Published in final edited form as:

Fish Res. 2015 August ; 168: 20–32. doi:10.1016/j.fishres.2015.03.013.

Evaluation of geostatistical estimators and their applicability to characterise the spatial patterns of recreational fishing catch rates

Eric N. Aidoo^{a,*}, Ute Mueller^a, Pierre Goovaerts^b, and Glenn A. Hyndes^c

^aSchool of Engineering, Edith Cowan University, 270 Joondalup Drive, Joondalup, WA 6027, Australia

^bBioMedware Inc., 121 W. Washington St., 4th floor-TBC, Ann Arbor, MI 48104, USA

^cCentre for Marine Ecosystems Research, School of Natural Sciences, Edith Cowan University, 270 Joondalup Drive, Joondalup, WA 6027, Australia

Abstract

Western Australians are heavily engaged in recreational fishing activities with a participation rate of approximately 30%. An accurate estimation of the spatial distribution of recreational catch per unit effort (catch rates) is an integral component for monitoring fish population changes and to develop strategies for ecosystem-based marine management. Geostatistical techniques such as kriging can provide useful tools for characterising the spatial distributions of recreational catch rates. However, most recreational fishery data are highly skewed, zero-inflated and when expressed as ratios are impacted by the small number problem which can influence the estimates obtained from the traditional kriging. The applicability of ordinary, indicator and Poisson kriging to recreational catch rate data was evaluated for three aquatic species with different behaviours and distribution patterns. The prediction performance of each estimator was assessed based on cross-validation. For all three species, the accuracy plot of the indicator kriging (IK) showed a better agreement between expected and empirical proportions of catch rate data falling within probability intervals of increasing size, as measured by the goodness statistic. Also, indicator kriging was found to be better in predicting the latent catch rate for the three species compared to ordinary and Poisson kriging. For each species, the spatial maps from the three estimators displayed similar patterns but Poisson kriging produced smoother spatial distributions. We show that the IK estimator may be preferable for the spatial modelling of catch rate data exhibiting these characteristics, and has the best prediction performance regardless of the life history and distribution patterns of those three species.

Keywords

Kriging estimators; Indicator kriging; Poisson kriging; Ordinary kriging; Catch rate estimation

*Corresponding author. Tel.: +61 405692203. eaidoo@our.ecu.edu.au (E.N. Aidoo).

1. Introduction

The application of geostatistical techniques has become a useful tool to describe the spatial distribution of fisheries catches and catch per unit effort (Mueller et al, 2008; Petitgas, 2009; Rivoirard et al, 2000). This tool can improve managers' understanding of trends in populations over space and time (Walters, 2003), which can facilitate specific management decisions, policy implementation and efficient distribution of resources with respect to geographical fishing areas. However, their application to recreational fisheries data is limited. With recreational fishing considered to be one of the major coastal zone leisure activities worldwide (Cowx, 2002), and an increasing rate of recreational fishing across the globe (Cooke and Cowx, 2004), there is a need for continuous spatial distributions of recreational catch rates (catch per unit of fishing effort).

The focus of fisheries management has traditionally been on commercial fisheries (Henry and Lyle, 2003), but the proportion of fish harvested from recreational fishing can be similar or even higher than that from commercial fishing in some areas (Coll et al, 2004). Indeed, in Western Australia, the participation in recreational fishing has changed from 19% in 1989/90 to 30% in 2011/12, indicating a relative growth of about 50% over that period (Department of Fisheries, 2012b; Lindner and McLeod, 1991). Such increases have highlighted the need to develop and implement resource allocations to both commercial and recreational sectors (Crowe et al, 2013). Furthermore, understanding the spatial distribution of recreational fishing catch rates can help identify high priority areas and guide the establishment of marine protected areas for targeted fish species. Thus, accurate estimation of recreational catch rate distributions in space is an integral component for fishery management and conservation (Wise et al, 2012).

The traditional way for obtaining continuous surfaces of estimates in a geostatistical framework is to use an interpolation technique known as ordinary kriging (Goovaerts, 1997). Catch rate at any unsampled location is simply computed as a weighted average of data falling within a search window; kriging weights account for spatial patterns in the data and their respective locations (e.g. closer data typically receive more weight in the estimation) (Goovaerts, 1997). However, most data sets collected in recreational fisheries include many zero values mixed with a few extreme values, leading to zero inflated, positively skewed distributions (O'Neill and Faddy, 2003; Taylor et al., 2011). In addition, when expressed as ratios, such data (i.e. catch rate) are impacted by the "small number problem" (Monestiez et al, 2006), that is, rates computed from low fishing effort are less reliable. These properties of the data influence the estimation and modelling of the spatial autocorrelation in the measurements (Kerry et al., 2010; Monestiez et al., 2006), and may lead to incomplete and misleading spatial patterns when the local mean and variance change (Rossi et al., 1992). Furthermore, extreme values in such data sets can strongly affect the characterisation of the spatial pattern and the prediction (Goovaerts, 2009). Hence, the use of the traditional kriging estimator may not be appropriate for such data.

Techniques are available to deal with outliers and skewness. One option is to transform each catch rate into an indicator that takes a value of zero or one depending on whether it exceeds or not a threshold specified by the user. These indicators are then analysed and interpolated

using the so-called indicator kriging (Rossi et al, 1992). By treating the data as outcomes of a Poisson process, Monestiez et al. (2006) recently developed a modified kriging estimator known as Poisson kriging where catch rate data based on high observation effort receive greater weight in the estimation and modelling of the spatial autocorrelation, as well as during the interpolation (Monestiez et al, 2006).

To employ geostatistical techniques in spatial catch rate estimation for recreational fisheries, suitable estimators need to be identified for such data. Also, there is a need to understand whether a single estimator can be applied to catch rate data for all species of interest, or whether a different estimator is required for each species with different behaviours (e.g. sedentary versus mobile) and distributional characteristics. In this study, ordinary, indicator and Poisson kriging are evaluated to determine their applicability to characterise the spatial distribution of recreational catch rates of three fishery species (snapper *Pagrus auratus*, baldchin groper *Choerodon rubescens* and blue swimmer crab *Portunus armatus*) with different behaviours and distribution patterns.

2. Material and methods

2.1. Study area and data

The study was limited to the West Coast bioregion of Australia, where about 85% of the recreational fishing in Western Australia takes place (Baharthah, 2008). The data used in this study were obtained from a 12-month survey on recreational fishing by the Western Australian Department of Fisheries (Ryan et al., 2013). The data were collected from March 1st, 2011 to February 29th, 2012 using an off-site phone-diary survey approach involving 2,977 boat-based recreational fishers in the state. Using the recreational fishing boat licence (RFBL) database, a sampling frame was obtained. The survey covered recreational fishers who were at least 5 years old, and held an RFBL for the 12-month period prior to the survey and intended to fish within the period of the survey (Ryan et al., 2013). The data collected from the fishers included for each fishing event: the fishing block visited, the fishing trip duration and the number of fish caught per species. The spatial resolution of the data collected was at the size of the fishing block, which is 10 by 10 nautical miles (NM). The blocks are identified via their centroids. Detailed information about the survey is documented in (Ryan et al, 2013).

The species considered for the analyses are caught by both recreational and commercial fishers and are used as indicator species to represent the status of all the exploited demersal finfish and invertebrate species in the West Coast bioregion of Australia (Department of Fisheries, 2012a; Fletcher and Santoro, 2012). Snapper has a wide distribution across the western Indo-Pacific region (Smallwood et al., 2013). In the study region, it can exhibit spawning aggregations and the juveniles occupy sheltered coastal waters prior to migrating to deeper, more offshore waters (Smallwood et al., 2013). Baldchin groper is endemic to the southern west coast of Australia. Unlike snapper, no spawning aggregations have been reported for this species, and juvenile and adult habitats overlap (Smallwood et al., 2013). The blue swimmer crab is distributed widely throughout the Indo-Pacific region, with adults and juveniles occurring in inshore marine and estuarine waters (Kailola et al., 1993).

