flounder

by

© Fatemeh Hatefi

A Thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of

Master of Science Department of Fisheries Science

Memorial University of Newfoundland

February 2021

St. John's, Newfoundland and Labrador, Canada

Abstract

A fundamental part of a fishery stock assessment is an abundance index that reflects stock trends and changes across time. An abundance index is a measurement that is proportional to stock size, and the constant of proportionality is assumed to be the same over time and space. The abundance index can be obtained from the information provided by research vessel surveys of fisheries resources. Changes in survey coverage are a common problem in such surveys, and these changes create uncertainty in interpreting the trend of abundance indices. This is because when there are changes in survey coverage, the changes in survey abundance indices over time may partially reflect changes in the areas surveyed in addition to changes in stock size. Thus, changes in survey coverage over time may lead to errors in fish stock assessment.

This thesis provides improved survey indices of stock size using spatiotemporal models that address changes in survey coverage for witch flounder on the east coast of Newfoundland and Labrador. I first estimate the design-based abundance indices to use as a baseline to compare the performance of indices derived from spatiotemporal models. Second, I develop spatiotemporal models and select the best-fitting model via model comparison to estimate the spatial variation of mean catches. These models can predict and interpolate the abundance indices for missing areas and years. These models enable us to fill the data gaps in some survey years to provide more reliable survey indices and longer time-series for the witch flounder stock assessment.

Acknowledgments

I would like to express my sincere gratitude to my supervisors Dr. Noel Cadigan and Dr. Jin Gao for guidance and help throughout my M.Sc study. I also thankfully acknowledge the funding for this research that was provided by the Ocean Frontier Institute, through an award from the Canada First Research Excellence Fund. Lastly, I thank my parents and siblings for their constant support and encouragement.

Contents

List of Tables

List of Figures

List of Abbreviations

AR(1)	Autoregressive process of order 1
DFO	Fisheries and Oceans Canada
GLMs	Generalized linear models
GLMMs	Generalized linear mixed models
GMRFs	Gaussian Markov random fields
MNPT	Mean catch number per tow
MWPT	Mean catch weight per tow
NAFO	Northwest Atlantic Fisheries Organization
PM	Precision matrix
SRS	Stratified random sampling
TMB	Template model builder

Chapter 1

Overview and Objectives

1.1 Introduction

Fisheries activities play a crucial role in providing food and growing economies throughout the world. FAO[1] statistics indicate that the yield of marine capture fisheries was around 80 million tonnes (t) in 2017 (FAO yearbook, 2017). In Canada, one of the leading provinces in producing fish and seafood is Newfoundland and Labrador (NL). Fish landings (i.e. the catches of marine fish landed in domestic or foreign ports) for NL in 2017 were 194,883 t (Seafood Industry Year in Review, 2017), which was about one-quarter of the total capture production for Canada in 2017 (Canada's Fisheries Fast Facts, 2018). Despite the development of new industries in NL, commercial fisheries are still counted as a significant economic generator. It is, therefore, vital to ensure the long-term sustainability of fisheries for this province.

Fisheries are managed by different stock units. A stock is a fish population with specific features in terms of migration pattern, spawning grounds, and a distinct fishery. The stock size is the total abundance or biomass of the fish. Stock size changes over time because of changes in many factors, such as recruitment, natural mortality, and

[1]Food and Agriculture Organization of the United Nations.

fishing mortality. It is very challenging to measure a stock size due to fish movement and distribution.

Fisheries management, generally, uses different techniques such as stock assessment to generate information about stock status. Stock status is quantified by the comparison between the stock abundance and productivity. For example, fish stock may be classified as over-exploited, fully exploited, or under-exploited if the stock abundance (e.g. biomass, which is total weight of all individuals in a stock or a population) is less than 40 percent of the unfished biomass, between 40-60 percent, or greater than 60 percent, respectively (Ye, 2011). We rely on stock assessments to infer the status of a stock and produce catch options advice for fisheries management. For example, Fisheries and Oceans Canada (DFO) stock assessment advice for the Atlantic Cod (*Gadus morhua*) stock in the Northwest Atlantic Fisheries Organization (NAFO), Subdivisions 4T and 4Vn (DFO, 2016b) was that the removals by all human sources should be kept to the lowest possible level because the stock size was estimated to be much less than the limit reference point. Another example is the assessment of redfish stocks (*Sebastes fasciatus* and *S.mentella*) in NAFO Subdivisions 3Pn and 4Vn in 2015, which determined that almost 50% of the fish in the 2011 cohort should be over 22 cm, the minimum harvest size, by 2018 (DFO, 2016a).

A stock assessment provides scientific support for fisheries managers by the determination of the stock status. Stock size indices are the primary input to fish stock assessments. A stock size index is a relative measure of abundance (i.e. total number) or biomass (i.e. total weight) of fish in a stock (Blackhart et al., 2006). An assessment of a stock and studying the stock size indices over a period of time facilitate descriptions of the past and current status of the stock as well as predictions for future stock status (Cooper, 2006). Based on the stock status indicators (e.g. spawning stock biomass, SSB), harvest control rules (HCRs) are provided to scientifically and sustainably manage fisheries. However survey indices are sometimes used as a direct input into HCRs (e.g. Greenland halibut

(Reinhardtius hippoglossoides) stock in Subarea 2 + Divisions 3KLMNO, see Annex I.F in NAFO, 2020). HCRs determine how much fishing is allowed based on stock status. It is generally classified as healthy, cautionary, and critical (Berger et al., 2012). HCRs can be used to set the TAC for a stock. Accurate abundance indices that are used in stock assessment reflect real changes in stock size that can lead to precise determination of stock status and the TAC. It is, therefore, crucial to obtain reliable stock size indices for stock assessments.

1.1.1 Stock size survey indices

Stock size indices are obtained from two primary data sources: fishery-independent data and fishery-dependent data. Fishery-independent data that are obtained from scientific research surveys are widely recognized to provide more reliable stock size indices compared to fishery-dependent data (e.g. commercial fishery data). A fundamental assumption of an abundance index in fish stock assessment is that the index is proportional to stock size, that is, index = q (catachability) times stock size. The q may also be assumed to be constant in time and space. This assumption usually requires that all other factors that affect catches (i.e. vessels, gears, fishing locations, etc.) are standardized to remove their effects on the index. In most scientific research surveys these factors are fixed by design, and changes in survey catch rates should directly indicate changes in stock size. Fishery-independent data can lead to effective stock monitoring and assessment. For example Dennis et al. (2015) showed for tropical rock lobster (Panulirus ornatus) that fishery independent surveys are worth the money. McDonald and Smith (1997) reviewed methods to evaluate cost and benefits for fisheries research, and they showed the benefit of conducting surveys to estimate absolute abundance. The results of Punt et al. (2002) in assessing the management benefit of fishery-independent surveys in Australia's shark fisheries also confirmed the value of conducting surveys compared with relying on commercial catch rate as an index of abundance.

Stock size survey indices that are derived from a statistical survey sampling design are preferable as they are unbiased (i.e. an estimate is unbiased when its expected value over all possible samples selected with the design is equal to the actual population value; Kimura and Somerton, 2006). However, sometimes we do not have access to survey data, because of specific situations such as wide-ranging and hard to survey stocks, or when economic or logistic constraints prevent implementation of research surveys. In such situations, stock assessment uses stock size indices derived from fishery-dependent data, such as catch per unit effort.

1.1.2 Witch flounder

In this thesis, I focus on witch flounder (*Glyptocephalus cynoglossus*) that have supported large-scale fisheries in the Northwest Atlantic in the past. Witch flounder are a long-lived and right-eye flounder that are found on both east and west sides of the North Atlantic Ocean. In the Northwest Atlantic, the distribution of witch flounder extends from Labrador (Hamilton Bank) to North Carolina. They live in soft bottom substrates, such as sand, clay or mud, within wide range of depths from 100 to 1200 metres (Bowering, 1987). Figure 1.1 is a map of NAFO Sub-areas 2 and 3, included Divisions 2J3KL, the distribution area of witch flounder in this study (i.e. the stock area). They are distributed along the shelf edge and in deeper channels around the banks, and primarily in Division 3K. Witch flounder are most abundant in warm waters within the bottom temperature range between 3.2°C and 3.6°C (Bowering, 1990).

The witch flounder fishery was established in the early 1940s in Newfoundland. In NAFO Divisions 2J3KL, the fishery began in the 1960s with a total landings from 1000 t in 1963 to over 24000 t in 1973 (Bowering, 1990), and then declined considerably to 137 t in 1994. As a result of this drastic decline in total landings, a moratorium was imposed on the directed witch flounder fishery within Canadian waters (within 200 nautical miles) since 1995. The moratorium was further extended to the NAFO regulatory

area in 1998. Since then, witch flounder have only been caught as by-catch in other groundfish (i.e. demersal fish that live near the bottom of seas) fisheries, particularly the Canadian Greenland halibut fishery. To assess the status of important fisheries species including witch flounder, DFO in the Newfoundland and Labrador region conducts an annual Autumn Research Vessel Stratified Random Multi-Species Bottom-trawl survey in NAFO Sub-areas 2 and 3.

1.2 Research trawl surveys

Bottom-trawl surveys are designed to sample demersal multi-species fish communities by towing a trawl along or close to the seafloor for a standard unit of time or distance (Kimura and Somerton, 2006). These surveys monitor fish populations in terms of annual changes in distribution, abundance, and population structure (e.g. age and length) over a specified time of a year. Survey indices of stock size are derived from data that were produced by these surveys, and the trends of these indices are used to evaluate the stock status via stock assessment. Bottom-trawl surveys have been conducted for about 70 years in the Northwest Atlantic to provide biological and statistical information on groundfish for stock assessment and management. In the Northwest Atlantic, stratified-random sampling (SRS) with proportional allocation to strata area is a standard design used to conduct research vessel (RV) surveys (Cadigan, 2011).

1.2.1 Stratified random sampling (SRS)

SRS is a powerful and flexible sampling design widely used in scientific surveys. In stratified sampling, a population is divided into non-overlapping subgroups, i.e. strata, such that the responses (i.e. catches for trawl surveys in this study) within a stratum are as similar as possible. Survey sets are typically distributed using a proportional allocation scheme such that the number of survey sets allocated to each stratum is proportional to

the stratum area (e.g. Rideout and Ings, 2018). However, at least two survey sets must be sampled from every stratum to allow computation of variance estimates (Cochran, 2007), which I describe in the next chapter. Survey sets are the standardized fishing hauls at the sampling sites (i.e. sampling units within each stratum) that are selected randomly. Random sampling within strata has several advantages over other sampling techniques. First, random sampling provides an unbiased estimate for stock size (Thompson, 2012). Second, for SRS, when catches in each stratum are more homogeneous (i.e. sampling units within a stratum are as similar as possible or the variability of sampling units within a stratum is small), then a relatively small sample in each stratum can provide a precise estimate of a stratum parameter. Therefore, an accurate estimate for the whole population can be obtained by combining the stratum-level estimates (Grosslein, 1969). Third, the variance of SRS with proportional allocation will be smaller than from a comparable simple random design if the mean square error within strata is smaller than the mean square error among strata (Cochran, 2007). This is a potential advantage of SRS, because a sample design generally gives a more efficient estimate than another design if its associated variance is smaller for the same total sample size. Moreover, the calculation of the population stratified mean (or total) and their associated variance for each combination of strata is straightforward. For these reasons, SRS has been the most preferred design for RV surveys (Grosslein, 1969).

In the Northwest Atlantic, RV surveys are mainly stratified by location and depth range. Depth is a precisely measurable variable and has been found to correlate with groundfish distribution (Grosslein, 1969). Strata can also be defined by other variables such as bottom type and geographical regions in addition to location and depth range. Strata maps for each NAFO division of the stock area (Divisions 2J3KL) are shown in Figures 1.2-1.4.

In the next chapter, I show how abundance indices (e.g. abundance, biomass and mean catch number/weight per tow) are estimated. Note that to estimate abundance

and biomass, a swept area is required. In a trawl survey, a swept area is the area which is swept by the trawl that is towed in a specified distance along the seafloor. The standard swept area is calculated by multiplying the distance between the wings of a trawl (i.e. wing spread) and the tow distance. For DFO trawl surveys in NAFO Divisions 2J3KL, the swept area is around 0.007274 square nautical miles or about 2.5 hectares. When RV and gear type change in surveys, the swept area changes. Therefore, many organizations have conducted scientific research trawl surveys using a standard RV and gear type since the 1970s (Chen et al., 2004).

1.2.2 Research problem

RV surveys are designed to be comparable from year to year, but some planned and unplanned changes occur over time. Planned changes include modifying the biological objectives of the survey. Unexpected changes may also happen because of vessel breakdowns and poor weather conditions. As a result, surveys may not be fully completed, which results in changes in coverage of surveys over time.

Changes in survey coverage may have a significant impact on the reliability of survey indices of stock size. Changes in survey coverage are a common problem in which there may be missed strata (i.e. non-sampled strata during the survey), and/or reduced number of survey sets per stratum as well as adding or removing some strata due to modifying the survey objectives (Brodie, 2005). For example, the stock assessment report of Capelin (*Mallotus villosus*) in the Barents Sea in 2014 indicated that ice coverage prevented the full survey coverage (ICES, 2015). Survey coverage changes may create uncertainty about the trend of abundance and biomass indices. Thus, changes in survey coverage over time may lead to a less accurate fish stock assessment. For example, Healey and Dwyer (2005) showed that the incomplete autumn survey in Divisions 3LNO in 2004 affected several stocks with analytical assessments, primarily Greenland Halibut and American Plaice (*Hippoglossoides platessoides*). Therefore, it is crucial to investigate the influences

of survey coverage changes on stock size indices and apply proper statistical models to estimate survey indices of stock size accounting for these survey defects.

Spatial distributions of witch flounder (i.e. catch number per set) during this time series (i.e. 1977-2017 in this study) are shown in Figures 1.5-1.15. There have been changes in survey coverage for this stock over time. For example, the surveys were extended from covering NAFO Divisions 2J3K to 2J3KL since 1983 (Figure 1.6).

The average catch per stratum for each NAFO division is illustrated in Figures 1.16-1.18. These figures better indicate the problems of changes in survey coverage over time compared to Figures 1.5-1.15. For example, in NAFO Division 2J, there were changes in survey coverage because of adding strata 237, 238 (depth range less than 200 metres), 239 (depth range 501-750 m) and 240 (depth range from 401 to 500 m) since 1993, as well as strata 220, 225, 232 (depth range 1001-1250 m), and 221, 226, 233 (depth range 1251-1500 m) since 1995 (Figure 1.16). Survey coverage was extended to inshore areas in NAFO Division 3K and areas deeper than 1000 metres (strata 648, 649, 653, 654 and 644) as well (Figure 1.17). In NAFO Division 3L, deep-water strata (737-751, in-depth range 732-1463 m), and inshore strata (784-800, in-depth range from 30 to 183 m) were added after 1995. However, in recent years, these strata have not been sampled (Figure 1.18).

The survey coverage percentage, derived by dividing the total surveyed area in Divisions 2J3KL each year by their maximum surveyed area (Figure 1.19), may not indicate the coverage issues' importance. The total area covered is one part of the picture because not all strata are of equal importance for witch flounder, and sometimes it is essential to account for the prevalence of witch flounder in particular strata. Therefore, models to adjust for changes in survey coverage during this period are required for successful stock assessment.

Two approaches have been widely used to address the problem of changes in the survey coverage: generalized linear models (GLMs) and spatiotemporal models. DFO scientists

obtained estimates of biomass and abundance for non-sampled strata using GLMs, i.e. multiplicative models (Lilly et al., 1998). A multiplicative model is simply a log-additive model in which the log of survey catch per tow is modelled as linear regression on year and strata effects for neighbouring years and strata. In this model, the mean catches in missed strata in incomplete coverage years can be inferred using the distribution of catch across strata in full coverage years (Lilly et al., 2005).

1.2.3 Spatiotemporal models

Spatiotemporal modelling is a more flexible approach to deal with many spatial issues in survey sampling. Over the last decade, fishery scientists have developed spatiotemporal models to account for different spatial problems associated with surveys. For example, they have utilized the knowledge of the spatial distribution of fish and the temporal changes to infer the catches in missing strata. Lewy and Kristensen (2009) applied the Log Gaussian Cox Process (LGCP) model to analyze the spatial distribution of cod in the North Sea. The LGCP model belongs to the family of parametric geostatistical models including Gaussian latent variable models. The LGCP is a flexible count data model with a multivariate Poisson log-normal distribution. More specifically, it is a mixture of observations that are Poisson-distributed with mean densities that follow a log-normal distribution. Lewy and Kristensen showed that this model could predict and interpolate unobserved densities at any location in the area. They estimated parameters using the maximum likelihood method. These parameters allow interpolation and prediction of unobserved densities at any point in space and time. Nielsen et al. (2014) modified the LGCP model to estimate correlations between observations based on time, space, and fish size. This modification facilitated expanding the LGCP models. Shelton et al. (2014) showed that incorporating relevant environmental covariates (e.g. bottom temperature, depth) into a statistical model could more accurately estimate abundance indices. They developed a general semi-parametric mixture model as a combination of measured

habitat variables and a non-parametric smoothing term. Applying the model for dark blotched rockfish (*Sebastes crameri*) survey data, they showed that the spatial habitat model yielded more precise, reasonable (in terms of biology), and interpretable estimates of stock size (i.e. abundance) than the classical design-based method. Thorson et al. (2015) utilized a geostatistical model to improve the estimation of abundance indices. Applying the geostatistical model, they showed that densities in non-sampled areas could be estimated by using habitat information and nearby samples. Cao et al. (2017) estimated the abundance index for northern shrimp (*Pandalus borealis*) in the Gulf of Maine using the spatiotemporal model developed by Shelton et al. (2014). They utilized Gaussian random fields to account for spatial variation in population densities. They showed that spatiotemporal index standardization could produce an abundance index that was more reliable than the index derived from the design-based method by incorporating spatial variation in densities. Therefore, they concluded that stock size indices could benefit from predicting non-sampled areas (i.e. filling spatial gaps), using a spatiotemporal model that explains the spatial variation in population densities.

The goal of this thesis is to improve stock size survey indices using spatiotemporal models, with application to witch flounder in NAFO Divisions 2J3KL. Indeed, I develop a spatiotemporal model that utilizes Gaussian Markov random fields to account for spatial variation in witch flounder caught in surveys from 1977 to 2017. In Chapter 2, I use a design-based approach to estimate stock size survey indices and show how changes in survey coverage can affect design-based indices. In Chapter 3, I will use a spatiotemporal model to address changes in survey coverage. The spatiotemporal model accounts for depth (i.e. strata as a proxy of depth) to more accurately estimate stock size indices (e.g. mean catch number per tow). I use Gaussian Markov random fields in the model to account for spatial variation in mean catches to predict stock size indices in strata that were not sampled during the survey. In Chapter 4, I first summarize what I did in the previous three chapters and describe the main finding of this research. I then give some

issues and recommendations that can be investigated in future research.

1.3 Figures

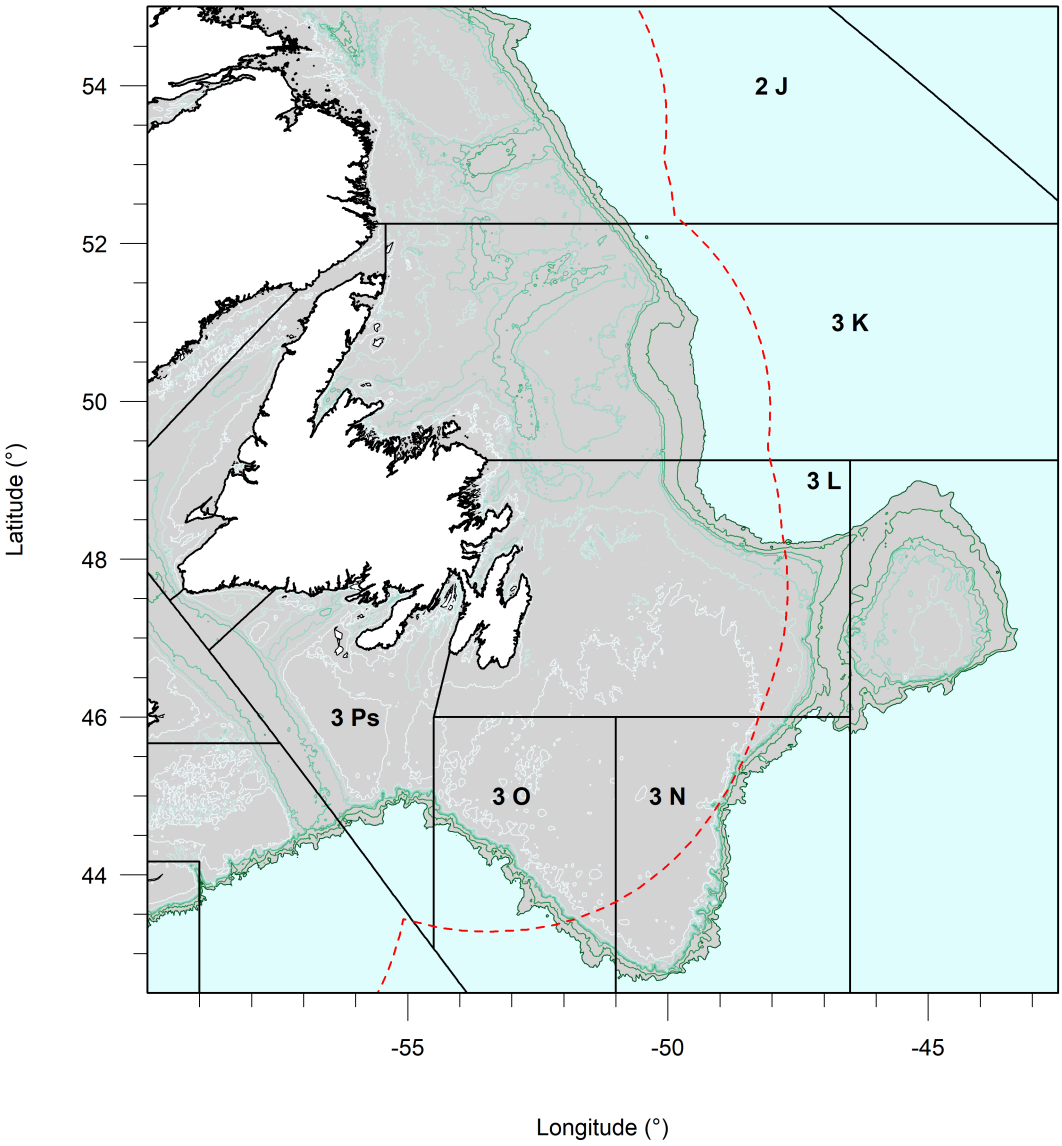

Figure 1.1: Map of NAFO Sub-areas 2 and 3. The red dashed line is Canada's nautical mile limit. White and green colour spectrum lines on the map show 100-200 and 201-1500 metre depth contours, respectively. A darker green colour indicates a greater depth.

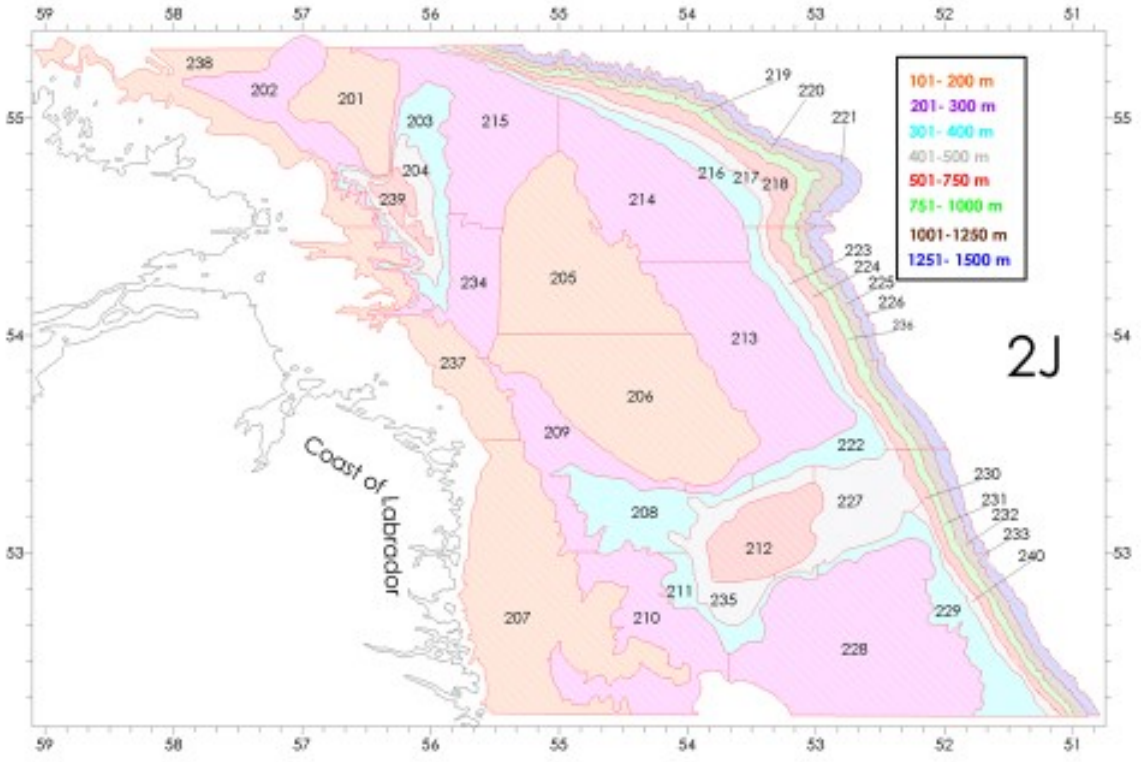

Figure 1.2: Strata map of NAFO Division 2J. Strata numbers and their depth ranges are indicated by the colour box on the right of the figure.

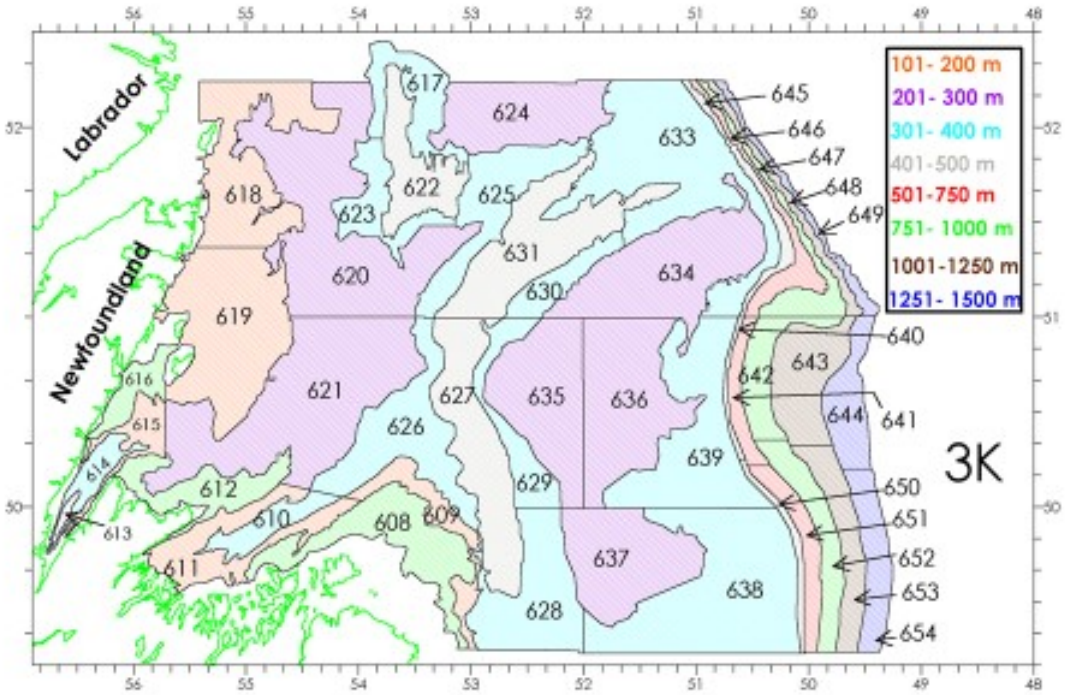

Figure 1.3: Strata map of NAFO Division 3K. See Figure 1.2 for additional description.

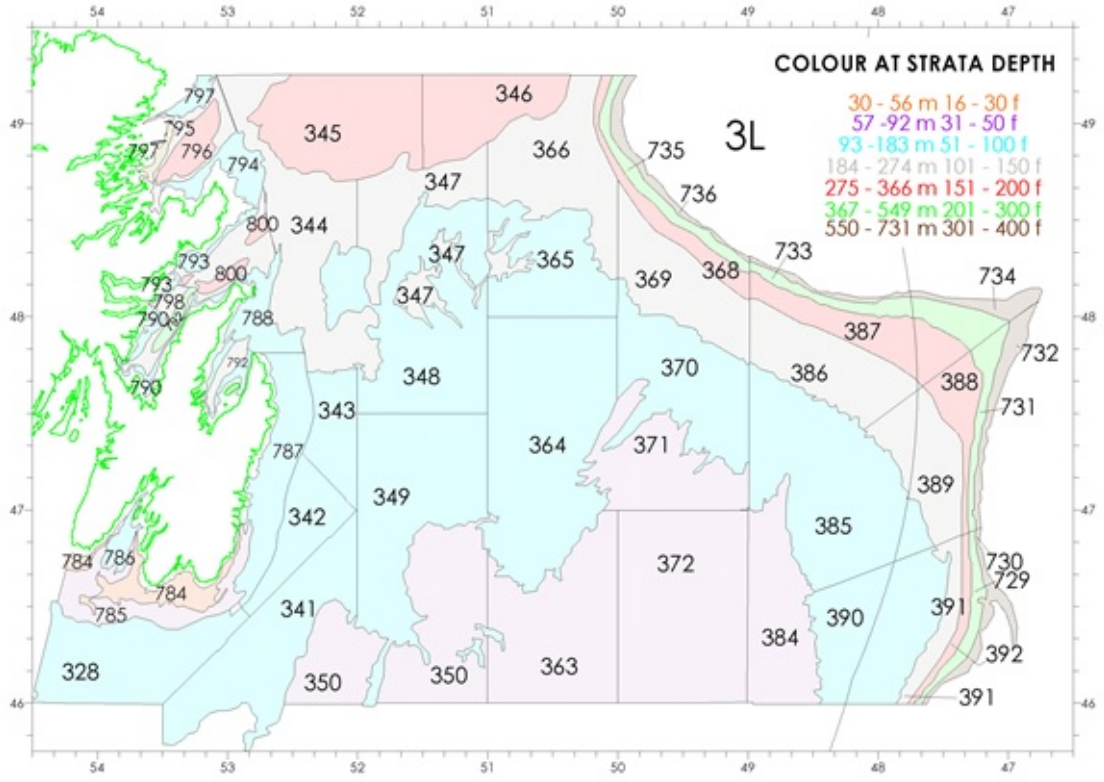

Figure 1.4: Strata map of NAFO Division 3L. Strata numbers and their depth ranges are indicated by the colour box on the right of the figure.

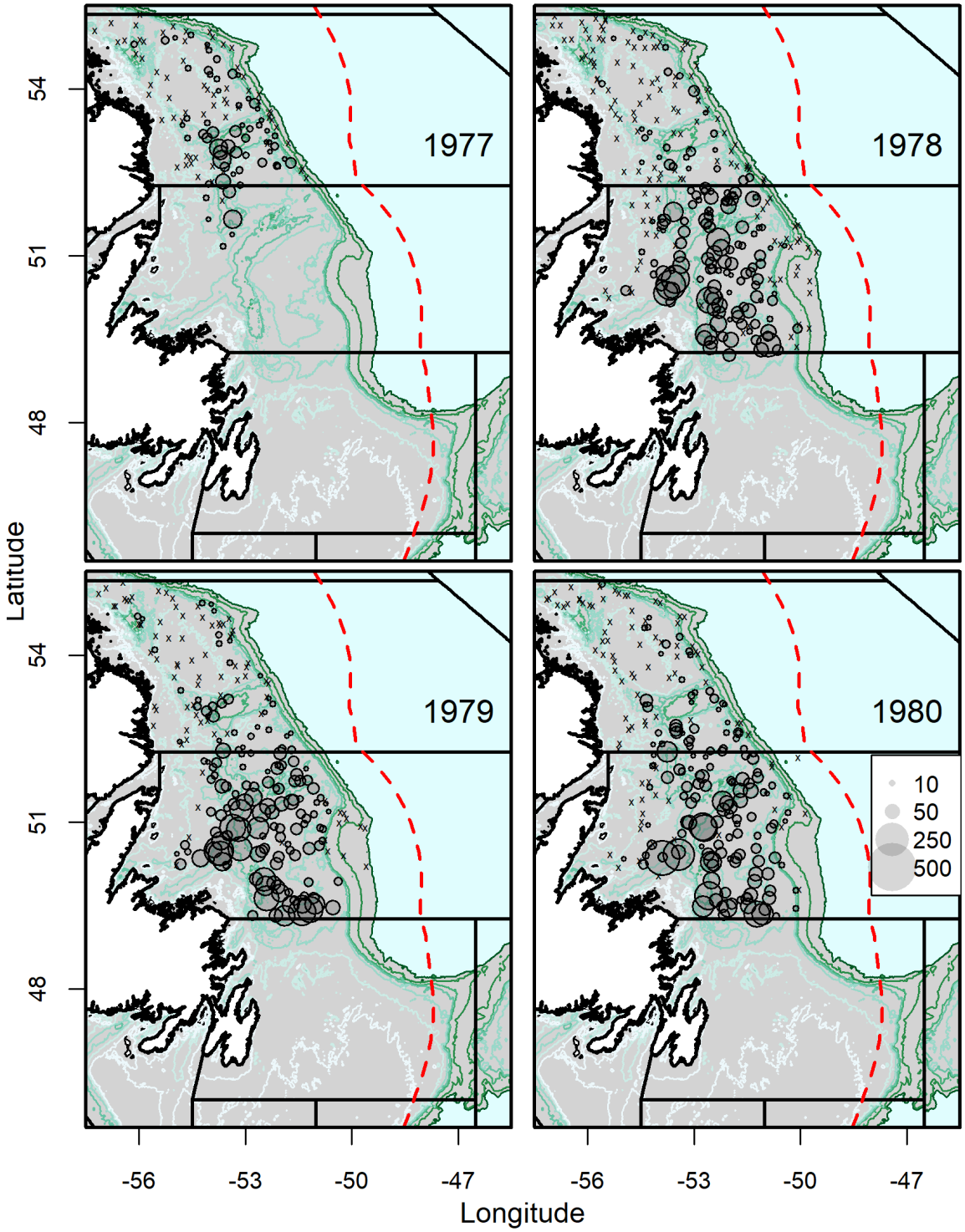

Figure 1.5: Set-by-set distribution of witch flounder catch in abundance (numbers) in DFO autumn RV surveys (1977-1980) in NAFO Divisions 2J3KL, where the × symbol means zero catch per set. A bubble indicates 5 × sqrt (catch/max) in each set in which max is the maximum number of witch flounder caught for all survey years, according to the size of grey bubbles in the box on the panel of 1980. The red dashed line is Canada's nautical mile limit. The longitude and latitude (the X and Y axes) show the geographic coordinates of each conducted survey set. The bathymetry around eastern Newfoundland and Labrador in depths ranging from 100 to 1500 metres are shown with coloured lines: White depth range 100-200 and green colour spectrum 201-1500. A darker green colour indicates a greater depth.

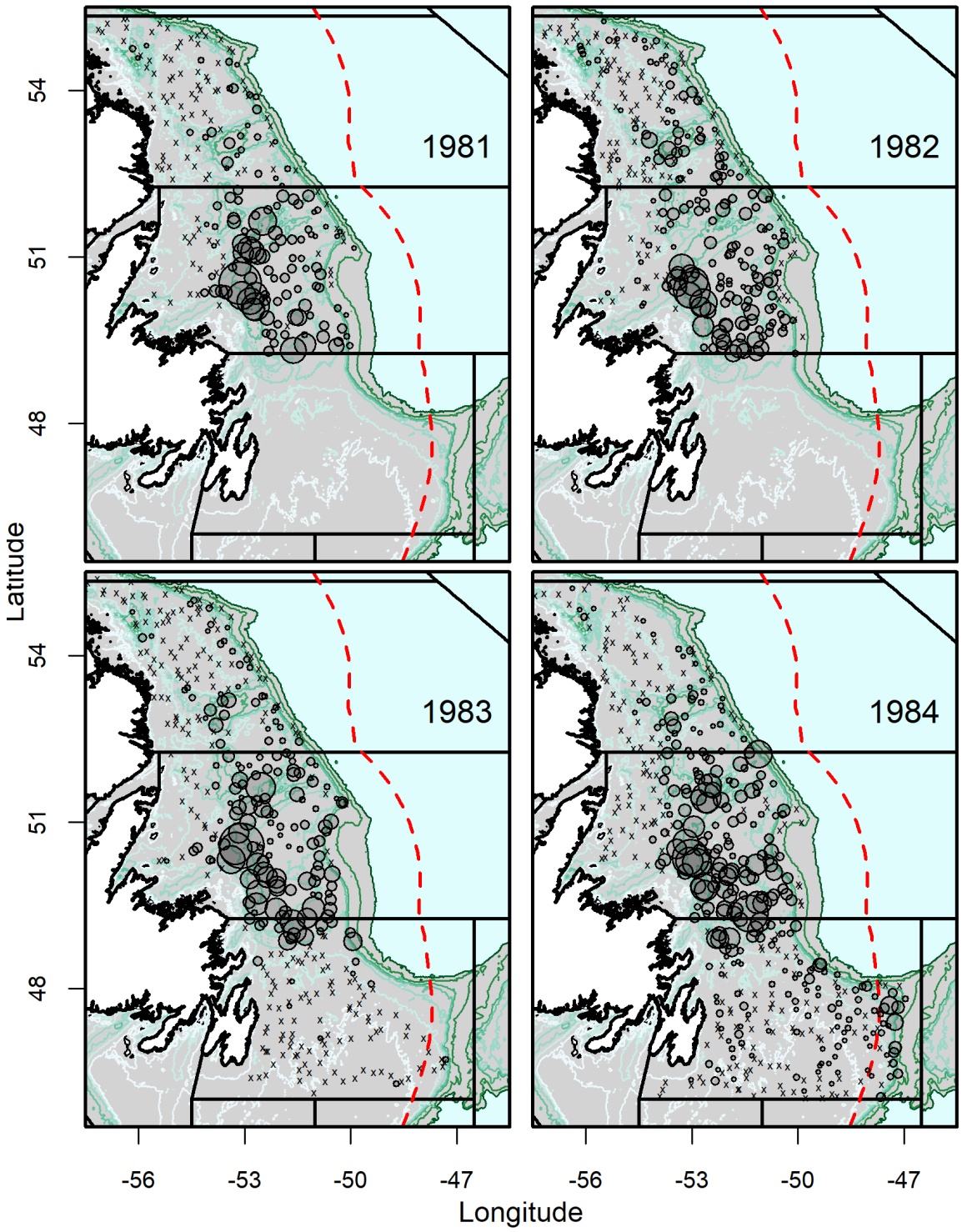

Figure 1.6: Set-by-set distribution of witch flounder catch in abundance (numbers) from DFO autumn RV surveys (1981-1984); see Figure 1.5 for additional description.

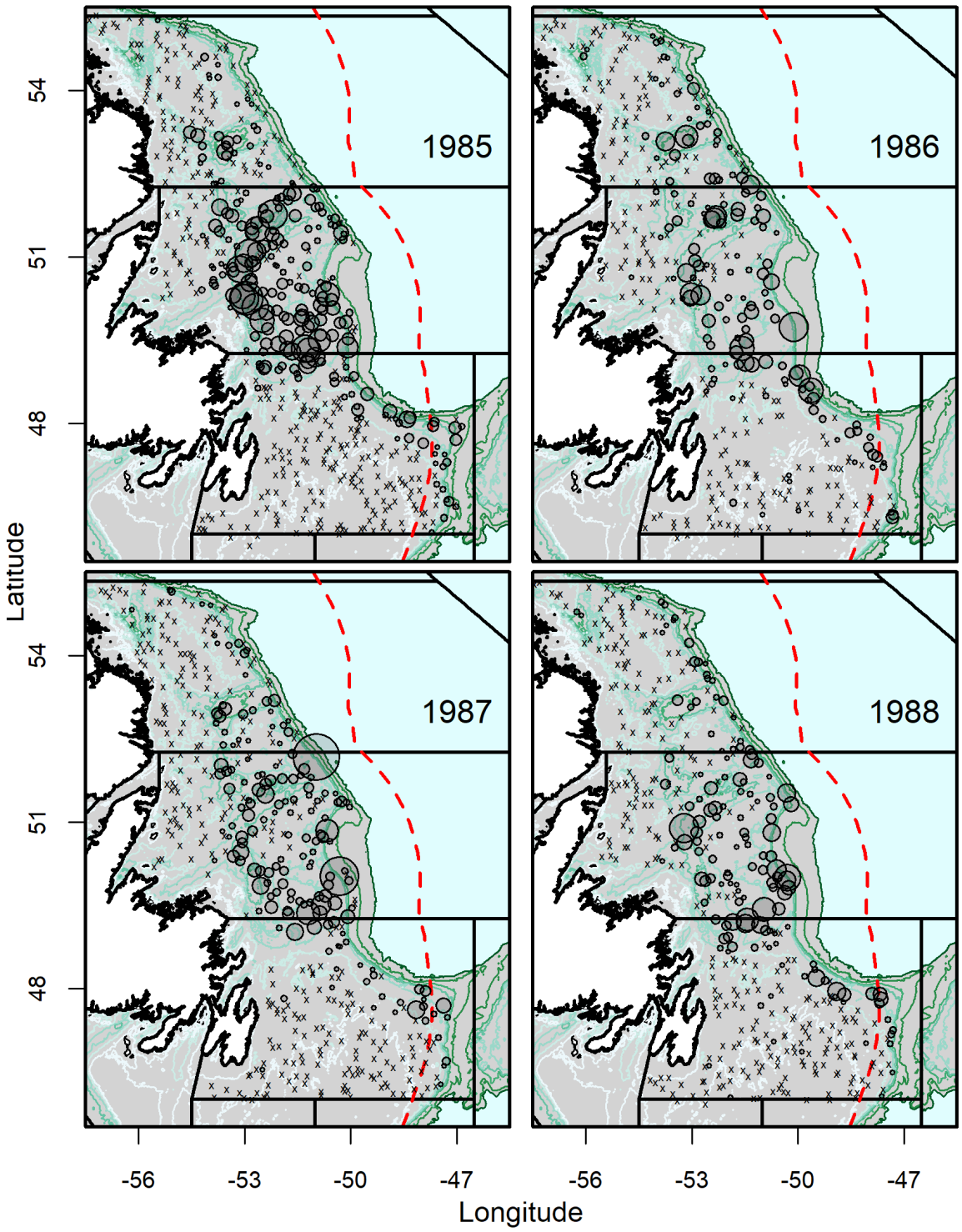

Figure 1.7: Set-by-set distribution of witch flounder catch in abundance (numbers) from DFO autumn RV surveys (1985-1988); see Figure 1.5 for additional description.

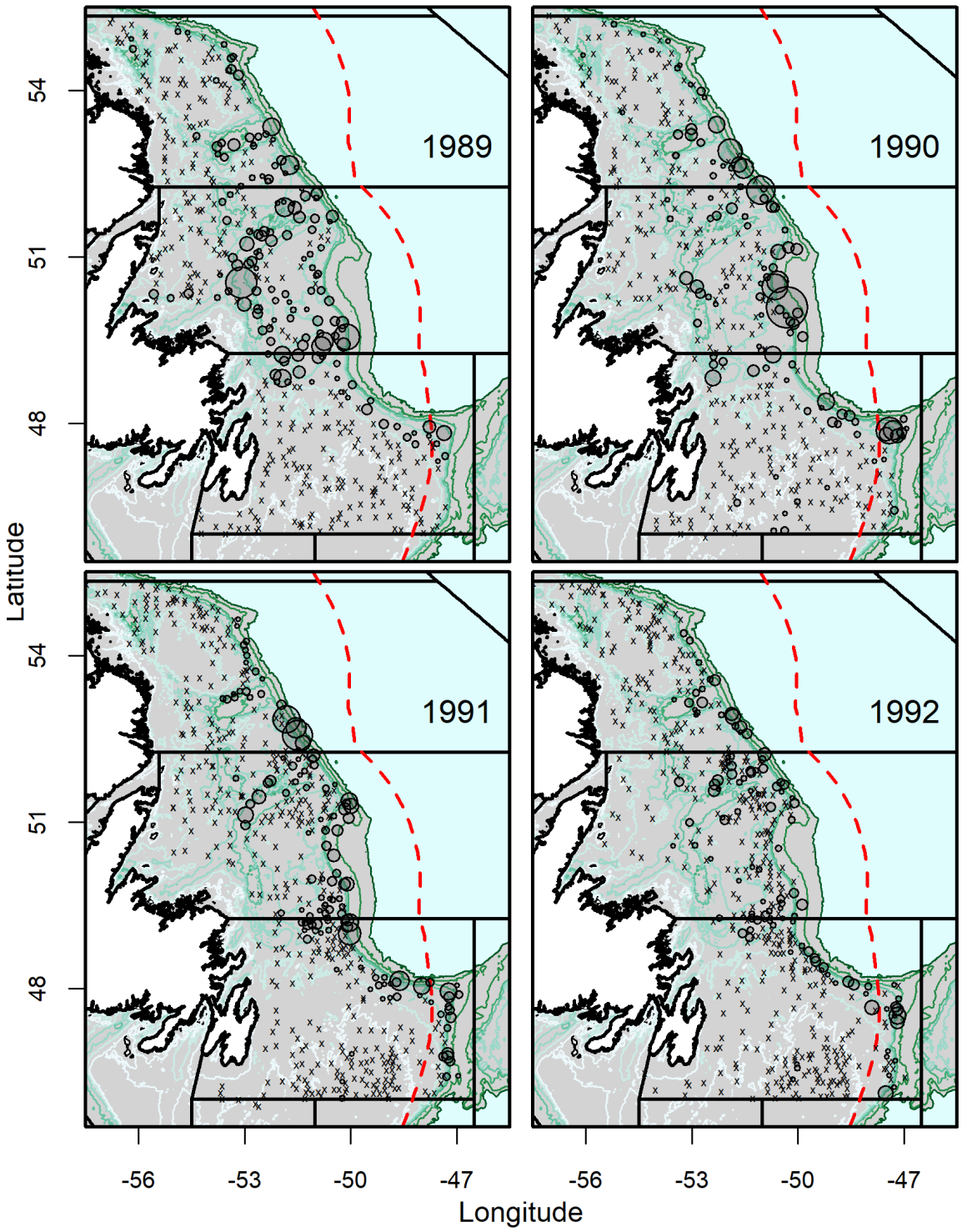

Figure 1.8: Set-by-set distribution of witch flounder catch in abundance (numbers) from DFO autumn RV surveys (1989-1992); see Figure 1.5 for additional description.

