Elsevier Editorial System(tm) for Fisheries Research

Manuscript Draft

Manuscript Number: FISH896R1

Title: Geostatistical Assessment of Sampling Designs for Portuguese Bottom Trawl Surveys

Article Type: Research Paper

Keywords: bottom trawl surveys; geostatistics; simulation; hake; horse mackerel; sampling design

Corresponding Author: Dr. Ernesto Jardim, MSc

Corresponding Author's Institution: IPIMAR - Instituto Nacional de Investigação Agrária e das Pescas

First Author: Ernesto Jardim, MSc

Order of Authors: Ernesto Jardim, MSc; Paulo Ribeiro, Jr, PhD

Manuscript Region of Origin: PORTUGAL

Geostatistical Assessment of Sampling Designs for Portuguese Bottom Trawl Surveys

Ernesto Jardim[1] <ernesto@ipimar.pt> and Paulo J. Ribeiro Jr[2] <paulojus@ufpr.br>

8th January 2007

[1]Instituto Nacional de Investigação Agrária e das Pescas, Av. Brasilia, 1449-006, Lisboa, Portugal
Tel: +351 213 027 093

[2]Departamento de Estatística, Universidade Federal do Paraná, C.P. 19.081 CEP: 81.531-990, Curitiba, Paraná, Brasil

Abstract

New sampling designs for the Autumn Portuguese bottom trawl survey (ptBTS) were investigated to explore alternative spatial configurations and possible increments on sample size. The currently used stratified random design and five proposals of systematic based designs were assessed by a simulation study, adopting a geostatistical approach based on likelihood methods of inference. The construction of the designs was based on "*informal*" method to reflect the practical constraints of bottom trawl surveys. The proposed designs were a regular design with 28 locations (S28), two regular designs with extra regular added locations with 44 (S44) and 47 (S47) locations, a design which overlaps the regular and stratified random design currently used with 45 locations (S45) and an high density regular design with 108 locations (S108), used just as a benchmark. The designs were assessed by computing bias, relative bias, mean square error and coverages of confidence intervals. Additionally a variance ratio statistic between each study designs and a corresponding random design with the same sample size was computed to separate out the effects of different sample sizes and spatial configurations. The best performance design was S45 with lower variance, higher coverage for confidence intervals and lower variance ratio. This result can be explained by the fact that this design combines good parameter estimation properties of the random designs with good prediction properties of regular designs. In general coverages of confidence intervals where lower than the nominal 95% level reflecting an underestimation of variance. Another interesting fact were the lower coverages of confidence intervals computed by sampling statistics for the random designs, for increasing spatial correlation and sample size. This result illustrates that in the presence of spatial correlation, sampling statistics will underestimate variances according to the combined effect of spatial correlation and sampling density.

Key-words: bottom trawl surveys; geostatistics; simulation; hake; horse mackerel; sampling design.

1 Introduction

Fisheries surveys are the most important sampling process to estimate fish abundance as they provide independent information on the number and weight of fish that exist on a specific area and period. Moreover this information can be disaggregated by several biological parameters like age, length, maturity status, etc. Like other sampling procedures the quality of the data obtained depends in part on the sampling design used to estimate the variables of interest.

For the last 20 to 30 years, bottom trawl surveys (BTS) have been carried out in Western European waters using design-based strategies (Anon., 2002, 2003). However, if one assumes that the number of fish in a specific location is positively correlated with the number of fish in nearby locations, then a geostatistical model can be adopted for estimation and prediction and a model-based approach can be considered to define and assess the sampling design. On the other hand geostatistical principles are widely accepted and can be regarded as a natural choice for modelling fish abundance (e.g. see Rivoirard et al., 2000; Anon., 2004).

Thompson (1992) contrasts design-based and model-based approaches considering that under the former one assumes the values of the variable of interest are fixed and the selection probabilities for inference are introduced by the design, whereas under the latter one consider the observed properties of interest as realisations of random variables and carries out inference based on their joint probability distribution. Hansen et al. (1983) points the key difference between the two strategies by stating that design-based inference does not need to assume a model for the population, the random selection of the sample provides the necessary randomisation, while the model-based inference is made on the basis of an assumed model for the population, and the randomisation supplied by nature is considered sufficient. If the model is appropriate for the problem at hand there will be an efficiency gain in inference and prediction with model-based approaches, however a model mis specification can produce inaccurate conclusions. In our context, with experience accumulated over 20 years of bottom trawls surveys within the study area, there is a fairly good idea of the characteristics of the population and the risk of assuming an unreasonable model should be small.

Portuguese bottom trawl surveys (ptBTS) have been carried out on the Portuguese continental waters since June 1979 on board the R/V Noruega, twice a year in Summer and Autumn. The main objectives of these surveys are: (i) to estimate indices of abundance and biomass of the most important commercial species; (ii) to describe the spatial distribution of the most important commercial species, (iii) to collect individual biological parameters as maturity, sex-ratio, weight, food habits, etc. (SESITS, 1999). The target species are hake (*Merluccius merluccius*), horse mackerel (*Trachurus trachurus*), mackerel (*Scomber scombrus*), blue whiting (*Micromessistius poutassou*), megrims (*Lepidorhombus boscii* and *L. whiffiagonis*), monkfish (*Lophius budegassa* and *L. piscatorius*) and Norway lobster (*Nephrops norvegi-*

cus). A Norwegian Campbell Trawl 1800/96 (NCT) with a codend of 20 mm mesh size, mean vertical opening of 4.8 m and mean horizontal opening between wings of 15.6 m has been used (Anon., 2002).

Between 1979 and 1980, a stratified random sampling design with 15 strata was adopted. Those strata were designed using depth and geographical areas. In 1981 the number of strata were revised to 36. In 1989 the sampling design was reviewed and a new stratification was defined using 12 sectors along the Portuguese continental coast subdivided into 4 depth ranges: 20-100m, 101-200m, 201-500m and 501-750 m, with a total of 48 strata. Due to constraints in the vessel time available the sample size was established in 97 locations, which were allocated equally splited to obtain 2 locations in each stratum. The locations' coordinates were selected randomly constraint by the historical records of clear tow positions and other information about the sea floor, avoiding places where the fishery engine was not able to trawl. This sampling plan was kept fixed over the years. The tow duration until 2001 was 60 minutes and since 2002 was set in 30 minutes, based on an experiment that showed no significant differences in the mean abundance and length distribution between the two tow duration.

The main objective of the present work is to investigated proposals of new sampling designs for the Autumn Portuguese bottom trawl survey (ptBTS). We aimed at explore new spatial configurations and possible increases on sample size, which could be achieved by e.g. reducing the hauling time (from 1 hour to 1/2 hour). Secondly, we aimed at describe a pragmatic procedure to build sampling designs for BTS, develop a statistical approach to compare sampling designs with different sample sizes and spatial configurations, and provide generalized results that could be used for other surveys and species. A simulation study was performed to compare the stratified random design which is currently used against five proposals of systematic based designs, which we called *the study designs*. A model based geostatistical approach (Diggle and Ribeiro, 2006) was adopted using likelihood based methods of inference and conditional simulations to estimate fish abundance on the study area.

Section 2 describes the framework for the simulation study starting with the model specifications followed by the description of the sampling designs and the setup for the simulation study, conducted in five steps as described in Section 2.3. The results of the simulation study comparing the study designs are presented in Section 3 and the findings are discussed in Section 4.

2 Methods

The survey area considered for this work corresponds to the Southwest of the Portuguese Continental EEZ (between Setubal's Canyon and S.Vicent Cape). Before any calculation the mercator projection was transformed into an orthonormal space by converting longitude by the cosine of the mean latitude (Rivoirard et al., 2000). At Portuguese latitude (38-42^o) $1^o lat \approx 60 nm$. The area has $\approx 1250 nm^2$ and the maximum distance between two locations was $\approx 81nm(1.35^{o}lat)$.

2.1 Geostatistical framework

The spatial model assumed here is a log-Gaussian geostatistical model. This is a particular case of the Box-Cox Gaussian transformed class of models discussed in Christensen et al. (2001). The data consists of the pair of vectors (x, y) with elements $(x_i, y_i) : i = 1, ..., n$, where x_i denote the coordinates of a spatial location within a study region $A \subset \mathbb{R}^2$ and y_i is the measurement of the abundance at this location. Denoting by z_i the logarithm of this measurement, the Gaussian model for the vector of variables Z can be written as:

$$Z(x) = S(x) + \varepsilon \tag{1}$$

where $S(x)$ is a stationary Gaussian process at locations x, with $E[S(x)] = \mu$, $Var[S(x)] = \sigma^2$ and an isotropic correlation function $\rho(h) = Corr[S(x), S(x')]$, where $h = \|x - x'\|$ is the Euclidean distance between the locations x and x'; and the terms ϵ are assumed to be mutually independent and identically distributed $\mathrm{Gau}(0, \tau^2)$. For the correlation function $\rho(h)$ we adopted the exponential function with algebraic form $\rho(h) = \exp\{-h/\phi\}$ where ϕ is the correlation range parameter such that $\rho(h) \simeq 0.05$ when $h = 3\phi$. Within the usual geostatistical *jargon* (Isaaks and Srivastava, 1989) $\tau^2 + \sigma^2$ is the (total) sill, σ^2 is the partial sill, τ^2 is the nugget effect and 3ϕ is the practical range.

Hereafter we use the notation $[\cdot]$ for *the distribution of* the quantity indicated within the brackets. The adopted model defines $[\log(Y)] \sim \mathrm{MVGau}(\mu\mathbf{1}, \Sigma)$, i.e $[Y]$ is multivariate log-Gaussian with covariance matrix Σ parametrised by (σ^2, ϕ, τ^2). Parameter estimates can be obtained by maximum likelihood (Diggle and Ribeiro, 2006). For spatial prediction consider first the prediction target $T(x_0) = \exp\{S(x_0)\}$, i.e. the value of the process in the original measurement scale at a vector of spatial locations x_0. Typically x_o defines a grid over the study area. From the properties of the model above the predictive distribution $[T(x)|Y]$ is log-Gaussian with mean μ_T and variance σ^2_T given by:

$$\begin{aligned}
\mu_T &= \exp\{\mathrm{E}[S(x_0)] + 0.5\ \mathrm{Var}[S(x_0)]\} \\
\sigma^2_T &= \exp\{2\ \mathrm{E}[S(x_0)] + \mathrm{Var}[S(x_0)]\}(\exp\{Var[S(x_0)]\} - 1)
\end{aligned}$$

with

$$\begin{aligned}
\mathrm{E}[S(x_0)] &= \mu + \Sigma_0'\Sigma^{-1}(Z - \mathbf{1}\mu) \\
\mathrm{Cov}[S(x_0)] &= \Sigma - \Sigma_0'\Sigma^{-1}\Sigma_0
\end{aligned}$$

where Σ_0 is a matrix of covariances between the variables at prediction locations x_0 and the data locations x and $\text{Var}[S(x_0)]$ is given by the diagonal elements of $\text{Cov}[S(x_0)]$. In practice, we replace the model parameters in the expressions above by their maximum likelihood estimates.

Under the model assumptions, $[T|Y]$ is multivariate log-Gaussian and inferences about prediction means and variances, or other properties of interest, can be drawn either analytically or, more generally, through conditional simulations. Prediction targets can be specified as functionals $\mathcal{F}(S)$ which are applied to the conditional simulations. For instance, inferences on the global mean of a particular realisation of the stochastic process over the area are obtained by defining x_0 as a grid covering the study area at which conditional simulations of $[S(x_0)|Y]$ are taken; the simulated values are then exponentiated and averaged.

2.2 Sampling designs

In general, survey sampling design is about choosing the sample size n and the sample locations x from which data Y can be used to predict any functional of the process. In the case of the ptBTS some particularities must be taken into account: (i) the survey targets several species which may have different statistical and spatial behaviours; (ii) for each species several variables are collected (weight, length, number, etc.); (iii) the sampling is destructive and replicates can not be obtained; (iv) the variability of observed fish abundance is typically high, (v) the planned sampling design may be unattained in practice due to unpredictable commercial fishing activity at the sampling area, bad sea conditions and other operational constraints.

Optimal designs can be obtained formally, by defining a criteria and finding the set of sampling locations which minimises some sort of loss function, as e.g. discussed in Diggle and Lophaven (2006). On the other hand, designs can be defined *informally* by arbitrarily defining locations which compromises between statistical principles and operational constraints. Both are valid for geostatistical inference as described in Section 2.1 provided that the locations x are fixed and stochastically independent of the observed variable Y. The above characteristics of the ptBTS makes it very complex to set a suitable criteria to define a loss function to be minimized with relation to the designs. Additionally, costs of a ship at sea are mainly day based and not haul based and increasing the sample size has to consider groups of locations instead of individual points. Therefore, our approach was to construct the proposed designs informally trying to accommodate: (i) historical information about hake and horse mackerel abundance distribution (Anon., 2002; Jardim, 2004), (ii) geostatistical principles about the estimation of correlation parameters (e.g. see Isaaks and Srivastava, 1989; Cressie, 1993; Muller, 2001) and (iii) operational constraints like known trawlable grounds and minimum distance between hauls.

The *study designs* included the design currently adopted for this survey, named "ACTUAL" with 20 locations, and five systematic based sampling designs. The systematic based designs were defined based on two possible increments in the sample size: a $\approx 40\%$ increment, which is expected to be achievable in practice by reducing haul time from 1 hour to 1/2 hour; and a $\approx 60\%$ increment, which could be achieved in practice by adding to the previous increment an allocation of higher sampling density to this area in order to cover the highest variability of hake recruits historically found within this zone. These designs are denoted by "S" followed by a number corresponding to the sample size. For the former increment a regular design named "S28" was proposed and for the latter three designs were proposed: "S45" overlaps the designs ACTUAL and S28, allowing direct comparison with the previous designs; "S44" and "S47" are two infill designs (Diggle and Lophaven, 2006) obtained by augmenting S28 with a set of locations positioned regularly at smaller distances, aiming to better estimate the correlation parameter and, in particular, the noise-to-signal ratio. S44 was built by defining a single denser sampling zone and S47 by adding three areas with denser sampling. A sixth design "S108" was defined to be used as reference with twice the density of S28.

The designs proposed differ in size and spatial configuration and a simple analysis of any statistics would confound these two effects. This situation motivated the development of a statistical approach to compare designs with different sample sizes and spatial configurations. We used a *ratio of variances* of the relevant estimators between pairs of study designs and random designs with the same sample size, isolating this way the spatial configuration effect. To carry out this analysis we built six additional designs with the same sample size as the study designs and with locations randomly chosen within the study area. We denote these by "R" followed by the number of corresponding locations. Each random design contains all the locations of the previous one such that the results are comparable without effects of the random allocation of the sampling locations.

The *study* and corresponding *random* designs are shown in Figure 1.

2.3 Simulation study

The simulation study was carried out in five steps as follows.

Step 1 **Define a set of study designs.** The sampling designs described in Section 2.2 are denoted by $\Lambda_d : d = 1, \ldots, 12$, with $d = 1, \ldots, 6$ for the study designs and $d = 7, \ldots, 12$ for the corresponding random designs, respectively.

Step 2 **Define a set of correlation parameters.** Based on the analysis of historical data of hake and horse mackerel spatial distribution and defining $\tau^2_{REL} = \tau^2/(\tau^2 + \sigma^2)$, a set of model parameters $\theta_p : p = 1, ..., P$ was defined by all combinations of $\tau^2_{REL} = \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$ and $\phi = \{0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4\}^o lat$. The values of σ^2 are given by setting $\sigma^2 + \tau^2 = 1$.

Step 3 **Simulate data.** For each parameter set θ_p we obtained S=200 simulations $Y_{ps} : s = 1, \ldots, S$ from $[Y]$ on a regular grid of 8781 locations under the model described in Section 2.1. Each simulation Y_{ps} approximates a possible realisation of the process within the study area from which we computed the mean value μ_{ps}. For each Y_{ps} we extracted the data Y_{pds} at the locations of the sampling designs Λ_d.

