

# Propagation of measurement uncertainty in spatial characterisation of recreational fishing catch rates using logistic transform indicator kriging

**E.N. Aidoo<sup>a</sup>, U. Mueller<sup>a</sup>, G.A. Hyndes<sup>b</sup> and K.L. Ryan<sup>c</sup>**

<sup>a</sup>*School of Engineering, Edith Cowan University, Joondalup, Australia*

<sup>b</sup>*Centre for Marine Ecosystem Research, School of Natural Sciences, Edith Cowan University, Joondalup, Australia*

<sup>c</sup>*Western Australian Fisheries and Marine Research Laboratories, Hillarys, Australia*  
Email: [en.aidoo@yahoo.com](mailto:en.aidoo@yahoo.com)

**Abstract:** Geostatistical estimation techniques (such as kriging) have been widely accepted and applied to characterise the spatial distribution of natural phenomena. In fisheries science, these techniques have been applied for computing indices including catch per unit effort (catch rate) used for stock assessment. The nonparametric kriging approach, known as indicator kriging, is particularly helpful for the estimation of catch rates using data observed from recreational fishing surveys because it can also handle other features of the data distribution such as zero-inflation, high skewness and class-specific spatial patterns. The problem considered in this paper is the use of indicator kriging for estimation of catch rates associated with an uncertainty due to multiple measurements observed at some locations. This measurement uncertainty is often non-negligible and needs to be propagated to produce accurate estimates. In addition, the uncertainty might be spatially autocorrelated and correlated with the data and so will restrict the use of a parametric approach of uncertainty propagation. Using catch rate data for Australian herring (*Arripis georgianus*) from a recreational fishing survey, this study presents a soft indicator kriging approach that uses a logistic function transformation to allow the propagation of measurement uncertainty in the estimation. The performance of the uncertainty propagated model was evaluated based on the leave-one-out cross-validation method. The accuracy plot and goodness statistic indicate agreement between the expected and empirical proportions of the observed catch rates falling within probability intervals of increasing size. These suggest a good estimation performance of the approach for Australian herring catch rate data. The spatial distributions of catch rate estimates with propagated uncertainty can be used for quantifying location-specific patterns to assist stock assessment or validation of policy supporting stock assessment models. The measurement uncertainty is considered to be potentially valuable for estimation of catch rate.

**Keywords:** *Logistic transform, soft indicator kriging, catch rate, spatial characterisation*

## 1. INTRODUCTION

Globally, populations of many fish species are either fully or overexploited, creating a challenge for managers of fisheries resources (Cooke and Cowx, 2004). Since it is becoming widely accepted that a significant proportion of the catch from some fisheries is taken by recreational sector (Cooke and Cowx, 2006), there is a need to establish appropriate techniques to assess the recreational sector. The ability to take a proactive approach towards fisheries management requires data collection and stock assessment including catch per unit effort (catch rate), which is frequently used as a relative measure of species abundance. Mapping the distribution of catch rates can assist in delineating high priority areas and improve our understanding of the spatial patterns of recreational fishing at appropriate spatial scales for fisheries management (Aidoo *et al.*, 2015).

Geostatistical estimation techniques (such as kriging) have been widely applied to characterise the spatial distribution of natural resources, including observed catch rates from recreational fishing (Aidoo *et al.*, 2015). The nonparametric kriging approach, known as indicator kriging, is particularly helpful for estimating catch rate because it can also handle other features of the data distribution such as zero-inflation, high skewness and class-specific spatial patterns. In the traditional indicator kriging approach, measured values are transformed into hard indicator variables with the assumption that the data are precise (Goovaerts, 1997). However, observed catch rates from recreational fishing are usually obtained through surveys and may be associated with high levels of uncertainty. For any particular location, observed catch rates are highly variable because skill levels of fishers and motivations for catching fish can differ greatly within any fishery (National Research Council, 1998). For example, in any given trip, only a few fishers catch many fish, while most fishers catch fewer fish (National Research Council, 1998); and the time of fishing within the year and the number of fishers at a particular location may vary resulting in the existence of measurement uncertainty in the observed catch rates. In addition, the uncertainty might be spatially autocorrelated and correlated with the observed catch rates and so will restrict the use of a parametric approach of uncertainty propagation.