For each species and fishing block with centroid u_a the total catch (in number of fish caught) and the fishing trip duration in hours (effort) within the survey year were converted to catch per hour $Z(u_a)$ (catch rate) using the ratio of mean estimator (Jones et al, 1995) to control for heterogeneity in numbers of catches posed by differences in fishing trip duration.

$$z(u_a) = \frac{\sum_{j=1}^{v(u_a)} c_j(u_a) / v(u_a)}{\sum_{j=1}^{v(u_a)} d_j(u_a) / v(u_a)} = \frac{\sum_{j=1}^{v(u_a)} c_j(u_a)}{\sum_{j=1}^{v(u_a)} d_j(u_a)} \quad (1)$$

where $c_j(u_a)$ and $d_j(u_a)$ are the number of fish caught and the boat fishing trip duration (hours) respectively by the j th fisher, and $v(u_a)$ is the total number of boats that visited the fishing block u_a .

2.2. Geostatistical estimation

2.2.1. Semivariogram estimators—A prerequisite for geostatistical interpolation is the definition of a model describing the spatial autocorrelation between the observations. A common tool is the semivariogram which measures the average dissimilarity among the observations at pairs of locations as a function of the separation distance vector h (lag). The semivariogram calculated from the sampled data is called experimental semivariogram. The traditional experimental semi-variogram denoted by $\hat{\gamma}_o(h)$ is estimated according to Matheron's method of moments defined as (Matheron, 1965):

$$\hat{\gamma}_o(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [Z(u_\alpha) - Z(u_\alpha + h)]^2 \quad (2)$$

where $Z(u_a)$ represents the catch rate of a particular species at location u_a , $Z(u_a + h)$ is the catch rate at a location separated from u_a by a distance vector h and $N(h)$ is the total number of pairs of catch rate samples separated by h (Goovaerts, 1997).

In the indicator approach, for each species a series of K thresholds $z_1, z_2, \dots, z_k$ is chosen to discretise the range of variation for the associated catch rate distribution. For each location u_a the catch rate is transformed into a binary vector of K indicators such that the k^{th} indicator represents the probability that the observed measurement does not exceed a specific threshold z_k (Goovaerts, 1997). If $z(u_a)$ is the catch rate of a particular species at location u_a the indicator variables are defined as (Goovaerts, 1997):

$$i(u_\alpha; z_k) = \begin{cases} 1 & \text{if } z(u_\alpha) \leq z_k \\ 0 & \text{otherwise} \end{cases} \quad k=1, \dots, K. \quad (3)$$

This approach allows the spatial continuity of the random variable $z(u)$ to vary between thresholds. The choice of the number and values of the thresholds z_k is to some extent subjective. According to Goovaerts (1997) a minimum of 5 and a maximum of 15 thresholds are appropriate for discretising the local distribution of the data. A simple way to determine the threshold values is to use the percentiles of the corresponding sample histogram (for example, 9 deciles for $K=9$). For skewed distributions more threshold values can be located

within the parts of the distribution that are of greater interest, e.g., the lower or upper tail of the local distribution (Goovaerts, 1997). Because of the high proportion of zeroes in the data sets, the use of regular percentile intervals is not appropriate and needs to be modified to take into account the zero-inflation. The first threshold z_1 was systematically set to a catch rate of 0. Then, an additional number of thresholds between four and nine (i.e. $5 \leq K \leq 10$) was evaluated, accounting for the shape of the individual catch rate distributions and ensuring that the occurrence of order relation deviations was minimised (see Goovaerts, 1997 for a discussion on order relation deviation). The experimental indicator semivariogram $\hat{\gamma}_I(h; z_k)$ is defined analogously to the traditional experimental semivariogram in (2) with $i(\cdot; z_k)$ replacing $z(\cdot)$:

$$\hat{\gamma}_I(h; z_k) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [i(u_\alpha; z_k) - i(u_\alpha + h; z_k)]^2 \quad (4)$$

Thus instead of a single experimental semivariogram serving as the basis for modelling the spatial continuity of the catch rate variable, K experimental indicator semivariograms are used. In the present case the use of a weighted semivariogram might be more appropriate as catch rates might become unreliable for small fishing effort, which is referred to as the “small number problem” (Susan, 1989). The experimental weighted semivariogram estimator $\hat{\gamma}_w(h)$ describing the spatial variability between catch rates recorded at two different locations separated by vector h is defined as (Monestiez et al., 2006):

$$\hat{\gamma}_w(h) = \frac{1}{2 \sum_{\alpha=1}^{N(h)} (d(u_\alpha)d(u_\alpha+h))/(d(u_\alpha)+d(u_\alpha+h))} \times \sum_{\alpha=1}^{N(h)} \left(\frac{d(u_\alpha)d(u_\alpha+h)}{d(u_\alpha)+d(u_\alpha+h)} [z(u_\alpha) - z(u_\alpha+h)]^2 - m^* \right) \quad (5)$$

where $N(h)$ is the number of pairs of locations separated by a vector h , $d(u_\alpha)$ and $d(u_\alpha + h)$ denote the total fishing trip durations at locations u_α and $u_\alpha + h$ respectively and m^* is the estimated mean catch rate. The weight $(d(u_\alpha)d(u_\alpha+h))/(d(u_\alpha)+d(u_\alpha+h))$ in Eq. (5) is introduced to homogenise the variance of differences terms $[z(u_\alpha) - z(u_\alpha + h)]$ and to attenuate the impact of noisy catch rates calculated from low fishing effort. Thus, pairs of catch rates corresponding to higher efforts receive more weight during the computation of the experimental semivariogram.

In all the semivariogram estimators, the separation vector h has both magnitude $|h|$ and direction. If the estimator depends on both the magnitude and direction; the spatial variability is referred to as anisotropic (i.e. direction-dependent) and it is called isotropic if it depends only on the magnitude of the separation vector.

Theoretical semivariogram model(s) are fitted to the experimental semivariograms to describe the spatial structure of the catch rate data for all separation distances and directions, including the ones that were not sampled. Possible semivariogram models include the nugget effect, spherical, exponential and Gaussian models (Goovaerts, 1997). In this paper,

linear combinations of the nugget effect and spherical models were used. The nugget effect model is defined as:

$$\gamma(h) = \begin{cases} 0 & \text{if } |h| = 0 \\ Q_0 & \text{if } |h| > 0 \end{cases} \quad (6)$$

where $Q_0 > 0$ is the sill of the model, usually called nugget effect. The nugget effect describes variability caused by measurement errors and occurring at distances smaller than the shortest sampling interval and is commonly modelled as isotropic.

