An Exploration of the Spatiotemporal Distribution of Snow Crab (*Chionoecetes opilio*) in the Eastern Bering Sea:

1982 – 2018

by

Bryna Michelle Mills

A Thesis Presented to the
FACULTY OF THE USC DORNSIFE COLLEGE OF LETTERS, ARTS AND SCIENCES
University of Southern California
In Partial Fulfillment of the
Requirements for the Degree
MASTER OF SCIENCE
(GEOGRAPHIC INFORMATION SCIENCE AND TECHNOLOGY)

May 2021

To Grandpa Conroy

Acknowledgements

Many thanks to my advisor, Dr. Bernstein, for guidance and support throughout the process of creating this document and for the feedback from Drs. Loyola and Sedano, who helped make it better. I am also grateful to all field biologists, survey technicians, and vessel crew who have reliably collected data for the eastern Bering Sea bottom trawl survey and other fisheries monitoring efforts in the region over decades, and the research councils and management agencies like the National Marine Fisheries Service for continuing to conduct the surveys and provide the data.

Table of Contents

List of Tables

List of Figures

Abbreviations

ADFG	Alaska Department of Fish and Game
AFSC	Alaska Fisheries Science Center
BSFEP	Bering Sea Fisheries Ecosystem Plan
CPUE	Catch per unit effort
CV	Coefficient of variation
EBM	Ecosystem-based management
EBS	Eastern Bering Sea
ERH	Environmental ratchet hypothesis
GAM	Generalized additive modeling
GIS	Geographic information system
GISci	Geographic information science
GLM	Generalized linear modeling
GLR	Generalized linear regression
GWR	Geographically weighted regression
ML	Machine learning
NMFS	National Marine Fisheries Service
NOAA	National Oceanic and Atmospheric Administration
NM	Nautical mile
NPFMC	North Pacific Fishery Management Council
OLS	Ordinary least squares
RACE	Resource Assessment and Conservation Engineering
RF	Random forest

SSI	Spatial Sciences Institute
TAC	Total allowable catch
USC	University of Southern California

Abstract

Snow crab, *Chionoecetes opilio*, is the largest commercial crab fishery in Alaska. Populations in the eastern Bering Sea have fluctuated over space and time, challenging statisticians attempting to model their distribution and predict stock trends to support sustainable management decisions. Climate change contributes to model uncertainty due to increased environmental variance and subsequent shifts in species assemblages adapting to changing conditions in the region. This research applied statistical toolkits and visualization techniques in GIS for spatiotemporal analysis of snow crab distribution in the eastern Bering Sea over thirty-seven years (1982 – 2018). The National Marine Fisheries Service standardized bottom trawl survey provided a robust dataset to statistically explore spatial and temporal patterns and relationships between snow crab abundance in terms of catch per unit of effort to sea temperatures, depth, and Pacific cod abundance. The temporal correlation in abundance patterns between snow crab year classes or cohorts was tested using exploratory regression and geographically weighted regression was used to visualize the nature and scale of relationships within the survey region. Overall spatial patterns of snow crab distribution in the eastern Bering Sea reflected large scale warming trends and contraction of the population to the north towards the Bering Strait. No significant relationship was found between snow crab and Pacific cod distributions on a global scale but there was evidence of a local scale inverse relationship in the southern survey region. In absence of favorable bottom temperatures in 2018, snow crab distribution displayed a greater depth dependence in the northernmost region. Temporal correlation was detected between age classes of snow crab, suggesting connectivity between maternal cohorts and progeny. These results identify local and global scale distribution trends which will support better predictive models for fisheries.

Chapter 1 Introduction

Snow crab, *Chionoecetes opilio (C. opilio)*, are widespread throughout the eastern Bering Sea (EBS) and are harvested in the largest commercial crab fishery in Alaska. Managers monitor the distribution and abundance of *C. opilio* and many other marine species of commercial and ecological significance in the region to prevent overfishing and maintain sustainable populations. Geographic Information Science (GIS) can be used to model these distributions spatially to support traditional stock assessments.

Climate change in the Bering Sea region has been imposing pressure on species' geographic ranges and the ecological structure of the EBS shelf habitat. Snow crab have retreated to the north with sea temperature rise and reduced sea ice (Orensanz 2004) while populations of groundfish species such as Pacific cod (*Gadus macrocephalus*) have increased (Windle et al. 2012; Kotwicki and Lauth 2013). The influx of predatorial gadids like Pacific cod further obfuscates the future of snow crab in the EBS with implications for both commercial fisheries. The significance of the impact of these ecological relationships has been measured and quantified in a variety of regression techniques with variable results to support predictive modeling of fisheries stock distributions. The goal of this project is to describe the spatiotemporal distribution of *C. opilio* distribution and abundance in relation to temperature, depth, and Pacific cod abundance in the EBS through GIS and geostatistical analysis, and in doing so, demonstrate how GIS can be applied towards marine ecology and fisheries science.

1.1. Study Area

The EBS shelf is a productive, sub-polar ecosystem supporting a diverse range of crab, flatfish, and groundfish fisheries. This region extends approximately 270 nautical miles (nm) seaward from the west coast of Alaska and breaks to the west near 200 m depth (Figure 1). The

main EBS shelf is relatively uniform in substrate and sea floor physiography, but rockier, heterogenous habitat is found along the shelf edge where mature snow crab tend to cluster (ADFG 2019). St. Lawrence Island (63°N 170°W) marks the northern entrance to the Bering Strait which connects the Bering to the Chukchi Sea and Arctic Ocean beyond. South of St. Lawrence is St. Matthew Island (60°N 172°W); continuing south the central region of the shelf near 57°N is flagged by the Pribilof Islands to the west shelf edge and Nunivak Island to the east nearer the coast. The Aleutian Islands form a southern border to the Bering Sea at about 54°N, extending from the mainland Alaska Peninsula and Bristol Bay region towards Kamchatka Peninsula and the east coast of Russia. St. Paul Island in the Pribilofs (57°N 170°W) is a main port for commercial snow crab deliveries and serves as a geographic reference point throughout this study.

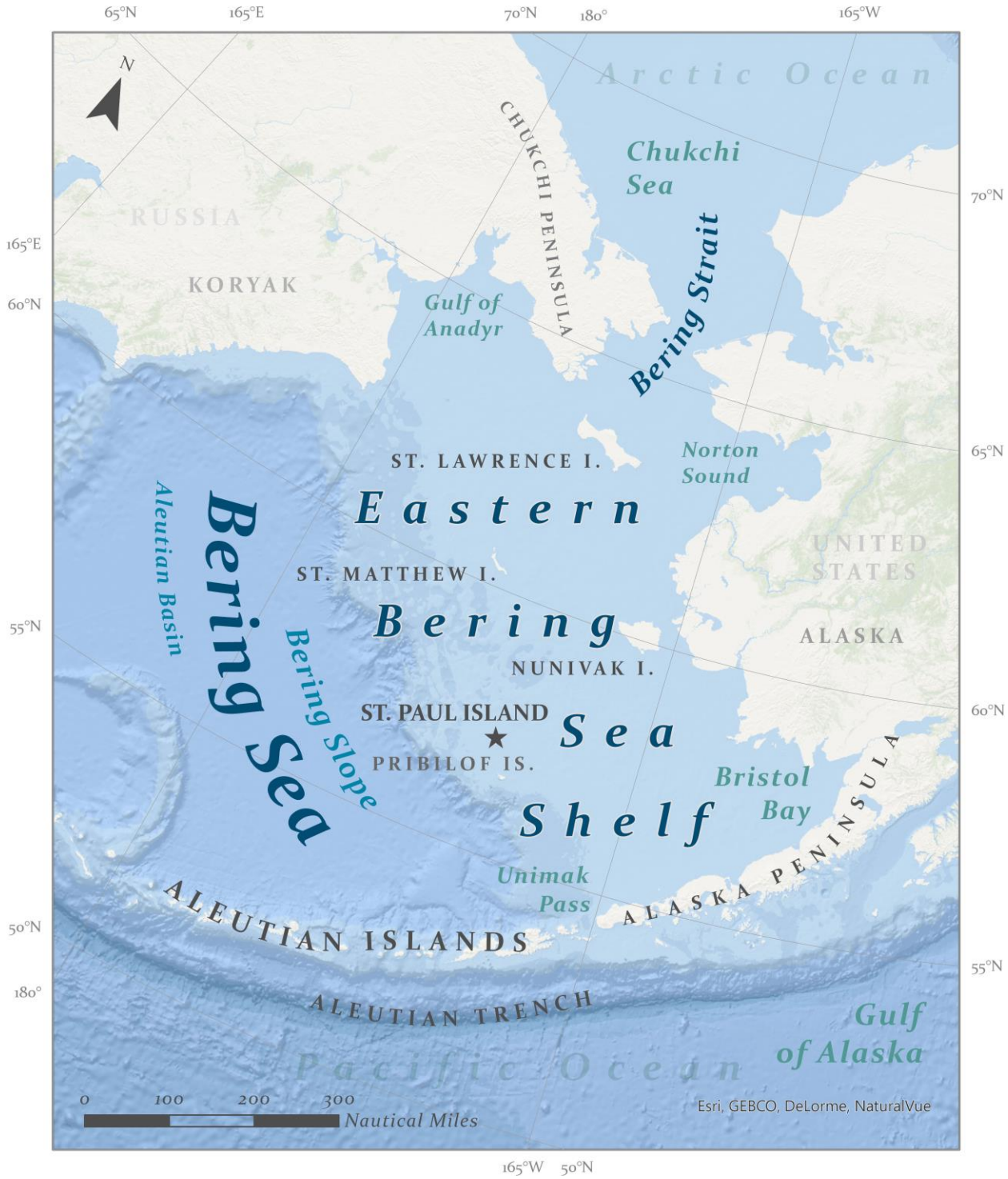


Figure 1. Study area, the eastern Bering Sea and EBS shelf

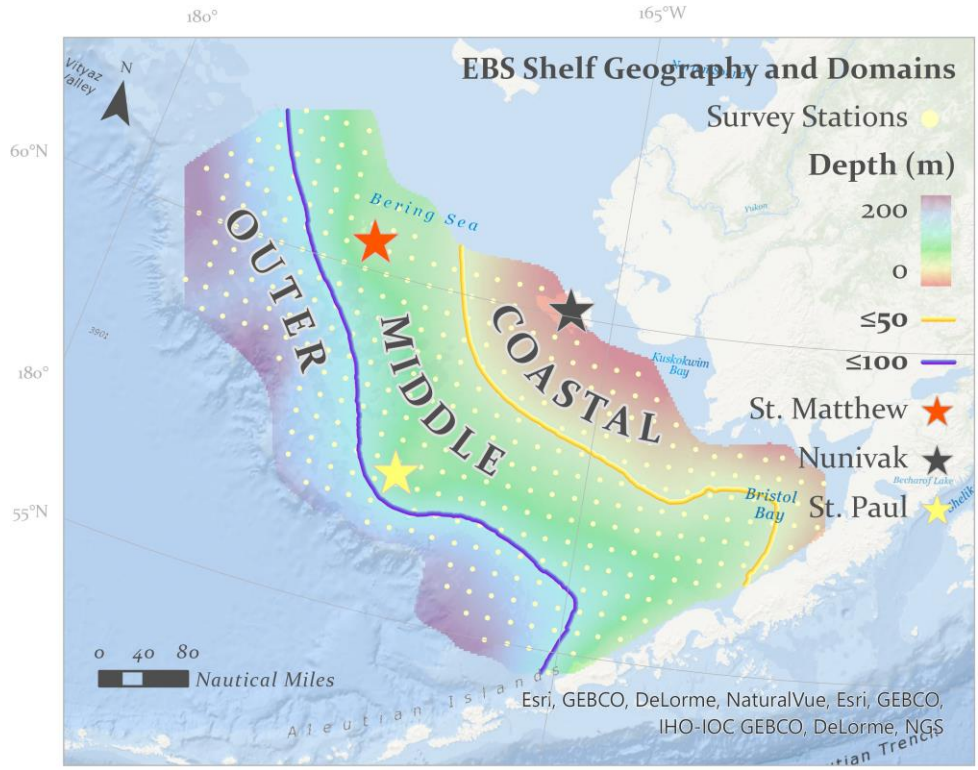
The Alaska Coastal Current is diverted into the Bering Sea from the Gulf of Alaska and Pacific Ocean. Much of the Coastal Current is directed through Unimak Pass, just west of the Alaska Peninsula, where it becomes the Bering Slope current as it continues north along the shelf's edge. Nutrients carried up from the Aleutian Trench along the south side of the island chain help to fuel a productive EBS ecosystem and form a productive front along the shelf edge where adult snow crab aggregate. The Pribilof Islands and St. Matthew Island divert flow from the Slope current. These island eddies provide a means of redistributing snow crab larvae and nutrients across the shelf as the main Slope current pushes north (Orensanz 2004; Parada et al. 2010).

Sea ice forms in the Bering Sea during winter months as polar currents from the Chukchi creep south over the shallow shelf region. Spring warming causes melt which sinks to the bottom forming a pool of colder bottom temperatures, typically under 2°C (NPFMC 2019). This cold pool (Appendix A) and temperature gradient that forms on the EBS shelf defines the ecosystem structure as it determines potential habitat for snow crab and other benthic marine species whose physiological function is adapted to specific thermal range limits (Molinos et al. 2018). Monitoring of temperature and climate tracking in the EBS is therefore vital to understanding patterns of species distributions and to anticipate ecosystem change scenarios in the future.

Ice can extend as far south as Bristol Bay and the Pribilof Islands in cold years, but sea ice formation and duration has decreased in recent years and the lowest recorded bottom temperature in the summer of 2018 was 1.6°C as reported by the National Marine Fisheries Service (NMFS) (NPFMC 2019). Managers are concerned the warming trends could have a detrimental impact on the snow crab fishery which may not be seen or detected for some years while the effects are borne out through the population life cycle (ADFG 2019). The Bering Sea

Fishery Ecosystem Plan (BSFEP) was formalized by the North Pacific Fishery Management Council (NPFMC) to begin development of ecosystem-based management (EBM) plans to supplement traditional fisheries stock assessments through studies that incorporate important variables like sea ice extent or sea temperatures along with spatially focused analyses of species distribution and relationships (Foy and Armistead 2012; NPFMC 2019). GIS enables integration, analysis, and visualization of spatiotemporal fisheries survey data and environmental variables of interest to better understand the ecological processes driving species distributions towards EBM goals.

Species abundance data gathered on standardized independent surveys designed by statisticians provide the bulk of the data used to model stocks for commercial fisheries. Since the 1970s, NMFS has conducted an annual bottom trawl survey to provide the necessary data for monitoring stocks and environmental conditions in the EBS. This trawl survey spans the shelf region (about 216,000 nm^2) from the Alaskan coast to the shelf edge as far west as 178°W and from the Alaska Peninsula north beyond St. Matthew Island to 62°N. The 50 and 100 m depth contours form the coastal, middle, and outer domains which describe the main geographic regions across the shelf (see Figure 2). Nunivak Island and St. Paul Island mark the central region of the shelf, defined by Zemchug Canyon to the north and Pribilof Canyon to the south.



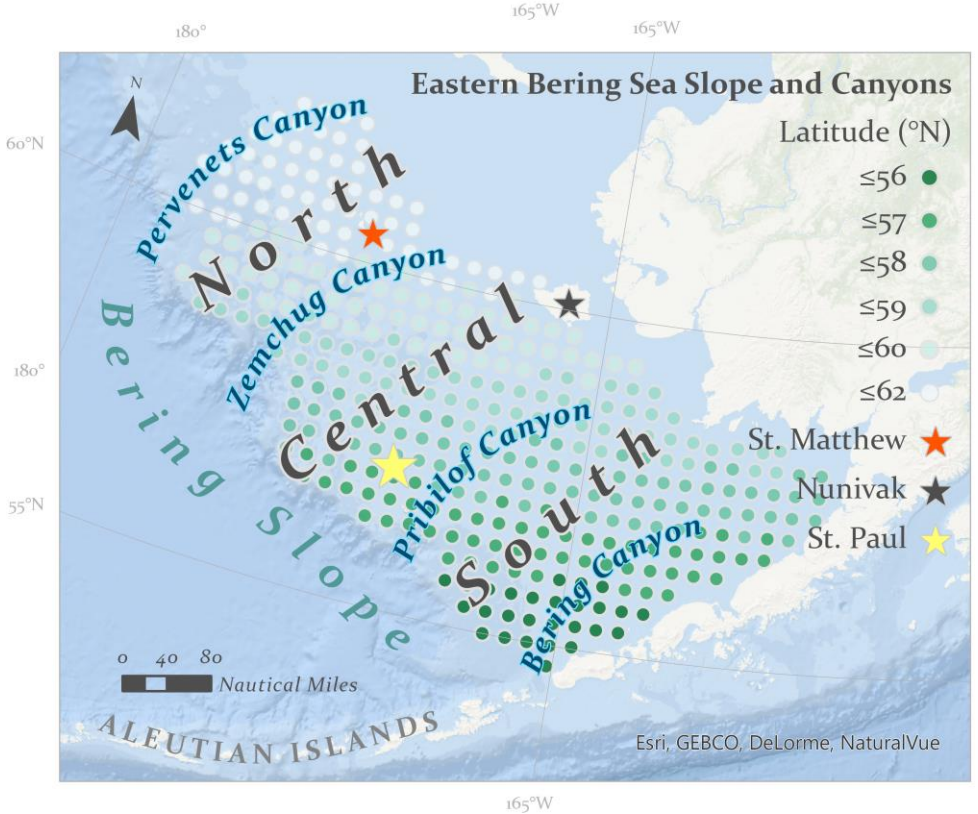


Figure 2. EBS survey coverage area with cross-shelf domains (top) and northern, central, and southern survey regions marked by canyons along the shelf edge (bottom)

1.2. Snow Crab Spatial Biology

Snow crab populations fluctuate in cyclical patterns where the frequency of pulse cycles of abundance reflects connectivity between maternal year classes (cohorts) and progeny year classes of immature snow crab (Ernst et al. 2005; Emond et al. 2015).

Different environmental conditions are preferred at each benthic life cycle stage of *C. opilio*, so that the population becomes spatially stratified according to age/sex demographics across the shelf. Snow crab begin their complex life cycle as larvae in the pelagic zone, transported by currents and subjected to prevailing surface temperatures for 3 to 5 months before settlement in the shallow and muddy coastal domain (Groβ et al. 2017). Immature snow crab migrate towards the middle domain, normally the coldest region of the EBS. Mature crab continue this migration towards the deeper outer domain and settle along the shelf edge in mature stages where sea temperature is typically warmer and reproductive energetics are more efficient (Orensanz et al. 2004).

Mature female age classes aggregate to the north of the main population and larger, commercially targeted males form dense patches along the shelf edge (Orensanz et al. 2004; Parada et al. 2010). Maternal cohorts release fertilized eggs in the outer domain near the edge; currents then carry the eggs and resulting pelagic larvae towards the shallower coastal domain where settlement occurs. Size and age frequency growth studies have shown that newly settled crab, or instars, take approximately four to six years of growth to reach the immature age class where it is large enough to be detected on survey, and reach maturity after another two to three years of growth. As immature age classes make up the largest proportion of the total population, this typically results in peaks in total abundance recurring every six to nine years. The contents of Pacific cod stomachs collected from EBS survey samples have shown that small and immature

snow crab are preferred prey and make up a substantial proportion of Pacific cod diet (Orensanz et al. 2004; Burgos et al. 2013; Groβ et al. 2017). This suggests that predation could be a major source of juvenile mortality and express a lagged detrimental impact on snow crab abundance, while the strength of maternal age classes would express a lagged positive correlation with future snow crab abundance and pulse cycles.

Many factors impact growth and survival of snow crab as age classes move across the shelf habitat in structured life history patterns, adapting to changing temperatures and species interactions. Spatial and temporal variability make it difficult to describe trends in snow crab abundance through global approaches to regression analysis alone (Ciannelli et al. 2008). Exploration of the temporal correlation based on life history characteristics and investigation of local scale relationships can help construct timelines of impact and describe regions where relationships may vary from the overall trends.

1.3. Fisheries Management

An ecosystem regime shift occurred in the eastern Bering Sea according to survey data in the late 1970s. Species assemblages and spatial distributions were shifting apparently in response to warming surface and bottom temperatures and the related decline in sea ice extent and duration. The temperature changes resulted in an influx of gadid fishes which began to tip the ecological balance of biomass away from sub-polar benthic invertebrates in favor of temperate groundfish species such as Pacific cod (*Gadus macrocephalus*) (Orensanz et al. 2004; Kotwicki and Lauth 2013).

Commercial landings for the 2018 Bering Sea Aleutian Islands (BSAI) snow crab fishery totaled 24,820,146 pounds at an ex-vessel price of $3.89 per pound and 130 million dollars for the industry (ADFG 2020). The fishery has fluctuated in biomass and landings over decades,

with low periods in the mid-1980s and historical lows in the early 2000s. A changing climate and shifting ecosystem contribute considerable uncertainty to stock assessment models which seek to describe the population dynamics to make predictions for future scenarios in terms of fishery productivity and sustainable fishing levels. Stock assessments are scrutinized by scientific review boards, government agencies, fish processing and seafood industry associations, fishing cooperatives, vessel owners, and permit holders prior to adoption of annual catch limits. Historical spatial records captured by standardized surveys are particularly well suited for analysis in GIS using statistical modeling developed for spatiotemporal datasets; and effective spatial representation in map visuals can help communicate complex results and engage stakeholders in the decision-making process for fishery management plans.

1.4. Summary

The overall goal of this study is to describe the spatiotemporal patterns of snow crab distribution and abundance in relation to environmental conditions (surface temperature, bottom temperature, depth) and predation (Pacific cod abundance). This study also seeks to demonstrate how GIS can be applied in marine fisheries ecology towards exploring and modeling relationships in space and time.

EBS bottom trawl survey data of snow crab distribution and abundance was gathered in ArcGIS Pro 2.6.1. The Space time Pattern Mining Toolbox and Modeling Spatial Relationships toolsets provided statistical modeling tools to analyze and visualize spatiotemporal trends across the EBS from 1982 to 2018. Concurrent predator abundance and temperature data were included for ecological context as two key explanatory variables impacting snow crab populations. Three-dimensional (3D) rendering of the dataset provided context for regression analysis which

explored the scale and significance of the ecological relationships with snow crab distribution in 2018.

This thesis is presented in five chapters, beginning with this introduction to snow crab spatial biology, significance of the fishery, and EBS ecosystem dynamics. Chapter 2 is a collection of related work on the spatiotemporal analysis of species distribution and abundance patterns in the EBS, including traditional regression techniques and more novel spatial approaches. Each of these works served as a guide in development of the methods outlined in Chapter 3, including dataset engineering, GIS integration, geostatistical tools, and analyses of spatiotemporal patterns. Chapter 4 presents the results of the analysis and main findings, and Chapter 5 expands on the results in a broader ecological context, discusses successes and limitations of the chosen methodology and potential for further development. Chapter 5 also presents the case for GIS as an effective analysis and visualization tool in marine fisheries ecology.

Chapter 2 Related Work

This chapter outlines previous research related to species distribution and climate in the eastern Bering Sea and provides examples of GIS as applied to spatiotemporal analysis and spatial regression techniques in marine fisheries and ecology. Recent environmental and biological trends are described for the EBS environment and *C. opilio*, and approaches to modeling expansive spatiotemporal datasets that extend over a large and dynamic environment like the EBS shelf are discussed. Examples from other regions and scientific domains which have utilized GIS for statistical analysis are also provided to supplement the relatively few examples of GIS and local regression analysis in marine fisheries studies.

Spatial non-stationarity is typical of species distributions in marine systems, but local scale variation is often masked by global scale trends. A better understanding of local variation can inform global regression model performance and development of hypotheses for the multi-scalar processes underlying variation in snow crab distribution and abundance. This project demonstrates the efficacy of GIS in performing exploratory spatiotemporal analysis and regression modeling of large datasets through visualization and geostatistical analysis. Spatially focused methods were structured to capture multi-scale patterns in snow crab distribution in the EBS and to introduce alternative methods for exploring temporal correlation as well as identifying local relationships in a large dataset.

2.1. Measuring Ecosystem Change in the Eastern Bering Sea

Temperature is a main determinate of marine species distribution and preferred habitat range as it controls physiological function (metabolism, growth, reproductive rate) (Molinas et al. 2018). Stevenson and Lauth (2018) have suggested that warming trends beginning in the 1970s coincide with a regime shift in which groundfish abundance began to increase and

overtake the ecosystem previously dominated by subpolar benthic invertebrates such as snow crab. The shift occurred as subpolar species retreated to the north and colder temperatures (Mueter and Litzow 2008; Stevenson and Lauth 2012; Kotwicki and Lauth 2013) but the significance of the change and the magnitude varies amongst the research depending on modeling approach and units of analysis. Though temperature has been identified as the most significant environmental determinate of wide scale distribution patterns in snow crab and other marine species, there is ongoing debate as to the significance of top-down predator-prey relationships between invertebrates and groundfish as populations are shifting (Orensanz et al. 2004; Zheng and Kruse 2006; Parada et al. 2010; Windle et al. 2010; Windle et al. 2012; Murphy 2020).

2.1.1. Spatial Units

Fisheries and species distributions are often modeled through some form of global regression analysis (Cianelli et. al 2008). For large study areas the region is usually divided into smaller spatial units prior to analysis to improve model performance as a single equation is fit to the spatial unit chosen. Some distribution and abundance studies divide the EBS according to oceanographic patterns (Parada et al. 2010) or physical characteristics like depth (Ernst et al. 2005; Emond et al. 2015). Burgos et al. (2013) divided the EBS according to geographic domain (coastal, middle, and outer as described in the introduction), a common reference system for the region that was adopted for this study. Burgos et al. (2013) further divided the EBS into transverse sections parallel with latitude, resulting in 13 spatial units in their analysis of snow crab distribution. Global results for each unit were compared to describe pseudo-local variation in distribution in relation to temperature and Pacific cod predation.

Snow crab distribution is often related to the extent of the cold pool, which can alternately be defined by the 1° or 2°C isotherms throughout literature. Kotwicki and Lauth (2013) calculated the change in area over a 30-year time series of EBS survey data adhering to the 1°C definition of the cold pool. This change variable (Δ) was an input parameter for a generalized additive model (GAM) to determine the impact on snow crab distribution and is one of the rare studies to report no significant relationship between temperature and species distributions, as no significant trend was detected in the cold pool extent over the study period. Trends in species distribution were attributed to temporal correlation, while environmental variables were found to be less significant. Other studies have defined the cold pool by the 2°C isotherm (Mueter and Litzow 2008; Marcello et al. 2012; Murphy 2020). Marcello et al. (2012) applied a similar GAM technique to describe snow crab distribution data from surveys in the northwest Atlantic and found significant correlation with lagged temperature variables.

The temporal units of analysis also vary from study to study. Year to year pairwise trends have been used to model temporal correlation at single locations (Kotwicki and Lauth 2013). Survey years have also been aggregated to investigate cumulative effects and large-scale processes (Orensanz et al. 2004; Marcello et al. 2012). Temporal lag from environmental impacts at various life history stages in the snow crab life cycle has been investigated to understand the cyclical patterns of abundance or temporal correlation and connectivity between year classes of snow crab (Ernst, Orensanz, and Armstrong, 2005; Marcello et al. 2012; Emond et al. 2015).

Spatiotemporal exploration and visualization of species distributions using a multiscale approach in GIS can lead to better developed regression models and therefore better prediction of species distributions and abundance. Geographically weighted regression (GWR) is used to visualize how the strength and nature of relationships vary spatially by performing the regression

at each location in the study area, which can identify regions where relationships are consistent and the dependent variable is predicted with higher accuracy – or regions where the model performs poorly indicating a missing variable (bias) or non-linear relationship (Mitchell 2009). In this way model results can help identify ecological regions and the conditions that shape species distributions. Global and local regression techniques are discussed in section 2.2.3.

2.1.2. Ecological Considerations

As snow crab populations shift north, Parada et al. (2010) postulated that circulation patterns in the EBS present a barrier to re-distribution into the southern EBS, even in years of favorable conditions (<2°C). A previous study by Orensanz et al. (2004) had termed this asymmetrical shift the 'environmental ratchet hypothesis' (ERH). In this case warming trends initially provided a bottom-up control of crab recruitment and potential range of habitat, but EBS currents, female migration patterns, and cod predation on juvenile crab prevented the southward expansion during more favorable cold years. This has resulted in a realized niche or limited extent of a species' potential habitat.

Based on their study of female distribution and immature cohort classes, Burgos et al. (2013) hypothesized that an extended cold period from 2006 to 2010 resulted in decreased abundance of cod, and therefore predation, which allowed for the observed increase in recruitment of immature crab to the middle domain in 2010. Pulses of high abundance have been noted throughout the literature and are believed to correlate with the strength of female parent cohorts, with some dampening effect of predation and the ERH proposed by Orensanz et al. (2004).

