Check for updates

Received: 7 August 2018 | Revised: 14 December 2018 | Accepted: 27 January 2019

DOI: 10.1111/ejss.12797

European Journal of Soil Science WILEY

ORIGINAL ARTICLE

Efficient sampling for geostatistical surveys

Alexandre M.J.-C. Wadoux[1] | Benjamin P. Marchant[2] | Richard M. Lark[3]

[1]Soil Geography and Landscape group, Wageningen University & Research, Wageningen, The Netherlands

[2]Environmental Science Centre, British Geological Survey, Keyworth, UK

[3]School of Biosciences, University of Nottingham, Sutton Bonington, UK

Correspondence
Alexandre M.J-C. Wadoux, Soil Geography and Landscape group, Wageningen University & Research, Droevendaalsesteeg 4, 6708 PB Wageningen, Netherlands.
Email: alexandre.wadoux@wur.nl

Funding information
European Union's Seventh Framework Programme for research, technological development and demonstration, Grant/Award Number: 607000

A geostatistical survey for soil requires rational choices regarding the sampling strategy. If the variogram of the property of interest is known then it is possible to optimize the sampling scheme such that an objective function related to the survey error is minimized. However, the variogram is rarely known prior to sampling. Instead it must be approximated by using either a variogram estimated from a reconnaissance survey or a variogram estimated for the same soil property in similar conditions. For this reason, spatial coverage schemes are often preferred, because they rely on the simple dispersion of sampling units as uniformly as possible, and are similar to those produced by minimizing the kriging variance. If extra sampling locations are added close to those in a spatial coverage scheme then the scheme might be broadly similar to one produced by minimizing the total error (i.e. kriging variance plus the prediction error due to uncertainty in the covariance parameters). We consider the relative merits of these different sampling approaches by comparing their mean total error for different specified random functions. Our results showed the considerable benefit of adding close-pairs to a spatial coverage scheme, and that optimizing with respect to the total error generally gave a small further advantage. When we consider the example of sampling for geostatistical survey of clay content of the soil, an optimized scheme based on the average of previously reported clay variograms was fairly robust compared to the spatial coverage plus close-pairs scheme. We conclude that the direct optimization of spatial surveys was only rarely worthwhile. For most cases, it is best to apply a spatial coverage scheme with a proportion of additional sampling locations to provide some closely spaced pairs. Furthermore, our results indicated that the number of observations required for an effective geostatistical survey depend on the variogram parameters.

Highlights

- We compared spatial coverage and spatial coverage plus a subset of 10% from the total sample as close-pairs.
- The objective function encompasses variogram uncertainty and prediction error variance.
- Spatial coverage schemes always performed poorly because of the lack of information at short distances.
- Using a scheme in which 10% of the sampling units are taken at short distances is a robust strategy.

KEYWORDS

geostatistics, ordinary kriging, pedometrics, sample size, spatial coverage scheme, uncertainty assessment

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2019 The Authors. European Journal of Soil Science published by John Wiley & Sons Ltd on behalf of British Society of Soil Science.

1 | INTRODUCTION

When mapping a continuous soil variable, geostatistical predictions at unobserved locations are made from a limited set of sampling units, called a sample. The spatial locations of those units, i.e. the sampling scheme, has a key role in determining the cost of the survey and the quality of the predictions. Often, limited resources are available and one must adopt efficient strategies for the soil sample collection.

Several solutions have been proposed to select additional sampling sites optimally using ordinary kriging, a basic technique in geostatistics. These often require prior knowledge about the correlation function (i.e. variogram) of the target property. For example, Van Groenigen, Siderius, and Stein (1999) proposed spatial simulated annealing (SSA) to optimize the sampling scheme so as to minimize the spatially averaged kriging variance as the objective function. This method leads to a space-filling distribution of observations, which are placed more or less evenly over the area of interest. A similar scheme can be obtained by the spatial coverage method described in Royle and Nychka (1998). They proposed a general geometric, space-filling criterion and published a point-swapping algorithm in S-plus to minimize this criterion. Brus, Spätjens, and De Gruijter (1999) proposed the mean of the squared shortest distance (MSSD) as a geometric minimization criterion, so that it can be minimized by the fast *k*-means algorithm. Later this was implemented in the R language by Walvoort, Brus, and De Gruijter (2010).

