Seasonal and Long term variability in species distribution of Atlantic cod (*Gadus morhua*) and Yellowtail Flounder (*Limanda ferruginea*) on Georges Bank

Keith D.M.a,b, Sameoto J.A.a, Keyser F.M.a, Ward-Paige C.A.c, Andrushchenko I.^

a Bedford Institue of Oceanography, b Dalhousie University, c eOceans

Sustainably managing marine fisheries has long been recognized as a global priority which has proven difficult to achieve. The reasons sustainable fisheries management goals have not been achieved include various socio-economic, political, and scientific factors. Scientifically, one of the major challenges has been understanding how environmental factors influence both the spatial and temporal dynamics of a stock. Fisheries science has spent a great deal of effort collecting data, both biological and environmental, which are inherently spatial and temporal in nature. Computational and statistical limitations have resulted in science products which do not fully utilize the spatio-temporal information contained in these data and tend to treat a stock as a homogeneous entities. Fortunately, computational advances coupled with more accessible statistical methods has resulted in new methodologies which can harness the spatio-temporal information contained in these fisheries data. Here we develop temporally variable species distribution models for yellowtail flounder (*Limanda ferruginea*) and Atlantic cod (*Gadus morhua*) on Georges Bank using a suite of static environmental covariates and presence-absence information from groundfish trawl surveys in Canada and the United States. These models indicate there are both seasonal and long term shifts in the distribution of both species. The average sea surface temperature (SST; average between XXXX and YYYY) and depth were significant predictors of the distribution of both species throughout the year. Significant shifts in the distribution of both species occurs relatively frequently, with the distribution of cod observed to differ approximately every 5 years, while the Yellowtail distribution appears to fluctuate every 3 years. These shifts in distribution are not random, with the center of gravity of the core areas for both species shifting to the north and east throughout the study period. Much of this shift is due to the loss of the species from southern and western portions of Georges Bank. Models for both species were also relatively successful at capturing the spatial dynamics of the stock up to 3 years into the future. The seasonal distribution of cod and yellowtail are relatively consistent throughout the late winter and spring. In the fall the distribution of cod shifts towards the edge of the bank, while it appears the catchability of yellowtail declines during the this time. Here we show how these models are able to provide novel insights into both seasonal and inter-annual variability in species distributions and identify environmental covariates which have a relationship with the distribution of these species. Incorporation of this kind of information into stock assessment processes will improve science advice and our ability to sustainably manage these stocks.

# Introduction

Sustainable management of marine fisheries has been recognized as a critical challenge facing society in the 21st century [@cbdAichiBiodiversityTargets2018]. The challenges facing sustainable fisheries management are multifaceted and include complex socio-economic, political, and scientific interactions (CITE). Environmental and ecological research has been at the heart of fisheries science for well over a century (CITE). From the early days of fisheries science it was recognized that an inability to fully account for spatial processes was potentially a serious issues to surmount (CITE). Many of the traditional fisheries methods developed, and still currently used to assess fisheries, required strong assumptions about the underlying spatial processes; during the development of these methods these assumptions were often identified as potentially problematic (CITE Ricker, BH, and maybe Carl and Ray’s book). A number of methods had subsequently been developed to attempt to account for spatial processes, but computational limitations had restricted the complexity of these models (CITE - Allowing for multiple fleets was one of these). In more recent years a flurry of computational and statistical advances have enabled the development of models in which spatial processes can be more rigorously addressed within these traditional fisheries modelling frameworks (CITE Anders, Kasper, Jim, Yihao, and Raph if ready too).)

One of the earliest modelling frameworks developed to explicitly account for spatial patterns and processes were species distribution models with (SDMs; CITE early SDM work both terrestrial and fisheries). These models use environmental data along information about the species to map the likelihood of encountering a species across some land(sea)-scape and originated with attempts to map terrestrial plant distributions (CITE). In the marine realm SDM’s have been used to assist with the development of Marine Protect Areas (MPAs), MPA networks, Species at Risk (SAR), name some other shit (CITE). In many cases SDM’s assume that there is no temporal change in the relationship between the environment cues and the response of the species to these cues; these SDM’s essentially provide a snapshot in time based on available data. To better predict future states more sophisticated SDM’s frameworks have been developed in which the underlying relationships can vary both in time and space (CITE).

*I need a segue paragraph which I’m not sure what it looks like, maybe something like this?*

The collection of data in fisheries science, both biological and environmental, is often inherently spatial and temporal in nature. In the past, computational and statistical limitations have resulted in science products which do not fully utilize the spatio-temporal information contained in these data, for example, SDM applications often average temporally where the year in which the biological data is collected is not accounted for (CITE), while stock assessment methods aggregate information spatially and treat stocks as homogeneous entities (dynamic pools?? CITE). Fortunately, the aforementioned computational advances coupled with more accessible statistical methods has resulted in new methodologies which can better harness the spatio-temporal information inherent in the data; this is especially relevant in regions in which there is a long history of high quality environmental, ecological, and fishery data collection.

Georges Bank has been home to some of the most productive fisheries in the world for centuries and is home to a wealth of natural resources. For this reason this area has been host to intensive scientific research programs for decades (CITE SOME STUFF). This research has resulted in a wealth of scientific knowledge about the geology, environment, and ecology of this area (CITE SOME STUFF). Historically numerous countries had large fisheries in the region but with the expansion of territorial seas to 200 miles offshore in the late 1970’s, resource exploitation (e.g. fisheries) on Georges Bank fell under the jurisdiction of Canada and the United States (CITE NAFO something). The final demarcation of the Canadian and United States territorial waters on Georges Bank were implemented with an International Court of Justice decision in 1984 (CITE). Within three years of this decision both countries had developed their own independent groundfish surveys which each covered the entirety of the bank at different times of the year.

Historically, Georges bank was home to large groundfish fisheries including Atlantic cod (*Gadus morhua*), Altantic haddock (*Melanogrammus aeglefinus*), Yellowtail flounder (*Limanda ferruginea*) and numerous other species (CITE). As observed throughout the northwest Atlantic, the biomass of Atlantic cod on Georges Bank declined significantly in early 1990’s and there have been little evidence for recovery of this stock since this collapse (CITE TRAC). Yellowtail flounder on Geroges Bank had been at low abundnace on the bank since the 1970’s, but evidence for a rapid recovery of this species was observed in the early 2000’s and resulted in directed fisheries for this species for several years, unfortunately this recovery was short lived and the abundance of this species has been near historical lows for the last decade (CITE). While the status of these two groundfish species remains bleak, other groundfish stocks (e.g. Atlantic haddock) have experienced large increases in biomass over the last decade, and currently the most lucrative fishery on Georges Bank has been an invertebrate sea scallop fishery which has experienced unprecedented productivity over the last two decades (CITE SABHU and DVORA).