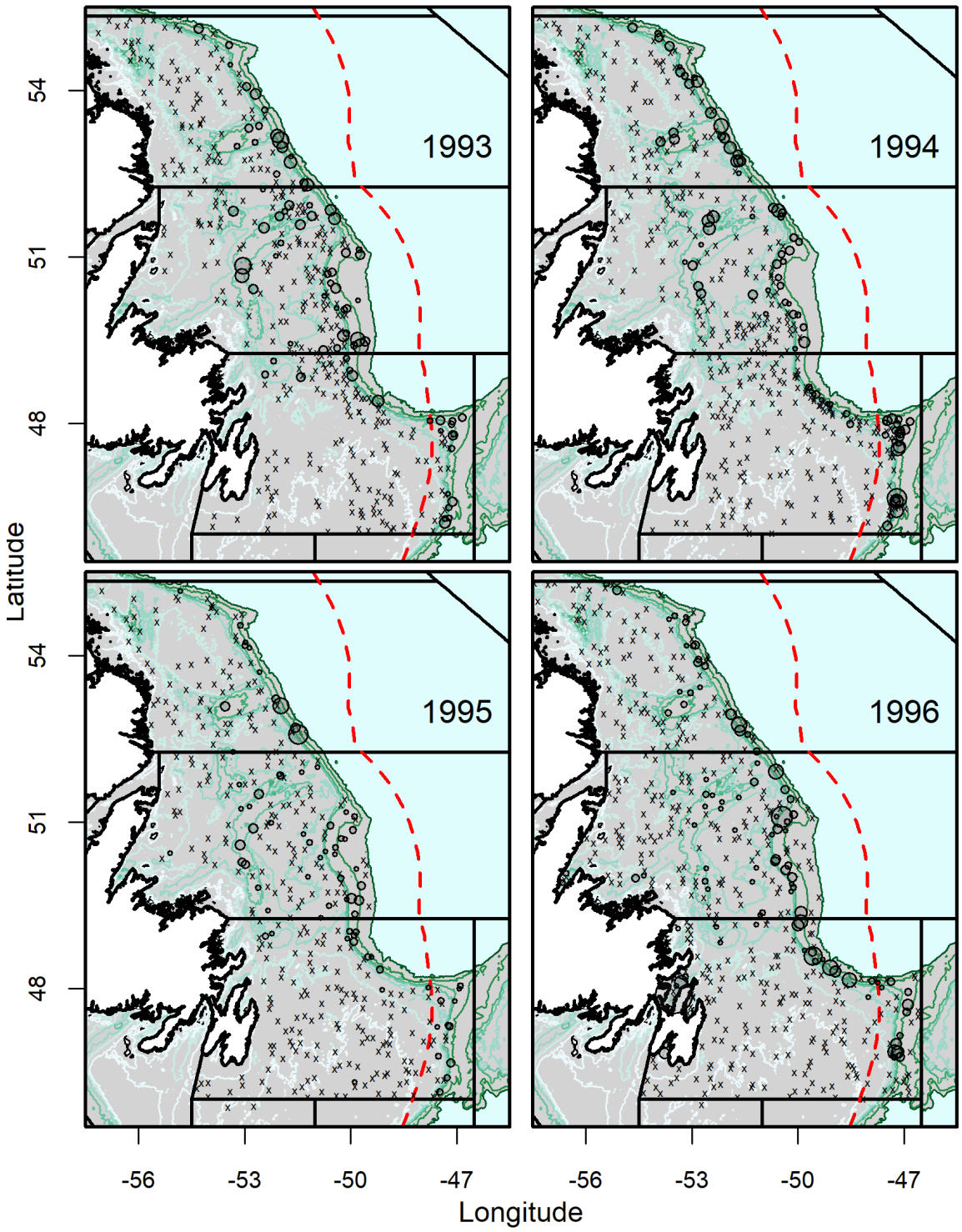

Figure 1.9: Set-by-set distribution of witch flounder catch in abundance (numbers) from DFO autumn RV surveys (1993-1996); see Figure 1.5 for additional description.

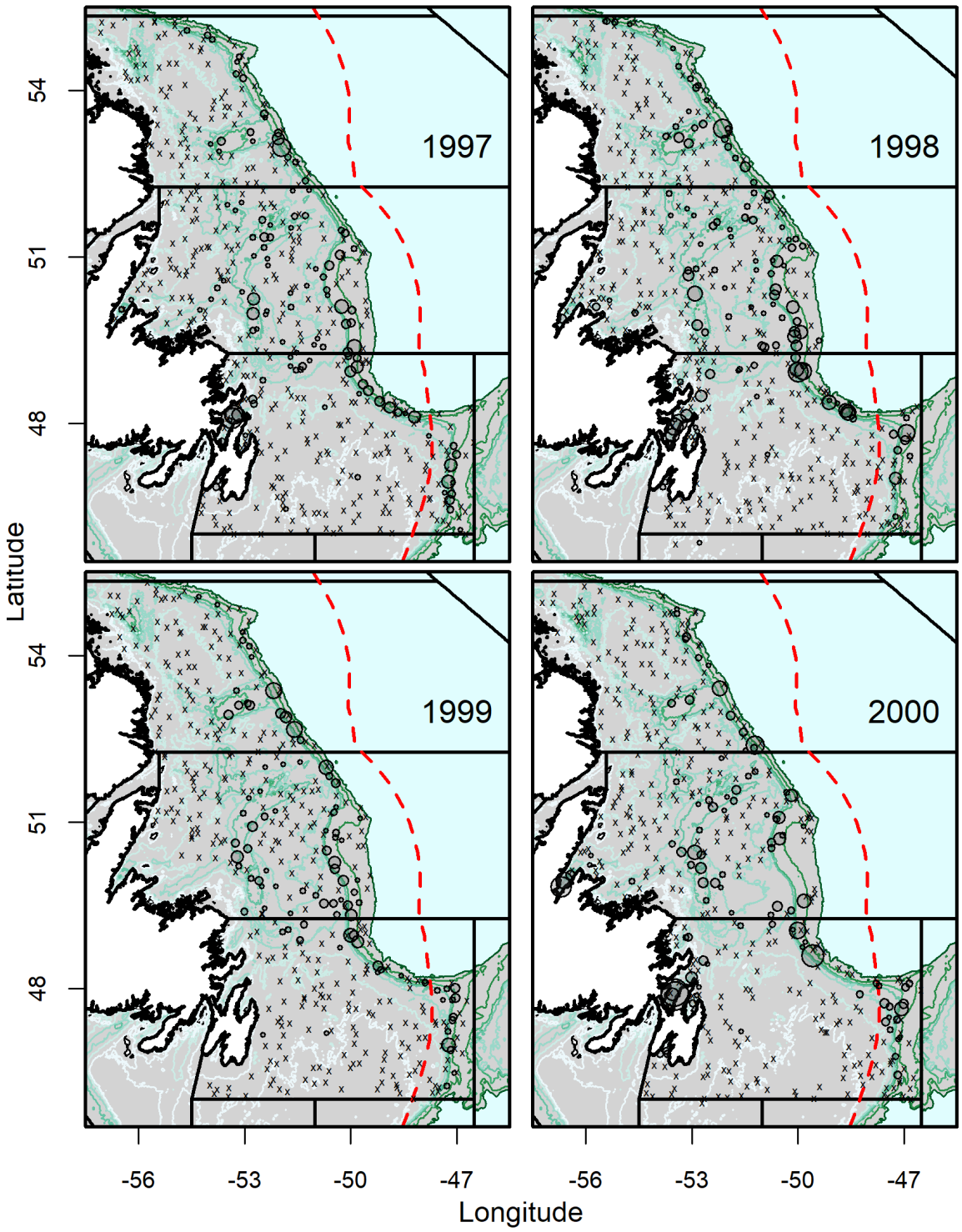

Figure 1.10: Set-by-set distribution of witch flounder catch in abundance (numbers) from DFO autumn RV surveys (1997-2000); see Figure 1.5 for additional description.

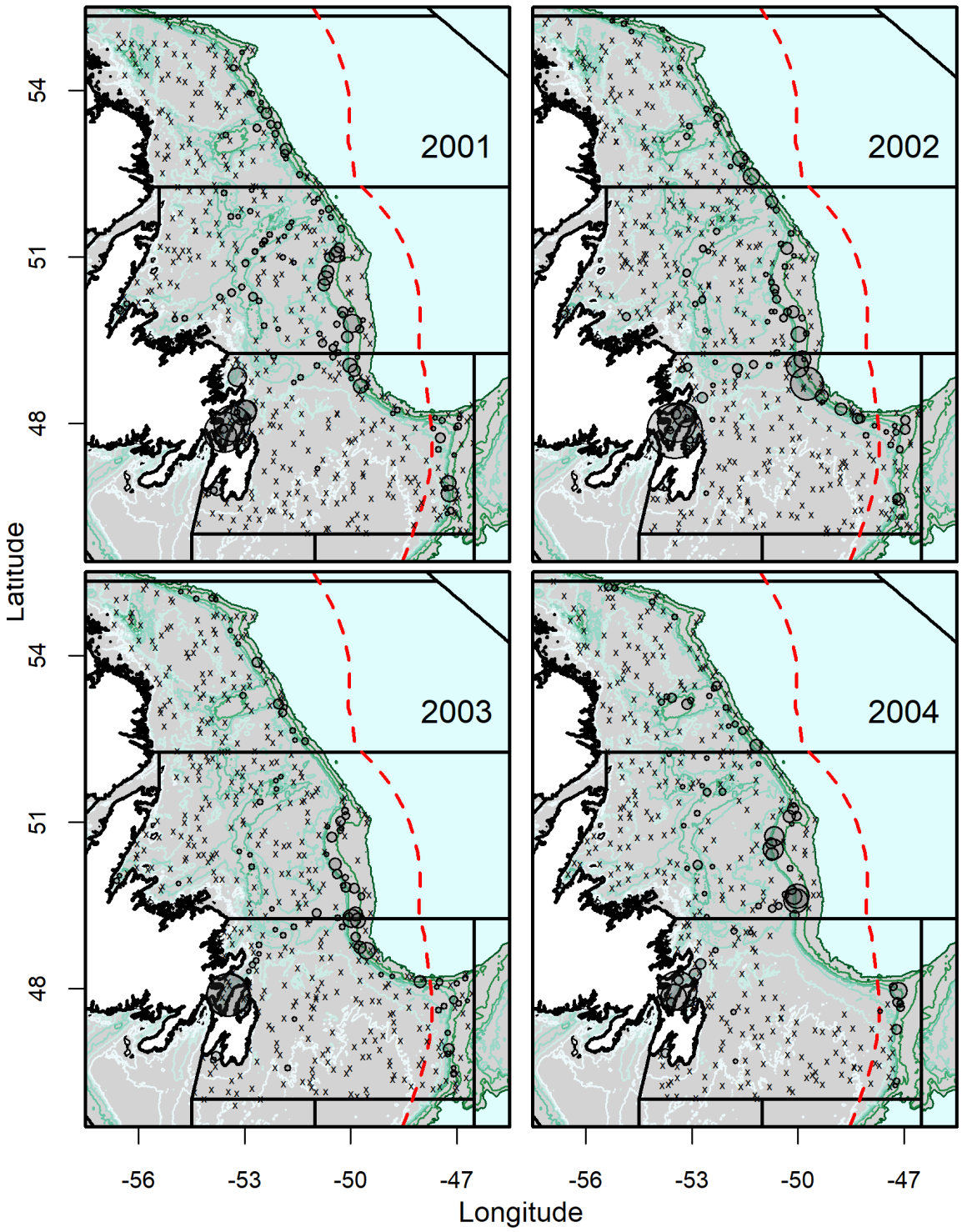

Figure 1.11: Set-by-set distribution of witch flounder catch in abundance (numbers) from DFO autumn RV surveys (2001-2004); see Figure 1.5 for additional description.

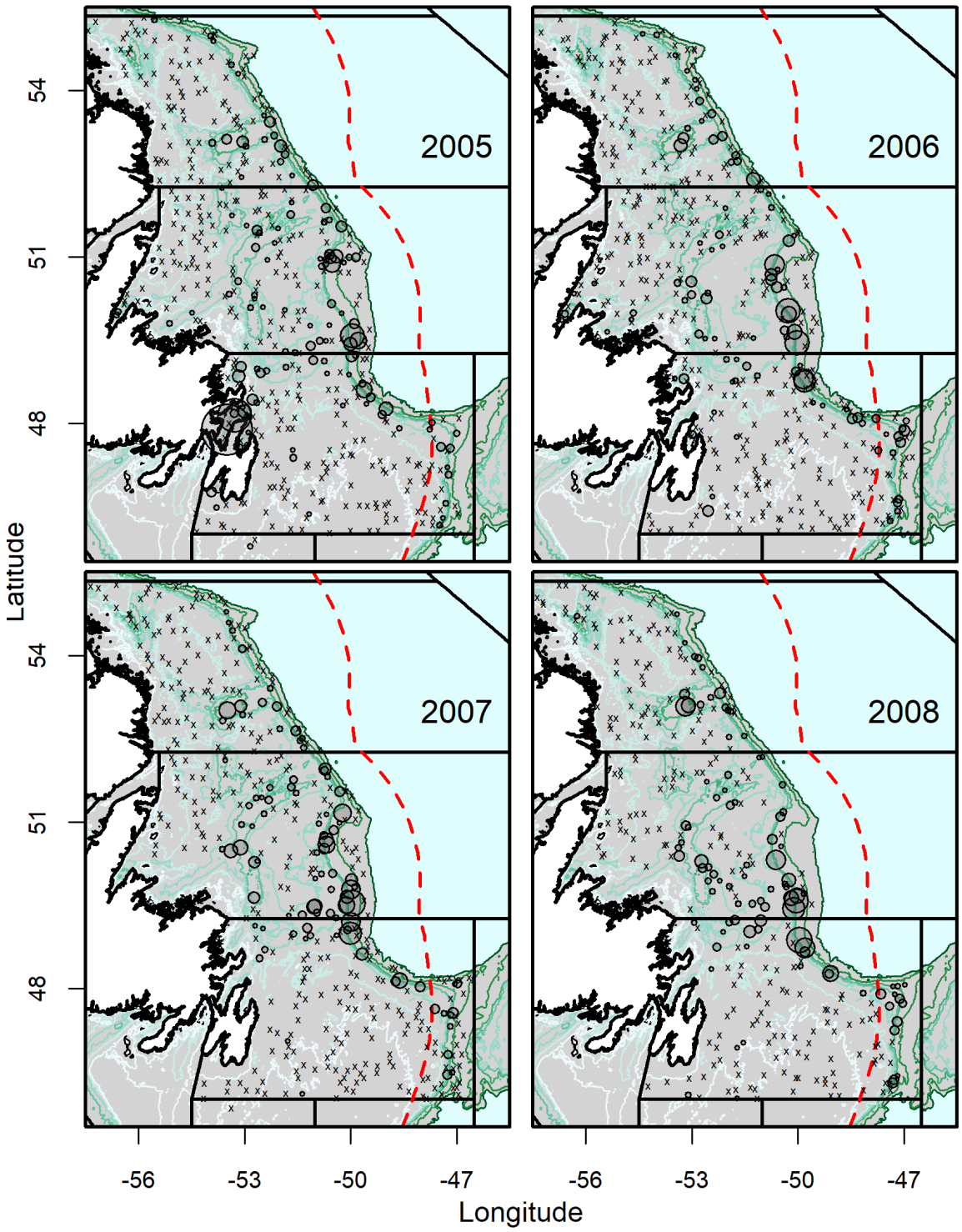

Figure 1.12: Set-by-set distribution of witch flounder catch in abundance (numbers) from DFO autumn RV surveys (2005-2008); see Figure 1.5 for additional description.

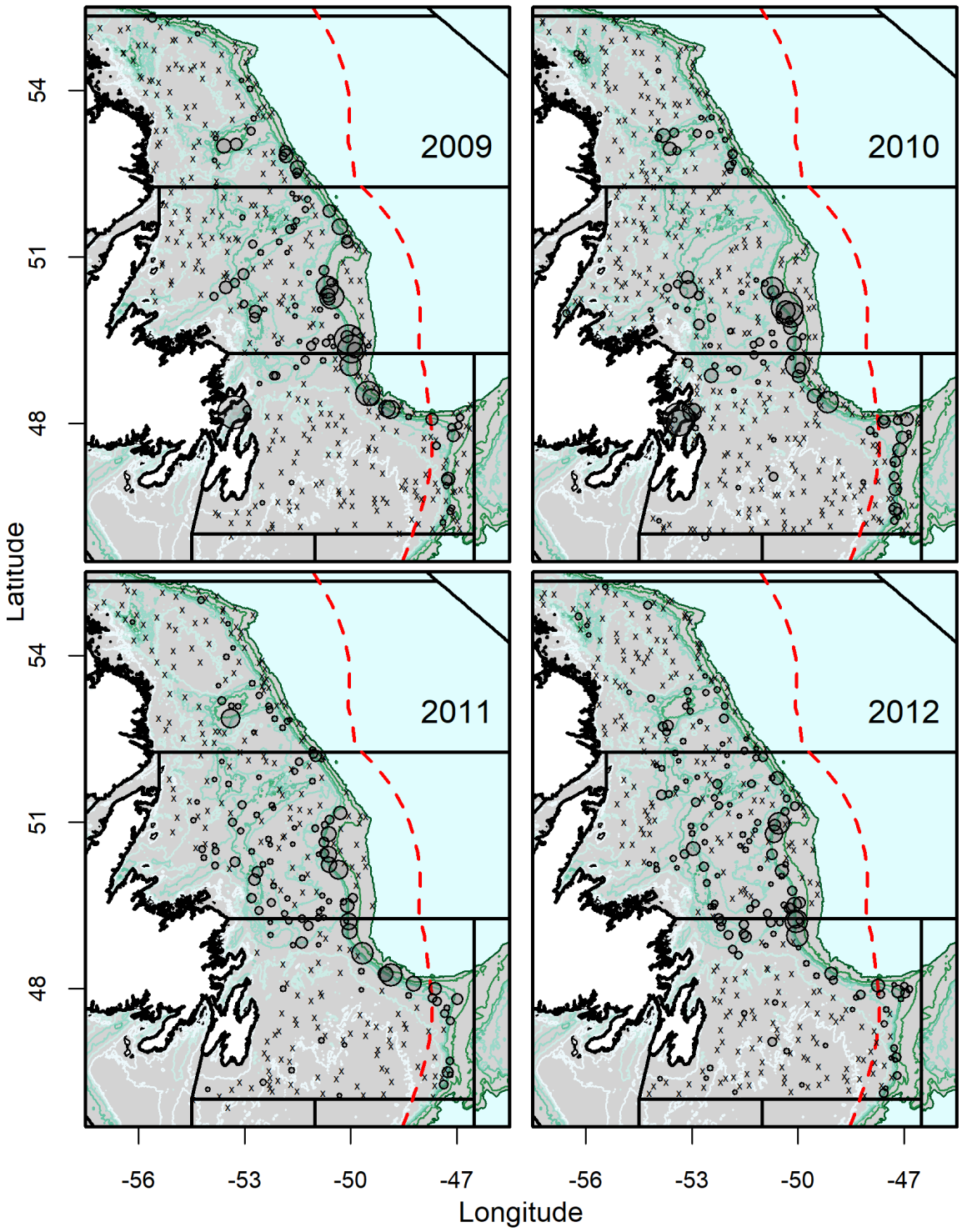

Figure 1.13: Set-by-set distribution of witch flounder catch in abundance (numbers) from DFO autumn RV surveys (2009-2012); see Figure 1.5 for additional description.

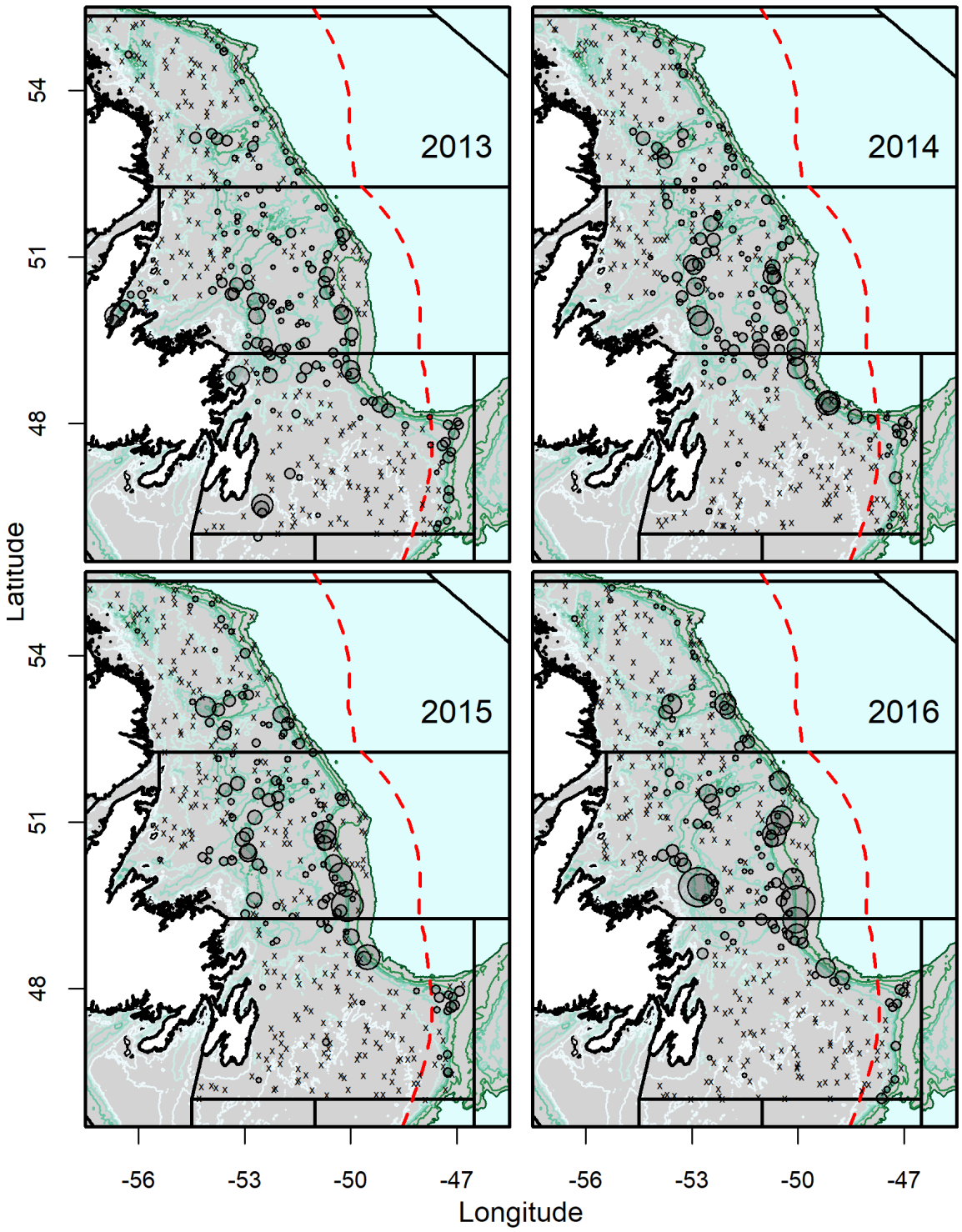

Figure 1.14: Set-by-set distribution of witch flounder catch in abundance (numbers) from DFO autumn RV surveys (2013-2016); see Figure 1.5 for additional description.

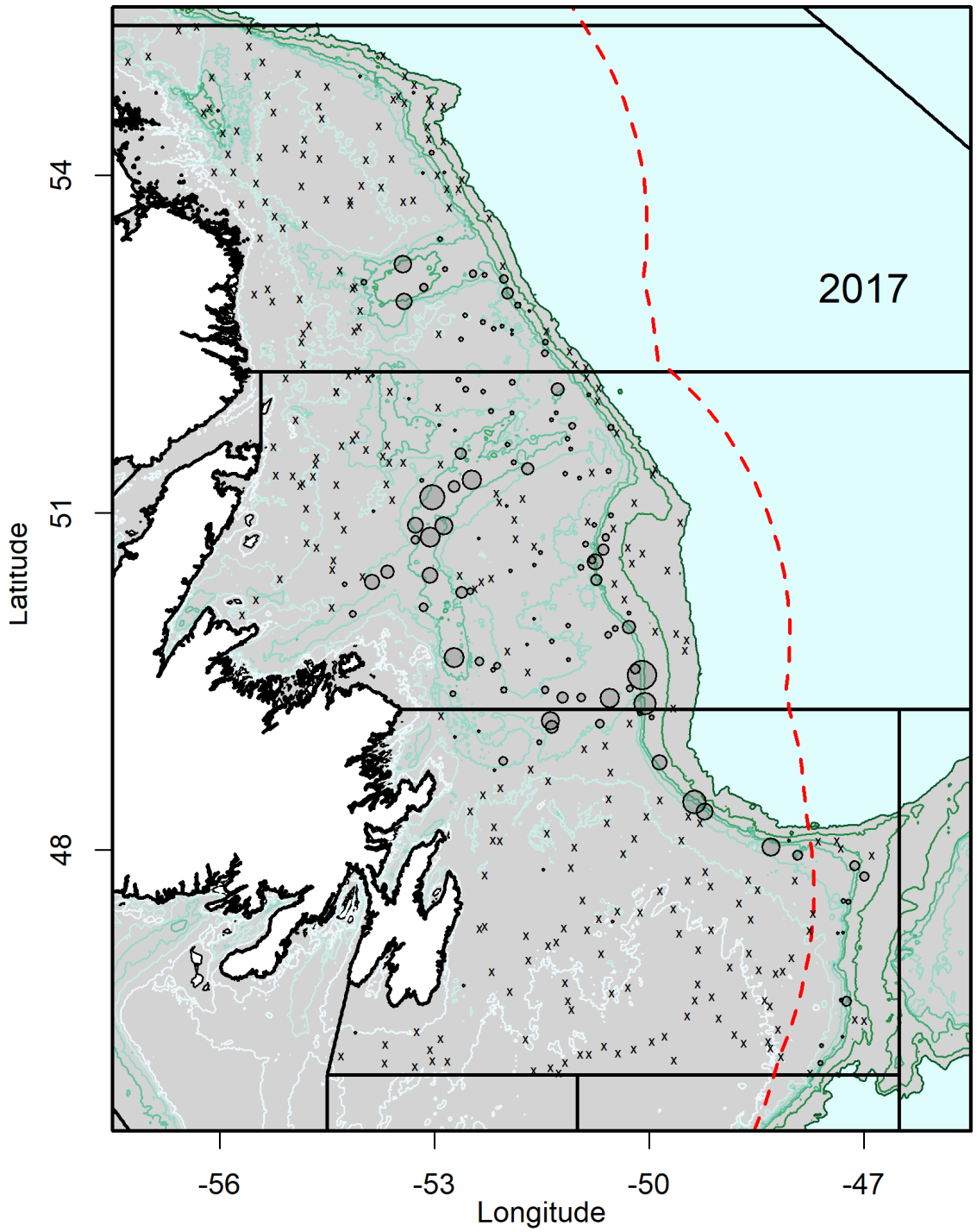

Figure 1.15: Set-by-set distribution of witch flounder catch in abundance (numbers) from DFO autumn RV surveys in 2017; see Figure 1.5 for additional description.

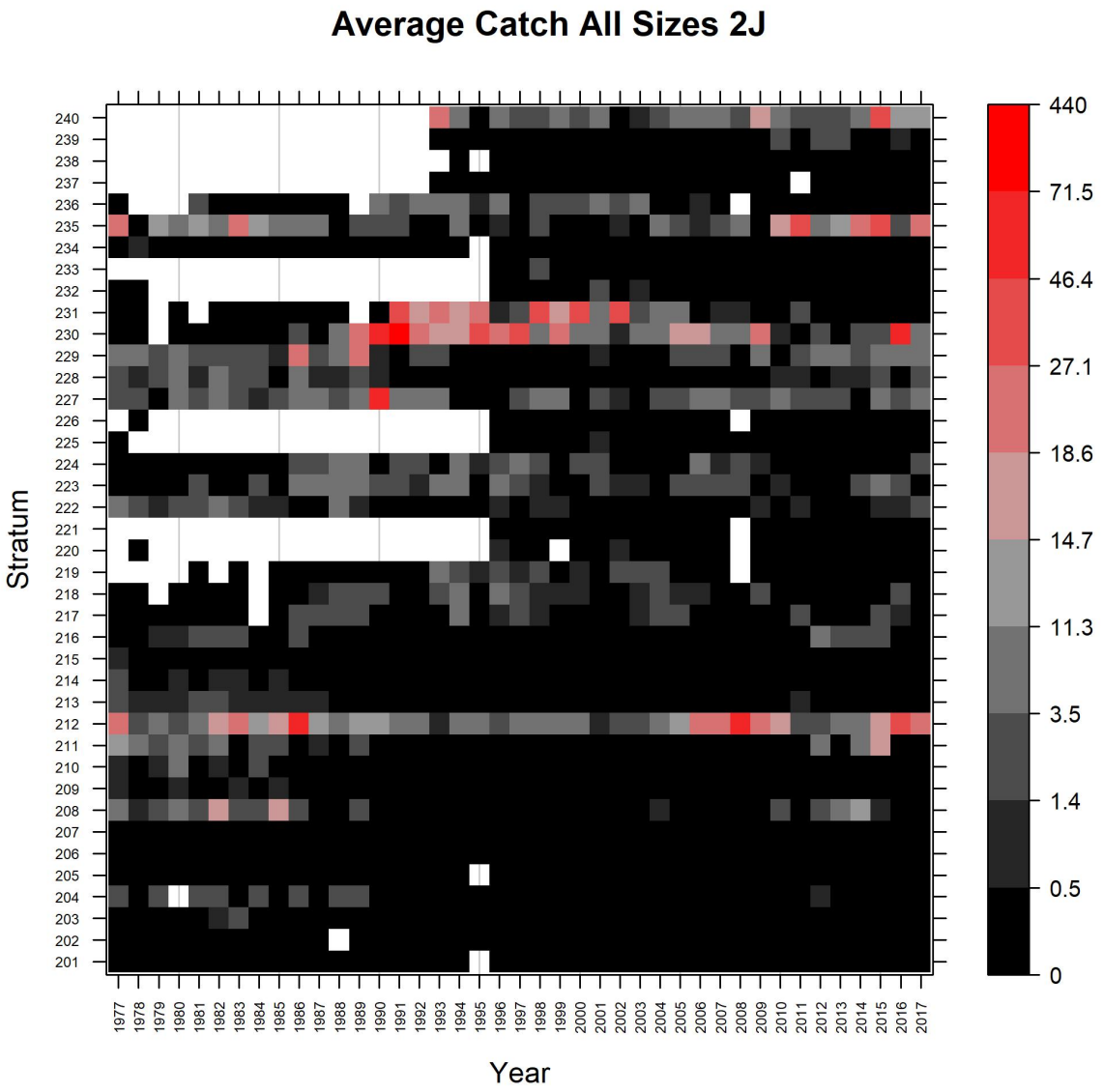

Figure 1.16: The average catch of witch flounder in abundance from the annual autumn DFO RV survey in NAFO Division 2J. Each cell's colour represents the value of average catch in the corresponding stratum and year based on the colour bar. The blank cells/areas indicate strata that were not sampled in survey years in this division.

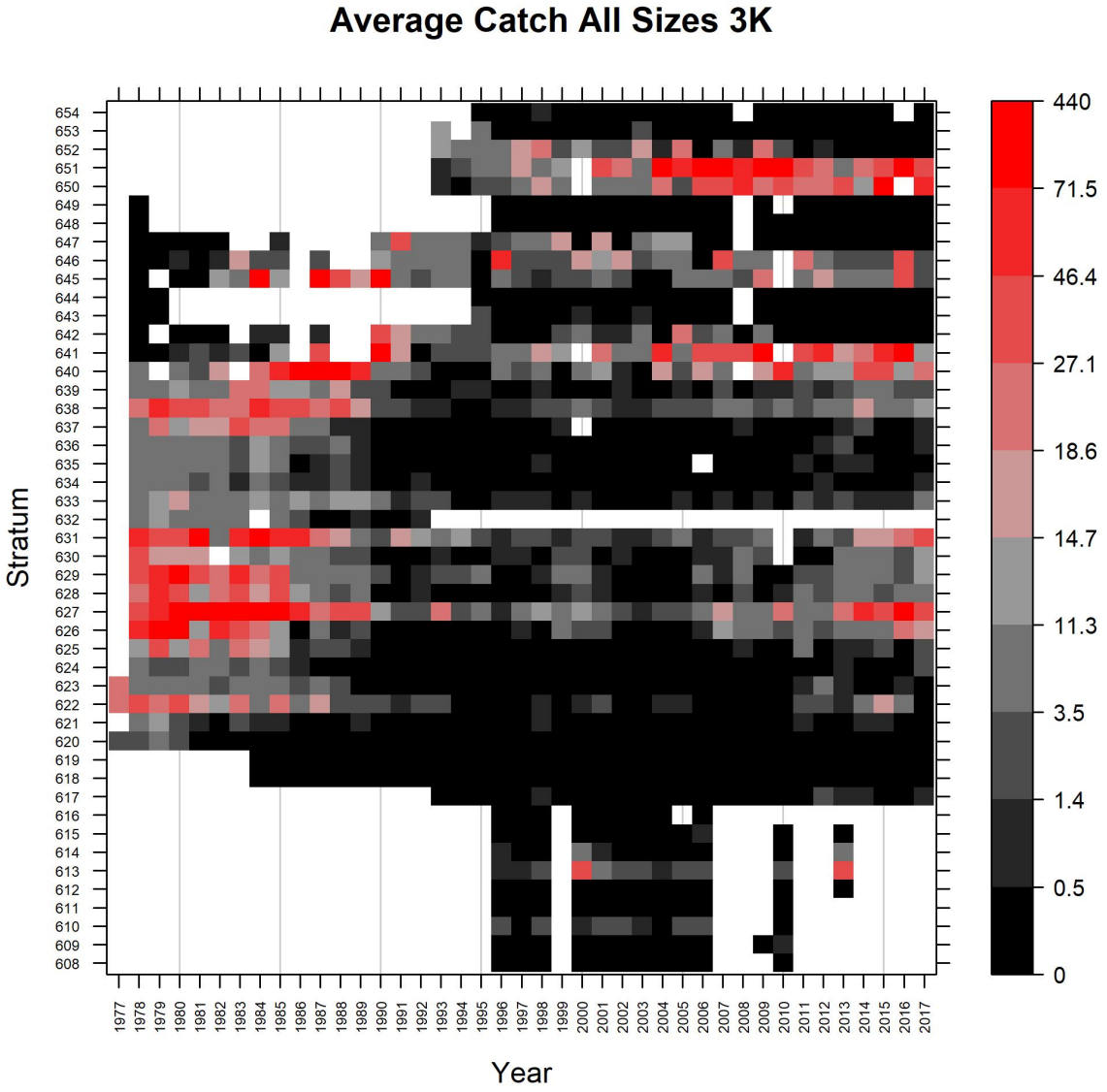

Figure 1.17: The average catch of witch flounder in abundance from the annual autumn DFO RV survey in NAFO Division 3K. See Figure 1.16 for additional description.

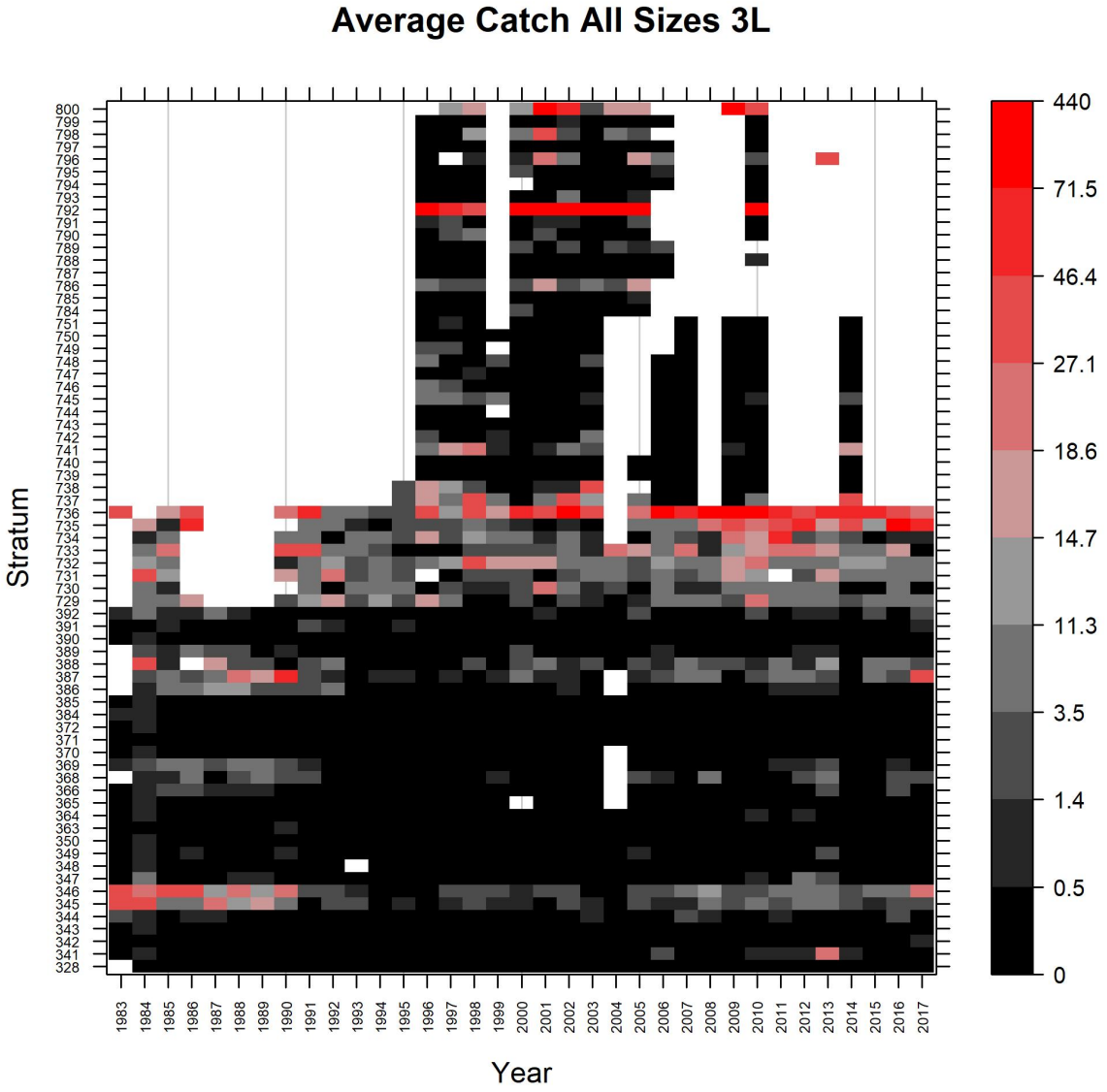

Figure 1.18: The average catch of witch flounder in abundance from the annual autumn DFO RV survey in NAFO Division 3L. See Figure 1.16 for additional description.

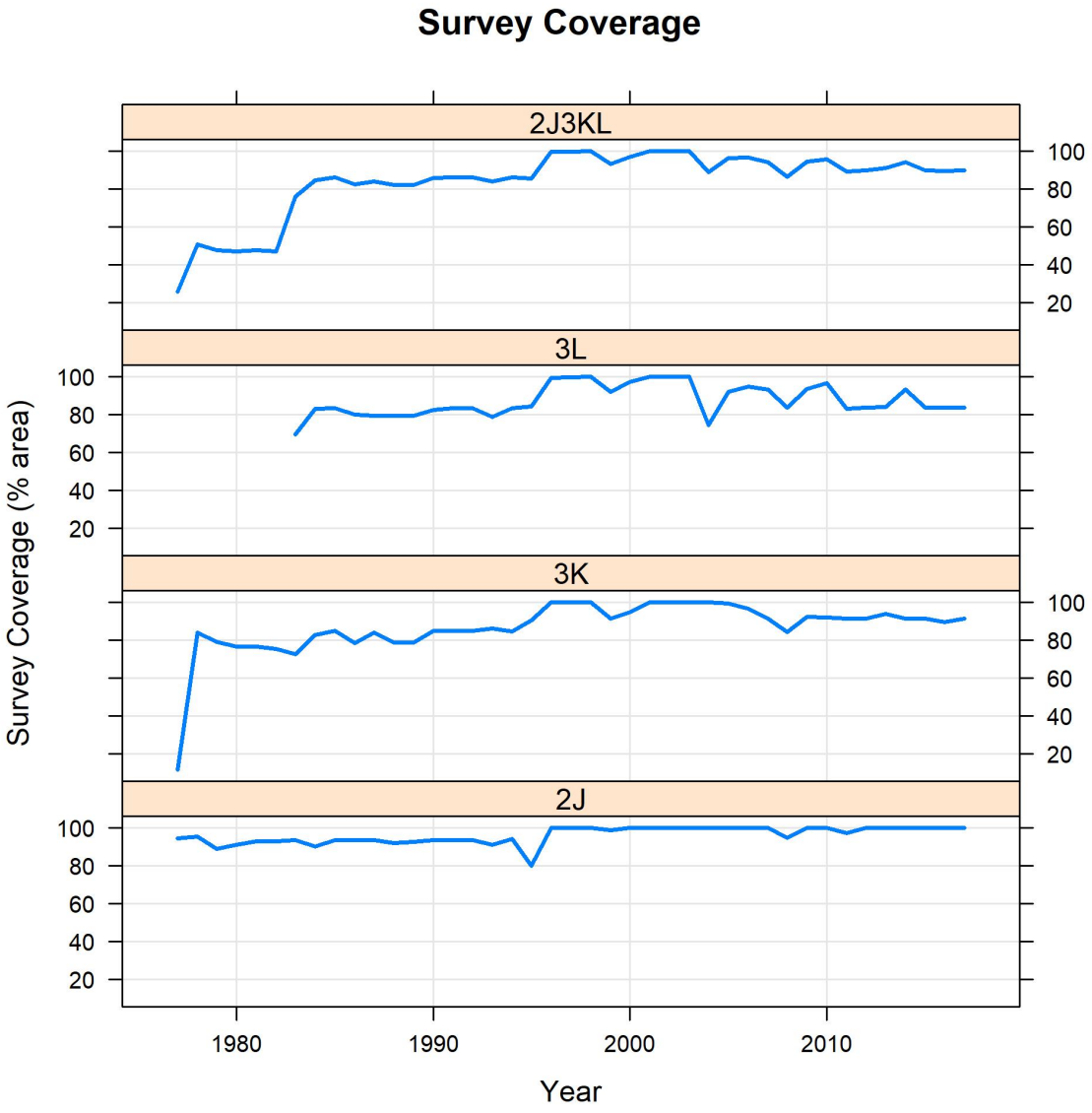

Figure 1.19: The percentage of survey area coverage for the entire area (2J3KL) and each NAFO division from annual autumn DFO RV survey.

Chapter 2

Design-based Survey Indices

2.1 Introduction

As described in Chapter 1, the goal of this thesis is to provide improved stock size survey indices. Stock size survey indices can be obtained by either model- or design-based approaches. Fisheries and Oceans Canada (DFO) and other fisheries management organizations (e.g. the US National Marine Fisheries Service, NMFS) around the world widely utilize design-based survey indices (Cadigan, 2011). The design-based approach has been used by DFO since 1981 to indicate the status of witch flounder on the east coast of Newfoundland and Labrador (Bowering, 1981) including the most recent stock assessment of witch flounder in the Northwest Atlantic Fisheries Organization (NAFO) Divisions 2J3KL (DFO, 2019). I use this approach as a baseline to compare the performance of a spatiotemporal model developed later in this thesis.

In Section 2.2, I first describe the design-based approach and reproduce the DFO survey indices for witch flounder abundance and biomass. I show that the design-based indices are not comparable between different years because of changes in survey coverage, and the indices may not reliably reflect trends in population dynamics. In addition, the strata have changed in terms of area and boundaries over time. Although these changes

in strata boundaries are not a problem in the design-based approach, they are a problem in the model-based approach that I pursue in later chapters. Therefore, in Section 2.3, I propose a re-stratification of the survey data based on currently defined strata boundaries. This re-stratification produces survey sets that are consistently stratified across the whole time-series. Finally, I compare the design-based survey indices for witch flounder based on the original and new strata I develop.

2.2 Design-based approach

The design-based approach utilizes the theory of statistical inference for finite populations. The survey sampling design-based approach is a conventional method used in analyzing fisheries survey data. In this approach, the survey catches (the response variable) at all possible sampling sites in the population are treated as unknown quantities, and the values observed at sampled sites in a survey are assumed to be measured without error. The design-based approach involves probability via the survey sampling design (i.e. the sample selection procedure) for statistical inference (e.g. Thompson, 2012). The design-based estimates are derived from n observations (i.e. the sample) of the population and the survey design, which specifies how every set of n observations are selected.

A fundamental advantage of the design-based approach is that it can provide unbiased estimators with no assumptions about the distribution of population values of the response variable (Thompson, 2012). This advantage is often important when there is little information about the underlying population, which may be the case for natural populations. An estimator is unbiased if the expected (i.e. average) value of the estimator taken over all possible samples that could be produced by the sampling design equals the true population parameter. A serious disadvantage is that the design-based approach requires a perfectly standard and consistent survey. Thus, changes in survey coverage is a problem in the design-based approach specifically when the goal is to derive a consistent

time series of survey indices. As illustrated in Chapter 1, not all strata are sampled each year in the 2J3KL survey. Design-based indices can only represent the strata surveyed each year. Therefore, design-based indices derived from surveys with changes in coverage may not be comparable from year to year. For example, the witch flounder abundance index for 2004 only represents the strata sampled from 89% (i.e. the percentage of survey coverage in 2004, Figure 2.1) of the entire stratified stock area while this index for 2005 represents the strata sampled from 96% of the total area. Therefore, the comparison of these two abundance indices may not make sense.

