

State Space Assessment on the Impact of Sea Surface Temperature on Fishing Effort in Continental US Exclusive Economic Zone

Mingzhao Hu *

Qutu Jiang[†]

Abstract

Recent changes in marine environments are expected to influence human fishing activities. In particular, the effect on fishing effort is essential to the public good. Fishing effort reflects the investment strategies and costs of the industry participants, therefore affects government environmental and fishery policies. Advances in satellite technology have enabled revolutionarily detailed worldwide fishing effort data. After integrating fishing effort with sea surface temperature averages, a dynamic spatio-temporal model in the state space framework is applied. Adopting a stochastic partial differential equation approach, parameter inference uses the integrated nested Laplace approximation. Since the focus is on fishing effort in the exclusive economic zone of continental United States, the model based on Pacific coast is compared against one based on the Atlantic coast. Furthermore, a forecast of fishing effort for one year is implemented using the proposed dynamic model.

Key Words: Dynamic spatio-temporal model, stochastic partial differential equation, integrated nested Laplace approximation, fishing effort, sea surface temperature, exclusive economic zone

1. Introduction

In today's world, fishery is of increasing concern as it is an industry at the intersection of environmental protection, national economy and global food security. With recent advances in satellite imaging and tracking technology, a consistent global monitor of fishing vessels on a daily level became available for the first time. We propose to consider the dynamic evolution in time of global fishing activities in response to environmental factors across different locations via a state space approach that handles large datasets in a computationally feasible manner. To this end, we model the spatio-temporal covariance with a Gaussian Markov Random Field (GMRF), which is constructed by stochastic partial differential equations (SPDE), and adopt integrated nested Laplace approximation (INLA) for parameter estimation. We apply this approach to an extensive novel dataset and investigate how fishing effort in the exclusive economic zone (EEZ) of continental United States are affected by the sea surface temperature. This not only demonstrates the flexibility of these spatio-temporal dynamic models on a larger and more detailed geographical scale, but more importantly, provides quantitative conclusions that assist in making future commercial fishing plans and environmental regulations.

The following sections are organized as follows. Section 1.1 introduces the datasets recording the fishing effort and the sea surface temperature. Section 1.2 outlines the main statistical challenges and the contributions of this project. Section 1.3 presents the research questions and objectives. Section 2 constructs the dynamic spatio-temporal model based on the general situation. Section 2.1 discusses the SPDE representation of the GMRF in modeling the spatial dependence of data and introduces the INLA approach for inference. Section 3 presents exploratory data analysis, where Section 3.1 discusses the projection of global data in coordinates using the Gall-Peters Equal Area Projection, Section 3.2 explains

*University of California at Santa Barbara, Department of Statistics and Applied Probability, Santa Barbara, CA 93106, USA

[†]Ocean College, Zhejiang University, Zhoushan, China

Table 1: Variables of interest.

Variable name	Definition
<i>date</i>	A string in format YYYY-MM-DD
<i>lat_bin</i>	The southern edge of the grid cell, in 100ths of a degree, 101 is the grid cell with a southern edge at 1.01 degrees north
<i>lon_bin</i>	The western edge of the grid cell, in 100ths of a degree, 101 is the grid cell with a western edge at 1.01 degrees east
<i>flag</i>	The flag state of the fishing effort, in iso3 value
<i>fishing_hours</i>	Hours that vessels of this gear type and flag were fishing in this gridcell on this day

the spherical distance and covariance, and Section 3.3 explores the general trends of global fishing effort and makes comparisons based on EEZ. Section 4 presents applications of the proposed methodology to the real-life fishing effort data. Section 5 discusses our findings as well as future research.

1.1 Introduction to Data

Using the automatic identification system (AIS) installed on individual fishing vessels for collision avoidance and convolutional neural networks (CNN) for identification, Global Fishing Watch (GFW, 2018, [5]) has compiled daily fishing effort data with corresponding location latitudes and longitudes from 2012 to 2018 (Kroodsman et al, 2018, [9]). The results are compiled in the dataset, “Daily Fishing Effort and Vessel Presence at 100th Degree Resolution by Flag State and GearType, 2012 to 2018”. This data offers fresh fisheries research potential, on unprecedented magnitude and in both spatial and temporal details.

Fishing effort is binned into grid cells of 0.01 degrees latitude and longitude, and measured in units of hours. The fishing hours is calculated by assigning to each AIS detection, and then summing all positions in each grid cell. Data is based on fishing detections of over 70,000 unique AIS devices on fishing vessels. Fishing vessels are identified via a neural network classifier and vessel registry databases.

The subset of recorded variables that are of special interest in our investigation of fishing effort in EEZ is listed in Table 1. Note that the latitudes and longitudes correspond to grid cells that are significantly smaller than most studies of fishing activities conducted on the global scale. This improves the accuracy of the models and brings more information, but increases the computational complexity. For simplicity, further dimensions such as gear type of the fishing vessel will be considered later as extensions of our current model.

The main focus is on fishing activities in the exclusive economic zone (EEZ) of each country, since this is one of the two main disciplines of fishing, the other being high seas fishery. In particular, the focus is on the fishing effort by each country in its own EEZ. Most commercial fishery data sources are hindered by concerns for private corporate information, and public data are mostly on species at annual and country level, such as the fisheries and aquaculture data published by the United Nations Food and Agriculture Organization (FAO) (1967 [3]). The few public sources that can link fisheries data to specific geographical locations are limited to specific regions. Moreover, the individual vessel information

recorded in the GFW (2018, [5]) data holds the potential for research into movement of fishing vessels, tracking, etc.

Global sea surface temperature is collected from the Optimum Interpolation Sea Surface Temperature dataset (Banzon et al, 2020, [1]) which is publicly available from National Centers for Environmental Information under NOAA. The original data in Optimum Interpolation Sea Surface Temperature (OISST) consists of 0.25 degrees latitude by 0.25 degrees longitude grid cells and is gathered from different platforms, including satellites, ships, buoys, etc., which has provided it with better coverage and cross-referenced accuracy. The missing geographical components are interpolated with considerations for bias adjustments. Using the daily sea surface temperature, we select the Advanced Very High Resolution Radiometer (AVHRR) data generated via infrared instruments for its higher resolution and inclusion of land-adjacent areas. The grid cells and actual geographical regions from the GFW (2018, [5]) and OISST (Banzon et al, 2020, [1]) datasets are carefully cleaned and matched before our explorations in Section 3 and Section 4.