Step 4 **Estimate correlation parameters.** For each Y_{pds} obtain maximum likelihood estimates (MLE's) $\tilde{\theta}_{pds}$ of the model parameter.

Step 5 **Simulating from the predictive distribution.** A prediction grid x_0 with 1105 locations and the estimates $\tilde{\theta}_{psd}$ were used to obtain C=150 simulations $\tilde{Y}_{pdsc} : c = 1, \ldots, C$ of the conditional distribution $[T(x_0)|Y]$ which were averaged to produce $\bar{\tilde{Y}}_{pdsc}$.

2.4 Analysis of simulation results

The simulation study requires maximum likelihood estimates for the model parameters which are obtained numerically. Therefore a set of summary statistics was computed in order to check the consistency of the results. We have recorded rates of non-convergence of the minimization algorithm; estimates which coincided with the limiting values imposed to the minimization algorithm ($\phi = 3$ and $\tau^2_{REL} = 0.91$); absence of spatial correlation ($\phi = 0$) and values of the parameter estimates which are considered atypical for the problem at hand ($\phi > 0.7$ and $\tau^2_{REL} > 0.67$).

The 48 parameters set (θ_p), 12 sampling designs (Λ_d), 200 data simulations (Y_{psd}) and 150 conditional simulations ($\tilde{Y}_{psdc}$) produced 17.28 million estimates of abundance which were used to compare the designs. For each design we have computed the estimator $\tilde{\mu}_{psd} = C^{-1} \sum_c \bar{\tilde{Y}}_{pdsc}$ of mean abundance μ_{ps} which has variance $\text{Var}(\tilde{\mu}_{psd}) = \bar{\tilde{\rho}}_{AA} + \sum_i^n \sum_j^n w_i w_j \tilde{\rho}_{ij} - 2 \sum_i^n w_i \bar{\tilde{\rho}}_{iA}$, where $\bar{\tilde{\rho}}_{AA}$ is the mean covariance within the area, estimated by the average covariance between the prediction grid locations (x_0); w are kriging weights; $\tilde{\rho}_{ij}$ is the covariance between a pair of data locations; and $\bar{\tilde{\rho}}_{iA}$ is the average covariance between each data locations and the area discretized by the prediction grid x_0 (Isaaks and Srivastava, 1989).

We used bias, relative bias, mean square error (MSE), confidence intervals coverage and ratio of variances to assess the simulation results, comparing the estimates of the abundance provided by the study designs. For each design these statistics were averaged over all the simulations (s) and parameter sets (p) or groups of parameters sets. Considering the difference between the abundance estimates $\tilde{\mu}_{psd}$ and simulated means μ_{ps}, bias was computed by the difference, relative bias was computed by the difference over the estimate $\tilde{\mu}_{ps}$ and MSE was computed by the square of the difference. For each estimate $\tilde{\mu}_{pds}$ a 95% confidence interval for μ_{ps}, given by $\text{CI}(\tilde{\mu}_{psd}) = \tilde{\mu}_{psd} \pm 1.96\sqrt{Var(\tilde{\mu}_{psd})}$, was constructed and the coverage of the confidence intervals δ were computed by the proportion of the intervals which contained the value of μ_{ps} over all the simulations. This statistic was introduced to help assessing the quality of the variance estimates. At least, we called *ratio of variances* a statistic ξ obtained by dividing the variance $Var(\tilde{\mu}_{psd})$ of each study design by the random design with the same size. Notice that the single difference among each pair of designs with the same size was the spatial configuration of the locations and ξ isolated this effect. Finally we used the results from the six random designs to contrast sampling design based and geostatistical based estimates.

All the analysis were performed with the R software (R Development Core Team, 2005) and the add-on packages geoR (Ribeiro Jr. and Diggle, 2001) and RandomFields (Schlather, 2001).

3 Results

Table 1 summarises the analysis of historical data showing parameter estimates for a sequence of years. This aims to gather information on reasonable values for the model parameters. Notice that units for ϕ are given in degrees and, for the adopted exponential correlation model, the practical range in nautical miles (r) is given by 3ϕ and also included in the table. The values of $\tau^2_{REL} = 1$ estimated in some years indicates an uncorrelated spatial process and for such cases estimates of ϕ equals to zero. For most of the cases τ^2_{REL} was estimated as zero due to the lack of nearby locations in the sampling plan and the behaviour of the exponential correlation function at short distances. Given that there is no information in the data about the spatial correlation at distances smaller than the smallest separation distance between a pair of location, this parameter can not be estimated properly and the results depend on the behaviour of the correlation function near the origin.

Table 2 present results used for checking the reliability of the parameter estimates once this could have an impact on the prediction results. The highest rate of lack of convergence was 0.6% for the designs ACTUAL and R20. Estimates of ϕ equals to the upper limit imposed to the algorithm were, in the worst case, 0.9% for R28 and R47 and for τ^2_{REL} it was 1.2% for R28. In general there was a slight worst performance of the random designs but this is irrelevant for the objectives of this study. The above simulations were not considered for subsequent analysis. Lack or weak spatial correlation given by $\phi = 0$ and/or $\tau^2_{REL} > 0.67$ were found in about 35% of the simulations for the designs with fewer number of locations, and this rate decreases as the sample size increases, down to below 10% for the largest designs. For both statistics the study designs showed slightly higher values than the corresponding random designs. Identification of weakly correlated spatial processes in part of the simulations was indeed expected to occur given the low values of ϕ (0.05 and 0.1) and high values of τ^2_{REL} (0.5) used in the simulations. The number of cases that presented $\phi > 0.7$ were slightly higher for random designs, with a maximum of 2.6% for R44 and R45, but were considered to be within an acceptable range given the high variability of the estimator. Our overall conclusion was that the estimation procedure and algorithms produced parameter estimates which can be trusted for subsequent analysis.

Figure 2 shows square bias, variance and MSE obtained from the estimates of correlation parameters ϕ and τ^2_{REL}. For τ^2_{REL} the majority of the designs presented similar patterns with a small contribution of bias to the MSE and increasing values of MSE for higher true parameter values. The designs ACTUAL, S28 and R20 behaved differently with higher values of bias at low values of τ^2_{REL} that pushed MSE to higher values. As an effect of the sample sizes, the absolute values of MSE defines 3 groups composed by designs with 20 and 28 locations, designs with 44, 45 and 47 locations, and designs with 108 locations; with decreasing values of MSE among them, respectively. MSE increases with the increase of the true value of ϕ and its absolute value decreases slightly with the increasing sample sizes. All designs presented a similar pattern with the variance contributing more than bias to the MSE. The study designs showed a slightly higher relative contribution of the variance to MSE compared with the random designs.

Table 3 shows geostatistical abundance estimates ($\tilde{\mu}$) and their bias, relative bias, variance, MSE and 95% confidence interval coverage for both sets of designs. Additionally the table also shows statistics based on sampling theory obtained for random designs. For subsequent analysis the designs S108 and R108 were regarded just as benchmarks since they are unrealistic for practical implementation. Bias were quite small in all situations and can be considered negligible with higher relative bias of 0.014 for S28. All random designs showed a negative bias whereas all study designs showed a positive one. Variances estimated by study designs were lower than the ones for the corresponding random designs. For random designs the variance decays with increasing sample sizes, whereas study designs behaved differently with S45 presenting the lowest variance followed by S47, S44, S28 and S20. MSE showed the same pattern since bias were small, supporting our claim that bias were not relevant for the purpose of this work. The coverages of confidence intervals (δ) were lower than the nominal level of 95% excepted for S108 and R108, reflecting an underestimation of the variance. Considering the designs individually it can be seen that ACTUAL, S28 and S45 showed a lower underestimation than the equivalent random designs. To better investigate this Figure 3 presents values of δ splitted by three levels of correlation (low=$\{0.05, 0.1\}$, med=$\{0.15, 0.20, 0.25\}$, high=$\{0.3, 0.35, 0.4\}$). For geostatistical estimates the coverages δ increases with higher true values of ϕ and larger sample sizes, whereas sampling statistics showed a different pattern, with maximum values for R44 for low and medium correlation levels and for R28 for high correlation levels. This behaviour is more noticeable for stronger spatial correlation, in particular, the largest designs showed lower confidence interval coverage pointing for a more pronounced underestimation of the variance.

Logarithms of the variance ratios between corresponding "S" and "R" designs are presented in Table 3.

Without considering S108 for the reasons stated before, the best result was found for S45 (-0.208) and the worst for S28 (-0.108). This must be balanced by the fact that S45 showed a lower variance underestimation than R45, with the opposite happening for S44/R44 and S47/R47, so, in reality, the value of ξ is smaller for S45 than for S44 and S47.

4 Discussion

The choice of sampling designs for BTS is subject to several practical constraints and this has motivated the adoption of *informally* defined designs which accommodated several sources of information like fishing grounds, haul duration, previous knowledge of the spatial distribution of hake and horse mackerel, among others; which could not be incorporated into a design criteria in an objective way. The fact that this can generate designs with different sample sizes is a drawback of this approach. However, implementing a systematic design on an irregular spatial domain is also to provide designs with different sample sizes, depending on the starting location. On the other hand costs of hauling are relatively small when compared with the fixed costs associated with a vessel's working day and increasing sample sizes for a BTS must consider sets of locations which can be sampled in one working day. For these reasons the different sample sizes of each design are not just a feature of the adopted approach but also a result of the BTS particularities.

The confounding effects of sample size and spatial configuration of the proposed designs jeopardized the comparison of their ability in estimating the abundance. To overcome this limitation a methodology to compare designs with different sample sizes and spatial configurations was required. To deal with this issue we've introduced a mean abundance variance ratio statistic, between the study designs and a corresponding simulated random design with the same sample size.

In fisheries science the main objective for the spatial analysis usually lies in predicting the distribution of the marine resource, aiming, for instance, to define marine protected areas and to compute abundance indices for stock assessment models (Anon., 2004). For such situations the model parameters are not the focus of the study, but just a device to better predict the abundance. Muller (2001) points that the optimality of spatial sampling designs depends on the objectives, showing that ideal designs to estimate covariance parameters of the stochastic process are not the same to predict the value of the stochastic process in a specific location and/or to estimate global abundance. We have not compared the study designs with respect to the estimation of the covariance parameters provided that our main concern was spatial prediction of abundance.

The choice of the parameter estimation method was a relevant issue in the context of this work. The absence of a formal criteria to identify the "best" design naturally led to the use of geostatistical simulations to compare the proposed designs. To carry out a simulation study it is useful to have an objective method capable of producing single estimates of the model parameters. Within traditional geostatistical methods (Isaaks and Srivastava, 1989; Cressie, 1993; Goovaerts, 1997; Rivoirard et al., 2000) the estimation entangles subjective analyst's intervention to define some empirical variogram parameters such as lag interval, lag tolerance and estimator for the empirical variogram. Likelihood based inference produces estimates of the covariance parameters without a subjective intervention of the data analyst, allowing for automatization of the estimation process, which is suitable for simulation studies. For the current work we have also used other methods such as restricted maximum likelihood (REML) and weighted least squares, but they have produced worse rates of convergence in the simulation study. In particular the REML presented an high instability with a high frequency of atypical results for ϕ. An aspect of parameter estimation for geostatistical models which is highlighted when using likelihood based methods is regarded to parameter identification due to over-parametrized or poorly identifiable models (see e.g. Zhang, 2004). To avoid over parametrization we used a log-transformation and the process was considered isotropic, avoiding the inclusion of three parameters on the model: the box-cox transformation parameter (Box and Cox, 1964) and the two anisotropy parameters, angle and ratio. The choice of the log transformation was supported by the analysis of historical data and does not impact the comparison of the designs, given that the relative performance of each design will not be affected by the transformation. A point of concern with the log transformation was the existence of zero values which, in the analysis of the historical data, were treated as measurement error and included in the analysis with a translation of the observed values, by adding a small amount to all observations. However, it must be noted this is not always recommended and, in particular, if the stock is concentrated on small schools that cause discontinuities on the spatial distribution, these transformations will not produce satisfactory results. Concerning anisotropy, a complete simulation procedure was carried out considering a fixed anisotropy angle on the north-south direction and an anisotropy ratio of 1, 1.5 or 2. As expected, the absolute values obtained were different but the overall relative performance was the same, supporting our decision to report results only for the isotropic model.

A major motivation for performing a simulation study was the possibility to use a wide range of covariance parameters that reflect different spatial behaviours. We used two species with different aggregation patterns, hake and horse mackerel, the first an ubiquitous species and the last a more scholastic species, to define the range of the parameters for simulation; suggesting results that can be extended for species with behaviour compatible with the covariance parameters used here.

From a space-time modeling perspective, one of the most interesting analysis for fisheries science is the fluctuation of the stochastic process over time contrasted with the specific realization in a particular time. Therefore the comparison with the mean of the realisations (μ_{ps}) was considered more relevant then to the mean of the underlying process (μ) for the computation of bias and variability. The results showed higher bias for study designs when compared with random designs, but in both cases showing low values which were considered negligible for the purposes of this work. This conclusion was also supported by the fact that MSE showed a similar relative behaviour as variance.

Apart from the design S108, which was introduced as a benchmark and not suitable for implementation, the design that performed better was S45 with lower variance, confidence interval coverage closer to the nominal level of 95% and lower variance ratio (Table 3). One possible reason is the balance between good estimation properties given by the random locations and good predictive properties given by the systematic locations, however the complexity of the BTS objectives makes it impossible to find a full explanation for this results. A possible indicator of the predictive properties is the average distance between the designs and the prediction grid locations, which reflects the extrapolation needed to predict over a grid. We found that S45 had an average of $2.61nm$ whereas for S47 the value is $2.72nm$, explaining in part the S45 performance.These results are in agreement with Diggle and Lophaven (2006) who showed that *lattice plus closed pairs* designs (similar to S45) performed better than *lattice plus in-fill* designs (similar to S44 and S47) for accurate prediction of the underlying spatial phenomenon. The combination of random and systematic designs like S45 is seldom considered in practice and we are not aware of recommendations of such designs for BTS.

It was interesting to notice that most designs presented a coverage of confidence intervals below the nominal level of 95% revealing the variances were underestimated. It was not fully clear how to use such results to correct variance estimation and further investigation is needed on the subject. Care must be taken when looking at variance ratios since underestimated denominators will produce higher ratios which can mask the results. This was the case of S45 when comparing to S47 and S44, supporting our conclusions about S45.

Another result of our work was the assessment of abundance estimates from random designs by sampling statistics, the most common procedure for fisheries surveys (Anon., 2004), under the presence of spatial correlation. In such conditions an increase in sample size may not provide a proportional increase in the quantity of information due to the partial redundancy of information under spatial correlation. Results obtained for coverages of confidence intervals illustrated this (Table 3 and Figure 3), with smaller coverages for larger sample sizes and higher spatial correlation. In our opinion this is due to an overestimation of the degrees of freedom that lead to an underestimation of prediction standart errors producing the smaller coverages. These findings support claims to consider geostatistical methods to estimate fish abundance, such that correlation between locations is explicitly considered in the analysis, and highlighting the importance of verifying the assumptions behind sampling theory before computing the uncertainty of abundance estimates.

5 Acknowledgements

The authors would like to thank the scientific teams evolved in the Portuguese Bottom Trawl Surveys, in particular the coordinator Fátima Cardador, and the comments by Manuela Azevedo. This work was carried out within the IPIMAR's project NeoMAv (QCA-3/MARE-FEDER, http://ipimar-iniap.ipimar.pt/neomav) and was co-financed by project POCTI/MATH/44082/2002.

References

Anon., 2002. Report of the International Bottom Trawl Survey Working Group. Tech. rep., International Council for the Exploitation of the Sea (ICES), ICES CM 2002/D:03.

