Ecological Informatics 36 (2016) 106–113

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Recruitment forecasting of yellowfin tuna in the eastern Pacific Ocean with artificial neuronal networks

Laura Karen Torres-Faurrieta [a], Michel J. Dreyfus-León [b,*], David Rivas [c]

[a] Posgrado en Ecología Marina, CICESE, Ensenada, Baja California, Mexico
[b] Centro Regional de Investigación Acuícola y Pesquera Ensenada, Instituto Nacional de la Pesca, Ensenada, Baja California, Mexico
[c] Departamento de Oceanografía Biológica, CICESE, Ensenada, Baja California, Mexico

ARTICLE INFO

Article history:
Received 1 July 2016
Received in revised form 12 October 2016
Accepted 14 October 2016
Available online 18 October 2016

Keywords:
Yellowfin tuna
Recruitment
NARX
Artificial neural network
Forecasting

ABSTRACT

The recruitment of yellowfin tuna in the eastern Pacific Ocean is modeled based on oceanographic as well as biological parameters, using two nonlinear autoregressive network models with exogenous inputs (NARX). In the first model (Model 1) the quarterly recruitment is modeled considering eastern Pacific global oceanographic conditions: the Southern Oscillation Index (SOI), the Pacific Decadal Oscillation (PDO), and spawners biomass. In Model 2, recruitment is predicted based on sea surface temperature, wind magnitude, and oceanic current magnitude of a smaller area within the eastern Pacific Ocean, considered as relevant for spawning and recruitment, and total spawners biomass. The correlation coefficient between the ANN recruitment estimate and the "real" recruitment is $r > 0.80$ in both models. Series of sensitivity analysis suggest that the SOI and the sea surface temperature are the most important variables for the recruitment in Model 1 and Model 2 also show that warm sea surface favors recruitment. A forecasting model under different climatological scenarios indicates that the recruitment of yellowfin tuna could be higher in the period 2015–2020 compared to the ones registered in the period 2009–2013.

1. Introduction

Yellowfin tuna (*Thunnus albacares*) is the second species in terms of its contribution to global tuna catches. The yellowfin tuna (YFT) fishery in the eastern Pacific Ocean (EPO) contributes yearly with >200,000 tons (IATTC, 2015). YFT in the EPO is managed by a regional fishery organization, the Interamerican Tropical Tuna Commission, as a single stock. Spawning occurs year-round, extending roughly from 26°N to 14°S, and from the coast to 140°W (Shaefer, 1998). Although YFT tuna is a multiple spawner, recruitment of the young fish to the fishery is highly variable, with a seasonal component. It has also experienced different productivity regimes (IATTC, 2013). These regime variations have been associated with oceanographic conditions, especially with sea surface temperature and other parameters affecting the feeding dynamics of the larvae. For example, wind modulates the encounter rates of the larvae with its prey through turbulence (Haury et al., 1990; Kimura et al., 2004), as well as primary production through upwelling (Cury and Roy, 1989). Moreover, oceanic currents must have effects on the dispersion and transport of larvae to areas with different concentration of prey biomass.

Recruitment (incorporation of juveniles to the fishing stock) is a very important variable as a biomass generator. Accurate prediction of recruitment to the fishery is an important tool for the management structure of any fish stock being exploited (Dreyfus-León and Schweigert, 2008), and it is useful with the aim of knowing the future status of a stock. However, there are difficulties in the yellowfin recruitment forecasting since no apparent relation between the amount of spawners and recruitment exists and it is quite possible that recruitment is shaped during early life stages, especially in the planktonic stage.

Thus, the recruitment may be influenced by biotic and abiotic factors; knowing these possible relations is of paramount importance and such relations are generally highly non-linear. A very suitable option to explore non-linear relations between certain variable and other predictive variables is the implementation of Artificial Neural Networks (ANNs). ANNs are a powerful tool to explore complex, nonlinear biological problems (Chen and Ware, 1999; Yáñez et al., 2010). ANNs are computer algorithms that simulate in a strongly simplified way the activity of neurons and information processing in the human brain (Jarre-Teichmann et al., 1995). They create their own organization and representation of information they receive through a learning stage called training. ANNs are capable to generalize the system as a whole and respond appropriately to data or situations they have not been exposed previously (Lek and Guégan, 1999; Zhang et al., 1998).

* Corresponding author at: 511 E. San Ysidro Blvd. 2430, San Ysidro, CA, USA.
E-mail address: dreyfus@cicese.mx (M.J. Dreyfus-León).

http://dx.doi.org/10.1016/j.ecoinf.2016.10.005

The purpose of this paper is to present ANNs, specifically Nonlinear Autoregressive models with exogenous Inputs (NARX), to predict recruitment of YFT in the EPO from oceanographic parameters used as proxy of conditions influencing natural mortality at pre-recruit stages, and to identify variables that are more important shaping recruitment.

2. Methods

2.1. Overview

YFT recruitment is modeled with two different NARX structures, according to the available data. In the first model (M1) the recruitment is forecasted from large scale oceanographic conditions in the EPO as explanatory variables. This model uses the spawners biomass data, the Southern Oscillation Index (SOI), and the Pacific Decadal Oscillation (PDO) index (Mantua et al., 1997) as inputs for the period from 1975 to 2012, a common period for all these variables.

The second model (M2) consists of spawners biomass data, sea surface temperature, wind magnitude, and magnitude of the sea water currents as inputs. In this case the shorter period from 1980 to 2012 was used, since the input oceanographic variables were available only for this time span. Notice that in this model (M2) these oceanographic variables represent conditions in the area where more juvenile catch is observed within the EPO. This area was chosen based on the map of catches by fishing-set developed by the Interamerican Tropical Tuna Commission (IATTC, 2015). We chose the area of the highest YFT catch associated with floating objects, assuming that there the highest recruitment occurs, since the smallest organisms (<60 cm) are obtained from fishing around floating objects (Dreyfus-León and Robles-Ruiz, 2008; IATTC, 2013). It extends from 1°N, 15°S and 75–95°W (Fig. 1).