Although temperature may be a main driver of species distributions and biogeography of the EBS, multiple factors influence the survival and distribution of snow crab at different stages

in its life cycle. Sea surface temperature (SST) will impact the growth and survival of pelagic larval stages, whose transport is controlled by surface currents in the EBS; bottom temperatures then exert more influence in benthic distribution as immature crab preferentially settle in the colder middle domain (Orensanz et al. 2004; Parada et al. 2010). Immature crab and small females are preferential prey for Pacific cod, so predation pressure effects are also focused on this segment of the population. Studies which break down the snow crab population into population demographic groups or sex age classes have captured variable distribution and abundance patterns that reflect sex and age class-specific preferences and ecological relationships (Ernst, Orensanz, and Armstrong 2005; Ernst et al. 2012; Emond 2015; Murphy 2020). Variable life history stages and a fluctuating environment in terms of temperature and predation suggests spatial non-stationarity, or locally variable relationships, that might contrast with global trends in snow crab distribution and abundance.

Emond et al. (2015) and Boudreau, Anderson, and Worm (2011) also studied female cohorts separately from the snow crab population total to describe temporal trends. They observed, in many cases, a correlation between mature female abundance and a lagged recruitment pulse approximately 4 years later as progeny presumably settled to the benthos. Murphy (2020) tracked immature females, mature females, and mature males separately in an analysis of snow crab and its cousin, tanner crab, in the EBS to flush out the relationships between each demographic with temperature and depth.

Pacific cod stomach contents from the EBS survey have been analyzed in various studies and indicate that snow crab is a main prey item (Lang et al. 2005; Boudreau, Anderson, and Worm 2011; Burgos et al. 2013). Predation has also been postulated as a top-down control of snow crab abundance, but global regression analyses have failed to capture any significant

relationship between predator species and snow crab. This may be due, in part, to significant differences in spatial distribution and overlap on the EBS as well as in scales of abundance. This scale factor and spatial variation between the two species drives much of the deviation in spatial units of analysis seen in previous studies.

An exploration of the spatiotemporal distribution of snow crab sex-age classes and historical environmental conditions can help visualize and define distribution patterns in space and time that can inform progressive statistical analysis and support further hypothesis development.

2.1.3. Space Time Exploration of Distribution

Bottom temperatures in the EBS have fluctuated between averages of .5 to 5°C for the shelf survey region since 1982. A recent warming trend began about 2011 and peaked in 2016; after three years of no sea ice formation over the shelf average temperatures remain near peak highs over 3°C (ADFG 2019). Charts for average bottom and surface temperatures and CPUE (number/ nm^2) for total snow crab, immature snow crab, mature female snow crab, and Pacific cod are shown in Figures 3 and 4.

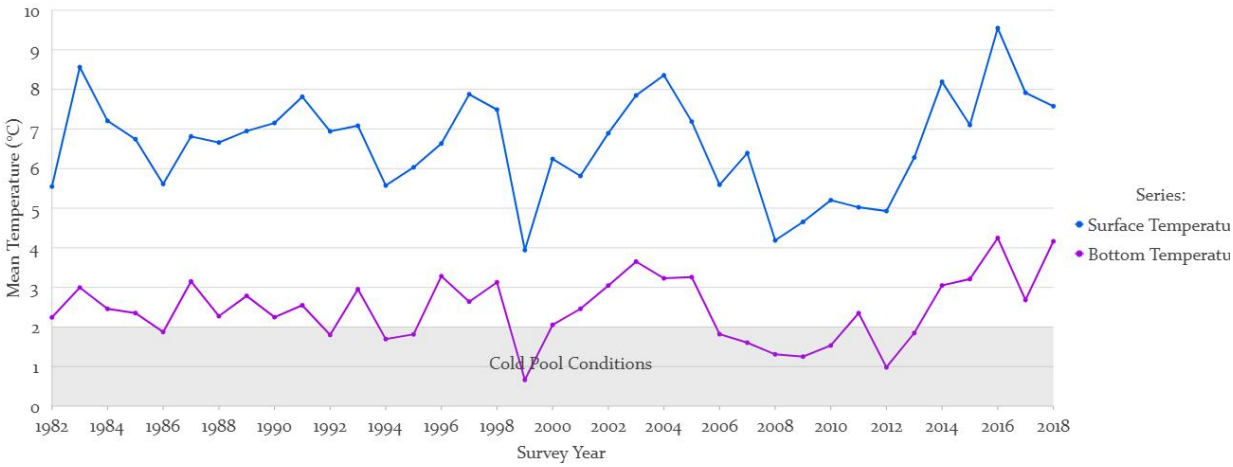


Figure 3. EBS average summer sea surface and bottom temperatures, 1982 – 2018

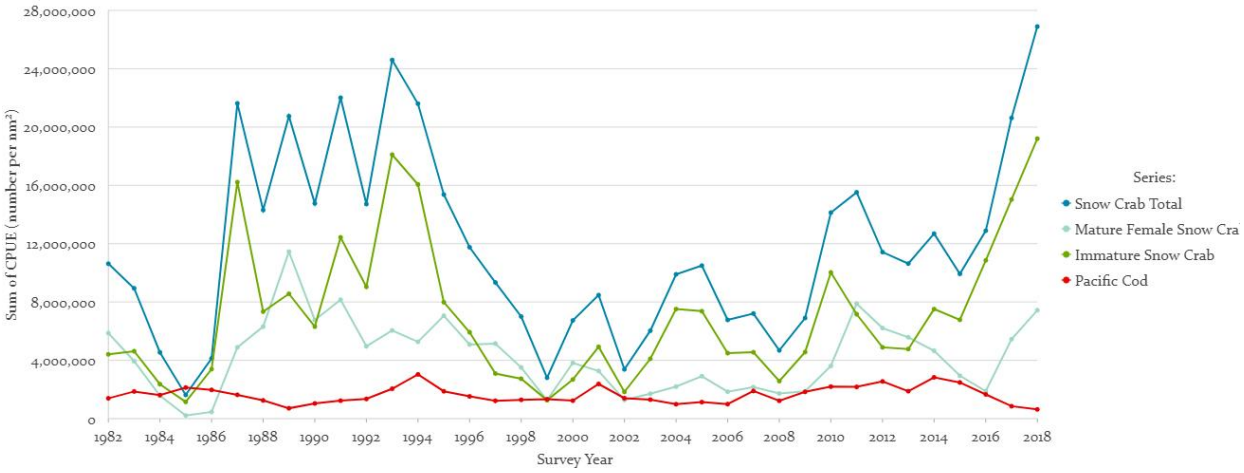


Figure 4. EBS annual total CPUE (number per nm square) for immature, mature female, and total snow crab age classes, and Pacific cod, 1982 – 2018

Snow crab abundance in the EBS over the time series was highest between 1986 and 1996. CPUE peaked at over 7 million in 1993, then dropped to under 500,000 by 1999. This was the first time Pacific cod CPUE values overcame those of snow crab since 1985. Since the sharp decline in 1998 and 1999 down to 500,000 CPUE, a small peak occurred in 2014 at just over 3,500,000 before CPUE again dropped to the historic low of 250,000 in 2016. Pacific cod abundance fluctuated at a smaller scale than snow crab over the series and was relatively more dispersed. Peak CPUE of Pacific cod on survey over the time series was just under 2,000,000 in 2014 and dropped to its lowest survey record in 2018 at 500,000 CPUE.

Prior to 1998 average bottom temperatures fluctuated between 2 and 3.5°C (a 1.5°C range) while post-1998 the average fluctuated between .5 and 5.5°C (a 5°C range). Average bottom temperatures rose from 2°C in 2006 to 4.5°C in 2018. After a second year of no sea ice formation over the EBS shelf, no cold pool formed in 2018. Only seven stations on the northeast fringe of the survey area reached a summer low of 1.6°C (bottom temperature maps for 2006 to 2018 are provided in Appendix A). In previous years when the cold pool formed it proliferated south along the middle domain (50 – 100 m), and immature snow crab clustered here.

2.2. GIS Modeling

GIS is a technology increasingly used for integrating, analyzing, and visualizing spatiotemporal data. Space time analysis and geostatistical methods have been used in various domains to explore, quantify, and build on established theories, and the mapping of spatial information and data visualization enhances communication. Visualizations can also promote engagement in the fisheries management and decision-making process (Kemp and Meaden 2002; Cianelli et al. 2008; Hardy et al. 2011).

Many of the previously mentioned studies apply basic GIS tools to interpolate data points and derive surface maps of temperature and abundance patterns, or to plot time series of population centers over time. These are simple yet effective methods of visualizing geographic shifts in species ranges and ecological relationships. GIS also provides more sophisticated tools for the analysis and visualization of spatiotemporal data, and cluster and outlier detection in spatiotemporal correlation tests. The suite of regression tools available in ArcGIS Pro has been expanded for global and local modeling techniques such as ordinary least squares (OLS) and GWR.

Predictive modeling and machine learning (ML) is also being developed in GIS and may provide fisheries managers with tools for making effective decisions for spatial quota allocations (Cianelli et al. 2008; Hardy et al. 2011). Extensive, robust datasets and repetitive testing are required to train models and accurately identify the scale of ecological processes in action, which can change over time. GIS enables manipulation and interchange of variables and analysis units (spatial or temporal) as input parameters in regression and ML algorithms towards better predictive modeling.

2.2.1. Spatiotemporal Analysis

Standardized fisheries surveys are designed to collect repeated measurements at regular frequency and locations to enable robust statistical analysis. This enables managers to measure change and estimate its significance over time with some amount of probability or confidence (Stamatopoulos 2002). Datasets with spatial locations and time stamps can be structured as a space time cube with netcdf file formatting to enable spatiotemporal pattern mining and statistical analysis in ArcGIS Pro. The cube structure enables visualization and analysis of change over time at each location by assigning location IDs and time step interval designations to each record. This makes space time cubes particularly well-suited for modeling ecological systems and managing station data like the EBS bottom trawl survey.

Spatiotemporal analysis in GIS differs from traditional statistics which focus on the attribute value in dataspace and assume independence between observations (Fotheringham 2002; Ciannelli et al. 2008). Spatial and temporal autocorrelation relate to Tobler's first law of geography in that nearby features are more similar than those located farther apart (Fotheringham 2002). Spatiotemporal analysis accounts for the autocorrelation of attribute values and accepts some degree of dependence between nearby observations by differentially weighting features (in this case individual survey station records) according to the distance between them (Mitchell 2009). For example, survey stations within a specified distance, or spatial neighborhood, are more heavily weighted in spatiotemporal analyses than those outside this distance since catch records of snow crab are likely comparable with catch records at nearby survey stations. As distance between survey locations increases, the correlation in attribute values is likely to decrease so the spatial neighborhood distance should represent the degree of interaction or dependency between features.

Without any spatiotemporal autocorrelation the attribute values would appear randomly distributed across the study area and over time (Mitchell 2009). GIS analyses quantify the level of clustering (positive correlation) or dispersion (negative correlation) and incorporate this aspect of the data in the calculation of statistics within the context of the spatial neighborhood to determine how significantly the patterns diverge from a random distribution (see Mitchell 2009 for a detailed explanation of the mathematical formulas, or Fotheringham 2002 in the case of GWR). A p-value is assigned in the statistical output to indicate whether the pattern is significantly different than random, and a z-score with a negative or positive designation to indicate if the trend is increasing or decreasing along a standard normal distribution curve (z-score of zero would be equal to the mean).

The Mann-Kendall statistic is automatically applied at every defined location during creation of a space time cube. This independent bin test summarizes the temporal trend in the attribute over time at each station location by summing each bin as an increase (+1) or decrease (-1), or tie (0) with the previous time step (Esri 2020). Time series cluster analysis in ArcGIS Pro compares these temporal trends for the characteristic or attribute of interest, and groups stations together based on correlation in the timing and proportional change in the value over time (attribute profile correlation). Time series clusters represent areas of similar population growth patterns which could be of interest to fisheries managers and ecologists.

Alternative spatiotemporal analyses are adapted from traditional methods to identify clusters of similar values. Cluster locations provide context for regression modeling and understanding relationships. The Anselin Local Moran's I statistic (clusters and outliers analysis in ArcGIS Pro) is calculated by comparing the target feature bin in the space time cube with nearby (local) bins using the spatial neighborhood concept, then comparing the local mean

against the global mean of features to identify outliers and/or correlation in the characteristic of interest (Esri 2020). Hot spot analysis calculates the Getis Ord Gi* statistic for each bin location by comparing the target feature within the spatial neighborhood; hot spots are high value features surrounded by other features with high values, and are considered statistically significant if locally (in space and time) the sum characteristic of interest is greater in proportion than the global sum. Hot and cold spots indicate areas of significant decline or growth in population. Emerging hot spot analysis in ArcGIS Pro further describes the trend by classifying each station location in the space time cube according to recent temporal patterns (for a full description of the emerging hot spot classification scheme see Esri 2020).

Whereas the spatial component of a data point and visualization are typically secondary to the quantitative results and the data attribute value, fisheries managers are increasingly calling for spatially focused analyses (ADFG 2019; NPFMC 2019). Kemp and Meaden (2002) developed a custom system to support decision making through spatiotemporal exploration, visualization, and multivariate modeling via GIS. Their goal was to present users (managers) with a tool for exploratory visualization through customized mining of spatiotemporal data. The GIS application allowed users to specify input variables (species and/or sex age classes), analysis units (spatiotemporal aggregation) and varied statistical tests. The development of customized software is beyond the scope of this project, but the project goal demonstrates the value to management of exploring spatiotemporal parameters to assess how choices in spatial neighborhood and spatial units affect the results. GIS also offers the efficacy of having immediate visual support of results.

There are few examples in related research which take advantage of the recently developed space time pattern mining tools in ArcGIS. A Master’s thesis by C. Steves (2017)

demonstrated the efficacy of the space time cube in detecting change in the Alaska bottom trawl fishery in terms of effort (number of trawl tows) and efficiency (weight per tow). Visualization of change in these parameters relative to the location of marine protected areas (MPAs) and sea ice extent in the EBS between 1993 and 2015 revealed clusters of increased (hot spots) or decreased (cold spots) bottom impact and fishery productivity. These spatial and temporal trends could help guide the decision-making process for identifying priority impact areas or low productivity fisheries that could benefit from temporary or established conservation areas.

Epidemiology is another domain which has taken advantage of the space time capabilities of ArcGIS software. Zulu, Kalipeni, and Johannes (2014) built a progressive, multi-scale statistical analysis based on a 7-year time series of HIV infection in Malawi to better understand spread of the disease. This case study applied spatiotemporal analysis and regression in GIS to analyze HIV prevalence in Malawi over a seven-year time period. Anselin Local Moran's I statistic identified clusters of similar prevalence rates (high positive autocorrelation) and outliers surrounded by much higher or lower prevalence rates (high negative autocorrelation). Clusters indicated areas experiencing similar disease trajectories which provided context for the regression analysis. OLS was used to measure drivers of infection such as population density and distance to population centers. By applying the regression to national and local level district administrative units, the results could provide a framework to implement intervention policy plans at district and national levels to predict and mitigate the spread of disease.

The progressive statistical analysis using GIS and visualization techniques for the study in Malawi serves as a model for this project's methods structure. The EBS bottom trawl survey provides a robust and extensive dataset to similarly explore space time trends and relationships that could help understand how fisheries species distributions are changing in space and time in

order to more efficiently allocate quota spatially. As in the Zulu, Kalipeni, and Johannes example (2014), preliminary space time explorations were used in this study to provide context to regression modeling at multiple scales.

2.2.2. Predictive Modeling

A unique and creative application of ML and the space time cube was implemented by Aydin and Butler (2019). In this case, random forest (RF) algorithms were used to determine the ocean conditions impacting the health of seagrass habitat and predict the expansion or degradation of these marine habitats globally. An effective map derived from this analysis depicts the results of an emerging hot spot analysis, showing areas of increasing or decreasing suitability for seagrass growth. A variation in the time cube 3D visualization structured rising temperature along the z-axis rather than time; hot or cold spots were predicted for each location depending on the magnitude of warming scenario as per degree of sea surface temperature increase.

Considering the northward shift in species distributions and the limited spatial coverage of the EBS survey, Hardy et al. (2011) developed a complex ensemble model for predicting the distribution of snow crab in areas outside the range of the survey grid by integrating data from the EBS with limited surveys conducted in the Chukchi and Beaufort Seas. Snow crab abundance and biomass were overlaid with 20 layers of environmental predictor variables which included typical ecological indicators such as sea surface temperature, nitrate concentration, salinity, chlorophyll-a, total organic carbon, infaunal biomass (food source), dissolved oxygen, and depth. The relative importance of each predictor was used to develop a quantitative model of the ecological niche and generate a predictive surface of the entire region.

Prediction was not the goal of this project, but the ML examples provide insight that can be incorporated in the current research to support choices of explanatory variables. The RF model ranked 3 variations of surface temperature as the most important predictor of snow crab distribution, supporting bottom temperature and SST as recorded at survey locations as significant environmental indicators. While Hardy et al.'s (2011) results were relatively successful in detecting the potential niche, the most successful algorithm still failed to accurately predict presence or absence in multiple southern regions of the EBS. This supports further investigation of factors other than temperature impacting snow crab distribution in the south through techniques like GWR which capture this non-stationarity.

2.2.3. Describing Relationships

As demonstrated by the seagrass study and the ensemble model of snow crab distribution, understanding ecological relationships is necessary to predict future scenarios for these vital marine resources in the face of a changing climate. Modeling the distribution of mobile organisms in a dynamic environment over time and deciphering the spatiotemporal correlations of a multitude of biotic and abiotic interactions presents challenges. Spatiotemporal visual exploration and analysis can be a strategic first step to identifying patterns and relationships to support regression and eventually predictive modeling.

Regression techniques such as ordinary least squares (OLS) and GWR have recently been developed for Esri's ArcGIS platform and are included in the Modeling Spatial Relationships toolset. Few examples exist in the literature for GWR as applied to fisheries, but the technique shows promise as an exploratory tool and is well suited for GIS as each station location is assessed individually, enabling visualizations of the results for each relationship in space.

2.2.3.1. Global regression

Many of the previously mentioned works apply global forms of regression such as GAM and generalized linear models (GLM) in their approach to modeling distribution of snow crab in relation to temperature, predation, and other variates. Global modeling results are highly sensitive to the areal unit chosen as the analysis is based on the entire dataset as a single solution is calculated for the intercept term, variable coefficients, and the model's goodness of fit across the study region.

Emond et al. (2015) investigated the cyclical fluctuation in the northwestern Atlantic snow crab populations in relation to environmental drivers. By tracking groups of early benthic instars over 23 years, this study was able to measure the strength of pseudocohorts (female year classes) over time in relation to multiple variables using global regression and ordinary least squares (OLS). Results suggested that intraspecies cannibalism and bottom water temperature had the strongest influence on distribution and survival for early instars (three years old, newly settled snow crab). This countered the hypothesis that historical predation or snow crab abundance values were the more significant variables determining the fluctuations in total snow crab abundance. Historical variables or lagged variables were incorporated as regression model independent variables as representative of the temporal correlation between current snow crab distribution patterns and past environmental conditions like sea temperatures, predator abundance, and maternal age class abundance or larval production. The global OLS modeling applied in this study may have been too large scale to capture local variation inherent in species relationships, particularly predator-prey. North Atlantic cod and snow crab ranges only partially overlap in this region so community scale relationships would have been dampened by OLS which smooths the local variation in favor of the global average.

This study in the northwest Atlantic and the works mentioned in relation to spatial analysis units highlight a limitation of global approaches to regression. An exploration of these processes at multiple scales using GIS (global OLS and local GWR) captures global trends and local variation, and spatiotemporal analysis provides context that may help with interpretation of scale and autocorrelation.

2.2.3.2. Local regression

GWR is a local technique that can be applied in an exploratory way to support the fine-tuning of global scale regression models (Fotheringham 2002; Foody 2004; Zhang and Shi 2004; Bevan and Connolly 2009; Windle 2010). GWR has the advantage of visual exploratory power through mapping of each independent variable's coefficient, calculated at each location in the study area. Clusters in residuals indicate collinearity or model misspecification and can therefore help identify missing variables and appropriate scale or spatial units. Global and local scale regression model diagnostics and the local variable coefficients in GWR can be compared to visualize local variation in the strength of relationships as compared to the overall trends described globally.

OLS model diagnostics include the Variance Inflation Factor (VIF), Joint F and Wald statistics, the Koenker (BP) statistic, and Jarque-Bera statistic, each with associated probabilities of significance. VIF scores greater than 7.5 indicate redundancy in the independent variables or multicollinearity. The Koenker statistic determines the level of spatial and value consistency in the relationship (nonstationarity). The null hypothesis for this test indicates a stationary process in space and CPUE variation, while a significant Koenker (p-value $< .05$) indicates spatial non-stationarity and further reasoning for GWR analysis. The Joint F statistic can be interpreted as a measure of overall model significance if the Koenker test is not significant. If the null hypothesis

is rejected and nonstationarity results in a significant Koenker test statistic, the Wald statistic should be used to determine model significance. A p-value < .05 for the Wald or F statistic indicates a significant model. Lastly, the Jarque-Bera statistic test indicates a normal distribution in the model residuals. If the Jarque-Bera test was statistically significant, this would indicate that residuals were skewed and the model was biased towards over- or underestimating CPUE values in certain regions on the map and/or in data space. Bias could be due to misspecification (missing model variables), nonlinearity in the relationships, extreme outliers, or spatial nonstationarity as indicated by the Koenker test.

Foody (2004) pioneered the application of GWR for ecological research on bird species distributions in Europe and found that the relationship between bird species biodiversity, temperature, precipitation, and NDVI varied spatially and at different scales in sub-Saharan Africa. Similar determinations of spatial non-stationarity or inconsistent relationships across space have been demonstrated in other domains in addition to ecology.

GWR outperformed global OLS models in Zhang and Shi's (2004) forestry productivity study that measured tree growth in relation to several local environmental parameters. In this case, mapping of model coefficients provided visualizations of the nature of the relationship between growth patterns, stand density, and timber yield for multiple tree species in New Hampshire. Local coefficient mapping illustrated the spatial processes under study and supported the development of established global models by identifying areas of poor model performance in low r-squared (r^2) values. GWR local r^2 values represent the proportion of variance in the dependent variable accounted for by variance in the independent variables. A low amount of variance explained indicated model misspecification or missing independent variables and differences between tree species.

Bevan and Conolly (2014) also demonstrated the utility of GWR and mapping local variable coefficients in the field of archaeology. Their investigation of pottery artifact density in relation to slope, geology, and other environmental variables across a small island in Greece enhanced the predictive ability in finding pottery deposits by identifying areas predicted to have similar densities and could help focus sampling efforts to maximize discovery and collection of artifacts and helped to focus sampling efforts in areas that transect sampling would have poorly covered. Maps of regression model residuals showed a high degree of spatial correlation in the pottery deposits and pockets of similar relationships, indicating what may be appropriate spatial units of analysis for future studies. This work uncovered spatial structure in pottery deposits at variable scales and enabled hypotheses of human settlement patterns and timelines, ultimately supporting more accurate global spatial models of ancient civilizations and the geomorphological processes which alter their archaeological record.

A single case of local modeling in fisheries was found in research which applied GWR. Windle et al. (2010) compared the performance of global and local regression modeling techniques to describe the spatial distribution of northern Atlantic cod off the coast of Newfoundland and Labrador, Canada in relation to temperature, distance from shore, and abundance of two key prey items (northern shrimp and snow crab). The predictive success (in terms of model error or residuals) of a traditional global logistics and binomial GAM were each compared to that of a logistic GWR. GWR outperformed the global GAM approach in terms of error, and spatial variation in the strength and nature of relationships were visualized through mapping of variable coefficients. Mapping of the GWR residuals in this northwestern Atlantic example also facilitated the detection of areas where the model was less effective in explaining

the observed variance, which supported inference as to model mis-specification or multi-scalar processes occurring in these areas.

In GWR and spatiotemporal analyses, the kernel bandwidth or spatial neighborhood chosen for the analysis is critical as it represents the window surrounding the sample location included in the local regression and spatially weighted statistical analysis. Windle et al. (2010) and Bevan and Connolly (2009) each devised a measure of spatial non-stationarity by iterating through increasing kernel bandwidths and comparing coefficient of variation (CV) scores to determine the scale at which relationships became heterogenous. Other methods of determining appropriate bandwidths include incremental adjustment of the distance until a minimum Akaike's Information Criterion (AIC) score is reached, indicating optimum model performance (Fotheringham 2002; Windle et al. 2012).

The insight gained from Windle et al.'s (2010) initial exploration of invertebrate/habitat/predator associations via GWR led to a second iteration in which snow crab and shrimp were examined as dependent variables, rather than cod (Windle et al. 2012). This study applied GWR to a 20-year time series but limited the spatial extent by first determining core habitat ranges for shrimp and snow crab in the northwest Atlantic. Windle et al. (2012) chose to highlight a warm (low abundance) and cold (high abundance) year from their time series to compare the spatial variability of the GWR coefficients, which proved an effective visualization technique. Model residuals were higher in shallower areas, indicating missing variables and possible grounds for partitioning of the dataset.

As in the previous study, the relationships between crab and cod were relatively weak and showed stronger dependence on depth and environmental factors. However, the species assemblages in this region exhibit variable oceanographic patterns and ecosystem structures

compared to the EBS. In the northwestern Atlantic, warmer waters are found at shallower depths, which is opposite that of the EBS. Species abundance for snow crab and cod are also much lower in the Atlantic and diet studies show that shrimp are preferred over snow crab. For lack of widespread presence Windle et al. (2012) restricted their study area to a 'core habitat' zone where snow crab was more prevalent in survey samples. Snow crab and Pacific cod are both widespread across the EBS and their ranges overlap to a greater extent than in the Atlantic. For these differences, GWR may produce variable results in the EBS as a significant relationship between shrimp and Atlantic cod was described in this study.

This second GWR study in the Atlantic also highlights an important point, or possible pitfall, of regression modeling. Relationships must be linear for OLS and GWR modeling, and scatterplots or histograms should be examined to determine if a data transformation is necessary. Windle et al. (2012) tested univariate relationships to determine significance and nonlinearity, then ran the regression with and without data transformations for the few variables that exhibited nonlinear relationships. An interesting result was that there was no significant difference between the analyses using transformed data. Methods were developed for this study with these factors in mind, and extensive data exploration was performed to understand data distributions and achieve a linear transformation. This process is discussed in the following chapters.

2.3. Summary

GIS enables integration, analysis, and visualization of complex, multivariate relationships, and the detection of spatial and temporal correlation. While GIS tools do not eliminate the pitfalls of the MAUP or scale when dealing with large datasets, exploration and visualization can help to understand these data characteristics and lead to better informed parameter choices.

This study seeks to build on these previous works of fisheries biologists and spatial analysts to describe the spatiotemporal distribution and abundance patterns of *C. opilio* in the EBS utilizing GIS visualization and statistical tools. Cues from other studies showing spatial stratification of the population and temporal correlation in abundance support the breakdown of the population into sex-age classes to investigate spatiotemporal patterns. The progressive spatiotemporal and regression analysis performed by Zulu, Kalipeni, and Johannes (2014) served as a basic methodology for space time and regression workflows while the approach of Windle et al. (2010, 2012) in the application of GWR was adapted to the EBS survey dataset.

Chapter 3 Methods

GIS can provide a means of integrating and manipulating complex spatiotemporal datasets, performing analysis, and visualizing results in three dimensions. The methods in this study were implemented to demonstrate the capabilities and benefits of GIS for these purposes with a specific application in fisheries surveys. Spatial and temporal explorations which describe where and when change is occurring in these historical spatial datasets facilitate deeper investigations into observed changes to understand the relationships and the mechanisms which shape species distributions. The methods developed for this study were intended to demonstrate the utility of GIS for spatiotemporal analysis and regression modeling, and the power of space time pattern mining and GWR in supplementing more traditional methods of fisheries stock assessments and statistical analyses.

3.1. Data Source: EBS Bottom Trawl Survey 1982-2018

A subset of the EBS bottom trawl survey conducted by National Oceanic and Atmospheric Administration's (NOAA) Alaska Fisheries Science Center (AFSC) and Resource Assessment and Conservation Engineering (RACE) division was developed for spatiotemporal analysis and regression modeling using ArcGIS Pro. A point feature class was created in ArcGIS Pro from the geographic coordinates of each standard survey station (provided by NMFS in decimal degrees) and projected to Alaska Albers Equal Area coordinate system. A space time cube was created from this feature class for visual exploration and analysis of species distributions and environmental conditions from 1982 - 2018.

This time series included 349 stations sampled annually from 1982 to 2018. EBS CPUE survey data was downloaded from NOAA's RACE division site (AFSC 2019). Catch per unit effort (CPUE) was adopted as the species abundance variable in the analysis as it is a

standardized measure often applied in fisheries research and management as a relative index of abundance (Orensanz et al. 2004; Parada et al. 2010; Zheng and Kruse 2006; ADFG 2019). The dataset was organized so that each survey station annual record included the CPUE for (total) snow crab, immature snow crab, mature female snow crab, and Pacific cod, as well as depth, near bottom temperature and surface temperature.