Fisheries management bodies in both Canada and the United States have implemented measures to protect cod and yellowtail on Georges Bank. While these measures vary between the countries, both countries have collaborated to develop a shared bycatch quota for these two species; this quota has declined substantially for both species over the last decade (CITE). There has been no directed fishery for cod on Georges bank for either country since XXXX. In addition to these regulations which attempt to directly limit the removals of these two species, both countries have implemented closures. In the United States two large closed area were implemented (Closed Area I and II) in the 1990’s, these closures were designed aid in the recover of groundfish and invertebrate stocks on Georges Bank (CHECK RATIONALE FOR CAI and CAII, Murawoski paper likely has good review of it). In Canada the groundfish fishery is not permitted to fish on Georges Bank during cod spawning, generally this closure runs from early February to the end of May since XXXX (CHECK!). The Canadian Offshore scallop fishery also faces rescritions on fishing during the peak spawning periods for both cod and yellowtail although the effect of these closures appears to be limitied (CITE OUR FIRST PAPER).

Here we use a recently developed statistical framework (CITE R-INLA) to develop spatio-temporal species distribution models for two depleted groundfish stocks on Georges Bank (Atlantic cod and Yellowtail flounder). Our objectives were; 1) Use a suite of static environmental layers to determine whether any of these environmental data informed the distribution of either species, 2) determine whether the species distributions changed over time and if so how rapidly changes in the distributions could be observed, 3) determine whether the species distributions change seasonally using data from groundfish surveys in the winter, spring, and fall, 4) determine how well these models can predict the distribution of these species 1, 2, and 3 years into the future.

# Methods

### Study area

Georges Bank, located in the northwest Atlantic straddling the US-Canada maritime border, is a 3-150 m deep plateau that covers approximately 40,000 and is characterized by high primary productivity, and historically high fish abundance [@townsendNitrogenLimitationSecondary1997]. It is an eroding bank with no sediment recharge, and covered with coarse gravel and sand that provides important habitat for many species [@valentineSeaFloorEnvironment1991]. Since 1984, Georges Bank has been divided between the US and Canada and, while some collaborative management exists, the US and Canadian portions are largely managed separately (Figure 1).

### Data

Survey data were obtained from the Fisheries and Oceans Canada (DFO) winter RV survey from 1987-2019 and the National Marine Fisheries Service (NMFS) spring and fall groundfish surveys from 1972-2019. The DFO-winter survey on Georges Bank typically occurs in February and early March, the NMFS-spring survey typically occurs in April and May, while the NMFS-fall survey generally takes place between September and November. For all surveys only tows deemed *successful* were used in this analysis. This resulted in 2590 tows from the DFO-winter survey, 2393 tows from the NMFS-spring survey, and 2506 tows from the NMFS-fall survey.

### Environmental covariates

A suite of 21 environmental variables with spatial information were obtained for this analysis (Table XX). To eliminate redundant variables, variance Inflation Factors (VIFs) were calculated for all variables and any variables with VIF scores > 3 were removed. This procedure was repeated until no variables remained with a VIF score > 3 (CITE ZUUR). Using the remaining 16 variables a Principle Component Analysis (PCA) was undertaken for each survey using the data from station locations for each survey these environmental data with the top 4 PCA components retained (these accounted for at least 80% of the variability in the data) and were included as covariates for the models that follow.

### Statistical Analysis

A Bayesian hierarchical methodology was implemented using the Integrated Nested Laplace Approximation approach available within the R Statistical Programming software R-INLA (CITE R and INLA). In recent years R-INLA has seen a rapid increase in use to model species distributions in both within the terrestrial and marine realms (CITE SOME PAPERS). This methodology solves stochastic partial differential equations on a spatial triangulated mesh; this mesh is typically based on the data available (CITE RUE). To avoid edge effects the mesh is extended beyond the boundaries of the data, the mesh used in this study included 6610 vertices (Figure 2).

For the INLA models data up to 2016 were used, while survey data from 2017-2019 were excluded from the main analysis and used only as testing data. For all analyses the response variable was presence absence of the species of interest () and a *Bernoulli* GLM was utilized within R-INLA.

Each variable retained after the VIF analysis along with the 4 PCA components were added to the model individually, all continuous covariates were modelled using the INLA random walk smoother which allows for non-linear relationships between the response and each variable (Cite ZUUR Vol 1). The continuous covariates were centred at their mean value and scaled by their standard deviation, covariates which were highly skewed (e.g. depth) were log transformed before being standardized. Due to low sample size of several of the levels the Sediment number (Sed) were amalgamated into one factor level which was represented by the ‘intercept’ term in models which included the Sediment number. This amalgamated level represented approximately 7% of survey tows across the three surveys (approximately 93% of the survey tows were on Sediment type 3 or 4).

Four spatial random fields () were compared for each species and each survey, these included a) a static random field (t = 1), b) independent random fields every 10 years, c) independent random fields every 5 years, and d) and independent random fields every 3 years. For b-d the random fields were set from the most recent year, so that when the time series was not a multiple of the time series length the first years of data had a shorter duration random field (e.g. the 10 year random fields for NMFS-spring survey were 2007-2016, 1997-2006, 1987-1996, 1977-1986, and 1972-1976). Models with the same covariate structure but different random fields were compared using WAIC and DIC; the results for both metrics were similar and only the WAIC results are discussed further. In all cases the static spatial field was an inferior model when compared to models with multiple random fields and the results discussed here use the 10/5/3 year random fields.

Initial model selection for the different covariate models was undertaken using a static random field (due to computational constraints) by adding individual covariates. For this first analysis covariates were retained if low WAIC scores were observed across multiple models. For cod this analysis identified depth (DEP) and the average sea surface temperature between 1997 and 2008 (SST) as having low WAIC scores in 2 of the 3 surveys. For yellowtail, depth (DEP) was the primary covariate observed, in addition sediment grain size (SED), and the average chlorophyll concentration between 1997 and 2008 (CHL) were retained due to their low scores in one survey. These variables were added pairwise (e.g. models included SST + DEP, DEP + CHL, and SST + CHL) for both species and again compared using WAIC. For cod a three term model including additive terms for SST, DEP, and CHL was the most complex model tested, while for yellowtail the most complex model included SST, DEP, and SED. For this step additional covariates were retained if the WAIC for that model resulted in an improvement of the WAIC of more than 2 when compared to the lowest WAIC more parsimonious model.