There have been limited attempts to propagate measurement uncertainty in kriging estimation such as kriging with measurement error (Hamzehpour *et al.*, 2013) and soft indicator kriging (Saito and Goovaerts, 2002), but only on physico-chemical properties of sediments. The latter approach is more suitable for catch rate data because of the distributional properties of those data. However, the choice of an indicator transformation based on a normal distribution may not be suitable for catch rate data. This paper aims to present a soft indicator kriging approach that uses a logistic function transformation to allow the propagation of the measurement uncertainty in the estimation. We illustrate this approach, with catch rates for Australian herring, *Arripis georgianus*, which is one of the popular nearshore species commonly targeted by recreational fishers (Smith *et al.*, 2013).

## 2. MATERIAL AND METHODS

### 2.1. Data

The data were collected through a phone-diary survey of boat-based recreational fishers in Western Australia (Ryan *et al.*, 2013) supplied by the Department of Fisheries, Western Australia. The study is limited to the West Coast bioregion of Australia extending from latitude 26°30'S to 115°30'E using observed catch rates for Australian herring. Estimates of catch rate and associated uncertainty were calculated for individual 10×10 nautical mile (NM) spatial blocks (location)  $u_\alpha$  using the ratio of mean estimator (Jones *et al.*, 1995). For each location  $u_\alpha$  (block centroid), the catch rate  $z(u_\alpha)$  was defined as:

$$z(u_\alpha) = \frac{\bar{c}(u_\alpha)}{\bar{d}(u_\alpha)} = \frac{\sum_{r=1}^{v(u_\alpha)} c_r(u_\alpha)/v(u_\alpha)}{\sum_{r=1}^{v(u_\alpha)} d_r(u_\alpha)/v(u_\alpha)} = \frac{\sum_{r=1}^{v(u_\alpha)} c_r(u_\alpha)}{\sum_{r=1}^{v(u_\alpha)} d_r(u_\alpha)} \quad (1)$$

and the standard error  $s(u_\alpha)$  of the estimator was approximated as:

$$s(u_\alpha) \approx \sqrt{\frac{1}{v(u_\alpha)\bar{d}^2(u_\alpha)} \cdot \frac{\sum_{r=1}^{v(u_\alpha)} [c_r(u_\alpha) - z(u_\alpha)d_r(u_\alpha)]^2}{v(u_\alpha) - 1}} \quad (2)$$

where  $c_r(u_\alpha)$ ,  $v(u_\alpha)$  and  $d_r(u_\alpha)$  denote respectively the number of fish caught by the  $r^{\text{th}}$  boat, the number of boats that visited the location and the effort in terms of the boat fishing trip duration (hours) multiplied by

the number of persons fishing on each boat. The mean catch and mean effort are represented by  $\bar{c}(u_\alpha)$  and  $\bar{d}(u_\alpha)$  respectively (Cochran, 1977; Sumner and Williamson, 1999).

## 2.2. Soft Indicator Kriging

The indicator kriging approach is a non-parametric modelling of the conditional cumulative distribution function (ccdf) of a random function (i.e. catch rate) using a vector of  $K$  indicator variables. At each location  $u_\alpha$  the indicator variables  $i(u_\alpha; z_k)$  are obtained by defining a set of  $K$  threshold values  $z_1, z_2, \dots, z_K$  and determining whether the threshold value is exceeded or not (Goovaerts, 1997):