The spherical model is defined as:

$$\gamma(|h|) = Q \begin{cases} 0 & \text{if } |h| = 0 \\ \left[\frac{3}{2} \left(\frac{|h|}{R} \right) - \frac{1}{2} \left(\frac{|h|}{R} \right)^3 \right] & \text{if } 0 < |h| \leq R \\ 1 & \text{if } |h| > R \end{cases} \quad (7)$$

where R is the range, Q is the sill and $|h|$ represents the separation distance. The range is the distance of spatial influence or the distance beyond which there exists no spatial autocorrelation between observations. If a combination of models is used, then the total sill is defined to be the sum of the sills of the individual models, including the nugget effect. The total sill represents the semivariance value at a very large distance $\gamma(\infty)$ and it also represents the expected variance of the random function (Goovaerts, 1997; Isaaks and Srivastava, 1989). The degree of spatial autocorrelation can be quantified using the ratio of the nugget effect to the total sill, called nugget to sill ratio (NSR). The degree of spatial autocorrelation is strong when the NSR is less than 0.25; it is moderate when the ratio ranges between 0.25 and 0.75; and a value greater than 0.75 indicates weak spatial autocorrelation (Cambardella et al, 1994). To limit complexity of the models, for ordinary and Poisson kriging, the semivariogram model was set to be a combination of a nugget effect and at most two spherical models, requiring the estimation of at most five parameters. The use of a spherical model is common in the geostatistical analysis of biological populations because it depicts variability increasing linearly with distance until it reaches a plateau, which corresponds with spatial variation that is often observed in nature (Kleisner et al, 2010). For the indicator semivariogram models combinations of a nugget effect and a spherical model were allowed, requiring the independent estimation of 3 parameters per threshold. The estimation of model parameters was performed using an automatic fitting procedure based on least-squares regression (Chiles and Delfiner, 2012; Goovaerts, 2005, 2009). Among the plausible alternative fitted models, the selected model was the one that minimises the sum of square differences between the experimental and model curve (Goovaerts, 2005, 2009). The semivariogram models corresponding to the estimators in Eqs. (3)–(5) are denoted by γ_o , $\gamma_I(\cdot; z_k)$ and γ_w respectively.

2.2.2. Kriging estimators—Following the determination of a semivariogram model for the catch rate data, geostatistical estimation is carried out. Three interpolation techniques corresponding to the semivariogram estimators introduced in Eqs. (3)–(5) were compared: ordinary, indicator and Poisson kriging.

The ordinary kriging (OK) estimator for computing the catch rate at location u is defined as¹:

$$z_{OK}^*(u) = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{OK}(u) z(u_{\alpha}), \quad \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{OK}(u) = 1 \quad (8)$$

where $\lambda_{\alpha}^{OK}(u)$ is the weight assigned to the observed value $z(u_{\alpha})$ and $n(u)$ is the number of closest observations used in the estimation (Goovaerts, 1997). To satisfy the unbiasedness constraint, the ordinary kriging weights must sum to one. The calculation of the optimal weights that minimise the estimation error variance is a constrained optimisation problem and can be solved by the method of Lagrange multipliers. The optimal weights $\lambda_{\alpha}^{OK}(u)$ are obtained as the solution of a system of linear equations with $[n(u) + 1]$ unknowns (Goovaerts, 1997):

$$\begin{cases} \sum_{\beta=1}^{n(u)} \lambda_{\beta}^{OK}(u) \gamma_0(u_{\alpha} - u_{\beta}) - \mu_{OK}(u) = \gamma_0(u_{\alpha} - u), & \alpha=1, 2, \dots, n(u) \\ \sum_{\beta=1}^{n(u)} \lambda_{\beta}^{OK}(u) = 1 \end{cases} \quad (9)$$

and the ordinary kriging error variance is obtained by the expression:

$$\sigma_{OK}^2(u) = E[z^*(u) - z(u)]^2 = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{OK}(u) \gamma_0(u_{\alpha} - u) - \mu_{OK}(u) \quad (10)$$

where $\gamma_0(u_{\alpha} - u_{\beta})$ is the semivariance between observations at locations u_{α} and u_{β} , $\gamma_0(u_{\alpha} - u)$ is the semivariance between the sampling location u_{α} and the unsampled location u being estimated, and $\mu_{OK}(u)$ is the Lagrange multiplier.

In the indicator approach, ordinary kriging is applied to each indicator variable. Thus indicator kriging does not provide a direct estimate of the catch rate but rather a set of K probability estimates (Goovaerts, 1997):

$$i^*(u; z_k) = \text{Prob} \{z(u) \leq z_k | (n)\} \quad k=1, \dots, K \quad (11)$$

The probability estimates allow the construction of a conditional cumulative probability distribution function (ccdf) $F_{IK}(u; z_k | (n))$ for the catch rate variable at each node of the estimation grid. Similar to ordinary kriging of catch rates (Eq. (8)) each probability is estimated as:

$$i^*(u; z_k) = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{IK}(u; z_k) i(u; z_k), \quad \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{IK}(u; z_k) = 1 \quad (12)$$

¹The subscript/superscript OK, IK and PK indicate that the estimator corresponds to ordinary kriging, indicator kriging and Poisson kriging respectively.

where $\lambda_{\alpha}^{IK}(u; z_k)$ is the weight assigned to the indicator value $i(u; z_k)$, $n(u)$ is the number of neighbouring indicator values to be used in the estimation. The optimal weights $\lambda_{\alpha}^{IK}(u; z_k)$ are obtained as the solution of a system of linear equations with $[n(u) + 1]$ unknowns (Goovaerts, 1997):

$$\begin{cases} \sum_{\beta=1}^{n(u)} \lambda_{\beta}^{IK}(u; z_k) \gamma_I(u_{\alpha} - u_{\beta}; z_k) - \mu(u; z_k) = \gamma_I(u_{\alpha} - u; z_k), & \alpha=1, 2, \dots, n(u) \\ \sum_{\beta=1}^{n(u)} \lambda_{\beta}^{IK}(u; z_k) = 1 \end{cases} \quad (13)$$

where $\gamma_I(u_{\alpha} - u_{\beta}; z_k)$ is the semivariance between indicator variables at location u_{α} and u_{β} , $\gamma_I(u_{\alpha} - u; z_k)$ is the semivariance between the sampling location u_{α} and the unsampled location u being estimated, and $\mu(u; z_k)$ is the Lagrange multiplier. Catch rate estimates were derived from the indicator kriging results following the approach described in (Goovaerts, 1997): the mean of the ccdf (E-type estimate) is used as the estimator of the catch rate while the variance of the ccdf is used as a measure of uncertainty. The expected value of the ccdf $z_{IK}^*(u)$ is approximated as (Goovaerts, 1997):

$$z_{IK}^*(u) \approx \sum_{k=1}^{K+1} [\bar{z}_k - z_{IK}^*(u; z_k | (n)) - F_{IK}(u; z_{k-1} | (n))] \quad (14)$$

and the conditional variance $\sigma_{IK}^2(u)$ is given by the expression:

$$\sigma_{IK}^2(u) \approx \sum_{k=1}^{K+1} [\bar{z}_k - z_{IK}^*(u)]^2 \cdot [F_{IK}(u; z_k | (n)) - F_{IK}(u; z_{k-1} | (n))] \quad (15)$$

where $\bar{z}_k$ is the mean of the class $(z_{k-1}, z_k]$ which depends on the within-class interpolation model (e.g. linear). The indicator kriging estimation is a non-parametric estimation method and does not require any distributional assumption (Goovaerts, 1997). On the other end, a ccdf can be constructed for ordinary kriging and Poisson kriging under the assumption of normality of prediction errors. In that case, the ccdf is defined by its mean and variance which corresponds to the kriging estimate and kriging variance.

