6-15-2004

Bayesian Geostatistical Design

Peter J. Diggle

Medical Statistics Unit, Lancaster University, UK & Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, p.diggle@lancaster.ac.uk

Soren Lophaven

Informatics and Mathematical Modelling, Technical University of Denmark

Bayesian Geostatistical Design

Peter Diggle[1] and Søren Lophaven[2*]

June 15, 2004

[1] Department of Mathematics and Statistics, Lancaster University
and Depetmant of Biostatistics, Johns Hopkins University
[2] Informatics and Mathematical Modelling, Technical University of Denmark

Abstract

This paper describes the use of model-based geostatistics for choosing the optimal set of sampling locations, collectively called the *design*, for a geostatistical analysis. Two types of design situations are considered. These are retrospective design, which concerns the addition of sampling locations to, or deletion of locations from, an existing design, and prospective design, which consists of choosing optimal positions for a new set of sampling locations. We propose a Bayesian desin criterion which focuses on the goal of efficient spatial prediction whilst allowing for the fact that model parameter values are unknown. The results show that in this situation a wide range of inter-point distances should be included in the design, and the widely used regular design is therefore not the optimal choice.

KEY WORDS: *model-based geostatistics, Bayesian inference, spatial design*

1 Introduction

A central concern of geostatistics is the prediction of a spatial surface over a region A, using data which consist of possibly imperfect measurements, $Y_i : i = 1, ..., n$, of the values on the surface at a finite collection of *sampling locations*, $x_i \in A : i = 1, ..., n$. A widely used stochastic model for such data is the linear Gaussian model. This assumes that the spatial surface of interest is a realisation of an unobserved stationary Gaussian process $\{S(x) : x \in \mathbb{R}^2\}$ such that, conditional on $S(\cdot)$ the Y_i are mutually independent, Normally distributed with means $\mathrm{E}[Y_i] = S(x_i)$ and common variance τ^2. The process $S(\cdot)$ has mean-value field $\mu(x) = d(x)'\beta$ for some spatially indexed set of explanatory variables $d(x)$, variance σ^2 and correlation function $\rho(u; \phi) = \mathrm{Corr}\{S(x), S(x - u)\}$, where ϕ may be vector-valued. The linear Gaussian model is implicitly assumed in the geostatistical technique known as *universal kriging*, whilst the special case in which $\mu(x) = \mu$ for all x corresponds, according to the treatment of unknown parameters in the model, to *simple* or *ordinary* kriging (Chiles and Delfiner, 1999).

Bayesian inference provides a natural framework for the incorporation of the effects of parameter uncertainty into spatial prediction. Bayesian methods for geostatistical analysis using the linear Gaussian model were proposed independently by Kitanidis (1986), Le and Zidek (1992) and Handcock and Stein (1993). Diggle, Moyeed and Tawn (1998) extended the Bayesian methodology to embrace generalized linear models (McCullagh and Nelder, 1989) with an unobserved Gaussian process $S(\cdot)$ in the linear predictor.

The set of sampling locations is collectively called the *design*. An important additional assumption in most geostatistical analyses is that the design is independent of the unobserved process $S(\cdot)$.

It is well known that designs which are efficient for parameter estimation are not necessarily efficient for spatial prediction given the model parameters (Müller, 2001; Martin, 2001). Our goal in this paper is to develop a Bayesian approach to design which integrates and balances these competing considerations in a natural way. Briefly, the idea is to find designs which are efficient for spatial prediction whilst making proper allowance for the effects of parameter uncertainty. Hence, parameter estimation is treated as a means to the primary end of spatial prediction, rather than as an end in itself. In our experience, most applications of geostatistics are of this kind. However, if in a particular application parameter estimation in itself were the scientific goal, our approach would be inappropriate.

We distinguish between two common situations. The first, which we call retrospective design, arises when locations are to be deleted from, or added to, an existing design. For example, the deletion case often arises in environmental monitoring, with the aim of economising on the long-term costs of maintaining a network of monitoring sites whilst minimising the consequent loss of information. The second situation, which we call prospective design, arises when we need to design in advance of data-collection.