### Model Validation and Prediction

Five fold cross validation was used to test the predictive performance of the models. The data were randomly divided into 5 subsets and trained using 4 of the subsets, the 5th dataset was treated as a testing dataset to determine how well the model was able to predict out of sample data. Model performance was measured by comparing the the model residuals from the training data to the prediction error from the testing data, the metrics used for this comparison were Root Mean Squared Error (RMSE), Mean Average Error (MAE), and the standard deviation (SD). For computational reasons the models compared using 5 fold cross validation were intercept only, DEP, DEP + SST, the 5 year random field was used for all model validation for both species.

In addition to the 5-fold cross validation the survey data from 2017-2019 was excluded from the models and was used as a testing dataset to determine how well the models were able to predict future species distributions and to quantify how the predictions of species distributions differ 1, 2, and 3 years after the final year of survey data. The predictive error was estimated using RMSE, MAE, and SD and this was compared to the residual error from each model.

# Results

### Model Selection

Initial model selection resulted in a significant reduction in the number of covariates in the model. For cod, the Winter (DFO) and Spring (NMFS) both identified SST as significant covariates, while the Spring survey also identified depth and stratification, the Fall (NMFS) survey did not indicate any covariates with an WAIC that were a significant improvement from the intercept only model, although again the inclusion of Depth did result in a slightly smaller WAIC (Figure @ref(fig:diag.1.fe). For yellowtail, inclusion of depth significantly improved the models in all 3 seasons (surveys), while Sediment grain size (Sed) and chlorophyll concentration in the Fall had a similar impact on the model WAIC as SST. As a result SST, Depth, Chl, and Sed were used to explore the development of more complex covariate models. For cod these more complex models resulted in an additive Dep + SST model being the preferred model in all 3 seasons (Figure @ref(fig:diag.2.fe). For yellowtail the best models with 2 covariates included some combination of Dep, SST, and Sed, further model selection indicated that the best model for yellowtail in all 3 seasons was an additive model including Dep, SST, and Sed (Figure @ref(fig:diag.2.fe and @ref(fig:diag.3.fe)). The cod the 5 year random field had the lowest WAIC in all seasons, whlie for yellowtail the 3 year field was prefered for Winter and Spring, while the 5 year field was preferred for Fall. (Figure @ref(fig:diag.rf))

### Environmental Variables

The depth relationship for cod (FIG: show cod depth for each survey) indicates…. The seasonal variation in this relationship…. (Figure X). The SST (FIG: show SST for every survey)… The seasonal variation in this…

For yellowtail say the same stuff

### Random Fields

For cod the preferred random field changes every 5 years (TABLE) and this was consistent for each of the 3 surveys.

For yellowtail the preferred model had a random field which changed every 3 years (TABLE) and this was consistent for each of the 3 surveys.

### Validation

The 5-fold cross validation indicated that the model was able to predict the distribution for all species and surveys without a significant loss of accuracy, the precision of the validation test set predictions was higher than the precision of the training data. In addition the prediction from the models using only the random fields was similar to the predictions from the models in which all covariates were used.

### Prediction

All of the models were able to predict the distribution of each species 1, 2, and 3 years into the future (RMSE and other stats). The 2018 Yellowtail NMFS-spring data had the lowest prediction accuracy, this is likely related to the unusually sparse coverage of this survey in the region in which the probability of encountering yellowtail was elevated (DASHBOARD?).

Inter-annual Variability

For both species their distributions shifted towards the north and east throughout the study period (COG + FIG). Much of this shift was driven by a decline in the probability of encounter in the western and southern portions of GB (FIG).

For cod the area in which the probability of encountering cod was high (> 75?) declined significantly throughout the study period with a rapid decline occurring in the early 1990s which has persisted thereafter (FIG or DASHBOARD). The cod population now is predominately found in the northeast corner of the bank, this becomes more pronounced during the fall which the highest probability of encounter is found along the edge of the northern portion of the survey area.

For yellowtail the population also shifted from the south and west portion of the bank and the population appeared to straddle the ICJ line separating Can/US. The size of the core area (PE >75%) has changed during the study period, with the largest areas of high probability of encounter observed in the early 2000s, the expansion of this core area occurred rapidly from a very small core area observed in the 1990s.

### Seasonal Variability

For cod the distribution of the species was similar for the DFO-winter (Feb-Mar) and NMFS-spring (April-May) surveys, while a clear shift in the distribution was observed for the NMFS-fall survey (Sept-Nov).

For yellowtail the distribution was observed to be similar for all three surveys with a clear core area consistently identified. In the fall the probability of encounter is generally lower that observed in the winter and spring despite the spatial distribution remaining similar (FIG, simple fig showing mean + SE for each survey from best model).

COMPARE THE DEPTH/SST relationships between surveys here, I believe they were very similar throughout but explore this here.

# Discussion

Here we have shown how models which incorporate both environmental and spatial information can be used to partition static environmental relationships from dynamic changes which occur both inter and intra-annually. This framework enables a better understanding of the magnitude of dynamical shifts along with identifying regions of consistently high and low probability of encounter throughout the study region.

If you want to protect something, knowing where it is going to be is kinda helpful. Implications for closures.

Cod is likely moving off the bank in the Fall now, implications for using the Fall as a survey index for cod.

Talk about YT protection efforts in Can/US and how the strategies don’t appear to be doing anything. The US has put in large protected areas, one of which is in the right spot (check the year that the closed area 1/2 was put in place), seemed to help initially, but population still declined after initial rebound.

Is yellowtail drop in encounter probability due to yellowtail being less susceptible to fishing gear during this period or due to a migration off the bank during this period?