The *Poisson kriging estimator* is defined as (Monestiez et al., 2006):

$$z_{PK}^*(u) = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{PK}(u) z(u_{\alpha}), \quad \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{PK}(u) = 1 \quad (16)$$

The optimal weights are obtained after solving a Poisson kriging system of $[n(u) + 1]$ linear equations:

$$\left\{ \begin{array}{l} \sum_{\beta=1}^{n(u)} \left[\lambda_{\beta}^{PK}(u) \gamma_w(u_{\alpha} - u_{\beta}) + \lambda_{\alpha}^{PK}(u) \delta_{\alpha\beta} \frac{m^*}{d(u_{\alpha})} \right] - \mu_{PK}(u) = \gamma_w(u_{\alpha} - u), \quad \alpha=1, 2, \dots, n(u) \\ \sum_{\beta=1}^{n(u)} \lambda_{\beta}^{PK}(u) = 1 \end{array} \right. \quad (17)$$

where $\delta_{\alpha\beta} = 1$ if $\alpha = \beta$ and 0 otherwise. The term $m^*/d(u_{\alpha})$ represents the error variance and increases as the fishing effort $d(u_{\alpha})$ decreases. The incorporation of this variance in the kriging system (Eq. (17)) results in assigning smaller weights $\lambda_{\alpha}^{PK}(u)$ to less reliable rates $z(u_{\alpha})$ based on smaller fishing efforts. The Poisson kriging error variance is obtained by:

$$\sigma_{PK}^2(u) = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{PK}(u) \gamma_w(u_{\alpha} - u) - \mu_{PK}(u) \quad (18)$$

where $\gamma_w(u_{\alpha} - u_{\beta})$ is the semivariance between observations at location u_{α} and u_{β} , $\gamma_w(u_{\alpha} - u)$ is the semivariance between the sampling location u_{α} and the unsampled location u being estimated, and $\mu_{PK}(u)$ is the Lagrange multiplier.

The choice of the number of observations $n(u)$ to be used in the kriging estimation is somewhat subjective. However, following the guidelines by Webster and Oliver (2007), all predictions in the present paper used a minimum of 3 observations (up to a maximum of 15) within a search neighbourhood of radius close to the range of the fitted semivariogram model.

2.3. Evaluation of estimators

Geostatistical models are designed to describe the variability in the data to facilitate estimation. The relative performance of the three kriging estimators was assessed using a leave-one-out (LOO) cross-validation procedure (Goovaerts, 1997; Isaaks and Srivastava, 1989), where each observation was removed from the data set in turn and the value at that location was re-estimated from the remaining data using each kriging estimators (Goovaerts, 1997; Isaaks and Srivastava, 1989). Because the removal of a single observation has little influence on semivariogram estimation the cross validation was performed for each of the three estimators based on the corresponding semivariogram model fitted to the entire data set. The procedure was repeated for all three data sets considered, and then final evaluation of the estimators was based on the goodness statistic (G) introduced by (Deutsch, 1996).

The goodness statistic provides a measure of the local accuracy of the estimates and is computed from the observation $z(u_{\alpha})$ and the corresponding cdfs $F(u_{\alpha}; z | (n))$ modelled from the LOO cross-validation using each kriging estimator. By construction there is a probability p ($0 < p < 1$) the true value falls into a given symmetric p -interval bounded by the $(1 - p)/2$ and $(1 + p)/2$ quantiles of the cdf (e.g. 0.5- p -interval is bounded by lower and upper quartiles). Looking at all N observations, one should thus expect that a proportion p of these observations fall into the p -interval. This “expected” proportion p can be compared to the “actual” proportion that is computed as:

$$\bar{\xi}(p) = \frac{1}{N} \sum_{\alpha=1}^N \xi(u_{\alpha}; p) \quad (19)$$

with

$$\xi(u_{\alpha}; p) = \begin{cases} 1 & \text{if } F^{-1}\left(u_{\alpha}, \frac{(1-p)}{2}\right) < z(u_{\alpha}) \leq F^{-1}\left(u_{\alpha}, \frac{(1+p)}{2}\right) \\ 0 & \text{otherwise} \end{cases}$$

A probability distribution is said to be accurate at p if for example, the 50% symmetric p -interval contains the true value 50% (or more) of the time, and so on for a series of p -probability intervals (Deutsch, 1996). The scatter plot of the calculated proportion $\bar{\xi}(p)$ versus the expected proportion p is called the accuracy plot. The departure from the 45° bisector line on the plot of $\bar{\xi}(p)$ versus p , which corresponds to a perfect prediction, can be numerically estimated by the goodness statistic (G):

$$G = 1 - \int_0^1 [3a(p) - 2][\bar{\xi}(p) - p] dp \quad (20)$$

where $a(p)$ is an indicator variable set to 1 if $\bar{\xi}(p) > p$ and 0 otherwise (Deutsch, 1996).

Maximum goodness corresponds to the case $G=1$ ($\bar{\xi}(p)=p$ for all p , $0 \leq p \leq 1$).

Global criteria for evaluating the performance of each estimator are the mean square error (MSE), mean absolute error (MAE), mean error (ME) and mean square standardised residual (MSSR) of prediction. They are computed as:

$$\text{MSE} = \frac{1}{N} \sum_{\alpha=1}^N [z^*(u_{\alpha}) - z(u_{\alpha})]^2 \quad (21)$$

$$\text{MAE} = \frac{1}{N} \sum_{\alpha=1}^N |z^*(u_{\alpha}) - z(u_{\alpha})| \quad (22)$$

$$\text{ME} = \frac{1}{N} \sum_{\alpha=1}^N [z^*(u_{\alpha}) - z(u_{\alpha})] \quad (23)$$

$$\text{MSSR} = \frac{1}{N} \sum_{\alpha=1}^N \left[\frac{z^*(u_{\alpha}) - z(u_{\alpha})}{\sigma^*(u_{\alpha})} \right]^2 \quad (24)$$

where $z(u_{\alpha})$ and $z^*(u_{\alpha})$ are the observed and predicted catch rates at the sampled location respectively and $\sigma^*(u_{\alpha})$ represents the estimation error standard deviation. The observed and kriged catch rates are estimators of the (unknown) latent catch rate at a location, but the observed rate is impacted by the underlying effort, while for kriging these are smoothed. As a consequence, the kriged catch rate estimates are expected to be closer to the observed

catch rates for large effort than for small effort. Since the observed catch rates can be unreliable because of the small number problem the MSE derived for each kriging estimator is calculated for different classes of the fishing effort (the deciles). A method is deemed effective for estimating the latent rate if the MSE is lower with increasing deciles (Clough, 2012). The MSSR is about one if the actual estimation error is equal on average to the error predicted by the model (Wackernagel, 2003).

3. Results

The distributions of catch rates for the three species were positively skewed with high proportions of zero catch rates, ranging from 18% to 46% for snapper, baldchin groper and blue swimmer crab, and the variation in catch rates was higher for baldchin groper than for snapper and blue swimmer crab (Table 1). Even though the sample size for blue swimmer crab was lower than for the other two species (Table 1), the number of observations are sufficient to infer semivariogram models (Antunes and Albuquerque, 2013; Figueira et al., 2009). Location maps of catch rates indicated high rates for snapper and baldchin groper in the northern part of the study area whilst those of blue swimmer crab were higher in the southern part (Fig. 1). The map also indicated high fishing effort for snapper and baldchin groper in the south-eastern part of the West Coast bioregion compared to more evenly distributed fishing effort for blue swimmer crab.

3.1. Spatial variability modelling

Due to the elongated shape of the study area (Fig. 1), it was not possible to model potential anisotropy in the data. Hence, isotropic behaviour was assumed, and for ordinary kriging omnidirectional experimental semivariograms were computed for each species based on Eq. (2). Fig. 2 shows the fitted semivariogram model(s) for snapper, baldchin groper and blue swimmer crab. The spatial structures of snapper and baldchin groper catch rates were described by a nugget effect, a short-range and a long range spherical model with spatial ranges of 20 NM and 98 NM for snapper and 57 NM and 150 NM for baldchin groper (Table 2). The spatial structure of blue swimmer crab catch rates was modelled by a nugget and a single spherical structure. The NSR was above 0.25, indicative of a moderate short scale spatial autocorrelation. The NSR was the lowest for the semivariogram model of the snapper catch rates, but at 20 NM the partial sill of the model already accounts for 80% of the total variance compared to 67% for baldchin groper and 71% for blue swimmer crab catch rates. Thus even though the models for baldchin groper and blue swimmer crab catch rates have higher NSRs, the ranges of autocorrelation are longer and the overall spatial autocorrelation is strongest for baldchin groper.