In Section 2 of the paper, we review classical approaches to the geostatistical design problem. In Section 3, we set out our Bayesian approach and discuss its implementation for retrospective and prospective design problems. In Section 4 we give two examples of retrospective design problems: a simulated example in which we compare the results obtained using classical and Bayesian criteria; and an application to the deletion of locations from a network of 70 monitoring stations measuring surface salinity in the Kattegat basin. In Section 5, we discuss two classes of prospective design, the lattice plus close pairs and lattice plus in-fill designs. Both are motivated by the idea that regular lattice designs are generally efficient for prediction when model parameters are known, whereas adding groups of closely spaced sample locations to the design is advantageous for parameter estimation. The design problem is then to find the optimum balance between lattice and non-lattice locations, for a fixed total number. The ideas in the paper are essentially independent of the chosen model, but the associated computations are routine only for the linear Gaussian model.

2 Classical design criteria

Traditionally, geostatistical design has involved the minimisation of a design criterion which is some suitable function, for example the mean or maximum, of the prediction variance, where predictions are computed by conventional kriging assuming that all parameters defining the covariance structure of the data are known. McBratney, Webster and Burgess (1981) and

2

McBratney and Webster (1981) addressed the problem of choosing the spacing of a regular rectangular or triangular lattice design to achieve an acceptable value of the maximum of the prediction variance over the region of interest. Winkels and Stein (1997) applied this approach so as to achieve a predefined accuracy in the determination of the spatial distribution of a number of contaminants in lake sediments. Spruill and Candela (1990) considered the effect on prediction accuracy, as measured by both the average and the maximum of the prediction variance, of deleting locations from an existing network for monitoring the concentration of chloride in groundwater. They concluded that the number of sampling locations could be reduced from 120 to 99 with only a marginal increase in their two design criteria. Ben-jemaa, Mariño and Loaiciga (1995) also considered the removal of locations from an existing network for monitoring mercury concentrations in lake sediments. They used ordinary cokriging for prediction, so as to exploit the correlation between the sediment concentration of mercury and a sediment grain size index.

Royle and Nychka (1998) describe a different approach, in which the design criterion is purely geometric, rather than involving an assumed geostatistical model. In qualitative terms, this approach favours a regular spatial distribution of sampling locations within the constraints imposed by the particular application. Hence, for example, in considering the addition or deletion of locations to or from an existing network, additional locations would tend to be placed so as to fill empty spaces, whilst prime candidates for removal would include one of a pair of near neighbours. Based on results in Johnson, Moore and Ylvisaker (1990) they gave a theoretical justification that the resulting designs are nearly optimal for spatial prediction.

In contrast to the above, Russo (1984), Warrick and Myers (1987) and Zimmerman and Homer (1991) considered the design problem from the point of view of semivariogram estimation. The *semivariogram* of a stationary spatial stochastic process $Y(x)$ is the quantity $V(u) = \frac{1}{2}\mathrm{E}[\{Y(x) - Y(x - u)\}^2]$, considered as a function of u. In traditional geostatistical practice, estimation of parameters of $V(u)$ is achieved by matching the semivariogram to its empirical counterpart, in which observed quantities $v_{ij} = \frac{1}{2}(Y_i - Y_j)^2$ are averaged within binned ranges of the distances, and angles if the assumed model includes directional effects, between corresponding pairs of sampling locations.