Knowing where firsh are seasonally would surely hlpe then trying to manage incidental mortality for these stocks

### Yellowtail

On the U.S. side closures were put in place in 1994 to assest with the rebuilding of stocks in the region (CITE LINK 2005 maybe? SOME MORE DETAILS OF THESE CLOSURES MAY BE USEFUL). Closed Area II straddels the ICJ line and includes much of the area identified as core yellowtail habitat on on the U.S. side of Georges Bank. Intreguingly, the expansion of the yellowtail populaiton occured shortly after the implementation of this closure, and the expansion of the core yellowtail habitat in the early 2000’s was centered around this closed area. The expansion of the yellowtail core area corresponded to a rapid, yet ephemerial increase in yellowtail biomass. This might suggest evidence of a positivie assocaition between this closure and yellowtail status, but unfortunately the yellowtail population has subsequently declined and is near historical low levels on the bank and the core area of yellowtail is similar in size to what was observed before this closure was put in place.

The recent declines may be tied to environmental change on Georges Bank (CITE!) and given the loss of Yellowtail from the warmer portions of the bank (West and South) it is possible that the remaining core area represents the last suitable habitat on Georges Bank for this species. Yellowtail prefer depths of 40-70 meters (CITE + these results) and the core area identifed here represents the most northern region of these depths on Georges Bank. If temperatures continue to increase as projected (CITE: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0231595#sec024>) the suitability of this habitat may continue to decline and may effectively lead to the extirpation of yellowtail from Georges Bank irrespective of any fisheries management action.

Abundances of yellowtail are low throughout the entirety of the U.S. (CITE the 3 assessments here <https://www.fisheries.noaa.gov/species/yellowtail-flounder#overview>) which represents the southern limit of the historical ranage of this species.

On the Canadian side of Georges Bank there has been no directed yellowtail fishery since XXXX (Freya) and Georges Bank is close to groundfish fishery from March 1 to May 31 each year. These closures are implemented to protect cod and haddock spawning aggregations but likely provide some benefit for yellowail spawners (OBrien, spanwing is ?May-July?). A seperate closure of the scallop fishery in June of the year closes a small variable subset of the bank to the scallop fishery to protect spawning yellowtail from bycatch in this fishery. The effectiveness of the yellowtail closures in achieving their management objective are questionable in large part due to the small size and short duration of these closures; this analysis supports the conclusion that the size of these time-area closures of the scallop fishery are too small given the size of the core area on the Canadian side of Georges Bank (Cite me).

### Cod

The distribtion of cod has steadily shifted throughout the duration of the study period. The depth preference of cod is more variable than yellowtail (6-400 meeters ish, find a real citation) but as observed with yellowtial the core area in the more southern and western reaches of the bank have declined over the course of the study period. The core area for cod collapsed rapidly in the early 1990’s in unison with the collapse of the stock throughout the Northwest Atlantic. Since the collapse the core area has remained relatively consistent but has continued to slowly shift to the north and east throughout the year, though the shift is more pronouced in the fall.

The fall distribution of cod is likely now located on the northeastern slope of the bank outside of the core survey domains of any of the surveys. This northeastern shift of the population over the course of this study suggests that the surveys are no longer sampling the entirety of this population throughout the course of the year (i.e. a higher proportion of the stock is now located outside of the survey domain). Each of the survey indicies are used as inputs to the cod stock assessment model for eastern Georges Bank cod (CITE). This assessment model suffered from such significant retrostpective patterns that this stock assessment model was eventually rejected; it is likely that the observed shift in the distribution of cod outside of the survey domain contributed to the model retrospective problems (CITE)

(CITE Link 2005 Am Fish, not yet downloaded The Effects of Area Closures on Georges Bank)

### Closures (May all be above…)

Both the U.S. and Canadian jurisdications have iplemented closures to protect these species. CA II aligns with the observed yellowtail distribution.

On the Canadian side several seasonal closures are put in place to protect spawning groundfish aggregations. The Canadian Groundfish fleet on GB predominately target Atlantic haddock (*Melanogrammus aeglefinus*) and is subjected to a closure to protect cod and haddock spawning aggregations from March 1 to May 31. The other major Canadian fishery on Georges Bank is the Offshore Scallop Fishery; this fishery is also subjected to closures of variable size and location (). These closures occur in February and March (protecting cod spawning aggregations) and June (protecting yellowtail spawning aggregations). For closures of limited size and variable location a means of predicting the spatial likelihood of encounter of the species to be protected is required for the upcoming year such as what has been developed here. These results indicate that for both species the core spawning area is significantly larger than the area protected by the scallop fishery time-area closures and as such these closures would not be able to provide complete protection of these species spawning aggregations.

# Conclusions

These models provides insight into how the distribution of both species changes both seasonally and inter-annually. This insight can support the provision of improved science advice (e.g. stock assessment and protected areas) and sustainable fisheries management.

