9th MEETING OF THE SCIENTIFIC COMMITTEE

Held virtually, 27 September to 2 October 2021

SC9-HM02

Jack Mackerel Interannual Habitat Variability
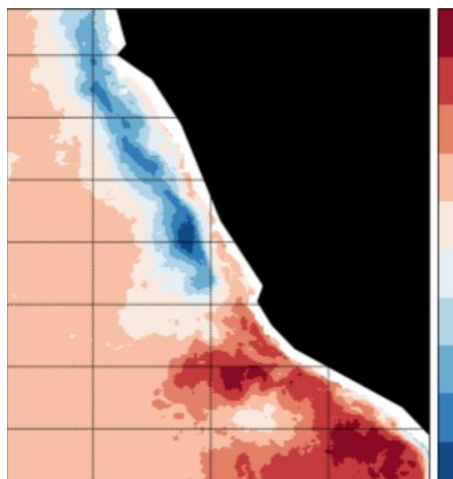
Peru

INSTITUTO DEL MAR DEL PERU
IMARPE

Sociedad Nacional de Pesquería
SNP

INSTITUTO HUMBOLDT
INVESTIGACIÓN MARINA Y ACUÍCOLA
IHMA

Interannual variability of the habitat suitability of jack mackerel in the Northern Peru Current System, 2011-2019

by

Daniel Grados[1], Jorge Csirke[1], Carlos Valdez[1], Salvador Peraltilla[2] and Mariano Gutierrez[3]

[1] Instituto del Mar del Perú (IMARPE), Callao, Perú
[2] Sociedad Nacional de Pesquería (SNP), Lima, Perú
[3] Instituto Humboldt de Investigación Marina y Acuícola (IHMA), Lima, Perú

This report contains information on the Jack mackerel fish stock and fishery in Peruvian jurisdictional waters that, we reiterate, the delegation of Peru, in use of its discretionary powers, voluntarily provides for the purpose of information and support to the scientific research work within the Scientific Committee of the SPRFMO. In doing so, while referring to Article 5 of the Convention on the Conservation and Management of High Seas Fishery Resources in the South Pacific Ocean and reiterating that Peru has not given the express consent contemplated in Article 20 (4) (a) (iii) of the Convention, Peru reaffirms that the decisions and conservation and management measures adopted by the SPRFMO Commission are not applicable within Peruvian jurisdictional waters.

2021

SUMMARY

Jack mackerel is a straddling species that is associated with relatively low temperatures and low levels of chlorophyll-a. To understand how jack mackerel responds to environmental changes one should also study its habitat. In this paper, we build a Species Distribution Model through Random Forest model (RF) by month between 2011 and 2019. The RF has a 98% capacity to predict the observed habitat of jack mackerel concentrations available to the Peruvian industrial fishing fleet. In addition, chlorophyll-a was the most recurrent variable in predicting the habitat in different seasons of the year. After chlorophyll, temperature and salinity were the most important variables. Based on the results of habitat prediction, a seasonal pattern was observed, where in warmer months the probability values of presence of the fish were higher than during colder months. In addition, a high internal variability of the habitat was observed, which was computed using the Empirical Orthogonal Function (EOF) analysis. This interannual variability is modulated by El Niño events, which were observed through the Oceanic Niño Index (ONI) for the Niño 1+2 region.

TABLE OF CONTENTS

1. INTRODUCTION

Jack mackerel (*Trachurus murphyi*, Nichols 1920) is a straddling species with a large spatial distribution in the South Pacific. The range of its spatial distribution goes from the equator to the austral region of Chile along the South American coastline, and from the coast of South America to New Zealand and Tasmania (Serra, 1991) along the subtropical convergence between 35° and 45° S in the South Pacific. Through this large spatial distribution, it can be found associated with a wide range of environmental parameters (Alegre *et al*., 2015). To study the relationship between the presence of jack mackerel and environmental parameters off Peru, in the northern part of the Peru Current system, or Northern Peru Current System (NPCS), also known as Northern Humboldt Current System, we use a Species Distribution Model (SDM).

Several statistical approaches are available to model the species distribution in relation to environmental parameters (Torrejon *et al*. 2019). Traditional methods such as GAM, GLM, etc., were used for several years. These statistical methods require that important assumptions be made, such as the independence of observations and normality in the dependent variable (Hengl *et al*. 2018), but these model assumptions are seldom true in the ecological context.

Nevertheless, machine learning algorithms such as Maximum Entropy Modelling (MaxEnt), Random Forests (RF), Classification and Regressions Trees, among others, have been shown to outperform the traditional regression-based approaches. Similar studies of species distribution modeling report the superior ability of RF algorithms in comparison to traditional regression-based algorithms like the logistic regression approach (Rather *et al.* 2020).