In this context, Russo (1984) suggested designing to minimise the variance of the distances included within each bin, whilst Warrick and Myers (1987) sought to achieve large numbers of contributions to each distance-angle bin together with small variances of the distances and angles within each. Zimmerman and Homer (1991) proposed an explicit design criterion for parametric estimation based on the empirical semivariogram. Their design criterion C is given by

$$C = \det\{D^T(V^TWV)^{-1}D\}$$

where V is a matrix of partial derivatives of the chosen semivariogram model with respect to its parameters θ, D is a matrix of partial derivatives of specified parametric functions g_j with respect to the parameters θ, and the g_j identify the semivariogram attributes which are of particular interest. The matrix W is diagonal, with non-zero elements equal to the reciprocal squared values of the assumed semi-variogram. The rationale for this is that under Gaussian assumptions for $Y(x)$, the sampling distributions of the empirical semivariogram ordinates v_{ij} are proportional to χ_1^2, with variance proportional to $V(u)^2$. Müller and Zimmerman

3

(1999) used a modification of the design criterion C in which W is non-diagonal whilst D was taken to be the identity matrix.

Lark (2002) considers the design problem from the point of view of maximum likelihood estimation for the Gaussian linear model. He considers the prediction variance for $S(x)$, say $V(x, \theta)$ at a target location x as a function of the model parameters θ. His design criterion is the variance of $V(x, \hat{\theta})$ induced by the asymptotically multivariate Normal sampling distribution of $\hat{\theta}$. The resulting designs typically combine a fairly regular array supplemented by groups of closely spaced locations. This is qualitatively similar to some of the results which we report below, although their theoretical basis is very different.

In many applications of geostatistical methods, spatial prediction is a primary goal whilst parameter estimation is largely a means towards this end. The goal of the present paper is therefore to construct designs which are efficient for spatial prediction whilst protecting against the consequences of grossly inefficient parameter estimation, rather than considering parameter estimation and spatial prediction as disjoint activities. Our Bayesian design approach achieves this by taking account of parameter uncertainty in a natural way.

3 A Bayesian approach

The specification of the classical linear model has the symbolic representation $[S, Y] = [S][Y|S]$ where $[\cdot]$ means "the distribution of", i.e. the model specifies the marginal distribution of the unobserved random field $S = \{S(x) : x \in \mathbb{R}^2\}$ and the conditional distribution of $Y = (Y_1, ..., Y_n)$ given S, with model parameters θ regarded as unknown, but fixed in value. Under the Bayesian paradigm, unknown model parameters are also treated as random variables and the symbolic representation is extended to

$$[S, Y, \theta] = [S|\theta][Y|S, \theta][\theta]$$

where $[\theta]$ denotes the prior for θ.

The required distribution for predictive inference is $[S|Y]$, which we evaluate as

$$[S|Y] = \int [S|Y, \theta][\theta|Y]d\theta. \tag{1}$$

Equation (1) shows that the Bayesian predictive distribution is a weighted average of classical, or plug-in, predictive distributions $[S|Y, \theta]$ with different values of θ weighted according to their posterior probabilities.

3.1 Retrospective design

We shall use as our design criterion the spatially averaged prediction variance,

$$\bar{v} = \int_A \text{Var}(S(x)|Y)dx. \tag{2}$$

The data from the pre-existing design provide partial knowledge of the unknown model parameters. Hence, for each candidate design we calculate the variance term in (2) from the

4

predictive distribution $[S|Y]$ given by (1). Typically, $[S|Y]$ will not be expressible analytically, and (2) is calculated approximately, using Monte Carlo methods. Diggle, Ribeiro and Christensen (2003) describe how to generate direct Monte Carlo samples from the posterior distribution for θ and hence from the predictive distribution $[S|Y]$ when the data are generated from the linear Gaussian model described in Section 1. In the examples below, we use their methods to calculate a Monte Carlo approximation to the retrospective design criterion (2) by repeated sampling from $[S|Y]$ where S represents the values of $S(x)$ at locations x on a regular lattice to cover A. We also approximate the spatial integral in (2) by a scaled sum over the lattice points x.