Following Eq. (3), the catch rate data were transformed into indicator variables. Deciding the number of thresholds was facilitated by the Auto-IK programme (Goovaerts, 2009). This programme provides a flexible implementation of the indicator kriging approach, including selection of thresholds for continuous data; computation and automatic modelling of indicator semivariograms, calculation of the mean and variance of the probability distribution as well as derivation of cross-validation statistics for model evaluation. Based on the goodness statistics computed from the cross-validation, MAE, ME, MSSR, the percentage of order relation violations and the structure of the semivariograms five

thresholds were found to be sufficient for modelling the conditional cumulative distribution of the catch rates for each species (see Supplementary Table 1 and Supplementary Figure 1), as the values of the goodness statistic, MAE, ME and MSSR vary with small differences with increasing number of thresholds for snapper, baldchin groper and blue swimmer crab. Thus, the use of more than five thresholds will not improve the performance of the models much. In addition the proportion of order relation violations was minimum using five thresholds for each species. The first threshold was determined naturally by the spikes of zero catch rates. For snapper catch rates the chosen thresholds correspond to the 18th, 36th, 52nd, 68th and 84th percentiles, whilst those of baldchin groper correspond to the 46th, 59th, 69th, 79th, 90th, and those of blue swimmer crab to the 40th, 54th, 65th, 77th and 90th percentiles of the cumulative distribution function (cdf) of the catch rate data (see Supplementary Figure 1). For each species, semivariogram models were fitted to the experimental indicator semivariograms for the various thresholds (Fig. 3). For all three species, the first indicator threshold corresponds to a catch rate of 0. For each of the three species, a NSR between 0.6 and 0.7 was observed for this threshold, indicating moderate spatial connectivity for zero catch rates. The parameters of the semivariogram models showed that the range of spatial influence of catch rate was low for snapper with ranges varying between 20 and 33 NM and high NSRs for the second and third thresholds. For baldchin groper the ranges varied from 48 to 88 NM and for blue swimmer crab from 91 to 120 NM (Table 3). For snapper and baldchin groper the ranges are lower than the modelled ranges for the traditional semivariogram estimator. Nugget effects and ranges of spatial continuity vary between thresholds, indicating changes in spatial structure for each threshold (Table 3). For none of the three species, the NSR and range show any clear trends as the threshold values increase.

Similarly, a model was fitted to each experimental weighted semivariogram computed according to Eq. (5) (Fig. 4). The spatial structures of catch rates for snapper and blue swimmer crab were described by a nugget and a single spherical model with ranges of spatial influence of 51 and 40 NM, respectively (Table 4). The spatial structure of baldchin groper catch rates was modelled by a spherical structure with a zero nugget effect (Table 4). Unlike the traditional and indicator semivariograms, the semivariograms weighted by fishing effort displayed lower sill and relative nugget effect. For each species, the nugget effect inferred from the weighted estimator was much smaller compared to the traditional semivariogram estimator (Table 3), which indicates that the short-range fluctuation displayed by the latter are caused by unreliable catch rates. Also, the NSR for the models indicates strong short-scale spatial correlation in the catch rates of baldchin groper followed by the blue swimmer crab and snapper respectively. However, the ranges of spatial autocorrelation were shorter when the weighted semivariogram estimator was used compared to the method of moment estimator.

3.2. Spatial estimation and mapping of catch rate

Based on the spatial structure of the fitted semivariogram models, catch rates for each species were interpolated at the nodes of a 10×10 NM spacing grid. For each species, maps of catch rates estimated by ordinary, indicator and Poisson kriging visually look fairly similar (Fig. 5). Much larger differences are observed for the variance maps (Fig. 6). As

indicated in Section 2.2.2, the variance obtained from ordinary and Poisson kriging are kriging error variances. They differ in the sense that the ordinary kriging variance is essentially a measure of data configuration (i.e. higher uncertainty in sparsely sampled areas), whereas the Poisson kriging variance also accounts for the fishing effort (i.e. higher uncertainty where fishing duration is low). Because both fishing effort and number of data increase near the shore of the bioregion (Fig. 1, bottom maps), the two maps share similar spatial patterns with lower variance near the shore. The magnitude of Poisson kriging variance is smaller as the lower impact of extreme rates on semivariogram estimation reduces the sill of the weighted semivariogram compared to the traditional semivariogram. In contrast to OK and PK variances, the IK-based variance represents the variance of the conditional cumulative distribution function (ccdf) and indicates higher uncertainty in areas where variation in catch rates was higher (Fig. 6).

3.3. Comparison of estimators

The performance of the kriging estimators was evaluated based on the prediction statistics computed from the cross-validation results. The accuracy plots indicated that the proportion of measured data values falling into different probability intervals was closer to the expected proportion when the IK estimator was used, compared to OK and PK (Fig. 7). For OK, the p -intervals are generally wider up to probabilities of 0.8, and the same is observed for PK in the case of baldchin groper and blue swimmer crab. In contrast, the PK estimator for snapper results has tighter than expected probability intervals. These results are also reflected in the goodness statistic, which is largest for the IK estimators (Fig. 7).

For snapper and baldchin groper, the MSE for the three estimators decreases towards 0 for an increase in fishing effort from the 2nd decile, whilst for blue swimmer crab, the MSE for the three estimators fluctuate for different classes of fishing effort (Fig. 8). Thus, for high deciles of fishing effort, the accuracy of the three estimators for estimating the latent catch rates of snapper and baldchin groper increases. For snapper and baldchin groper, the mean square error (MSE) of IK remains lower for an increasing fishing effort as compared to OK and PK estimators. This indicates that IK performed better in predicting the latent catch rate for snapper and baldchin groper compared to OK and PK. However, the differences between OK and PK MSE only occur at the low effort deciles. For blue swimmer crab, there is no clear cut as which of the three estimators perform better. Similarly, for all the three species the global accuracy of the three modelling approaches as indicated by the MAE and MSSR shows no clear indication as to which of the three estimators performs better (Table 5). Based on the MAE, PK performs best for estimating the catch rates for all the three species. However, the difference among estimators is negligible. Based on the ME, OK performs best for estimating the catch rate of snapper and blue swimmer crab. On the average, the OK prediction variance has the ability to capture the actual magnitude of the prediction error for baldchin groper and blue swimmer crab catch rates whilst IK does best for snapper catch rates as indicated by the MSSR value. Since these global statistics can be influenced by extreme values as well as the magnitude of the extreme error (Li and Heap, 2008), the comparison in general will be based on the goodness statistics and the accuracy plots. In general, IK performs best compared to the other two approaches for estimating the catch rates for the three species based on the goodness statistics.

4. Discussion and conclusions

Unlike catch rate data for commercial fish species whose spatial patterns have been successfully modelled using geostatistics (Kleisner et al., 2010; Petitgas et al., 2003), the recreational catch rate data for all three fisheries species were highly skewed with a high proportion of zero values. The data have these properties because many recreational fishing trips fail to catch any fish (O'Neill and Faddy, 2003), which would at least partly reflect the distribution and abundances of the fish species, and the choice of fishing locations and fishing ability of fishers. Despite this, we show that geostatistical techniques are practicable for estimating catch rates from recreational fishing data. However, of the three estimators used for kriging and the description of spatial patterns in catch rates, IK provides the best prediction performance for the three species, regardless of their life history and distribution patterns.

