Geostatistical tools for assessing sampling designs applied to a Portuguese bottom trawl survey field experience

ERNESTO JARDIM [1] and PAULO J. RIBEIRO JR. [2]

[1] Instituto Nacional de Investigação Agrária e das Pescas, Av. Brasilia, 1449-006, Lisboa, Portugal.
E-mail: ernesto@ipimar.pt
[2] Laboratório de Estatística e Geoinformação, Universidade Federal do Paraná, C.P. 19.081 CEP: 81.531-990, Curitiba, Paraná, Brasil.

SUMMARY: We present a bottom trawl survey (BTS) field experience carried out off the Portuguese Continental shelf to test two sampling designs proposals previously analysed by simulation which implement a hybrid random-systematic and a systematic sampling strategy. We used a common base regular grid covering the survey area and overlapped it with the existent random design to build the hybrid design while the systematic design added a set of regular locations at smaller distances creating four denser sampling areas. We use hake (*Merluccius merluccius*) abundance and model-based geostatistics to compute measures such as mean abundance, μ, and the 95% percentile, p_{95}, which summarise the areal behaviour; coverage of the prediction confidence interval, ξ, to assess the adequacy of the model; and a modified generalised cross validation index, π, to evaluate prediction precision. The hybrid design showed a lower coefficient of variation for μ (11.89% against 13.25%); a slightly higher coefficient of variation for p_{95} (11.31% against 11.09%); similar ξ (0.94); and lower π (16.32 against 18.82). We conclude that the hybrid design performs better, our procedure for building it can be used to adjust BTS designs to modern geostatistical techniques, and the statistics used constitute valuable tools for assessing BTS performance.

Keywords: model-based geostatistics, hake, sampling design, bottom trawl survey.

RESUMEN: Uso de herramientas geostadísticas para evaluar diseños de muestreo aplicado a una experiencia de campo de campaña de arrastre demersal en Portugal. – Presentamos una experiencia de campo mediante campañas de arrastre demersal (CAD) llevada a cabo en la plataforma continental portuguesa con el fin de evaluar dos propuestas de diseño de muestreo. Las dos propuestas se analizaron previamente mediante técnicas de simulación implementando dos estrategias de muestreo: estrategia híbrida aleatoria-sistemática y estrategia sistemática. Se utilizó una rejilla de celdas regular común cubriendo el área de estudio y se superpuso con el diseño aleatorio existente para construir el diseño híbrido, mientras que el diseño sistemático añade un conjunto de puntos de muestreo regulares a menor distancia para crear cuatro áreas de muestreo más densas. Se usa la abundancia de merluza (*Merluccius merluccius*) y técnicas geostadísticas basadas en modelo para calcular medidas como: abundancia media, μ, y el percentil 95%, p_{95}, que resumen el comportamiento areal; la cobertura del intervalo de confianza de predicción, ξ, para evaluar la bondad del modelo; y un índice modificado de validación cruzada generalizada, π, para evaluar la precisión de la predicción. Se obtuvo un coeficiente de variación menor para μ con el diseño híbrido (11.89% frente a 13.25%); un coeficiente de variación ligeramente más alto para p_{95} (11.31% frente a 11.09%); una ξ similar (0.94); y menor π (16.32 contra 18.82). se concluye que el diseño híbrido funciona mejor y nuestro procedimiento para construirlo puede ser usado para ajustar diseños CAD a técnicas geostadísticas modernas, y que los estadísticos usados constituyen herramientas valiosas para evaluar el rendimiento de CAD.

Palabras clave: geostadística basada en modelo, merluza, diseño de muestreo, campaña de arrastre demersal.

INTRODUCTION

Designs for bottom trawl surveys (BTS) rely on previous knowledge of the spatial distribution and population structure of the target species combined with statistical analysis of preliminary data (e.g. Ault *et al.*, 1999; Hata and Berkson, 2004) or simulation procedures (e.g. Schnute and Haigh, 2003; Anon., 2005b). In defining the BTS sampling design, one is confronted with operational constraints such as the existence of trawlable grounds and vessel time available. The survey design is typically reviewed from time to time to adjust the stratification (e.g. Smith and Gavaris, 1993; Folmer and Pennington, 2000), tow duration (e.g. Cerviño and Saborido-Rey, 2006; Wieland and Storr-Paulsen, 2006), technical issues such as gear changes (e.g. Zimmermann *et al.*, 2003; Cooper *et al.*, 2004) and other factors which may change over the years.

