Ecological Modelling 348 (2017) 125–134

Contents lists available at ScienceDirect

Ecological Modelling

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

Bayesian spatial predictive models for data-poor fisheries

Marie-Christine Rufener[a,b,*], Paul Gerhard Kinas[a], Marcelo Francisco Nóbrega[b], Jorge Eduardo Lins Oliveira[b]

[a] Laboratório de Estatística Ambiental, Instituto de Matemática, Estatística e Física, Universidade Federal do Rio Grande, Avenida Itlália km 8, Carreiros, CEP: 96201-900, Rio Grande, Rio Grande do Sul, Brazil

[b] Laboratório de Biologia Pesqueira, Departamento de Oceanografia e Limnologia, Universidade Federal do Rio Grande do Norte, Avenida Costeira s/n, Mãe Luiza, CEP: 59014-002, Natal, Brazil

ARTICLE INFO

Article history:
Received 1 March 2016
Received in revised form 6 January 2017
Accepted 20 January 2017
Available online 31 January 2017

Keywords:
Bayesian geostatistical models
Integrated Nested Laplace Approximations
Stochastic Partial Differential Equations
Essential fish habitats
Fisheries ecology

ABSTRACT

Understanding the spatial distribution and identifying environmental variables that drive endangered fish species abundance are key factors to implement sustainable fishery management strategies. In the present study we proposed hierarchical Bayesian spatial models to quantify and map sensitive habitats for juveniles, adults and overall abundance of the vulnerable lane snapper (*Lutjanus synagris*) present in the northeastern Brazil. Data were collected by fishery-unbiased gillnet surveys, and fitted through the Integrated Nested Laplace Approximations (INLA) and the Stochastic Partial Differential Equations (SPDE) tools, both implemented in the R environment by the R-INLA library (http://www.r-inla.org). Our results confirmed that the abundance of juveniles and adults of *L. synagris* are spatially correlated, have patchy distributions along the Rio Grande do Norte coast, and are mainly affected by environmental predictors such as distance to coast, chlorophyll-*a* concentration, bathymetry and sea surface temperature. By means of our results we intended to consolidate a recently introduced Bayesian geostatistical model into fisheries science, highlighting its potential for establishing more reliable measures for the conservation and management of vulnerable fish species even when data are sparse.

1. Introduction

Fisheries action as well as environmental fluctuations may induce many changes in fish stocks, which are commonly related to their abundance, size and spatial distribution (Haddon, 2001; King, 2007). In this sense, detecting the main environmental factors driving natural abundance fluctuation and understanding how these associations vary over space and time are key concepts in ecology and in fisheries sciences (Ross et al., 2012).

The links between fish stock dynamics and its surrounding environment are therefore of fundamental importance in order to improve sustainable fisheries' management and conservation strategies (Babcok et al., 2005; Valavanis et al., 2008). The way in which we may access the species-environment relationships are commonly treated in the specialized literature as Species Distribution Models (SDMs), and are frequently analyzed in the context of statistical tools to evaluate a species distribution with respect to environmental variables (Franklin, 2010). The main goal of SDMs relies on the prediction, identification and understanding of a species spatial distribution, and may be considered as those habitats where it fulfills some of its biological process, such as reproduction, spawning and feeding. When dealing with marine species, and particularly with fishes, the output provided by the SDMs designate the term *Essential Fish Habitat* (EFH), which constitute areas that promote the fishes most favorable habitats for spawning, feeding, or growth to maturity.

Over the past three decades a huge effort has been spent in the development of powerful statistical models to explore more realistic scenarios regarding the species-environmental relationships. Artificial Neural Networks (ANN, e.g., SPECIES), Maximum Entropy (ME, e.g., MAXENT), Climatic Envelops (CE, e.g., BIOCLIM), Classification and Regression Trees (CART, e.g., BIOMOD) and regression models such as Generalized Linear and Additive (Mixed) Models (GLM/GLMM/GAM/GAMM) are among the many methodological tools that have been proposed for modelling the species distribution (Franklin, 2010; Guisan and Thuiller, 2005).

* Corresponding author at: Laboratório de Estatística Ambiental, Instituto de Matemática, Estatística e Física, Universidade Federal do Rio Grande, Avenida Itlália km 8, Carreiros, CEP: 96201-900, Rio Grande, Rio Grande do Sul, Brazil.

E-mail addresses: machris55@hotmail.com (M.-C. Rufener), paulkinas@furg.br (P.G. Kinas), marnobrega@hotmail.com (M.F. Nóbrega), jorgelins@ufrnet.br (J.E. Lins Oliveira).

http://dx.doi.org/10.1016/j.ecolmodel.2017.01.022
0304-3800/

Although these applications are considered robust, they usually address only explanatory models that simply aim to verify the relationship between the response variable (e.g. presence or abundance of the species) and some environmental predictors (e.g. bathymetry and chlorophyll-*a* concentration in marine context), without considering explicitly the spatial component (Ciannelli et al., 2008). This may result in a poor characterization of the species response to environmental factors by underestimating the degree of uncertainty in its predictors (Latimer et al., 2006).

Considering that the marine environment is an extremely heterogeneous domain, and that a given marine species has biological and ecological constraints, it is commonly noted that biological resources like most fish species present a gregarious distribution. Hence, ignoring spatial autocorrelation violates the main assumption of classical inference, which assumes that the data are independent and identically distributed, and thus can lead to biased estimates. Spatial correlation should therefore be considered since species are generally subject to similar environmental factors (Muñoz et al., 2013). Such spatial models allow constructing powerful predictive models that not only provide the estimation of processes that influence species distribution, but also promote the possibility of predicting their occurrence in unsampled areas (Chakraborty et al., 2010).