Anon., 2003. Report of the International Bottom Trawl Survey Working Group. Tech. rep., International Council for the Exploitation of the Sea (ICES), ICES CM 2003/D:05.

Anon., 2004. Report of the Workshop on Survey Design and Data Analysis. Tech. rep., International Council for the Exploitation of the Sea (ICES), ICES CM 2004/B:07.

Box, G., Cox, D., 1964. An Analysis of Transformations. Journal of the Royal Statistical Society Series B 26, 211–243.

Christensen, O., Diggle, P., Ribeiro Jr, P., 2001. Analysing positive-valued spatial data: the transformed gaussian model. In: Monestiez, P., Allard, D., Froidevaux (Eds.), GeoENV III - Geostatistics for Environmental Applications. Vol. 11 of Quantitative Geology and Geostatistics. Kluwer, pp. 287–298.

Cressie, N., 1993. Statistics for spatial data - Revised Edition. John Wiley and Sons, New York.

Diggle, P. J., Lophaven, S., 2006. Bayesian geostatistical design. Scandinavian Journal of Statistics 33, 55–64.

Diggle, P., Ribeiro, P., 2006. Model-based Geostatistics. Springer, New York, in press.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.

Hansen, M., Madow, W., Tepping, B., 1983. An Evaluation of Model-Dependent and Probability-Sampling Inferences in Sample Surveys. Journal of the American Statistical Association 78 (384), 776–793.

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

Jardim, E., 2004. Visualizing hake recruitment - a non-stationary process. In: Sanchez-Vila, X., Carrera, J., Gómez-Hernandéz, J. J. (Eds.), geoENV IV - Geostatistics for Environmental Applications. Vol. 13 of Quantitative Geology and Geostatistics. Kluwer Academic Publishers, London, pp. 508–509.

Muller, W., 2001. Collecting Spatial Data - Optimum Design of Experiments for Random Fields, 2nd Edition. Contributions to statistics. Physica-Verlag, Heidelberg.

Ribeiro Jr., P., Diggle, P., June 2001. geoR: a package from geostatistical analysis. R-NEWS 1 (2), 15–18. URL http://cran.R-project.org/doc/Rnews

Rivoirard, J., Simmonds, J., Foote, K., Fernandes, P., Bez, N., 2000. Geostatistics for Estimating Fish Abundance. Blackwell Science, London, England.

R Development Core Team, 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
URL http://www.R-project.org

Schlather, M., June 2001. Simulation and analysis of random fields. R News 1 (2), 18–20.
URL http://CRAN.R-project.org/doc/Rnews/

SESITS, 1999. Evaluation of Demersal Resources of Southwestern Europe from Standardised Groundfish Surveys - Study Contract 96/029. Final report, DG XIV / EC.

Thompson, S., 1992. Sampling. Statistics. John Wiley & Sons, INC, New York.

Zhang, H., 2004. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. Journal of the American Statistical Association 99 (465), 250 – 261.

Table 1: Exponential covariance function parameters ($\phi, \tau^2_{\mathrm{REL}}$) and the geostatistical range (r) estimated yearly (1990-2004) for hake and horse mackerel abundance. The values of ϕ are presented in degrees of latitude and range in nautical miles. The maximum distance between pairs of locations was 63nm.

	Hake			Horse mackerel		
	ϕ(^{o}lat)	r(nm)	τ^2_{REL}	ϕ(^{o}lat)	r(nm)	τ^2_{REL}
1990	0.05	9.1	0.01	0.42	76.4	0.00
1991	0.14	24.4	0.63	0.49	88.9	0.43
1992	0.00	0.0	1.00	0.22	39.3	0.05
1993	0.05	9.3	0.00	0.00	0.0	1.00
1995	0.05	8.8	0.00	0.08	14.4	0.00
1997	0.14	24.8	0.00	0.21	38.6	0.42
1998	0.02	3.4	0.00	0.09	16.5	0.00
1999	0.10	17.8	0.00	0.09	16.0	0.00
2000	0.03	4.6	0.00	0.16	29.5	0.00
2001	0.07	12.9	0.00	0.42	75.7	0.06
2002	0.00	0.0	1.00	0.05	8.9	0.00
2003	0.33	59.0	0.00	0.34	62.0	0.00
2004	0.09	15.4	0.00	0.09	17.0	0.00

Table 2: Statistics to provide simulation quality assessment (in percentages) for both design sets and all sample sizes: non-convergence of the minimization algorithm (non-conv); cases truncated by the limits imposed to the minimization algorithm ($\phi = 3$ and $\tau^2_{\mathrm{REL}} = 0.91$); uncorrelated cases ($\phi = 0$); and atypical values of the correlation parameters ($\phi > 0.7$ and $\tau^2_{\mathrm{REL}} > 0.67$).

statistic	design	sample size					
		20	28	44	45	47	108
non-conv	study	0.6	0.5	0.2	0.2	0.2	0.1
	random	0.6	0.4	0.2	0.2	0.2	0.1
$\phi = 3$	study	0.7	0.5	0.7	0.7	0.5	0.2
	random	0.6	0.9	0.8	0.8	0.9	0.1
$\tau^2_{\mathrm{REL}} = 0.91$	study	0.7	0.7	1.0	0.9	0.8	0.4
	random	0.8	1.2	1.1	1.1	1.1	0.2
$\phi = 0$	study	36.3	33.0	20.7	20.6	18.0	5.3
	random	32.8	28.5	18.1	17.2	16.2	3.3
$\phi > 0.7$	study	1.3	1.6	1.9	1.9	1.8	1.4
	random	1.8	2.2	2.6	2.6	2.4	1.7
$\tau^2_{\mathrm{REL}} > 0.67$	study	38.5	35.8	24.2	24.7	21.8	10.0
	random	35.0	31.6	22.1	21.1	20.3	7.6

Table 3: Summary statistics per sets of sampling designs and sample size. Geostatistical abundance estimates ($\tilde{\mu}$) in kg/hour, bias (bias($\tilde{\mu}$)), relative bias (bias$_r$($\tilde{\mu}$)), variance (var($\tilde{\mu}$)), mean square error (MSE) and 95% confidence interval coverage ($\delta(\tilde{\mu})$). Mean log variance ratios per sampling design type (ξ) measures the relative log effect of the systematic based designs configuration with relation to the random designs. The last six rows present the same statistics estimated for random designs by sampling statistics.

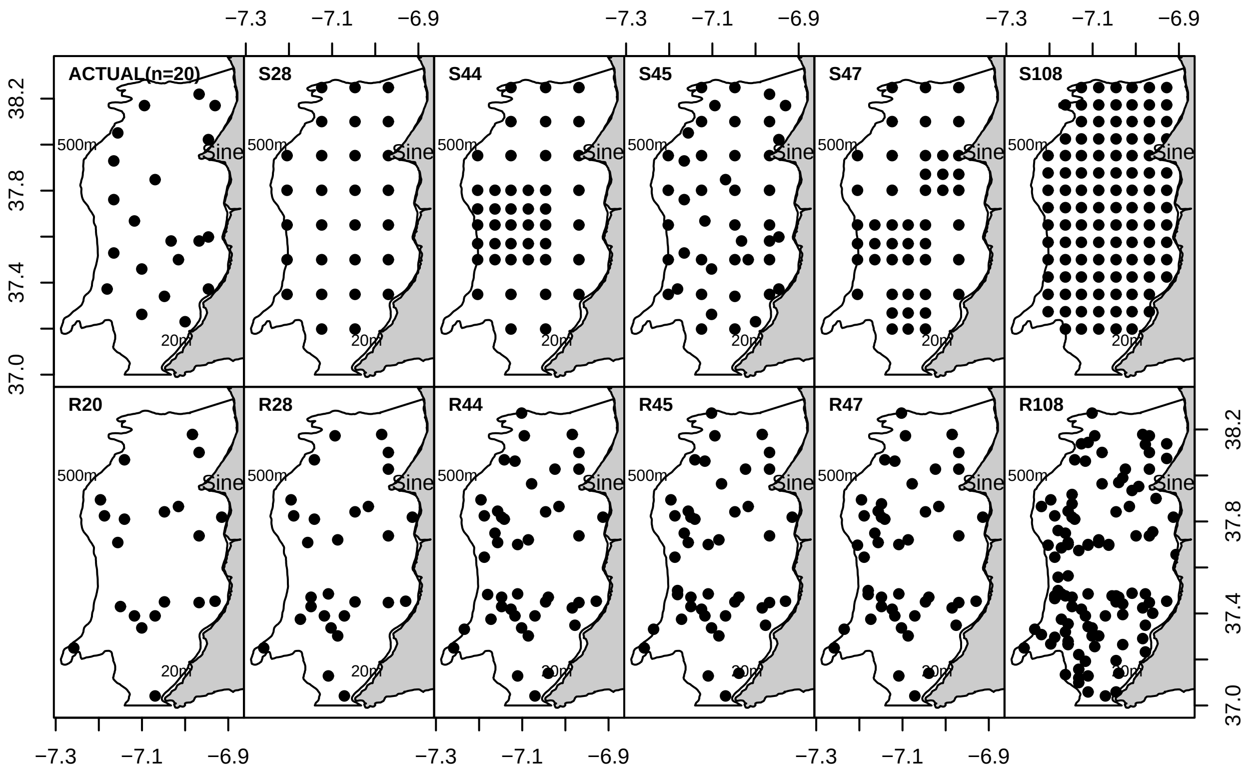
method	statistic	design	number of locations					
			20	28	44	45	47	108
geostatistics	$\tilde{\mu}$	study	1.658	1.662	1.649	1.657	1.651	1.641
		random	1.631	1.624	1.625	1.624	1.625	1.625
	bias($\tilde{\mu}$)	study	0.025	0.030	0.016	0.026	0.019	0.008
		random	-0.001	-0.008	-0.007	-0.009	-0.008	-0.007
	bias$_r$($\tilde{\mu}$)	study	0.012	0.014	0.003	0.012	0.005	0.001
		random	-0.004	-0.008	-0.005	-0.006	-0.005	-0.005
	var($\tilde{\mu}$)	study	0.136	0.108	0.092	0.086	0.089	0.081
		random	0.168	0.129	0.113	0.112	0.112	0.097
	MSE($\tilde{\mu}$)	study	0.272	0.196	0.164	0.144	0.154	0.104
		random	0.321	0.230	0.173	0.171	0.171	0.124
	$\delta(\tilde{\mu})$	study	0.908	0.922	0.907	0.939	0.920	0.960
		random	0.895	0.909	0.937	0.934	0.934	0.954
	ξ	stu/rnd	-0.128	-0.107	-0.150	-0.208	-0.179	-0.228
sampling statistics	$\bar{Y}$	random	1.615	1.619	1.618	1.616	1.618	1.622
	bias($\bar{Y}$)	random	-0.017	-0.014	-0.014	-0.017	-0.015	-0.010
	bias$_r$($\bar{Y}$)	random	-0.017	-0.014	-0.013	-0.014	-0.014	-0.006
	var($\bar{Y}$)	random	0.197	0.146	0.091	0.088	0.085	0.037
	MSE($\tilde{\mu}$)	random	4.133	4.238	4.109	4.083	4.090	4.073
	$\delta(\bar{Y})$	random	0.900	0.910	0.908	0.900	0.896	0.840

Figure 1: Sampling designs and the study area (southwest of Portugal). Each plot shows the sample locations, the bathymetric isoline of 500m and 20m and the coast line. The sampling design name is presented on the top left corner of the plots. The top row shows the *study* designs and the bottom row the random designs.

Figure 2: Summary statistics for the covariance parameters estimation by sampling design as a function of the true parameter values. bias^2 ($\circ$), variance ($\triangle$) and mean square error ($+$). Top figure presents τ^2_{REL} results and bottom figure ϕ.

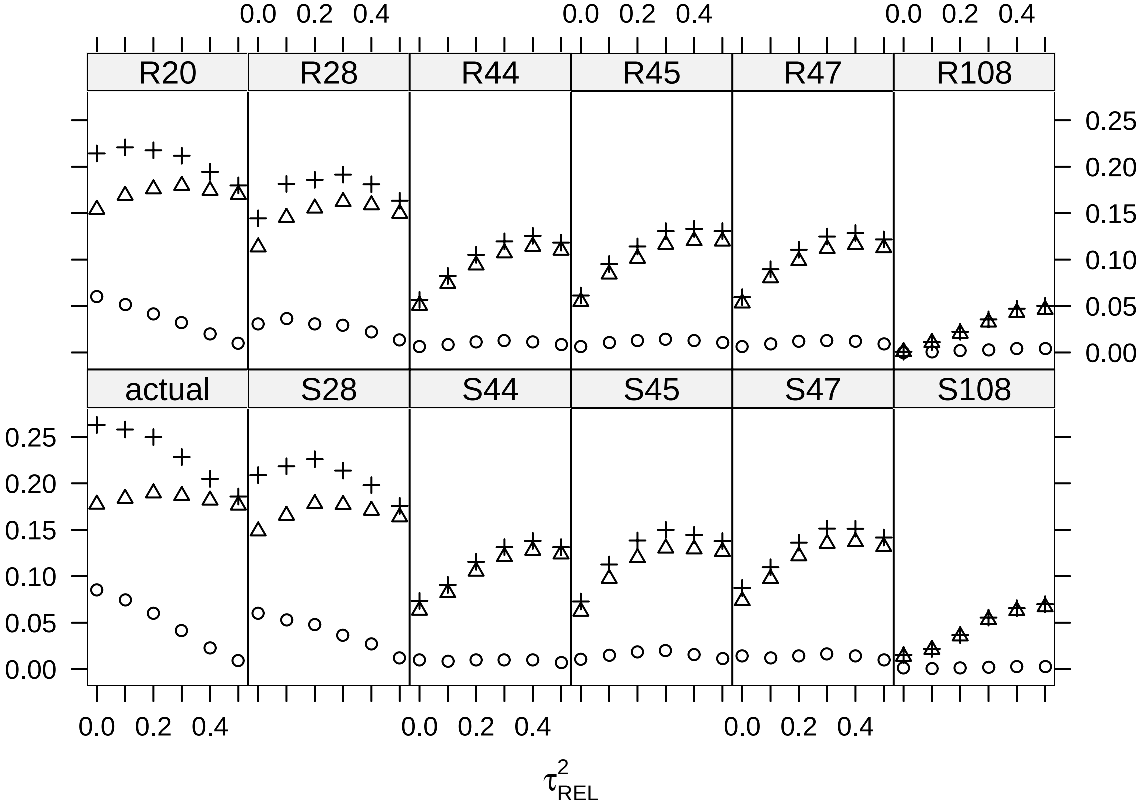
Figure 3: Coverage of the confidence intervals (δ) for different ϕ levels (low $=$ $\{0.05, 0.1\}$, med$\{0.15, 0.20, 0.25\}$ high $=$ $\{0.30, 0.35, 0.40\}$) for estimates of abundance by sampling statistics for the random designs ($+$) and by geostatistics for the study (o) and random designs ($*$).