3.2   Prospective design

For the prospective design problem, the data Y do not yet exist. We therefore propose to use as our design criterion the expectation of the spatially averaged prediction variance (2) with respect to the marginal distribution of Y. The design criterion is now $\mathrm{E}[\bar{v}]$ where $\bar{v}$ is given by (2) and the expectation it taken with respect to the distribution of Y at the true parameter values, θ_0 say. Hence,

$$\mathrm{E}(\bar{v}) = \int_A \mathrm{E}_{Y|\theta_0}[\mathrm{Var}(S(x)|Y)]dx. \tag{3}$$

We evaluate (3) using Monte Carlo methods as follows. We first simulate s independent data-sets, $Y_k : k = 1, ..., s$ from the model with parameter values $\theta_0 = (\mu_0, \sigma_0^2, \phi_0, \tau_0^2)$. From each simulated data-set, we calculate the corresponding value, $\bar{v}_k$, of the design criterion (2) as described in Section 3.1 above. We then use the sample average of the $\bar{v}_k$ as an approximation to $\mathrm{E}[\bar{v}]$

4   Retrospective design: optimum deletion of design points

In this Section, we first compare designs obtained by applying classical and Bayesian approaches to a simulated data-set, where the goal is retrospective reduction of the number of locations from 50 to 20 with minimal loss of predictive accuracy. We then apply the Bayesian approach to an existing monitoring network for measurement of salinity in the Kattegat basin, with the aim of reducing the number of locations from 70 to 20. In both cases, consideration of all possible designs would be computationally prohibitive. We therefore adopt the pragmatic strategy of sequentially removing one location at a time so as to minimise the spatially averaged prediction variance at each stage.

Our simulation model, and assumed model for the Kattegat salinity data, is the linear Gaussian model with exponential correlation function, $\rho(u; \phi) = \sigma^2 \exp(-|u|/\phi)$.

5

4.1 Simulated data

The simulated data were generated at the 50 sampling locations shown in Figure 1, using the linear Gaussian model with a spatially constant mean, $\beta = 0$, scale parameter $\sigma^2 = 1$ and exponential correlation function with parameter $\phi = 0.3$. For the noise-to-signal variance ratio, $\nu^2 = \tau^2/\sigma^2$, we used each of the values $\nu^2 = 0$, 0.3 and 0.6.

Our design criterion was the average prediction variance over a regular 6×6 grid. For the classical design approach we evaluated the prediction variance using ordinary kriging for prediction. This is equivalent to a Bayesian approach using an improper uniform prior for β but assuming that the remaining parameters ϕ, σ^2 and ν^2 are known. For the Bayesian approach our prior for ϕ was uniform on $(0, 2.35)$, whilst for $(\beta, \sigma^2|\phi)$ we used a diffuse prior proportional to $1/\sigma^2$. For the ratio ν^2 we compared results obtained by assuming known ν^2 and by assigning to ν^2 a prior uniform distribution on $(0, 1)$. Posteriors were computed by direct simulation as described in Diggle, Ribeiro and Christensen (2003).

The resulting designs when we reduce the number of locations from 50 to 20 are shown in Figure 2, for three different values of ν^2. In all three cases, the designs based on the classical design criterion consist of well-separated monitoring sites, as a consequence of their focussing on spatial prediction as if the geostatistical model were known. In contrast, the designs constructed using the Bayesian approach retain some close pairs of sites, representing a compromise between designing for prediction and for parameter estimation. Within the Bayesian approach, the designs show differences in detail according to the value of ν^2 and whether or not we treat ν^2 as known or unknown, but in all cases they exhibit less spatial regularity than their classical counterparts.

Table 1 compares the achieved values of the Bayesian design criterion (2), treating all model parameters as unknown, for each of the designs shown in Figure 2. On their own terms, the Bayesian designs are between five and ten times more efficient than the classical designs, whereas the differences between the Bayesian designs which treat ν^2 as known or unknown are small. When the number of sampling locations in the design is large, we would expect there to be less need to compromise between prediction and estimation, leading to smaller differences between the classical and Bayesian designs.