Another disadvantage of this approach is that the design-based estimates (e.g. abundance and biomass indices) derived from trawl surveys sometimes have large variance estimates (Smith, 1990), partly because the trawl produces an uncertain measurement of stock size at every tow site (Cadigan, 2011). This latter observation error is also assumed to not exist in typical design-based statistical inference. In this chapter, we only focus on the design-based means and the problems introduced by inconsistent survey coverage.

2.2.1 Estimating design-based indices

Table 2.1 contains some symbols and notations commonly used for describing design-based estimators for a stratified random sampling (SRS) design to estimate survey indices of stock size (Thompson, 2012). In SRS an unbiased estimator of the stratum total is

$$\widehat{\tau}_h = N_h \bar{y}_h. \tag{2.1}$$

An unbiased estimator for the population total is

$$\widehat{\tau}_{st} = \sum_{h=1}^{H} N_h \bar{y}_h. \tag{2.2}$$

The finite-population variance of this estimator is given by equation (2.3) in which ${\sigma_h}^2$ in equation (2.4) is the variance from stratum h.

$$var(\widehat{\tau}_{st}) = \sum_{h=1}^{H} N_h(N_h - n_h)\frac{{\sigma_h}^2}{n_h}, \tag{2.3}$$

$${\sigma_h}^2 = (N_h - 1)^{-1}\sum_{h=1}^{N_h}(y_{hi} - \mu_h)^2. \tag{2.4}$$

An unbiased estimator of $var(\widehat{\tau}_{st})$ is

$$\widehat{v}ar(\widehat{\tau}_{st}) = \sum_{h=1}^{H} N_h(N_h - n_h)\frac{{s_h}^2}{n_h}, \tag{2.5}$$

where ${s_h}^2$ is the sample variance from stratum h,

$${s_h}^2 = (n_h - 1)^{-1}\sum_{h=1}^{n_h}(y_{hi} - \bar{y}_h)^2. \tag{2.6}$$

Similar to the above equations, an unbiased estimator for the population mean per trawl unit (i.e. the stratified sample mean) is given by equation (2.7). In a stratified trawl survey, this estimator is also called mean number (or weight) per tow (MNPT). The finite population variance of this estimator is given by equation (2.8), and an unbiased estimator of $var(\bar{y}_{st})$ is given in equation (2.9).

$$\bar{y}_{st} = N^{-1}\sum_{h=1}^{H} N_h\bar{y}_h, \tag{2.7}$$

$$var(\bar{y}_{st}) = \sum_{h=1}^{H} \left(\frac{N_h}{N}\right)^2 \left(\frac{N_h - n_h}{N_h}\right) \frac{{\sigma_h}^2}{n_h}, \tag{2.8}$$

$$\widehat{var}(\bar{y}_{st}) = \sum_{h=1}^{H} \left(\frac{N_h}{N}\right)^2 \left(\frac{N_h - n_h}{N_h}\right) \frac{{s_h}^2}{n_h}. \tag{2.9}$$

Equation (2.2) is used to estimate abundance and biomass (i.e. total abundance and total biomass) indices each year. My results are similar to the indices obtained in the DFO report (e.g. Figure 2.2; DFO, 2019), while DFO produced abundance and biomass indices from 1983 onwards. In this study, I focus on MNPT estimated by equation (2.7).

MNPT for each NAFO division and the entire stock area are shown in Figure 2.3. This index has declined significantly from around 13 in 1979 to 4 in 2017 in the entire stock area. Most witch flounder are found in Division 3K. For example, in 2016, witch flounder mean weight per tow (MWPT) was about 2.2 Kg in Division 3K, while it was 0.7 in 3L and about 0.8 Kg in 2J (Figure 2.4).

2.3 Re-stratification

Another complication in analyzing the trawl survey data for 2J3KL is that many changes have occurred to some strata, especially since 1996. For example, stratum number 632 in Division 3K was removed from the surveys conducted since 1993. Most of these changes in strata (i.e. strata area and strata boundaries) were mainly due to inaccuracies in depth contours discovered with improved bottom mapping. Bishop (1994) described the revisions to stratification schemes in NAFO Subarea 2 and 3. Other changes included the addition of inshore and deep-water strata in the 1996 survey.

These changes in the strata area and boundaries are not a problem in the design-based approach; however, they are a problem in the model-based approach I develop in later

chapters, which involves a stratum-effects model in which the strata are covariates. It is important in these models that strata be defined consistently over time. In stratified random trawl surveys, the strata are convenient spatial units where the depth range (i.e. a stratification variable) is less variable within strata than between strata. Therefore, a re-stratification of the survey data is proposed to address this problem.

I dealt with this problem by re-stratifying the survey catch data based on current strata boundaries. I re-stratified using the current strata shape-files provided by DFO. These shape-files are the spatial polygon data defining strata boundaries in NAFO Divisions 2J3KL. These shape-files were imported to R (R Core Team, 2013), and then I used the shape-files to assign each sample observation in the survey time-series to the current stratum definitions (i.e. the new strata). However, due to inaccuracies in the shape-files, some observations were assigned to multiple strata, and in this case I defined the original strata as the new strata. A small number of observations (i.e. 22) were located in strata with missing identification numbers (i.e. not available, NA). I plotted the locations of these observations to determine which stratum they belonged. Also, six observations occurred at positions that were not covered by the shape-file polygons for NAFO Divisions 2J3KL. I also plotted the locations of these six observations to determine which stratum they should be assigned to. For example, the blue point in Figure 2.5 was not contained within 2J3KL strata polygons (i.e. it was on the boundary of Division 2J) and I assigned it to stratum 216. Note that there is no re-stratification after 1996 since there has been no change in the stratification since then.

The impact of the re-stratification is illustrated in Figures 2.6-2.8 for each NAFO division. In Division 2J (Figure 2.6), significant changes occurred for strata 205, 206, 209 and 210. For example, after the re-stratification, only about 55% of observations from the original stratum 210 were assigned to the new stratum 210, while about 40% and less than 5% of its observations were assigned to new strata 228 and 207, respectively. Changes due to re-stratification were less in Division 3L (Figure 2.8) than in 2J or 3K.

I compared the design-based survey indices based on the original and new strata. Figure 2.9 shows the differences between the mean catch number per tow (MNPT) as well as mean catch weight per tow (MWPT) for witch flounder before and after re-stratification are small.

In this section, re-stratification provided consistent strata during the entire time series. In the next chapter, I will apply a generalized linear mixed-effect model, with strata and year effects to adjust for changes in survey coverage to produce survey indices that are more comparable over time.

2.4 Tables

Table 2.1: List of symbols for estimation based on a stratified random sampling design

Name	**Symbol**
Number of strata	H
Number of sampling units in stratum h=$\frac{\text{Stratum area}}{\text{Swept area}}$	N_h
Number of sampling units in survey	$N{=}\sum_{h=1}^{H} N_h$
Number of units sampled in the stratum h	n_h
ith observation in the stratum h	y_{hi}
Sample mean in the stratum h	$\bar{y}_h = n_h^{-1} \sum_{i=1}^{n_h} y_{hi}$
Population mean of stratum h	$\mu_h = N_h^{-1} \sum_{i=1}^{N_h} y_{hi}$
Population total of stratum h	$\tau_h = \sum_{i=1}^{N_h} y_{hi} = N_h \mu_h$
Population total of survey	$\tau = \sum_{h=1}^{H} \tau_h$

2.5 Figures

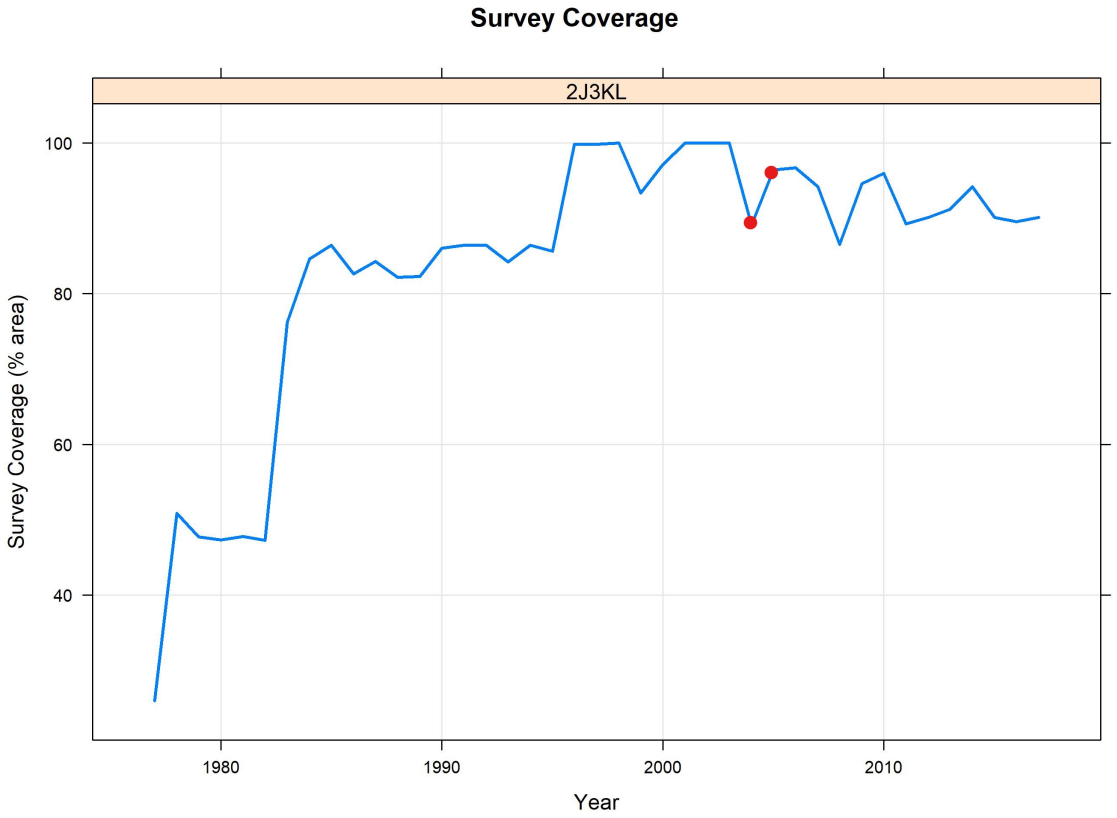

Figure 2.1: The percentage of survey coverage for witch flounder from the annual autumn DFO RV survey in NAFO Divisions 2J3KL. Red points show the survey coverage percentages in 2004 and 2005.

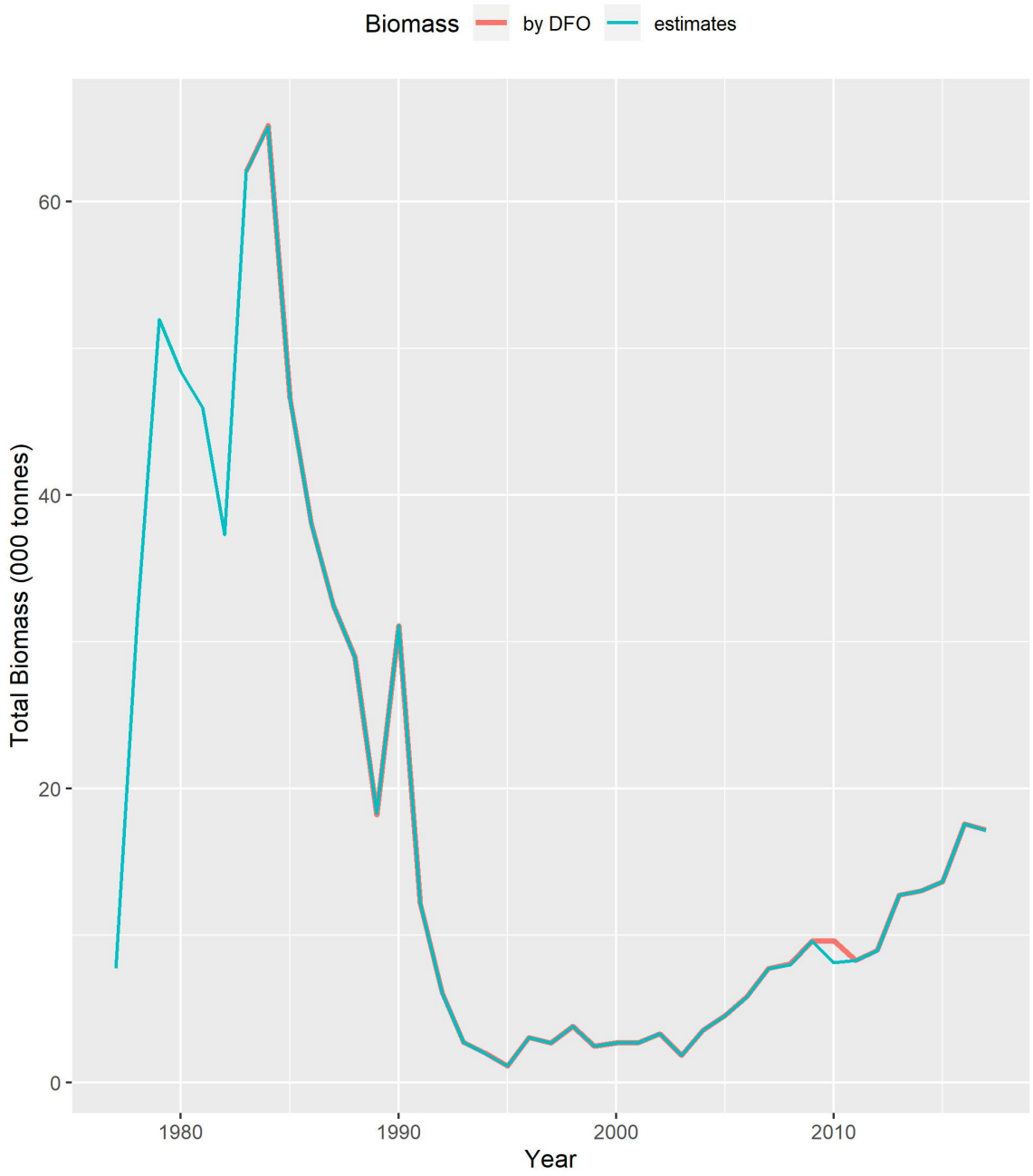

Figure 2.2: The comparison of estimated total biomass (blue line) with total biomass obtained by DFO (red line since 1983) in NAFO Divisions 2J3KL. They overlapped from 1983 to 2017, except in 2010.

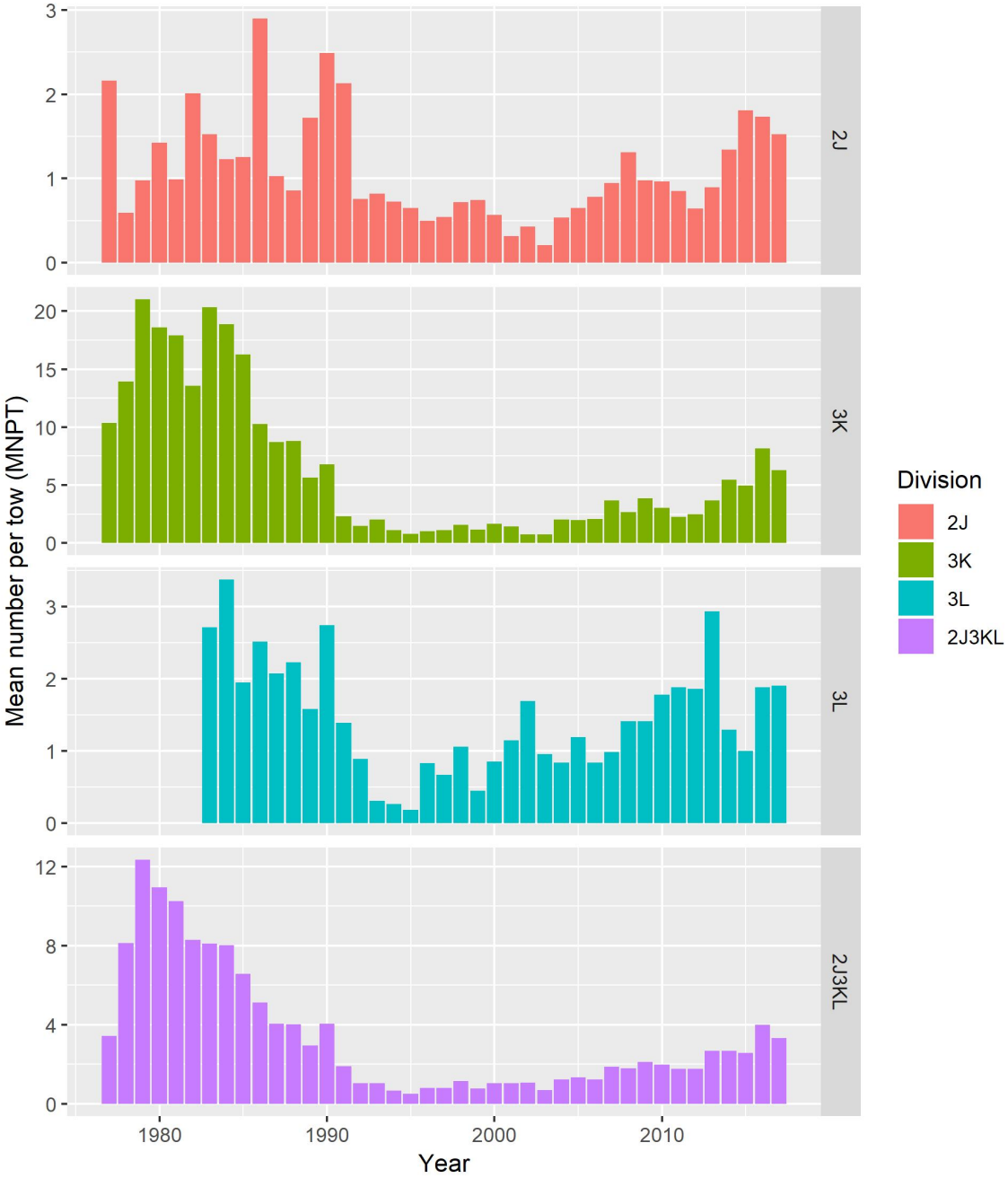

Figure 2.3: Mean catch number per tow (MNPT) in each NAFO division and stock area for witch flounder from annual autumn DFO RV survey in NAFO Divisions 2J3KL.

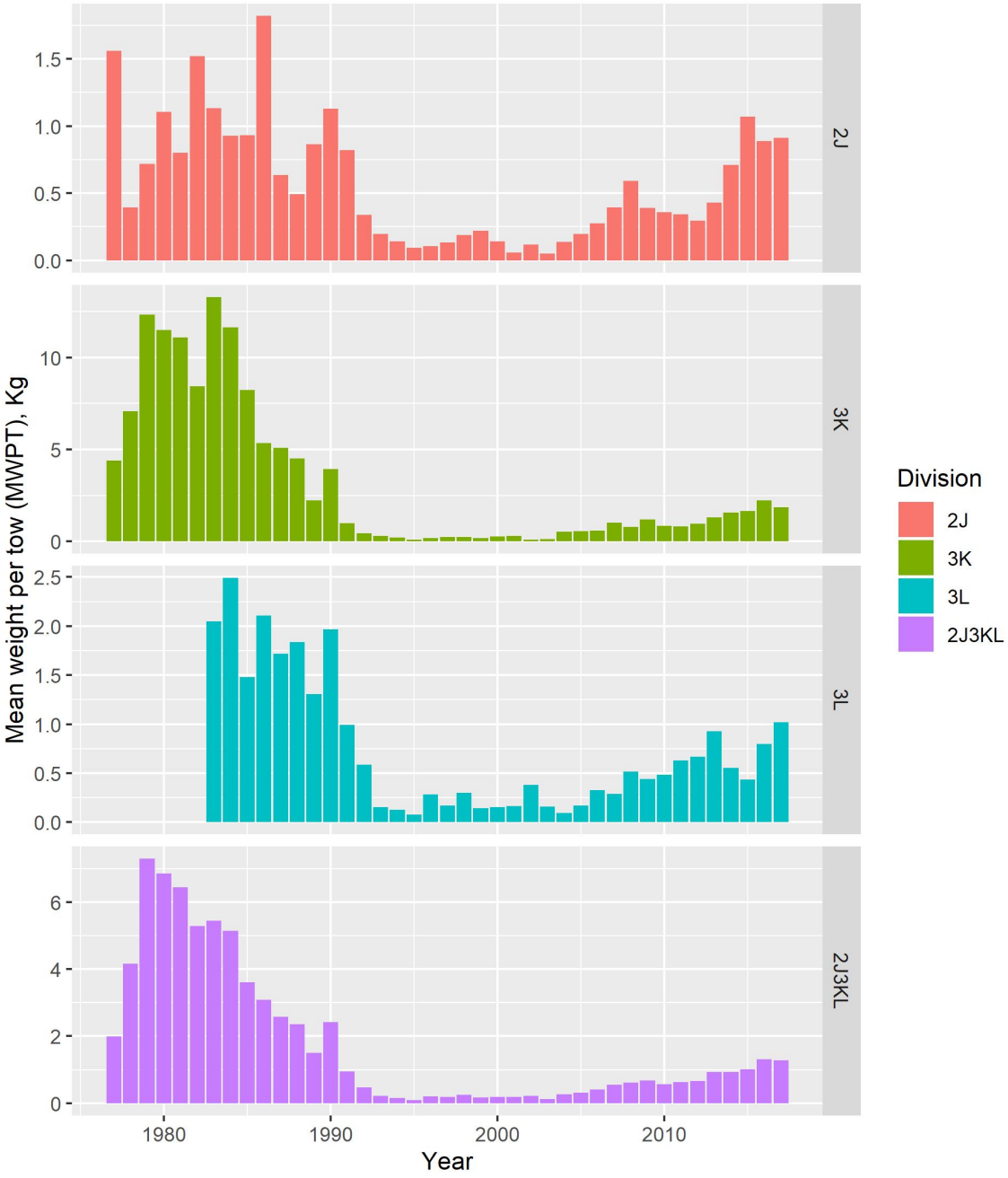

Figure 2.4: Mean catch weight per tow (MWPT) in each NAFO division and stock area for witch flounder from annual autumn DFO RV survey in NAFO Divisions 2J3KL.

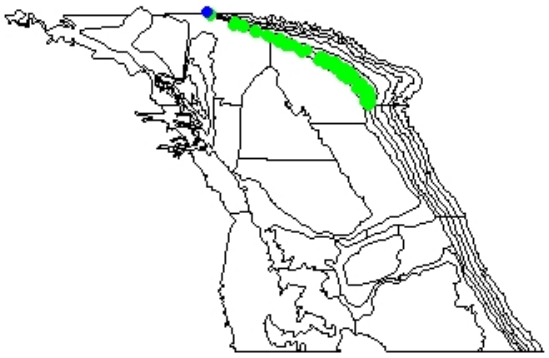

Figure 2.5: The blue point is an observation from the survey data, which was not contained within 2J3KL strata polygons. The green points indicate all observation in the stratum 216 in NAFO Division 2J. Therefore, the blue point was assigned to stratum 216.

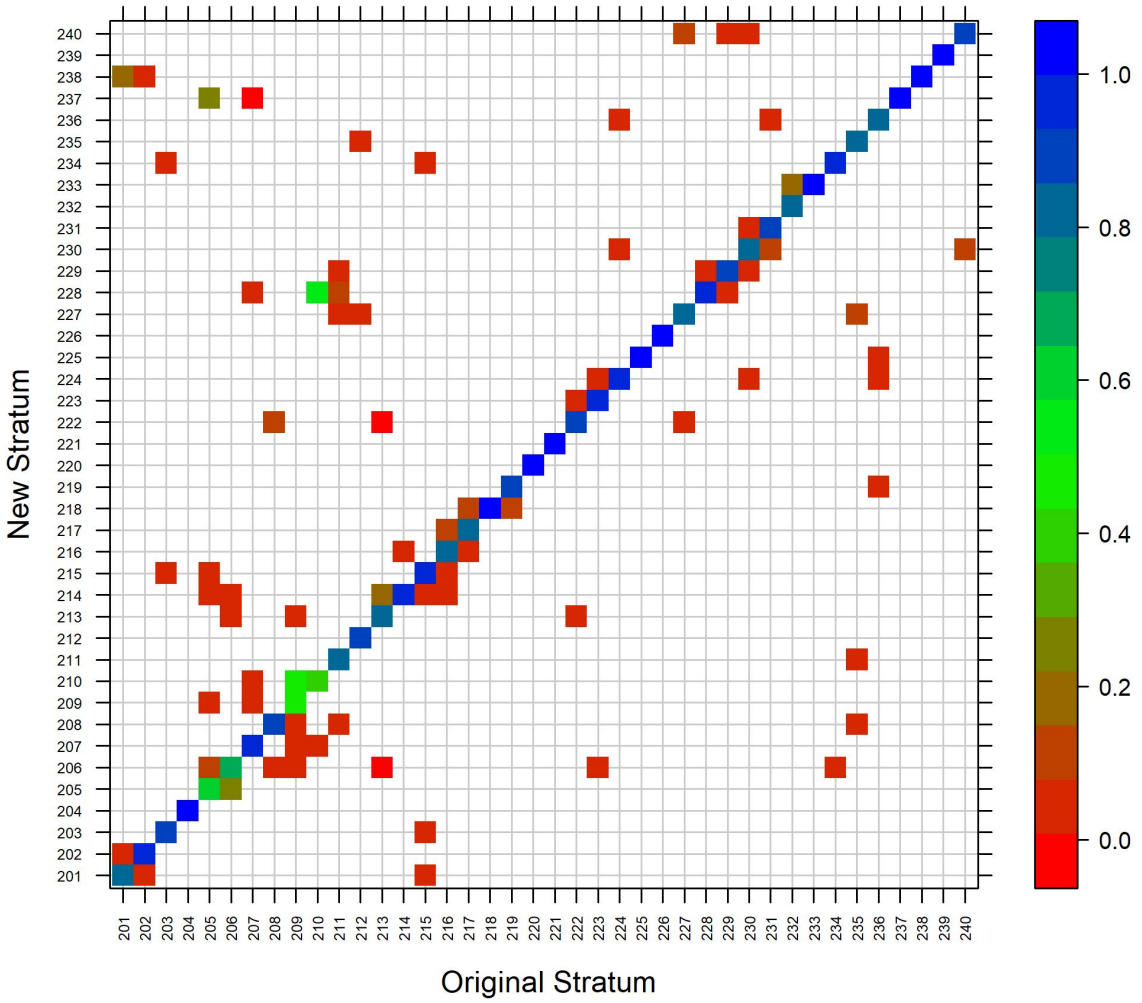

Figure 2.6: The proportion of observations (i.e. survey sets) from an original stratum which was assigned to new strata after the re-stratification in NAFO Division 2J. The colour bar at the right indicates the percentage of observations from original stratum which were assigned to new strata. For example, after the re-stratification, only about 65 percent (i.e. green grid square) of observations of the original stratum 206 were assigned to the new stratum 206. The rest of its observations, about 20 percent (i.e. dark green grid square), and less than 10 percent (i.e. red grid square) were assigned to new strata 205, and 213 and 214, respectively.

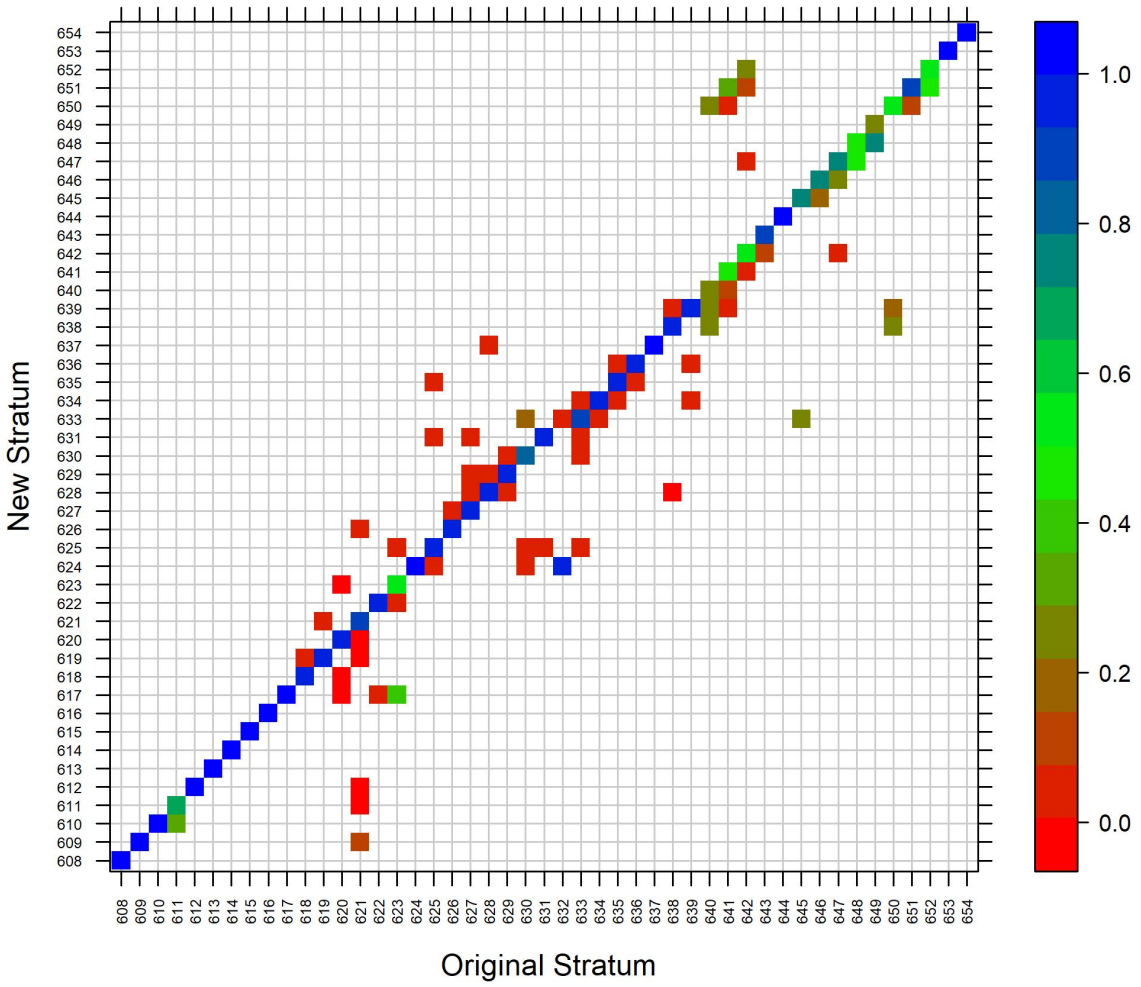

Figure 2.7: The proportion of observations (i.e. survey sets) from an original stratum which was assigned to new strata after the re-stratification in NAFO Division 3K. For additional description, see Figure 2.6.

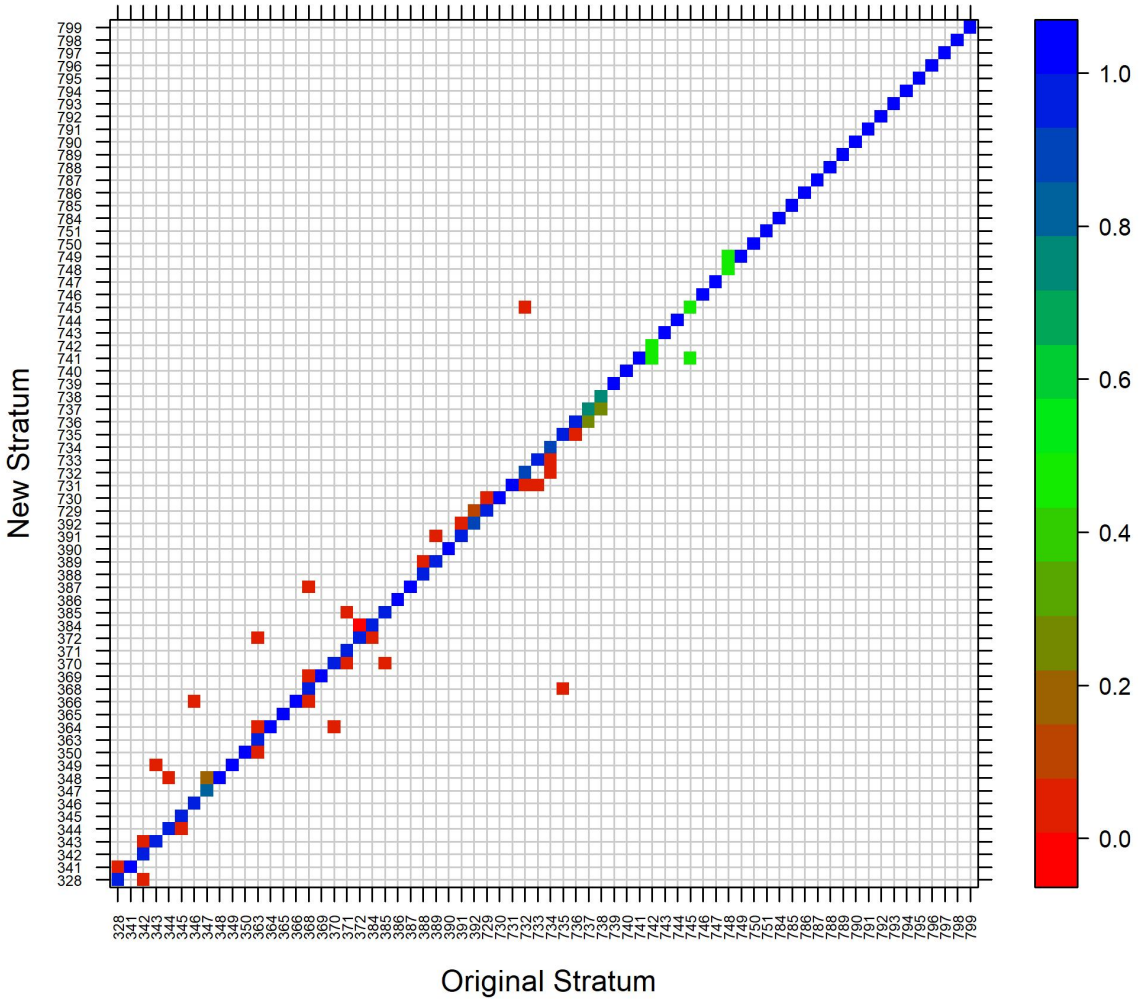

Figure 2.8: The proportion of observations (i.e. survey sets) from an original stratum which was assigned to new strata after the re-stratification in NAFO Division 3L. For additional description, see Figure 2.6.

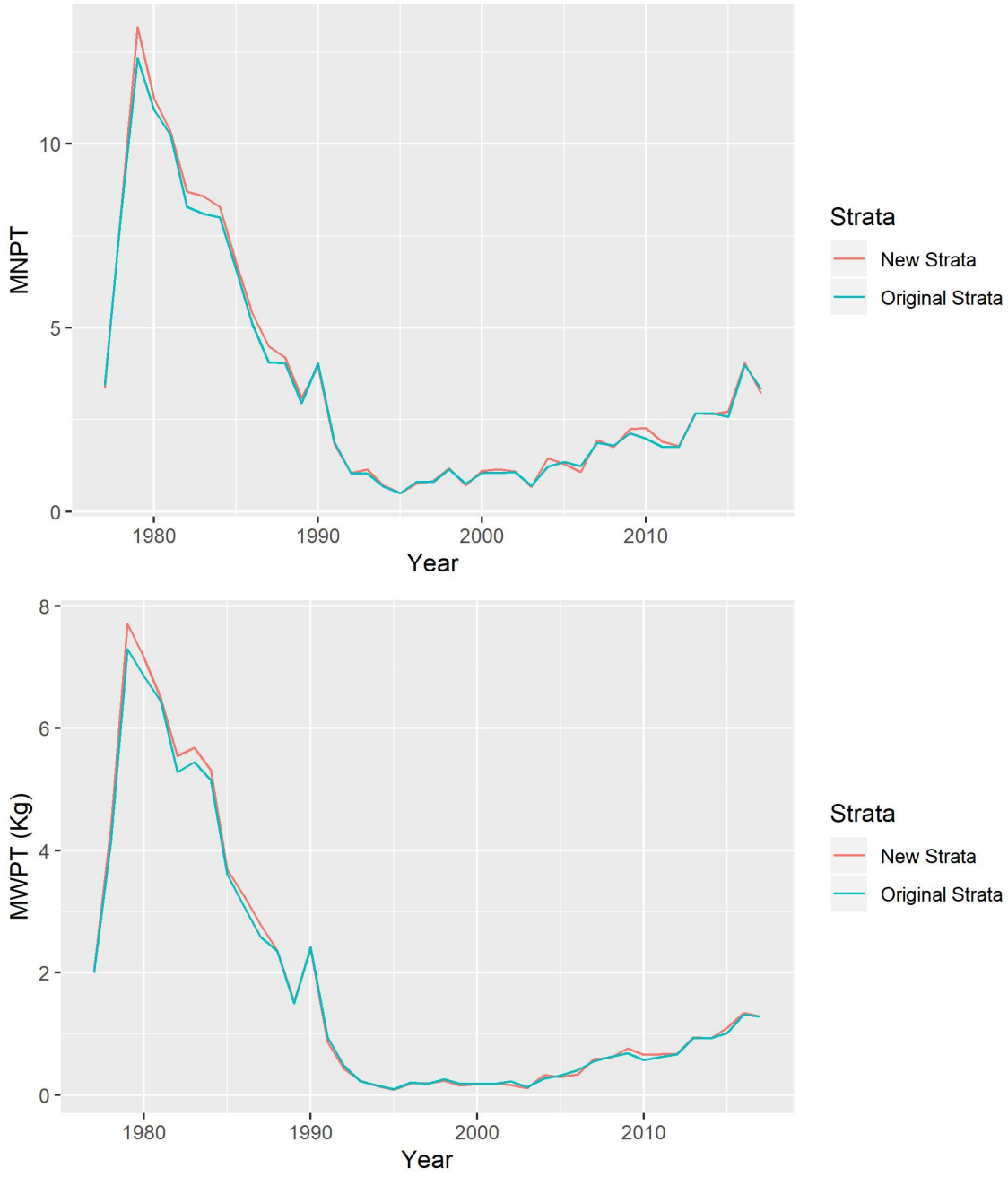

Figure 2.9: Mean catch number per tow (MNPT, top panel) and mean catch weight per tow (MWPT, bottom panel) based on original and re-stratified (i.e. new strata) witch flounder data from annual autumn DFO RV surveys in NAFO Divisions 2J3KL.

Chapter 3

Model-based Survey Indices

3.1 Introduction

In Chapter 2, I used the design-based approach to estimate the abundance indices for witch flounder survey data. According to Smith (1990), the design-based estimates from bottom-trawl surveys have large variance estimates, which may result in unreliable estimates of abundance and biomass. Smith suggested using statistical models to estimate more precise abundance indices. A model-based approach and a model-assisted survey sampling are two approaches that have been used to analyze fishery survey data to more accurately estimate abundance indices.

3.1.1 Model assisted survey sampling

The model-assisted survey sampling approach combines the design- and model-based approaches. Specifically, the model-assisted approach assists survey sampling using the best available model to provide an efficient estimator for a survey quantity. However, the model-assisted statistical inferences are based on the survey design. Model-assisted estimators of the finite population mean or total are approximately design-unbiased (i.e.

the design bias tends to be zero with large sample sizes). The estimators are efficient if the assumed model well describes the population data (Thompson, 2012). The generalized difference estimator, generalized regression estimator and pseudo-empirical likelihood method are three main model-assisted estimators proposed in the survey sampling literature (Chen et al., 2004). Moreover, the model-assisted approach can utilize auxiliary information available in a survey (e.g. bottom temperature) and suggest a regression estimator approach. The variance of the regression estimator is smaller than the variance of a design-based estimator (e.g. Horvitz-Thompson estimator; Horvitz and Thompson, 1952) when the population is well described by an assumed model. It is suggested that even when the population is not well described by the model, the regression estimator is still approximately unbiased (Särndal et al., 2003). Despite those advantages, the model-assisted approach assumes observations with no measurement error similar to the design-based approach.

Measurement error occurs during the data collection phase of a survey when the observed value of a variable of interest for a sampled unit differs from the true value (i.e. the value of the variable of interest in a population unit; Särndal et al., 2003). Ignoring the measurement error may have considerable effects on the survey estimates by inflating the variance estimates and biasing the estimates. Therefore, in this chapter, I use a model-based approach to include measurement error in estimating abundance indices such as mean catch number per tow (MNPT). The primary goal of this chapter is to use a mixed-effect model to predict the mean catches for missing strata to fill missing areas in the witch flounder survey data, which leads to a longer and hopefully more reliable time series of the MNPT.

In Section 3.2, I explain the model-based approach in detail and discuss the advantages of using statistical models. I then describe the generalized linear mixed models as a class of statistical models that I will use. Also, I introduce the template model builder package that can facilitate the required computations. In Section 3.3, I first explain Gaussian

Markov random fields that account for spatial and spatiotemporal correlation in witch flounder caught in the surveys. Second, I show how a model-based MNPT can be derived for all and sampled-only strata in a survey that are described further in this section. I then present and describe six mixed-effect models to find the best-fitting model for the survey data. Third, I use the bias-correction method implemented in the template model builder package to reduce the transformation bias of model MNPT. I then examine the performance of the model chosen by comparing model-based (for all and sampled-only strata) and design-based estimation of divisional MNPT for diagnostic purposes. I also use cluster analysis to summarize spatiotemporal variation in survey catches. In Section 3.4, I first use residuals diagnostics and goodness-of-fit tests to investigate if the model used was correctly specified, and then I address temporal and spatial autocorrelation tests. Finally, in Section 3.5, I briefly address the main findings of the mixed-effect model developed in this chapter.

3.2 Model-based approach

In a model-based approach, the value of the variable of interest (i.e. the amount of catch, Y) is considered as a random variable that follows a parametric probability distribution (Thompson, 2012). In this approach, stock size indices are estimated via the estimation of model parameters (Chen et al., 2004). Model-based estimates of stock size and their variances are based on the survey observations and their assumed probability distribution (Smith, 1990).

There are various advantages associated with the use of a model-based approach. The first advantage is that it can incorporate the measurement error and missing survey data, thus improving the reliability of the estimates and adjust for biases of stock size estimates due to missing data (Thompson, 2012). Moreover, a model-based approach allows for the assessment of the efficiency of proposed estimators as well as population

assumptions (Thompson, 2012). Statistical models, therefore, can produce estimators that use the survey (i.e. sample) data efficiently. Finally, it is straightforward to include auxiliary variables in such models. For example, environmental variables such as bottom temperature or depth information that were measured during the sampling can be incorporated into a statistical model to improve estimation of the stock size index (Shelton et al., 2014). However, statistical models for survey data may produce a biased estimate of the finite population mean, when the expected value of an estimator of a population quantity does not equal to the finite population mean (Smith, 1990). An obvious reason for this is if the statistical model is mis-specified or poorly specified so that the model does not adequately describe response values throughout the population. Models with simple parametric covariate effects may suffer from this problem. This is why less restrictive and more flexible non-parametric regression models or spatial random effects models are often preferred when modelling complex survey data. In this chapter, I explore the model-based approach to account for the measurement error associated with witch flounder survey data to improve the reliability of abundance indices. More importantly, I use mixed-effect models to address changes in witch flounder survey coverage using the prediction of abundance indices for years with missing strata and to increase the length of abundance indices time series.

Mixed-effect models are statistical models that include both fixed effects and random effects. Fixed effects are non-random or fixed parameters of the models, while random effects are model parameters that are considered as random variables. Random effects resulting from a random process (e.g. variation over time, space; Thorson and Minto, 2015) can account for data (i.e. survey data) that are not statistically independent. Generalized linear mixed models (GLMMs) with both fixed and random effects are suitable to account for the correlation in witch flounder survey catch data.

3.2.1 Generalized linear mixed models (GLMMs)

GLMMs are an expansion of generalized linear models (GLMs) that are a generalization of classic linear models (i.e. regression models). GLMs allow fitting a relationship between linear predictors ($X\beta$) and a response variable (dependant variable, Y) via a nonlinear function (link function). In these models,

$$E(Y) = \mu = g^{-1}(X\beta) \tag{3.1}$$

where $E(Y)$ is the expected value of Y, denoted as μ. The distribution of Y is assumed to belong to the exponential family, which includes the normal, Poisson and the negative binomial distributions. In equation (3.1), $X\beta$ is the linear predictor in which X is a model matrix for fixed effects (β) and $g(.)$ is the link function. The linear predictor is the combination of unknown parameters β as fixed effects that can be estimated using the maximum likelihood method. In GLMs, the distribution of response variables can follow a distribution other than the normal (Mccllagh and Nelder, 1989).

GLMMs are GLMs but also include random effects and have been widely used in recent years. GLMMs are first formulated by conditioning on the random effects (u),

$$E(Y|u) = \mu = g^{-1}(X\beta + Zu) \tag{3.2}$$

where X is a model matrix for fixed effects, β is a vector of parameters (fixed effects), Z are random effects model matrix and u are the random effects.

One benefit of including random effects is to better model variation. For example, a GLMM can model temporal variation by incorporating year as a random effect. Random effects can account for different types of correlation in the response variable, including spatial and temporal correlation. Random effects can be used to analyze data with lots of levels or groups (i.e. strata in this study) and/or when there is an inconsistent sampling

across levels. For example, witch flounder survey data includes lots of strata (155 strata), and some of them were not sampled for several years. Therefore, I apply a GLMM with random effects discussed in the following section to analyze the survey data.

I use the maximum likelihood estimation method to fit a series of GLMMs and the Akaike information criterion, AIC (Sakamoto et al., 1986), and Bayesian information criterion, BIC (Schwarz et al., 1978) to select the best GLMM for the witch flounder survey data. A maximum likelihood estimate is the value in the parameter space that maximizes the likelihood function, which is based on the probability model for the response variable. Maximizing a likelihood function is equivalent to minimizing the negative logarithm of the likelihood function (negative log-likelihood, NLL). Due to the computational complexity (e.g. a large number of random effects) involved in the survey models, the estimation of parameters requires extensive computational efforts. Therefore, advanced tools such as template model builder (TMB) have been developed to facilitate the computational process of fitting these statistical models and the estimation of model parameters.