The first main objective of this study is to implement a SDM model based on RF algorithms, and the second one is to study the interannual variability of the habitat suitability of jack mackerel concentrations available to the industrial fishing fleet off Peru, in the NPCS. To implement the RF model, we used georeferenced catch data by fishing sets made by the pelagic Peruvian industrial fleet between 2011 and 2019. In addition, to study the spatial and temporal variability we carried out an Empirical Orthogonal Function (EOF) type of analysis (Halldor and Venegas 1997).

2. MATERIALS AND METHODS

2.1. Data

Environmental data

Several environmental parameters for the period from January 2011 to December 2019 were gathered from different sources, and were used in the present study. This included oceanographic data from Hycom (https://www.hycom.org/), such as on sea surface temperature (SST), sea surface salinity (SSS), sea surface height (SSH), U and V geostrophic currents. Then, using fields of SST we computed the thermal gradient (Gsst), which is a proxy of the oceanographic fronts (Bowman and Iverson 1978, Fedorov 1986). Also, using U and V geostrophic components we computed the eddy kinetic energy (EKE). Due to the high cloud coverage over the area of the NPCS, modeled chl-a data was used from the Copernicus Marine Service (https://marine.copernicus.eu/). All data was downloaded at a daily scale. Every jack mackerel fishing set was linked to each environmental parameter according to its position and date.

Jack mackerel data

Records of a total of 5 506 positive jack mackerel fishing sets from the Peruvian industrial purse-seine fleet collected from 2011 to 2019 were used. The data was resampled monthly using grids of 5x5 nautical miles (nm).

Pseudo-absence modelling

To cope with the lack of confirmed jack mackerel absence data, we built a set of pseudo-absence data using Ecological Niche Factor Analysis (ENFA, Hirzel *et al*. 2002). ENFA is a technique that compares the distribution in the cells with presence of the species of interest, with the distribution in the whole set of cells.

To generate pseudo-absence data, we used the environmentally - and geographically - weighted method proposed by Hengl *et al*. (2009). This method is based on both the Habitat Suitability Index (HSI, derived through ENFA) and the distance from the observations that are subsequently used to weight pseudo-absence points selection. In this way, pseudo-absences are located both in areas of low HSI (unsuitable habitat) and further away from the locations of positive occurrence.

2.2. Modelling approach

RF constructs a regression or classification tree by successively splitting the data based on single predictors. Each split forms a branch in the decision tree and trees are grown without pruning. RF uses bagging (bootstrap aggregation) that builds a large number of trees and the model output is obtained by averaging the aggregated tress or by maximum vote. During bagging, a bootstrap sample is randomly drawn to build each tree, and the data not included in the bootstrap sample is termed as 'out-of-bag' (OOB) which is used to estimate an unbiased error rate and to rank the variable importance. Before running the RF, the data was split in two subgroups, one for training data (the 80% of the whole data), and the other 20% for testing. We ran the RF for training data in order to evaluate the results and the performance by comparing the obtained model with the testing data.

Model assessment

We used the area under the receiver (AUC), the operating characteristic curve (ROC), Out-of-Bag estimate (OOB) and confusion matrix as means to assess the model performance.

Model Prediction

To predict the jack mackerel habitat suitability, we run the RF model with the whole environmental data. The runs were made by each month, from 2011 to 2019.

2.3. Interannual variability

By running the RF model for each month from 2011 until 2019 we built 3D maps (longitude by latitude by time). Then, using these results, we computed the monthly climatology of the habitat suitability, and used the monthly climatology to calculate the anomalies of the habitat suitability.

Using the anomalies of the habitat suitability, we apply an empirical orthogonal function (EOF) to obtain the spatial and temporal variability. We used the Oceanic Niño Index for the Niño 1+2 region (ONI 1+2) to compare the results obtained.