In total the EBS bottom trawl survey samples nearly 216,000 nm^2 (400,000 km^2) of the shelf. Vessels typically tow for 30 minutes at a standard 3 knots (1.54 m/s), starting in the coastal domain in late May/early June and ending in August or September with the outer domain stations. The survey is designed so that established coordinate locations or stations are sampled annually; 349 standard stations are stratified across the shelf in a 20 by 20 nm grid aligned with latitude and longitude. Certain areas of high catch rates are fished in a denser grid by adding a station at the corner of the 20 x 20 nm grid cell but these stations were excluded from the analysis to maintain spatial and temporal consistency in the CPUE index. EBS survey data prior to 1982 was excluded due to a change in trawl gear specifications, which likely impacted catch efficacy. This resulted in over 12,913 data points over 37 years.

A bathymetric surface layer was interpolated from the depth of survey stations using the geostatistical method kriging. Ordinary kriging defaults were accepted as the layer was for visualization purposes only. The default cell size is $1/250^{th}$ of the lesser extent, height or width, in the output coordinate system linear reference – in this case Alaska Albers and meters. This resulted in a raster of cell size 2 x 2 nm (3763 m), which was then clipped to the survey extent within a 10 nm buffer. Depth contours at 50 and 100 m were derived from this raster layer as a second visual aid to delineate the shelf domains (coastal, middle, outer). The cold pool expands south from the Arctic along the wide and flat middle domain, between 50 and 100 m. Immature

snow crab are stenothermic and thrive within a narrow range of temperatures around 2°C so the cold pool extent is critical in shaping snow crab population spatial structure (Orensanz et al. 2004).

One limitation of the EBS survey design is that smaller and more fragile animals (early settlement phase snow crab and youngest immature crab) can be missed in the larger net mesh or destroyed in the trawling/sampling process. Some survey bias also likely results from timing. The survey begins in May/June in the coastal domain and vessels fish towards the shelf edge and deeper stations in the outer domain to finish in August/September every year. Climate change may be affecting the timing of spring phytoplankton blooms which could affect the timing of snow crab migration and reproductive cycles, and the sex-age class spatial structures observed during summer months (Orensanz et al. 2004; Parada et al. 2010). Despite these timing and catch efficiency biases, the length and consistency of the EBS bottom trawl survey provides a reliable dataset that supports rigorous statistical analyses. A complete list of data variables and definitions from the survey data used in this analysis is provided in Table 1.

Results of the spatiotemporal analyses and regression are highly dependent on the spatial neighborhood distance or kernel bandwidth. A preliminary test of GWR was performed to determine the optimal distance band that would minimize the AIC score. This was done to synchronize the distance band with the spatial neighborhood distance applied in the spatiotemporal analysis, to maintain scale for comparison. The Gaussian kernel bandwidth in the GWR was first tested at 25 nm, or just over the distance between two survey stations. The distance was incrementally increased by 20 nm to 45 and 65 nm (just over the distance between three stations). The AIC scores dipped to a minimum at 45 nm before increasing again, so 45 nm (85 km) was the adopted spatial neighborhood distance for the spatiotemporal analysis and

kernel bandwidth in the regression modeling. Snow crab CPUE in 2018 was elected as the dependent variable for the regression analysis as this was the most current year available.

Table 1. EBS Spatiotemporal Data: 1982 – 2018

Attribute	Definition	Unit, resolution	Data description
YEAR	Year survey conducted	Year, annual	1982 – 2018 surveys
STATION	Survey station ID	Nominal ID, unique	349 standard shelf stations 20 x 20 nm stratified grid
LATITUDE, LONGITUDE	Station location	decimal degrees, $1e^{-05}$ °N, °W	Average geographic coordinate location per station (approximate centroid of grid cell)
BOT_DEPTH	Bottom depth	meters, .1 m	Weighted average depth for area swept
BOT_TEMP	Near bottom temperature	degrees Celsius, .1 °C	Weighted average temperature measured at maximum depth of trawl headrope
SURF_TEMP	Surface temperature	degrees Celsius, .1 °C	Temperature measured at surface
CPUE	Catch number per area swept	number/nm^2	Total Snow Crab Immature Snow Crab Mature Female Snow Crab Pacific Cod

3.1.1. Data Distribution and Exploration

CPUE data distributions were explored, and a preliminary local regression test was performed on snow crab CPUE to determine an appropriate spatial neighborhood or kernel bandwidth for analysis. The presence of many records of zero CPUE interspersed throughout the dataset by nature of the patchy distribution of snow crab resulted in a skewed data distribution. To prevent the loss of any data and maintain a continuous model of distribution, a log (x +1)

transformation was applied to all CPUE values to normalize the data and enable regression analysis. Twenty-nine stations were removed from the coastal domain which recorded zero CPUE for snow crab the entire period (discussed further in the results for time series cluster analysis).

Abundance data for snow crab remained slightly skewed following transformation (Appendix A). Linear regression requires normally distributed data for optimal performance. However, previous work of Windle et al. (2012) in their second GWR study showed that the regression results did not significantly change by transforming the abundance data from a non-normal distribution. Considering these results using similar survey data from the NW Atlantic, the log (x +1) transformation sufficed for the exploratory intent of this study.

3.2. Space Time Cube Exploration of Distribution

A space time cube was created which organized the CPUE station data into bins for each year and station location, and time represented in the vertical t dimension so that 37 bins stacked represent a time series of CPUE at that location. The space time cube was input to Emerging Hot Spots, Local Clusters and Outliers (Anselin Local Moran's), and Time Series Clusters. Each analysis is further explained below. The distance band discovered in the preliminary GWR exploration was used to define the spatial neighborhood for these analyses. The results of each analysis are written back to the space time cube .nc file in ArcGIS Pro. These spatiotemporal trends were then visualized using the Space Time Cube Explorer extension.

The space time cube 3D visualization encompasses a large dataset of over 12,000 data points. Many data classification methods are available in ArcGIS Pro for binning and visualizing the range in CPUE values. Quantile classification is typically applied for linear datasets where equal number of data values are assigned to each class; this can cause distortion between

adjacent classes when the data is not distributed normally. In this case many data values would be spread into different classes although they were similar in value due to the skewed distribution towards zero or low CPUE values. The geometric interval classification method was chosen to accommodate the continuous but skewed CPUE data in this study. For a detailed description of the mathematics involved with the class scheme see Esri's online help (Esri 2020).

3.2.1. Hot Spots

The space time cube created from defined locations was input to emerging hot spot analysis to identify hot and cold spots in snow crab CPUE for each age class (immature, mature female, and total), and for Pacific cod. The time step interval or temporal neighborhood was left as one year for this exploratory analysis, so that the temporal neighborhood consisted of three years (one year before and one year after the year of the target feature). The spatial neighborhood was set at 45 nm to coincide with the Gaussian kernel bandwidth used in the GWR analysis.

The emerging hot spots analysis categorizes bins as hot or cold then assesses each location's cumulative temporal series in a modified Mann Kendall test of snow crab and pacific cod CPUE. The combination of temporal trend and hot spot classification was used to further categorize each station location in the space time cube to form a 2D summary visualization of the hot and cold spot results.

3.2.2. Local Clusters and Outliers

Clusters and outliers analysis identified significant spatiotemporal clustering of high or low CPUE values and outliers of high CPUE stations surrounded by low CPUE stations or low CPUE stations surrounded by high CPUE stations. This analysis was applied to each age class of snow crab (immature, mature female, and total CPUE) and Pacific cod, using the same spatial neighborhood distance as applied in the hot spot analysis and gaussian kernel bandwidth (45

nm). A local Moran's I index value of correlation was calculated for each bin in the space time cube and a cluster category was assigned if the pattern was statistically significant with at least 95% confidence (pseudo p-value < .05).

3.2.3. Time Series Clusters

Time series cluster analysis was applied to identify station locations with similar profiles of CPUE for total snow crab and Pacific cod. Results from the default/initial pseudo-F permutations analyzing snow crab CPUE time series were explored by specifying the number of clusters to identify in the analysis, prior to running the test. Profile correlation in CPUE values was selected to determine similarity between station locations.

3.3. Analysis of Relationships

Time cube visualization and spatiotemporal analysis of CPUE data helped describe the distribution of snow crab and Pacific cod over the time series and provide context for regression analysis. Exploratory regression was first used to identify significant temporal correlation with historical conditions or abundance patterns. These lagged years of delayed impact on snow crab distribution in 2018 were included in the next regression test to compare a global scale ordinary least squares technique with and without lagged impact years included in the analysis as independent variables. The survey period between 2006 and 2018 was selected to represent an average snow crab life span which enabled the exploration of lagged temporal effects related to life history stages and age classes, such as mature female snow crab abundance at time of likely egg extrusion (maternal cohort connectivity), surface temperatures at time of pelagic larval phase of development, and bottom temperatures or predation during settlement and early instar or immature life history stages.

Model performance and accuracy was compared for GWR and OLS, interpreted through Akaike's Information Criterion (AIC_c) score and r^2 or explained variance. Model residuals were mapped and compared for spatial autocorrelation which would indicate misspecification, bias, and/or multicollinearity and lack of variation in model variables. For GWR, additional mapping of model parameters for each explanatory variable enabled visualizations of the strength and scale of relationships of snow crab abundance in space.

3.3.1. Exploratory Regression

The final twelve years of the time series were extracted for exploration of lagged impact. Each survey station was represented by a single point feature with attribute fields pertaining to annual CPUE records from 2006 to 2018. Exploratory regression was run individually for every independent variable (bottom temperature, surface temperature, immature snow crab CPUE, mature female snow crab CPUE, and Pacific cod CPUE) to determine which lagged years correlated most significantly with snow crab distribution in 2018. A full description of the variables and the lagged relationship with current snow crab distribution is provided in Table 2.

Table 2. Exploratory regression variables and lagged impact on 2018 snow crab CPUE

Independent Variables		Lagged Impact on Distribution
Climate Pressure	Surface Temperature	Egg extrusion/hatching and pelagic larval stages
	Bottom Temperature	Settlement phase to maturity
Environmental Variable	Depth	Immature to mature phase migration (no temporal lag, does not vary in time)
Age Class Connectivity	Mature female snow crab CPUE	Abundance of maternal cohort year class is reflected in progeny
	Immature snow crab CPUE	Immature snow crab represent the surviving progeny of matrernal cohort
Predation Pressure	Pacific cod CPUE	Vulnerable immature age classes, small females

3.3.2. Global Regression

After testing each lagged variable, the top three most significant lagged years between 2006 and 2018 were selected as independent variables for the OLS regression using ArcGIS Pro's generalized linear regression (GLR, equivalent to OLS). This global regression approach was also implemented using only 2018 variables (no lagged years as independent variables). The strength and significance of the relationship with snow crab distribution per independent variable was interpreted through the single, global variable coefficients. Model accuracy and performance with and without lagged variables was compared through r^2 and AIC_c values. Spatial autocorrelation in model residuals and the regression statistical diagnostics described in Chapter 2 were used to interpret results between global models, and between the global OLS and local GWR models.

3.3.3. Local Regression

As in the OLS regression, 2018 snow crab CPUE was input as the dependent variable in the GWR. Each station location was analyzed using a Gaussian (continuous) model type with a distance band of 85 km or 45 nm (just over the distance of 2 survey grid cells, each 20 nm). GWR local r^2 results and residuals were mapped to assess model performance spatially. Local variable coefficients were mapped to visualize change across space in terms of relationship strength and consistency. Local model coefficients were divided by the local standard error to estimate a scaled magnitude of error, similar to a t-statistic (Esri 2020). The same regression modeling diagnostics described in Chapter 2 and interpreted for global OLS tests were interpreted for the results of the GWR.

Chapter 4 Results

This chapter outlines the patterns identified in the spatiotemporal analysis and statistical diagnostics and of the regression analyses. Local GWR modeling more accurately modeled snow crab distribution patterns observed in 2018 than the global OLS regression. Global regression techniques were effective in detecting temporal correlation and lagged impact from environmental variables but variance in the GWR local variable coefficients and spatiotemporal patterns of snow crab CPUE suggest spatial non-stationarity and heteroskedasticity across the EBS. Alternate linear transformations of the snow crab CPUE data should be explored to minimize the effect of the skewed abundance patterns. Alternately, GWR derivative results comparing local variable coefficients to local error identified transition zones in the relationships which could be used to break the study area into smaller, ecologically defined units for further spatiotemporal analysis and regression modeling.

4.1. Space Time Cube and Snow Crab Distribution

Raw CPUE data can be rendered in the space time cube to visualize overall distribution patterns. After running each of the spatiotemporal analyses (hot spots, clusters and outliers) the space time explorer can also render the cube according to the statistical results of each test. ArcGIS Pro 3D scenes enable the user to explore the cube in any rendering scheme in 360° and from adjustable heights and perspectives to view more detail. Snow crab CPUE in the EBS from 1982 to 2018 is shown in the space time cube from multiple angles in Figures 5 and 6.

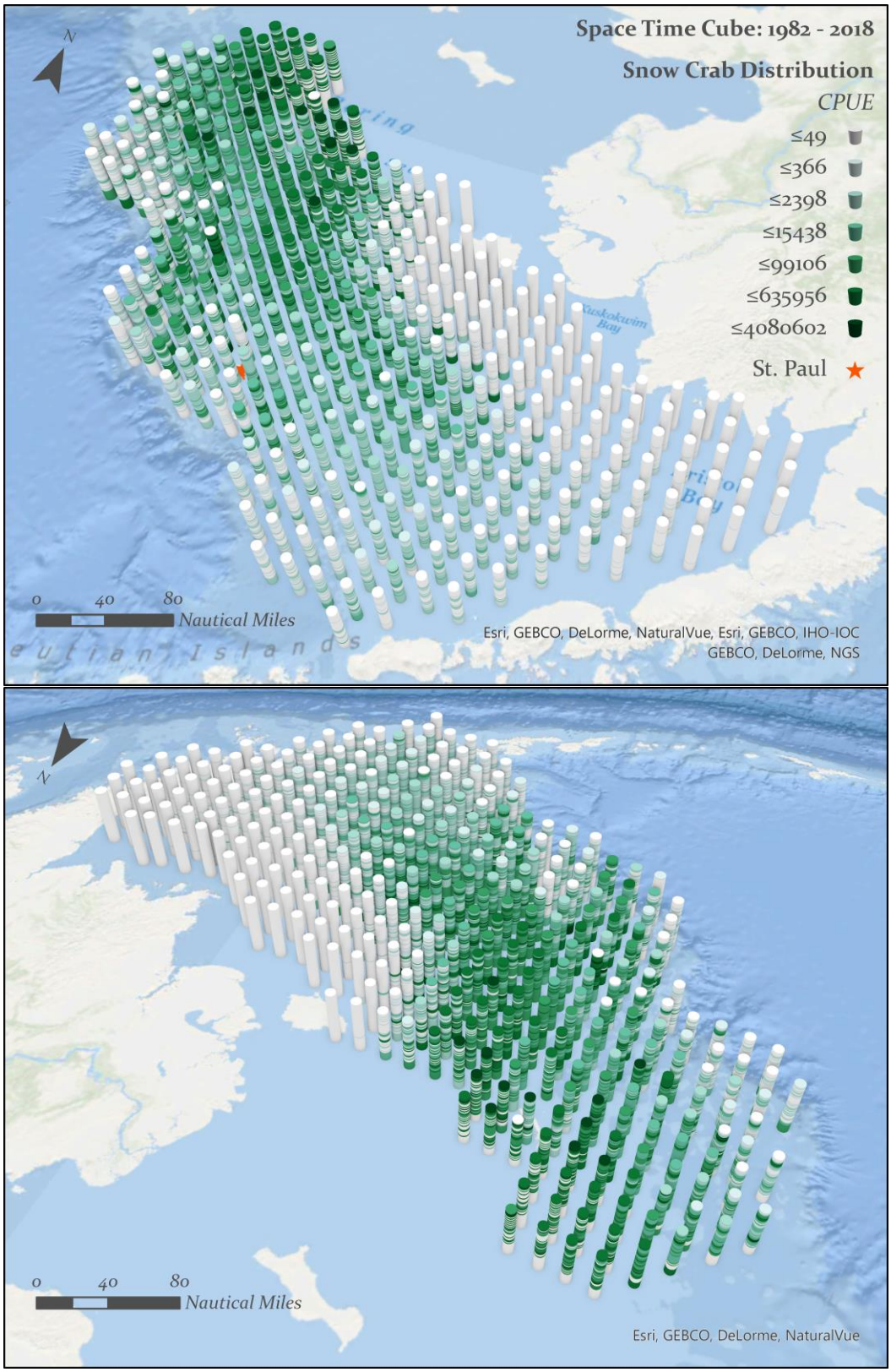


Figure 5. Space time cube views of snow crab CPUE due north (top) and south (bottom)

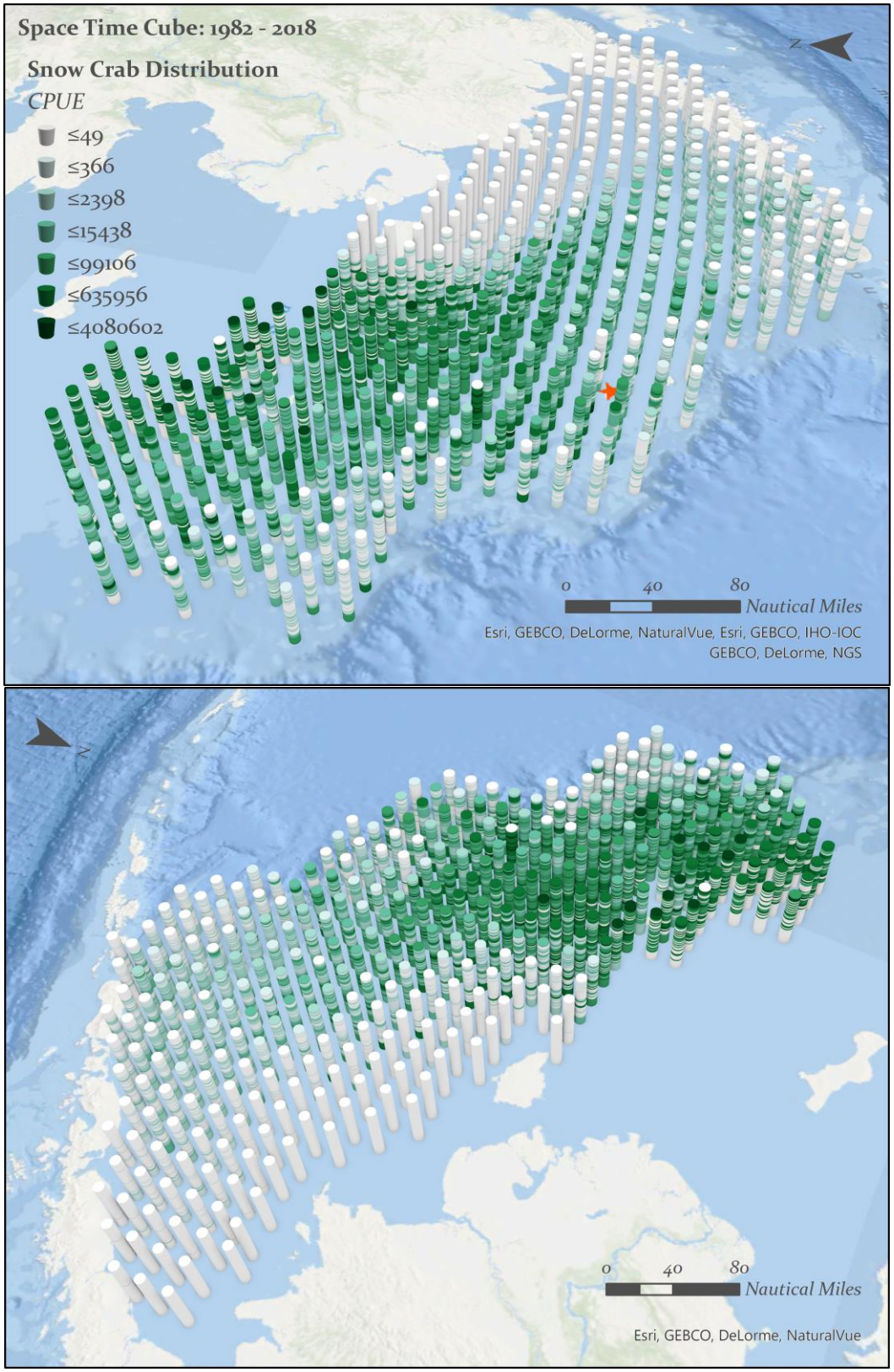


Figure 6. Space time cube views of snow crab CPUE due east (top) and west (bottom)

The highest concentration of snow crab was located north of the Pribilof Islands along the middle domain (50-100 m). Snow crab CPUE decreased from north to south along the middle domain and was lowest along the coastal domain (<50 m) and Bristol Bay region in the southeastern shelf (left foreground of bottom cube in Figure 6). The most recent survey showed that snow crab CPUE has increased in the northeastern region and continues to increase in the stations nearest the Bering Strait from west to east. These abundance gradients are inverse to bottom temperature gradients in the EBS. As snow crab abundance has increased along the northeastern front of the survey, the time cube view looking towards the east (top cube in Figure 6) shows a clearer view of CPUE trends decreasing over time along the outer domain and shelf edge.

Despite these visible spatial trends, aspatial temporal analysis of snow crab CPUE showed no significant global pattern of change in CPUE for any of the age class groups overall, or in Pacific cod (see Table 3). The Mann-Kendall statistic did show that immature snow crab numbers have increased slightly (1.2687 trend statistic) while the mature female age class has declined (-.3270 trend statistic). The total population of snow crab has increased slightly (trend statistic .5362), bolstered by the growth of the immature age class.

Table 3. Mann-Kendall data trends for CPUE, 1982 - 2018

CPUE Group	Trend Direction	Trend Statistic	p-Value
Total Crab	Not Significant	0.5362	.5918
Immature Crab	Not Significant	1.2687	.2046
Mature Female Crab	Not Significant	-.3270	.7437
Pacific Cod	Not Significant	.1962	.8445

Since 1982 there has been a weakly positive trend in Pacific cod CPUE (.1962 trend statistic) but the trend statistic for snow crab is stronger and also positive. Despite management concerns of warming sea temperatures and the ecological implications of an influx of predatorial Pacific cod populations in the EBS, the temporal trends in CPUE do not suggest a shift in ecosystem structure between invertebrate and groundfish communities. Pacific cod and snow crab do not show an inverse abundance relationship that might indicate top-down predation control of the population at this scale of analysis.

Temporal patterns of CPUE vary spatially along two axes, from north to south and from the coastal to outer domains. These gradients can be visualized through comparison of the banding patterns amongst a stratified subsample of time cube stacks spanning the shelf geographic regions (north, central, south) and domains (outer, middle, coastal). A group of stacks spanning the shelf per each northern, central, and southern survey region are shown in Figure 7. Each trio group includes one stack from the outer (>100 m), middle (50 to 100 m), and coastal domain (<50 m).

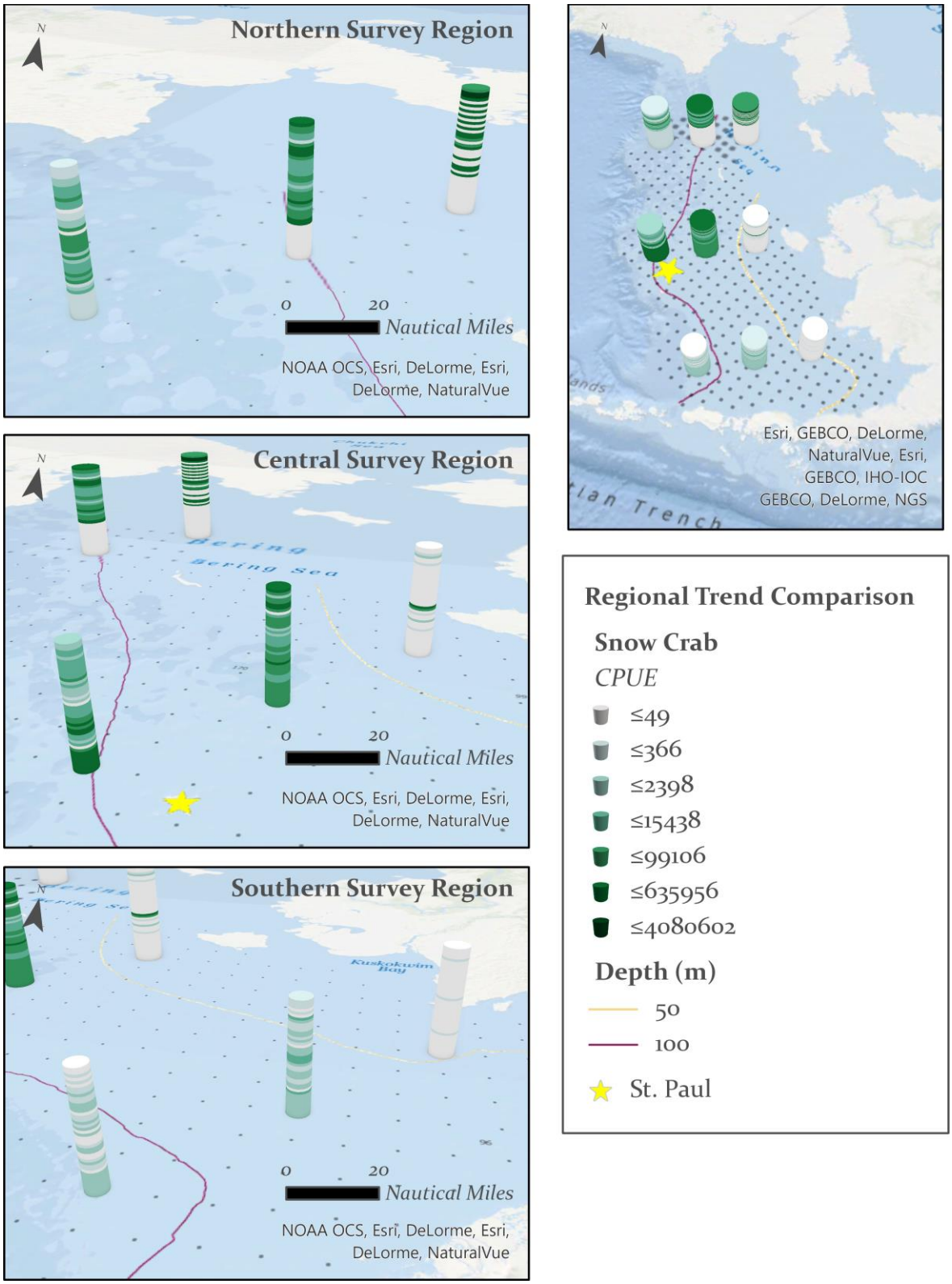


Figure 7. Stratified sample of survey station time cube stacks showing regional variation in snow crab CPUE, 1982 – 2018

The regional time cube stacks featured in Figure 7 are displayed individually and labeled with station ID and survey year in Figure 8 to make further detailed temporal comparisons. In the northern region, snow crab CPUE increased across-shelf from west to east, or from the outer to middle domain. Abundance began increasing sequentially at these stations across the shelf starting in the west or outer domain in 1985. The middle station in this northern region subsample then began to increase in 1986, and the easternmost station lagged another year before beginning to increase in 1987. This indicates a progression towards colder waters nearer the Bering Strait and concentration of the population in the northern region of the survey.

Snow crab CPUE average over the time series increases from 488 in the south (station F-07) to 80,542 in the north (station S-28), and 297,772 in the northeastern-most station closest the Bering Strait (V-25). By contrast, the two coastal locations (<50 m) recorded an average CPUE of 39 at the station closest Nunivak Island in the central region (N-01) and average CPUE of 5 in the southernmost coastal station (I-10). Temporal profiles are revisited in the time series cluster analysis results.

The peak and sustained high CPUE records of snow crab on survey from the mid-1980s (bottom half of the stack) until the steep drop in 1998 (midway up the stacks) can be seen in the banding patterns at each location. It is also of note from these individual stack visualizations that the most variation in CPUE of snow crab occurs along the oceanographic fronts of the EBS: the shelf edge along the outer domain where the slope current flows and the northeastern survey region nearest the Bering Strait where Arctic currents approach from the north. Snow crab CPUE in each of these areas fluctuate on an annual basis while the middle domain experiences more gradual change in CPUE over time. This would seem another spatial indicator that bottom temperatures maintain influence over snow crab distribution on the EBS.