# Conclusions

## Acknowledgements

I’m not sure who to thank…

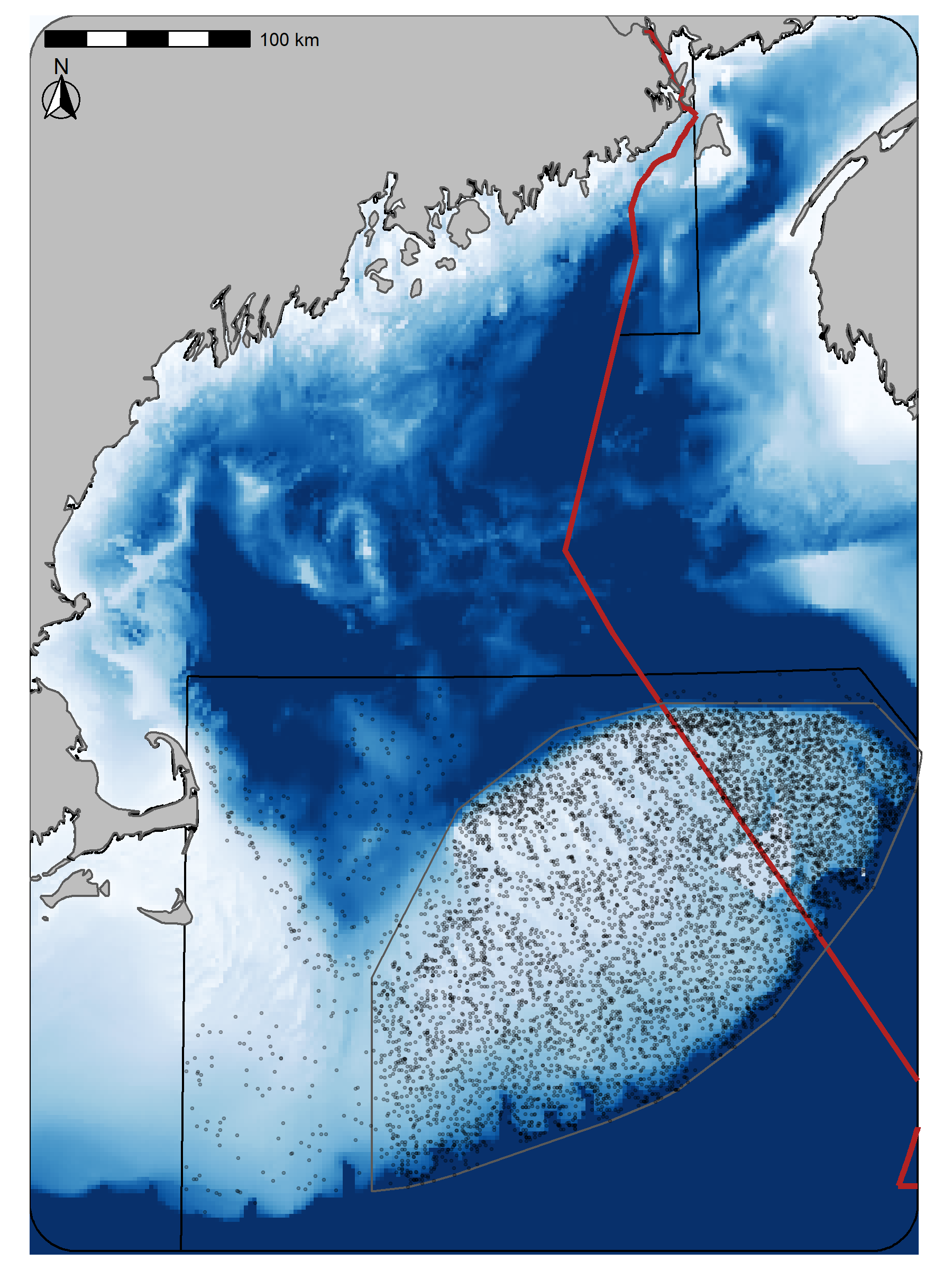
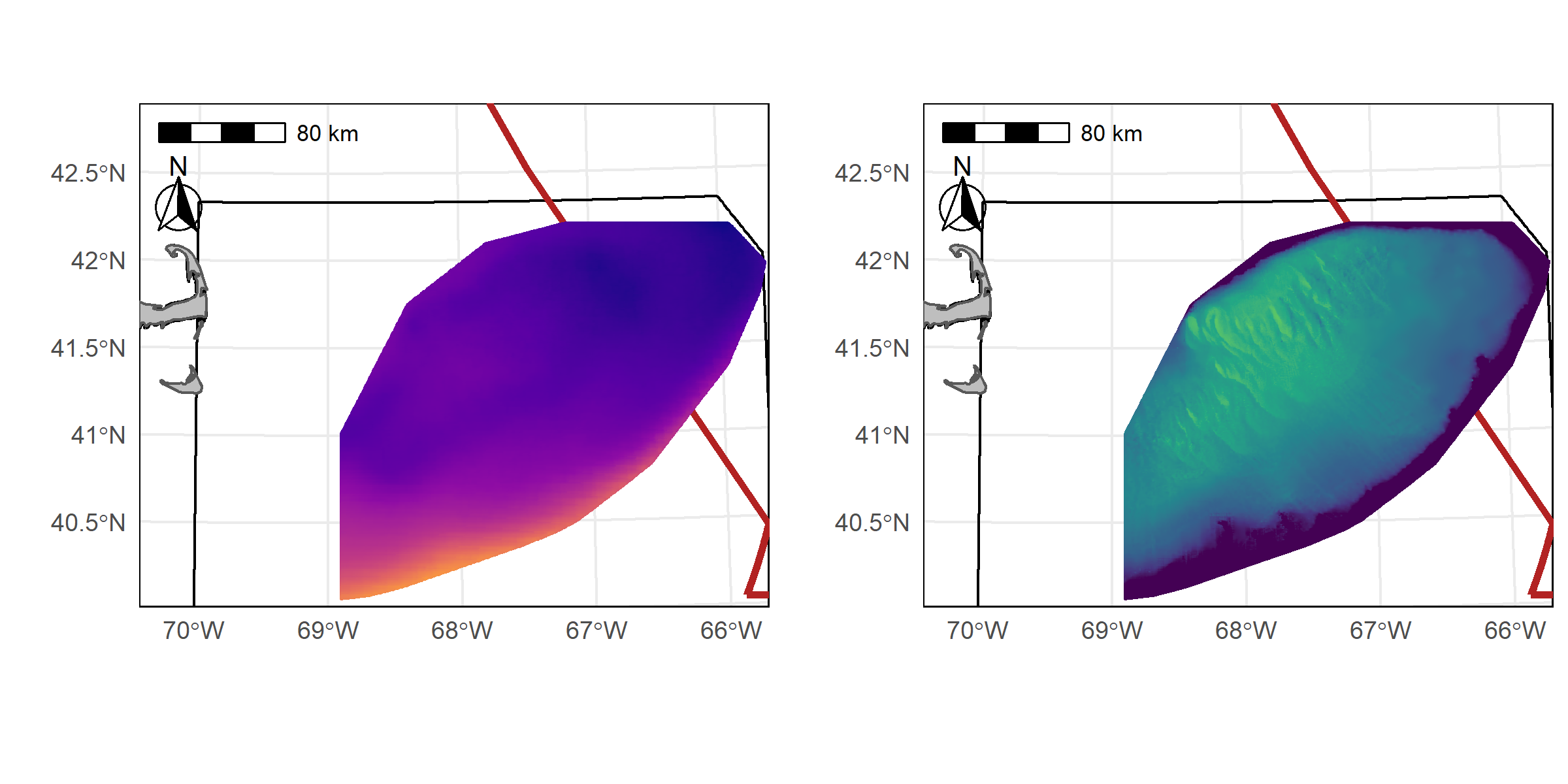


Figure 1: Georges Bank study area



(#fig:SST\_and\_Depth)SST (199X - 20XX average field) and Depth fields on Georges Bank