FIGURE 01



ACTUAL(n=20)
S28
S44
S45
S47
S108
R20
R28
R44
R45
R47
R108
500m
Sine
20m
−7.3
−7.1
−6.9
37.0
37.4
37.8
38.2

FIGURE 02


R20
R28
R44
R45
R47
R108
actual
S28
S44
S45
S47
S108
0.0 0.2 0.4
0.00 0.05 0.10 0.15 0.20 0.25
τ^2_{REL}

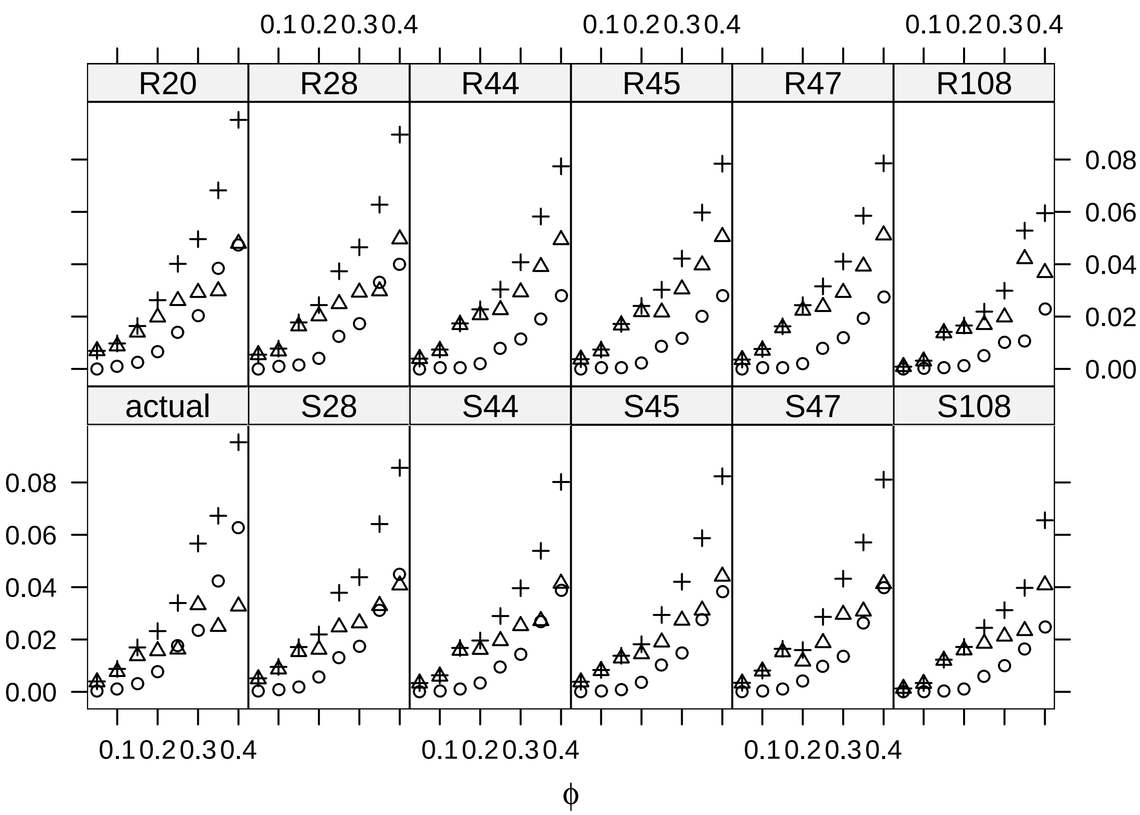

R20
R28
R44
R45
R47
R108
actual
S28
S44
S45
S47
S108
0.1 0.2 0.3 0.4
0.00 0.02 0.04 0.06 0.08
ϕ

FIGURE 03

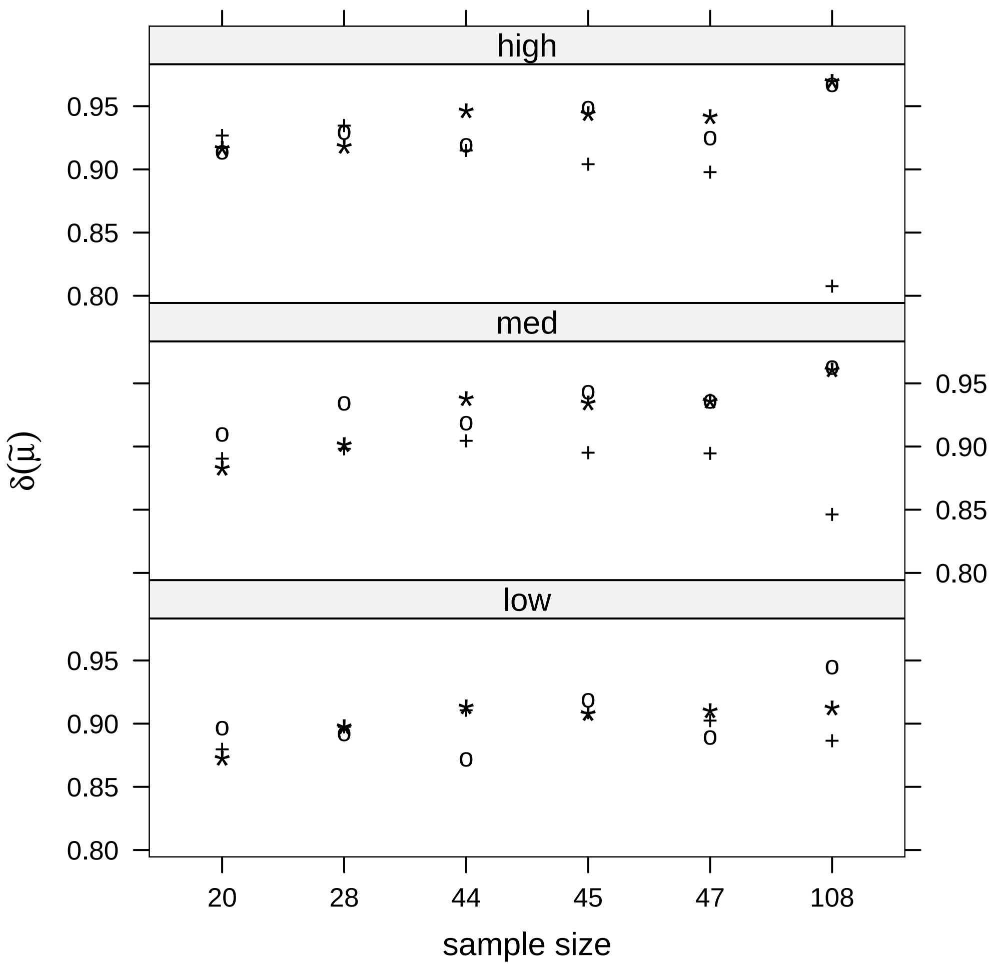
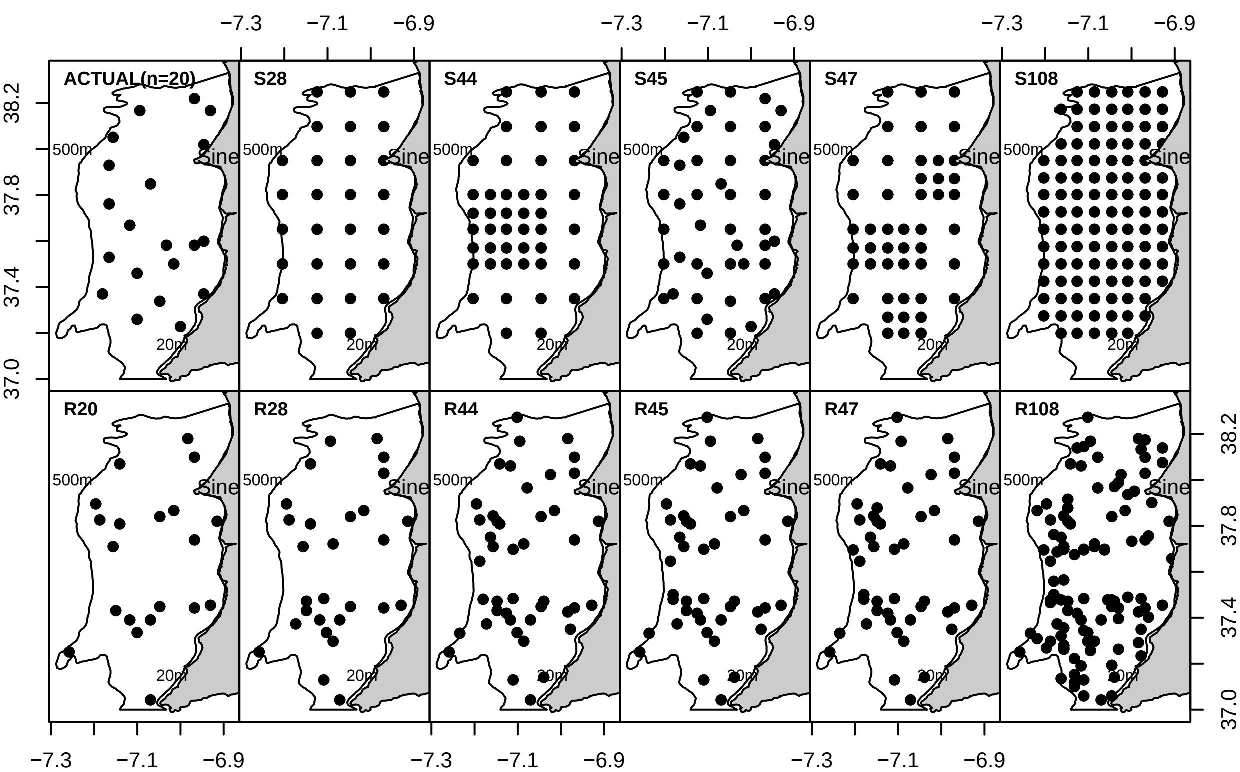
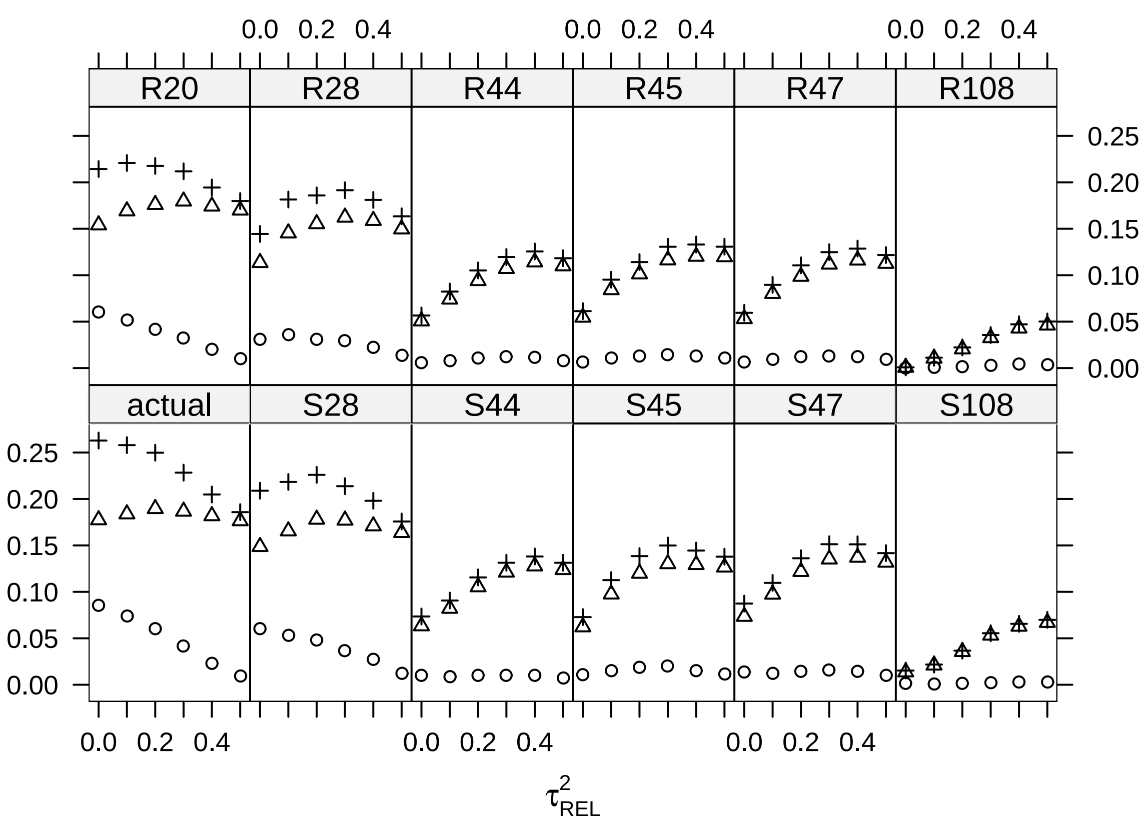


Figure01



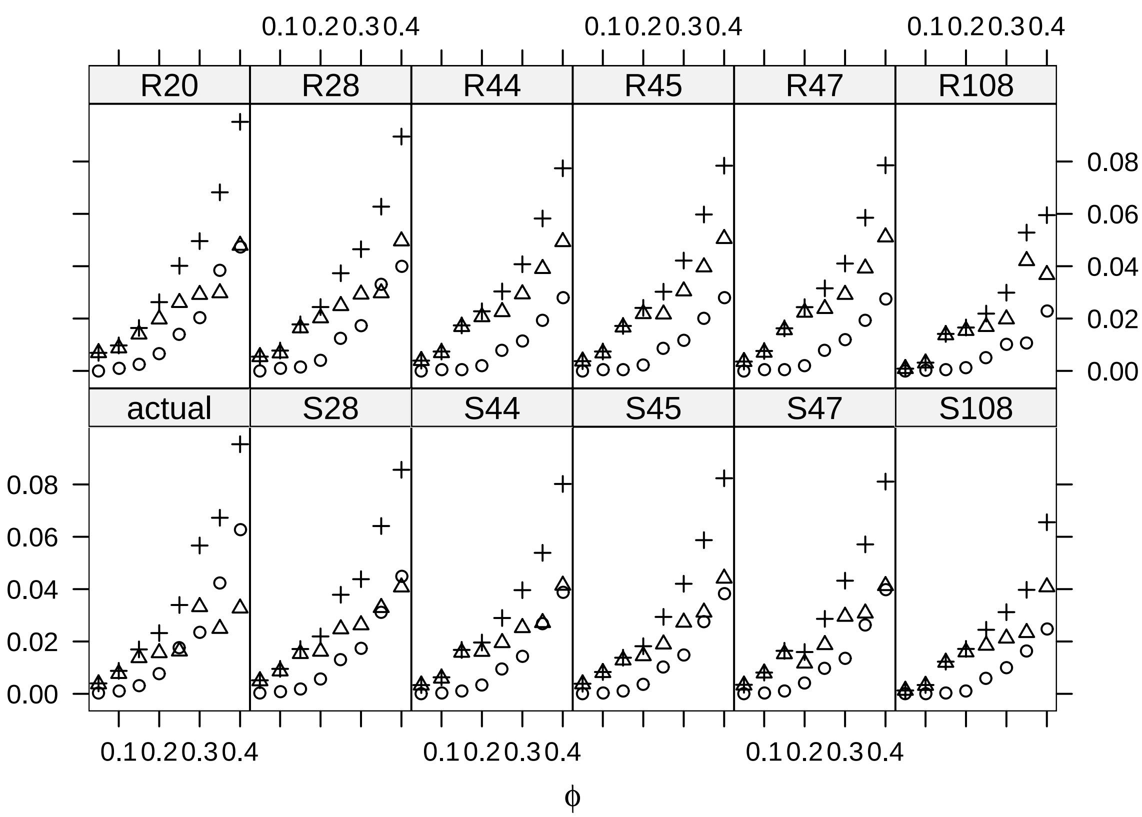
ACTUAL(n=20)
S28
S44
S45
S47
S108
R20
R28
R44
R45
R47
R108
500m
20m
Sine
−7.3
−7.1
−6.9
38.2
37.8
37.4
37.0

