Fisheries Research 100 (2009) 200–209

Contents lists available at ScienceDirect

Fisheries Research

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

Comparing three indices of catch per unit effort using Bayesian geostatistics

Júlio César Pereira[a,b,*], Roseli Aparecida Leandro[a], Miguel Petrere Jr.[c], Tom Nishida[d]

[a] UFSCAR, Campus Sorocaba, Rodovia João Leme dos Santos, km 110, CEP 18052-780, Sorocaba (SP), Brazil
[b] ESALQ/USP, Departamento de Ciências Exatas, CP 9, 13418-900 Piracicaba (SP), Brazil
[c] UNESP, Departamento de Ecologia, CP 199, 13506-900 Rio Claro (SP), Brazil
[d] National Research Institute of Far Seas Fisheries, Fisheries Research Agency, 5-7-1, Orido, Shimizu-Ward, Shizuoka, 424-8633, Japan

ARTICLE INFO

Article history:
Received 10 February 2009
Received in revised form 26 July 2009
Accepted 28 July 2009

Keywords:
Simulation
Geostatistics
CPUE
Estimation
Linear coregionalization

ABSTRACT

In assessing a fish stock, indices based on catch per unit effort (CPUE) are frequently used. Estimates of three indices of catch per unit effort were compared here ($CPUE_1$, $CPUE_2$ and $CPUE_3$), considering the fitting of two models: (i) a bivariate geostatistical model for catch and effort; (ii) a bivariate model where catch and effort were considered spatially independent. For comparing the estimates of the three indices after the fitting of the two models, catch and effort data were simulated in different scenarios. The simulation study showed that, in general, the estimates of $CPUE_1$ expressed by the ratio of the means of catch and effort, present better results for different scenarios and that the estimates from (i) are better than (ii), mainly when there is a correlation between catch and effort and an additional spatial correlation.

1. Introduction

To evaluate a fish stock, data of catch and effort resulting from commercial fishing are usually used in heuristic relationships. Based on catch and effort data, indices of relative abundance are calculated in order to supply information about the stock. In a given inhabited area by a given stock, if the density (or concentration) of fish (biomass/volume) is constant for the whole area the CPUE is proportional to stock abundance (strict proportionality) (Clark, 1985). In some cases this relationship might not be linear. The examination of this relationship is not the main theme of this paper. However, in the light of this paper for any supposed model the relationship CPUE × abundance, it is necessary to estimate the CPUE in order to evaluate the abundance.

Detailed records, with information on the geographic coordinates where fishing occurred allow a spatial analysis of fishing. Normally a point of reference is given for each quadrat (sub-regions delimited by parallels and meridians) where fishing occurred (ICCAT, 2007).

* Corresponding author. UFSCAR, Campus Sorocaba, Rodovia João Leme dos Santos, km 110, CEP 18052-780, Sorocaba (SP), Brazil. Tel.: +55 15 3229 5971; fax: +55 15 3229 5940.
E-mail addresses: julio-pereira@ufscar.br (J.C. Pereira), raleandr@esalq.usp.br (R.A. Leandro), mpetrere@rc.unesp.br (M. Petrere Jr.), tnishida@affrc.go.jp (T. Nishida).

0165-7836/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.fishres.2009.07.010

Based on catch × effort data, three CPUE indices may be defined for a whole area:

$$CPUE_1 = \frac{1}{n}\sum_{i=1}^{n}\frac{C_i}{E_i}; \quad \text{mean of the ratios catch by effort;}$$

$$CPUE_2 = \frac{\sum_{i=1}^{n} C_i}{\sum_{i=1}^{n} E_i}; \quad \text{ratio of total catch by total effort;}$$

$$CPUE_3 = \frac{\sum_{i=1}^{n} C_i E_i}{\sum_{i=1}^{n} E_i^2}; \quad \text{ratio estimator}$$

as proposed by Snedecor and Cochran (1967),where C_i, $i = 1, 2, \ldots, n$, represents the catch in the ith quadrat and E_i the respective effort, n is the total number of quadrats superposed as an artificial grid in the fishing area.