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

In equation (3), the observed catch rate  $z(u_\alpha)$  is assumed to be precise without any measurement uncertainty leading to a binary (hard) indicator variable  $i(u_\alpha; z_k)$  for each threshold  $z_k$ . To account for measurement uncertainty, a logistic function was used to transform the observed catch rate  $z(u_\alpha)$  and associated uncertainty  $s(u_\alpha)$  into soft indicators defined as:

$$j(u_\alpha; z_k) = \frac{1}{1 + \exp\left(-\frac{[z_k - z(u_\alpha)]}{s(u_\alpha)}\right)}, \quad s(u_\alpha) > 0 \quad (4)$$

At locations with zero measurement uncertainty hard indicators were used. The spatial variability of each of the soft indicator variables is described by the indicator semivariogram  $\hat{\gamma}(h; z_k)$  describe as:

$$\hat{\gamma}(h; z_k) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [j(u_\alpha; z_k) - j(u_\alpha + h; z_k)]^2 \quad (5)$$

where  $j(u_\alpha; z_k)$  and  $j(u_\alpha + h; z_k)$  represent the soft indicators at locations  $u_\alpha$ , and  $u_\alpha + h$  and  $N(h)$  is the number of pairs of samples separated by the distance vector  $h$  (Goovaerts, 1997).

The soft indicator semivariograms were modelled as linear combinations of a nugget and a spherical model (Goovaerts, 1997) to describe the spatial continuity. The models were fitted using an ordinary least squares procedure. The estimates of the  $K$  soft indicators at each unsampled location were obtained via ordinary kriging of the neighbouring data:

$$j^*(u; z_k) = \sum_{\alpha=1}^{n(u)} \lambda_\alpha(u; z_k) j(u_\alpha; z_k), \quad \sum_{\alpha=1}^{n(u)} \lambda_\alpha(u; z_k) = 1 \quad (6)$$

The optimal kriging weights  $\lambda_\alpha(u; z_k)$  were obtained as the solution of a system of linear equations with  $[n(u)+1]$  unknowns (Goovaerts, 1997). The estimates allow the construction of a conditional cumulative distribution function (ccdf)  $F(u; z_k | (n))$  at each grid node with expected value of the ccdf (E-type estimate)  $z^*(u)$ , an estimator for catch rate, was approximated by as (Goovaerts, 1997):

$$z^*(u) \approx \sum_{k=1}^{K+1} \bar{z}_k [F(u; z_k | (n)) - F(u; z_{k-1} | (n))] \quad (7)$$

and the spread of the distribution was measured by the inter-quartile range:

$$q_R(u) = F^{-1}(u; 0.75 | (n)) - F^{-1}(u; 0.25 | (n)) \quad (8)$$

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). Linear extrapolation was used to obtain ccdf values beyond the largest threshold.

### 2.3. Model Evaluation

The performance of the spatial model was evaluated via the goodness statistic computed from leave-one-out cross-validation (Deutsch, 1996). At any location  $u_\alpha$ , a series of symmetric  $p$ -probability intervals (PI) bounded by the  $(1-p)/2$  and  $(1+p)/2$  quantiles was calculated from the known ccdf  $F(u_\alpha; z|(n))$  for all  $p \in [0,1]$ . The proportion of the true values falling into the symmetric  $p$ -PI was defined as (Deutsch, 1996):

$$\bar{\xi}(p) = \frac{1}{N} \sum_{\alpha=1}^N \xi(u_\alpha; p) \quad \text{for all } p \in [0,1] \quad (9)$$

with

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

The accuracy of the model was assessed visually using ‘‘accuracy plot’’, a scattergram of the estimated versus theoretical fractions. The closeness of the estimated and theoretical fractions was quantified by the goodness statistics ( $G$ ) defined as (Deutsch, 1996):