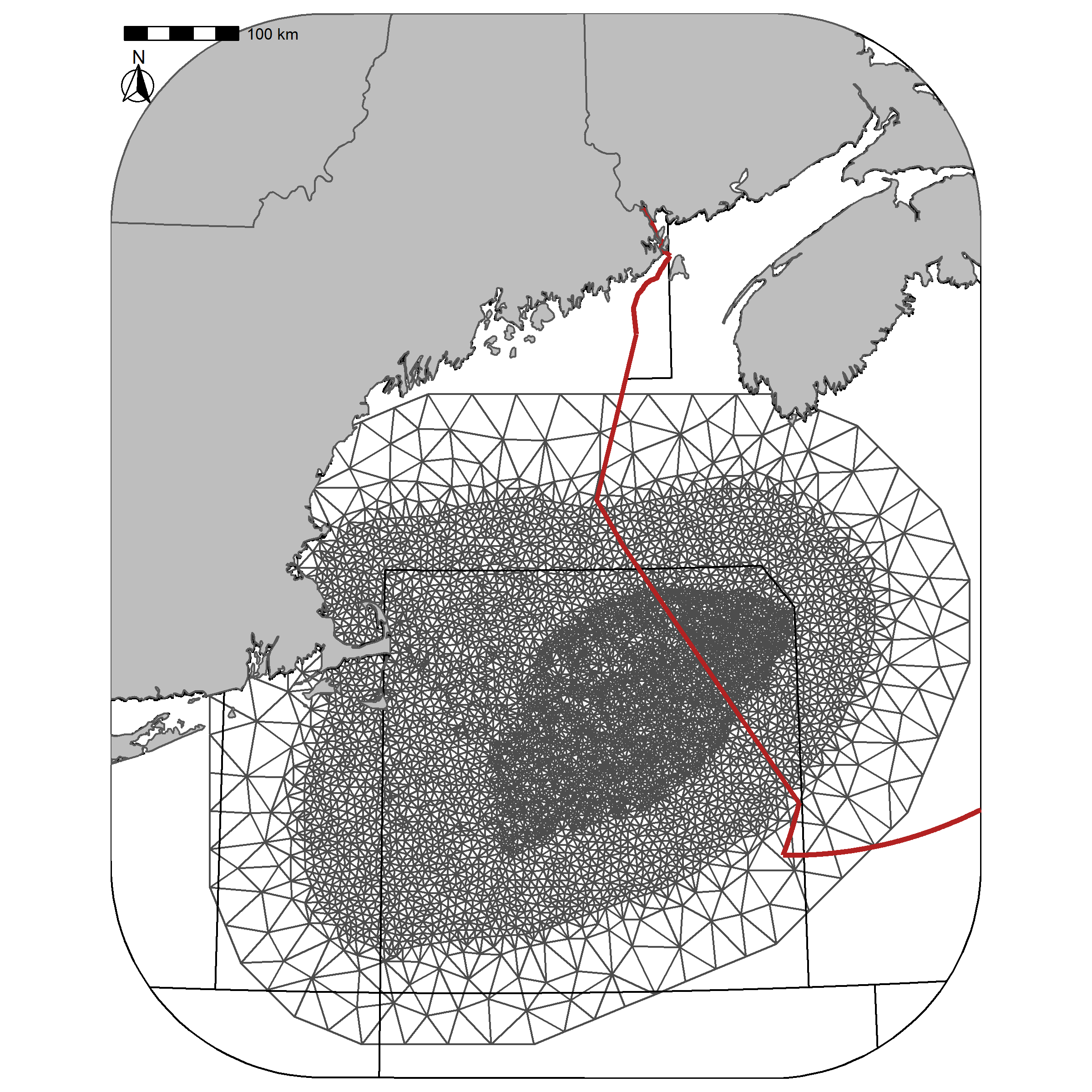
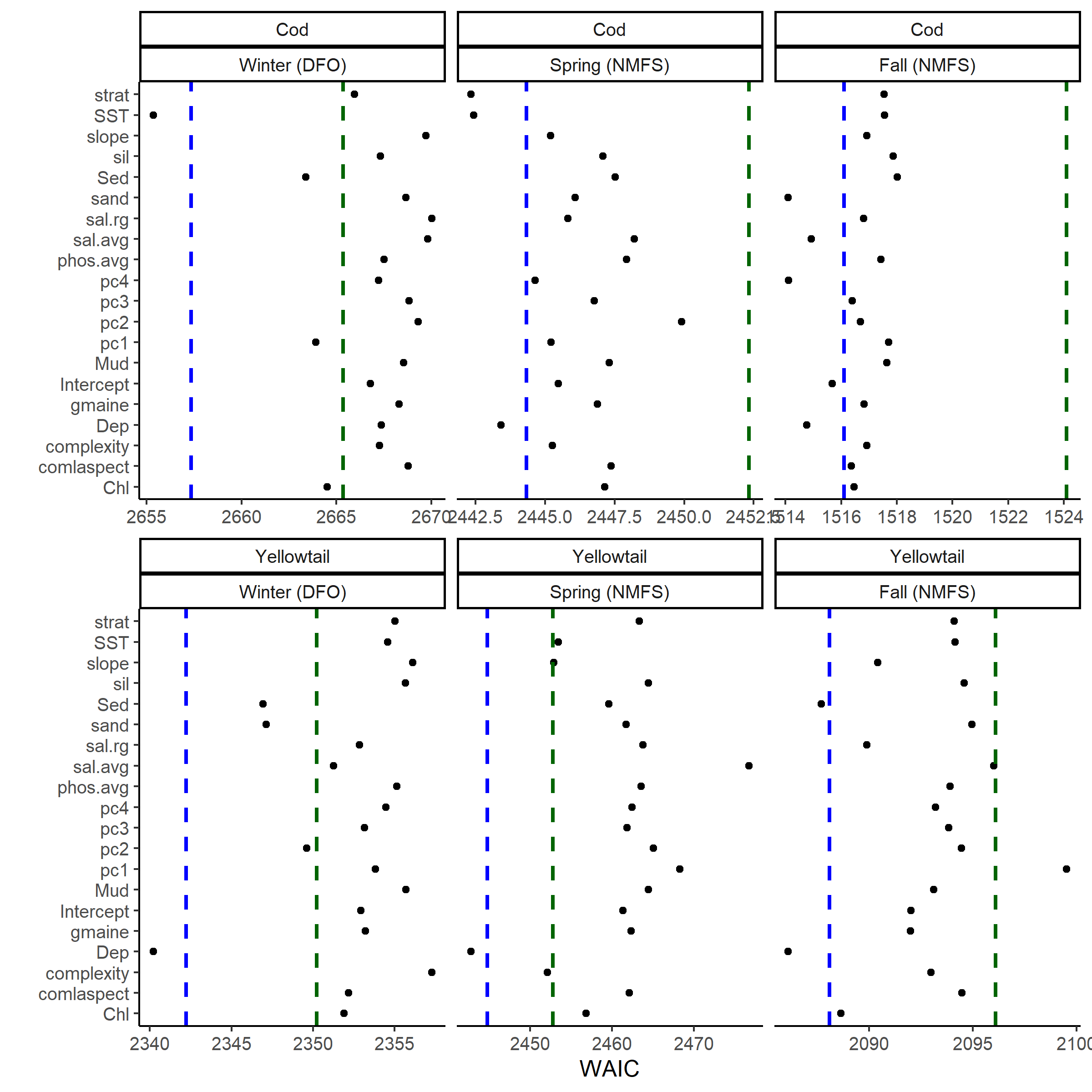
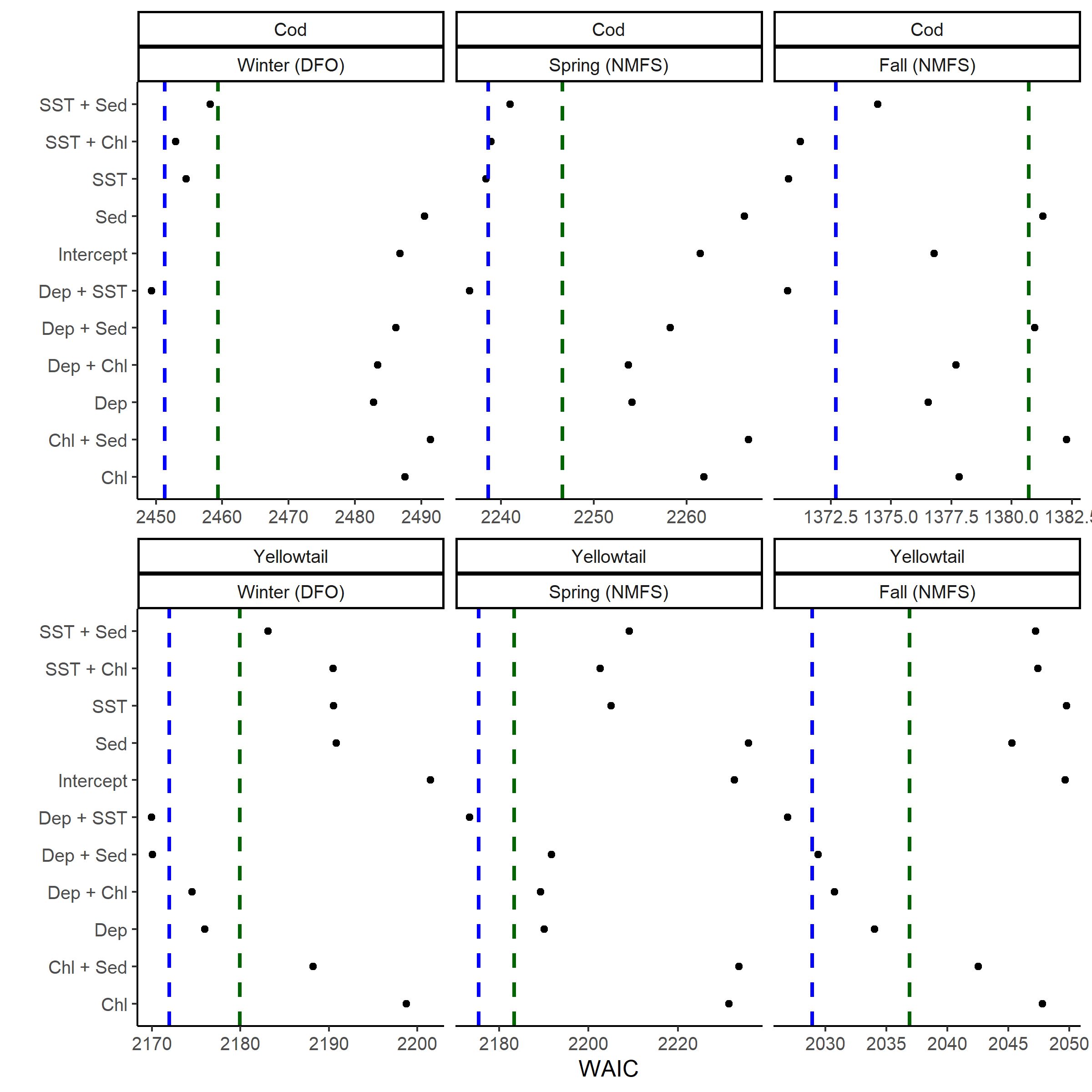