Figure02_1



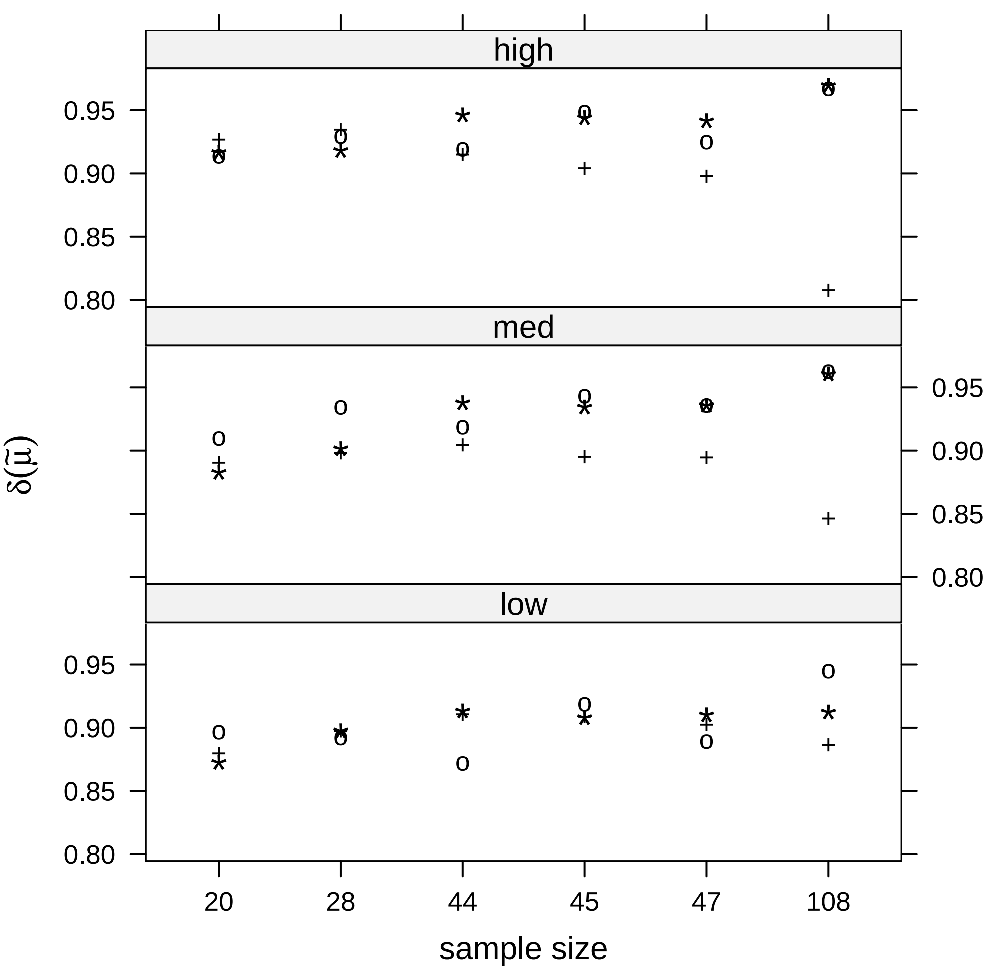
0.0 0.2 0.4
0.0 0.2 0.4
0.0 0.2 0.4
R20
R28
R44
R45
R47
R108
0.25
0.20
0.15
0.10
0.05
0.00
actual
S28
S44
S45
S47
S108
0.25
0.20
0.15
0.10
0.05
0.00
0.0 0.2 0.4
0.0 0.2 0.4
0.0 0.2 0.4
τ^2_{REL}

Figure02_2



R20
R28
R44
R45
R47
R108
actual
S28
S44
S45
S47
S108
0.00
0.02
0.04
0.06
0.08
0.1 0.2 0.3 0.4
ϕ

Figure03



FISH896 REVISION NOTES
ERNESTO JARDIM
08/Jan/2007

The revision was carried out to accommodate the comments of the Reviewers. Below you can find the reviewers comments followed by our answers. A pdf file with all the corrections made was also uploaded to the system, named "ejpj.ptBTSgeosim.revisions.pdf".

- Reviewer #1

General evaluation: Acceptable after minor revision

This is an interesting, straightforward manuscript assessing the effect of sample size and spatial configuration of Portuguese bottom trawl surveys in fish abundance estimates through geostatistical methods. The writing is clear and the figures and tables appropriate. The simulations are carefully designed, including the simulated data and the set of correlation parameters with their respective maximum likelihood estimates.

General Comments

1. Line 100. "The spatial model assumed here is a Log-Gaussian geostatistical model".

In the discussion section, the authors justify the use of isotropic models (lines 338 to 341) but no explanation and/or justification about the log-Gaussian geostatistical model selected are given. Further explanation about the reasons of the model selection will clarify the results.

2. Line 241. "Table 2 summarizes the checks of the results of the parameter estimates which were considered satisfactory and coherent".

It is not thoroughly clear in the text what the authors mean with satisfactory and coherent. More detail will be relevant to better understand the sampling design and survey processes

3. Line 347. "Furthermore, the results can be retained for all species with a spatial behaviour covered by these parameters".

It seems like the authors assume all the species surveyed have similar spatial behavior. This is not necessarily true, especially if the survey is targeting species with different life history traits and aggregation behaviors under different spatial scales (i.e. demersal fishes vs. sedentary invertebrates). Furthermore, the autocorrelation structure in the data is not explicitly mentioned or described. Additional information and discussion on the effect of spatial correlation for the different stocks targeted by the trawl survey on the model selection will improve the robustness of this study.

Minor comments

Line 123: repeated word: the the

Line 125: Unnecessary word: are

Lines 183 and 206: different notation for sampling designs <LAMBDA>d and ?d

Line 234: confusing sentence/notation: "...and also included in the Table Â®..."

Table 3: Summary statistics units are not specified.

Figure 1: X and Y axis legends should be specified (i.e. Longitude West and Latitude North respectively).

Figure 2: Variables in the X axis are specified in the legend but not in the figure

- Answers to Reviewer #1

1. There's a paragraph (lines 325-332) justifying the use of a log transform, in particular in lines 330-332 is mentioned that the log was found on previous analysis of the historical data.

2. We agree with the reviewer comment and adjusted the text to clarify it. The key issue was that the convergence was good and the parameters estimates were within the range of the initial parameters, so the simulations could be trusted for the following work.

3.1 We generalized our results for all species that fit in the range of the covariance parameters used. This may not apply to invertebrates but certainly apply for most demersal species, which are the target of our survey. This sentence was revised to clarify it's aim.

3.2 The autocorrelation structure in the data is presented in Table 1 where all the correlation parameters estimated are shown and in lines 231-240 we describe them and the most important particularities found.

3.3 We used two different species with very different aggregation behaviors, hake an ubiquitous species and horse mackerel a more scholastic species, and both species present quite different life traits. We believe these two entangle characteristics that are quite extreme within our target species, although we can not guarantee that other species in specific years would not present correlation structures that are outside the range choose.

- Reviewer #3

I propose rejecting this submission because it is overly detailed on the simulation results (1), gives little insight how the simulations relate to the original Portuguese survey data (2), of which little is spoken, and because it is not clear why this is to be considered more than an exercise confirming what already has been stated in Diggle and Lophaven (3). The authors do show an understanding of the issues involved in simulation and did not, in my mind, make any errors. Some of the results are technical and issues of isotropy, parameter estimation and the like are discussed at a more technical level than would be understood by a general reader. The one significant result is that when there is autocorrelation in the underlying data it is better to use a combindation of regular survey with paired random additions (to provide points close to each other and better estimate autocorrelation I presume) than a pure random design for fisheries surveys. If this is indeed a new result (I'm really not sure whether it is) then this could be acceptable as a greatly reduced in size 'note' that gives the results and refers to a web document or report for details of the simulations (4). Certainly the geostatistical equations are not needed and are better found elsewhere (5). They are not new to the fisheries literature. Finally, in simulation work like this I am left unsure how general the results are to other areas (6). This the authors discussed some and think the results are general (maybe they are). There is little need in that case to focus on the real system (7). Otherwise, some evaluation using actual data would be useful (if there were a year when higher sampling intensity was used â€“ it could be subsampled to see how much the estimates changed) (8). In fairness to the authors I did not study the results in detail. Maybe someone who does will find gold in it. I did not think it was worth looking.

- Answers to reviewer #3

(1) The detailed simulation results were included to allow readers to understand the scope of our work and have enough information to judge if their own situation is inside the range of our work.

(2) The historical data was used to condition the simulation work using the covariance parameters obtained with it to define the range of the parameters used for simulation.

(3) The results obtained by Diggle and Lophaven were theoretical and not applied to a real situation, like we did. On the other hand their work compares two specific ways of building sampling designs, "lattice plus close pairs" and "lattice plus infill", and never include a pure random or regular design, which we did. Also they use only geostatistical methods and we also included a comparison of the designs performance using sampling theory estimators. We included anisotropy and log transformation on our analysis. More important of all, we describe an easy way of building a sampling design that has the characteristics of "lattice plus close pairs", by overlapping the random design with a regular design that can be applicable to most European Bottom Trawl Surveys. However, this comment called our attention to the fact that the achievements may not be clearly described on the paper and made the necessary changes.

(4) This results are new at least in Fisheries Science once that there is no reporting of surveys using such sampling strategy. The authors can not guarantee that the theoretical results of Diggle and Lophaven were not implemented already in other scientific areas, but the bibliographic search did not show any papers about its implementation. Also there are secondary results that are new in this work (i) the approach to build the sampling designs, (ii) the approach to compare sampling designs with different sample sizes, (iii) the result about the underestimation of abundance variance by the variance of the sampling mean. However, this comment called our attention to the fact that the achievements were not clearly highlighted and we introduced the necessary revisions.

(5) Section 2.1 was included to make the paper self contained and to introduce our notation, providing information so that readers clearly understand the scope of the work. However, we partially agree with the referee and revised and decreased the presentation of the geostatistical framework to a minimum necessary for the readers to follow the paper.

(6) The results are generalized by the spatial behavior of the resource (see answer 3.1 to reviewer #1). If in another area someone exploring the spatial correlation of a resource finds parameters that fit inside the range of parameters used for our simulations, there is a good chance that the sampling

design of the survey collecting its data will gain by adopting a mixed random/regular design.

(7) As said in point (2) the focus on the real system is just enough to provide information for conditioning the simulations so that the results are applicable to the real world. There was not the intention of explore deeply the data or completely ignore it.

(8) This would be a valid approach if the spatial correlation is ignored, once that the removal of a location would not only reduce the sample size but also the configuration of the sampling design with and impact extremely difficult to assess.

Geostatistical Assessment of Sampling Designs for Portuguese Bottom Trawl Surveys

Ernesto Jardim[1] <ernesto@ipimar.pt> and Paulo J. Ribeiro Jr[2] <paulojus@ufpr.br>

8th January 2007

[1]Instituto Nacional de Investigação Agrária e das Pescas, Av. Brasilia, 1449-006, Lisboa, Portugal
Tel: +351 213 027 093

[2]Departamento de Estatística, Universidade Federal do Paraná, C.P. 19.081 CEP: 81.531-990, Curitiba, Paraná, Brasil

Abstract

New sampling designs for the Autumn Portuguese bottom trawl survey (ptBTS) were investigated to explore alternative spatial configurations and possible increments on sample size. The currently used stratified random design and five proposals of systematic based designs were assessed by a simulation study, adopting a geostatistical approach based on likelihood methods of inference. The construction of the designs was based on "*informal*" method to reflect the practical constraints of bottom trawl surveys. The proposed designs were a regular design with 28 locations (S28), two regular designs with extra regular added locations with 44 (S44) and 47 (S47) locations, a design which overlaps the regular and stratified random design currently used with 45 locations (S45) and an high density regular design with 108 locations (S108), used just as a benchmark. The designs were assessed by computing bias, relative bias, mean square error and coverages of confidence intervals. Additionally a variance ratio statistic between each study designs and a corresponding random design with the same sample size was computed to separate out the effects of different sample sizes and spatial configurations. The best performance design was S45 with lower variance, higher coverage for confidence intervals and lower variance ratio. This result can be explained by the fact that this design combines good parameter estimation properties of the random designs with good prediction properties of regular designs. In general coverages of confidence intervals where lower than the nominal 95% level reflecting an underestimation of variance. Another interesting fact were the lower coverages of confidence intervals computed by sampling statistics for the random designs, for increasing spatial correlation and sample size. This result illustrates that in the presence of spatial correlation, sampling statistics will underestimate variances according to the combined effect of spatial correlation and sampling density.

Key-words: bottom trawl surveys; geostatistics; simulation; hake; horse mackerel; sampling design.

1 Introduction

Fisheries surveys are the most important sampling process to estimate fish abundance as they provide independent information on the number and weight of fish that exist on a specific area and period. Moreover this information can be disaggregated by several biological parameters like age, length, maturity status, etc. Like other sampling procedures the quality of the data obtained depends in part on the sampling design used to estimate the variables of interest.

For the last 20 to 30 years, bottom trawl surveys (BTS) have been carried out in Western European waters using design-based strategies (Anon., 2002, 2003)~~()~~. However, if one assumes that the number of fish in a specific location is positively correlated with the number of fish in nearby locations, then a geostatistical model can be adopted for estimation and prediction and a model-based approach can be considered to define and assess the sampling design. On the other hand geostatistical principles are widely accepted and can be regarded as a natural choice for modelling fish abundance (~~see~~ e.g. see Rivoirard et al., 2000; Anon., 2004).

Thompson (1992) contrasts design-based and model-based approaches considering that under the former one assumes the values of the variable of interest are fixed and the selection probabilities for inference are introduced by the design, whereas under the latter one consider the observed properties of interest as realisations of random variables and carries out inference based on their joint probability distribution. Hansen et al. (1983) points the key difference between the two strategies by stating that design-based inference does not need to assume a model for the population, the random selection of the sample provides the necessary randomisation, while the model-based inference is made on the basis of an assumed model for the population, and the randomisation supplied by nature is considered sufficient. If the model is appropriate for the problem at hand there will be an efficiency gain in inference and prediction with model-based approaches, however a model mis specification ~~misspecification~~ can produce inaccurate conclusions. In our context, with experience accumulated over 20 years of bottom trawls surveys within the study area, there is a fairly good idea of the characteristics of the population and the risk of assuming an unreasonable model should be small.

Portuguese bottom trawl surveys (ptBTS) have been carried out on the Portuguese continental waters since June 1979 on board the R/V Noruega, twice a year in Summer and Autumn. The main objectives of these surveys are: (i) to estimate indices of abundance and biomass of the most important commercial species; (ii) to describe the spatial distribution of the most important commercial species, (iii) to collect individual biological parameters as maturity, sex-ratio, weight, food habits, etc. (SESITS, 1999)~~(SESITS 1999)~~. The target species are hake (*Merluccius merluccius*), horse mackerel (*Trachurus trachurus*), mackerel (*Scomber scombrus*), blue whiting (*Micromessistius poutassou*), megrims (*Lepidorhombus boscii* and *L. whiffiagonis*), monkfish (*Lophius budegassa* and *L. piscatorius*) and Norway lobster (*Nephrops* *norvegicus*). A Norwegian Campbell Trawl 1800/96 (NCT) with a codend of 20 mm mesh size, mean vertical opening of 4.8 m and mean horizontal opening between wings of 15.6 m has been used (Anon., 2002)~~()~~.

Between 1979 and 1980, a stratified random sampling design with 15 strata was adopted. Those strata were designed using depth and geographical areas. In 1981 the number of strata were revised to 36. In 1989 the sampling design was reviewed and a new stratification was defined using 12 sectors along the Portuguese continental coast subdivided into 4 depth ranges: 20-100m, 101-200m, 201-500m and 501-750 m, with a total of 48 strata. Due to constraints in the vessel time available the sample size was established in 97 locations, which were allocated equally splited to obtain 2 locations in each stratum. The locations' coordinates were selected randomly constraint by the historical records of clear tow positions and other information about the sea floor, avoiding places where the fishery engine was not able to trawl. This sampling plan was kept fixed over the years. The tow duration until 2001 was 60 minutes and since 2002 was set in 30 minutes, based on an experiment that showed no significant differences in the mean abundance and length distribution between the two tow duration.