One advantage of coverage schemes is that they do not depend on the variogram of the soil property to be sampled. Coverage schemes are created by minimizing a criterion that is simply a function of the distance between sampling locations. Brus, De Gruijter, and Van Groenigen (2007) showed that using a spatial coverage scheme led to only marginally larger mean ordinary kriging variances (MKV) than schemes where this quantity was minimized directly. The authors endorsed early geostatistical practice in soil science where sampling units were located on a regular grid (Yfantis, Flatman, & Behar, 1987).

However, regularly-spaced sampling schemes are inadequate to model the short-range variation of the soil property, which is critical for geostatistical analyses (Starks, 1986). A practical solution, as suggested for instance by De Gruijter, Brus, Bierkens, & Knotters (2006, pp. 166-168), is to supplement the spatial coverage sample by a few additional units, located at short distances from the existing units. Recently, Lark and Marchant (2018) demonstrated that including such a short-distance subset markedly decreased the uncertainty of the kriging prediction for little additional effort in field data collection. Over a contrasting set of random variables, the authors proposed a simple rule that about 10% of the total sample size should be devoted to short-distance units.

Using a more formal expression of the total error in a geostatistical survey, Marchant and Lark (2007) optimized a sampling scheme by minimization of the sum of error contributions from the kriging variance and the effects of uncertainty in the variogram estimate. We refer to this objective function as the total error. The authors showed that the configuration of the optimized scheme varied according to the variogram, which was unknown prior to sampling, and used a Bayesian framework to account for a set of plausible values of variogram parameters. A similar approach was applied by Zhu and Stein (2006) for redesigning an air monitoring network. The authors noted that estimates of the variogram parameters were uncertain. They approximated the error covariance matrix of the parameters by the inverse of the Fisher information matrix, and used a Taylor series approximation of its effect on the prediction variance to account for it in their sampling objective function. For both studies, the resulting optimized schemes closely resembled the spatial coverage scheme with a small number of close-pairs of locations included, which are useful for estimating the spatial correlation over short distances. They showed that the number of close-pair locations depended largely on the variogram parameter values, and especially the variogram distance parameter.

However, the optimization procedure using a formal criterion for minimization of the total error is complex and time consuming. The formula for the total prediction error depends on the variogram and therefore it cannot be calculated exactly prior to sampling. Instead it must be approximated by using either a variogram estimated from a reconnaissance survey or a variogram estimated for the same soil property in similar conditions. Schemes based on approximate variograms are likely to be suboptimal. In such cases, spatial coverage sampling schemes (possibly with additional close-pairs) offer a viable and relatively simple alternative to plan a soil survey with little or no prior information.

Surveyors must also consider the number of sampling units that are required to produce effective geostatistical predictions. The sample must be sufficient to estimate an accurate variogram function. Kerry and Oliver (2007) noted that it is generally accepted that 100 units are required to produce a reliable method of moments estimate of the variogram. This advice stems from a study of simulated random functions conducted by Webster and Oliver (1992). Kerry and Oliver (2007) subsampled four field-scale surveys of clay content and determined that a reliable residual maximum likelihood (REML) estimate of the variogram could be attained with fewer than 50 sampling units.

In summary, the sampling scheme affects the uncertainty in the variogram parameters, which can have an impact on the prediction error variance. Supplementing a spatial coverage sample by a simple rule of thumb reduces the prediction error variance, but the overall distribution of sample points

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

in a scheme can be optimized, although this is laborious and requires some prior information. Whether a practical sampling strategy is markedly better when based on optimization rather than the simple rule remains an open question, and past work has not compared the approaches directly. That is what we address here.

In this research we examined empirically the difference between spatial coverage sampling schemes (sc), spatial coverage schemes supplemented with close-pairs of points (sc_+) and schemes optimized to reduce the total error. We compared these schemes with respect to the sample size required to obtain comparable results. Our objective was to show whether formal optimization is generally worthwhile, given the computational demands and the challenges of specifying prior values of variance parameters, and whether spatial coverage sampling with supplementary points is a robust practical strategy.

In our first scenario we minimized this error for a known hypothetical variogram and a given sample size. Then we determined the size of a spatial coverage scheme that would be required to achieve the same total prediction error. Similarly, we considered the size of a spatial coverage scheme plus 10% close-pairs that would also achieve the same total prediction error.