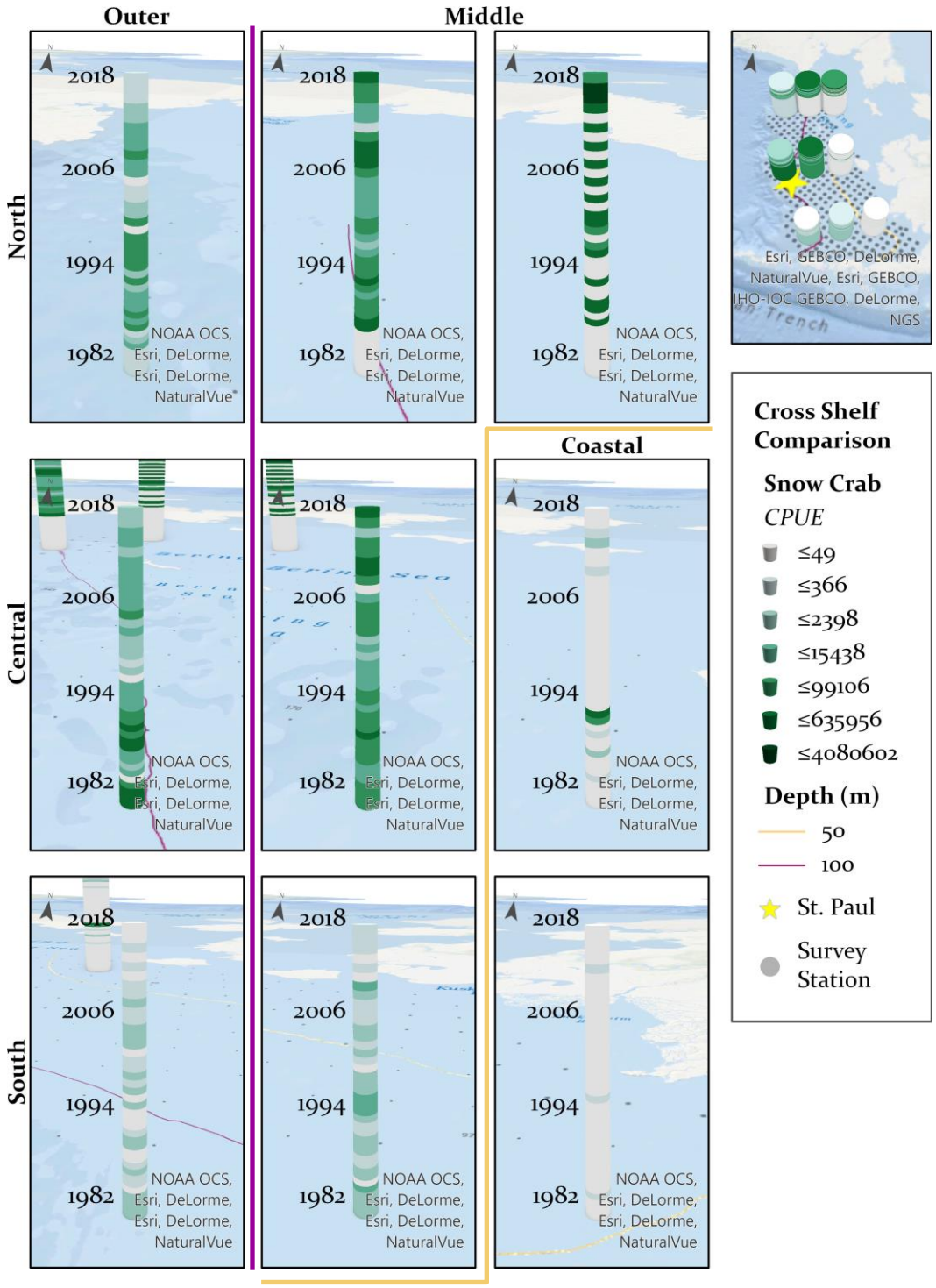


Figure 8. Individual time cube stacks from north to south and from outer to coastal domain showing range of temporal profiles of snow crab CPUE, 1982 to 2018

Space time cubes of immature and mature female snow crab distributions are shown in Figure 9. The CPUE patterns for these age classes reflect the same north-to-south gradient described in the general population but female snow crab are restricted to smaller clusters in the north flanking either side of St. Matthew Island. The most recent bins along the northwestern region of the survey in the outer domain have decreased in the latest time step in each age class.

Pacific cod do not follow the same environmental gradients observed in snow crab distributions. Figure 10 shows the space time cube for Pacific cod CPUE. Stations with higher CPUE (>7500) of Pacific cod are clustered along the coastal domain from Bristol Bay to Nunivak Island, and additional clustering of high CPUE occurred around St. Matthew Island and northeast of the Pribilof Islands.

Regional trends in CPUE of Pacific cod are shown in Figure 11 for the same subsample of time cube stacks stratified across the survey region as described for snow crab CPUE. Pacific cod abundance was historically low in the north and nearly absent in the northeast (station V-25) but abundance has increased here recently in 2016 and 2018. Pacific cod abundance was highest throughout the central survey region, but the southeast stack in Bristol Bay (station I-10) recorded the highest average CPUE of Pacific cod over the study period (8863). Pacific cod CPUE in the outer domain has decreased over the time series, similar to snow crab patterns although this down trend was only visually apparent in the southern and northern survey region time stacks in this example for Pacific cod.

This dataset and the time cubes encompass a large spatial and temporal range and the raw CPUE values can be difficult to decipher. Spatiotemporal analysis of the magnitude scale of change in hot spot Getis Ord Gi* and Anselin Local Moran's I clusters and outliers tests can better summarize this CPUE data.

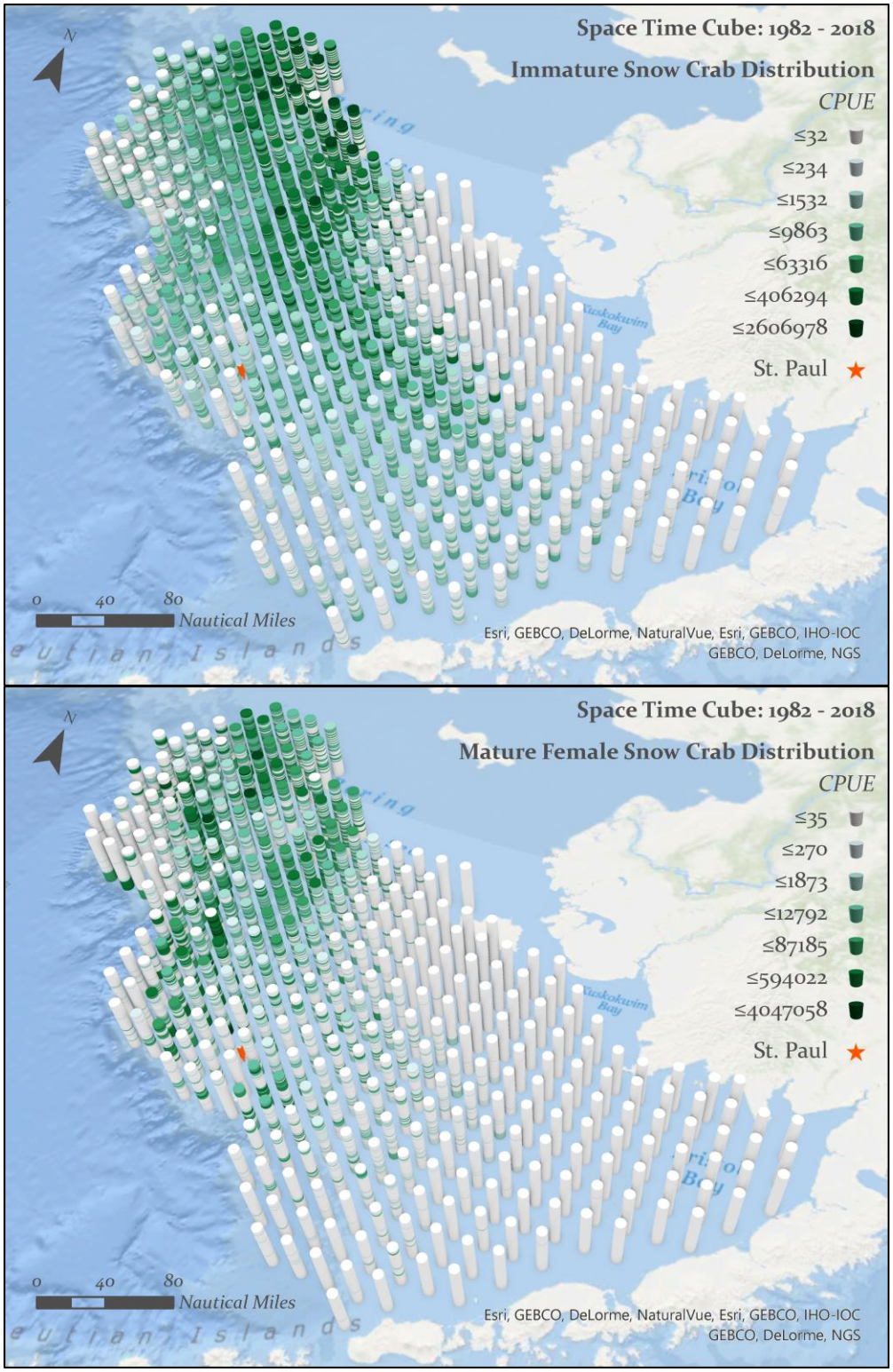


Figure 9. Space time cube views of immature snow crab (top) and mature female snow crab (bottom) CPUE, 1982 to 2018

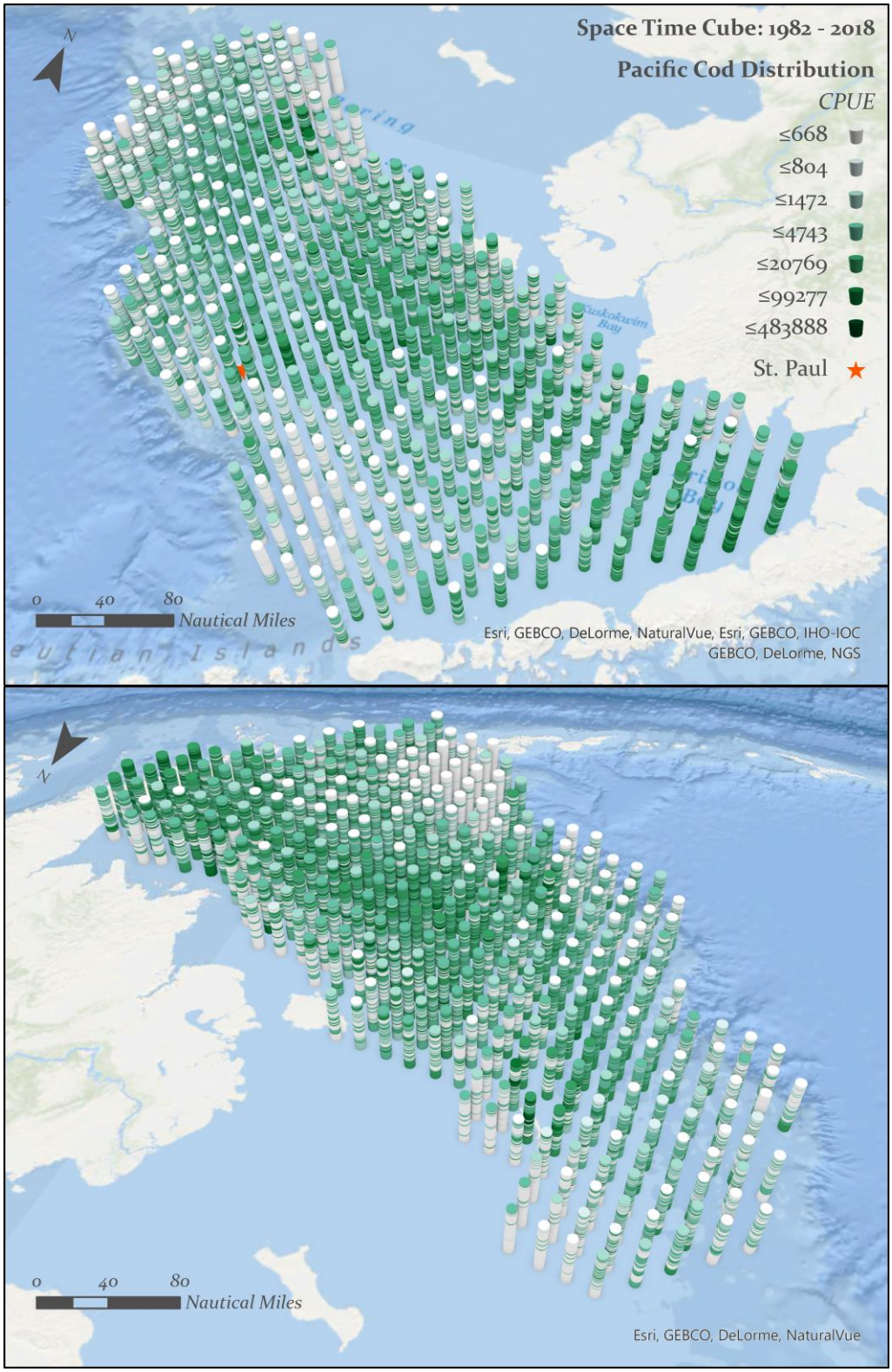


Figure 10. Space time cube views of Pacific cod CPUE due north (top) and south (bottom)

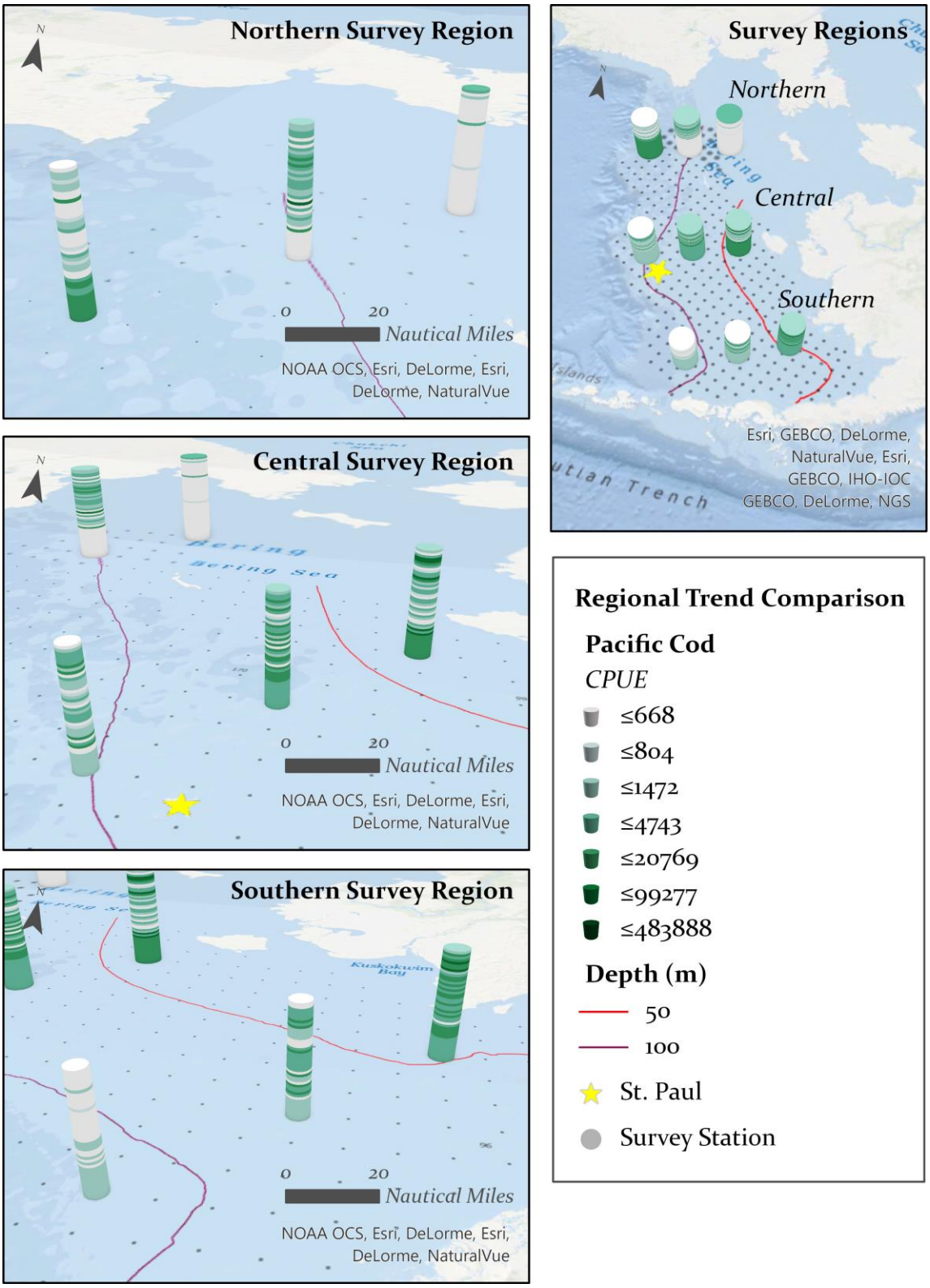


Figure 11. Stratified sample of survey station time cube stacks showing regional variation in Pacific cod CPUE, 1982 – 2018

4.1.1. Hot and Cold Spots

Survey stations in the northeastern region of the shelf showed significantly higher CPUE values recently, as seen in the cluster of hot spots (red bins) in the time cube in Figure 12 (top). Snow crab abundance at these 75 stations was significantly higher than the survey average, and CPUE patterns were sporadic or intermittently high throughout the series. No cold spots were detected at this scale of analysis applying a 45 nm spatial neighborhood.

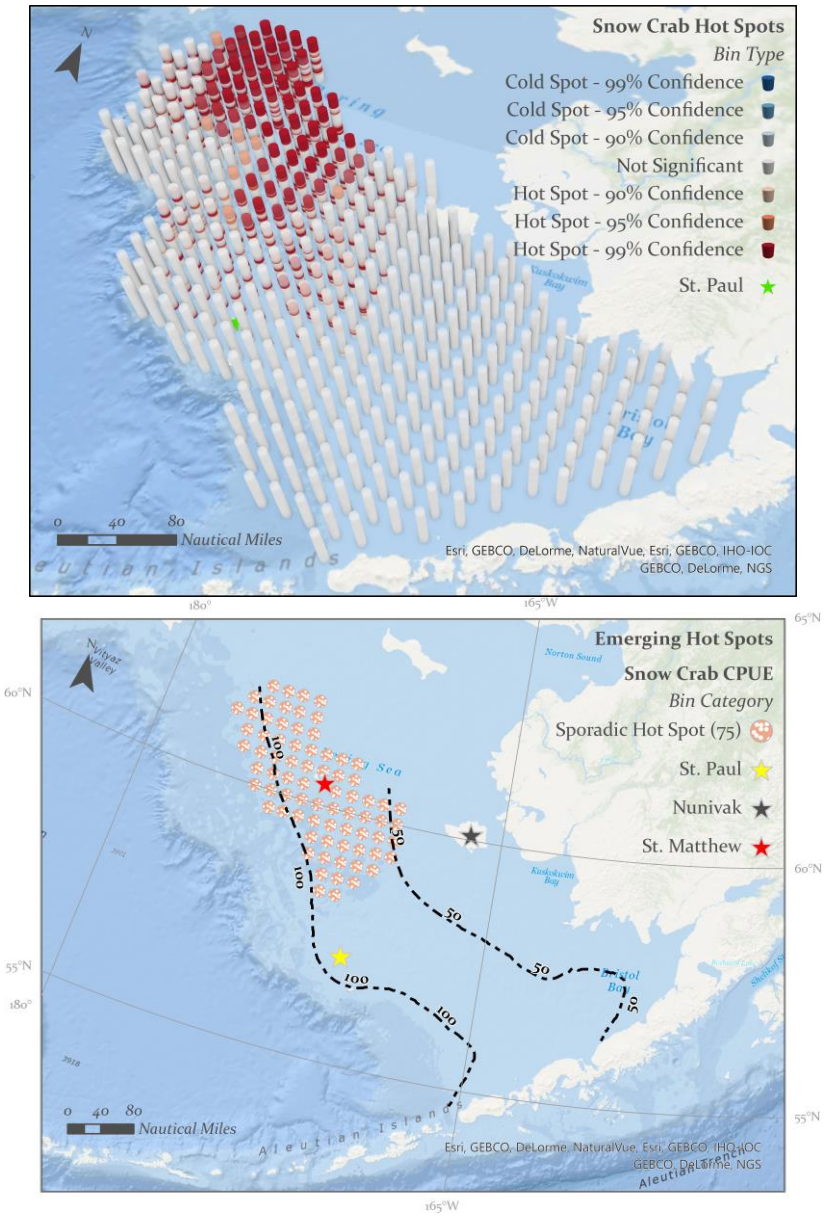


Figure 12. Space time cube 1982 - 2018 snow crab CPUE hot spots (top), with emerging hot spot trend summary (bottom)

Space time cubes showing CPUE hot spots for mature female and immature snow crab age classes are shown alongside their corresponding 2D emerging hot spot summary in Figure 13. The immature age class results closely resemble the patterns described by the total population, but the cluster is reduced to 64 stations. Two clusters of sporadic hot spots can be seen in mature female snow crab CPUE that flank either side of St. Matthew, and new hot spots have emerged only recently in this age class on the eastern flank of these hot spot clusters, nearest the Bering Strait. These patterns reflect the spatial stratification described in previous research that arises from settlement and migration patterns and movement from east to west so that immature crab move to the colder domain and eventually migrate to the west.

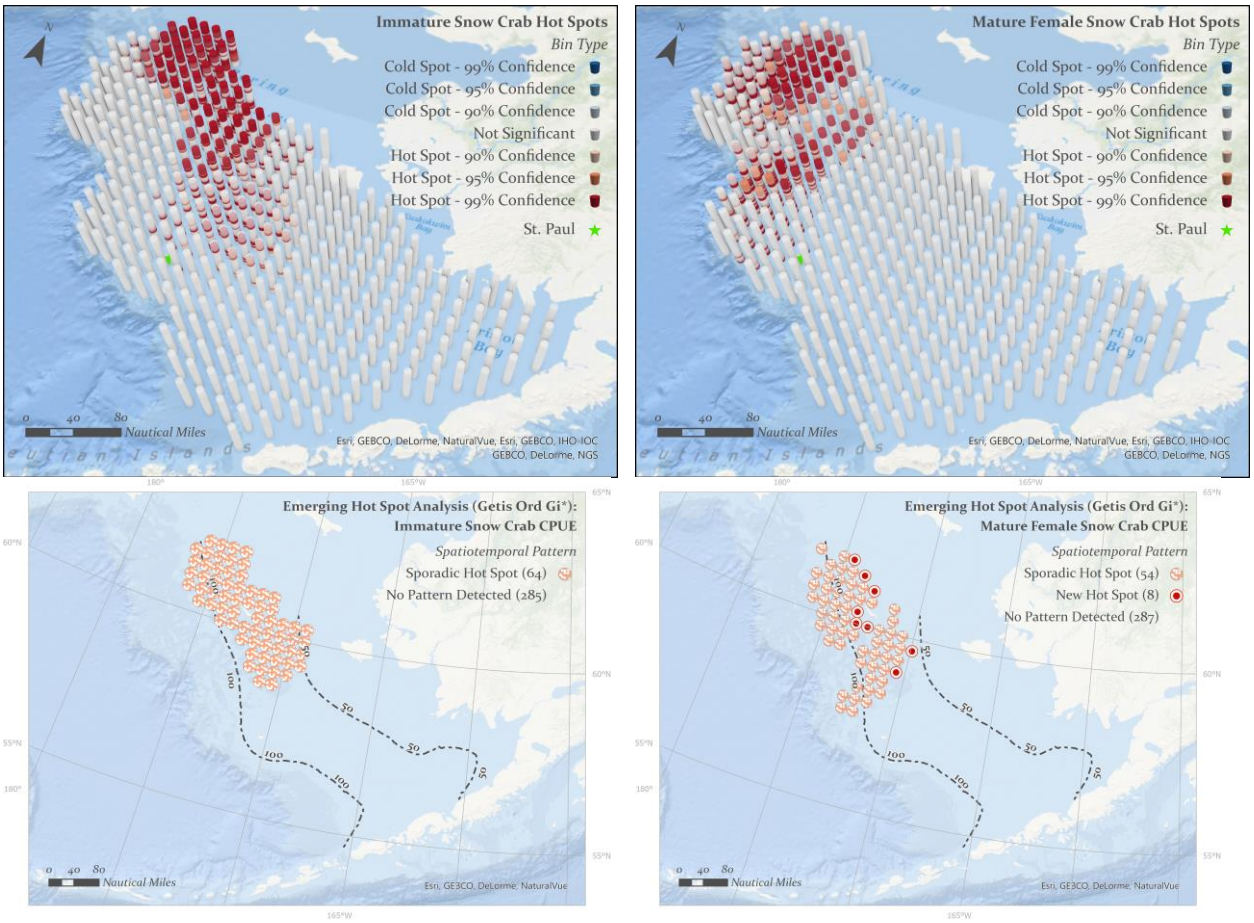


Figure 13. Snow crab CPUE hot spots for immature and mature female age classes, time cube (top) and corresponding 2D emerging hot spots temporal summary (bottom), 1982 – 2018

No cold spots were identified despite significant down-trends in CPUE captured by the Mann-Kendall temporal trend test. Reasons are likely related to the heteroskedasticity in snow crab abundance patterns, or the variation in CPUE variance between southern and northern regions. Snow crab CPUE in the north could reach in the millions while in the south the range was in the hundreds and thousands. So, despite significantly down-trending CPUE records in the south, the intensity of this decrease was too weak to be detected as a cold spot due to extremely high CPUE fluctuations in the northeastern hot spots. If exploring spatiotemporal patterns further, the spatial neighborhood should be expanded to enable detection of more subtle variation in the south. A second option would be to use the results of the GWR to explore spatial analysis units so that the study area was broken up into ecological units that reflect the relationships patterns described by local variable coefficients. These possibilities for further development are revisited in the local clusters and outliers analysis and GWR results, and in the discussion.

Pacific cod CPUE hot spots were detected throughout the central region surrounding Nunivak Island but were mostly restricted to a parallel band along the southern coastal domain and Bristol Bay region in the southeast (see Figure 14). No significant temporal trends were detected in the emerging hot spot analysis of Pacific cod CPUE and there were no hot spots in the last time step (2018).

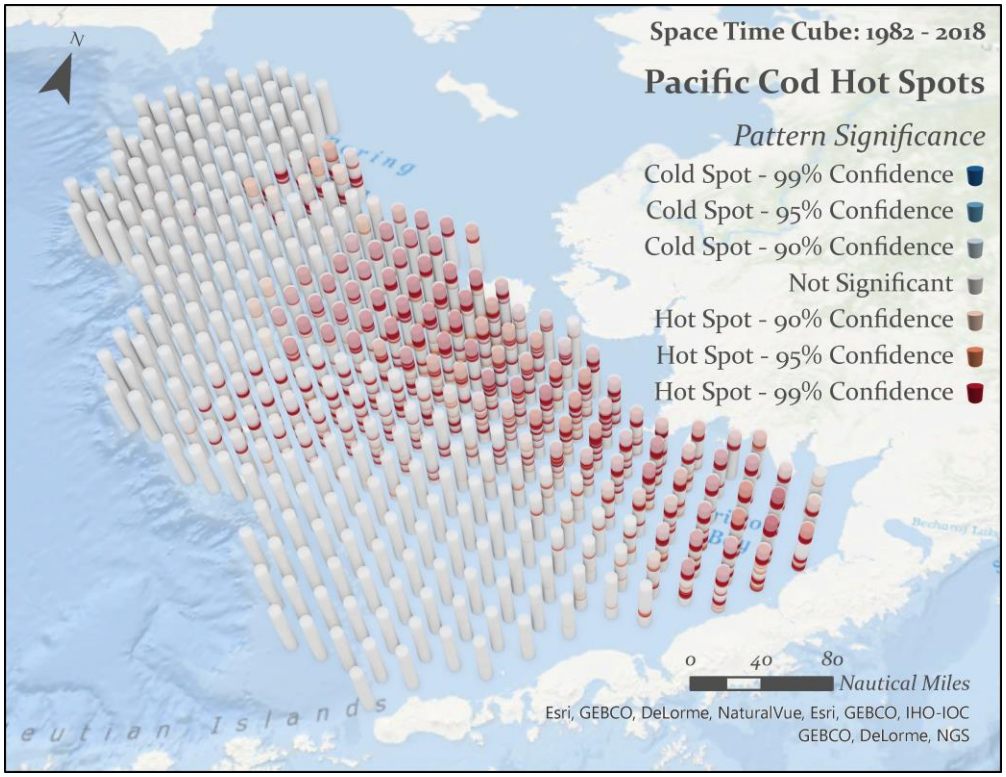


Figure 14. Space time cube showing Pacific cod CPUE hot spots, 1982 to 2018

As described previously in reference to the lack of cold spots detected in snow crab CPUE over the time series, a lack of cold spots in Pacific cod CPUE may be accurate or the analysis may have failed to identify relatively weak negative trends, or an inappropriate spatial neighborhood may have been specified. The distance band (45 nm) in this study was developed to optimize analysis of the total snow crab age class CPUE patterns and demonstrate the method applied in GIS, but each group could be investigated independently to determine a more appropriate spatial neighborhood for the species or age class of interest.

The spatial variation and timing of Pacific cod CPUE hot spots were visualized in the regional subsample of survey stations (Figure 15). There were no hot spots for snow crab CPUE in any of the southern region survey stations (C-04, F-07, I-10). One hot spot occurred near the start of the series in 1983, in the central region stack nearest St. Paul in the outer domain (H-23).

Hot spots then appear in 1990 in the central-middle stack (K-20), closely followed by hot spots in 1993 in the northern stations of the middle domain (S-28 and V-25) and the central-coastal station nearest Nunivak (N-01). The central-middle stack was identified as a hot spot again in 2005, and hot spots occurred throughout the northern stacks about 2011. The two northeastern stations of the middle domain (S-28 and V-25) have both been classified as snow crab CPUE hot spots over the last few time steps (since 2016 and 2015). These trends are captured in the sporadic hot spots in the northeast survey region and describe the northward shift and contraction of the snow crab population towards colder temperatures.