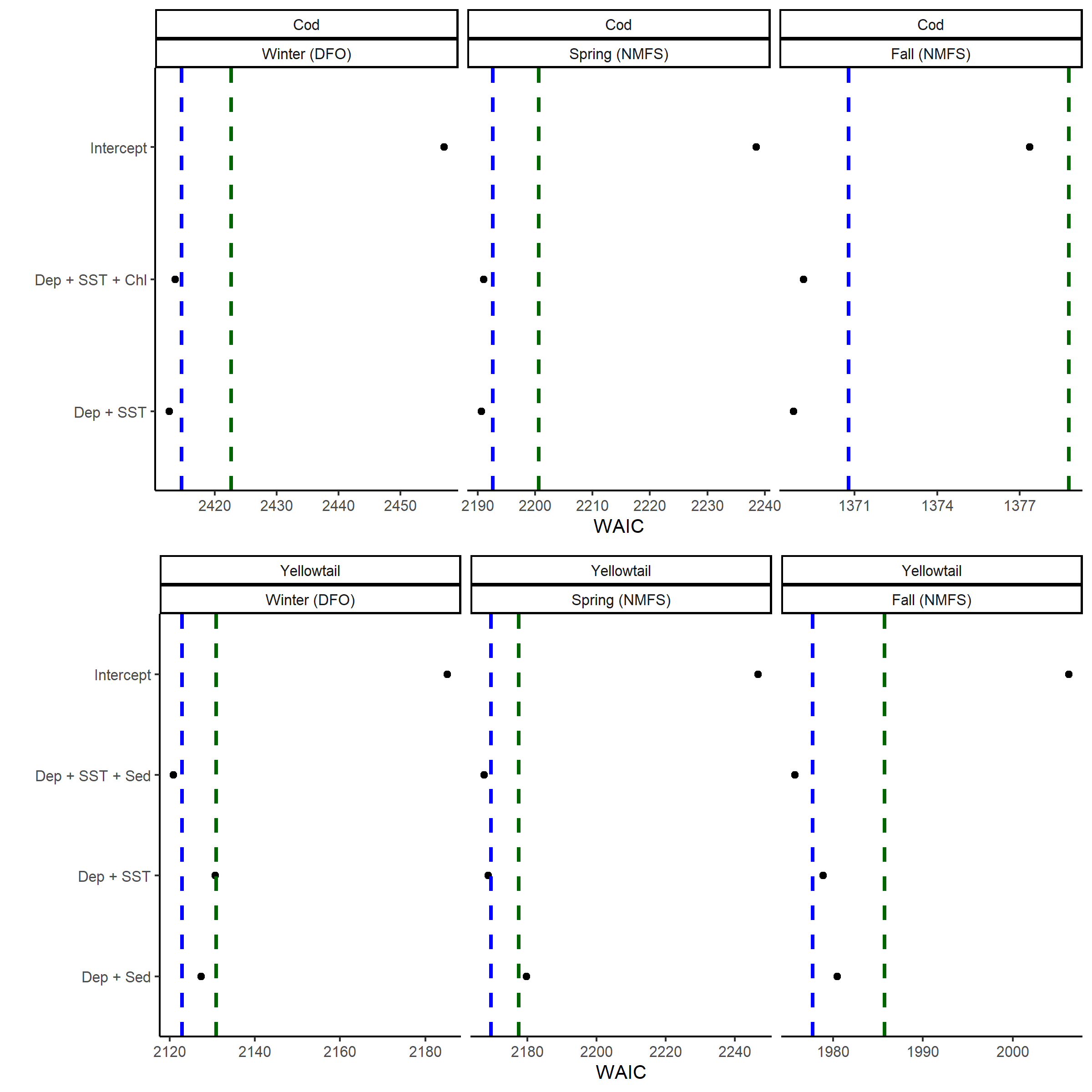
Figure 2: Mesh used for INLA.