Additionally, it is advantageous to incorporate Bayesian inference in predictive models, given that it is possible to integrate all types of uncertainties using exclusively the probability as its metric. Combining the uncertainty in the data (expressed by the likelihood) with extra-data information (expressed by prior distributions), posterior probability distributions are built for all unknown quantities of interest using Bayes' theorem (Kinas and Andrade, 2010).

Hierarchical Bayesian Models (HBMs) are very suitable for such situations, as they allow to introduce sequentially the uncertainties associated with the entire fishery phenomenon, as well as a spatial random effect in the form of a Gaussian random field (GRF) (Cosandey-Godin et al., 2015). Traditionally HBMs relied on simulation techniques such as Markov Chain Monte Carlo (MCMC). However, with increased model complexity the computational time required to approximate the posterior distributions became unfeasible. To sidestep this limitation, Rue et al. (2009) proposed an alternative numerical computation to obtain posterior distributions, called Integrated Nested Laplace Approximations (INLA), and which is currently implemented in the R environment by the *R-INLA* package (http://www.r-inla.org). Rather than using stochastic simulation techniques, INLA uses numerical approximations by means of the Laplace operator, which revealed to be much faster, flexible and accurate than MCMC whenever applicable.

Thanks to Illian et al. (2013) and Muñoz et al. (2013), spatio-temporal models were introduced to the ecological community in order to fit point process and point-referenced data through the INLA approach. In regard to marine ecology and fisheries research, more studies have slowly emerged since then using exclusively INLA for spatial and temporal purposes (Cosandey-Godin et al., 2015; Muñoz et al., 2013; Paradinas et al., 2015; Pennino et al., 2013, 2014; Quiroz et al., 2015; Roos et al., 2015; Ward et al., 2015; Bakka et al., 2016; Damasio et al., 2016; Paradinas et al., 2016; Pennino et al., 2016).

However, most of these studies relied on "data-rich" fisheries, where certainly any quantitative method would have had good performance. On the other hand, in developing countries such as Brazil, fisheries tend to be poorly documented and inadequately managed due to lack of research funding for monitoring and analysis (Honey et al., 2010). Conventional analytical fisheries stock assessment tools that demand big data sets may not be applicable in "data-poor" situations like these. Therefore, flexible and reliable statistical tools with good performance albeit limited information are paramount (Bentley, 2014).

In order to expand the use of such tools in data limited fisheries, we will demonstrate the flexibility and usefulness of the Bayesian modelling approach for some important fisheries issues, such as delimiting fish stocks into age groupings and evaluating the spatial distribution of these age groupings. Specifically, our main objectives are: (i) predict abundance and age groupings of the target lane snapper (*Lutjanus synagris*) along a fraction of the northeastern coast of Brazil; and (ii) identify environmental drivers that affect lane snapper's abundance and so, provide important insights of its spatial distribution.

This paper is organized as follows: firstly we describe the importance of our study subject, how the main dataset was obtained and how we achieve age groupings using Bayesian logistic regression. Thereafter, we apply hierarchical Bayesian spatial models into EFH's, and discuss the entire modelling procedures used in this study. Finally, we describe and discuss our results, outlining opportunities for future spatial fisheries management.

2. Material & methods

2.1. Case study

As a case study, we modelled the spatial occurrence of *Lutjanus synagris* (Linnaeus, 1758), popularly known as lane snapper and which is considered one of the most important fishing resource caught within the Lutjanidae family (Luckhurst et al., 2000). Lane snappers inhabit a variety of habitats from coastal waters to depths up to 400 m, often occurring in coral reefs and vegetation on sandy bottoms (Carpenter, 2002), and are widely distributed throughout the tropical region of the western Atlantic, from North Carolina to southeastern Brazil, including the Gulf of Mexico and the Caribbean Sea (McEachran and Fechhelm, 2005).

Given its high commercial and recreational value, this species is one of the mainstays of artisanal fisheries not only in Venezuela and the Caribbean Sea, but also in northeastern Brazil (Gómez et al., 2001; Luckhurst et al., 2000; Rezende et al., 2003). According to Lessa et al. (2009), catches of lane snapper in northeastern Brazil have been recorded since the late 1970′s where it is suffering strong fishing pressure which, despite their high abundance, is leading to its decline over the past few decades.

In general, most information about *L. synagris* relies on its biology, whereas its population dynamics and spatial distribution remain poorly understood (Cavalcante et al., 2012; Freitas et al., 2011, 2014). Therefore, knowing their spatial distribution and identifying environmental variables that drive their abundance are key factors to implement sustainable fishery management strategies.

2.2. Study area and data survey

Situated in the northeast of Brazil, the Rio Grande do Norte state (RN) is located in an important coastline transition zone which abruptly changes its direction from south-north to east-west. Between July 2012 and June 2014, about three experimental fisheries were monthly conducted by fishing vessels of the artisanal fleets which operate with bottom gillnets along the RN coast. Throughout this period, 126 fishing events were reported whose depths ranged from 5 to 50 m (Fig. 1). Biological sampling was recorded along with extra information for each fishery including geolocation (latitude and longitude), bathymetry (m), sea surface temperature (°C), distance to coast (km), month (from January to December), gill net length (m) and height (m), as well as its soaking time (h).

It is noteworthy that data collected from artisanal fleets would usually characterize a clear example of preferential sampling, since fleets are commercially driven to catch target species hotspots.