1.2 Challenges and Novelties

First and foremost, the novel data from GFW (2018, [5]) presents challenges in data manipulations. The fishing hours are aggregated into grids of 0.01 degree of latitude by 0.01 degree of longitude, hence the first challenge in the following analysis is recognizing that these should not be treated as point referenced data. For each entry in *fishing hours*, it gives the fishing effort measured in hours by all vessels of that country while inside that grid. Furthermore, the grids are irregular since they are measured in degrees of longitude and latitude. Because the Earth is an ellipsoid, the longitudes change from 111 km for 1 degree at the equator to 0 at the poles. The latitudes also vary from 110.567 km at the equator to 111.699 km for one degree at the poles. For simplicity, in the following analysis we will assume the Earth is a perfect sphere.

The second challenge is that all data are recorded in spherical coordinates since this is a global data set. Unlike a regional data, for example, one limited to California, in this case one cannot ignore the surface curvature of the Earth. The critical choice of an appropriate projection is discussed in the third section.

Some other challenges for data analysis are due to the size of the data which leads to higher cost in manipulation. Part of the solution is to carefully filter the data in preparation so as to discard unnecessary information such as vessel presence and integrate to suitable spatial and temporal scale. Furthermore, as a relatively recent source, Global Fishing Watch (2018, [5]) only started in 2012. The lack of historical data will be a concern in later research on the spatiotemporal scale.

Modeling the dynamic evolutions along time now needs to consider the spatial covariance structure, and the computational feasibility is challenged not only by the filtering process in the dynamic models, but also the complexity of the covariance which is considered in general as a Gaussian field. Finding efficient ways for the inference of the parameters in our dynamic models is also a vital component in this regard.

Despite these obstacles, this is an unparalleled opportunity to consider fishery at specific locations on a global and daily scale. To begin with, the fisheries data is collected at detailed, specific locations on a global and daily scale over continuous time for each day in the year. This provides the potential for a dynamic approach on a larger geographical scale and with more details. To the best of our knowledge, there has been no application of dynamic models in fisheries studies on such a comprehensive spatial and temporal level. In addition, each fishing vessel is identified with AIS, allowing for future research into movement, tracking, etc.

More importantly, the state space model captures the dynamic evolution, therefore in considering the influence of sea surface temperature on fishing efforts, the year-over-year and month-over-month comparisons are more informative. The proposed modeling framework is available for advising commercial fishing operations on predicting peak fishing period and hotspots to maximize revenue. In addition, it provides assistance in making governmental guidelines on fishing activities to minimize overfishing and other illegal fishing activities to preserve the marine living resources and ecosystem balance.

1.3 Research Questions

This project aims at addressing the following concerns. First, we propose to model the dynamic influence of sea surface temperature on fishing efforts when both spatial locations and temporal progressions are considered. This is one of our main contributions and provides quantitative support for fishing-related decision-making with new details revealed from the global activities.

Second, we aim to develop a computationally feasible model for this purpose given a large dataset. This is the focus of Section 2.

Third, we interpret the economic and environmental revelations from studying how fishing efforts in exclusive economic zone (EEZ) along the Atlantic coast of US dynamically relates to sea surface temperature, versus that of the Pacific coast. This extra factor of comparison in addition to sea surface temperature is reflected in our design of implementation.

Moreover, prediction and interpolation based on geographical location can be found for guiding future fishing activities. While the effort in fishing at different locations is very much due to subjective human decision of the crew, it is still influenced by factors such as cost, revenue, stock assessment and government policy such as subsidy and moratorium. Having a means to estimate future fishing effort at a location serves as another factor in helping participants in the fishing industry make the most beneficial plan to allocate their resources and effort.

2. Dynamic Spatio-temporal Model

Longitudinal variables are ones that contain repeated measurements of time. To model the dynamic process of a longitudinal variable changing with respect to another longitudinal variable over a spatial region, the state space models, also known as the dynamic linear models, are especially advantageous and offer a unified methodology on a system assumed to be determined by unobserved state vectors. More specifically, the dynamic evolution of fishing effort in response to sea surface temperature is a special application. The proposed modeling framework in Section 2 accommodates all spatio-temporal data analysis and is not limited to research topics in fisheries. There have been discussions on implementing dynamic linear models for spatio-temporal data (Stroud et al, 2001 [19], Lindstrom et al, 2014 [10], Hefley et al, 2017 [8]), but to accommodate the scale of global and daily data in a computationally feasible manner, we follow the framework of Godana et al (2019, [6]). We will now introduce the state space model setup.

Let $Y(s, t)$ be the spatiotemporal process of the longitudinal variable of interest at time point t and location s , and assume the model:

$$Y(s, t) = \beta Z(s, t) + \Phi(s, t) + \epsilon(s, t), \quad (1)$$

which is spatially and temporally independent, $Z(s, t)$ is the covariates, $\Phi(s, t)$ is the latent states, and the white noise $\epsilon(s, t)$ is serially and spatially independent, following $N(0, \sigma_\epsilon^2)$.

The latent states evolves dynamically via:

$$\Phi(s, t) = \lambda\Phi(s, t - 1) + \gamma(s, t), \tag{2}$$

which is spatially and temporally independent. $\Phi(s, 1)$ follows $N(0, \frac{\sigma_\gamma^2}{1-\lambda^2})$ and $\gamma(s, t)$ independently follows $N(0, \sigma^{2*})$.

The parameters are given by $\Theta = (\beta, \lambda, \sigma_\epsilon^2, \sigma_\gamma^2, \kappa)$. Furthermore, the joint posterior of the parameters, $P(\Theta, \Phi|Y)$, is the product of the data model, process model, and parameter model, $P(Y|\Theta, \Phi)P(\Phi|\Theta)P(\Theta)$. This can be determined under a Markovian process assumption and independent hyperparameters assumption.

Note that for equation (2), σ^{2*} is a spatio-temporal covariance that is sensitive to the number of geographical locations and time points. When the data is comprehensive, to avoid issues with computational feasibility, we introduce the following stochastic partial differential equations approach.

2.1 Stochastic Partial Differential Equation (SPDE) Approach

We shall first define the spatio-temporal covariance σ^{2*} explicitly:

$$\begin{aligned} \sigma^{2*} &= Cov(\gamma(s, t), \gamma(s', t')) \\ &= \sigma_\gamma^2 \mathbf{1}_{(t=t', s=s')} + \sigma_\gamma^2(r) \mathbf{1}_{(t=t', s \neq s')}, \end{aligned} \tag{3}$$

where $\mathbf{1}_{(\cdot)}$ is the indicator function, and σ_γ^2 is defined by Matern on the spatial distance r :

$$C(r) = \frac{1}{\Gamma(\nu)2^{\nu-1}} (kr)^\nu \kappa_\nu(kr), \tag{4}$$

and κ_ν is the modified Bessel function of the second kind, ν is the smoothness parameter, k is the scaling parameter and range $\rho = \frac{\sqrt{8\nu}}{k}$. In the following, ν is specified to 1.