Several authors have discussed the advantages of systematic designs over random designs for sampling spatial correlated variables such as fish abundance (Cochran, 1960; Ripley, 1981; Thompson, 1992; Cressie, 1993; Chiles and Delfiner, 1999; Kimura and Somerton, 2006; Diggle and Ribeiro Jr, 2007). Nevertheless, in the case of spatial correlated variables there are two conflicting objectives that cannot be combined in a single criterion: estimation of the covariance function parameters and prediction (Müller, 2001). In the first situation it is important to have locations at short distances to inspect the behaviour of the correlation function close to the origin, and locations at distances close to the limit of spatial correlation to estimate the correlation range (Müller, 2001). In the second situation the best predictions will result from the design with the highest covariance with the locations to be predicted (Thompson, 1992). In the case of predicting fish abundance it is common to require a complete map of the study area and the best choice will be a design that covers the area evenly. However, when the covariance function is unknown (a common characteristic of fish abundance analysis), it must be estimated from the data before predicting and the two objectives must be combined. Several authors suggest designs that mix a set of locations covering the area with additional locations at short distances (Müller, 2001; Diggle and Lophaven, 2006; Zhu and Stein, 2006) to balance the two objectives. Such designs applied to bottom trawl surveys have received limited attention (Selzenmuller *et al.*, 2005), although fish abundance

characteristics fit well in the assumptions of these proposals.

Our analysis adopts a model-based geostatistical method (Diggle *et al.*, 1998; Diggle and Ribeiro Jr, 2007) to explicitly take into account spatial patterns of abundance and provide a flexible modelling framework. It is worth mention previous work by Simard *et al.* (1992) and Petitgas (2001) wich also addressed the issue of sampling design, although not using model-based geostatistcs.

The designs are accessed by a set of statistics to provide information about aspects of the data that are relevant for modelling fish abundance. In a global perspective, referring to the entire study region, we use mean abundance and the 95% percentile to summarise the areal abundance, commonly used for studying time trends and building abundance indices for stock assessment. In a local perspective, referring to particular locations within the study area, we use the observed values to assess the suitability of the model by computing the coverage of the prediction confidence interval, and the prediction precision by computing a modified generalised cross-validation index. Note that the assessment of the model suitability and the prediction precision are extremely valuable statistics, once that kriging is a linear predictor and the maps produced with it will be used to estimate the spatial distribution of abundance and the abundance index mentioned above. With relation to the analysis reported here we rely on our experience with bottom trawl surveys (Anon., 2002, 2003, 2004, 2005a, 2006; Sousa *et al.*, 2005; Mendes *et al.*, 2007; Sousa *et al.*, 2007) to provide contextual information which supports the adoption of a particular class of model, and avoid model mis-specification as far as possible.

The work described on this paper aims to: (i) report a BTS field experience for testing sampling designs, and (ii) describe geostatistical tools for assessing the performance of sampling designs.

MATERIAL

The Portuguese BTS started in June 1979, covering the continental shelf and following a stratified random design. In 1989 the stratification was defined by 12 sectors along the coast subdivided into 4 depth ranges: 20-100 m, 101-200 m, 201-500 m and 501-750 m, with a total of 48 strata. Due to constraints in the vessel time available the sample size was set

to 97 locations evenly allocated to each stratum. The coordinates of the sampling locations were selected randomly, albeit constrained by the historical records of clear tow positions and other information about the sea floor, thus avoiding places where trawling was not possible. During this period haul duration was set to one hour but recent experiments proved that half-hour hauls provide the same information about length distributions (Cardador, pers. comm.). In light of this findings haul duration was reduced to half an hour and an additional set of hauls were available which led to a revision of the sampling design. The revision was split into a preliminary phase using simulations and geostatistical analysis (Jardim and Ribeiro Jr, 2007) and a second phase during which a field test was performed to provide real information about the proposed sampling designs. In the third phases the decision will be based on the scientific data provided and the existing financial and administrative constraints.

The field experience was carried out during the summer of 2001, with R/V Noruega off the southwest of the Portuguese Continental shelf (Fig. 1) using a Norwegian Campbell Trawl 1800/96 (NCT) with a codend of 20 mm, a mean vertical opening of 4.8 m and a mean horizontal opening between wings of 15.5 m. The survey performed two sampling designs selected from the simulation study reported by Jardim and Ribeiro Jr (2007). The survey area was limited in the south by the cape of S.Vicente (37.00º north), in the north by Setubal's Canyon (38.30º north), in the east by the 20 m depth isoline and in the west by the 500 m isoline. The survey area was approximately 4300 km^2 in size and the maximum distance within the area was approximately 150 km. The data collected in both designs and considered here consisted of date/time, geographical location and hake (*Merluccius merluccius*) catch in weight (kg). Geographical coordinates were transformed into UTM units and hake abundance was computed in kg/km and assigned to the haul starting coordinates. The area swept was computed using the haul start and ending positions to correct haul speed variations.