4.2 Monitoring salinity in the Kattegat basin

Our second example deals with optimum deletion of monitoring stations measuring salinity in the Kattegat basin. The current monitoring network consists of 70 stations located as shown in Figure 3. We now examine which monitors should be removed in order to reduce the network to 20 stations. A geostatistical analysis showed a north-south trend in the salinity data, and for the design evaluations we therefore included in the assumed model a linear trend surface, $\mu(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ where $x = (x_1, x_2)$, and write $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)$. We fixed the ratio $\nu^2 = \tau^2/\sigma^2$ at its estimated value, $\nu^2 = 0.42$, partly to economise on computation but also because we have found in simulated examples such as the one reported above that fixing ν^2 generally has a small impact on the chosen design. For ϕ we used a uniform prior on the interval from 10 to 100 kilometres, whilst for $(\boldsymbol{\beta}, \sigma^2|\phi)$ we again used a diffuse prior proportional to $1/\sigma^2$. We computed the spatially averaged prediction variance (2) from 95 locations in a regular grid covering the Kattegat area at a spacing of 15 kilometres. The

6

resulting network of 20 retained monitoring stations is shown in Figure 3. It consists largely of well-separated stations, but with some pairs of closely located stations, again reflecting a compromise between designing for prediction and for estimation.

5 Prospective design: augmented lattice designs

We now consider two classes of prospective geostatistical design which, in different ways, reflect the need to compromise between regularly spaced sampling locations for predictive accuracy under a known model, and irregularly spaced locations for better estimation of unknown model parameters.

The *lattice plus close pairs* design consists of locations in a regular $k \times k$ lattice at spacing Δ together with a further m points, each of which is located uniformly at random within a disc of radius $\delta = \alpha\Delta$ whose centre is at a randomly selected lattice location. We use the notation $(k \times k, m, \alpha)$, noting that from the design point of view, the choice of the distance scale is arbitrary, hence Δ is irrelevant. Figure 4A shows a $(7 \times 7, 15, 0.5)$ lattice plus close pairs design on the unit square.

The *lattice plus in-fill* design consists of locations in a regular $k \times k$ lattice at spacing Δ together with further locations in a more finely spaced lattice within m randomly chosen cells of the primary lattice. We use the notation $(k \times k, m, r \times r)$, where each in-filled lattice cell consists of an $r \times r$ lattice and therefore involves $r^2 - 4$ additional locations. Figure 4B shows a $(7 \times 7, 3, 3 \times 3)$ lattice plus in-fill design on the unit square.

We now illustrate the efficiency of lattice plus close pairs and lattice plus in-fill designs relative to a regular lattice design when the total number of locations is fixed at N=64. Specifically, we compare the regular 8×8 lattice with the $(7 \times 7, 15, 0.5)$ lattice plus close pairs design and the $(7, 3, 3 \times 3)$ lattice plus in-fill design All designs were constructed on a unit square region, with the lattice spacing adjusted accordingly.

Although designs of the same type vary because of the random selection of the secondary locations, we found that this had only a small impact on the spatially averaged prediction variance. In evaluating the design criterion (3) for either the lattice plus close pairs or the lattice plus in-fill design, we therefore averaged over 5 independent replicates. For each replicate design, we considered the linear Gaussian model with constant mean β and exponential correlation function, $\rho(u; \phi) = \sigma^2 \exp(-|u|/\phi)$.

We specified the priors for the model parameters as follows. For ϕ we used a uniform prior on $(0, 1.3)$. For $(\beta, \sigma^2|\phi)$ we used a diffuse prior proportional to $1/\sigma^2$. For the ratio $\nu^2 = \tau^2/\sigma^2$ we used a uniform prior on $(0, 1)$.

We evaluated the design criterion (3) for each of the three candidate designs with true parameter values $\beta = 0$, $\sigma^2 = 1$, $\phi = 0.2, 0.4, 0.6, 0.8, 1.0$ and $\nu^2 = 0.0, 0.2, 0.4, 0.6, 0.8$. The results are summarised in Figure 5. We see that the lattice plus close pairs design results in lower values of the design criterion compared to the lattice plus in-fill design and the regular 8×8 lattice, meaning that predictions are computed more accurately from the lattice plus close pairs design. In contrast, the performance of the lattice plus in-fill design is only slightly better than that of the regular lattice.