Spatial dependence of areal units could be modeled by a Gaussian Markov random field (GMRF), which is a multivariate $N(0, Q)$ distribution, with sparse precision matrix Q . This sparsity gives a significant computational advantage. For a Matern field $X(s)$ with domain D_s , i.e. a second order stationary and isotropic Gaussian field with Matern covariance, we adopt the SPDE approach to define the Matern field as a linear combination of the basis based on a triangulation of the domain. If given a pseudodifferential operator $(\kappa^2 - \Delta)^{\frac{\alpha}{2}}$ with $\alpha = \nu + \frac{m}{2}$, $\kappa > 0$, $\nu > 0$, $\Delta = \sum_{i=1}^m \frac{d^2}{dx^2}$, then the Matern Field $X(s)$ is a solution to the stochastic partial differential equation (SPDE):

$$(\kappa^2 - \Delta)^{\frac{\alpha}{2}} X(s) = \omega(s), \tag{5}$$

where $\omega(s)$ is white noise following spatial $N(0, 1)$. Furthermore, given a triangulation, basis functions defined on the triangulation ψ_j , and the Matern field as the solution to the SPDE $X(s)$, we have:

$$X(s) = \sum_{j=1}^k w_j \psi_j(s), \tag{6}$$

where the weights (w_1, \dots, w_k) are a GMRF. γ_t in the dynamic spatiotemporal models is a Matern Field that can now be specified with a GMRF with precision matrix Q_s via the SPDE approach. The dynamic spatiotemporal model (2) becomes:

$$\begin{aligned} Y(s, t) &= \beta Z(s, t) + \sum_{i=1}^n H_{ij} \Phi(s, t) + \epsilon(s, t), \\ \Phi(s, t) &= \lambda\Phi(s, t - 1) + \gamma^*(s, t), \end{aligned} \tag{7}$$

where $H_{ij} = \mathbf{1}_{(\text{Triangulation vector } j \text{ is placed at location } i)}$, and $\gamma^*(s, t)$ now follows $N(0, Q_s^{-1})$, $\Phi(s, 1)$ follows $N(0, \frac{Q_s^{-1}}{1-\lambda^2})$.

We implement the integrated nested Laplace approximation (INLA) to compute the dynamic model. INLA (Rue et al, 2009, [16]) is a computational Bayesian procedure to approximate the posterior marginals of the latent states and the hyperparameters. In implementing INLA, we write the observation equation in matrix form, then stack the simple model building blocks into large complicated models. The details of the procedure may be found in Rue et al (2009, [16]). In Section 4, we adopt R-INLA (Blangiardo & Cameletti, 2015 [2]) in model computations.

The framework is specifically designed to be computationally efficient and flexible for accommodating different spatio-temporal datasets. In the following sections, we consider the novel GFW (2018, [5]) dataset as a special case for implementing the proposed dynamic model and investigate the influence of the covariate, sea surface temperature.

3. Exploratory Data Analysis

We begin with some necessary assumptions for the GFW (2018, [5]) dataset. Gaussianity and isotropy are assumed along with the previous assumption that the Earth is a sphere. These are not perfect representations of the real-life situation, but simplify the discussion and reduces the number of obstacles in the following analysis.

3.1 Projection and Coordinates

The original gridded fishing effort and oceans were initially recorded in degrees of longitude and latitude under GPS coordinates projection. Hence, the grids are irregular and this increases difficulty in spatial analysis. For completeness and to ensure the appropriate projection of spherical data onto 2D plots, they are reprojected using a Gall-Peters Equal Area Projection. Gall-Peters Equal Area Projection is a map projection that was first introduced by Gall (1855 [4]). As a rectangular map projection, it sacrifices surface features of the Earth in the tropics and poles but preserves the relative sizes. For given the Earth radius R , longitude (from central meridian) λ , and latitude θ , the projection is:

$$\begin{aligned} x &= \frac{R\pi\lambda}{180^\circ\sqrt{2}}, \\ y &= \sqrt{2}R\sin(\theta). \end{aligned} \tag{8}$$

The construction of plots in the following sections are all in Gall–Peters Equal Area Projection, but for ease of understanding, the resulting transformed coordinates are still referred to as longitude and latitudes.

3.2 Spherical Distance and Covariance

While Gall-Peters Equal Area Projection provides a method for mapping the spherical data onto 2D plots, in analysis it remains an issue to correctly calculate the spherical distances and covariances. We look for the great-circle distance, which is the shortest length between two points along the surface of the sphere, in kilometers for each pair of locations of nonzero fishing effort in corresponding EEZ by country. When given radius R , longitude λ_1, λ_2 , latitude θ_1, θ_2 , the great-circle distance is:

$$d = R\arccos(\sin\theta_1\sin\theta_2 + \cos\theta_1\cos\theta_2\cos(|\lambda_1 - \lambda_2|)), \tag{9}$$

where using *Rdist.earth* function from R package “fields” (Nychka et al, 2017 [14]), we found that in the GFW (2018, [5]) dataset the minimum spherical distance of each location to other locations with nonzero fishing effort is from 1.1 km to 410.4 km. This discovery suggests that these locations with observed fishing activity are relatively close, partly due to the constraint of exclusive economic zones. Also there is considerable variations between these spherical distances, with some nearest-location pairs next to each other, others hundreds of kilometers away. Therefore, a quantitative tool for predicting the fishing effort at any given location and time would help ship crew determine the allocation of fishing resources.

Several spherical covariance functions have been recently proposed, such as by Gribov and Esri (2018, [7]), and by Porcu et al (2018, [15]). One way is to compute the isotropic spherical covariance on a compact support (Wiens and Krock, 2017, [21]). Given the spherical distance d and the range of support ϕ , the covariance is:

$$C(d) = \left(1 - \frac{3d}{2\phi} + \frac{1}{2}\left(\frac{d}{\phi}\right)^3\right)1_{\{0 \leq d \leq \phi\}}, \quad (10)$$

where ϕ is the range. This isotropic spherical covariance is defined on a compact support, which is similar to the covariance proposed by Porcu et al.(2018, [15]). Constraining to an appropriately small area within the range ϕ suggests that the surface curvature can be ignored for simplicity.

3.3 Results

The following figure gives the intersection of grid cells in GFW (2018, [5]) dataset of nonzero fishing hours and the single ocean polygon in 10m Oceans file from Natural Earth (2019, [13]).

