Modeling Complex Climate Change Effects on Fluctuating Populations of Fish Communities in the Northern Pacific Ocean

by

Hannah E. Correia

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
August 3, 2019

Keywords: generalized additive models, rank estimation, variable selection, Granger causality, convergent cross mapping, groundfish population dynamics

Approved by

F. Stephen Dobson, Professor, Department of Biological Sciences
Asheber Abebe, Professor, Department of Mathematics and Statistics
Jack Feminella, Professor, Department of Biological Sciences
Matthew Catalano, Assistant Professor, School of Fisheries, Aquaculture, and Aquatic Sciences
George Flowers, Dean of the Graduate School

Abstract

The main purpose of this study was to bring new mathematical and statistical methods to the ecological community and highlight novel application of these methods to practical scientific questions. Deepwater marine systems provide a challenge to understanding population dynamics, and commercial fisheries are particularly interested in understanding how valuable fish populations change with variable climate conditions. Several interdecadal climate modes and their effects on marine systems have been studied, and many of those climate modes are found to contribute to variability in the northern Pacific Ocean system. Studies on how climate and marine environment affect commercially important fish populations in this region often lean too heavily on the potential effect of climate modes. The importance of climate variability induced by unknown sources in marine systems must also be considered when understanding population changes in marine vertebrates. This study aimed to incorporate a variety of marine and atmospheric variables to model population-level changes in seven groundfish species. The response of these commercially valuable fishes to changes in climate and their marine environment are not well understood, as they live in deep waters (> 300 m) making experimental studies on adults difficult. Such models involved considering population responses over space and time for multiple variables, increasing model complexity beyond the capacity of basic statistical methods.

Several useful methods in statistics and mathematics allow for modeling of high-dimensional data without assumptions on population distributions. While previous fisheries research relied heavily on time series analysis, the past decade has seen a move to generalized additive models (GAMs) as a nonlinear method of modeling fish populations using smooth coefficient functions. The method provides the flexibility of fitting high-dimensional functions to allow incorporation of space and time in a non-additive way. The single-index model combines the interpretability of generalized linear models (GLMs) with GAMs, encasing a GLM in an unspecified link function. An extension of this is the varying-coefficient single-index model, which models

several common covariates within a single-index model and uses varying smooth coefficients of the single-index model to quantify the relationship between other additive covariates and a response. Quantifying the relationship of individual species' responses to changes in climate using these methods provides a starting framework to consider how these groundfish species interact over time and space. A new mathematical tool called convergent cross mapping (CCM) can factor in multiple variables and map dynamic causal relationships extracted from time series data. So far the method has been applied to a two-species sardine-anchovy system off the coast of California, but there is the potential for this method to be expanded to measuring spatiotemporal effects involving more than two species.

The goal of this study was to (1) quantify groundfish responses to climate and ecosystem fluctuations using multiple indicators of ocean variability through improved statistical methods described herein, and (2) detect causal factors and complex interactions involved in changes in populations numbers found in Alaskan groundfish species by applying CCM to this complex groundfish ecosystem. A combination of multiple sources of ecological, environmental, climatological, and geographic data were used to investigate potential causal factors and attempt to explain visible changes in groundfish populations. It is hypothesized that changes in fish populations will be best explained by multiple interacting variables that, when modeled correctly, will provide a more accurate system for monitoring and predicting the health of fish communities. Not only does this understanding have the potential to contribute to the development of more informed management practices for wild fish populations, it also provides greater insight for modeling changing ecosystems over time and creating more accurate prediction models describing complex systems.

Over twenty years of fishery longline surveys collected by the NOAA were merged with corresponding climate data from ICOADS, COPEPOD, WODA13, and WOD13. Multivariate analysis methods were used to discover potentially significant clustering within the data, examine relationships and dependence among variables, and uncover potential correlations for further investigation. Robust and efficient nonparametric statistical procedures provided inference for small samples. Single-index models (SIMs) including a host of environmental variables were used to determine the most important environmental effects on groundfish catch

rates. Single-index varying coefficient models (SIVCMs) provided framework for including latitude and longitude to environmental variables that vary spatially, allowing for estimation and prediction of groundfish catch over space and time without the 'curse of dimensionality'. SIMs and SIVCMs also permitted exploration of the effects of trophic and habitat interactions by co-occurring groundfish on the species of interest. CCM was applied to the ecological system to search for signals that indicate potential causal effects of common environmental forces on the fish populations. All statistical analyses were performed using the free statistical computing software environment **R**. Results from the analyses and modeling of interactions were evaluated with the following main questions: (1) Do certain sources of variation influence fish population dynamics more heavily? (2) Which model(s) tested herein most accurately predict future ecological fluctuations? (3) Can CCM be applied to different complex systems efficiently and precisely? Application of causal analyses such as CCM to marine ecosystem data has the potential to provide explanations for changes in catch and improve prediction models with applications to broader ecological modeling that can inform wildlife policies and fisheries management. Considering that the first application of the CCM analytical protocol was on a complex sardine-anchovy system over multi-year scales (Sugihara et al., 2012), the longline surveys on groundfish populations in the northern Pacific Ocean and related environmental data from sources such as ICOADS are ideal candidates for CCM and comparisons of successful method deployment.

Based on the results of CCM and statistical analyses, I will create models to explain the dynamic changes in the fish populations and test model prediction capabilities and limitations utilizing cross-validation as well as data from future longline surveys. The models will include three main factors theorized to affect population dynamics: time, climate, and species interactions. CCM techniques will be refined to improve the detection accuracy of causal relationships. Further research into the connections behind the complex ecological interactions of the system may be explored using graph theoretic techniques. Future plans for modifications to the CCM technique will increase accuracy of quantifying causes and creating models to predict future population responses to potential ecosystem variations.

Acknowledgments

I would like to sincerely thank Ash Abebe, Steve Dobson, Jack Feminella, and Matt Catalano for taking the time to serve on my committee and provide helpful feedback on my research. I am particularly appreciative of Steve's encouragement and support of the direction of my research. I would also like to thank Alan Wilson for serving as outside reader and providing comments and feedback on this dissertation.

Thanks also to Pete Johnson for planting the idea of pursuing graduate research in the 2010 REU and first drinks that have facilitated the research process since then.

Much appreciation to Overtoun Jenda for inviting me to participate in the Masamu Workshops in conjunction with SAMSA. The experience has helped push this research much further in a much shorter timeline than anticipated and has added to my global cognizance.

No small amount of acknowledgment goes to the NSF GRFP that I was awarded in 2015 and the subsequent NSF GROW in 2017, which provided me the opportunity to focus on my PhD research and also facilitated novel research avenues and future collaborations.

I am thankful for my dog, Harry, whose friendly disposition and antics always put a smile on my face and kept me sane no matter how difficult previous hours or days had been.

Finally, I thank my family for the influences that brought me to this point in life. Your free-thinking and cultivation of strong reasoning has developed my expansive thought process and provided me a basis on which to build my independent and unique research process.

Table of Contents

List of Figures

List of Tables

Chapter 1

Modeling Environmental Effects on Groundfish Catch in the Northern Pacific

1.1 Northern Pacific Ocean Marine System Dynamics

Evidence of global climate change has generated strong interest in studying environmental effects on animal populations. Current evolutionary theory models of species' responses to changing climate focus on single subspecies and local adaptation, but observations suggests a more complex network of species interactions and ecological dynamics (De Mazancourt et al., 2008; Urban et al., 2012). Ecological theories on adaptation consider species in isolation, but the basis for emergence of biodiversity underpinning these theories are not typically representative of observed ecosystems (Scheffer et al., 2018). There is also a fundamental lack of theoretical studies supported by empirical evidence on the impact of global environmental change on ecological interactions (e.g. complex trophic interactions, positive interactions such as mutualism, competition's effect on species coexistence, novel community assemblages) and how these could affect evolutionary mechanisms and feedbacks on population dynamics and range shifts (Lavergne et al., 2010). These studies are critical for generating appropriate community models for predicting and understanding a wide range of potential eco-evolutionary dynamics under climate change (Urban et al., 2012).

Studies have revealed potential amplified warming effects in the northern-high latitude region (60°N) relative to overall global warming trends (Holland and Bitz, 2003; Serreze and Francis, 2006). This trend has brought about changes in Pacific marine systems and is predicted to affect future fish diversity and population sizes (Brander, 2007; Cheung et al., 2013). As a major intersection for several fluctuating climate systems (Anderson and Piatt, 1999; Bakun,

1999), a diverse and complex marine system (Livingston et al., 1999), and a significant source of fishing income and food for the United States and Japan (National Marine Fisheries Service, 2014), Alaska's surrounding waters are an ideal environment to study the effects of a changing climate on an ecosystem with a dynamic structure, known anthropogenic pressures, and crucial role in economics and the global food supply. Present research on the ecology of Alaskan groundfish, a major commercial fishery for the area, is particularly limited due to the difficulty of accurately studying and managing deepwater fishes. Modeling responses of adult groundfish to persistent fluctuations in their marine environment while also accounting for their dynamic system will contribute significantly to the knowledge of groundfish population dynamics and behavior.

A major surge in research of the Pacific marine system came in the 1980s after a strong regime shift was observed by fisheries along the northeastern Pacific coast of North America and in the Alaskan waters in the late 1970s. This shift was found ecosystem-wide, characterized by a strong increase in the catch of Alaskan salmon and several groundfish and flatfish species, a dramatic change in patterns of phytoplankton biomass in the Gulf of Alaska and Californian waters, and a steep drop in the abundance of many forager species such as shrimp and Pacific northwest salmon (Anderson and Piatt, 1999; Francis et al., 1998). These substantial variations prompted fisheries management experts and marine scientists to consider the environmental forces behind these biological changes. Analysis of water temperatures and pressure anomalies as related to marine populations stimulated further questions into the sources of atmospheric and oceanic processes.

A primary focus of marine studies on the environmental effects on fish populations has been to determine the processes behind changing temperatures and assess whether these processes adhere to temporal cycles. Specifically for the northern Pacific, several hypothesized climate patterns have been suggested for interdecadal anomalies in sea surface temperature patterns and the North Pacific Pressure Index (Francis et al., 1998). An issue with these hypothesized patterns is that these climate systems are harder to model and predict, and there is little physical evidence to support one hypothesized pattern over all others. One climate pattern conceptualized and modeled by Hollowed and Wooster (1992) suggested warm and cool

periods based on the winter atmospheric circulation of the north Pacific, each period lasting 6 to 12 years. The pattern appeared to be linked with the El Nino-Southern Oscillation (ENSO) phenomenon, and Hollowed and Wooster (1992) showed that groundfish recruitment success in the northeast Pacific was associated with this pattern of warming and cooling eras. There is a debate on whether this pattern is discernible in the California Current System during some time periods, leading some scientists to question the model's validity (Francis et al., 1998). The more well-known measure of interdecadal climate variability is the Pacific Decadal Oscillation (PDO), which is also based on the pressure changes in the north Pacific (Mantua and Hare, 2002; Mantua et al., 1997). PDO was also shown to affect SST and North Pacific Pressure Index in northern Pacific waters, although it was not found to be linked with ENSO. Fluctuations in salmon stocks appear to follow this regime more closely than the ENSO-related model of Hollowed and Wooster (1992), however the connection between PDO and groundfish stocks is less supported. More recently, another climate cycle for variations in salinity, nitrate, phosphate, silicate, dissolved oxygen (DO), and chlorophyll has been described and termed the Northeast Pacific Gyre Oscillation (NPGO) by Di Lorenzo et al. (2008). The NPGO is claimed to be a previously unrealized factor more predictive of dramatic ecosystem shifts than the PDO by capturing important interannual and decadal biological patterns such as nutrient cycling that are not explained by ENSO or PDO. The interplay between the PDO, ENSO, and NPGO drive both low and high-frequency climate variability throughout the North Pacific Ocean (Di Lorenzo et al. (2013), Figure 1.1). Large-scale indicators such as these often account for more variance within an ecosystem than a single environmental variable, because these indicators integrate several environmental variables such as temperature, sea ice levels, and current patterns (Drinkwater et al., 2010).

1.2 Environmental Variables Relevant to Marine Ecosystems in the North Pacific

Large changes in climate and associated oceanic conditions can be predicated by shifts in environmental variables, which, in turn, can have cascading effects on primary production, trophic structures, and population recruitment in marine life (Bakun, 1999). Temperature changes have

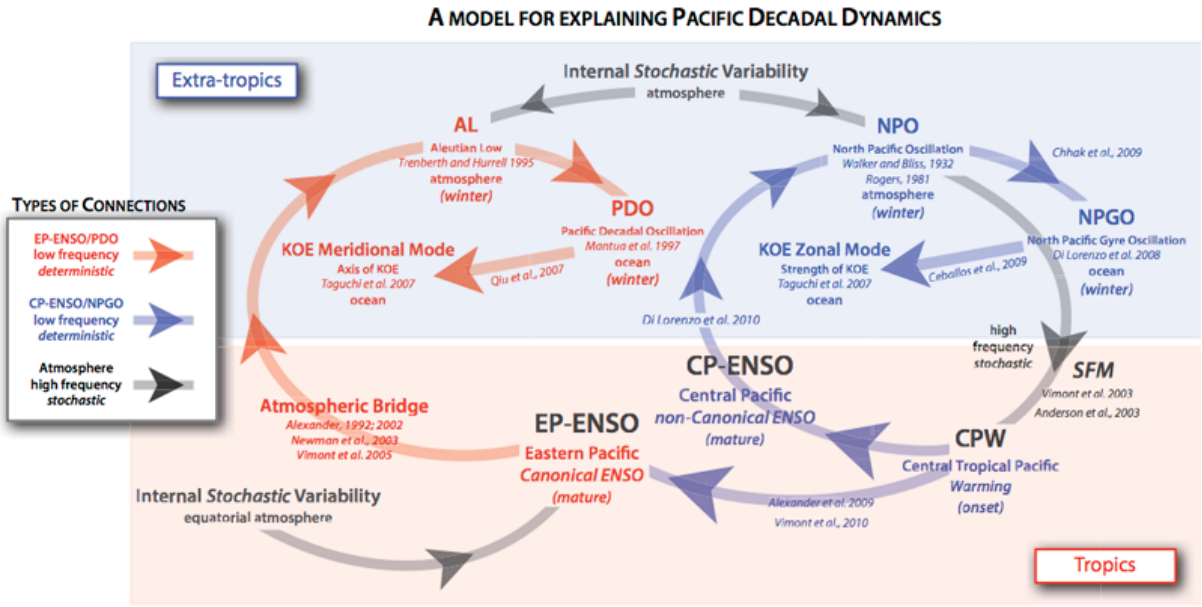


Figure 1. Synthesis of Pacific climate dynamics and teleconnections. The Pacific Decadal Oscillation (PDO; red path) and North Pacific Gyre Oscillation (NPGO; blue path) outline teleconnections at low-frequency time scales. The gray path shows how sources of high-frequency stochastic variability in the atmosphere energize the Aleutian Low (AL), North Pacific Oscillation (NPO), and El Niño-Southern Oscillation (ENSO) systems. In the schematic, NPO low-frequency variability is linked to Central Pacific (CP)-El Niño; however, processes internal to the North Pacific atmosphere appear to drive its high-frequency variability (gray path).

Figure 1.1: Figure 1 from Di Lorenzo et al. (2013) illustrating the interplay between the Pacific Decaldal Oscillation, El Niño-Southern Oscillation, and North Pacific Gyre Oscillation in the North Pacific.

been a central focus in the climate change discussion, although with increasing global temperature comes changes in winds, pressure systems, ocean transport systems, and coastal runoff. Together, alterations in these factors can transform the composition of oceans (Brodeur et al., 1999; Sadorus, 2012) and contribute to ecosystem stress (Livingston et al., 1999). Some of the more common oceanic variables being collected for use in studying climatic effects on marine life include temperature, DO, salinity, acidity, chlorophyll, and nitrogen. More generally, northward contraction in the distribution of several species has been observed as temperatures increased in a 25-year study (Sadorus, 2012).

Salinity is a good indicator of the level of mixing in oceans and plays an important role in stratification (Kakehi et al., 2017). Increased salt-stratification of upper ocean waters from increased sea-ice melt and changes in the hydrologic cycle induced by global warming is likely to constrain the flow of nutrients in the photic zone which will reduce primary production and potentially favor webs dominated by low-energy predators (Carmack and McLaughlin, 2011). Oceans are becoming less saline as warmer polar conditions promotes the melting and

runoff of ice-trapped freshwater into oceans (Sadorus, 2012). In the subarctic North Pacific, decreased ocean salinity indicates increased freshwater concentrations from continental run-off, increasing the supply of iron and thereby stimulating diatom production. Increased freshwater also changes concentrations of silicate and nitrate concentrations in surface waters, limiting or promoting diatom production (Dugdale et al., 1995; Wong and Matear, 1999).

Increasing acidity and decreasing DO in oceans are linked to rising levels of carbon dioxide and other greenhouse gases. Oceans are a major sink for anthropogenic carbon, which has already resulted in a reduction of oceanic pH by 0.1 (acidification) and is predicted to decrease by another 0.3 by 2100 (Caldeira and Wickett, 2003). Ocean acidification can cause metabolic suppression in fishes and decreased ability to form shells in calcifying organisms, potentially leading to drastic changes in trophic systems. Increased carbon in oceans can also promote plankton blooms that vastly exceed consumption by zooplankton, leading to larger and more persistent hypoxic zones which can harm marine life and obstruct growth, metabolism, and predatory behavior of marine organisms (Sadorus, 2012).

The North Pacific Ocean is experiencing increasing upper ocean acidification at rates closely following atmospheric CO_2 rises (Byrne et al., 2010). Rates of acidification in the upper 500 m of the ocean have been significantly increasing since 1991, and increases in acidification of near-surface and mixed-layer depth waters are primarily attributed to anthropogenic carbon uptake (Caldeira and Wickett, 2003). Increased acidity of marine waters reduces the calcification of carbonate-forming organisms such as bivalve molluscs and crustaceans (Wakita et al., 2013). Not only does this affect these large shell-forming species, but it also affects calcifying marine plankton such as abundant planktic foraminifera and pteropods in the North Pacific (Bednaršek et al., 2014; Taylor et al., 2018). Negative impacts on these planktonic organisms are likely to create cascading effects in trophic systems of the Northern Pacific (Wootton et al., 2008). Acidification of oceans also resulted in over 10% decreased low-frequency sound absorption in most Pacific waters above 400 m and projected to decrease further (Hester et al., 2008), which may alter predator-prey interactions of marine fishes through changes in abilities to detect prey and avoid predators.

Alkalinity is part of the ocean's carbonate system and is often quantified through pH, dissolved inorganic carbon (DIC), or total alkalinity (TA) (Hagens and Middelburg, 2016). While TA contributes to pH, a change in alkalinity does not necessarily translate to a change in pH. Changes in temperature, salinity, silicate, and nitrate can also affect pH levels, and freshwater inputs mainly control TA (Lee et al., 2006). Denitrification and consumption of nitrate by biota increase alkalinity while use of calcium carbonate by calcifying organisms decreases alkalinity in surface waters (Wong et al., 2002a). Increased TA in surface waters (or decrease of TA in deeper waters) is typically brought about by vertical mixing, which brings deeper waters rich in alkalinity to the surface; in contrast, TA decreases in surface waters occur from calcification by primary producers, but it is not a smooth and continuous decrease (Fiadeiro, 1980; Wong et al., 2002a). The carbonate system is rapidly changing due to anthropogenic atmospheric CO_2 increases (Caldeira and Wickett, 2003; Feely et al., 2004; Orr et al., 2005). While alkalinity is mainly decreasing due to increased atmospheric CO_2 from anthropogenic sources, anthropogenic increases in atmospheric nitrate and sulfur deposition also contribute to decreasing alkalinity in waters of the Northern Hemisphere by altering saltwater chemistry (Doney et al., 2007). Increased variability in the interannual carbonate system in the Pacific Ocean are also attributed to variability in the PDO and ENSO (Fry et al., 2015).

Silicate is another important tracer for the ocean's calcium carbonate system and an important nutrient for diatom growth (Wong and Matear, 1999). In the North Pacific, diatom production and plankton productivity is limited by iron, which adversely affects diatom utilization of silicate in the region (Wong and Matear, 1999). As diatoms in this region require higher levels of silicon in order to grow, silicate is likely to become a limiting factor in plankton biomass accumulation in the region (Wong and Matear, 1999; Wong et al., 2002a). Silicate limitation is likely to co-occur with or may be induced by iron limitation in the Northern Pacific Ocean. The process is initiated by an influx of iron, often through runoff from freshwater systems into the ocean, favoring increased diatom production. Increased diatom growth leads to depletion of iron, silicate, and nitrate. Silicate and iron are therefore important regulators of phytoplankton growth in Alaskan marine waters (Harrison et al., 2004).

Seasonal cycles of nitrate are important for assessing effective removal of atmospheric CO_2 by the world's oceans, and nitrate is an essential nutrient contributing to primary production. Measuring seasonal nitrate depletion at the ocean's surface can be used to determine the intensity of carbon and nutrient removal from surface waters to be sequestered to the deep ocean in the form of dissolved and particulate organic matter (Wong et al., 2002b). The Northern Pacific Ocean at higher latitudes has high concentrations of surface nitrate (Fry et al., 2015). The subarctic North Pacific is considered a high-nitrate low-chlorophyll (HNLC) region, where it is rare for surface nitrogen to be depleted and seasonal depletion occurs in cycles of approximately 4-6 months in duration (Wong et al., 2002a) Similar to effects of iron on silicate utilization, iron-limited diatoms in the North Pacific do not utilize nitrate as efficiently as diatoms less limited by iron in other regions (Wong and Matear, 1999). Large decreases in nitrate levels may therefore be due to increased diatom production, however El Niño can impact seasonal nitrate patterns which can alter seasonal primary production and nutrient depletion rates (Wong and Matear, 1999; Wong et al., 2002b).

Phosphorus is another major nutrient contributing to primary production (Desmit et al., 2015). Phosphorus is important for growth of plankton and bacteria especially for the creation of cell membranes and DNA. However, excess phosphorus can result in overabundant algal blooms that then create hypoxic zones in the ocean (Childers et al., 2005; Yoshimura et al., 2007).

The distribution of DO in ocean waters is a primary habitat indicator for most marine organisms. Photosynthetic activity generates oxygen in the photic zone, and benthic communities rely on strong mixing to deliver these oxygen-rich waters from the surface. Respiration by algae and other organisms also removes oxygen from both pelagic and deep waters. DO concentrations in the deep ocean and its relation to biological productivity are therefore strong indicators of marine ecosystem function (Boyer et al., 1999; Loubere, 1994). A consequence of increasing ocean temperatures is that less oxygen is contained in warmer sea water, thereby decreasing DO availability to marine organisms (Boyer et al., 1999). Marine DO concentrations have been decreasing in subsurface (100–1000 m depth) waters in most of the world's mid- and high-latitude ocean basins over the past 50 years (Deutsch et al., 2005; Ono et al., 2001).

Some of the most well-documented changes are in the North Pacific, where variations in the Pacific Decadal Oscillation propel changes in basin-level circulation rates that affect biological production and distribution of oxygen and other nutrients (Deutsch et al., 2005; Whitney et al., 2007).

It is predicted that deepwater DO concentrations will decrease by 20-40% globally over the next 100 years. Declines in ocean DO concentrations precipitate changes in primary and secondary production in regional waters and has been linked to declines in pelagic and deepwater fish production (Koslow et al., 2011). For example, benthic and schooling rockfish communities were severely effected by inner-shelf hypoxia events in the California Current System in the early 2000s (Grantham et al., 2004). In all areas of the northern Pacific, declines in DO levels are being observed with decreasing trends at all depths up to 400 m for 50 years of data from 1956-2006, and these decreasing trends remain significant at up to 1000 m depths. Based on approximate calculations of oxygen consumption rates and declining oxygen levels, it would take roughly 20 years for the Northern Pacific waters to reach hypoxic levels. A few deepwater species tolerant of low oxygen, such as sablefish and rockfish, may increase their ranges under these conditions. Most other species dependent on high DO levels will move to shallower waters, increasing competition and interactions with predators (Whitney et al., 2007).

Transport of nutrients from deep waters towards surface waters are important for enriching phytoplankton communities. Since all phytoplankton contain chlorophyll, chlorophyll levels in oceans are important for estimating phytoplankton biomass. Chlorophyll *a* strongly correlates with winter levels of salinity, nitrogen, and phosphorus through plankton bloom formation and peak biomass (Desmit et al., 2015). A reduction in deepwater plankton biomass and phytoplankton concentrations being confined to shelf and slope regions of the North Pacific basin are characteristic of El Niño effects; these characteristics indicate reduced carbon transport to the deep ocean (Sackmann et al., 2004). Spatiotemporal variation in plankton are important for energy transfer to higher tropic levels and are also important to the transport of DO and nutrients via respiration and decomposition to deepwater communities (Smith Jr and Baldwin, 1984). Dominant species of marine phytoplankton and zooplankton are being affected globally by increased atmospheric CO_2 concentrations through the decreased ability to form calcareous

skeletons, decreasing chlorophyll concentrations over recent decades (Benson and Trites, 2002; Riebesell et al., 2000). Given these global trends in atmospheric CO_2, there is growing interest in how changes in phytoplankton dynamics will translate to commercially and ecologically important fishes.

Wind and wave movements, especially from intense winter storm systems, are important for mixing of oxygen-rich waters from the euphotic zone with benthic waters to distribute nutrients from the deep ocean. However, changes in global climate are precipitating shifts in wind and wave patterns (Harrison and Wallace, 2005). Positive correlations between winter wind speed and summertime plankton biomass have been observed globally (Feng et al., 2015; Kahru et al., 2010). Primary and secondary production can therefore be affected by atmospheric pressure changes that alter wind and wave circulation patterns and intensities (Benson and Trites, 2002). Strong interannual changes in winds driven by pressure systems can affect the duration of wind-mixing events and change the depth of the mixed-layer, affecting biological production and resulting in strong ecosystem regime shifts, such as the ones noted in the Northern Pacific Ocean (Polovina et al., 1995).

1.3 Fish of the Northern Pacific

Sablefish (*Anoplopoma fimbria*), also commonly known as black cod, has been managed by the US government since 1988. These fish occupy the ocean floor at depths of 300 to 900 m along the continental slope and live to 50 years or more. Females can grow to 110 cm in length and weigh 14 kg, while males only reach 88 cm and 6.8 kg. Main predators on sablefish juveniles are adult salmon, while as adults their main predator is Pacific halibut. Sablefish are opportunistic feeders and will prey on pollock, herring, and Pacific cod, among other fishes (Alaska Fisheries Science Center, 2019a). Several laboratory studies have been conducted on juvenile sablefish to understand how they may respond to temperature and light. Current research suggests that sablefish prefer temperatures ranging from 2 to 8°C (Stoner and Sturm, 2004), while juveniles prefer warmer temperatures from 9 to 14°C (Sogard and Olla, 1998b). Sogard and Olla (1998b) found during nighttime-simulated light conditions, juvenile sablefish activity significantly increased in the presence of colder (3°C) temperatures, indicating potential avoidance

behavior in low light when sablefish are typically not as active. These temperature preferences also can affect sablefish feeding behavior. Stoner and Sturm (2004) found increases in attack rates, consumption rates, and amount of food consumed by juvenile sablefish in 5°C and 8°C waters when compared to those in 2°C. Time to locate, attack, and consume bait decreased with increasing temperature, indicating that sablefish may be able to detect olfactory cues of bait more easily in warmer waters. Marine Ecology and Stock Assessment (MESA) longline surveys bait with squid, therefore lab findings on juvenile sablefish such as those of Stoner and Sturm (2004) may provide important insight into behavior of adult sablefish. Juvenile sablefish only enter colder waters when strongly motivated by food (Sogard and Olla, 1998a). Longer dives by small juvenile sablefish in sharp thermal gradients (approximately 10°C difference between surface and bottom temperatures) resulted in loss of equilibrium and mortality. Larger juveniles were able to sustain longer dives to cold waters, but showed preferences for 12°C waters. Increasing temperature appears to have a positive effect on juvenile sablefish growth, with Sogard and Olla (2001) reporting some of the highest growth rates among all juvenile teleosts in water temperatures up to 14°C. Schirripa and Colbert (2006) considered oceanographic variables in relation to sablefish recruitment in the California Current System and identified significant relationships with monthly mean sea level and north and east flow of wind-driven water movement known as the Ekman transport. This southern population is considered separate from the northern population present in the Gulf of Alaska and Bering Sea, thus it is unknown if the northern sablefish population responds to the same oceanic conditions as the southern population.

Pacific cod (*Gadus macrocephalus*) generally reach 130-140 cm in length and weigh up to 25 kg, and they inhabits waters of up to 250 m in depth and migrate to shallower waters in the summer. As a commercially important and relatively short-lived (up to 18 years) species, Pacific cod are susceptible to overfishing pressures. Pacific cod is also an important prey species for Pacific halibut and endangered Stellar sea lions (Aydin et al., 2015; Goen and Erikson, 2017). Therefore, there is strong interest in understanding the species' response to increasing temperatures. Like sablefish, Pacific cod have higher growth rates with increasing water temperatures. Laurel et al. (2016) found highest growth rates for juvenile Pacific cod

at 8°C, whereas Hurst et al. (2010) found 11°C to be most conducive for high growth rates and potentially persistent temperature-dependent growth rates from time of hatching. There is a significant effect of temperature on the size and survival of Pacific cod larvae from egg stage to hatching, with warmer temperatures precipitating early hatchings with smaller larvae Laurel et al. (2008). Early work by Forrester and Alderdice (1966) and Alderdice and Forrester (1971) demonstrated Pacific cod sensitivity to temperature, salinity, and DO in embryonic development. Wild Pacific cod cohorts in colder-than-average years had significantly higher growth rates at colder temperatures, but also higher mortality at warmer temperature treatments than those cohorts from a warmer-than-average year, indicating potential phenotypic plasticity (Hurst et al., 2010, 2012b). Cold temperatures also negatively affect shoaling behavior in juvenile Pacific cod (Davis and Ottmar, 2009), which can reduce foraging and protection from predators. These temperature responses lead to concerns that the increased fluctuations in weather conditions predicted to occur from climate change may make it difficult for Pacific cod to adapt sufficiently to these variations. Results from Hurst et al. (2012a) hint at these temperature effects producing changes in wild populations: range contraction of Pacific cod in the Bering Sea was evident in years where poor recruitment coincided with cold conditions.

Pacific halibut (*Hippoglossus stenolepis*) are one of the largest known flatfish at lengths of 2.5 m and weights of 300 kg and thus one of the top marine fish predators in the northern Pacific ecosystem. Pacific halibut are believed to remain in deep water (to 450 m) for much of their adult life (Seitz et al., 2007), however pop-up tags have revealed much more vertical movement and variations in migration behavior than previously estimated for Pacific halibut in the Gulf of Alaska (Loher and Seitz, 2006). As with sablefish and Pacific cod, laboratory experiments on juvenile Pacific halibut indicate responses to temperature and density. Temperature-related activity and food motivation, such as food-absent movement and food searching and locating frequency increasing with water temperature, have been observed (Stoner et al., 2006). Pacific halibut were significantly slower in finding, attacking, and consuming prey in 2°C, therefore catchability of Pacific halibut may be influenced by temperature. The relationship between catch and abundance has been explored more directly by considering size and density of young Pacific halibut in finding bait (Stoner and Ottmar, 2004). Grouped fish were significantly faster

at finding bait, and larger fish were more likely to consume bait by either reaching the bait first or through theft or bullying tactics on smaller fish. Absence of Pacific halibut being caught may therefore be due to low densities where fish are unable to detect bait or are disinterested, rather than absence of fish entirely (Stoner and Ottmar, 2004).

The Pacific halibut fishery is also one of the most valuable U.S. fishery resources, with Pacific halibut landings in the US worth $115 million in 2014; of that, over $106 million comes from Alaska landings alone (National Marine Fisheries Service, 2014). This has motivated a considerable amount of new research into Pacific halibut distribution and catch rates with technical advances in data collection to overcome the difficulties associated with tracking and studying adult deepwater fish in the wild. The International Pacific Halibut Commission has more recently prioritized research on climate change's effect on Pacific halibut distribution by collecting oceanographic data with catch data since 2000. The proposed effects of some of these variables have been examined, and preliminary research identified minimum DO (0.9 ml/l) and temperature values (0.5°C) for Pacific halibut numbers (Sadorus, 2012; Sadorus et al., 2014). The relationship between DO and catch was significant, where catch increased for DO values from 0.9 ml/l up to 3 ml/l. Sufficient DO is important for survival of marine life, although deepwater species tend to be more tolerant of hypoxic conditions than pelagic species.

Giant grenadier (*Albatrossia pectoralis*) is a little-studied but abundant fish in the northern Pacific. It is a deep-sea fish commonly occurring in 700 to 1100 m depths, but can be found at depths greater than 2000 m (Rodgveller et al., 2010). The giant grenadier is the largest of all the grenadier with average female weights in excess of 14 kg. Because of difficulties in determining ages of giant grenadier, their maximum age is unknown; however, some grenadier have been estimated to be 58 years old and estimated to mature around 23 years of age (Rodgveller et al., 2010) . Despite its lack of appeal to commercial fisheries due to its poor taste and texture, the giant grenadier is likely of great ecological importance due to its biomass dominance and position as an apex predator on the continental shelf (Drazen et al., 2001; Rodgveller et al., 2010). Beaked whales will eat grenadier, however the species is an insignificant portion of the whale diets (Walker et al., 2002). Fish, squid and scavenged material are a major portion of

the diet of California Current giant grenadier, with the diet composition varying by size, habitat depth, and latitude (Drazen et al., 2001). These results appear agree with giant grenadier feeding habits in the North Pacific and Bering Sea. Due to its lack of commercial importance, studies on the giant grenadier's response to environmental variables and changing oceanic conditions have not been the focus of deep-sea research. However, the giant grenadier diet appears to overlap with sablefish and shortspine thornyhead (discussed below) and could exert considerable pressure on these fish via competition for prey (Drazen et al., 2001; Rodgveller et al., 2008).

Shortspine thornyhead (*Sebastolobus alascanus*), shortraker rockfish (*Sebastes borealis*), and rougheye rockfish (*Sebastes aleutianus*) belong to the Sebastidae family. The rougheye and shortraker rockfishes prefer steep, rocky habitat at 300-500 m depths and are long-lived: the rougheye rockfish may be the longest-lived fish with a maximum reported age of 200 years. These two species vary drastically in size, with the rougheye rockfish adults averaging 40 cm in length, and the shortraker rockfish is the largest of all *Sebastes* with lengths up to 120 cm. Shortspine thornyheads, which reach 80 cm in length, are similarly long-lived with estimated maximum ages of more than 160 years (Andrews et al., 1999) and depths of 150-450 m. Large rougheye rockfish, shortraker rockfish, and shortspine thornyhead prefer average bottom temperatures of 3.6°C to 4.1°C (Reuter and Spencer, 2007). All three Sebastidae primarily consume shrimp and fish as adults. Shortspine thornyhead and shortraker rockfish are estimated to be tertiary carnivores based on nitrogen isotope ratios Kline Jr (2007). Shortspine thornyhead co-occur with sablefish on the western US continental slope, which may relate to ontogenetic shifts in depth distribution common to both species Tolimieri and Levin (2006). The energy transfer provided by these shifts has the capacity to indirectly link shallow water effects of fishing and climate change to deepwater assemblages (Tolimieri and Levin, 2006). Shortraker and rougheye rockfishes may be under-represented in longline surveys because of hook competition with more aggressive predators such as sablefish, Pacific cod, and Pacific halibut (Rodgveller et al., 2008). However, the current rougheye rockfish population model used to estimate rockfish abundance may be overestimating the efficacy of longline equipment, thereby underestimating current rockfish populations in the northern Pacific (Rodgveller et al., 2011).

An alternative best-fitting, two-part model to describe and predict presence and abundance of shortspine thornyhead for 13 years of trawl data included depth, local bottom slope, thermocline temperature, predation refuge, and prey abundance (Rooper and Martin, 2009). Unlike many other studies on marine fish, temperature was only important in predicting presence of shortspine thornyhead and found to be one of the least critical variables for abundance. Water temperature varies little for adult shortspine thornyhead since they have a narrow preferred depth range, therefore the effect of temperature on these fish may be limited (Rooper and Martin, 2009).

Walleye pollock (*Gadus chalcogramma*), also commonly known as Alaskan pollock, are important prey species for Pacific halibut, sablefish, and Pacific cod in the Northern Pacific Ocean (Hollowed et al., 2000; Livingston, 1993). Walleye pollock also serves as an important food source for Arctic marine mammals including spotted seals (*Phoca largha*) and ribbon seals (*Phoca fasciata*) (Bluhm and Gradinger, 2008). Cannibalism on age-0 pollock by adult pollock is common throughout the Gulf of Alaska and Bering Sea, which influences interannual recruitment variability through temperature changes that impact distribution overlap between adults and juveniles (Livingston, 1993; Mueter et al., 2006; Wespestad, 1996; Wespestad et al., 2000). Walleye pollock is also a commercially important species, composing over 40% of the global whitefish production (Ianelli et al., 2011). Pollock are considered short-lived species, typically living about 12 years. Pollock typically reach a maximum length of 50 cm long and weight of 1.3 kg, although pollock up to 90 cm long have been reported (Dorn et al., 2017; Smith, 1979). Walleye pollock commonly occur along the continental shelf at depths of 150-900 m, and adults move vertically through the water column diurnally to forage. Ideal temperatures for walleye pollock are between 3 and 7.5°C, with growth rates increasing with temperature, however under limited food availability, more rapid growth is observed at colder temperatures (Smith et al., 1986; Sogard and Olla, 2000). Temperature also affects predator avoidance behavior in walleye pollock . Small juvenile walleye pollock avoided colder waters to reduce interactions with adults that concentrate above the thermocline, thereby reducing cannibalism. Juveniles are also particularly sensitive to thermal stratification which may affect their responses to predators (Sogard and Olla, 1993). Lower temperatures were also associated

with faster swim speeds, increased group cohesiveness, and decreased path sinuosity, which have varying effects on encounters with predators and their outcomes (Hurst, 2007). The North Pacific regime shift has already drastically affected recruitment of walleye pollock, from being largely dependent on environmentally controlled larval survival to being strongly controlled by juvenile mortality dictated by abundance of predatory groundfish (Bailey, 2000). Hence, increasing ocean temperatures are intensifying biotic effects on walleye pollock recruitment by influencing pollock feeding conditions and exposure to predators, and these impacts are expected to substantially reduce recruitment in the future, leading to steep declines in walleye pollock production (Mueter et al., 2011).

1.4 Motivation and Implications for Analyses of Wild Groundfish Populations in the Northern Pacific Ocean

Typical investigations about marine fish species have revolve around commercial fishing techniques, improving catches, and mechanisms of reproduction that improve population numbers. However, commercial fishing organizations are becoming increasingly aware of the role climate change plays on determination of suitable fishing quotas designed to maintain healthy fish populations. It is now widely considered that marine fishes may not recover from population collapses as quickly as previously thought (Hutchings, 2000), and that fishing amplifies fluctuations in harvested populations that can predicate stock declines (Anderson et al., 2008). There are four major types of response mechanisms by which climate can affect fish populations: physiological, behavioral, population-level, and ecosystem-level responses (Rijnsdorp et al., 2009). Most of the studies described for the above eight target fishes focus on organism-level physiological and behavioral responses, and the results of such studies are difficult to transfer for understanding population and ecosystem-level responses. Commercial exploitation further complicates the effects of climate change on these eight northern Pacific marine species (Rijnsdorp et al., 2009). Interest in more flexible modeling techniques have attempted to quantify temporal population-level changes in marine fishes, although duration of studies and confounding environmental variables collected have limited conclusions from these efforts. Further, considerations for trophic interactions and three-dimensional spatial movements

have been difficult to incorporate using the historical modeling methods. More dynamic modeling structures are needed to describe and predict complicated marine systems in order to derive useful inferences for management and scientific understanding of hard-to-observe adult demersal fishes. Connecting specific environmental changes to fluctuations in the populations of these species will represent major progress in discerning the effect of climate change on marine ecosystem health and food security.

1.5 Effect of environmental changes in the subarctic Pacific Ocean ecosystem

Under all scenarios of anthropogenic climate change, dramatic changes in the ecology of Arctic and subarctic oceans have been detected (Orr et al., 2005; Overland et al., 2014; Walsh, 2008; Wassman et al., 2011). Shifts in species composition and northward range expansions have been observed with changing climate patterns and reduced sea ice persistence. Warming waters and reduced sea ice affect timing and intensity of primary and secondary production as well as impact pelagic-benthic nutrient cycling, the combination of which impact marine benthic species and both marine and terrestrial higher trophic predators (Grebmeier, 2012).

Environmental properties have strong effects on specific characteristics of marine organisms that can affect survival and reproductive success. The effect of temperature on marine organisms is one of the most commonly studied environmental features. Temperature optimizations and tolerances of marine organisms are theorized to enable optimal oxygen uptake by species, maximizing individual growth and function. Changes in temperature outside the limit of tolerances for marine animals have been shown to affect growth, swim speeds and activity, reproduction, phenology, distribution, and recruitment. Verticial stratification and mixed-layer depth are also affected by temperature as well as by salinity. Therefore temperature impacts primary production via ocean stratification modifying light availability and mixture of nutrients necessary for production (Drinkwater et al., 2010). Increase in the biomass of jellyfish, a low-energy predator, and decline in benthic biomass in the Bering Sea have been attributed to warming waters in the region (Grebmeier et al., 2006). With increasing atmospheric temperatures also comes rising sea levels as melting of glaciers increases the volume of freshwater entering marine systems. Species relying on specific nearshore habitats may be in danger of

losing these habitats due to both increasing sea levels and decreasing salinity from the influx of glacial freshwater (Drinkwater et al., 2010). Long-term shifts in atmospheric temperature and pressure cause alterations in climate systems, changing intensity and frequency of wind and wave activity over large bodies of water (Harrison and Wallace, 2005; Hemer et al., 2013; Weisse, 2010). Increased turbulence increases contact between plankton predators and prey, while current patterns change nutrient fluxes across marine basins, affecting the dispersion of fish eggs, larvae, and zooplankton as critical energy inputs for fish populations. (Drinkwater et al., 2010)

Timing of sea ice retreat is another common environmental indicator studied for its effect on subarctic and Arctic marine ecosystems. Early sea ice retreat is associated with late phytoplankton blooms and concomitant decreased zooplankton biomass (Drinkwater et al., 2010). Extreme sea ice retreats in 2007–2009 have precipitated intrusion of Pacific zooplankton into the Arctic waters, followed by appearances of walleye pollock and Pacific cod beyond their typical ranges, indicating that these species' distributions are expanding northward (Grebmeier et al., 2010). Such range expansions often lead to changes in tropic interactions, including both direct and indirect bottom-up forcing indicated in the North Pacific Ocean (Drinkwater et al., 2010). Northward invasion of Pacific cod associated with sea ice changes and warming waters reduced crab abundance in the Bering Sea (Orensanz et al., 2004); however Pacific cod spawning biomass has reduced in the Bering Sea in response to ocean warming and sea ice changes (Overland and Stabeno, 2004). Walleye pollock, on the other hand, are experiencing northward range shifts and increased biomass due to similar temperature and sea ice changes (Overland and Stabeno, 2004; Wassman et al., 2011).

Large-scale climate indices are often favored when studying regime shifts of the Northern Pacific marine system, but the underlying mechanisms for such effects on a population or community are difficult to identify. Furthermore, many common modeling methods are unable to handle numerous nonlinear environmental properties and their myriad interactions with each other along with accounting for differing responses of individual populations within a community to the same variable. The aim of my dissertation research is to reconcile some of these issues by investigating effects of specific environmental factors on individual fish populations

as well as predator-prey systems while simultaneously accounting for spatial and temporal changes in these populations brought about by regime shifts.

1.6 The Data

The major source of data for fish populations in the northern Pacific are derived from marine fisheries catch surveys conducted or overseen by the National Oceanic and Atmospheric Administration (NOAA), which are used to assess population health of major commercial fish populations for management purposes. The Gulf of Alaska and Bering Sea are the focus of commercial fisheries operations for important fish stocks such as salmon, cod, and flatfish. The Marine Ecology and Stock Assessment (MESA) Program is the central focus of NOAA's Alaska Fisheries Science Center (AFSC) for marine stock assessment and habitat research in the northern Pacific Ocean. The MESA Program includes annual longline surveys which have been conducted for over three decades. This long-term dataset is particularly useful for considering climatic effects on fish populations, as the study is performed at over one hundred locations along the shelf of the Alaskan coast surveyed at the same time each year. Seven commercially important fishes are regularly sampled in the annual longline survey: sablefish, Pacific cod, Pacific halibut, giant grenadier, shortspine thornyhead, rougheye rockfish, and shortraker rockfish. For each species, the number of fish caught and mean weight are recorded at each station, and catch per unit effort (CPUE), relative population number, and relative populate weight have been calculated within each management area.

The Groundfish Assessment Program administered by the Resource Assessment and Conservation Engineering (RACE) Division of the AFSC is another important assessment survey conducted regularly to understand groundfish distribution and abundance in marine waters off the Alaskan coast. Bottom trawl surveys from 1982–2015 were obtained with catch weights per area and catch numbers per area provided for all species caught in trawls. Location, surface temperature, bottom temperature, and bottom depth were also reported for each haul. RACE data were filtered for four species of interest: sablefish, Pacific cod, Pacific halibut, and walleye pollock. Water temperatures recorded during surveys were useful measures of environmental variation in locations in close proximity to sampled locations of the MESA longline survey.

Due to the inconsistent sampling years of the RACE trawl surveys, records of fish catches and weights were not useful to examine population dynamics between walleye pollock as an important prey species and its groundfish predators.

More comprehensive climate data collection has only been prioritized in the last two decades by the NOAA through satellite deployment and advanced buoy systems, however some environmental data are available as far back as 1800 and have been digitized for public use (NOAA, 2016). For my research, four sources of environmental data were of primary use in determining the responses of the MESA fish species to climatic variables: ICOADS, COPEPOD, WOD13, and WOA13. The ICOADS database is one of the most comprehensive collection of marine surface data from a variety of measurement technologies dating back to 1800, including ships, buoys, coastal stations, and drilling rigs (NOAA, 2016). Gridded summary data in 1° by 1° latitude-longitude boxes are available from 1960 onwards. COPEPOD is a global plankton database containing quality-controlled plankton biomass and abundance data along with any chemical and biological oceanographic variables included during collection (National Marine Fisheries Service, NOAA, 2014). As discussed in section 1.2, plankton can be a significant indicator of climate conditions affecting the oceans, and as primary producers they are a vital part of marine trophic systems. The WOD13 dataset contains quality-controlled historical and recent physical, chemical and biological oceanographic data at both standard and observed depth levels at collection sites across the world's oceans (Boyer et al., 2013). WOA13 is a set of long-term climatological means in 1° by 1° latitude-longitude grids at annual, seasonal, and monthly periods. The dataset contains water temperature, salinity, alkalinity, chlorophyll, DO, phosphate, silicate, and nitrate at standard depth levels from ocean's surface to the sea floor (Garcia et al., 2014a,b; Locarnini et al., 2013; Zweng et al., 2013). The WOA13 set uses WOD13 data to build climatologies for the past six decades as a baseline for understanding how the world's oceans are changing over space and time. Large-scale climate patterns with potential effects on the northern Pacific Ocean were considered, including PDO, multivariate ENSO, and NPGO indices (Di Lorenzo, 2018; Mantua and JISAU, University of Washington, 2016; NOAA Earth Systems Research Lab, 2017).

1.7 Modeling of Fisheries Data: The Present

Most of the previous studies discussed in section 1.3 used correlation tests, regression analysis, analysis of variance, or generalized linear models (GLMs), mainly because the studies were purely exploratory, relationships were already known from previous research, or the studies were conducted in a controlled environment. In contrast, field studies are less predictable, and relationships between catch and environmental variables more nuanced. Currently, time-series analyses and generalized additive models (GAMs) are the most common statistical tools used to analyze marine fish catch datasets for responses to environmental factors.

The generalized additive model is an extension of the GLM that uses estimated nonlinear functions instead of real numbers as coefficients, providing a more flexible method for describing nonlinear relationships between the response and predictors (Hastie and Tibshirani, 1990). GAMs therefore provide a method for modeling ecological systems where the relationships are unknown or cannot be assumed to be linear. This is particularly true for the effects of climate change on marine systems, as the systems are usually too large or complex to be tested experimentally or in controlled laboratory settings. Swartzman et al. (1992) was one of the first to apply GAMs to fish survey data not long after Hastie and Tibshirani (1990) had published the method. Since then, fisheries research has relied heavily on GAMs to describe spatial distributions of marine populations ranging from single species of shrimp and squid to large assemblages of fauna (Augustin et al., 2013; Denis et al., 2002; Hulson et al., 2013; Katsanevakis et al., 2009; Sohn et al., 2016; Sousa et al., 2006). Recently, research has focused more heavily on environmental effects on marine populations (Mourato et al., 2014; Ortega-García et al., 2015; Phillips et al., 2014; Stige et al., 2014). A few unique applications of GAMs have made use of developments in statistics to address specific problems. Arcuti et al. (2013) modeled shrimp catch using the Tweedie distribution, which is more suitable for zero-inflated marine catch data than other distributions. The Tweedie distribution has since become more common for modeling catch data in fisheries surveys. Townhill et al. (2015) used historical catch data of Atlantic cod to explain present and future population changes. Since many new statistical techniques and computational efficiencies were unavailable when management and survey

implementation began, a renewed effort to digitize logbooks and clean historical data has provided scientists with long-term data required to determine patterns of abundance and responses to long-term climate changes in fisheries stocks.

The GAM estimation algorithm has been improved several times, with the most recent improvement based on penalized least squares (Wood, 2006). There are still other options for improving estimation of these models, including optimization for zero-inflated data (Barry and Welsh, 2002) common in marine surveys and M-estimators designed to make estimation more robust (Alimadad and Salibian-Barrera, 2011; Croux et al., 2012; Wong et al., 2014). Correia and Abebe (2017) proposed a robust GAM estimation method using an R-estimator, and applied this method to describe sablefish and Pacific cod catch rates in relation to space, time, and sea surface temperature (SST). A more thorough discussion of GAM estimation and robust extensions can be found in Correia and Abebe (2017), and a summary of their results can be found in section 1.7.1.

A major problem with using GAMs in modeling environmental data is that models become overly complex for data with large numbers of covariates. Some environmental variables are also known to be related to location (such as temperature decreasing as distance from the equator increases), which requires higher-dimensional smoothers as illustrated in Arcuti et al. (2013), Augustin et al. (2013), and Stige et al. (2014). These factors thus complicate fitting of GAMs, reducing model precision, affecting model specification, and necessitating a different approach to avoid the "curse of dimensionality".

1.7.1 Results from Prior Research

Two papers (Correia and Abebe, 2017; Sun et al., 2019) explore MESA datasets using two modeling techniques: GAMs and single-index varying coefficient models. Environmental data used for modeling groundfish catches in these papers included sea surface temperature (SST), wind speed and direction, wave height, wave period, and sea level pressure. These oceanographic variables are known to change water movements and affect primary production (Bakun, 1999; Brodeur et al., 1999), which are likely to impact recruitment of groundfish (Schirripa and Colbert, 2006) and thus population size and structure.

Correia and Abebe (2017) explored the spatio-temporal and environmental effects of SST on sablefish and Pacific cod in the Gulf of Alaska and Aleutian Islands from 1979 to 2013 using MESA longline survey data. To account for spatio-temporal variation, SST was modeled using a three-dimensional smoother. Sablefish were primarily caught in the Gulf of Alaska, though heavier sablefish were found both in the central gulf and Bering Sea. Pacific cod were instead primarily caught in the Bering Sea and waters surrounding the Aleutian Islands, with heavier fish in the Bering Sea and southeast Alaska near the Dixon Entrance. It was also found that modeling catch over space and time was more effective in parsing out patterns of catch that could be separated into two main areas: catch patterns common to the Gulf of Alaska and those in the Bering Sea and Aleutian Islands region.

Sun et al. (2019) pioneered understanding how groundfish species interact in their ecosystem while also considering oceanographic influences such as wind, wave, pressure, and temperature on these fish. The three fish considered in Sun et al. (2019) (Pacific cod, Pacific halibut, and sablefish) are hypothesized to have interconnecting trophic relationships based on inspections of stomach contents (Best and St-Pierre, 1986; Gaichas et al., 2010; Moukhametov et al., 2008). No known trophic relationship models have been established for these species, therefore the single-index varying coefficient model seemed appropriate to incorporate environmental variables with fish catch rates without assuming linear relationships. It was shown that Pacific cod catches had a negative effect on Pacific halibut catch when below a threshold of environmental variables but had an increasing effect when above the threshold, while sablefish had a positive relationship on the catch of Pacific halibut when between lower and upper thresholds. The estimated environmental matrix differs for Pacific halibut, where catches were found to be decreasing independent of other fish effects. While the findings in Sun et al. (2019) are not able to elucidate which environmental variables affect the differing species, it is clear that species at similar trophic levels and with analogous habitats can still respond in opposing ways to the same environmental variables.

1.8 Modeling of Fisheries Data: The Future

One criticism of the GLMs and GAMs is their capacity as predictive models are sometimes limited, as they natively incorporate biotic interactions and stochastic effects that may vary regionally (Guisan et al., 2002). The GAMs in Correia and Abebe (2017) were compared for their predictive capability and showed improved results when using the proposed iteratively reweighted rank quasi-likelihood (IRRQL) estimation over the least squares (LS) based GAM estimation. However, fisheries researchers desire to improve predictions of population and catch estimates for future years. A preliminary extension of the GAM research will be to use $t-1$ previous years to build a spatial GAM for catch in year t for which data will already be available but not used to build the model. The predictive capacity of the model for year t can be compared to actual data for year t, then the model refit including year t data to predict year $t+1$. The process will continue until the last year of available MESA data is reached. Approaches related to this concept are surveyed in Arlot and Celisse (2010). This cross-validation procedure will enhance fisheries management models and provide more accurate forecasting of managed populations for informed regional policy-making.

The GAM formulations analyzed in Correia and Abebe (2017) only considered the effect of SST on groundfish catch. However, there are many environmental predictors believed to affect marine fish, as discussed in section 1.1.As there are several climatic and oceanic predictors that should be considered when modeling groundfish catch, there is a need for more complex models better equipped to handle dimensional expansions while also providing robust predictions. One solution for dimension reduction is the single-index model (Hardle et al., 1993; Ichimura, 1993; Stoker, 1986) given as

$$E(Y) = g(\theta^T X)$$

where Y is the response, X is the vector of predictor variables, g is an unknown functional coefficient, and θ is the unknown index vector. In this model, it is feasible to model the catch of sablefish (Y) in response to multiple environmental predictors (X), such as SST, salinity,

DO, chlorophyll a, nitrate, silicate, etc. The focus is on estimating g and θ, which is typically achieved through a least-squares approach based on local linear smoothing estimation that simultaneously estimates the functional coefficient and the index vector (Xia et al., 2002). Rank-based estimation for the single-index model has been more recently developed by Feng et al. (2012) using a Wilcoxon rank-based objective function, which is an R-estimator. Advantages of R-estimators for robust model estimation and prediction specifically for environmental data are examined in Correia and Abebe (2017). A full exploration of the nonparametric relationships of multiple environmental variables on six of the groundfish species from MESA longline surveys using rank-estimated GAMs are discussed in Chapter 2. Additionally, Chapter 3 presents a method incorporating the ability of GAMs to accommodate spatiotemporal data into a forecasting procedure and presents an application of this method to MESA longline data.

An important factor that should be considered in modeling large-scale mobile populations is location (Austin, 2002; Fisher et al., 2014; Guisan et al., 2002). For example, Suda et al. (2015) reported temperature responses in Pacific cod catch varying considerably among geographically distinct regions of the waters surrounding Japan, indicating that a single population of fish may have contrasting responses to environmental factors based on location. In GAMs, modeling these location-varying responses is achieved using higher-dimensional smoothers (Correia and Abebe, 2017; Wood, 2006; Wood and Augustin, 2002). As mentioned previously, inclusion of more covariates using higher-dimensional smoothers complicates GAM fitting. However, single-index models are unable to incorporate spatial smoothers for the predictors. However, there are model alternatives that allow for many environmental effects to be modeled over space and time. For example, the modeling of SST in the GAM formulation of Correia and Abebe (2017) as

$$Y_{(u,v,t)} = \beta_0 + z_1(u,v,t) + z_2(u,v,t) \cdot SST_{(u,v,t)} + \varepsilon_{(u,v,t)}$$

with (u, v, t) being a matrix of longitude, latitude, and time, respectively, can be considered a varying-coefficient model of the form

$$g(E(Y)) = \beta_0 + \beta_1(R_1) + \beta_2(R_1)x_1,$$

where β_0, β_1 are functional coefficients; R_1 is the matrix of longitude, latitude, and time; and x_1 is SST. As the relationships between environmental variables and groundfish catch are unlikely to be linear, a more flexible model is required for these analyses. The previous model can therefore be translated into a single-index varying coefficient model (SIVCM) (Xia and Li, 1999):

$$g(E(Y)) = f_0(\boldsymbol{\theta}^T Z) + f_1(\boldsymbol{\theta}^T Z)x_1 \tag{1.1}$$

where Z is a matrix of latitude, longitude, and time with its own coefficient θ^T; x_1 is SST. Estimation of $\boldsymbol{\theta}$ and f is least-squared-based (Xue and Pang, 2013; Xue and Wang, 2012), however rank estimation for SIVCMs has been presented in Sun et al. (2019) with an application on the MESA data focused on trophic relationships between sablefish, Pacific cod, and Pacific halibut. Biological results of Sun et al. (2019) are discussed in section 1.7.1. As mentioned in sections 1.1 and 1.3, changes in trophic interactions and species compositions, such as a decrease in foragers with simultaneous increase in predators, can signal strong regime shifts in ocean ecosystems (Anderson and Piatt, 1999). Models such as the application in Sun et al. (2019) allow for such examination of trophic relationships with potential for variable selection of environmental variables most contributive to the systemic shift. My dissertation research identifies these environmental variables and determine their contributions to variability in catches of groundfish species, while also considering the interfacing relationships of fishes with ecological interactions, such as the subgroup studied in Sun et al. (2019).