3.2.2 Template model builder (TMB)

TMB is a computational R software (R Core Team, 2013) package developed by Kristensen et al. (2016) at the Danish Technical University. The TMB package is based on the automatic differentiation model builder (ADMB, Fournier et al., 2012) package that was designed for large, complex mixed-effects models. ADMB and TMB use automatic differentiation (AD) to assist with estimating the parameters for non-linear statistical models. The TMB package assists users to efficiently estimate and compute complex random effect models using C++ templates (Eddelbuettel and Balamuta, 2017). The TMB package is capable of combining different templates in R to efficiently implement the Laplace approximation of the marginal likelihood where the random effects are integrated to produce the marginal likelihood. TMB also provides exact first-order derivatives of the marginal loglikelihood. Moreover, the TMB package provides features, including,

for instance, automatic sparseness detection and parallel user templates. The R package glmmTMB (Magnusson et al., 2017) can also be used to fit simple GLMMs. As the name suggests, glmmTMB is based on TMB. This package also uses marginal maximum likelihood estimation via TMB to estimate model parameters. In the glmmTMB package, random effects follow normal or multivariate normal distributions.

3.3 Mixed-effect Models

In this section, I present several mixed-effect models for the witch flounder survey data. The models allow examination of the spatial and spatiotemporal variation in witch flounder caught over time. Note that the witch flounder survey data is a spatiotemporal data set since each observation of the survey is characterized by a geographic position (i.e. longitude and latitude) in a specific year. Gaussian Markov random fields (GMRFs) are widely used to approximate the spatial and spatiotemporal covariance structures. The covariance function of a GMRF for such data is a function of the relative position of two locations. Therefore, I estimate the random effects as GMRFs in GLMMs to analyze the witch flounder spatiotemporal data set. A model with spatial and temporal correlations can be used to estimate mean catches in unsampled areas by using catch information in neighbouring (in both space and time) areas.

The Gaussian (i.e. normal) distribution is one of the most important probability distributions in statistics. An n-dimensional random vector Z has a multivariate normal distribution in which probability distribution function of that is given by

$$f_Z(z) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2} z^\top \Sigma^{-1} z\right\}, \quad \text{for} \quad z \in \mathcal{R}^n \tag{3.3}$$

where Σ is the covariance matrix of random vector Z, and $f_Z(.)$ is a probability distribution function of Z (Hogg et al., 2005). The random vector Z that satisfies (3.3) is written

$Z \sim \text{MN}(0, \boldsymbol{\Sigma})$.

3.3.1 Gaussian Markov random fields (GMRFs)

A GMRF is a finite-dimensional random vector that follows a multivariate normal distribution with conditional independence assumptions that are commonly called the Markov property. To illustrate what the conditional independence means, let $X = (X_1, X_2, X_3)^T$ be a random vector; if X_1 and X_2 are conditionally independent given $X_3 = x_3$ then exploring X_2 reveals nothing new about the distribution of X_1 that implies:

$$p\left(X_1 = x_1 | X_2 = x_2, X_3 = x_3\right) = p\left(X_1 = x_1 | X_3 = x_3\right)$$

where p is a conditional distribution (Rue and Held, 2005). A Markov chain is a discrete-time stochastic process that has the Markov property. A sequence of random variables $X_1, X_2, \ldots$ forms a Markov chain (Gagniuc, 2017) if for $x_1, x_2, \ldots, x_{n+1}$

$$P\{X_{n+1} = x_{n+1} | X_1 = x_1, ..., X_n = x_n\} = P\{X_{n+1} = x_{n+1} | X_n = x_n\}. \tag{3.4}$$

In a Markov random field (e.g. GMRF), the Markov property can be extended to two or more dimensions. For example, in a two-dimensional GMRF, the Markov property represents that a random variable s in the site i (or stratum i in this study) depends only on a set of neighbours, Δ_i to each site i. Two sites are neighbours when they share a common border such that if $i \in \Delta_j$ then also $j \in \Delta_i$. Therefore, the conditional distribution $p\left(s_i | s_{-i}\right)$, where s_{-i} denotes all elements but s_i, depends only on the $\Delta_i = \{s_g, s_h, s_j, s_m, s_n, s_q, s_r, s_t\}$, the set of neighbours to s_i, not on the rest of the s's (grey cells in Figure 3.1). Note that sometimes the neighbours are only defined to be $\{s_h, s_m, s_n, s_r\}$.

Later in this section, I consider a random effect, the main stratum effect, as a GMRF to account for spatial catch dependency (or correlations in catches among neighbouring

strata) in the survey data. Therefore, the GMRF covariance matrix, $\mathbf{\Sigma}$, needs to be specified to find the correlations (i.e. neighbours). However, in practice, the matrix inverse of the covariance matrix, the precision matrix (PM) Q, is often specified as follows:

$$\mathbf{Q}_{ij} = \begin{cases} -q & \text{if strata } i \text{ and } j \text{ are neighbours} \\ q(n_i + d) & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

where n_i is the number of neighbours of stratum i, and q and d are parameters that are estimated. A simple expression for the covariance based on the q and d parameters does not exist. However, in Section 3.3.3 I illustrate this. The reason for using the PM, Q, is that the Markov property makes most elements of the PM zero. For example, when x_d and x_s of a multivariate normal are conditionally independent (i.e. Markov property) given the other variables $\{x_k : k \neq d\ , s\}$, the Q_{ds} and Q_{sd} elements of the PM will be zero and vice versa. This means that the PM is a sparse matrix (i.e. a matrix in which most of the elements are zero) when many of the variables are conditionally independent. The sparse structure of the PM relates to the concept of neighbours: $Q_{ij} \neq 0$ if and only if i and j are neighbours. The sparse structure of the PM also accelerates computations of GMRF because we do not need to compute terms that are zero while using the covariance matrix involves many computer operations (Banerjee et al., 2008). Also, when using the PM, we do not need to do a matrix inversion in equation (3.3) every time the density function is evaluated.

An autoregressive (AR) model is used to describe specific time-varying processes. The autoregressive process of order 1, AR (1), is the simplest case of a GMRF. In an AR (1) process, the output variable depends linearly on its own previous value and on a

stochastic term:

$$x_t = \Phi x_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2), \qquad |\Phi| < 1 \tag{3.5}$$

where Φ is a model parameter, the index t represents time and ϵ_t is a stochastic term, which follows a normal distribution (Rue and Held, 2005). Regarding conditional independence assumptions, equation (3.5) can be written:

$$x_t | x_1, \ldots, x_{t-1} \ \sim N(\Phi x_{t-1}, \sigma_\epsilon^2), \tag{3.6}$$

for $t = 2, ..., n$, while for $1 < i < t$, x_i depends both on x_{i-1} and x_{i+1}. In Section 3.3.3, I first consider the main year effect and the year effect within each stratum as AR (1) to present the temporal variation and annually spatial variation in the response variable. However, in other models, the year effects in each stratum will include both temporal and spatial correlations.

I explain the GLMMs fitted to the survey data below (see the symbols and data used in Table 3.1). Random effects of the models represent spatial and temporal variations of witch flounder caught during surveys. The estimated model parameters (fixed and random effects) can be used to predict the mean catches for each survey observation.

3.3.2 Estimating the model-based MNPT

The amount of catch (i.e. the number of caught witch flounder), Y, is assumed to follow a negative binomial (NB) distribution, which is parameterized with mean μ , and scale parameter k, such that μ is the expected value of Y and $\mu + \frac{\mu^2}{k}$ is the variance of Y. The NB distribution is often found to be suitable for modelling catches from bottom-trawl surveys (e.g. Cadigan, 2011). Here I describe a model-based estimate of the population mean per trawl unit (i.e. model MNPT) for all and sampled-only strata in a survey. As

illustrated in equation (2.7) in Chapter 2, the design-based MNPT for year t is calculated by

$$\bar{y}_t = \frac{\sum_{h \in H_{sample}} N_h \bar{y}_{h,t}}{\sum_{h \in H_{sample}} N_h}, \tag{3.7}$$

where H_{sample} represents sampled-only strata in survey year t (see design-based notations in Table 2.1 in which H=H_{sample}), $\bar{y}_{h,t}$ is sample mean in the stratum h for year t, and $\bar{y}_t$ is an estimate of $\bar{Y}_{t,sample}$ that represents population mean per trawl unit for sampled strata in year t,

$$\bar{Y}_{t,sample} = \frac{\sum_{h \in H_{sample}} N_h \bar{Y}_{h,t}}{\sum_{h \in H_{sample}} N_h}. \tag{3.8}$$

A model-based estimate of $\bar{Y}_{t,sample}$ is

$$\bar{\mu}_{t,sample} = \frac{\sum_{h \in H_{sample}} N_h \hat{\mu}_{h,t}}{\sum_{h \in H_{sample}} N_h}, \tag{3.9}$$

where $\hat{\mu}_{h,t}$ is an estimated mean catch for stratum h in year t using a model, and $\bar{\mu}_{t,sample}$ as a strata size-weighted average, is a model-based MNPT for sampled-only strata. I estimate model-based MNPT (i.e. model MNPT) for sampled-only strata (3.9) for diagnostic purposes. Since the design-based MNPT is unbiased for the sampled strata, we expect the model MNPT for sampled-only strata to be similar to design-based MNPT. Also, the estimation of model MNPT for sampled-only strata allows a separate examination of the impact of a model on MNPT for sampled strata, and the impact of a model for interpolation of non-sampled strata.

A stock assessment problem with using the MNPT in equations (3.7) or (3.9) as an index of stock size is that the sampled strata may not be the same each year, and H_{sample} may also differ. Hence, $\bar{y}_t$ may not be an estimate of the same part of the stock each

year. Ideally, we would like to have a MNPT estimate for the same strata every year, and for as many strata as possible to ensure that the entire stock is covered by the survey indices. For this reason, we apply H_{all}, which represents all strata that are in the target area in a survey. We use a spatiotemporal model to compute a model-based estimate of $\bar{Y}_t$,

$$\bar{Y}_t = \frac{\sum_{h \in H_{all}} N_h \bar{Y}_{h,t}}{\sum_{h \in H_{all}} N_h} \tag{3.10}$$

which is

$$\bar{\mu}_t = \frac{\sum_{h \in H_{all}} N_h \hat{\mu}_{h,t}}{\sum_{h \in H_{all}} N_h}, \tag{3.11}$$

a model-based MNPT for all strata in the survey. I will then use the best-fitting model detailed below to determine model-based MNPT for all (equation 3.11) and sampled-only (equation 3.9) strata in the witch flounder survey.

3.3.3 Examination of fitted models

Six models described below were examined to find the best-fitting model. In these models, I considered strata that have average catch greater than 0.1 because only strata with non-zero average catch represent the spatial and spatiotemporal variation in the survey data. The means for strata with average catch less than 0.1 were simply fixed at zero when computing MNPT. The best-fitting model can efficiently model random effects (Table 3.2) to explain the spatial and spatiotemporal variations of the witch flounder survey data.

First, I fitted a model with the main stratum effect as a fixed effect and the strata-year effect as a random effect (3.12). The mean catch for observation i of stratum s in year t

was approximated by

$$\log \mu_i = \log \mu_s + \log \gamma_{s(i),t(i)} + \text{offset}(i) \tag{3.12}$$

where μ_s is the main stratum effect (fixed effect) and γ is the strata-year effect (random effect; see the definition of symbols, $s(i)$ and $t(i)$ in Table 3.1), which is AR(1) over years within each stratum such that they are independent across strata. The offset term (in equations 3.12-3.16) represents a function of distance towed for each observation. It is used to standardize the data to a common tow distance. The witch flounder data were collected using two surveys gears, the Engel 145 otter trawl and the Campelen 1800 shrimp trawl. However, I used the converted Engel catch data, that were converted to be equivalent to Campelen catches (see, e.g. Warren, 1996 and Warren et al., 1997). I used a function of distance towed (e.g. log (distance towed/0.8), in which 0.8 is the targeted distance towed (nautical miles, nm) of the Campelen trawl to adjust for variations in the actual tow distance from the target of 0.8 nm. Usually the offset will be zero unless the tow distance is different than 0.8 nm.

I used the TMB package to fit this model and estimate model parameters (Tables 3.2 and 3.3; the TMB and C++ codes for the model shown in Appendix A). I also used the glmmTMB package to fit this model, which is an easier way to fit simple GLMMs than the TMB package. Indeed, glmmTMB can estimate model parameters with only one-line of R code:

```
glmmTMB(number ~ 0 + fstrat + ou(times+0|fstrat) +
        offset(offset),data=subd, family=nbinom2)
```

where "fstrat" is the fixed effect μ_s, and "times" is the random effect of year within the strata group. For fitting the model with this package, I used the Ornstein–Uhlenbeck (OU, Magnusson et al., 2017) covariance structure instead of AR(1) for the random effect. The advantage of OU is that it can manage irregular time points (i.e. year gaps

in strata) in a data set. The R code is another simple representation of the model. The values of the model parameter estimates and AIC, BIC, and NLL were the same for both TMB and glmmTMB. The TMB and the glmmTMB packages allow using different covariance structures and provide standard deviation estimates of all estimated parameters. However, the glmmTMB package cannot be easily used to estimate standard errors for MNPT and other derived parameters. In contrast, the TMB package can easily estimate standard errors for derived parameters using the 'ADREPORT' feature (see Appendix A).

In the second model, I added a main year effect, ε, that is AR(1) and common for all strata, in addition to the strata-year effect γ in equation (3.12). The second model is:

$$\log \mu_i = \log \mu_s + \log \varepsilon_{t(i)} + \log \gamma_{s(i),t(i)} + \text{offset}(i). \tag{3.13}$$

I fitted this model again using the TMB and the glmmTMB packages to estimate the model parameters. The values of model AIC, BIC, and NLL were the same for these two packages (Table 3.6).

In the third model, I incorporated another random effect, the main stratum effect (or main grid effect, ω) as a GMRF. I also considered the main strata effect, m, as the only fixed effect in the model.

$$\log \mu_i = \log m + \log \omega_{s(i)} + \log \varepsilon_{t(i)} + \log \gamma_{s(i),t(i)} + \text{offset}(i) \tag{3.14}$$

The random effect ω as a GMRF follows a MN$(0,\Sigma_\omega)$, where Σ_ω is a covariance matrix. As I mentioned before, to improve the computational process the inverse covariance matrix $\Sigma_\omega^{-1} = Q_\omega$ is computed, so that if stratum i and stratum j are neighbours, they are correlated (see Figure 3.2) and $Q_{\omega_{ij}} = Q_{\omega_{ji}} \neq 0$, otherwise $Q_{\omega_{ij}} = 0$ (as Figure 3.1 shows). This model was fitted using only the TMB package. It was not obvious how this model could be fit with glmmTMB. The values of the model comparison criteria are

shown in Table 3.6.

In the fourth model, I investigated another strata-year effect that is GMRF correlated across strata and AR(1) correlated across years (3.15). This model differs from equation (3.14) where the γ random effects are assumed to be correlated across years but independent between strata, whereas in equation (3.15) the δ's are correlated across years and strata. For example, in the random effect δ, two neighbouring strata, stratum i in year t_1 and stratum j in year t_2, are GMRF correlated and also AR(1) correlated for two different years t_1 and t_2.

$$\log \mu_i = \log m + \log \omega_{s(i)} + \log \varepsilon_{t(i)} + \log \delta_{s(i),t(i)} + \text{offset}(i) \tag{3.15}$$

where the random effect ω as a GMRF follows a MN(0, Q_ω) for $Q_\omega = {\Sigma_\omega}^{-1}$, and the strata-year effect, δ as an AR(1) has a MN(0, Q_δ) distribution, $Q_\delta = {\Sigma_\delta}^{-1}$; that is, if δ_t is the vector of stratum effects for year t then $\delta_t - \rho_\delta \delta_{t-1} \sim MN(0, Q_\delta)$. Note that Q_ω and Q_δ are the PMs of the covariance matrices.

For the fifth model, I removed the main year effect, ε, in the fourth model, because the estimated standard deviation of the main year effect was close to zero in the fourth model (Table 3.6). In this model (3.16), I considered only two random effects ω and δ explained in the third and fourth models.

$$\log \mu_i = \log m + \log \omega_{s(i)} + \log \delta_{s(i),t(i)} + \text{offset}(i) \tag{3.16}$$

In the sixth model, I considered the stratum effect, ω as a GMRF based on the distance among strata, GMRF$_D$. Indeed, I produced a spatio-correlation matrix using the distance inverse between the centers of neighbour strata. For those strata that are not neighbours, elements of this matrix are zero. The PM of GMRF$_D$ that is computed based on the spatio-correlation matrix, is a sparse matrix and represents the distance factor (i.e. $(\frac{1}{distance})^\tau$, precision proportion) between neighbour strata. In this model,

ω is the mean stratum effect (mean grid effect) that has a MN(0,Σ_ω) distribution and it differs from ω in the fifth model. Note that the PMs of the stratum effect in these two models are different, because the PM in the sixth model depends on the distance between strata. In model six, the strata-year effect, δ is also different from the model five due to including the distance factor in GMRF; thus, their PMs are also different. The model parameter τ determines the strength of the distance factor among strata. When an estimate of τ is zero or close to zero, it means that the distance between strata does not affect the correlation between them, and then the model becomes a regular GMRF.

Table 3.6 shows that based on NLL and two information criteria (or ΔAIC and ΔBIC) the last two models fit better than the first four models. In the sixth model, the estimate of τ is close to zero (= -0.155), which indicates that the distance is not a significant factor in this model. Therefore, the fifth model is the best-fitting model selected among those I examined. This model accounts for correlation between catches in neighbour strata and adjacent years to estimate abundance indices and predict these indices for those areas that were not sampled in some survey years. The model R and C++ codes are shown in Appendix B.

The computer program automatically predicts mean catches in areas that were not sampled using the estimated correlations in nearby sampled areas. For unsampled strata, MNPTs were derived from neighbours via the model. MNPT (as a quantity derived from the model, see equation 3.11) is calculated by a nonlinear function (i.e. an exponential function in this study) of fixed and random effect estimates. This nonlinear transformation of random effects can produce bias in the derived MNPT. Thorson and Kristensen (2016) introduced a bias-correction estimator for quantities (e.g. MNPT) derived from mixed-effect models. The bias-correction estimator was implemented in the TMB package. Using the bias-correction method, Figure 3.3 shows the differences between the predicted model MNPTs and the bias-corrected model MNPTs for each division and the entire stock area. Although the differences between bias-corrected estimates and non-

corrected estimates are usually small, the bias-corrected model MNPTs are greater than the model MNPTs, particularly in the first years of the time series.

Next, I compared the model-based MNPT (for all and sampled-only strata) and design-based MNPT to show the performance of the model and design. To make the comparisons, the divisional model MNPT for sampled-only (3.9) and all (3.11) strata and the divisional design-based MNPT were plotted against each other (Figures 3.4-3.7). Note that for these comparisons, I used the bias-corrected model MNPT for all and sampled-only strata.

I compared model MNPT for sampled-only strata and design-based MNPT to diagnose the model used and show how the model performed in terms of index estimation compared to the design-based approach, which provides an unbiased index for sampled strata. I expected that the model MNPT for sampled-only strata to be similar to the design-based MNPT (in the top panel of Figures 3.4-3.7). For example, the top panel in Figure 3.7 (for the entire stock area, 2J3KL) shows that the bias-corrected model MNPTs for sampled-only strata are a little greater than the design-based MNPTs (i.e. those from the re-stratified data) from 1977 to 1990. Since then, they are almost equal, as I expected. This suggests that the model used is appropriate to predict MNPT, particularly after 1990.

The model MNPT for all strata were compared with the design-based MNPT to diagnose the design-based method used in Chapter 2 and show how the design performed in terms of index estimation compared to the model-based approach used (the middle panel of Figures 3.4-3.7). For example, in the middle panel of Figure 3.7, the design-based MNPTs are greater than the bias-corrected model MNPTs for all strata from 1977 to 1983. Since then, they are almost equal, except for some years that the model MNPTs are a little greater than those design-based MNPTs.

Also, the model MNPT derived from all strata were compared with the model MNPT derived from sampled-only strata to show the impact of model interpolation (the bottom

panel of the figures). For example, in Figure 3.7, the bottom panel shows that the bias-corrected model MNPTs for sampled-only strata are greater than those for all strata in years before 1990, particularly from 1977 to 1983. However, they are almost equal after 1990, which indicate that the impact of the model for interpolation of unsampled strata is small after 1990.

3.3.4 Cluster analysis

I used clustering as a way to summarize the spatiotemporal variation in the survey data. I clustered strata based on the spatial correlation matrix derived from the estimated strata-year effect matrix. Clustering is a method to group objects which are similar to each other based on a feature. Here I used cluster analysis to detect the similarities between strata based on the spatiotemporal variation of the mean catches among strata. I used the R package "clValid" (Brock et al., 2008) to determine the most appropriate clustering method and the optimum clusters for the data set. I clustered strata based on a hierarchical method that the cluster validation test suggested using the package. The test also indicated that the optimum clusters (i.e. groups) for strata are two clusters. However, I cut the dendrogram at three clusters (one cluster more than optimum clusters), because three clusters show better the similarities of strata with each other than two clusters (Figure 3.8). Regarding cutting the strata dendrogram, strata clusters were plotted on the map (Figure 3.9) to visually show which strata are similar to each other based on the hierarchical method used. Note that in Figure 3.9, the pink regions represent strata with the mean catch was fixed to be zero and were not used in the clustering process. Also, the model MNPT over years by cluster is shown in Figure 3.10 to demonstrate the overall spatiotemporal variation in MNPT.

3.4 Model diagnostics

3.4.1 Residual diagnostics and goodness-of-fit tests

I first examined the model via residual diagnostics. I used the DHARMa package in R (Hartig, 2020) to detect non-randomness in the model residuals to investigate if the model was correctly specified. Specifically, the DHARMa package for GLMMs creates interpretable residuals standardized to values between 0 and 1 (i.e. quantile residuals) by a simulation-based approach. The DHARMa package first simulates new data from the fitted model for each observation. Second, it calculates the empirical cumulative density function (ECDF) for the simulated observations. This ECDF for the simulated observations characterizes the values (and their probability) at the predictor combination of the observed values, assuming that the fitted model is correct. The package then considers the residual as the value of the empirical density function at the value of the observed data. When a residual equals zero, it means that all simulated values are greater than the observed value, and the quantile residual is equal to 0.5 when half of the simulated values are greater than the observed values. If a GLMM is correctly specified, the distribution of residuals has a uniform (0,1) distribution.

For the selected model, Figure 3.11 shows a pattern (dashed red line, i.e. the 0.5 quantile regression line) in residuals around the 0.5 quantile residuals (i.e. the median of residuals, solid red line). This pattern indicates deviations of residuals from a uniform distribution (uniformity in y-direction). Despite the pattern in the residuals, the model results are use-able. However, they should be used with caution as there may be model-misspecification. In a GLMM, particularly when the maximum likelihood estimation method is used to fit the model, random effect estimates are often slightly biased. For this reason, DHARMa often shows a slight pattern in the residuals, even if the model is correctly specified, and tests on residuals can get significant for large sample sizes. Therefore, it is suggested to further investigate the model in future research.

Next, I examined the model using goodness-of-fit tests on the quantile residuals, such as a qq-plot, outliers, dispersion, and zero-inflation tests. The qq-plot test explores the uniformity of these quantile residuals such that their trend must match with a straight line (a red line). In Figure 3.12, the qq-plot trend follows the red line, except in the tails, which denote deviations from the expected distribution. The DHARMa package also provides the Kolmogorov-Smirnov test to check the uniformity of the residuals. The uniformity test shows a lack of fit in this model (the test p-value = 0.01404<0.05 then the test is significant). Also, the package provides the deviation of data points outside the range of model-simulated values (i.e. outliers that were shown as red stars in Figure 3.11). The outliers test p-value is greater than 0.05 (= 0.42166), which indicates that there are no outliers.

Over-dispersion, under-dispersion and zero-inflation are the most common considerations for GLMMs. Over/under-dispersion refers to the situation when the residual variance is larger/smaller than expected under the fitted model. Over/under-dispersion can be found for any statistical distribution with fixed variance. However, over-dispersion is more common than under-dispersion. Specifically, over-dispersion is very common for count data such that it underestimates variability and standard errors and thus reduces p-values. A mis-specified model is a common reason for over-dispersion. Over-dispersion can result from missing predictors or a mis-specified model. The DHARMa package performs the over-dispersion test by comparing the dispersion (i.e. variability or variance) of simulated residuals to the observed residuals. I used the "testDispersion" function of the package to over-dispersion test (i.e. a non-parametric test on the simulated residuals). This test indicates that there is over-dispersion (the test p-value=2.2e-16 then the test is statistically significant; Figure 3.13).

The zero-inflation test denotes the situation that the number of observed zeros is more than the expected number of zeros under the fitted model. Figure 3.14 shows that the zero-inflation test (p-value =0) is significant. This means that observed zeros are more

than the simulated zeros under the fitted model. The result suggests that a better model for dealing with zeros is needed. Also, using the 'test Generic' function in this package, I tested for the generic excess of values, including 1, 2, 3, 4 and greater than 4. Figure 3.14 also shows that only the test for values greater than 4 is significant.

3.4.2 Temporal and spatial autocorrelation tests

Second, regarding temporal and spatial variables in the model, it is necessary to plot residuals against time and space. I performed a temporal autocorrelation test using the DHARMa package. The package performs a Durbin-Watson test from the package "lmtest" (Zeileis and Hothorn, 2002) on the quantile residuals. Since there were lots of residuals, I used the function "recalculate Residuals" in the DHARMa package to group the residuals by year. I then used the function "testTemporalAutocorrelation" to test for temporal autocorrelation in the residuals and plot the residuals against time (Figure 3.15). The Durbin-Watson test p-value (=0.8565) indicates that there is no temporal autocorrelation in the residuals.

It is difficult to investigate spatial autocorrelation in the residuals with only one test because spatial autocorrelation can be tested on different time scales such as short or long scales. Therefore, I examined the spatial autocorrelation using "Moran.I" test from the package "ape" (Paradis and Schliep, 2019) for each year. Figure 3.16 shows the spatial autocorrelation test p-values for each year. These p-values indicate that spatial autocorrelation in the residuals is significant (p-value < 0.05) for the years 1978-1979, 1981-1982 and 1984-1985. The quantile residuals were transformed to normality (a vector of quantiles with positive and negative values) to plot residuals against space (i.e. longitude and latitude). Figures 3.17-3.27 show the spatial distribution of residuals for every survey set on the map in each year. The spatial distribution of residuals helps detect any correlation between the residuals near each other. For example, in Figure 3.17 (Division 3K in 1979) and Figure 3.19 (Divisions 3K and 3L in 1985), there are

positive correlations among the residuals (i.e. when residuals are followed by neighbour residuals of the same symbol and about the same size). Note that residuals are negatively correlated when neighbour residuals of the different symbol follow residuals.

3.5 Conclusion

In this chapter, I selected a spatiotemporal model (a log linear mixed-effect model) to examine the spatial variation of mean catches during the time series. The model can estimate the MNPT index of population abundance, and importantly can predict the MNPT for unsampled strata and years. In this model, spatial variation is included in two components; the first component is constant across time and the second component changes between years. These two components represent changes in the spatial distribution of witch flounder over the time series. Figures 3.28-3.38 show the spatial distribution of predicted mean catch for the survey data in each year.

The model can increase the length of MNPT time series by estimation of mean catches for areas that were not sampled in the witch flounder survey using catch information in neighbouring (in both space and time) areas. For example, the model can predict MNPT for the first six years of time series (1977-1982) in Division 3L (as shown in the middle panel of Figure 3.6). Therefore, the model enables us to fill missing areas (or missing strata, Figures 3.39-3.41) in the survey time-series.

The bias of model MNPT, arising from the nonlinear transformation of random effects, was reduced using the bias-correction method implemented in the TMB package. The spatiotemporal model used is appropriate to predict MNPT because model MNPT for sampled-only strata are similar to design-based MNPT that are unbiased (as shown in the top panel of Figures 3.4-3.7). Also, the impact of the model interpolation for unsampled strata is slight (as shown in the bottom panel of Figures 3.4-3.7).

3.6 Tables

Table 3.1: List of all indices and data that used in model description and during parameter estimation.

Name	**Symbol**	**Type**
Observation	i	Index
Year	t	Index
Stratum	s	Index
Catch (in number)	Y	Data
Mean catch for observation i	μ_i	Derived
Year for observation i	$t(i)$	Data
Stratum for observation i	$s(i)$	Data
Offset for observation i	offset(i)	Derived

Table 3.2: List of all random effects (i.e. random parameters) incorporated in the models.

Name	**Symbol**	**Process**	**Model**
Stratum (or grid) effect	ω	GMRF	3,4,5
Year effect	ε	AR (1)	2,3,4
Strata-year effect	γ	YEAR(AR 1)|STRAT (year effect is independent across strata)	1,2,3
Strata-year effect	δ	YEAR(AR1)*STRAT(GMRF) (year effect is not independent across strata)	4,5
Stratum effect	ω	GMRF_D (based on the distance between strata)	6
Strata-year effect	δ	YEAR(AR1)*STRAT(GMRF_D) (year effect is not independent across strata)	6

Table 3.3: List of all fixed parameters of the models.

Parameter name	**Symbol**	**Model**
Scale parameter of NB distribution	k	1,2,3,4,5,6
The parameter of the main stratum effect (spatial variation among strata, fixed effect)	μ_s	1,2
The parameter of the main strata effect (average spatial variation among strata, fixed effect	m	3,4,5,6
The parameters to create precision matrix for grid effect	q_ω, d_ω	3,4,5
The parameters to create precision matrix for annual grid effect	q_δ, d_δ	4,5
The correlation coefficient of strata annual effect	ρ_δ	4,5
The standard deviation of strata annual effect	σ_γ	1,2,3
The correlation coefficient of strata annual effect	ρ_γ	1,2,3
The standard deviation of the year effect	σ_ε	2,3,4
The correlation coefficient of the year effect	ρ_ε	2,3,4
The standard deviation of the grid effect	σ_{ω_d}	6
The parameter to create the precision matrix for the grid effect	q_{ω_d}	6
The standard deviation of strata annual effect	σ_{δ_d}	6
The correlation coefficient of strata annual effect	ρ_{δ_d}	6
The parameter to create precision matrix for annual grid effect	q_{δ_d}	6
The parameter to approximate precision proportion among strata	τ	6

Table 3.4: The comparison of the six models.

No. Model	**Model (k constant)**	**No. of Fixed Effects**	**No. of Random Effects**	**NLL**	**AIC**	**BIC**	**ΔAIC**	**ΔBIC**
1	STRAT (fixed) + YEAR($AR1$)| STRAT	112	4469	19458.99	39141.98	39968.18	548.43	1322.81
2	STRAT (fixed) + YEAR (AR1) + YEAR(AR1)| STRAT	114	4510	19401.06	39030.11	39871.06	436.56	1225.69
3	INTERCEPT + STRAT(GMRF) + YEAR(AR1) + YEAR(AR1)| STRAT	8	4619	19513.12	39042.23	39101.25	448.68	455.88
4	INTERCEPT + STRAT(GMRF) + YEAR(AR1) + YEAR(AR1)*STRAT(GMRF)	9	4619	19289.86	38597.73	38664.12	4.18	18.75
5	INTERCEPT + STRAT(GMRF) + YEAR(AR1)*STRAT(GMRF)	7	4578	19289.86	38593.73	38645.37	0.18	0
6	INTERCEPT+ STRAT(GMRF_D) + YEAR(AR1)*STRAT(GMRF_D)	8	4578	19288.78	38593.55	38652.57	0	7.2

$$\Delta AIC = AIC - minAIC,\ \Delta BIC = BIC - minBIC$$

3.7 Figures

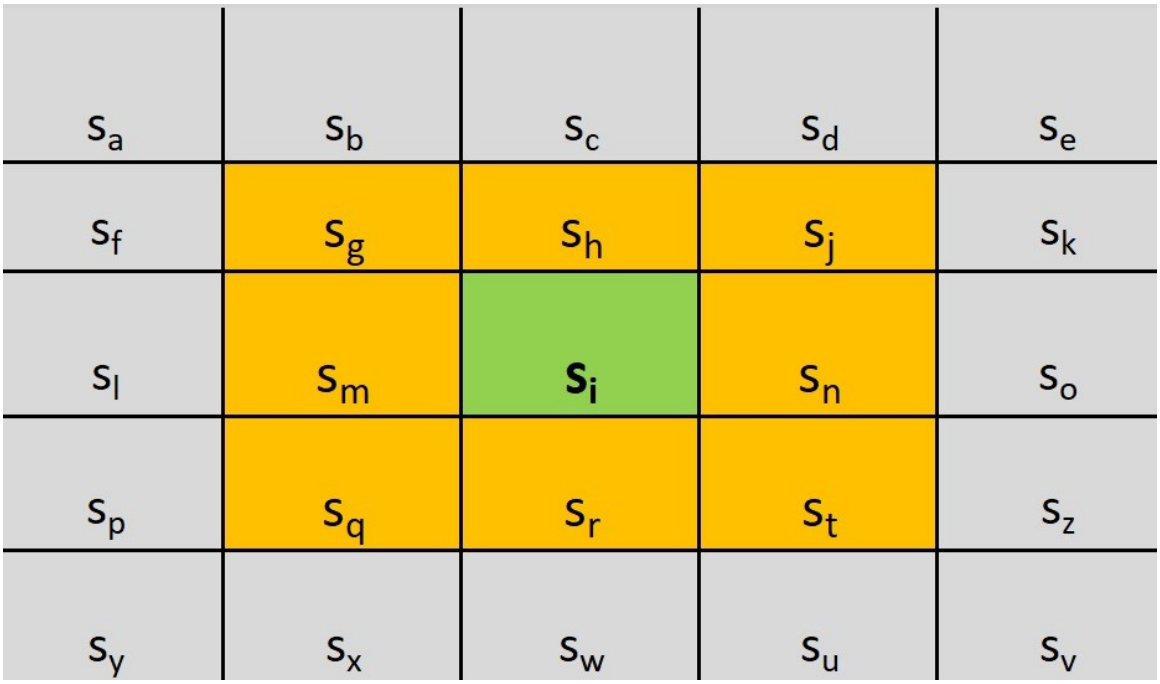

Figure 3.1: A random variable s in the site i (green cell) depends only on a set of its neighbours (yellow cells), $\Delta_i = \{s_g, s_h, s_j, s_m, s_n, s_q, s_r, s_t\}$, not on the rest of the s's (i.e. variables in grey cells).

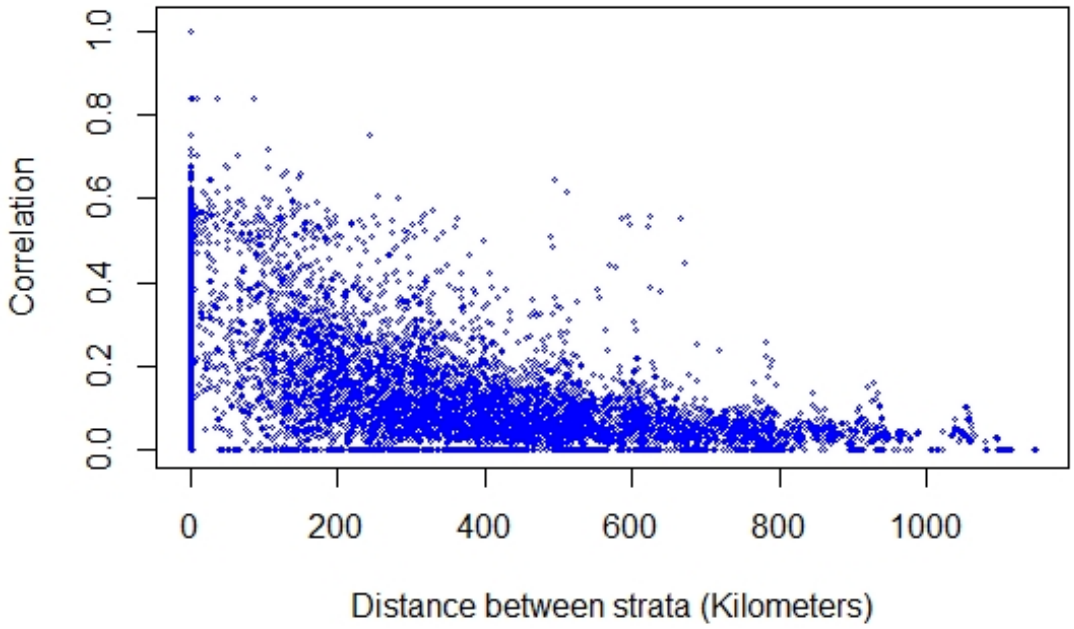

Figure 3.2: The plot shows the correlation between neighbour strata.

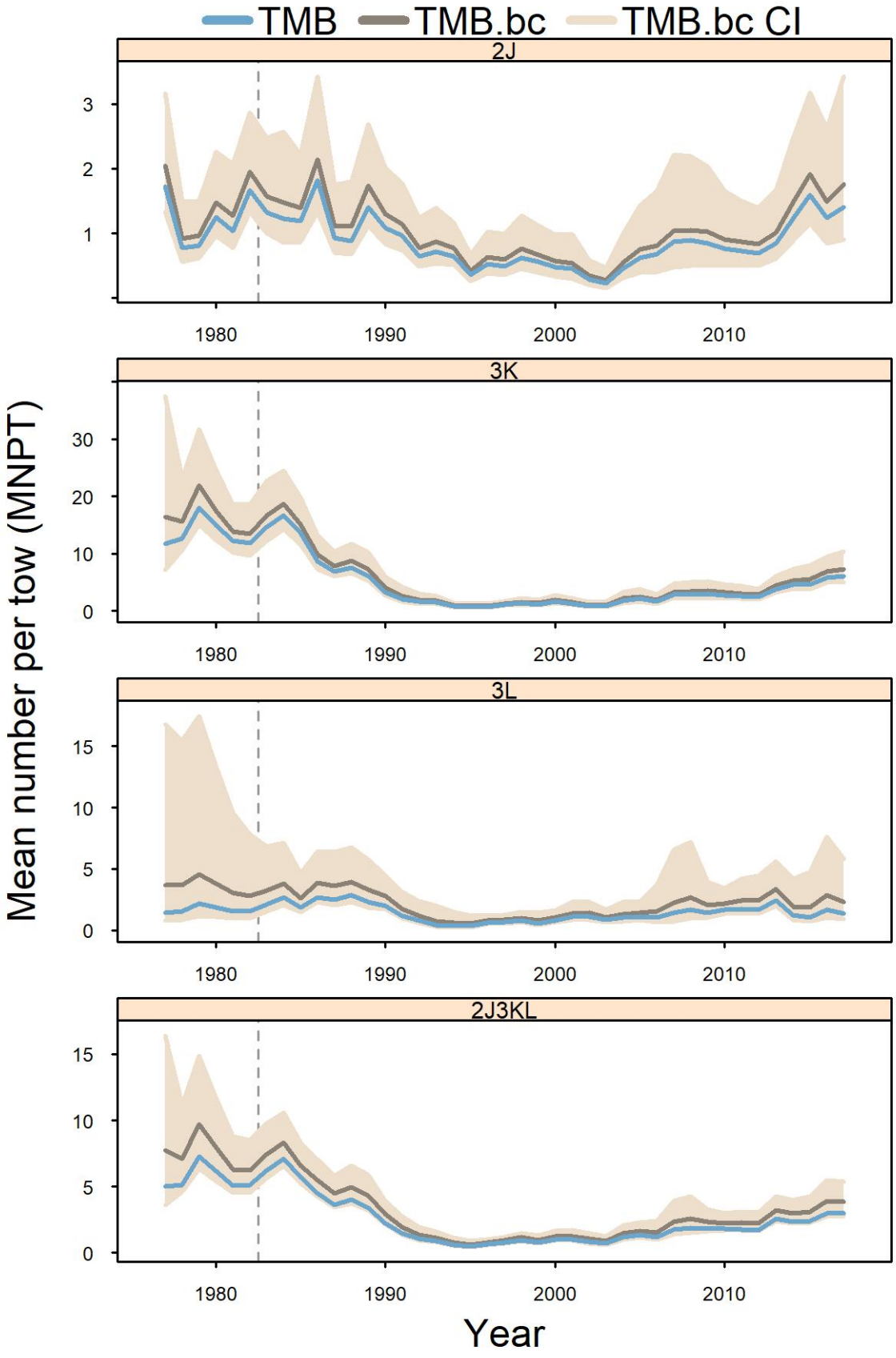

Figure 3.3: The model MNPT (TMB, blue line) and the bias-corrected model MNPT (TMB.bc, dark grey line) and its 95% intervals (TMB.bc CI, shaded cream areas) compared by each division and the whole stock area (2J3KL) for witch flounder data from annual autumn DFO RV survey. The vertical dashed line shows 1983, the year that surveys began in NAFO Division 3L.

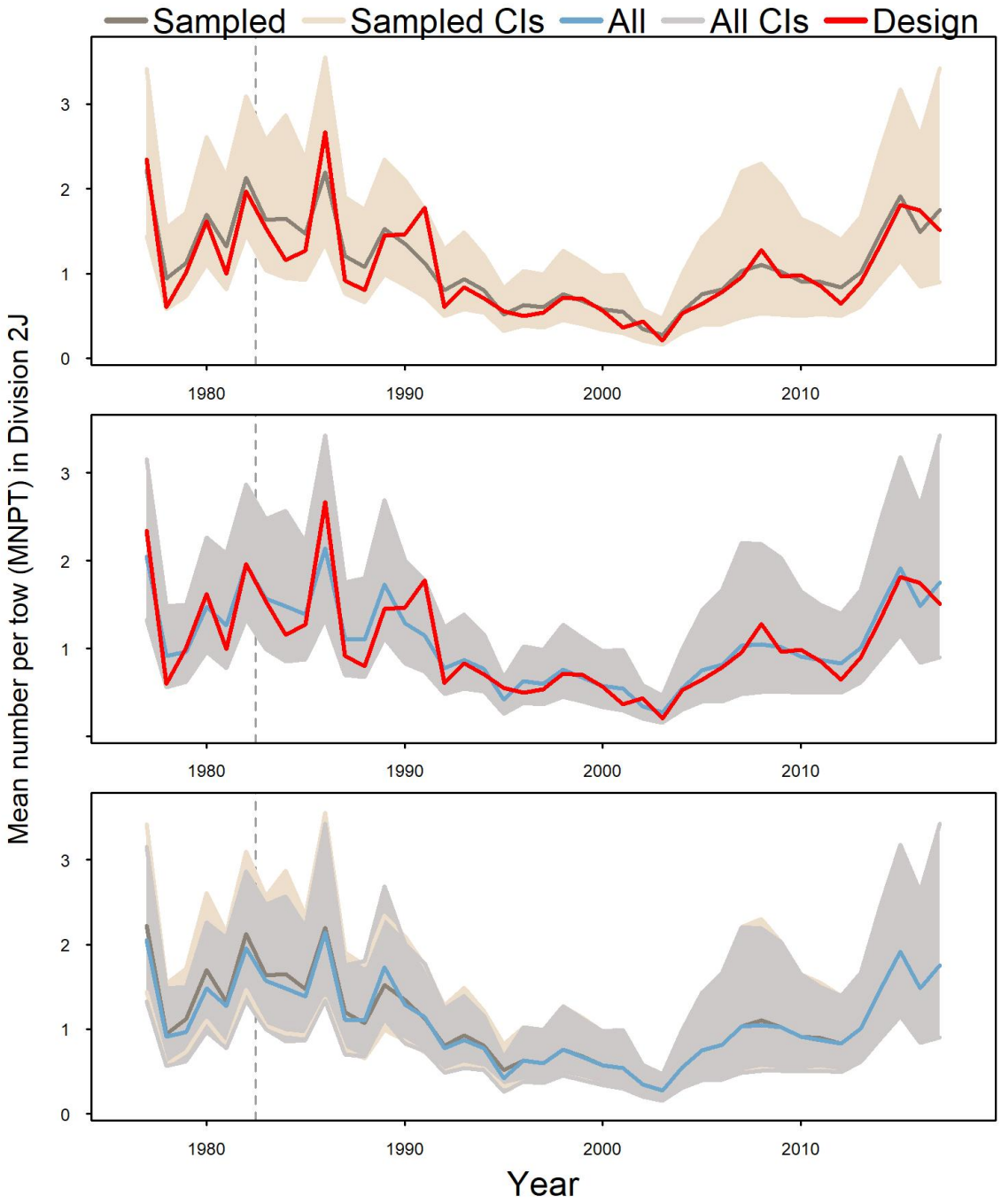

Figure 3.4: A comparison of estimated MNPTs for witch flounder survey data in NAFO Division 2J. The top panel compares the bias-corrected model MNPT (dark grey line) based only on sampled strata and its 95% intervals (shaded cream areas) with the design-based MNPT (red line). In the middle panel, the design-based MNPT (red line) is compared with the bias-corrected model MNPT (blue line) based on all strata and 95% intervals for model MNPT (shaded grey areas). The comparison between the bias-corrected model MNPT based only on sampled strata (dark grey line) and the bias-corrected model MNPT based on all strata (blue line) associated with their 95% intervals (shaded cream and grey areas, where light grey is their overlap) is shown in the bottom panel. The vertical dashed line shows 1983, the year that surveys began in NAFO Division 3L.