METHODS

This section describes the sampling designs to be tested and how they were built. It also describes the geostatistical modelling framework and the ad-
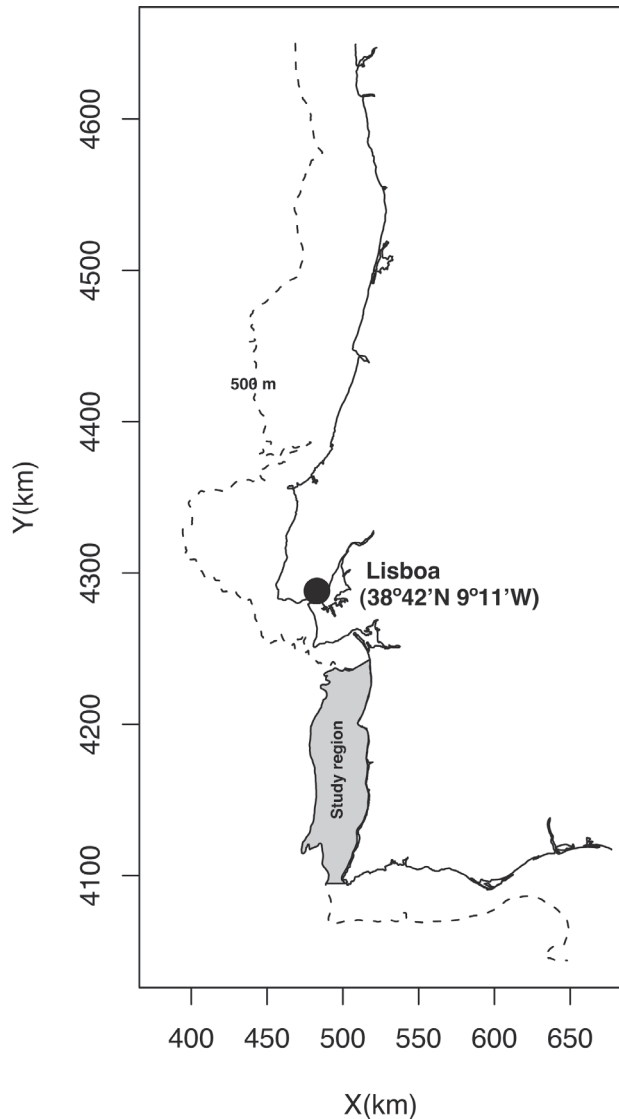

Fig. 1. – Survey area on the southwest of the Portuguese Continental shelf between 20 m and 500 m.

justments considered to cope with the small dataset available—a common characteristic of BTS due to its high price. Finally, we describe the technical details of the performance statistics chosen.

Sampling designs

Our sampling designs were built by mixing a set of operational constraints with the geostatistical principles described above and the need to keep the continuity of the survey history. In particular, the two designs tested were built to distinguish between a hybrid random-systematic sampling strategy and a systematic strategy.

The sampling effort available for the candidate design was 36 locations. We built two candidate de-
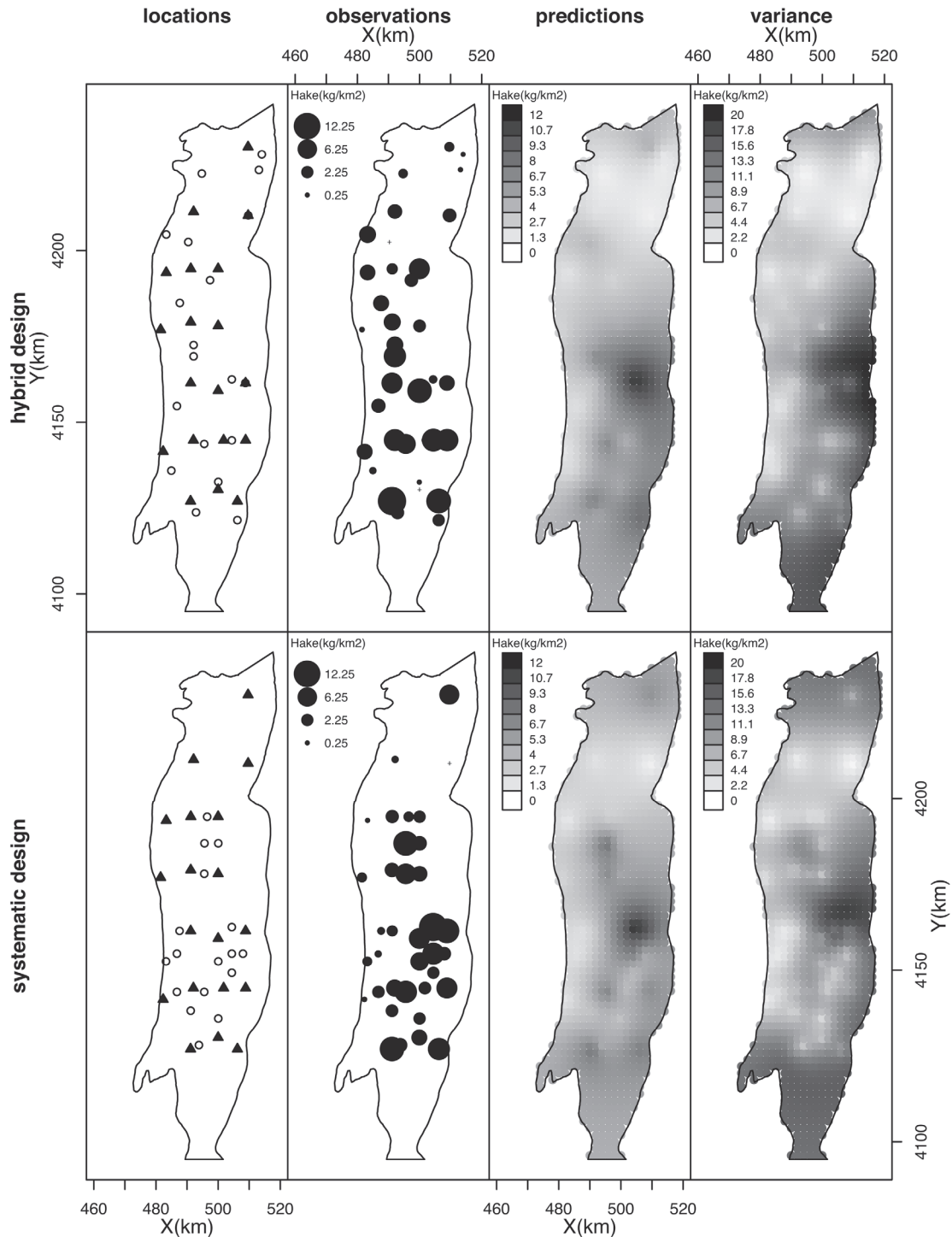
FIG. 2. – Study area on the Portuguese southwest coast. The top panels show information about the hybrid random-systematic design and the bottom panels about the systematic design. The leftmost plots show the sampling designs locations, the black triangles represent the regular grid common to both designs, and the open circles the additional locations. Follows the observations of hake abundance (kg/km²) and the predictions obtained by kriging, both on the square root scale. The rightmost plots present the kriging variance