Motivated by the discussion above, one may further consider a no-intercept model of (1.1) with only one functional coefficient

$$g(E(Y)) = f_1(\boldsymbol{\theta}^T Z)\mathbf{x}$$

but a whole suite of environmental variables in the vector $\mathbf{x} = (x_1, x_2, \ldots, x_q)$ measured over space and time via the Z matrix. This idea has motivated the concept of a single-index model with lasso-type selection on the environmental variables in vector $\mathbf{x}$ across a latitude-longitude-year matrix to identify the most significant elements affecting catch rates of the seven groundfish species from a suite of oceanic and environmental factors. The application to the MESA data is presented in Chapter 4.

The models discussed previously will substantially contribute to the understanding of groundfish ecology in the northern Pacific, but there is an inherent connection among all species in the system that share abiotic driving factors such as temperature, upwelling, and chemical composition which can lead to ephemeral correlations between species, known in population ecology as the Moran effect (Sugihara et al., 2012). The Moran effect is a recently popularized concept (Engen and Sæther, 2005; Koenig, 2002; Massie et al., 2015; Ranta et al., 1997) that explains the synchronization of separated populations by shared correlations in environmental variation (Moran, 1953). Simple correlation studies often find a relationship between these populations and can prove difficult to identify for some forms of causation (Sugihara et al., 2012). A recently developed technique called 'convergent cross mapping' (CCM) extracts causes from visible effects in complex systems (Sugihara et al., 2012). The technique avoids the difficulty of "correlation does not imply causation" common to basic statistical methods by invoking causal approaches to dynamic systems. While the issue of causality has been addressed previously (Granger, 1969), a key requirement of separability for these methods is not possible in nonlinear dynamic systems common in nature. CCM separates out the effects of shared environmental variables and can correctly identify whether those variables are causally affecting a species without misidentifying couplings between unrelated species. Unlike Granger causality, which can only determine if variable X causes Y by measuring if the complete removal of variable X from the universe of all potential causal factors decreases the predictability of Y, CCM claims to be able to determine if X is causally linked to Y even if information about X relevant to predicting Y cannot be completely removed from the system by eliminating X, as is the case in deterministic dynamic systems. Therefore, CCM states that variable X causes Y (or similarly, X drives Y) if information about past states of X can be recovered from the time

series of Y, but not vice versa (Sugihara et al., 2012). This version of causality is achieved in ecological time series data by measuring the ability of the historical record of Y to reliably estimate states of X with increasing predictability as the time-series length increases, which occurs only if Y is causally influenced by X. CCM can also determine the strength of the causal relationship, and by the transitive nature of causality, indirect causal links can also be established in large interaction networks such as marine systems. This strength of CCM can also be used to determine whether a predator-prey system is more heavily top-down (predator) controlled or bottom-up (prey) controlled, as illustrated in the application by Sugihara et al. (2012) on the classical *Didinium-Paramecium* system.

An interesting aspect of CCM is the incorporation of multiple species into the model to determine causal effects. Sugihara et al. (2012) successfully applied CCM to a laboratory predator-prey system and sardine-anchovy-SST dataset modeling the California current ecosystem. They also showed the method was effective in a simulated five-species model and could correctly distinguish between correlation and causality in a system with two non-interacting populations sharing a common environmental forcing variable (e.g. SST). Therefore, not only would CCM be suitable for determining if an environmental variable X causes changes in species Y, but it would also be able to determine causal relationships among species and correctly identify spurious correlations previously believed to be causal for specific populations. The northern Pacific groundfish system and related environmental variables determined by GAMs, SIMs, and SIVCMs will test the flexibility and accuracy of CCM by attempting to incorporate seven fish species and multiple environmental factors into a dynamical systems model while accurately determining causal effects within the system. Recent extensions to CCM for short time series highly replicated in space (Clark et al., 2015) and distinguishing between synchrony from strong unidirectional forcing and bidirectional cross mapping (Ye et al., 2015) will be particularly useful in tackling the unique nature of the MESA data and associated climatic variables. The application of spatial CCM to MESA groundfish survey data is provided in Chapter 5.

Chapter 2

Effects of physical, chemical, and biological covariates affecting adult groundfish catches and mean weights in the Gulf of Alaska

Abstract

The North Pacific Ocean has undergone several regime shifts in which the environment and community structure have changed substantially. It is also home to many commercially important fisheries for the United States. I investigated the influence of multiple relevant environmental variables on the catches per unit effort (CPUE) and mean weights of six groundfish species in the Gulf of Alaska from 1979–2013 using generalized additive models. Comprehensive analyses of effects of chemical and physical ocean attributes on fishes in the wild has not been performed on such a scale before. Common significant environmental variables for all six groundfish CPUEs were sea surface temperature and average wave period. Bottom temperature, wind speed, and nutrients phosphate, silicate and nitrate in deeper waters (600-900m) were significant contributors to CPUE for most study species. Deepwater (900m) salinity was a common significant variable to all rockfish CPUEs. Knowing the specific relationship of these variables to commercially and ecologically important fish species outside of laboratory conditions is crucial to understanding how these fish are responding to current changes in the marine system and in planning effective management strategies for future environmental regime shifts.

Keywords and phrases: additive models, sablefish, rockfish, Pacific cod, Pacific halibut, environmental effects

2.1 Introduction

The Northern Pacific Ocean is a main focus in conversations about climate change effects on marine systems, but limited research has been done on the effects of changes in physical and chemical oceanic factors on major fish species in the wild. Laboratory studies have considered the effects of extreme temperature changes on some species, but effects of key variables such as salinity, plankton concentrations, and dissolved oxygen in natural landscapes are much less well known, particularly in the Pacific Ocean (Pörtner et al., 2001; Rouyer et al., 2014). Typically temperature is the simplest variable to measure over large regions, and because of its effect on individuals movement and behavior, temperature is most often considered the primary causal agent of change in marine fish populations (Baumann et al., 2006; Monllor-Hurtado et al., 2017; Suda et al., 2015). Effects of large-scale climate patterns are instead used as proxies for variations in individual environmental conditions, but it is difficult to disentangle which environmental variables are primarily responsible for changes in marine populations (Stige et al., 2006).

The North Pacific marine system, like most temperate and polar regional ocean systems, exhibits seasonal fluctuations in nutrients and plankton concentrations. Wind-induced upwellings in late spring and summer lead to increased accumulation of phytoplankton biomass, peaking in September. Shelf and slope concentrations of chlorophyll exhibit marked seasonality, whereas deepwater concentrations vary much less (Sackmann et al., 2004). Upper-water column concentrations of nitrate, silicate, and phosphate are at their maximums in March, before being utilized by phytoplankton, but after strong winter mixing has brought these nutrients up from the deep. Annual minimums for these nutrients in the surface waters usually occur in August and September, after they have been depleted by summer plankton growth and blooms (Childers et al., 2005; Wong et al., 2002a). Silicate is particularly important in the North Pacific, as the region is rich in diatoms that require higher than average silicate levels to grow (Wong and Matear, 1999; Wong et al., 2002a). Silicate reaches maximal levels in March and April and decreases over the summer from diatom utilization in the Alaskan Gyre, reaching

minimum levels in August (Harrison et al., 2004; Wong and Matear, 1999). Dissolved oxygen (DO) concentrations on the surface waters of the North Pacific also maximize in spring after winter mixing events and photosynthetic activity of spring phytoplankton growth. DO concentrations decrease on the surface over the summer from respiration by algae and pelagic organisms, reaching a minimum in September (Boyer et al., 1999; Emerson, 1987). Salinity levels also peak in the early spring (March and April) in the northeastern North Pacific and Alaskan Gyre. Late spring and summer glacial melts push freshwater into nearshore ocean waters, decreasing salinity to minimum levels in September and October (Bingham et al., 2010). Minimal pH is observed from January to March for waters near Alaska, as acidic deep waters are brought to the surface by upwellings, whereas pH peaks in August due to photosynthetic utilization of CO_2 during the spring and summer (Fietzke et al., 2015; Takahashi et al., 2014).

Deeper waters (up to 275 m) exhibit contrasting maximal and minimal cycles from the surface-level waters. Nitrate is at a maximum in August as plankton die and break down, whereas silicate and phosphate peak in April. Low-pressure winter storms with high wind and wave activity replenish surface water nitrate levels from decomposing material in deep waters (Wong et al., 2002a). Nitrate, silicate, and phosphate all show minimums around February in deeper waters (Childers et al., 2005). Eddies are also important transport systems for chlorophyll and phytoplankton distribution in surface waters and carry excess nutrients from decomposition upward from deep waters, where they are subject to wind-forced mixing that enriches plankton communities (Crawford et al., 2007). Highest concentrations of chlorophyll have been observed in summer for the Aleutian Islands (Childers et al., 2005; Fietzke et al., 2015). DO in and below the thermocline peaks in August; however maximum DO concentrations are observed in the mixed layer in May and minimum concentrations in August (Emerson, 1987).

The aim of this study is to determine the shape of relationships of seasonal variations in ocean nutrients and climatic activity on groundfish species in the Northern Pacific. I hypothesized that effects of well-studied attributes such as water and temperature on groundfish will follow similar trends in the wild as are seen in experimental and laboratory studies, and environmental variables will exhibit nonlinear trends on groundfish catches and weights.

2.2 Methods

I considered the effects of environmental covariates on six focal groundfish species in the Northern Pacific Ocean: sablefish (*Anoplopoma fimbria*), Pacific cod (*Gadus macrocephalus*), Pacific halibut (*Hippoglossus stenolepis*), shortspine thornyhead (*Sebastolobus alascanus*), short-raker rockfish (*Sebastes borealis*), and rougheye rockfish (*Sebastes aleutianus*). I obtained catch and weight data for these groundfish species for 1979-2013 from the NOAA Marine Ecology Assessment (MESA) program (Alaska Fisheries Science Center, 2019a). Defined locations (hereafter stations) were sampled annually during the summer months using longlines, and groundfish catch per unit effort (CPUE) and mean weights per survey in kilograms were recorded. Only mature individuals were included in the survey data. Field collection methods and calculation of summary statistics are detailed in Sigler and Lunsford (2009). During the years of 1988-1994, both the US and Japan conducted longline surveys in the Bering Sea and waters of the Aleutian Islands, resulting in repeated measures of groundfish CPUE and mean weights for stations in those regions. For stations with repeated values for a given year, the mean of the two survey values was taken.

Water temperatures at the bottom of the water column (hereafter bottom temperature) for locations in close proximity to the stations where groundfish were sampled were obtained from the NOAA Resource Assessment and Conservation Engineering (RACE) Division's groundfish assessment program (Alaska Fisheries Science Center, 2019b; von Szalay and Raring, 2016). Nutrients limited to nitrate, phosphate, and silicate at depths of 75 m and 900 m, along with chemical and biological tracers including alkalinity, chlorophyll, DO, and salinity for depths of 75, 400, and 900 m were obtained from the NOAA's 2013 World Oceans Database (WOD) (Boyer et al., 2013). A depth of 75 m was chosen to best represent the mixed layer depth (MLD), as MLD can range from 25 to 100 m throughout the year along the coast of Alaska (Ohno et al., 2009). Most effort (i.e. percent of longlines deployed at the station at a given depth range) across all areas of the MESA longline survey was concentrated at the 400 to 600 m depths (Echave et al., 2013), thus chemical and biological variables likely to affect adult groundfish of spawning age were obtained at 400 m depths. Finally, 900 m depths were chosen

to represent bottom ocean nutrient levels, as 75% of station depths reached 1000 m (Echave et al., 2013).

Other environmental variables of interest were obtained from the National Data Buoy Center, a collection of marine environment observations from buoy, ship, and land stations (National Data Buoy Center, 2018a). Included variables were air temperature (ATMP) in degrees Celsius, sea level pressure (PRES) in hectopascals (hPa), wind speed in m/s (WSPD), surface water temperature in degrees Celcius (WTMP), and significant wave height in meters (WVHT). Formal definitions and details of these summary variables are given in National Data Buoy Center (2018b). Concentrations of zooplankton biomass in number per cubic meter for September 1979 through August 2007 were obtained from all available samples across the Gulf of Alaska and the Bering Sea maintained in the NOAA's global plankton database known as COPEPOD (National Marine Fisheries Service, NOAA, 2014).

Since all environmental factors are measured on differing temporal scales and spatial points to each other and to the fish data, interpolation of monthly averages for environmental variables to the fish station locations for years 1979-2013 was performed via spatiotemporal inverse distance weighting (Li et al., 2014a) implemented in R using the package **geosptdb** (Melo and Melo, 2015). In order to reduce issues with multicollinearity, backward variable elimination of explanatory environmental variables using computed variance inflation factors (VIF) was performed. The VIF for explanatory variable i was calculated as $1/(1 - R_i^2)$ where R_i^2 was the correlation coefficient of the linear model with variable i regressed against all other explanatory variables. VIFs for all variables were computed, and the variable with largest VIF that exceeded a value of 5 was removed (Neter et al., 1996; O'Brien, 2007a). The process was repeated after each removed variable, until no variables had a VIF < 5.

A seasonal amplitude was used to summarize the monthly means of the selected explanatory variables into meaningful yearly measures (Takahashi et al., 2014). For chemical and nutrient tracers, seasonal amplitude was calculated as the August-September-October mean value minus the March-April-May mean value. For physical variables, the seasonal amplitude was calculated as the June-July-August mean minus the December-January-February mean. Alternative calculations for physical versus chemical variables was optimized to capture the

difference in maximal and minimal seasons of each variable type. The sign and magnitude of seasonal amplitude indicated the phase of a given variable from winter to summer, where a positive value indicated that the environmental variable increased from winter to summer, while a negative value indicated that the environmental variable was maximal in winter and decreased throughout the spring and summer. Values close to zero indicated little change between summer and winter monthly means for that environmental factor.

Since the fish targeted in the MESA survey are adults, environmental variables were lagged based on the sexual maturity of the groundfish. Sablefish and Pacific cod female maturity is reached around 5 years old (Mason et al., 1983; Stark, 2007). Average age of Pacific halibut females reaching sexual maturity is 12 years whereas maturity for males is 8 years (Clark et al., 1999). Rockfish sexual maturity is typically attained from 3-7 years of age (Echeverria, 1987). Therefore, lags of 5 years were used for environmental variables in models with sablefish, rockfish, or Pacific cod CPUEs and weights as responses. Lags of 10 years were used for environmental variables in models with responses of Pacific halibut CPUE and weight.

A generalized additive model (GAM) of the form

$$E(y) = f_0 + \sum_{j=1}^{p} f_j(x_j) + z_1(u, v, t) + \varepsilon \ ,$$

was fit for each of the six groundfish, where y is the CPUE or mean weight of a species for a given year and location, x_j are the seasonal amplitudes of environmental factors, f_0 is the intercept function, f_j are the smooths relating the environmental variables to the response, z_1 is a three-dimensional tensor smooth, u, v is the longitude-latitude pair at which each fish station is located, and t is the year for which measures were recorded. Since the distribution of some species' CPUEs and mean weights were skewed or heavy-tailed, rank estimation of GAMs as described in Correia and Abebe (2017) was employed. Significant explanatory variables were plotted to visualize their effects on the specified responses. All analyses were performed in the R open-source software environment (R Core Team, 2017).

2.3 Results

All physical and chemical marine variables remained after backwards selection using the VIF $<$ 5 criterion. The following variables were therefore used in the models and abbreviated as follows throughout the text: ATMP; PRES; WSPD; WTMP; WVHT; bottom temperature; alkalinity at 75, 400, and 900 m (Alk_75m, Alk_400m, Alk_900m, respectively); chlorophyll at 75, 400, and 900 m (Chl_75m, Chl_400m, Chl_900m, respectively); nitrate at 75 and 900 m (NO_3_75m, NO_3_900m, respectively); DO at 75, 400, and 900 m (Oxy_75m, Oxy_400m, Oxy_900m, respectively); phosphate at 75 and 900 m (Phos_75m, Phos_900m, respectively); salinity at 75, 400, and 900 m (Sal_75m, Sal_400m, Sal_900m, respectively); and silicate at 75 and 900 m (Sil_75m, sil_900m, respectively).

All models for fish CPUEs and mean weights had significant spatiotemporal terms, indicating that groundfish CPUEs and mean weights varied significantly over space and time (Tables A.1 to A.12). Sablefish catches increased in the western Gulf of Alaska from the mid 1980s (Figure A.7), whereas Pacific cod catch decreased during the 1980s and early 1990s in the Bering Sea (Figure A.8). Catches of Pacific halibut and the three rockfish species increased throughout the Gulf of Alaska and Bering Sea over the 35 year period of the study (Figures A.10 to A.12). Sablefish and Pacific cod weights became more spatially homogenous and increased over the period of record, whereas weights of Pacific halibut became more spatially heterogeneous over time (Figures A.13 to A.15). Pacific halibut CPUE and weights of all three rockfish species remained relatively consistent over space and throughout the study period (Figures A.9 and A.16 to A.18).

I first examined the environmental predictors of the GAMs with CPUEs of the six groundfish as responses. The GAM for sablefish CPUE had a good fit (adjusted $R^2 = 0.928$) and had 17 significant environmental variables (Table A.1). Decreasing sablefish CPUE was related to positive values of WVHT, bottom temperature, Sal_400m, and Sil_75m, with Sal_400m and Sil_75m exhibiting approximately linear relationships to sablefish CPUE. Positive values of ATMP, WSPD, WTMP, NO_3_75m, NO_3_900m, and Phos_900m are related to increasing sablefish CPUE, whereas the relationships of PRES, plankton, Chl_75m, Chl_400m, Chl_900m,

Oxy_75m, and Sil_900m to sablefish CPUE exhibited fluctuations with no directional trend (Figure 2.1). Fourteen variables were significant to Pacific cod CPUE and the model fit the data well (adjusted $R^2 = 0.914$, Table A.2). Pacific cod CPUE increased with positive values of ATMP, PRES, WTMP, plankton, Oxy_75m, Oxy_900m, Phos_75m, and Sil_75m, whereas positive values of Phos_900m were associated with decreasing CPUE of Pacific cod. Positive values of WSPD, WVHT, Alk_900m, Sal_75m, and Sil_900m did not contribute significant overall changes to Pacific cod CPUE, however nonlinear fluctuations in their effects on cod CPUE were apparent. The relationships of PRES, plankton, and Phos_900m to Pacific cod CPUE were approximately linear (Figure 2.2). The GAM for Pacific halibut CPUE had 17 significant variables and a good fit to the data (adjusted $R^2 = 0.807$, Table A.3). Positive values of WTMP, WVHT, Alk_75m, Chl_400m, Chl_900m, NO_3_75m, and Oxy_75m were associated with increasing values of Pacific halibut CPUE, whereas positive values of PRES, bottom temperature, Alk_900m, Oxy_400m, Oxy_900m, and Sal_400m were related to decreasing values of halibut CPUE. Bottom temperature, Alk_75m, Chl_75m, Oxy_900m, and Sal_400m had approximately linear relationships with Pacific halibut CPUE (Figure 2.3).

GAMs for the CPUEs of the three rockfish species had some variables in common, however few shared the same relationships to rockfish CPUEs over all three species' models. The GAM for shortspine thornyhead CPUE was a good fit to the data (adjusted $R^2 = 0.902$) and had 16 significant environmental factors (Table A.4). Positive values of ATMP, WVHT, plankton, Alk_75m, NO_3_75m, and Sal_900m were related to decreasing values of shortspine thornyhead CPUE, whereas increasing CPUE for shortspine thornyhead was associated with positive values of PRES, WSPD, WTMP, Alk_900m, Chl_400m, Oxy_75m, Oxy_400m, and Oxy_900m (Figure 2.4). The relationships of Alk_75m, Alk_900m, and Oxy_75m to shortspine thornyhead CPUE were approximately linear, whereas Chl_75m and bottom temperature showed nonlinear patterns that had negative impacts on shortspine thornyhead CPUE when their values were negative. Fifteen variables were significantly related to rougheye rockfish CPUE, and the GAM was a good fit to the data (adjusted $R^2 = 0.777$, Table A.5). Generally decreasing values of rougheye rockfish CPUE were related to positive values of PRES, WTMP, bottom temperature, Oxy_75m, and Sil_900m, whereas increasing rougheye rockfish CPUE was related to

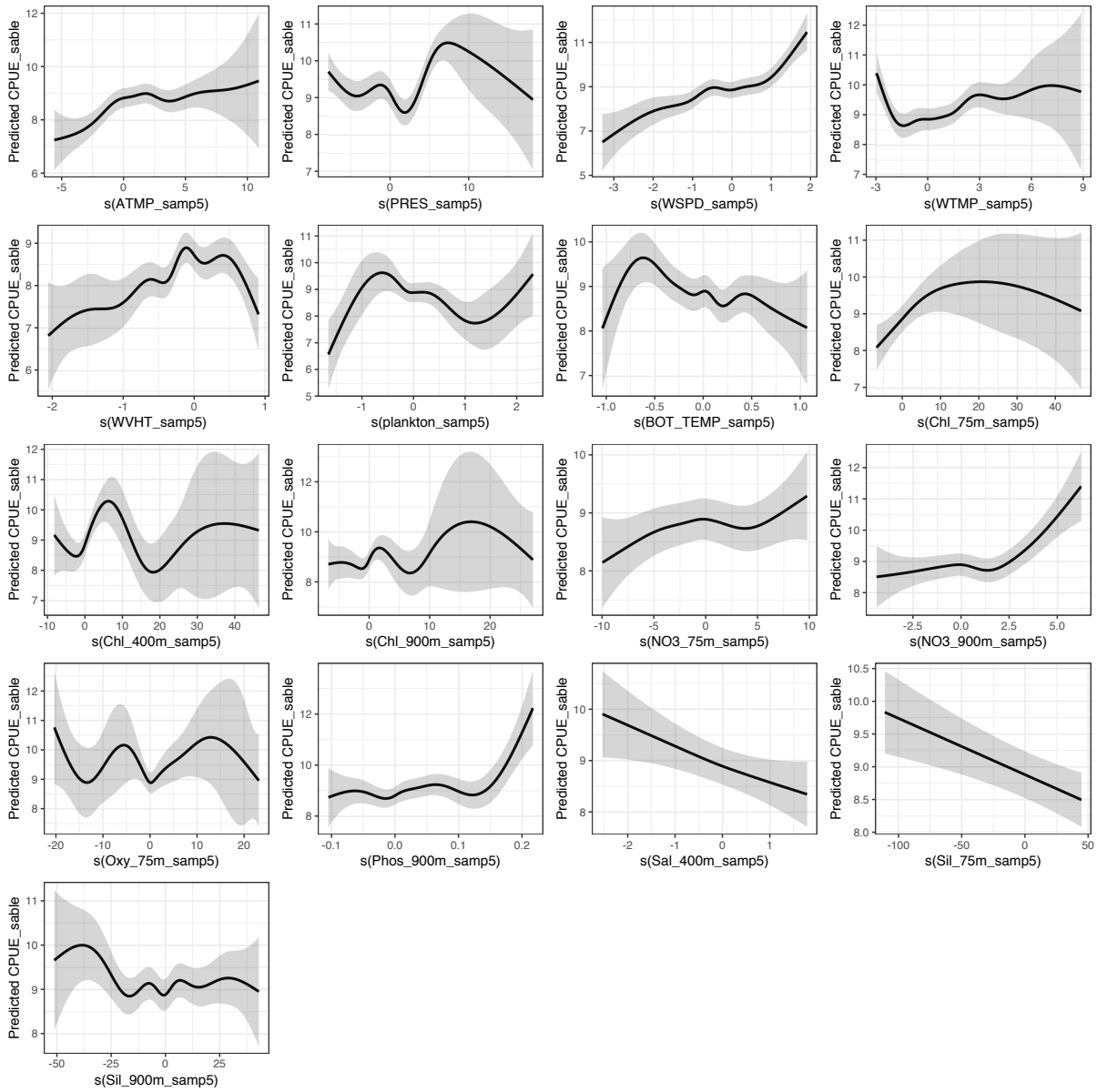


Figure 2.1: Significant smooths in rank GAM for sablefish CPUE.

positive values of NO_3_900m and Oxy_400m. Positive values of ATMP, WSPD, WVHT, plankton, Alk_900m, Chl_400m, Chl_900m and Sal_400m did not heavily impact rougheye rockfish CPUE, however negative values had an impact on rougheye rockfish CPUE and nonlinear relationships were apparent. Negative, approximately linear relationships between rougheye rockfish CPUE and bottom temperature and NO_3_900m were observed (Figure 2.5). The GAM for shortraker rockfish CPUE fit the data well (adjusted $R^2 = 0.838$) and 14 variables were found to be significant in the model (Table A.5). Positive values of WTMP, plankton, Alk_900m, and Chl_75m were associated with approximately increasing values of shortraker rockfish CPUE,

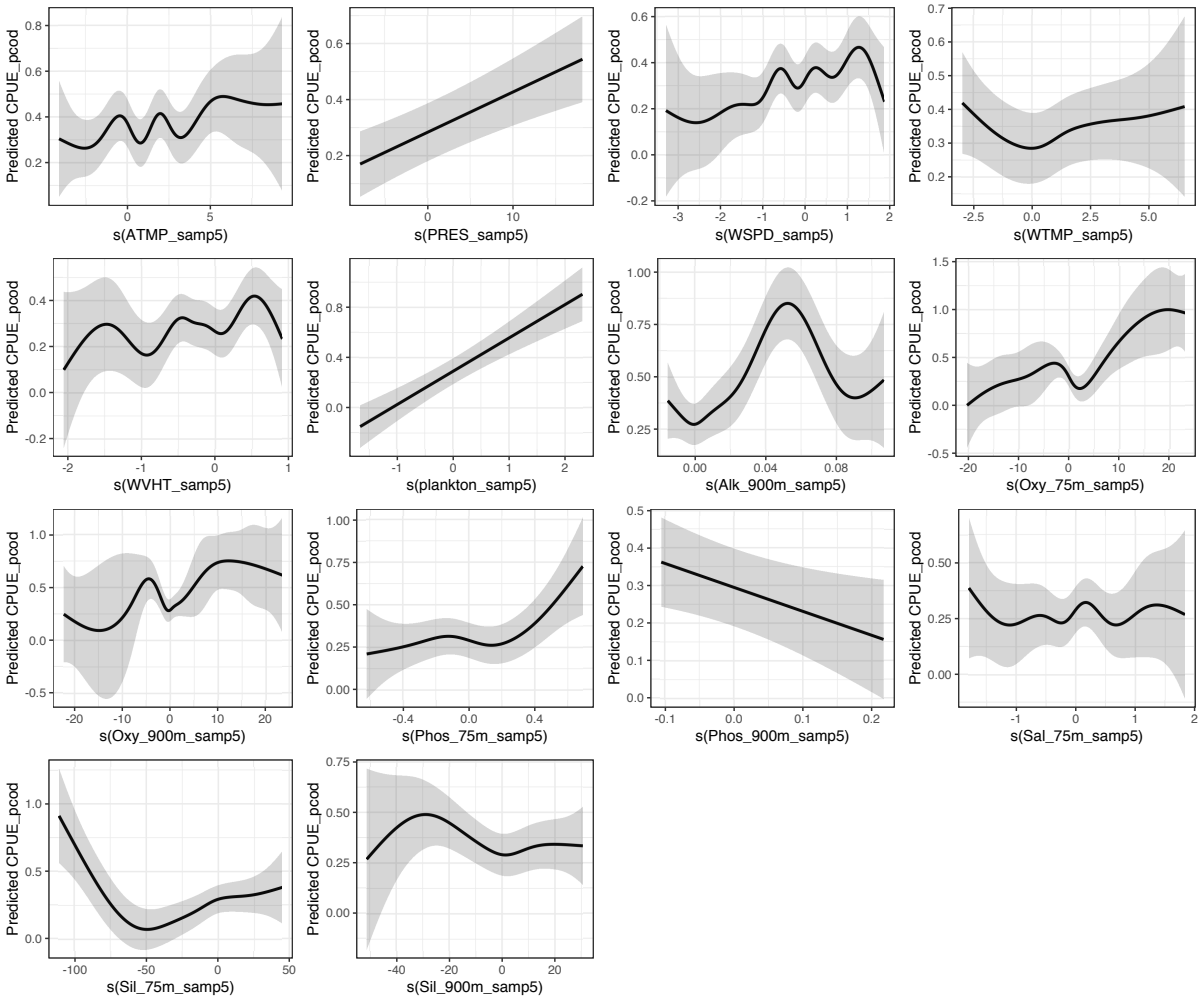


Figure 2.2: Significant smooths in rank GAM for Pacific cod CPUE.

whereas positive values of PRES, WSPD, bottom temperature, Alk_75m, Sal_75m, Sal_400m, Sil_75m, and Sil_900m were related to decreases in shortraker rockfish CPUE (Figure 2.6). Bottom temperature exhibited an approximately linear relationship to shortraker rockfish CPUE, whereas negative values of WHVT and Chl_400m resulted in nonlinear decreased in shortraker rockfish CPUE.

I then looked at the environmental predictors of the GAMs with mean weights of the six groundfish as responses. The GAM for sablefish mean weights moderately fit the data (adjusted $R^2 = 0.604$) and four environmental variables were significant to sablefish weight (Table A.7). Decreasing values of sablefish weight were associated with positive values of ATMP, WTMP and bottom temperature, whereas overall increasing sablefish weight was associated with negative values of WSPD. Nine environmental variables were significantly related to Pacific cod

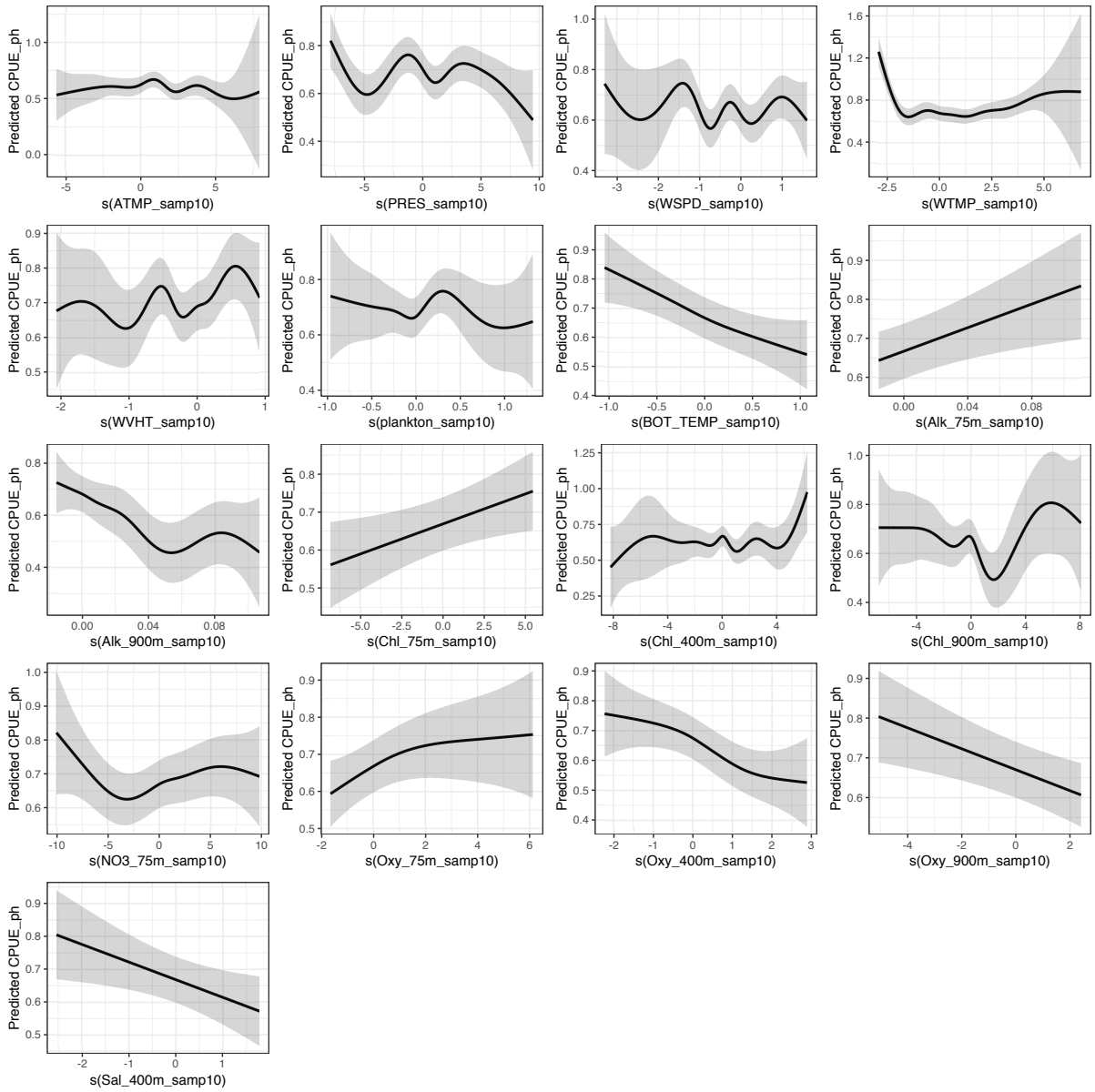


Figure 2.3: Significant smooths in rank GAM for Pacific halibut CPUE.

weight, and the GAM fit the data moderately well (adjusted $R^2 = 0.460$, Table A.8). Negative, approximately linear relationships between WVHT, NO_3_75m, Sal_75m, and Sil_75m and weights of Pacific cod were observed, while there was a positive, approximately linear relationship between Alk_400m and Pacific cod weight. Large positive values of PRES (> 10) were associated with increasing weights of Pacific cod, whereas negative values of bottom temperature and positive values of Sil_900m were associated with decreasing Pacific cod weights (Figure A.2). The model for Pacific halibut weight moderately fit the data (adjusted $R^2 = 0.468$), and eight variables were significant to Pacific halibut weight (Table A.9). Positive values of

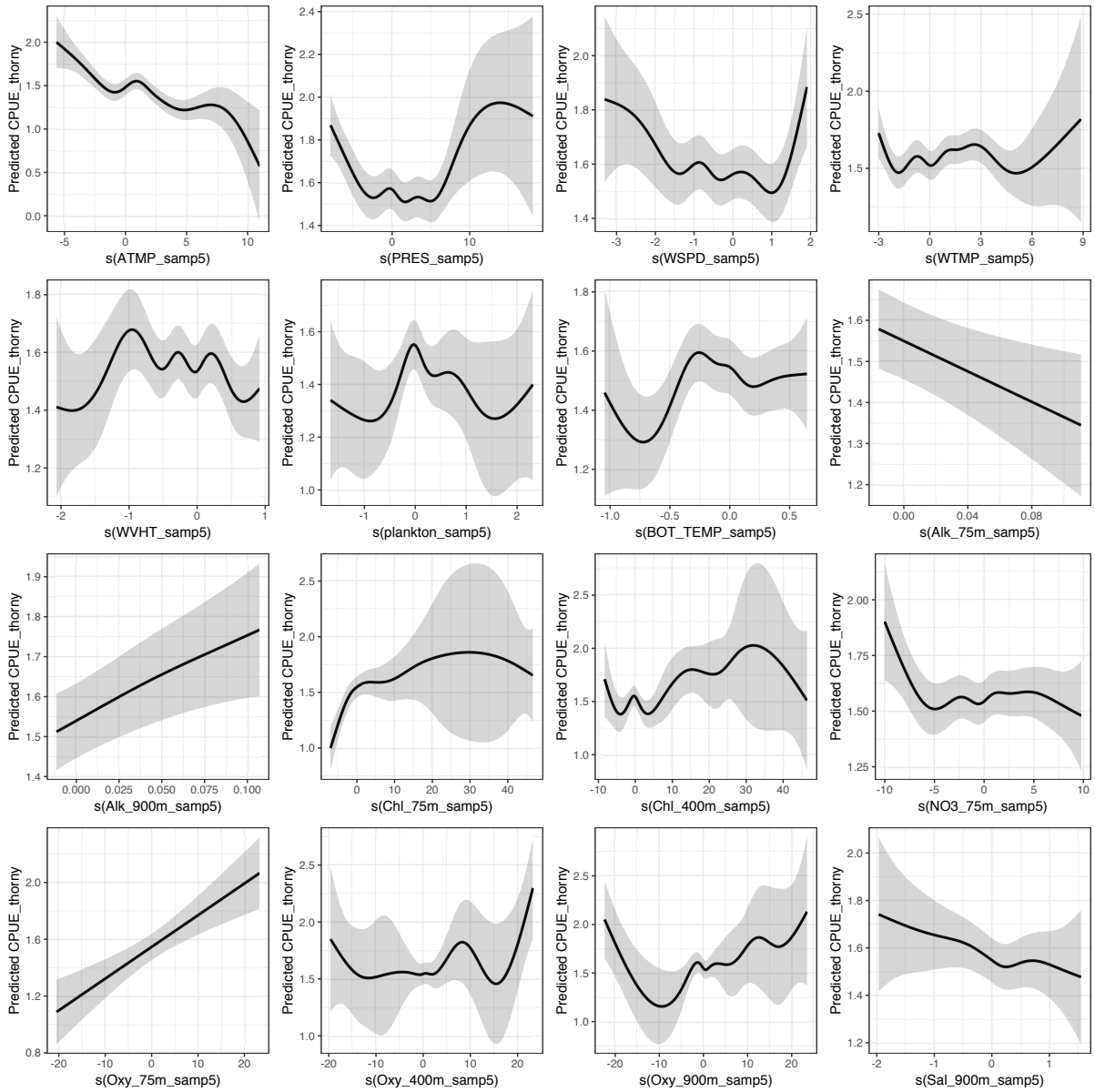


Figure 2.4: Significant smooths in rank GAM for shortspine thornyhead CPUE.

Chl_75m were related to an increase in halibut weight, whereas positive values of Sal_900m were associated with halibut weight decrease. As plankton and Alk_900m values became more negative, halibut weight increased. WSPD and Phos_900m had positive, approximately linear relationships to halibut weight, whereas Alk_75m and Oxy_400m had negative linear relationships to halibut weight (Figure A.3).

The models for rockfish weights shared DO as common significant variables. Three environmental factors were significantly related to shortspine thornyhead weights, and the model fit the data well (adjusted $R^2 = 0.844$, Table A.10). Positive values of Oxy_75m, Oxy_400m, and

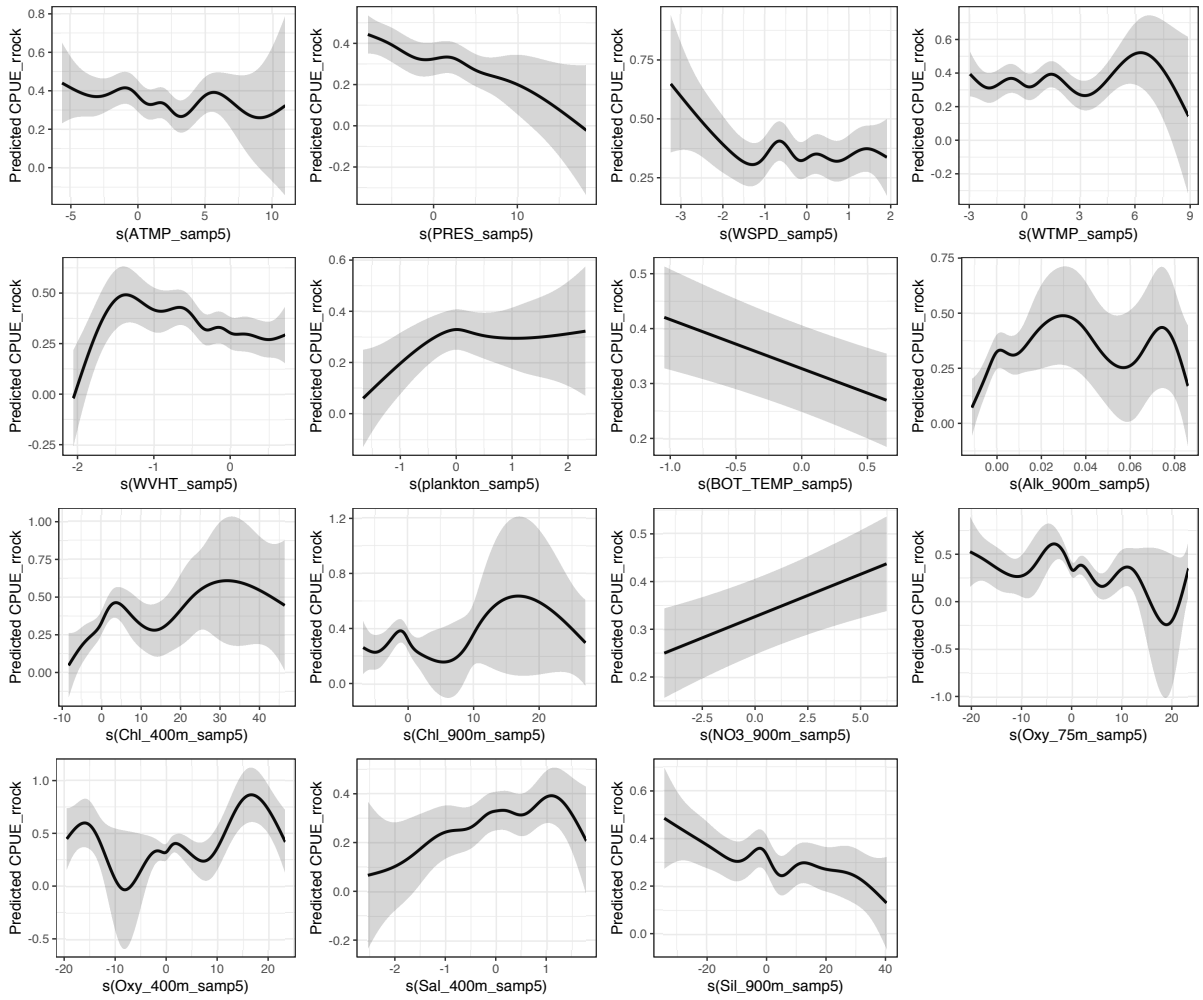


Figure 2.5: Significant smooths in rank GAM for rougheye rockfish CPUE.

Oxy_900m were associated with decreasing shortspine thornyhead weights, however the relationship for positive values of Oxy_400m was parabolic (Figure A.4). The GAM for rougheye rockfish weights fit the data moderately (adjusted $R^2 = 0.638$). Five variables were significant to rougheye rockfish weights, of which two were approximately linear, WSPD and NO_3_900m (Table A.11). Positive values of WSPD and NO_3_900m were associated with decreasing rougheye rockfish weights, while positive values of Chl_75m and Chl_400m and negative values of Oxy_75m were related to increasing weights (Figure A.5). The GAM for shortraker rockfish weights moderately fit the data (adjusted $R^2 = 0.604$). Five variables were significant to shortraker rockfish weights, with positive values of ATMP, Alk_75m, and Sal_400m increasing weight and positive values of Oxy_75m and Sal_75m associated with decreasing weights (Table A.12, Figure A.6).

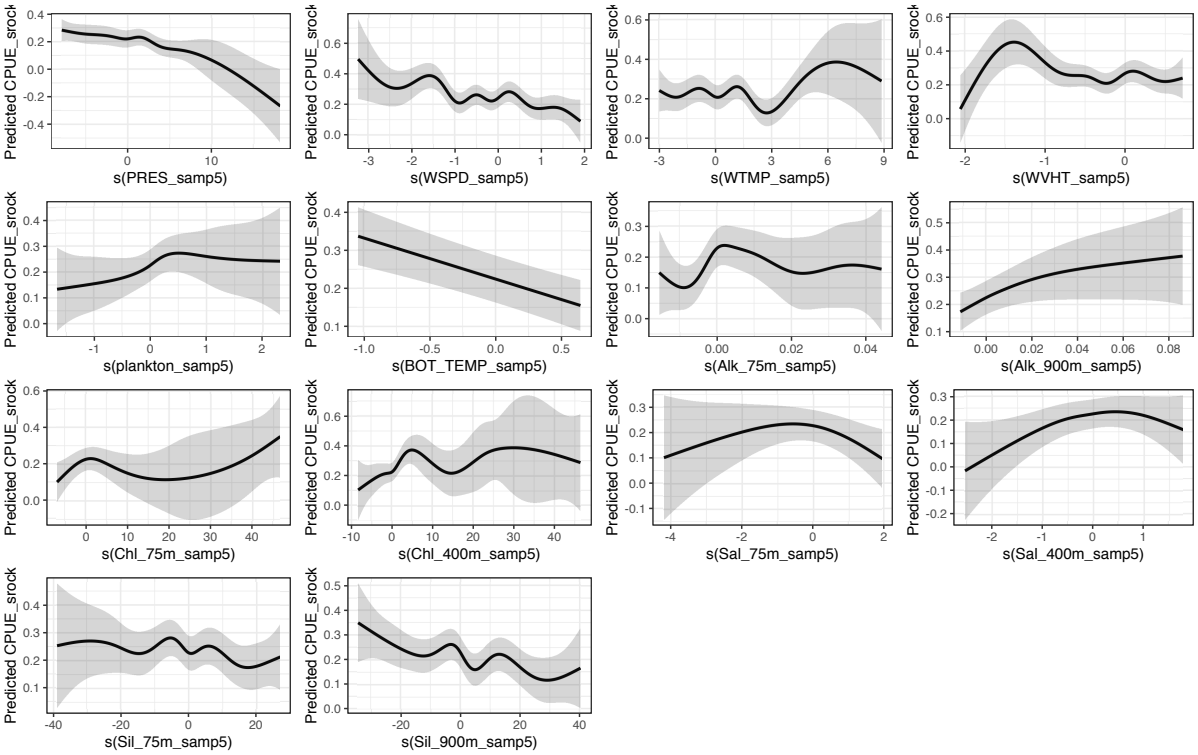


Figure 2.6: Significant smooths in rank GAM for shortraker rockfish CPUE.

2.4 Discussion

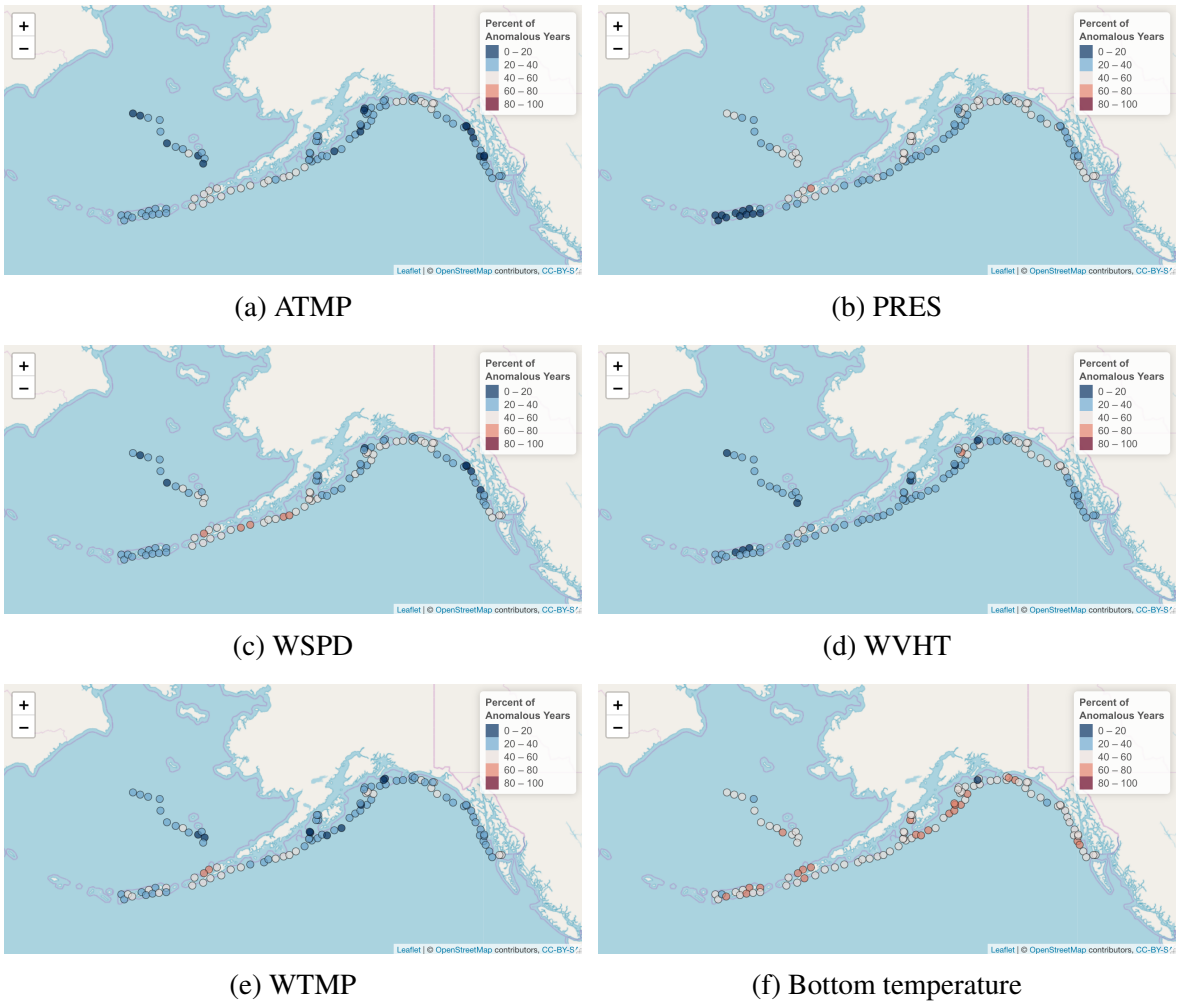
Nonlinear dynamics in marine ecosystems both between species and of external environmental effects on marine populations are well known, and more recent studies have attempted to account for such nonlinearity (Frainer et al., 2017; Hewitt et al., 2016; Liu et al., 2012). It is clear from the smooths of GAMs used in these analyses that many of the environmental variables collected in the Gulf of Alaska and Bering Sea are also likely to exhibit nonlinear relationships to groundfish catches and weights. Nonlinear relationships between water temperature and weight, prey consumption, and time to reach 50% maturity have been observed for sablefish and rockfish species in the California Current (Harvey, 2009). In the Gulf of Alaska, wind, air temperature, and predation were also observed to have nonparametric effects on the survival of pollock, an important prey species for the sablefish, Pacific cod, and Pacific halibut considered in this study (Ciannelli et al., 2004). It is therefore unsurprising that nonlinear environmental effects on groundfish catches and weights were evident in these analyses, and it is likely that these are manifested both directly and indirectly.

Before discussing more specific inference regarding some of the environmental and physical variables in the GAMs, I will first focus on some of the spatial and temporal trends of the seasonal amplitudes of these variables in the North Pacific region. Negative seasonal amplitudes indicated that summer or fall means were lower than winter or spring means, whereas positive seasonal amplitudes indicated higher means in the latter part than the early part of the year. The magnitude of these seasonal amplitudes indicated the amount by which the fall and spring means (or in the case of physical variables, summer and winter means) varied. Based on the minimal concentrations of chemical variables observed in the months of August, September, and October and maximal concentrations observed in March, April, and May in the Gulf of Alaska and Bering Sea regions, alkalinity nitrate, DO, phosphate, salinity, and silicate were all expected to show negative seasonal amplitudes in surface and near-surface waters. In many cases, seasonal patterns hold for deeper waters, although nitrate was at a maximum in August and minimum in February for waters >250 m. From known cycles of biological activity in the waters around Alaska, zooplankton and chlorophyll were generally expected to exhibit positive seasonal amplitudes. Winter storms along the coast of Alaska create intense downwelling winds, promoting eddies and wave activity that transports nutrients across the ocean and from the deep to surface waters (Whitney et al., 2005; Whitney and Robert, 2002). Seasonal amplitudes for wind and wave activity were therefore expected to be negative under the conditions of strong winter storms and calmer summer weather. Air pressure, which is typically low during storm activity, was expected to have a positive seasonal amplitude. Air temperature, water temperature, and bottom temperature being typically maximal in the summer months and minimal in the winter months resulted in a positive expected seasonal amplitude for these measures.

To summarize these seasonal amplitudes over all stations for the period of record, I took the sign of the value for each year and location, then calculated the percentage of values that did not exhibit the expected signs as indicated from observed patterns in biology or climatology. For simplicity, only 75-m depths of chemical and biological variables were plotted. Bottom temperature and plankton had more than half of the years exhibit opposing trends (negative values rather than the expected positive values) for many stations along the coast of Alaska (Figures 2.7f and 2.8a). Oxygen, phosphate, and silicate at 75 m also saw at least half the years

exhibiting opposite trends (positive values rather than negative expected seasonal amplitudes) for most stations in the MESA longline survey (Figures 2.8e, 2.8g and 2.8h). More than 60% of the seasonal amplitudes for chlorophyll, alkalinity, and nitrate were opposite their expected trends for a large portion of the Alaskan stations (Figures 2.8b, 2.8c and 2.8f). It should be noted that silicate and phosphate may only exhibit short periods of interannual variability in deep waters with minimum concentrations in February and maximums in April (Childers et al., 2005). The "unexpected" seasonal amplitudes in 75-m nutrient concentrations and biological activity may be partially explained by changes in the MLD and surface water nutrient cycling (Ohno et al., 2009). Alkalinity and salinity exhibited some spatial trends in the anomalous seasonal amplitudes. Stations demonstrating mostly negative chlorophyll (opposite the positive values expected) were clustered around the central and eastern Gulf of Alaska (Figure 2.8d). Locations with mostly positive seasonal amplitudes for alkalinity (opposite the expected negative values) were concentrated along the Aleutian Islands and the western Gulf of Alaska (Figure 2.8c). Changes in the North Pacific Current and Alaskan Gyre circulation over several decades have altered the upwelling system in the Gulf of Alaska (Buil and Di Lorenzo, 2015). These changes are likely to modify the transport of nutrients and plankton across the Alaskan shelf, where distinct spatial clustering of mostly positive or negative seasonal amplitudes such as those seen in our analyses could be expected.