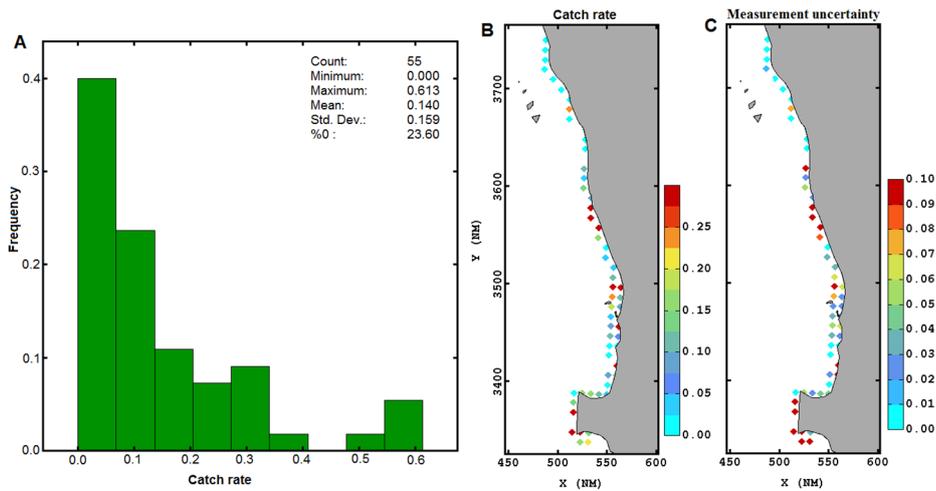
$$G = 1 - \int_0^1 [3a(p) - 2][\bar{\xi}(p) - p] dp \quad (10)$$

where  $a(p)$  is equal to 1 if  $\bar{\xi}(p) \geq p$  and to 0 otherwise. For maximum goodness,  $G = 1$  (i.e.  $\bar{\xi}(p) = p$  for all  $p \in [0,1]$ ) and  $G = 0$  when  $\bar{\xi}(p) = 0$  for all  $p \in [0,1]$ .

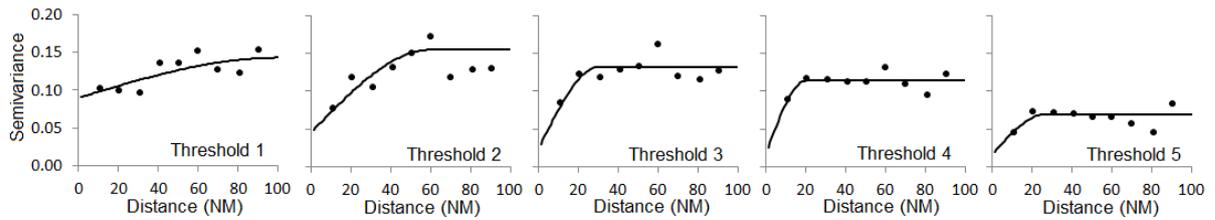
### 3. RESULTS AND DISCUSSION

The distribution of observed herring catch rates was positively skewed and ranged from 0 to 0.613 (fish per angler hour), with a mean of 0.14 (Figure 1A). The data showed spatial variation with a greater prevalence of higher catch rate estimates in the south and south-central areas and lower estimates in the north of the study area (Figure 1B). The measurement uncertainty associated with catch rate varied across the study area (Figure 1C) and showed a distribution similar to that of the observed catch rates.

The observed catch rate and associated uncertainty were transformed into soft indicator variables using equation (4) with 5 threshold values. Fewer threshold values were used due to small sample size and to ensure that the occurrence of order relation deviations was minimised (see Goovaerts, 1997 for discussion). The first threshold was set to zero to account for observed catch rates equal to zero (i.e. nil catch) and the remaining threshold values correspond to 20<sup>th</sup>, 40<sup>th</sup>, 60<sup>th</sup> and 80<sup>th</sup> percentiles of the non-zero observed catch rates. These threshold values are catch rates of 0, 0.051, 0.108, 0.166 and 0.299.



**Figure 1.** Histogram of observed catch rates (A), and spatial distribution of observed catch rate (B) and measurement uncertainty (C) across 10×10 NM locations.