Of the three kriging estimators, it was anticipated that ordinary kriging (OK) would perform worse in estimating the recreational catch rates than IK or PK, since the data used in this study were highly skewed with zeros, and such properties influence the traditional variogram (Kerry et al, 2010; Monestiez et al., 2006). Based on the accuracy plots and the goodness statistics, IK performed better than OK and PK for all three species. The accuracy plot of IK for all three species showed a good correspondence between expected and empirical proportions of catch rate data falling within probability intervals of increasing size, as measured by the goodness statistic and may reflect the property of IK to provide a better estimate of the local distribution for these highly skewed, zero-inflated data. The performance of PK was poor, which might be a result of the estimates being weighted by the fishing effort to remove the noise in the original data. Such estimates might look different in most cases, compared to the true values, leading to low accuracy and goodness statistics. The mean square error as a function of the effort decile also confirmed the superiority of IK for predicting the latent catch rate for the three species compared to OK and PK. For snapper and baldchin groper, the prediction performance of IK was better for all effort deciles whilst for blue swimmer crab the MSE for IK was smaller from the 4th decile upwards and the MSE for PK was smaller from the 3rd decile downwards. Also, for high deciles of fishing effort, the accuracy of the three estimators for estimating the latent catch rates of snapper and baldchin groper increased. This suggests that IK provides an accurate estimate of catch rates of all three species, regardless of their life history characteristics and distribution patterns. Snapper can exhibit spawning aggregations, with juveniles occupying more sheltered coastal waters prior to migrating to deeper, more offshore waters, while baldchin groper appears to have no spawning aggregations and the habitats of juvenile and adult habitats overlap (Smallwood et al, 2013). In comparison, the adults and juveniles of the blue swimmer crab occupy inshore marine and estuarine waters (Kailola et al, 1993). Variability in the distribution of life history stages of the various species across the fishing region would presumably influence the spatial patterns in effort and catch rates, particularly in the frequency and distribution of zero catches, and thereby influence the performance of the kriging estimators. However, IK appeared to overcome these issues and provide accurate estimates of catch rates for all three species. In the IK approach, more spatial information is considered through multiple indicator semivariograms. By using multiple indicator

semivariograms, more information about class-specific spatial behaviour in the catch rate data is obtained compared to a single semivariogram as used by OK and PK. For none of the species was there a clear cut conclusion as to what method performs best from the perspective of global accuracy performance as measured by the MAE, ME and MSSR. Based on the MAE, PK performs best for estimating the catch rates for all the three species. However the differences between the three estimators are negligible. Based on the ME, OK performs better in estimating the catch rate of snapper and blue swimmer crab whilst PK performs better for estimating baldchin groper catch rate. On the average, the OK prediction variance has the ability to capture the actual magnitude of the prediction error whilst IK does best for snapper catch rates as indicated by the MSSR value. Since these global statistics can be influenced by extreme values as well as the magnitude of the extreme error (Li and Heap, 2008), the comparison in general will be based on the goodness statistics and the accuracy plots.

Our results differ from those of (Kerry et al., 2010), who found the performance of IK to be poor compared to PK when mapping the spatial pattern of abundances of herbivore species (e.g. zebra and giraffe) in South Africa, which would also be highly skewed and zero-inflated. A possible explanation is that the indicator semivariograms of catch rate data for recreational fish species were better structured than the semivariograms for rare large counts of herbivore species. Another difference between the two datasets was based on the level of spatial support. The spatial support was highly heterogeneous for the count data for herbivore species, but this was not a major problem for the fisheries catch rate data since all data were measured on the same spatial support.

For each species, the map of spatial estimates produced by ordinary kriging was globally similar to those obtained by IK and PK. However, in some areas, PK filtered the noise in the catch rate data caused by low fishing effort. The kriging error variances from OK and PK showed similar patterns, but were lower in magnitude for PK than for OK. This result is consistent with the findings of Monestiez et al. (2006), who argued that the PK variance is modulated by the effort and does not solely depend on the geometric configuration of the data, but rather is proportional to the inverse of the effort when available. The map of OK and PK variances indicated greater uncertainty in the western part of the study area where data are sparse, and smaller uncertainty in areas where data are much more concentrated. In contrast, the map of conditional variance from IK showed a different pattern and indicated greater uncertainty in areas with high variations in catch rates (i.e. the northern and central part of the bioregion). Indeed, the IK variance depends on the data values, whereas the error variance for OK and PK depends on the data configuration and this makes a direct comparison quite difficult. The two types of variance thus highlight different aspects of uncertainty and are both important for the spatial characterisation of catch rates. However, the choice of one over the other will depend on the objectives of the study. For instance, the estimation error variance is useful if the objective is to evaluate the sampling design of the survey (Petitgas, 2001, 2009). In contrast, if the objective is to evaluate local precision, then the conditional variance from IK will be more suitable (Goovaerts, 1997; Lloyd and Atkinson, 2001). The conditional variance from the IK is also a useful measure of uncertainty for data with proportional effect (Goovaerts, 1997).

In conclusion, geostatistical techniques are practicable for estimating catch rates from recreational fishing and mapping patterns in those rates. While mapping of catch rates has been used successfully for commercial species (Kleisner et al., 2010; Petitgas et al., 2003), recreational catches suffer from zero-inflated data in addition to highly skewed distributions. However, we show that IK may be preferable for the spatial modelling of catch rate data exhibiting these characteristics, and has the best prediction performance regardless of the life history and distribution patterns of those three species. Compared to other known catch rate estimation techniques such as the generalised linear model (GLM) and generalised additive modelling approach (GAM), kriging allows one to estimate catch rates even in the absence of explanatory variables such as environmental factors (Yu et al., 2013). Also, kriging does not share the shortcomings of spatial GAM/GLM and delta-GLM approaches, including the requirement for explanatory variables, error structure assumptions and interaction effect problems (Yu et al., 2013).

Understanding spatial distributions of catch rates through mapping the kriged estimates can be a useful tool for fisheries management. For example, it can be applied to detect “hotspots” of catch rates and spatial shifts in those hotspots or declines in catch rates over time. This could be used to target boundaries for spatial closures to prevent or mitigate overfishing in certain regions. Another application is the comparison of fishing hotspots between fishing sectors where resource allocation issues exist. Allocating resources between conflicting fishing sectors can be difficult and complex (Arlinghaus, 2005). Two of the species (snapper and baldchin groper) analysed in this study are part of the West Coast Demersal Scalefish fisheries in Western Australia, for which there has been a 34 and 66% allocation of the resource to the recreational and commercial fishery, respectively, through the Integrated Fisheries Management Policy (Crowe et al., 2013). As part of that policy, management measures have been implemented to reduce catches in both sectors through input and output controls due to concerns regarding overfishing. For example, a spatial closure for the commercial sector was implemented in 2005/06 (Crowe et al., 2013), but there was no information on spatial patterns in recreational catch rates at the time. Such information could, therefore, be used to define boundaries for spatial closures, and areas for segregation of fishing activities of both sectors. Since resource allocation is an issue in a range of countries, including USA (Wallace et al., 1994), Canada (Pearse, 2010) and New Zealand (Pearse, 2010), the geostatistical approaches we have used here for recreational fisheries have wide application.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors are grateful to the staff of the Western Australian Department of Fisheries for providing the data for the study and also to K.L. Ryan, B.S. Wise, and N.G. Hall for their useful comments and suggestions. Any remaining errors are our own responsibility. The research conducted by the third author was funded by the grant 1R21 ES021570-01A1 from the National Cancer Institute (NCI). Views stated in this publication do not necessarily represent the official views of the NCI.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.fishres.2015.03.013>.