signs using as a basis a regular grid with 19 locations covering the survey area, and added seventeen additional ones at shorter distances. Two candidate designs were built, a *hybrid design* that allocated the additional locations randomly and a *systematic*

design that allocated them at regular locations. The hybrid design overlaps the regular grid with the existent random design, keeping some continuity with the survey historical records (top-left plot in Fig. 2). The systematic design includes regular locations at

smaller distances, creating 4 denser sampling areas (bottom-left plot in Fig. 2). The vessel time available did not allow other possibilities to be tested.

Geostatistical model

Geostatistical observations consist of pairs (x,y) with elements $(x_i,y_i) : i = 1,...,n$, where x_i denotes the coordinates of each of the n spatial locations within a study region $A \subset R^2$ and y_i the measurement of the corresponding observable study variable. We adopted the Box-Cox transformed Gaussian model with transformation parameter λ as presented in Christensen *et al.* (2001). Denoting by z_i the transformed values, such that $g_\lambda(y_i) = z_i$, the model for the vector of variables Z observed at locations x can be written as a linear model $Z(x) = S(x) + \varepsilon$, where S is a stationary Gaussian stochastic process, with $E[S(x)] = \mu$, $Var[S(x)] = \sigma^2$ and an isotropic correlation function $\rho(h) = Corr[S(x),S(x')]$, where $h = \|x\text{-}x'\|$ is the Euclidean distance between locations x and x'. The terms ε are assumed to be mutually independent and identically distributed, $\varepsilon \sim Gau(0,\tau^2)$. For the correlation function $\rho(h)$ we adopt the exponential function with algebraic form $\rho(h) = \exp\{-h/\phi\}$ where ϕ is the *range* parameter such that $\rho(h) \simeq 0.05$ when $h = 3\phi$. Following usual geostatistical terminology (Isaaks and Srivastava, 1989) we call $\sigma_\tau^2 = \tau^2 + \sigma^2$ the total sill, σ^2 the partial sill, τ^2 the nugget effect and 3ϕ the practical range. Geometric anisotropy (Isaaks and Srivastava, 1989; Cressie, 1993) is considered an extension of this model with an extra parameter $\psi = \{\psi_A,\psi_R\}$, where ψ_A is the anisotropic angle and ψ_R is the anisotropic ratio.