The main objective of the present work is to ~~present work~~ investigated proposals of new sampling designs for the Autumn Portuguese bottom trawl survey (ptBTS). We aimed at explore new spatial configurations and possible increases on sample size, which could be achieved by e.g. reducing the hauling time (from 1 hour to 1/2 hour). Secondly, we aimed at describe a pragmatic procedure to build sampling designs for BTS, develop a statistical approach to compare sampling designs with different sample sizes and spatial configurations, and provide generalized results that could be used for other surveys and species. A simulation study was performed to compare the stratified random design which is currently used against five proposals of systematic based designs, which we called *the study designs*. A model based geostatistical approach (Diggle and Ribeiro, 2006) was adopted using likelihood based methods of inference and conditional simulations to estimate fish abundance on the study area.

Section 2 describes the framework for the simulation study starting with the model specifications followed by the description of the sampling designs and the setup for the simulation study, conducted in five steps as described in ~~(~~Section 2.3~~)~~. The results of the simulation study comparing the study designs are presented in Section 3 and the findings are discussed in Section 4.

2 Methods

The survey area considered for this work corresponds to the Southwest of the Portuguese Continental EEZ (between Setubal's Canyon and S.Vicent Cape). Before any calculation the mercator projection was transformed into an orthonormal space by converting longitude by the cosine of the mean latitude

(Rivoirard et al., 2000). At Portuguese latitude (38-42°) $1^o lat \approx 60nm$. The area has $\approx 1250nm^2$ and the maximum distance between two locations was $\approx 81nm(1.35^o lat)$.

2.1 Geostatistical framework

~~Fish in a certain area interact with each other looking for food, reproductive conditions, etc. Therefore it is natural to consider that the abundance of fish between spatial locations is positively correlated such that the correlation decays with increasing separation distances. This conjecture justifies adopting the spatial model as defined in geostatistics (see e.g. , Part 1) to describe and obtain predictions of fish abundance over an area. This approach contrasts with the *sampling theory* (see e.g.) where the correlation between observations is not taken into account. Additionally, within the geostatistical approach it is possible to estimate the abundance variance from systematic designs and the parameters of the correlation function allows for the definition of different phenomena. Sampling theory estimates would be obtained as the particular case, in the absence spatial correlation. Possible concerns includes the extra complexity given by the model choice and eventual difficulties in estimating the model parameters.~~

The spatial model assumed here is a log-Gaussian geostatistical model. This is a particular case of the Box-Cox Gaussian transformed class of models discussed in Christensen et al. (2001). The data consists of the pair of vectors (x, y) with elements $(x_i, y_i) : i = 1, ..., n$, where x_i denote the coordinates of a spatial location within a study region $A \subset \mathbb{R}^2$ and y_i is the measurement of the abundance at this location. Denoting by z_i the logarithm of this measurement, the Gaussian model for the vector of variables Z can be written as:

$$Z(x) = S(x) + \varepsilon \tag{1}$$

where $S(x)$ is a stationary Gaussian process at locations x, with $E[S(x)] = \mu$, $Var[S(x)] = \sigma^2$ and an isotropic correlation function $\rho(h) = Corr[S(x), S(x')]$, where $h = \|x - x'\|$ is the Euclidean distance between the locations x and x'; and the terms ϵ are assumed to be mutually independent and identically distributed $\mathrm{Gau}(0, \tau^2)$. For the correlation function $\rho(h)$ we adopted the exponential function with algebraic form $\rho(h) = \exp\{-h/\phi\}$ where ϕ is the correlation range parameter such that $\rho(h) \simeq 0.05$ when $h = 3\phi$. Within the usual geostatistical *jargon* (Isaaks and Srivastava, 1989) $\tau^2 + \sigma^2$ is the (total) sill, σ^2 is the partial sill, τ^2 is the nugget effect and 3ϕ is the practical range.

Hereafter we use the notation $[\cdot]$ for *the distribution of* the quantity indicated within the brackets. The adopted model defines $[\log(Y)] \sim \mathrm{MVGau}(\mu\mathbf{1}, \Sigma)$, i.e $[Y]$ is multivariate log-Gaussian with covariance matrix Σ parametrised by (σ^2, ϕ, τ^2). Parameter estimates can be obtained by maximum likelihood (Diggle and Ribeiro, 2006). ~~maximising the log-likelihood for this model, given by:~~

~~$l(\mu, \sigma^2, \phi, \tau^2) = -\sum_{i=1}^{n} \log(y_i) - 0.5n\log(2\pi) + \log|\Sigma| + (z_i - \mathbf{1})'\Sigma^{-1}(z_i - \mathbf{1})$.~~

~~Likelihood based methods for geostatistical models are discussed in detail in .~~ For spatial prediction consider first the prediction target $T(x_0) = \exp\{S(x_0)\}$, i.e. the value of the process in the original measurement scale at a vector of spatial locations x_0. Typically x_o defines a grid over the study area. From the properties of the model above the predictive distribution $[T(x)|Y]$ is log-Gaussian with mean μ_T and variance σ_T^2 given by:

$$
\begin{aligned}
\mu_T &= \exp\{\mathrm{E}[S(x_0)] + 0.5\ \mathrm{Var}[S(x_0)]\} \\
\sigma_T^2 &= \exp\{2\ \mathrm{E}[S(x_0)] + \mathrm{Var}[S(x_0)]\}(\exp\{Var[S(x_0)]\} - 1)
\end{aligned}
$$

with

$$
\begin{aligned}
\mathrm{E}[S(x_0)] &= \mu + \Sigma_0'\Sigma^{-1}(Z - \mathbf{1}\mu) \\
\mathrm{Cov}[S(x_0)] &= \Sigma - \Sigma_0'\Sigma^{-1}\Sigma_0
\end{aligned}
$$

where Σ_0 is a matrix of covariances between the ~~the~~ variables at prediction locations x_0 and the data locations x and $\mathrm{Var}[S(x_0)]$ is given by the diagonal elements of $\mathrm{Cov}[S(x_0)]$. In practice, we replace the model parameters in the expressions above ~~are~~ by their maximum likelihood estimates.

Under the model assumptions, $[T|Y]$ is multivariate log-Gaussian and inferences ~~it is therefore possible to make inferences not only~~ about prediction means and variances, or ~~but also about~~ other properties of interest, can be drawn either analytically or, more generally, through conditional simulations. Prediction ~~. Although analytical expressions can be obtained for some particular properties of interest, in general, we use conditional simulationsto compute them. Simulations from $[T|Y]$ are obtained by simulating from the multivariate Gaussian $[S(x_0)|Y]$, and then exponentiating the simulated values. Possible prediction~~ targets can be specified as functionals ~~denoted as functional~~ $\mathcal{F}(S)$ which are applied to the ~~, for which inferences are obtained by computing the quantity of interest on each of the~~ conditional simulations. For instance, inferences on the ~~a functional of particular interest in the present work was the~~ global mean of a ~~the~~ particular realisation of the stochastic process over the area are obtained ~~, which can be predicted~~ by defining x_0 as a grid covering the study area at which conditional simulations of $[S(x_0)|Y]$ are taken; the simulated values are then exponentiated and averaged ~~over the area, obtaining the conditional simulations and computing the mean value for each conditional simulation. More generally other quantities of possible interest as, for instance, the percentage of the area for which the abundance is above a certain threshold, can be computed in a similar manner~~.

2.2 Sampling designs

In general, survey sampling design is about choosing the sample size n and the sample locations x from which data Y can be used to predict any functional of the process. In the case of the ptBTS some particularities must be taken into account: (i) the survey targets several species which may have different statistical and spatial behaviours; (ii) for each species several variables are collected (weight, length, number, etc.); (iii) the sampling is destructive and replicates can not be obtained; (iv) the variability of observed fish abundance is typically high, (v) the planned sampling design may be unattained in practice due to unpredictable commercial fishing activity at the sampling area, bad sea conditions and other ~~possible~~ operational constraints.

Optimal designs can be obtained formally, by defining a criteria and finding the set of sampling locations which minimises some sort of loss function, as e.g. discussed in Diggle and Lophaven (2006). On the other hand, designs can be defined *informally* by arbitrarily defining locations which compromises between statistical principles and operational constraints. Both are valid for geostatistical ~~geostatiscal~~ inference as described in Section 2.1 provided that the locations x are fixed and stochastically independent of the observed variable Y. The above characteristics of the ptBTS makes it very complex to set a suitable criteria to define a loss function to be minimized with relation to ~~w.r.t.~~ the designs. Additionally, costs of a ship at sea are mainly day based and not haul based and increasing the sample size ~~sizes~~ has to consider groups of locations instead of ~~samples instead of the addition of~~ individual points. Therefore, our approach was to construct the proposed designs informally trying to accommodate: (i) historical information about hake and horse mackerel abundance distribution (Anon., 2002; Jardim, 2004)~~()~~, (ii) geostatistical principles about the estimation of correlation parameters (e.g. see Isaaks and Srivastava, 1989; Cressie, 1993; Muller, 2001) and (iii) operational constraints like known trawlable grounds and minimum distance between hauls.

The *study designs* included the design currently adopted for this survey, named "ACTUAL" with 20 locations, and five systematic based sampling designs. The systematic based designs were defined based on two possible increments in the sample size: a $\approx 40\%$ increment, which is expected to be achievable in practice by reducing haul time from 1 hour to 1/2 hour; and a $\approx 60\%$ increment, which could be achieved in practice by adding to the previous increment an allocation of higher sampling density to this area in order to cover the highest variability ~~density~~ of hake recruits historically found within this zone. These designs are denoted by "S" followed by a number corresponding to the sample size. For the former increment a regular design named "S28" was proposed and for the latter three designs were proposed~~for the latter~~: "S45" overlaps the designs ACTUAL and S28, allowing direct comparison with the previous designs; "S44" and "S47" are two infill designs (Diggle and Lophaven, 2006) obtained by augmenting S28 with a set of locations positioned regularly at smaller distances, aiming to better estimate the correlation parameter and, in particular, the noise-to-signal ratio. S44 was built by defining a single denser sampling zone and S47 by adding three areas with denser sampling. A sixth design "S108" was defined to be used as reference with twice the density of S28.

The designs proposed differ in size and spatial configuration and a simple analysis of any statistics would confound these two effects. This situation motivated the development of a statistical approach to compare designs with different ~~A feature of these choices is the possible confounding between the effect of~~ sample sizes and spatial configurations. We used a *ratio of variances* of the relevant estimators between pairs of study designs and random designs with the same sample size, isolating this way the spatial configuration effect. To carry out this analysis we built ~~configuration. We circunvect this problem by building~~ six additional designs with the same sample size as the study designs and with locations randomly chosen within the study area. We denote these by "R" followed by the number of corresponding locations. Each random design contains all the locations of the previous one such that the results are comparable without effects of the random allocation of the sampling locations.

The *study* and corresponding *random* designs are shown in Figure 1.

2.3 Simulation study

The simulation study was carried out in five steps as follows.

Step 1 **Define a set of study designs.** The sampling designs described in Section 2.2 are denoted by $\Lambda_d : d = 1, \ldots, 12$, with $d = 1, \ldots, 6$ for the study designs and $d = 7, \ldots, 12$ for the corresponding random designs, respectively.

Step 2 **Define a set of correlation parameters.** Based on the analysis of historical data of hake and horse mackerel spatial distribution and defining $\tau^2_{REL} = \tau^2/(\tau^2 + \sigma^2)$, a set of model parameters $\theta_p : p = 1, ..., P$ was defined by all combinations of $\cancel{\phi = \{0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4\}^o lat}$ ~~and~~ $\tau^2_{REL} = \{0, 0.1, 0.2, 0.3, 0.4, 0.5\}$ and $\phi = \{0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4\}^o lat$. The values of σ^2 are given by setting $\sigma^2 + \tau^2 = 1$.

Step 3 **Simulate data.** For each parameter set θ_p we obtained S=200 simulations $Y_{ps} : s = 1, \ldots, S$ from $[Y]$ on a regular grid of 8781 locations under the model described in Section 2.1. Each simulation Y_{ps} approximates a possible realisation of the process within the study area from which we computed the mean value μ_{ps}. For each Y_{ps} we extracted the data Y_{pds} at the locations of the sampling designs Λ_d.

Step 4 **Estimate correlation parameters.** For each Y_{pds} obtain maximum likelihood estimates (MLE's) $\tilde{\theta}_{pds}$ of the model parameter.

Step 5 **Simulating from the predictive distribution.** A prediction grid x_0 with 1105 locations and the estimates $\tilde{\theta}_{psd}$ were used to obtain C=150 simulations $\tilde{Y}_{pdsc} : c = 1, \ldots, C$ of the conditional distribution $[T(x_0)|Y]$ which were averaged to produce $\bar{\bar{\tilde{Y}}}_{pdsc}$.

2.4 Analysis of simulation results

The simulation study requires maximum likelihood estimates for the model parameters which are obtained numerically. Therefore a set of summary statistics was computed in order to check the consistency of the results. We have recorded rates of non-convergence of the minimization algorithm; estimates which coincided ~~coincides~~ with the limiting values imposed to the minimization algorithm ($\phi = 3$ and $\tau^2_{REL} = 0.91$); absence of spatial correlation ($\phi = 0$) and values of the parameter estimates which are considered atypical for the problem at hand ($\phi > 0.7$ and $\tau^2_{REL} > 0.67$).

The 48 parameters set (θ_p), 12 sampling designs ($\Lambda_d \cancel{\Delta_d}$), 200 data simulations (Y_{psd}) and 150 conditional simulations ($\tilde{Y}_{psdc}$) produced 17.28 million estimates of abundance which were used to compare the designs. For each design we have computed the estimator $\tilde{\mu}_{psd} = C^{-1} \sum_c \bar{\bar{\tilde{Y}}}_{pdsc}$ of mean abundance μ_{ps} which has variance $\text{Var}(\tilde{\mu}_{psd}) = \bar{\bar{\tilde{\rho}}}_{AA} + \sum_i^n \sum_j^n w_i w_j \tilde{\rho}_{ij} - 2 \sum_i^n w_i \bar{\tilde{\rho}}_{iA}$, where $\bar{\bar{\tilde{\rho}}}_{AA}$ is the mean covariance within the area, estimated by the average covariance between the prediction grid locations (x_0); w are kriging weights; $\tilde{\rho}_{ij}$ is the covariance between a pair of data locations; and $\bar{\tilde{\rho}}_{iA}$ is the average covariance between each data locations and the area discretized by the prediction grid x_0 (Isaaks and Srivastava, 1989).

We used bias, relative bias, mean square error (MSE), confidence intervals coverage and ratio of variances to assess the simulation results, comparing the estimates of the abundance provided by the study designs. For each design these statistics were averaged over all the simulations (s) and parameter sets (p) or groups of parameters sets. Considering the difference between the abundance estimates $\tilde{\mu}_{psd}$ and simulated means μ_{ps}, bias was computed by the difference, relative bias was computed by the difference over the estimate $\tilde{\mu}_{ps}$ and MSE was computed by the square of the difference. For each estimate $\tilde{\mu}_{pds}$ a 95% confidence interval for μ_{ps}, given by $\text{CI}(\tilde{\mu}_{psd}) = \tilde{\mu}_{psd} \pm 1.96\sqrt{Var(\tilde{\mu}_{psd})}$, was constructed and the coverage of the confidence intervals δ were computed by the proportion of the intervals which contained the value of μ_{ps} over all the simulations. This statistic was introduced to help assessing the quality of the variance estimates. At least, we called *ratio of variances* a statistic ξ obtained by dividing the variance $Var(\tilde{\mu}_{psd})$ of each study design by the random design with the same size. Notice that the single difference among each pair of designs with the same size was the spatial configuration of the locations and ξ isolated this effect. Finally we used the results from the six random designs to contrast sampling design based and geostatistical based estimates.

All the analysis were performed with the R software (R Development Core Team, 2005) and the add-on packages geoR (Ribeiro Jr. and Diggle, 2001) and RandomFields (Schlather, 2001).