References

- Antunes IMHR, Albuquerque MTD. Using indicator kriging for the evaluation of arsenic potential contamination in an abandoned mining area (Portugal). *Sci Total Environ*. 2013; 442:545–552. [PubMed: 23220092]
- Arlinghaus R. A conceptual framework to identify and understand conflicts in recreational fisheries systems, with implications for sustainable management. *Aquat Resour Cult Dev*. 2005; 1:145–174.
- Baharthah, T. Department of Fisheries Community Survey 2007. Department of Fisheries; Western Australia: 2008. Fisheries Occasional Paper No. 47 http://www.fish.wa.gov.au/Documents/occasional_publications/fop047.pdf
- Cambardella CA, Moorman TB, Novak JM, Parkin TB, Karlen DL, Turco RF, Konopka AE. Field-scale variability of soil properties in central Iowa soils. *Soil Sci Soc Am J*. 1994; 58:1501–1511.
- Chiles, JP.; Delfiner, P. *Geostatistics: Modelling spatial uncertainty*. John Wiley & Sons, Inc; Hoboken, NJ: 2012.
- Clough, L. *Disease Mapping of Syphilis in Forsyth County, North California with Enhanced Geoprivacy and Spatial Resolution*. University of North Carolina; Chapel Hill: 2012. <http://dc.lib.unc.edu/cdm/ref/collection/etd/id/5411>
- Coll J, Linde M, Garcia-Rubies A, Riera F, Grau AM. Spear fishing in the Balearic Islands west central Mediterranean: species affected and catch evolution during the period 1975–2001. *Fish Res*. 2004; 70:97–111.
- Cooke SJ, Cowx IG. The role of recreational fishing in global fishing crises. *Bioscience*. 2004; 54:857–859.
- Cowx, IG. Recreational fishing. In: Hart, P.; Reynolds, JD., editors. *Handbook of Fish Biology and Fisheries*. Blackwell Science; Oxford: 2002.
- Crowe FM, Longson IG, Joll LM. Development and implementation of allocation arrangements for recreational and commercial fishing sectors in Western Australia. *Fish Manag Ecol*. 2013; 20:201–210.
- Department of Fisheries. *Monitoring the Status of the West Coast Demersal Scalefish Resources*. Department of Fisheries Newsletter. WA Fisheries and Marine Research Laboratories; P.O. Box 20, North Beach, WA 6920: 2012a.
- Department of Fisheries. *A Resourced-Based Management Approach for Recreational Fishing in Western Australia 2012–2017: State-wide Management Proposals for Finfish, Crustaceans, Molluscs and Other Invertebrates Fisheries Management Paper, No 252*. Department of Fisheries; Western Australia: 2012b. http://www.fish.wa.gov.au/Documents/management_papers/fmp252.pdf
- Deutsch, CV. Direct assessment of local accuracy and precision; 5th International Geostatistics Congress; Wollongong' 96. 1996.
- Figueira R, Tavares PC, Palma L, Beja P, Sergio C. Application of indicator kriging to the complementary use of bioindicators at three trophic levels. *Environ Pollut*. 2009; 157:2689–2696. [PubMed: 19477568]
- Fletcher, WJ.; Santoro, K., editors. *Status Reports of the Fisheries and Aquatic Resources of Western Australia 2011/12: The State of the Fisheries*. Department of Fisheries, Western Australia; 2012. http://www.fish.wa.gov.au/Documents/sofar/status_reports_of_the_fisheries_and_aquatic_resources_2011-12.pdf
- Goovaerts, P. *Geostatistics for Natural Resources Evaluation*. Oxford University Press; New York: 1997.

- Goovaerts P. Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging. *Int J Health Geogr.* 2005; 4:31. <http://dx.doi.org/10.1186/1476-072X-4-31>. [PubMed: 16354294]
- Goovaerts P. AUTO-IK: a 2D indicator kriging program for the automated non-parametric modeling of local uncertainty in earth sciences. *Comput Geosci.* 2009; 35:1255–1270. [PubMed: 20161335]
- Henry, WG.; Lyle, MJ. Final Report for FRDC Project, No 99/158. Australian Government Department of Agriculture, Fisheries and Forestry; Canberra: 2003. The National Recreational and Indigenous Fishing Survey. <http://www.environment.gov.au/system/files/pages/5886ef12-7b4a-40fd-aea5-c08e7621ecbf/files/co62nationalrecreationalfishingurvey.pdf>
- Isaaks, EH.; Srivastava, RM. An Introduction to Applied Geostatistics. Oxford University Press, Inc; New York, USA: 1989.
- Jones CM, Robson DS, Lakkis HD, Kressel J. Properties of catch rate used in analysis of angular survey. *Trans Am Fish Soc.* 1995; 124:911–928.
- Kailola, PJ. Fisheries Research and Development Corporation (Australia), Australia. Bureau of Resource Sciences. Australian Fisheries Resources. Bureau of Resource Sciences and the Fisheries Research and Development Corporation; Canberra, ACT: 1993.
- Kerry, R.; Goovaerts, P.; Smit, IPJ.; Ingram, BR. Accuracy 2010 Symposium. Leicester, UK: 2010. A comparison of indicator and poisson kriging of herbivore species abundance in Kruger National Park, South Africa.
- Kleisner KM, Walter JF III, Diamond SL, Die DJ. Modelling the spatial autocorrelation of pelagic fish abundance. *Mar Ecol Prog Ser.* 2010; 411:203–213.
- Li, J.; Heap, AD. A Review of Spatial Interpolation Methods for Environmental Scientists. Geoscience Australia, Record 2008/23. 2008. http://www.ga.gov.au/webtemp/image_cache/GA12526.pdf
- Lindner, RK.; McLeod, PB. An Economic Impact of Recreational Fishing in Western Australia. Department of Fisheries, Western Australia, WA Fisheries and Marine Research Laboratories; PO Box 20, North Beach, WA 6920: 1991. Fisheries Management Paper, No. 38
- Lloyd CD, Atkinson PM. Assessing uncertainty in estimates with ordinary and indicator kriging. *Comput Geosci.* 2001; 27:929–937.
- Matheron, G. La théorie des variables régionalisées et ses applications. Masson et Cie; Éditeurs, Paris: 1965.
- Monestiez P, Dubroca L, Bonnin E, Durbec JP, Guinet C. Geostatistical modelling of spatial distribution of *Balaenoptera physalus* in the Northwestern Mediterranean Sea from sparse count data and heterogeneous observation effort. *Ecol Model.* 2006; 193:615–628.
- Mueller U, Dickson J, Kangas M, Caputi N. Geostatistical modelling of the scallop density distribution in Shark Bay. *Geostats.* 2008; 2008:971–980.
- O'Neill MF, Faddy MJ. Use of binary and truncated negative binomial modelling in the analysis of recreational catch data. *Fish Res.* 2003; 60:471–477.
- Pearse P. Allocation of catches among fishing sectors: opportunities for policy development. *FAO Fish Aquac Proc.* 2010; 15:123–135.
- Petitgas P. Geostatistics in fisheries survey design and stock assessment: models, variances and applications. *Fish Fish.* 2001; 2:231–249.
- Petitgas, P. Geostatistics and their application to fisheries survey data: a history of ideas, 1990–2007. In: Megrey, BA.; Moksness, E., editors. *Computers in Fisheries Research*. Springer Science +Business Media B.V.; 2009.
- Petitgas P, Poulard JC, Biseau A. Comparing commercial and research survey catch per unit of effort: megrim in the Celtic Sea. *ICES J Mar Sci.* 2003; 60:66–76.
- Rivoirard, J.; Simmonds, J.; Foote, KG.; Fernandes, P.; Bez, N. *Geostatistics for Estimating Fish Abundance*. Blackwell Science; United Kingdom: 2000.
- Rossi RE, Mulla DJ, Journel AG, Franz EH. Geostatistical tools for modelling and interpreting ecological spatial dependence. *Ecol Monogr.* 1992; 62:277–314.
- Ryan, KL.; Wise, BS.; Hall, NG.; Pollock, KH.; Sulin, EH.; Gaughan, DJ. An Integrated System to Survey Boat-Based Recreational Fishing in Western Australia 2011/2012. Department of Fisheries; Western Australia: 2013. Fisheries Research Report No. 249