For most GAMs where plankton and chlorophyll were both significant variables, which include those for sablefish, Pacific halibut, rougheye rockfish, and shortspine thonyhead CPUEs, zooplankton and chlorophyll exhibited similar patterns within each model. Increasingly negative differences between fall and spring means translated to decreased CPUEs for sablefish, rougheye rockfish, and shortspine thornyhead. A reasonable explanation of these mirrored patterns is that chlorophyll indicates phytoplankton biomass, and the relation of zooplankton distribution to its phytoplankton prey's distribution means that increased chlorophyll will approximately translate to accumulation of zooplankton biomass (Kang and Ohman, 2014; Longhurst, 1976; Moeller et al., 2019; Taniguchi, 1973). A positive correlation between summer plankton biomass and winter wind speeds has also been reported in the Bering Sea, though



(a) ATMP (b) PRES

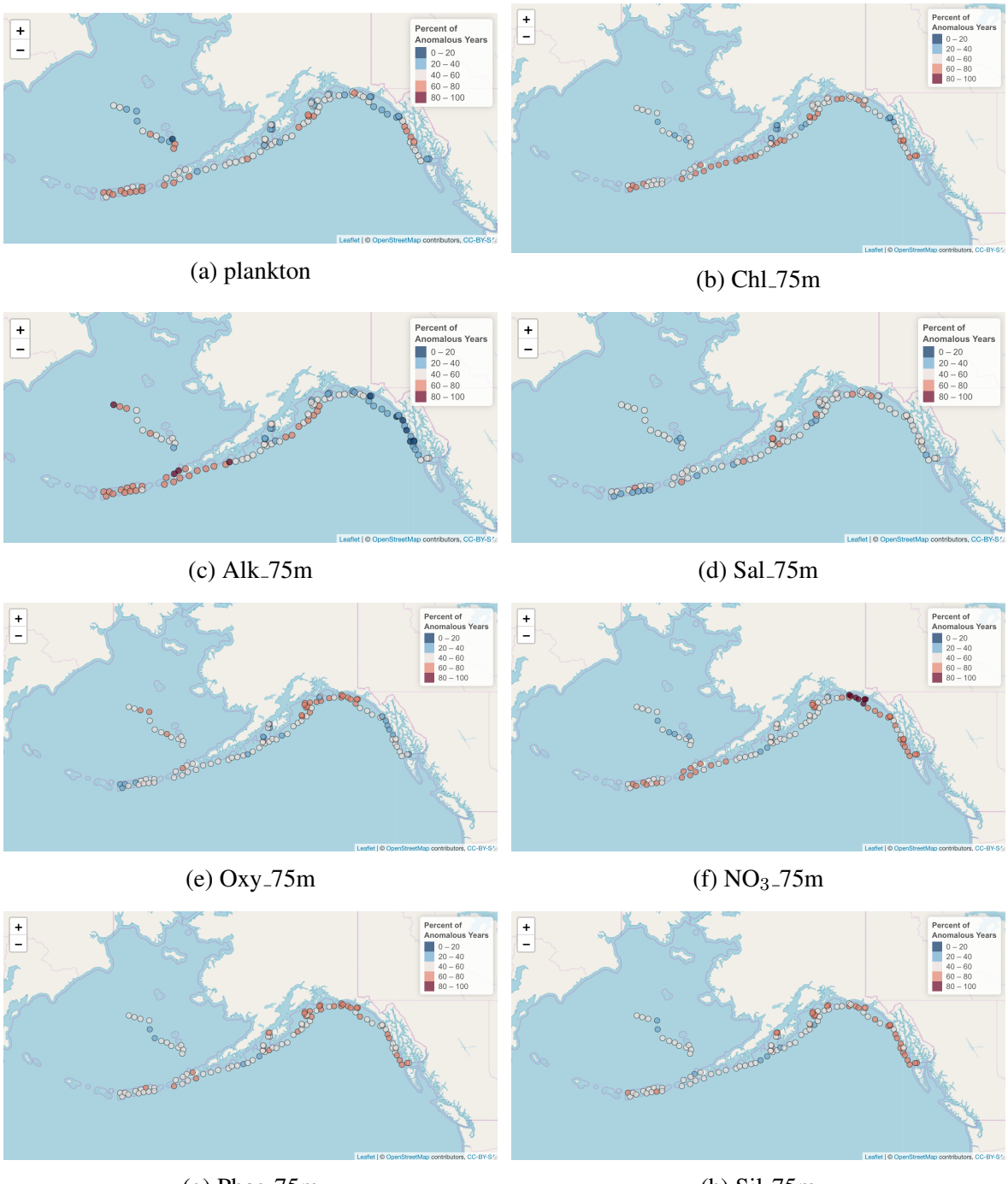
(c) WSPD (d) WVHT

(e) WTMP (f) Bottom temperature

Figure 2.7: Percent of years in which each of the physical attributes' seasonal amplitudes exhibited trends opposite their expected sign. ATMP, PRES, WTMP, and bottom temperature were expected to be positive, while WSPD and WVHT were expected to be negative. Each circle represents a station, with blue circles indicating <40% of the years for which measures were available for that location were opposite of expectations, and red circles indicating >60% of years were opposite of expectations.

parallel relationships between wind speed and shallow water chlorophyll or zooplankton were only apparent in the sablefish CPUE model (Sugimoto and Tadokoro, 1997).

Atmospheric pressure, wind speed, water temperature, zooplankton biomass, and bottom temperature were significant variables for all six species' CPUEs. Wind speeds that were modestly higher in the summer that in the winter resulted in high predicted Pacific halibut CPUE, while wave heights approximately 1.5 m higher in winter than in summer resulted in the largest increase in halibut CPUE. Indeed, most groundfish CPUE maximized when mean



(a) plankton

(b) Chl_75m

(c) Alk_75m

(d) Sal_75m

(e) Oxy_75m

(f) NO_3_75m

(g) Phos_75m

(h) Sil_75m

Figure 2.8: Percent of years in which each of the chemical attributes' seasonal amplitudes exhibited trends opposite their expected sign. Plankton and chlorophyll were expected to be positive, while alkalinity, salinity, DO, nitrate (NO_3), phosphate, and silicate were expected to be negative. The most selected depth for each chemical variable was plotted to be representative of the parameter's general pattern. Each circle represents a station, with blue circles indicating <40% of the years for which measures were available for that location were opposite of expectations, and red circles indicating >60% of years were opposite of expectations.

wave heights were around 1.5 m higher in winter than summer. A notable exception was sablefish CPUE, which was instead at a minimum for wave heights 1.5 m higher in winter than summer. Rockfish CPUE generally maximized when mean surface water temperatures in summer were around 6°C higher than winter. Anything more than that difference often resulted in an appreciable decline in predicted CPUEs. Cod, sablefish, and halibut CPUEs increased for more negative seasonal amplitudes. As the difference between summer and winter bottom temperatures increased from negative to positive, most groundfish CPUEs decreased and had maximums when summer bottom temperatures were lower than those in winter. Such occurrences can be explained by the lag in deep-water temperatures in relation to air and surface-water temperatures of anywhere from two to five months (Turner et al., 2017). Deep ocean waters can take up significantly more heat than the ocean at less than 300 m depths (Meehl et al., 2011). With global air temperatures increasing substantially due to increased atmospheric CO_2, deep water warming could therefore outpace sea surface temperature increases. As the North Pacific has recorded trends of increasing water temperatures, these results can generally be interpreted as most groundfish being intolerant of warmer waters, which agrees with available laboratory studies on these species (Fukasawa et al., 2004; Hurst et al., 2010, 2012b; Laurel et al., 2008; Levitus et al., 2000). One exception is sablefish, which may be slightly more tolerant of higher water temperatures than other North Pacific fish (Leeuwis et al., 2019). This result may also be borne out in our analyses, as sablefish CPUEs did not see any decline at higher positive differences between summer and winter mean water temperatures. In past deglaciation periods, sudden increases in subsurface water temperatures have led to widespread hypoxic events in the Gulf of Alaska (Praetorius et al., 2015). These events can be disruptive to ocean ecosystems through mass die-offs of fish and invertebrate populations (Grantham et al., 2004). Notably, DO did not play a significant role in sablefish mean weights, and only DO at 75 m was significant to sablefish CPUE. This may reflect sablefish tolerance of low DO environments (Leeuwis et al., 2019; Mandic et al., 2008; Rummer et al., 2010).

Models for groundfish weights were inconstant in the number and type of variables deemed significant for each species. As only groundfish that had reached maturity (at approximately

three to five years) were surveyed in these analyses, and given that weight of adult fish corresponds to growth rate, it is likely that conditions during juvenile development of these fishes permanently reduced or enhanced fish growth rates (Hurst et al., 2010; Laurel et al., 2008, 2016; Sogard and Olla, 2000). I attempted to take into account effects of environmental variables on recently mature adults, but it is likely that multiple lags are necessary to capture several years of recently recruited adults in the population. However, because of the limitations of additive models, it would be difficult to include multiple lagged covariates for all physical and chemical variables considered in this study. Given the inconsistent results in this paper, I would recommend future research to incorporate multiple lags of environmental predictors on adult groundfish weights in a model structure able to accommodate so many variables given the modest size of the data presently available for the North Pacific.

A consensus in the field of statistics is that p-values are poorly understood and abundantly misused (Wasserstein et al., 2019). Effect sizes for nonparametric models fit using splines have not been developed, and the traditional effect size measures do not appear to be effective for such models (Correia and Abebe, 2017). Strong caution should therefore be given to making inference about individual variables' relations to the catches or weights of specific groundfish species or putting heavy weight on only the variables selected by GAMs from this study. However, the results of my study and others like it emphasize the need to consider the potentially large effects of multiple environmental variables, even surface-level and atmospheric conditions, on population dynamics of not only pelagic fish species but also deep-ocean species (Leverette and Metaxas, 2005; Levin et al., 2001). The effects of changes in such variables to biological populations can also be used to trace temporal shifts in the environment to both anthropogenic and ecological changes in connected ecosystems (Whitney et al., 2005).

It would be more appropriate to be able consider a suite of physical and chemical variables holistically, and even consider these environmental effects on individual species as part of larger networks. A limitation of many current methods is that they assume linear or simplistic nonlinear relationships between putative effects and responses (e.g. structural equation models); such assumptions are not realistic assumptions for marine systems as evident from the results presented in this study. It is also surmised that the effects of environment on groundfish catches

in the Gulf of Alaska vary over space and time (Correia and Abebe, 2017). A GAM structure can accommodate predictors changing over space and time, but the models can quickly run into problems with excessive dimensionality for relatively few spatiotemporal predictors. A model structure related to the GAM but able to incorporate many spatiotemporal predictors and select important variables from these is the single index model (SIM) (Xia and Li, 1999; Xue and Pang, 2013). I suggest that SIM selection will provide a more realistic representation of nonparametric effects of environmental conditions on groundfish catches in the North Pacific Ocean and allow me to consider whether the subset of variables selected as important to groundfish catches change over time. However, despite the obvious weaknesses of the GAM analyses presented here, my inference focused on overall trends and motivated the suitability of more complex, but more flexible modeling structures to better capture the nature of climatic and environmental effects on groundfish populations and on ocean communities in general.

Chapter 3

Spatio-temporally explicit model averaging for forecasting of Alaskan groundfish catch

Abstract

1. Fisheries management is dominated by the need to forecast catch and abundance of commercially and ecologically important species. The influence of spatial information and environmental factors on forecasting error is not often considered. I propose a forecasting method called spatio-temporally explicit model averaging (STEMA) to combine spatial and temporal information through model averaging.

2. I examine the performance of STEMA against two popular forecasting models and a modern spatial prediction model: the autoregressive integrated moving averages with explanatory variables (ARIMAX) model, the Bayesian hierarchical model, and the varying coefficient model. I focus on applying the methods to four species of Alaskan groundfish for which catch data are available.

3. My method reduces forecasting errors significantly for most of the tested models when compared to ARIMAX, Bayesian, and varying coefficient methods. I also consider the effect of sea surface temperature (SST) on the forecasting of catch, as multiple studies reveal a potential influence of water temperature on the survival and growth of juvenile groundfish. For most of the preferred models, inclusion of SST in the model improved forecasting of catch.

4. It is advisable to consider both spatial information and relevant environmental factors in forecasting models to obtain more accurate projections of population abundance. The STEMA method is capable of accounting for spatial information in forecasting and can be applied to various types of data because of its flexible varying coefficient model structure. It is therefore a suitable forecasting method for application to many fields including ecology, epidemiology, and climatology.

Keywords and phrases: Forecast, model averaging, multi-model inference, spatio-temporal.

3.1 Introduction

Forecasting is a vital component of fisheries management and furnishes necessary input for management decisions. However, forecasting models are often based on simplistic time series trend analyses, which do not capture spatial information. Parametric time series methods such as autoregressive integrated moving averages (ARIMA) models provide accurate forecasts when the trend is consistent or many time points are supplied (Box and Jenkins, 1970). These time series methods often fail when there are insufficient measures over time or the response fluctuates with little apparent trend (Koutroumanidis et al., 2006; Stergiou and Christou, 1996). Another issue with these models is the inclusion of covariates. Often the covariates of interest must be forecast individually and the relationship between those predictors and the response are assumed to be linear (Box et al., 2013). This is an unrealistic assumption for recent climate data that exhibit anthropogenic-driven trends where the pattern of the relationship between predictors and response can change over time. Additionally, time series methods are unable to consider covariates that also vary over space.

Varying coefficient models offer a flexible modeling structure which allows for nonlinear relationships between predictors and the response; these models are also capable of handling covariates that change over space and time (Augustin et al., 2013; Hastie and Tibshirani, 1993; Phillips et al., 2014). Unlike time series models, prediction in varying coefficient models is limited to the model structure and cannot produce forecasts beyond the time range of the data. It would thus be ideal to combine the nonlinear spatial information from varying coefficient models with the forecasting capabilities of time series methods. Model averaging allows the combination of information from multiple models to inform the predictions of a response while accounting for model uncertainty. The models are often weighted according to best fit using the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) (Buckland et al., 1997) in the aggregation process, however other criteria for weighting are also possible (Hansen, 2007; Raftery et al., 2005). I propose a methodology based on a combination of information gained through the flexibility of varying coefficient models with the trend analyses of ARIMA models to obtain predictions for catch rates at specific locations, thereby creating

predicted species distributions. The method takes advantage of varying coefficient models and model averaging while refining these for use in spatio-temporal forecasting. My technique is the result of complex leave-one-out construction procedures used to create forecast stability. This procedure provides prediction errors that are used as weights for model averaging.

Water temperature has effects on the growth, survival, and behavior of juvenile fish. Survival of juveniles to reproductive age is a key indicator of population maintenance and growth. Fisheries management and restoration strategies are keen to monitor recruitment and abundance for target species. However, information on reproductive success and recruitment to model population abundance is often lacking for deepwater marine species. Catch rates are related to population size and are commonly used in fisheries management as an index of abundance (Battaile and Quinn, 2004; Council, 2000; Ricker, 1975). Many deepwater species move inshore to reproduce, and thus offspring are affected by the temperature of surface waters in the first few years of life when they are sensitive to environmental extremes. Changes in SST affect plankton availability, distribution, and composition, which are an important nutrition resource for deepwater species and act as a carbon sink (Brierley and Kingsford, 2009). Additionally, SST acts as a proxy for many other oceanic processes, affecting currents, ocean mixing, and sea ice retreat, all of which have effects on both fish biomass (Bouchard and Fortier, 2008; Hunt et al., 2002) and catch (Cheung et al., 2009; Kim et al., 2012; Monllor-Hurtado et al., 2017). Further, it is evident that management strategies must now consider temperature trends in order for managers to provide accurate long-term advice (Biswas et al., 2005; Ianelli et al., 2011; Vaidyanathan, 2017). Still, current management implementation rarely includes ecosystem processes that have been shown to affect fish stock productivity (Skern-Mauritzen et al., 2015). It is therefore meaningful to consider SST in predicting catch, particularly in studies where recruitment information is not available or difficult to assess. While adding additional variables will improve model fit, it will not necessarily improve prediction. Thus, I assess whether prediction of catch is improved through the inclusion of SST.

3.2 Background

Prediction of fish abundance and catch is crucial in creating management strategies for commercially important species, particularly in oceans with high system variability. The northern Pacific Ocean has undergone several recent regime shifts that affect marine groups differently (Napp and Hunt, 2001). Changes in abundance and population dynamics of various marine fishes, including salmon, cod, halibut, and sardines, have manifested in response to climatic regime shifts in the past century (Benson and Trites, 2002; Möllmann and Diekmann, 2012; Noakes and Beamish, 2009). Abrupt changes in climatic cycles via persistent, area-specific shifts in trends of water temperature, ocean currents, and primary production create profound changes in the marine ecosystem, though the precise mechanisms through which these changes occur are still not well understood (Anderson and Piatt, 1999; Francis et al., 1998). Sea surface temperature (SST) is a simple measure to obtain. It acts as an easily identifiable representative for more complex relationships between oceanic and atmospheric conditions that precede or accompany marine regime shifts (deYoung et al., 2008; Möllmann et al., 2015). Consequently, SST is the most commonly used environmental variable considered when modeling fish catch and abundance in the wild. However, not many studies consider the effect of sea surface temperature (SST) on the forecasting of fish catch, especially in management settings. Water temperature has also been shown to affect the feeding motivation, metabolism, reproduction, and behavior of many fish species (Donelson et al., 2010; Pörtner et al., 2001), which in turn influences recruitment and abundance. Increased water temperatures due to climate change are therefore likely to affect the amount and composition of aquatic species in northern latitudes (Pörtner and Knust, 2007; Sharma et al., 2007). Along with increased water temperatures, climate variability is expected to increase as a result of climate change (Easterling et al., 2000; Timmermann et al., 1999). Anomalous oceanic conditions brought about by persistent changes in atmospheric patterns, such as the warm SST anomaly known as “the Blob” in the Northern Pacific Ocean (Tseng et al., 2017), have effects on regional weather and impacts on coastal and deepwater fisheries operations as well as the composition of ecosystems (Bond et al., 2015). Extreme changes in the marine environment that often accompany ocean anomalies are more

detrimental to juvenile fish and can affect their recruitment into the adult population (Baumann et al., 2006; Beaugrand et al., 2003; Stige et al., 2006).

Three commercially important groundfish species, Pacific cod (*Gadus macrocephalus*), Pacific halibut (*Hippoglossus stenolepis*), and sablefish (*Anoplopoma fimbria*), and the most abundant groundfish species, giant grenadier (*Albatrossia pectoralis*), are located within the north Pacific ecosystem. The commercial fisheries of Pacific cod, Pacific halibut, and sablefish are predominantly or solely longline, and catch of giant grenadier is predominantly through bycatch on sablefish longlines (Goen and Erikson, 2017; NPFMC, 2017; Rodgveller et al., 2008). Pacific halibut, Pacific cod, and sablefish also have management guidelines in effect that would likely benefit from new and more accurate prediction techniques. Winter ocean conditions in the northeast Pacific Ocean have been linked to recruitment in groundfish stocks (Hollowed and Wooster, 1992; Schirripa and Colbert, 2006). Studies on juveniles of these four species show that increased water temperatures affect behavioral responses, growth, and survival (Laurel et al., 2016; Sogard and Olla, 2001; Stoner et al., 2006; Stoner and Sturm, 2004). No laboratory studies have been conducted on the temperature tolerances of giant grenadier. Sablefish and giant grenadier are known to compete for baited hooks in longline surveys (Rodgveller et al., 2008). These results indicate that giant grenadier may inhabit similar temperature zones as sablefish. This highlights the need to understand the relationship of temperature to an apex deepwater predator likely to be the most abundant fish in the northern Pacific (Rodgveller and Hulson, 2014).

Climate change is characterized in many areas of the globe as a consistent warming trend which favors acclimation in fishes (Crozier and Hutchings, 2014). Variability in global climate systems is also increasing the occurrence of extreme climate events and changing marine ecosystems dramatically and suddenly (Hoegh-Guldberg and Bruno, 2010; Walther et al., 2002). If oceanic conditions continue to experience increased variability and instability, persistent changes to the physiology of fishes as a result of acclimatisation are likely to translate into reduced phenotypic plasticity (Reed et al., 2011; Seebacher et al., 2014). Pacific cod displayed "cold-adapted" responses in hatching, growth rates, and mortality when sampled from the coldest cohort in three decades (Hurst et al., 2012b). This illustrates that groundfish from a

cohort experiencing more extreme temperature changes, either anomalously cold or warm, may be at a disadvantage when experiencing the opposing extreme conditions to which they experienced when hatching. The effect is likely to be pronounced if an intensely cold year during the hatching of a cohort is followed by an extremely warm year (or vice versa) when those fish are still in their vulnerable juvenile state. It is therefore important to gain a greater understanding of the effects of temperature on commercially and ecologically important species such as those discussed here.

3.3 Data

The data for this study were collated from two data sets provided by the National Oceanic and Atmospheric Administration (NOAA). Of primary use were the annual longline survey data of the Marine Ecology and Stock Assessment (MESA) Program conducted by the Auke Bay Laboratories in Alaska (Alaska Fisheries Science Center, 2019a). The MESA Program has performed longline surveys independently since 1979, dropping baited lines at specific locations ("stations") off the coast of Alaska to collect information on groundfish species. Seven major groundfish species are surveyed in the MESA Program by the Alaska Fisheries Science Center (AFSC), of which four (sablefish, Pacific cod, Pacific halibut, and giant grenadier) will be considered in these analyses. The AFSC records number of fish per species collected at each location and calculates a catch per unit effort (CPUE) within each management area from the total number of fish caught divided by the total number of skates, 100-meter longlines with 45 evenly spaced hooks per line, deployed each day (Sigler and Lunsford, 2009). The CPUE is therefore a standardized measure of catch at each location. Longline surveys recording CPUE have been shown to be an accurate fishery-independent index of abundance for sablefish (Sigler, 2000) and Pacific halibut (Monnahan and Stewart, 2018) when properly accounting for hook spacing and spatial stratification.

Daily global SST readings, available for dates starting in 1981 through 2012, were obtained from the National Centers for Environmental Information (NOAA, 2015). The data were interpolated and optimized from satellites, buoys, and ships on $1/4^{\circ}$ latitude-longitude

grids using a method devised by Richard W. Reynolds at the National Centers for Environmental Prediction. A coefficient of variation for SST was derived for each $1/4^{\circ}$ latitude-longitude grid for the winter season (November through April), because the groundfish studied in the MESA surveys undergo reproductive activity in the winter months in the waters surrounding Alaska. In addition, evidence has suggested that winter conditions have the greatest influence on groundfish populations (Hollowed and Wooster, 1992). The winter coefficient of variation for SST was calculated as

$$c_v = \frac{\sigma}{\mu}$$

at each latitude-longitude pairing, with σ being the winter seasonal standard deviation and μ the winter season's mean of SST. The coefficient of variation is an improved measure of seasonal SST over the mean, because it standardizes scale and allows us to consider the changes in variation of SST with the changes in mean over time.

Fluctuations in CPUE are likely to be linked to changes across cohorts which are often determined by survival in the first year of life. Water temperature has been found to affect the MESA groundfish covered by my analyses, and juvenile fish are more susceptible to environmental changes than their adult counterparts. Therefore, CPUE for a given year is likely to be linked to the winter SST encountered at the juvenile state by fish entering the adult population. Since the MESA survey targets waters where adults reside during the summer, and the four species covered in my analyses reach maturity at five to eight years, SST was lagged for years one through five to allow us to capture the effect of SST on the juvenile stages and recruitment. All five lagged SST measures were included for modeling.

I focused on determining the spatio-temporal catch predictions for four of the species in the MESA study area known as the Gulf of Alaska which ranges from the Dixon Entrance west to Chuginadak Island. The fisheries data were matched with winter SST data from 1982 to 2012. With lagged winter SST included, this created a dataset of CPUE for four groundfish species spanning 23 years from 1990 to 2012. There are 1679 observations each for sablefish, Pacific cod, giant grenadier, and Pacific halibut.

3.4 Methods

My proposed forecasting method consisted of two parts, a model averaging technique made up of a spatially variant coefficient model with prediction obtained via an ARIMA model and a temporally-varying coefficient model with prediction incorporated via an ARIMA model. The proposed method was applied to the Alaska groundfish data. I then compared three main methods of forecasting to my proposed method: a simple ARIMA model with covariates (ARIMAX) with lagged winter SST from one to five years used as predictors, a naïve spatially varying coefficient model in which the fitted values for the current year were considered the predicted values for the next year, and a hierarchical Bayesian forecasting procedure. The ARIMAX and Bayesian implementations are linear models, while the naïve spatially varying coefficient model and the proposed forecasting method make use of nonlinear models.

The distribution of CPUE values for Pacific cod and Pacific halibut were right-skewed and were accommodated in the model fitting, while sablefish and giant grenadier CPUE values were Gaussian distributed.

All of the methods were subjected to a leave-one-out procedure. This allowed us to determine if the success of the proposed technique was mainly due to its predictions being verified and adjusted using the leave-one-out procedure. Since the naïve spatially varying coefficient model is not a typical forecasting procedure, only the leave-one-out setting was considered for this model. For the naïve model, a station was removed from the dataset and the spatial model fitting was performed on the $\{1, 2, \ldots, b-1, b+1, \ldots, n\}$ stations. The spatial forecast obtained after each leave-one-out operation is denoted $\widetilde{Y}_J^{sp(-b)}$, where b is the removed station. The ARIMAX and hierarchical Bayesian models go through a similar leave-one-out procedure, where a year c was removed from the dataset and the ARIMAX (or Bayesian) modeling procedure was performed on the $\{1, 2, \ldots, c-1, c+1, \ldots, J-1\}$ years. The ARIMAX forecast obtained after each leave-one-out-procedure is similarly denoted $\widetilde{Y}_J^{t(-c)}$ for the removed year c. A mean and standard deviation of the leave-one-out forecasts for each station's ARIMAX, Bayesian, and naïve spatial models were calculated. The means of the leave-one-out forecasts for each station were used as the final forecast CPUE values for the leave-one-out versions of

the ARIMAX, Bayesian, and naïve spatial models. Weights for the model averaging of the two component predictions for the proposed method were determined by the standard error of the leave-one-out predictions for each. A minimum of ten observations per location was considered to provide a sufficient number of points to obtain a trend over time; that is, a minimum of ten yearly observations per station were included in the training datasets that were used to predict the subsequent year. For each of the following models, let J be the year for which prediction is sought, where $J = 2000, 2001, \ldots, 2012$.

In Section 3.4.1, I introduce and describe the basic ARIMAX model, the spatially varying coefficient model, and the Bayesian forecasting method to which I compared my proposed forecasting technique. I then show how the ARIMA model and spatially varying coefficient model were combined using model averaging to produce my proposed forecasting method in Section 3.4.2.

3.4.1 Some existing forecasting procedures

3.4.1.1 ARIMAX model

An ARIMAX of order p, d, q in the form

$$\left(1 - \sum_{m=1}^{p} \phi_m L^m\right)(1-L)^d Y_i = \delta + \left(1 + \sum_{m=1}^{q} \theta_m L^m\right)\epsilon_i + \left(1 - \sum_{m=1}^{p} \phi_m L^m\right)(1-L)^d \mathbf{X}_i^T \boldsymbol{\beta} \,,$$

was fit for each location. The lag operator is denoted by L, ϕ_m are the autoregressive parameters, θ_m are the moving average parameters, $\boldsymbol{\beta}$ is the predictor coefficient matrix, and ϵ_i are the error terms (Box et al., 2013). The predictor vector $\mathbf{X}_i = (SST_{i-1}, SST_{i-2}, SST_{i-3}, SST_{i-4}, SST_{i-5})^T$ includes lagged winter SST values for one to five years. The order (p, d, q) with drift $\delta/(1 - \Sigma\phi_m)$ for the ARIMAX model is automatically determined using minimization of AIC and MLE to determine the best ARIMAX model using the function `auto.arima` in the **forecast** package (Hyndman, 2017; Hyndman and Khandakar, 2008) in R (R Core Team, 2017). The right-skewed distributions of Pacific cod and Pacific halibut CPUEs does not affect the fitting of the ARIMAX models, as each location is fit individually. At most, only one outlier was identified per station for Pacific cod and Pacific

halibut when modeling the entire range of training data for one station representative of each of the four management areas (Figs S1 and S2 in Appendix A). The fitted ARIMAX model was then used to predict year J using known winter SST values from years $J-5$ to $J-1$,

$$\left(1-\sum_{m=1}^{p}\widehat{\phi}_m L^m\right)(1-L)^d\widehat{Y}_J = \delta + \left(1+\sum_{m=1}^{q}\widehat{\theta}_m L^m\right)\epsilon_J + \left(1-\sum_{m=1}^{p}\widehat{\phi}_m L^m\right)(1-L)^d\mathbf{X}_J^T\widehat{\beta}\ ,$$

with the predicted value for year J denoted $\widetilde{Y}_J^A$, employing the `forecast` function from the **forecast** package.

3.4.1.2 Naïve spatially varying coefficient model

A spatially varying coefficient model for CPUE of a given species including lagged winter SST for five years to one year that varies over space is fit for year $J-1$,

$$Y_{J-1} = \mathbf{X}_{J-1}^T\mathbf{G}_{J-1}(U) + \varepsilon_{J-1}\ ,$$

where $\mathbf{X}_{J-1} = (1, SST_{(J-1)-1}, SST_{(J-1)-2}, SST_{(J-1)-3}, SST_{(J-1)-4}, SST_{(J-1)-5})^T$ and $\mathbf{G}_{(J-1)}(U) = (g_{0,J-1}(U), g_{1,J-1}(U), g_{2,J-1}(U), g_{3,J-1}(U), g_{4,J-1}(U), g_{5,J-1}(U))^T$ is the functional coefficient vector of winter SSTs with U being the longitude-latitude pairs representing sampled locations along the Gulf of Alaska. The fitted CPUE values for year $J-1$ given as

$$\widehat{Y}_{J-1} = \mathbf{X}_{J-1}^T\widehat{\mathbf{G}}_{J-1}(U) = \widehat{g}_{0(J-1)} + \sum_{k=1}^{5}\widehat{g}_{k(J-1)}(U)SST_{(J-1)-k}$$

were considered to be the predicted values for year J. The predicted values from this naïve spatial model are denoted $\widetilde{Y}_J^N$. The spatially varying coefficient models used rank-based estimation as described in Correia (2018). Rank estimation techniques are more suitable than least squares estimation for reducing the influence of outliers and contamination common in fisheries and ecological data on prediction. The rank-based estimation for varying coefficient models was coded as a modification to the `gam` function in the **mgcv** package (Wood, 2006) in R (R Core Team, 2017) and is included as supplemental material in Correia (2018). To accommodate the right-skewed distributions of Pacific cod and Pacific halibut CPUE values, I used

the Gaussian distribution and weights given by the bent score function (Kloke and McKean, 2014) in the varying coefficient model fitting process.

3.4.1.3 Hierarchical Bayesian forecasting

To implement Bayesian forecasting methods, I chose a hierarchical independent Gaussian process model. Let $\mathbf{Z}_i$ denote the observed data, and $\mathbf{O}_i$ be the corresponding true values for station $s_r, r = 1, \ldots, n$ at time $i = 1, \cdots, J-1$. Also let $\mathbf{Z}_i = (Z(s_1, i), \ldots, Z(s_n, i))^T$, $\mathbf{O}_i = (O(s_1, i), \ldots, O(s_n, i))^T$, and $N = n \times (J-1)$ be the total number of observations modeled. The Gaussian process model is specified as

$$\mathbf{Z}_i = \mathbf{O}_i + \boldsymbol{\epsilon}_i \qquad \text{and} \qquad \mathbf{O}_i = \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\eta}_i \ ,$$

where $\boldsymbol{\beta}$ is the regression coefficient vector, and $\boldsymbol{\epsilon}_i = (\epsilon_{s_1,i}, \ldots, \epsilon_{s_n,i}) \sim N(\mathbf{0}, \sigma_\epsilon^2 \mathbf{I}_n)$ is the pure error term, σ_ϵ^2 is the unknown variance and $\mathbf{I}_n$ is the identity matrix of order n. The spatio-temporal random effects are denoted $\boldsymbol{\eta}_i = (\eta(s_1, i), \ldots, \eta(s_n, i))^T \sim N(\mathbf{0}, \Sigma_\eta)$, where $\Sigma_\eta = \sigma_\eta^2 S_\eta$ is composed of the spatial variance, σ_η^2, and the spatial correlation matrix, S_η. The spatial correlation matrix is derived from the Matérn correlation function

$$\kappa(s_i, s_j : \phi, \nu) = \frac{1}{2^{\nu-1}\gamma(\nu)}(2\sqrt{\nu}||s_i - s_j||\phi)^\nu K_\nu(2\sqrt{\nu}||s_i - s_j||\phi), \quad \phi > 0, \quad \nu > 0 \ ,$$

where ϕ controls the correlation decay rate as distance between two spatial points $||s_i - s_j||$ increases, K_ν is the modified Bessel function of order ν, and ν controls the smoothness of the random field. Let all of the parameters of the model be denoted $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_\epsilon^2, \sigma_\eta^2, \phi, \nu)$, and let $\pi(\boldsymbol{\theta})$ denote the prior distributions. The prior distribution for the inverse variance model parameters is given as

$$\left(\frac{1}{\sigma_\epsilon^2}, \frac{1}{\sigma_\nu^2}\right) \sim \Gamma\left(\frac{a}{b}, \frac{a}{b^2}\right) \ ,$$

where $a = 2$ and $b = 1$, while the prior distributions for the mean parameter $\boldsymbol{\beta}$ is $N(\mu_{\boldsymbol{\beta}}, \delta_{\boldsymbol{\beta}}^2)$, where $\mu_{\boldsymbol{\beta}} = 0$ and $\delta_{\boldsymbol{\beta}}^2 = 10^{10}$. The logarithm of the joint posterior distribution for this Gaussian

process model is

$$\log \pi(\boldsymbol{\theta}, \mathbf{O}, \mathbf{z}^*|\mathbf{z}) \propto -\frac{N}{2}\log \sigma_\epsilon^2 - \frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^{J-1}(\mathbf{Z}_i - \mathbf{O}_i)^T(\mathbf{Z}_i - \mathbf{O}_i)$$
$$-\frac{1}{2}\log|\sigma_\eta^2 S_\eta| - \frac{1}{2\sigma_\eta^2}\sum_{i=1}^{J-1}(\mathbf{O}_i - \mathbf{X}_i\mathbf{B})^T S_\eta^{-1}(\mathbf{O}_i - \mathbf{X}_i\mathbf{B}) + \log \pi(\boldsymbol{\theta}) \ .$$

The Bayesian forecasting method was implemented via the R package **spTimer** (Bakar et al., 2015).

3.4.2 Spatio-temporally explicit model averaging

The spatio-temporally explicit model averaging (STEMA) technique was derived from the combination of a spatially varying coefficient model (Section 3.4.1.2) and a yearly varying coefficient model where latitude-longitude pairs in the spatially varying coefficient model were replaced by year. Each model's fitted values were then used to fit an ARIMA and forecast the year for which prediction was sought. The separate model forecasts were averaged using weights based on standard deviations of the leave-one out procedure, giving more weight to the model with lower standard deviation to produce a final forecast for each station. The components of the STEMA forecasting procedure are described in the following three subsections, with example code of the procedure given in the Supplementary Materials.

3.4.2.1 Spatial model with ARIMA

A spatially varying coefficient model as described above was fit over space for each year $i = 1990, 1991, \ldots, J-1$. The fitted CPUE values for year i given as

$$\widehat{Y}_i^{sp} = \mathbf{X}_i^T \widehat{\mathbf{G}}_i(U) = \widehat{g}_{0i} + \sum_{k=1}^{5} \widehat{g}_{ki}(U) SST_{i-k}$$

were then used to fit an ARIMA model for each station,

$$\left(1 - \sum_{m=1}^{p} \phi_m L^m\right)(1-L)^d \widehat{Y}_i^{sp} = \delta + \left(1 + \sum_{m=1}^{q} \theta_m L^m\right)\epsilon_i \ ,$$

yielding the fitted ARIMA values $\widetilde{Y}_J^{sp}$, where the order (p, d, q) with drift $\delta/(1 - \Sigma\phi_m)$ is automatically determined using the `auto.arima` function as described in section 3.4.1.1. The two-step process allows for inclusion of multiple lagged winter SST variables smoothed over space in the varying coefficient model setting while providing a method for future prediction which is not available in these models.

3.4.2.2 Temporal model with ARIMA

A time varying coefficient model of the form

$$Y_{J-1} = g_{0J}(t) + g_{1J}(t)SST_{t-5} + \varepsilon_{J-1}$$

was fit for each location, where $i = 1990, 1991, \ldots, J-1$; $t = \{1990, 1991, \ldots, J-1\}$; and SST_{t-5} is the winter SST for that location lagged by five years. The fitted values from this model,

$$\widehat{Y}_i^t = \widehat{g}_{0J}(i) + \widehat{g}_{1J}(i)SST_{i-5} \ ,$$

were then used to fit an ARIMA model

$$\left(1 - \sum_{m=1}^{p} \phi_m L^m\right)(1-L)^d \widehat{Y}_i^t = \delta + \left(1 + \sum_{m=1}^{q} \theta_m L^m\right)\epsilon_i \ ,$$

yielding the fitted ARIMA values $\widetilde{Y}_J^t$, where the order (p, d, q) with drift $\delta/(1 - \Sigma\phi_m)$ is automatically determined as described for the spatial model. The coefficient model smooths the CPUEs for each location, thereby allowing the time series model to determine a more accurate trend despite highly variable CPUE values.

3.4.2.3 Model averaging

The STEMA method underwent the same leave-one-out procedures as described for the ARIMA and naïve spatial models, where the temporal model with ARIMA used the temporal leave-one-out procedure and the naïve spatial model utilized the spatial leave-one-out steps. For the STEMA technique, the means of the spatial and temporal leave-one-out procedures ($\bar{\widetilde{Y}}_J^{sp}$ and

$\bar{\widetilde{Y}}_J^t$, respectively) were weighted for each location using a ratio of the spatial (σ_{sp_n}) and temporal (σ_{t_n}) standard deviations from the leave-one-out predictions,

$$\omega_{sp} = \frac{\sigma_t}{\sigma_{sp} + \sigma_t} \qquad \text{and} \qquad \omega_t = \frac{\sigma_{sp}}{\sigma_{sp} + \sigma_t} \ ,$$

where $\omega_{sp} + \omega_t = 1$. The final spatio-temporally explicit model averaged prediction was obtained for each location by

$$\widetilde{Y}_J = \omega_{sp}\bar{\widetilde{Y}}_J^{sp} + \omega_t\bar{\widetilde{Y}}_J^t \ .$$

The standard error of the spatio-temporally explicit model averaged predictions is given as

$$SE\left(\widetilde{Y}_J\right) = \sqrt{\omega_{sp}^2 SE\left(\bar{\widetilde{Y}}_J^{sp}\right) + \omega_t^2 SE\left(\bar{\widetilde{Y}}_J^t\right)}$$

3.5 Assessment of forecast performance via cross-validation

3.5.1 Model comparison

A time series cross-validation based on one-step forecasts was performed on the ARIMAX model (A), the hierarchical Bayesian model (B), the naïve spatially varying coefficient model (N[1]), the spatio-temporally explicit model averaging technique (STEMA), and the leave-out-out versions of the ARIMAX (A[1]) and Bayesian (B[1]) models. I consider h to be the minimum number of years needed to create a reliable forecast and proceed as follows: for $f = 1, 2, \ldots, T - h$ where T is the total number of years available and $j = h + f$, train on $F_h, \ldots, F_{j-1}$, and forecast and validate on F_j. For each estimation technique, a forecast $\widehat{Y}_{aj}$, $a =$ A, A[1], B, B[1], N[1], STEMA was computed from the training sets, and the error on the validation set was recorded as

$$e_{aj} = Y_j - \widehat{Y}_{aj}.$$

To compare estimation techniques, I used a Friedman rank sum test on the absolute errors to determine if there were significant differences among methods (Friedman, 1937). If

the Friedman test indicated significant differences, I then performed pairwise multiple comparisons on the differences between the absolute errors for each pair of methods (Bretz et al., 2016; Tukey, 1949). In order to control for the effect of location, a generalized linear mixed model was fit with the stations set as random effects. P-values calculated for the pairwise tests were adjusted using the Benjamini-Hochberg procedure to control the false discovery rate (Benjamini and Hochberg, 1995). If the difference was significantly less than zero, the first of the two compared methods was the method that produced smaller errors; if the difference was significantly greater than zero, the second method produced smaller errors. The mixed model was fit using the `glmer` function in the **lme4** package (Bates et al., 2015), while the pairwise multiple comparisons were performed in the **multcomp** package (Hothorn et al., 2008) using the `glht` function in R (R Core Team, 2017).

3.5.2 Forecast performance in the presence of an environmental covariate

In order to determine if adding SST to the models improved forecasting, a null model for each of the four techniques was fit, subjected to the same leave-one-out procedure as described previously, and a forecast obtained for each. The null A and A^1 models are of the form

$$\left(1 - \sum_{m=1}^{p} \phi_m L^m\right)(1-L)^d \widehat{Y}_i = \delta + \left(1 + \sum_{m=1}^{q} \theta_m L^m\right)\epsilon_i,$$

where the order p, d, q is determined as before. The fitted CPUE values for year $J-1$ from the null naïve spatially varying coefficient model given as

$$\widehat{Y}_{J-1} = \widehat{G}_{J-1}(U) = \widehat{g}_{0(J-1)}$$

are considered to be the predicted values for year J. The spatio-temporally model averaged forecasts were derived from the two null varying-coefficient models

$$Y_i^{sp_0} = g_{0i} + \varepsilon_i \qquad \text{and} \qquad Y_{J-1}^{t_0} = g_{0J}(t) + \varepsilon_{J-1},$$

of which the fitted values were each used to fit ARIMA models following the steps in section 3.4.2. The forecasts obtained from those fitted ARIMA models were averaged as described in section 3.4.2.3 to form the final prediction. The null model errors were compared to the errors of their model counterparts which include winter SST for each method using one-sided Wilcoxon signed-rank tests, where the errors are matched by station and year. If SST is important to forecasting, the inclusion of SST in the model will significantly reduce forecasting error.

In order to obtain the magnitude of the effect of winter SST on prediction using the preferred methods, I calculate a rank-correlation r statistic using the asymptotic normal distribution of the Wilcoxon signed-rank statistic W on the absolute error differences D_i between the null model and the SST model. W is calculated as

$$W = \sum_{i=1}^{N} Rank(|D_i|) \times I(D_i > 0) \ ,$$

where N is the total number of calculated errors. Under the hypothesis that winter SST has no impact on prediction, W is asymptotically normal as

$$Z = \frac{W - \frac{N(N+1)}{4}}{\sqrt{\frac{N(N+1)(2N+1)}{24}}}$$

(Hollander and Wolfe, 1999). The rank-correlation is given by

$$r = \frac{Z}{\sqrt{N}}$$

with estimated variance $(1 - r^2)/(N - 2)$ (Rosenthal et al., 1994). Small, medium, and large effect sizes are .10, .30, and .50, respectively (Cohen, 1992).

3.5.3 Control of Bayesian parameter ϕ in cross-validation

One issue that arises from using the time series cross-validation on the Bayesian forecasting method is the fluctuation in spatial point acceptance rate as the available years of data change. While the spatial decay parameter ϕ can be chosen by the user to obtain optimal acceptance rate of spatial points for the calculation of the spatial correlation matrix, the appropriate value

of ϕ changes given varying data structure. There is also insufficient guidance on how forecast values are affected by misspecification of ϕ. According to Bakar et al. (2015), the choice of ϕ is obtained with acceptance rates between 20% and 40%, which is justified by Gelman et al. (2004).

I chose to apply a search similar to Paci et al. (2013) for the optimal ϕ value by fitting the current data set with values of ϕ starting at 10 and decreasing by an order of magnitude of 1 for each subsequent fitting. Once the model achieved an acceptance rate closest to 32%, that model was then used to obtain forecasting estimates for year J. This ensured that ϕ was selected for each model fitting step to always obtain an optimal acceptance rate despite the changing size of training data. Variable training data that occurs when using the temporal cross-validation affects the spatial information available, making a fixed value of ϕ unsuitable for accurate forecasting using the Bayesian method with temporal cross-validation.

3.6 Results

Friedman tests for all four species revealed significant differences across model techniques (sablefish: $\chi^2 = 74.022$, $p <= 0.0001$; Pacific cod: $\chi^2 = 365.501$, $p <= 0.0001$; Pacific halibut: $\chi^2 = 152.471$, $p <= 0.0001$; giant grenadier: $\chi^2 = 460.030$, $p <= 0.0001$). The STEMA method had lowest mean absolute errors for Gaussian distributed species (sablefish and giant grenadier) when ignoring station and year effects (Table 3.1). Pairwise multiple comparisons were therefore performed on method pairings for all species to determine the best methods of forecasting for each species. The results of these pairwise comparison tests of the different methods are given in Tables 3.2 to 3.5. Based on the differences in mean absolute errors that are significant, the following methods had the lowest significant absolute errors: the STEMA method for sablefish; the naïve and STEMA methods for Pacific cod; the naïve method for Pacific halibut; and the STEMA method for giant grenadier. STEMA did not significantly improve forecasting over the N^1 model in the case of Pacific cod, and the N^1 method outperformed STEMA in lowering forecasting errors for Pacific halibut.

The results of the one-sided Wilcoxon signed-rank tests for comparing models including winter SST to those without for all forecasting methods are summarized in Table 3.6. For all

Table 3.1: Mean absolute error of each method for four species, ignoring station effects. Lowest mean absolute errors for each species are in bold.

	A	A[1]	N[1]	B	B[1]	STEMA
Sablefish	1.771	1.688	1.612	1.558	1.562	**1.369**
Pacific cod	0.280	0.267	**0.136**	0.180	0.175	0.145
Pacific halibut	0.293	0.285	**0.199**	0.205	0.204	0.205
Giant grenadier	1.599	1.480	1.614	1.844	1.839	**1.086**

Table 3.2: Pairwise multiple comparisons of absolute errors of forecasting methods with winter SST included in the models for **sablefish**. P-values are adjusted using false discovery rate method. A p-value < 0.05 indicates the difference in absolute errors of the comparison are significant (in bold). Differences significantly less than zero indicate the first of the two compared methods was the method that produced smaller errors; estimates significantly greater than zero indicate the second method produced smaller errors.

Linear Hypotheses	Estimate	Std. Error	z value	Pr(>|z|)
A[1] - A	-0.050	0.038	-1.323	0.253
B - A	-0.130	0.038	-3.435	**0.002**
B - A[1]	-0.080	0.038	-2.115	0.057
B[1] - A	-0.128	0.038	-3.369	**0.002**
B[1] - A[1]	-0.078	0.038	-2.049	0.061
B[1] - B	0.002	0.038	0.066	0.947
N[1] - A	-0.090	0.038	-2.378	**0.033**
N[1] - A[1]	-0.040	0.038	-1.057	0.335
N[1] - B	0.040	0.038	1.058	0.335
N[1] - B[1]	0.038	0.038	0.993	0.344
STEMA - A	-0.254	0.038	-6.707	**0.000**
STEMA - A[1]	-0.204	0.038	-5.386	**0.000**
STEMA - B	-0.124	0.038	-3.264	**0.002**
STEMA - B[1]	-0.126	0.038	-3.331	**0.002**
STEMA - N[1]	-0.164	0.038	-4.330	**0.000**

[1] Leave-one-out procedure used

Table 3.3: Pairwise multiple comparisons of absolute errors of forecasting methods with winter SST included in the models for **Pacific cod**. P-values are adjusted using false discovery rate method. A p-value < 0.05 indicates the difference in absolute errors of the comparison are significant (in bold). Differences significantly less than zero indicate the first of the two compared methods was the method that produced smaller errors; differences significantly greater than zero indicate the second method produced smaller errors.

Linear Hypotheses	Estimate	Std. Error	z value	Pr(>|z|)
A[1] - A	-0.033	0.040	-0.835	0.454
B - A	-0.275	0.040	-6.828	$\mathbf{0.000}$
B - A[1]	-0.241	0.040	-6.004	$\mathbf{0.000}$
B[1] - A	-0.307	0.040	-7.626	$\mathbf{0.000}$
B[1] - A[1]	-0.273	0.040	-6.802	$\mathbf{0.000}$
B[1] - B	-0.032	0.040	-0.800	0.454
N[1] - A	-0.525	0.040	-12.996	$\mathbf{0.000}$
N[1] - A[1]	-0.492	0.040	-12.185	$\mathbf{0.000}$
N[1] - B	-0.251	0.040	-6.226	$\mathbf{0.000}$
N[1] - B[1]	-0.219	0.040	-5.434	$\mathbf{0.000}$
STEMA - A	-0.550	0.040	-13.699	$\mathbf{0.000}$
STEMA - A[1]	-0.516	0.040	-12.879	$\mathbf{0.000}$
STEMA - B	-0.275	0.040	-6.855	$\mathbf{0.000}$
STEMA - B[1]	-0.243	0.040	-6.061	$\mathbf{0.000}$
STEMA - N[1]	-0.024	0.040	-0.609	0.542

[1] Leave-one-out procedure used

Table 3.4: Pairwise multiple comparisons for absolute errors of forecasting methods with winter SST included in the models for **Pacific halibut**. P-values are adjusted using false discovery rate method. A p-value < 0.05 indicates the difference in absolute errors of the comparison are significant (in bold). Differences significantly less than zero indicate the first of the two compared methods was the method that produced smaller errors; differences significantly greater than zero indicate the second method produced smaller errors.

Linear Hypotheses	Estimate	Std. Error	z value	Pr(>|z|)
A[1] - A	-0.034	0.006	-5.413	**0.000**
B - A	-0.384	0.006	-61.003	**0.000**
B - A[1]	-0.350	0.009	-39.580	**0.000**
B[1] - A	-0.387	0.006	-61.442	**0.000**
B[1] - A[1]	-0.353	0.009	-39.895	**0.000**
B[1] - B	-0.003	0.009	-0.334	0.739
N[1] - A	-0.423	0.006	-67.557	**0.000**
N[1] - A[1]	-0.389	0.009	-44.229	**0.000**
N[1] - B	-0.040	0.009	-4.494	**0.000**
N[1] - B[1]	-0.037	0.009	-4.160	**0.000**
STEMA - A	-0.391	0.006	-62.332	**0.000**
STEMA - A[1]	-0.357	0.009	-40.516	**0.000**
STEMA - B	-0.008	0.009	-0.860	0.450
STEMA - B[1]	-0.005	0.009	-0.526	0.642
STEMA - N[1]	0.032	0.009	3.634	**0.000**

[1] Leave-one-out procedure used

Table 3.5: Pairwise multiple comparisons for absolute errors of forecasting methods with winter SST included in the models for **giant grenadier**. P-values are adjusted using false discovery rate method. A p-value < 0.05 indicates the difference in absolute errors of the comparison are significant (in bold). Differences significantly less than zero indicate the first of the two compared methods was the method that produced smaller errors; differences significantly greater than zero indicate the second method produced smaller errors.

Linear Hypotheses	Estimate	Std. Error	z value	Pr(>|z|)
A[1] - A	-0.075	0.040	-1.887	0.068
B - A	0.274	0.040	6.797	$\mathbf{0.000}$
B - A[1]	0.349	0.040	8.661	$\mathbf{0.000}$
B[1] - A	0.271	0.040	6.733	$\mathbf{0.000}$
B[1] - A[1]	0.346	0.040	8.595	$\mathbf{0.000}$
B[1] - B	-0.002	0.040	-0.058	0.954
N[1] - A	0.010	0.040	0.255	0.856
N[1] - A[1]	0.085	0.040	2.134	$\mathbf{0.041}$
N[1] - B	-0.263	0.041	-6.500	$\mathbf{0.000}$
N[1] - B[1]	-0.261	0.041	-6.435	$\mathbf{0.000}$
STEMA - A	-0.367	0.040	-9.218	$\mathbf{0.000}$
STEMA - A[1]	-0.292	0.040	-7.333	$\mathbf{0.000}$
STEMA - B	-0.641	0.040	-15.895	$\mathbf{0.000}$
STEMA - B[1]	-0.639	0.040	-15.817	$\mathbf{0.000}$
STEMA - N[1]	-0.378	0.040	-9.478	$\mathbf{0.000}$

[1] Leave-one-out procedure used

four groundfish species, the STEMA method of forecasting had significantly reduced absolute errors when lagged winter SSTs were included as covariates. For Pacific cod, Pacific halibut, and giant grenadier, the models using the naïve method of forecasting benefited significantly from the addition of SST. The A and A[1] methods for all species had higher absolute errors upon the addition of SST to the models.

Table 3.6: One-sided Wilcoxon signed-rank test comparing the absolute errors of the rank-estimated GAMs including winter SST with the absolute errors of the null model. Mean absolute errors and standard deviations in parentheses are given. A p-value < 0.05 indicates the absolute errors of the models including winter SST are significantly smaller than the absolute errors of the null models. Methods with lowest forecast errors as determined by the pairwise multiple comparisons in Tables 3.2 to 3.5 are in bold.

Species	Method	Abs. Errors w/o SST	Abs. Errors w/ SST	P-value
Sablefish	A	1.455 (1.017)	1.771 (1.291)	1.000
	A[1]	1.396 (1.055)	1.688 (1.237)	1.000
	B	1.626 (1.053)	1.558 (1.034)	0.000
	B[1]	1.619 (1.059)	1.562 (1.046)	0.000
	N[1]	1.585 (1.186)	1.612 (1.161)	0.845
	STEMA	**1.386 (1.057)**	**1.369 (1.093)**	**0.006**
Pacific cod	A	0.157 (0.180)	0.280 (0.314)	1.000
	A[1]	0.165 (0.186)	0.267 (0.289)	1.000
	B	0.222 (0.146)	0.180 (0.120)	0.000
	B[1]	0.215 (0.138)	0.175 (0.116)	0.000
	N[1]	**0.142 (0.136)**	**0.136 (0.134)**	**0.000**
	STEMA	**0.153 (0.159)**	**0.145 (0.162)**	**0.000**
Pacific halibut	A	0.242 (0.153)	0.293 (0.286)	1.000
	A[1]	0.227 (0.147)	0.285 (0.277)	1.000
	B	0.193 (0.177)	0.205 (0.176)	1.000
	B[1]	0.193 (0.177)	0.204 (0.176)	1.000
	N[1]	**0.207 (0.181)**	**0.199 (0.179)**	**0.001**
	STEMA	0.208 (0.165)	0.205 (0.164)	0.011
Giant grenadier	A	1.041 (1.126)	1.599 (1.505)	1.000
	A[1]	1.055 (1.131)	1.480 (1.393)	1.000
	B	1.847 (1.381)	1.844 (1.572)	0.020
	B[1]	1.817 (1.391)	1.839 (1.590)	0.168
	N[1]	1.747 (1.474)	1.614 (1.386)	0.001
	STEMA	**1.149 (1.165)**	**1.086 (1.109)**	**0.000**

[1] Leave-one-out procedure used

3.7 Discussion

I propose a model averaging forecasting technique to capture both spatial and temporal information in ecological time series data. The method incorporates a flexible model capable of handling spatially-dependent covariates with the familiarity and forecasting ability of ARIMA models for time series analysis. I applied my method to catch data of four ecologically and commercially important species of groundfish where information regarding juvenile survival is often difficult to obtain and life history data are sparse or unknown, thereby making projections of population and catch challenging.

The A and A^1 models were inadequate for forecasting annual catch by location for any of the four species in the analysis. Previous studies indicated that ARIMA models outperform other linear time series methods when forecasting monthly data (Stergiou et al., 1997), however ARIMA is less suited to yearly data (Stergiou and Christou, 1996) and non-linear time series (Koutroumanidis et al., 2006). Spatial information is therefore an important component of modeling and forecasting catch in mobile marine species. A more flexible model, such as the varying coefficient model I employed, is also more desirable for capturing unknown nonlinear relationships between the response and predictors in complex systems.

My proposed STEMA method was always chosen as a preferred method for forecasting over ARIMAX and Bayesian models. It should be noted that catches for the two species in which STEMA did not significantly outperform the N^1 model were right-skewed, as noted in section 3.4. For these two species, the N^1 and STEMA methods which employed the rank-based estimation of Correia (2018) using a Gaussian distribution with bent score function outperformed the A, A^1, B, and B^1 techniques for forecasting. Correia (2018) showed that a bent score function in the estimation of generalized additive models (GAMs) improved model fit for Pacific cod catch over modeling with a Gamma distribution using a log link function. This Gamma distribution is one of the typical methods employed in fisheries research to deal with skewed catch data. However, the bent score function more appropriately accounted for skewness in the distribution of Pacific cod catch. The lower absolute forecasting errors for models using rank-based estimation (N^1 and STEMA) for Pacific cod and Pacific halibut data indicate

that the success of the bent score function to accommodate skewness also reaches to forecasting applications of varying coefficient models, which are an extension of GAMs. The application of the estimation techniques of Correia (2018) to the varying coefficient models used in STEMA takes advantage of the improved fit for heavy-tailed distributions common in fisheries data.

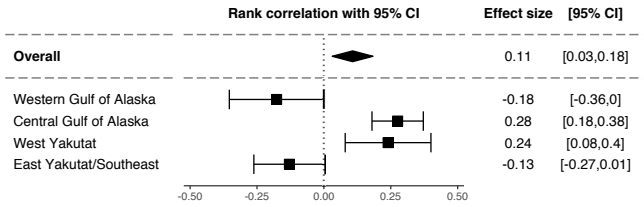
While the STEMA method did not beat the N^1 method in two of the species, naïve methods are notoriously difficult to beat in time-series forecasting, particularly for annual data (Athanasopoulos et al., 2011; Kilian and Taylor, 2003). The bent score function used in the estimation of the varying coefficient models in the N^1 and STEMA methods for Pacific cod and Pacific halibut reduces the effect of extreme values on estimation. This dampens large deviations in Pacific cod and Pacific halibut CPUE and produces fitted values closer to the mean CPUE. Naïve methods will invariably do better for very short term forecasts, because responses close to their mean values behave more like a random walk (Kilian and Taylor, 2003). The fact that STEMA was better than or equal to the N^1 method for short term forecasts in three out of the four species despite the known strengths of the naïve method illustrates the effectiveness of the STEMA method.

A statistically significant reduction in forecasting errors was discernible when winter SST was included for all preferred forecasting methods with lowest absolute errors. Thus, adding covariates relevant to the ecology of the species under consideration can significantly improve the forecasting power of a model. The inclusion of SST in the A and A^1 models increases the absolute forecasting errors for all species. Covariates in the ARIMAX and Bayesian settings are incorporated linearly, however the effect of winter SST on groundfish catch is likely to be nonlinear (Laurel et al., 2008; Rooper and Martin, 2009; Sadorus et al., 2014), which is apparent in Fig. S3 (Appendix A); therefore the effect's nonlinear shape is not being taken into account in the A, A^1, B, and B^1 forecasting methods.