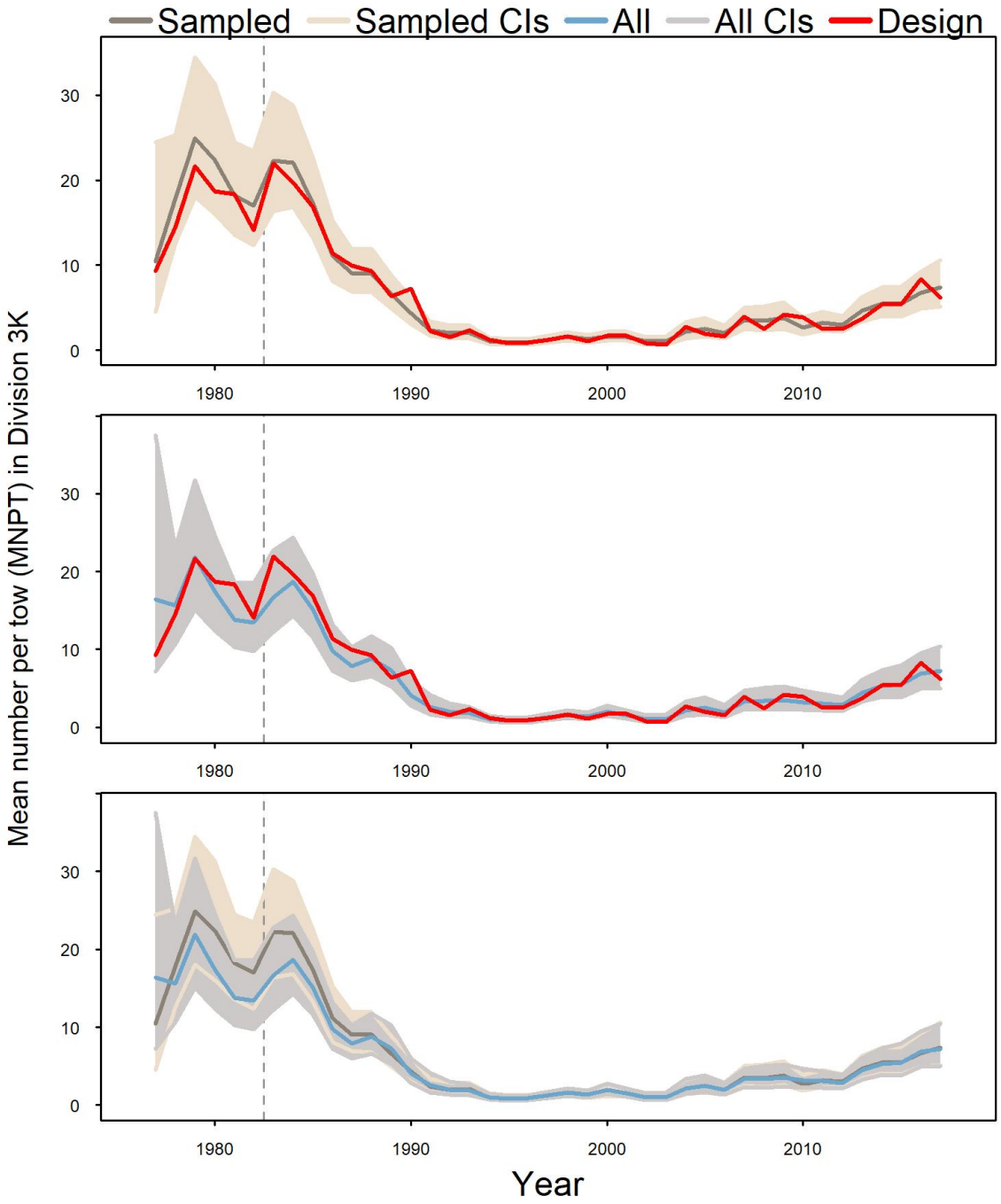

Figure 3.5: A comparison of estimated MNPTs for witch flounder survey data in NAFO Division 3K. The top panel compares the bias-corrected model MNPT (dark grey line) based only on sampled strata and its 95% intervals (shaded cream areas) with the design-based MNPT (red line). In the middle panel, the design-based MNPT (red line) is compared with the bias-corrected model MNPT (blue line) based on all strata and 95% intervals for model MNPT (shaded grey areas). The comparison between the bias-corrected model MNPT based only on sampled strata (dark grey line) and the bias-corrected model MNPT based on all strata (blue line) associated with their 95% intervals (shaded cream and grey areas, where light grey is their overlap) is shown in the bottom panel. The vertical dashed line shows 1983, the year that surveys began in NAFO Division 3L.

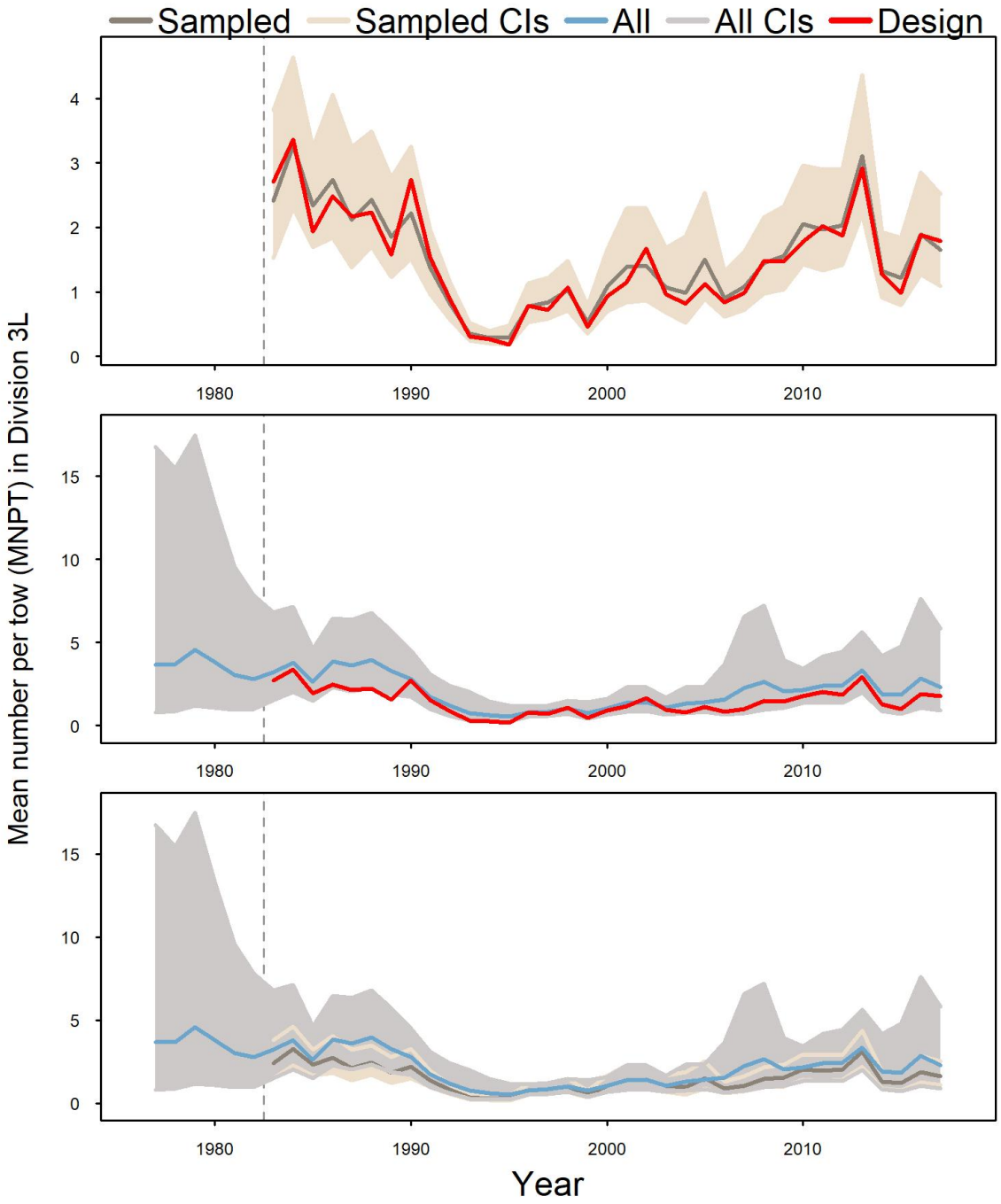

Figure 3.6: A comparison of estimated MNPTs for witch flounder survey data in MAFO Division 3L. The top panel compares the bias-corrected model MNPT (dark grey line) based only on sampled strata and its 95% intervals (shaded cream areas) with the design-based MNPT (red line). In the middle panel, the design-based MNPT (red line) is compared with the bias-corrected model MNPT (blue line) based on all strata and 95% intervals for model MNPT (shaded grey areas). The comparison between the bias-corrected model MNPT based only on sampled strata (dark grey line) and the bias-corrected model MNPT based on all strata (blue line) associated with their 95% intervals (shaded cream and grey areas, where light grey is their overlap) is shown in the bottom panel. The vertical dashed line shows 1983, the year that surveys began in NAFO Division 3L.

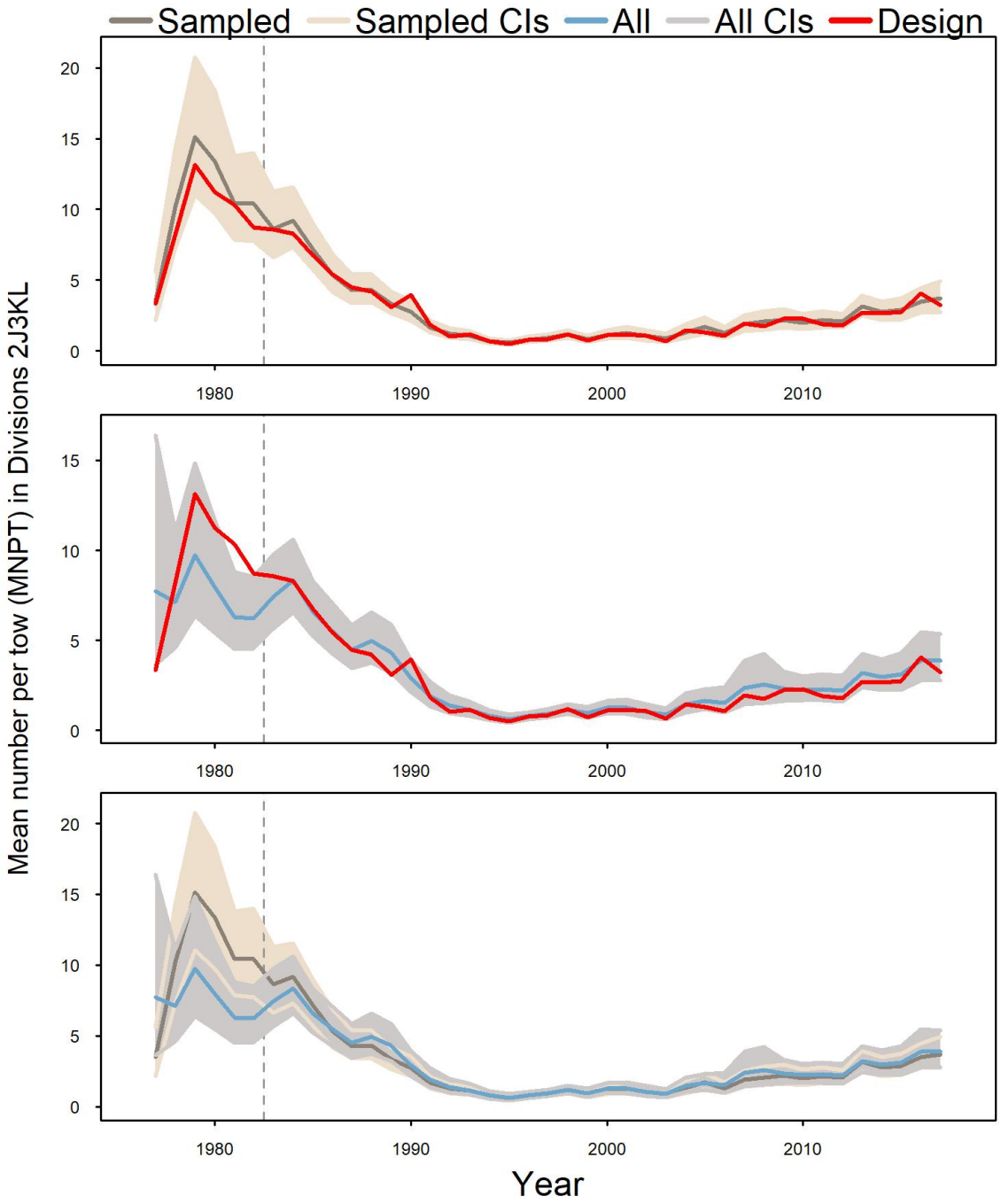

Figure 3.7: A comparison of estimated MNPTs for witch flounder survey data in NAFO Divisions 2J3KL. The top panel compares the bias-corrected model MNPT (dark grey line) based only on sampled strata and its 95% intervals (shaded cream areas) with the design-based MNPT (red line). In the middle panel, the design-based MNPT (red line) is compared with the bias-corrected model MNPT (blue line) based on all strata and 95% intervals for model MNPT (shaded grey areas). The comparison between the bias-corrected model MNPT based only on sampled strata (dark grey line) and the bias-corrected model MNPT based on all strata (blue line) associated with their 95% intervals (shaded cream and grey areas, where light grey is their overlap) is shown in the bottom panel. The vertical dashed line shows 1983, the year that surveys began in NAFO Division 3L.

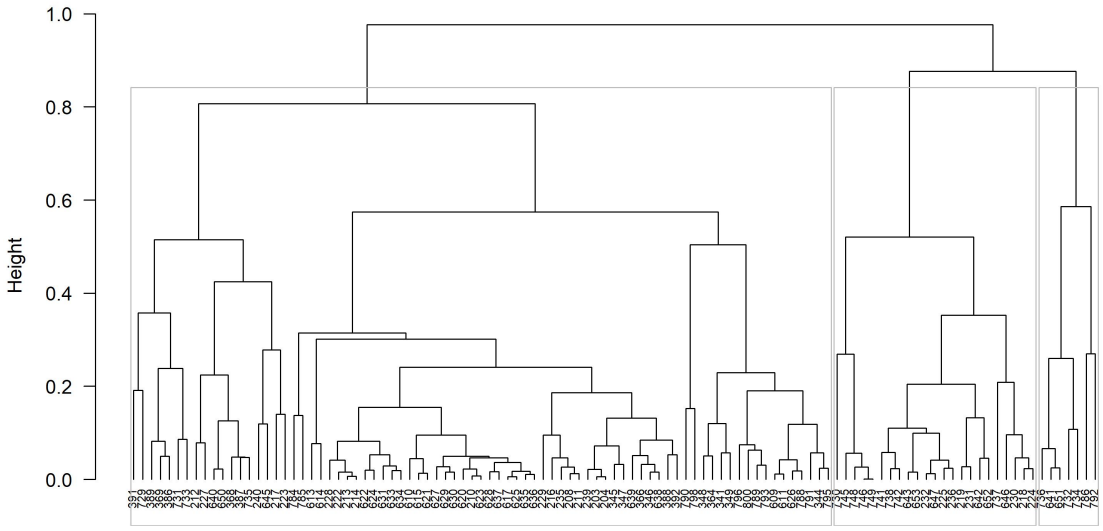

Figure 3.8: The strata dendrogram, cut into three clusters by vertical grey lines.

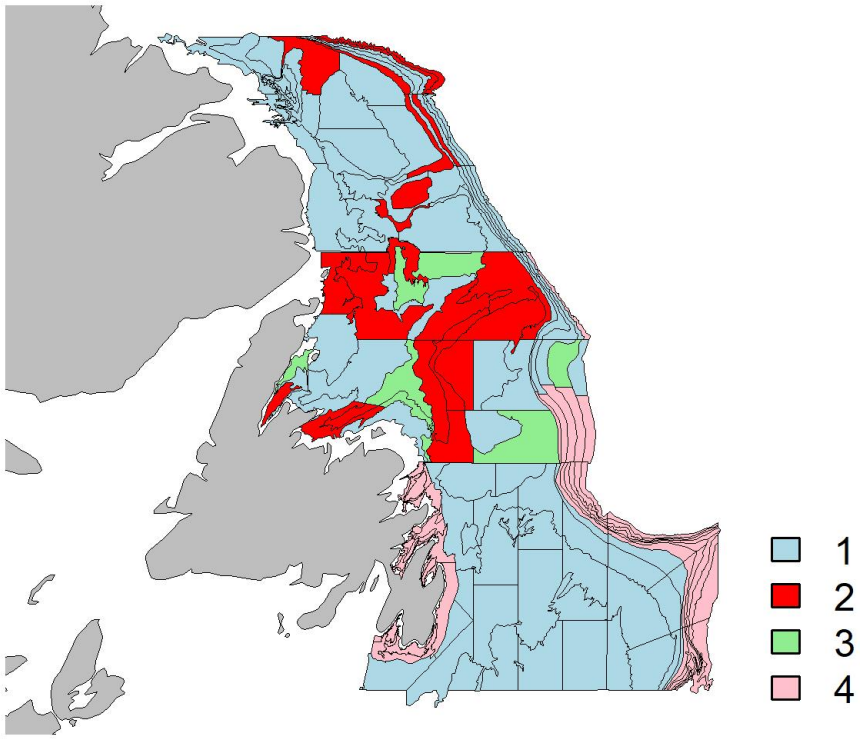

Figure 3.9: The stock area divided into three coloured regions (blue, red and green regions) represents three strata clusters and indicates which strata are more similar to each other. The pink regions represent strata with the mean catch about zero that were not used in the clustering process.

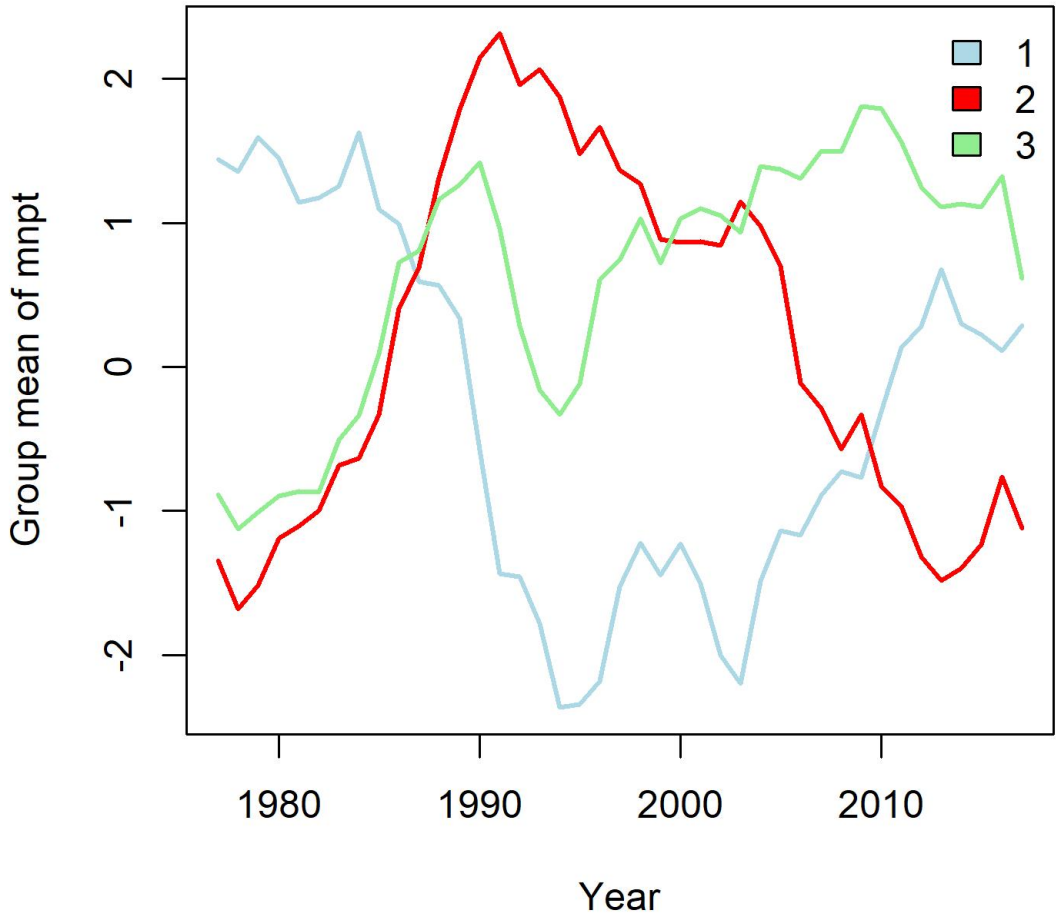

Figure 3.10: The trend of model MNPT (log MNPT) for each strata cluster over the years. Red and blue seem negatively correlated. This can happen because of fish movements between red and blue strata.

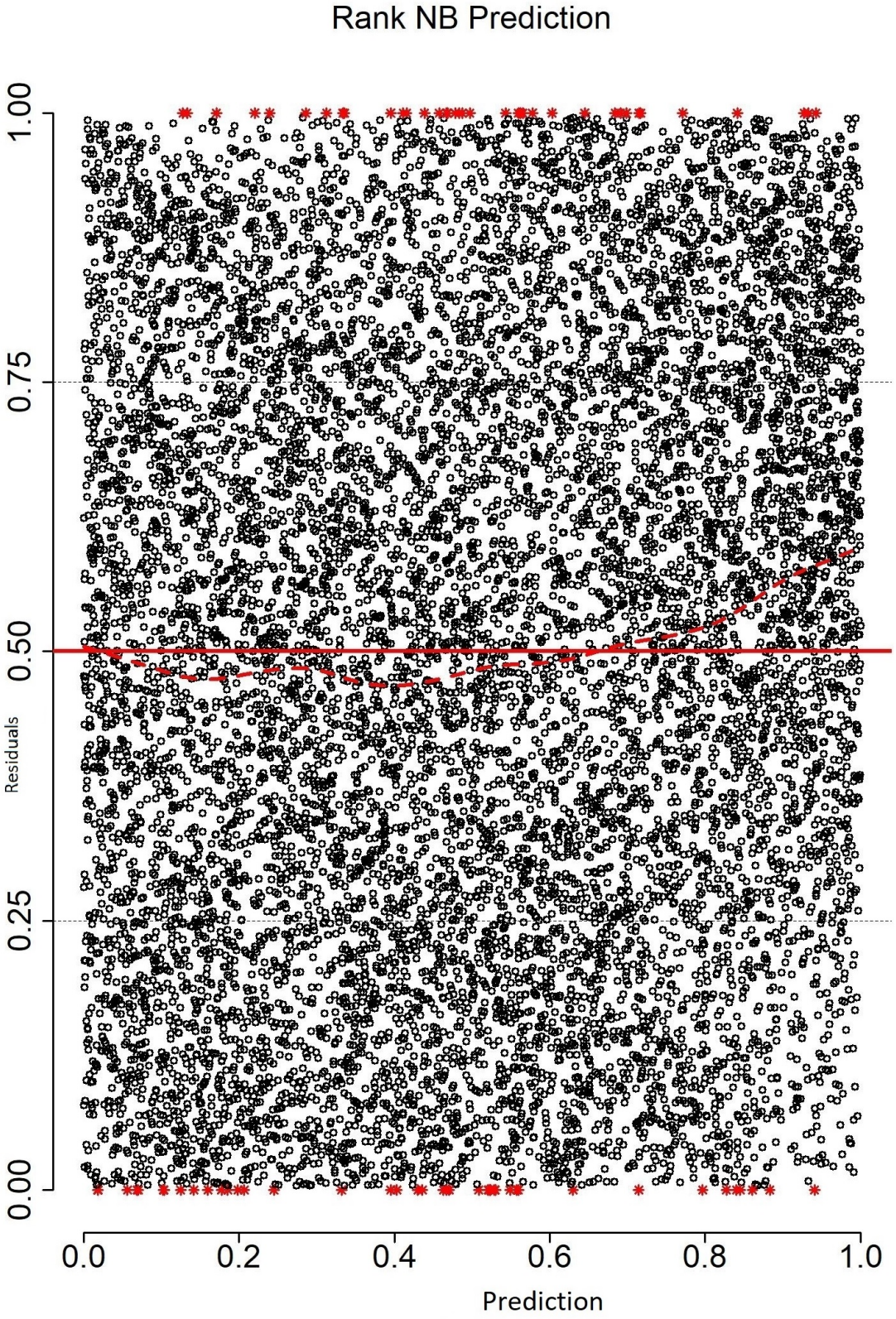

Figure 3.11: A plot of residuals created from the DHARMa package against rank negative binomial predictions. The solid red line shows the 0.5 quantile residuals (i.e. the median of residuals), and the dashed red line represents the 0.5 quantile regression residuals. The red stars are outliers.

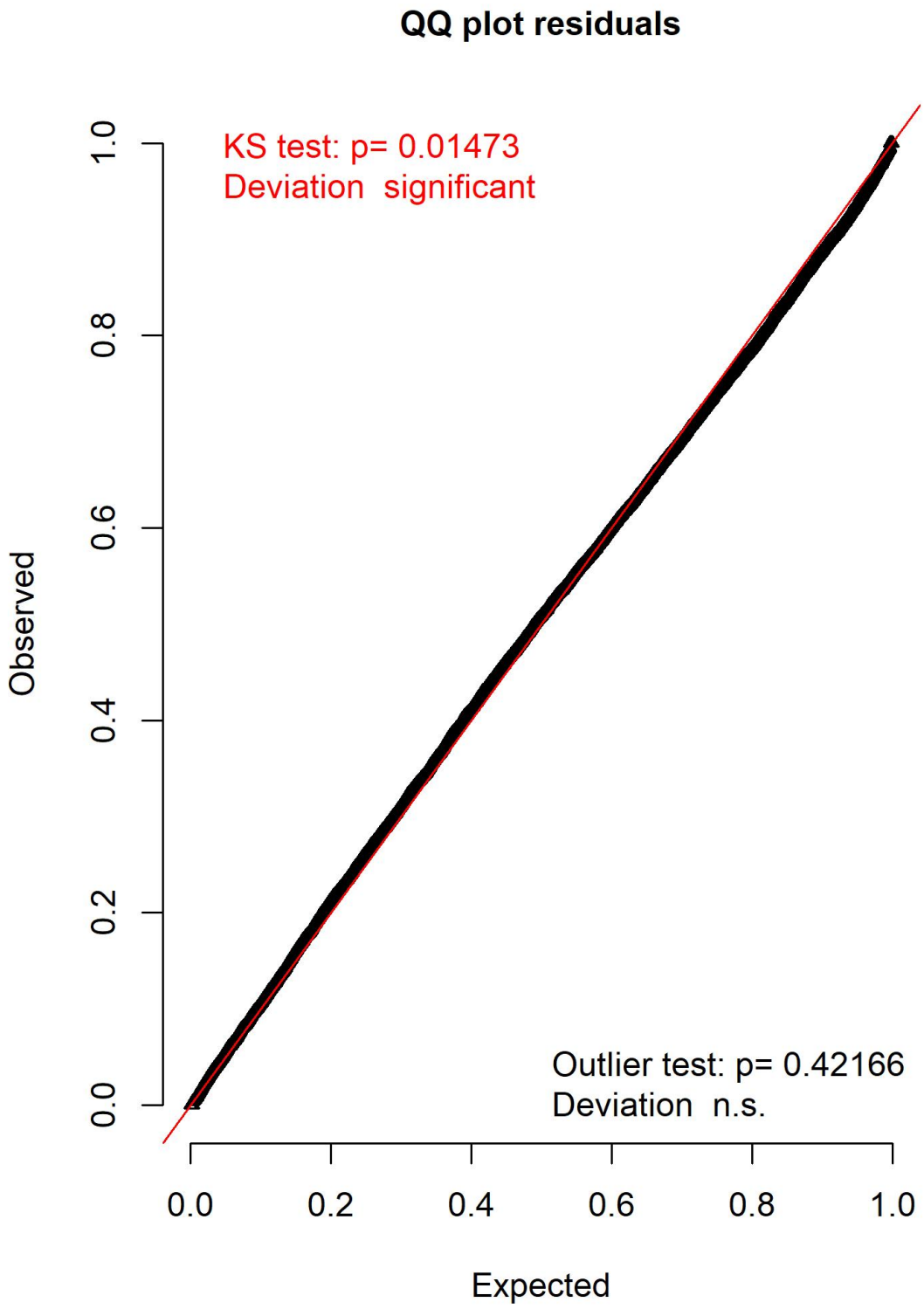

Figure 3.12: QQ-plot shows the uniformity of the quantile residuals such that their trend (bold black line) matches with a straight line (solid red line). The Kolmogorov-Smirnov test indicates significant deviation from the expected distribution. The outlier test p-value indicates there are no outliers.

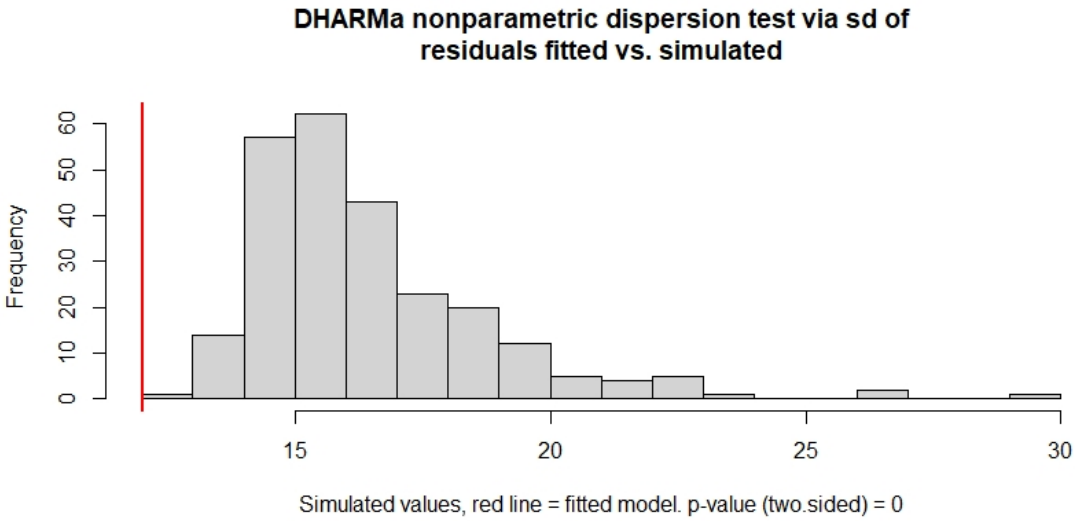

Figure 3.13: Frequency distribution of the standard deviation of residuals simulated (grey bars) in comparison with the standard deviation of fitted model residuals (red line).

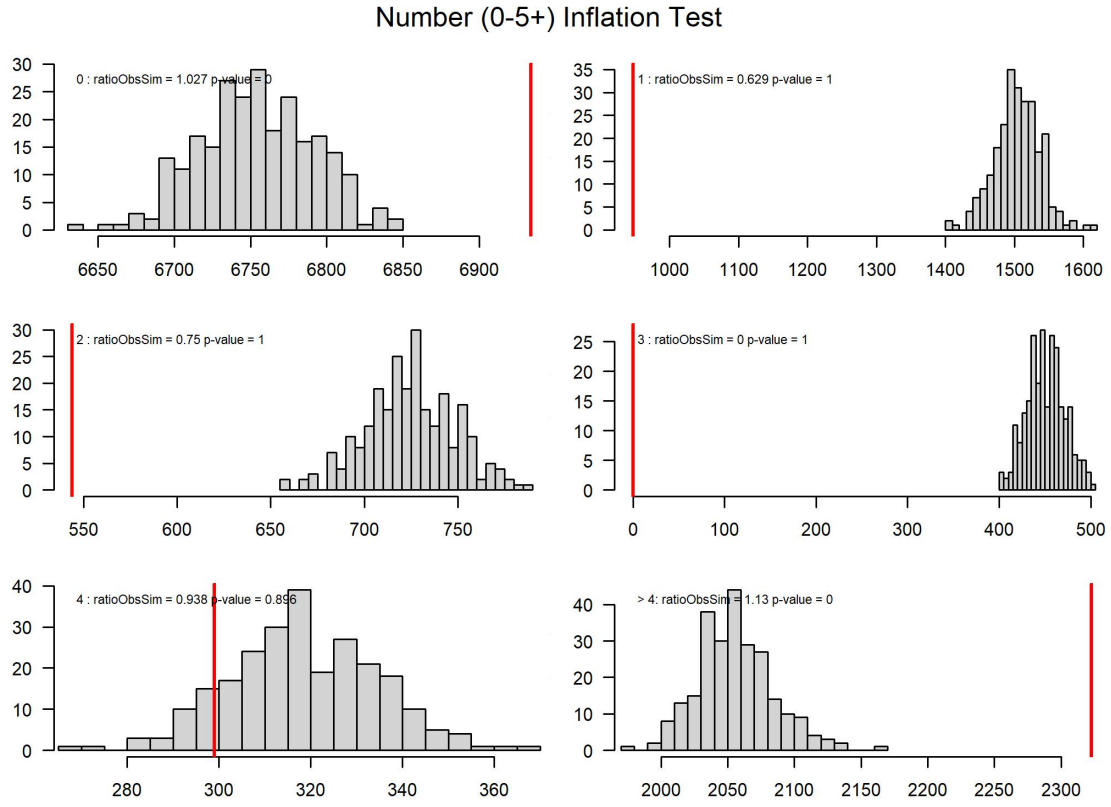

Figure 3.14: The zero-inflation test (the top left panel) compares the observed zeros (in the fitted model, red vertical line) with the simulated distribution of zero's (grey bars), which is significant (p-value =0). Also, tests for the generic excess of values (1, 2, 3, 4, and 4+) that compare the simulated distribution of values (grey bars) with the observed values (in the fitted model, red line). The test p-values indicate that only the test for values greater than 4 is significant.

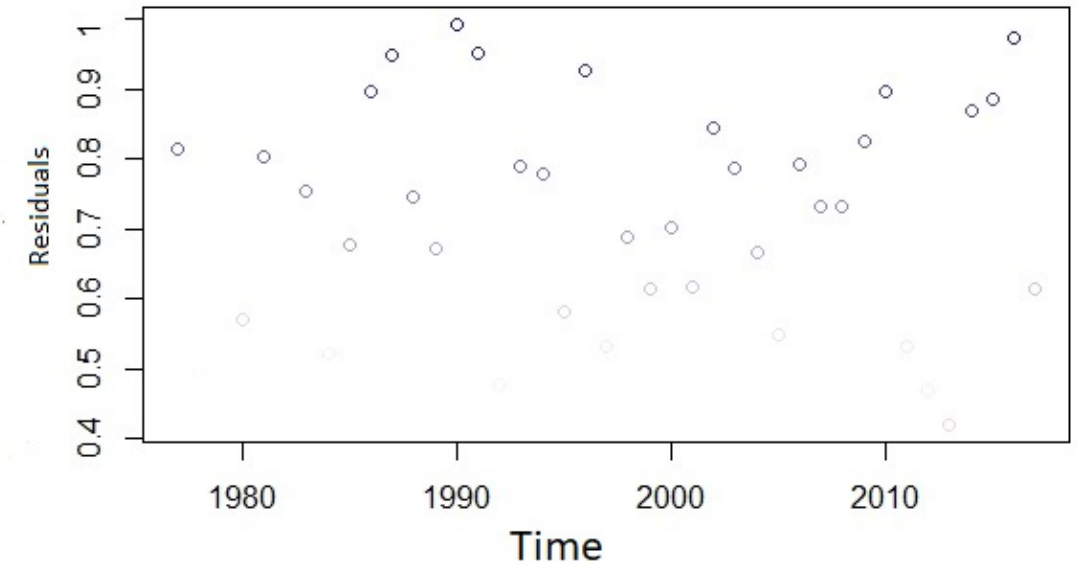

Figure 3.15: The plot of grouped quantile residuals against time (i.e. survey years). The circles represent the mean quantile residuals each year. These circles are bold when their value is close to 1.

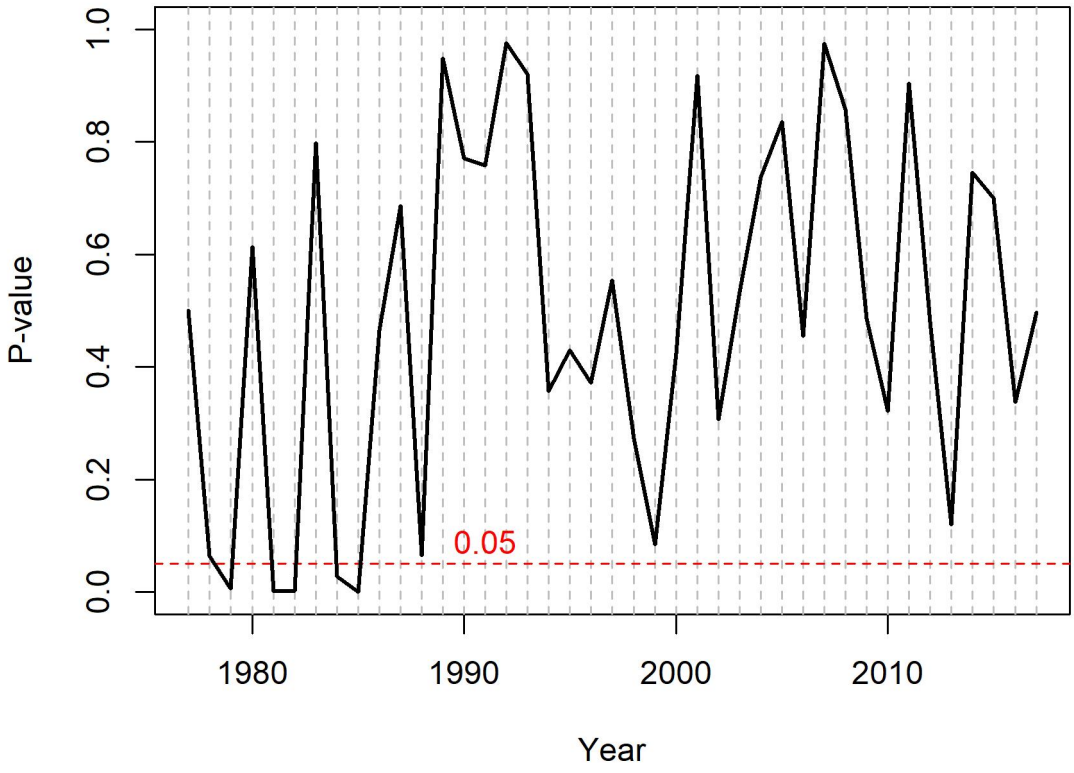

Figure 3.16: The spatial autocorrelation Moran.I's test p-values in each year of the time series. The dashed red line represents the statistical significance level of the test(p-value=0.05).

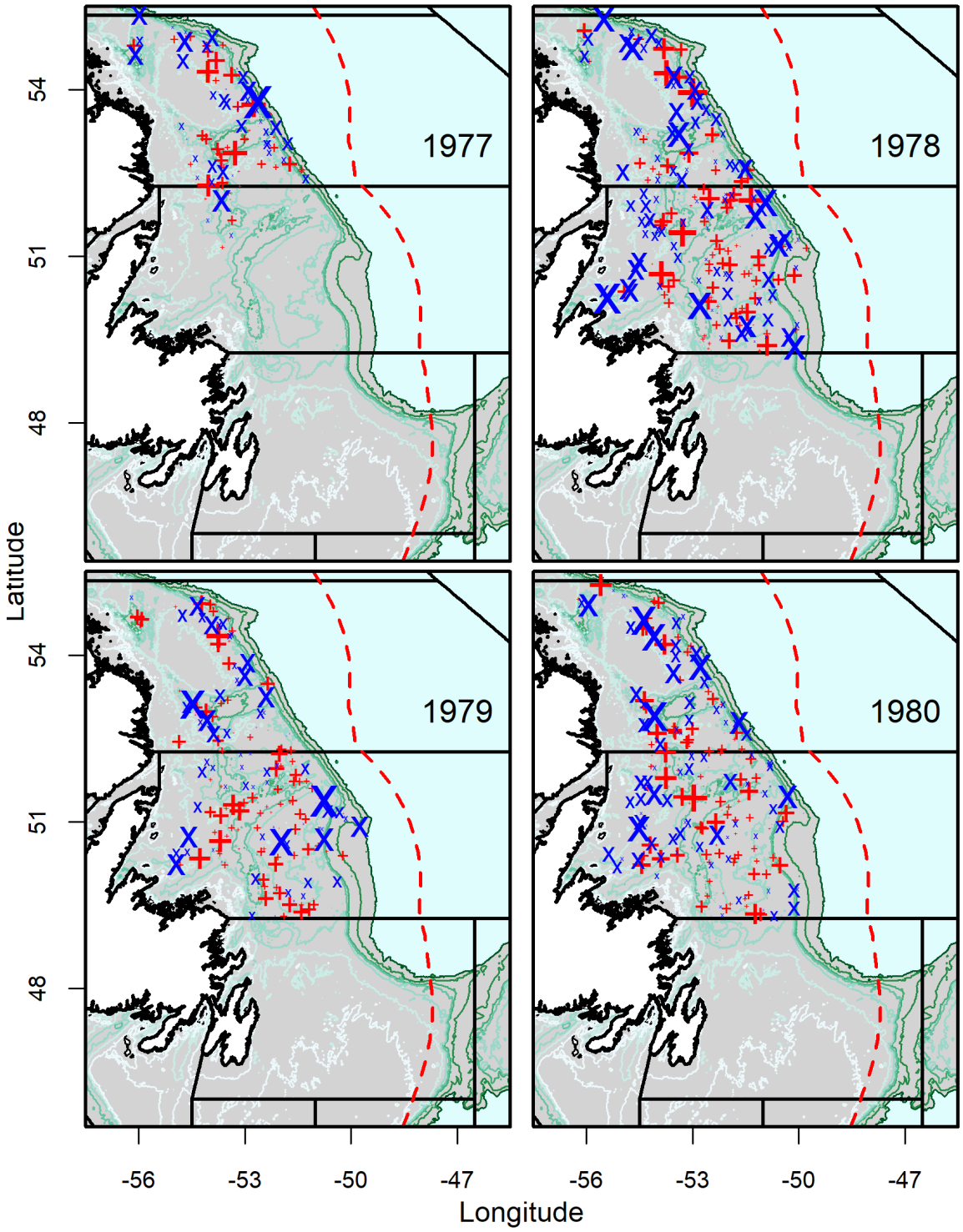

Figure 3.17: The spatial distribution of the simulated residuals transformed to normality in every set of NAFO Divisions 2J3KL from 1977 to 1980. Red +'s and blue ×'s represent positive and negative residuals, respectively, and the size of these symbols indicates the absolute value of the residuals. See Figure 1.5 for other figure details.

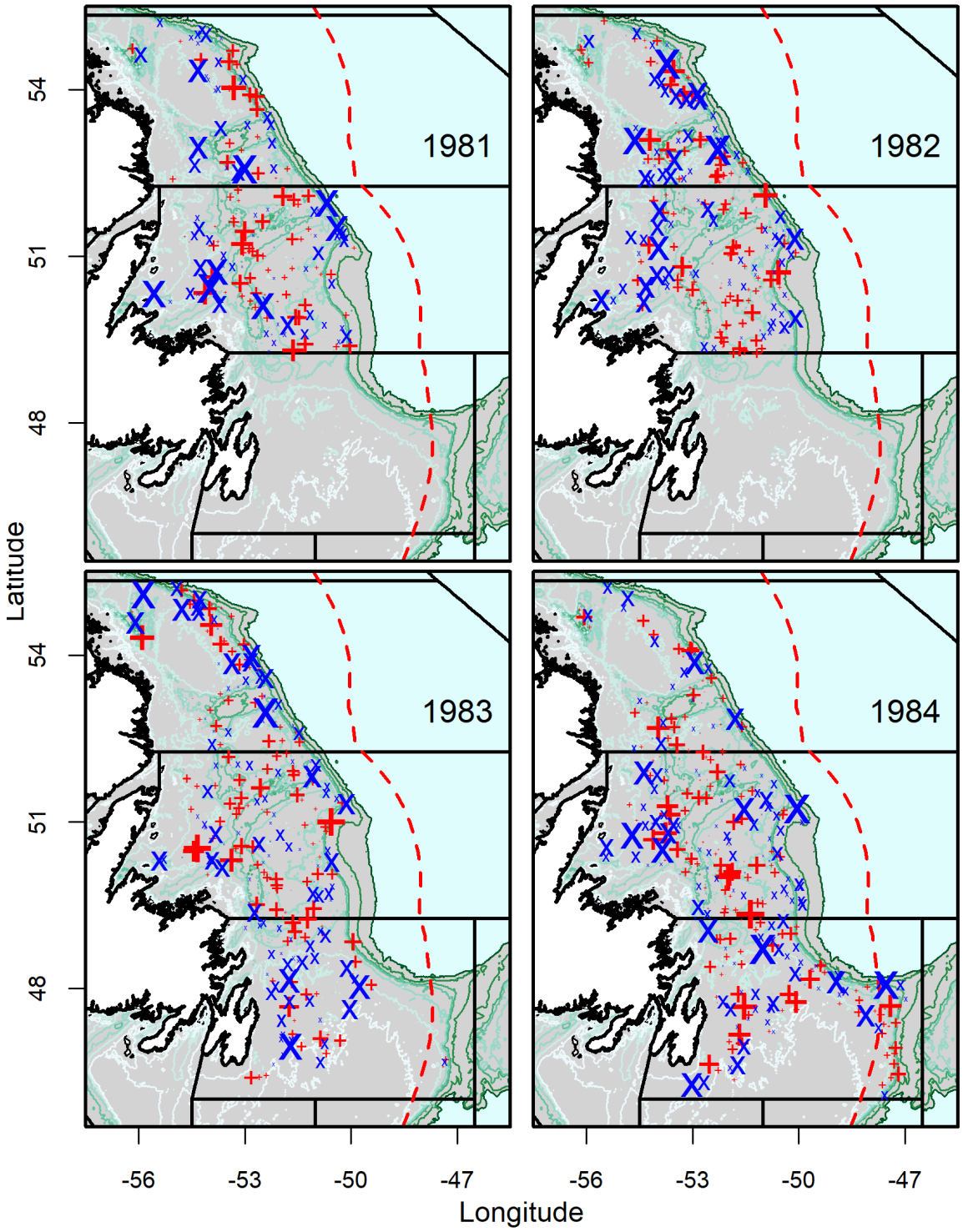

Figure 3.18: The spatial distribution of the simulated residuals transformed to normality in every set of NAFO Divisions 2J3KL from 1981 to 1984. Red +'s and blue ×'s represent positive and negative residuals, respectively, and the size of these symbols indicates the absolute value of the residuals. See Figure 1.5 for other figure details.