3. RESULTS

From the oceanographic data associated to each of the 5 506 positive jack mackerel fishing sets we estimated and analyzed the several seasonal patterns in the presence and availability of fishable concentrations of this species to the fishing fleet, and some clear seasonal differences were observed. During warmer months of summer and fall, catch records indicating the presence and availability of jack mackerel concentrations to the fishing fleet were associated with higher values of SST, SSS, Gsst, chl-a and SSH than those observed during the colder months of winter and, in most cases, also spring (Figure 1). However, the seasonal pattern for SSS is not very clear and for EKE there appears to be is no seasonal pattern.
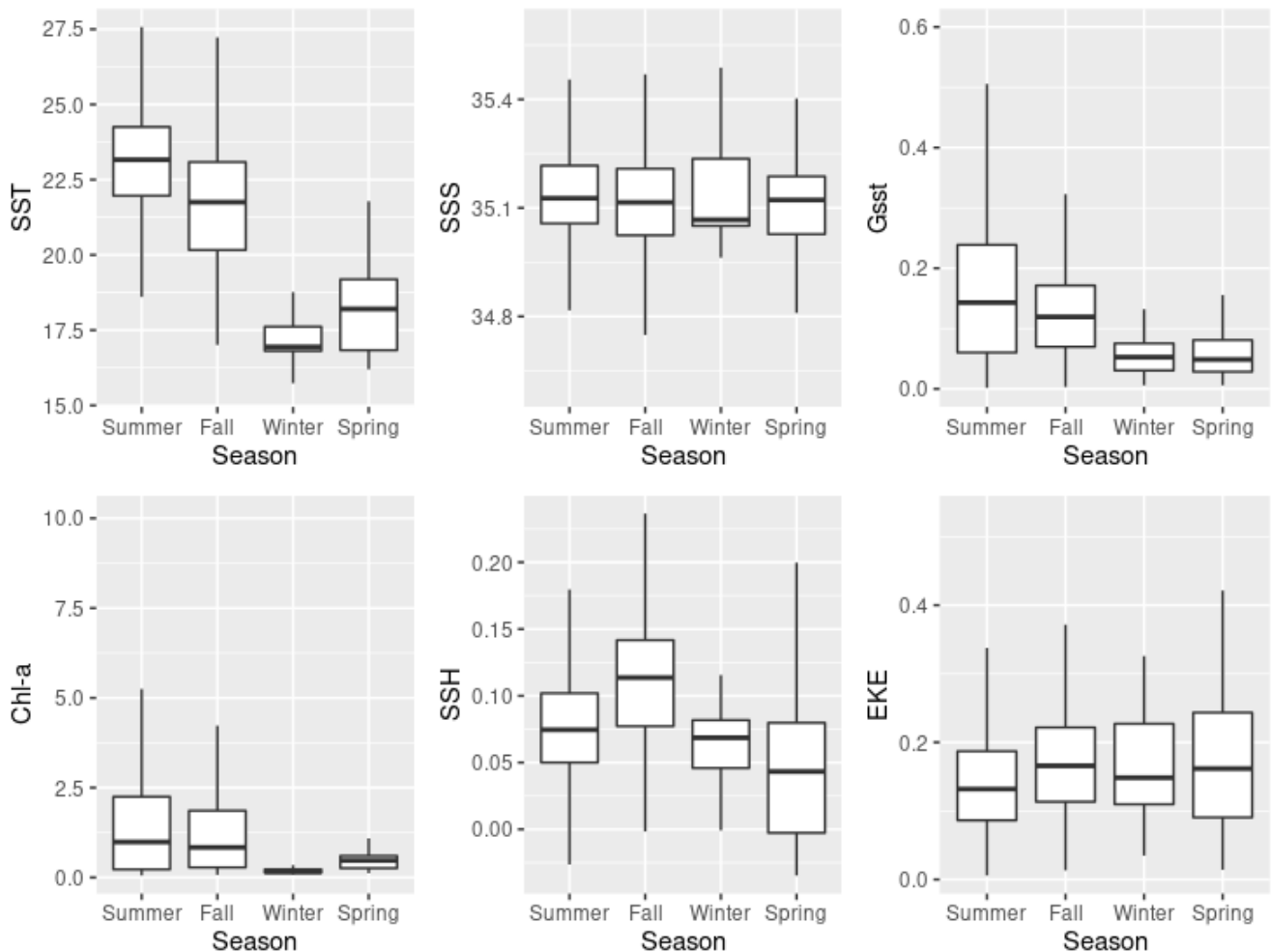

Figure 1.- Seasonal pattern of oceanographic variables associated to the presence and availability of jack mackerel to the fishing fleet for the period 2011-2019. The y-axis indicates each environmental covariate used in the study.

Random Forest and the importance of variables

Using the presence and pseudo-absence data, and their association with environmental parameters, we fitted a RF model. We split the data in two datasets, the training data (80% of the whole data) and the validation data (20% of the data).

The fitted model has an AUC of 0.98, denoting a 98% chance of predicting the presence/absence of jack mackerel concentrations available to the fishing fleet, and the Out-of-Bag estimate (OOB) value was 6.2%. These values indicate a good performance of the model to predict the presence of jack mackerel in suitable concentrations for the fishing fleet.