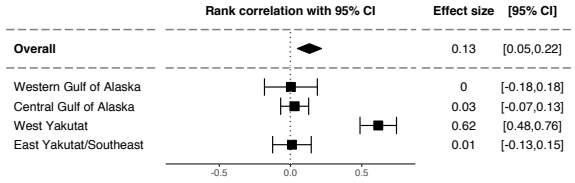
I broke down size of the effect of winter SST on groundfish catch for the preferred forecasting methods by management area as defined by the Alaska Fisheries Science Center for the MESA survey in Figure 3.1, where the order of the areas is from west to east along the coast of Alaska. Effect size of winter SST on the forecasting errors varies from none to large as defined by Cohen (1992), depending upon species and management area. It is likely that a given

species will respond to SST differently in different locations (Rouyer et al., 2014), which is evident by the variable effect sizes of winter SST by station provided in Figs S4 to S8 (Appendix A). Other factors such as habitat, prey availability, and proximity to other individuals of the same species may influence the effects of SST on survival of juveniles. For example, while all these groundfish are not schooling species, Stoner and Ottmar (2004) found that young Pacific halibut were more likely to locate and attack baits in groups than when solitary. Therefore Pacific halibut, which experience reduced ability to locate bait in low temperatures, may instead successfully find bait in the presence of other individuals. SST may also be a proxy for other environmental variables, such as dissolved oxygen (DO) levels, ocean mixing, and plankton availability, that may affect these groundfish to varying degrees. Sadorus et al. (2014) found a significant relationship between DO and catch rates of Pacific halibut. Primary production (plankton) concentration and distribution and subsequent changes in secondary production levels have also been linked to groundfish abundance (Francis et al., 1998; McGowan et al., 1998). Correia (2018) found improved prediction when adding winter SST to models for sablefish and Pacific cod catches, however model fit did not substantially improve with the addition of winter SST. Therefore the link of SST to groundfish catches is likely complex and difficult to quantify directly in wild populations.

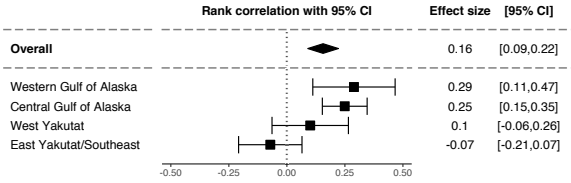
I have shown that spatial information is crucial to forecasting in large-scale data, and my spatio-temporally explicit model averaging technique is successful in reducing forecasting errors. Additionally, the inclusion of environmental covariates can improve forecasting in many cases. As is the case with forecasting and prediction techniques, predictions outside the range of observed covariates (i.e. extrapolation) is ill-advised (Conn et al., 2015; Steyerberg et al., 2010). Forecasts more than one time point ahead can be achieved for the STEMA technique via the `forecast` function after fitting the ARIMA models in the spatial model with ARIMA (Section 3.4.2.1) and temporal model with ARIMA (Section 3.4.2.2). The leave-one-out procedure and model averaging would be performed as described (Section 3.4.2.3) for each time point for which forecasts were estimated. While the proposed technique is only suitable to forecast future, regular time points for the same locations, this is typically desirable for many ecological and epidemiological analyses where predicting the status of a fixed population at



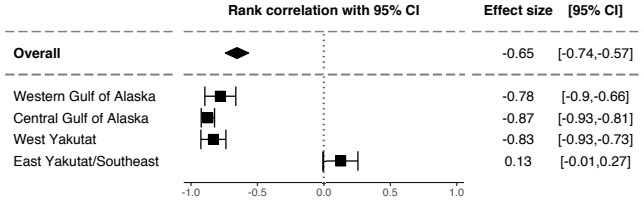
(a)



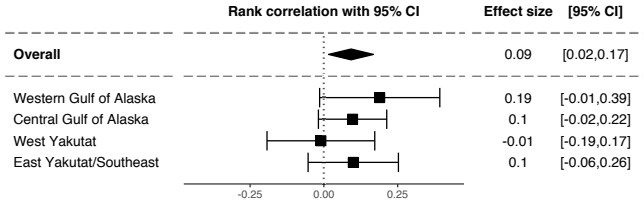
(b)



(c)



(d)



(e)

Figure 3.1: Effect size of lagged winter SST on CPUE of each species broken down by management area using best forecasting method as determined by the results of pairwise multiple comparisons: (a) Sablefish (STEMA), (b) Pacific cod (N^1), (c) Pacific cod (STEMA), (d) Pacific halibut (N^1), and (e) Giant grenadier (STEMA). The rank correlation r statistic is given as the effect size.

future time points is desired. It would be feasible to extend STEMA-generated forecasts to new locations by using any of several spatial interpolation methods including inverse distance weighting, kriging, and smoothing splines. Migratory and irregular population values can also be forecast provided seasonality is appropriately accounted for in the ARIMA model structure portion of the STEMA method. The STEMA technique is also as intuitive, accessible, and simpler to deploy than other forecasting methods compared in this paper, making it a suitable forecasting method for population ecology, fisheries and wildlife management, vector-borne disease research and monitoring, and econometrics.

Chapter 4

Selecting environmental covariates affecting adult groundfish catches in the Gulf of Alaska

Abstract

Fisheries management organizations are interested in monitoring economically important marine fish species and using this information to inform management strategies. Determining environmental factors that best predict changes in these populations are therefore a priority to ecologists. I illustrate the application of the least squares-based spline estimation and group LASSO (LSSGLASSO) procedure for selection of coefficient functions in single index varying coefficient models (SIVCMs) on an ecological data set including environmental covariates suspected to play a role in the catches and weights of six groundfish species. Temporal trends in variable selection were apparent, though the selection of variables was unrelated to common North Pacific climate indices. These results indicate that the strength of an environmental variable's effect on a groundfish population may change over time, and not necessarily in-step with known low-frequency patterns of ocean-climate variability commonly attributable to large-scale regime shifts in the North Pacific.

Keywords and phrases: Single index model, group LASSO selection

4.1 Introduction

It is well-understood that the northern Pacific system is controlled by multiple interdecadal patterns of climate variability that stem from different physical sources (Francis et al., 1998). Groundfish populations in the northeastern Pacific Ocean followed the six- to 12-year warming and cooling periods of the El Niño-Southern Oscillation (ENSO) (Hollowed and Wooster, 1992). Sea surface temperature and pressure changes in the North Pacific are captured by the Pacific Decadal Oscillation (PDO), which is mainly separate from ENSO behavior in the region (Mantua and Hare, 2002). A third climate cycle described recently by Di Lorenzo et al. (2008) and termed the Northeast Pacific Gyre Oscillation (NPGO) follows variations in ocean nutrient cycling and phytoplankton abundance and plays a role in the larger system of climate variability with ENSO and PDO (Di Lorenzo et al., 2013).

Dramatic, permanent changes in marine species compositions in response to shifts in climate modes (commonly referred to as regime shifts) such as the strong one observed in 1976-1977 may likely be the convergence of several climate patterns switching phases within the same time period. This switching of regimes makes it difficult to identify which specific patterns are culprits in affecting distinct marine populations, particularly in deepwater populations where relationships between the marine environment and atmospheric trends are more nuanced and may involve complex lagged effects. Many studies on fisheries systems continue to focus on these interdecadal climate modes as primary sources of variability in population sizes of marine fishes. Litzow et al. (2014) highlighted the insufficiency of climate modes alone to accurately describe variability found in many commercially valuable marine populations on multi-year scales. Other sources of oceanic variability not directly linked to climate modes exist and should be considered when attempting to create accurate models to describe and predict changes in marine populations, even in areas that appear to be dominated by shifts in climate regimes. The complicated interplay of ocean-climate systems in the North Pacific region makes it difficult to identify which and how specific indices are culprits in affecting distinct marine populations.

Many of the marine fishes in the Northern Pacific Ocean are commercially important species that contribute significantly to the economy of the United States and are an important source of food both domestically and internationally (Council, 2016; Goen and Erikson, 2017; Johnson et al., 2016). Several of these populations are managed by international or regional fishing commissions to control commercial harvests and monitor population health (Goen and Erikson, 2017; NPFMC, 2017; Pennoyer and Balsiger, 1998; Rodgveller et al., 2008). These organizations are becoming increasingly concerned about the role climate plays in maintaining healthy fish populations, especially as marine fishes do not recover from population collapses as quickly as previously believed (Hutchings, 2000). Fishing activities are becoming increasingly concentrated on deeper-dwelling species (Moore and Mace, 1999; Moore, 1999). While focus on the effect of various climate modes has dominated ecological research on fishes in the North Pacific region, relationships between the marine environment and atmospheric trends are nuanced and may involve complex lagged effects, particularly for deepwater populations (Rijnsdorp et al., 2009). It is also problematic to study organisms that inhabit the deep ocean, as they are not adapted to surface-level conditions and prove difficult to sample and keep alive, making experiments in laboratory conditions impossible or prohibitively expensive. Determining which specific environmental variables contribute to fluctuations in the populations of these species from observational data would represent major progress in discerning the impact of climate variability on marine ecosystem health and how those changes affect the economy and food security. A model structure able to accommodate a suite of environmental variables that vary spatiotemporally would be necessary to examine effects of many environmental covariates on deepwater marine populations simultaneously.

Consider the single index variable coefficient model (SIVCM) of the form

$$y_i = \{G(\boldsymbol{\theta}_0^T Z_i)\}^T X_i + \varepsilon_i \quad i = 1, \ldots, n \ , \tag{4.1}$$

where $\boldsymbol{\theta}_0$ is a vector of unknown coefficients representing the single-index direction, $G(\cdot) = \{g(\cdot)_0, \ldots, g(\cdot)_p\}^T$ are nonparametric coefficient functions, and ε are the random errors. The

SIVCM is a convenient structure for incorporating spatiotemporal effects for multiple environmental predictors. Forward selection, backwards elimination, and stepwise selection methods are unstable for models with many predictors and even with advancements to the algorithms, these methods are considered sub-optimal for variable selection, particularly for high-dimensional models (Cai et al., 2009). Penalty-based regression procedures, such as ridge regression and least absolute shrinkage and selection operator (LASSO) estimation, penalize large regression coefficients to reduce overfitting. LASSO additionally performs variable selection by penalizing small regression coefficients to zero, effectively removing these coefficients from the model. LASSO works particularly well for models with many predictors because it shrinks large coefficients to zero rather than minimizing them, and it is computationally efficient (Ledolter, 2013). Group LASSO incorporates information about groupings of variables into the penalty function, which is particularly important for categorical predictor variables (Meier et al., 2009). While selection for varying coefficient models (VCMs), a lower-order relative of the SIVCM, have built on both the smoothly clipped absolute deviation (SCAD) and LASSO approaches (Fan et al., 2003; Matsui and Misumi, 2015; Wang and Xia, 2009; Xue and Qu, 2012), selection procedures of the single-index direction coefficients or the functions in SIVCMs has so far exclusively used SCAD penalties (Feng and Xue, 2013; Song et al., 2016; Yang and Yang, 2017). SCAD procedures are unbiased, but they are sensitive to initial estimation and parameter tuning (Xue and Qu, 2012). LASSO procedures are typically simpler to implement than SCAD, and group LASSO has been shown to correctly select important variables for VCM where the number of dimensions far exceeds the number of observations (Wei et al., 2011). Here, I will use a combination of least squared-based spline estimation and group LASSO proposed by Sun (2017) to select coefficient functions and estimate the index parameters in a SIVCM of spatiotemporally-varying environmental covariates potentially contributing to changes in groundfish populations in the North Pacific Ocean. With this application, I aimed to establish relevant environmental conditions that affect populations of focal groundfish species in this region.

4.2 Methods

Annual surveys of several groundfish species are taken at established locations in the waters along the coast of Alaska by the Alaska Fisheries Science Center (AFSC), a division of the National Oceanic and Atmospheric Administration (NOAA). Catch per unit effort (CPUE) and mean weight in kg of six groundfish species determined at each location for each survey year were obtained for years 1979–2013 (Alaska Fisheries Science Center, 2019a; Sigler and Lunsford, 2009). Air temperature in degrees Celcius (ATMP), sea level pressure in hPa (PRES), wind speed in meters per second averaged over eight-minute periods (WSPD), sea surface temperature in degrees Celsius (WTMP), and the average height in meters of the highest one-third of all waves in 20-minute sampling periods (WVHT) measured daily from buoys in the Gulf of Alaska were obtained from the National Data Buoy Center and summarized by monthly means (National Data Buoy Center, 2018a). Temperature in degrees Celsius measured at the sea floor (hereafter bottom temperature) was obtained from the AFSC Resource Assessment and Conservation Engineering (RACE) Division's bottom trawl surveys (Alaska Fisheries Science Center, 2019b). Zooplankton biomass volume given in number per cubic meter were obtained from the NOAA's Coastal and Oceanic Plankton Ecology, Production, and Observation Database (O'Brien, 2007b). Alkalinity (Alk), chlorophyll (Chl), nitrate (NO_3), dissolved oxygen (Oxy), phosphate (Phos), and silicate (Sil) concentrations at depths of 75, 400, and 900 meters were obtained from the NOAA's World Ocean Database (Boyer et al., 2013).

Since measurements of environmental variables were not measured at the same locations across all years for which groundfish surveys were performed, spatiotemporal interpolation via inverse distance weighting was used to obtain environmental measures at exact locations where MESA surveys were conducted (Li et al., 2014b). For all environmental variables, a seasonal amplitude was calculated for each survey year. For physical variables ATMP, PRES, WSPD, WTMP, WVHT and bottom temperature, seasonal amplitude was defined as the mean of June, July, and August minus the mean of December, January, and February because of temperatures maximizing in the summer and minimizing in winter and winter storms enabling strong mixing of ocean nutrients. Seasonal amplitude for chemical and biological variables

including plankton, chlorophyll, alkalinity, nitrate, dissolved oxygen (DO), phosphate, salinity, and silicate was calculated as the mean of August, September, and October minus the mean of March, April, and May. Zooplankton biomass in May is increasing after winter vertical mixing of deepwater nutrients and benefiting from spring plankton blooms but has not yet been depleted by grazing from summer-migrating pelagic fishes and cephalopods (Brodeur and Ware, 1992; Chiba et al., 2006). Therefore nutrients are generally maximal in the spring and minimal in fall after depletion by phytoplankton, whereas zooplankton and chlorophyll would be maximal in fall after nutrient consumption and growth over the summer in the North Pacific region (Childers et al., 2005; Sackmann et al., 2004; Wong et al., 2002a).

A SIVCM of the form given in (4.1) was fit, where y_i was the CPUE or mean weight of a groundfish at each location for each year, $X = (\mathrm{x}_{0i}, \ldots, \mathrm{x}_{pi})^T$ with $\mathrm{x}_{0i} = 1$ and $\mathrm{x}_{1i}, \ldots, \mathrm{x}_{pi}$ being the seasonal amplitudes of environmental variables described previously, and $Z = (z_{1i}, z_{2i}, z_{3i})^T$ are the longitude, latitude, and year for each observation. Heat maps were used to visualize selection of variables per year and over all years for each groundfish species. The dissimilarity between variables was calculated, where distances between variable x and y was calculated as the L_2 norm, $\sqrt{\sum_i (x_i - y_i)^2}$. Dendrograms were added to the heat maps to visualize dissimilarities. After selection, models for all years were refit with scaled $Y, X,$ and Z using only variables selected by LSSGLASSO to make the coefficient functions comparable between different responses.

To further explore if regional climate conditions had an effect on the variables being selected each year, I considered potential relationships between the selection of a variable over time and corresponding climate indices. The Pacific Decadal Oscillation (PDO) monthly index, multivariate El Niño/Southern Oscillation bi-monthly index (MEI), and North Pacific Gyre Oscillation (NPGO) monthly index for 1979–2013 were obtained from Di Lorenzo (2018), NOAA ESRL Physical Sciences Division (2019) and Mantua and JISAU, University of Washington (2016). Annual signs for climate indices were defined for each year as the sign of the annual mean, where $+1$ indicated more positive monthly indices than negative indices in a year, -1 indicated more negative than positive monthly indices in a year, and 0 indicated equal number of positive and negative indices in a year. Logistic regression was then used to fit the model

$g(E(y_i)) = \beta_0 + \beta_k x_{ki}$ for years $i = 1979, \ldots, 2013$ for each variable within a groundfish response, where $g(\cdot)$ was a log-link function; β_0 was the intercept term; β_k were linear coefficients, $k \in \{\text{PDO, MEI, NPGO}\}$; x_{ki} were annual signs of k index; and y_i were binary indicators of whether a variable was or was not selected for each year. A Chi-square test was used to compare the models to an intercept-only model. P-values from the Chi-square test were adjusted within groundfish response to control for the false discovery rate (FDR) of multiple testing using the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995).

4.3 Results

Variables consistently selected by LSSGLASSO as important to Pacific cod CPUE were Alk_75m, plankton, Chl_75m, Sil_75m, bottom temperature, Sil_900m, and Chl_400m, which usually were in years 1985, 1990, 1992, 1995, and 2010 (Figure 4.1). Variables consistently selected in 1992, 1996, and 2002 for Pacific cod weight were bottom temperature, plankton, Alk_75m, Alk_400m, Alk_900m, Chl_75m, Chl_400m, Oxy_75m, and Sil_900m (Figure 4.2). Alk_400m, Alk_900m, Chl_75m, Chl_900m, Sal_400m, Sal_900m, WTMP, PRES, and WVHT in years 1991-1992, 1997-1999, and 2013 were selected variables for Pacific halibut CPUE, whereas variables important for Pacific halibut weight included Alk_75m, Chl_400m, Chl_900m, plankton, Oxy_75m, Oxy_400m, and Sil_75m for years 1989, 1991, 1993, 1994 (Figures 4.3 and 4.4). Chl_900m, plankton, NO_3_900m, and Oxy_400m were consistently selected together in years 1984 and 1992 for sablefish CPUE, while WSPD and ATMP were the only variables selected as important to sablefish weight more than once over the 30-year period of record (Figures 4.5 and 4.6). A wide array of environmental variables were selected for rougheye rouckfish CPUE consistently and almost exclusively in years 1991, 1995, 2001, 20011, and 2013, whereas 1990 was the only year when more than one variable was selected for rougheye rockfish weight (Figures 4.7 and 4.8). This was also the case for shortraker rockfish, where most variables were selected in 1992, 1998, 2001, 2001, 2005, 2007, and 2009 for CPUE and only selected in 1994, 2003, and 2008 for weight (Figures 4.9 and 4.10). Chl_900m, Alk_900m, Oxy_400m, Sal_75m, and Sil_900m as a suite were selected for shortspine thornyhead CPUE in 1984 and 1986, however many environmental variables contributed to thornyhead weight in 1984-1986, 2000, and

2004 (Figures 4.11 and 4.12). Only sablefish CPUE, rougheye rockfish CPUE and weight, and Pacific cod CPUE had variables selected as important predictors of groundfish CPUE or weights when analyzing selection across all available years. Alk_900m was the main selected variable for Pacific cod and sablefish CPUEs for all years included in the selection procedure. Annual signs of PDO and NPGO climate indices in the North Pacific were not significantly related to the selection of any environmental variables; however, MEI was related to the selection of several environmental variables in the Pacific cod weight SIVCM (Figure 4.13 and Tables C.1 to C.3).

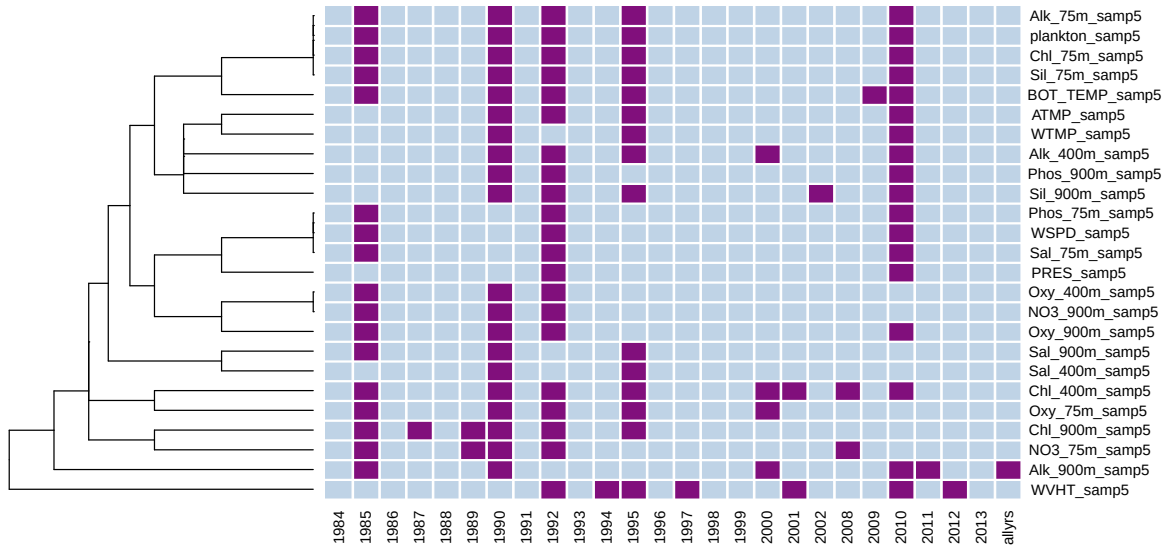


Figure 4.1: Variable selection analyzed for each available year and all years together for the SIVCM with Pacific cod CPUE as the response. X-axis represents years for which selection was performed, and all available years of data used in selection labeled as "allyrs". Y-axis contains all variables in SIVCM from which selection was performed. Dark purple indicates the selected variable(s). Dendrograms represent the dissimilarity between selected variables calculated using the L_2 norm.

I plotted the selected coefficient functions of WTMP, Alk_400m, Alk_900m, Sil_75m, and Sil_900m for sablefish CPUE using all years of data to provide a detailed example of interpretation of the SIVCM functions (Figure 4.14). Parameters representing the single-index direction estimated by the LSSGLASSO procedure were $\widehat{\boldsymbol{\theta}} = (0.9372, -0.3296, -0.1146)$. For a given location specified by a fixed longitude-latitude pair, trends of sablefish catch for a given coefficient function are opposite what is shown because of the negative value of the year coefficient

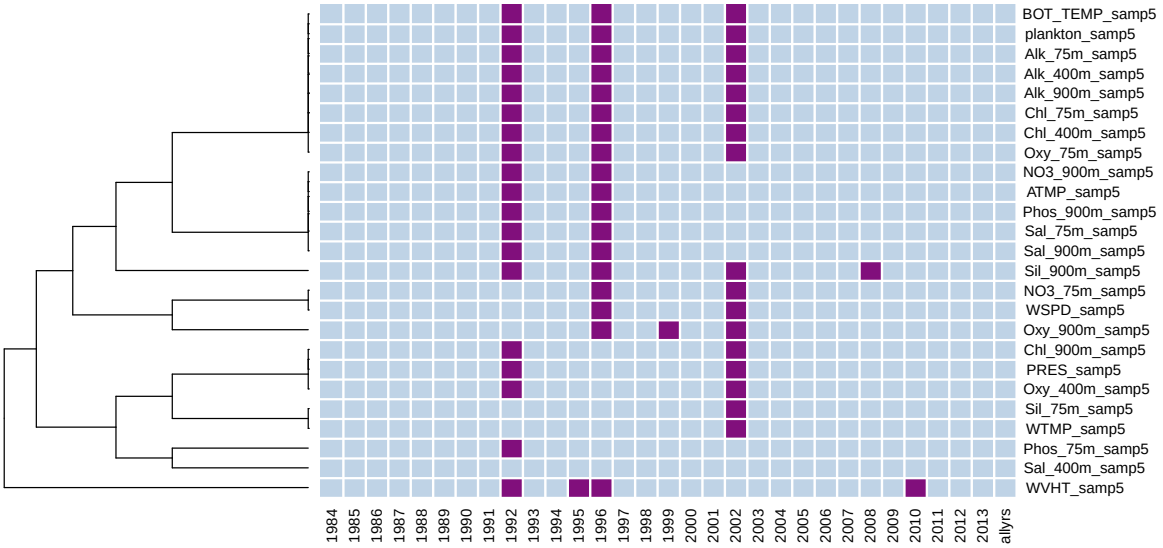


Figure 4.2: Variable selection analyzed for each available year and all years together for the SIVCM with Pacific cod weight as the response. Axes, colors, and dendrograms are as described in Figure 4.1.

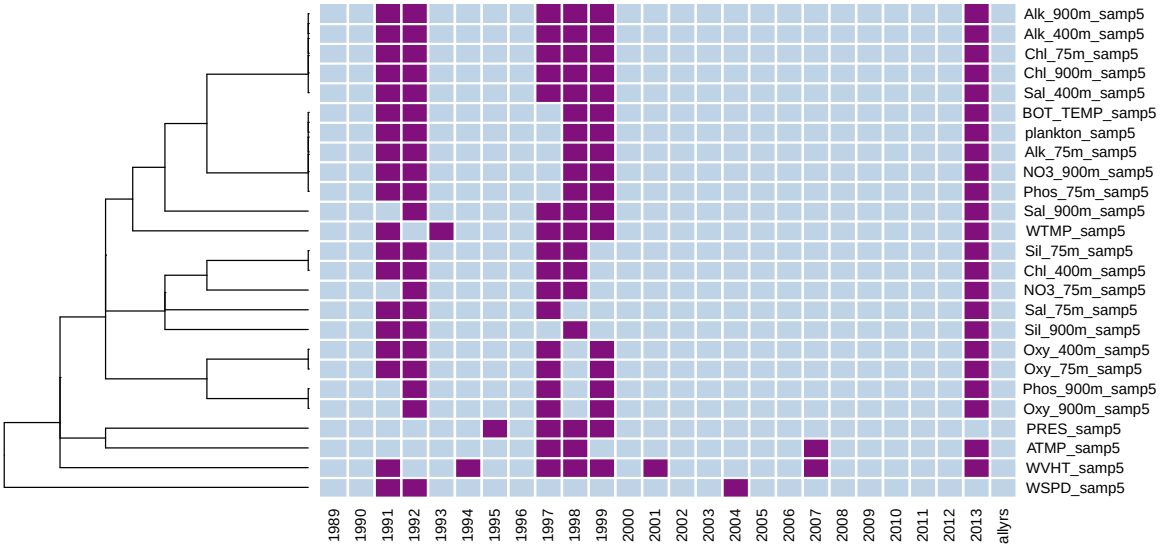


Figure 4.3: Variable selection analyzed for each available year and all years together for the SIVCM with Pacific halibut CPUE as the response. Axes, colors, and dendrograms are as described in Figure 4.1.

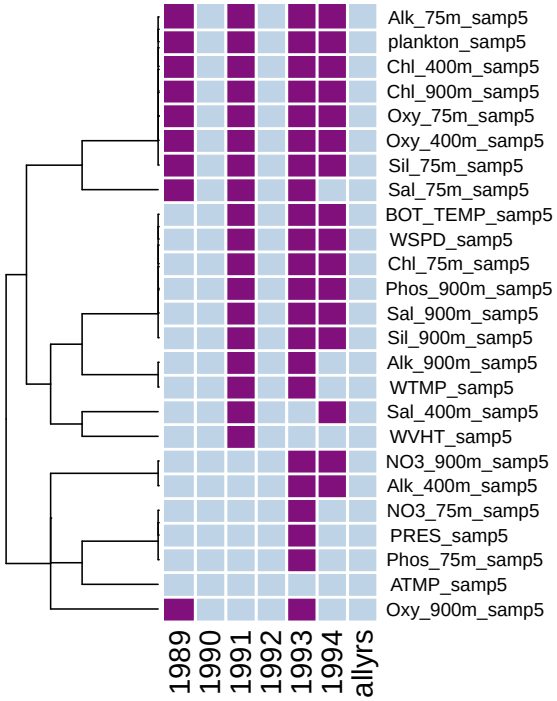


Figure 4.4: Variable selection analyzed for each available year and all years together for the SIVCM with Pacific halibut weight as the response. Axes, colors, and dendrograms are as described in Figure 4.1.

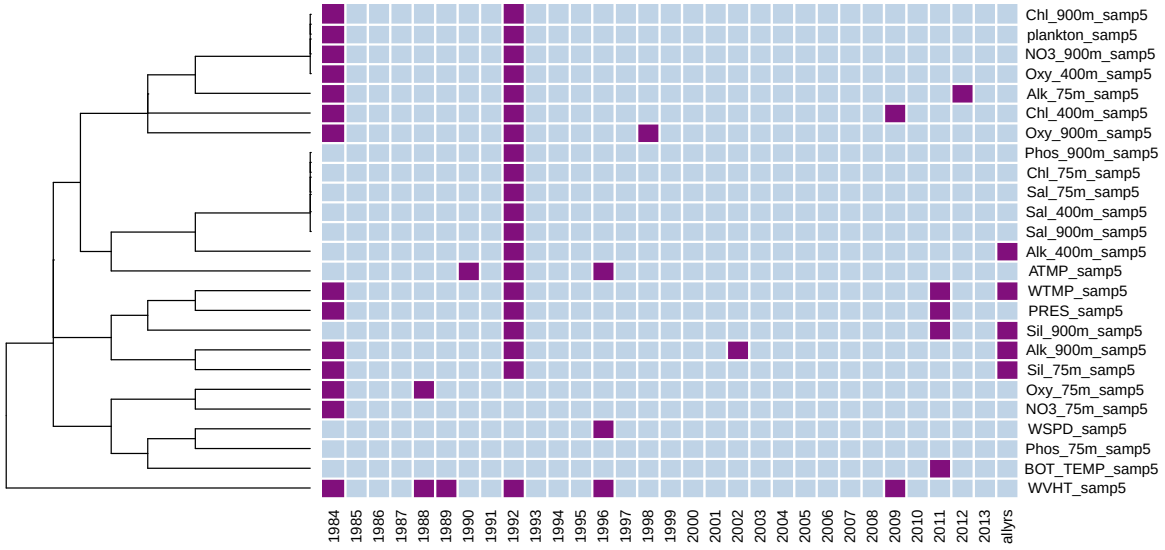


Figure 4.5: Variable selection analyzed for each available year and all years together for the SIVCM with sablefish CPUE as the response. Axes, colors, and dendrograms are as described in Figure 4.1.

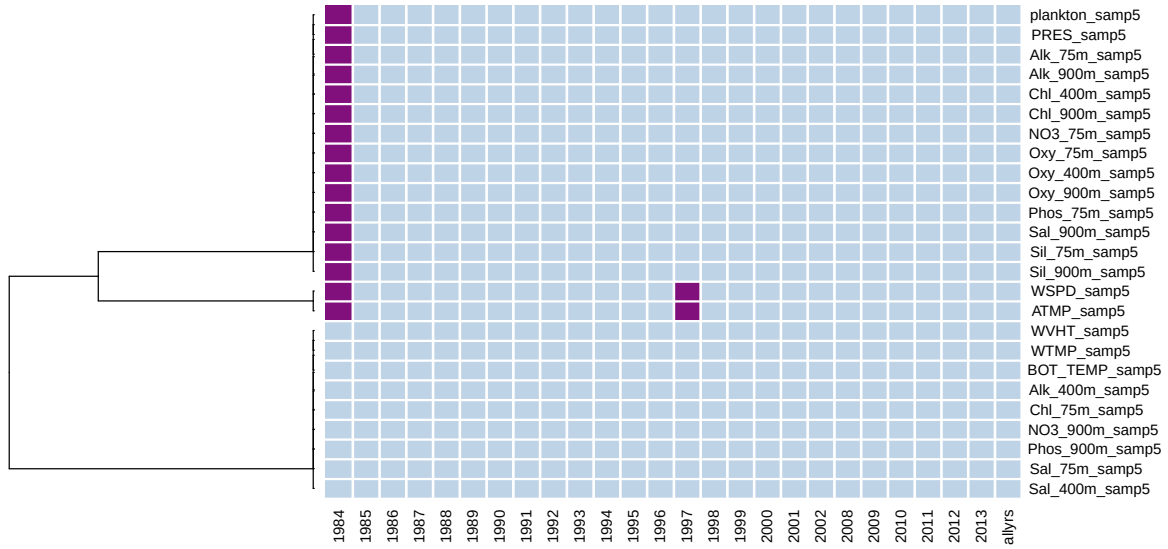


Figure 4.6: Variable selection analyzed for each available year and all years together for the SIVCM with sablefish weight as the response. Axes, colors, and dendrograms are as described in Figure 4.1.

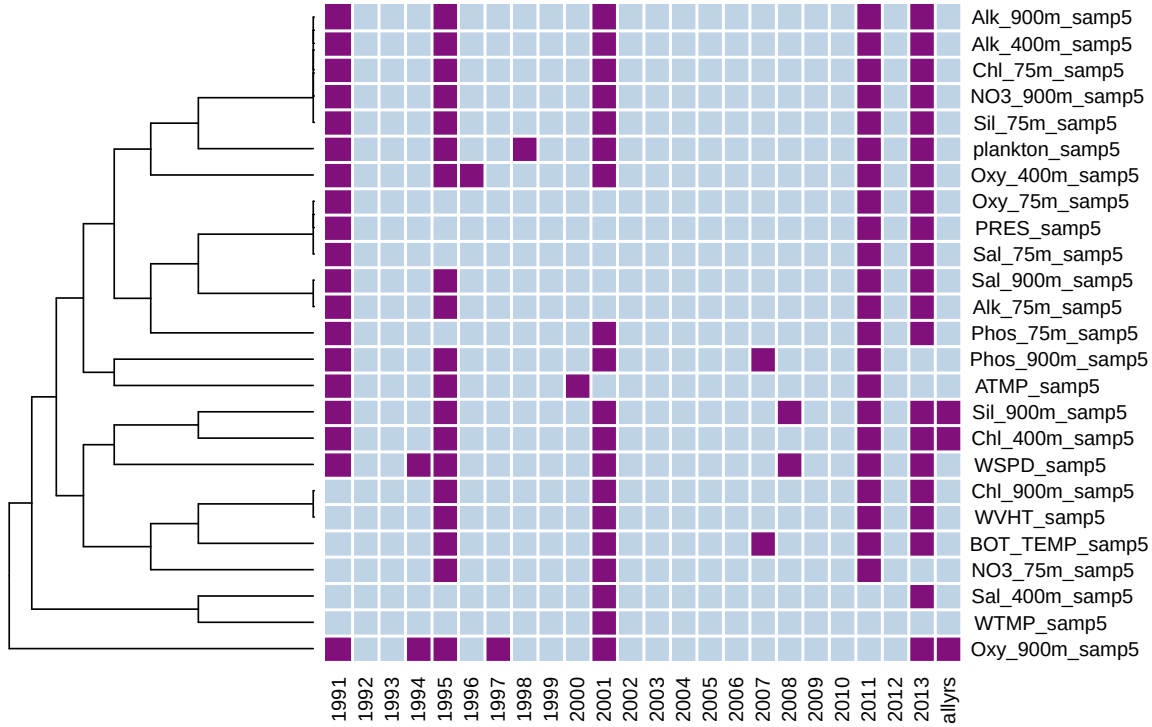


Figure 4.7: Variable selection analyzed for each available year and all years together for the SIVCM with rougheye rockfish CPUE as the response. Axes, colors, and dendrograms are as described in Figure 4.1.

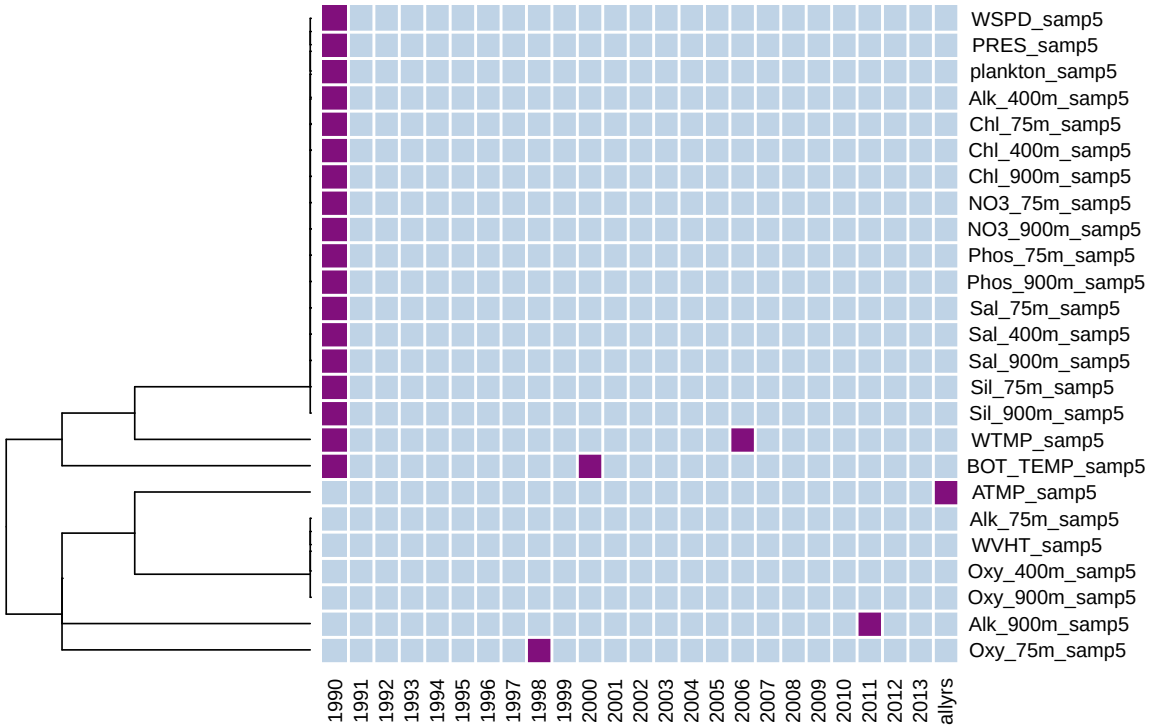


Figure 4.8: Variable selection analyzed for each available year and all years together for the SIVCM with rougheye rockfish weight as the response. Axes, colors, and dendrograms are as described in Figure 4.1.

Figure 4.9: Variable selection analyzed for each available year and all years together for the SIVCM with shortraker rockfish CPUE as the response. Axes, colors, and dendrograms are as described in Figure 4.1.

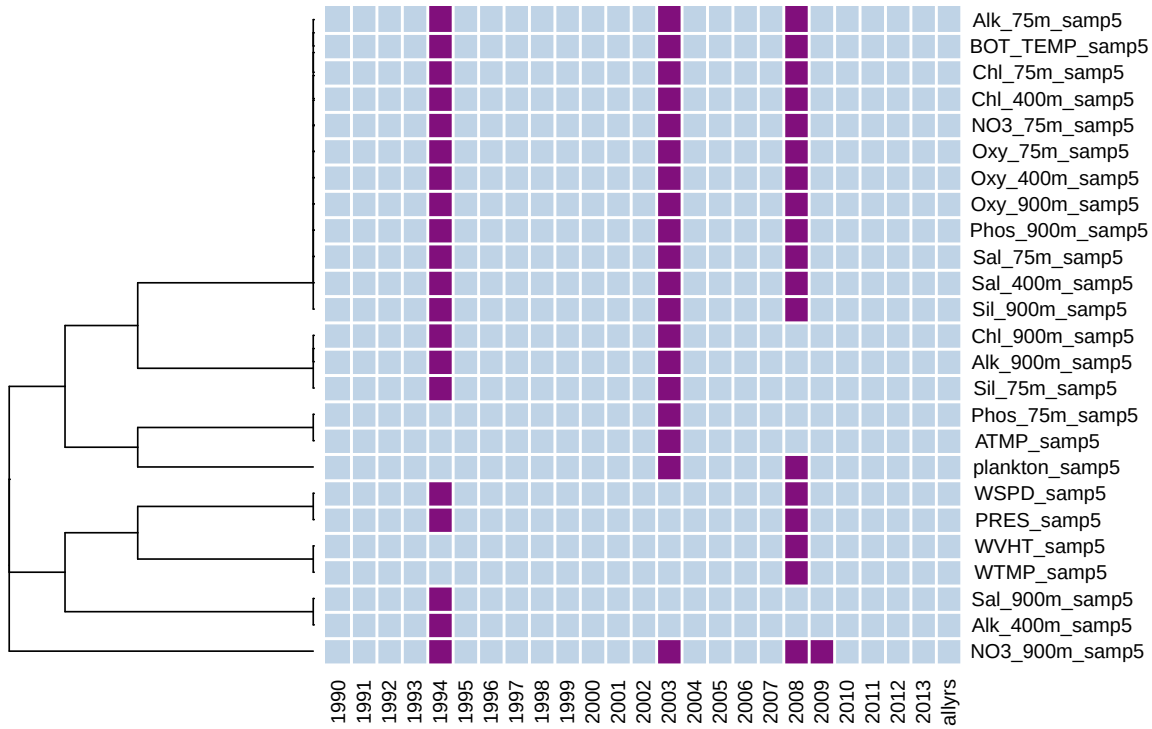


Figure 4.10: Variable selection analyzed for each available year and all years together for the SIVCM with shortraker rockfish weight as the response. Axes, colors, and dendrograms are as described in Figure 4.1.

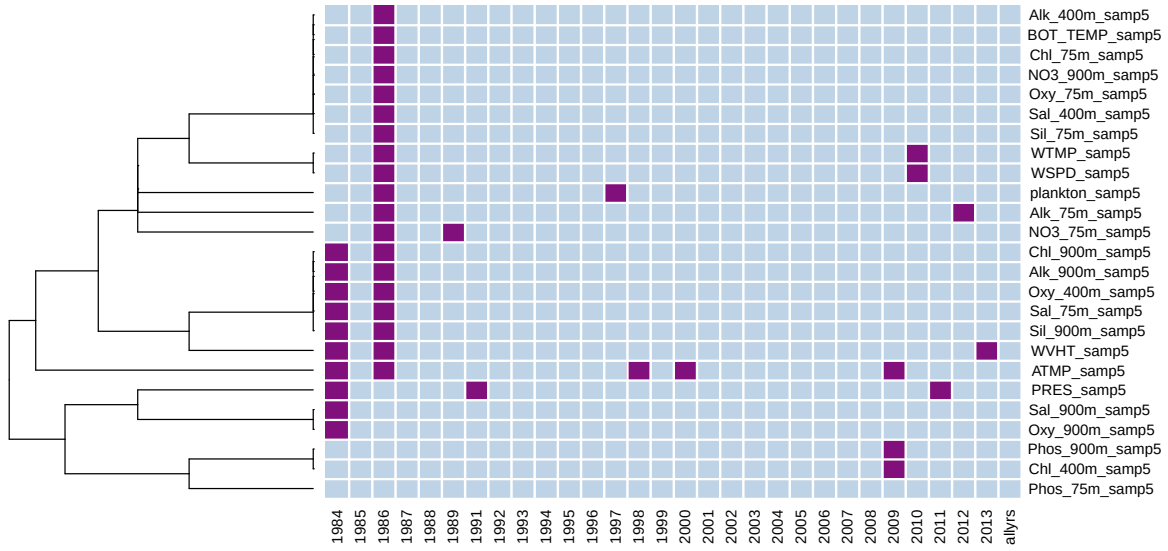


Figure 4.11: Variable selection analyzed for each available year and all years together for the SIVCM with shortspine thornyhead CPUE as the response. Axes, colors, and dendrograms are as described in Figure 4.1.

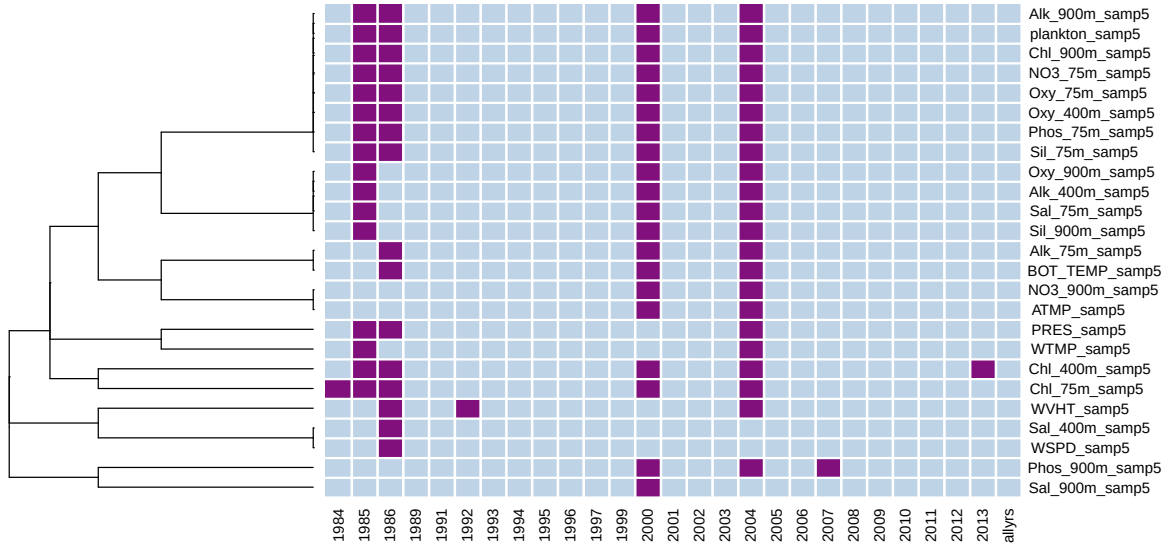
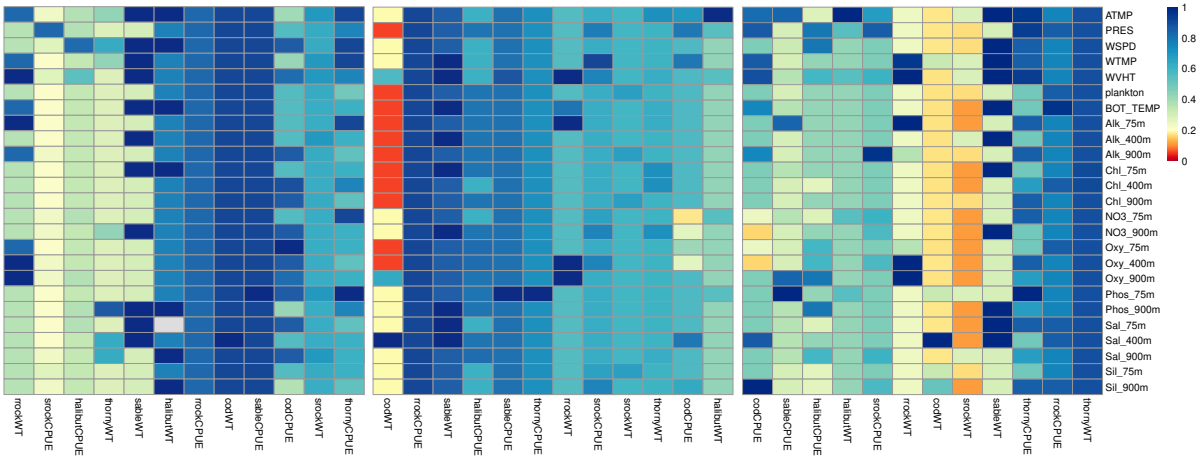


Figure 4.12: Variable selection analyzed for each available year and all years together for the SIVCM with shortspine thornyhead weight as the response. Axes, colors, and dendrograms are as described in Figure 4.1.



(a) PDO seasonal amplitude (b) MEI seasonal amplitude (c) NPGO seasonal amplitude

Figure 4.13: FDR-corrected p-values for relationship between selection of each environmental variable contributing to groundfish CPUEs or WTs each year and yearly signs of climate indices PDO (a), MEI (b), and NPGO (c) fitted using logistic regression. Abbr: cod = Pacific cod; halibut = Pacific halibut; sable = sablefish; rrock = rougheye rockfish; srock = shortraker rockfish; thorny = shortspine thornyhead; CPUE = catch per unit effort; WT = mean weight in kg. $P < 0.1$ are indicated by red colors. Grey indicates a p-value could not be calculated due to insufficient sample size ($n = 6$) for the model.

in the single-index direction. Sablefish catch initially decreased sharply, followed by a moderate increase over more recent years (approximately 2003–2013), as indicated by the intercept function in Figure 4.14. The coefficient functions for WTMP and Sil_900m showed these having decreasing effects on mean sablefish CPUE in earlier (1984–1993) and later (2005-2013) years and an increasing effect on sablefish catch in 1994–2004. The effect of Alk_400m at a given location had a decreasing effect on sablefish catch through around 1993 before increasing rapidly. Alk_900m shared a similar trend for the latter years of the data, but tended towards a more cubic relationship on sablefish catch for a given location. Sil_75m showed an overall negative effect on sablefish CPUE for increasing time at a given location until about 1997, when the effect on sablefish increased before resuming its negative relationship from around 1999 onward.

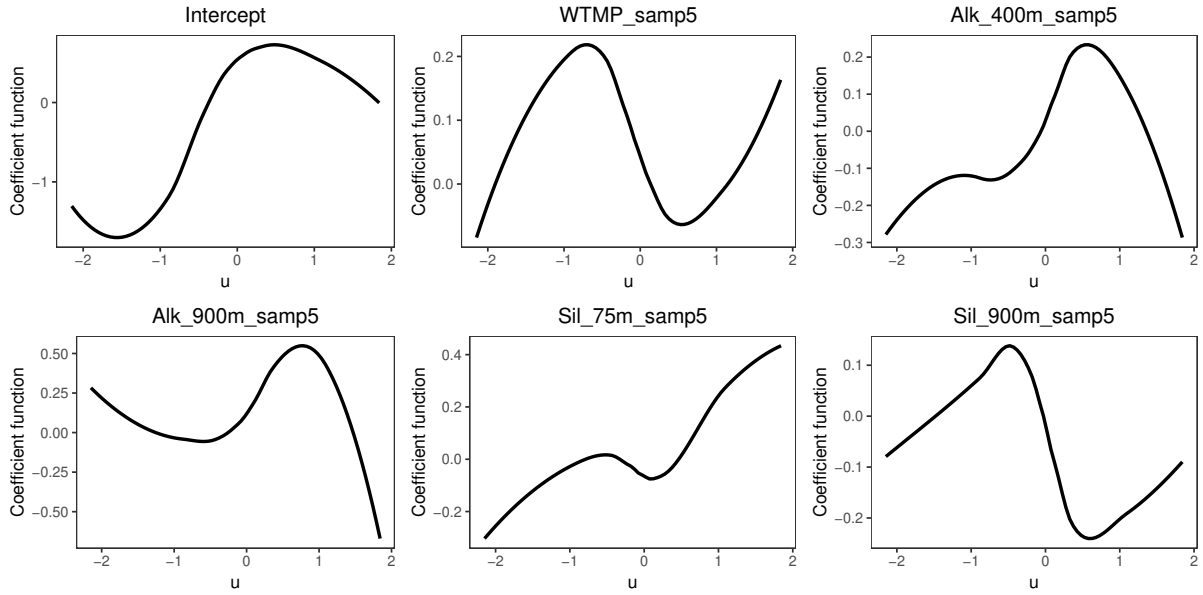


Figure 4.14: Coefficient functions selected and estimated by LSSGLASSO as significant predictors of sablefish CPUE.

4.4 Discussion

Selection of environmental variables appeared to be unrelated to annual signs of both the PDO and NPGO climate indices. However, large-scale climate patterns have previously been shown to be associated with groundfish responses (Francis et al., 1998; Hollowed and Wooster, 1992). To explore whether PDO, MEI, or NPGO were good estimators of groundfish catches and

weights and if the effect of a climate index on the selection of variables is related to that index's effect on the groundfish populations directly, I fit GAMs with seasonal amplitudes of PDO, MEI, and NPGO (lagged by five years for all groundfish except Pacific halibut, which was lagged 10 years) as three additive smooth predictors and groundfish CPUEs or weights as responses (Wood, 2006). P-values were estimated for each smooth predictor in each model and plotted in a heatmap for efficient visualization (Figure 4.15). Lagged PDO and MEI were good estimates of sablefish CPUE and provided a moderate fit for the data (adjusted R^2=0.445). Pacific cod weight was mostly explained by lagged PDO, with the GAM fitting the data well (adjusted R^2=0.643). Lagged PDO was also a good estimator of rougheye rockfish weight and provided a good fit to the data (adjusted R^2=0.642). Therefore, climate indices are not universally ideal for modeling groundfish CPUE and mean weights along the coast of Alaska.

Interestingly, Pacific cod weight was best predicted by PDO annual signs and not those of MEI, which was related to the selection of variables affecting cod weight. This result suggests that the selection of variables affecting cod weight was not the result of an existing, strong effect of MEI on cod populations, but rather the existence of several pathways of climate and/or other environmental effects on groundfish responses. These indices could be useful predictors of fish catches and weights, although there is no way to ascertain what specific changes in the atmosphere and ocean summarized by these indices affect these fish. Therefore, these climate indices may be suitable for prediction or describing long-term temporal trends, but are less useful for inference and motivating specific further study. For select groundfish whose catches or weights can be reasonably predicted by one or more of the climate indices, separate pathways of environmental effects on responses of groundfish populations may exist.

The coefficient functions of the SIVCM for sablefish CPUE confirmed the nonlinear relationships of physical and environmental covariates to groundfish catch in Chapter 2 even when spatiotemporal effects are accounted for across all predictors. As shown by my examination of atmospheric oscillations to the selection of variables and the groundfish responses, temporal patterns in specific groups of variables being selected can be used to consider associations of large-scale ecosystem variations to variable selection behavior. Deepwater species such as

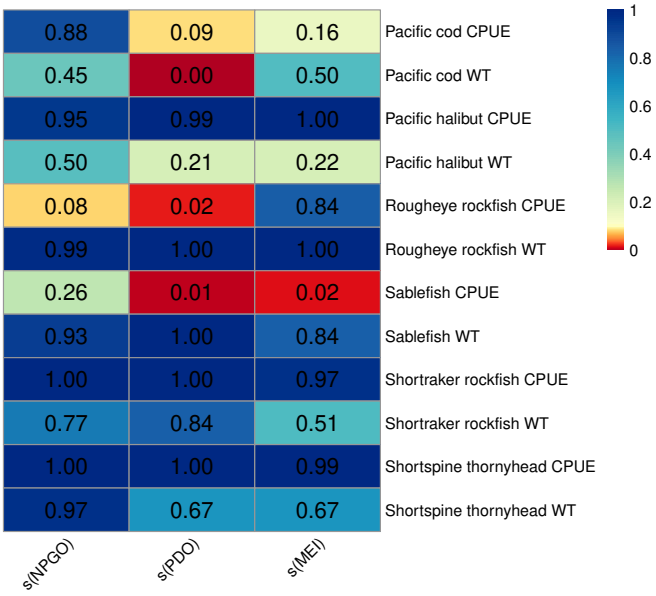


Figure 4.15: Heatmap of p-values for nonparametric relationship between groundfish CPUEs or WTs each year and lagged yearly seasonal amplitudes of climate indices PDO, MEI, and NPGO fitted using GAMs. Colors and abbr. are as described in Figure 4.13.

the groundfish examined in this study are difficult to monitor regularly, so determining environmental variables that most accurately predict population health of deep water fishes and incorporating these into forecasting models are useful for management organizations to make more appropriate recommendations.

Chapter 5

Spatial convergent cross mapping for a predator-prey system in the North Pacific

Abstract

Historically, causal dynamics in ecology have been difficult to analyze and reproduce, especially in field data, although recently causality-based applications have been advanced. I explored whether one such method, convergent cross mapping (CCM), could be used on realistic ecological data to identify potential causal dynamics and produce clearly interpretable results. Spatial CCM was applied to a dataset containing population data of North Pacific deepwater sablefish, Pacific cod, and Pacific halibut, along with lagged environmental variables. Some results were intuitively expected, but in other cases there was no clear intuition and false positives were suspected when prior knowledge of putative causal relationships were not already well-established for the system. It is clear that the technique is only suitable for confirming causal dynamics in well-studied systems or as part of exploratory analyses for unknown systems, as CCM can produce conclusions that may be no more helpful in understanding population dynamics in ecological systems than observational field studies analyzed using more traditional statistical methods.

Keywords and phrases: causal analysis, convergent cross mapping, groundfish population dynamics

5.1 Introduction

Distinguishing between causal effects and correlations in ecological research continues to be a challenge, particularly for examination of relationships between climate fluctuations and population changes in long-term studies. Causal inference attempts to distinguish causal relationships from statistical associations. Two main versions of causal analysis exist The most predominant version and the one with a statistical formulation is based in the probabilities of causes and counterfactuals, and was popularized by Rubin (1974) and Robins (1986). It is also most heavily used in biostatistics. Establishment of causality in this framework requires observation of both potential outcomes, the causal outcome and the counterfactual outcome, at once with all other factors held constant (i.e. comparison of different states of the same world). The impossibility of actually observing both states simultaneously is known as the 'fundamental problem of causal inference' (Holland, 1986). For causal modeling, this dilemma is typically treated as a missing data problem, where blocking in experimental studies and matching in observational studies are methods used to account for it. However even in randomized controlled trials, which are considered the gold standard for causal analyses, these causal analysis methods requires untestable, highly restrictive identifying assumptions that may not realistically hold, often leading to sub-optimal analyses (Blalock, 2018; Rubin, 2008).