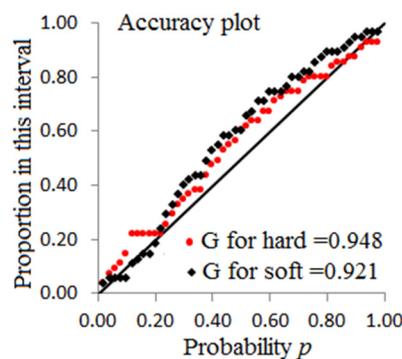
**Figure 2.** Indicator semivariogram (dots) and the fitted model (solid line).

The semivariogram models (Figure 2) of the soft indicator variables were fitted using an ordinary least squares procedure. Due to the elongated shape of the study area, an isotropic process was assumed. The range parameter representing the distance of spatial influence varied between 26 and 100 NM (Table 1). The nugget and sill parameters represent the semivariance values at a short distance and very large distance, respectively. The degree of spatial correlation, as quantified by the ratio (NSR) of the nugget to the total sill (i.e. including the nugget) (Cambardella *et al.*, 1994), varied between 0.167 and 0.629 (Table 1). The spatial autocorrelation for thresholds 3 to 5 is strong relative to thresholds 1 and 2 which show moderate spatial autocorrelation. The differences in the semivariograms models and the NSR indicate the presence of class-specific spatial patterns in the observed catch rates. Hence, using an indicator approach and different models for distinct thresholds is appropriate for catch rate data.

The accuracy plot suggests that the model is satisfactory and so suitable for prediction (Figure 3). The plot indicates that the model is accurate for probability intervals above 0.2, but less accurate below 0.2. In comparison to hard indicator kriging, the deviation from the 45° bisector line was higher for soft IK with smaller goodness statistic.

**Table 1.** Threshold values of observed catch rate and corresponding cumulative percentages with the parameters of the fitted spherical models and corresponding nugget to sill ratio (NSR).

Percentage	Threshold	nugget	Spherical model		NSR
			range	sill	
23.6	0.000	0.090	100	0.053	0.629
40.0	0.051	0.046	60	0.109	0.297
54.5	0.108	0.025	30	0.107	0.189
69.0	0.166	0.019	20	0.095	0.167
83.6	0.299	0.016	26	0.053	0.232