In addition to the spatial arrangement of sampling units we also considered the minimum number of units that were required to produce useful geostatistical predictions. For sample sizes larger than this minimum sample size the ordinary kriging predictor outperformed the simple random sample mean as a predictor of the values at points. For sample sizes smaller than this minimum there was no benefit from a geostatistical approach for mapping. We assumed that a geostatistical survey should, as a basic minimum requirement, ensure that local spatial predictions have an average prediction error variance that is smaller than the prediction error variance of the regional mean, estimated by design-based sampling. Webster and Lark (2013) discussed how the design-based mean can be treated statistically as a point prediction. We assumed that this design-based survey was the same size as the geostatistical survey, that the sampling units were selected according to a simple random scheme and that the corresponding design-based estimate of the mean was used as the prediction at each location.

In our second scenario we considered a geostatistical survey of soil clay content and the effect of using the average variogram of a set presented by Paterson, McBratney, Minasny, and Pringle (2018) as a basis for a sampling scheme. We minimized the total prediction error variance given a sample size based on the average variogram and then repeated the tests conducted in the first scenario to find the size of the sc and sc_+ schemes that would be required to achieve the same total error as the optimized scheme for each of the clay variograms.

2 | MATERIALS AND METHODS

2.1 | Formulation of the objective function

Using the ordinary kriging formulation, we consider the situation in which the soil property (which is assumed to be a realization of a random function Z) has been measured at n locations $\mathbf{s}_i (i = 1,\ldots,n; \mathbf{s}_i \in \mathcal{A})$. The measurements $z(\mathbf{s}_i)$ are treated as realizations of $Z(\mathbf{s}_i)$ and prediction is done for Z at unobserved locations $\mathbf{s}_0$, with a known covariance parameter vector $\boldsymbol{\theta}$. Stacking the $z(\mathbf{s}_i)$ in a vector $\mathbf{z}$ and changing to matrix notation yields the ordinary kriging prediction equation (Webster & Oliver, 2007):

$$\tilde{Z}(\mathbf{s}_0|\boldsymbol{\theta}) = \boldsymbol{\lambda}^\top \mathbf{z}, \tag{1}$$

where $\boldsymbol{\lambda}^\top$ is the vector of kriging weights, obtained from the kriging equation:

$$\boldsymbol{\lambda} = \mathbf{A}^{-1}\mathbf{d}, \tag{2}$$

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \psi \end{pmatrix} = \begin{bmatrix} C(\mathbf{s}_1-\mathbf{s}_1|\boldsymbol{\theta}) & C(\mathbf{s}_1-\mathbf{s}_2|\boldsymbol{\theta}) & \ldots & C(\mathbf{s}_1-\mathbf{s}_n|\boldsymbol{\theta}) & 1 \\ C(\mathbf{s}_2-\mathbf{s}_1|\boldsymbol{\theta}) & C(\mathbf{s}_2-\mathbf{s}_2|\boldsymbol{\theta}) & \ldots & C(\mathbf{s}_2-\mathbf{s}_n|\boldsymbol{\theta}) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C(\mathbf{s}_n-\mathbf{s}_1|\boldsymbol{\theta}) & C(\mathbf{s}_n-\mathbf{s}_2|\boldsymbol{\theta}) & \ldots & C(\mathbf{s}_n-\mathbf{s}_n|\boldsymbol{\theta}) & 1 \\ 1 & 1 & \ldots & 1 & 0 \end{bmatrix}^{-1} \times \begin{bmatrix} C(\mathbf{s}_0-\mathbf{s}_1|\boldsymbol{\theta}) \\ C(\mathbf{s}_0-\mathbf{s}_2|\boldsymbol{\theta}) \\ \vdots \\ C(\mathbf{s}_0-\mathbf{s}_n|\boldsymbol{\theta}) \\ 1 \end{bmatrix}, \tag{3}$$

where ψ is the Lagrange multiplier introduced to allow minimization of the kriging variance subject to the constraint that the n weights $\lambda_1, \lambda_2, \cdots, \lambda_n$ sum to one. The covariance between the ith and jth locations is denoted by $C(\mathbf{s}_i - \mathbf{s}_j|\boldsymbol{\theta})$. The term $C(s_i - s_i)$ is the sill variance (*a priori* variance). Note that while $\mathbf{A}$ needs to be derived (and inverted) once if all observations are used for prediction at every target site, $\mathbf{d}$ must be computed for every prediction location $\mathbf{s}_0$.