The second version of causality was primarily introduced by Wiener (1956), Sims (1972), and Granger (1969) for use in econometric time series. Often referred to as "predictive causality", it is based on hypothesis testing derived from correlations between time series of observed outcomes. Granger causality was proposed by Granger (1969) as a method for testing causality by determining if past observations of one time series could be used to accurately predict future observations in a different time series. For two stationary time series X_t, Y_t with zero means, the general causal model defined by Granger (1969) is

$$\begin{aligned} X_t &= \sum_{j=1}^{m} a_j X_{t-j} + \sum_{j=1}^{m} b_j Y_{t-j} + \epsilon_t \\ Y_t &= \sum_{j=1}^{m} c_j X_{t-j} + \sum_{j=1}^{m} d_j Y_{t-j} + \eta_t \ , \end{aligned} \tag{5.1}$$

where ϵ_t and η_t are uncorrelated white noise. Causality in this framework implies Y_t causes X_t if $b_j \neq 0$, and X_t causes Y_t if $c_j \neq 0$. Intuitively, Y_t is said to cause X_t if past observations Y_{t-j} contain information that improves prediction of X_t more than past observations X_{t-j} do alone. If both $b_j \neq 0$ and $c_j \neq 0$, a feedback relationship (bidirectional causality) where X_t is causing Y_t and Y_t is also causing X_t is possible, though synchrony, where X_t and Y_t are being forced by an unmeasured external mechanism, may also be suspected. As seen in (5.1), causality in this setting is only suitable to test for linear causal relationships; however extensions for nonlinear systems have been developed (Bell et al., 1996; Diks and Panchenko, 2006; Hiemstra and Jones, 1994; Marinazzo et al., 2008).

Despite the limitations of causal methods and debates surrounding their role and interpretation in science (reviewed in Gelman (2011)), the promise of being able to detect and measure causal dynamics in populations is attractive to ecologists, particularly those with the need to understand how changing global climate affect ecosystem dynamics. Granger causality and its extensions are the most often used causal analysis variant in large-scale biological and climate systems (Arjas and Eerola, 1993; Elsner, 2007; Wang et al., 2004). A newer causality detection method related to Granger causality was developed by Sugihara et al. (2012) and coined convergent cross mapping (CCM). Sugihara et al. (2012) asserted that Granger causality had a considerable weakness if applied to nonlinear dynamic systems that needed to be addressed if the approach was to be useful for exploring biological systems' dynamics. Granger causality requires separability, i.e. if Y causes X, information about Y should not be contained in the time series of X, which is typically not satisfied in general dynamic systems. Sugihara et al. (2012) claims CCM is able to distinguish causal interactions from effects of shared driving variables, known as the Moran effect. Several ecological applications of CCM have been published since the original paper, including applications in atmospheric science (van Nes et al., 2015), climate effects on marine fish populations (Deyle et al., 2013), and nutrient and plankton interactions in deep lakes (Frossard et al., 2018).

Modeling animal responses to climate variability using causal analysis techniques such as CCM has the potential to improve understanding of intricate environmental processes involved

in ecosystem dynamics. As a major intersection for several fluctuating climate systems (Anderson and Piatt, 1999; Bakun, 1999), a diverse and complex marine system (Livingston, 1993), and a significant source of fishing income and food for the United States and Japan (National Marine Fisheries Service, 2014), the North Pacific marine system is an ideal environment to study the effects of a changing climate on an ecosystem with nonlinear, dynamic interactions, known anthropogenic pressures, and a crucial role in economics and the global food supply. The North Pacific marine system is influenced by multiple interacting climate modes, including the Pacific Decadal Oscillation (PDO), El Niño-Southern Oscillation (ENSO), and North Pacific Gyre Oscillation (NPGO) (Di Lorenzo et al., 2013). Present research on the influences of environment on Alaskan groundfish, stocks of which support major commercial fisheries for the area, is particularly limited due to the difficulty of accurately studying and managing deepwater fishes. Determining potential causal effects of environmental fluctuations on adult groundfish populations while accounting for their nonlinear, dynamic system can contribute significantly to the knowledge of groundfish population dynamics and behavior. Identifying environmental variables causing changes in groundfish populations would be useful in focusing the collection of such variables during groundfish surveys and predicting commercial catch for management purposes using forecasts of such variables from climate models.

Here I applied CCM to time series of three deepwater groundfish populations and related environmental variables in the Northern Pacific Ocean. I focused on the catches of sablefish, Pacific cod, and Pacific halibut along the coast of Alaska, as they have interacting trophic relationships. All three species share a common prey, walleye pollock (Hollowed et al., 2000; Livingston, 1993). Additionally, Pacific cod and sablefish serve as an important prey species for Pacific halibut, and sablefish opportunistically prey on Pacific cod (Best and St-Pierre, 1986; Moukhametov et al., 2008; Yang et al., 2006). No known trophic models or networks of community dynamics have been established for these commercially important fishes, thus it is of interest to examine potential causal relationships between groundfish catches and their environment as well as community dynamics among the three species. Applicability and reliability of CCM to systems with time delays and nonlinear influences are not known (BozorgMagham et al., 2015), along with CCM's known potential to produce false positives in simulations (Clark

et al., 2015); thus, along with interpretation of results, potential issues with the CCM application and similar causal analysis methods to ecological data were discussed.

5.2 Methods

5.2.1 Convergent cross mapping and the multispatial extension

CCM relies on the construction of shadow manifolds, say $\mathbf{M}_X$ and $\mathbf{M}_Y$, built from lagged coordinates of time series $X = \{x_t\}_{t=1}^{L}$ and $Y = \{y_t\}_{t=1}^{L}$ respectively with length $L \in \mathbb{N}$, which according to Takens' theorem act as proxies for underlying dynamic processes (Takens, 1981). More specifically, the attractor manifold $\mathbf{M}_X = \{\mathbf{x}_t\}_{t=1+(E-1)\tau}^{L}$ is a set of E-dimensional vectors where $\mathbf{x}_t = [x_t, x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-(E-1)\tau}]^T$ for $t = 1 + (E-1)\tau$ to $t = L$, E is the embedding dimension and τ is the time lag. Under CCM, a cross-mapped estimate of y_t is formed by locating the time-corresponding vector on $\mathbf{M}_X$ and determining its $E+1$ nearest-neighbor points in $\mathbf{x}_t$. The time indices of these nearest neighbor points ordered from nearest to farthest are $t_1, \ldots, t_{E+1}$ and are then used to identify neighbors in y_t, by which $\widehat{y}_t$ is obtained using a locally weighted mean,

$$\widehat{y}_t|\mathbf{M}_X = \sum_{i=1}^{E+1} w_i y_{t_i} \ , \tag{5.2}$$

which is known as simplex projection (Sugihara and May, 1990). Here, w_i are distance-based weights,

$$w_i = \frac{k(u_i)}{\sum_{i=1}^{E+1} k(u_i)}, \qquad \text{where } u_i = \frac{\|\mathbf{x}_t - \mathbf{x}_{t_i}\|}{\|\mathbf{x}_t - \mathbf{x}_{t_1}\|} \tag{5.3}$$

and $k(u_i)$ represents the Gaussian kernel and $\|\cdot\|$ is the Euclidean norm on the E-dimensional real space. Equation (5.2) is a linear smoother also known as the Nadaraya-Watson estimator which is based on the method of least squares. As L increases, $\widehat{y}_t|\mathbf{M}_X$ converges to $Y(t)$. The efficacy of cross mapping is the Pearson's correlation coefficient ρ between $\{\widehat{y}_t|\mathbf{M}_X\}_{t=1+(E-1)\tau}^{L}$ and $\{y_t\}_{t=1+(E-1)\tau}^{L}$. A cross-mapped estimate of x_t and its cross mapping efficacy are defined analogously. If Y causes X (unidirectional causality), Y can be estimated

from X, but the reverse is not possible. In CCM, this is represented as a high values of ρ between $\{\widehat{y}_t|\mathbf{M}_X\}_{t=1+(E-1)\tau}^{L}$ and $\{y_t\}_{t=1+(E-1)\tau}^{L}$ compared to ρ between $\{\widehat{x}_t|\mathbf{M}_Y\}_{t=1+(E-1)\tau}^{L}$ and $\{x_t\}_{t=1+(E-1)\tau}^{L}$.

Construction of shadow manifolds by simplex projection as described in Sugihara et al. (2012) typically requires one long time series, which is not common in ecological data. Hsieh et al. (2008) was able to show that simplex projection could be successfully performed using several short time series considered equivalent to create one longer composite time series (dewdrop regression). Clark et al. (2015) adopted this technique by using spatial replicates as versions of multiple short time series to create the single long time series. Incorporation of spatially-replicated short time series was achieved by sampling from all spatial replicates assumed to come from the same dynamical system, then a weighted average of observations with similar historical dynamics from across all samples is used to perform simplex projection. Bootstrapping of the estimation procedure is used to average multiple combinations of spatial data (Clark et al., 2015).

5.2.2 Application of CCM to Alaskan groundfish populations

The MESA data was prepared as described in Section 4.2. I tested for causal relations between catches of each groundfish (Pacific halibut, sablefish or Pacific cod) and the following environmental variables: surface temperature, bottom temperature, chlorophyll at 75 and 150 m, and zooplankton biomass. I then tested for a causal relationship between the catches of Pacific halibut, and sablefish, Pacific halibut and Pacific cod, and sablefish and Pacific cod. The spatial version of CCM assumes some level of homogeneity across spatial replicates, which is unlikely to hold for my data. I therefore reran CCM for each of the 6 management areas defined by the Alaska Fisheries Science Center and for each of 13 regions defined by hierarchical clustering using complete linkage to compare results of CCM at different spatial scales. Regions created through clustering were defined with a distance threshold of 400 km, as effects of environmental and physical covariates and metapopulation dynamics are most influential in the population dynamics of deep-sea marine fish at regional scales of thousands of kilometers (Levin et al., 2001; Williams et al., 2010).

To appropriately test for causal relationships using CCM, several prior steps were taken before testing for significant causal relationships between processes A and B. First, all missing values and zeros were removed from the data, as zeros do not provide any information in time series. Next, stations with incomplete time series or inconsistent sampling schedules (e.g. switching from annual surveys to every two years during the study) were removed, since CCM is only suitable for time series with equal sampling intervals fixed throughout the series (Chang et al., 2017). Incomplete time series could be imputed, however imputation is likely to create artifacts that may bias any potential causal relationships. An optimal embedding dimension E is then optimized (see Clark et al. (2015)), with the caveat that τ, the length of time steps used for lagged components in the attractor space, was also optimized. For the analysis done on the whole data together, E being an integer between 2 and 6 (i.e. $E \in [2\ ..\ 6]$) and $\tau \in [1\ ..\ 5]$ were tested for each pair of variables. Due to the decreasing size of data included in analyses done for each area or clustered region, $E \in [2\ ..\ 5]$ and $\tau \in [1\ ..\ 4]$ were tested for the area analyses, whereas only $E \in [2\ ..\ 4]$ and $\tau \in [1\ ..\ 3]$ were tested for analyses on each of the small regions defined by hierarchical clustering. Using the optimal E and tau, I then tested for nonlinearity, since CCM is mainly suited for nonlinear systems. Only once significant nonlinearity was established was CCM then used to calculate the ability of process A to predict the dynamics of process B, and vice versa, using Pearson's correlation coefficient ρ. Bootstrapping sampled spatial replicates with replacement was performed to allow for information from multiple spatial locations to be incorporated into the CCM procedure. For my analyses, the number of bootstrapping iterations was set at 500 for CCM done by region and area, whereas 100 iterations were done when performing CCM on the full data. Last, significance of the causal relationship (i.e. convergence) was ascertained if both $\rho > 0$ was significantly monotonically increasing determined by Kendall's τ test and $\rho_{L_{max}} > \rho_{L_{min}}$ was significant by Fisher's $\Delta\rho\ Z$ test for at least 95% of the bootstrapped iterations, where L_{max} and L_{min} are the minimum and maximum library sizes, respectively, used in the analysis. The implementation of the multispatial CCM algorithm was carried out using the package **multispatialCCM** in R (Clark et al., 2015). Spearman's rank correlation coefficient was also calculated between process A and process B for each region, area, and for the full data set.

5.3 Results

Examination of potential causal relationships between Pacific halibut CPUE and environmental variables revealed some signals that were immediately apparent. For regions or areas with sufficient complete time series, unidirectional forcing of zooplankton on halibut CPUE was consistent for most regions (Figure 5.1). Chlorophyll at 75 m lagged five years causally forced halibut CPUE in the West Yakutat, however significant cross-map correlations for both chlorophyll forcing halibut and halibut forcing chlorophyll were observed in the nearby region 10. Likewise, five-year lagged bottom water temperature had significant cross-map skills for both causal directions across regions and areas. CCM cross-map skills for lagged SST and halibut catch, as well as lagged chlorophyll at 150 m and halibut CPUE, were inconsistent across areas and regions. Cross-map correlations for sablefish versus Pacific halibut CPUEs were high and significant for both causal directions, which was also true for Pacific cod and halibut CPUEs, except at region 11. Spearman's correlation coefficients between halibut CPUE and five-year lagged environmental covariates were low for all cases (Figure 5.1).

CCM analysis of sablefish and lagged environmental variables generally reflected similar patterns as those of Pacific halibut. Cross-map correlations with sablefish were inconsistent across regions and areas for five-year lagged SST and chlorophyll at 150 m, with higher significant Pearson's ρ values for environment forcing CPUE for some regions, but in other regions higher ρ for CPUE forcing environment (Figure 5.2). Two regions observed unidirectional forcing for chlorophyll at 150 m on sablefish catch, whereas the results in region 7 indicated sablefish causing chlorophyll at 150 m unilaterally. Zooplankton biomass lagged five years unlaterally forced sablefish CPUE in regions 6 and 10, whereas sablefish catch was unidirectionally forced by lagged bottom water temperature at regions 4 and 7. In cases where CCM was done for all spatial locations together, no clear forcing direction was apparent. Spearman's correlation coefficients between sablefish CPUE and five-year lagged environmental covariates were low for all cases, whereas moderate correlation was observed between cod and sablefish CPUEs for the West Yakutat and Central Gulf of Alaska management areas and regions 3, 6, and 10 (Figure 5.2).

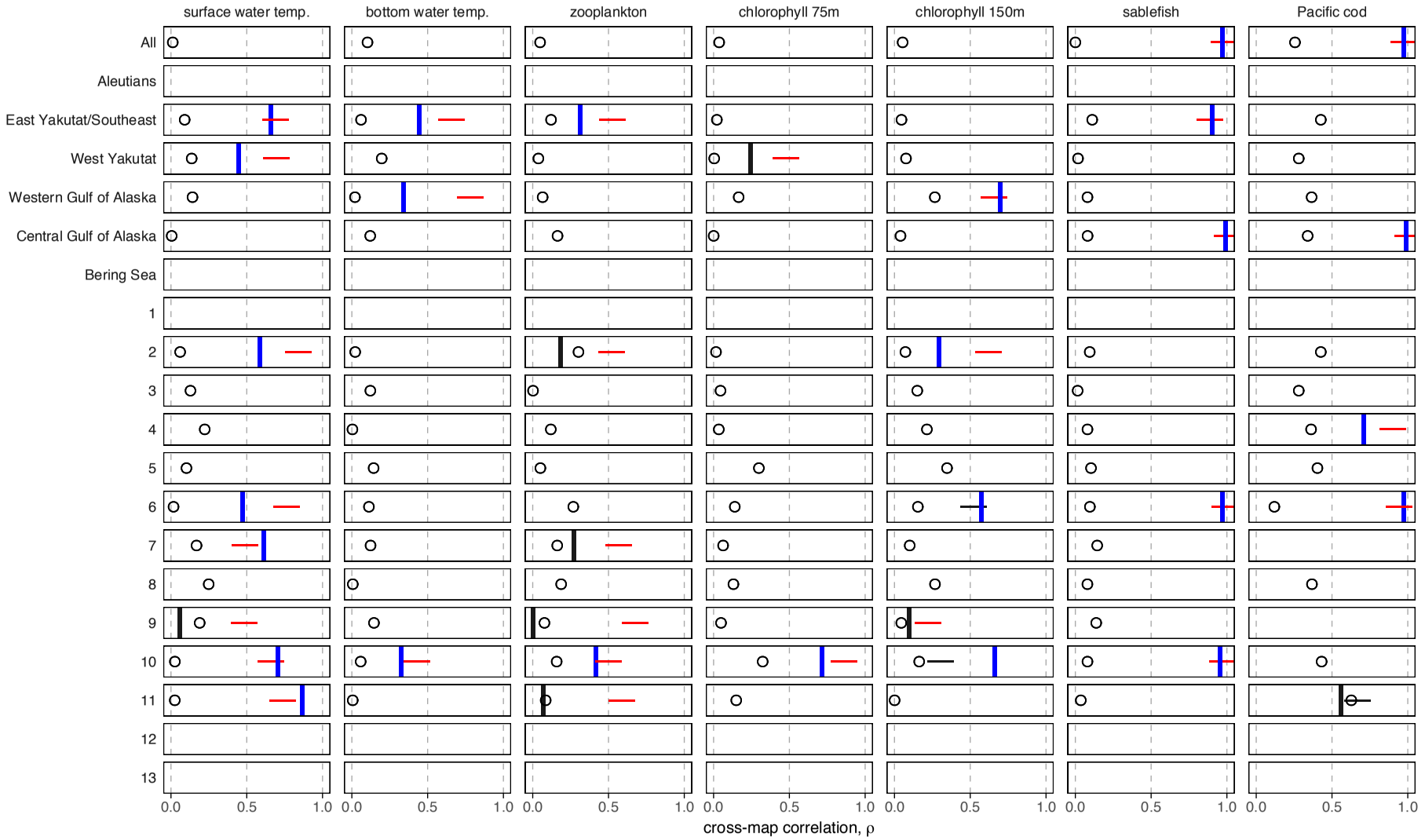


Figure 5.1: Maximum correlation of cross-mapped versus observed values as a function of time series length for comparisons between environment, cod, and sablefish (processes A, as columns) and Pacific halibut CPUE (process B). Significant forcing of A on B is represented as red horizontal bars, while significant forcing of B on A are blue vertical bars. Non-significant relationships are indicated as dark grey horizontal bars (A→B) and vertical bars (B→A). Spearman's rank correlation coefficient ρ is represented as open circles. Empty plots indicate that there were insufficient observations from a complete time series for the multispatial CCM algorithm to begin. Plots with only Spearman's ρ indicate that the predictive ability of one or both processes did not significantly decrease with increasing time distance, so CCM was not performed.

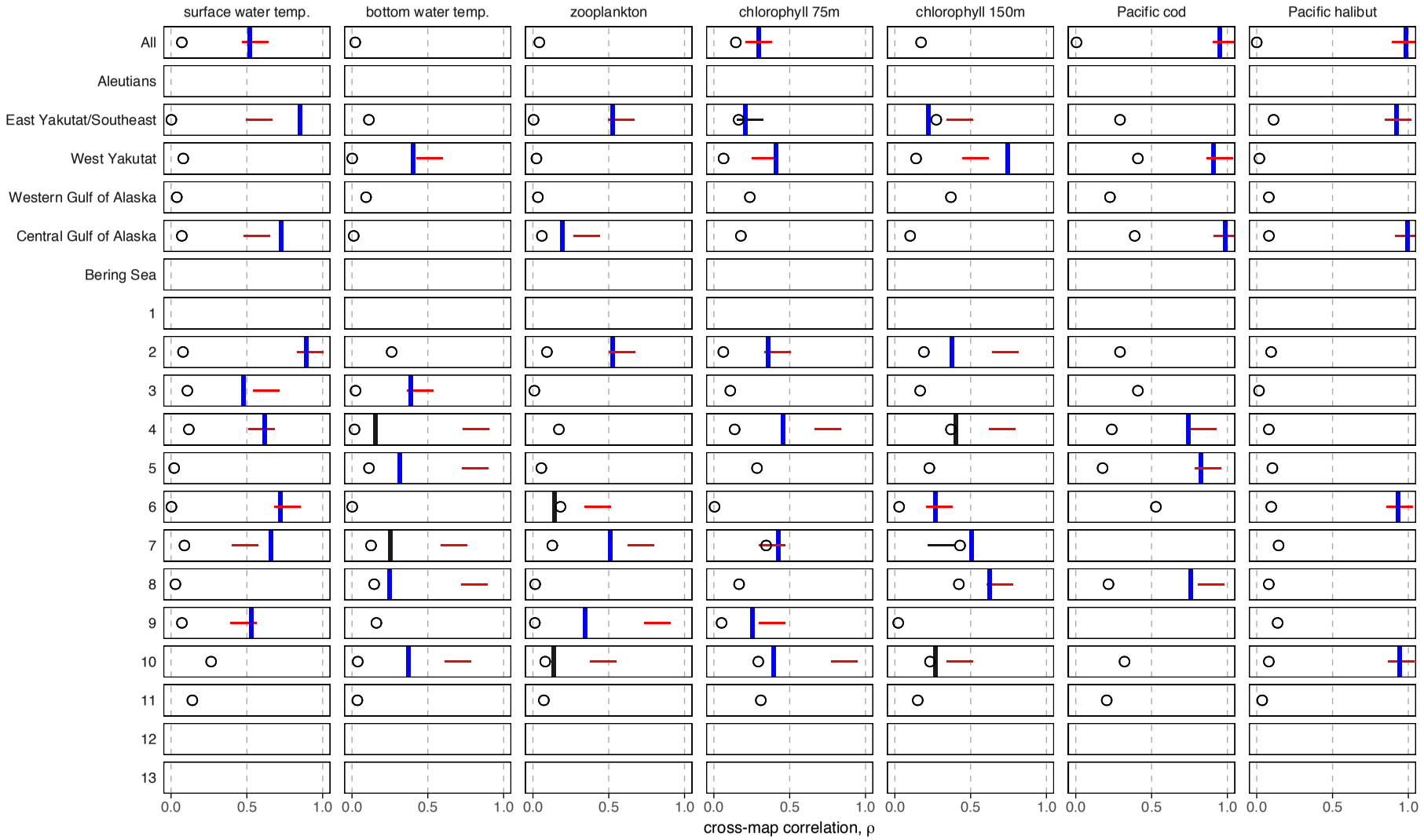


Figure 5.2: Maximum correlation of cross-mapped versus observed values as a function of time series length for comparisons between environment, cod, and halibut (processes A, as columns) and sablefish CPUE (process B). Significant forcing of A on B is represented as red horizontal bars, while significant forcing of B on A are blue vertical bars. Non-significant relationships are indicated as dark grey horizontal bars (A→B) and vertical bars (B→A). Spearman's rank correlation coefficient ρ is represented as open circles. Empty plots indicate that there were insufficient observations from a complete time series for the multispatial CCM algorithm to begin. Plots with only Spearman's ρ indicate that the predictive ability of one or both processes did not significantly decrease with increasing time distance, so CCM was not performed.

For causal dynamics of Pacific cod CPUE and bottom water temperature, significant unidirectional forcing of temperature on cod catch was consistently detected across most regions and areas (Figure 5.3). Significant nonlinear dynamics for lagged zooplankton and chlorophyll at 75 m when considering cod CPUE were not detected in most areas and regions, so CCM was not conducted. For the one region where nonlinear dynamics were significant for zooplankton (region 6), unilateral forcing of zooplankton on cod CPUE was suggested. Moderate Spearman's correlation coefficient values between cod CPUE and both depths of chlorophyll were observed for the West Yakutat and Western Gulf of Alaska management areas and regions 3, 4, 5, 8, 10, and 11, but Spearman's ρ values between CPUE and both water temperatures along with plankton were low (Figure 5.3).

5.4 Discussion

Regions determined by hierarchical clustering approximately followed larger management areas defined by the MESA survey (Figure 5.4) (Sigler and Lunsford, 2009). Cross-mapping correlations were often high and significant in both causal directions for areas, but these tended to separate and become indicators of unidirectional forcing when broken down into the smaller spatial units of regions. A main issue highlighted here that is likely to hamper the use of CCM in large-scale ecological data is its inability to adequately accommodate spatially heterogeneous dynamics. As illustrated with my CCM analysis on all locations across the Gulf of Alaska simultaneously, potential signals within time series may be swamped by high noise of heterogeneous time series. Tests for spatial homogeneity based on distances should be considered before applying the multispatial CCM algorithm to data at large spatial scales (Anderson, 2006). CCM on smaller homogeneous clusters of data may be required to obtain accurate causal inference (Wang et al., 2018).

Another concern with the CCM approach is the interpretation of more ambiguous cross-map correlation results seems to vary depending on the application. Several possible variations of patterns between cross-mapping correlation ρ and library size L for my data are given in Figure 5.5. Unidirectional causality that agrees with prior knowledge of the system is relatively straightforward - A causes B has a high and significant ρ compared to B causing A with a low ρ

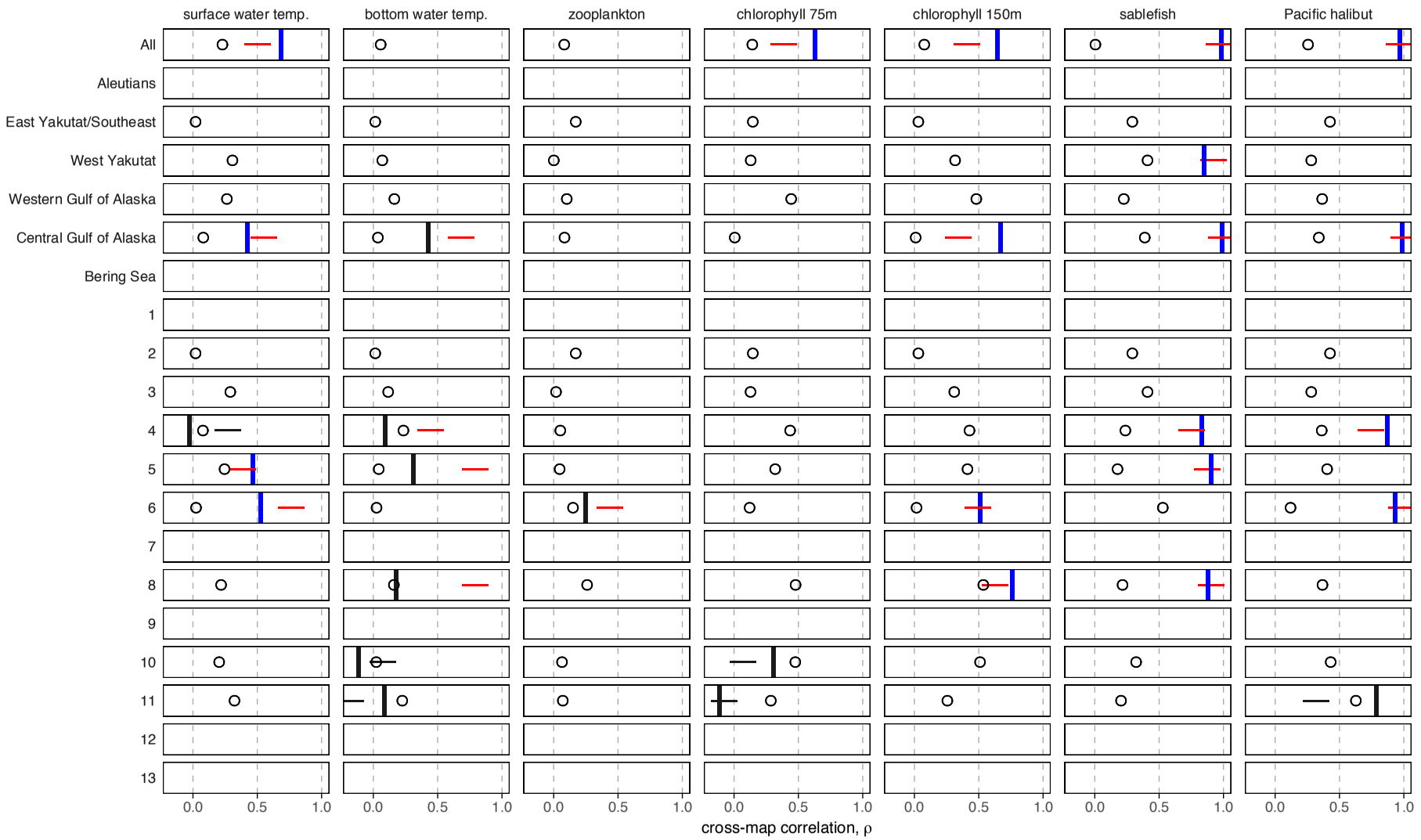


Figure 5.3: Maximum correlation of cross-mapped versus observed values as a function of time series length for comparisons between environment, halibut, and sablefish (processes A, as columns) and Pacific cod CPUE (process B). Significant forcing of A on B is represented as red horizontal bars, while significant forcing of B on A are blue vertical bars. Non-significant relationships are indicated as dark grey horizontal bars (A→B) and vertical bars (B→A). Spearman's rank correlation coefficient ρ is represented as open circles. Empty plots indicate that there were insufficient observations from a complete time series for the multispatial CCM algorithm to begin. Plots with only Spearman's ρ indicate that the predictive ability of one or both processes did not significantly decrease with increasing time distance, so CCM was not performed.

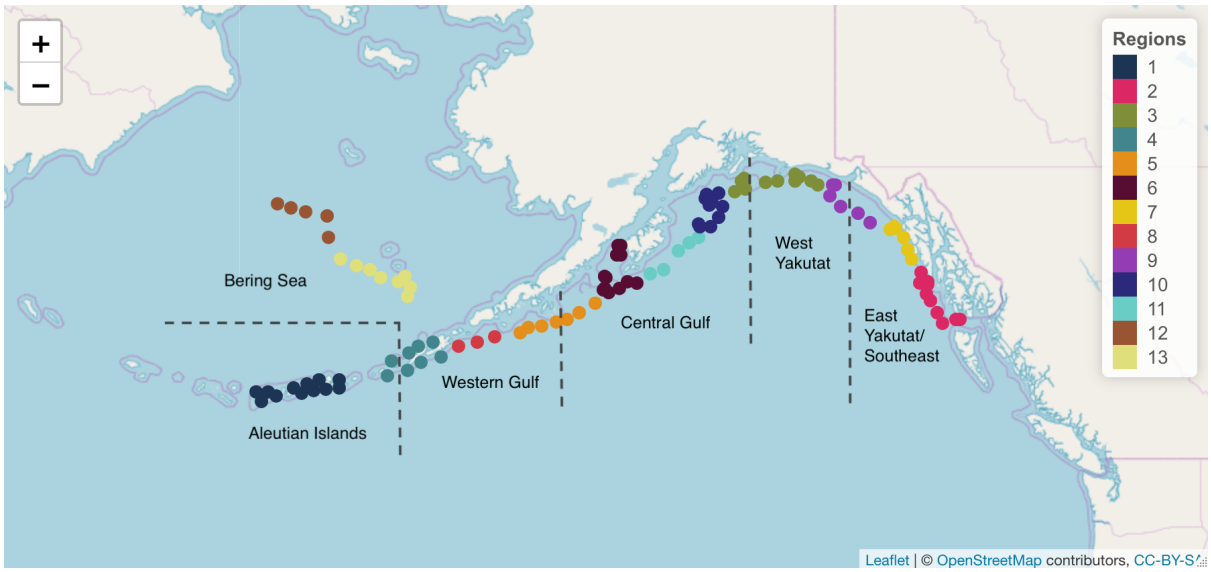


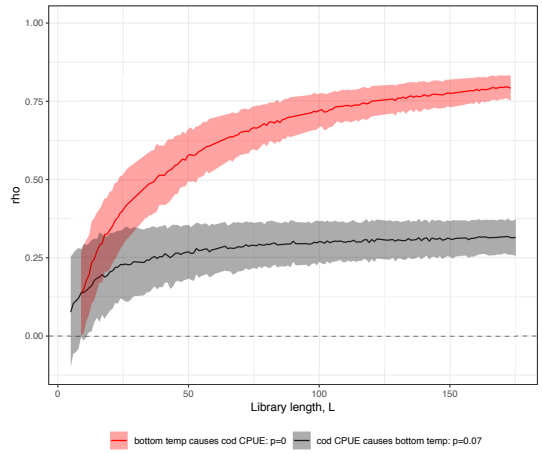
Figure 5.4: Regions of MESA stations defined by hierarchical clustering with MESA management areas labeled and borders indicated with dashed lines. Regions were determined using a 400 km distance threshold. Stations are colored by region.

that is not significant (Figures 5.5a and 5.5b). Unidirectional causality that does not agree with prior knowledge, such as the forcing of sablefish CPUE on SST in Figure 5.5c, is interpreted similarly; however, conclusions to be drawn from this case are unclear. The question remains if such results should be ignored/discarded as uninteresting as suggested by Sugihara et al. (2017), or viewed as legitimate signals that CCM behaves inconsistently in its ability to detect causal mechanisms (Baskerville and Cobey, 2017; McCracken and Weigel, 2014). Cases of synchrony, when forcing of one process is so dominant that response dynamics are similar to dynamics of the causal force, are also not well-defined. In Sugihara et al. (2012) and Clark et al. (2015), a rapid rise in cross-map skill ρ with L such as that of Figure 5.5e is likely an indication of synchrony, and any inference of a causal link is spurious. Wang et al. (2018) noted similar behavior as strong coupling but not synchrony, since optimal cross-map lags τ for their cases are all negative. This situation may also be the case for Figure 5.5e; however, analysis of cross-map lags as outlined in Ye et al. (2015) would be required to confirm strong coupling for such a case. It should be noted that discerning optimal cross-map lags is wholly different from identifying the lag τ for optimum cross-map correlation ρ between time series (van Nes et al., 2015). Another version of results with several possible interpretations is shown in Figure 5.5d, where both causal directions are significant and which could be bidirectional coupling despite large

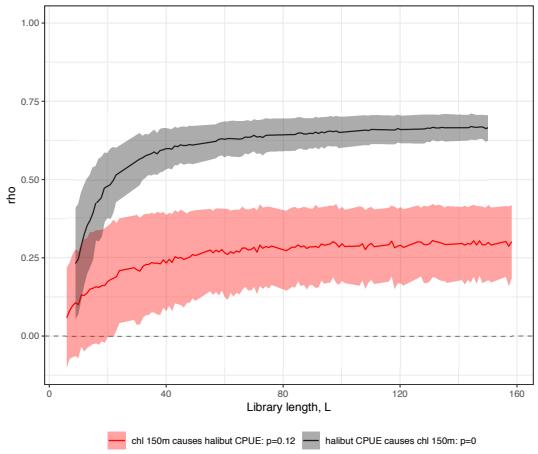
differences in cross-map correlation ρ between the two causal directions. In a similar result in Mønster et al. (2017), however, the direction with the largest ρ is believed to be the true driving factor and synchronization is increasing in the system over time. Equations for the system are already known in the case presented in Mønster et al. (2017), making inference simpler than observational data applications. Last, the version of CCM results in Figure 5.6 has not been seen or addressed in any CCM publications to my knowledge. In the case of Figure 5.6a, Spearman's rank correlation between the original time series observations for CPUE and five year lagged chlorophyll at 150m was moderate ($\rho = 0.422$), which may indicate the presence of the Moran effect.

Sugihara et al. (2012) claimed that CCM can accurately distinguish between true causal effects and correlation, however this claim may be untrue in systems with weak to moderate coupling (Mønster et al., 2017). The two time series in Figure 5.6b were uncorrelated ($\rho = 0.028$), so weak bidirectional forcing may be present in this system. In the case of Figure 5.6c, Spearman's correlation between the two time series was also low ($\rho = 0.026$), however bidirectional forcing does not fit with common knowledge of this system: it is farcical to conclude that Pacific halibut force SST at such spatial scales. That leaves synchrony, although the cross-map Pearson's correlation does not rise sharply at low values of L as would be expected of synchrony cases. From these select exemplar result of CCM application to groundfish system dynamics, along with the varying conclusions derived from other applications of CCM, inference from anything other than clear results such as those in Figures 5.5a and 5.5b should be viewed with considerable skepticism.

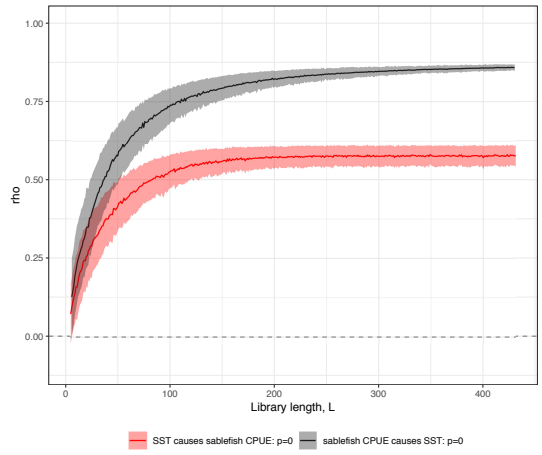
Legitimacy of some comparisons made with CCM that produce illogical results has also been argued. Sugihara et al. (2017) argued that directionality of causation is not a pertinent question when it is counterintuitive to strongly established directions of causality (e.g. flu cannot "cause" humidity). One could argue that this sort of behavior from CCM, while obvious to discard when prior knowledge or intuition about the system is clear, could easily apply to systems where minimal to no prior knowledge exists. This situation would lead to CCM identifying a forcing that is illogical for the system and be a false positive of causal dynamics present (Baskerville and Cobey, 2017; McCracken and Weigel, 2014). This would lead to spurious



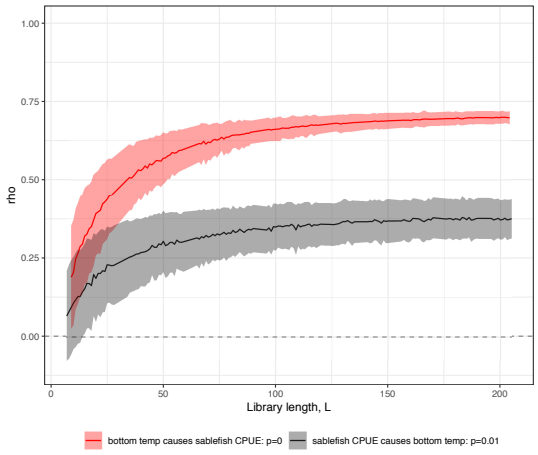
(a) Unidirectional forcing of process A (red) on process B (black).



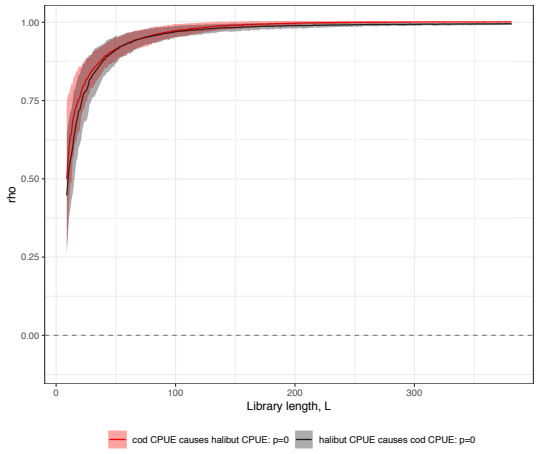
(b) Unidirectional forcing of process B (black) on process A (red).



(c) Illogical forcing of process B (black) on process A (red).

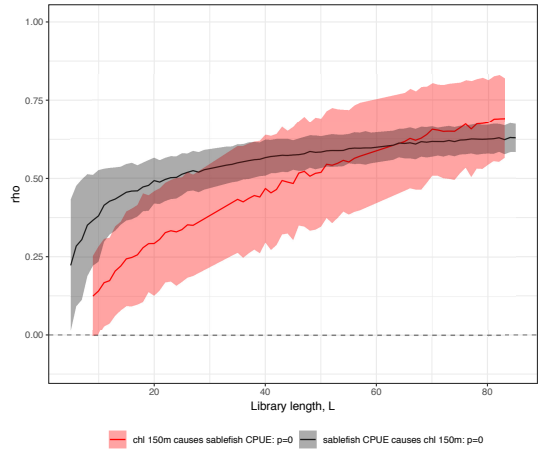


(d) Bidirectional forcing or stronger forcing of A on B and becoming synchronized.

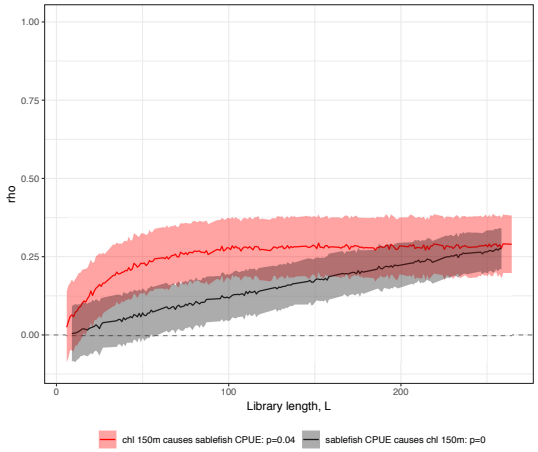


(e) Synchrony or strong bidirectional coupling.

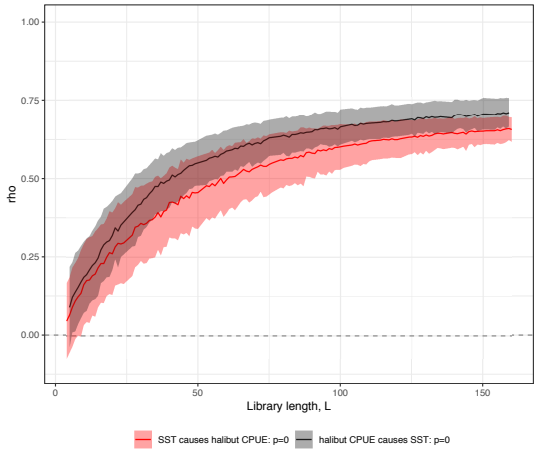
Figure 5.5: Variations of possible CCM ρ values changing with library size L.



(a) Moran effect, synchrony, or bidirectional coupling in region 8.



(b) Bidirectional coupling or synchrony in region 6.



(c) Synchrony, bidirectional coupling, or Moran effect.

Figure 5.6: Variations of possible CCM ρ values changing with library size L with overlap.

conclusions that may persist if CCM analysis is accepted as a strong analytical technique for nonlinear dynamics.

Granger (2004) cautioned that applications outside of econometrics were likely to conclude unlikely causal relationships and the usefulness in empirical areas has yet to emerge. It should be noted that Granger causality and its variants have come under scrutiny as not necessarily true causality in the strictest and most intuitive sense. They are instead typically considered causal discover techniques used for exploratory causal analysis. Though success of CCM in a few well-behaved and well-studied dynamical systems has precipitated many applications to observational data from several nonlinear dynamical systems, these studies are highlighting

concerns about whether CCM can accurately detect causal relationships, even after all requirements for CCM have been met (Baskerville and Cobey, 2017; McCracken and Weigel, 2014; Mønster et al., 2017). The use of CCM for any policy or management decisions is therefore strongly cautioned against given the number of open questions about the precise conditions for which CCM are suitable (Mønster et al., 2017). As mentioned above, many recept applications of CCM have been applied to well behaved dynamics and mostly served as confirmatory analyses for these systems(Liu et al., 2019; van Nes et al., 2015). Until further simulations and empirical studies can clearly identify when and how weaknesses in CCM occur and how to minimize or understand them, this method should only be considered as exploratory analysis of unknown systems where a starting point is needed for efficient experimental designs and not necessarily evidence of causal dynamics.

5.5 Proposed Future Extensions for CCM

The construction of shadow manifolds in Sugihara et al. (2012) and Clark et al. (2015) are nonlinear difference equations, however these lagged models could theoretically be more complex and take the form of a GAM or SIM to more suitably capture nonlinearity and spatial variation in the time series. This is the first suggested improvement I propose for CCM as applied to MESA groundfish, since it has already been shown that groundfish can be appropriately modeled with GAMs (Correia and Abebe, 2017) and SIVCMs (Sun et al., 2019). For example, consider a system where a species' abundance for year $t+1$ is affected by the same species' abundance the previous year t and the species' main predator abundance in the same year $t+1$. An appropriate GAM model for this system would be

$$g(E(Y_{t+1})) = s_1(Y_t) + s_2(X_{t+1}),$$

where Y_{t+1} is the species' abundance in year $t+1$, Y_t is abundance the following year, X_t is the predator's abundance in the following year, and $s(\cdot)$ are unknown smooth functions. Sugihara et al. (2012) illustrated the successful application of CCM in a simulated five-species model using five difference equations (see (S5) in the Sugihara et al. (2012) supplementary material).

More than one X that may causally influence Y can instead be considered using a SIM similar to the incorporation of multiple environmental variables discussed previously.

I propose a second improvement to CCM: In the estimation of $X(t)$ and $Y(t)$, the weights given by Sugihara et al. (2012) in (5.3) use Gaussian kernel functions. These weights are known to have a complex bias, which may lead to difficulty with certain bandwidth choices (Gasser and Engel, 1990). Therefore, the weighting scheme for (5.2) can be improved, which provides an opportunity to test use of robust estimators in CCM.

References

Alaska Fisheries Science Center (2019a). AFSC/ABL: Longline Sablefish Survey. https://noaa-fisheries-afsc.data.socrata.com/dataset/AFSC-ABL-Longline-Sablefish-Survey/itxd-qjvg/data. Accessed: 2014-04-14.

Alaska Fisheries Science Center (2019b). AFSC/RACE/GAP: RACEBASE Database. http://www.afsc.noaa.gov/RACE/groundfish/survey_data/default.htm.

Alderdice, D. F. and Forrester, C. R. (1971). Effects of salinity, temperature, and dissolved oxygen on early development of the pacific cod (gadus macrocephalus). *Journal of the Fisheries Research Board of Canada*, 28(6):883–902.

Alimadad, A. and Salibian-Barrera, M. (2011). An outlier-robust fit for generalized additive models with applications to disease outbreak detection. *J. Amer. Statist. Assoc.*, 106(494):719–731.

Anderson, C. N. K., Hsieh, C.-h., Sandin, S. A., Hewitt, R., Hollowed, A., Beddington, J., May, R. M., and Sugihara, G. (2008). Why fishing magnifies fluctuations in fish abundance. *Nature*, 452(7189):835–839.

Anderson, M. J. (2006). Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, 62(1):245–253.

Anderson, P. J. and Piatt, J. F. (1999). Community reorganization in the Gulf of Alaska following ocean climate regime shift. *Marine Ecology Progress Series*, 189:117–123.

Andrews, A., Coale, K., Nowicki, J., Lundstrom, C., Palacz, Z., Burton, E., and Cailliet, G. (1999). Application of an ion-exchange separation technique and thermal ionization mass spectrometry to 226ra for radiometric age determination of long-lived fishes. *Canadian Journal of Fisheries and Aquatic Sciences*, 56(8):1329–1338.

Arcuti, S., Calculli, C., Pollice, A., D'Onghia, G., Maiorano, P., and Tursi, A. (2013). Spatio-temporal modelling of zero-inflated deep-sea shrimp data by Tweedie generalized additive. *Statistica*, 73(1):87.

Arjas, E. and Eerola, M. (1993). On predictive causality in longitudinal studies. *Journal of Statistical Planning and Inference*, 34(3):361–386.

Arlot, S. and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statist. Surv.*, 4:40–79.

Athanasopoulos, G., Hyndman, R. J., Song, H., and Wu, D. C. (2011). The tourism forecasting competition. *International Journal of Forecasting*, 27(3):822–844.

Augustin, N. H., Trenkel, V. M., Wood, S. N., and Lorance, P. (2013). Space-time modelling of blue ling for fisheries stock management. *Environmetrics*, 24(2):109–119.

Austin, M. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, 157(2–3):101–118.

Aydin, K., Barbeaux, S., Barnard, D., Chilton, L., Clark, B., Conners, M., Conrath, C., Dalton, M., Echave, K., Fritz, L., Furuness, M., Hanselman, D., Haynie, A., Hoff, J., Honkalehto, T., Hulson, P., Ianelli, J., Kotwicki, S., Lauth, R., Lowe, S., Lunsford, C., McGilliard, C., McKelvey, D., Nichol, D., Norcross, B., Ormseth, O., Palsson, W., Rodgveller, C., Rooper, C., Spencer, P., Spies, I., Stockhausen, W., Stram, D., TenBrink, T., Thompson, G., Tribuzio, C., Wilderbuer, T., and Williamson, N. (2015). Stock assessment and fishery evaluation report for the groundfish resources of the Bering Sea/Aleutian Islands regions. Technical report, North Pacific Fishery Management Council.

Bailey, K. M. (2000). Shifting control of recruitment of walleye pollock theragra chalcogramma after a major climatic and ecosystem change. *Marine Ecology Progress Series*, 198:215–224.

Bakar, K. S., Sahu, S. K., et al. (2015). sptimer: Spatio-temporal bayesian modelling using R. *Journal of Statistical Software*, 63(15):1–32.

Bakun, A. (1999). A dynamic scenario for simultaneous regime-scale marine population shifts in widely separated large marine ecosystems of the pacific. In Sherman, K. and Qisheng, T., editors, *Large Marine Ecosystems of the Pacific Rim: Assessment, Sustainability and Management*. Blackwell Science Ltd.

Barry, S. C. and Welsh, A. (2002). Generalized additive modelling and zero inflated count data. *Ecological Modelling*, 157(2–3):179–188.

Baskerville, E. B. and Cobey, S. (2017). Does influenza drive absolute humidity? *Proceedings of the National Academy of Sciences*, 114(12):E2270–E2271.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Battaile, B. C. and Quinn, T. J. (2004). Catch per unit effort standardization of the eastern Bering Sea walleye pollock (*Theragra chalcogramma*) fleet. *Fisheries Research*, 70(2):161–177. Models in Fisheries Research: GLMs, GAMS and GLMMs.

Baumann, H., Hinrichsen, H. H., Möllmann, C., Köster, F. W., Malzahn, A. M., and Temming, A. (2006). Recruitment variability in Baltic Sea sprat (*Sprattus sprattus*) is tightly coupled to temperature and transport patterns affecting the larval and early juvenile stages. *Canadian Journal of Fisheries and Aquatic Sciences*, 63(10):2191–2201.

Beaugrand, G., Brander, K. M., Alistair Lindley, J., Souissi, S., and Reid, P. C. (2003). Plankton effect on cod recruitment in the North Sea. *Nature*, 426:661 EP.

Bednaršek, N., Feely, R. A., Reum, J. C. P., Peterson, B., Menkel, J., Alin, S. R., and Hales, B. (2014). ¡i¿limacina helicina¡/i¿ shell dissolution as an indicator of declining habitat suitability owing to ocean acidification in the california current ecosystem. *Proceedings of the Royal Society B: Biological Sciences*, 281(1785):20140123.

Bell, D., Kay, J., and Malley, J. (1996). A non-parametric approach to non-linear causality testing. *Economics Letters*, 51(1):7–18.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300.

Benson, A. J. and Trites, A. W. (2002). Ecological effects of regime shifts in the Bering Sea and eastern North Pacific Ocean. *Fish and Fisheries*, 3(2):95–113.

Best, E. and St-Pierre, G. (1986). Pacific halibut as predator and prey. Technical Report 21, International Pacific Halibut Commission.

Bingham, F. M., Foltz, G. R., and McPhaden, M. J. (2010). Seasonal cycles of surface layer salinity in the pacific ocean. *Ocean Science*, 6(3):775–787.

Biswas, B. K., Svirezhev, Y. M., and Bala, B. K. (2005). A model to predict climate-change impact on fish catch in the world oceans. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 35(6):773–783.

Blalock, H. M. (2018). *Causal inferences in nonexperimental research*. University of North Carolina Press.

Bluhm, B. A. and Gradinger, R. (2008). Regional variability in food availability for arctic marine mammals. *Ecological Applications*, 18(sp2):S77–S96.

Bond, N. A., Cronin, M. F., Freeland, H., and Mantua, N. (2015). Causes and impacts of the 2014 warm anomaly in the NE Pacific. *Geophysical Research Letters*, 42(9):3414–3420. 2015GL063306.

Bouchard, C. and Fortier, L. (2008). Effects of polynyas on the hatching season, early growth and survival of polar cod *Boreogadus saida* in the Laptev Sea. *Marine Ecology Progress Series*, 355:247–256.

Box, G. E. and Jenkins, G. M. (1970). *Time Series Models for Forecasting and Control*. Holden-Day, San Francisco, CA.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (2013). *Time Series Analysis*, chapter Seasonal Models, pages 353–411. Wiley-Blackwell, 4th edition.

Boyer, T., Antonov, J. I., Baranova, O. K., Coleman, C., Garcia, H. E., Grodsky, A., Johnson, D. R., Locarnini, R. A., Mishonov, A. V., O'Brien, T., Paver, C., Reagan, J., Seidov, D., Smolyar, I. V., and Zweng, M. M. (2013). World Ocean Database 2013. Technical report, National Oceanographic Data Center, Ocean Climate Laboratory, NOAA.

Boyer, T., Conkright, M., and Levitus, S. (1999). Seasonal variability of dissolved oxygen, percent oxygen saturation, and apparent oxygen utilization in the atlantic and pacific oceans. *Deep Sea Research Part I: Oceanographic Research Papers*, 46(9):1593–1613.

BozorgMagham, A. E., Motesharrei, S., Penny, S. G., and Kalnay, E. (2015). Causality analysis: Identifying the leading element in a coupled dynamical system. *PLOS ONE*, 10(6):1–17.

Brander, K. M. (2007). Global fish production and climate change. *Proceedings of the National Academy of Sciences*, 104(50):19709–19714.

Bretz, F., Westfall, P., and Hothorn, T. (2016). *Multiple comparisons using R*. Chapman and Hall/CRC.

Brierley, A. S. and Kingsford, M. J. (2009). Impacts of climate change on marine organisms and ecosystems. *Current Biology*, 19(14):R602 – R614.

Brodeur, R. D., Frost, B. W., Hare, S. R., Francis, R. C., and Ingraham Jr., W. J. (1999). Interannual variations in zooplankton biomass in the gulf of alaska, and covariation with

california current zooplankton biomass. In Sherman, K. and Qisheng, T., editors, *Large Marine Ecosystems of the Pacific Rim: Assessment, Sustainability and Management*. Blackwell Science Ltd.

Brodeur, R. D. and Ware, D. M. (1992). Long-term variability in zooplankton biomass in the subarctic pacific ocean. *Fisheries Oceanography*, 1(1):32–38.

Buckland, S. T., Burnham, K. P., and Augustin, N. H. (1997). Model selection: An integral part of inference. *Biometrics*, 53(2):603–618.

Buil, M. P. and Di Lorenzo, E. (2015). Decadal changes in gulf of alaska upwelling source waters. *Geophysical Research Letters*, 42(5):1488–1495.

Byrne, R. H., Mecking, S., Feely, R. A., and Liu, X. (2010). Direct observations of basin-wide acidification of the north pacific ocean. *Geophysical Research Letters*, 37(2).

Cai, A., Tsay, R. S., and Chen, R. (2009). Variable selection in linear regression with many predictors. *Journal of Computational and Graphical Statistics*, 18(3):573–591.

Caldeira, K. and Wickett, M. E. (2003). Oceanography: anthropogenic carbon and ocean pH. *Nature*, 425(6956):365.

Carmack, E. and McLaughlin, F. (2011). Towards recognition of physical and geochemical change in subarctic and arctic seas. *Progress in Oceanography*, 90(1-4):90–104.

Chang, C.-W., Ushio, M., and Hsieh, C.-h. (2017). Empirical dynamic modeling for beginners. *Ecological Research*, 32(6):785–796.

Cheung, W. W. L., Lam, V. W. Y., Sarmiento, J. L., Kearney, K., Watson, R., Zeller, D., and Pauly, D. (2009). Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Global Change Biology*, 16(1):24–35.

Cheung, W. W. L., Sarmiento, J. L., Dunne, J., Frolicher, T. L., Lam, V. W. Y., Deng Palomares, M. L., Watson, R., and Pauly, D. (2013). Shrinking of fishes exacerbates impacts of global ocean changes on marine ecosystems. *Nature Clim. Change*, 3(3):254–258.

Chiba, S., Tadokoro, K., Sugisaki, H., and Saino, T. (2006). Effects of decadal climate change on zooplankton over the last 50 years in the western subarctic north pacific. *Global Change Biology*, 12(5):907–920.

Childers, A. R., Whitledge, T. E., and Stockwell, D. A. (2005). Seasonal and interannual variability in the distribution of nutrients and chlorophyll a across the gulf of alaska shelf: 1998–2000. *Deep Sea Research Part II: Topical Studies in Oceanography*, 52(1):193–216. U.S. GLOBEC Biological and Physical Studies of Plankton, Fish and Higher Trophic Level Production, Distribution, and Variability in the Northeast Pacific.

Ciannelli, L., Chan, K.-S., Bailey, K. M., and Stenseth, N. C. (2004). Nonadditive effects of the environment on the survival of a large marine fish population. *Ecology*, 85(12):3418–3427.

Clark, A. T., Ye, H., Isbell, F., Deyle, E. R., Cowles, J., Tilman, G. D., and Sugihara, G. (2015). Spatial convergent cross mapping to detect causal relationships from short time series. *Ecology*, 96(5):1174–1181.

Clark, W. G., Hare, S. R., Parma, A. M., Sullivan, P. J., and Trumble, R. J. (1999). Decadal changes in growth and recruitment of pacific halibut (hippoglossus stenolepis). *Canadian Journal of Fisheries and Aquatic Sciences*, 56(2):242–252.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1):155–159.

Conn, P. B., Johnson, D. S., and Boveng, P. L. (2015). On extrapolating past the range of observed data when making statistical predictions in ecology. *PLOS ONE*, 10(10):1–16.

Correia, H. and Abebe, A. (2017). Rank-based estimation for generalized additive models with an application to fisheries data. In *Joint Statistical Meetings 2017*. American Statistical Association.

Correia, H. E. (2018). Spatiotemporally explicit model averaging for forecasting of alaskan groundfish catch. *Ecology and Evolution*, 8(24):12308–12321.