3 Results

Table 1 summarises the analysis of historical data showing parameter estimates for a sequence of years. This aims to gather information on reasonable values for the model parameters. Notice that units for ϕ are given in degrees and, for the adopted exponential correlation model, the practical range in nautical miles (r) is given by 3ϕ and also included in the table~~Table (r) with units in nautical miles~~. The values of $\tau^2_{REL} = 1$ estimated in some years indicates an uncorrelated spatial process and for such cases estimates of ϕ equals to zero. For most of the cases τ^2_{REL} was estimated as zero due to the lack of nearby locations in the sampling plan and the behaviour of the exponential correlation function at short distances. Given that there is no information in the data about the spatial correlation at distances smaller than the smallest separation distance between a pair of location, this parameter can not be estimated properly and the results depend on the behaviour of the correlation function near the origin.

Table 2 present results used for checking the reliability of the parameter estimates once this could have an impact on the prediction results~~summarizes the checks of the results of the parameter estimates which were considered satisfactory and coherent~~. The highest rate of lack of convergence was 0.6% for the designs ACTUAL and R20. Estimates of ϕ equals to the upper limit imposed to the algorithm were, in the worst case, 0.9% for R28 and R47 and for τ^2_{REL} it was 1.2% for R28. In general there was a slight worst performance of the random designs but this is irrelevant for the objectives of this study. The above ~~Those~~ simulations were not considered for subsequent analysis. Lack or weak spatial correlation given by $\phi = 0$ and/or $\tau^2_{REL} > 0.67$ were ~~was~~ found in about 35% of the simulations for the designs with fewer number of locations, and this rate decreases as the sample size increases, down to below 10% for the largest designs. For both statistics the study designs showed slightly higher values than the corresponding random designs. Identification of weakly correlated spatial processes in part of the simulations was indeed expected to occur given the low values of ϕ (0.05 and 0.1) and high values of τ^2_{REL} (0.5) used in the simulations. The number of cases that presented $\phi > 0.7$ ~~atypical estimates for ϕ~~ were slightly higher for random designs, with a maximum of 2.6% for R44 and R45, but were considered to be within an acceptable range given the high variability of the estimator. Our overall conclusion was that the estimation procedure and algorithms produced parameter estimates which can be trusted for subsequent analysis.

Figure 2 shows square bias, variance and MSE obtained from the estimates of correlation parameters ϕ and τ^2_{REL}. For τ^2_{REL} the majority of the designs presented similar patterns with a small contribution of bias to the MSE and increasing values of MSE for higher true parameter values. The designs ACTUAL, S28 and R20 behaved differently with higher values of bias at low values of τ^2_{REL} that pushed MSE to higher values. As an effect of the sample sizes, the absolute values of MSE defines 3 groups composed by designs with 20 and 28 locations, designs with 44, 45 and 47 locations, and designs with 108 locations; with decreasing values of MSE among them, respectively. MSE increases with the increase of the true value of ϕ and its absolute value decreases slightly with the increasing sample sizes. All designs presented a similar pattern with the variance contributing more than bias to the MSE. The study designs showed a slightly higher relative contribution of the variance to MSE compared with the random designs.

Table 3 shows geostatistical abundance estimates ($\tilde{\mu}$) and their bias, relative bias, variance, MSE and 95% confidence interval coverage for both sets of designs. Additionally the table also shows statistics based on sampling theory obtained for random designs. For subsequent analysis the designs S108 and R108 were regarded just as benchmarks since they are unrealistic for practical implementation. Bias were quite small in all situations and can be considered negligible with higher relative bias of 0.014 for S28. All random designs showed a negative bias whereas all study designs showed a positive one. Variances estimated by study designs were lower than the ones for the corresponding random designs. For random designs the variance decays with increasing sample sizes, whereas study designs behaved differently with S45 presenting the lowest variance followed by ~~with greater differences between S44, S45 and~~ S47~~and R44~~, S44, S28 and S20. MSE showed the same pattern since ~~R45 and R47. The same is valid for MSE , since the~~ bias were small, ~~however with higher absolute values~~ supporting our claim that bias were not relevant for the purpose of this work. The coverages of confidence intervals (δ) were lower than the nominal level of 95% excepted for S108 and R108, reflecting an underestimation of the variance. Considering the designs individually it can be seen that ACTUAL, S28 and S45 showed a lower underestimation than the equivalent random designs. To better investigate this Figure 3 presents values of δ splitted by three levels of correlation (low=$\{0.05, 0.1\}$, med=$\{0.15, 0.20, 0.25\}$, high=$\{0.3, 0.35, 0.4\}$). For geostatistical estimates the coverages δ increases with higher true values of ϕ and larger sample sizes, whereas sampling statistics showed a different pattern, with maximum values for R44 for low and medium correlation levels and for R28 for high correlation levels. This behaviour is more noticeable for stronger spatial correlation, in particular, the largest designs showed lower confidence interval coverage pointing for a more pronounced underestimation of the variance.

Logarithms of the variance ratios between corresponding "S" and "R" designs are presented in Table 3. Without considering S108 for the reasons stated before, the best result was found for S45 (-0.208) and the worst for S28 (-0.108). This must be balanced by the fact that S45 showed a lower variance underestimation than R45, with the opposite happening for S44/R44 and S47/R47, so, in reality, the value of ξ is smaller for S45 than for S44 and S47.

4 Discussion

The choice of sampling designs for BTS is subject to several practical constraints and this has motivated the adoption of *informally* defined designs which accommodated several sources of information like fishing grounds, haul duration, previous knowledge of the spatial distribution of hake and horse mackerel, among others; ~~,~~ which could not be incorporated into a design criteria in an objective way. The fact that this can generate designs with different sample sizes is a drawback of this approach. However, implementing a systematic design on an irregular spatial domain is also ~~likely~~ to provide designs with different sample sizes, depending on the starting location. On the other hand costs ~~Costs~~ of hauling are relatively small when compared with the fixed costs associated with a vessel's working day and increasing sample sizes for a BTS must consider sets of locations which can be sampled in one working day. For these reasons the different sample sizes of each design are not just a feature of the adopted approach but also a result of the BTS particularities.

The confounding effects of sample size and spatial configuration of the proposed designs jeopardized the comparison of their ability in estimating the abundance. To overcome ~~circunvect~~ this limitation a methodology to compare designs with different sample sizes and spatial configurations was required. To deal with this issue we've introduced a mean abundance variance ratio statistic, between the study designs and a corresponding simulated random design with the same sample size.

In fisheries science the main objective for the spatial analysis usually lies in predicting the distribution of the marine resource, aiming, for instance, to define marine protected areas and to compute abundance indices for stock assessment models (Anon., 2004)~~()~~. For such situations the model parameters are not the focus of the study, but just a device to better predict the abundance. Muller (2001) points that the optimality of spatial sampling designs depends on the objectives, showing that ideal designs to estimate covariance parameters of the stochastic process are not the same to predict the value of the stochastic process in a specific location and/or to estimate global abundance. We have not compared the study designs with respect to the estimation of the covariance parameters provided that our main concern was spatial prediction of abundance.

The choice of the parameter estimation method was a relevant issue in the context of this work. The absence of a formal criteria to identify the "best" design naturally led to the use of geostatistical simulations to compare the proposed designs. To carry out a simulation study it is useful to have an objective method capable of producing single estimates of the model parameters. Within traditional geostatistical methods (Isaaks and Srivastava, 1989; Cressie, 1993; Goovaerts, 1997; Rivoirard et al., 2000) ~~(e.g. ,);~~ the estimation entangles subjective analyst's intervention to define some empirical variogram parameters such as lag interval, lag tolerance and estimator for the empirical variogram. Likelihood based inference produces estimates of the covariance parameters without a subjective intervention of the data analyst, allowing for automatization of the estimation process, which is suitable for simulation studies. For the current work we have also used other methods such as restricted maximum likelihood (REML) and weighted least squares, but they have produced worse rates of convergence in the simulation study. In particular the REML presented an high instability with a high frequency of atypical results for ϕ. An aspect of parameter estimation for geostatistical models which is highlighted when using likelihood based methods is regarded to parameter identification due to over-parametrized or poorly identifiable models (see e.g. Zhang, 2004). To avoid over parametrization we used a log-transformation and the process was considered isotropic, avoiding the inclusion of three parameters on the model: the box-cox transformation parameter (Box and Cox, 1964) and the two anisotropy parameters, angle and ratio. The choice of the log transformation was supported by the analysis of historical data and does not impact the comparison of the designs, given that the relative performance of each design will not be affected by the transformation. A point of concern with the log transformation was the existence of zero values which, in the analysis of the historical data, were treated as measurement error and included in the analysis with a translation of the observed values, by adding a small amount to all observations. However, it must be noted this is not always recommended and, in particular, if the stock is concentrated on small schools that cause discontinuities on the spatial distribution, these transformations will not produce satisfactory results. Concerning anisotropy, a complete simulation procedure was carried out considering a fixed anisotropy angle on the north-south direction and an anisotropy ratio of 1, 1.5 or 2. As expected, the absolute values obtained were different but the overall relative performance ~~the designs~~ was the same, supporting our decision to report results only for the isotropic model.

~~Overall, maximum likelihood estimation of the model parameters was considered satisfactory and checks of the consistence of simulation analysis did not reveal major problems with the parameters estimates showing the designs performed equally well and with similar patterns on bias and MSE.~~

A major motivation for performing a simulation study was the possibility to use a wide range of covariance parameters that reflect different spatial behaviours~~, reflecting different possible spatial behaviourswhich implicitly evaluates robustness. Furthermore, the results can be retained for all species with a spatial behaviour covered by these parameters~~. We used two species with different aggregation patterns, hake and horse mackerel, the first an ubiquitous species and the last a more scholastic species, to define the range of the parameters for simulation; suggesting results that can be extended for species with behaviour compatible with the covariance parameters used here.

From a space-time modeling perspective, one of the most interesting analysis for fisheries science is the fluctuation of the stochastic process over time contrasted with the specific realization in a particular time. Therefore the comparison with the mean of the realisations (μ_{ps}) was considered more relevant then to the mean of the underlying process (μ) for the computation of bias and variability. The results showed higher bias for study designs when compared with random designs, but in both cases showing low values which were considered negligible for the purposes of this work. This conclusion was also supported by the fact that MSE showed a similar relative behaviour as variance.

Apart from the design S108, which was introduced as a benchmark and not suitable for implementation, the design that performed better was S45 with lower variance, confidence interval coverage closer to the nominal level of 95% and lower variance ratio (Table 3). One possible reason is the balance between good estimation properties given by the random locations and good predictive properties given by the systematic locations, however the complexity of the BTS objectives makes it impossible to find a full explanation for this results. A possible indicator of the predictive properties is the average distance between the designs and the prediction grid locations, which reflects the extrapolation needed to predict over a grid. We found that S45 had an average of $2.61nm$ whereas for S47 the value is $2.72nm$, explaining in part the S45 performance.

These results are in agreement with Diggle and Lophaven (2006) who showed that *lattice plus closed pairs* designs (similar to S45) performed better than *lattice plus in-fill* designs (similar to S44 and S47) for accurate prediction of the underlying spatial phenomenon. The combination of random and systematic designs like S45 is seldom considered in practice and we are not aware of recommendations of such designs for BTS.

It was interesting to notice that most designs presented a coverage of confidence intervals below the nominal level of 95% revealing the variances were underestimated. It was not fully clear how to use such results to correct variance estimation and further investigation is needed on the subject. Care must be taken when looking at variance ratios since underestimated denominators will produce higher ratios which can mask the results. This was the case of S45 when comparing to S47 and S44, supporting our conclusions about S45.

Another result of our work was the assessment of abundance estimates from random designs by sampling statistics, the most common procedure for fisheries surveys (Anon., 2004)~~()~~, under the presence of spatial correlation. In such conditions an increase in sample size may not provide a proportional increase in the quantity of information due to the partial redundancy of information under spatial correlation. Results obtained for coverages of confidence intervals illustrated this (Table 3 and Figure 3), with smaller coverages for larger sample sizes and higher spatial correlation. In our opinion this is due to an ~~, reflecting an over estimation of the degrees of freedom. The~~ overestimation of the degrees of freedom that lead ~~led~~ to an underestimation of prediction standart errors producing the smaller coverages. These findings ~~fundings~~ support claims to consider geostatistical methods to estimate fish abundance, such that correlation between locations is explicitly considered in the analysis, and highlighting the importance of verifying the assumptions behind ~~behing~~ sampling theory before computing the uncertainty of abundance estimates.

5 Acknowledgements

The authors would like to thank the scientific teams evolved in the Portuguese Bottom Trawl Surveys, in particular the coordinator Fátima Cardador, and the comments by Manuela Azevedo. This work was carried out within the IPIMAR's project NeoMAv (QCA-3/MARE-FEDER, http://ipimar-iniap.ipimar.pt/neomav) and was co-financed by project POCTI/MATH/44082/2002.

References

Anon., 2002. Report of the International Bottom Trawl Survey Working Group. Tech. rep., International Council for the Exploitation of the Sea (ICES), ICES CM 2002/D:03.

Anon., 2003. Report of the International Bottom Trawl Survey Working Group. Tech. rep., International Council for the Exploitation of the Sea (ICES), ICES CM 2003/D:05.

Anon., 2004. Report of the Workshop on Survey Design and Data Analysis. Tech. rep., International Council for the Exploitation of the Sea (ICES), ICES CM 2004/B:07.

Box, G., Cox, D., 1964. An Analysis of Transformations. Journal of the Royal Statistical Society Series B 26, 211–243.

Christensen, O., Diggle, P., Ribeiro Jr, P., 2001. Analysing positive-valued spatial data: the transformed gaussian model. In: Monestiez, P., Allard, D., Froidevaux (Eds.), GeoENV III - Geostatistics for Environmental Applications. Vol. 11 of Quantitative Geology and Geostatistics. Kluwer, pp. 287–298.

Cressie, N., 1993. Statistics for spatial data - Revised Edition. John Wiley and Sons, New York.

Diggle, P., Ribeiro, P., 2006. Model-based Geostatistics. Springer, New York, in press.

Diggle, P. J., Lophaven, S., 2006. Bayesian geostatistical design. Scandinavian Journal of Statistics 33, 55–64.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York.

Hansen, M., Madow, W., Tepping, B., 1983. An Evaluation of Model-Dependent and Probability-Sampling Inferences in Sample Surveys. Journal of the American Statistical Association 78 (384), 776–793.

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

Jardim, E., 2004. Visualizing hake recruitment - a non-stationary process. In: Sanchez-Vila, X., Carrera, J., Gómez-Hernandéz, J. J. (Eds.), geoENV IV - Geostatistics for Environmental Applications. Vol. 13 of Quantitative Geology and Geostatistics. Kluwer Academic Publishers, London, pp. 508–509.

Muller, W., 2001. Collecting Spatial Data - Optimum Design of Experiments for Random Fields, 2nd Edition. Contributions to statistics. Physica-Verlag, Heidelberg.

R Development Core Team, 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.
URL `http://www.R-project.org`

Ribeiro Jr., P., Diggle, P., June 2001. geoR: a package from geostatistical analysis. R-NEWS 1 (2), 15–18.
URL `http://cran.R-project.org/doc/Rnews`

Rivoirard, J., Simmonds, J., Foote, K., Fernandes, P., Bez, N., 2000. Geostatistics for Estimating Fish Abundance. Blackwell Science, London, England.

Schlather, M., June 2001. Simulation and analysis of random fields. R News 1 (2), 18–20.
URL `http://CRAN.R-project.org/doc/Rnews/`

SESITS, 1999. Evaluation of Demersal Resources of Southwestern Europe from Standardised Groundfish Surveys - Study Contract 96/029. Final report, DG XIV / EC.