Furthermore, in order to produce YFT recruitment forecasting we developed a variant of model M2, which includes only the oceanographic parameters and excludes the spawners biomass. These oceanographic parameters were taken for future climate-change scenarios, described in Section 2.2, but no predictions for spawners biomass were available, hence this biotic variable was excluded in this model. Nonetheless, as will be shown in the following sections, including only the abiotic, oceanographic variables in a NARX-based model results in an adequate estimate of recruitment in the EPO.

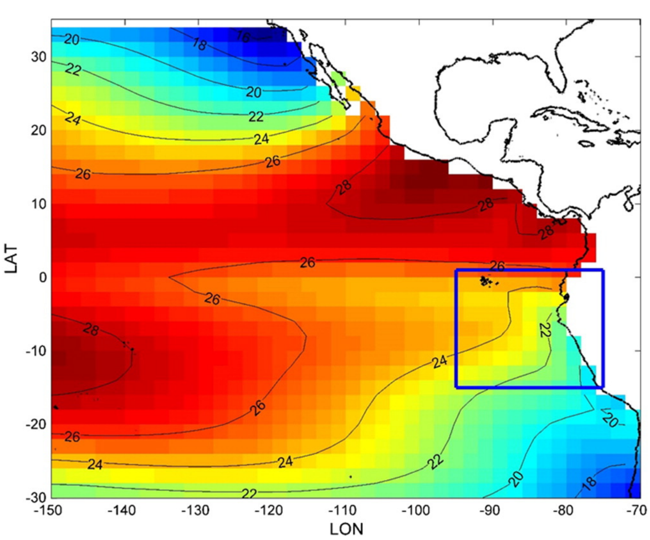

Fig. 1. Area of the highest YFT catch associated with floating objects, on the map of the long-term annual mean sea surface temperature (in °C).

2.2. Data source

The Stock expected Recruitment and spawners biomass data were provided by the IATTC. These estimates come from the *Stock Synthesis model* 3.23b (Methot, 2011), a statistical age-structure model for stock assessment. SOI data come from the National Center for Atmospheric Research (NCAR), available on http://www.cgd.ucar.edu/cas/catalog/climind/soi.html. PDO index data was obtained from the Join Institute for the Study of the Atmosphere and Ocean (JISAO) from its website: http://jisao.washington.edu/pdo/. Wind data were obtained from a monthly reanalysis product with a 2.5° spatial resolution, held by the Physical Sciences Division (PSD) of the National Oceanic and Atmospheric Administration (NOAA) available through its website: http://www.esrl.noaa.gov/psd/. Currents velocity data was taken from a monthly reanalysis product with an 1/3° resolution in latitude and 1° in longitude, provided by Global Ocean Data Assimilation System (GODAS), also available in the previous website. Sea surface temperature monthly data were obtained with a 2° resolution, derived from an analysis of the Comprehensive Ocean-Atmosphere Data Set (ICOADS) provided by NOAA in its website: http://icoads.noaa.gov/products.html.

Recruitment forecast was carried out by using oceanographic parameters (sea surface temperature, wind magnitude, magnitude of the sea water currents) from four different climate-change scenarios as inputs. These scenarios are numerical-modeling future simulations forced by specific emissions of greenhouse gases, referred as representative concentration pathway (RCP): RCP2.6, RCP4.5, RCP6.0, and RCP8.5 (IPCC, 2013). RCP labels then provide rough estimates of radiative forcing that will be achieved through the year 2100. Numerous General Circulation Model (GCM) outputs, available in the Coupled Model Intercomparison Project Phase 5 (CMIP5) website (http://cmip-pcmdi.llnl.gov/cmip5/), provide atmospheric and oceanic variables which are used for the assessment of the climate change impacts carried out by the Intergovernmental Panel on Climate Change (IPCC; http://www.ipcc.ch/). Herein the GCM we chose is the Geophysical Fluid Dynamics Laboratory (GFDL) Coupled Physical Model CM3, available in the GFDL website: http://data1.gfdl.noaa.gov. We decided to use this GCM (GFDL CM3) because since its previous versions (GFDL 2CM.0 and GFDL CM2.1) it has proved to reproduce the oceanic dynamics associated with the El Niño (Lin, 2007) and the PDO (Overland and Wang, 2007).

2.3. NARX configuration

NARX models were configured using *nnstart* tool in Matlab R2012b. A neural network consists of an input layer composed of processing elements (neurons) that receive input signals at the start of the network, one hidden layer, and an output layer with one output neuron. NARX is a recurrent dynamic network, with feedback connections, where both the input layer and the hidden layer receive inputs signals (x) at time t but also at time t-n, where n is the number of lags. In the same way receive as input signals, the output (y) calculated at time t-n (Beale et al., 2014).

The network configuration consists of determining the number of neurons in the hidden layer and the number of lags in the input signals to maximize modeling ability. There is no default criterion to determine the best structure, therefore we evaluated the performance of the network with different structures, after a training phase. In this training phase a set of input/target pairs are used for the back propagation algorithm training (Rumelhart et al., 1986). The best models were selected according to the criteria in Beale et al. (2014). In addition, we calculated the correlation coefficient between the estimated recruitment and the expected recruitment (estimated recruitment provided by IATTC) for the entire data set (General CC), and the correlation coefficient between the recruitment calculated for those input signals not used during the training (test data) and the recruitment expected for that dataset (Test CC).