Council, N. R. (2000). *Improving the Collection, Management, and Use of Marine Fisheries Data*. The National Academies Press, Washington, DC.

Council, P. F. M. (2016). Pacific coast groundfish fishery management plan. Technical report, NOAA.

Crawford, W. R., Brickley, P. J., and Thomas, A. C. (2007). Mesoscale eddies dominate surface phytoplankton in northern Gulf of Alaska. *Progress in Oceanography*, 75(2):287–303. Time Series of the Northeast Pacific.

Croux, C., Gijbels, I., and Prosdocimi, I. (2012). Robust estimation of mean and dispersion functions in extended generalized additive models. *Biometrics*, 68(1):31–44.

Crozier, L. G. and Hutchings, J. A. (2014). Plastic and evolutionary responses to climate change in fish. *Evolutionary Applications*, 7(1):68–87.

Davis, M. W. and Ottmar, M. L. (2009). Vertical distribution of juvenile Pacific cod *Gadus macrocephalus*: potential role of light, temperature, food, and age. *Aquatic Biology*, 8(1):29–37.

De Mazancourt, C., Johnson, E., and Barraclough, T. G. (2008). Biodiversity inhibits species'evolutionary responses to changing environments. *Ecology Letters*, 11(4):380–388.

Denis, V., Lejeune, J., and Robin, J. P. (2002). Spatio-temporal analysis of commercial trawler data using general additive models: patterns of Loliginid squid abundance in the north-east Atlantic. *ICES Journal of Marine Science: Journal du Conseil*, 59(3):633–648.

Desmit, X., Ruddick, K., and Lacroix, G. (2015). Salinity predicts the distribution of chlorophyll a spring peak in the southern North Sea continental waters. *Journal of Sea Research*, 103:59–74.

Deutsch, C., Emerson, S., and Thompson, L. (2005). Fingerprints of climate change in north pacific oxygen. *Geophysical Research Letters*, 32(16).

Deyle, E. R., Fogarty, M., Hsieh, C.-h., Kaufman, L., MacCall, A. D., Munch, S. B., Perretti, C. T., Ye, H., and Sugihara, G. (2013). Predicting climate effects on pacific sardine. *Proceedings of the National Academy of Sciences*, 110(16):6430–6435.

deYoung, B., Barange, M., Beaugrand, G., Harris, R., Perry, R. I., Scheffer, M., and Werner, F. (2008). Regime shifts in marine ecosystems: detection, prediction and management. *Trends in Ecology & Evolution*, 23(7):402–409.

Di Lorenzo, E. (2018). North pacific gyre oscillation. NPGO index.

Di Lorenzo, E., Combes, V., Keister, J. E., Strub, P. T., Thomas, A. C., Franks, P. J., Ohman, M. D., Furtado, J. C., Bracco, A., Bograd, S. J., Peterson, W. T., Schwing, F. B., Chiba, S., Taguchi, B., Hormazabal, S., and Parada, C. (2013). Synthesis of pacific ocean climate and ecosystem dynamics. *Oceanography*, 26(4):68–81.

Di Lorenzo, E., Schneider, N., Cobb, K. M., Franks, P. J. S., Chhak, K., Miller, A. J., McWilliams, J. C., Bograd, S. J., Arango, H., Curchitser, E., Powell, T. M., and Rivière, P. (2008). North pacific gyre oscillation links ocean climate and ecosystem change. *Geophysical Research Letters*, 35(8).

Diks, C. and Panchenko, V. (2006). A new statistic and practical guidelines for nonparametric granger causality testing. *Journal of Economic Dynamics and Control*, 30(9):1647–1669. Computing in economics and finance.

Donelson, J., Munday, P., McCormick, M., Pankhurst, N., and Pankhurst, P. (2010). Effects of elevated water temperature and food availability on the reproductive performance of a coral reef fish. *Marine Ecology Progress Series*, 401:233–243.

Doney, S. C., Mahowald, N., Lima, I., Feely, R. A., Mackenzie, F. T., Lamarque, J.-F., and Rasch, P. J. (2007). Impact of anthropogenic atmospheric nitrogen and sulfur deposition on ocean acidification and the inorganic carbon system. *Proceedings of the National Academy of Sciences*, 104(37):14580–14585.

Dorn, M., Aydin, K., Fissel, B., Jones, D., McCarthy, A., Palsson, W., and Spalinger, K. (2017). Gulf of alaska stock assessments.

Drazen, J. C., Buckley, T. W., and Hoff, G. R. (2001). The feeding habits of slope dwelling macrourid fishes in the eastern North Pacific. *Deep Sea Research Part I: Oceanographic Research Papers*, 48(3):909–935.

Drinkwater, K. F., Beaugrand, G., Kaeriyama, M., Kim, S., Ottersen, G., Perry, R. I., Pörtner, H.-O., Polovina, J. J., and Takasuka, A. (2010). On the processes linking climate to ecosystem changes. *Journal of Marine Systems*, 79(3):374–388. Impact of climate variability on marine ecosystems: A comparative approach.

Dugdale, R. C., Wilkerson, F. P., and Minas, H. J. (1995). The role of a silicate pump in driving new production. *Deep Sea Research Part I: Oceanographic Research Papers*, 42(5):697–719.

Easterling, D. R., Meehl, G. A., Parmesan, C., Changnon, S. A., Karl, T. R., and Mearns, L. O. (2000). Climate extremes: Observations, modeling, and impacts. *Science*, 289(5487):2068.

Echave, K., Rodgveller, C., and Shotwell, S. (2013). Calculation of the geographic area sizes used to create populationindices for the alaska fisheries science center longline survey. Technical Report NOAA Tech. Memo. NMFS-AFSC-253, U.S. Department of Commerce.

Echeverria, T. W. (1987). Thirty-four species of california rockfishes: Maturity and seasonality of reproduction. *Fishery Bulletin*, 85(2):229–250.

Elsner, J. B. (2007). Granger causality and atlantic hurricanes. *Tellus A: Dynamic Meteorology and Oceanography*, 59(4):476–485.

Emerson, S. (1987). Seasonal oxygen cycles and biological new production in surface waters of the subarctic Pacific Ocean. *Journal of Geophysical Research: Oceans*, 92(C6):6535–6544.

Engen, S. and Sæther, B.-E. (2005). Generalizations of the moran effect explaining spatial synchrony in population fluctuations. *The American Naturalist*, 166(5):603–612.

Fan, J., Yao, Q., and Cai, Z. (2003). Adaptive varying-coefficient linear models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 65(1):57–80.

Feely, R. A., Sabine, C. L., Lee, K., Berelson, W., Kleypas, J., Fabry, V. J., and Millero, F. J. (2004). Impact of anthropogenic co2 on the caco3 system in the oceans. *Science*, 305(5682):362–366.

Feng, J., Durant, J. M., Stige, L. C., Hessen, D. O., Hjermann, D. Ø., Zhu, L., Llope, M., and Stenseth, N. C. (2015). Contrasting correlation patterns between environmental factors and chlorophyll levels in the global ocean. *Global Biogeochemical Cycles*, 29(12):2095–2107.

Feng, L., Zou, C., and Wang, Z. (2012). Rank-based inference for the single-index model. *Statistics & Probability Letters*, 82(3):535–541.

Feng, S. and Xue, L. (2013). Variable selection for single-index varying-coefficient model. *Frontiers of Mathematics in China*, 8(3):541–565.

Fiadeiro, M. (1980). The alkalinity of the deep pacific. *Earth and Planetary Science Letters*, 49(2):499–505.

Fietzke, J., Ragazzola, F., Halfar, J., Dietze, H., Foster, L. C., Hansteen, T. H., Eisenhauer, A., and Steneck, R. S. (2015). Century-scale trends and seasonality in ph and temperature for shallow zones of the Bering Sea. *Proceedings of the National Academy of Sciences*, 112(10):2960–2965.

Fisher, J. A. D., Casini, M., Frank, K. T., Möllmann, C., Leggett, W. C., and Daskalov, G. (2014). The importance of within-system spatial variation in drivers of marine ecosystem regime shifts. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370(1659).

Forrester, C. and Alderdice, D. (1966). Effects of salinity and temperature on embryonic development of the Pacific cod (*Gadus macrocephalus*). *Journal of the Fisheries Research Board of Canada*, 23(3):319–340.

Frainer, A., Primicerio, R., Kortsch, S., Aune, M., Dolgov, A. V., Fossheim, M., and Aschan, M. M. (2017). Climate-driven changes in functional biogeography of arctic marine fish communities. *Proceedings of the National Academy of Sciences*, 114(46):12202–12207.

Francis, R. C., Hare, S. R., Hollowed, A. B., and Wooster, W. S. (1998). Effects of interdecadal climate variability on the oceanic ecosystems of the NE Pacific. *Fisheries Oceanography*, 7(1):1–21.

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200):675–701.

Frossard, V., Rimet, F., and Perga, M.-E. (2018). Causal networks reveal the dominance of bottom-up interactions in large, deep lakes. *Ecological Modelling*, 368:136–146.

Fry, C. H., Tyrrell, T., Hain, M. P., Bates, N. R., and Achterberg, E. P. (2015). Analysis of global surface ocean alkalinity to determine controlling processes. *Marine Chemistry*, 174:46–57.

Fukasawa, M., Freeland, H., Perkin, R., Watanabe, T., Uchida, H., and Nishina, A. (2004). Bottom water warming in the north pacific ocean. *Nature*, 427(6977):825–827.

Gaichas, S. K., Aydin, K. Y., and Francis, R. C. (2010). Using food web model results to inform stock assessment estimates of mortality and production for ecosystem-based fisheries management. *Canadian Journal of Fisheries and Aquatic Sciences*, 67(9):1490–1506.

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O., Zweng, M., Reagan, J., and Johnson, D. (2014a). World Ocean Atlas 2013, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. Technical report, National Oceanographic Data Center, Ocean Climate Laboratory, NOAA.

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O., Zweng, M., Reagan, J., and Johnson, D. (2014b). World Ocean Atlas 2013, Volume 4: Dissolved Inorganic Nutrients (phosphate, nitrate, silicate). Technical report, National Oceanographic Data Center, Ocean Climate Laboratory, NOAA.

Gasser, T. and Engel, J. (1990). The choice of weights in kernel regression estimation. *Biometrika*, 77(2):377.

Gelman, A. (2011). Causality and statistical learning. *American Journal of Sociology*, 117(3):955–966.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian data analysis*. Chapman & Hall/CRC, Boca Raton, FL,, 2nd edition.

Goen, J. and Erikson, L. (2017). Fishery statistics. Technical Report IPHC-2018-AM094-05, International Pacific Halibut Commission.

Granger, C. W. (2004). Time series analysis, cointegration, and applications. *American Economic Review*, 94(3):421–425.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Grantham, B. A., Chan, F., Nielsen, K. J., Fox, D. S., Barth, J. A., Huyer, A., Lubchenco, J., and Menge, B. A. (2004). Upwelling-driven nearshore hypoxia signals ecosystem and oceanographic changes in the northeast pacific. *Nature*, 429(6993):749–754.

Grebmeier, J. M. (2012). Shifting patterns of life in the pacific arctic and sub-arctic seas. *Annual Review of Marine Science*, 4(1):63–78. PMID: 22457969.

Grebmeier, J. M., Moore, S. E., Overland, J. E., Frey, K. E., and Gradinger, R. (2010). Biological response to recent pacific arctic sea ice retreats. *Eos, Transactions American Geophysical Union*, 91(18):161–162.

Grebmeier, J. M., Overland, J. E., Moore, S. E., Farley, E. V., Carmack, E. C., Cooper, L. W., Frey, K. E., Helle, J. H., McLaughlin, F. A., and McNutt, S. L. (2006). A major ecosystem shift in the northern Bering Sea. *Science*, 311(5766):1461–1464.

Guisan, A., Jr, T. C. E., and Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, 157(2–3):89–100.

Hagens, M. and Middelburg, J. J. (2016). Attributing seasonal ph variability in surface ocean waters to governing factors. *Geophysical Research Letters*, 43(24):12,528–12,537.

Hansen, B. E. (2007). Least squares model averaging. *Econometrica*, 75(4):1175–1189.

Hardle, W., Hall, P., and Ichimura, H. (1993). Optimal smoothing in single-index models. *Ann. Statist.*, 21(1):157–178.

Harrison, G. P. and Wallace, A. R. (2005). Climate sensitivity of marine energy. *Renewable Energy*, 30(12):1801 – 1817.

Harrison, P. J., Whitney, F. A., Tsuda, A., Saito, H., and Tadokoro, K. (2004). Nutrient and plankton dynamics in the ne and nw gyres of the subarctic pacific ocean. *Journal of Oceanography*, 60(1):93–117.

Harvey, C. J. (2009). Effects of temperature change on demersal fishes in the california current: a bioenergetics approach. *Canadian Journal of Fisheries and Aquatic Sciences*, 66(9):1449–1461.

Hastie, T. and Tibshirani, R. (1993). Varying-coefficient models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(4):757–796.

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized additive models*, volume 43 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, Ltd., London.

Hemer, M. A., Fan, Y., Mori, N., Semedo, A., and Wang, X. L. (2013). Projected changes in wave climate from a multi-model ensemble. *Nature Climate Change*, 3:471 EP –.

Hester, K. C., Peltzer, E. T., Kirkwood, W. J., and Brewer, P. G. (2008). Unanticipated consequences of ocean acidification: A noisier ocean at lower ph. *Geophysical research letters*, 35(19).

Hewitt, J. E., Ellis, J. I., and Thrush, S. F. (2016). Multiple stressors, nonlinear effects and the implications of climate change impacts on marine coastal ecosystems. *Global Change Biology*, 22(8):2665–2675.

Hiemstra, C. and Jones, J. D. (1994). Testing for linear and nonlinear granger causality in the stock price-volume relation. *The Journal of Finance*, 49(5):1639–1664.

Hoegh-Guldberg, O. and Bruno, J. F. (2010). The impact of climate change on the world's marine ecosystems. *Science*, 328(5985):1523–1528.

Holland, M. M. and Bitz, C. M. (2003). Polar amplification of climate change in coupled models. *Climate Dynamics*, 21(3):221–232.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960.

Hollander, M. and Wolfe, D. A. (1999). *Nonparametric statistical methods*. Wiley Series in Probability and Statistics: Texts and References Section. John Wiley & Sons, Inc., New York, second edition. A Wiley-Interscience Publication.

Hollowed, A. B., Ianelli, J. N., and Livingston, P. A. (2000). Including predation mortality in stock assessments: a case study for Gulf of Alaska walleye pollock. *ICES Journal of Marine Science*, 57(2):279–293.

Hollowed, A. B. and Wooster, W. S. (1992). Variability of winter ocean conditions and strong year classes of northeast pacific groundfish. In *ICES Mar. Sci. Symp*, volume 195, pages 433–444.

Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50(3):346–363.

Hsieh, C., Anderson, C., and Sugihara, G. (2008). Extending nonlinear analysis to short ecological time series. *The American Naturalist*, 171(1):71–80. PMID: 18171152.

Hulson, P.-J. F., Quinn, T. J., Hanselman, D. H., and Ianelli, J. N. (2013). Spatial modeling of Bering Sea walleye pollock with integrated age-structured assessment models in a changing environment. *Canadian Journal of Fisheries and Aquatic Sciences*, 70(9):1402–1416.

Hunt, Jr, G. L., Stabeno, P., Walters, G., Sinclair, E., Brodeur, R. D., Napp, J. M., and Bond, N. A. (2002). Climate change and control of the southeastern bering sea pelagic ecosystem. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(26):5821–5853. Ecology of the {SE} Bering Sea.

Hurst, T. P. (2007). Thermal effects on behavior of juvenile walleye pollock (theragra chalcogramma): implications for energetics and food web models. *Canadian Journal of Fisheries and Aquatic Sciences*, 64(3):449–457.

Hurst, T. P., Laurel, B. J., and Ciannelli, L. (2010). Ontogenetic patterns and temperature-dependent growth rates in early life stages of Pacific cod (*Gadus macrocephalus*). *Fishery Bulletin*, 108(4):382–392.

Hurst, T. P., Moss, J. H., and Miller, J. A. (2012a). Distributional patterns of 0-group Pacific cod (*Gadus macrocephalus*) in the eastern Bering Sea under variable recruitment and thermal conditions. *ICES Journal of Marine Science: Journal du Conseil*, 69(2):163–174.

Hurst, T. P., Munch, S. B., and Lavelle, K. A. (2012b). Thermal reaction norms for growth vary among cohorts of Pacific cod (*Gadus macrocephalus*). *Marine Biology*, 159(10):2173–2183.

Hutchings, J. A. (2000). Collapse and recovery of marine fishes. *Nature*, 406(6798):882–885.

Hyndman, R. J. (2017). *forecast: Forecasting functions for time series and linear models*. R package version 8.0.

Hyndman, R. J. and Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*, 26(3):1–22.

Ianelli, J. N., Hollowed, A. B., Haynie, A. C., Mueter, F. J., and Bond, N. A. (2011). Evaluating management strategies for eastern Bering Sea walleye pollock (*Theragra chalcogramma*) in a changing environment. *ICES Journal of Marine Science*, 68(6):1297–1304.

Ichimura, H. (1993). Semiparametric least squares (sls) and weighted sls estimation of single-index models. *Journal of Econometrics*, 58(1):71–120.

Johnson, K. F., Rudd, M. B., Pons, M., Akselrud, C. A., Lee, Q., Hurtado-Ferro, F., Haltuch, M. A., and Hamel, O. S. (2016). Status of the u.s. sablefish resource in 2015. Technical report, Pacific Fishery Management Council.

Kahru, M., Gille, S. T., Murtugudde, R., Strutton, P. G., Manzano-Sarabia, M., Wang, H., and Mitchell, B. G. (2010). Global correlations between winds and ocean chlorophyll. *Journal of Geophysical Research: Oceans*, 115(C12).

Kakehi, S., Ito, S.-i., and Wagawa, T. (2017). Estimating surface water mixing ratios using salinity and potential alkalinity in the kuroshio-oyashio mixed water regions. *Journal of Geophysical Research: Oceans*, 122(3):1927–1942.

Kang, Y.-S. and Ohman, M. D. (2014). Comparison of long-term trends of zooplankton from two marine ecosystems across the North Pacific: Northeastern Asian Marginal Sea and Southern California Current System. *CalCOFI Report*, 55:169–182.

Katsanevakis, S., Maravelias, C. D., Damalas, D., Karageorgis, A. P., Tsitsika, E. V., Anagnostou, C., and Papaconstantinou, C. (2009). Spatiotemporal distribution and habitat use of commercial demersal species in the eastern Mediterranean Sea. *Fisheries Oceanography*, 18(6):439–457.

Kilian, L. and Taylor, M. (2003). Why is it so difficult to beat the random walk forecast of exchange rates? *Journal of International Economics*, 60(1):85–107.

Kim, S., Kang, S., Zhang, C.-I., Seo, H., Kang, M., and Kim, J. J. (2012). Comparison of fisheries yield and oceanographic features at the southern boundaries of the western and eastern Subarctic Pacific Ocean. *ICES Journal of Marine Science*, 69(7):1141–1147.

Kline Jr, T. C. (2007). Rockfish trophic relationships in Prince William Sound, Alaska, based on natural abundances of stable isotopes. In Heifetz, J., DiCosimo, J., Gharrett, A., Love, M., O'Connell, V., and Stanley, R., editors, *Biology, Assessment, and Management of North Pacific Rockfishes*. Alaska Sea Grant College Program. AK-SG-07-01.

Kloke, J. and McKean, J. W. (2014). *Nonparametric statistical methods using R*. CRC Press.

Koenig, W. D. (2002). Global patterns of environmental synchrony and the moran effect. *Ecography*, 25(3):283–288.

Koslow, J. A., Goericke, R., Lara-Lopez, A., and Watson, W. (2011). Impact of declining intermediate-water oxygen on deepwater fishes in the california current. *Marine Ecology Progress Series*, 436:207–218.

Koutroumanidis, T., Iliadis, L., and Sylaios, G. K. (2006). Time-series modeling of fishery landings using ARIMA models and fuzzy expected intervals software. *Environmental Modelling & Software*, 21(12):1711–1721.

Laurel, B. J., Hurst, T. P., Copeman, L. A., and Davis, M. W. (2008). The role of temperature on the growth and survival of early and late hatching Pacific cod larvae (*Gadus macrocephalus*). *Journal of Plankton Research*, 30(9):1051–1060.

Laurel, B. J., Spencer, M., Iseri, P., and Copeman, L. A. (2016). Temperature-dependent growth and behavior of juvenile Arctic cod (*Boreogadus saida*) and co-occurring north Pacific gadids. *Polar Biology*, 39(6):1127–1135.

Lavergne, S., Mouquet, N., Thuiller, W., and Ronce, O. (2010). Biodiversity and climate change: Integrating evolutionary and ecological responses of species and communities. *Annual Review of Ecology, Evolution, and Systematics*, 41(1):321–350.

Ledolter, J. (2013). *Penalty-Based Variable Selection in Regression Models with Many Parameters (LASSO)*, chapter 6, pages 71–82. John Wiley & Sons, Ltd.

Lee, K., Tong, L. T., Millero, F. J., Sabine, C. L., Dickson, A. G., Goyet, C., Park, G.-H., Wanninkhof, R., Feely, R. A., and Key, R. M. (2006). Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans. *Geophysical Research Letters*, 33(19).

Leeuwis, R. H., Nash, G. W., Sandrelli, R. M., Zanuzzo, F. S., and Gamperl, A. K. (2019). The environmental tolerances and metabolic physiology of sablefish (anoplopoma fimbria). *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 231:140–148.

Leverette, T. L. and Metaxas, A. (2005). *Predicting habitat for two species of deep-water coral on the Canadian Atlantic continental shelf and slope*, pages 467–479. Springer Berlin Heidelberg, Berlin, Heidelberg.

Levin, L. A., Etter, R. J., Rex, M. A., Gooday, A. J., Smith, C. R., Pineda, J., Stuart, C. T., Hessler, R. R., and Pawson, D. (2001). Environmental influences on regional deep-sea species diversity. *Annual Review of Ecology and Systematics*, 32(1):51–93.

Levitus, S., Antonov, J. I., Boyer, T. P., and Stephens, C. (2000). Warming of the world ocean. *Science*, 287(5461):2225–2229.

Li, L., Losser, T., Yorke, C., and Piltner, R. (2014a). Fast inverse distance weighting-based spatiotemporal interpolation: A web-based application of interpolating daily fine particulate matter pm2.5 in the contiguous u.s. using parallel programming and k-d tree. *International Journal of Environmental Research and Public Health*, 11(9):9101–9141.

Li, L., Losser, T., Yorke, C., and Piltner, R. (2014b). Fast inverse distance weighting-based spatiotemporal interpolation: A web-based application of interpolating daily fine particulate matter pm2.5 in the contiguous u.s. using parallel programming and k-d tree. *International Journal of Environmental Research and Public Health*, 11(9):9101–9141.

Litzow, M. A., Mueter, F. J., and Hobday, A. J. (2014). Reassessing regime shifts in the North Pacific: incremental climate change and commercial fishing are necessary for explaining decadal-scale biological variability. *Global Change Biology*, 20(1):38–50.

Liu, H., Fogarty, M., Glaser, S., Altman, I., Hsieh, C.-h., Kaufman, L., Rosenberg, A., and Sugihara, G. (2012). Nonlinear dynamic features and co-predictability of the georges bank fish community. *Marine Ecology Progress Series*, 464:195–207.

Liu, H., Lei, M., Zhang, N., and Du, G. (2019). The causal nexus between energy consumption, carbon emissions and economic growth: New evidence from china, india and g7 countries using convergent cross mapping. *PLOS ONE*, 14(5):1–18.

Livingston, P., Low, L.-L., and Marasco, R. (1999). Eastern bering sea ecosystem trends. In Sherman, K. and Qisheng, T., editors, *Large Marine Ecosystems of the Pacific Rim: Assessment, Sustainability and Management*. Blackwell Science Ltd.

Livingston, P. A. (1993). Importance of predation by groundfish, marine mammals and birds on walleye pollock theragra chalcogramma and pacific herring clupea pallasi in the eastern bering sea. *Marine Ecology Progress Series*, 102:205–215.

Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., Zweng, M. M., Paver, C. R., Reagan, J. R., Johnson, D. R., Hamilton, M., and Seidov, D. (2013). World ocean atlas 2013, volume 1: Temperature. Technical report, National Oceanographic Data Center, Ocean Climate Laboratory, NOAA.

Loher, T. and Seitz, A. (2006). Seasonal migration and environmental conditions of Pacific halibut *Hippoglossus stenolepis*, elucidated from pop-up archival transmitting (PAT) tags. *Marine Ecology Progress Series*, 317:259–271.

Longhurst, A. R. (1976). Interactions between zooplankton and phytoplankton profiles in the eastern tropical pacific ocean. *Deep Sea Research and Oceanographic Abstracts*, 23(8):729–754.

Loubere, P. (1994). Quantitative estimation of surface ocean productivity and bottom water oxygen concentration using benthic foraminifera. *Paleoceanography*, 9(5):723–737.

Mandic, M., Todgham, A., and G Richards, J. (2008). Mechanisms and evolution of hypoxia tolerance in fish. *Proceedings. Biological sciences / The Royal Society*, 276:735–44.

Mantua, N. J. and Hare, S. R. (2002). The Pacific decadal oscillation. *Journal of Oceanography*, 58(1):35–44.

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., and Francis, R. C. (1997). A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the American Meteorological Society*, 78(6):1069–1079.

Mantua, N. J. and JISAU, University of Washington (2016). The Pacific decadal oscillation. Online: `http://research.jisao.washington.edu/pdo/`.

Marinazzo, D., Pellicoro, M., and Stramaglia, S. (2008). Kernel method for nonlinear granger causality. *Phys. Rev. Lett.*, 100:144103.

Mason, J. C., Beamish, R. J., and McFarlane, G. A. (1983). Sexual maturity, fecundity, spawning, and early life history of sablefish (anoplopoma fimbria) off the pacific coast of canada. *Canadian Journal of Fisheries and Aquatic Sciences*, 40(12):2126–2134.

Massie, T. M., Weithoff, G., Kuckländer, N., Gaedke, U., and Blasius, B. (2015). Enhanced moran effect by spatial variation in environmental autocorrelation. *Nature Communications*, 6:5993 EP.

Matsui, H. and Misumi, T. (2015). Variable selection for varying-coefficient models with the sparse regularization. *Computational Statistics*, 30(1):43–55.

McCracken, J. M. and Weigel, R. S. (2014). Convergent cross-mapping and pairwise asymmetric inference. *Phys. Rev. E*, 90:062903.

McGowan, J. A., Cayan, D. R., and Dorman, L. M. (1998). Climate-ocean variability and ecosystem response in the Northeast Pacific. *Science*, 281(5374):210–217.

Meehl, G. A., Arblaster, J. M., Fasullo, J. T., Hu, A., and Trenberth, K. E. (2011). Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Climate Change*, 1(7):360.

Meier, L., van de Geer, S., and Bühlmann, P. (2009). High-dimensional additive modeling. *Ann. Statist.*, 37(6B):3779–3821.

Melo, C. and Melo, O. (2015). *geosptdb: Spatio-Temporal Inverse Distance Weighting and Radial Basis Functions with Distance-Based Regression*. R package version 0.5-0.

Moeller, H. V., Laufkötter, C., Sweeney, E. M., and Johnson, M. D. (2019). Light-dependent grazing can drive formation and deepening of deep chlorophyll maxima. *Nature communications*, 10(1):1978–1978.

Möllmann, C. and Diekmann, R. (2012). Chapter 4 - Marine Ecosystem Regime Shifts Induced by Climate and Overfishing: A Review for the Northern Hemisphere. In Woodward, G., Jacob, U., and O'Gorman, E. J., editors, *Global Change in Multispecies Systems Part 2*, volume 47 of *Advances in Ecological Research*, pages 303–347. Academic Press.

Möllmann, C., Folke, C., Edwards, M., and Conversi, A. (2015). Marine regime shifts around the globe: theory, drivers and impacts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1659):20130260.

Monllor-Hurtado, A., Pennino, M. G., and Sanchez-Lizaso, J. (2017). Shift in tuna catches due to ocean warming. *PLoS ONE*, 12(6):e0178196.

Monnahan, C. C. and Stewart, I. J. (2018). The effect of hook spacing on longline catch rates: Implications for catch rate standardization. *Fisheries Research*, 198:150–158.

Mønster, D., Fusaroli, R., Tylén, K., Roepstorff, A., and Sherson, J. F. (2017). Causal inference from noisy time-series data — testing the convergent cross-mapping algorithm in the presence of noise and external influence. *Future Generation Computer Systems*, 73:52–62.

Moore, J. and Mace, P. (1999). Challenges and prospects for deep-sea finfish fisheries. *Fisheries*, 24(7):22–23.

Moore, J. A. (1999). Deep-sea finfish fisheries: Lessons from history. *Fisheries*, 24(7):16–21.

Moran, P. (1953). The statistical analysis of the canadian lynx cycle. *Australian Journal of Zoology*, 1(3):291–298.

Moukhametov, I., Orlov, A., and Leaman, B. (2008). Diet of pacific halibut (hippoglossus stenolepis) in the northwestern pacific ocean. Technical Report 52, International Pacific Halibut Commission.

Mourato, B. L., Hazin, F., Bigelow, K., Musyl, M., Carvalho, F., and Hazin, H. (2014). Spatio-temporal trends of sailfish, *Istiophorus platypterus* catch rates in relation to spawning ground and environmental factors in the equatorial and southwestern Atlantic Ocean. *Fisheries Oceanography*, 23(1):32–44.

Mueter, F. J., Bond, N. A., Hollowed, A. B., and Ianelli, J. N. (2011). Expected declines in recruitment of walleye pollock (Theragra chalcogramma) in the eastern Bering Sea under future climate change. *ICES Journal of Marine Science*, 68(6):1284–1296.

Mueter, F. J., Ladd, C., Palmer, M. C., and Norcross, B. L. (2006). Bottom-up and top-down controls of walleye pollock (theragra chalcogramma) on the eastern bering sea shelf. *Progress in Oceanography*, 68(2):152–183. Marine Ecosystem Structure and Dynamics.

Napp, J. M. and Hunt, G. L. (2001). Anomalous conditions in the south-eastern Bering Sea 1997: linkages among climate, weather, ocean, and biology. *Fisheries Oceanography*, 10(1):61–68.

National Data Buoy Center (2018a). Meteorological and oceanographic data collected from the national data buoy center coastal-marine automated network (c-man) and moored (weather) buoys. `https://accession.nodc.noaa.gov/NDBC-CMANWx`.

National Data Buoy Center (2018b). Meteorological and oceanographic data collected from the national data buoy center coastal-marine automated network (c-man) and moored (weather) buoys: Measurement descriptions and units. `https://www.ndbc.noaa.gov/measdes.shtml`.

National Marine Fisheries Service (2014). Fisheries of the United States, 2014. NOAA Current Fishery Statistics 2014, U.S. Department of Commerce.

National Marine Fisheries Service, NOAA (2014). The coastal & oceanic plankton ecology, production, & observation database. Online: `http://www.st.nmfs.noaa.gov/copepod/`.

Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. (1996). *Applied linear statistical models*, volume 4. Irwin Chicago.

NOAA (2015). National Data Buoy Center - Optimum Interpolation Sea Surface Temperature. Online: `https://www.ncdc.noaa.gov/oisst`. Accessed: 2015-09-18.

NOAA (2016). International comprehensive ocean-atmosphere data set. Online: `http://icoads.noaa.gov`.

NOAA Earth Systems Research Lab (2017). Multivariate ENSO index (MEI). Online: `https://www.esrl.noaa.gov/psd/enso/mei/`.

NOAA ESRL Physical Sciences Division (2019). Multivariate enso index version 2 (mei.v2). ENSO index.

Noakes, D. J. and Beamish, R. J. (2009). Synchrony of marine fish catches and climate and ocean regime shifts in the North Pacific Ocean. *Marine and Coastal Fisheries*, 1(1):155–168.

NPFMC (2017). Fishery management plan for groundfish of the Gulf of Alaska. Technical report, North Pacific Fishery Management Council.

O'Brien, R. M. (2007a). A caution regarding rules of thumb for variance inflation factors. *Quality & quantity*, 41(5):673–690.

O'Brien, T. (2007b). Copepod: The global plankton database. a review of the 2007 database contents and new quality control methodology. Technical Report NOAA Tech. Memo. NMFS-F/ST-34, U.S. Dep. Commerce.

Ohno, Y., Iwasaka, N., Kobashi, F., and Sato, Y. (2009). Mixed layer depth climatology of the north pacific based on argo observations. *Journal of Oceanography*, 65(1):1–16.

Ono, T., Midorikawa, T., Watanabe, Y. W., Tadokoro, K., and Saino, T. (2001). Temporal increases of phosphate and apparent oxygen utilization in the subsurface waters of western subarctic pacific from 1968 to 1998. *Geophysical Research Letters*, 28(17):3285–3288.

Orensanz, J. M., Ernest, B., Armstrong, D. A., Stabeno, P., and Livingston, P. (2004). Contraction of the geographic range of distribution of snow crab (chionoecetes opilio) in the eastern bering sea: an environmental ratchet? *CalCOFI Report*, 45.

Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney, S. C., Feely, R. A., Gnanadesikan, A., Gruber, N., Ishida, A., Joos, F., et al. (2005). Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature*, 437(7059):681.

Ortega-García, S., Camacho-Bareño, E., and Martínez-Rincón, R. O. (2015). Effects of environmental factors on the spatio-temporal distribution of striped marlin catch rates off Cabo San Lucas, Baja California Sur, Mexico. *Fisheries Research*, 166:47–58. Proceedings of the 5th International Billfish Symposium.

Overland, J. E. and Stabeno, P. J. (2004). Is the climate of the bering sea warming and affecting the ecosystem? *Eos, Transactions American Geophysical Union*, 85(33):309–312.

Overland, J. E., Wang, M., Walsh, J. E., and Stroeve, J. C. (2014). Future arctic climate changes: Adaptation and mitigation time scales. *Earth's Future*, 2(2):68–74.

Paci, L., Gelfand, A. E., and Holland, D. M. (2013). Spatio-temporal modeling for real-time ozone forecasting. *Spatial Statistics*, 4:79–93.

Pennoyer, S. and Balsiger, J. (1998). Groundfish total allowable catch specifications and prohibited species catch limits under the authority of the fishery management plans for the groundfish fishery of the Bering Sea and Aleutian Islands area and groundfish of the Gulf of Alaska : final supplemental environmental impact statement. Technical report, United States National Marine Fisheries Service Alaska Regional Office, Juneau, Alaska.

Phillips, A. J., Ciannelli, L., Brodeur, R. D., Pearcy, W. G., and Childers, J. (2014). Spatio-temporal associations of albacore CPUEs in the northeastern Pacific with regional SST and climate environmental variables. *ICES Journal of Marine Science: Journal du Conseil*, 71(7):1717–1727.

Polovina, J. J., Mitchum, G. T., and Evans, G. T. (1995). Decadal and basin-scale variation in mixed layer depth and the impact on biological production in the central and north pacific, 1960-88. *Deep Sea Research Part I: Oceanographic Research Papers*, 42(10):1701–1716.

Pörtner, H. O., Berdal, B., Blust, R., Brix, O., Colosimo, A., De Wachter, B., Giuliani, A., Johansen, T., Fischer, T., Knust, R., Lannig, G., Naevdal, G., Nedenes, A., Nyhammer, G., Sartoris, F. J., Serendero, I., Sirabella, P., Thorkildsen, S., and Zakhartsev, M. (2001). Climate induced temperature effects on growth performance, fecundity and recruitment in marine fish: developing a hypothesis for cause and effect relationships in Atlantic cod (*Gadus morhua*) and common eelpout (*Zoarces viviparus*). *Continental Shelf Research*, 21(18–19):1975–1997.

Pörtner, H. O. and Knust, R. (2007). Climate change affects marine fishes through the oxygen limitation of thermal tolerance. *Science*, 315(5808):95.

Praetorius, S. K., Mix, A. C., Walczak, M. H., Wolhowe, M. D., Addison, J. A., and Prahl, F. G. (2015). North pacific deglacial hypoxic events linked to abrupt ocean warming. *Nature*, 527:362 EP –.

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raftery, A. E., Gneiting, T., Balabdaoui, F., and Polakowski, M. (2005). Using bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, 133(5):1155–1174.

Ranta, E., Kaitala, V., Lindström, J., and Helle, E. (1997). The moran effect and synchrony in population dynamics. *Oikos*, 78(1):136–142.

Reed, T. E., Schindler, D. E., and Waples, R. S. (2011). Interacting effects of phenotypic plasticity and evolution on population persistence in a changing climate. *Conservation Biology*, 25(1):56–63.

Reuter, R. F. and Spencer, P. D. (2007). Characterizing aspects of rockfish (*Sebastes* spp.) assemblages in the Aleutian Islands, Alaska. In J. Heifetz, J. D., Gharrett, A., Love, M.,

O'Connell, V., and Stanley, R., editors, *Biology, Assessment, and Management of North Pacific Rockfishes*, pages 383–409. Alaska Sea Grant College Program.

Ricker, W. (1975). *Computation and Interpretation of Biological Statistics of Fish Populations*. Number 191 in Bulletin - Fisheries Research Board of Canada. Department of the Environment, Fisheries and Marine Service.

Riebesell, U., Zondervan, I., Rost, B., Tortell, P. D., Zeebe, R. E., and Morel, F. M. M. (2000). Reduced calcification of marine plankton in response to increased atmospheric co2. *Nature*, 407(6802):364–367.

Rijnsdorp, A. D., Peck, M. A., Engelhard, G. H., Möllmann, C., and Pinnegar, J. K. (2009). Resolving the effect of climate change on fish populations. *ICES Journal of Marine Science: Journal du Conseil*, 66:1570–1583.

Robins, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Mathematical Modelling*, 7(9):1393–1512.

Rodgveller, C. and Hulson, P. (2014). Assessment of the grenadier stock complex in the Gulf of Alaska, Eastern Bering Sea, and Aleutian Islands. Technical report, North Pacific Fishery Management Council.

Rodgveller, C. J., Clausen, D. M., Nagler, J. J., and Hutchinson, C. (2010). Reproductive characteristics and mortality of female giant grenadiers in the northern pacific ocean. *Marine and Coastal Fisheries*, 2(1):73–82.

Rodgveller, C. J., Lunsford, C. R., and Fujioka, J. T. (2008). Evidence of hook competition in longline surveys. *Fishery Bulletin*, 106(4):364–374.

Rodgveller, C. J., Sigler, M. F., Hanselman, D. H., and Ito, D. H. (2011). Sampling efficiency of longlines for shortraker and rougheye rockfish using observations from a manned submersible. *Marine and Coastal Fisheries*, 3(1):1–9.

Rooper, C. and Martin, M. (2009). Predicting presence and abundance of demersal fishes: a model application to shortspine thornyhead *Sebastolobus alascanus*. *Marine Ecology Progress Series*, 379:253–266.

Rosenthal, R., Cooper, H., and Hedges, L. (1994). Parametric measures of effect size. *The handbook of research synthesis*, pages 231–244.

Rouyer, T., Fromentin, J.-M., Hidalgo, M., and Stenseth, N. C. (2014). Combined effects of exploitation and temperature on fish stocks in the Northeast Atlantic. *ICES Journal of Marine Science*, 71(7):1554–1562.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688.

Rubin, D. B. (2008). For objective causal inference, design trumps analysis. *Ann. Appl. Stat.*, 2(3):808–840.

Rummer, J. L., Roshan-Moniri, M., Balfry, S. K., and Brauner, C. J. (2010). Use it or lose it? Sablefish, *Anoplopoma fimbria*, a species representing a fifth teleostean group where the βNHE associated with the red blood cell adrenergic stress response has been secondarily lost. *Journal of Experimental Biology*, 213(9):1503–1512.

Sackmann, B., Mack, L., Logsdon, M., and Perry, M. J. (2004). Seasonal and inter-annual variability of seawifs-derived chlorophyll a concentrations in waters off the washington and vancouver island coasts, 1998–2002. *Deep Sea Research Part II: Topical Studies in Oceanography*, 51(10-11):945–965.

Sadorus, L. L. (2012). The influence of environmental factors on halibut distribution as observed on the IPHC stock assessment survey: A preliminary examination. Technical report, International Pacific Halibut Commission.

Sadorus, L. L., Mantua, N. J., Essington, T., Hickey, B., and Hare, S. (2014). Distribution patterns of Pacific halibut (*Hippoglossus stenolepis*) in relation to environmental variables

along the continental shelf waters of the US West Coast and southern British Columbia. *Fisheries Oceanography*, 23(3):225–241.

Scheffer, M., van Nes, E. H., and Vergnon, R. (2018). Toward a unifying theory of biodiversity. *Proceedings of the National Academy of Sciences*, 115(4):639–641.

Schirripa, M. J. and Colbert, J. J. (2006). Interannual changes in sablefish (anoplopoma fimbria) recruitment in relation to oceanographic conditions within the california current system. *Fisheries Oceanography*, 15(1):25–36.

Seebacher, F., White, C. R., and Franklin, C. E. (2014). Physiological plasticity increases resilience of ectothermic animals to climate change. *Nature Climate Change*, 5:61 EP –.

Seitz, A., Loher, T., and Neilson, J. (2007). Seasonal movements and environmental conditions experienced by Pacific halibut in the Bering Sea, examined by pop-up satellite tags. Technical report, International Pacific Halibut Commission.

Serreze, M. C. and Francis, J. A. (2006). The Arctic amplification debate. *Climatic Change*, 76(3):241–264.

Sharma, S., Jackson, D. A., Minns, C. K., and Shuter, B. J. (2007). Will northern fish populations be in hot water because of climate change? *Global Change Biology*, 13(10):2052–2064.

Sigler, M. F. (2000). Abundance estimation and capture of sablefish (*Anoplopoma fimbria*) by longline gear. *Canadian Journal of Fisheries and Aquatic Sciences*, 57(6):1270–1283.

Sigler, M. F. and Lunsford, C. R. (2009). Survey protocol for the alaska sablefish longline survey. Technical report, Alaska Fisheries Science Center.

Sims, C. A. (1972). Money, income, and causality. *The American economic review*, 62(4):540–552.

Skern-Mauritzen, M., Ottersen, G., Handegard, N. O., Huse, G., Dingsør, G. E., Stenseth, N. C., and Kjesbu, O. S. (2015). Ecosystem processes are rarely included in tactical fisheries management. *Fish and Fisheries*, 17(1):165–175.

Smith, G. B. (1979). The biology of walleye pollock. *Fish. oceanogr. eastern Bering Sea shelf.-Seattle, Northwest and Alaska Fish. Center*, pages 213–279.

Smith, R. L., Paul, A. J., and Paul, J. M. (1986). Effect of food intake and temperature on growth and conversion efficiency of juvenile walleye pollock (Theragra chalcogramma (Pallas)): a laboratory study . *ICES Journal of Marine Science*, 42(3):241–253.

Smith Jr, K. and Baldwin, R. (1984). Seasonal fluctuations in deep-sea sediment community oxygen consumption: central and eastern north pacific. *Nature*, 307(5952):624.

Sogard, S. M. and Olla, B. L. (1993). Effects of light, thermoclines and predator presence on vertical distribution and behavioral interactions of juvenile walleye pollock, theragra chalcogramma pallas. *Journal of Experimental Marine Biology and Ecology*, 167(2):179–195.

Sogard, S. M. and Olla, B. L. (1998a). Behavior of juvenile sablefish, Anoplopoma fimbria (Pallas), in a thermal gradient: Balancing food and temperature requirements. *Journal of Experimental Marine Biology and Ecology*, 222(1–2):43–58.

Sogard, S. M. and Olla, B. L. (1998b). Contrasting behavioral responses to cold temperatures by two marine fish species during their pelagic juvenile interval. *Environmental Biology of Fishes*, 53(4):405–412.

Sogard, S. M. and Olla, B. L. (2000). Endurance of simulated winter conditions by age-0 walleye pollock: effects of body size, water temperature and energy stores. *Journal of Fish Biology*, 56(1):1–21.

Sogard, S. M. and Olla, B. L. (2001). Growth and behavioral responses to elevated temperatures by juvenile sablefish *Anoplopoma fimbria* and the interactive role of food availability. *Marine Ecology Progress Series*, 217:121–134.

Sohn, D., Ciannelli, L., and Duffy-Anderson, J. (2016). Distribution of early life Pacific halibut and comparison with Greenland halibut in the eastern Bering Sea. *Journal of Sea Research*, 107, Part 1:31–42. Proceedings of the Ninth International Symposium on Flatfish Ecology.

Song, Y., Jian, L., and Lin, L. (2016). Robust exponential squared loss-based variable selection for high-dimensional single-index varying-coefficient model. *Journal of Computational and Applied Mathematics*, 308:330–345.

Sousa, P., Azevedo, M., and Gomes, M. C. (2006). Species-richness patterns in space, depth, and time (1989–1999) of the Portuguese fauna sampled by bottom trawl. *Aquatic Living Resources*, 19:93–103.

Stark, J. W. (2007). Geographic and seasonal variations in maturation and growth of female pacific cod (gadus macrocephalus) in the gulf of alaska and bering sea. *Fishery Bulletin*, 105(3):396–407.

Stergiou, K. and Christou, E. (1996). Modelling and forecasting annual fisheries catches: comparison of regression, univariate and multivariate time series methods. *Fisheries Research*, 25(2):105–138.

Stergiou, K., Christou, E., and Petrakis, G. (1997). Modelling and forecasting monthly fisheries catches: comparison of regression, univariate and multivariate time series methods. *Fisheries Research*, 29(1):55–95.

Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M. J., and Kattan, M. W. (2010). Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology (Cambridge, Mass.)*, 21(1):128–138.

Stige, L. C., Dalpadado, P., Orlova, E., Boulay, A.-C., Durant, J. M., Ottersen, G., and Stenseth, N. C. (2014). Spatiotemporal statistical analyses reveal predator-driven zooplankton fluctuations in the Barents Sea. *Progress in Oceanography*, 120:243–253.

Stige, L. C., Ottersen, G., Brander, K., Chan, K.-S., and Stenseth, N. C. (2006). Cod and climate: effect of the North Atlantic Oscillation on recruitment in the North Atlantic. *Marine Ecology Progress Series*, 325:227–241.

Stoker, T. M. (1986). Consistent estimation of scaled coefficients. *Econometrica*, 54(6):1461–1481.

Stoner, A. W. and Ottmar, M. L. (2004). Fish density and size alter pacific halibut feeding: implications for stock assessment. *Journal of Fish Biology*, 64(6):1712–1724.

Stoner, A. W., Ottmar, M. L., and Hurst, T. P. (2006). Temperature affects activity and feeding motivation in pacific halibut: Implications for bait-dependent fishing. *Fisheries Research*, 81(2–3):202–209.

Stoner, A. W. and Sturm, E. A. (2004). Temperature and hunger mediate sablefish (Anoplopoma fimbria) feeding motivation: implications for stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, 61(2):238–246.

Suda, A., Suzuki-Ohno, Y., Nagata, N., Sato, M., Narimatsu, Y., and Kawata, M. (2015). Different responses to water temperature in two distinct groups of pacific cod (*Gadus macrocephalus*) inhabiting around japan. In *Proceedings of the North Pacific Marine Science Organization (PICES) 2015 Annual Meeting*. The North Pacific Marine Science Organization (PICES).

Sugihara, G., Deyle, E. R., and Ye, H. (2017). Reply to baskerville and cobey: Misconceptions about causation with synchrony and seasonal drivers. *Proceedings of the National Academy of Sciences of the United States of America*, 114(12):E2272–E2274.

Sugihara, G., May, R., Ye, H., Hsieh, C.-h., Deyle, E., Fogarty, M., and Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, 338(6106):496–500.

Sugihara, G. and May, R. M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344(6268):734–741.

Sugimoto, T. and Tadokoro, K. (1997). Interannual–interdecadal variations in zooplankton biomass, chlorophyll concentration and physical environment in the subarctic pacific and bering sea. *Fisheries Oceanography*, 6(2):74–93.

Sun, W. (2017). *Rank-Based Methods for Single-Index Varying Coefficient Models*. PhD thesis, Auburn University.

Sun, W., Bindele, H. F., Abebe, A., and Correia, H. (2019). General local rank estimation for single-index varying coefficient models. *Journal of Statistical Planning and Inference*, 202:57 – 79.

Swartzman, G., Huang, C., and Kaluzny, S. (1992). Spatial analysis of Bering Sea groundfish survey data using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences*, 49(7):1366–1378.

Takahashi, T., Sutherland, S., Chipman, D., Goddard, J., Ho, C., Newberger, T., Sweeney, C., and Munro, D. (2014). Climatological distributions of ph, pco2, total co2, alkalinity, and caco3 saturation in the global surface ocean, and temporal changes at selected locations. *Marine Chemistry*, 164:95–125.

Takens, F. (1981). Detecting strange attractors in turbulence. In Rand, D. and Young, L.-S., editors, *Dynamical Systems and Turbulence, Warwick 1980*, pages 366–381, Berlin, Heidelberg. Springer Berlin Heidelberg.

Taniguchi, A. (1973). Phytoplankton-zooplankton relationships in the western pacific ocean and adjacent seas. *Marine Biology*, 21(2):115–121.

Taylor, B. J., Rae, J. W., Gray, W. R., Darling, K. F., Burke, A., Gersonde, R., Abelmann, A., Maier, E., Esper, O., and Ziveri, P. (2018). Distribution and ecology of planktic foraminifera in the North Pacific: Implications for paleo-reconstructions. *Quaternary Science Reviews*, 191:256–274.

Timmermann, A., Oberhuber, J., Bacher, A., Esch, M., Latif, M., and Roeckner, E. (1999). Increased El Niño frequency in a climate model forced by future greenhouse warming. *Nature*, 398(6729):694–697.

Tolimieri, N. and Levin, P. S. (2006). Assemblage structure of Eastern Pacific groundfishes on the U.S. Continental Slope in relation to physical and environmental variables. *Transactions of the American Fisheries Society*, 135(2):317–332.

Townhill, B. L., Maxwell, D., Engelhard, G. H., Simpson, S. D., and Pinnegar, J. K. (2015). Historical Arctic logbooks provide insights into past diets and climatic responses of cod. *PLoS ONE*, 10(9):1–15.

Tseng, Y.-H., Ding, R., and meng Huang, X. (2017). The warm blob in the northeast Pacific—the bridge leading to the 2015/16 El Niño. *Environmental Research Letters*, 12(5):054019.

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 5(2):99–114.

Turner, R. E., Rabalais, N. N., and Justić, D. (2017). Trends in summer bottom-water temperatures on the northern gulf of mexico continental shelf from 1985 to 2015. *PLOS ONE*, 12(9):1–12.

Urban, M. C., De Meester, L., Vellend, M., Stoks, R., and Vanoverbeke, J. (2012). A crucial step toward realism: responses to climate change from an evolving metacommunity perspective. *Evolutionary Applications*, 5(2):154–167.

Vaidyanathan, G. (2017). Inner workings: Climate change complicates fisheries modeling and management. *Proceedings of the National Academy of Sciences*, 114(32):8435–8437.

van Nes, E. H., Scheffer, M., Brovkin, V., Lenton, T. M., Ye, H., Deyle, E., and Sugihara, G. (2015). Causal feedbacks in climate change. *Nature Climate Change*, 5:445 EP.

von Szalay, P. G. and Raring, N. W. (2016). Data report: 2015 gulf of alaska bottom trawl survey. Technical Report NOAA Tech. Memo. NMFS-AFSC-325, U.S. Dep. Commerce.

Wakita, M., Watanabe, S., Honda, M., Nagano, A., Kimoto, K., Matsumoto, K., Kitamura, M., Sasaki, K., Kawakami, H., Fujiki, T., et al. (2013). Ocean acidification from 1997 to 2011 in the subarctic western North Pacific Ocean. *Biogeosciences*, 10(12):7817–7827.

Walker, W. A., Mead, J. G., and Brownell, R. L. (2002). Diets of Baird's beaked whales, *Berardius bairdii*, in the southern Sea of Okhotsk and off the Pacific coast of Honshu, Japan. *Marine Mammal Science*, 18(4):902–919.

Walsh, J. E. (2008). Climate of the arctic marine environment. *Ecological Applications*, 18(sp2):S3–S22.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J.-M., Hoegh-Guldberg, O., and Bairlein, F. (2002). Ecological responses to recent climate change. *Nature*, 416:389 EP.

Wang, H. and Xia, Y. (2009). Shrinkage estimation of the varying coefficient model. *Journal of the American Statistical Association*, 104(486):747–757.

Wang, W., Anderson, B. T., Kaufmann, R. K., and Myneni, R. B. (2004). The relation between the north atlantic oscillation and ssts in the north atlantic basin. *Journal of Climate*, 17(24):4752–4759.

Wang, Y., Yang, J., Chen, Y., De Maeyer, P., Li, Z., and Duan, W. (2018). Detecting the causal effect of soil moisture on precipitation using convergent cross mapping. *Scientific Reports*, 8(1):12171.

Wasserstein, R. L., Schirm, A. L., and Lazar, N. A. (2019). Moving to a world beyond "p $<$ 0.05". *The American Statistician*, 73(sup1):1–19.

Wassman, P., Duarte, C. M., Agustí, S., and Sejr, M. K. (2011). Footprints of climate change in the arctic marine ecosystem. *Global Change Biology*, 17(2):1235–1249.

Wei, F., Huang, J., and Li, H. (2011). Variable selection and estimation in high-dimensional varying-coefficient models. *Statistica Sinica*, 21(4):1515–1540.

Weisse, R. (2010). *Marine climate and climate change: storms, wind waves and storm surges*. Springer Science & Business Media.

Wespestad, V. (1996). Importance of cannibalism in the population dynamics of walleye pollock, theragra chalcogramma. Technical report, NOAA NMFS.

Wespestad, V. G., Fritz, L. W., Ingraham, W. J., and Megrey, B. A. (2000). On relationships between cannibalism, climate variability, physical transport, and recruitment success of Bering Sea walleye pollock (Theragra chalcogramma). *ICES Journal of Marine Science*, 57(2):272–278.

Whitney, F., Crawford, W., and Harrison, P. (2005). Physical processes that enhance nutrient transport and primary productivity in the coastal and open ocean of the subarctic ne pacific. *Deep Sea Research Part II: Topical Studies in Oceanography*, 52(5):681–706. Linkages between coastal and open ocean ecosystems.

Whitney, F. and Robert, M. (2002). Structure of haida eddies and their transport of nutrient from coastal margins into the ne pacific ocean. *Journal of Oceanography*, 58(5):715–723.

Whitney, F. A., Freeland, H. J., and Robert, M. (2007). Persistently declining oxygen levels in the interior waters of the eastern subarctic Pacific. *Progress in Oceanography*, 75(2):179–199. Time Series of the Northeast Pacific.

Wiener, N. (1956). The theory of prediction. *Modern mathematics for engineers*.

Williams, A., Althaus, F., Dunstan, P. K., Poore, G. C. B., Bax, N. J., Kloser, R. J., and McEnnulty, F. R. (2010). Scales of habitat heterogeneity and megabenthos biodiversity on an extensive australian continental margin (100–1100 m depths). *Marine Ecology*, 31(1):222–236.

Wong, C. and Matear, R. (1999). Sporadic silicate limitation of phytoplankton productivity in the subarctic ne pacific. *Deep Sea Research Part II: Topical Studies in Oceanography*, 46(11-12):2539–2555.

Wong, C., Waser, N., Nojiri, Y., Whitney, F., Page, J., and Zeng, J. (2002a). Seasonal cycles of nutrients and dissolved inorganic carbon at high and mid latitudes in the north pacific ocean

during the skaugran cruises: determination of new production and nutrient uptake ratios. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(24-25):5317–5338.

Wong, C. S., Waser, N. A., Nojiri, Y., Johnson, W. K., Whitney, F. A., Page, J. S., and Zeng, J. (2002b). Seasonal and interannual variability in the distribution of surface nutrients and dissolved inorganic carbon in the northern north pacific: Influence of el niño. *Journal of Oceanography*, 58(2):227–243.

Wong, R. K. W., Yao, F., and Lee, T. C. M. (2014). Robust estimation for generalized additive models. *J. Comput. Graph. Statist.*, 23(1):270–289.

Wood, S. N. (2006). *Generalized additive models*. Texts in Statistical Science Series. Chapman & Hall/CRC, Boca Raton, FL. An Introduction with R.

Wood, S. N. and Augustin, N. H. (2002). GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. *Ecological Modelling*, 157(2–3):157–177.

Wootton, J. T., Pfister, C. A., and Forester, J. D. (2008). Dynamic patterns and ecological impacts of declining ocean ph in a high-resolution multi-year dataset. *Proceedings of the National Academy of Sciences*, 105(48):18848–18853.

Xia, Y. and Li, W. K. (1999). On single-index coefficient regression models. *Journal of the American Statistical Association*, 94(448):1275–1285.

Xia, Y., Tong, H., Li, W. K., and Zhu, L.-X. (2002). An adaptive estimation of dimension reduction space. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(3):363–410.

Xue, L. and Pang, Z. (2013). Statistical inference for a single-index varying-coefficient model. *Statistics and Computing*, 23(5):589–599.