Also, we studied the importance of the variables to assess their seasonal differences (Figure 2). The observed results indicate that the chl-a is the most important variable for all seasons except winter. The other important variables are SST and SSS. This shows that water masses and chl-a are important to predict the presence or absence of jack mackerel concentrations.
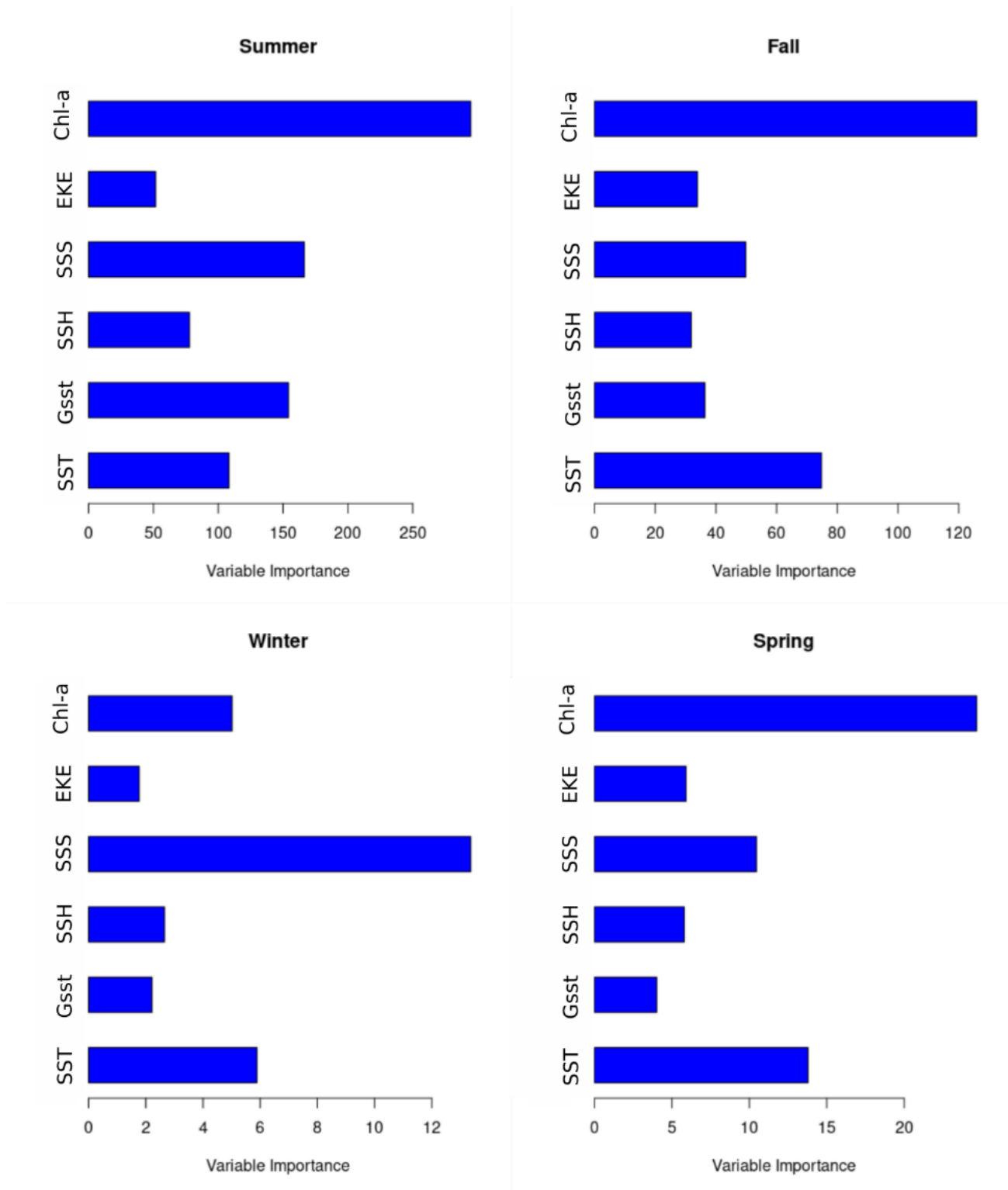

Figure 2.- Seasonal importance of the oceanographic variables in predicting the presence of jack mackerel concentrations available to the industrial fleet, for the years 2011-2019.