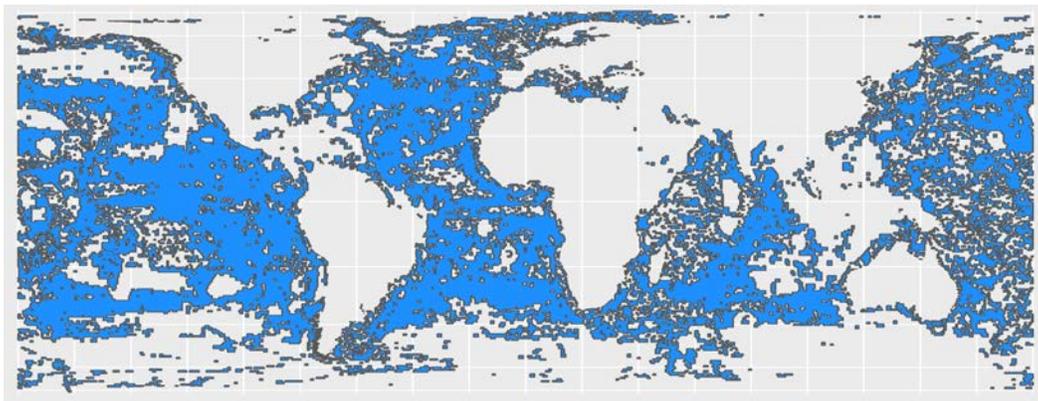


Figure 1: Projected global fishing area, 2018.

All grid cells where fishing activities took place are represented in Figure 1. Next, to appreciate the full potential of the original data, we shall look at fishing effort on a global scale. In Figure 2, lower fishing effort is labeled as yellow and higher fishing effort as darker colors. It is apparent that high fishing effort is constrained to several fishing grounds, such as the North Sea and Newfoundland in the Atlantic, the Adriatic Sea and the Aegean Sea in the Mediterranean, and East Sea and Yellow Sea in Eastern Asia. Also, there is a significant difference between the regions close to the coastline and the high seas. This is because commercial fishing in the high seas and in EEZ are very different in many aspects,

such as size of vessels, profit per fishing trip, etc. (Sumaila et al, 2015, [20]). Large areas in the high seas have negligible fishing effort, such as in Northern and Southern Pacific.

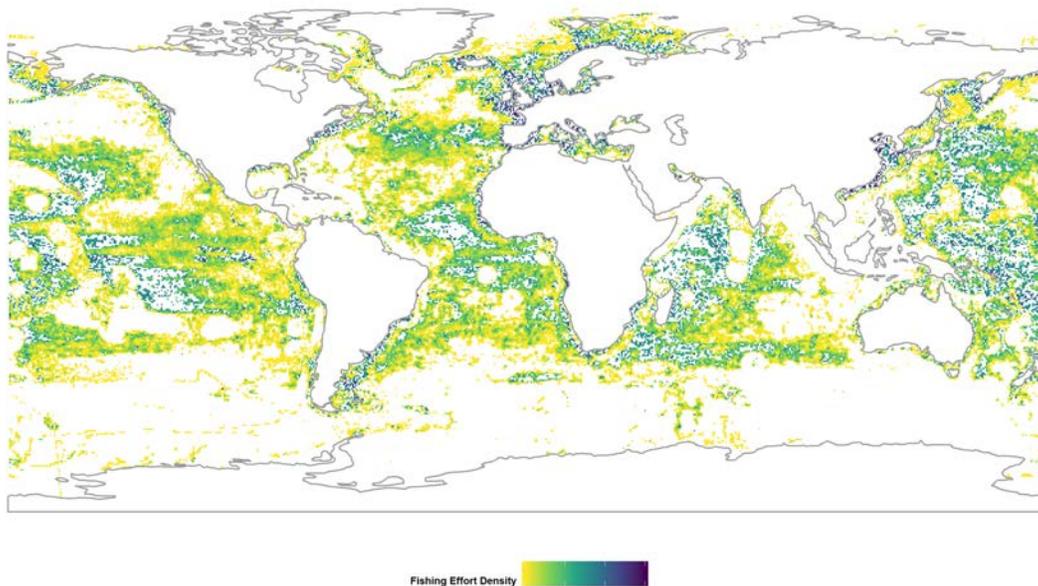


Figure 2: Global fishing effort per sq km, 2018.

For our project, the interest is on the fishing effort within the exclusive economic zone (EEZ), which is the area within 200 nautical miles from the territorial coastline. The EEZ boundaries are obtained from Marineregions.org (2019, [11]).

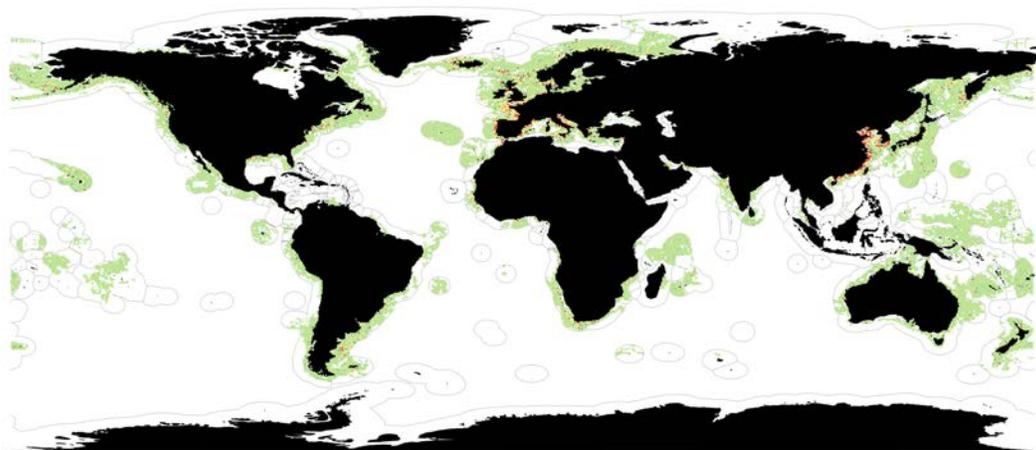


Figure 3: Global fishing effort in EEZ, 2018.

Figure 3 suggests that the fishing activities in the United States EEZ seem to concentrate in the Pacific Northwest, the Mexican Gulf, and New England. Also, the color of grid cells are predominantly light green, with occasional instances of warmer colors indicating higher fishing effort per square km. The concentration of fishing effort values in a relatively short interval resulted in most grid cells sharing similar color, which indicates that the grid-wise variation is insignificant relative to the maximum fishing effort. On the other hand, popular fishing grounds, such as Newfoundland, North Sea and the Japanese Sea are more

distinguishable. For a comparison between countries, we refer to Figure 4 and Figure 5 below.