Thompson, S., 1992. Sampling. Statistics. John Wiley & Sons, INC, New York.

Zhang, H., 2004. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. Journal of the American Statistical Association 99 (465), 250 – 261.

Table 1: Exponential covariance function parameters ($\phi, \tau^2_{\mathrm{REL}}$) and the geostatistical range (r) estimated yearly (1990-2004) for hake and horse mackerel abundance. The values of ϕ are presented in degrees of latitude and range in nautical miles. The maximum distance between pairs of locations was 63nm.

	Hake			Horse mackerel		
	ϕ(^{o}lat)	r(nm)	τ^2_{REL}	ϕ(^{o}lat)	r(nm)	τ^2_{REL}
1990	0.05	9.1	0.01	0.42	76.4	0.00
1991	0.14	24.4	0.63	0.49	88.9	0.43
1992	0.00	0.0	1.00	0.22	39.3	0.05
1993	0.05	9.3	0.00	0.00	0.0	1.00
1995	0.05	8.8	0.00	0.08	14.4	0.00
1997	0.14	24.8	0.00	0.21	38.6	0.42
1998	0.02	3.4	0.00	0.09	16.5	0.00
1999	0.10	17.8	0.00	0.09	16.0	0.00
2000	0.03	4.6	0.00	0.16	29.5	0.00
2001	0.07	12.9	0.00	0.42	75.7	0.06
2002	0.00	0.0	1.00	0.05	8.9	0.00
2003	0.33	59.0	0.00	0.34	62.0	0.00
2004	0.09	15.4	0.00	0.09	17.0	0.00

Table 2: Statistics to provide simulation quality assessment (in percentages) for both design sets and all sample sizes: non-convergence of the minimization algorithm (non-conv); cases truncated by the limits imposed to the minimization algorithm ($\phi = 3$ and $\tau^2_{\mathrm{REL}} = 0.91$); uncorrelated cases ($\phi = 0$); and atypical values of the correlation parameters ($\phi > 0.7$ and $\tau^2_{\mathrm{REL}} > 0.67$).

statistic	design	sample size					
		20	28	44	45	47	108
non-conv	study	0.6	0.5	0.2	0.2	0.2	0.1
	random	0.6	0.4	0.2	0.2	0.2	0.1
$\phi = 3$	study	0.7	0.5	0.7	0.7	0.5	0.2
	random	0.6	0.9	0.8	0.8	0.9	0.1
$\tau^2_{\mathrm{REL}} = 0.91$	study	0.7	0.7	1.0	0.9	0.8	0.4
	random	0.8	1.2	1.1	1.1	1.1	0.2
$\phi = 0$	study	36.3	33.0	20.7	20.6	18.0	5.3
	random	32.8	28.5	18.1	17.2	16.2	3.3
$\phi > 0.7$	study	1.3	1.6	1.9	1.9	1.8	1.4
	random	1.8	2.2	2.6	2.6	2.4	1.7
$\tau^2_{\mathrm{REL}} > 0.67$	study	38.5	35.8	24.2	24.7	21.8	10.0
	random	35.0	31.6	22.1	21.1	20.3	7.6

Table 3: Summary statistics per sets of sampling designs and sample size. Geostatistical abundance estimates ($\tilde{\mu}$) in kg/hour, bias (bias($\tilde{\mu}$)), relative bias (bias$_r$($\tilde{\mu}$)), variance (var($\tilde{\mu}$)), mean square error (MSE) and 95% confidence interval coverage ($\delta(\tilde{\mu})$). Mean log variance ratios per sampling design type (ξ) measures the relative log effect of the systematic based designs configuration with relation to the random designs. The last six rows present the same statistics estimated for random designs by sampling statistics.

method	statistic	design	number of locations					
			20	28	44	45	47	108
geostatistics	$\tilde{\mu}$	study	1.658	1.662	1.649	1.657	1.651	1.641
		random	1.631	1.624	1.625	1.624	1.625	1.625
	bias($\tilde{\mu}$)	study	0.025	0.030	0.016	0.026	0.019	0.008
		random	-0.001	-0.008	-0.007	-0.009	-0.008	-0.007
	bias$_r$($\tilde{\mu}$)	study	0.012	0.014	0.003	0.012	0.005	0.001
		random	-0.004	-0.008	-0.005	-0.006	-0.005	-0.005
	var($\tilde{\mu}$)	study	0.136	0.108	0.092	0.086	0.089	0.081
		random	0.168	0.129	0.113	0.112	0.112	0.097
	MSE($\tilde{\mu}$)	study	0.272	0.196	0.164	0.144	0.154	0.104
		random	0.321	0.230	0.173	0.171	0.171	0.124
	$\delta(\tilde{\mu})$	study	0.908	0.922	0.907	0.939	0.920	0.960
		random	0.895	0.909	0.937	0.934	0.934	0.954
	ξ	stu/rnd	-0.128	-0.107	-0.150	-0.208	-0.179	-0.228
sampling statistics	$\bar{Y}$	random	1.615	1.619	1.618	1.616	1.618	1.622
	bias($\bar{Y}$)	random	-0.017	-0.014	-0.014	-0.017	-0.015	-0.010
	bias$_r$($\bar{Y}$)	random	-0.017	-0.014	-0.013	-0.014	-0.014	-0.006
	var($\bar{Y}$)	random	0.197	0.146	0.091	0.088	0.085	0.037
	MSE($\tilde{\mu}$)	random	4.133	4.238	4.109	4.083	4.090	4.073
	$\delta(\bar{Y})$	random	0.900	0.910	0.908	0.900	0.896	0.840

Figure 1: Sampling designs and the study area (southwest of Portugal). Each plot shows the sample locations, the bathymetric ~~bathimetric~~ isoline of 500m and 20m and the coast line. The sampling design name is presented on the top left corner of the plots. The top row shows the *study* designs and the bottom row the random designs.

Figure 2: Summary statistics for the covariance parameters estimation by sampling design as a function of the true parameter values. bias^2 ($\circ$), variance ($\triangle$) and mean square error ($+$). Top figure presents τ^2_{REL} results and bottom figure ϕ.

Figure 3: Coverage of the confidence intervals (δ) for different ϕ levels (low $= \{0.05, 0.1\}$, med$\{0.15, 0.20, 0.25\}$ high $= \{0.30, 0.35, 0.40\}$) for estimates of abundance by sampling statistics for the random designs ($+$) and by geostatistics for the study (o) and random designs ($*$).

FIGURE 01

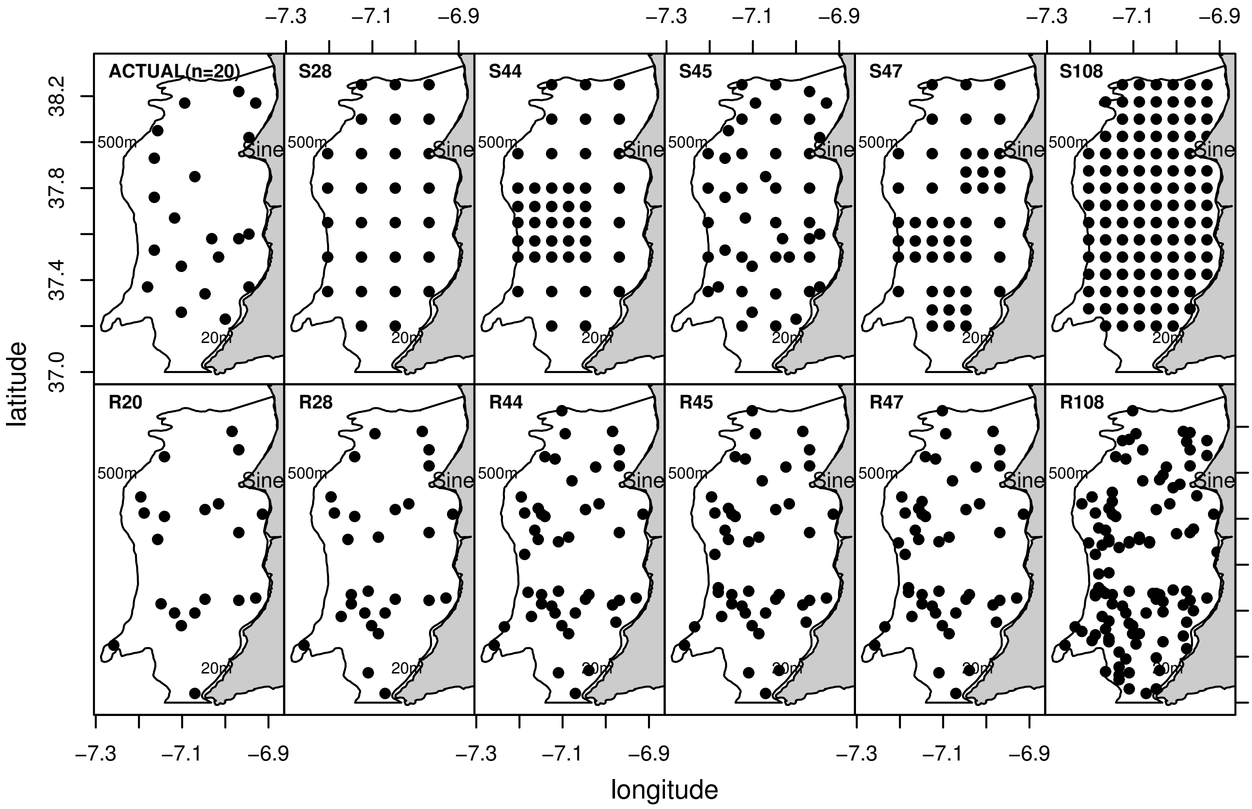
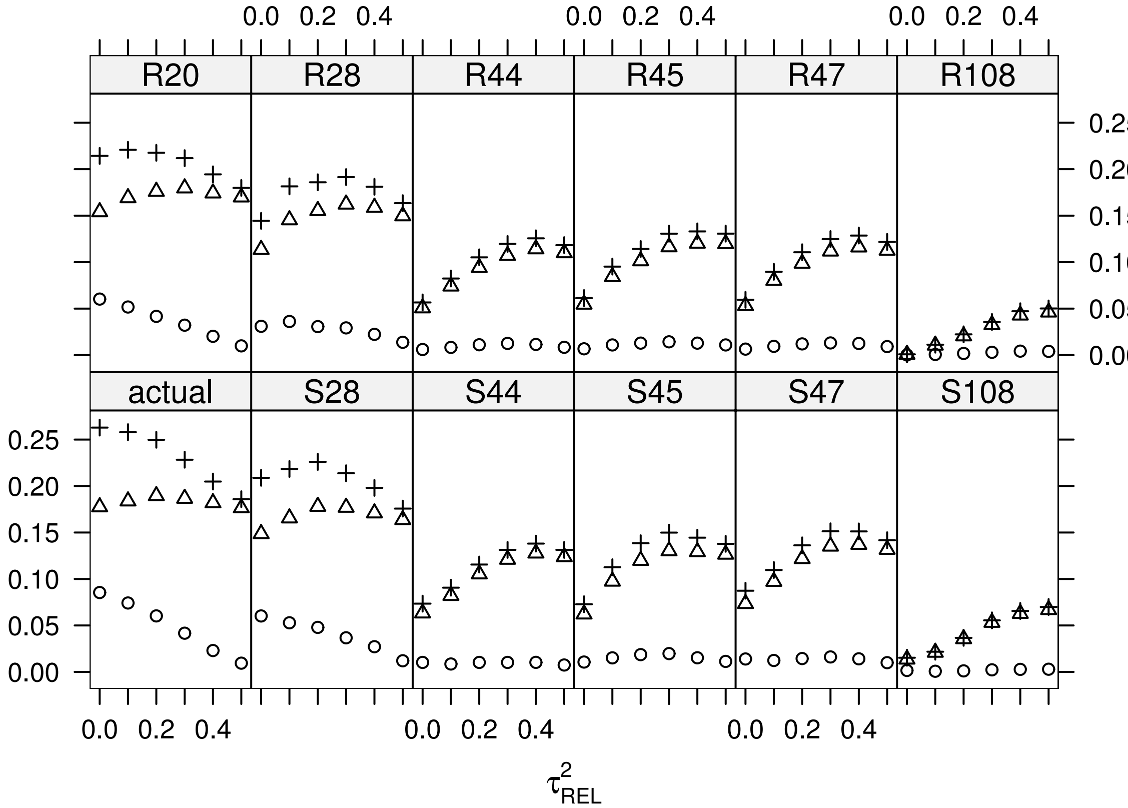

ACTUAL(n=20)
S28
S44
S45
S47
S108
R20
R28
R44
R45
R47
R108
500m
Sine
20m
latitude
longitude
−7.3
−7.1
−6.9
37.0
37.4
37.8
38.2

FIGURE 02


R20
R28
R44
R45
R47
R108
actual
S28
S44
S45
S47
S108
τ^2_{REL}

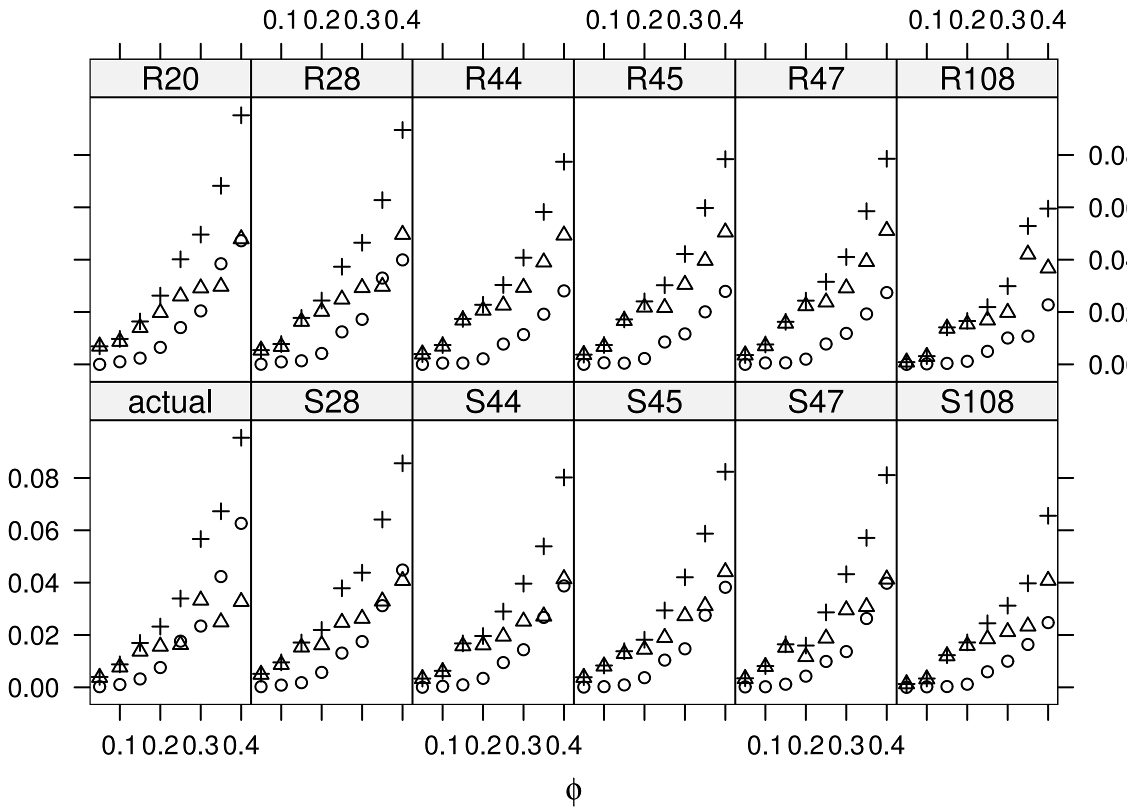

R20
R28
R44
R45
R47
R108
actual
S28
S44
S45
S47
S108
ϕ

FIGURE 03