Projection of the habitat

Using the fields of the different environmental parameters we proceed to project the habitat at a monthly scale. We projected maps from 2011 to 2019 and then computed the climatology of these maps. Looking at the plots of the climatology (Figure 3) we observe higher probabilities of presence of jack mackerel concentrations available to the industrial fleet between January and May, with lower probabilities during the other months. These results are consistent with the lower catches of this species by the industrial feet during colder periods and also with their lower catch rates, although it is noted that due to fishery regulations and interaction with other pelagic fisheries the industrial fleet tends to take most if not all its jack mackerel catch allocation earlier in the year. Another important observation is that the higher probabilities of presence of jack mackerel concentrations are located off the center-south part of the Peruvian coast.
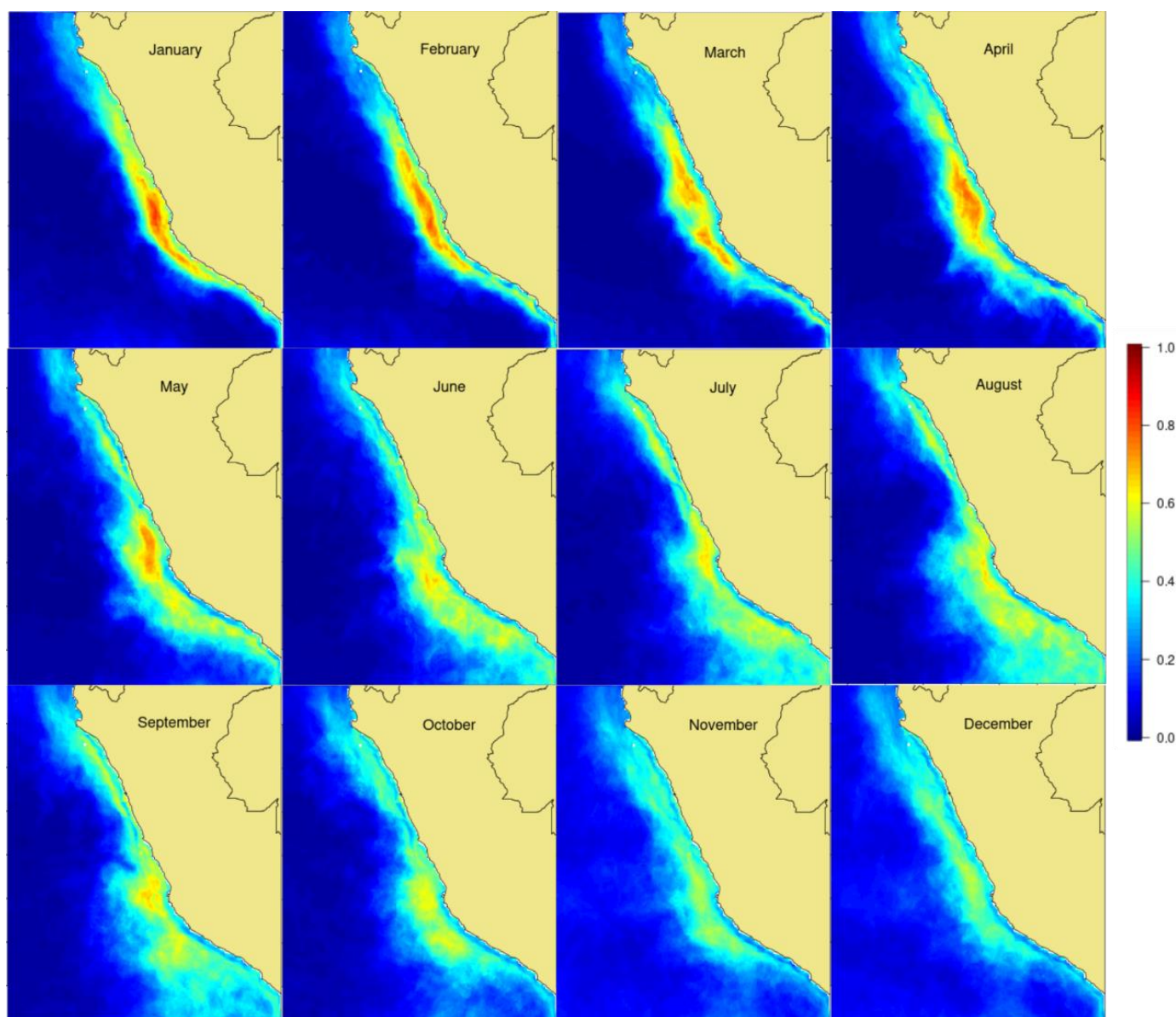

Figure 3.- Mean monthly climatology of the habitat suitability of jack mackerel concentrations available to the industrial fleet off Peru, in the Northern Peru Current System, for the years 2011-2019.

Interannual variability

Finally, using the monthly results of the projected maps, an EOF analysis was carried out, which allows to decompose the variability into spatial and temporal components. The projection of the first EOF analysis (Figure 4) shows that the spatial variability from 14°S to the south is much higher than to the north of 14°S, which indicates that in areas to the south of 14°S the jack mackerel habitat suitability is highly variable and fluctuates more than further north.