Hereafter we use $[\cdot]$ to denote *the distribution of* the quantity indicated within brackets. Following the adopted model, $[g_\lambda(Y)] \sim MVGau(\mu 1,\Sigma)$, i.e. $[Y]$ is multivariate trans-Gaussian with expected value μ and covariance matrix Σ parametrised by $\{\sigma^2,\phi,\tau^2\}$. Parameter estimates can be obtained by maximum likelihood (Cressie, 1993; Diggle *et al.*, 1998; Diggle and Ribeiro Jr, 2007) and used for spatial prediction. In its simplest format, spatial prediction given by the *kriging predictor* consists in obtaining expected values and associated variances at unsampled locations. More generally, the *predictive distribution* of quantities of interest can be obtained analytically, if possible, or by sampling from this distribution. Considering a prediction target $T(x_0) = g_\lambda^{-1}(S(x_0))$, the realised value of the process in the original measurement scale at spatial locations x_0. Simulations from the conditional distribution $[T(x_0)|Y(x)]$ are obtained by simulating from the multivariate Gaussian $[S(x_0)|Y(x)]$ and back transforming the simulated values to the original scale of measurement (Chiles and Delfiner, 1999; Diggle and Ribeiro Jr, 2007). These simulations are called *conditional simulations* referring to the fact they are obtained from the distribution of the quantity of interest conditioned to the observed values $Y(x)$.

We split inference into two steps. First the Box-Cox transformation parameter λ and the anisotropy parameter ψ_R are investigated by pooling all the observations in a single dataset and computing profile likelihoods (Diggle and Ribeiro Jr, 2007). We consider the north-south coastal orientation of the study region as the direction of greatest spatial continuity and fix ψ_A in 0 degrees azimuthal angle. Afterward, having estimated these two parameters we regard their point estimates as constants in the model and proceed by computing, for each design, the maximum likelihood estimates of the remaining model parameters. The reasoning for the two-step procedures is twofold. Pragmatically, it overcomes the difficulty of identifying all parameters with a small dataset. With regards to modelling assumptions it considers the transformation and anisotropy parameters as part of the model specification reflecting the nature of the data and contextual information, therefore not requiring to be identified by the designs. Thereafter, we compute kriging predictions on a 2 × 2 km grid within the study area, x_0, with a total of 1070 locations, and obtain 1,000 conditional simulations from $[Y(x_0)|Y]$ for each design.

Performance statistics

Consider $E[Z(x_i)]$ and $\sigma_z^2(x_i)$ the kriging predictor and its variance on the Gaussian scale at location $x_i \in x_0$. For a value of 0.5 of the transformation parameter λ, the back transformation to the original scale gives $E[Y(x_i)] = (1 + 0.5E[Z(x_i)])^2 + 0.25\sigma_z^2(x_i)$ and the global mean is estimated by averaging the predicted values $\hat{\mu} = m^{-1}\sum_{i=0}^{m}\hat{E}[Y(x_i)]$. The variance of $\hat{\mu}$, denoted by $\hat{\sigma}_\mu^2$, is computed by the mean of all terms in the covariance matrix $\Sigma_Y(x_0) = Var[Y(x_0)|Y(x)]$, back transformed by $\Sigma_Y(x_0) = \Sigma_Z(x_0)[8^{-1}\Sigma_Z(x_0) + (1 + 0.5E[Z(x)])^2]$, where $\Sigma_Z(x_0)$ is the covariance matrix of $[S(x_0)|Z(x)]$. More generally, inferences on other quantities of interest $T(x_0)$ are obtained from the conditional simulations. Denote by $t_s(x_0)$, $s = 1,...,S = 1,000$ conditional simulations from $[T(x_0)|Y(x)]$. For

example, an α-th percentile is estimated by $\hat{p} = S^{-1}\Sigma_s$ $\hat{p}_s$ where $\hat{p}_s = p_\alpha(t_s(x_0))$, the average of the empirical distribution $\hat{p}$ obtained from the conditional simulations. The variance of $\hat{p}$ is given by $\hat{\sigma}_p^2 = (S-1)^{-1}$ $\Sigma_s(\hat{p}_s - \hat{p})^2$.

The coverage of the prediction confidence interval, ξ, and the generalised cross validation index, π, were computed using cross-validation statistics (Hastie *et al.*, 2001) combined with conditional simulations as follows. First, create a new data set by leaving one observation out at a location x_i, simulate 1,000 values of the variable at that location, and repeat this procedure visiting all data locations. Subsequently, consider $y(x_i)$ an observation of the process Y at location x_i, $i = 1,...,n$; $y(x_{(i)})$ the observed data set without the observation $y(x_i)$ and $t_s(x_i)$ a conditional simulation $s = 1,...,S$ of $[T(x_i)|Y = y(x_{(i)})]$ at location x_i. The predictive confidence interval is given by $CI(x_i) = [p_{2.5}(t_s(x_i)), p_{97.5}(t_s(x_i))]$ and the proportion of observations lying inside the intervals $\xi = n^{-1} \Sigma_i(y(x_i) \in CI(x_i))$ provides the *coverage* of the prediction confidence interval. The cross validation index is given by $\pi = n^{-1} \Sigma_i(S^{-1} \Sigma_s(t_s(x_i) - y(x_i))^2)$, the average of the mean quadratic error at each location estimated using the full set of conditional simulations.