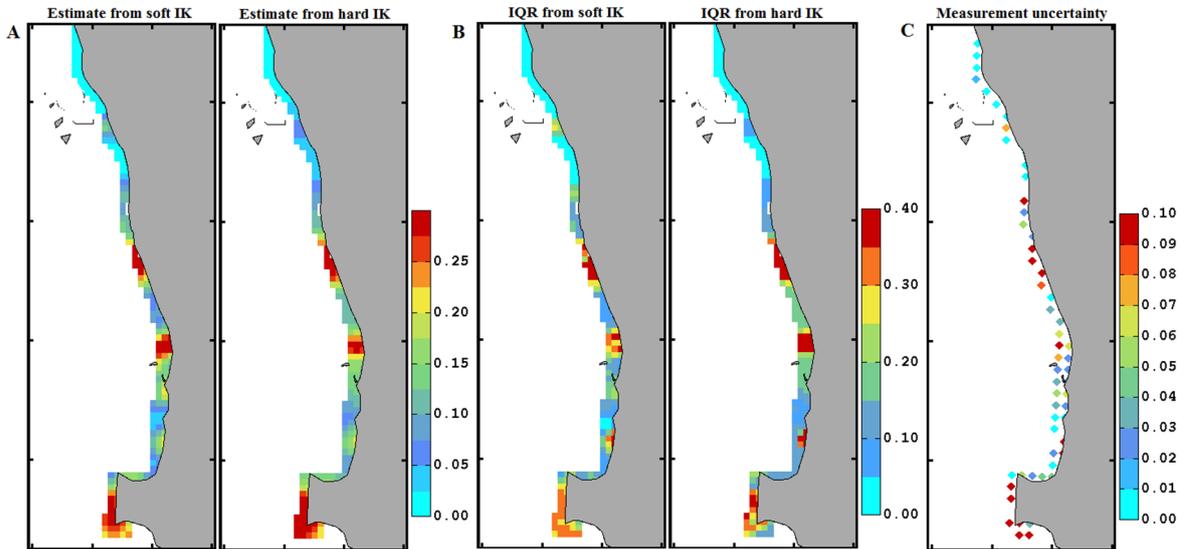


**Figure 3.** Accuracy plot obtained from cross-validation based on the soft and the traditional hard indicator kriging approaches.

### 3.1. Spatial Distribution of Estimated Catch Rates

Based on the fitted semivariogram models, ordinary kriging estimates of the indicators were obtained at nodes of size 5×5 NM. Possible order relation deviations were corrected using an upward/downward averaging method (Goovaerts, 1997). For soft and hard indicators cdfs the expected value was used as the estimator of catch rate (Figure 4). The estimates from both approaches show globally similar spatial distribution patterns but differ locally, especially in areas where measurement uncertainty is high. These differences may be attributed to the uncertainty propagated in the model. High catch rate estimates are concentrated in the south and there is some patchiness in the central of the study region. The predicted values

at the unsampled locations have lower mean and variance than the observed samples (Table 2), with hard indicator estimates having a slightly higher mean. Part of these differences may be attributed to the uncertainty propagation in the model. The uncertainty associated with the predicted catch rates was indicated by high inter-quartile range (IQR) in areas where high and low estimates of catch rates were intermingled, and the occurrence was more pronounced in the south of the study region (Figure 4B). The spatial patterns of the IQR are similar for both approaches but differ in areas where measurement uncertainty is high.



**Figure 4.** Maps of estimated catch rate (A) and inter-quartile range (B) from soft and hard indicator kriging and measurement uncertainty (C).

**Table 2.** Summary statistics of the observed catch rates and the E-type estimates from the soft indicator kriging.

Estimator	Count	Min	Max	Mean	Variance	Skewness
Observed	55	0.000	0.613	0.140	0.159	1.422
Estimate from soft IK	234	0.007	0.367	0.126	0.009	0.522
Estimate from hard IK	234	0.008	0.372	0.135	0.009	0.712

#### 4. CONCLUSIONS

Mapping the distribution of catch rates from recreational fishing is essential to fisheries managers for delineating high priority areas (Aidoo *et al.*, 2015). Incorporation of measurement uncertainty is particularly important to produce accurate and reliable results (Woiillez *et al.*, 2009). In this paper, soft indicator kriging was shown to be appropriate for incorporating measurement uncertainty into maps of estimated catch rates. This method does not require any assumption(s) regarding the measured uncertainty in contrast to other approaches (Hamzhepour *et al.*, 2013; Saito and Goovaerts, 2002). It is a promising for informing stock assessment and management, taking into account the class-specific spatial patterns and highly skewed distributions common to the catch rates from recreational fishing. The results show that inclusion of measurement error provides better discrimination in areas where uncertainty is high. For Australian herring highest catch rates occur near the metropolitan area, where the majority of the Western Australian population resides. Understanding the spatial patterns in catch rates of this species is important, as there are specific sustainability concerns related to resource sharing between commercial and recreational fishers, potential shift of fishing activity towards targeting near shore species and increasing human population growth (Smith *et al.*, 2013), and the predominance of immature fish in the catches (Smith and Brown, 2014). This study only considers catch rates for boat-based fishers, as shore-based fishers who also target the species (Smith and Brown, 2014; Smith *et al.*, 2013) were out of scope of the phone-diary survey (Ryan *et al.*, 2013), nor are seasonal differences in fishing effort accounted for (Ryan *et al.*, 2013; Smith *et al.*, 2013).