The three indices may be described as C_i/E_i averages distinguished by the weighing criteria, that is $\sum_{i=1}^{n}(C_i/E_i)w_i$. In $CPUE_1$ the weighing factor is $w_i = 1/n$; in $CPUE_2$ it is $w_i = E_i/(\sum_{j=1}^{n} E_j)$ and in $CPUE_3$ we have $w_i = E_i^2/(\sum_{j=1}^{n} E_j^2)$.

Whenever C (Capture) be proportional to E (Effort), the regression line between them statistically goes through the origin, and can be fitted by the simple model $C_i = \beta E_i + \varepsilon_i$. $CPUE_3$, together with $CPUE_1$ and $CPUE_2$, are all unbiased estimates of the population ratio β in normally distributed populations. The choice among the three is a matter of precision: the most precise among the three is $CPUE_3$, $CPUE_2$ and $CPUE_1$ if the variance of ε (error term) is constant, proportional to E or to E^2, respectively. If the variance of ε increases moderately with E, $CPUE_2$ is still expected to perform well (Petrere Jr. et al., 2008). Data on catch and effort are usually not available for all quadrats to establish these indices, in other words, not all quadrats are fished. In this situation, a possibility of estimation is simply to use the observed data. This might not be the best option. According to Walters (2003), if a spatial correlation structure between catch rates is found, spatial statistics can be used for extrapolation to grid squares where no fishing took place, before calculating any abundance index. Moreover, this author claims that covariates such as water temperature can be used to provide estimates of the spatial trend. The oceanic surface temperature is an important covariate, as several studies show that it is correlated with CPUE (Dow et al., 1975; Dow, 1980; Evans et al., 1995; Fonteneau, 1995; Lima et al., 2000; Goodyear, 2003).

In this work a model was utilized for the variables catch and effort, whose covariance structure is described by a model of linear coregionalization (Gelfand et al., 2004), from now on this model will be called special bivariate model (SBM). This was one of the models investigated in the spatial analysis of catch and effort data. The choice was based on observations of Walters (2003) on the use of spatial statistical techniques and covariates and also taking into consideration that in practice the observed data are bivariate (catch and effort) (ICCAT, 2007), besides other characteristics that are inherent to certain fishing data sets, e.g., the existence of spatial correlation (Swain and Wade, 2004; Walters, 2003), relationship between catch and effort and that the effort can be considered random since it may depend, for example, on the commercial value of the target species, time of year, climate conditions, sea surface temperature, perception of fishermen of a fish stock (observation or non-observation), information from other fishermen (Sanchéz et al., 2004; Walters and Martell, 2004; Hilborn and Walters, 1992). Besides describing the structure of covariance between variables catch and effort and the spatial correlation, the model SBM shapes the structure of the cross covariance, i.e., the covariance between the effort at any location s_i and catch at location s_j, and vice versa. In other words, the observations (E_i, C_i) are treated as a sample of a bivariate spatial process. A fitting of the proposed model makes the extrapolation of catch and effort possible to quadrats that were not observed. Besides the linear coregionalization model, it was fitted a bivariate model was fitted in which catch and effort are considered spatially independent, from now on called bivariate model without spatial component (BMWSC).

Since the CPUE indices are used in the assessment of fish stocks it is important to assess the performance of the three indices ($CPUE_1$, $CPUE_2$ or $CPUE_3$) in different scenarios. Above all, it is essential to use methods that estimate each index accurately. Petrere Jr. et al. (2008) conducted a simulation study to compare $CPUE_1$, $CPUE_2$ and $CPUE_3$. However, these indices were not studied in the presence of spatial correlation.

We suggest that the indices $CPUE_1$, $CPUE_2$ and $CPUE_3$ be estimated as follows: by extrapolation of catch and effort to unfished quadrats; for this purpose one of the above models was adjusted, according to the Bayesian approach; after the extrapolation the indices $CPUE_1$, $CPUE_2$ and $CPUE_3$ were estimated based on the data set consisting of observed and predicted values.