On the other hand, the estimated values of the amplitude of the principal components (PC) of the EOF (Figure 5a) indicate that there was a higher temporal variability of the jack mackerel habitat suitability between 2014 and 2017, which matches the higher environmental values of the ONI 1+2 index (Figure 5b). In addition, rather similar phases are observed in both time series (Figs. 5a and 5b).
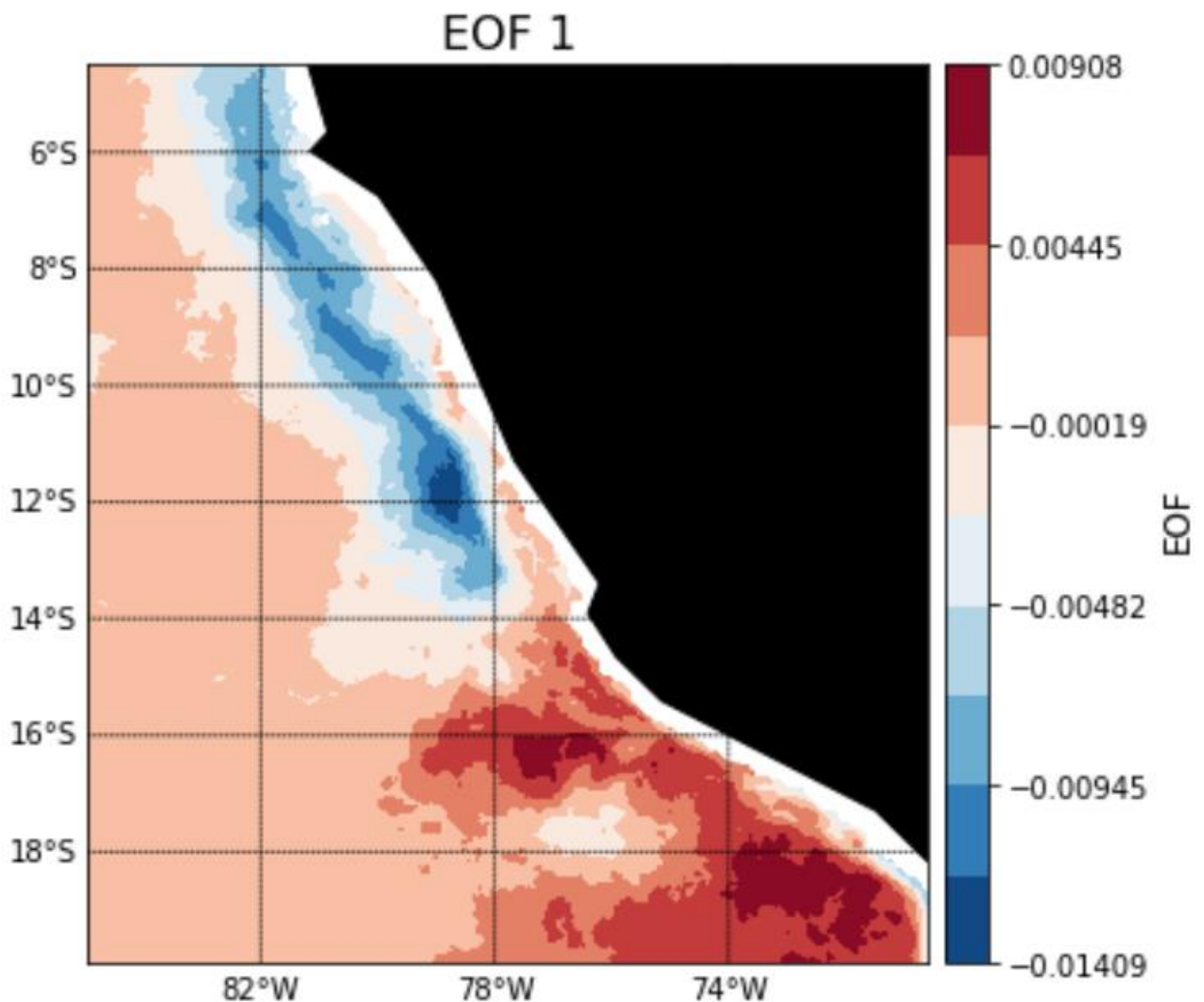