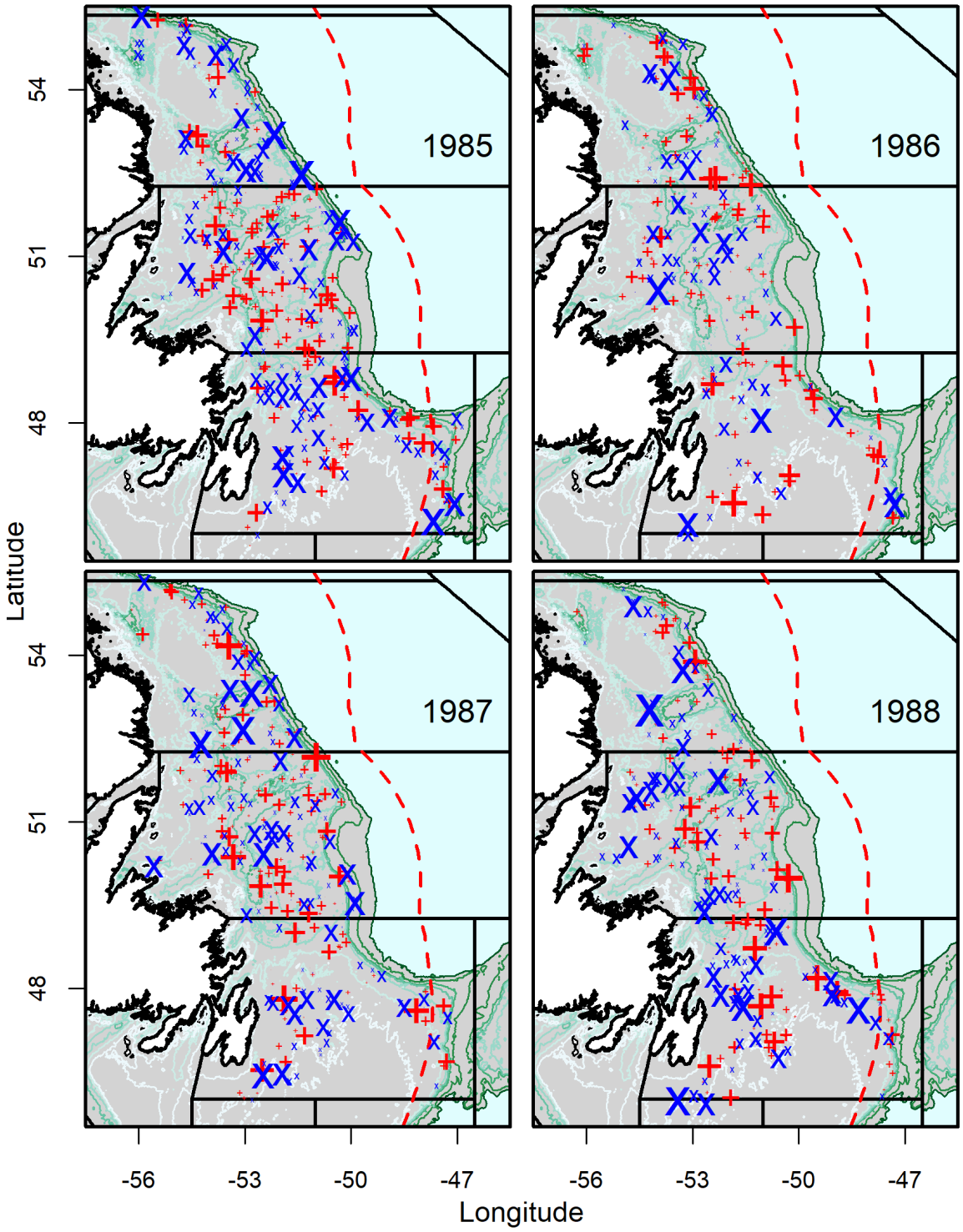

Figure 3.19: The spatial distribution of the simulated residuals transformed to normality in every set of NAFO Divisions 2J3KL from 1985 to 1988. Red +'s and blue ×'s represent positive and negative residuals, respectively, and the size of these symbols indicates the absolute value of the residuals. See Figure 1.5 for other figure details.

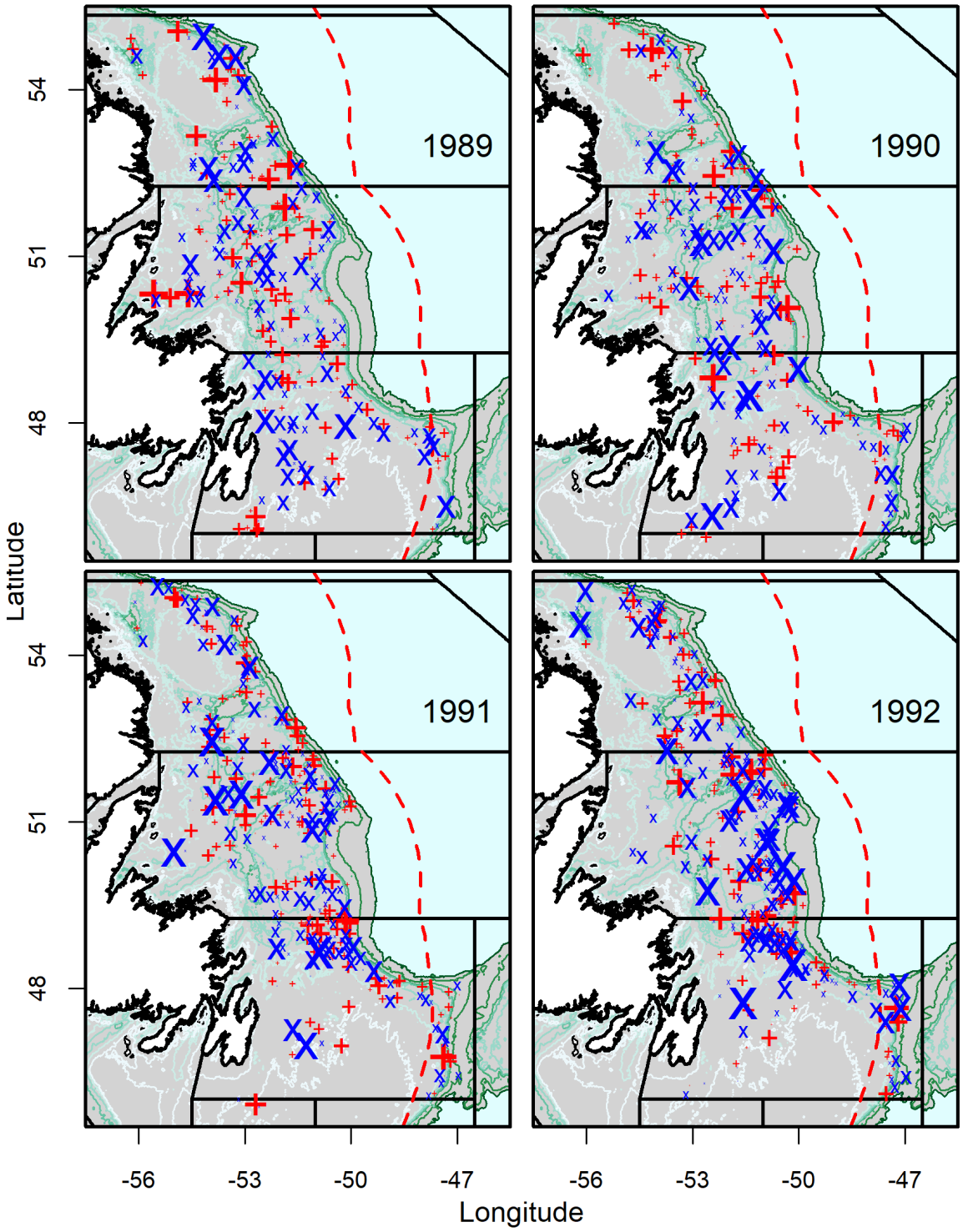

Figure 3.20: The spatial distribution of the simulated residuals transformed to normality in every set of NAFO Divisions 2J3KL from 1989 to 1992. Red +'s and blue ×'s represent positive and negative residuals, respectively, and the size of these symbols indicates the absolute value of the residuals. See Figure 1.5 for other figure details.

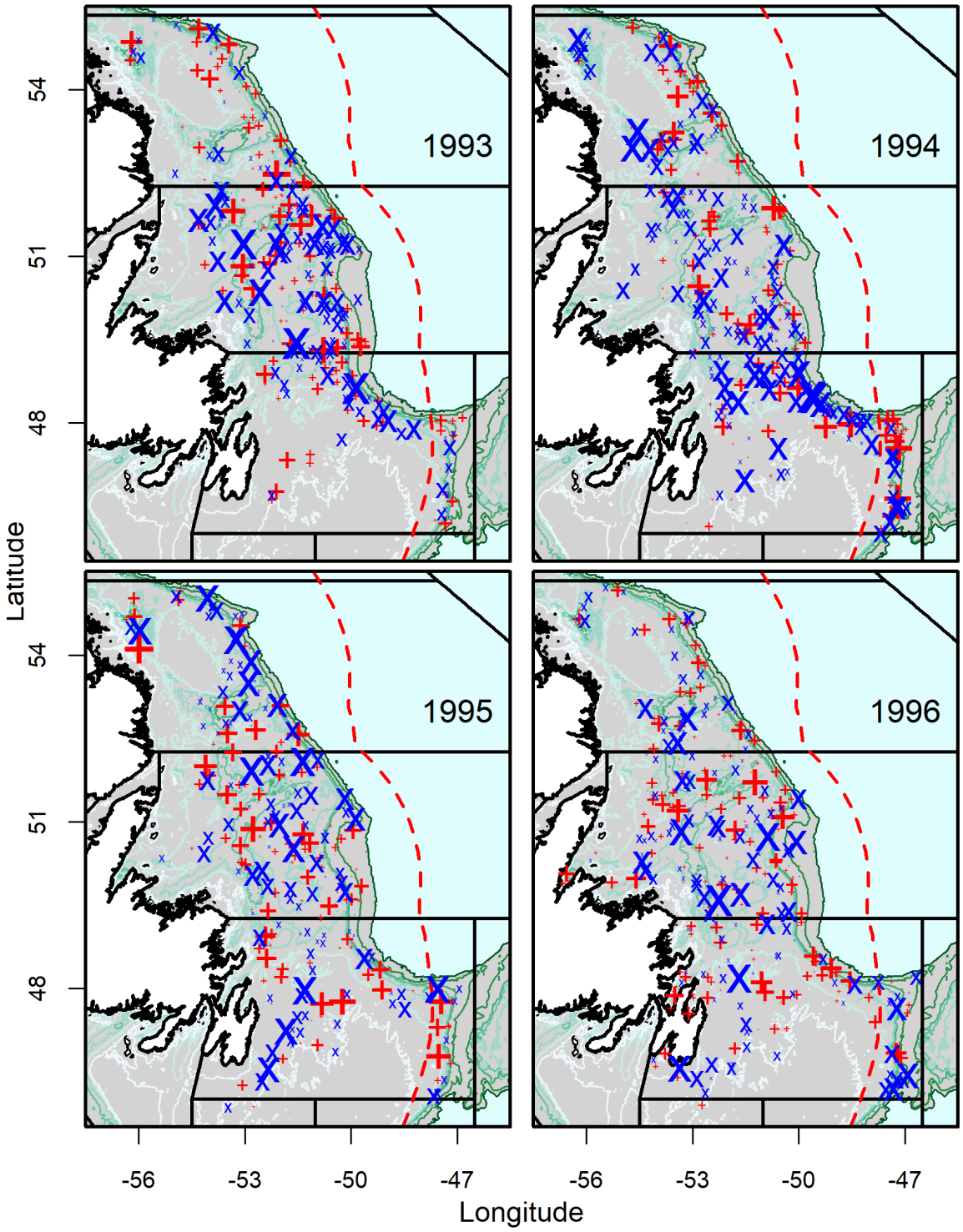

Figure 3.21: The spatial distribution of the simulated residuals transformed to normality in every set of NAFO Divisions 2J3KL from 1993 to 1996. Red +'s and blue ×'s represent positive and negative residuals, respectively, and the size of these symbols indicates the absolute value of the residuals. See Figure 1.5 for other figure details.

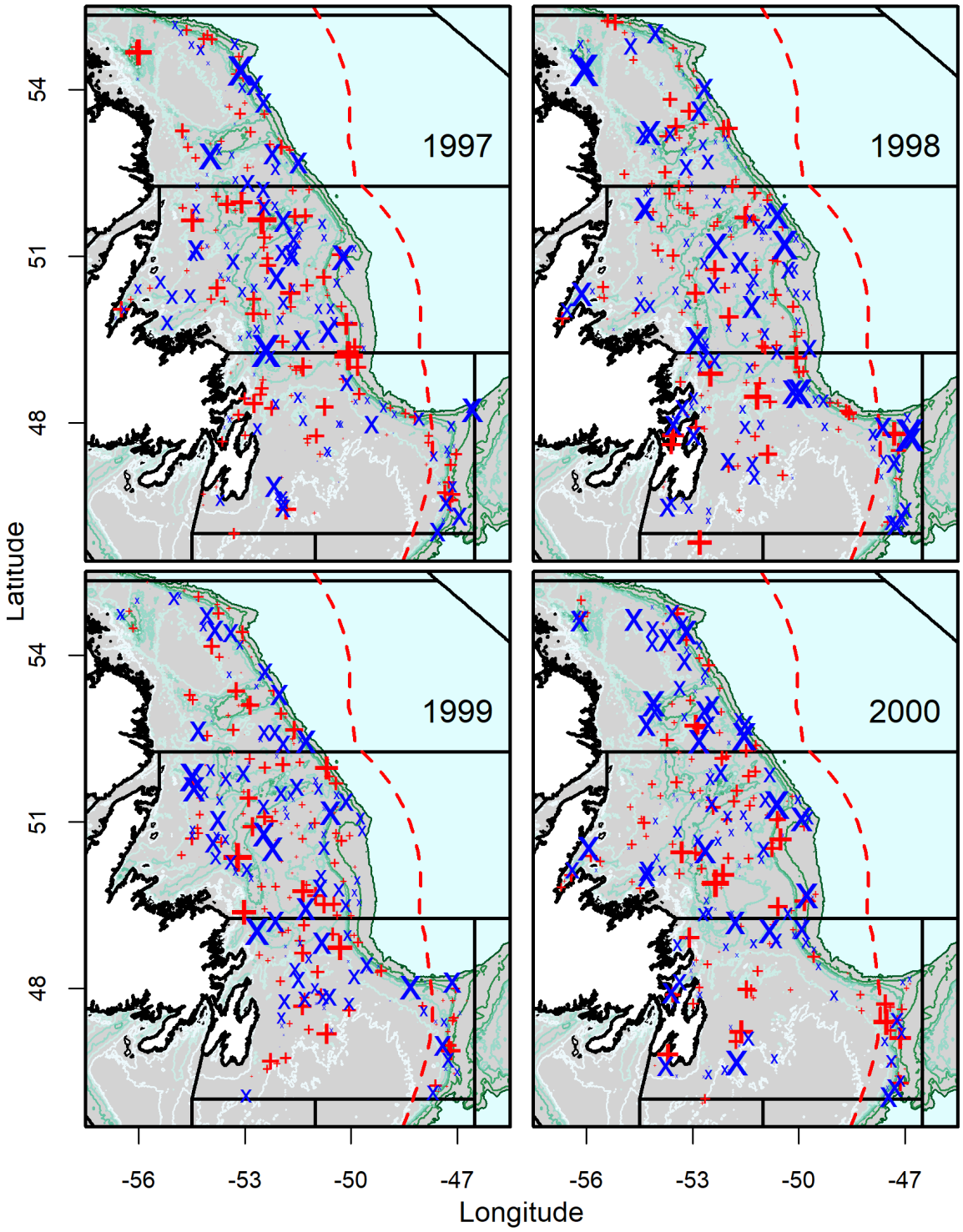

Figure 3.22: The spatial distribution of the simulated residuals transformed to normality in every set of NAFO Divisions 2J3KL from 1997 to 2000. Red +'s and blue ×'s represent positive and negative residuals, respectively, and the size of these symbols indicates the absolute value of the residuals. See Figure 1.5 for other figure details.

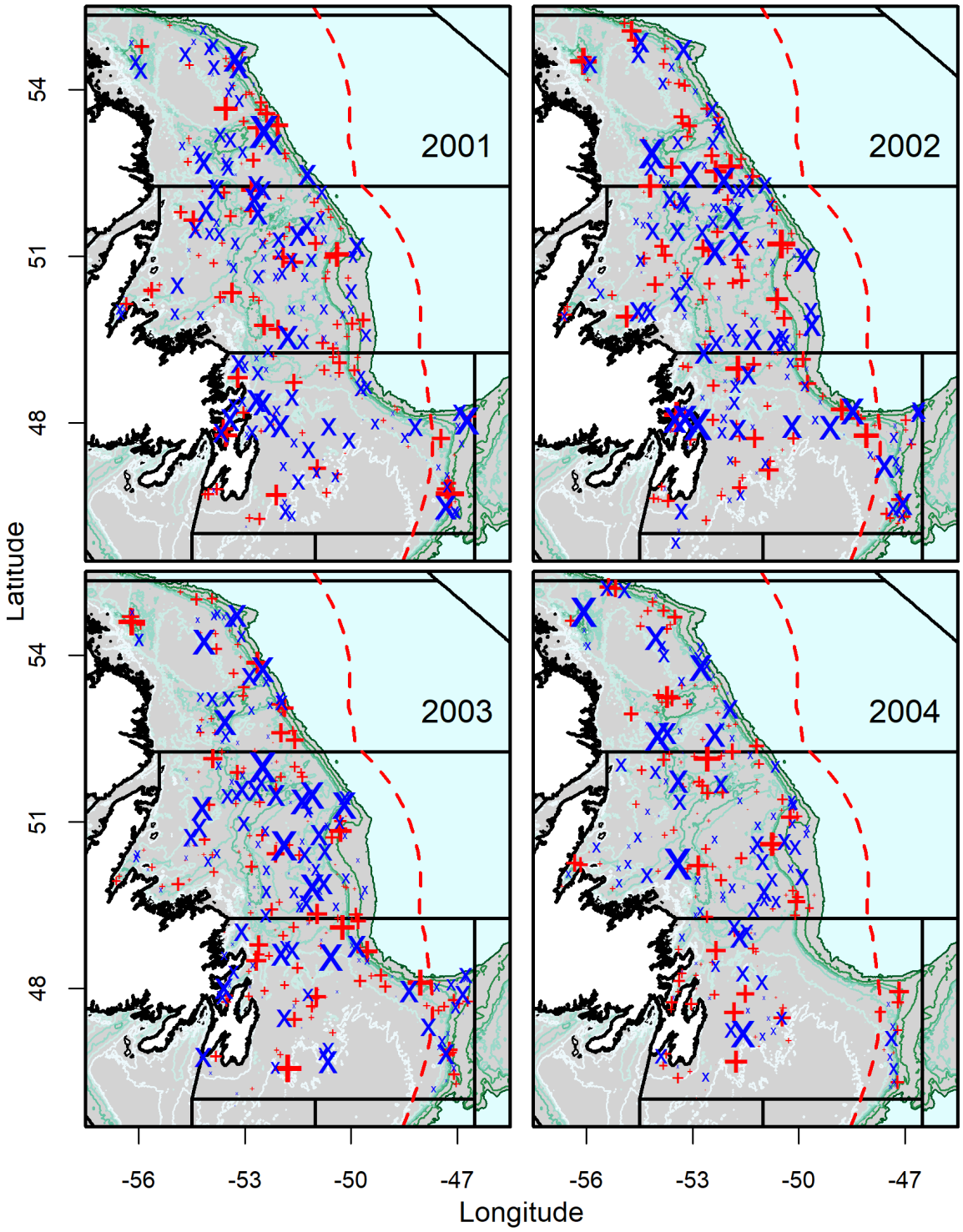

Figure 3.23: The spatial distribution of the simulated residuals transformed to normality in every set of NAFO Divisions 2J3KL from 2001 to 2004. Red +'s and blue ×'s represent positive and negative residuals, respectively, and the size of these symbols indicates the absolute value of the residuals. See Figure 1.5 for other figure details.

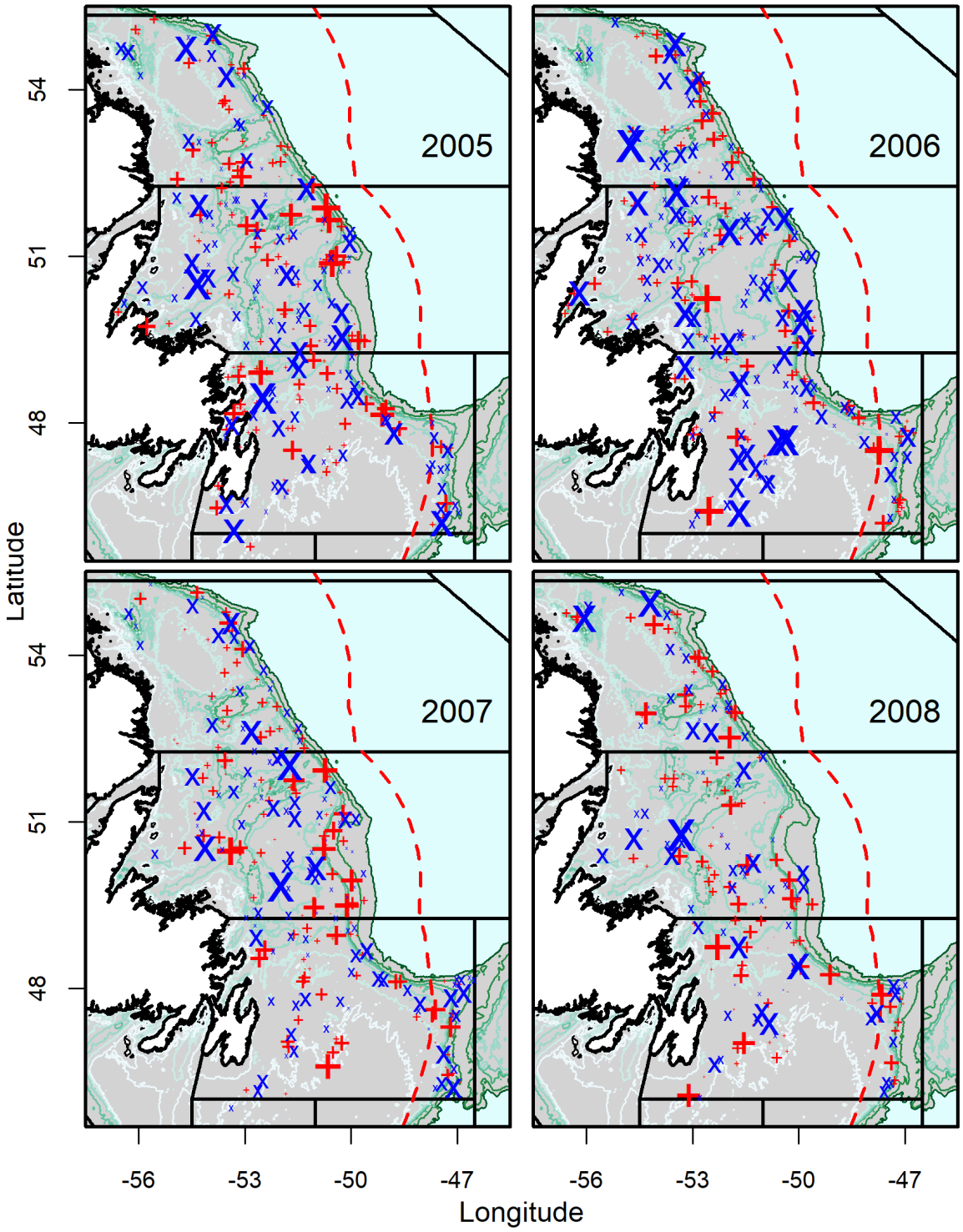

Figure 3.24: The spatial distribution of the simulated residuals transformed to normality in every set of NAFO Divisions 2J3KL from 2005 to 2008. Red +'s and blue ×'s represent positive and negative residuals, respectively, and the size of these symbols indicates the absolute value of the residuals. See Figure 1.5 for other figure details.

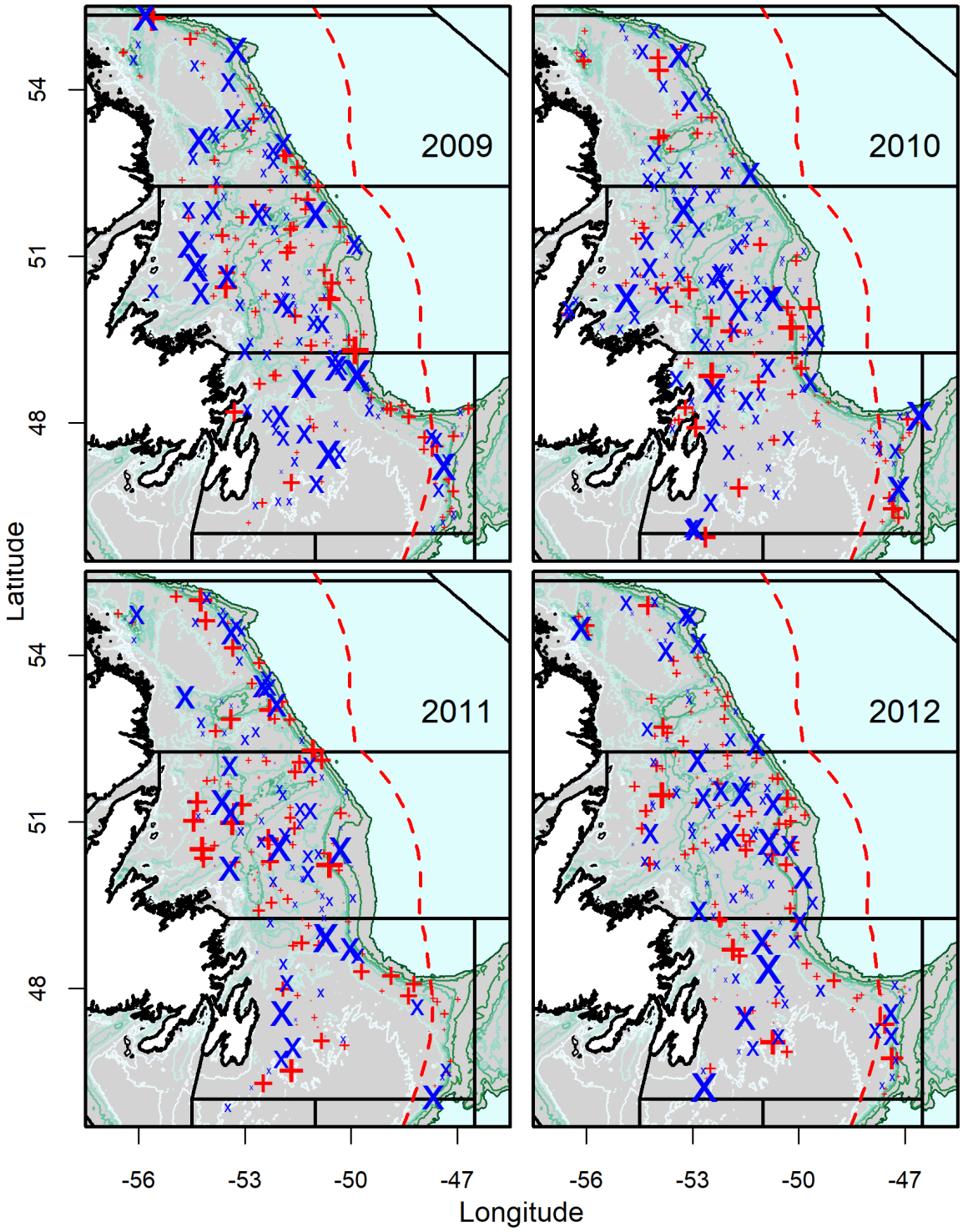

Figure 3.25: The spatial distribution of the simulated residuals transformed to normality in every set of NAFO Divisions 2J3KL from 2009 to 2012. Red +'s and blue ×'s represent positive and negative residuals, respectively, and the size of these symbols indicates the absolute value of the residuals. See the Figure 1.5 for other figure details.

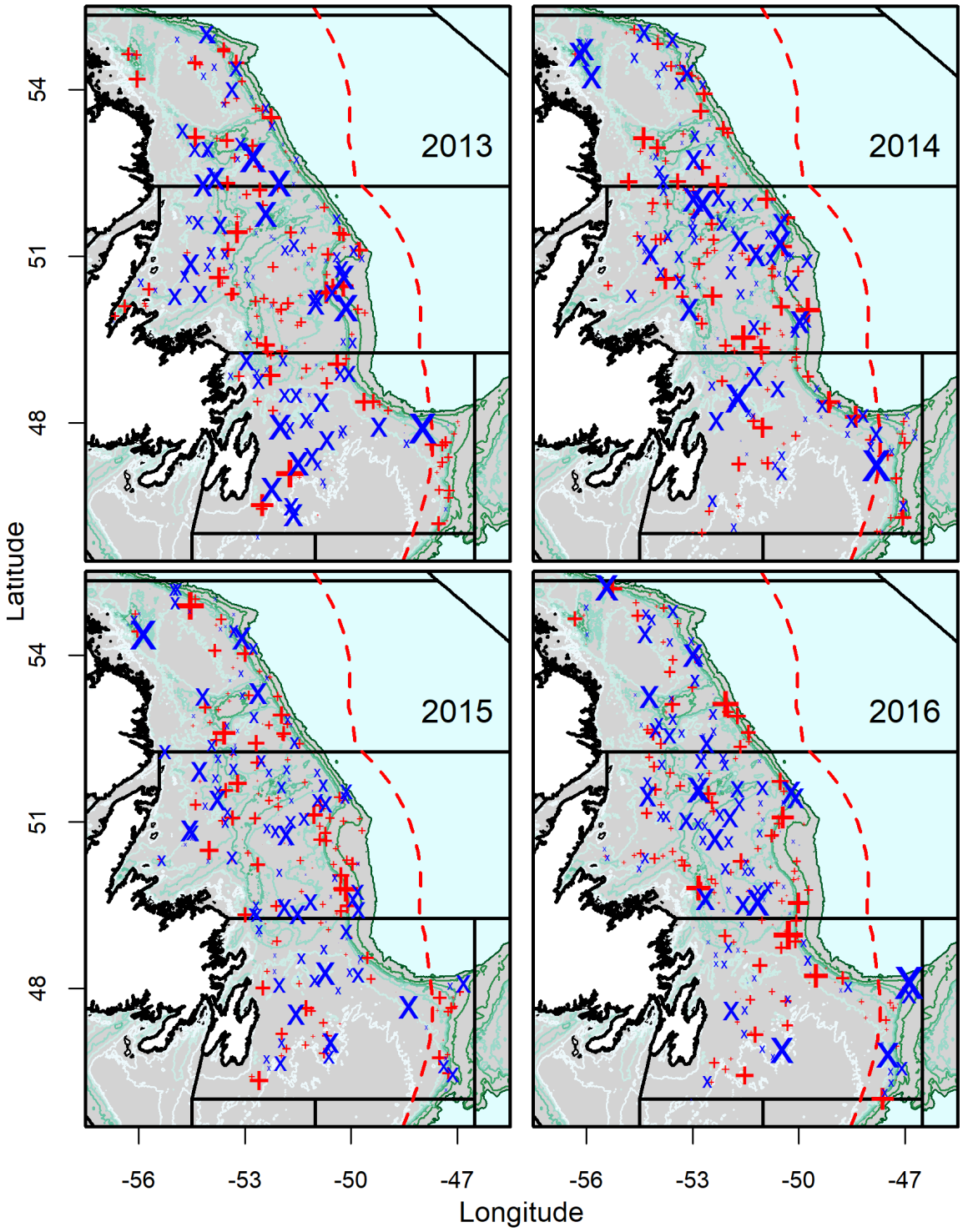

Figure 3.26: The spatial distribution of the simulated residuals transformed to normality in every set of NAFO Divisions 2J3KL from 2013 to 2016. Red +'s and blue ×'s represent positive and negative residuals, respectively, and the size of these symbols indicates the absolute value of the residuals. See Figure 1.5 for other figure details.

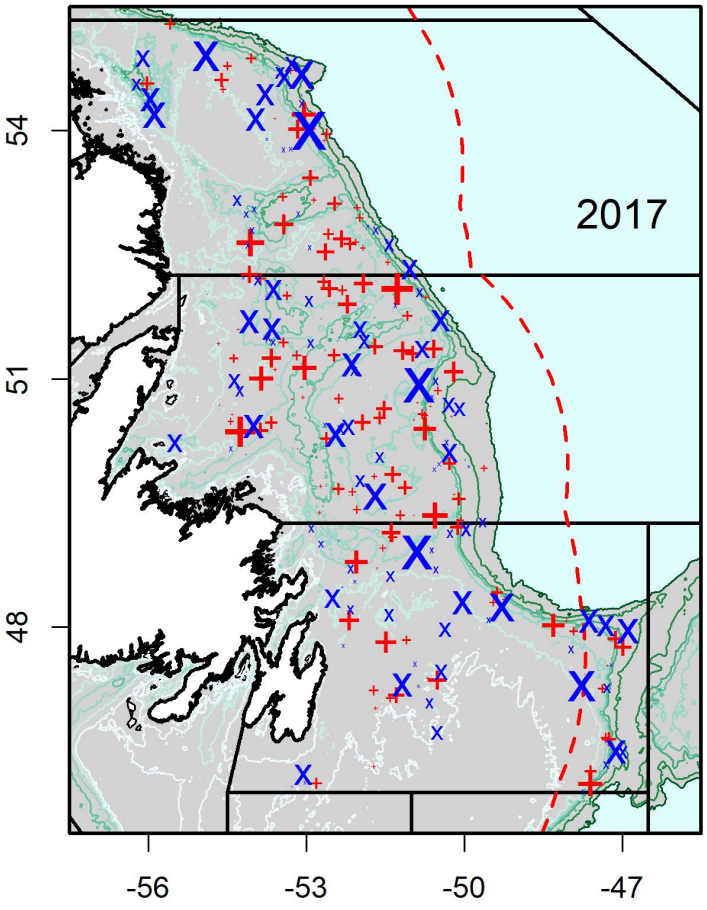

Figure 3.27: The spatial distribution of the simulated residuals transformed to normality in every set of NAFO Divisions 2J3KL in 2017. Red +'s and blue ×'s represent positive and negative residuals, respectively, and the size of these symbols indicates the absolute value of the residuals. See Figure 1.5 for other figure details.

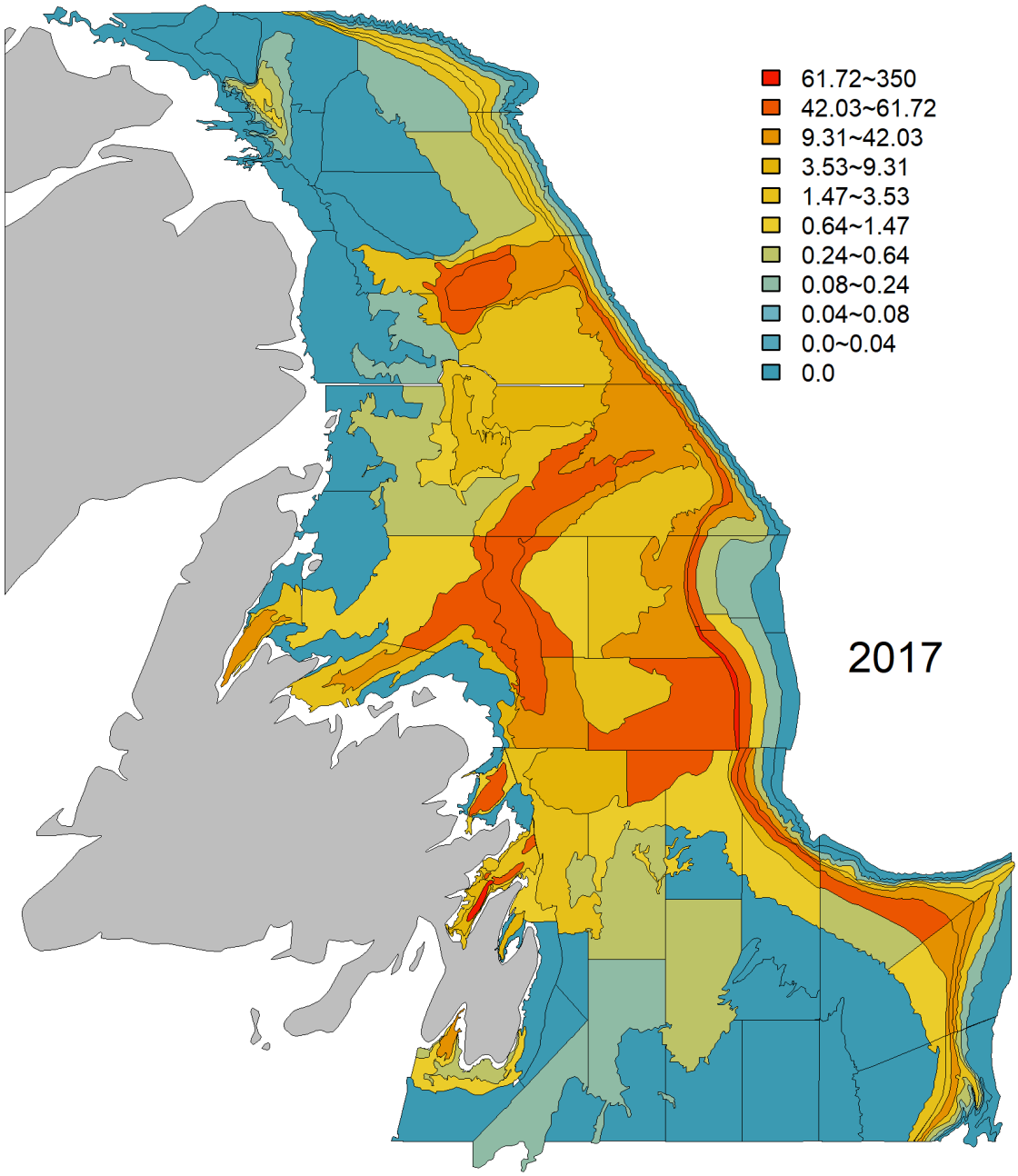

Figure 3.28: The spatial distribution of predicted mean catch in each stratum for the witch flounder survey data in 2017. The NB predicted mean catch for each stratum is indicated by the colour box on the right of the figure.

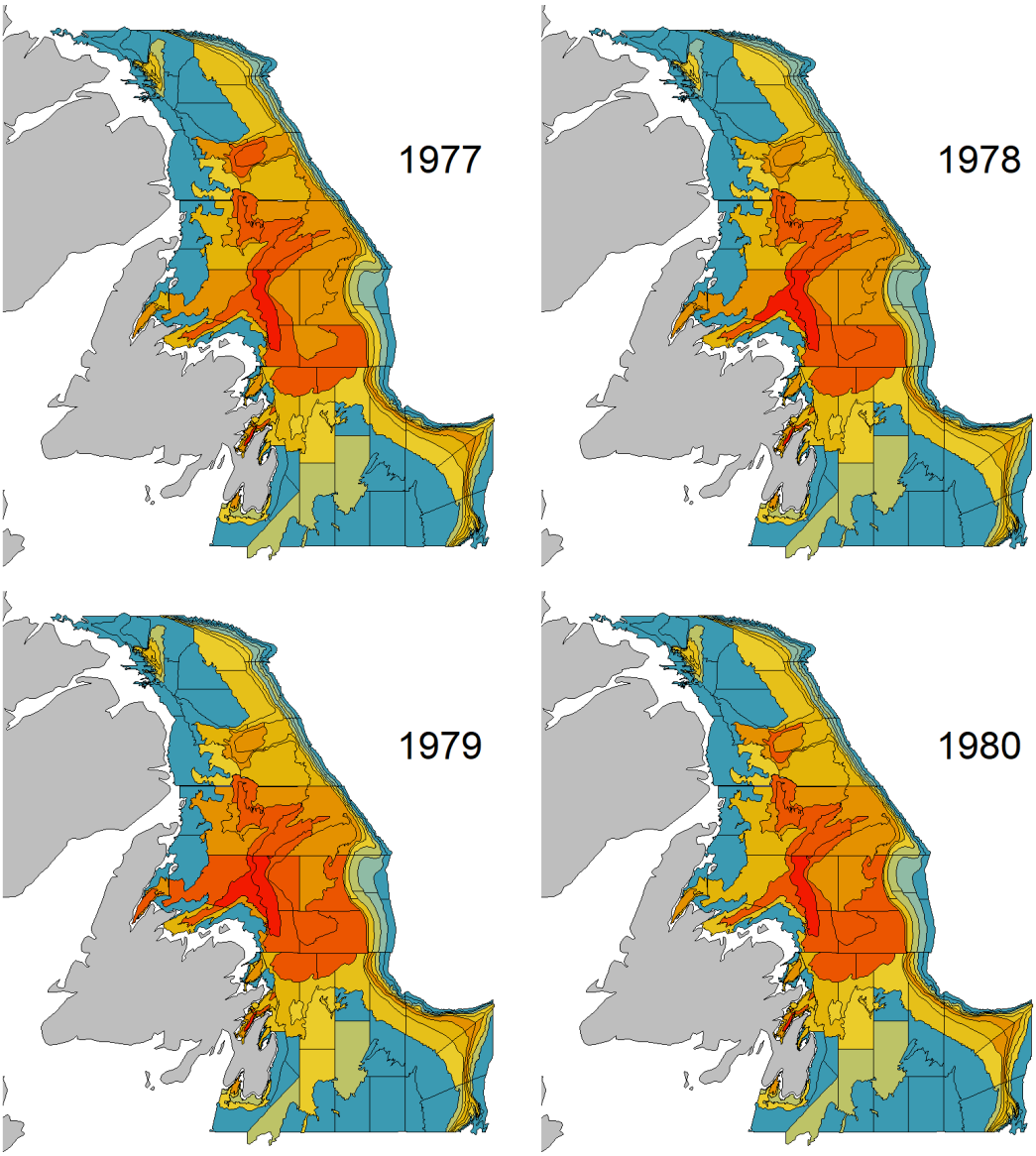

Figure 3.29: The spatial distribution of predicted mean catches in each stratum for the witch flounder survey data from 1977 to 1980. See Figure 3.28 for additional description.

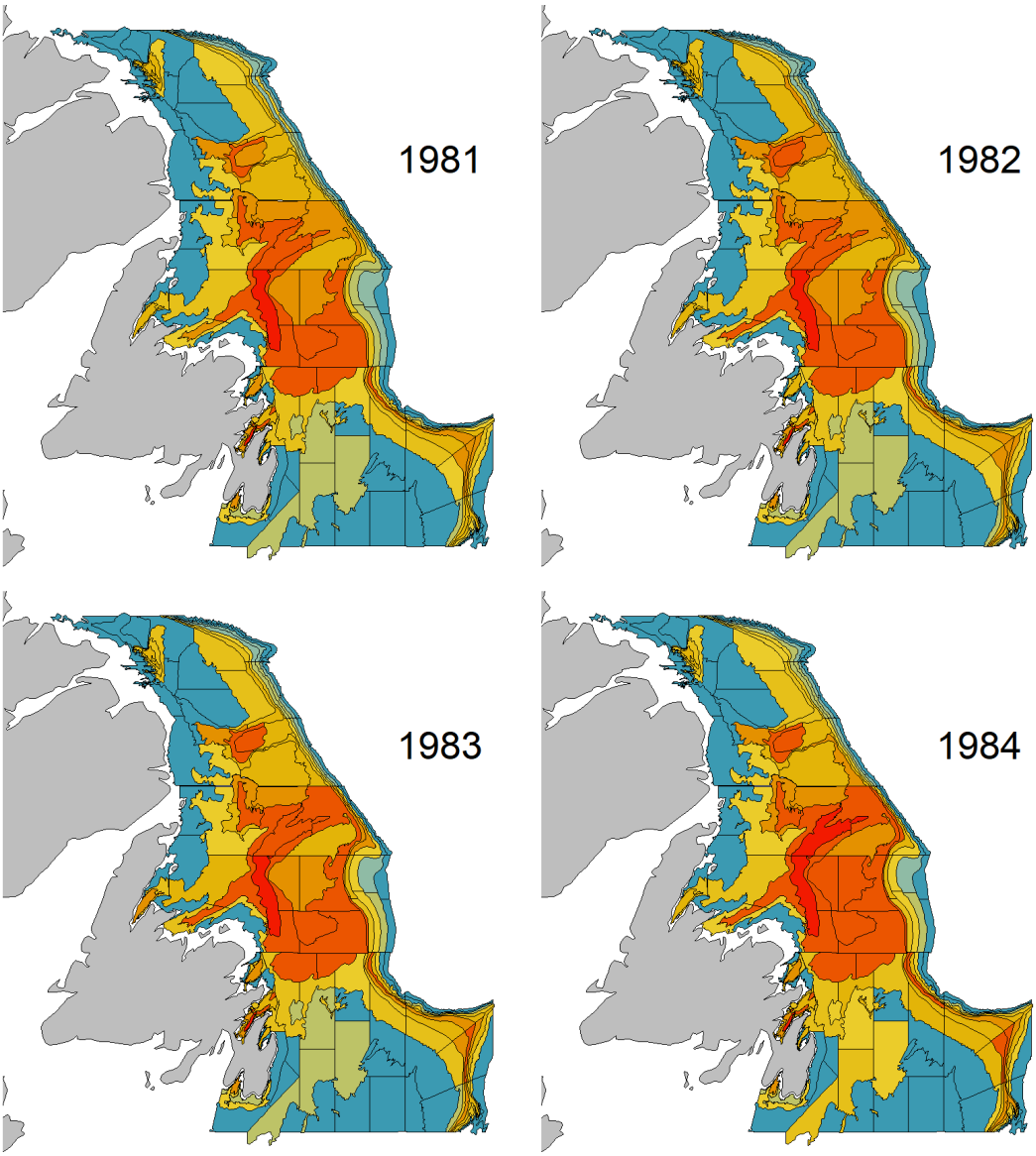

Figure 3.30: The spatial distribution of predicted mean catches in each stratum for the witch flounder survey data from 1981 to 1984. See Figure 3.28 for additional description.

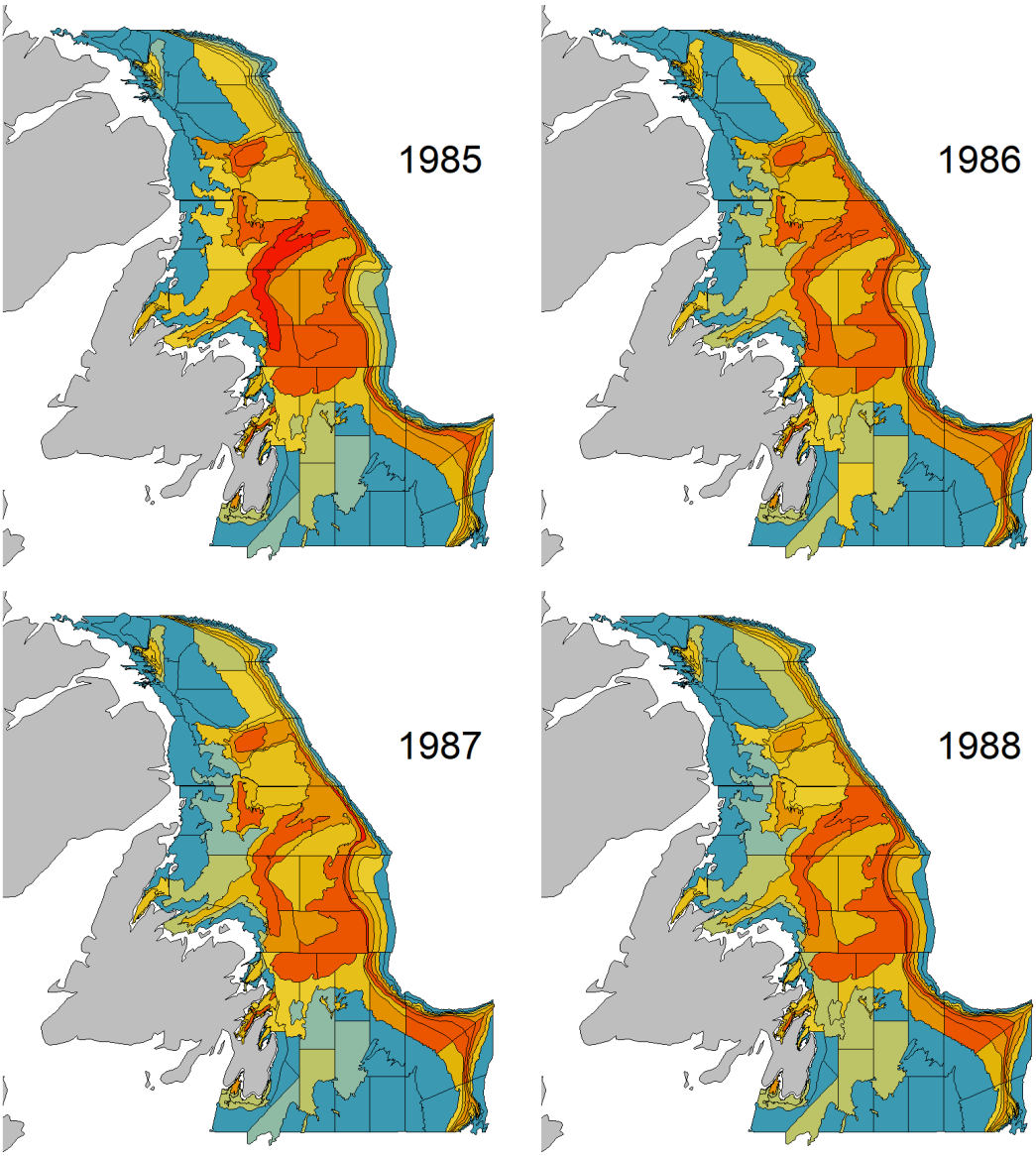

Figure 3.31: The spatial distribution of predicted mean catches in each stratum for the witch flounder survey data from 1985 to 1988. See Figure 3.28 for additional description.