Xue, L. and Qu, A. (2012). Variable selection in high-dimensional varying-coefficient models with global optimality. *Journal of Machine Learning Research*, 13(Jun):1973–1998.

Xue, L. and Wang, Q. (2012). Empirical likelihood for single-index varying-coefficient models. *Bernoulli*, 18(3):836–856.

Yang, J. and Yang, H. (2017). Robust modal estimation and variable selection for single-index varying-coefficient models. *Communications in Statistics - Simulation and Computation*, 46(4):2976–2997.

Yang, M.-S., Dodd, K., Hibpshman, R., and Whitehouse, A. (2006). Food habits of groundfishes in the gulf of alaska in 1999 and 2001. NOAA Technical Memorandum 164, U.S. Department of Commerce, NOAA, NMFS, AFSC.

Ye, H., Deyle, E. R., Gilarranz, L. J., and Sugihara, G. (2015). Distinguishing time-delayed causal interactions using convergent cross mapping. *Scientific Reports*, 5:14750.

Yoshimura, T., Nishioka, J., Saito, H., Takeda, S., Tsuda, A., and Wells, M. L. (2007). Distributions of particulate and dissolved organic and inorganic phosphorus in North Pacific surface waters. *Marine Chemistry*, 103(1):112–121.

Zweng, M., Reagan, J., Antonov, J., Locarnini, R., Mishonov, A., Boyer, T., Garcia, H., Baranova, O., Johnson, D., D.Seidov, and Biddle, M. (2013). World Ocean Atlas 2013, Volume 2: Salinity. Technical report, National Oceanographic Data Center, Ocean Climate Laboratory, NOAA.

Appendices

Appendix A

Chapter 2 Rank GAM Supplementary Material

Table A.1: Rank GAM parameter estimates for sablefish CPUE using n=2398 observations. $R^2_{adj} = 0.928$.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	7.2254	0.0174	415.5173	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	95.3729	101.6192	194.8051	$<$ **0.0001**
s(ATMP_samp5)	6.9385	8.0600	10.6985	$<$ **0.0001**
s(PRES_samp5)	7.8014	8.5915	30.1157	$<$ **0.0001**
s(WSPD_samp5)	7.3308	8.3118	20.9802	$<$ **0.0001**
s(WTMP_samp5)	8.2409	8.8246	9.0155	$<$ **0.0001**
s(WVHT_samp5)	8.7219	8.9716	15.4024	$<$ **0.0001**
s(plankton_samp5)	6.4871	7.6038	4.1760	**0.0001**
s(BOT_TEMP_samp5)	8.1180	8.7933	5.3048	$<$ **0.0001**
s(Alk_75m_samp5)	3.1184	3.8901	0.9029	0.4012
s(Alk_400m_samp5)	1.6382	2.0681	1.3967	0.2375
s(Alk_900m_samp5)	2.5165	3.2245	1.2233	0.3017
s(Chl_75m_samp5)	2.4834	3.0127	6.1409	**0.0004**
s(Chl_400m_samp5)	6.2628	7.3231	8.6350	$<$ **0.0001**
s(Chl_900m_samp5)	6.6054	7.6737	3.7224	**0.0003**
s(NO3_75m_samp5)	4.3167	5.4085	2.4268	**0.0326**
s(NO3_900m_samp5)	4.8212	6.0040	6.8127	$<$ **0.0001**
s(Oxy_75m_samp5)	7.2372	8.1097	2.4270	**0.0117**
s(Oxy_400m_samp5)	3.0222	3.6969	1.1343	0.3826
s(Oxy_900m_samp5)	5.7075	6.7157	1.1203	0.3095
s(Phos_75m_samp5)	1.0009	1.0017	0.1946	0.6596
s(Phos_900m_samp5)	7.0527	8.1404	5.6123	$<$ **0.0001**
s(Sal_75m_samp5)	1.0002	1.0003	2.1582	0.1419
s(Sal_400m_samp5)	1.4661	1.8274	5.7839	**0.0116**
s(Sal_900m_samp5)	1.0002	1.0003	0.1776	0.6735
s(Sil_75m_samp5)	1.0001	1.0002	13.0594	**0.0003**
s(Sil_900m_samp5)	8.2496	8.8407	3.3021	**0.0004**

Table A.2: Rank GAM parameter estimates for Pacific cod CPUE using n=1913 observations. $R^2_{adj} = 0.914$.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.8180	0.0052	158.6443	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	104.0860	110.6271	90.7087	$<$ **0.0001**
s(ATMP_samp5)	8.4778	8.9170	4.8940	$<$ **0.0001**
s(PRES_samp5)	1.0001	1.0002	19.0707	$<$ **0.0001**
s(WSPD_samp5)	8.7041	8.9698	5.0235	$<$ **0.0001**
s(WTMP_samp5)	3.6058	4.5886	2.6911	**0.0239**
s(WVHT_samp5)	7.8638	8.6520	3.0774	**0.0040**
s(plankton_samp5)	1.0000	1.0000	41.2428	$<$ **0.0001**
s(BOT_TEMP_samp5)	2.8713	3.7540	1.5875	0.1576
s(Alk_75m_samp5)	1.0000	1.0000	0.0010	0.9746
s(Alk_400m_samp5)	1.5980	2.0128	2.2699	0.1033
s(Alk_900m_samp5)	6.4916	7.6758	10.6967	$<$ **0.0001**
s(Chl_75m_samp5)	1.0000	1.0000	2.0350	0.1539
s(Chl_400m_samp5)	1.0000	1.0000	2.4281	0.1194
s(Chl_900m_samp5)	1.0000	1.0000	1.0658	0.3020
s(NO3_75m_samp5)	1.5864	1.9989	0.8693	0.4280
s(NO3_900m_samp5)	1.0000	1.0000	2.6490	0.1038
s(Oxy_75m_samp5)	6.2843	7.3519	6.5191	$<$ **0.0001**
s(Oxy_400m_samp5)	1.0000	1.0000	3.2147	0.0732
s(Oxy_900m_samp5)	7.6227	8.4716	4.5539	$<$ **0.0001**
s(Phos_75m_samp5)	4.0950	5.1702	3.6168	**0.0028**
s(Phos_900m_samp5)	1.0000	1.0000	5.0976	**0.0241**
s(Sal_75m_samp5)	7.0594	8.1341	2.4860	**0.0105**
s(Sal_400m_samp5)	2.5947	3.3615	2.4974	0.0530
s(Sal_900m_samp5)	2.8908	3.7389	0.9288	**0.3729**
s(Sil_75m_samp5)	4.7415	5.9042	6.5153	$<$ **0.0001**
s(Sil_900m_samp5)	4.7344	5.9042	3.6886	**0.0015**

Table A.3: Rank GAM parameter estimates for Pacific halibut CPUE using n=1848 observations. $R^2_{adj} = 0.807$.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.5471	0.0035	154.9622	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	96.5370	104.3250	52.2808	$<$ **0.0001**
s(ATMP_samp10)	7.7749	8.5159	4.6300	$<$ **0.0001**
s(PRES_samp10)	7.7202	8.5403	8.1055	$<$ **0.0001**
s(WSPD_samp10)	8.4696	8.9105	7.5199	$<$ **0.0001**
s(WTMP_samp10)	8.4365	8.8659	11.6876	$<$ **0.0001**
s(WVHT_samp10)	7.9253	8.6884	3.0946	**0.0012**
s(plankton_samp10)	5.6423	6.8262	3.8428	**0.0005**
s(BOT_TEMP_samp10)	1.6476	2.1084	17.5220	$<$ **0.0001**
s(Alk_75m_samp10)	1.0000	1.0000	7.4190	**0.0065**
s(Alk_400m_samp10)	5.8171	6.9878	1.9821	0.0584
s(Alk_900m_samp10)	5.2359	6.4368	4.3486	**0.0001**
s(Chl_75m_samp10)	1.0000	1.0000	5.3250	**0.0211**
s(Chl_400m_samp10)	8.3827	8.8804	3.1097	**0.0010**
s(Chl_900m_samp10)	7.3932	8.3578	2.7981	**0.0039**
s(NO3_75m_samp10)	4.4938	5.5923	3.3164	**0.0035**
s(NO3_900m_samp10)	1.0000	1.0000	3.7890	0.0518
s(Oxy_75m_samp10)	1.8553	2.3108	4.1154	**0.0141**
s(Oxy_400m_samp10)	2.9386	3.7809	13.2686	$<$ **0.0001**
s(Oxy_900m_samp10)	1.0000	1.0000	8.8952	**0.0029**
s(Phos_75m_samp10)	1.0000	1.0000	0.0302	0.8621
s(Phos_900m_samp10)	1.0000	1.0000	2.6976	0.1007
s(Sal_75m_samp10)	1.0000	1.0000	0.5784	0.4471
s(Sal_400m_samp10)	1.0000	1.0000	5.4792	**0.0194**
s(Sal_900m_samp10)	1.0000	1.0000	1.4096	0.2353
s(Sil_75m_samp10)	4.8343	6.0144	1.7151	0.1135
s(Sil_900m_samp10)	1.0000	1.0000	0.0831	0.7732

Table A.4: Rank GAM parameter estimates for shortspine thornyhead CPUE using n=2243 observations. $R^2_{adj} = 0.902$.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	1.3154	0.0042	311.8674	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	91.6437	99.5099	112.4463	$<$ **0.0001**
s(ATMP_samp5)	7.9823	8.7326	18.8092	$<$ **0.0001**
s(PRES_samp5)	8.0947	8.7483	6.1895	$<$ **0.0001**
s(WSPD_samp5)	7.7099	8.5628	2.8806	**0.0022**
s(WTMP_samp5)	8.6577	8.9560	6.1947	$<$ **0.0001**
s(WVHT_samp5)	8.4758	8.9102	3.6489	**0.0003**
s(plankton_samp5)	6.8943	7.9585	4.4118	$<$ **0.0001**
s(BOT_TEMP_samp5)	7.3882	8.3859	6.1857	$<$ **0.0001**
s(Alk_75m_samp5)	1.0000	1.0000	7.2954	**0.0070**
s(Alk_400m_samp5)	7.0167	8.0832	1.2468	0.2382
s(Alk_900m_samp5)	1.2366	1.4362	9.6023	**0.0015**
s(Chl_75m_samp5)	5.2171	6.1523	6.6079	$<$ **0.0001**
s(Chl_400m_samp5)	8.0296	8.6962	4.1306	**0.0001**
s(Chl_900m_samp5)	1.0000	1.0000	0.9615	0.3269
s(NO3_75m_samp5)	7.3448	8.3521	2.5239	**0.0099**
s(NO3_900m_samp5)	3.8175	4.8525	1.9601	0.0868
s(Oxy_75m_samp5)	1.0000	1.0000	19.0360	$<$ **0.0001**
s(Oxy_400m_samp5)	6.7665	7.6687	2.2766	**0.0272**
s(Oxy_900m_samp5)	7.6209	8.4204	2.8485	**0.0036**
s(Phos_75m_samp5)	3.1096	4.0208	0.8525	0.4991
s(Phos_900m_samp5)	1.0000	1.0000	0.2975	0.5855
s(Sal_75m_samp5)	3.5007	4.4495	2.1180	0.0688
s(Sal_400m_samp5)	5.8388	7.0112	2.0170	0.0518
s(Sal_900m_samp5)	4.7618	5.8937	2.3864	**0.0244**
s(Sil_75m_samp5)	1.0000	1.0000	0.0953	0.7575
s(Sil_900m_samp5)	1.0000	1.0000	2.3049	0.1291

Table A.5: Rank GAM parameter estimates for rougheye rockfish CPUE using n=1731 observations. $R^2_{adj} = 0.777$.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.5455	0.0032	170.3336	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	90.7873	98.7796	37.1464	$<$ **0.0001**
s(ATMP_samp5)	8.4175	8.8982	5.9154	$<$ **0.0001**
s(PRES_samp5)	5.3917	6.5186	5.3192	$<$ **0.0001**
s(WSPD_samp5)	8.4091	8.8898	5.4173	$<$ **0.0001**
s(WTMP_samp5)	8.4476	8.8991	5.6473	$<$ **0.0001**
s(WVHT_samp5)	8.2977	8.8588	6.4698	$<$ **0.0001**
s(plankton_samp5)	3.3387	4.1922	2.7785	**0.0223**
s(BOT_TEMP_samp5)	1.0000	1.0000	13.5388	**0.0002**
s(Alk_75m_samp5)	1.0000	1.0000	2.1800	0.1400
s(Alk_400m_samp5)	1.0000	1.0000	0.8320	0.3619
s(Alk_900m_samp5)	7.5266	8.4256	6.3466	$<$ **0.0001**
s(Chl_75m_samp5)	4.6916	5.6664	1.5070	0.2468
s(Chl_400m_samp5)	6.2099	7.2647	6.3643	$<$ **0.0001**
s(Chl_900m_samp5)	6.5118	7.5620	3.9499	**0.0003**
s(NO3_75m_samp5)	1.0000	1.0000	0.5681	0.4511
s(NO3_900m_samp5)	1.0000	1.0000	11.1183	**0.0009**
s(Oxy_75m_samp5)	8.6466	8.9018	4.5532	$<$ **0.0001**
s(Oxy_400m_samp5)	8.5038	8.8413	4.9083	$<$ **0.0001**
s(Oxy_900m_samp5)	1.0000	1.0000	0.1580	0.6910
s(Phos_75m_samp5)	2.9951	3.8372	1.4937	0.2550
s(Phos_900m_samp5)	1.0000	1.0000	1.1963	0.2742
s(Sal_75m_samp5)	1.0000	1.0000	0.1071	0.7436
s(Sal_400m_samp5)	7.5601	8.4934	3.1115	**0.0023**
s(Sal_900m_samp5)	1.0000	1.0000	0.2126	0.6448
s(Sil_75m_samp5)	4.3872	5.4680	1.5260	0.1932
s(Sil_900m_samp5)	8.3852	8.8886	7.4664	$<$ **0.0001**

Table A.6: Rank GAM parameter estimates for shortraker rockfish CPUE using n=1738 observations. $R^2_{adj} = 0.838$.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.4363	0.0029	150.7043	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	74.6394	84.6966	61.1331	$<$ **0.0001**
s(ATMP_samp5)	3.4919	4.4888	1.5848	0.1760
s(PRES_samp5)	6.7978	7.8477	7.0194	$<$ **0.0001**
s(WSPD_samp5)	8.7698	8.9751	9.4908	$<$ **0.0001**
s(WTMP_samp5)	8.2760	8.8292	8.0358	$<$ **0.0001**
s(WVHT_samp5)	7.9906	8.7211	4.9348	$<$ **0.0001**
s(plankton_samp5)	4.0199	4.9514	3.5680	**0.0037**
s(BOT_TEMP_samp5)	1.0000	1.0000	25.8542	$<$ **0.0001**
s(Alk_75m_samp5)	6.3648	7.5427	9.0931	$<$ **0.0001**
s(Alk_400m_samp5)	1.0000	1.0000	0.0076	0.9307
s(Alk_900m_samp5)	2.0435	2.5163	4.2013	**0.0126**
s(Chl_75m_samp5)	3.8181	4.6600	2.3642	**0.0281**
s(Chl_400m_samp5)	6.5707	7.6031	3.3546	**0.0013**
s(Chl_900m_samp5)	6.6084	7.6501	1.5082	0.1914
s(NO3_75m_samp5)	4.2631	5.3322	1.7806	0.1202
s(NO3_900m_samp5)	1.0000	1.0000	0.1829	0.6690
s(Oxy_75m_samp5)	1.0000	1.0000	0.2132	0.6443
s(Oxy_400m_samp5)	2.1535	2.6589	1.6637	0.2561
s(Oxy_900m_samp5)	2.7380	3.3410	0.8778	0.4757
s(Phos_75m_samp5)	1.0000	1.0000	0.0121	0.9126
s(Phos_900m_samp5)	1.0000	1.0000	0.8050	0.3697
s(Sal_75m_samp5)	2.3780	3.0440	2.6762	**0.0472**
s(Sal_400m_samp5)	3.0211	3.8655	2.5857	**0.0325**
s(Sal_900m_samp5)	1.0000	1.0000	1.1457	0.2846
s(Sil_75m_samp5)	7.7498	8.5676	2.8813	**0.0032**
s(Sil_900m_samp5)	8.0671	8.7554	7.9205	$<$ **0.0001**

Table A.7: Rank GAM parameter estimates for sablefish mean weight (kg) using n=2385 observations. $R^2_{adj} = 0.604$.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	2.9552	0.0074	397.4390	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	91.8757	99.4588	26.3834	$<$ **0.0001**
s(ATMP_samp5)	3.1427	4.0676	3.0505	**0.0159**
s(PRES_samp5)	1.5165	1.8942	1.2482	0.3464
s(WSPD_samp5)	8.5225	8.9228	5.0531	$<$ **0.0001**
s(WTMP_samp5)	7.0264	8.0926	5.1804	$<$ **0.0001**
s(WVHT_samp5)	5.8098	7.0317	1.9633	0.0550
s(plankton_samp5)	1.0000	1.0000	1.3904	0.2385
s(BOT_TEMP_samp5)	2.4347	3.1841	2.7054	**0.0382**
s(Alk_75m_samp5)	1.0000	1.0000	0.0898	0.7645
s(Alk_400m_samp5)	3.1981	4.0425	1.3086	0.2595
s(Alk_900m_samp5)	1.5756	1.9715	1.0634	0.3738
s(Chl_75m_samp5)	1.0000	1.0000	0.0834	0.7727
s(Chl_400m_samp5)	1.0000	1.0000	2.5716	0.1089
s(Chl_900m_samp5)	1.3341	1.5949	0.4125	0.4891
s(NO3_75m_samp5)	2.4542	3.1829	1.6885	0.1741
s(NO3_900m_samp5)	1.0000	1.0001	1.8071	0.1790
s(Oxy_75m_samp5)	1.0001	1.0003	0.3519	0.5531
s(Oxy_400m_samp5)	1.7460	2.2022	2.0640	0.1297
s(Oxy_900m_samp5)	1.0001	1.0002	0.0050	0.9434
s(Phos_75m_samp5)	6.2562	7.4676	1.6029	0.1212
s(Phos_900m_samp5)	1.0000	1.0000	0.3379	0.5611
s(Sal_75m_samp5)	4.1734	5.1842	1.3524	0.2383
s(Sal_400m_samp5)	1.0000	1.0000	0.2073	0.6489
s(Sal_900m_samp5)	1.2529	1.4640	0.0531	0.8403
s(Sil_75m_samp5)	3.3933	4.3039	1.4801	0.1975
s(Sil_900m_samp5)	1.0000	1.0000	0.4509	0.5020

Table A.8: Rank GAM parameter estimates for Pacific cod mean weight (kg) using n=1892 observations. $R^2_{adj} = 0.460$.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	2.9375	0.0102	288.5356	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	89.3338	96.7584	14.0390	$<$ **0.0001**
s(ATMP_samp5)	1.0000	1.0001	2.9858	0.0842
s(PRES_samp5)	8.4050	8.8826	5.6667	$<$ **0.0001**
s(WSPD_samp5)	2.0402	2.6416	1.2792	0.2451
s(WTMP_samp5)	1.6118	2.0256	3.5598	**0.0284**
s(WVHT_samp5)	1.0000	1.0000	25.6873	$<$ **0.0001**
s(plankton_samp5)	1.0000	1.0000	2.0771	0.1497
s(BOT_TEMP_samp5)	3.4141	4.4298	8.4963	$<$ **0.0001**
s(Alk_75m_samp5)	1.3024	1.5435	0.3389	0.7436
s(Alk_400m_samp5)	1.0000	1.0000	4.1426	**0.0420**
s(Alk_900m_samp5)	4.7524	5.8810	1.9887	0.0653
s(Chl_75m_samp5)	1.3690	1.6356	0.3576	0.7326
s(Chl_400m_samp5)	1.3036	1.5410	0.2787	0.5880
s(Chl_900m_samp5)	1.0000	1.0000	2.1520	0.1426
s(NO3_75m_samp5)	1.0000	1.0000	6.8731	**0.0088**
s(NO3_900m_samp5)	1.0000	1.0000	0.5033	0.4782
s(Oxy_75m_samp5)	1.0000	1.0000	0.4636	0.4960
s(Oxy_400m_samp5)	1.0000	1.0001	2.5303	0.1119
s(Oxy_900m_samp5)	1.0000	1.0000	0.1998	0.6550
s(Phos_75m_samp5)	5.5218	6.7462	1.1261	0.3431
s(Phos_900m_samp5)	2.9359	3.7662	1.2331	0.3315
s(Sal_75m_samp5)	1.0000	1.0000	4.3804	**0.0365**
s(Sal_400m_samp5)	1.0000	1.0000	0.4454	0.5046
s(Sal_900m_samp5)	1.0000	1.0000	0.0297	0.8632
s(Sil_75m_samp5)	1.0000	1.0000	8.0143	**0.0047**
s(Sil_900m_samp5)	3.5048	4.4759	2.8198	**0.0200**

Table A.9: Rank GAM parameter estimates for Pacific halibut mean weight (kg) using n=331 observations. $R^2_{adj} = 0.468$. Due to low number of observations, the number of knots per spline was reduced from the default of 10 to 5 for fitting of the GAM.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	2.6543	0.0303	87.5738	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	18.5680	22.4492	5.3606	$<$ **0.0001**
s(ATMP_samp10)	3.6880	3.9210	1.5502	0.1455
s(PRES_samp10)	1.7873	2.1805	1.3292	0.2885
s(WSPD_samp10)	1.0000	1.0000	4.7954	**0.0294**
s(WTMP_samp10)	1.0000	1.0000	3.4762	0.0633
s(WVHT_samp10)	1.4979	1.7639	1.4641	0.2135
s(plankton_samp10)	3.1523	3.6391	3.9262	**0.0168**
s(BOT_TEMP_samp10)	1.4500	1.7614	0.5407	0.6496
s(Alk_75m_samp10)	1.0000	1.0000	6.7142	**0.0101**
s(Alk_400m_samp10)	1.0000	1.0000	0.2147	0.6435
s(Alk_900m_samp10)	3.7967	3.9660	5.0430	**0.0007**
s(Chl_75m_samp10)	3.7019	3.9420	10.3416	$<$ **0.0001**
s(Chl_400m_samp10)	1.0000	1.0000	2.9501	0.0870
s(Chl_900m_samp10)	1.0000	1.0000	0.4532	0.5014
s(NO3_75m_samp10)	3.2386	3.7003	0.9307	0.2985
s(NO3_900m_samp10)	1.0000	1.0000	0.2906	0.5903
s(Oxy_75m_samp10)	1.0000	1.0000	2.5562	0.1110
s(Oxy_400m_samp10)	1.0000	1.0000	4.7042	**0.0310**
s(Oxy_900m_samp10)	1.8010	2.2872	0.9010	0.5108
s(Phos_75m_samp10)	1.0000	1.0000	0.0096	0.9222
s(Phos_900m_samp10)	1.0000	1.0000	3.9448	**0.0480**
s(Sal_75m_samp10)	1.0000	1.0000	0.0494	0.8243
s(Sal_400m_samp10)	1.0000	1.0000	1.8463	0.1754
s(Sal_900m_samp10)	4.0000	4.0000	3.6574	**0.0064**
s(Sil_75m_samp10)	2.7769	3.3122	2.4872	0.0593
s(Sil_900m_samp10)	2.7939	3.2986	1.9542	0.0969

Table A.10: Rank GAM parameter estimates for shortspine thornyhead mean weight (kg) using n=2222 observations. $R^2_{adj} = 0.844$.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	0.6936	0.0024	292.7378	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	121.7052	122.7425	87.8403	$<$ 0.0001
s(ATMP_samp5)	1.0004	1.0008	1.0367	0.3089
s(PRES_samp5)	1.0011	1.0022	2.8464	0.0915
s(WSPD_samp5)	1.0001	1.0003	1.0980	0.2948
s(WTMP_samp5)	1.4612	1.8130	0.2810	0.7024
s(WVHT_samp5)	1.0003	1.0006	0.6043	0.4372
s(plankton_samp5)	1.0002	1.0004	2.6903	0.1011
s(BOT_TEMP_samp5)	1.0004	1.0009	0.3481	0.5557
s(Alk_75m_samp5)	1.0002	1.0003	0.0700	0.7913
s(Alk_400m_samp5)	2.9380	3.7292	1.2783	0.2659
s(Alk_900m_samp5)	1.0000	1.0001	1.4570	0.2275
s(Chl_75m_samp5)	1.0000	1.0000	1.8945	0.1688
s(Chl_400m_samp5)	1.0000	1.0000	2.3143	0.1283
s(Chl_900m_samp5)	1.0000	1.0001	0.8982	0.3434
s(NO3_75m_samp5)	1.0011	1.0022	0.2708	0.6036
s(NO3_900m_samp5)	1.0007	1.0013	1.4006	0.2369
s(Oxy_75m_samp5)	1.0001	1.0002	10.1491	**0.0015**
s(Oxy_400m_samp5)	8.7545	8.9713	3.7499	**0.0003**
s(Oxy_900m_samp5)	2.6087	3.2114	5.9168	**0.0004**
s(Phos_75m_samp5)	1.0002	1.0003	2.1469	0.1430
s(Phos_900m_samp5)	1.0001	1.0002	0.5581	0.4551
s(Sal_75m_samp5)	1.0005	1.0010	0.7886	0.3745
s(Sal_400m_samp5)	1.0002	1.0005	2.6166	0.1059
s(Sal_900m_samp5)	1.0002	1.0004	1.0494	0.3057
s(Sil_75m_samp5)	1.0004	1.0007	0.0536	0.8174
s(Sil_900m_samp5)	1.0002	1.0004	0.7257	0.3944

Table A.11: Rank GAM parameter estimates for rougheye rockfish mean weight (kg) using n=1660 observations. $R^2_{adj} = 0.638$.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	1.4499	0.0055	262.4880	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	47.6426	48.6173	47.0077	$<$ **0.0001**
s(ATMP_samp5)	1.0001	1.0001	0.3490	0.5548
s(PRES_samp5)	1.0000	1.0001	1.9676	0.1609
s(WSPD_samp5)	1.0000	1.0000	6.3452	**0.0119**
s(WTMP_samp5)	1.4784	1.8375	0.4470	0.6774
s(WVHT_samp5)	1.0000	1.0001	0.3366	0.5619
s(plankton_samp5)	5.5740	6.7099	1.6557	0.1103
s(BOT_TEMP_samp5)	1.0000	1.0001	2.4451	0.1181
s(Alk_75m_samp5)	2.2781	2.9020	0.8651	0.5448
s(Alk_400m_samp5)	1.0000	1.0001	1.3441	0.2465
s(Alk_900m_samp5)	1.0001	1.0001	2.4015	0.1214
s(Chl_75m_samp5)	2.2949	2.8310	5.1131	**0.0033**
s(Chl_400m_samp5)	3.3233	4.1110	11.6975	$<$ **0.0001**
s(Chl_900m_samp5)	1.1052	1.1980	2.8355	0.0956
s(NO3_75m_samp5)	1.0001	1.0001	0.8282	0.3629
s(NO3_900m_samp5)	1.0000	1.0000	6.4465	**0.0112**
s(Oxy_75m_samp5)	4.5590	5.5398	2.3118	**0.0392**
s(Oxy_400m_samp5)	1.4954	1.8343	0.3831	0.5849
s(Oxy_900m_samp5)	1.0000	1.0001	0.0448	0.8325
s(Phos_75m_samp5)	1.3484	1.6254	0.3355	0.7385
s(Phos_900m_samp5)	1.0000	1.0001	3.0987	0.0785
s(Sal_75m_samp5)	1.4805	1.8512	0.3705	0.6886
s(Sal_400m_samp5)	1.0001	1.0001	0.4107	0.5217
s(Sal_900m_samp5)	1.0000	1.0001	0.4135	0.5203
s(Sil_75m_samp5)	1.2753	1.5025	0.1271	0.7310
s(Sil_900m_samp5)	1.0000	1.0001	0.7093	0.3998

Table A.12: Rank GAM parameter estimates for shortraker rockfish mean weight (kg) using n=1547 observations. $R^2_{adj} = 0.604$.

Parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	1.4491	0.0073	197.7515	$<$ **0.0001**
Smooth terms	edf	Ref.df	F-value	p-value
te(Longitude,Latitude,Year)	42.5210	48.0894	37.3158	$<$ **0.0001**
s(ATMP_samp5)	2.1302	2.7591	3.6608	**0.0122**
s(PRES_samp5)	1.0000	1.0000	0.0702	0.7911
s(WSPD_samp5)	1.0000	1.0000	0.0019	0.9653
s(WTMP_samp5)	1.0000	1.0000	0.0839	0.7721
s(WVHT_samp5)	1.0000	1.0000	2.0169	0.1558
s(plankton_samp5)	1.0000	1.0000	0.7302	0.3930
s(BOT_TEMP_samp5)	1.7390	2.2387	1.6841	0.1875
s(Alk_75m_samp5)	1.0000	1.0000	0.2042	0.6514
s(Alk_400m_samp5)	1.0000	1.0000	1.3139	0.2519
s(Alk_900m_samp5)	5.5674	6.7252	2.0580	**0.0448**
s(Chl_75m_samp5)	1.0000	1.0000	1.6130	0.2043
s(Chl_400m_samp5)	1.0000	1.0000	0.1568	0.6922
s(Chl_900m_samp5)	1.0000	1.0000	1.5636	0.2113
s(NO3_75m_samp5)	5.5463	6.7237	1.5914	0.1538
s(NO3_900m_samp5)	1.0064	1.0127	0.6245	0.4274
s(Oxy_75m_samp5)	3.7585	4.6434	3.8596	**0.0037**
s(Oxy_400m_samp5)	6.3236	7.4316	0.6980	0.6035
s(Oxy_900m_samp5)	1.0000	1.0000	0.2895	0.5906
s(Phos_75m_samp5)	1.9761	2.5524	0.8646	0.5292
s(Phos_900m_samp5)	2.6623	3.4091	1.1441	0.3020
s(Sal_75m_samp5)	2.5124	3.2307	3.7728	**0.0089**
s(Sal_400m_samp5)	1.0000	1.0000	12.5844	**0.0004**
s(Sal_900m_samp5)	1.0000	1.0000	0.0340	0.8537
s(Sil_75m_samp5)	1.6222	2.0358	0.4334	0.6728
s(Sil_900m_samp5)	1.0000	1.0000	0.1690	0.6811

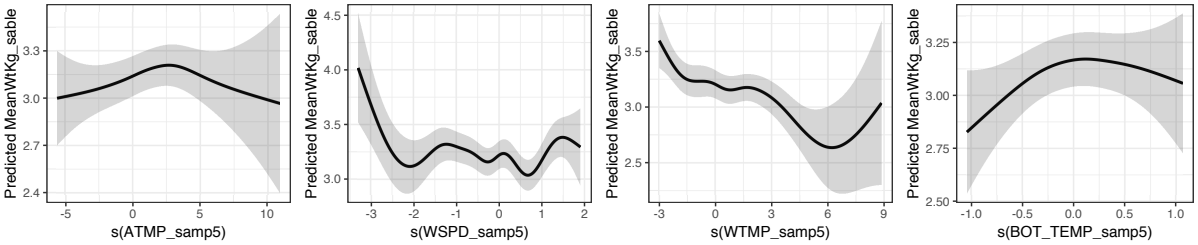


Figure A.1: Significant smooths in rank GAM for sablefish mean weight (kg).

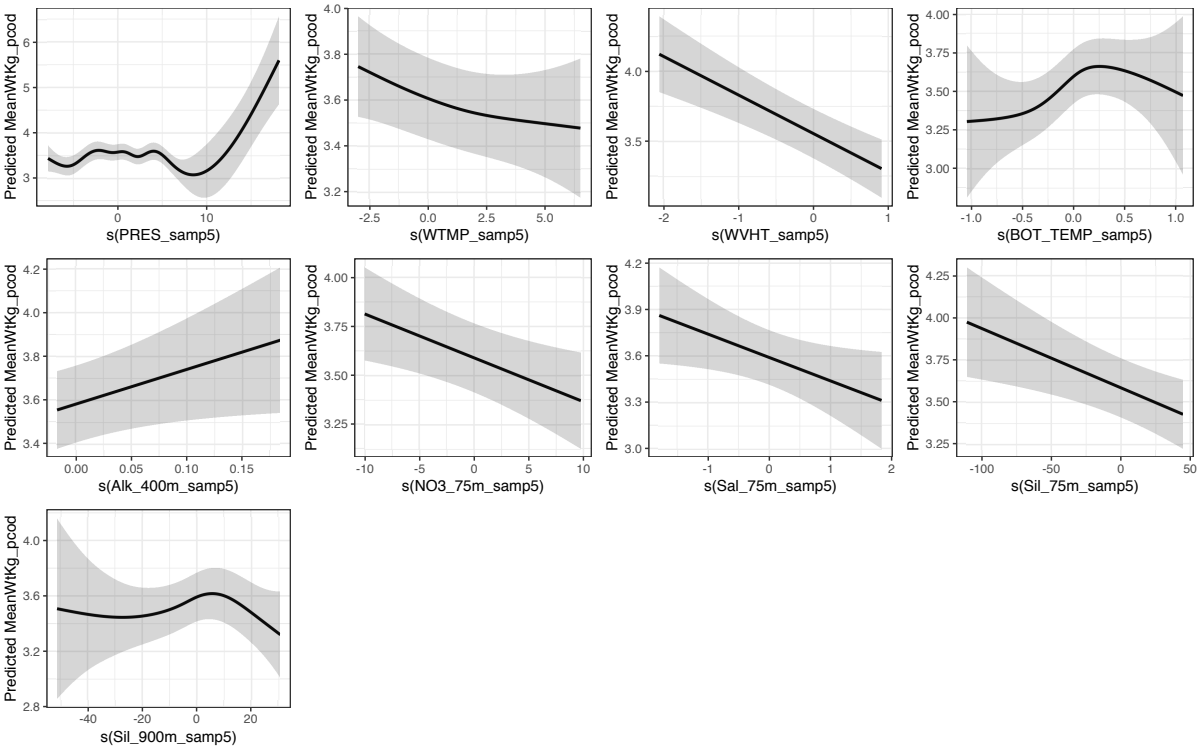


Figure A.2: Significant smooths in rank GAM for Pacific cod mean weight (kg).

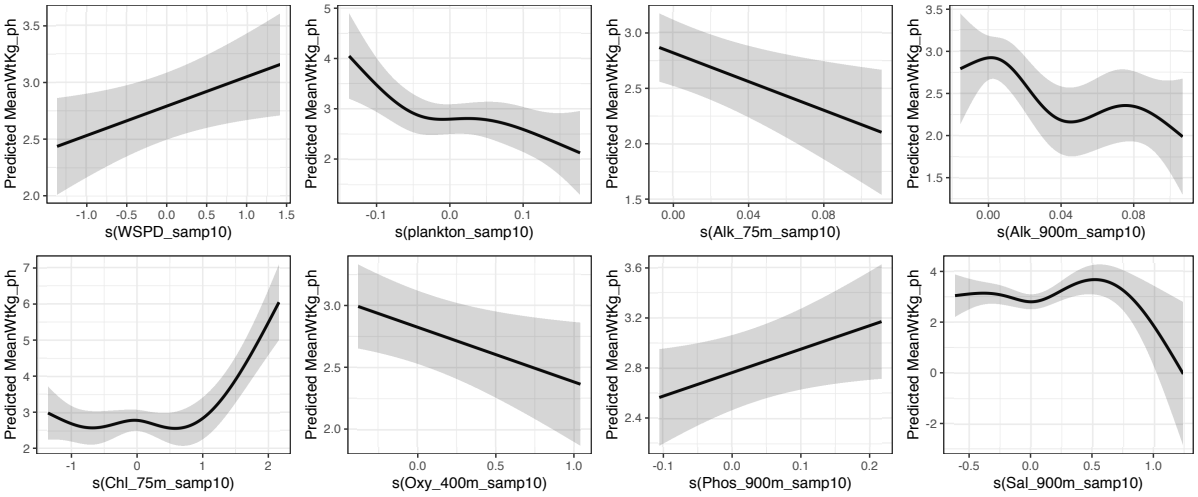


Figure A.3: Significant smooths in rank GAM for Pacific halibut mean weight (kg).

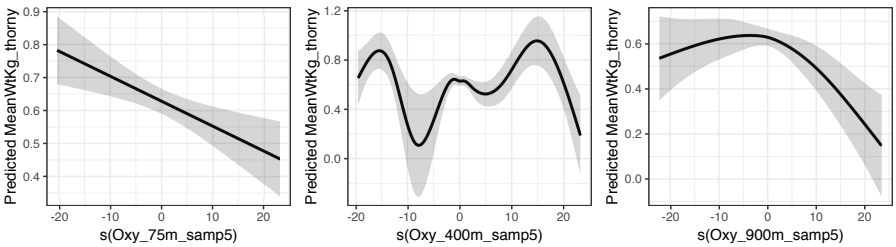


Figure A.4: Significant smooths in rank GAM for shortspine thornyhead mean weight (kg).

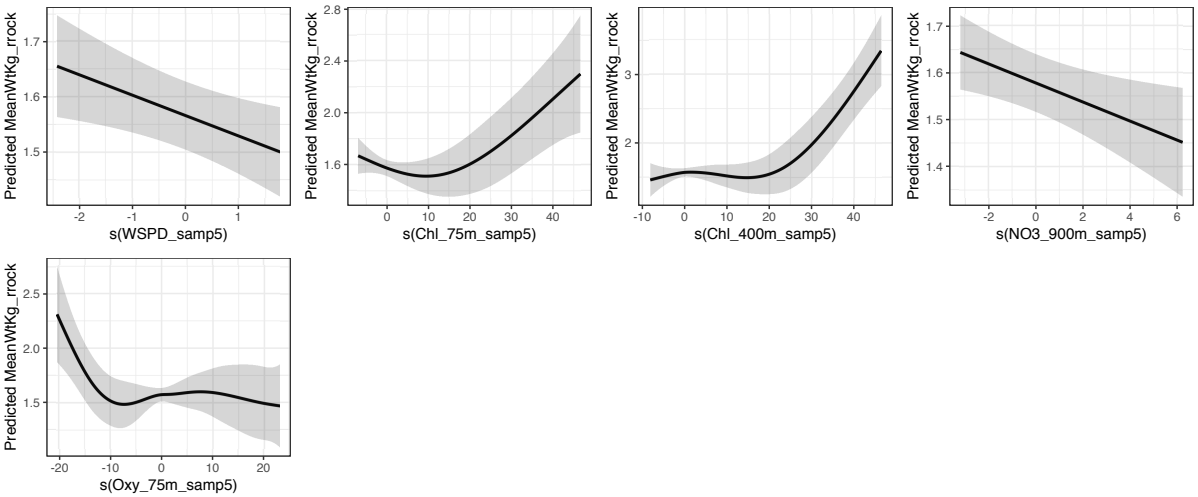


Figure A.5: Significant smooths in rank GAM for rougheye rockfish mean weight (kg).

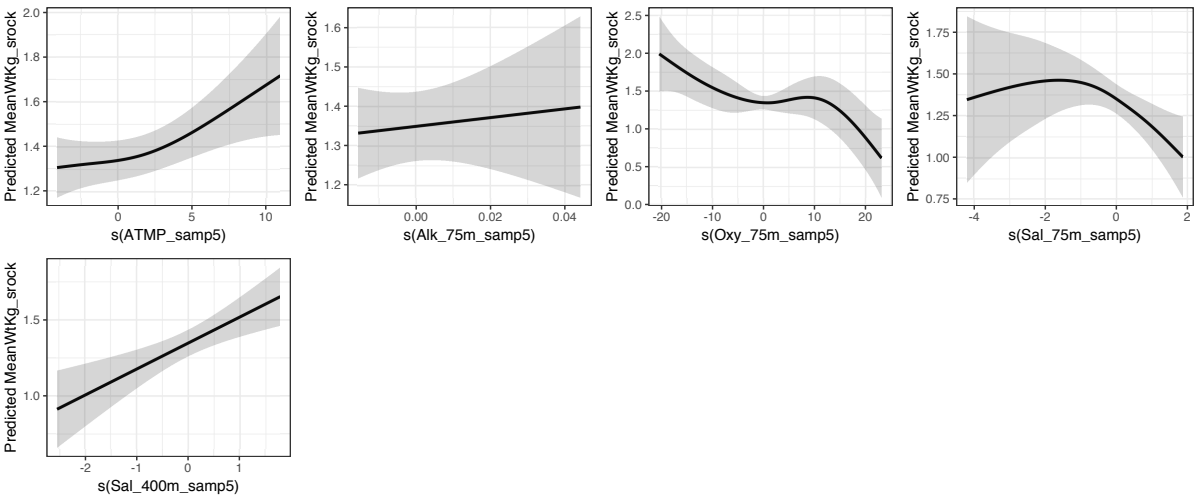


Figure A.6: Significant smooths in rank GAM for shortraker rockfish mean weight (kg).

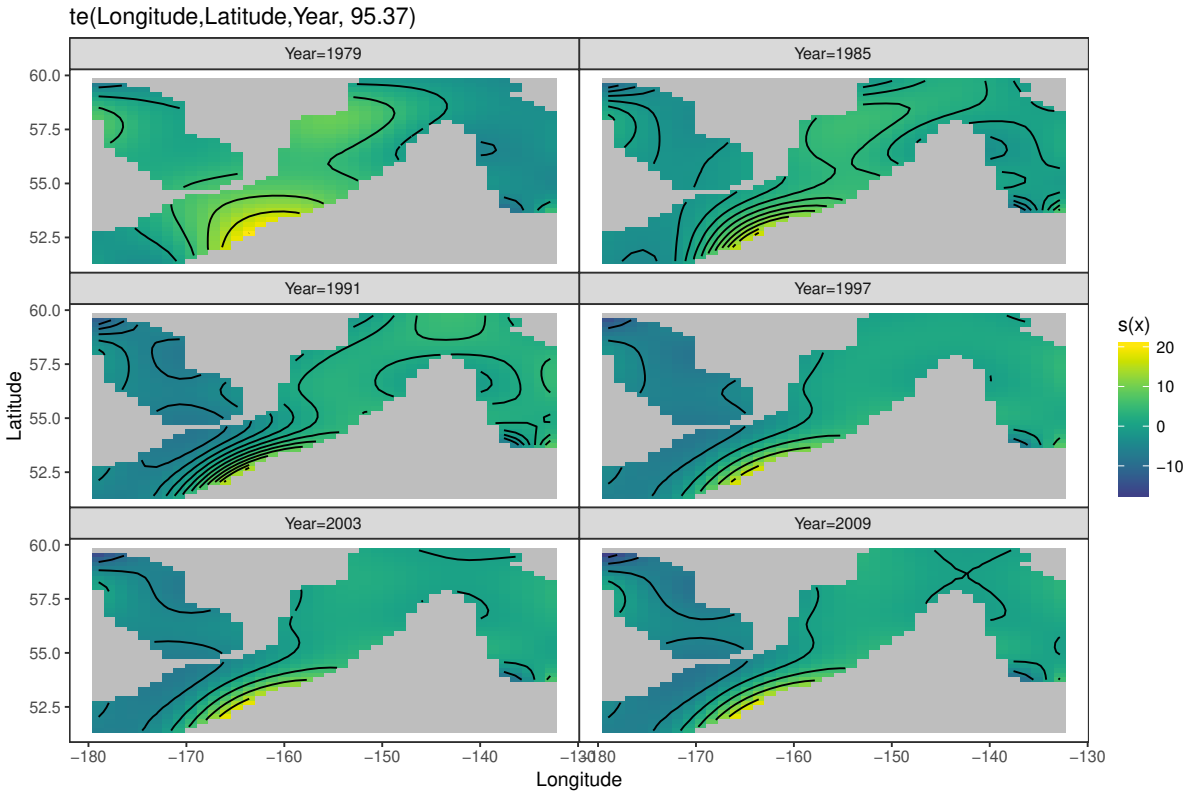


Figure A.7: Spatial distribution of sablefish CPUE for selected years between 1979 and 2013 from rank GAM.

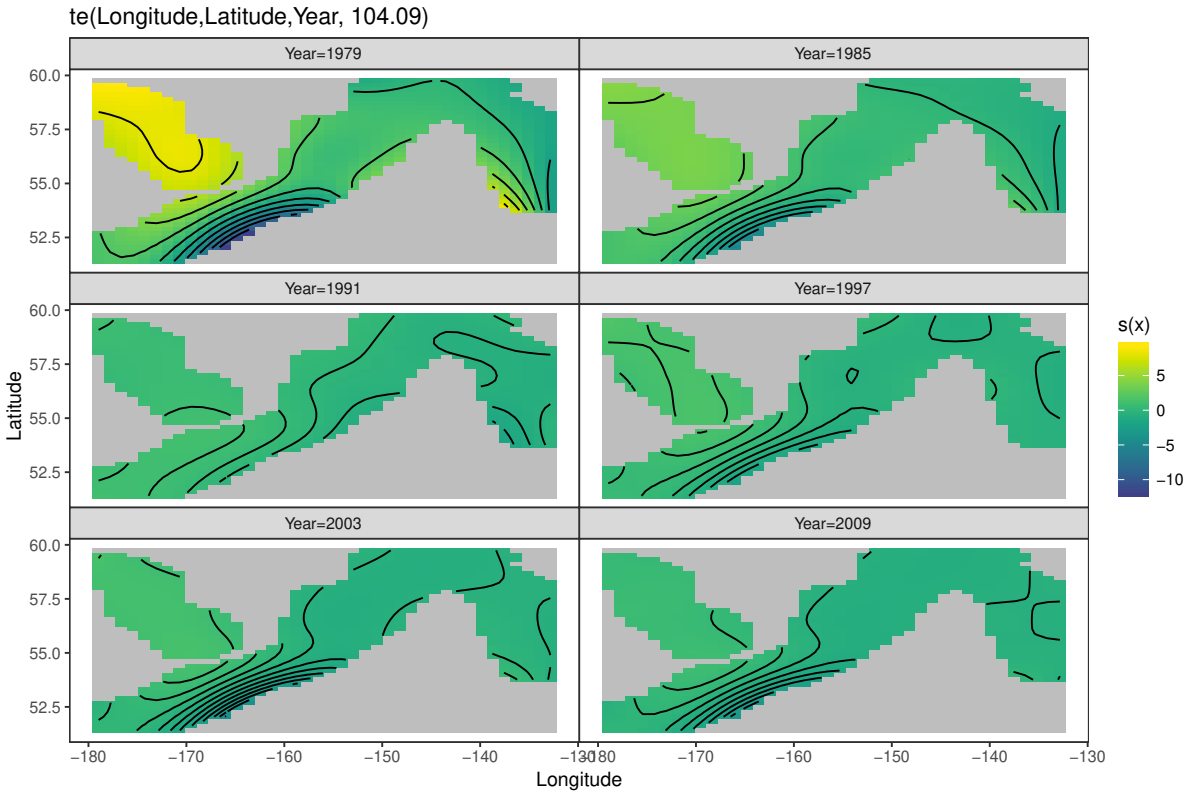


Figure A.8: Spatial distribution of Pacific cod CPUE for selected years between 1979 and 2013 from rank GAM.

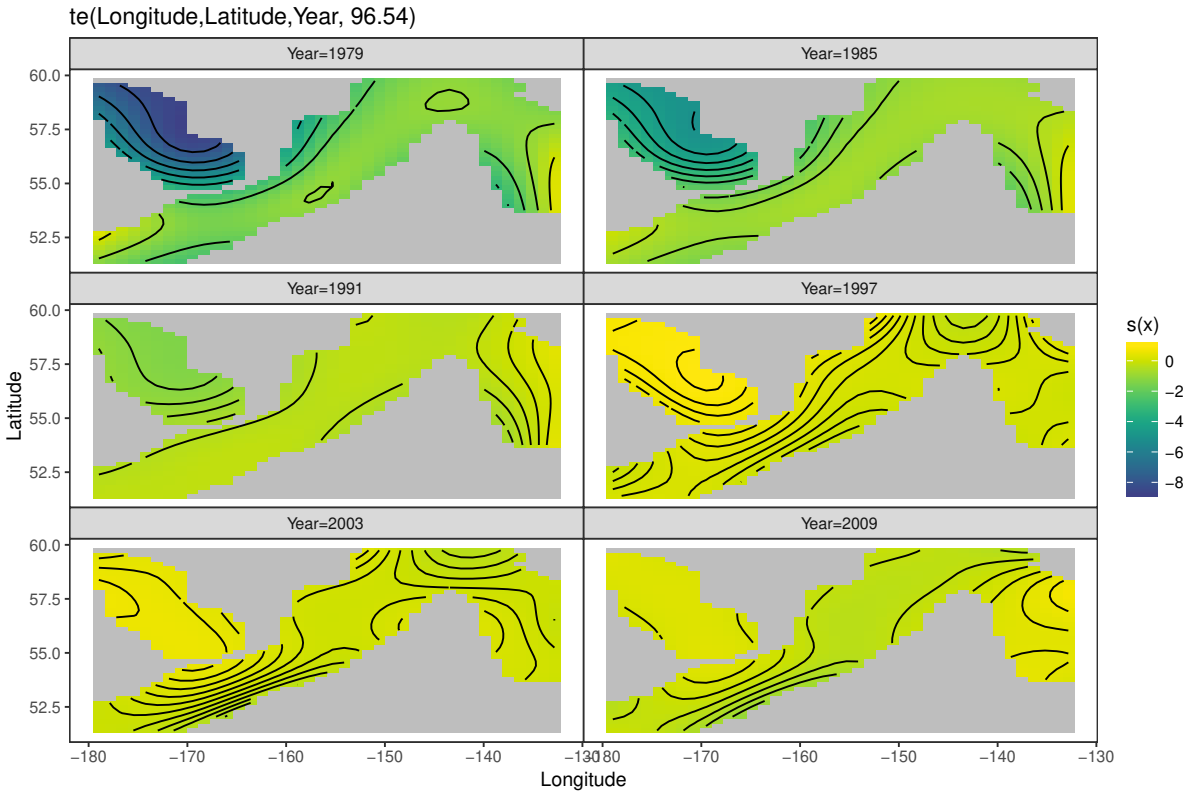


Figure A.9: Spatial distribution of Pacific halibut CPUE for selected years between 1979 and 2013 from rank GAM.

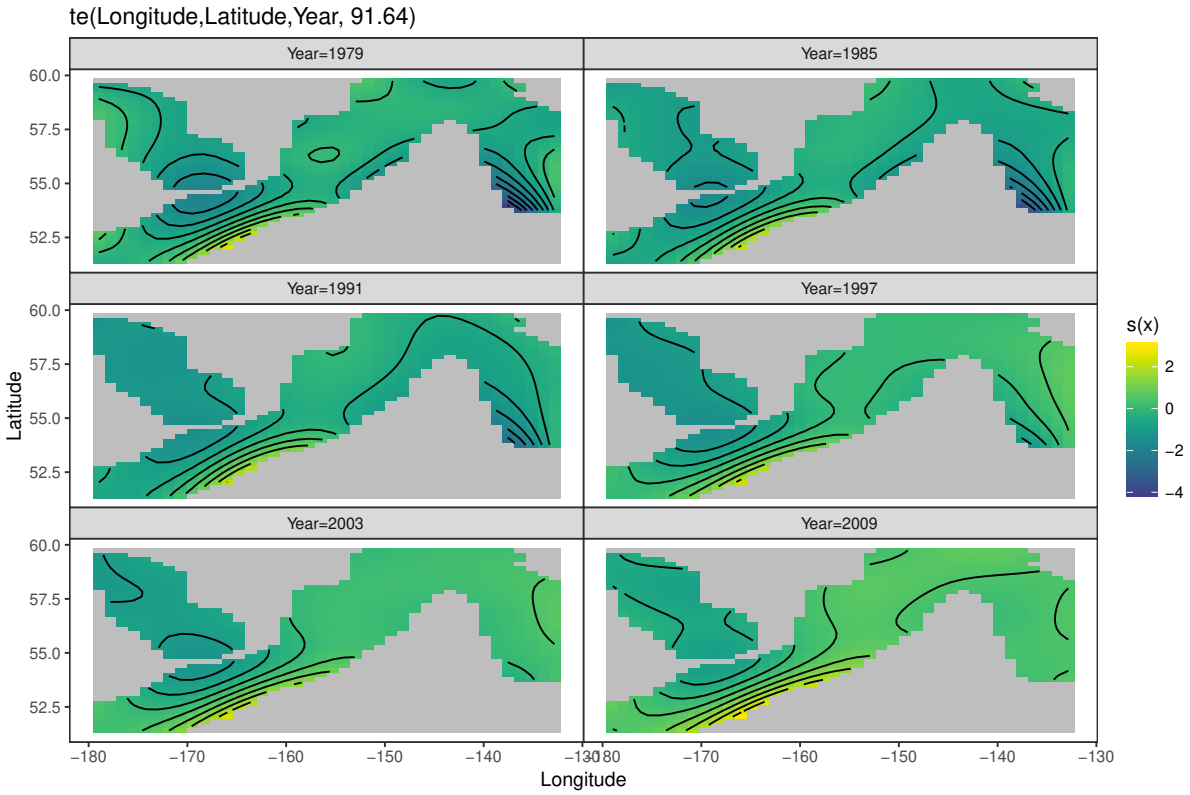


Figure A.10: Spatial distribution of shortspine thornyhead CPUE for selected years between 1979 and 2013 from rank GAM.

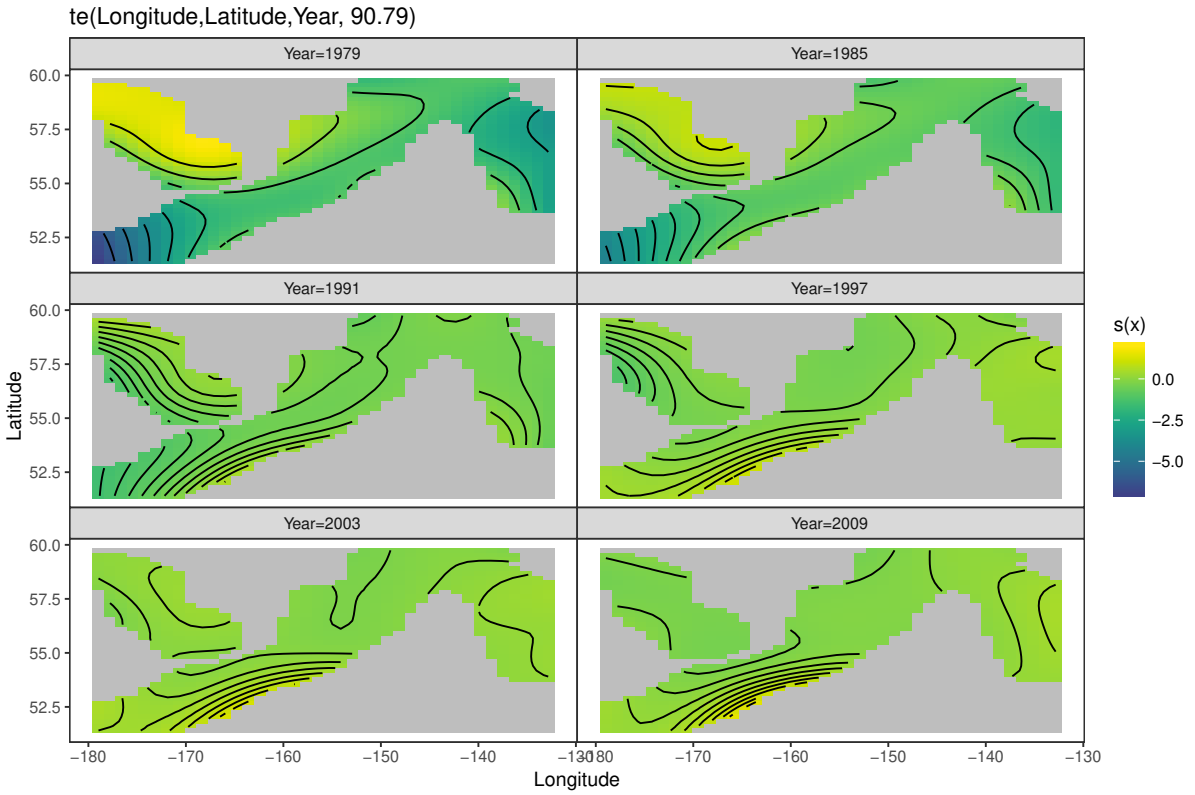


Figure A.11: Spatial distribution of rougheye rockfish CPUE for selected years between 1979 and 2013 from rank GAM.