Pacific cod CPUE hot spots are shown for the sample of survey stations in Figure 16. There were no hot spots in any outer domain locations, or in the northern survey region. The first hot spot occurred in the central-coastal domain (N-01) in 1982, followed by a short bloom over two years from 1993 to 1994 in the southeastern Bristol Bay region (I-10) and central-middle stack (K-20). The next hot spot of Pacific cod CPUE appeared in 2001 in the southeastern Bristol Bay region again (I-10). A prolonged hot spot was identified between 2011 and 2016 spread amongst the coastal domain stations (N-01 and I-10) and the central-middle stack to a lesser extent (2014 to 2016).

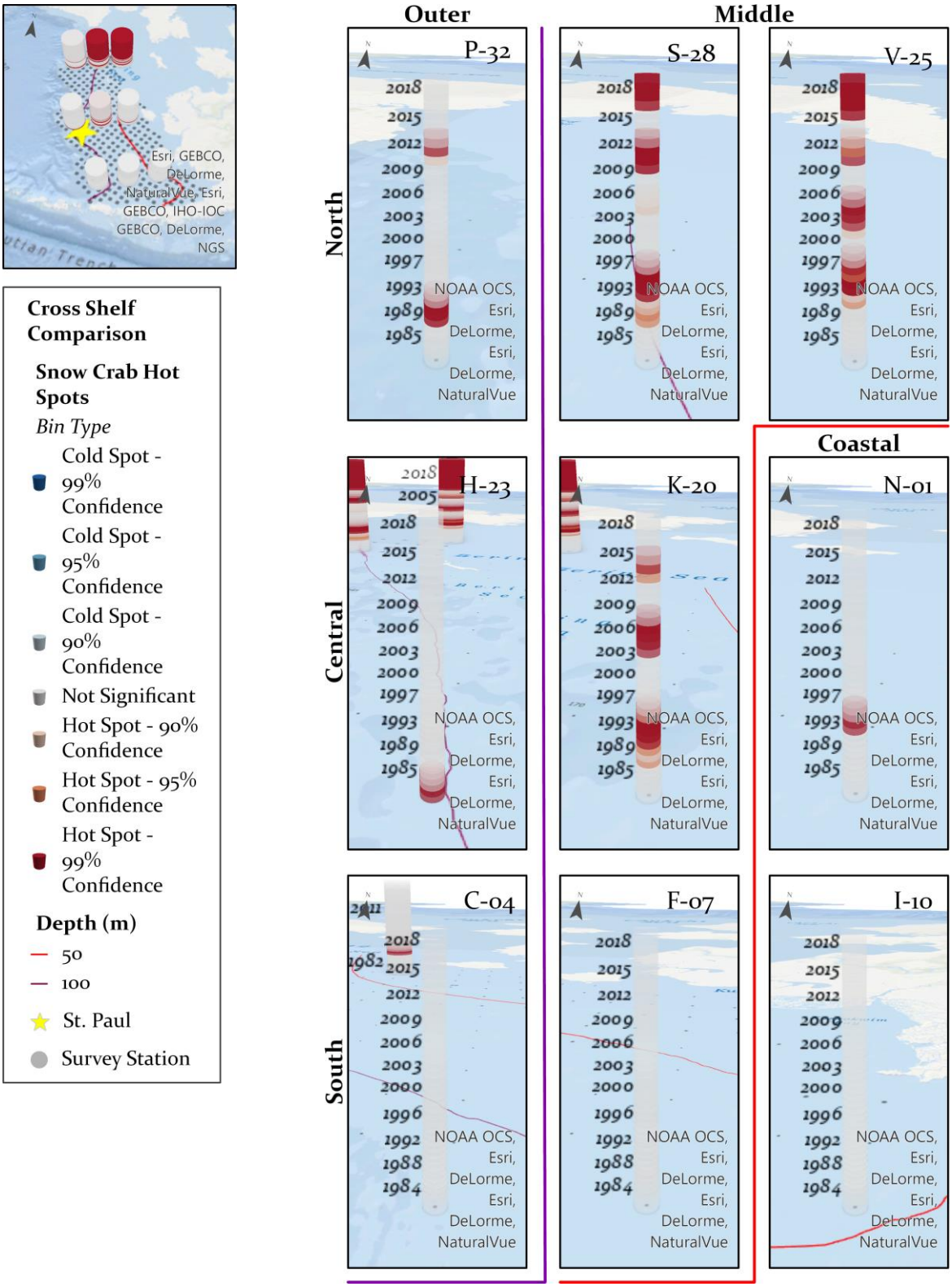


Figure 15. Stratified sample of time cube stacks across the EBS shelf regions and domains showing hot spots of snow crab CPUE, 1982 - 2018

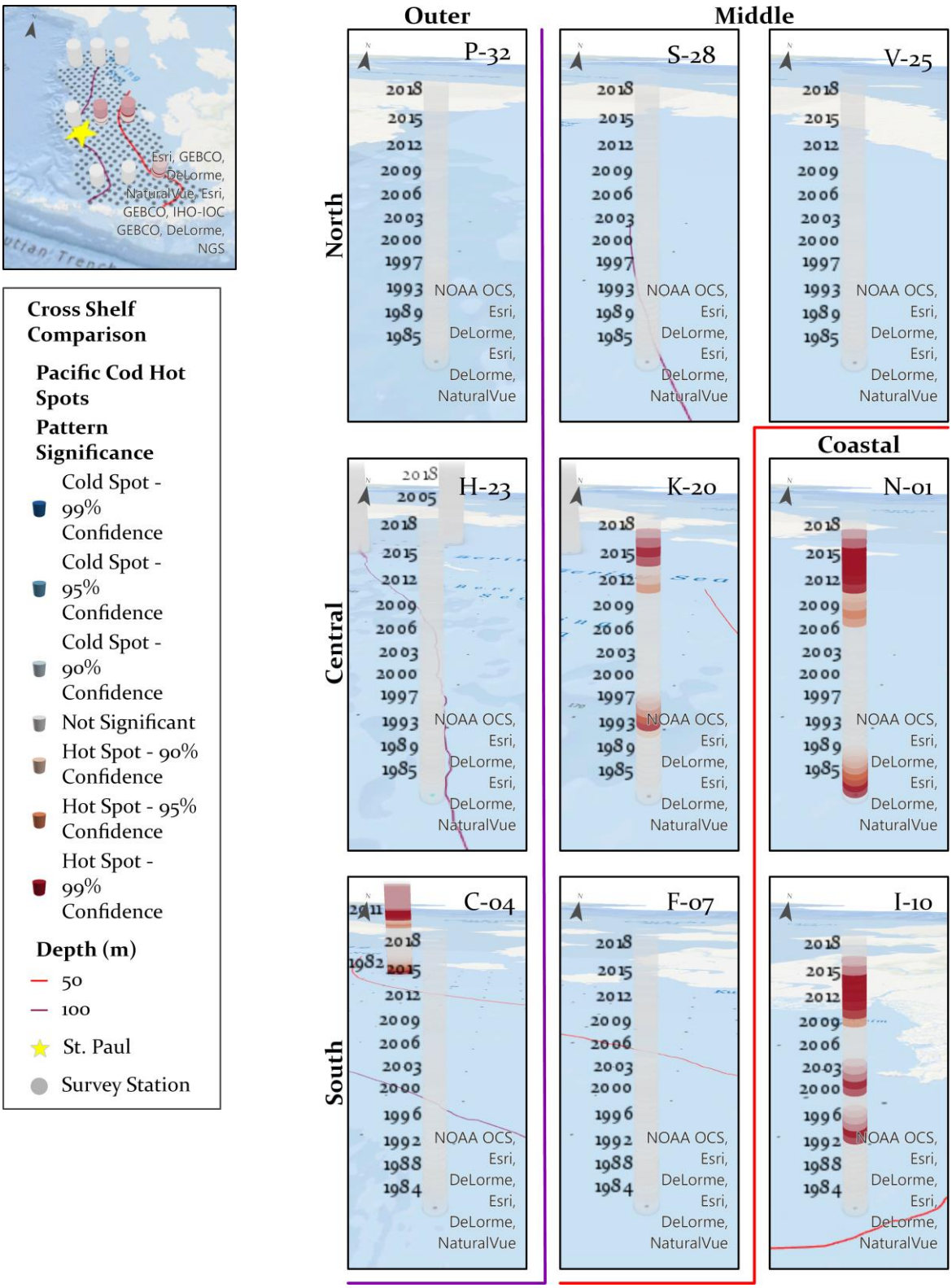


Figure 16. Stratified sample of time cube stacks across the EBS shelf regions and domains showing hot spots of Pacific cod CPUE, 1982 - 2018

The Mann-Kendall trend tests for each snow crab age class and Pacific cod CPUE over the time series revealed regions of up- and down-trending abundance. Figure 17 shows the 2D summary of temporal trends in the time cube at each station location for each age class of snow crab and Pacific cod.

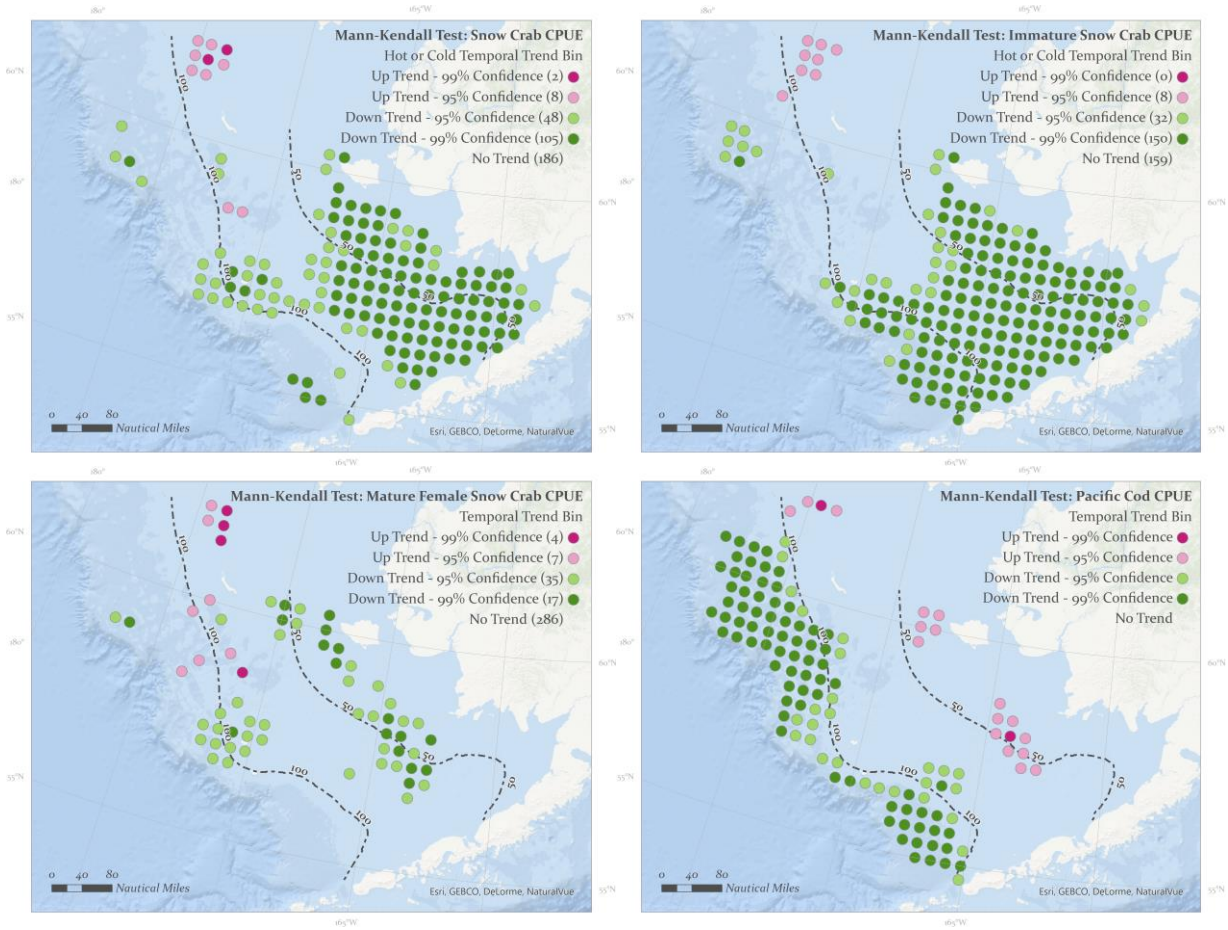


Figure 17. Temporal trends in CPUE for total snow crab (top left), immature snow crab (top right), mature female snow crab (bottom left), and Pacific cod (bottom right), 1982 to 2018

Down trends were detected with 99% confidence in 150 of the 349 total survey stations for immature crab CPUE and in 105 of 349 stations for the total population of snow crab, yet overall, the global statistics reported previously in Table 3 were positive for both groups (.5362 and p=.5918 for total, 1.2687 and p=.2046 for immature class). This correlation test does not

reflect the magnitude of the trend, which was slight because CPUE of snow crab has been historically and uniformly low in the southern region. The decreasing trend here was relatively insignificant compared to the global dataset trends that are more heavily influenced by the hot spots in the northeast. The down trends in snow crab CPUE in the south were too small in scale to be identified in the hot spot analysis as significantly cold.

Pacific cod temporal trends were also spatially variable and disparate between the Mann-Kendall trend test and hot spot analysis. The outer domain was classified entirely as down-trending in Pacific cod CPUE despite a lack of cold spots and no significant global trend (trend statistic for Pacific cod was slightly positive, .1962 with p=.8445). A group of four stations along the northern fringe of the survey region was categorized as up-trending; these may represent the most recent trends of increased CPUE of Pacific cod in the north that is not intense enough to be identified as hot spots. Snow crab also showed up-trends in the north corresponding to hot spot locations which further supports the ecological hypothesis of northward species shift.

4.1.2. Time Series Clusters

Time series correlation revealed four spatial clusters of survey stations with similar temporal profiles of CPUE of snow crab in terms of the value and proportionate change over time. Cluster trend statistics are provided in Table 4 with corresponding clusters mapped in Figure 18. Cluster 4 showed a significant increase in snow crab CPUE over the time series (1.6872, p=.0916) and was located in the northeast region of the survey where hot spots were detected and the Mann-Kendall tests identified up-trends. This was the only significant trend detected in any of the time series cluster groups (tested up to 7 clusters). Clusters 2 and 3 were both decreasing but the trends were not significant. These stations were located to the east of Cluster 4 along the outer domain and south across the southern shelf. Twenty-nine stations

identified as Cluster 1 in Figure 18 recorded zero CPUE of snow crab every year of the time series. These stations along the easternmost coastal domain were removed from the dataset prior to regression analysis as described in the Methods section. The issue of data transformation and analysis units is revisited in the Discussion.

Table 4. Time series cluster trend statistics for snow crab CPUE temporal profile correlation
**Indicates statistically significant trend*

Cluster ID	Direction	Statistic	p-Value	Locations
1	Not Significant	0.0000	1.0000	29
2	Not Significant	-1.5302	0.1260	125
3	Not Significant	-0.6932	0.4882	102
4	Increasing*	1.6872	0.0916	93

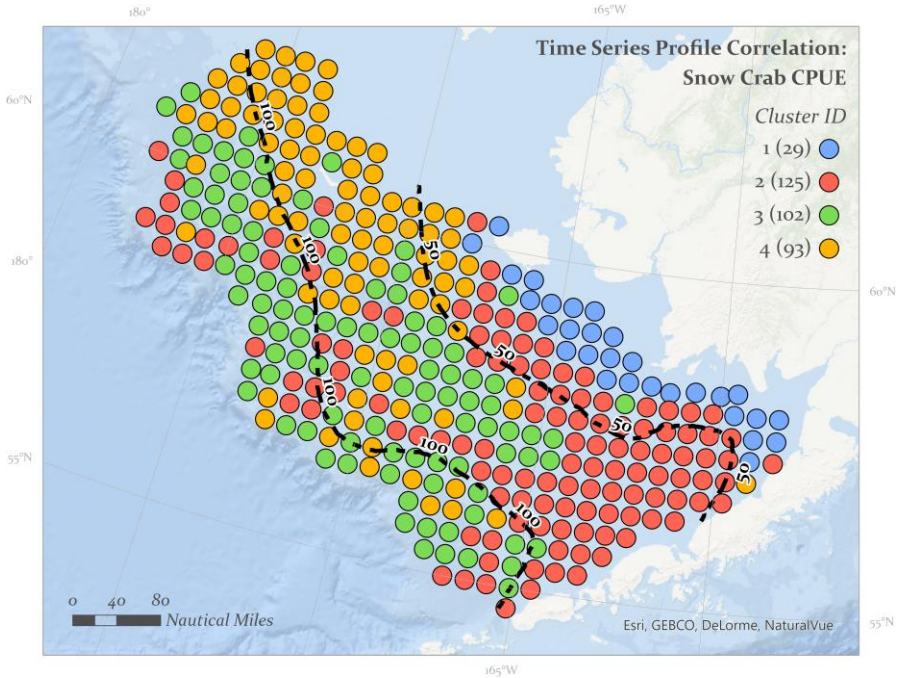


Figure 18. Time series clusters, four groups of survey stations with correlating temporal profiles of snow crab CPUE, 1982 to 2018

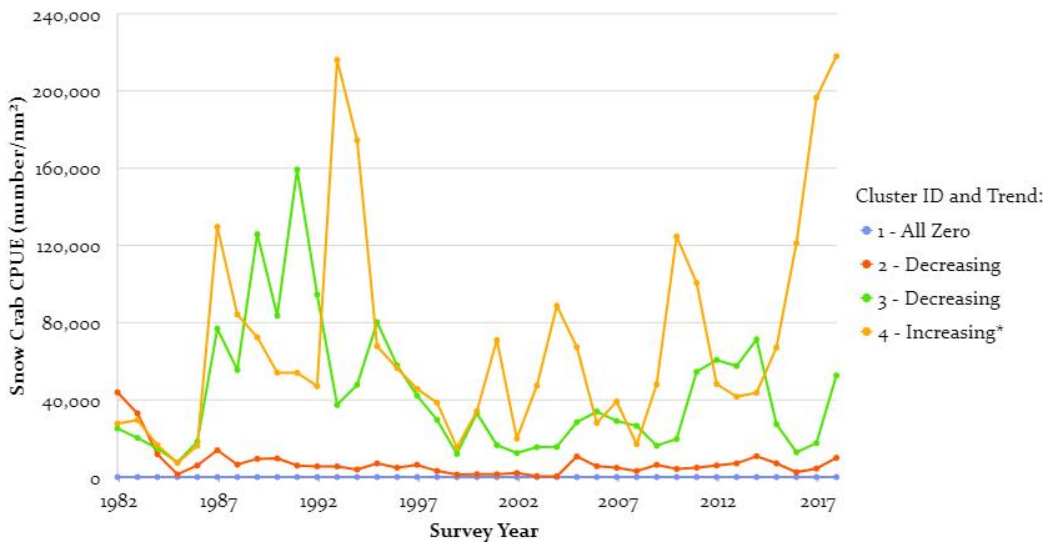


Figure 19. Average snow crab CPUE for each time series cluster group
* Indicates significant trend

Time series cluster analysis identified two clusters of similar temporal profiles in Pacific cod CPUE (see trend statistics in Table 5). The spatial clustering pattern reflects the trends seen in the hot spot analysis and Mann-Kendall temporal trends, with decreasing CPUE along the outer domain (Cluster 1 in Figure 18). Although Cluster 2 was increasing over the time period, these stations were not identified as hot spots due to the recent decline in Pacific cod CPUE which has brought the average for each cluster closer together nearest a historical low of 2,000 CPUE in 2018.

Table 5. Time series cluster trend statistics for Pacific cod CPUE temporal profile correlation
* Indicates statistically significant trend

Cluster ID	Direction	Statistic	p-Value	Locations
1	Decreasing*	-2.8120	0.0049	210
2	Increasing*	2.5242	0.0116	139

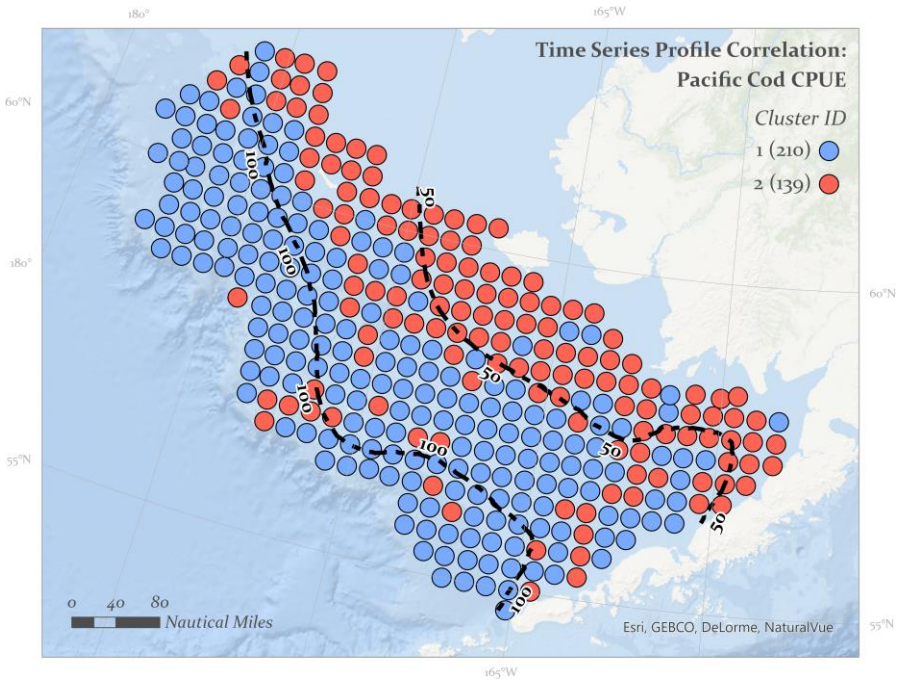


Figure 20. Time series correlation in temporal profile for Pacific cod CPUE, 1982 to 2018

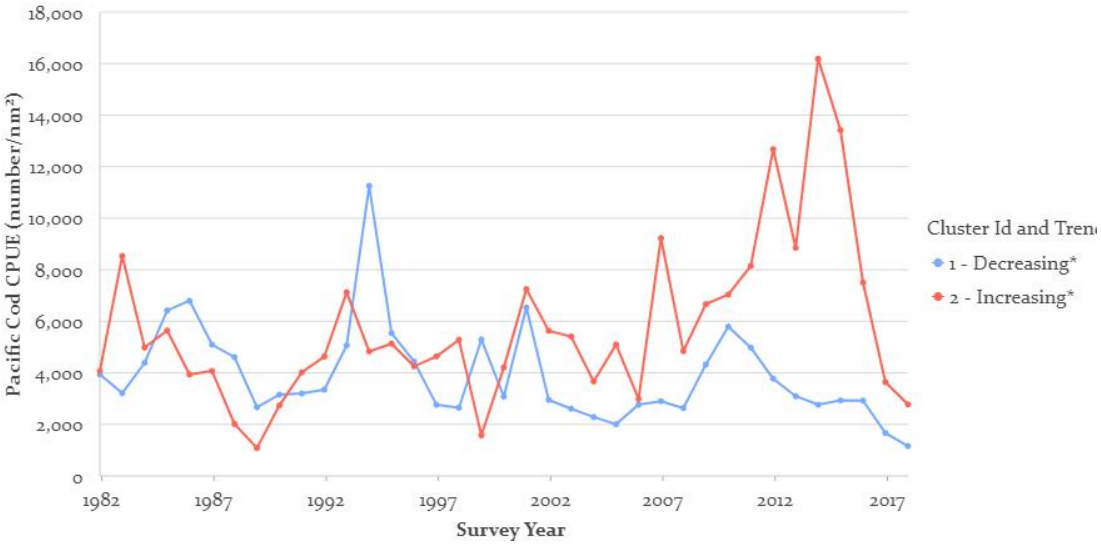


Figure 21. Average Pacific cod CPUE for each time series cluster *(*indicates significant trend)*

Pacific cod abundance peaked in Cluster 2 between 2012 and 2016, years identified in the coastal domain as hot spots. These time series clusters describe spatial heteroskedasticity in the data, or unequal change in the variables (CPUE) across space in both the snow crab and Pacific cod CPUE data.

4.1.3. Local Clusters and Outliers

A large cluster of low CPUE of snow crab was identified by the Anselin local Moran's I statistic for nearly all survey stations located south of St. Paul and Nunivak. Time cube views for all clusters and outliers results for age class of snow crab and Pacific cod are shown in Figure 22.

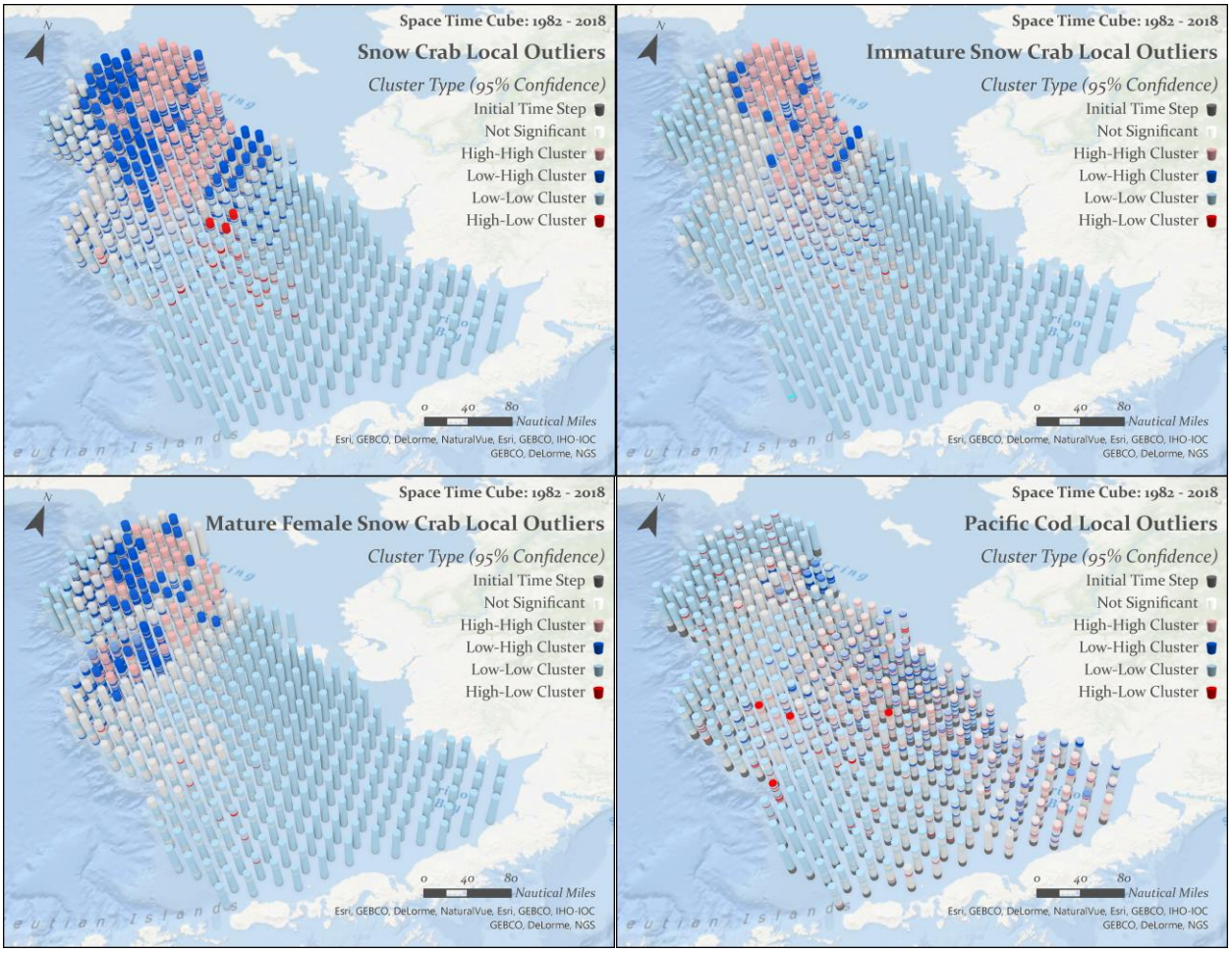


Figure 22. Time cube views of CPUE clusters and outliers for snow crab (top left), immature snow crab (top right), mature female snow crab (bottom left), and Pacific cod, 1982 to 2018

Low-low clusters (low CPUE stations near other low CPUE stations) made up 49% of the snow crab time cube bins overall (6,031 out of 12,913), and 50% of all survey station locations (175 of 349). Snow crab CPUE within these clusters of stations decreased over the time series,

but this was an insignificant trend overall for the reasons related to scale and magnitude of change as previously described in the hot spot and Mann-Kendall comparisons. The local clusters and outliers analysis for snow crab and Pacific cod is summarized in 2D in Figure 23. The southern trio of individual time cube stacks ranged from 0 to 18,817 CPUE. Variance in this region ranged from 360 to 535,936 and increased by several orders of magnitude in the northern region of the survey (up to 60,480,131,868 at station V-25) where CPUE fluctuated between 25,204 and 673,285. This heteroscedasticity has been visualized in the hot spot analysis regional comparison in Figure 15 in which the timing, extent, and frequency of hot spots varies spatially across the shelf domains and from north to south.