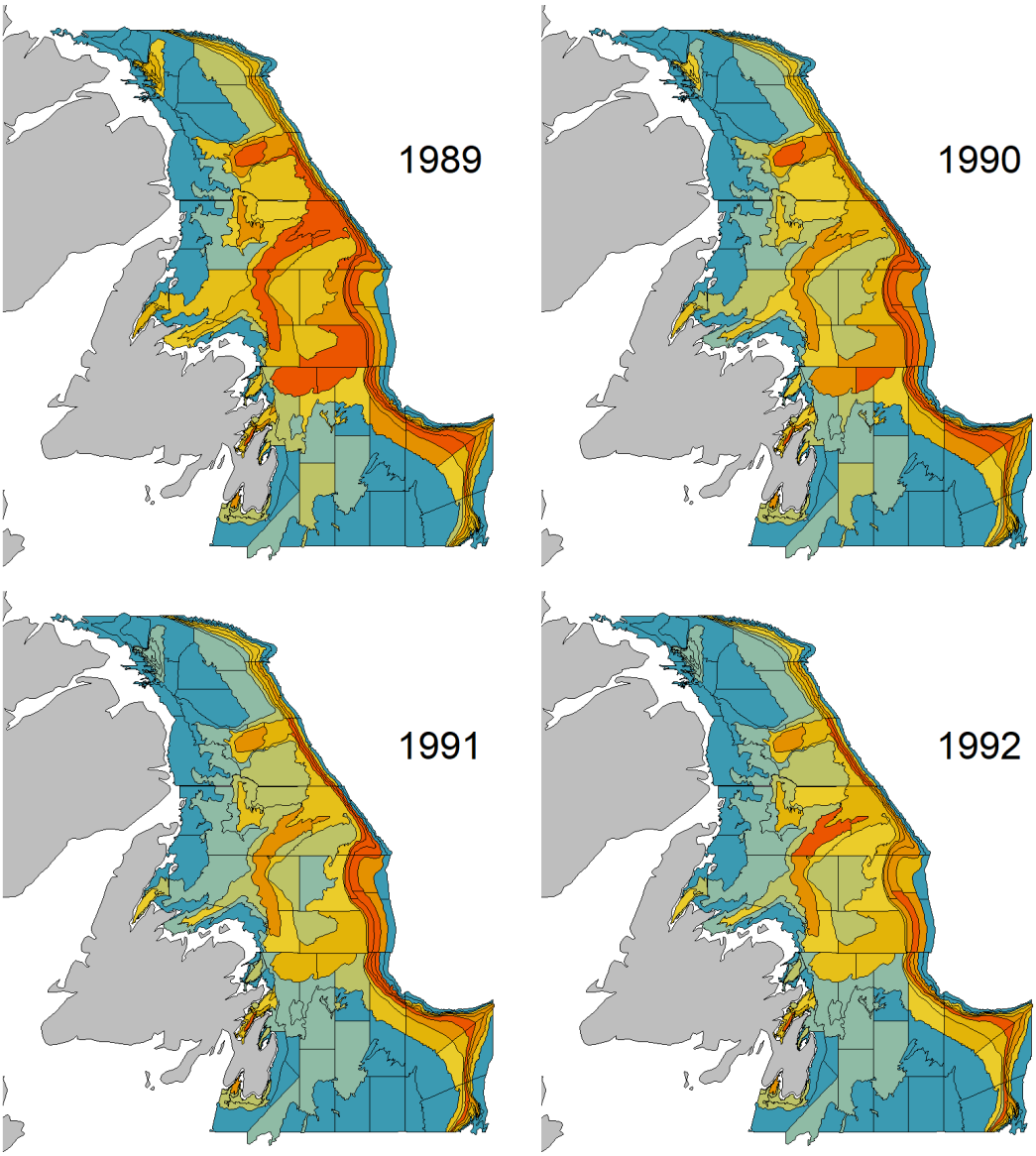

Figure 3.32: The spatial distribution of predicted mean catches in each stratum for the witch flounder survey data from 1989 to 1992. See Figure 3.28 for additional description.

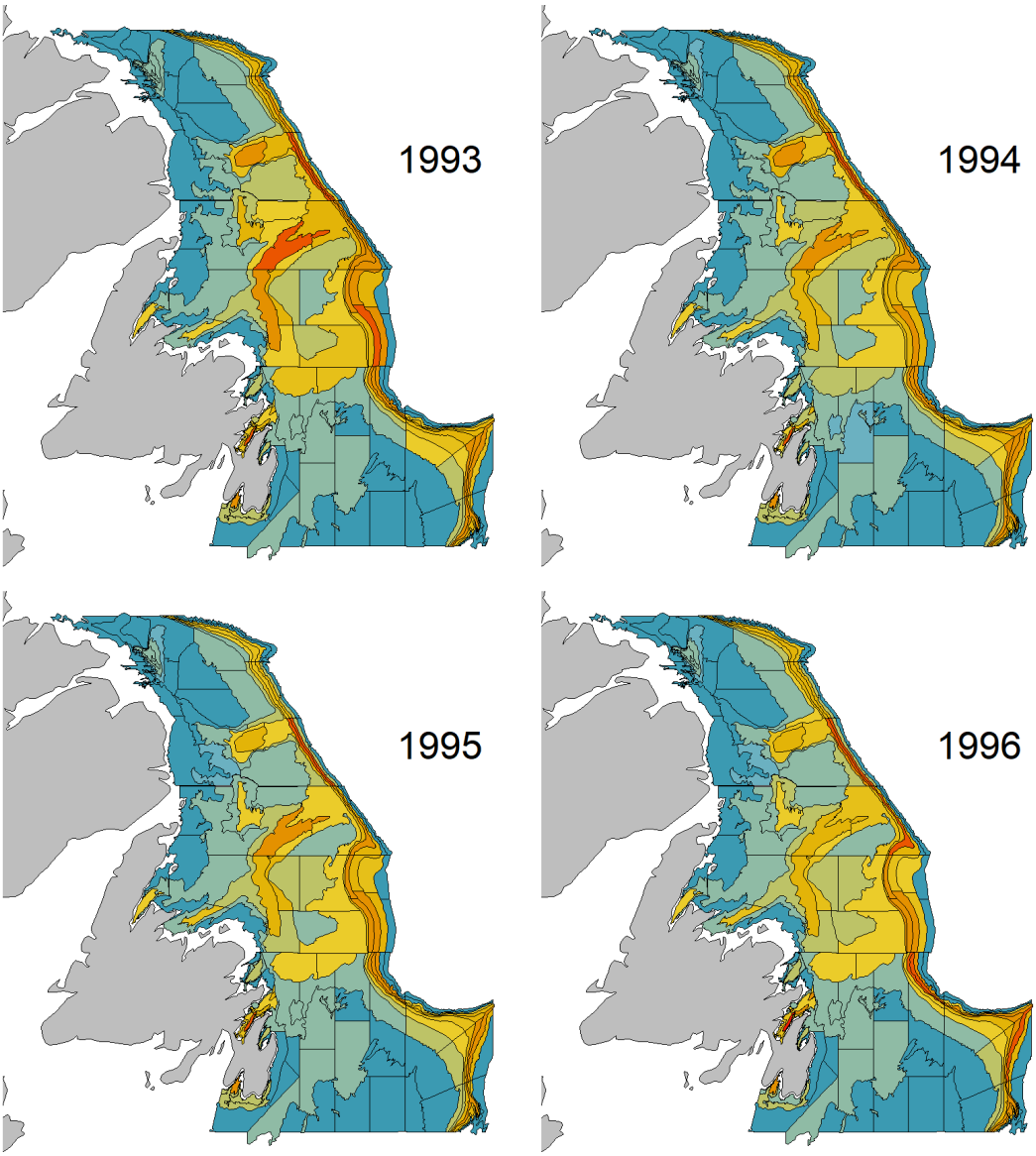

Figure 3.33: The spatial distribution of predicted mean catches in each stratum for the witch flounder survey data from 1993 to 1996. See Figure 3.28 for additional description.

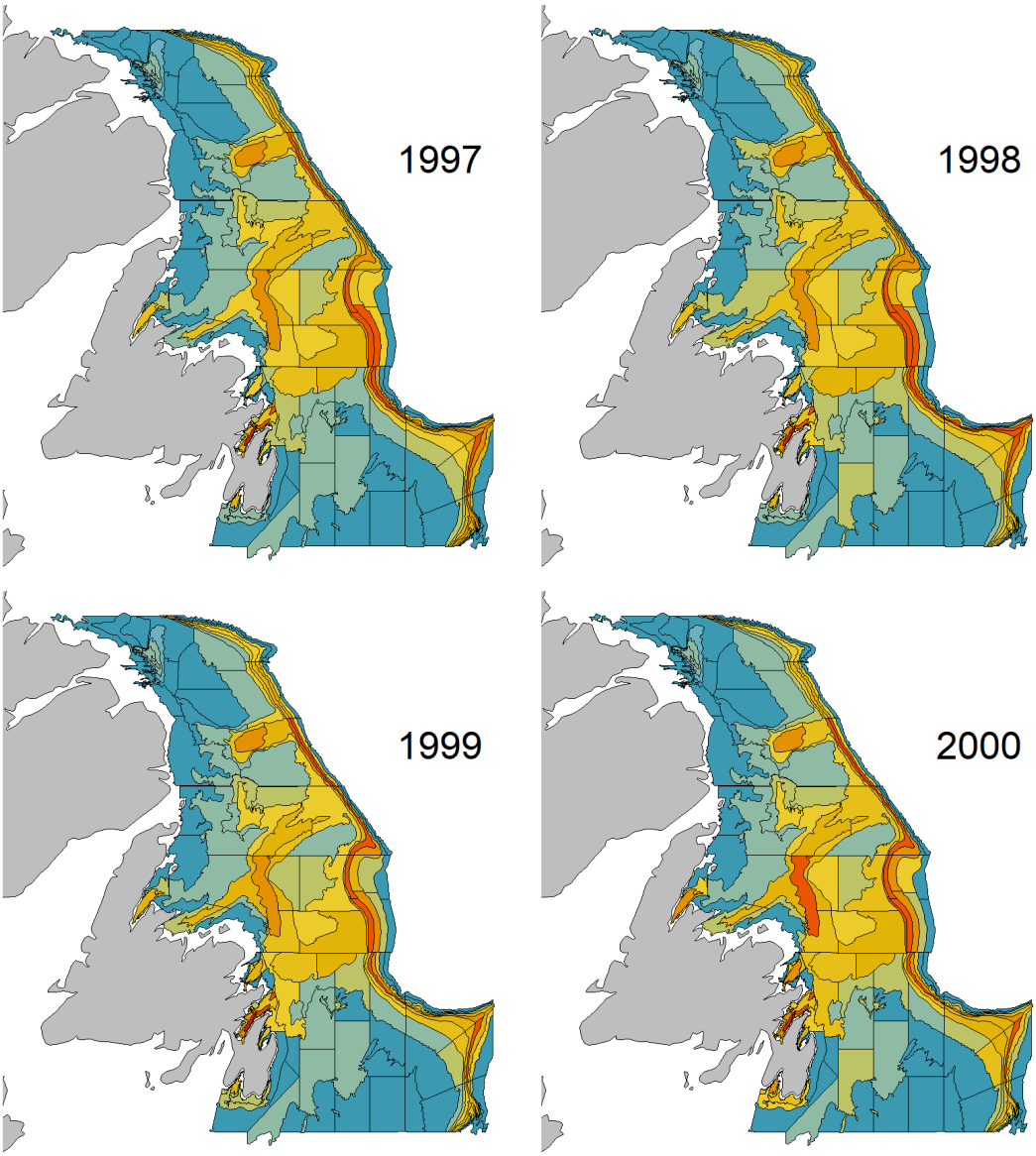

Figure 3.34: The spatial distribution of predicted mean catches in each stratum for the witch flounder survey data from 1997 to 2000. See Figure 3.28 for additional description.

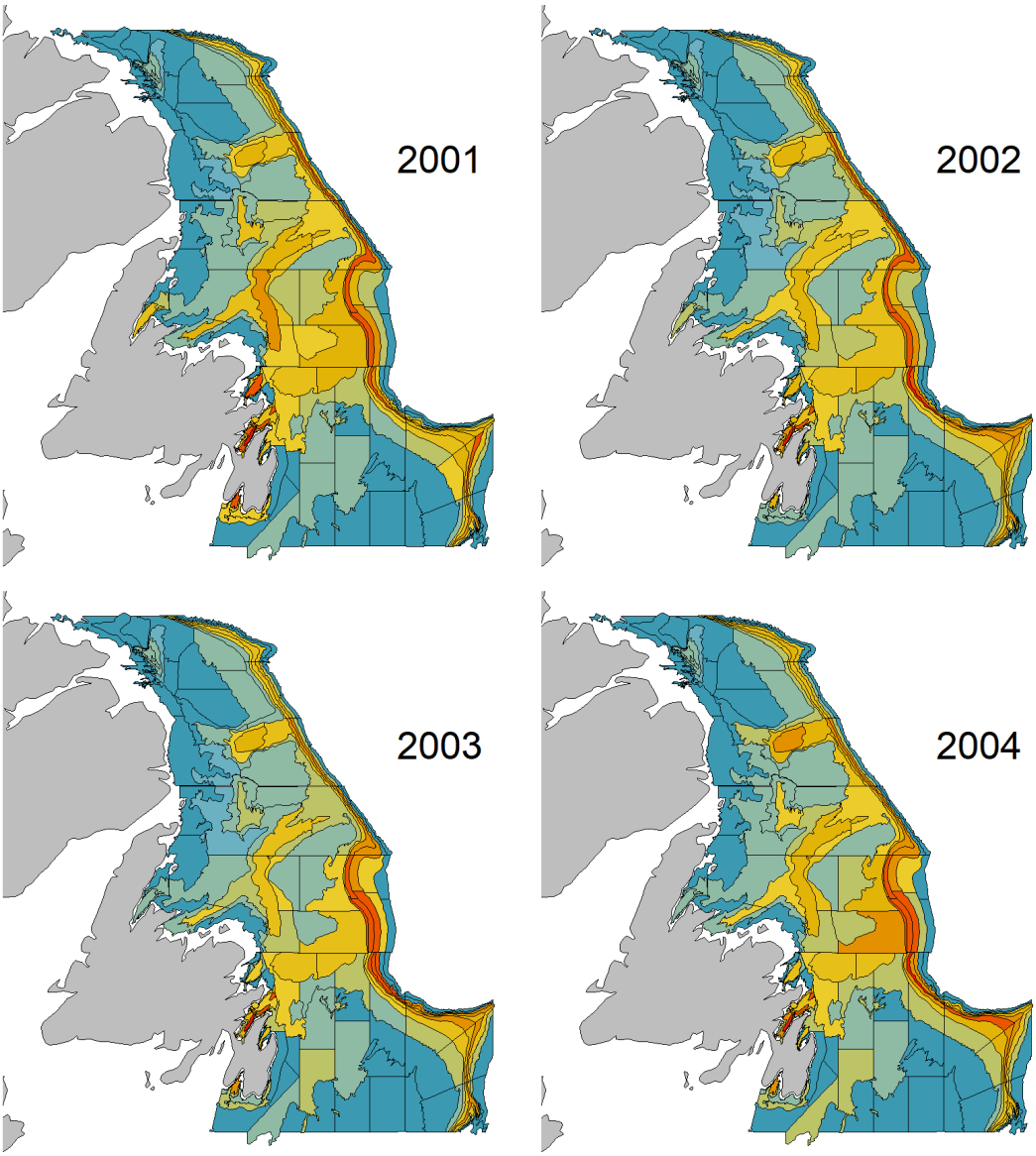

Figure 3.35: The spatial distribution of predicted mean catches in each stratum for the witch flounder survey data from 2001 to 2004. See Figure 3.28 for additional description.

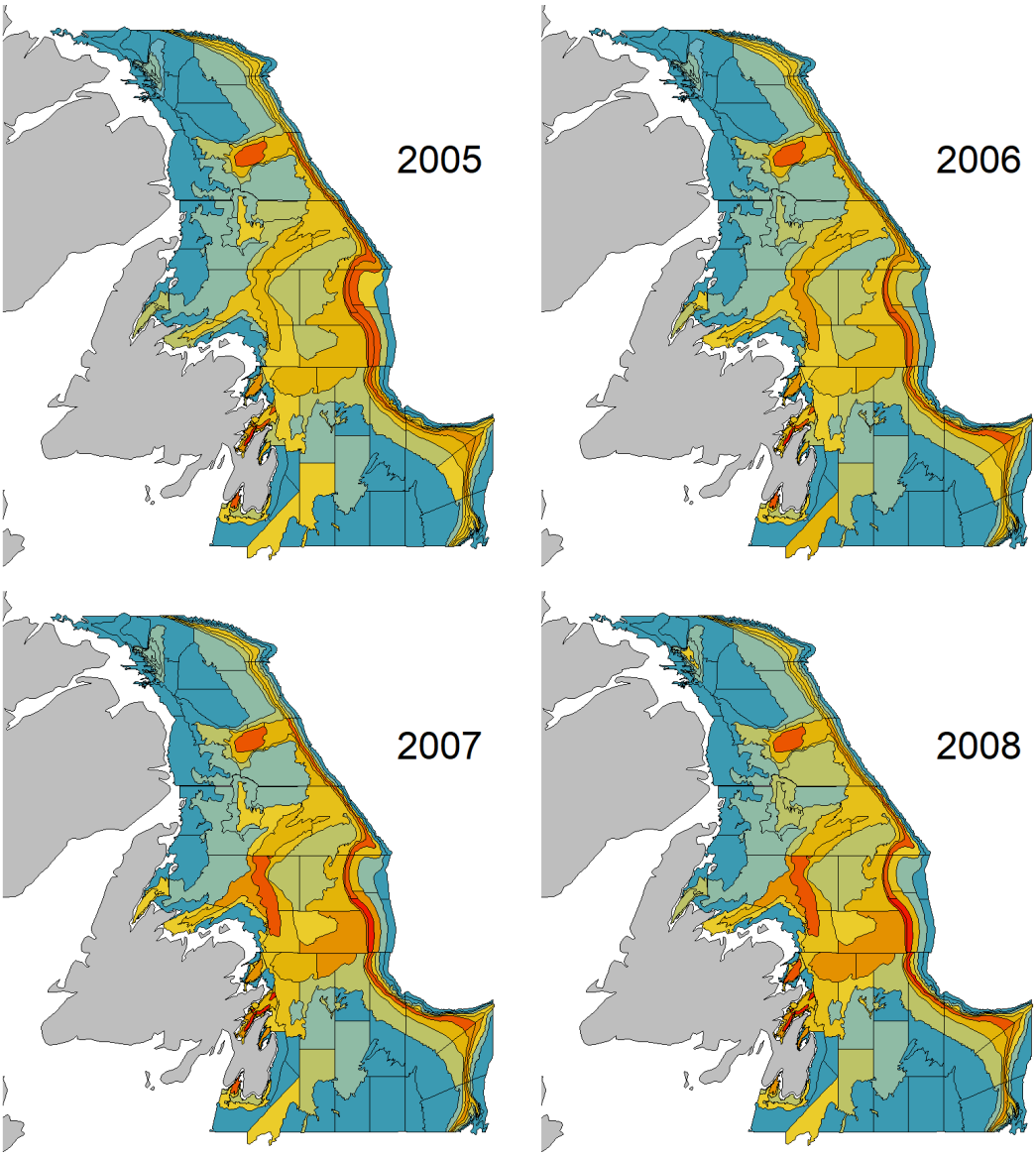

Figure 3.36: The spatial distribution of predicted mean catches in each stratum for the witch flounder survey data from 2005 to 2008. See Figure 3.28 for additional description.

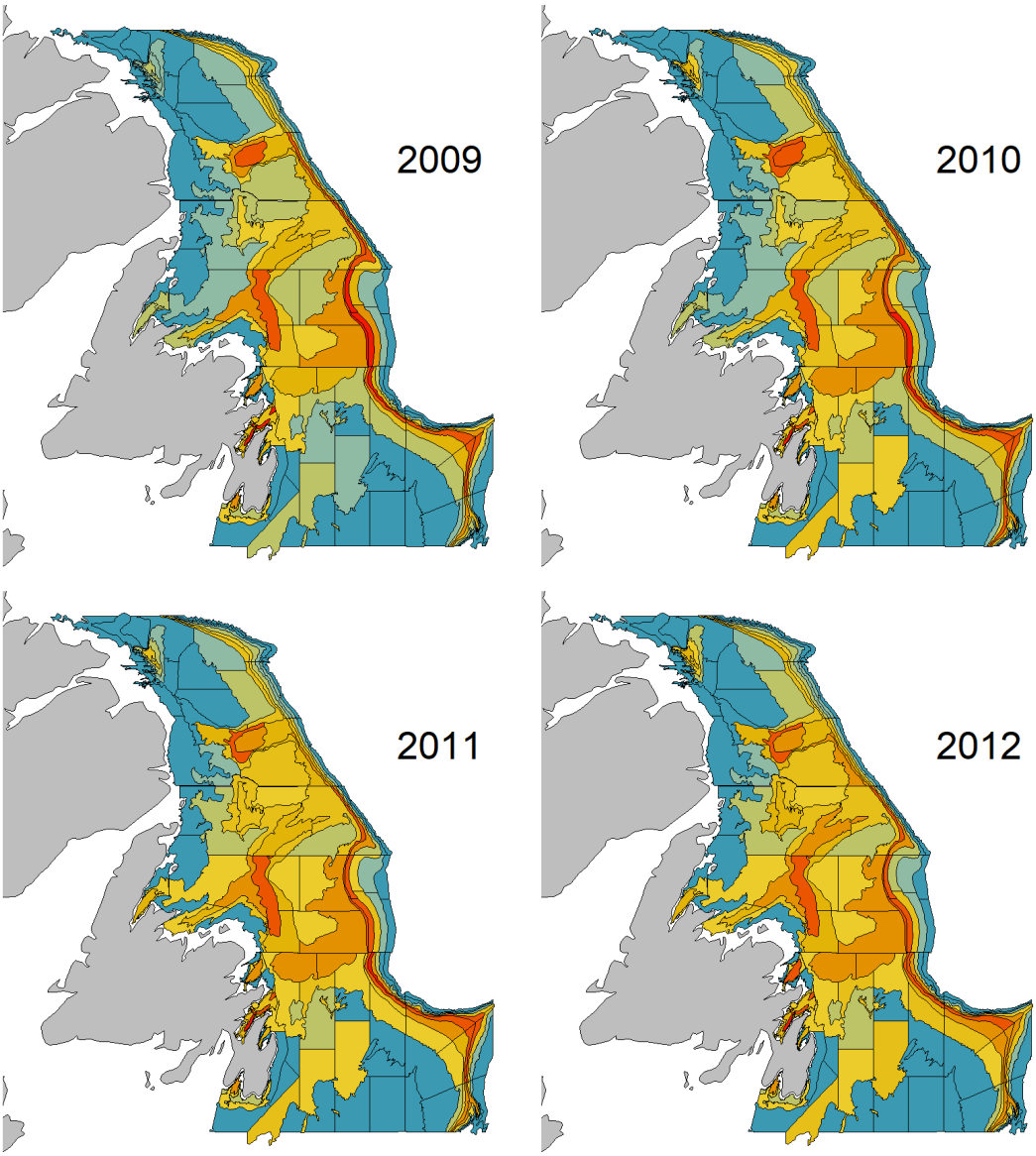

Figure 3.37: The spatial distribution of predicted mean catches in each stratum for the witch flounder survey data from 2009 to 2012. See Figure 3.28 for additional description.

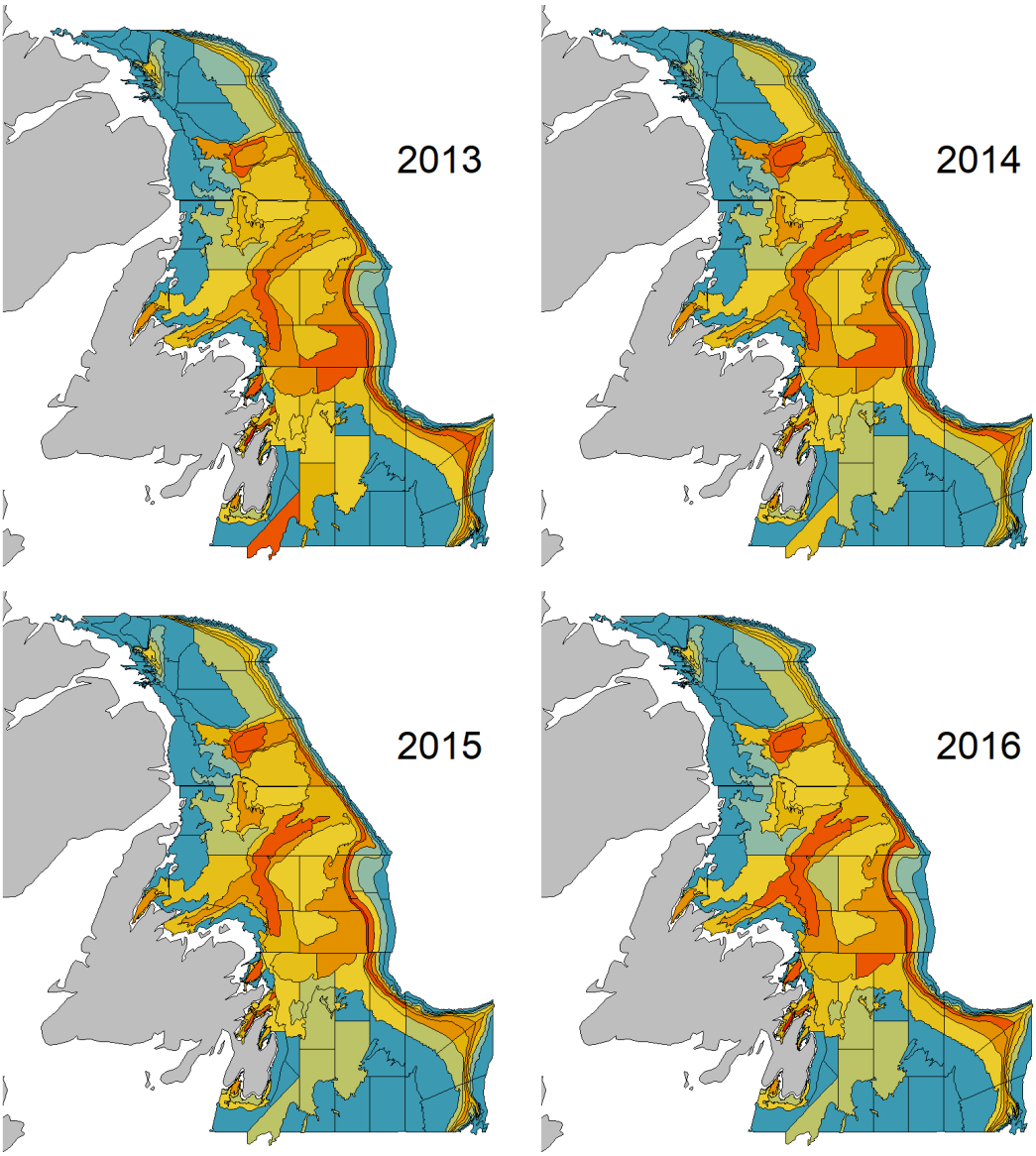

Figure 3.38: The spatial distribution of predicted mean catches in each stratum for the witch flounder survey data from 2013 to 2016. See Figure 3.28 for additional description.

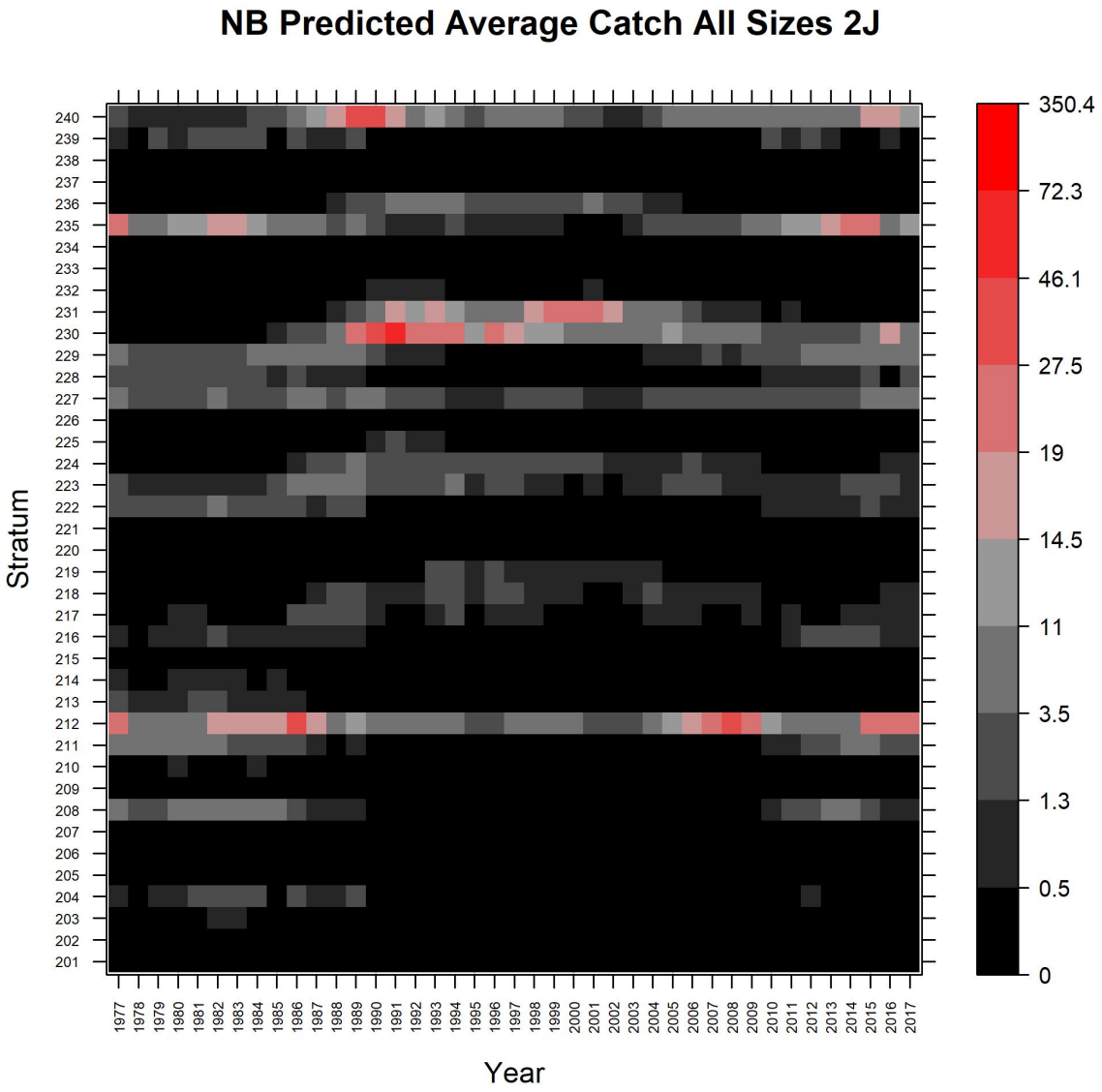

Figure 3.39: The negative binomial predicted mean catch for each stratum in Division 2J during 1977-2017. Each cell's colour represents the value of mean catch in the corresponding stratum and year based on the colour bar.

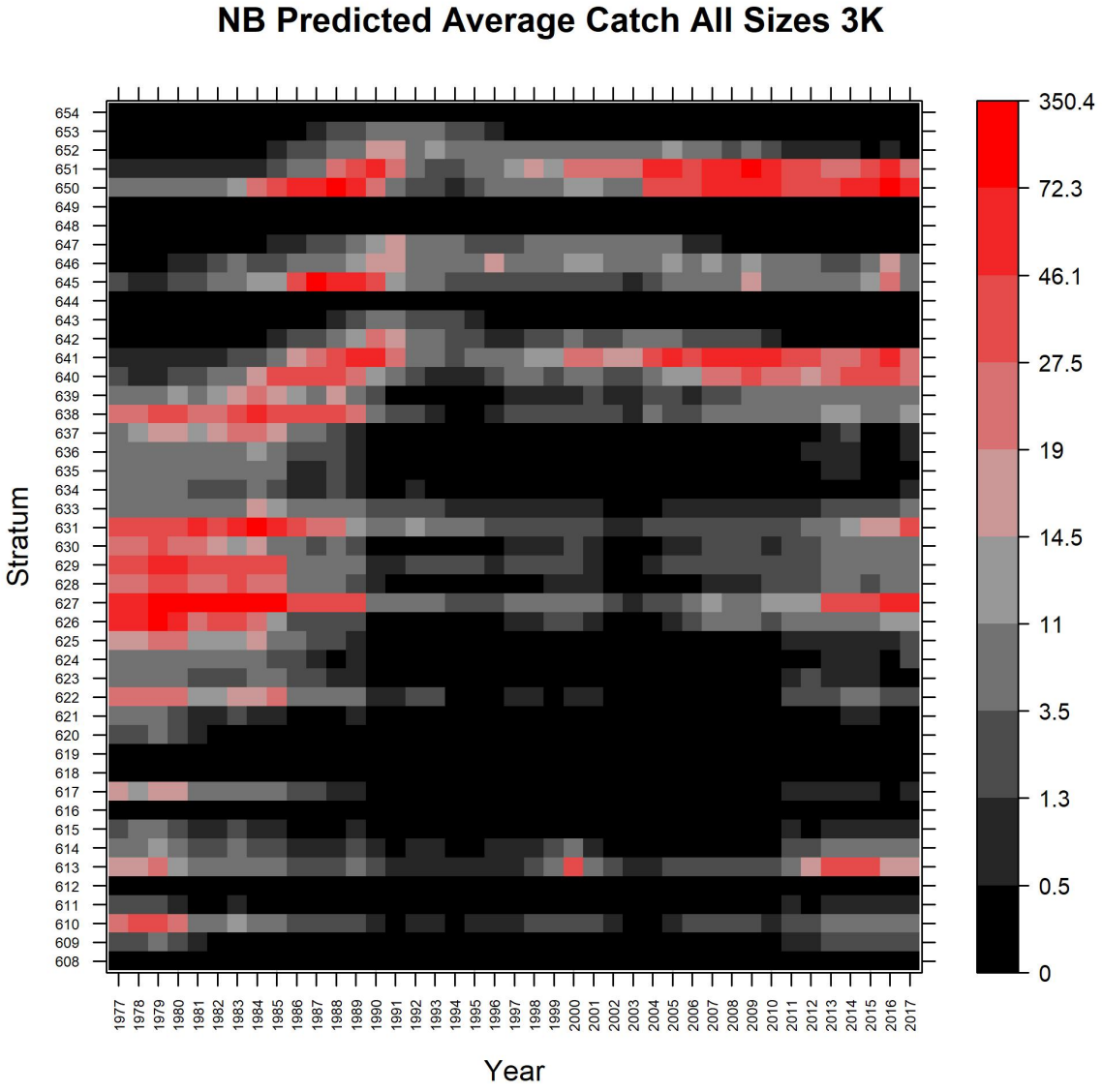

Figure 3.40: The negative binomial predicted mean catch for each stratum in Division 3K during 1977-2017. See Figure 3.39 for additional description.

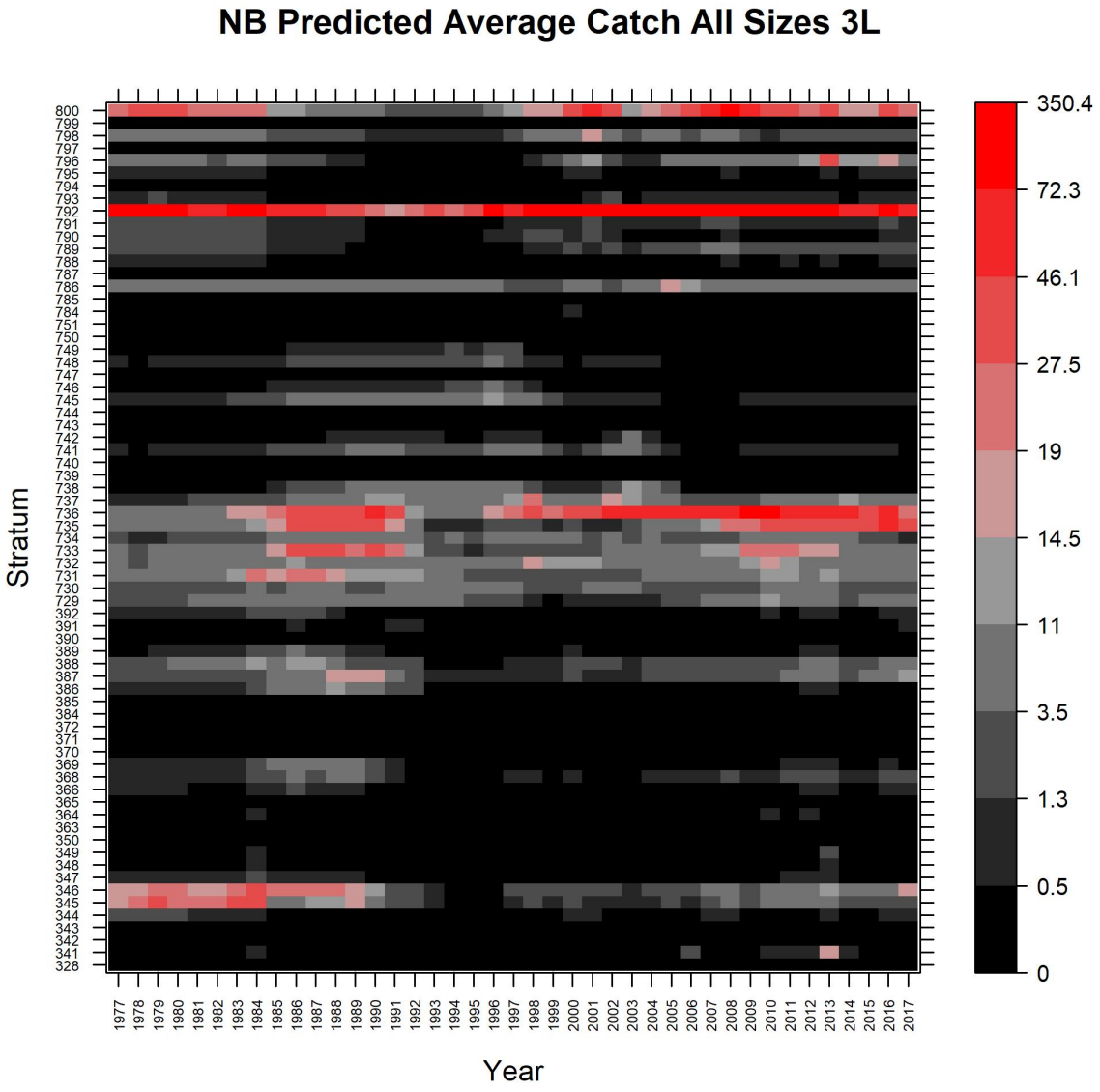

Figure 3.41: The negative binomial predicted mean catch for each stratum in Division 3L during 1977-2017. See Figure 3.39 for additional description.

Chapter 4

Summary and Future Research

4.1 Summary

Stock size survey indices are the primary input to stock assessment. More reliable indices should result in more reliable stock assessments. One issue that may cause an unreliable trend of the indices is changes in survey coverage during a time series. The primary goal of this research was to develop a spatiotemporal model to address changes in survey coverage to improve stock size survey indices with an application to witch flounder in the Northwest Atlantic Fisheries Organization (NAFO) Divisions 2J3KL off the eastern coasts of Newfoundland and Labrador.

In Chapter 1, I described the Fisheries and Oceans Canada (DFO) trawl survey in the Northwest Atlantic and details of this survey design (i.e. stratified random sampling design). I then described the research problem in this thesis, which is changes in coverage of surveys conducted during the time series. This is a common problem in research vessel (RV) surveys due to missing strata (e.g. strata not sampled because of poor weather conditions, vessel break-downs, etc) and/or adding strata (i.e. expanding survey region because of modifying the survey objectives). Since not all strata are sampled each year in the surveys, the estimated indices of stock size (e.g. the design-based indices

of abundance) only represent the strata surveyed each year. Therefore, survey coverage changes create uncertainty in the trend of stock size indices and have a significant impact on the reliability of stock size index time series. One approach to deal with changes in survey coverage is spatiotemporal modelling. The spatiotemporal models can predict and interpolate mean catches at missing areas in the stock area.

In Chapter 2, I used the design-based approach as a baseline to compare the performance of a spatiotemporal model developed in this thesis. Therefore, I estimated design-based survey indices for witch flounder that were similar to the indices in the most recent stock assessment by DFO (2019). Detailed investigation of the witch flounder survey data showed, apart from changes in survey coverage, many changes have occurred to some strata in NAFO Divisions 2J3KL, especially since 1996. These changes in the strata area and boundaries are a problem in spatiotemporal models that consider strata as covariates. Indeed, strata as spatial units in stratified random trawl surveys needed to be redefined consistently over time. Therefore, I did a re-stratification of the survey data based on currently defined strata boundaries to address this problem and use consistent strata in a spatiotemporal model.

In Chapter 3, I used a model-based approach (i.e. a statistical model) to address changes in coverage of witch flounder surveys to improve the reliability of abundance indices (e.g. mean number per tow, MNPT) for this stock. In fact, I used a mixed-effect model to predict the mean catches for missing strata and fill missing areas.

The best-fitting model I selected included two random effects, the main stratum effect, ω, assumed to be a Gaussian Markov random field (GMRF), and strata-year effect, δ, that is GMRF correlated across strata and autoregressive of order 1, AR (1), correlated across years. Thus, the spatial variation in the model contains two components: one component that is constant across time, and another component that changes among years and represents changes in the spatial distribution of witch flounder over the time series. This model accounts for the correlation between catches in neighbour strata and adjacent

years to estimate abundance indices and predict these indices for those areas that were not sampled in some survey years. Indeed, the computer program automatically predicts mean catches in areas that were not sampled using nearby sampled areas. Therefore, the spatiotemporal model I developed can predict the abundance index (i.e. MNPT) for missing areas using neighbouring information and increase the length of index time series.

A stock assessment problem with using the design-based MNPT as an index of stock size is that the sampled strata may not be the same each year, and the total area sampled may also differ. Therefore, using the spatiotemporal model, I calculated model-based MNPT for all strata in each NAFO division and in each year. The model enabled us to fill the data gaps in some survey years to provide more reliable survey indices for the witch flounder stock assessment.

For diagnostic purposes, I compared the model MNPT for sampled-only strata with the design-based MNPT and the model MNPT for all strata. As the design-based approach provides an unbiased index for sampled strata, the comparison of model MNPT for sampled-only strata and design-based MNPT was considered as the first potential step in assessing the appropriateness of the model to predict missing observations (unsampled strata). However, further steps to assess the model's ability to predict missing observations can be cross-validation and simulation testing. These comparisons showed that the model used is appropriate for predicting MNPT because model MNPT for sampled-only strata are similar to the design-based MNPT and the model impact for interpolation of unsampled strata is small. For these comparisons, I used the bias-corrected model MNPT for all and sampled-only strata because a nonlinear function of fixed and random effect estimates used to calculate model MNPT can produce biased estimates of MNPT.

I also clustered strata based on the spatial correlation matrix of the strata-year effects to detect the similarities between strata in terms of the spatiotemporal variation of the witch flounder survey data. I plotted a strata dendrogram cut into three clusters to show which strata are more similar to each other based on the hierarchical method. The model

MNPT for each cluster was plotted to show the MNPT trend of each cluster during the time series, as a simple way to demonstrate the overall spatiotemporal variation in MNPT.

I performed spatial and temporal autocorrelation tests on residuals using the DHARMa package. The test for temporal autocorrelation indicated that there is no temporal autocorrelation in the residuals. Lack of temporal autocorrelation is good because it suggests the model has adequately accounted for temporal trends in the survey data. The spatial autocorrelation test was performed for each year that was significant for only six years. Significant spatial autocorrelation in these six years indicates there may be some potential of improved fitting with further refinement of the spatial effects (i.e. strata) in the model. However, we do not anticipate that this will produce much different standardized abundance indices for stock assessment.

4.2 Future Research

Since some residual patterns were apparent and some goodness-of-fit tests were significant (see Section 3.4.1), I suggest further investigation of the model is warranted. Interpreting residuals for generalized linear mixed models (GLMMs) is often challenging and complicated because of random effects. For this reason, it is necessary to examine the model and its estimation further.