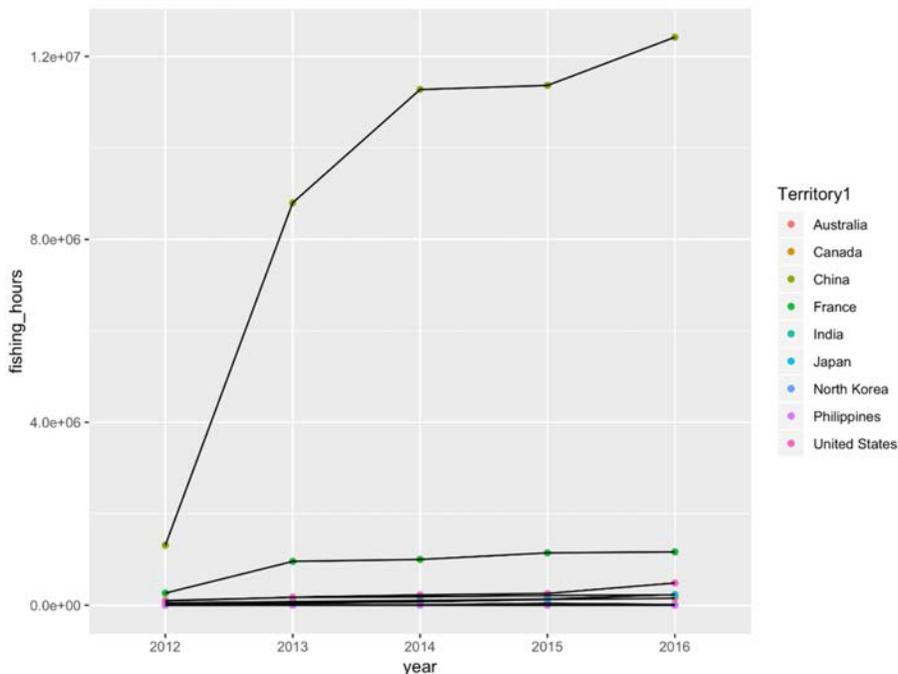


Figure 4: Total annual fishing effort in EEZ by country, 2012-2016.

In Figure 4, China is reported to have significantly higher total fishing hours than any other country by a large margin, and France also stands out from the rest. All of these top countries have increasing fishing effort.

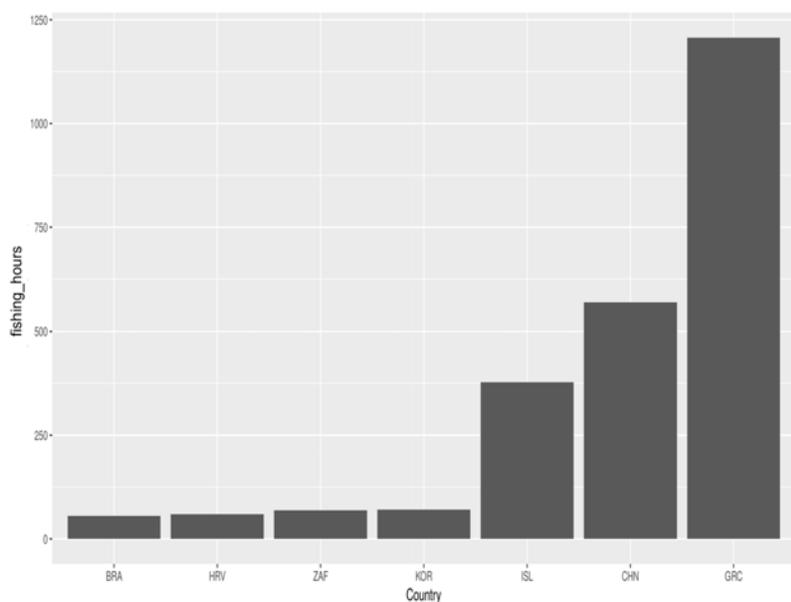


Figure 5: Highest fishing effort in EEZ per sq km by country, 2018.

From top down, the countries listed in Figure 5 are Greece, China, Iceland, South Korea, South Africa, Croatia and Brazil. The fishing effort in EEZ by country is very

different from that of total fishing effort. Even within the top seven countries, the per square km fishing hours is hugely different, ranging from over 1200 to below 100.

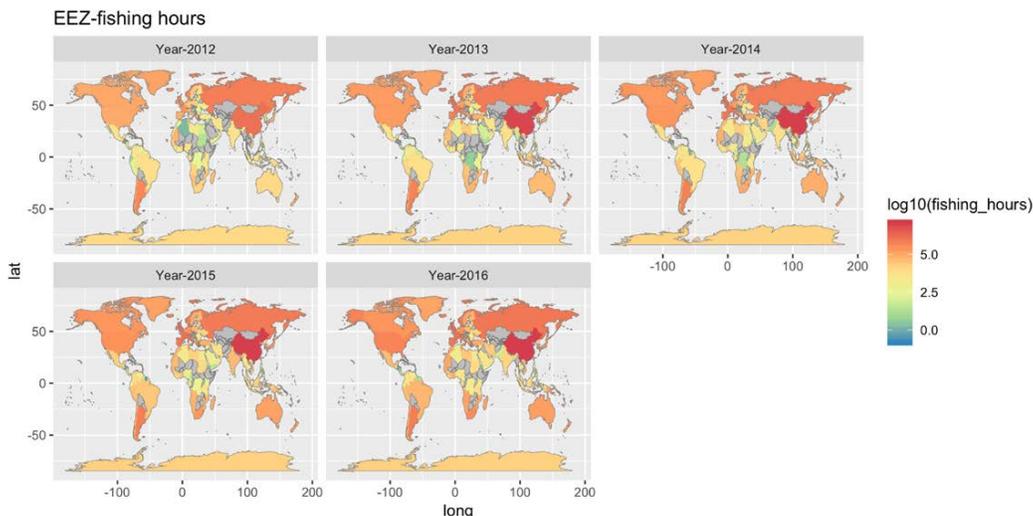


Figure 6: Change in fishing effort by country in EEZ, 2012 to 2016.

In Figure 6, annual changes are vastly different for different countries. By adding a temporal dimension to the spatial analysis of fishing effort, a dynamic point of view is achievable. This calls for establishing a dynamic model that is more appropriate for the purpose of guiding ongoing and future commercial fishery activities and related economic policies. Together, Figures 4 to 6 underline the increasing global trend of fishing effort in EEZ by year from 2012 to 2016.

4. Implementation Results

In this section, we apply the models from Section 2 to the GFW (2018, [5]) dataset. We begin with more data wrangling for continental US EEZ. We first select fishing hours from 2012 to 2018 in 0.01 by 0.01 degree grids, then import EEZ boundaries for east coast and west coast of continental US separately as polygons. After importing sea surface temperature data from NOAA, it is matched with fishing hours by location into grids of 1 degree latitude by 1 degree longitude, for each week from 2012 to 2018. This reduces computational complexity yet retains sufficient information to reflect geographical and seasonal variations. Furthermore, to address the third research question and compare between the EEZ on the Atlantic coast and the Pacific coast, we divide the data into two subsets.