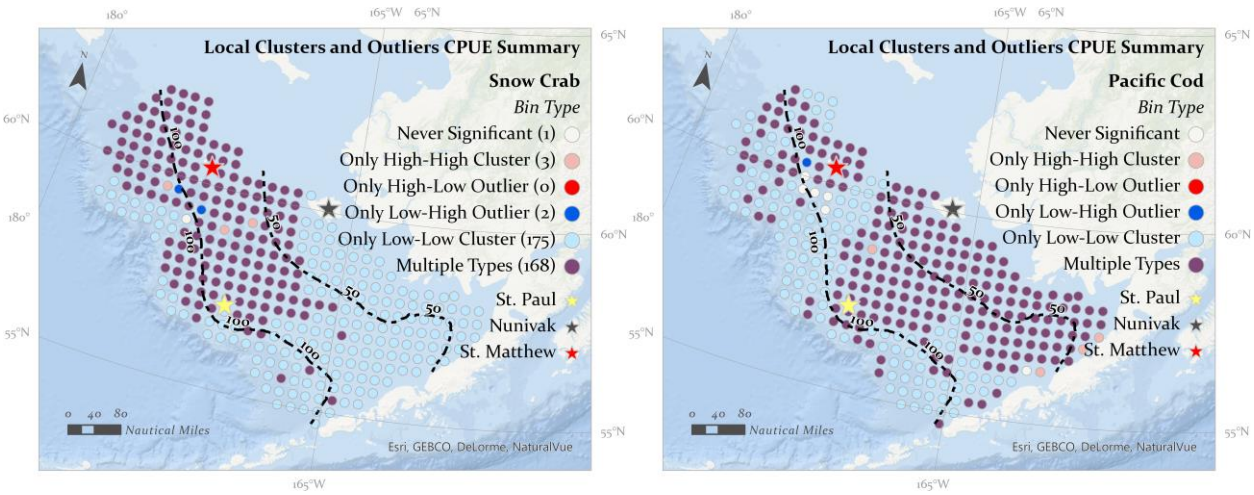


Figure 23. Summary of CPUE clusters and outliers for snow crab (left) and Pacific cod (right), 1982 – 2018

There were 175 survey stations in the southern region categorized as only low-low clusters for snow crab CPUE, and 168 as multiple type where the correlation was weaker and CPUE variance was higher. Correlation in CPUE of Pacific cod was highest along the outer domain in 119 stations categorized as low-low clusters and 219 as multiple type clusters. Low-low clusters formed a transverse corridor across the shelf of correlated CPUE in stations that

stretched from the south side of St. Matthew in the east to the southern flank of Zemchug Canyon in the west. These low-low clusters divided the survey area into two clusters of multiple type category that follow the cross-shelf depth gradient as opposed to north-south distribution patterns seen in snow crab. These visual comparisons of CPUE trends and relationships are explored quantitatively in the regression analysis.

4.2. Lagged Relationships

The top five variables from past years in terms of lagged correlation with 2018 CPUE of snow crab since 2006 are listed in the vertical timeline in Table 6, based on the exploratory regression (OLS) testing all independent variables separately. The top three most significant years of impact per each independent variable (bottom and surface temperatures, immature and mature female age classes CPUE, and Pacific cod CPUE) are outlined in Table 6 and were included in the global OLS regression that follows. Positive correlation is presented to the right of the timeline, and negative correlation to the left.

The highest significance in lagged mature female snow crab CPUE was identified for the 2006, 2008, and 2017 classes (100% associations). This suggests connectivity between maternal age classes from 2006 to 2008 and progeny that have grown to and now constitute the total snow crab population in 2018. The more recent correlation in 2017 is likely temporal correlation, which was similarly identified in the immature age classes in 2017.

Table 6. Top three most significant lagged years (2006 and 2018)

Lagged Independent Variables	Snow Crab Population Group (*approximate life history stage*)
2017 Mature Female Snow Crab CPUE 2017 Immature Snow Crab CPUE 2016 Immature Snow Crab CPUE 2016 Surface Temperature	Total Population *(mature reproductive stage)*
2015 Bottom Temperature 2014 Pacific Cod CPUE 2014 Bottom Temperature 2013 Bottom Temperature 2012 Pacific Cod CPUE 2011 Surface Temperature 2010 Surface Temperature	Immature Age Classes *(growth and development stage)*
2008 Mature Female Snow Crab CPUE 2007 Immature Snow Crab CPUE 2006 Mature Female Snow Crab CPUE 2006 Pacific Cod CPUE	Mature Female Age Classes *(egg extrusion stage)*

The years 2013 likely represent the transition period of growth from early benthic stages to mature adults in 2018. Immature snow crab are highly stenothermic and typically aggregate in the middle domain where there are colder temperatures (1°C). These were the most significant impact years in the timeline for bottom temperature, further supporting evidence outlined in previous research into the life history cycle and ecological niche differences between snow crab age classes.

Lagged results for surface temperature showed positive and negative associations and are more difficult to interpret but the significance of this variable is clustered between 2010 and 2016. These years correspond to the warming phase described in the time series charts of average EBS sea temperatures at the beginning of this chapter. The exploratory regression also produced mixed results for lagged impact of predation but the top three most significant years in the timeline for the Pacific cod CPUE variable (2006, 2012 and 2014) expressed an inverse

correlation with snow crab CPUE. These years correspond to maternal age classes and the early (more vulnerable) benthic stages of snow crab development.

Snow crab CPUE was most significantly correlated with bottom temperature in the lagged years that would have coincided with immature life history stages (2013 – 2016) .

The shape of the cold pool (<2°C) in 2016 and 2017 reflects the spatial distribution of hot spots observed in snow crab CPUE in Figure 11 and the similarity in these years likely contributes to temporal correlation detected in immature and mature female age classes. There was no cold pool formation in 2018. Only seven stations along the northeastern edge of the survey reached a summer low of 1.6°C, which may impact the results of the regression analysis if bottom temperatures drive snow crab distribution and the gradient has broken down. EBS bottom temperatures were mapped for each year included in the lagged regression analysis, 2006 to 2018, and figures are provided in the Appendix.

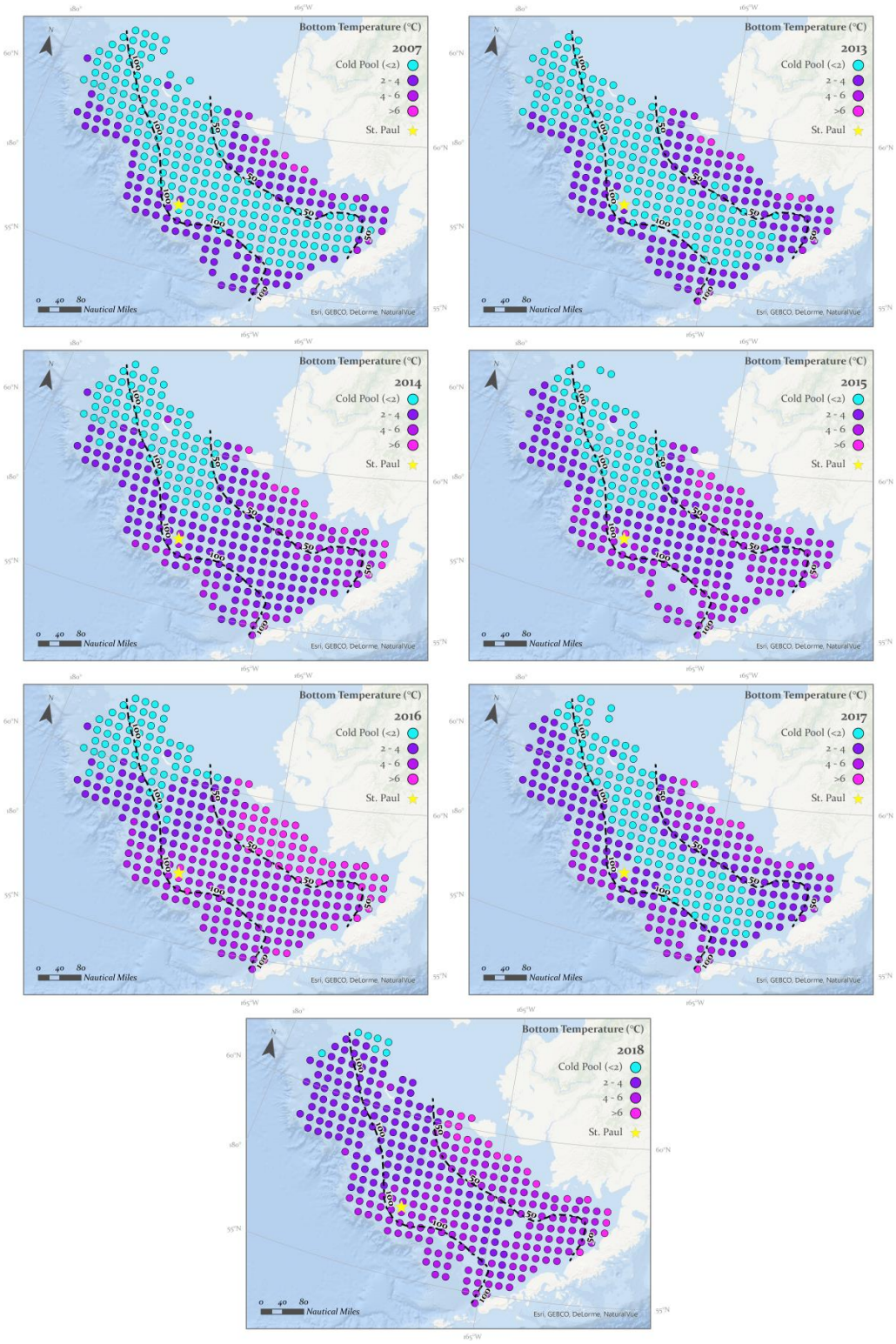


Figure 24. EBS bottom temperatures for the most significant lagged impact years for 2018 snow crab CPUE since 2006, with 2018 bottom temperature as a reference (bottom)

4.3. Global Relationship Trends

The global regression model was first tested without including lagged independent variables identified in the exploratory regression. The OLS restricted to 2018 variables identified depth, bottom temperature, and surface temperature as significantly related to snow crab CPUE. The relationship with Pacific cod CPUE was positive and not significant, contrary to a presumed negative impact. Summary results for each of the independent variables in the OLS (excluding lag) are presented in Table 7 and show that bottom temperature has the strongest (negative) relationship.

Table 7. Summary of OLS results model of 2018 snow crab CPUE, excluding lagged variables
** Indicates a statistically significant relationship*

Variable	Coefficient	Std. Error	t-Statistic	p-Value	VIF
Intercept	5.5649	3.0341	1.8341	0.0676	-
Depth	-0.0658	0.0088	-7.4375	0.0000*	2.0837
Bottom Temperature	-2.6118	0.2627	-9.9405	0.0000*	1.6161
Surface Temperature	1.8950	0.2685	7.0573	0.0000*	2.5039
Pacific Cod CPUE	0.2084	0.1998	1.0430	0.2978	1.2226

OLS model diagnostics in Table 8 include results of the global regression test with and without including the lagged independent variables identified in the exploratory regression. Interestingly, the Koenker (BP) statistic was not significant when lagged variables were excluded, so the relationships between snow crab CPUE in 2018 and the independent variables were determined to be spatially consistent. The BP test statistic was significant when lagged independent variables were included, indicating inconsistent relationships. For this reason the robust probability and Wald Statistic values were relied upon to determine coefficient significance for the regression results including these lagged variables.

Table 8. OLS model diagnostics for 2018 snow crab CPUE, with and without lagged variables
** Indicates a statistically significant statistic*

Statistic	No Lag	With Lag
Number of Observations	308	248
AIC_c	1654.8338	1108.1404
Multiple r^2	.4922	.7834
F/Wald	73.4247 (F)*	2056.1117 (Wald)*
Koenker (BP)	9.3174	39.4512*
Jarque-Bera	0.8949	5.3024

Jarque-Bera statistics were not significant in either regression (with or without lagged variables). Therefore, the model residuals were normally distributed or not clustered or significantly biased. Model residuals when excluding lagged variables were relatively small (range of 19 from -6 to 13) with some underpredicting in the middle domain where snow crab CPUE was higher. The range in residuals was further reduced including lagged variables (range of 9, from -3 to 6). Standardized residuals for each of the OLS regression models are shown in Figure 25.

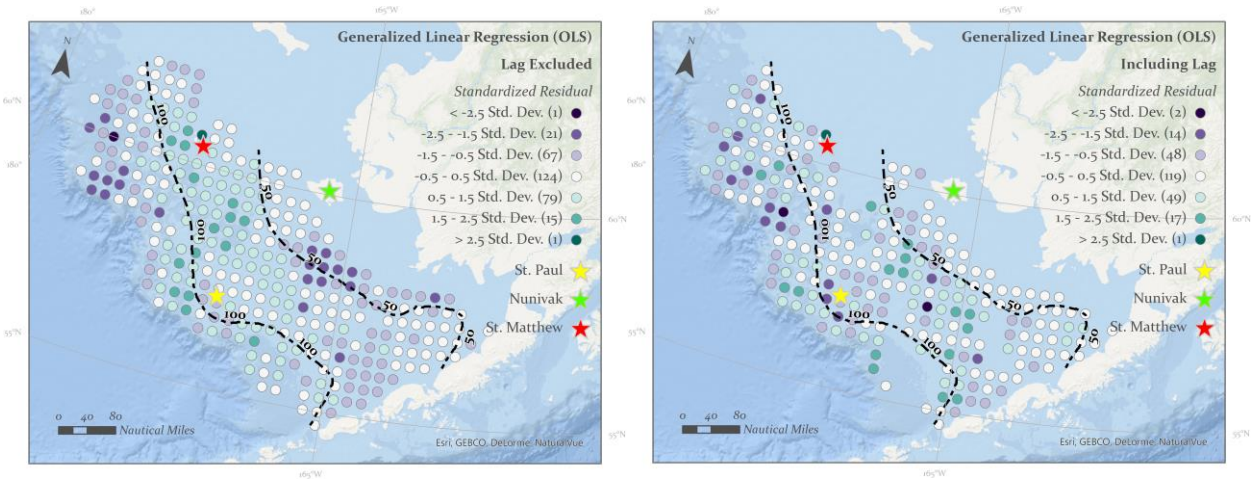


Figure 25. OLS standardized residuals for snow crab 2018 CPUE without (left) and with (right) lagged independent variables since 2006

As a measure of model performance including lag significantly reduced the AIC_c score (from 1655 to 1108) and increased model accuracy and r^2 from .49 to .78 (% variance explained). Due to missing surface and bottom temperature records in the dataset the predicted results when including lagged variables were reduced to 248 survey stations; this has limited the efficacy of including lagged variables in the regression despite the improved model accuracy and performance despite the heavier dependency detected in temporal correlation with historical CPUE. An alternate source of surface and/or bottom temperatures would improve the results of the regression analysis. A summary of the OLS results including lagged year variables is provided in Table 9.

Bottom temperature in 2018 was not significant according to OLS when including lagged variables. Surface temperatures in 2011 and 2016 were the only significant temperature variables identified. As previously stated, 2018 was an historically warm summer and no sea ice formed the prior winter. Typical temperature associations are likely confounded by this change but the lagged impact of the spatial and temporal correlation in previous years distributions of snow crab populations supersedes the bottom temperature association when including lagged variables. One other weakness in the model was a significant VIF for depth (> 7.5), indicating multicollinearity. As there was very little variation in the high bottom temperatures observed in 2018, these two variables likely expressed greater collinearity than average years.

Table 9. Summary of OLS regression variable coefficients including top 3 lagged independent variables from exploratory regression
** Indicates a statistically significant relationship*

*** Indicates redundant variable*

Variable	Coefficient	Robust SE	Robust t	Robust Pr	VIF
Intercept	-2.6273	3.1735	-0.8279	0.4086	--------
Depth	-0.0236	0.0104	-2.2616	0.024653*	7.7238**
2018 Surface Temperature	0.3714	0.1839	2.0198	0.044564*	3.5858
2018 Pacific Cod CPUE	-0.1129	0.1558	-0.7245	0.4695	1.5141
2018 Bottom Temperature	-0.0335	0.3167	-0.1058	0.9159	4.7847
2017 Mature Female Snow Crab CPUE	0.2074	0.0632	3.2798	0.001213*	2.9622
2017 Immature Snow Crab CPUE	0.2441	0.1025	2.3806	0.018094*	4.7612
2016 Surface Temperature	0.3517	0.1392	2.5276	0.012153*	3.1770
2016 Immature Snow Crab CPUE	0.1567	0.0688	2.2766	0.023725*	2.8841
2015 Bottom Temperature	-0.2630	0.1771	-1.4853	0.1388	8.9154**
2014 Pacific Cod CPUE	-0.0896	0.1460	-0.6133	0.5403	2.4020
2014 Bottom Temperature	0.2759	0.2411	1.1445	0.2536	8.8868**
2013 Bottom Temperature	-0.2672	0.1832	-1.4584	0.1461	4.1236
2012 Pacific Cod CPUE	-0.1690	0.1335	-1.2658	0.2069	1.8920
2011 Surface Temperature	0.3612	0.1967	1.8368	0.0675	6.2667
2010 Surface Temperature	-0.2321	0.1132	-2.0501	0.041492*	3.9222
2008 Mature Female Snow Crab CPUE	0.1157	0.0743	1.5568	0.1209	3.7557
2007 Immature Snow Crab CPUE	0.1702	0.0694	2.4512	0.014980*	3.1556
2006 Pacific Cod CPUE	0.3697	0.1385	2.6690	0.008150*	1.8548
2006 Mature Female Snow Crab CPUE	0.0581	0.0542	1.0726	0.2846	2.2235

4.4. Local Relationships

The local form of regression in GWR including only 2018 variables (bottom and surface temperatures, depth, Pacific cod CPUE) performed better than the global form and resulted in an AIC_c score of 1410 compared to 1655 in the OLS (see Table 10). The amount of variance explained also increased from 49% to 83%. These model results are comparable to the OLS including lagged independent variables.

Table 10. GWR model performance and diagnostics

GWR Diagnostics	
r^2	0.8328
Adjusted r^2	0.7913
AICc	1409.7732
Sigma-Squared	4.9775
Sigma-Squared MLE	3.9922
Effective Degrees of Freedom	247.0269

Local r^2 for the GWR is mapped in Figure 26. Model accuracy was poorest along the southern edge of the survey along the Alaska Peninsula as well as the western edge along the outer domain, south of Zemchug Canyon. GWR model accuracy was highest (local r^2 = .94) in the central region and middle domain nearest Nunivak Island but averaged 78%, well above that of the OLS (49%). This area coincides with higher snow crab CPUE and Pacific cod CPUE values; locally weighted regression requires a certain amount of spatial variation in the independent variables, which may explain the pattern of poor performance in other areas.

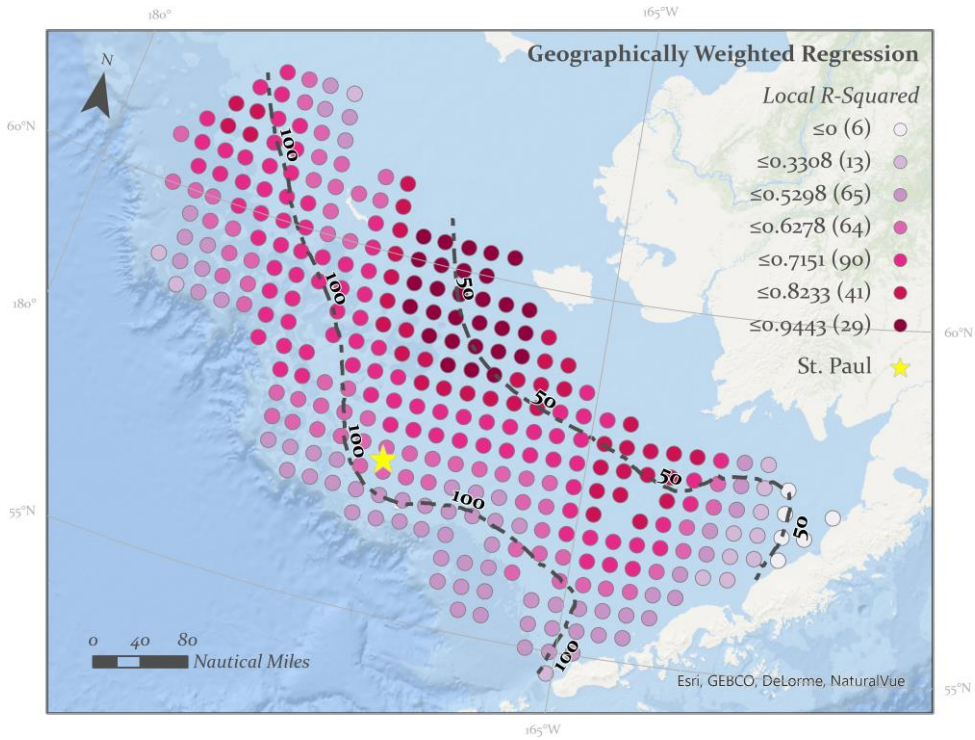


Figure 26. GWR model accuracy (local r^2)

The poorest model performance was seen in Bristol Bay where snow crab CPUE was consistently low without variance. Depth and temperature variables are spatially uniform across the southeastern shelf region which may have contributed to the poor performance considering multicollinearity or redundancy was detected in the OLS model when including lagged variables (see Table 9, VIF>7.5 for depth and 2014, 2015 bottom temperatures).

Figure 26 shows that the western shelf edge was more difficult to model using GWR compared to the middle domain and central region of the EBS, although the local r^2 along the outer domain was still over 53%. Larger error residuals accompanied the locations with poorer performance along the shelf edge (see Figure 27). There was no significant autocorrelation in the model residuals but the map in Figure 27 (top) does show clustering in the Bristol Bay region. Snow crab CPUE was perhaps too consistently low or depth and temperature variables lacked

enough variation for the regression to accurately model CPUE in the southeastern shelf at this scale of analysis.

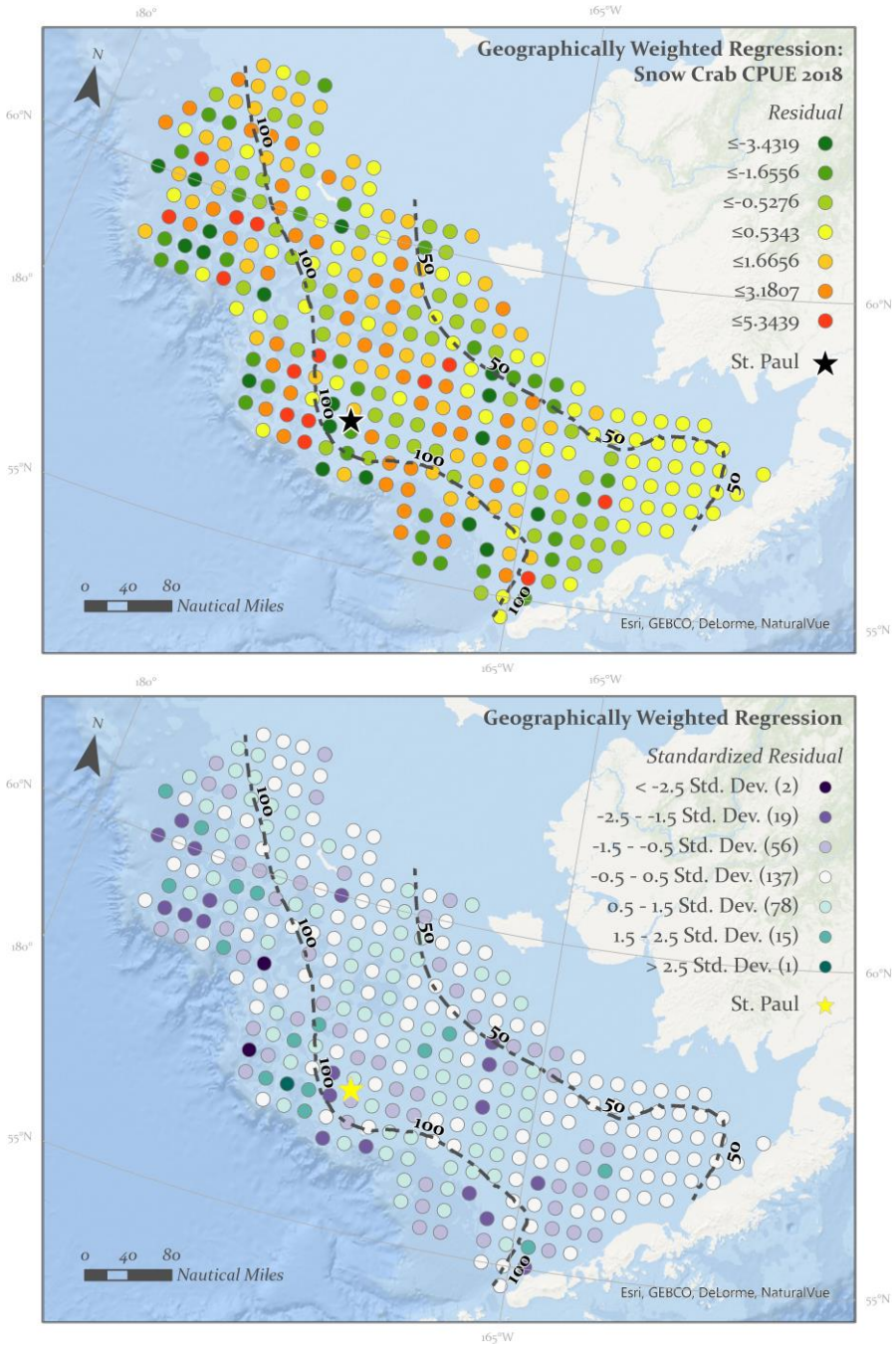


Figure 27. GWR model residuals and standardized residuals showing spatial performance in modeling 2018 snow crab CPUE

The scaled magnitude of error, similar to a t-statistic, was calculated at each survey station and the results are presented alongside the local variable coefficients in Figures 28 and 29. Areas of low coefficient to error ratios were identified as transition zones, where the variable was not effectively modeled in the GWR (Esri 2020). These areas are symbolized as yellow survey stations in Figures 28 and 29, and regions of higher coefficient to error ratios and consistent strength in the coefficient are highlighted in red for each variable.