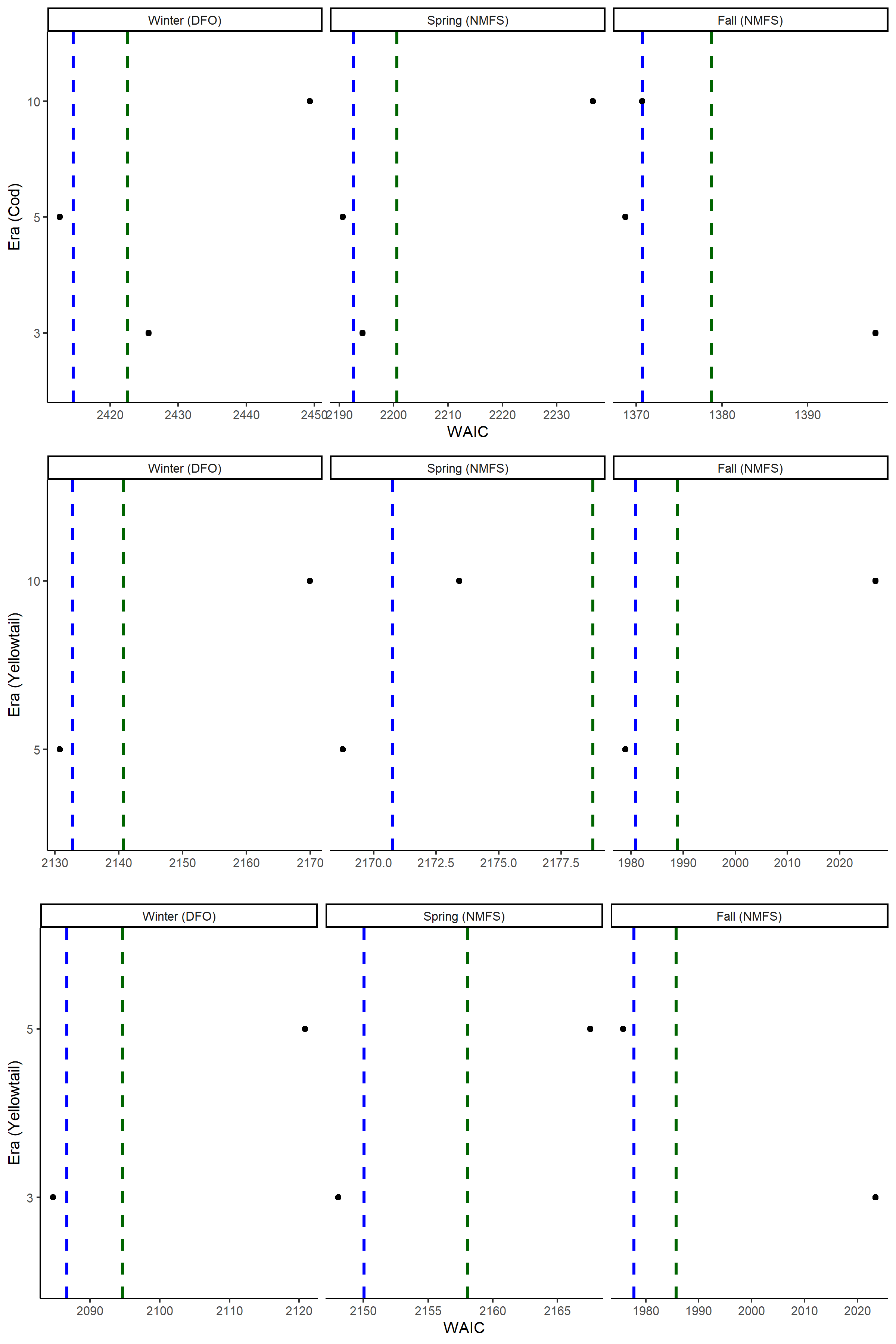
(#fig:diag.1.fe)Initial model selection



(#fig:diag.2.fe)Step 2 model selection



(#fig:diag.3.fe)Step 3 model selection



(#fig:diag.rf)Random field model selection. For cod the model used is Dep + SST for all of the random fields. For Yellowtail the 5 and 10 fields were compared using the Dep + SST model, while the 5 and 3 fields were compared using the full Dep + SST + Sed model

# References