Another issue to examine in future research is if changes in the timing of the surveys each year could be affecting the survey indices. I examined if the survey timing has changed during the time series in each NAFO division. Large changes in survey timing within NAFO divisions can be a problem because of fish movement during the fall and winter when the surveys typically occur. Figure 4.1 shows that most surveys were conducted in the fall in 2J3KL. I used box plots to show the distribution of survey days within months for NAFO Divisions 2J, 3K and 3L separately (Figures 4.2-4.4). Therefore,

there are no large changes in survey timing for witch flounder. However, the dates for some surveys, within each division, could change as much as 6 weeks (e.g. a few survey sets sampled in winter, Jan. and Feb., see Figures 4.5-4.7).

Future studies should extend the model to use length information to develop spatiotemporal models and length-based survey indices. In the witch flounder survey data, the lengths of fish were assigned to length bins. The bins are 2 centimeters (cm) in width such that the smallest witch flounder in the survey was in a 2.5 cm bin, and the longest was in the 70.5 cm bin. For future research, the length effect can be modelled as a GMRF (or AR (1) process) with consecutive length bins being neighbours. A problem with length-based models is the high dimension of the model, which is in four dimensions (space $\times$ time $\times$ length). Modelling such high-dimensional correlations is computationally difficult and very slow.

An approach to investigate in future research to partially address the high dimension problem is to use a mixture model with age. The idea is to replace the L length classes with a mixture of A length distributions at age such that $A \ll L$ (i.e. A is much less than L). Therefore, in this approach, rather than have $L \sim 70$ (the maximum length of witch flounder) length classes, it is possible to have a model with $A \sim 14$ (i.e. the maximum age of witch flounder; Bowering, 1990) age classes.

Fisheries scientists study body size and growth information because growth in body size is a vital component of biological production. Time-series of length-frequencies have been used to estimate growth for more than a century (e.g. Petersen, 1892). Fisheries scientists have developed methods to use length-frequency information to study growth and estimate growth parameters. Time-series of length-frequencies can be used to provide an estimate of the growth rate and natural mortality rate, which is used for yield-per-recruit analysis.

Batts et al. (2019) developed a method to model fish growth with length-frequency data, based on the expectation-maximization algorithm (EM algorithm, Dempster et al.,

1977). In their method, they incorporated the von Bertalanffy growth structure (Von Bertalanffy, 1938) and accounted for variability in growth between cohorts or years. They then extended the method to include bivariate random effects (i.e. developing a hierarchical model) to capture some growth variations in length-frequency data. The hierarchical model allowed cohorts to be modelled over the years. Indeed, they modelled growth by time × length. Fish growth can spatially and temporally vary between cohorts, between individuals, and within individuals (Quinn and Deriso, 1999; Morrongiello and Thresher, 2015). Therefore, another future research is to extend the model developed by Batts et al., which can be a model with time × length × space.

The last issue that needs to be discussed here is that a random variable as AR(1) for the year effect in models might cause issues when estimating the model-based abundance index used as input data in a subsequent stock assessment model. Specifying an AR(1) process implies that the index is smoothed during the modelling process, which will induce a partial correlation in the estimated index between adjacent years. This estimation covariance is rarely passed to the stock assessment fitting the index because assessment models usually assume that an abundance index is independent among years. Smoothing the index might cause a hyper-stable index, where it responds slowly and has positive retrospective bias during a stock collapse. Therefore, it will be important to account for temporal autocorrelation in model-based standardized indices when fitting a stock assessment model.

4.3 Figures

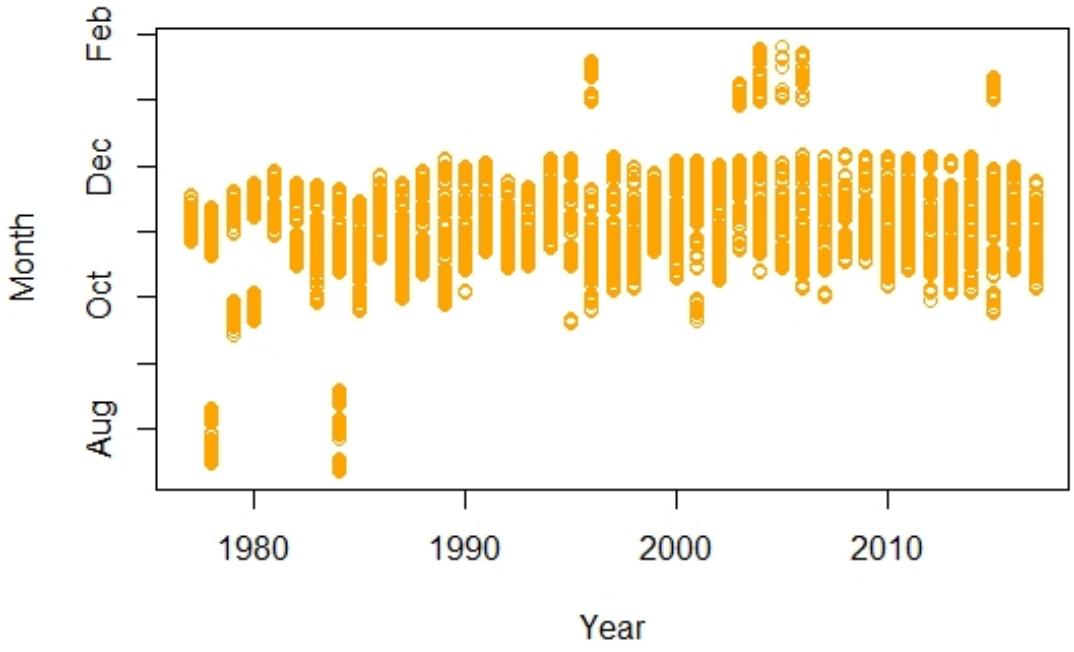

Figure 4.1: The annual survey sets sampled by month in NAFO Divisions 2J3KL. Orange circles show the most survey sets sampled from Aug. to Dec., a few survey sets sampled in Jan. and Feb.

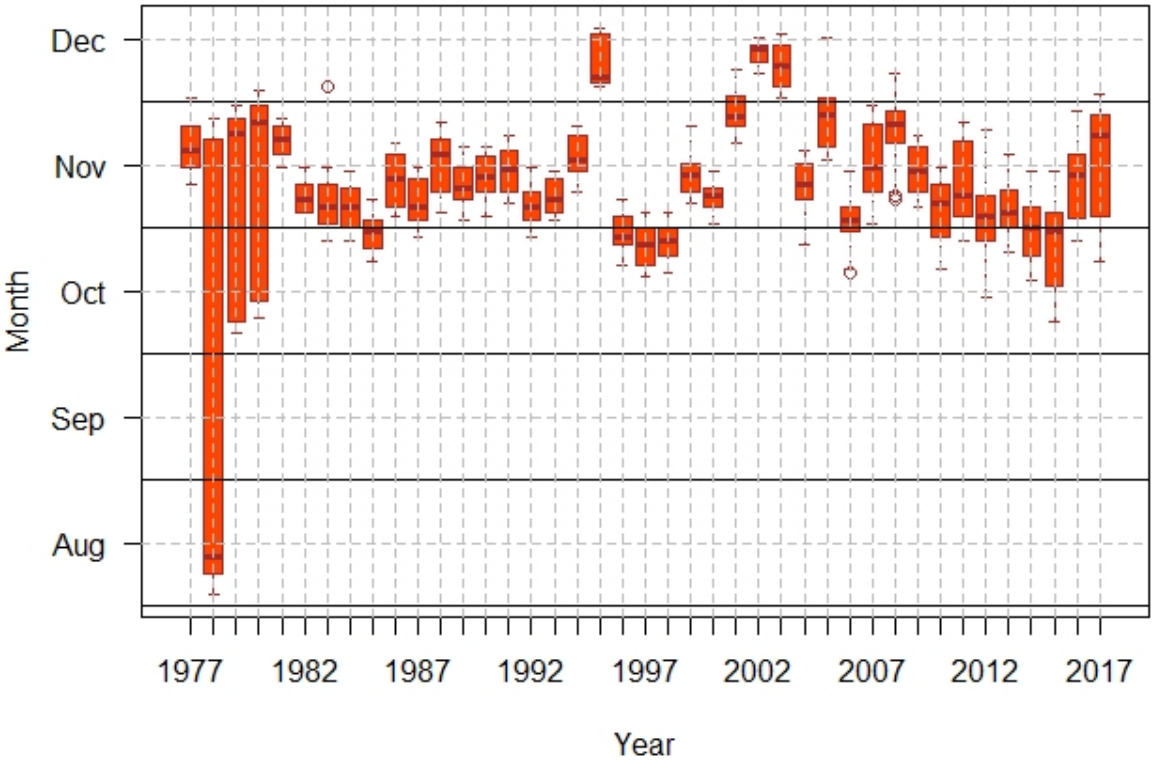

Figure 4.2: Annual survey timing in NAFO Division 2J. Box plots indicate the distribution of days within months from Aug. to Dec.

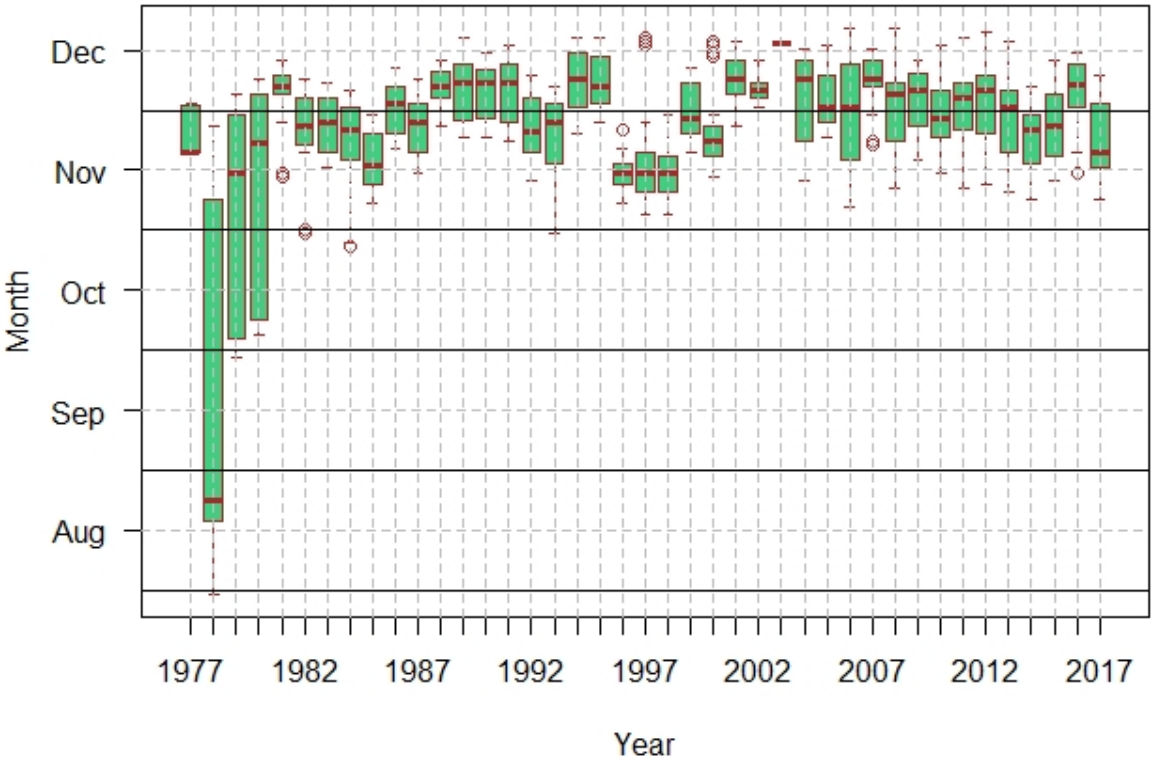

Figure 4.3: Annual survey timing in NAFO Division 3K. Box plots indicate the distribution of days within months from Aug. to Dec.

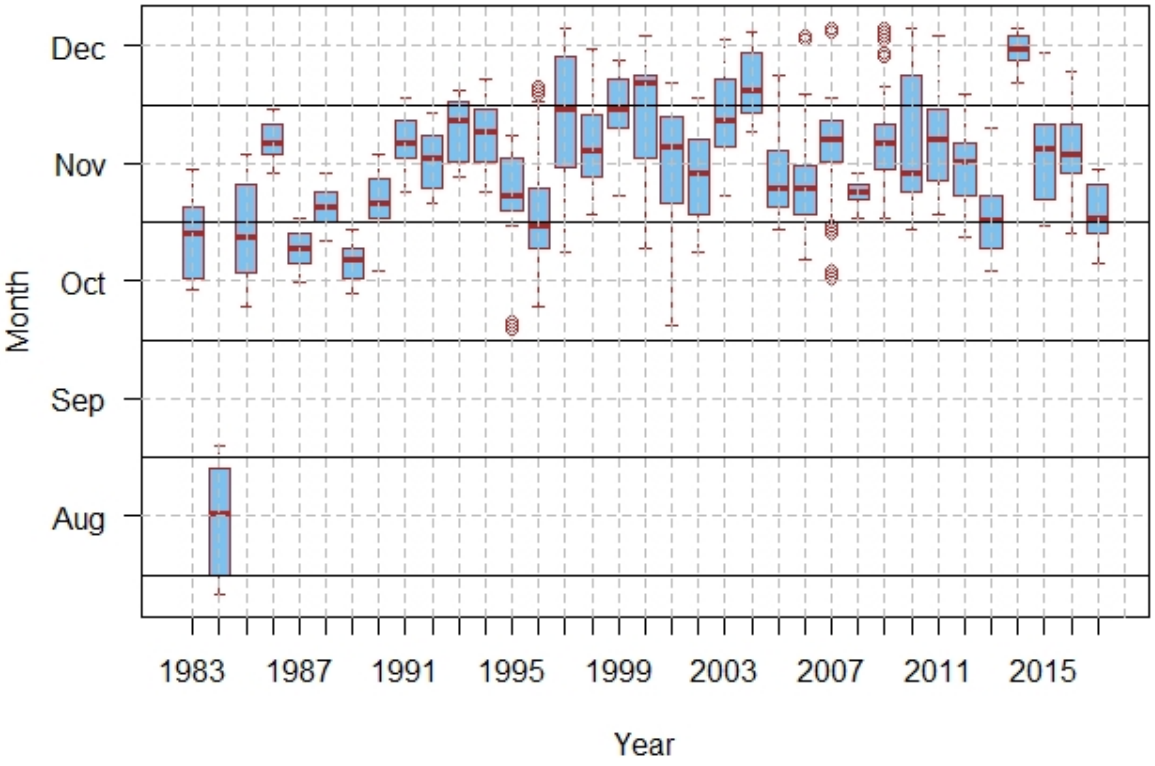

Figure 4.4: Annual survey timing in NAFO Division 3L. Box plots indicate the distribution of days within months from Aug. to Dec.

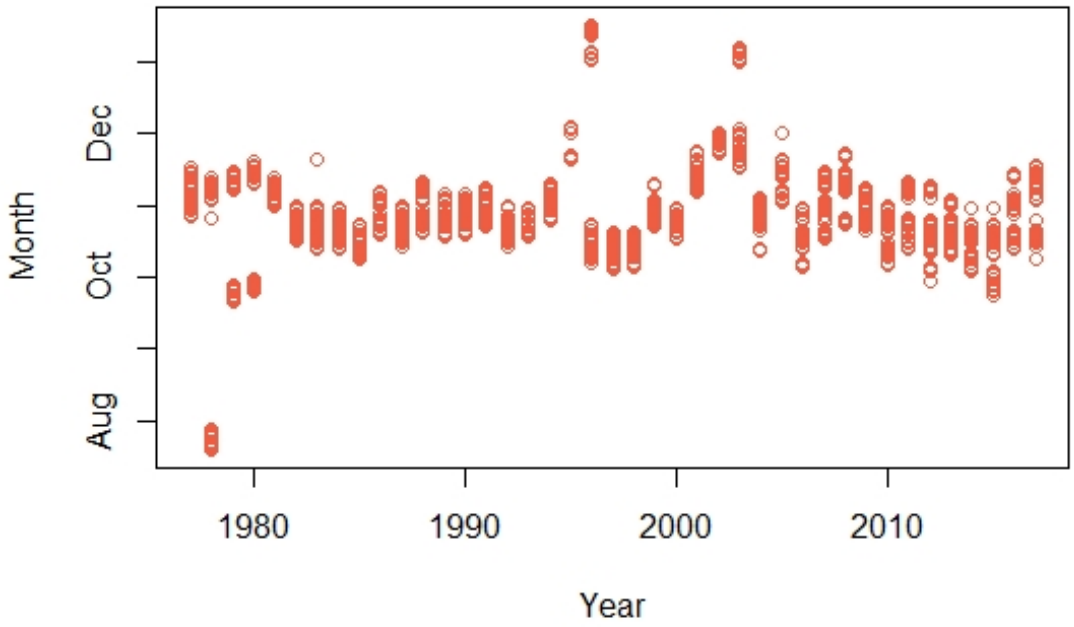

Figure 4.5: The annual survey sets sampled by month in NAFO Division 2J, which shows a few survey sets sampled in Jan. in this division.

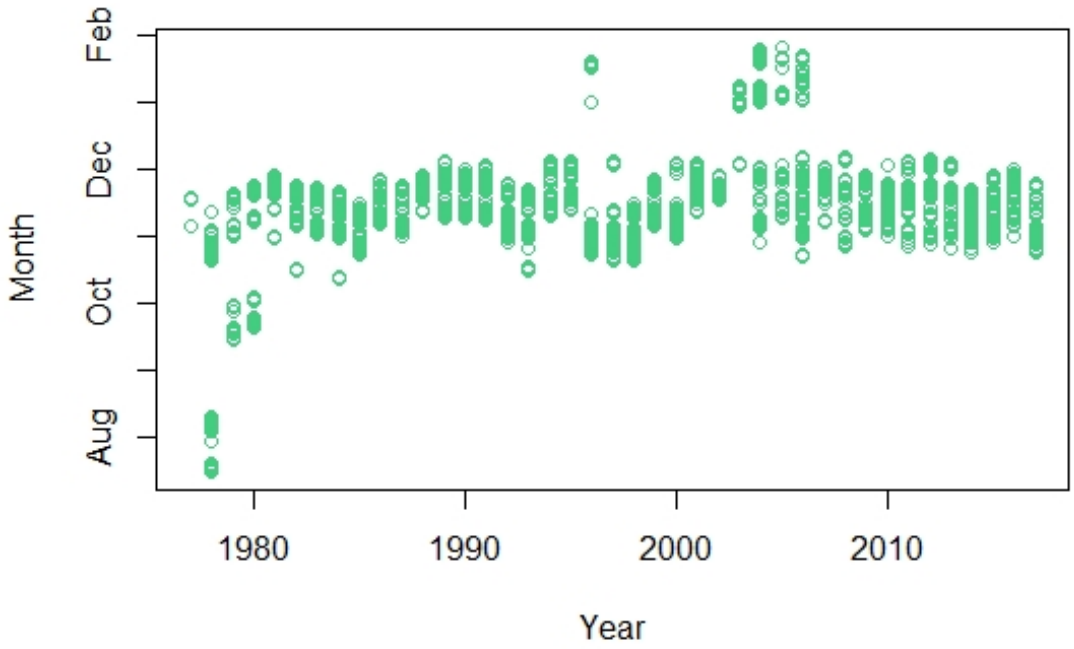

Figure 4.6: The annual survey sets sampled by month in NAFO Division 3K, which shows a few survey sets sampled in Jan. and Feb. in this division.

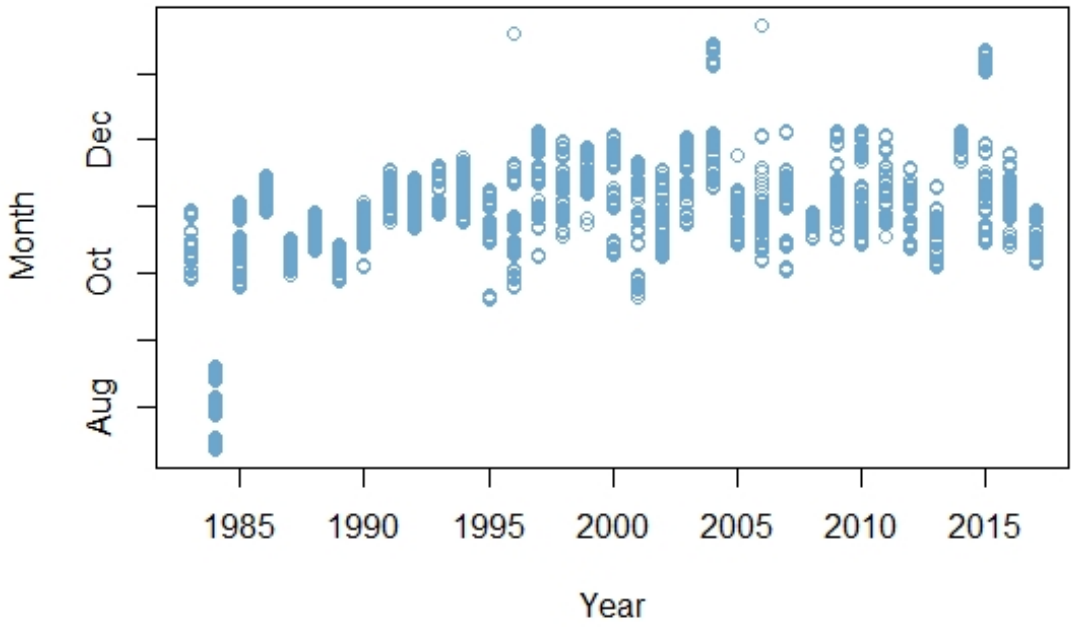

Figure 4.7: The annual survey sets sampled by month in NAFO Division 3L, which shows a few survey sets sampled in Jan. in this division.

Bibliography

Banerjee, S., A. E. Gelfand, A. O. Finley, and H. Sang (2008). Gaussian predictive process models for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 70*(4), 825–848. (Cited on page 56.)

Batts, L., C. Minto, H. Gerritsen, and D. Brophy (2019). Estimating growth parameters and growth variability from length frequency data using hierarchical mixture models. *ICES Journal of Marine Science 76*(7), 2150–2163. (Cited on pages 118 and 119.)

Berger, A., S. Harley, G. Pilling, N. Davies, and J. Hampton (2012). Introduction to harvest control rules for WCPO Tuna fisheries mow1-ip/06 14 nov 2012. (Cited on page 3.)

Bishop, C. (1994). Revisions and additions to stratification schemes used during research vessel surveys in NAFO Subareas 2 and 3. *NAFO SCR Doc 94*(43), 10. (Cited on page 35.)

Blackhart, K., D. G. Stanton, and A. Shimada (2006). *NOAA fisheries glossary*. United States Department of Commerce, National Oceanic and Atmospheric. (Cited on page 2.)

Bowering, W. (1981). Witch Flounder in NAFO Divisions 2J and 3KL. *Canadian Atlantic Fisheries Scientific Advisory Committee*. (Cited on page 31.)

Bowering, W. (1987). Distribution of Witch Flounder in southern Labrador and eastern Newfoundland area and changes in certain biological parameters after 20 years of exploitation. *Fishery Bulletin 85*(3). (Cited on page 4.)

Bowering, W. (1990). The Witch Flounder. *Communications Directorate Department of Fisheries and Oceans* (12). (Cited on pages 4 and 118.)

Brock, G., V. Pihur, S. Datta, and S. Datta (2008). clValid: An R package for cluster validation. *Journal of Statistical Software 25*(4), 1–22. (Cited on page 65.)

Brodie, W. (2005). A description of the fall multispecies surveys in SA2+Divisions 3KLMNO from 1995-2004. *NAFO SCR Doc 5*(008). (Cited on page 7.)

Cadigan, N. G. (2011). Confidence intervals for trawl-able abundance from stratified-random bottom trawl surveys. *Canadian Journal of Fisheries and Aquatic Sciences 68*(5), 781–794. (Cited on pages 5, 31, 33 and 57.)

Cao, J., J. T. Thorson, R. A. Richards, and Y. Chen (2017). Spatiotemporal index standardization improves the stock assessment of northern shrimp in the Gulf of Maine. *Canadian Journal of Fisheries and Aquatic Sciences 74*(11), 1781–1793. (Cited on page 10.)

Chen, J., M. E. Thompson, and C. Wu (2004). Estimation of fish abundance indices based on scientific research trawl surveys. *Biometrics 60*(1), 116–123. (Cited on pages 7, 49 and 50.)

Cochran, W. G. (2007). *Sampling techniques*. John Wiley & Sons. (Cited on page 6.)

Cooper, A. B. (2006). *A guide to fisheries stock assessment: from data to recommendations*. University of New Hampshire, Sea Grant College Program. (Cited on page 2.)

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological) 39*(1), 1–22. (Cited on page 118.)

Dennis, D., É. Plagányi, I. Van Putten, T. Hutton, and S. Pascoe (2015). Cost benefit of fishery-independent surveys: Are they worth the money. *Marine Policy 58*, 108–115. (Cited on page 3.)

DFO (2016a). Assessment of redfish stocks (*Sebastes mentella and S. fasciatus*) in Units 1 and 2 in 2015. *Can. Sci. Advis. Sec* (047). (Cited on page 2.)

DFO (2016b). Assessment of Atlantic cod (*Gadus morhua*) in the southern Gulf of St. Lawrence (NAFO Div. 4T-4Vn (Nov. – April)) to 2014. *DFO Can. Sci. Advis. Sec. Sci. Advis. Rep* (2015/061). (Cited on page 2.)

DFO (2017). Seafood industry year in review. (Cited on page 1.)

DFO (2018). Canada's fisheries fast facts. (Cited on page 1.)

DFO (2019). Stock assessment of Witch Flounder (*Glyptocephalus cynoglossus*) in NAFO Divisions 2J3KL. *DFO Can. Sci. Advis. Sec. Sci. Advis. Rep 053.* (Cited on pages 31, 35 and 115.)

Eddelbuettel, D. and J. J. Balamuta (2017, aug). Extending R with C++: A Brief Introduction to Rcpp. *PeerJ Preprints 5*, e3188v1. (Cited on page 53.)

FAO (2017). *FAO yearbook.* Fishery and Aquaculture Statistics. (Cited on page 1.)

Fournier, D. A., H. J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M. N. Maunder, A. Nielsen, and J. Sibert (2012). Ad Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software 27*(2), 233–249. (Cited on page 53.)

Gagniuc, P. A. (2017). *Markov chains: from theory to implementation and experimentation.* John Wiley & Sons. (Cited on page 55.)

Grosslein, M. (1969). *Groundfish survey methods.* Bureau of Commercial Fisheries, Biological Laboratory. (Cited on page 6.)

Hartig, F. (2020). *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models.* R package version 0.3.1. (Cited on page 66.)

Healey, B. and K. Dwyer (2005). A simple examination of canadian autumn survey trends in NAFO Divisions 3LNO for Greenland halibut and American plaice: the impact of incomplete coverage of this survey in 2004. (Cited on page 7.)

Hogg, R. V., J. McKean, and A. T. Craig (2005). *Introduction to mathematical statistics.* Pearson Education. (Cited on page 54.)

Horvitz, D. G. and D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association 47*(260), 663–685. (Cited on page 49.)

ICES, C. (2015). Report of the benchmark workshop on arctic stocks (WKARCT). (Cited on page 7.)

Kimura, D. K. and D. A. Somerton (2006). Review of statistical aspects of survey sampling for marine fisheries. *Reviews in Fisheries Science 14*(3), 245–283. (Cited on pages 4 and 5.)

Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. M. Bell (2016). TMB: Automatic differentiation and Laplace approximation. *Journal of Statistical Software 70*(5), 1–21. (Cited on page 53.)

Lewy, P. and K. Kristensen (2009). Modelling the distribution of fish accounting for spatial correlation and overdispersion. *Canadian Journal of Fisheries and Aquatic Sciences 66*(10), 1809–1820. (Cited on page 9.)

Lilly, G., J. Brattey, N. G. Cadigan, B. Healey, and E. Murphy (2005). An assessment of the cod (*Gadus morhua*) stock in NAFO Divisions 2J3KL in March 2005. (Cited on page 9.)

Lilly, G. R., P. Shelton, J. Brattey, N. Cadigan, E. Murphy, D. Stansbury, M. Davis, and M. Morgan (1998). An assessment of the cod stock inNAFO Divisions 2J+ 3KL. *Sci. Counc. Res. Doc. NAFO* (98), 94. (Cited on page 9.)

Magnusson, A., H. Skaug, A. Nielsen, C. Berg, K. Kristensen, M. Maechler, K. van Bentham, B. Bolker, M. Brooks, and M. M. Brooks (2017). Package 'glmmTMB'. *R Package Version 0.2. 0*. (Cited on pages 54 and 60.)

Mccllagh, P. and J. Nelder (1989). Generalized linear models. (Cited on page 52.)

McDonald, A. D. and A. D. Smith (1997). A tutorial on evaluating expected returns from research for fishery management using bayes'theorem. *Natural Resource Modeling 10*(3), 185–216. (Cited on page 3.)

McDonald, A. D., A. D. Smith, A. E. Punt, G. N. Tuck, and A. J. Davidson (1997). Empirical evaluation of expected returns from research on stock structure for determination of total allowable catch. *Natural Resource Modeling 10*(1), 3–29. (Not cited.)

Morrongiello, J. R. and R. E. Thresher (2015). A statistical framework to explore on to genetic growth variation among individuals and populations: a marine fish example. *Ecological Monographs 85*(1), 93–115. (Cited on page 119.)

NAFO (2020). Conservation and enforcement measures 2020. *NAFO 20-01*. (Cited on page 3.)

Nielsen, J. R., K. Kristensen, P. Lewy, and F. Bastardie (2014). A statistical model for estimation of fish density including correlation in size, space, time and between species from research survey data. *PLoS One 9*(6), e99151. (Cited on page 9.)

Paradis, E. and K. Schliep (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics 35*, 526–528. (Cited on page 68.)

Petersen, C. (1892). On the eggs and breeding of our Gobiidae. *Report of the Danish Biological Station 2*. (Cited on page 118.)

Punt, A. E., T. I. Walker, and J. D. Prince (2002). Assessing the management-related benefits of fixed-station fishery-independent surveys in Australia's Southern Shark Fishery. *Fisheries Research 55*(1-3), 281–295. (Cited on page 3.)

Quinn, T. J. and R. B. Deriso (1999). *Quantitative fish dynamics*. oxford university Press. (Cited on page 119.)

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. (Cited on pages 36 and 53.)

Rideout, R. and D. Ings (2018). Temporal and spatial coverage of Canadian (Newfoundland and Labrador region) spring and autumn multi-species RV bottom trawl surveys, with an emphasis on surveys conducted in 2017. *NAFO SCR Doc 18*, 017. (Cited on page 6.)

Rue, H. and L. Held (2005). *Gaussian Markov random fields: theory and applications*. CRC press. (Cited on pages 55 and 57.)

Sakamoto, Y., M. Ishiguro, and G. Kitagawa (1986). Akaike information criterion statistics. *Dordrecht, The Netherlands: D. Reidel 81*. (Cited on page 53.)

Särndal, C.-E., B. Swensson, and J. Wretman (2003). *Model assisted survey sampling*. Springer Science & Business Media. (Cited on page 49.)

Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics 6*(2), 461–464. (Cited on page 53.)

Shelton, A. O., J. T. Thorson, E. J. Ward, and B. E. Feist (2014). Spatial semiparametric models improve estimates of species abundance and distribution. *Canadian Journal of Fisheries and Aquatic Sciences 71*(11), 1655–1666. (Cited on pages 9, 10 and 51.)

Smith, S. J. (1990). Use of statistical models for the estimation of abundance from groundfish trawl survey data. *Canadian Journal of Fisheries and Aquatic Sciences 47*(5), 894–903. (Cited on pages 33, 48, 50 and 51.)

Thompson, S. K. (2012). *Simple random sampling*. John Wiley & Sons, Inc. (Cited on pages 6, 32, 33, 49, 50 and 51.)

Thorson, J. T. and K. Kristensen (2016). Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. *Fisheries Research 175*, 66–74. (Cited on page 63.)

Thorson, J. T. and C. Minto (2015). Mixed effects: a unifying framework for statistical modelling in fisheries biology. *ICES Journal of Marine Science 72*(5), 1245–1256. (Cited on page 51.)

Thorson, J. T., A. O. Shelton, E. J. Ward, and H. J. Skaug (2015). Geostatistical delta-generalized linear mixed models improve precision for estimated abundance indices for West Coast groundfishes. *ICES Journal of Marine Science 72*(5), 1297–1310. (Cited on page 10.)

Von Bertalanffy, L. (1938). A quantitative theory of organic growth (inquiries on growth laws. ii). *Human biology 10*(2), 181–213. (Cited on page 119.)

Warren, W., W. Brodie, D. Stansbury, S. Walsh, J. Morgan, and D. Orr (1997). Analysis of the 1996 comparative fishing trial between the Alfred Needler with the Engel 145 trawl and the Wilfred Templeman with the Campelen 1800 trawl. *NAFO SCR Doc 97*, 68. (Cited on page 60.)

Warren, W. G. (1996). Report on the comparative fishing trial between the Gadus Atlantica and Teleost. *NAFO SCR Doc 96*, 28. (Cited on page 60.)

Ye, Y. (2011). Assessment methodology. *FAO Fisheries and Aquaculture Circular*. (Cited on page 2.)

Zeileis, A. and T. Hothorn (2002). Diagnostic checking in regression relationships. *R News 2*(3), 7–10. (Cited on page 68.)

Appendices

Appendix A

R and C++ codes for the first Model

```
###### TMB codes for the first Model
compile("ts_strap.cpp")
dyn.load("ts_strap")
  parms <- list(
  log_k=0,
  logit_strat_rho=0,
  log_strat_sd=-2,
  log_strat_mean = log(unlist(tapply(tmb.data$y,tmb.data$ist,mean))),
  log_strat_dev = matrix(0,tmb.data$nstrat,tmb.data$nyears)
)
parms.L <- list(
  log_k=-2,
  logit_strat_rho=-10,
  log_strat_sd=-5,
  log_strat_mean = rep(-Inf,length(parms$log_strat_mean))
)
 parms.U <- list(
  log_k=2,
  logit_strat_rho=10,
```

```
  log_strat_sd=5,
  log_strat_mean = rep(Inf,length(parms$log_strat_mean))
)
lower = unlist(parms.L)
upper = unlist(parms.U)
  rname = c("log_strat_dev")
obj <- MakeADFun(tmb.data,parms,random=rname,DLL="ts_strap",
  inner.control = list(trace=TRUE),
)
length(lower)
length(upper)
length(obj$par)
obj$gr(obj$par)
system.time(
opt<-nlminb(obj$par,obj$fn,obj$gr,lower=lower,upper=upper,
      control = list(iter.max=1000,eval.max=1000))
)
rep=obj$report()
sd.rep<-sdreport(obj)


###### C++ codes for the first Model
template<class Type>
Type objective_function<Type>::operator() ()
{
  DATA_INTEGER(nyears);
  DATA_INTEGER(nstrat);
  DATA_INTEGER(nstrat_full);
```

```
  DATA_VECTOR(y);
  DATA_VECTOR(offset);
  DATA_IVECTOR(iy);
  DATA_IVECTOR(ist);
  DATA_VECTOR(area);
  DATA_IVECTOR(ist_full_est);
  DATA_IVECTOR(idiv);
  DATA_MATRIX(total_area_div);
  PARAMETER(log_k);
  PARAMETER(logit_strat_rho);
  PARAMETER(log_strat_sd);
  PARAMETER_VECTOR(log_strat_mean);
  PARAMETER_MATRIX(log_strat_dev);
  using namespace density;
//  using namespace Eigen;
  int nobs = y.size();
  Type k = exp(log_k);
  Type strat_rho = exp(logit_strat_rho)/(1.0 +
  exp(logit_strat_rho));
  Type strat_sd = exp(log_strat_sd);
  Type nll = 0.0;
  vector<Type> log_mu(nobs);
  for(int i=0; i<nobs; i++){
    log_mu(i) = log_strat_mean(ist(i)) +
    log_strat_dev(ist(i),iy(i)) + offset(i);
  }
  vector<Type> mu = exp(log_mu);
  vector<Type> vary = mu + mu*mu/k;
```

```
// negative loglikelihood for y;
  nll -= sum(dnbinom2(y, mu, vary,true));
// nll for log_strat_dev;
  vector<Type> temp(nyears);
  for(int i = 0;i < nstrat;++i){
    for(int j = 0;j < nyears;++j){
      temp(j) = log_strat_dev(i,j);
    }
    nll += SCALE(AR1(strat_rho),strat_sd)(temp);
  }
  vector<Type> resid = y-mu;
  vector<Type> std_resid = resid/sqrt(vary);
// create mnpt for each div and then all divisions;
matrix<Type> strat_eff(nstrat_full,nyears);
  for(int i=0;i<nstrat_full;i++){
    for(int j=0;j<nyears;j++){
      strat_eff(i,j)=0.0;
    }
  }
  for(int i=0;i<nstrat;i++){
    for(int j=0;j<nyears;j++){
      strat_eff(ist_full_est(i),j)= exp(log_strat_mean(i) +
      log_strat_dev(i,j));
    }
  }
  matrix<Type> num(nyears,4);
  for(int i=0;i<4;i++){
    for(int j=0;j<nyears;j++){num(j,i)=0.0;}}
```

```
    for(int i=0;i<nstrat_full;i++){
    for(int j=0;j<nyears;j++){
      num(j,idiv(i)) += area(i)*strat_eff(i,j);
      num(j,3) += area(i)*strat_eff(i,j);
  }}
  matrix<Type> mnpt(nyears,4);
  matrix<Type> log_mnpt(nyears,4);
  for(int i=0;i<4;i++){
    for(int j=0;j<nyears;j++){
       mnpt(j,i) =  num(j,i)/total_area_div(j,i);
       log_mnpt(j,i) =  log(mnpt(j,i));
  }}
  REPORT(log_mu);
  REPORT(mu);
  REPORT(resid);
  REPORT(std_resid);
  REPORT(strat_eff);
  REPORT(num);
  REPORT(mnpt);
  REPORT(log_mnpt);
  REPORT(strat_rho);
  REPORT(strat_sd);
  REPORT(log_strat_mean);
  REPORT(log_strat_dev);
  REPORT(k);
  ADREPORT(log_mnpt);
  return nll;
}
```

Appendix B

R and C++ codes for the selected Model

```
#### R(TMB) codes for the fifth Model
compile("ts_strap_new2.cpp")
dyn.load("ts_strap_new2")
LS_mean = as.vector(log(unlist(tapply(tmb.data$y,
tmb.data$ist,mean))))
parms <- list(
  LS_main=mean(LS_mean),
  log_k=0,
  logit_YE_rho=0,
  log_YE_sd=-2,
  logit_LS_rho=0,
  log_q = 0,
  log_d = -2,
  log_q_dev = 0,
  log_d_dev = -2,
  LS_eff = LS_mean - mean(LS_mean),
  log_YE = rep(0,tmb.data$nyears),
  LS_year_dev = matrix(0,tmb.data$nstrat,
  tmb.data$nyears)
```

```
)
parms.L <- list(
  LS_main=-Inf,
  log_k=-2,
  #  logit_YE_rho=-10,
  #  log_YE_sd=-5,
  logit_LS_rho=-10,
  log_q = -Inf,
  log_d = -Inf,
  log_q_dev = -Inf,
  log_d_dev = -Inf
)
parms.U <- list(
  LS_main=Inf,
  log_k=5,
  #  logit_YE_rho=10,
  #  log_YE_sd=5,
  logit_LS_rho=10,
  log_q = Inf,
  log_d = Inf,
  log_q_dev = Inf,
  log_d_dev = Inf
)
lower = unlist(parms.L)
upper = unlist(parms.U)
## map to take out some parameters
map = list(logit_YE_rho=factor(NA),log_YE_sd=factor(NA),
           log_YE = factor(rep(NA,tmb.data$nyears))
```

```
)
rname = c("LS_eff","LS_year_dev")
## turn off nll for log_YE;
tmb.data$est_ye=0
obj <- MakeADFun(tmb.data,parms,map=map,random=rname,DLL="ts_strap_new2",
                 inner.control = list(trace=TRUE),
)
rep=obj$report()
length(lower)
length(upper)
length(obj$par)
obj$gr(obj$par)
system.time(
  opt.M5Sampledn<-nlminb(obj$par,obj$fn,obj$gr,lower=lower,upper=upper,
                         control = list(iter.max=1000,eval.max=1000))
)
######
opt.M5Sampledn$AIC = 2*opt.M5Sampledn$objective + 2*length(opt.M5Sampledn$par)
opt.M5Sampledn$BIC = 2*opt.M5Sampledn$objective +
log(length(tmb.data$y))*length(opt.M5Sampledn$par)
rep.M5Sampledn=obj$report()
rownames(rep.M5Sampledn$mnpt) = sort(unique(subd$survey.year))
colnames(rep.M5Sampledn$mnpt) = c('2J','3K','3L','2J3KL')
sd.rep.M5Sampledn<-sdreport(obj)


############## C++ codes for the fifth Model
template<class Type>
```

```
Type objective_function<Type>::operator() ()
{
  DATA_INTEGER(nyears);
  DATA_INTEGER(nstrat);
  DATA_INTEGER(nstrat_full);
  DATA_VECTOR(y);
  DATA_VECTOR(offset);
  DATA_IVECTOR(iy);
  DATA_IVECTOR(ist);
  DATA_VECTOR(area);
  DATA_IVECTOR(ist_full_est);
  DATA_IVECTOR(idiv);
  DATA_MATRIX(samplset);
  DATA_MATRIX(total_area_div);
  DATA_MATRIX(total_areaSampled_div);
// matrices for GMRFs;
  DATA_SPARSE_MATRIX(Dm);
  DATA_SPARSE_MATRIX(Im);
  DATA_INTEGER(est_ye);
// note LS indicated log_stratum
  PARAMETER(LS_main);
  PARAMETER(log_k);
  PARAMETER(logit_YE_rho);
  PARAMETER(log_YE_sd);
  PARAMETER(logit_LS_rho);
  PARAMETER(log_q);
  PARAMETER(log_d);
  PARAMETER(log_q_dev);
```

```
  PARAMETER(log_d_dev);
// random effects below
  PARAMETER_VECTOR(LS_eff);
  PARAMETER_VECTOR(log_YE);
  PARAMETER_ARRAY(LS_year_dev);
  using namespace density;
  using namespace Eigen;
  int nobs = y.size();
  Type k = exp(log_k);
  Type LS_rho = exp(logit_LS_rho)/(1.0 + exp(logit_LS_rho));
  Type YE_rho = exp(logit_YE_rho)/(1.0 + exp(logit_YE_rho));
  Type YE_sd = exp(log_YE_sd);
  Type nll = 0.0;
  // The precision matrix for LS_eff;
  Eigen::SparseMatrix<Type> LS_eff_PM = exp(log_q)*Dm + exp(log_q + log_d)*Im;
  Eigen::SparseMatrix<Type> LS_dev_PM = exp(log_q_dev)*Dm +
  exp(log_q_dev + log_d_dev)*Im;
  vector<Type> log_mu(nobs);
  for(int i=0; i<nobs; i++){
    log_mu(i) = LS_main + LS_eff(ist(i)) +
    log_YE(iy(i)) + LS_year_dev(ist(i),iy(i)) + offset(i);
  }
  vector<Type> mu = exp(log_mu);
  vector<Type> vary = mu + mu*mu/k;
// negative loglikelihood for y;
  nll -= sum(dnbinom2(y, mu, vary,true));
// nll for log_YE;
  if(est_ye==1){nll += SCALE(AR1(YE_rho),YE_sd)(log_YE);}
```

```
// nll for LS_eff;
  nll += GMRF(LS_eff_PM)(LS_eff);
// nll for LS_year_dev;
  nll += SEPARABLE(AR1(LS_rho),GMRF(LS_dev_PM))(LS_year_dev);
  vector<Type> resid = y-mu;
  vector<Type> std_resid = resid/sqrt(vary);
  // create mnpt for each div and then all divisions;
  matrix<Type> strat_eff(nstrat_full,nyears);
  strat_eff.setZero();
  for(int i=0;i<nstrat;i++){
    for(int j=0;j<nyears;j++){
      strat_eff(ist_full_est(i),j)= exp(LS_main +
      LS_eff(i) + log_YE(j) + LS_year_dev(i,j));
    }}
  matrix<Type> num(nyears,4);
  num.setZero();
  for(int i=0;i<nstrat_full;i++){
    for(int j=0;j<nyears;j++){
      num(j,idiv(i)) += area(i)*strat_eff(i,j);
      num(j,3) += area(i)*strat_eff(i,j);
  }}
  matrix<Type> mnpt(nyears,4);
  matrix<Type> log_mnpt(nyears,4);
  for(int i=0;i<4;i++){
    for(int j=0;j<nyears;j++){
       mnpt(j,i) =  num(j,i)/total_area_div(j,i);
       log_mnpt(j,i) =  log(mnpt(j,i));
  }}
```

```
matrix<Type> sample(nstrat_full,nyears);
sample.setZero();
for(int i=0;i<nstrat_full;i++){
  for(int j=0;j<nyears;j++){
    if (samplset(i,j)==1){sample(i,j)=strat_eff(i,j);}
  }}
num.setZero();
for(int i=0;i<nstrat_full;i++){
  for(int j=0;j<nyears;j++){
    num(j,idiv(i)) += area(i)*sample(i,j);
    num(j,3) += area(i)*sample(i,j);
  }}
matrix<Type> mnpt_sampled(nyears,4);
matrix<Type> log_mnpt_sampled(nyears,4);
log_mnpt_sampled.setZero();
mnpt_sampled.setZero();
for(int i=0;i<4;i++){
  for(int j=0;j<nyears;j++){
    Type tot_area = total_areaSampled_div(j,i);
    if(tot_area>0){
      mnpt_sampled(j,i) =  num(j,i)/tot_area;
      log_mnpt_sampled(j,i) =  log(mnpt_sampled(j,i));}
    if(tot_area==0){
      mnpt_sampled(j,i) =  -9.0;
      log_mnpt_sampled(j,i) =  -9.0;}
  }}
REPORT(log_mu);
REPORT(mu);
```

```
  REPORT(resid);
  REPORT(std_resid);
  REPORT(strat_eff);
  REPORT(sample);
  REPORT(num);
  REPORT(mnpt);
  REPORT(log_mnpt);
  REPORT(mnpt_sampled);
  REPORT(log_mnpt_sampled);
  REPORT(LS_rho);
  REPORT(LS_eff);
  REPORT(LS_year_dev);
  REPORT(log_YE);
  REPORT(k);
  REPORT(LS_eff_PM);
  REPORT(LS_dev_PM);
  ADREPORT(log_mnpt);
  ADREPORT(log_mnpt_sampled);
  return nll;
}
```