Mesh constructs the triangulation of the domain for the basis functions and the figures below shows the results. There is a tradeoff between computation cost and estimation accuracy. Having more triangles achieves higher accuracy, but would take longer to compute. The desired mesh would have smaller triangles where the original data is dense and larger triangles where the original is more sparse.

In Figure 7 and Figure 8, the blue lines show the boundaries of the EEZ and the red dots indicate the grid cells. The mesh is desirable and achieves a good balance between computation cost and estimation accuracy. INLA produced the following estimated observation equations for east coast and west coast:

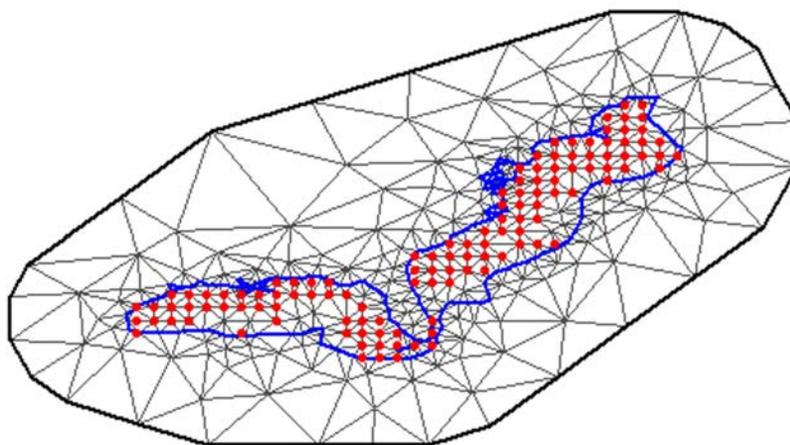


Figure 7: Triangulation of east coast EEZ.

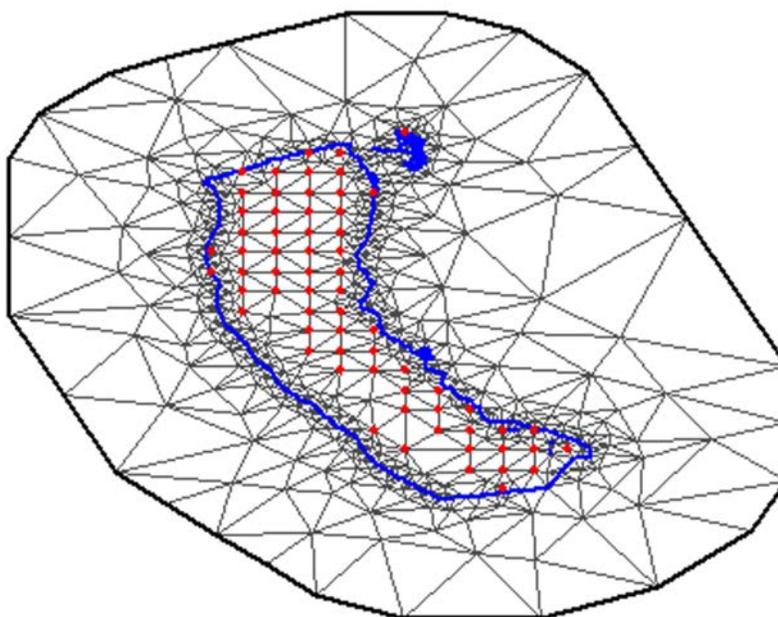


Figure 8: Triangulation of west coast EEZ.

$$\begin{aligned}
 Y(s, t) &= 59.93 + 4.32Z(s, t) + \Phi(s, t) + \epsilon(s, t), \text{ for east coast,} \\
 Y(s, t) &= 27.57 + 3.59Z(s, t) + \Phi(s, t) + \epsilon(s, t), \text{ for west coast.}
 \end{aligned}
 \tag{11}$$

Overall both observation equations show a positive relationship between the sea surface temperature and fishing effort, which fits the intuition that in general warmer waters and seasons correspond to more fishing activities along both the east and west coast of United States. At the same time, east coast would have higher fishing hours if holding all else constant. The rest of the posterior estimates of parameters are given in the following tables:

The most significant difference between Table 2 and Table 3 is that the variation for the white noise in the observation equation for east coast is considerably higher than for west coast. The rest of the estimates look reasonable. These posterior estimates and densities for the parameters in Table 2 and Table 3 can be summarized in Figure 9 and Figure 10 below,

Table 2: Posterior estimates of parameters for east coast.

Parameter	Posterior estimates
σ_ϵ^2	11212.49
σ_γ^2	17.45
ρ	1.04
λ	0.85

Table 3: Posterior estimates of parameters for west coast.

Parameter	Posterior estimates
σ_ϵ^2	2952.089
σ_γ^2	14.30
ρ	1.06
λ	0.89

where the red line indicates the expected value of the parameter.

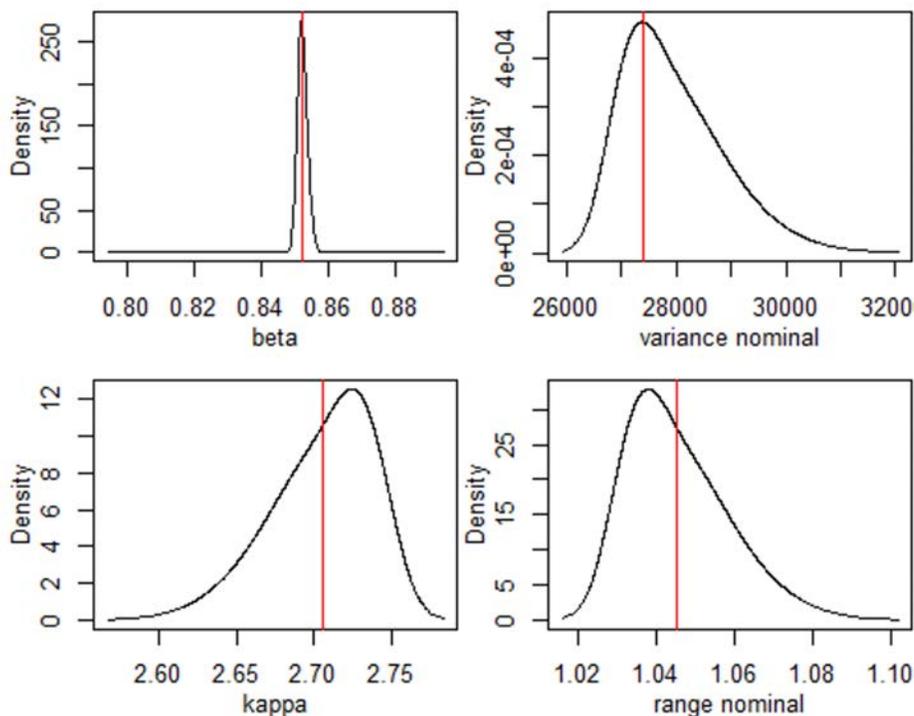


Figure 9: Posterior densities for east coast EEZ.