RESULTS

The two sampling designs and the observations of hake abundance are presented in the leftmost panels of Figure 2, where the base regular design is represented by the black triangles. The abundance of hake observed showed that the distribution of abundance was spread over the area, with lower values in the north and a small number of zeros.

The 95% confidence interval obtained for the Box-Cox transformation parameter was [0.12, 0.55] and we set $\hat{\lambda} = 0.5$, corresponding to a square root transformation. The profiled log-likelihood of the anisotropy ratio showed no evidence of anisotropy. Nevertheless, we carried out analysis using different values of ψ_P to check the sensitivity of the results, which proved negligible.

Covariance parameter estimates showed higher values for the hybrid design than the corresponding ones given by the systematic design (Table 1). The total variance $\hat{\sigma}_T^2$ was 3.75, with $\hat{\sigma}^2 = 0.75$ and $\hat{\tau}^2 = 3.00$; and $\hat{\phi} = 16.64$. The systematic design estimates were $\hat{\sigma}_T^2 = 3.20$, with $\hat{\sigma}^2 = 0.61$ and $\hat{\tau}^2 = 2.59$; and $\hat{\phi} = 10.21$. Considering $\hat{\tau}^2_{REL}$, which computes

the relationship between the random variability and the spatial structure variability, and $\sigma^2\phi^{-1}$, which give information about the "size" of the spatial process. The two designs showed similar relative nuggets. However, the hybrid design showed a lower ratio between sill and range, reflecting a higher spatial structure of the stochastic process. Notice that the practical range, 3ϕ, was ≈ 50 km for the hybrid and ≈ 30 km for the systematic design.

The rightmost panels of Figure 2 show the abundance maps predicted and their variance for each design. Both predictions are similar and the spatial pattern of variance reflects the influence of the observations, showing lower variability near the observed locations and higher variability in areas where extrapolation was further extended. The hybrid design had higher variance in the centre-east of the study area and lower variance in the north due to a better coverage in this area.

The estimates of μ and p_{95} were similar, although the hybrid design showed slightly lower values. The hybrid design showed a lower coefficient of variation for μ, $CV_\mu = 11.89\%$, than the systematic design, $CV_\mu = 13.25\%$. Sampling statistics computed for these designs showed a similar pattern (Table 1). The p_{95} variance was slightly lower for the systematic design, $CV_{p95} = 11.09\%$, than for the hybrid design, $CV_{p95} = 11.31\%$. The coverage of the prediction confidence intervals was 0.94 for both designs. These results reinforce our modelling choices given that if the model was wrong we would expect ξ to be different from the nominal value of the confidence interval. The generalised cross validation index showed a lower estimate with the hybrid design, 16.32, than with the systematic design, 18.82, indicating a higher prediction precision of the hybrid design. The above results show that the higher spatial structure of the stochastic process estimated for the hybrid design exceeded its higher total variability with relation to the estimation of these performance statistics.

DISCUSSION

Assessing sampling designs for BTS raises interesting questions about appropriate methodologies for analysing data and obtaining statistics of interest, which are particularly relevant considering the multipurpose/multispecies nature of BTS and the small sample sizes.

TABLE 1. – Sampling statistics, estimates of model parameters and performance statistics by design. Sampling statistics are: n, the sample size; $\bar{Y}$, the sampling mean; $s_{\bar{Y}}^2$, the variance of the sampling mean. Model parameters are: τ^2, the short distance variance or nugget effect; σ^2 the variance of the spatial process; σ_T^2 the total variance; ϕ the correlation range parameter; and the transformation parameters λ, the Box-Cox parameter and the anisotropy parameters $\{\psi_A, \psi_R\}$. The relative nugget, τ_{REL}^2, and the ratio between relative sill and range $\sigma^2\phi^{-1}$, were computed to give more insights about the spatial process. Performance statistics are: $\hat{\mu}$ and $\hat{\sigma}_\mu^2$, the mean and variance of the global abundance; $\hat{p}_{95}$ and $\hat{\sigma}_p^2$, the mean and variance of the 95th percentile of the global abundance; π, the generalised cross validation index and ξ, the coverage of the prediction confidence interval with nominal level of 0.95.