Figure 4.- Spatial variability of the jack mackerel habitat suitability as indicated by the projection of the first Empirical Orthogonal Function (EOF 1)
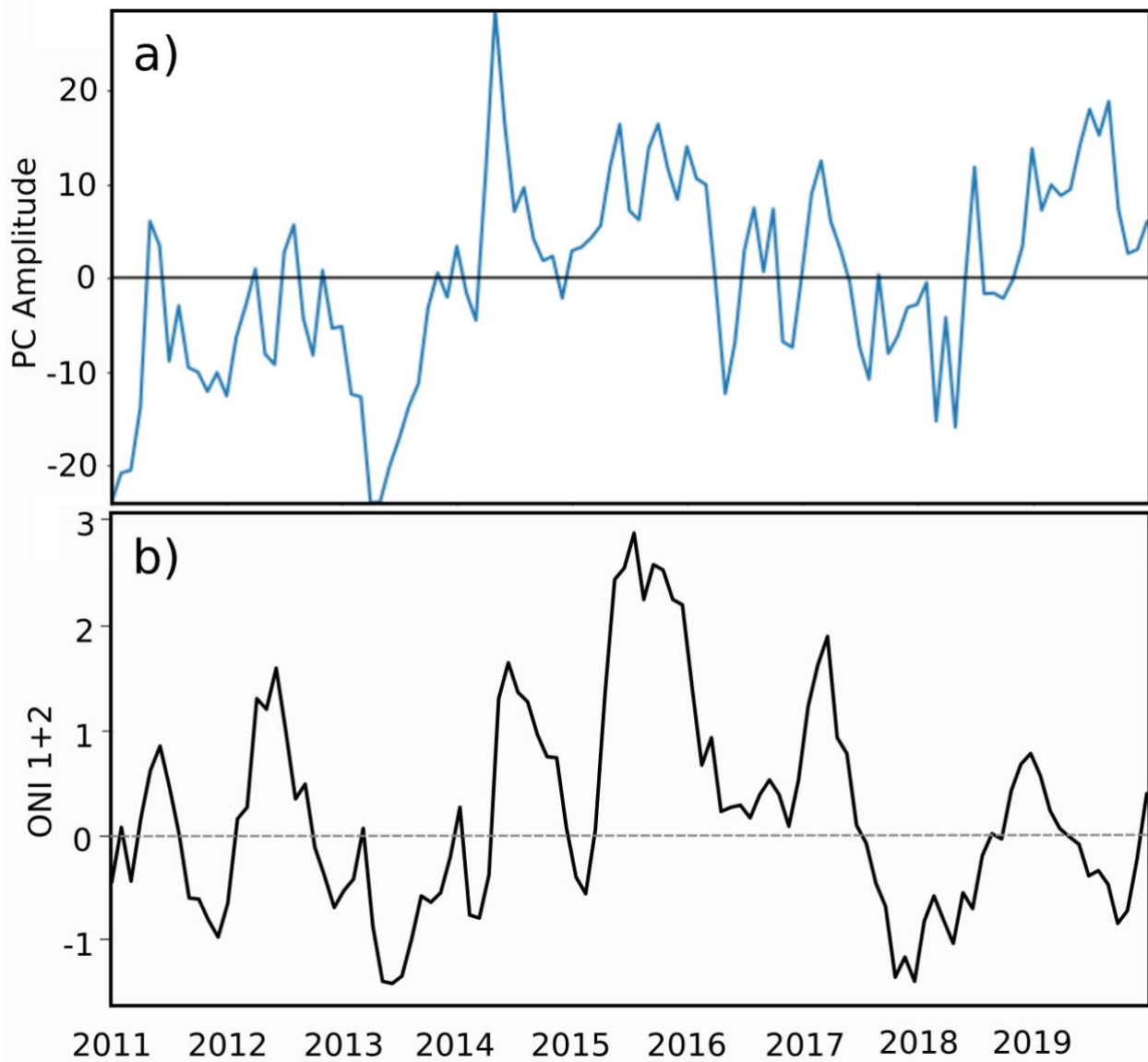
Figure 5.- a) Temporal variability of the jack mackerel habitat suitability as indicated by the first EOF (or principal component PC) amplitude; and, (b) Temporal variability of the Oceanic Niño Index for the Niño 1+2 region (ONI 1+2). Years 2011-2019.

4. DISCUSSION AND CONCLUSIONS

Data between 2011 and 2019 from the Peruvian pelagic industrial fishing fleet was used to analyze the interannual variability of jack mackerel concentrations' habitat suitability to the fishing fleet. In total, 5 506 positive jack mackerel fishing sets data were used, grouped by month. The RF model used had a capacity of 98% to identify areas with presence of jack mackerel concentrations, which gives us an acceptable level of confidence when using it. Also, other studies where machine learning models have been used to analyze changes in the habitat of wild terrestrial animals have shown that RF models perform better and with higher precision in the modeling of the habitat in comparison with more traditional statistical approaches such as GAM, GLM, etc. (*e.g.,* Rather *et al.* 2020).