7

6 Discussion

Our aim in this paper has been to propose and illustrate an approach to geostatistical design when the scientific goal is accurate prediction of the underlying spatial phenomenon, while acknowledging that parameter values in the assumed geostatistical model are unknown. This necessarily involves a compromise, which the Bayesian paradigm achieves in a natural way through its usual device of specifying prior distributions for unknown model parameters, and by including posterior parameter uncertainty in the evaluation of predictive accuracy.

For the retrospective design problem, in which locations are to be deleted from an existing design with minimally deleterious effects, we have shown that the Bayesian approach leads to designs which are materially different, in an intuitively sensible way, from designs derived on the assumption that model parameters are known.

For the prospective design problem, we have defined two classes of design, the lattice plus close pairs design and the lattice plus in-fill design. Each of these formalises the widely accepted view that in order to compromise between prediction accuracy and efficient parameter estimation, good geostatistical designs should include close pairs or groups of sampling locations within an otherwise spatially regular arrangement (Martin, 2001; Müller, 2001; Lark, 2002). Our simulation experiments suggest that the lattice plus close pairs design brings a clear benefit, whereas the performance of the lattice plus in-fill design is only slightly better than that of the regular lattice. In this context, it is noteworthy that the principle example in Diggle, Moyeed and Tawn (1998) used a lattice plus in-fill design on the recommendation of the first author who, with the benefit of hindsight, would have favoured a lattice plus close pairs design.

Our approach is computationally intensive. For this reason, we have illustrated the method using the simplest non-trivial class of geostatistical models, with four parameters representing a mean, two variance components and a rate of decay of spatial correlation with distance. However, the computations would remain feasible in any applications requiring a model with a larger number of parameters.

The software currently available to us (http://www.r-project.org/; Ribeiro, Christensen and Diggle, 2003) would encounter serious difficulties in evaluating designs with numbers of locations larger than a few hundred. However, work by Rue (2001) and Rue and Tjelmeland (2002) is leading to very substantial gains in the speed of computation for geostatistical models.

Acknowledgements

8

References

Ben-Jemaa, F., Mariño, M. and Loaiciga, H. (1995). Sampling design for contaminant distribution in lake sediments. *Journal of Water Resources Planning and Management*, **121**, 71–79.

Chiles, J-P and Delfiner, P. (1999). *Geostatistics.* New York : Wiley.

Diggle, P.J., Moyeed, R.A. and Tawn, J.A. (1998). Model-based Geostatistics (with Discussion). *Applied Statistics*, **47** 299–350.

Diggle, P.J., Ribeiro, P.J. and Christensen, O. (2003). An introduction to model-based geostatistics. In *Spatial Statistics and Computational Methods. Lecture Notes in Statistics 173*, ed J. Møller, 43–86. New York: Springer.

Handcock, M.S. and Stein, M.L. (1993). A Bayesian analysis of kriging. *Technometrics*, **35**, 403–10.

Johnson, M.E., Moore L.M. and Ylvisaker, D. (1990). Minimax and maximin distance designs. *Journal of Statistical Planning and Inference*, **26**, 131–148.

Kitanidis, P.K. (1986). Parameter uncertainty in estimation of spatial functions: Bayesian analysis. *Water Resources Research*, **22**, 499–507.

Lark, R.M. (2002). Optimized spatial sampling of soil for estimation of the variogram by maximum likelihood. *Geoderma*,**105**, 49–80.

Le, N.D. and Zidek, J.V. (1992). Interpolation with uncertain covariances: a Bayesian alternative to Kriging. *Journal of Multivariate Analysis*, **43**, 351–74.

Martin, R.J. (2001). Comparing and contrasting some environmental and experimental design problems. *Environmetrics*, **12**, 303–317.