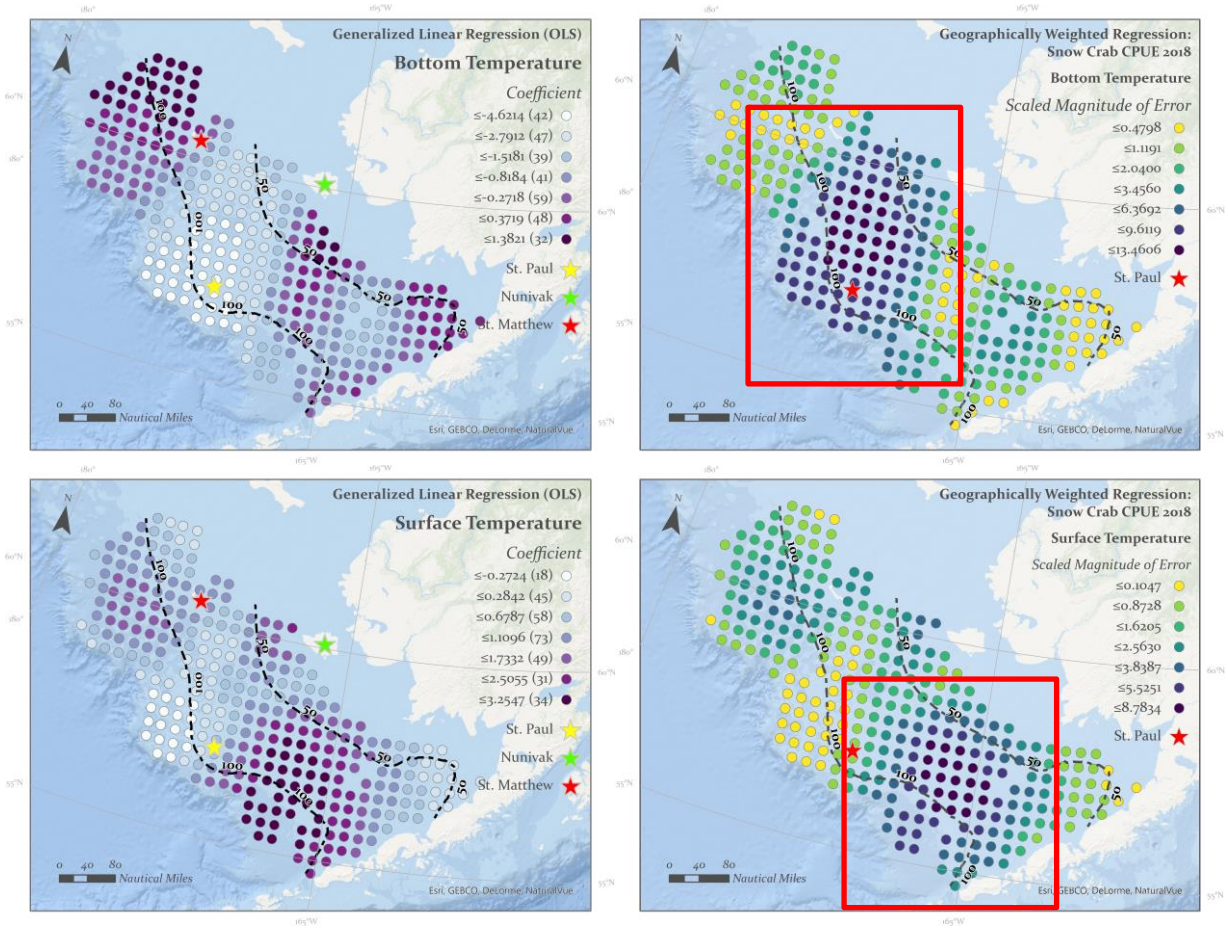


Figure 28. GWR local variable coefficients and scaled error for bottom temperature (top) and surface temperature (bottom)

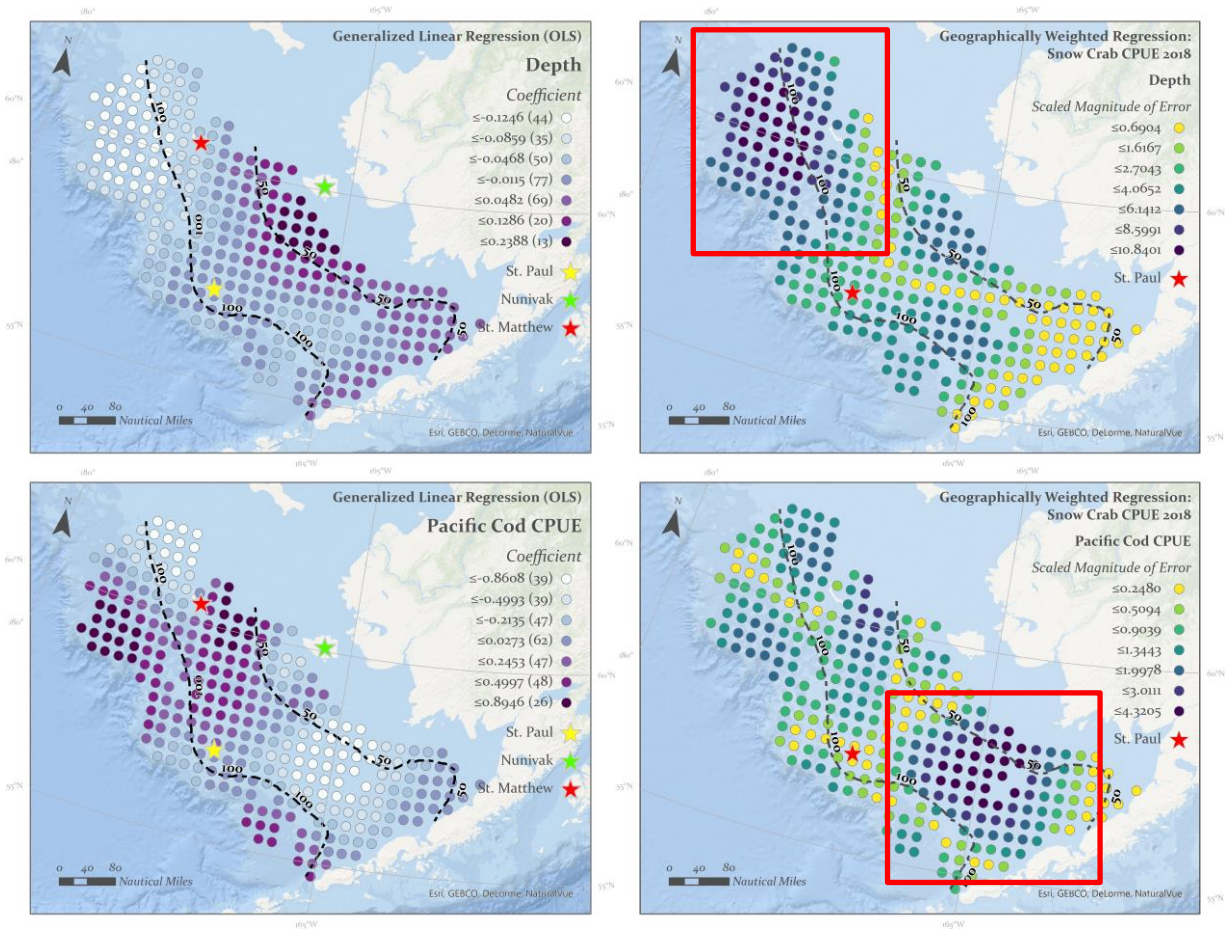


Figure 29. GWR local variable coefficients and scaled error for depth (top) and Pacific cod CPUE (bottom)

Figure 28 shows the strongest local coefficient for the bottom temperature variable (-4.4) occurred throughout the central region of the EBS, between Zemchug and Pribilof Canyons. The scaled magnitude of error outside this region decreases, indicating a shift in the relationship where other variables gain influence. Outside the central region, snow crab distribution correlated (negatively) with depth to the north of Zemchug Canyon according to the highest coefficient to error ratio. This matches the patterns identified in the time cube and spatiotemporal analysis which showed a gradient of increasing CPUE of snow crab in this area moving from west to east towards the Bering Strait and colder temperatures. By comparison snow crab CPUE

in the southern region of the EBS shelf appears to be dually influenced by a positive relationship with surface temperature and a negative relationship with Pacific cod CPUE (Figure 29). The transition zones delineated in the scaled magnitude of error maps represent possible boundaries for spatial units of analysis, discussed further in the last chapter.

GWR was also applied to the same lagged independent variables tested in OLS regression. The model diagnostics are listed in Table 11. Including lag decreased model performance and AIC increased from 1108 to 1339. Local r^2 was high, 94%, or .84 adjusted for the addition of extra explanatory variables (this increases the numerator for the GWR including lag. The increased AIC score suggests that including lagged independent variables in a locally weighted regression may be less appropriate than this approach using a global form of regression. The local variable coefficient results are provided in Appendix B for the GWR with lag included, but further research should be done prior to developing this model and is discussed in the final chapter.

Table 11. GWR model performance and diagnostics including lag

GWR Diagnostics	
r^2	0.9479
Adjusted r^2	0.8370
AICc:	1339.4155
Sigma-Squared:	3.1642
Sigma-Squared MLE:	1.0192
Effective Degrees of Freedom:	79.8827

Chapter 5 Discussion

Snow crab abundance patterns in terms of CPUE and ecological relationships in the EBS were explored through spatiotemporal analyses and a multi-scale combination of global and local regression techniques. The results confirmed basic findings of recent research using comparable NMFS bottom trawl survey data and showed that snow crab were shifting north and east towards the source of the cold pool and the Bering Strait. The varied methods and analyses applied here demonstrated the versatility of GIS for performing biostatistical analysis and visualization of species distributions from standardized fisheries surveys. Large and complex datasets like the EBS trawl survey are easily and effectively modeled by the space time cube data structure.

GIS spatial analytics and visual explorations of snow crab distribution across space and time in relation to key environmental variables like sea temperature and depth support ecosystem-based fisheries management and ecological monitoring efforts. Global methods indicated spatial autocorrelation or clustering of similar values. Quantifying local relationships and visualizing how these variable coefficients varied in space helped to identify ecological regions and transition zones that could be applied towards development of an improved global regression model to support fisheries statistical analysis. This type of approach can supplement traditional stock assessments that rely on purely statistical analyses which do not account for the spatial and temporal correlation inherent in natural systems. As species distributions and by extension fisheries in the EBS shift, managers can benefit from GIS and exploratory techniques using the space time cube data structure and local regression analysis.

This chapter first summarizes the results, shortcomings, and solutions for improvement to the spatiotemporal analysis section of the study. Regression results and suggestions are discussed

similarly, followed by a discussion of opportunities for further development or model adaptations.

5.1. Spatiotemporal Explorations

The discrepancy in results observed between purely temporal (Mann-Kendall) and spatial (Getis Ord Gi* or Anselin Moran's I) statistics of snow crab abundance patterns highlights the utility of performing this type of dual space time exploration to test each aspect of autocorrelation in ecological and fisheries survey datasets. No significant global trends were identified by the Mann Kendall temporal test but there was obvious regional spatial correlation in snow crab CPUE trends. By incorporating spatial autocorrelation and neighborhood context, the Getis Ord Gi* and Local (Anselin) Moran's I tests were able to confirm significantly different CPUE trends between northern and southern survey regions over the study period. The snow crab population was shifting towards the Bering Strait according to hot spots in the northeast and a slow but consistent decline in the south. The difference between observed snow crab CPUE space time trends in northern and southern survey regions corroborated previous reports of a northern shift in benthic species distributions, based on similar variations of the EBS survey data (Orensanz et al. 2004; Parada et al. 2010; Stevenson and Lauth 2018).

Exploratory regression revealed temporal correlation in snow crab CPUE or age class connectivity between the total snow crab population and maternal and immature age classes as laid out by Ernst et al. (2012) and Emond et al. (2015). Snow crab CPUE exhibited a greater dependency on historical abundance and the timing between life history cycles than to external biological (Pacific cod) or environmental (bottom temperature) variables, historic or prevailing. Age classes showed spatial stratification similarly described by Orensanz et al. (2004), evident in the time cube visualizations of immature snow crab CPUE (clustered along the middle domain)

and mature female snow crab CPUE (aggregated to the north and west of the main population/immature age classes).

The survey dataset was extensive and covered a wide spatial, temporal, and attribute range. Its resulting space time cube contained 12,913 individual space time bins. A simpler view of individual time cube stacks was more effective for regional comparisons of snow crab CPUE (and Pacific cod CPUE). This stratified sampling approach to visualization also allowed space for labeling with information such as survey year which helped to pinpoint the timing of hot spot blooms and CPUE change. The spatial variation of temporal trends in snow crab CPUE was effectively summarized in the time series profile correlation analysis and helped to visually divide the survey region into zones exhibiting variable CPUE patterns, or spatial nonstationarity. The arrangement and trend direction of each of the time series cluster zones showed that snow crab CPUE varied differentially along each axis of the shelf: numbers decreased to the east and west of the middle domain and even more drastically from north to south.

5.2. Regression Exploration

Following spatiotemporal analysis and visualization, progressive regression tests allowed for the exploration of snow crab historical (lagged) and contemporary relationships with ecological factors, just as Zulu et al. (2014) developed a spatiotemporal context throughout their analysis of infection spread in Malawi. GWR performed better than the OLS according to AIC and r^2 values, just as Windle et al. (2010) showed in their studies of snow crab in the north Atlantic. The locally weighted regression became unstable when incorporating lagged independent variables but the technique should be studied further in concert with OLS development.

Extremely low, near-zero CPUE records occurred in Bristol Bay while near-millions snow crab CPUE were recorded in the northeast survey region. These heteroskedastic abundance patterns in the snow crab population highlight the limitation of relying solely on a global OLS for regression modeling, which smoothed each of these trends to fit a single regression line; the result might not reflect the northern or southern distribution trends accurately (Fotheringham 2002). Though the OLS model residuals were not clustered and there was no significant amount of bias or model misspecification, the model diagnostics in the JB statistic identified significantly inconsistent relationships, confirming the heteroskedastic attribute scale for snow crab CPUE. GWR and an exploration of the local variation in snow crab relationships could then help to pinpoint the independent variables which contribute more significantly to shaping the distribution and ecologically consistent zones where relationships were stable according to the t-statistic in the local variable coefficients.

The GWR local variable coefficient map of bottom temperature (excluding lag) showed a large and consistent cluster of stations in the central survey region where the temperature relationship was strongest and most stable. Other explanatory variables gained influence at either end of the survey region: the snow crab CPUE relationship with depth was stronger and more consistent in the north, while surface temperatures and Pacific cod CPUE were more influential in the south. This stands to reason as Pacific cod CPUE increased in shallower areas near the main islands in the EBS and the coastal domain, while in the coldest northern regions of the survey range distribution varied more according to depth.

Results of the GWR and OLS comparison in this study are reflective of those discovered by Windle et. al (2010) in the north Atlantic, in that environmental and biological relationships varied locally, due to spatial dependence and spatial autocorrelation in the data. For these

ecological characteristics the locally weighted regression technique was better able to explain local variation in snow crab CPUE and generate a better-fit model. Windle et al. (2010) have taken the technique applied in this study one step further by applying a k-means cluster analysis of local variable coefficient t-values (coefficient to error ratio) to spatially distinguish consistent relationship zones. Multivariate clustering could be applied to the GWR local variable coefficient/error in a similar approach to divide the study area into a pre-defined number of clusters. Survey stations would be grouped according to likeness of CPUE values through an unsupervised ML algorithm, so the attribute or analysis field is standardized to account for the stronger influence of variables with large variances. To accomplish this standardization the global mean of the attribute is subtracted from each attribute value, then divided by the standard deviation for all values (Esri 2020). To investigate the north/south differential and where these patterns in CPUE diverge, it may be appropriate to begin with the designation of two clusters in a multivariate clustering analysis. GIS enables simple parameter adjustments so that additional numbers of clusters could be easily explored. Significantly different clusters would identify areas of significantly distribution patterns that would need to be managed according to separate standards and/or regulations.

Though bottom and surface sea temperature were identified as significant independent variable in the first OLS test, the OLS regression test when including lag showed that when considering both temporal and spatial correlation throughout snow crab life history, bottom temperature was not significant in any year. Temporal correlation and age class connectivity were more significant at this scale of analysis, which again emphasizes how useful it can be to explore both spatial and temporal autocorrelation, and to consider the dependent variable variation at local and global scales to compare results and gain a better understanding of the scale

of the attribute as well as the spatial and temporal range. Applying the GWR to each year of snow crab CPUE included in the exploratory regression similar to the work of Windle et al. (2012) could also show how the ecological relationships have changed in space and/or strength over time as climate conditions have shifted.

5.3. Further Development

The results of the spatiotemporal analysis showed that the snow crab population trends were divergent at either end of the EBS survey range. The study area should be divided into smaller spatial units of analysis that would more aptly represent this observed spatial structure in the snow crab population. Previous research has parsed EBS survey data spatially in various ways prior to analysis to achieve improved results of OLS and other global approaches to regression (Ciannelli et al. 2008; Kotwicki and Lauth 2013). OLS regression showed that temporal correlation and age class connectivity was the strongest determinant of snow crab CPUE in 2018. Rather than use the results of spatiotemporal analyses it may prove more representative to model the spatial analysis units after the spatial results of the temporal trend test (Mann-Kendall); however, the location of the negative temporal trends in CPUE coincide with the low-low CPUE cluster from the cluster and outlier analysis. This would divide the study area between northern and southern analysis units and likely increase the accuracy of the OLS model. Dividing the shelf between north and south according to where the statistically significant downward trends in snow crab CPUE or the low-low cluster of survey stations were identified would divide the shelf into northern and southern spatial analysis units at 167°W longitude (just west of Nunivak).

Snow crab CPUE data remains skewed and requires a more effective linear transformation. The log (x + 1) transformation improved but did not fully normalize the data.

Other work applying GWR alongside a global scale regression by Windle et al. (2012) showed the significance of the relationships in their results did not vary whether using transformed or raw data for shrimp, snow crab and Pacific cod abundance variables. For this study the log (x + 1) transformation was accepted with the acknowledgement that the significance of the results should be interpreted carefully and only in an exploratory nature. An alternative transformation such as box cox should be developed to increase the reliability of the global or local regression modeling results.

One crucial statistical parameter used in this study that deserves further exploration was the spatial neighborhood definition in the spatiotemporal analysis. The EBS survey dataset for the standardized 20x20nm stations was left intact and treated as a single spatial unit, and the neighborhood distance band remained fixed at 45 nm throughout the spatiotemporal analysis to match the gaussian kernel bandwidth applied in the GWR regression. However, the fixed neighborhood distance (45 nm) was optimized for the GWR, not necessarily the time cube. This parameter should be further tested towards representing the spatial and temporal autocorrelation inherent in the entire dataset or space time cube frame. The most common method using ArcGIS statistical analysis to determine an appropriate spatial neighborhood distance involves measuring the level of spatial autocorrelation at increasing distance band intervals using the Global Moran's I statistic and selecting the distance at which spatial autocorrelation or the I-statistic peaks (Mitchell 2009; Steves 2017; Esri 2020). Preliminary exploration following the conclusion of these results indicated that peak z-scores in spatial autocorrelation (clustering) occurred at a greater neighborhood distance than found for the optimal Gaussian kernel (45 nm). The spatial neighborhood applied in the spatiotemporal analysis and the spatial bandwidth of the kernel applied in the GWR should each now be fine-tuned according to optimized AIC scores in the

case of GWR, and optimal autocorrelation in the case of spatiotemporal analysis. This will likely result in varied results for the hot spot analysis, which failed to identify any cold spots in the entire series using the 45 nm bandwidth for neighborhood context despite wide-ranging downward temporal trends in snow crab CPUE in the southern survey region.

Other adjustments to the spatiotemporal parameters should be tested, including the temporal neighborhood. A single year/annual time step was selected for this preliminary exploration of the space time cube and the EBS survey data. Expanding the temporal neighborhood by two or three survey years or aggregating the data into multi-year bins should be explored in tandem with adjustments to the spatial neighborhood and spatial analysis units.

As in previous studies (Orensanz et al. 2004; Windle et al. 2009), regression results (excluding lagged independent variables) suggested that snow crab distribution was more dependent on bottom-up environmental pressures and climate-scale processes rather than top-down predation by cod. However, climate change and ecological shifts could result in shifts in diet. Including CPUE catch records of other predatory fish such as yellowfin sole or Pacific halibut in the predation pressure index by aggregating these attributes into a single predation index might provide a better representation of this type of impact on snow crab.

There are many biological and environmental variables that could be further explored using these methods. One key factor which could significantly impact all benthic species in the EBS is commercial bottom trawling. C. Steves (2015) showed that gear impact was increasing in certain areas on the shelf through a spatiotemporal analysis of trawl fishing effort in Alaska. This trend is likely to continue as groundfish and other trawled species increase in the EBS as a result of the rising temperatures. In addition to habitat damage, trawling can quickly deplete populations of non-target species. Vessels are required carry observers to measure and report

commercial bycatch of snow crab and other prohibited or managed quota species to state or federal regulatory agencies. Annual bycatch estimates could shed light on a missing model parameter.

Oceanographic data to supplement or replace the bottom and surface temperature data from the EBS bottom trawl survey may improve regression model performance, particularly when including lag as many survey station bins were missing values at some point in the time series, and these locations could not be included in the results. Satellite and remote sensing datasets could provide a means of supplementing surface temperature data, which was significant in the relationship with snow crab distribution in the OLS with and without lag variables. Ocean color satellite imagery can also be analyzed to measure primary production as an independent variable (Brody, Lozier, and Dunne 2013). Care should be taken to assess the resolution, range, and overall quality or reliability of any data external to the fishery standardized survey.

5.4. Conclusion

The goal of this project was to explore the spatiotemporal distribution of snow crab in the EBS using GIS to support marine fisheries and ecosystem-based management decisions. Spatial analytics incorporate autocorrelation and, in some cases, reveal trends masked by traditional statistical testing or global regression modeling. Spatiotemporal analysis also provides context to regression modeling to better understand the strength and significance of key ecological relationships. The space time cube was an effective data structure for modeling standardized surveys and enabled pattern mining and regression analyses supportive of traditional regression modeling and statistical assessments. Spatiotemporal analysis and regression modeling in GIS were effective and complimentary approaches to fisheries monitoring and ecosystem-based spatial management based on this study. The exploratory regression of lagged environmental

variables shows that temporal correlation of snow crab abundance can reveal age-class connections between maternal parent classes, immature classes, and total population in a timeline. Exploratory temporal correlation can also identify significant past events in the life history of snow crab such as climate pressure from sea temperature warming or predation from Pacific cod.

GIS is a versatile platform that can manage large and complex datasets typical of standardized biological surveys and should be explored further to support traditional single species stock assessments. With further development these techniques could be developed in an ecosystem-based approach towards fisheries management. Spatiotemporal autocorrelation can identify homogenous areas of the attribute of interest, and zones of consistent ecological relationships that could be applied towards determining or allocating quota.

References

Alaska Department of Fish and Game (ADFG). 2019. "Review of TACs Bering Sea Crab: 2019/20 Season (Part 1 and 2)." Alaska Department of Fish and Game, Juneau, Alaska. http://www.adfg.alaska.gov/static/fishing/PDFs/commercial/bering_aleutian/part1_industry_tac_meeting.pdf; http://www.adfg.alaska.gov/static/fishing/PDFs/commercial/bering_aleutian/part2_industry_tac_meeting.pdf (Accessed 7 June 2020).

Alaska Department of Fish and Game (ADFG). 2020. "Commercial Shellfish Fishery Ex-Vessel Prices." Alaska Department of Fish and Game, Juneau, Alaska. https://www.adfg.alaska.gov/index.cfm%3Fadfg%3DCommercialByFisheryshellfish.shellfishcatch_exvessel_crab

Alaska Fisheries Science Center (AFSC). 2019. "Groundfish Survey Data." Alaska Fisheries Science Center, National Oceanic Atmospheric Administration Resource Assessment and Conservation Engineering Division, Seattle, Washington.https://apps-afsc.fisheries.noaa.gov/RACE/groundfish/survey_data/data.htm, https://apps-afsc.fisheries.noaa.gov/maps/crabmap/crabmap.html (Accessed 7 June 2020).

Aydin, O., and K. Butler. 2019. "Predicting Global Seagrass Habitats." In *GIS for Science: applying mapping and spatial analytics*, edited by D. Wright and C. Harder, 58-69. Redlands, CA: Esri Press.

Bevan, A., and J. Conolly. 2009. "Modelling Spatial Heterogeneity and Nonstationarity in Artifact-Rich Landscapes." *Journal of Archaeological Science* 36 (4): 956–64. https://doi.org/10.1016/j.jas.2008.11.023.

Boudreau, S., Anderson, S., & B. Worm. 2011. Top-down interactions and temperature control of snow crab abundance in the northwest Atlantic Ocean. *Marine Ecology Progress Series*, *429* (December 2015), 169–183. https://doi.org/10.3354/meps09081.

Brody, S., M. Lozier, and J. Dunne. 2013. "A comparison of methods to determine phytoplankton bloom initiation." *Journal of Geophysical Research: Oceans* 118: 2345–2357. https://doi.org/10.1002/jgrc.20167.

Burgos, J., B. Ernst, D. Armstrong, and J. Orensanz. 2013. "Fluctuations in Range and Abundance of Snow Crab (*Chionoecetes Opilio*) from the Eastern Bering Sea: What Role for Pacific Cod (*Gadus Macrocephalus*) Predation?" *Bulletin of Marine Science* 89 (1): 57–81. https://doi.org/10.5343/bms.2011.1137.

Ciannelli, L., P. Fauchald, K. Chan, V. Agostini, and G. Dingsør. 2008. "Spatial Fisheries Ecology: Recent Progress and Future Prospects." *Journal of Marine Systems* 71 (3–4): 223–36. https://doi.org/10.1016/j.jmarsys.2007.02.031.

Emond, K., B. Sainte-Marie, P. Galbraith, and J. Bety. 2015. "Top-down vs. Bottom-up Drivers of Recruitment in a Key Marine Invertebrate: Investigating Early Life Stages of Snow Crab." *ICES Journal of Marine Science* 72 (May): 1336–48.

Ernst, B., J. Orensanz, and D. Armstrong. 2005. "Spatial Dynamics of Female Snow Crab (*Chionoecetes Opilio*) in the Eastern Bering Sea." *Canadian Journal of Fisheries and Aquatic Sciences* 62 (2): 250–68. https://doi.org/10.1139/f04-201.

Ernst, B., D. Armstrong, J. Burgos, and J. Orensanz. 2012. "Life History Schedule and Periodic Recruitment of Female Snow Crab (*Chionoecetes Opilio*) in the Eastern Bering Sea." *Canadian Journal of Fisheries and Aquatic Sciences* 69 (3): 532–50. https://doi.org/10.1139/F2011-173.

Esri. 2020. "Data Classification Methods." Symbolization, Esri. Redlands, California. https://pro.arcgis.com/en/pro-app/help/mapping/layer-properties/data-classification-methods.htm.

Esri. 2020. "How Geographically Weighted Regression (GWR) Works." Modeling Spatial Relationships toolset concepts, Esri. Redlands, California. https://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/how-geographicallyweightedregression-works.htm#:~:text=Geographically%20Weighted%20Regression%20(GWR)%20is,every%20feature%20in%20the%20dataset.

Foody, G. M. 2004. Spatial nonstationarity and scale-dependency in the relationship between species richness and environmental determinants for the sub-Saharan endemic avifauna. *Global Ecology and Biogeography*, *13*(4), 315–320. https://doi.org/10.1111/j.1466-822X.2004.00097.x.

Fotheringham, S., C. Brunsdon, and M. Charlton. 2002. *Geographically Weighted Regression: the analysis of spatially varying relationships*. University of Newcastle, UK: John Wiley and Sons, Ltd.

Foy, R., and C. Armistead. 2012. "The 2012 Eastern Bering Sea Continental Shelf Bottom Trawl Survey: Results for Commercial Crab Species." U.S. Department of Commerce, NOAA Technical Memorandum NMFS-AFSC-242.

Groß, J., B. Konar, T. Brey, and J. Grebmeier. 2017. "Size-Frequency Distribution, Growth, and Mortality of Snow Crab (*Chionoecetes opilio*) and Arctic Lyre Crab (*Hyas Coarctatus*) in the Chukchi Sea from 2009 to 2013." *Deep-Sea Research Part II: Topical Studies in Oceanography* 144: 142–55. https://doi.org/10.1016/j.dsr2.2017.04.021.

Hardy, S., M. Lindgren, H. Konakanchi, and F. Huettmann. 2011. "Predicting the Distribution and Ecological Niche of Unexploited Snow Crab (*Chionoecetes Opilio*) Populations in Alaskan Waters: A First Open-Access Ensemble Model." *Integrative and Comparative Biology* 51 (4): 608–22. https://doi.org/10.1093/icb/icr102.

Kotwicki, S. and R. Lauth. 2013. "Detecting Temporal Trends and Environmentally-Driven Changes in the Spatial Distribution of Bottom Fishes and Crabs on the Eastern Bering

Sea Shelf." *Deep-Sea Research Part II: Topical Studies in Oceanography* 94: 231–43. https://doi.org/10.1016/j.dsr2.2013.03.017.

Marcello, L., F. Mueter, E. Dawe, and M. Moriyasu. 2012. "Effects of temperature and gadid predation on snow crab recruitment: Comparisons between the Bering Sea and Atlantic Canada." *Marine Ecology Progress Series*, 469, 249–261. https://doi.org/10.3354/meps09766.

Mitchell, A. 2009. *The Esri Guide to GIS Analysis Volume 2: Spatial Measurements and Statistics*. Redlands, CA: Esri Press.

Murphy, J. 2020. "Climate Change, Interspecific Competition, and Poleward vs. Depth Distribution Shifts: Spatial Analyses of the Eastern Bering Sea Snow and Tanner Crab (*Chionoecetes Opilio* and *C. Bairdi*)." *Fisheries Research* 223: 105417. https://doi.org/10.1016/j.fishres.2019.105417.

North Pacific Fishery Management Council (NPFMC). 2019. "Bering Sea Fishery Ecosystem Plan." Report presented at 2019 meeting of North Pacific Fishery Management Council and Alaska Fisheries Science Center. https://meetings.npfmc.org/CommentReview/DownloadFile?p=c334ad33-4139-4b5a-b205-a8b7c5028562.pdf&fileName=D6%20Final%20BS%20FEP%20Jan%202019.pdf (Accessed 7 June 2020).

Orensanz, J., B. Ernst, D. Armstrong, P. Stabeno, and P. Livingston. 2004. "Contraction of the Geographic Range of Distribution of Snow Crab (*Chionoecetes opilio*) in the Eastern Bering Sea: An Environmental Ratchet?" *CalCOFI Report* 45: 63–79.

Parada, C., D. Armstrong, B. Ernst, S. Hinckley, and J. Orensanz. 2010. "Spatial Dynamics of Snow Crab (*Chionoecetes opilio*) in the Eastern Bering Sea-Putting Together the Pieces of the Puzzle." *Bulletin of Marine Science* 86 (2): 413–37.

Stamatopoulos, C. 2002. "Sample-based Fishery Surveys: A technical handbook." Food and Agriculture Organization of the United Nations (FAO) Fisheries Technical Paper. 425: 132p. Rome, FAO. http://www.fao.org/3/a-y2790e.pdf.

Stevenson, D., and R. Lauth. 2019. "Bottom Trawl Surveys in the Northern Bering Sea Indicate Recent Shifts in the Distribution of Marine Species." *Polar Biology* 42 (2): 407–21. https://doi.org/10.1007/s00300-018-2431-1.

Steves, C. 2017. "Trends in the Alaskan Bottom-Trawl Fishery from 1993-2015: A GIS-based Spatiotemporal Analysis." Master's thesis, University of Southern California.