	Hybrid	systematic
sampling statistics		
n	36	36
$\bar{Y}$	4.21	4.41
$s_{\bar{Y}}^2$	0.35	0.35
cv	14.04	13.41
model parameters		
τ^2	0.75	0.61
σ^2	3.00	2.59
σ_T^2	3.75	3.20
ϕ	16.64	10.21
τ_{REL}^2	0.20	0.19
$\sigma^2\phi^{-1}$	0.18	0.25
ψ_A	0.00	0.00
ψ_R	1.00	1.00
λ	0.5	0.5
performance statistics		
$\hat{\mu}$	4.07	4.20
$\hat{\sigma}_\mu^2$	0.23	0.31
cv	11.89	13.25
$\hat{p}_{95}$	11.01	10.78
$\hat{\sigma}_p^2$	1.55	1.43
cv	11.31	11.09
ξ	0.94	0.94
π	16.32	18.82

The adoption of a formal criteria and loss function to find an optimum design seems unrealistic in practice due to the multidimensionality of the data and the conflicting objectives of inference and prediction. Here we followed a pragmatic approach to sampling design, starting with a design that joins a regular grid with the old random design, following with a design that uses the same regular grid but reallocates the random locations in a regular shape. Considering the wide literature that supports the use of systematic designs for spatial correlated variables, these designs implement the two most promising strategies and test the possibility of keeping the continuity with the historical time series. To compare these proposals we rely on spatial modelling to compute statistics of primary interest and look for consistency among them, exploring several aspects of the same dataset. We advocate that the approach described above will provide valuable information to support the decision-making process.

Although the results obtained are constrained by the characteristics of the area and the species analysed, we believe that the methodology defined by our approach can be applied to other areas and species, providing an important source of information for revising sampling design. It would not be surprising if similar results are found for other species, because the principles behind the construction of the sampling designs tested are quite generic and can be applied to most fish species.

The performance statistics were selected to reflect relevant characteristics and different aspects of spatial prediction. The global mean is the most widely used index of abundance, often estimated by the sample average. We favour the geostatistical estimator presented and its variance as a measure of uncertainty, because it takes into account the spatial dependency within the area and insights about the spatial process. The 95th percentile estimated by conditional simulations can be used to identify areas of high abundance, giving information about candidate areas to protect. The coverage of the prediction confidence intervals is a diagnostic tool. A small coverage reflects an underestimation of the variance or the inadequacy of the model to explain the available data. The cross-validation index combined with conditional simulations incorporates the prediction precision in the index, which is not taken into account by the traditional cross-validation. For example, if a location has the same value predicted by different designs but with different prediction variances, this index will distinguish the two situations.

Our results showed that the hybrid design performed better in all cases except for σ_p^2. A clear parallel can be established with the *lattice plus closed pairs* designs of Diggle and Lophaven (2006), the *EK-optimal* designs of Zimmerman (2006) and the D_{EA} designs of Zhu and Stein (2006). Although following different construction, these designs cover the entire study areas in study, include a set of positions at small distance and performed better than their random or systematic competitors. Common to all these studies and our work is the fact that the analyses were carried out in situations in which the model parameters were considered unknown and needed to be estimated from the data, which made it clear that both parameter estimation and prediction are important for the precision of the prediction target.

In conclusion, we consider that our results indicate that keeping the old random design and adding

a regular grid to build a new design can be a good and pragmatic solution for adjusting BTS designs to modern geostatistical techniques. Secondly, the performance statistics described above seem to capture the most important features of the data with relation to abundance estimation, constituting good measures for assessing BTS performance.

ACKNOWLEDGEMENTS

REFERENCES

Anonymous. – 2002. Report of the International Bottom Trawl Survey Working Group. *ICES CM 2002/D:03.*

Anonymous. – 2003. Report of the International Bottom Trawl Survey Working Group. *ICES CM 2003/D:05.*

Anonymous. – 2004. Report of the International Bottom Trawl Survey Working Group. *ICES CM 2004/D:05.*

Anonymous. – 2005a. Report of the International Bottom Trawl Survey Working Group. *ICES CM 2005/D:05.*

Anonymous. – 2005b. Report of the Workshop on Survey Design and Data Analysis. *ICES CM 2005/ACFM:11.*

Anonymous. – 2006. Report of the International Bottom Trawl Survey Working Group. *ICES CM 2006/D:05.*

Ault, J., G. Diaz, S. Smith, J. Luo and J. Serafy. – 1999. An Efficient Sampling Survey Design to Estimate Pink Shrimp Population Abundance in Biscayne Bay, Florida. *North. Am. J. Fish. Manage.*, 19: 696-712.

Cerviño, S. and F. Saborido-Rey. – 2006. Using bootstrap to investigate the effects of varying tow lengths and catch sampling schemes in fish survey. *Fish. Res.*, 79: 294-302.

Chiles, J.P. and P. Delfiner. – 1999. *Geostatistics: Modeling Spatial Uncertainty.* Wiley, New York.

Christensen, O., P. Diggle and P. Ribeiro Jr. – 2001. Analysing positive-valued spatial data: the transformed Gaussian model. In: P. Monestiez, D. Allard and Froidevaux (eds.), *GeoENV III - Geostatistics for Environmental Applications*, pp. 287-298.

Cochran, W. – 1960. *Sampling Techniques.* John Wiley and Sons, New York.