In Figure 9 and Figure 10, β , the variance, κ and the range ρ are represented here and the posterior densities are reasonable. Based on these, we continue to carry out the predictions for fishing effort separately for the east coast and the west coast EEZ. To save space and for better presentation, only the first week of each month is selected in construction of Figure 11 and Figure 12.

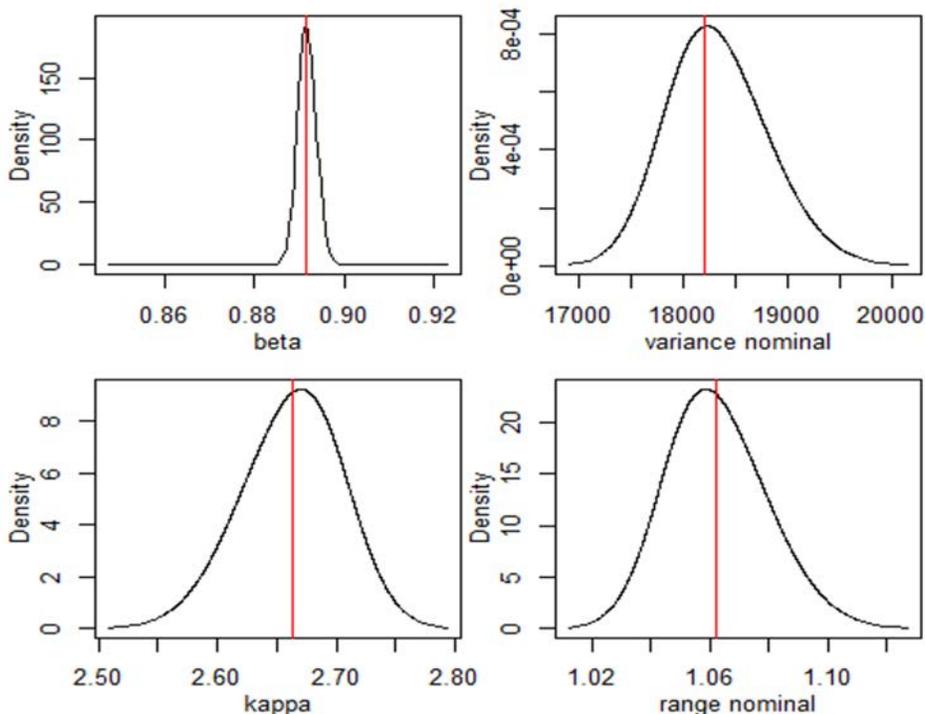


Figure 10: Posterior densities for west coast EEZ.

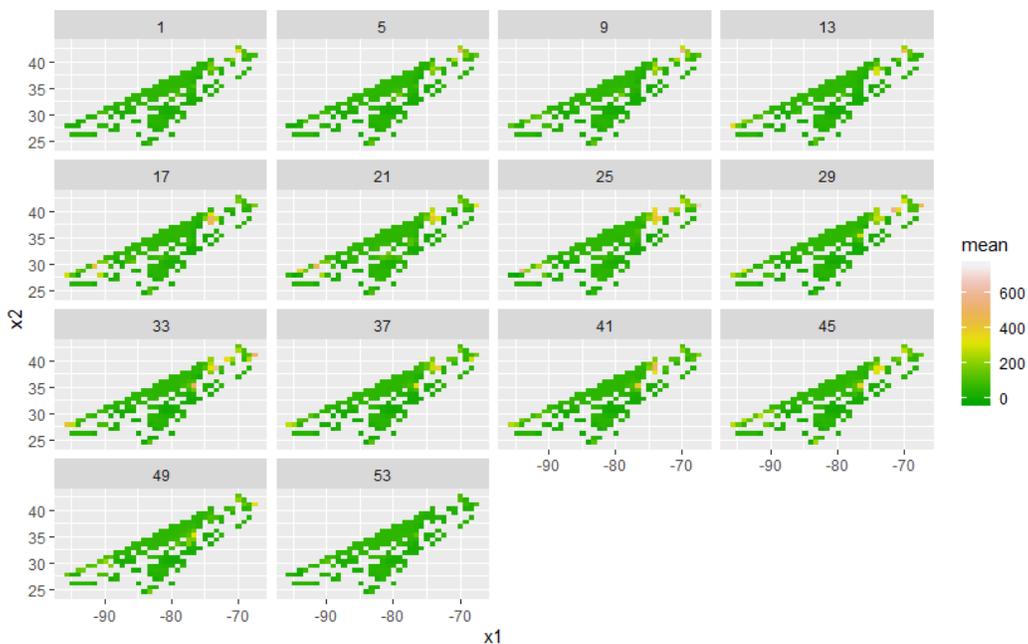


Figure 11: Predicted weekly fishing effort in east coast EEZ.

In Figure 11 featuring the one-year-ahead prediction for east coast fishing effort based on location, time and sea surface temperature, there is more lighter color in grid cells in the summer months, while activities concentrate in the Northeast and Southwest, corresponding to New England and the Mexican Gulf. For the west coast, fishing activities are also

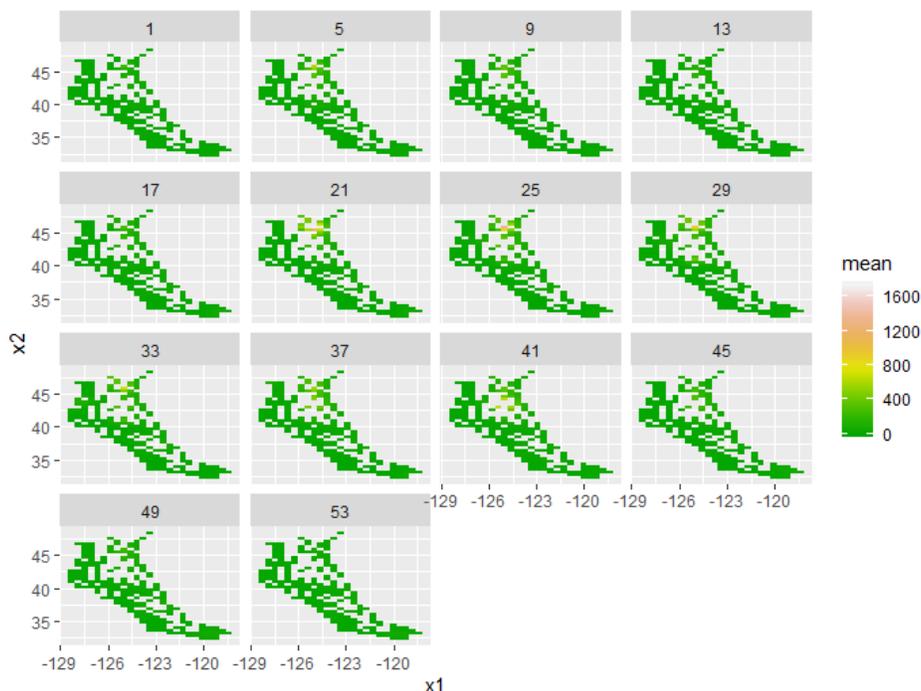


Figure 12: Predicted weekly fishing effort in west coast EEZ.

more common in the summer, while concentration is in the Pacific Northwest. These results echo general intuition and demonstrate the effectiveness of the dynamic spatio-temporal model in providing useful predictions based on the novel GFW (2018, [5]) dataset.