The objectives of this study were:

(i) to compare the statistical behavior of the estimates of three indices ($CPUE_1$, $CPUE_2$ and $CPUE_3$), estimated when using the SBM model, based on simulated data sets of different scenarios;
(ii) to compare the estimates calculated through the interpolation of catch and effort in those not observed quadrats using the SBM and BMWSC models.

2. Materials and methods

The geostatistical techniques used here assume that the variables to be modelled follow normal distribution. The distributions of variables catch and effort are generally asymmetrical and in many cases the logarithmic transformation is sufficient to correct the lack of normality (Abuabara, 1996; Sanchéz et al., 2004). To perform the simulation study it was therefore assumed that catch and effort follow normal distribution in the logarithmic scale. In the following we describe the utilized models in the simulation studies and the inference procedures.

2.1. Spacial bivariate model (SBM)

When using Gelfand et al. (2004) model, the catch and effort observations across a region are treated as a sample from a bivariate spatial process. The proposed model, easily interpretable and computer processable, creates a structure of flexible covariance, where the ranges (i.e., the distance beyond which there is practically no spatial correlation between data points) associated with the variate are not necessarily the same. The authors show that there is an equivalence, based on reparametrization of the conditional specification given by Eq. (1) and the unconditional specification of the model.

Clearly there is a cause/effect between effort /catch. So when conditioning the model effort comes first, then capture. So, the logarithm of the fishing effort (Y_1) is modelled first and then the logarithm of catch, given by the logarithm of effort:

$$\begin{aligned} Y_1(s) &= \beta_{01} + \beta_{11} temp(s) + \sigma_1 w_1(s) \\ Y_2(s)\,\big|\,Y_1(s) &= \beta_{02} + \beta_{12} temp(s) + \alpha Y_1(s) + \sigma_2 w_2(s) + \tau_2 u_2(s), \end{aligned} \tag{1}$$

where $temp(s)$ represents the temperature at location s, $w_1(s)$ and $w_2(s)$ are Gaussian spatial processes with mean zero and variance 1, independent, but not identically distributed, and $u_2(s)$ has distribution $N(0,1)$. The term $\beta_{01} + \beta_{11} temp(s)$ in Eq. (1) determines the expected value of the logarithm of effort for a location s and $\beta_{02} + \beta_{12} temp(s) + \alpha(\beta_{01} + \beta_{11} temp(s))$, determines the expected value of the logarithm of capture. $\sigma_1 w_1(s)$ and $\alpha\sigma_1 w_1(s) + \sigma_2 w_2(s)$ accounts for spatial correlation in these quantities (effort and catch, respectively). The term $\tau_2 u_2(s)$ is responsible for microescale variation (nugget effect).The adopted correlation function was the exponential $\rho(d)$=exp$(-\phi d)$, with parameter ϕ_1 for Y_1 and ϕ_2 for Y_2, where d is the distance between any two points s, s'. The parameter ϕ, expresses how quickly the correlation drops to zero.

The use of conditional specification; the model for $Y_1(s)$ must not have a white noise component to ensure equivalence. The model written in its conditional form may have a pure nugget effect in the first equation. This occurs when spatial correlation is practically null, albeit in its presence, that is, when the correlation function parameter of the first equation (effort equation) is positive, we have for this equation a pure spatial effect. The remaining variation is inherited from the second equation, as in this equation $Y_2(s)$ is written in function of $Y_1(s)$.

Considering the model in its conditional form we have a large computational advantage, since instead of dealing with one covariance matrix $2n \times 2n$, two covariance matrices $n \times n$ are used (Gelfand et al., 2004).

Given any location s, it may be shown that the correlation between $Y_1(s)$ (logarithm of effort) and $Y_2(s)$ (logarithm of catch) is given by

$$\rho_{Y_1,Y_2} = \frac{\alpha\sigma_1^2}{\sqrt{\sigma_1^2(\alpha^2\sigma_1^2 + \sigma_2^2)}} \tag{2}$$

in which $\alpha\sigma_1^2$ is the covariance between $Y_1(s)$ and $Y_2(s)$, σ_1^2 is the variance of $Y_1(s)$ and $\alpha^2\sigma_1^2 + \sigma_2^2$ is the variance of $Y_2(s)$.