The measurement uncertainty is considered to be potentially valuable input for estimating catch rates since it can influence the estimation when ignored. In addition to measurement uncertainty, it must be noted that variogram and spatial uncertainty may be potentially valuable. Future research will be focused on the extension of the approach to other fish species in a simulation study and to assess the effect of measurement uncertainty on catch rate mapping and area delineation.

## ACKNOWLEDGEMENTS

This study would not be possible without contributions from the recreational fishers who participated in the recreational fishing surveys and the Survey Research Centre (Edith Cowan University) for facilitating data collection and entry of the phone-survey data.

## REFERENCES

- Aidoo, E. N., Mueller, U., Goovaerts, P., and Hyndes, G. A. (2015). Evaluation of geostatistical estimators and their applicability to characterise the spatial patterns of recreational fishing catch rates. *Fisheries Research*, 168, 20-32.
- Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, D. L., Turco, R. F., et al. (1994). Field-scale variability of soil properties in central Iowa soils. *Soil Science Society of America Journal*, 58, 1501-1511.
- Cochran, W. G. (1977). *Sampling Techniques* (3 ed.). New York: Wiley.
- Cooke, S. J., and Cowx, I. G. (2004). The role of recreational fishing in global fishing crises. *Bioscience*, 54(9), 857-859.
- Cooke, S. J., and Cowx, I. G. (2006). Contrasting recreational and commercial fishing: searching for common issues to promote unified conservation of fisheries resources and aquatic environments. *Biological Conservation*, 128(1), 93-108.
- Deutsch, C. V. (1996). *Direct assessment of local accuracy and precision*. Paper presented at the 5th International Geostatistics Congress, Wollongong' 96.
- Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. New York: Oxford University Press.
- Hamzehpour, N., Eghbal, M. K., Bogaert, P., Toomanian, N., and Sokouti, R. S. (2013). Spatal prediction of soil salinity using kriging with measurement errors and probabilistic soft data. *Arid Land Research and Management*, 27, 128-139.
- Jones, C. M., Robson, D. S., Lakkis, H. D., and Kressel, J. (1995). Properties of catch rate used in analysis of angular survey. *Transactions of the American Fisheries Society*, 124, 911-928.
- National Research Council. (1998). *Review of Northeast fishery stock assessments*. Washinton, DC: National Academy Press.
- Ryan, K. L., Wise, B. S., Hall, N. G., Pollock, K. H., Sulin, E. H., and Gaughan, D. J. (2013). *An integrated system to survey boat-based recreational fishing in Western Australia 2011/2012*. Fisheries Research Report No. 249, Department of Fisheries, Western Australia.
- Saito, H., and Goovaerts, P. (2002). Accounting for measurement error in uncertainty modeling and decision-making using indicator kriging and p-field simulation: application to a dioxin contaminated site. *Environmetrics*, 13, 555-567.
- Smith, K., and Brown, J. (2014). *Biological synopsis of Australian herring (Arripis georgianus)*. Fisheries Research Report No. 251. Department of Fisheries, Western Australia.
- Smith, K., Brown, J., Lewis, P., Dowling, C., Howard, A., Lenanton, R., et al. (2013). *Status of nearshore finfish stocks in south-western Western Australia Part 1: Australian herring*. Fisheries Research Report No. 246. Department of Fisheries, Western Australia.
- Sumner, N. R., and Williamson, P. C. (1999). A 12-month survey of coastal recreational boat fishing between Augusta and Kalbarri on the west coast of Western Australia during 1996-97. *Fisheries Research Report No. 117, Department of Fisheries, Western Australia*, 52 p.
- Woillez, M., Rivoirard, J., and Fernandes, P. G. (2009). Evaluating the uncertainty of abundance estimates from acoustic surveys using geostatistical simulations. *ICES Journal of Marine Science*, 66(6), 1377-1383.