Cooper, A., A. Rosenberg, B. Stefánson and M. Mangel. – 2004. Examining the importance of consistency in multi-vessel trawl survey design based on the U.S. west coast groundfish bottom trawl survey. *Fish. Res.*, 70: 239-250.

Cressie, N. – 1993. *Statistics for spatial data*. John Wiley and Sons, New York.

Diggle, P. and S. Lophaven. – 2006. Bayesian geostatistical design. *Scand. J. Stat.*, 33: 55-64.

Diggle, P. and P. Ribeiro Jr. – 2007. *Model-based Geostatistics.* Springer, New York.

Diggle, P., J. Tawn and R. Moyeed. – 1998. Model-based geosta-

tistics (with discussion). *J. R. Stat. Soc. Ser. C-Appl. Stat.*, 47: 299-350.

Folmer, O. and M. Pennington. – 2000. A statistical evaluation of the design and precision of the shrimp trawl survey off West Greenland. *Fish. Res.*, 49: 165-178.

Hastie, T., R. Tibshirani and J. Friedman. – 2001. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction.* Springer, New York.

Hata, D. and J. Berkson. – 2004. Factors Affecting Horseshoe Crab *Limulus polyphemus* Trawl Survey Design. *Trans. Am. Fish. Soc.*, 133: 292-299.

Isaaks, E. and M. Srivastava. – 1989. *An Introduction to Applied Geostatistics.* Oxford University Press, New York.

Jardim, E. and P. Ribeiro Jr. – 2007. Geostatistical Assessment of Sampling Designs for Portuguese Bottom Trawl Surveys. *Fish. Res.*, 85: 239-247.

Kimura, D.K. and D.A. Somerton. – 2006. Review of Statistical Aspects of Survey Sampling for Marine Fisheries. *Rev. Fish. Sci.*, 14, 3: 245-283.

Mendes, J.M., K.F. Turkman and E. Jardim. – 2007. A Bayesian hierarchical model for over-dispersed count data: a case study for abundance of hake recruits. *Environmetrics*, 18, 1: 27-53.

Müller, W. – 2001. *Collecting Spatial Data - Optimum Design of Experiments for Random Fields.* Physica-Verlag, Heidelberg, 2nd ed.

Petitgas, P. – 2001. Geostatistics in fisheries survey design and stock assessment models, variance and applications. *Fish Fish.*, 2: 231-249.

Ripley, B. – 1981. *Spatial Statistics.* Wiley, New Jersey.

Schnute, J. and R. Haigh. – 2003. A simulation model for designing groundfish trawl surveys. *Can. J. Fish. Aquat. Sci.*, 60: 640-656.

Simard, Y., P. Legendre, G. Lavoie and D. Marcotte. – 1992. Mapping, Estimating Biomass, and Optimizing Sampling Programs for Spatially Autocorrelated Data: Case Study of the Northern Shrimp (*Pandalus borealis*). *Can. J. Fish. Aquat. Sci.*, 49: 32-45.

Smith, S. and S. Gavaris. – 1993. Improving the Precision of Abundance Estimates of Eastern Scotian Shelf Atlantic Cod from Bottom Trawl Surveys. *North Am. J. Fish. Mange.*, 13: 35-47.

Sousa, P., M. Azevedo and M. Gomes. – 2005. Demersal assemblages off Portugal: Mapping, seasonal, and temporal patterns. *Fish. Res.*, 75: 120-137.

Sousa, P., R. Lemos, M. Gomes and M. Azevedo. – 2007. Analysis of horse mackerel, blue whiting, and hake catch data from Portuguese surveys (1989-1999) using an integrated GLM approach. *Aquat. Living Resour.*, 20: 105-116.

Stelzenmüller, V., S. Ehrich and G. Zauke. – 2005. Impact of additional small-scale survey data on the geostatistical analyses of demersal fish species in the North Sea. *Sci. Mar.*, 69: 587-602.

Thompson, S.K. – 1992. *Sampling.* Wiley, New York.

Wieland, K. and M. Storr-Paulsen. – 2006. Effect of tow duration on catch rate and size composition of Northern shrimp (*Pandallus borealis*) and Greenland halibut (*Reinhardtius hippoglossoides*) in the West Greenland Bottom Trawl Survey. *Fish. Res.*, 78: 276-285.

Zhu, Z. and M. Stein. – 2006. Spatial Sampling Design for Prediction With Estimated Parameters. *J. Agric. Biol. Environ. Stat.*, 11: 24-44.

Zimmerman, D. – 2006. Optimal network design for spatial prediction, covariance parameter estimation, and empirical prediction. *Environmetrics*, 17: 635-652.

Zimmermann, M., M. Wilkins, K. Weinberg, R. Lauth and F. Shaw. – 2003. Influence of improved performance monitoring on the consistency of a bottom trawl survey. *ICES J. Mar. Sci.*, 60: 818-826.