9

Our results indicate that the chlorophyll-a is a variable that acquires a high importance in explaining the presence or absence of jack mackerel concentrations available to the fishing fleet, and it is only in winter that chlorophyll doesn't have the highest importance (Figure 2). However, it is noted these chlorophyll values associated with the presence of jack mackerel concentrations correspond to relatively low concentrations of chlorophyll (*i.e.*, mean values under 1.3 mg/m^3, Figure 1), which are associated with the edges of areas with high chlorophyll concentrations (*i.e.*, of 5 mg/m^3 or higher) found in Peruvian waters. The effect of chlorophyll on the presence of jack mackerel concentrations is an indirect one, since it has to transit up the food-web to at least the euphausids, which are one of the main jack mackerel preys (Alegre et al., 2013). Other variables that are important for this study are the SST and SSS, which are fundamental variables to classify water masses regarding the spatial distribution of many marine species (Swartzman et al. 2008).

A strong seasonal pattern is observed in the jack mackerel habitat suitable for the industrial purse-seine fishery, particularly when defined in relation to variables such as SST and Gsst. The higher probability values of presence and availability of jack mackerel to the industrial purse-seine fishing fleet occur during the warmer seasons (summer and fall) and are lower during the colder seasons (winter and spring). These results are coherent with the seasonality of the jack mackerel fishery by the industrial purse-seine fleet, that tends to concentrate most of its jack mackerel fishing activities during summer.

Finally, it is noted that the monthly and interannual variability of the climatology corresponding to the jack mackerel habitat suitability calculated by applying an empirical orthogonal function (EOF) shows that there is a wide temporal thermal variability, fluctuating with trends and phases that have close similarity with those of the monthly Oceanic Niño Index for the Niño 1+2 region (ONI 1+2 index). This similarity between the EOF and the ONI 1+2 index are particularly noticeable from 2014 to 2017. It is also noted that these strong thermal anomalies in the climatology of the jack mackerel habitat suitability contributes to explaining the wide fluctuations in the availability of suitable jack mackerel concentrations for the industrial purse-seine fishery.

5. BIBLIOGRAPHIC REFERENCES

Alegre, A., Espinoza, P., Espino, M. 2013. Trophic ecology of jack mackerel *Trachurus murphyi* in Peru, 1977-2011. Revista Peruana de Biología, 20, 075-082.

Alegre, A., Bertrand, A., Espino, M., Espinoza, P., Dioses, T., Ñiquen, M., Navarro, I., Simier, M., Ménard, F., 2015. Diet diversity of jack and chub mackerels and ecosystem changes in the northern Humboldt Current system: a long-term study. Progress in Oceanography, 137, 299–313.

Bowman, M. J., Iverson, R. 1978. Estuarine and plume fronts, in Oceanic Fronts in Coastal Processes, edited by M. J. Bowman and W. E. Esaias, pp. 87–104, Springer, Berlin.

Fedorov, K. 1986. The Physical Nature and Structure of Oceanic Fronts. Springer-Verlag, Amsterdam, Netherlands, p. 333.

Halldor, B., Venegas, S. "A manual for EOF and SVD analyses of climate data", McGill University, CCGCR Report No. 97-1, Montréal, Québec, 52pp., 1997.

Hengl, T., Nussbaum, M., Wright, M., Heuvelink, G., Graler, B. 2018. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ, 6:e5518.

Hengl, T., Sierdsema, H., Radovic, A., Dilo, A., 2009. Spatial prediction of species' distributions from occurrence-only records: combining point pattern analysis, ENFA and regression-kriging. Ecological Modelling, 220, 3499–3511.

Hirzel, A., Hausser, J., Chessel, D., Perrin, N., 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? Ecology, 83, 2027–2036.

Rather, T;. Kumar, S.; Khan, J. 2020. Multi-scale habitat modelling and predicting change in the distribution of tiger and leopard using random forest algorithm. Scientific Reports, 10, 1-19.

Serra, R., 1991. Important life history aspects of the Chilean jack mackerel, *Trachurus symmetricus murphyi*. Investigación pesquera, 36, 67-83.

Swartzman, G. Bertrand, A. Gutiérrez, M. Bertrand, S. Vasquez, L. 2008. The relationship of anchovy and sardine to water masses in the Peruvian Humboldt Current System from 1983 to 2005. Progress in Oceanography, 79, 228-237.

Torrejón-Magallanes, J. Grados, D. Lau-Medrano, W. 2019. Spatio-temporal distribution modeling of dolphinfish (*Coryphaena hippurus*) in the Pacific Ocean off Peru using artisanal longline fishery data. Deep-Sea Research Part II: Topical Studies in Oceanography, 169-170, 104665.