Windle, M., G. Rose, R. Devillers, and M. Fortin. 2010. "Exploring Spatial Non-Stationarity of Fisheries Survey Data Using Geographically Weighted Regression (GWR): An Example from the Northwest Atlantic." *ICES Journal of Marine Science* 67 (1): 145–54. https://doi.org/10.1093/icesjms/fsp224.

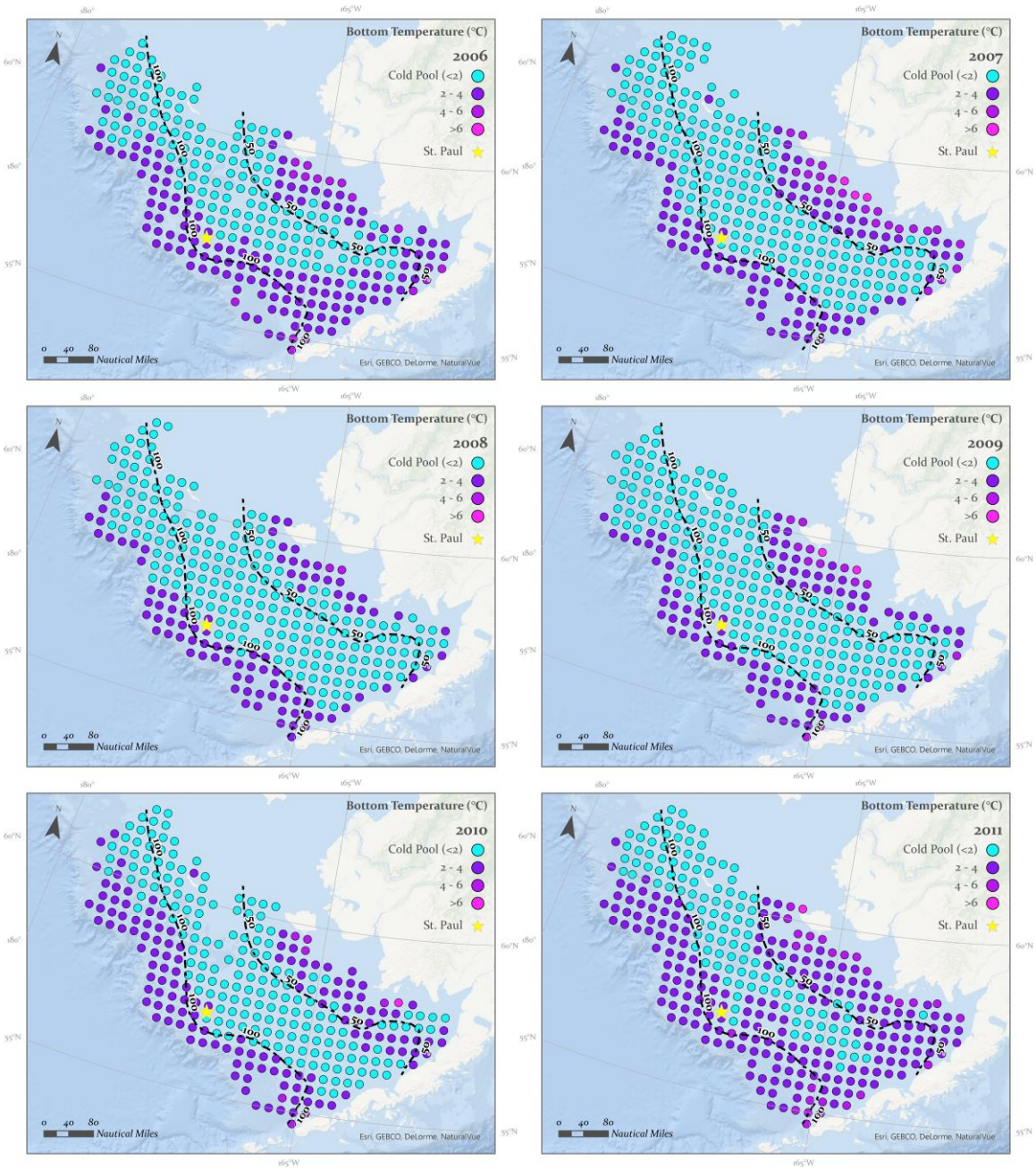
Windle, M., G. Rose, R. Devillers, and M. Fortin. 2012. Spatio-temporal variations in invertebrate-cod-environment relationships on the Newfoundland-Labrador Shelf, 1995-2009. *Marine Ecology Progress Series*, *469*, 263–278. https://doi.org/10.3354/meps10026.

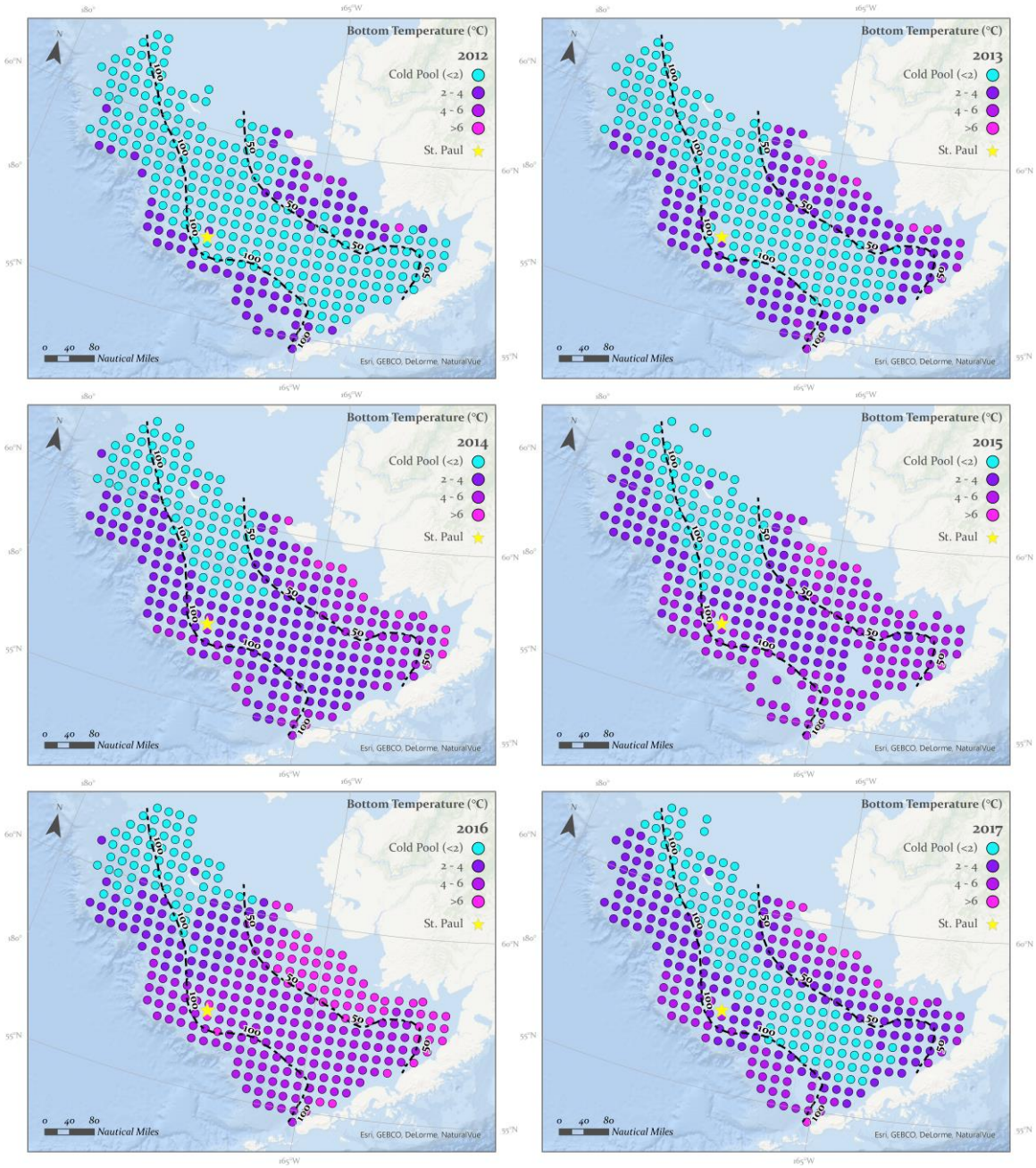
Zhang, L., and H. Shi. 2004. "Local Modeling of Tree Growth by Geographically Weighted Regression." *Forest Science* 50 (2): 225–44. https://doi.org/10.1093/forestscience/50.2.225.

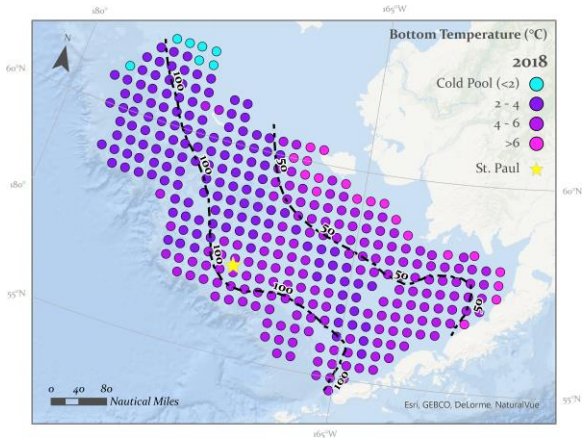
Zheng, J., and G. Kruse. 2006. "Recruitment Variation of Eastern Bering Sea Crabs: Climate-Forcing or Top-down Effects?" *Progress in Oceanography* 68 (2–4): 184–204. https://doi.org/10.1016/j.pocean.2006.02.002.

Zulu, L., E. Kalipeni, and E. Johannes. 2014. "Analyzing Spatial Clustering and the Spatiotemporal Nature and Trends of HIV/AIDS Prevalence Using GIS: The Case of Malawi, 1994-2010." *BMC Infectious Diseases* 14 (1): 1–21. https://doi.org/10.1186/1471-2334-14-285.

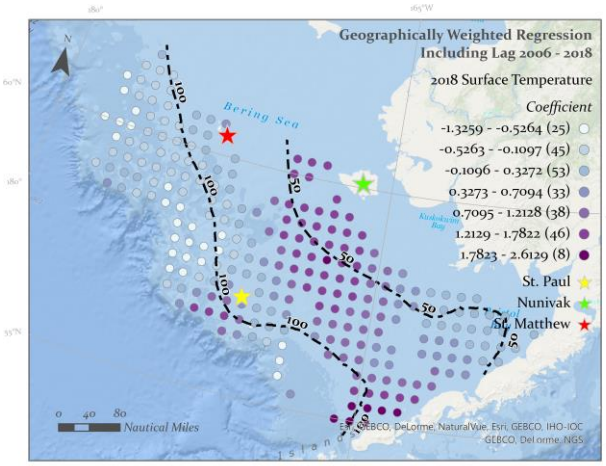
Appendix A EBS Bottom Temperatures, 2006 to 2018

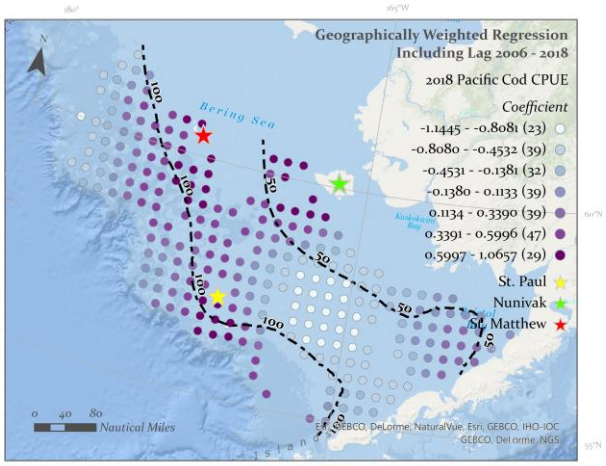


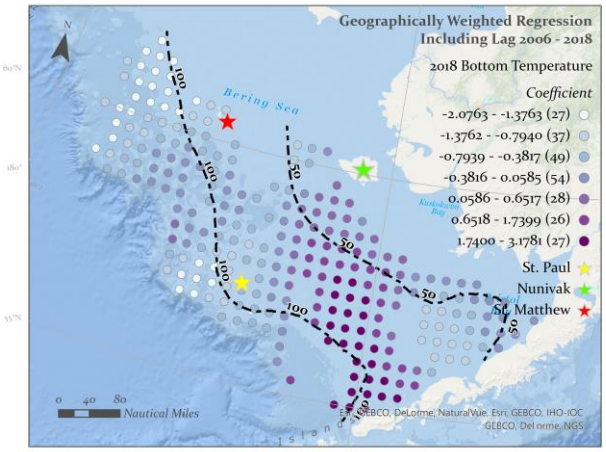

Bottom Temperature (°C)
2012
2013
2014
2015
2016
2017
Cold Pool (<2)
2 - 4
4 - 6
>6
St. Paul
0 40 80
Nautical Miles
Esri, GEBCO, DeLorme, NaturalVue
180°
165°W
60°N
55°N
N
100
50

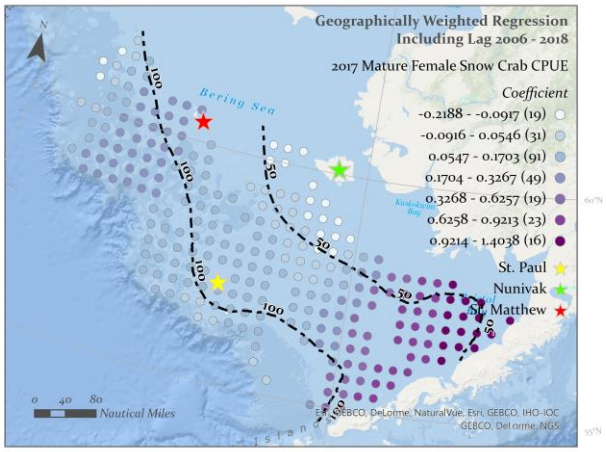

Bottom Temperature (°C)
2018
Cold Pool (<2)
2 - 4
4 - 6
>6
St. Paul
0 40 80
Nautical Miles
Esri, GEBCO, DeLorme, NaturalVue
N
180°
165°W
60°N
55°N
100
50
30
60
70

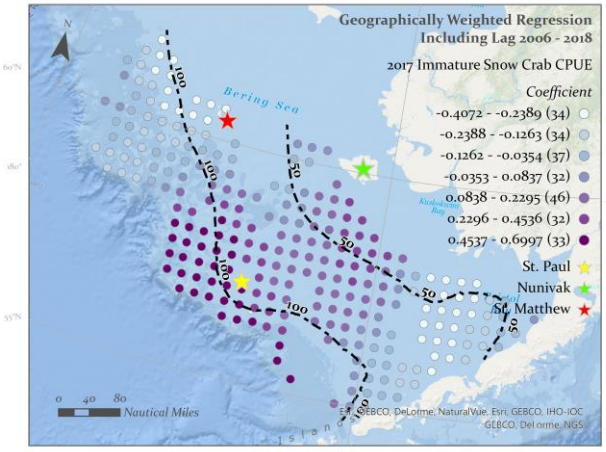
Appendix B GWR Local Variable Coefficients (including lag)

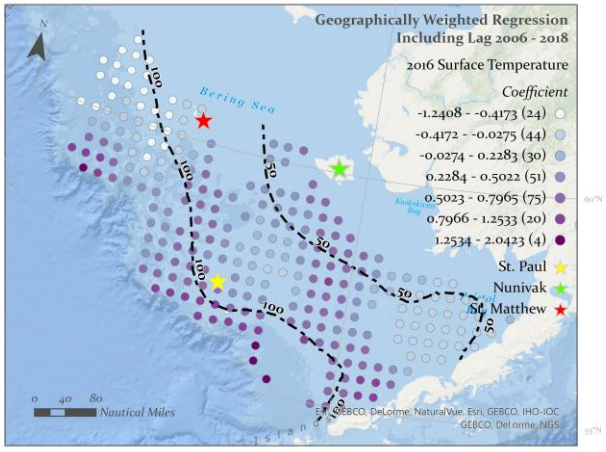


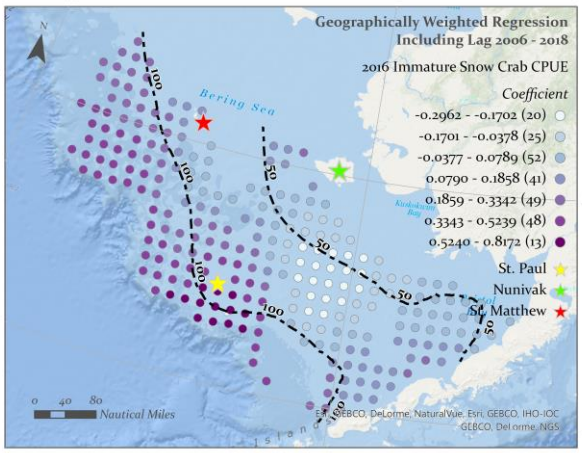


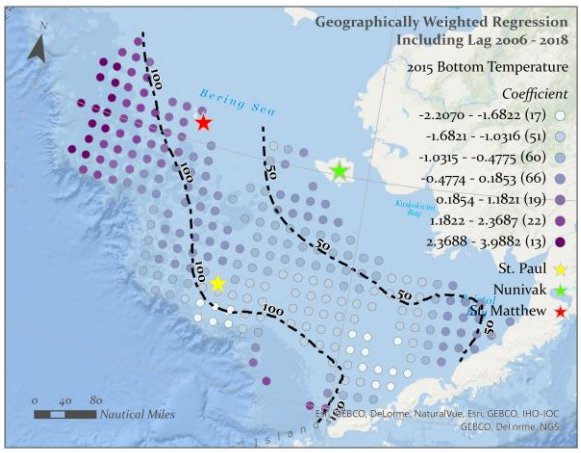


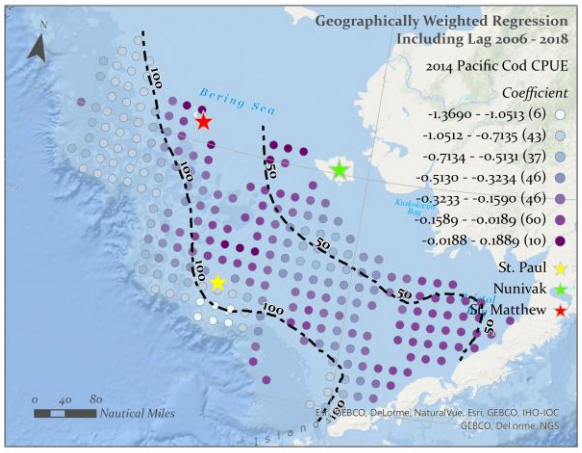


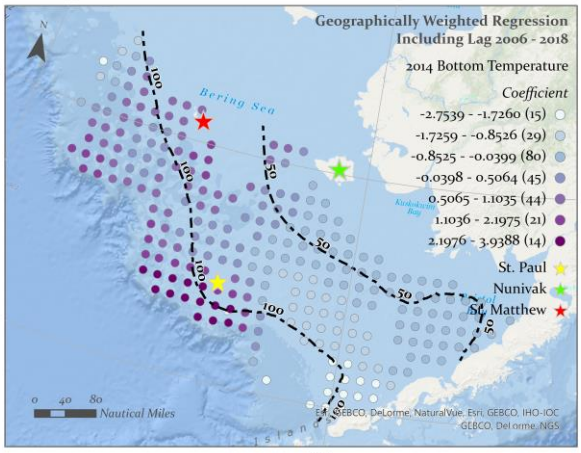


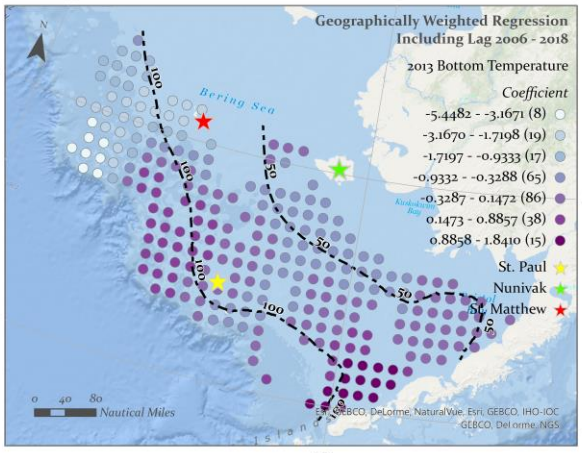


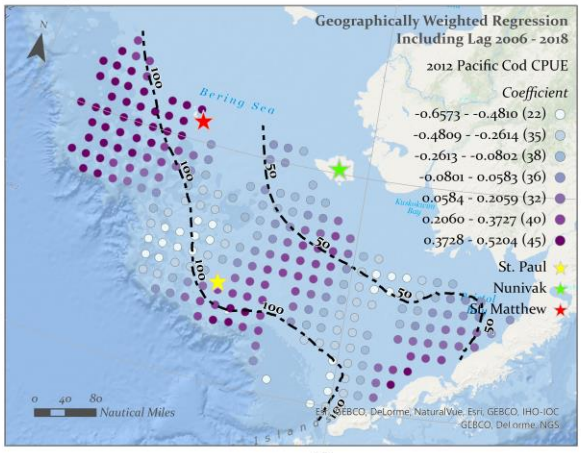

Geographically Weighted Regression
Including Lag 2006 - 2018
2016 Immature Snow Crab CPUE
Coefficient
-0.2962 - -0.1702 (20)
-0.1701 - -0.0378 (25)
-0.0377 - 0.0789 (52)
0.0790 - 0.1858 (41)
0.1859 - 0.3342 (49)
0.3343 - 0.5239 (48)
0.5240 - 0.8172 (13)
St. Paul
Nunivak
St. Matthew
Bering Sea
0 40 80
Nautical Miles
Esri, GEBCO, DeLorme, NaturalVue, Esri, GEBCO, IHO-IOC
GEBCO, Delorme, NGS

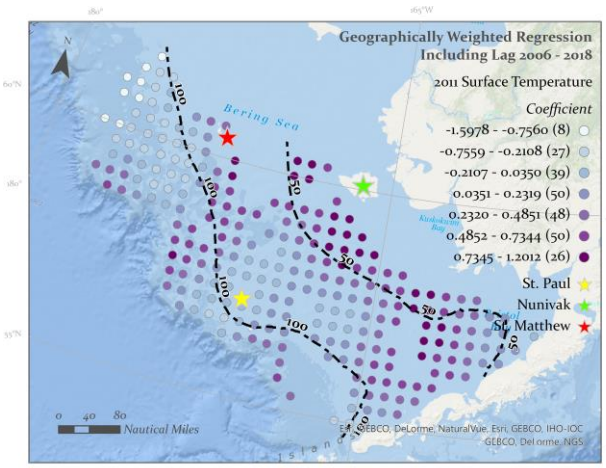

Geographically Weighted Regression
Including Lag 2006 - 2018
2015 Bottom Temperature
Coefficient
-2.2070 - -1.6822 (17)
-1.6821 - -1.0316 (51)
-1.0315 - -0.4775 (60)
-0.4774 - 0.1853 (66)
0.1854 - 1.1821 (19)
1.1822 - 2.3687 (22)
2.3688 - 3.9882 (13)
St. Paul
Nunivak
St. Matthew
Bering Sea
0 40 80
Nautical Miles
Esri, GEBCO, DeLorme, NaturalVue, Esri, GEBCO, IHO-IOC
GEBCO, Delorme, NGS

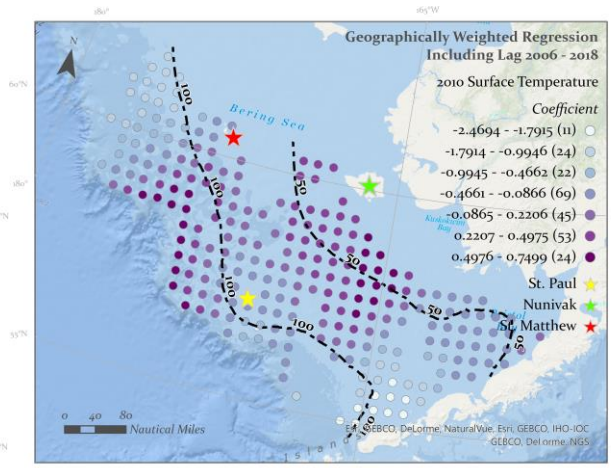

Geographically Weighted Regression
Including Lag 2006 - 2018
2014 Pacific Cod CPUE
Coefficient
-1.3690 - -1.0513 (6)
-1.0512 - -0.7135 (43)
-0.7134 - -0.5131 (37)
-0.5130 - -0.3234 (46)
-0.3233 - -0.1590 (46)
-0.1589 - -0.0189 (60)
-0.0188 - 0.1889 (10)
St. Paul
Nunivak
St. Matthew
Bering Sea
0 40 80
Nautical Miles
Esri, GEBCO, DeLorme, NaturalVue, Esri, GEBCO, IHO-IOC
GEBCO, Delorme, NGS

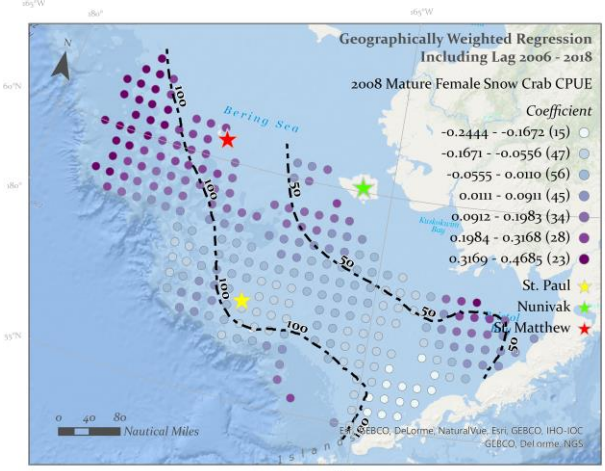

Geographically Weighted Regression
Including Lag 2006 - 2018
2014 Bottom Temperature
Coefficient
-2.7539 - -1.7260 (15)
-1.7259 - -0.8526 (29)
-0.8525 - -0.0399 (80)
-0.0398 - 0.5064 (45)
0.5065 - 1.1035 (44)
1.1036 - 2.1975 (21)
2.1976 - 3.9388 (14)
St. Paul
Nunivak
St. Matthew
Bering Sea
0 40 80
Nautical Miles
Esri, GEBCO, DeLorme, NaturalVue, Esri, GEBCO, IHO-IOC
GEBCO, Delorme, NGS


Geographically Weighted Regression
Including Lag 2006 - 2018
2013 Bottom Temperature
Coefficient
-5.4482 - -3.1671 (8)
-3.1670 - -1.7198 (19)
-1.7197 - -0.9333 (17)
-0.9332 - -0.3288 (65)
-0.3287 - 0.1472 (86)
0.1473 - 0.8857 (38)
0.8858 - 1.8410 (15)
St. Paul
Nunivak
St. Matthew
Bering Sea
0 40 80
Nautical Miles
Esri, GEBCO, DeLorme, NaturalVue, Esri, GEBCO, IHO-IOC
GEBCO, Delorme, NGS


Geographically Weighted Regression
Including Lag 2006 - 2018
2012 Pacific Cod CPUE
Coefficient
-0.6573 - -0.4810 (22)
-0.4809 - -0.2614 (35)
-0.2613 - -0.0802 (38)
-0.0801 - 0.0583 (36)
0.0584 - 0.2059 (32)
0.2060 - 0.3727 (40)
0.3728 - 0.5204 (45)
St. Paul
Nunivak
St. Matthew
Bering Sea
0 40 80
Nautical Miles
Esri, GEBCO, DeLorme, NaturalVue, Esri, GEBCO, IHO-IOC
GEBCO, Delorme, NGS


Geographically Weighted Regression
Including Lag 2006 - 2018
2011 Surface Temperature
Coefficient
-1.5978 - -0.7560 (8)
-0.7559 - -0.2108 (27)
-0.2107 - 0.0350 (39)
0.0351 - 0.2319 (50)
0.2320 - 0.4851 (48)
0.4852 - 0.7344 (50)
0.7345 - 1.2012 (26)
St. Paul
Nunivak
St. Matthew
Bering Sea
Nautical Miles


Geographically Weighted Regression
Including Lag 2006 - 2018
2010 Surface Temperature
Coefficient
-2.4694 - -1.7915 (11)
-1.7914 - -0.9946 (24)
-0.9945 - -0.4662 (22)
-0.4661 - -0.0866 (69)
-0.0865 - 0.2206 (45)
0.2207 - 0.4975 (53)
0.4976 - 0.7499 (24)
St. Paul
Nunivak
St. Matthew
Bering Sea
Nautical Miles


Geographically Weighted Regression
Including Lag 2006 - 2018
2008 Mature Female Snow Crab CPUE
Coefficient
-0.2444 - -0.1672 (15)
-0.1671 - -0.0556 (47)
-0.0555 - 0.0110 (56)
0.0111 - 0.0911 (45)
0.0912 - 0.1983 (34)
0.1984 - 0.3168 (28)
0.3169 - 0.4685 (23)
St. Paul
Nunivak
St. Matthew
Bering Sea
Nautical Miles