5. Discussion

We have incorporated covariate effects in a dynamic spatio-temporal model and studied how fishing effort is related to sea surface temperature. The dynamic perspective is more suitable for capturing the nature of shifting temperature in different time and place, therefore it is a more accurate representation of this relationship.

The dynamic models are implemented via a computationally efficient approach which consists of solving SPDE and implementing INLA in R. In particular, we introduced GMRF to model the spatio-temporal covariance and the resulting sparsity improved the computations. INLA further simplifies the inference process.

Using novel global fishing dataset at a more detailed geographical and temporal level than previous fishery studies, this project reveals that fishing effort is positively related to sea surface temperature, with the east coast EEZ having more fishing hours if under the same conditions. The model provides reasonable predictions for fishing effort and highlights the environmental influence for commercial fishing.

Given the variations in fishing hours reflected by the model and illustrations, future conservation efforts regarding fishing regulations may take into account not only the specific location and time, but also the sea surface temperature when analyzing expected fishing activities. A realistic prognosis is vital to determine the balance between economic profit and sustainable development.

For future research, computational efficiency is key in dealing with large spatial datasets. We hope to explore alternative strategies to outperform the current SPDE approach, and compare the performances. As another possible improvement, some assumptions do not

adequately reflect reality, and further attempts should be made to relax them. For example, we would like to consider extending the dynamic model to the nonGaussian case. This requires numerical approximations of the nonGaussian case as a Gaussian, therefore adding to the computational burden. Furthermore, the model has the potential to incorporate more covariates such as vessel types, fish species, etc. The GFW (2018, [5]) dataset also includes further fishery variables for consideration. Actual policy making calls for a more extensive consideration based on the current dynamic model. We may expand to more longitudinal variables in addition to fishing effort.

Cutting-edge technology has greatly improved global monitoring data. In particular, records of fishing activities in near real-time are vital for preserving our oceans and promoting sustainability. In this project, we have introduced and implemented statistical tools that offers a dynamic perspective on human fishing effort, providing a quantitative platform for the prediction of fishing hours and a comparison between patterns in continental United States EEZ. These measures contribute to environmental protection efforts that not only enable sustainability, but also promote the public good .

References

- [1] Banzon, V., T. M. Smith, M. Steele, B. Huang, and H. Zhang (2020). Improved Estimation of Proxy Sea Surface Temperature in the Arctic. *Journal of Atmospheric and Oceanic Technology*, 37, 341349.
- [2] Blangiardo, M., & Cameletti M. (2015). *Spatial and spatio-temporal Bayesian models with R-INLA*. John Wiley & Sons.
- [3] Fisheries and Aquaculture Department, Food and Agriculture Organization of the United Nations (1967). Retrieved from <http://www.fao.org/fishery/statistics/en>
- [4] Gall, J. (1885). Use of cylindrical projections for geographical, astronomical, and scientific purposes. *Scottish Geographical Magazine*. 1 (4): 119123.
- [5] Global Fishing Watch. (2018). Retrieved from <https://github.com/GlobalFishingWatch/paper-global-footprint-of-fisheries>.
- [6] Godana, A., Mwalili, S., and Orwa, G. (2019). Dynamic spatiotemporal modeling of the infected rate of visceral leishmaniasis in human in an endemic area of Amhara regional state, Ethiopia. *PLOS ONE*, 14(3).
- [7] Gribov, A., and Krivoruchko, K. (2017). New Flexible Compact Covariance Model on a Sphere. Retrieved from <https://arxiv.org/abs/1701.03405>.
- [8] Hefley, T.J., Hooten, M.B., Hanks, E.M., Russell, R.E., Walsh, D.P. (2017). Dynamic Spatio-temporal Models for Spatial Data. *Spatial Statistics*, 20, 206 - 220.
- [9] Kroodsma, D., Mayorga, J., et al. (2018). Tracking the global footprint of fisheries. *Science*, (359)6378, 904 - 908
- [10] Lindstrom, J., Szpiro, A. A., Sampson, P. D., Oron, A. P., Richards, M., Larson, T. V., & Sheppard, L. (2014). A Flexible Spatio-Temporal Model for Air Pollution with Spatial and Spatio-Temporal Covariates. *Environmental and ecological statistics*, 21(3), 411433.
- [11] Marine Regions (2019). Retrieved from <http://www.marineregions.org/downloads.php>.

- [12] Martins, T., Simpson, D., Lindgren F., and Rue, H. (2013). Bayesian computing with INLA: new features. *Computational Statistics & Data Analysis*, 67, 68 - 83.
- [13] Natural Earth. (2019). Retrieved from <http://www.naturalearthdata.com/downloads/>.
- [14] Nychka, D., Furrer, R., Paige, J., and Sain, S. (2017). *fields: Tools for spatial data*. doi: 10.5065/D6W957CT (URL:<http://doi.org/10.5065/D6W957CT>), R package version 10.0, <https://github.com/NCAR/Fields>.
- [15] Porcu, E., Alegria, A., Furrer, R. (2017). Modeling Temporally Evolving and Spatially Globally Dependent Data. Retrieved from <https://arxiv.org/abs/1706.09233>.
- [16] Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society, Series B*, 71(2), 319 - 392.
- [17] Sea Around Us. (2016). Retrieved from <http://www.seaaroundus.org/data>.
- [18] Simons, R.A. (2019). ERDDAP. Retrieved from https://upwell.pfeg.noaa.gov/erddap/info/noaa_esrl_cf70_04df_9df1/index.html. Monterey, CA. NOAA/NMFS/SWFSC/ERD.
- [19] Stroud, J., Muller, P., & Sanso, B. (2001). Dynamic Models for Spatiotemporal Data. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 63(4), 673 - 689.
- [20] Sumaila, U., Lam, V., Miller, D. et al. (2015). Winners and losers in a world where the high seas is closed to fishing. *Sci Rep* 5, 8481
- [21] Wiens, A., and Krock, M. (2017). *fields* vignette. Retrieved from <https://github.com/NCAR/fields/blob/master/fieldsVignette.pdf>