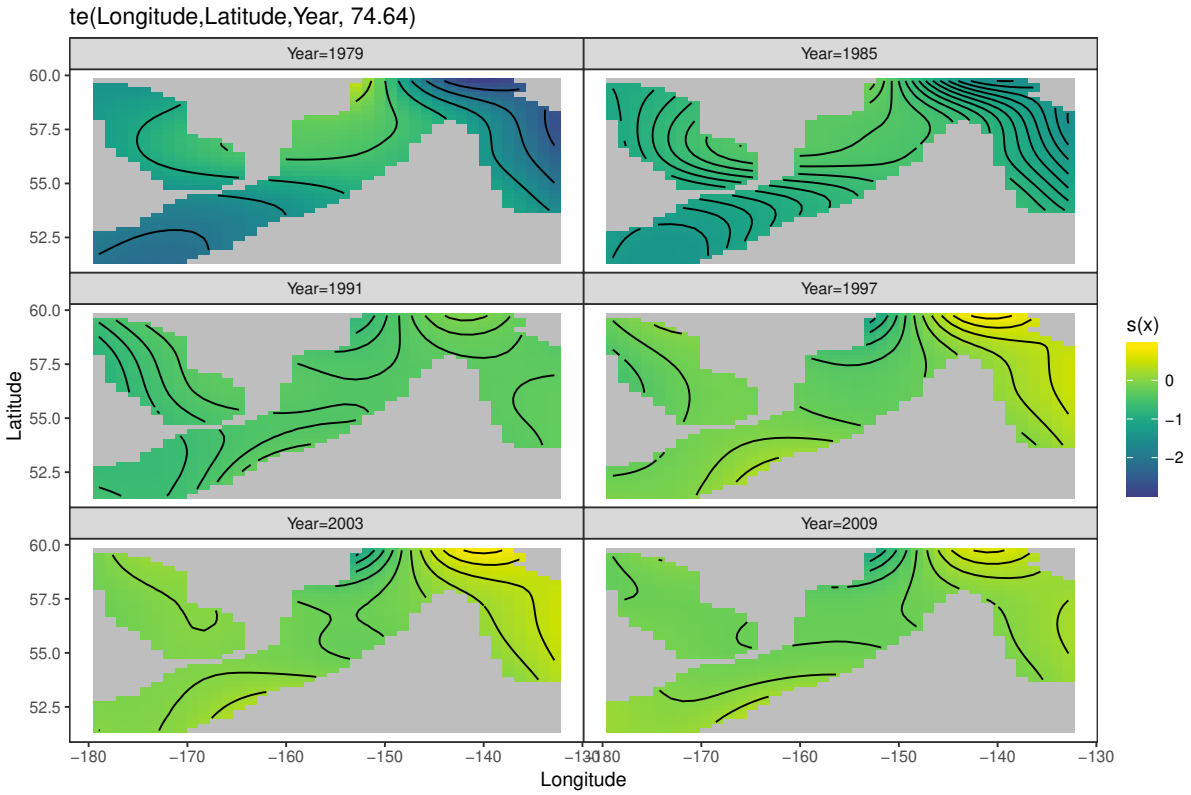


Figure A.12: Spatial distribution of shortraker rockfish CPUE for selected years between 1979 and 2013 from rank GAM.

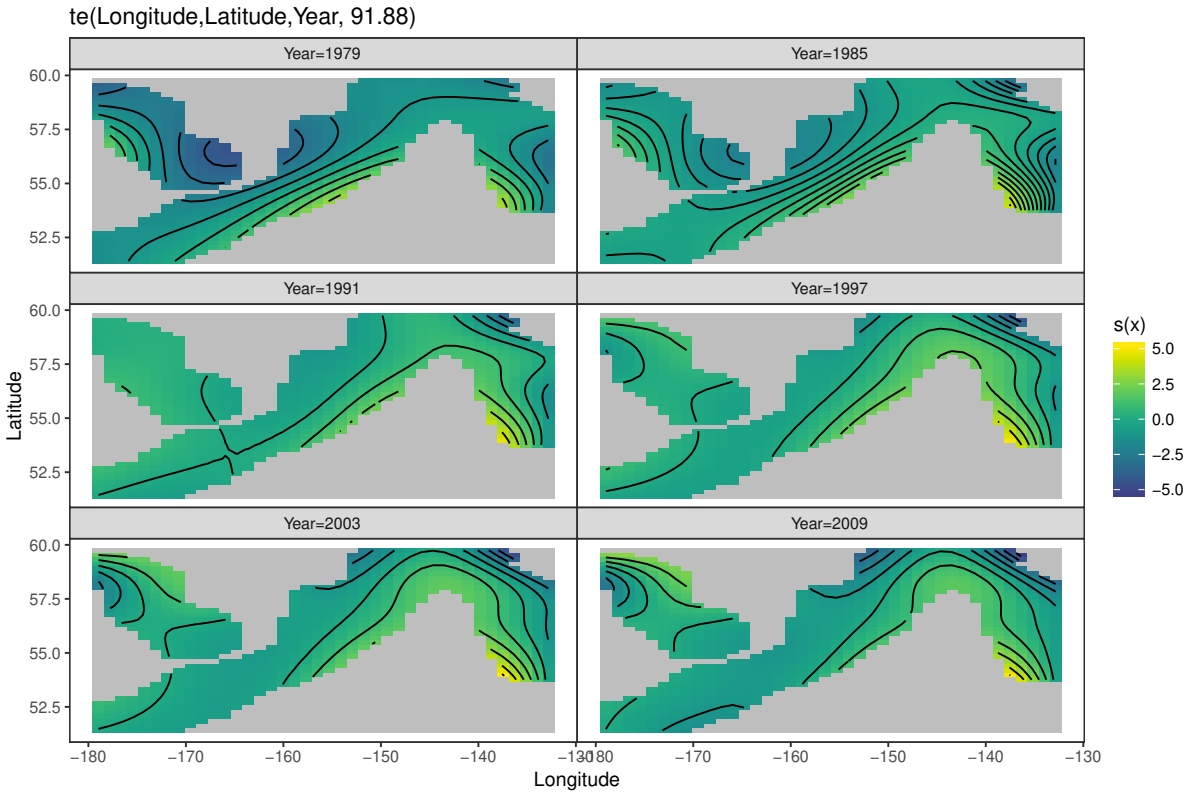


Figure A.13: Spatial distribution of sablefish mean weight (kg) for selected years between 1979 and 2013 from rank GAM.

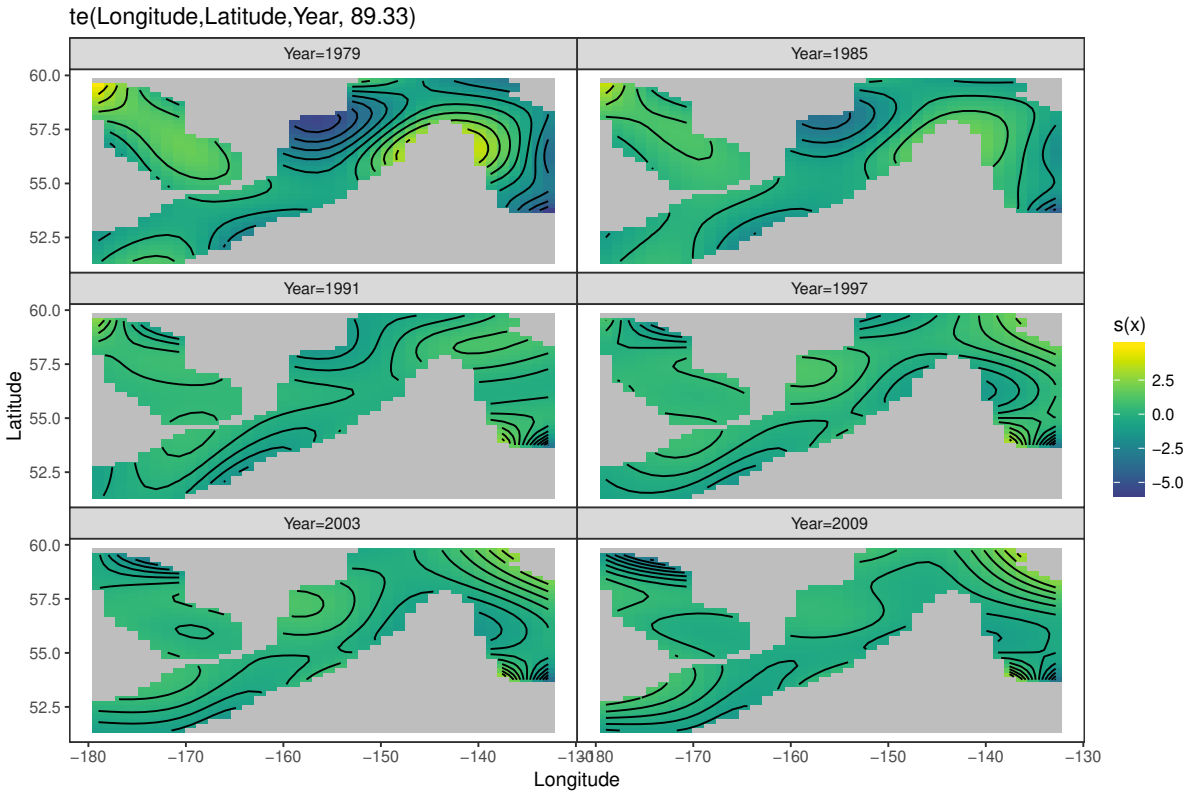


Figure A.14: Spatial distribution of Pacific cod mean weight (kg) for selected years between 1979 and 2013 from rank GAM.

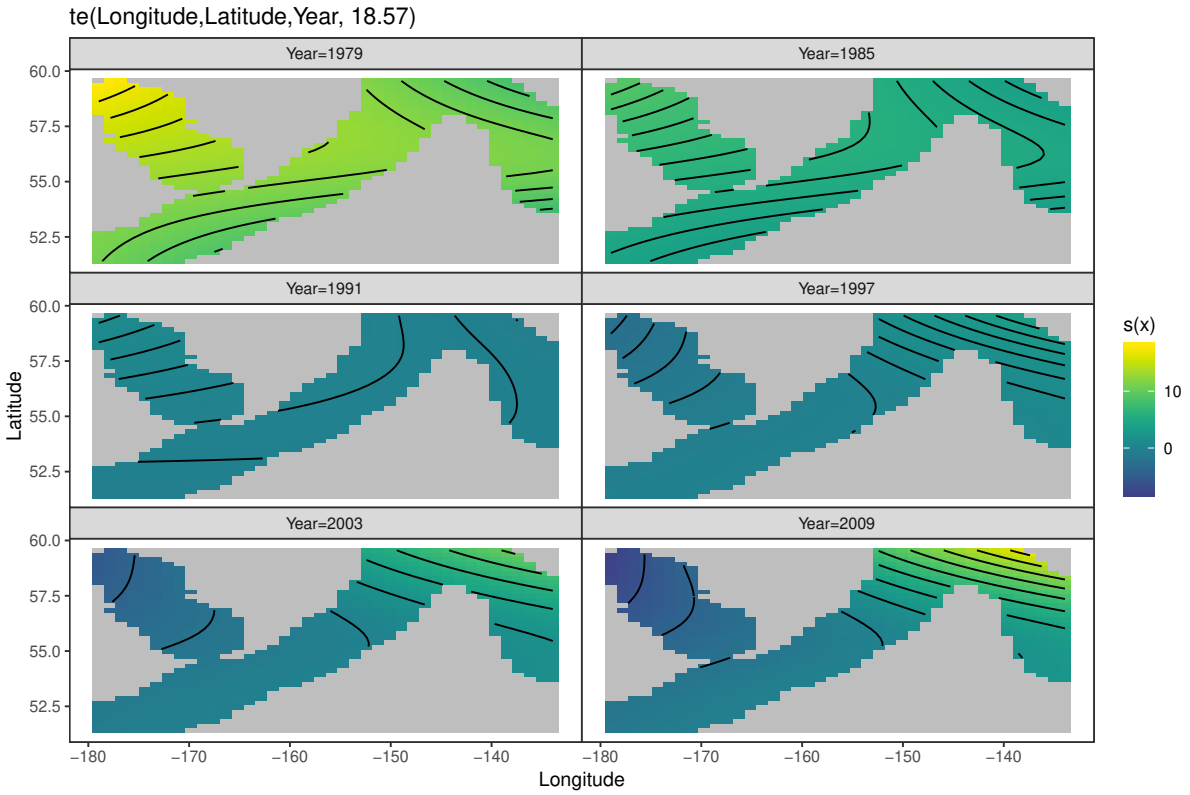


Figure A.15: Spatial distribution of Pacific halibut mean weight (kg) for selected years between 1979 and 2013 from rank GAM.

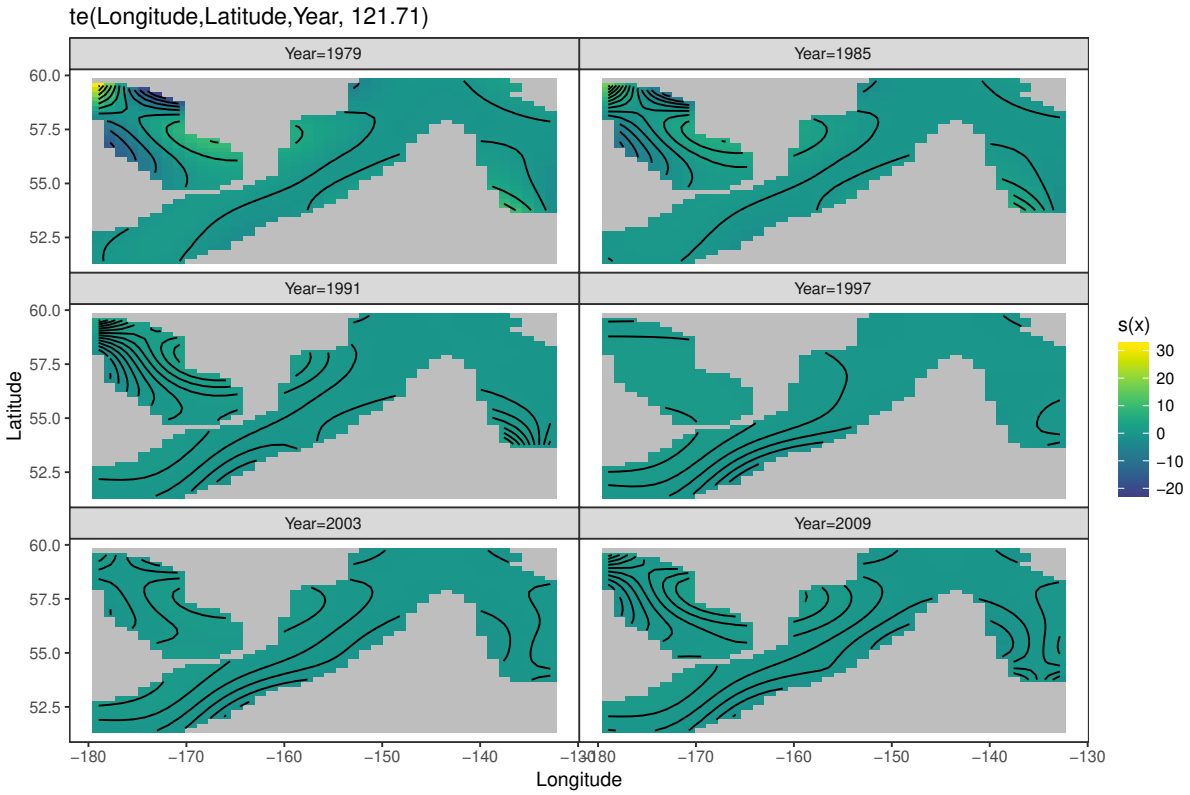


Figure A.16: Spatial distribution of shortspine thornyhead mean weight (kg) for selected years between 1979 and 2013 from rank GAM.

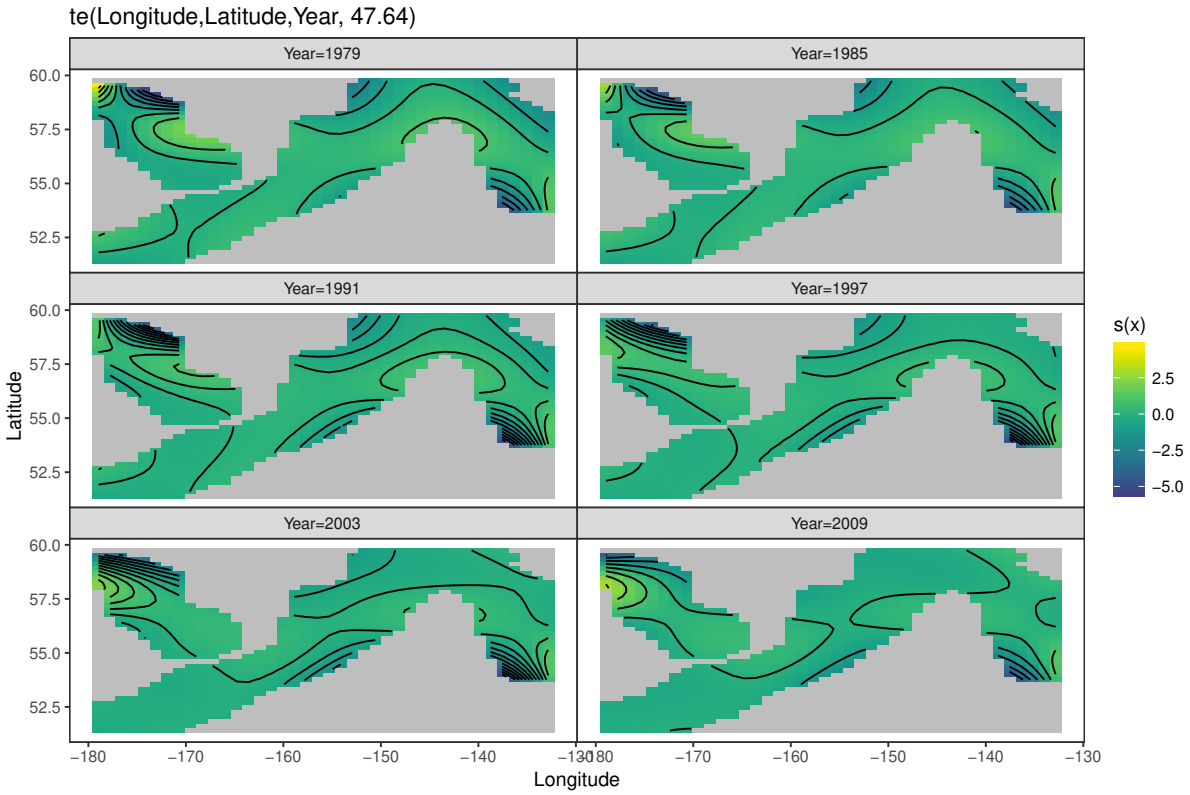


Figure A.17: Spatial distribution of rougheye rockfish mean weight (kg) for selected years between 1979 and 2013 from rank GAM.

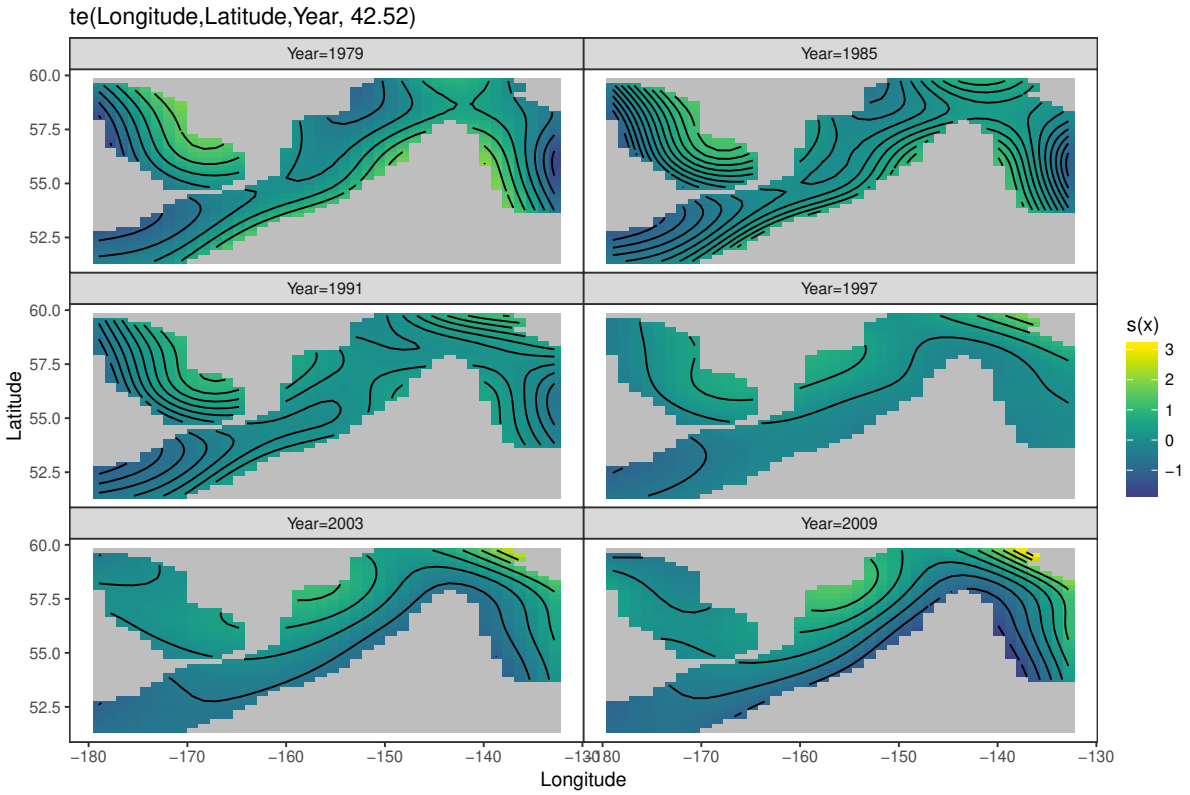


Figure A.18: Spatial distribution of shortraker rockfish mean weight (kg) for selected years between 1979 and 2013 from rank GAM.

Appendix B

Chapter 3 STEMA Supplementary Material

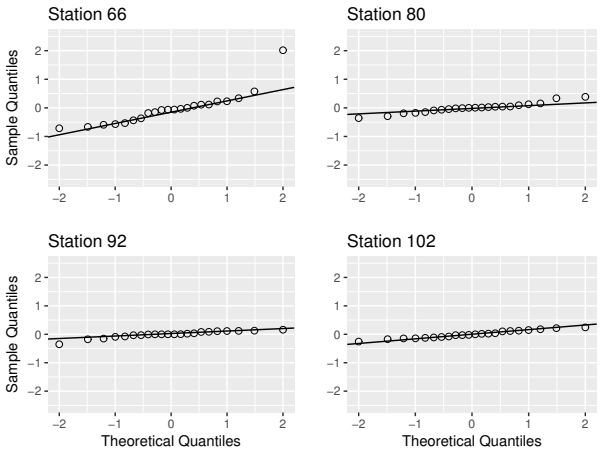


Figure B.1: Residuals for **Pacific cod** CPUE for given stations and years $1981 - 2011$ fit using the ARIMA model.

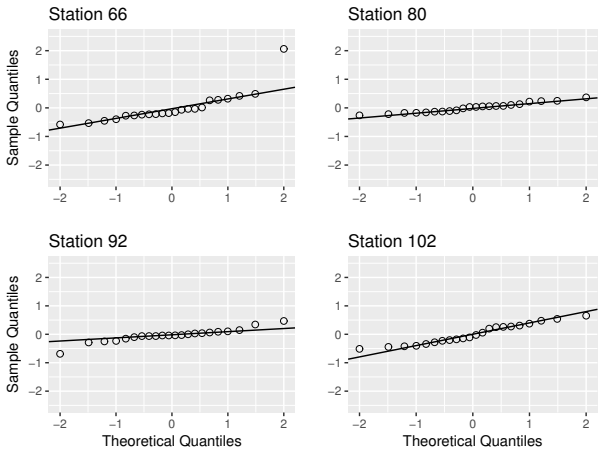


Figure B.2: Residuals for **Pacific halibut** CPUE for given stations and years $1981-2011$ fit using the ARIMA model.

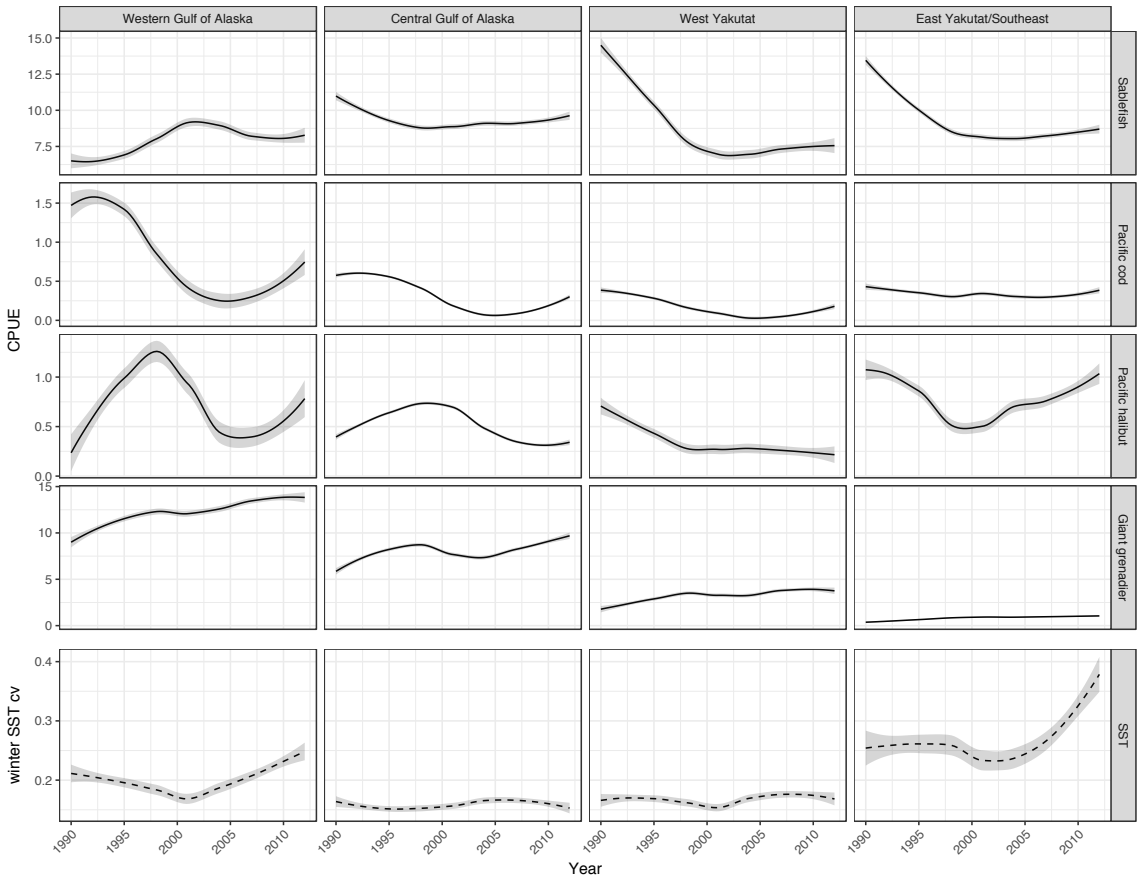


Figure B.3: Loess smooths of CPUE and winter SST by management area over time for each of the four species. Solid line is CPUE; dashed line is the coefficient of variation of winter SST; shaded regions are confidence intervals for each smooth.

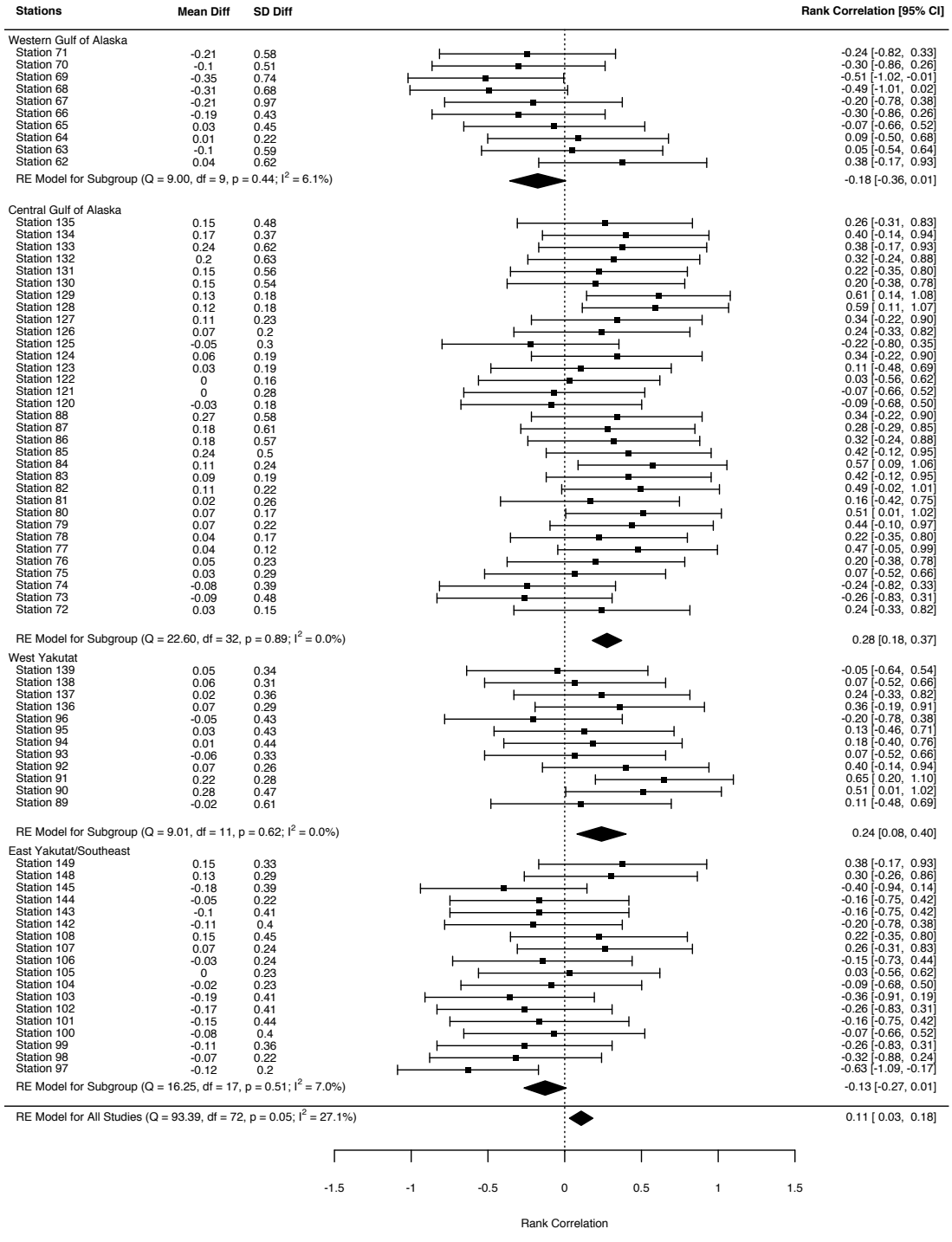


Figure B.4: Forest plot of SST effect on sablefish CPUE by station using the STEMA forecasting method. The rank correlation r statistic is given as the effect size.

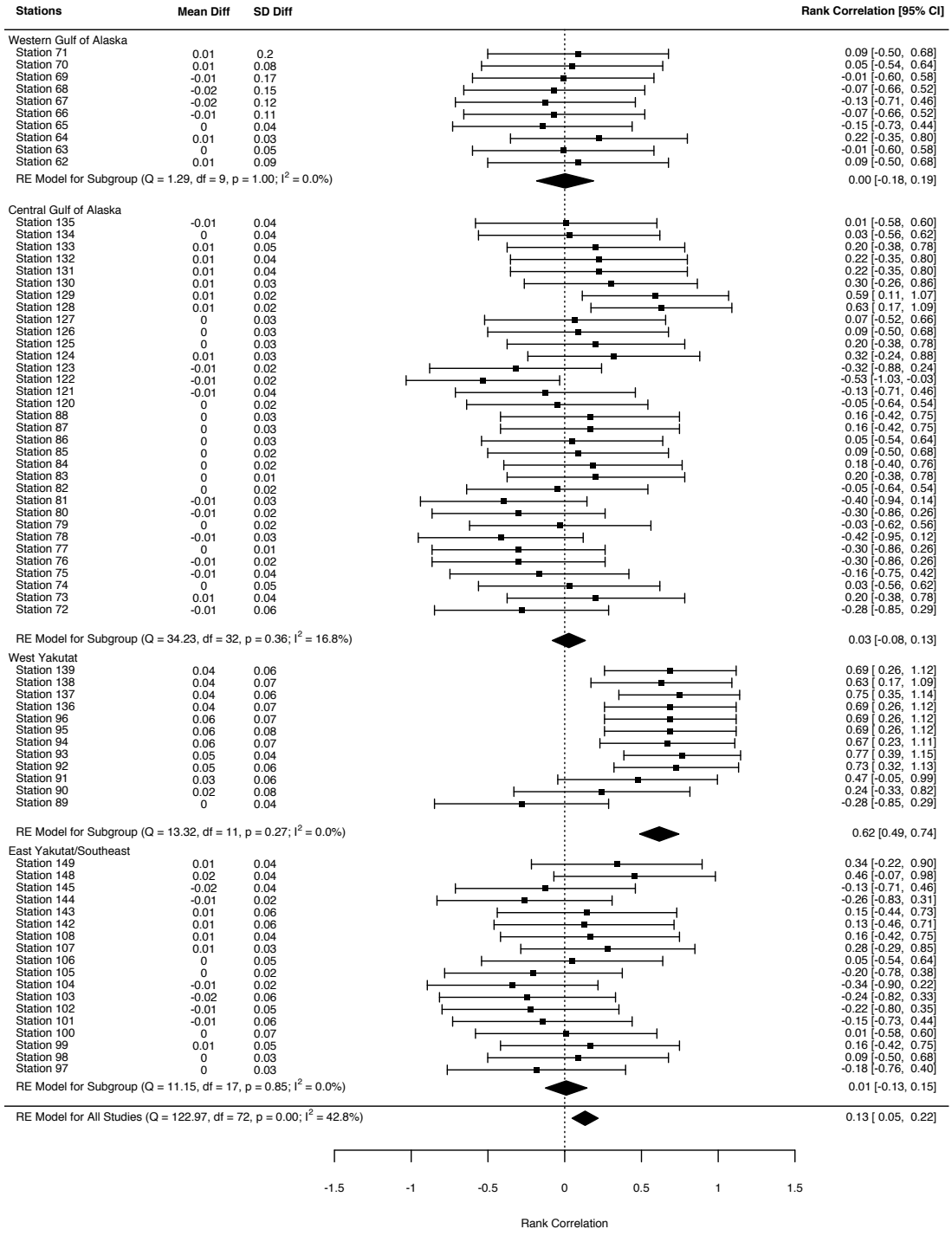


Figure B.5: Forest plot of SST effect on Pacific cod CPUE by station using the naïve forecasting method. The rank correlation r statistic is given as the effect size.

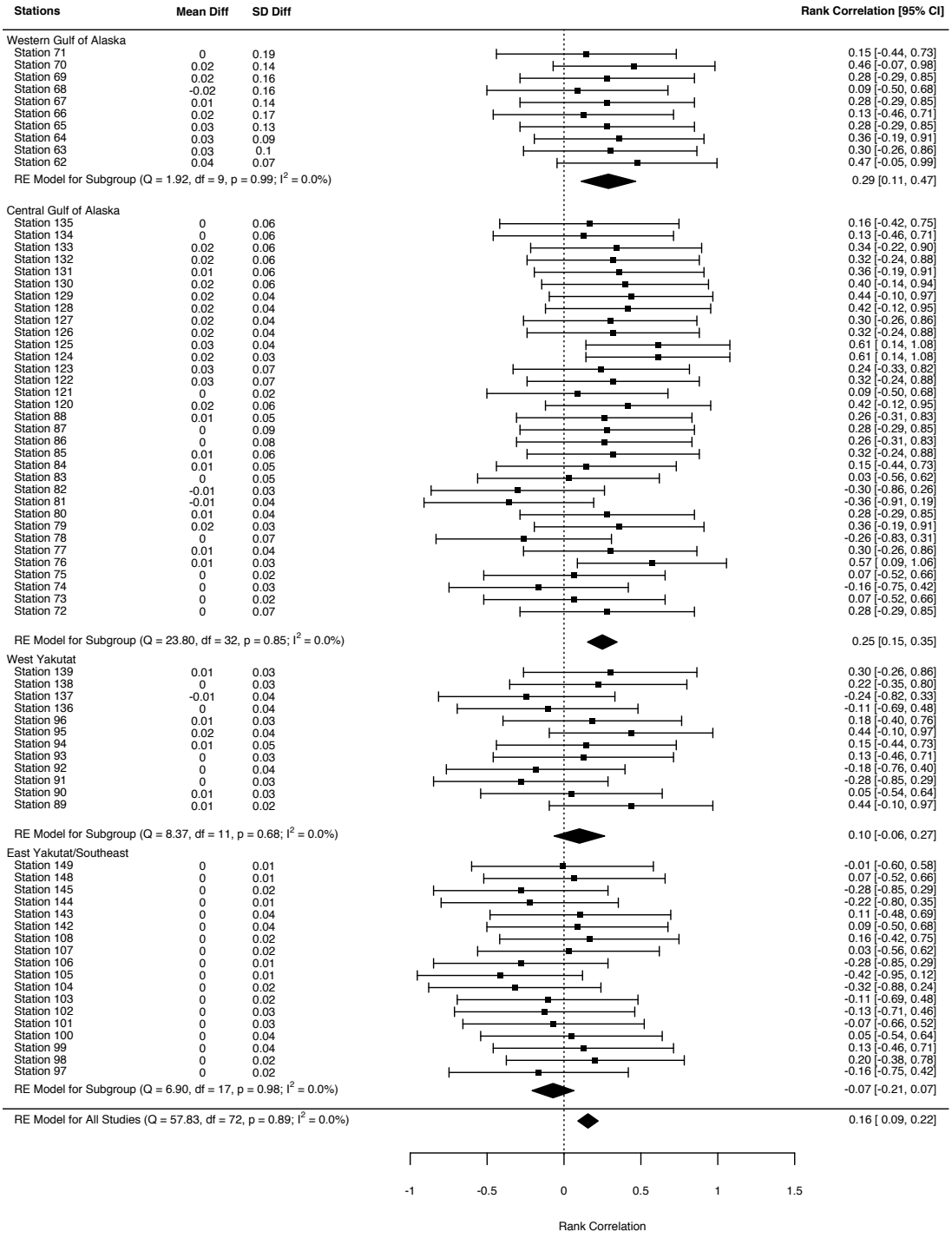


Figure B.6: Forest plot of SST effect on Pacific cod CPUE by station using the STEMA forecasting method. The rank correlation r statistic is given as the effect size.

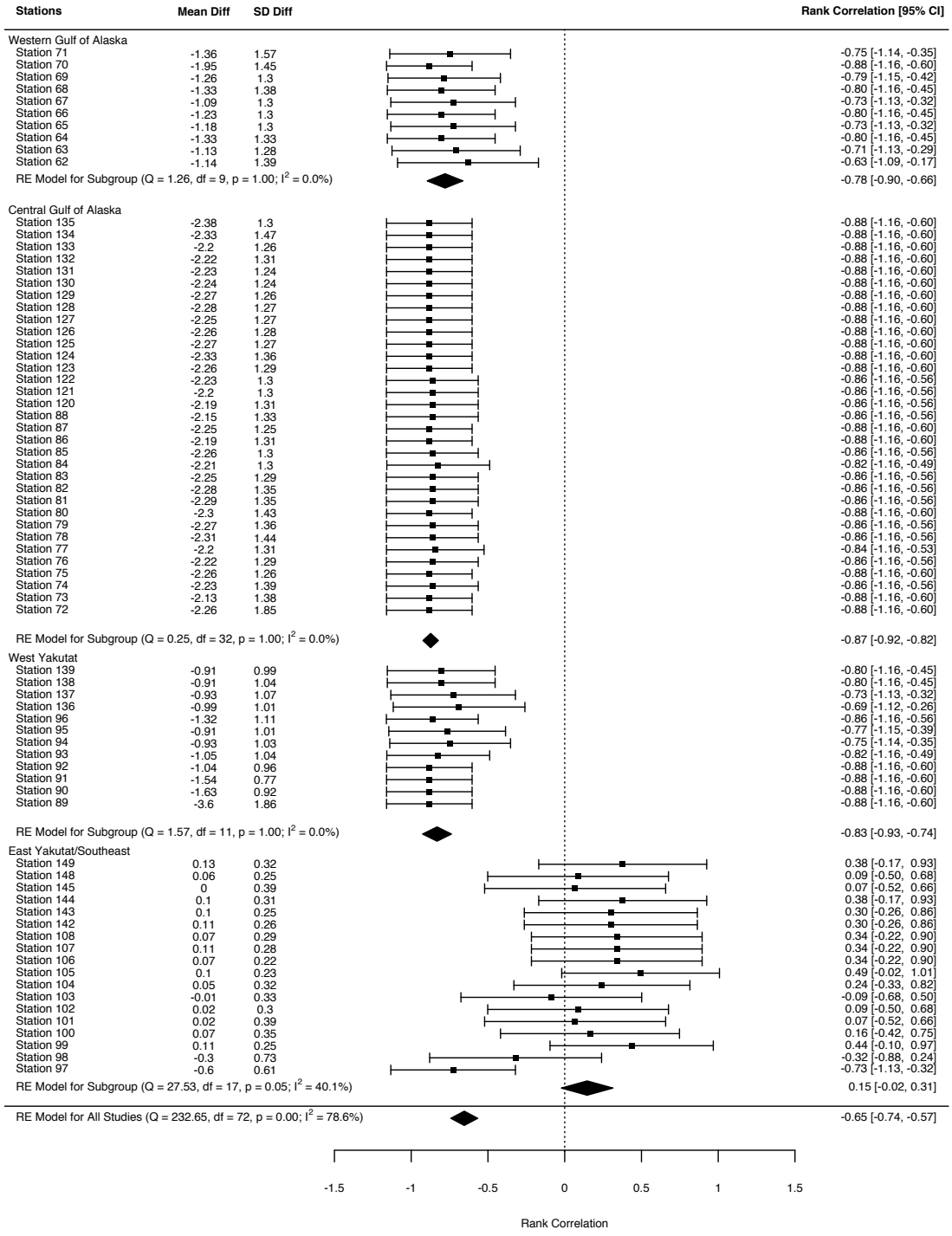


Figure B.7: Forest plot of SST effect on Pacific halibut CPUE by station using the naïve forecasting method. The rank correlation r statistic is given as the effect size.

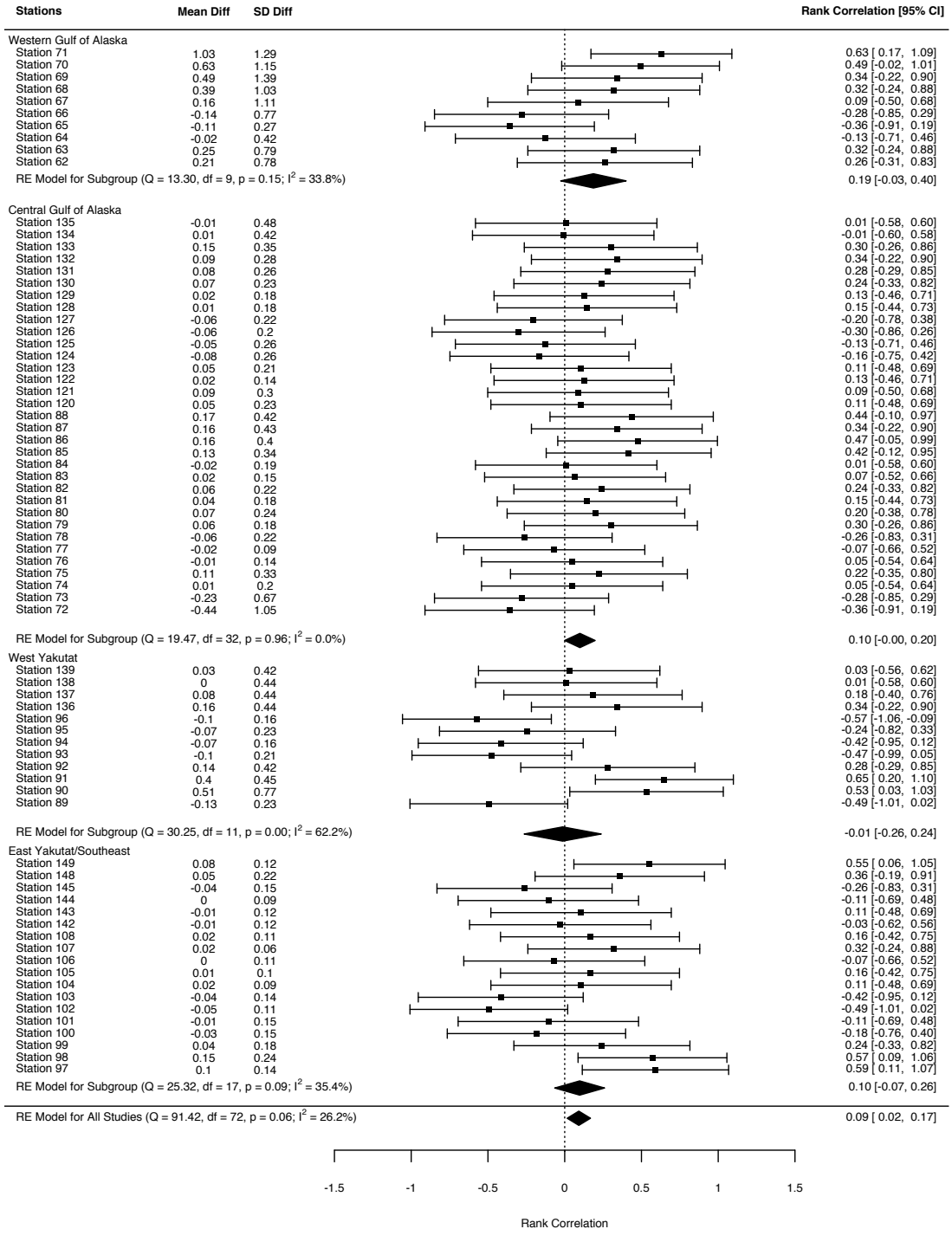


Figure B.8: Forest plot of SST effect on giant grenadier CPUE by station using the STEMA forecasting method. The rank correlation r statistic is given as the effect size.

Appendix C

Chapter 4 SIVCM Selection Supplementary Material

Table C.1: FDR-corrected p-values for relationship between selection of each environmental variable contributing to groundfish CPUE or mean weight each year and annual sign of PDO climate index fitted using GAMs.

	Pacific cod		Pacific halibut		Rougheye rockfish		Sablefish		Shortraker rockfish		Shortspine thornyhead	
Variable	CPUE	WT	CPUE	WT	CPUE	WT	CPUE	WT	CPUE	WT	CPUE	WT
ATMP	0.403	0.922	0.375	1.000	0.840	0.836	0.922	1.000	0.236	0.702	0.921	0.428
PRES	0.557	0.922	0.375	0.769	0.840	0.365	0.922	0.301	0.840	0.636	0.762	0.287
WSPD	0.869	0.922	0.826	1.000	0.840	0.365	0.922	1.000	0.279	0.636	0.921	0.633
WTMP	0.424	0.922	0.343	0.769	0.840	0.836	0.922	1.000	0.214	0.702	0.921	0.428
WVHT	0.822	0.922	0.537	0.769	0.840	1.000	0.922	1.000	0.279	0.702	0.762	0.287
plankton	0.557	0.922	0.343	0.769	0.840	0.365	0.922	0.301	0.214	0.636	0.493	0.287
BOT_TEMP	0.557	0.922	0.343	1.000	0.840	0.836	0.922	1.000	0.214	0.627	0.614	0.287
Alk_75m	0.557	0.922	0.343	0.769	0.840	1.000	0.922	0.301	0.214	0.627	0.921	0.287
Alk_400m	0.557	0.922	0.343	0.769	0.840	0.365	0.922	1.000	0.214	0.702	0.614	0.287
Alk_900m	0.869	0.922	0.343	0.769	0.840	0.836	0.922	0.301	0.236	0.636	0.53	0.287
Chl_75m	0.557	0.922	0.343	1.000	0.840	0.365	0.922	1.000	0.214	0.627	0.614	0.287
Chl_400m	0.897	0.922	0.343	0.769	0.840	0.365	0.922	0.301	0.214	0.627	0.762	0.287
Chl_900m	0.869	0.922	0.343	0.769	0.840	0.365	0.922	0.301	0.214	0.636	0.53	0.287
NO3_75m	0.557	0.922	0.375	0.769	0.840	0.365	0.922	0.301	0.236	0.627	0.921	0.287
NO3_900m	0.869	0.922	0.343	0.769	0.840	0.365	0.922	1.000	0.236	0.627	0.614	0.428
Oxy_75m	1.000	0.922	0.343	0.769	0.840	0.836	0.922	0.301	0.214	0.627	0.614	0.287
Oxy_400m	0.869	0.922	0.343	0.769	0.840	1.000	0.922	0.301	0.236	0.627	0.53	0.287
Oxy_900m	0.822	0.922	0.375	0.769	0.840	1.000	0.922	0.301	0.236	0.627	0.614	0.287
Phos_75m	0.869	0.922	0.343	0.769	0.840	0.365	1.000	0.301	0.214	0.702	1.000	0.287
Phos_900m	0.424	0.922	0.375	1.000	0.840	0.365	0.922	1.000	0.214	0.627	0.762	0.878
Sal_75m	0.869	0.922	0.375	NA	0.840	0.365	0.922	1.000	0.214	0.627	0.53	0.287
Sal_400m	0.557	1.000	0.343	0.769	0.840	0.365	0.922	1.000	0.214	0.627	0.614	0.633
Sal_900m	0.869	0.922	0.343	1.000	0.840	0.365	0.922	0.301	0.236	0.702	0.614	0.633
Sil_75m	0.557	0.922	0.343	0.769	0.840	0.365	0.922	0.301	0.279	0.636	0.614	0.287
Sil_900m	0.373	0.922	0.343	1.000	0.840	0.365	0.922	0.301	0.236	0.627	0.53	0.287

Table C.2: FDR-corrected p-values for relationship between selection of each environmental variable contributing to groundfish CPUE or mean weight each year and annual sign of MEI climate index fitted using GAMs. P-values < 0.1 are in bold.

	Pacific cod		Pacific halibut		Rougheye rockfish		Sablefish		Shortraker rockfish		Shortspine thornyhead	
Variable	CPUE	WT	CPUE	WT	CPUE	WT	CPUE	WT	CPUE	WT	CPUE	WT
ATMP	0.701	0.198	0.609	1.000	0.924	0.558	0.814	0.861	0.627	0.599	0.732	0.734
PRES	0.797	**0.060**	0.807	0.548	0.924	0.558	0.814	0.861	0.778	0.599	0.732	0.608
WSPD	0.576	0.198	0.609	0.426	0.924	0.558	0.814	0.861	0.627	0.599	0.732	0.608
WTMP	0.797	0.198	0.609	0.426	0.924	0.558	0.814	1.000	0.924	0.599	0.732	0.608
WVHT	0.576	0.572	0.609	0.548	0.924	1.000	0.886	1.000	0.766	0.599	0.732	0.630
plankton	0.576	**0.060**	0.807	0.426	0.924	0.558	0.814	0.861	0.627	0.675	0.732	0.608
BOT_TEMP	0.576	**0.060**	0.807	0.426	0.924	0.803	0.814	1.000	0.627	0.599	0.732	0.630
Alk_75m	0.576	**0.060**	0.807	0.426	0.924	1.000	0.814	0.861	0.627	0.599	0.732	0.630
Alk_400m	0.576	**0.060**	0.807	0.426	0.924	0.558	0.814	1.000	0.627	0.599	0.732	0.608
Alk_900m	0.576	**0.060**	0.807	0.426	0.924	0.558	0.814	0.861	0.627	0.675	0.732	0.608
Chl_75m	0.576	**0.060**	0.807	0.426	0.924	0.558	0.814	1.000	0.627	0.599	0.732	0.734
Chl_400m	0.576	**0.060**	0.609	0.426	0.924	0.558	0.814	0.861	0.627	0.599	0.732	0.734
Chl_900m	0.576	**0.060**	0.807	0.426	0.924	0.558	0.814	0.861	0.627	0.675	0.732	0.608
NO3_75m	0.179	0.198	0.609	0.548	0.924	0.558	0.814	0.861	0.669	0.599	0.732	0.608
NO3_900m	0.266	0.198	0.807	0.426	0.924	0.558	0.814	1.000	0.766	0.599	0.732	0.734
Oxy_75m	0.410	**0.060**	0.822	0.426	0.924	0.558	0.814	0.861	0.627	0.599	0.732	0.608
Oxy_400m	0.266	**0.060**	0.822	0.426	0.924	1.000	0.814	0.861	0.766	0.599	0.732	0.608
Oxy_900m	0.576	0.646	0.807	0.426	0.924	1.000	0.814	0.861	0.627	0.599	0.732	0.608
Phos_75m	0.576	0.198	0.807	0.548	0.924	0.558	1.000	0.861	0.627	0.599	1.000	0.608
Phos_900m	0.576	0.198	0.807	0.426	0.924	0.558	0.814	1.000	0.627	0.599	0.732	0.608
Sal_75m	0.576	0.198	0.609	0.426	0.924	0.558	0.814	1.000	0.627	0.599	0.732	0.608
Sal_400m	0.797	1.000	0.807	0.426	0.924	0.558	0.814	1.000	0.627	0.599	0.732	0.608
Sal_900m	0.576	0.198	0.822	0.426	0.924	0.558	0.814	0.861	0.627	0.599	0.732	0.630
Sil_75m	0.576	0.198	0.609	0.426	0.924	0.558	0.814	0.861	0.669	0.675	0.732	0.608
Sil_900m	0.644	0.198	0.609	0.426	0.924	0.558	0.814	0.861	0.766	0.599	0.732	0.608

Table C.3: FDR-corrected p-values for relationship between selection of each environmental variable contributing to groundfish CPUE or mean weight each year and annual sign of NPGO climate index fitted using GAMs.

	Pacific cod		Pacific halibut		Rougheye rockfish		Sablefish		Shortraker rockfish		Shortspine thornyhead	
Variable	CPUE	WT	CPUE	WT	CPUE	WT	CPUE	WT	CPUE	WT	CPUE	WT
ATMP	0.823	0.180	0.283	1.000	0.755	0.248	0.835	1.000	0.682	0.292	0.955	0.891
PRES	0.869	0.180	0.800	0.548	0.858	0.248	0.299	0.301	0.858	0.172	0.936	0.891
WSPD	0.470	0.180	0.800	0.426	0.755	0.248	0.327	1.000	0.469	0.172	0.856	0.891
WTMP	0.869	0.333	0.426	0.426	0.755	0.957	0.299	1.000	0.469	0.292	0.856	0.891
WVHT	0.892	0.175	0.587	0.548	0.755	1.000	0.365	1.000	0.601	0.292	0.936	0.891
plankton	0.470	0.175	0.426	0.426	0.858	0.248	0.299	0.301	0.469	0.172	0.492	0.891
BOT_TEMP	0.748	0.175	0.426	0.426	0.964	0.248	0.365	1.000	0.469	0.115	0.492	0.891
Alk_75m	0.470	0.175	0.426	0.426	0.755	1.000	0.835	0.301	0.469	0.115	0.856	0.891
Alk_400m	0.470	0.175	0.426	0.426	0.755	0.248	0.299	1.000	0.469	0.292	0.492	0.891
Alk_900m	0.748	0.175	0.426	0.426	0.755	0.365	0.299	0.301	0.964	0.172	0.856	0.891
Chl_75m	0.470	0.175	0.426	0.426	0.755	0.248	0.365	1.000	0.469	0.115	0.492	0.907
Chl_400m	0.458	0.175	0.269	0.426	0.858	0.248	0.299	0.301	0.469	0.115	0.674	0.907
Chl_900m	0.458	0.180	0.426	0.426	0.755	0.248	0.299	0.301	0.469	0.172	0.856	0.891
NO3_75m	0.251	0.180	0.283	0.548	0.755	0.248	0.365	0.301	0.601	0.115	0.856	0.891
NO3_900m	0.166	0.180	0.426	0.426	0.755	0.248	0.299	1.000	0.575	0.115	0.492	0.891
Oxy_75m	0.251	0.175	0.587	0.426	0.858	0.365	0.299	0.301	0.469	0.115	0.492	0.891
Oxy_400m	0.166	0.180	0.587	0.426	0.755	1.000	0.299	0.301	0.575	0.115	0.856	0.891
Oxy_900m	0.458	0.175	0.800	0.426	0.755	1.000	0.835	0.301	0.469	0.115	0.674	0.891
Phos_75m	0.470	0.333	0.426	0.548	0.755	0.248	1.000	0.301	0.469	0.292	1.000	0.891
Phos_900m	0.470	0.180	0.800	0.426	0.755	0.248	0.365	1.000	0.469	0.115	0.674	0.891
Sal_75m	0.470	0.180	0.283	0.426	0.858	0.248	0.365	1.000	0.469	0.115	0.856	0.891
Sal_400m	0.869	1.000	0.426	0.426	0.858	0.248	0.365	1.000	0.469	0.115	0.492	0.891
Sal_900m	0.470	0.180	0.587	0.426	0.755	0.248	0.365	0.301	0.601	0.292	0.674	0.891
Sil_75m	0.470	0.333	0.269	0.426	0.755	0.248	0.299	0.301	0.469	0.172	0.492	0.891
Sil_900m	1.000	0.520	0.269	0.426	0.858	0.248	0.299	0.301	0.575	0.115	0.856	0.891