- Smallwood, CB.; Hesp, SA.; Beckley, LE. Biology, Stock Status and Management Summaries for Selected Fish Species in South-Western Australia. Department of Fisheries; Western Australia: 2013. Fisheries Research Report No. 242
- Susan, K. The small number problem and the accuracy of spatial databases. In: Goodchild, MF.; Gopal, S., editors. The Accuracy of Spatial Databases. Taylor & Francis Ltd; 1989.
- Taylor SM, Webley JAC, Mayer DG. Improving the precision of recreational fishing harvest estimates using two-part conditional general linear models. *Fish Res.* 2011; 110:408–414.
- Wackernagel, H. Multivariate Geostatistics. Springer; Germany: 2003.
- Wallace, RK.; Hosking, W.; Szedlmayer, ST. Fisheries Management for Fishermen: A Manual for Helping Fishermen to understand the Federal Management Process. Auburn University Marine Extension & Research Center Sea Grant Extension; Mobile: 1994.
- Walters CJ. Folly and fantasy in the analysis of spatial catch rate data. *Can J Fish Aquat Sci.* 2003; 60:1433–1436.
- Webster, R.; Oliver, MA. Geostatistics for Environmental Scientists. John Wiley & Sons, Ltd; 2007.
- Wise BS, Telfer CF, Lai EKM, Hall NG, Jackson G. Long-term monitoring of boat-based recreational fishing in Shark Bay. Western Australia: providing scientific advice for sustainable management in a World Heritage Area. *Mar Freshw Res.* 2012; 63:1129–1141.
- Yu H, Jiao Y, Carstensen LW. Performance comparison between spatial interpolation and GLM/GAM in estimating relative abundance indices through a simulation study. *Fish Res.* 2013; 147:186–195.

Fig. 1. Location maps of catch rates (top panel) and effort (bottom panel) for snapper (left), baldchin groper (middle) and blue swimmer crab (right). Snapper and baldchin groper catch rates share the same colour scale.

Fig. 2. Omnidirectional experimental semivariogram (dots) for snapper (left), baldchin groper (middle) and blue swimmer crab (right) catch rate with the fitted isotropic model (solid line).

Fig. 3.

Omnidirectional experimental indicator semivariogram (dots) with the fitted models (solid line) for snapper (left panel), baldchin groper (middle panel) and blue swimmer crab (right panel) computed for the five thresholds. For each panel, the semivariogram and the fitted model are presented for the various thresholds from top to bottom.

Fig. 4. Omnidirectional experimental weighted semivariogram (dots) for snapper (left), baldchin groper (right) and blue swimmer crab (bottom) catch rate with the fitted isotropic models (solid line).

Fig. 5.

Maps of ordinary (left), indicator (middle) and Poisson (right) kriging estimates of catch rate. The top, middle and bottom panels present the estimates for snapper, baldchingroperand blue swimmer crab respectively. Snapper and baldchin groper share the same colour scale.

Fig. 6.

Maps of ordinary (left), indicator (middle) and Poisson variance. The top, middle and bottom panels present the variance estimates for snapper, baldchin groper and blue swimmer crab respectively. Snapper and baldchin groper share the same colour scale.

Fig. 7.

Accuracy plot and goodness statistic obtained from leave-one-out cross-validation for snapper (top), baldchin groper (middle) and blue swimmer crab (bottom). The left, middle and right panels present the results from ordinary, indicator and Poisson kriging estimation respectively.

Fig. 8.

The MSE for OK, IK and PK for different classes of fishing effort (the decile). The left, middle and right panels present the results for snapper, baldchin groper and blue swimmer crab, respectively. The MSE has been standardised by the maximum MSE of 0.135, 0.215 and 9.881 for snapper, baldchin groper and blue swimmer crab respectively.

Table 1

Summary statistics of catch rates for snapper, baldchin groper and blue swimmer crab.

Variable	Count	Min	Max	Mean	LQ	Median	UQ	CV	% zeroes
Snapper	142	0.000	1.263	0.223	0.042	0.127	0.297	1.162	18
Baldchin groper	130	0.000	1.600	0.128	0.000	0.009	0.118	2.176	46
Blue swimmer crab	57	0.000	13.375	2.205	0.000	0.248	3.636	1.522	40

Min, Max, LQ, UQ and CV represent the minimum, maximum, lower quartile, upper quartile and coefficient of variation respectively.

Table 2

Parameters of the fitted method of moment semivariogram models for snapper, baldchin groper and swimmer crab and their corresponding nugget to total sill ratio (NSR) computed as $Q_0/(Q_0 + Q_1 + Q_2)$. The models for snapper and baldchin groper are made up of a nugget and two spherical models.

Species	Q_0	<u>Spherical model 1</u>		<u>Spherical model 2</u>		NSR
		R_1	Q_1	R_2	Q_2	
Snapper	0.023	20	0.027	98	0.020	0.329
Baldchin Groper	0.036	57	0.028	150	0.015	0.456
Blue swimmer crab	6.600	95	4.800	-	-	0.579

Table 3

Parameters of the fitted spherical models for the indicator data for snapper, baldchin groper and blue swimmer crab and their corresponding nugget to sill ratios (NSR). The table also provides the threshold values in terms of catch rate and their corresponding cumulative percentages.

Percentage	Threshold	Q_0	<u>Spherical model</u>		NSR
			R	Q	
<i>Snapper</i>					
18	0.000	0.103	31	0.050	0.675
36	0.082	0.204	33	0.022	0.904
52	0.132	0.221	30	0.020	0.918
68	0.237	0.138	20	0.064	0.682
84	0.418	0.056	23	0.079	0.417
<i>Baldchin groper</i>					
46	0.000	0.154	69	0.101	0.603
59	0.024	0.103	75	0.137	0.429
69	0.078	0.101	72	0.116	0.466
79	0.143	0.100	88	0.076	0.568
90	0.312	0.036	48	0.055	0.400
<i>Blue swimmer crab</i>					
40	0.000	0.174	120	0.087	0.667
54	0.268	0.185	102	0.059	0.758
65	1.289	0.146	102	0.074	0.663
77	3.906	0.135	118	0.052	0.723
90	7.423	0.071	91	0.046	0.606

Table 4

Parameters of the fitted weighted semivariogram models for snapper, baldchin groper, blue swimmer crab and their corresponding NSR. The models consist of a nugget and a spherical model.

Structure	Q_0	<u>Spherical model</u>		$Q_0/(Q_0 + Q)$
		R	Q	
Snapper	0.005	51	0.018	0.217
Baldchin groper	0.000	68	0.034	0.000
Blue swimmer crab	1.366	40	11.507	0.106

Table 5

Global prediction accuracy statistics based on the cross-validation results for snapper, baldchin groper and blue swimmer crab.

Species/estimator	MAE	ME	MSSR
Snapper			
OK	0.175	−0.013	1.063
IK	0.177	0.025	1.053
PK	0.174	−0.011	4.110
Baldchin groper			
OK	0.124	0.002	1.255
IK	0.131	0.022	1.968
PK	0.118	−0.002	4.265
Blue swimmer crab			
OK	2.085	0.026	1.005
IK	1.974	0.237	1.863
PK	1.869	0.148	1.173