McBratney, A., Webster, R. and Burgess, T. (1981). The design of optimal sampling schemes for local estimation and mapping of regionalised variables-I. Theory and methods. *Computers and Geosciences*, **7**, 331-334.

McBratney, A. and Webster, R. (1981). The design of optimal sampling schemes for local estimation and mapping of regionalised variables-II. Program and examples. *Computers and Geosciences*, **7**, 335-365.

McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models* (second edition). London: Chapman and Hall.

Müller, W.G. and Zimmerman, D.L. (1999). Optimal designs for variogram estimation. *Environmetrics*, **10**, 23–37.

Müller, W.G. (2001). *Collecting spatial data.* Heidelberg: Springer-Verlag.

http://www.r-project.org/. *The R Project for Statistical Computing.*

Ribeiro, P.J., Christensen, O. and Diggle, P.J. (2003). Geostatistical Software - geoR and geoRglm. *Proceedings of the Distributed Statistical Computing (DSC-2003) Conference.*

Royle, J.A. and Nychka, D. (1998). An algorithm for the construction of spatial coverage

designs with implementation in S-PLUS. *Computers and Geosciences*, **24**, 479–488.

Rue, H. (2001). Fast sampling of Gaussian Markov random fields. *Journal of the Royal Statistical Society*, B **63**, 325–338.

Rue, H. and Tjelmeland, H. (2002). Fitting Gaussian Markov random fields to Gaussian fields. *Scandinavian Journal of Statistics*, **29**, 31–49.

Russo, D. (1984). Design of an optimal sampling network for estimating the variogram. *Soil Science Society of American Journal*, **52**, 708–716.

Spruill, T.B. and Candela, L. (1990). Two approaches to design of monitoring networks. *Ground Water*, **28**, 430–442.

Warrick, A. and Myers, D. (1987). Optimization of sampling locations for variogram calculations. *Water Resources Research*, **23**, 496–500.

Winkels, H. and Stein, A. (1997). Optimal cost-effective sampling for monitoring and dredging of contaminated sediments. *Journal of Environmental Quality*, **26**, 933–946.

Zimmerman, D.L. and Homer, K.E. (1991). A network design criterion for estimating selected attributes of the semivariogram. *Environmetrics*, **4**, 425–441.
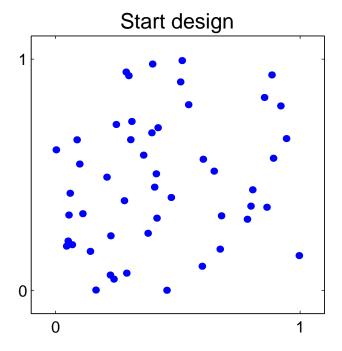
Figure 1: *Simulation example of retrospective design: the initial design consisting of 50 locations in the unit square.*
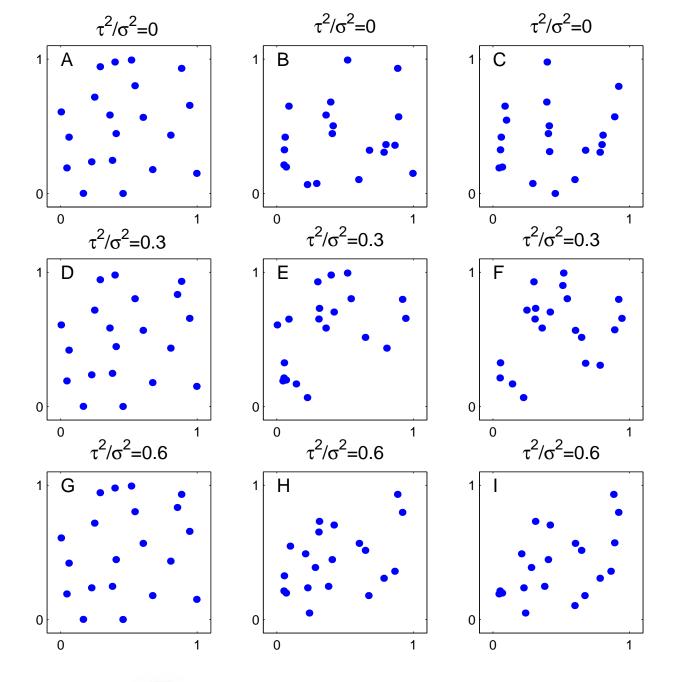
Figure 2: *Simulation example of retrospective design: reducing the initial design in Figure 1 to 20 locations. A, D and G) Classical designs for models with $\tau^2/\sigma^2 = 0$, 0.3, 0.6 respectively. B, E and H) Bayesian designs for models with $\tau^2/\sigma^2 = 0$, 0.3, 0.6 respectively with the ratio τ^2/σ^2 considered known. C, F and I) Bayesian designs for models with $\tau^2/\sigma^2 = 0$, 0.3, 0.6 respectively with the ratio τ^2/σ^2 is considered unknown with a uniform prior.*
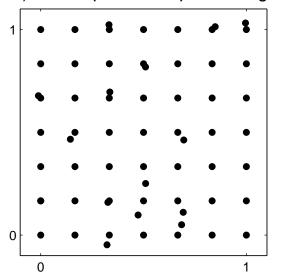
Figure 3: *Locations of the 70 monitoring stations (● and ✕) measuring salinity in the Kattegat basin, and of the 20 stations (✕) which constitute the final design.*
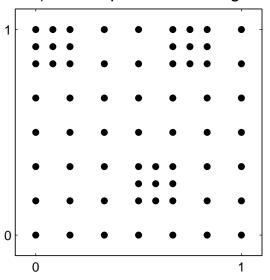
Figure 4: *Examples of a* $(7 \times 7, 15, 0.5)$ *lattice plus close pairs design, and a* $(7 \times 7, 3, 3 \times 3)$ *lattice plus in-fill design.*
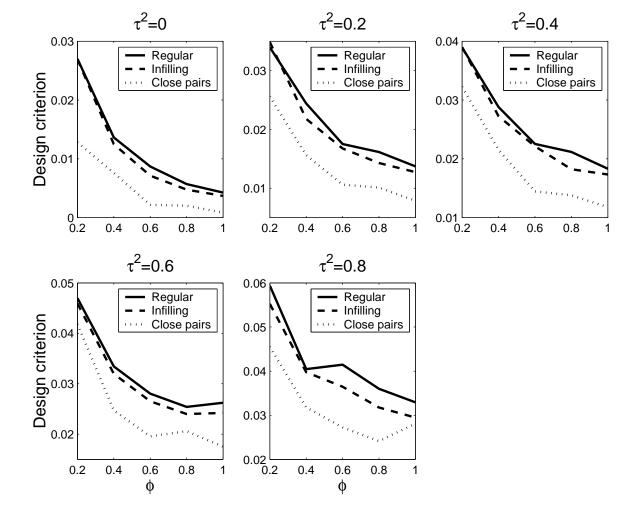
14

Figure 5: *Prospective design results comparing the efficiencies of the regular 8×8 lattice, the $(7 \times 7, 15, 0.5)$ lattice plus close pairs design, and the $(7 \times 7, 3, 3 \times 3)$ lattice plus in-fill design.*

τ^2/σ^2-ratio	Sites in Bayesian design τ^2/σ^2 random	Sites in Bayesian design τ^2/σ^2 fixed	Sites in classical design
0	0.022	0.025	0.22
0.3	0.094	0.21	0.59
0.6	0.074	0.091	0.93

Table 1: *Bayesian design criteria for the 9 designs in Figure 2.*

Addresses:

Peter Diggle
Department of Mathematics and Statistics
Lancaster University
Lancaster LA1 4YF
UK
email: p.diggle@lancaster.ac.uk

Søren Lophaven
Informatics and Mathematical Modelling
Richard Petersens Plads, Building 321, Room 011
Technical University of Denmark
DK-2800 Kgs. Lyngby
Denmark
email: snl@imm.dtu.dk

17