From Equations (1) and (3), the expected squared error of the prediction is given by:

$$\begin{aligned} \sigma^2_{\mathrm{OK}}(\mathbf{s}_0) &= \mathrm{var}\left(Z(\mathbf{s}_0) - \tilde{Z}(\mathbf{s}_0|\boldsymbol{\theta})\right) \\ &= C(\mathbf{s}_0-\mathbf{s}_0|\boldsymbol{\theta}) - \boldsymbol{\lambda}^\top \mathbf{d}. \end{aligned} \tag{4}$$

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In addition to the squared error of the prediction, Marchant and Lark (2007) and Zhu and Stein (2006) considered the effect of uncertainty in the estimated spatial model (variogram) parameters by a Taylor series approximation:

$$\mathrm{E}\left[\tau^2(\mathbf{s}_0)\right] = \sum_{i=1}^{q}\sum_{j=1}^{q}\mathrm{cov}\left(\theta_i,\theta_j\right)\frac{\partial\boldsymbol{\lambda}^\top}{\partial\theta_i}\mathbf{C}\frac{\partial\boldsymbol{\lambda}}{\partial\theta_j}, \quad (5)$$

where $\mathrm{cov}(\theta_i, \theta_j)$ is the covariance between the ith and jth parameters. This requires the variogram parameters $\theta_i(i, j = 1,\ldots, q)$ to be known so that Equation (5) can be approximated prior to sampling. The n-vector of partial derivatives of the kriging weights with respect to the ith variance parameter is denoted by $\frac{\partial\boldsymbol{\lambda}^\top}{\partial\theta_i}$ and can be obtained by (Marchant & Lark, 2007):

$$\frac{\partial\boldsymbol{\lambda}}{\partial\theta_i} = \mathbf{A}^{-1}\left(\frac{\partial\mathbf{d}}{\partial\theta_i} - \frac{\partial\mathbf{A}}{\partial\theta_i}\mathbf{A}^{-1}\mathbf{d}\right). \quad (6)$$

The covariance between the variogram parameters can be approximated using the inverse of the Fisher information matrix $\mathbf{F}$ (Kitanidis, 1987):

$$\mathrm{cov}\left(\theta_i,\theta_j\right) \approx \mathbf{F}^{-1}\left(\theta_i,\theta_j\right) = \left(\frac{1}{2}\mathrm{Tr}\left[\mathrm{C}^{-1}\frac{\partial\mathrm{C}}{\partial\theta_i}\mathrm{C}^{-1}\frac{\partial\mathrm{C}}{\partial\theta_j}\right]\right)^{-1}, \quad (7)$$

where Tr[·] denotes the trace of the matrix. The total error at locations $\mathbf{s}_0$, $\sigma_{\mathrm{P}}^2(\mathbf{s}_0)$ is given by the sum of the squared prediction error $\sigma_{\mathrm{OK}}^2(\mathbf{s}_0)$ and the spatial model parameter uncertainty $\mathrm{E}[\tau^2(\mathbf{s}_0)]$:

$$\sigma_{\mathrm{P}}^2(\mathbf{s}_0) = \sigma_{\mathrm{OK}}^2(\mathbf{s}_0) + \mathrm{E}\left[\tau^2(\mathbf{s}_0)\right], \quad (8)$$

where subscript P stands for parameter. This can be aggregated to obtain a spatial average:

$$\bar{\sigma}_{\mathrm{P}}^2 = \frac{1}{\mathcal{A}}\int_{s\in\mathcal{A}}\left(\sigma_{\mathrm{OK}}^2(\mathbf{s}) + \mathrm{E}\left[\tau^2(\mathbf{s})\right]\right)\mathrm{d}s. \quad (9)$$

In practice, the integral $\bar{\sigma}_{\mathrm{P}}^2$ is numerically approximated by a discrete summation over a spatial grid.

2.2 | Optimization of the sampling schemes

We start with an initial random set of sampling locations of size N, lying within the boundaries of study area $\mathcal{A}$. We assume that $Z(\mathbf{s}_i)$ is a stationary isotropic normally distributed random field, characterized by a constant mean and fitted correlation function $\rho(h)$ (h is the spatial lag or separation distance). The aim is to find the optimal sampling scheme, which minimizes the objective function (Equation (9)), given the parameters of $\rho(h)$. Many algorithms have been developed for solving optimization problems. We use simulated annealing (Kirkpatrick, Gelatt, and Vecchi (1983)), extended for spatial optimization by Van Groenigen et al. (1999) for generating sequences of new possible schemes. A new sampling scheme is created by randomly shifting a randomly selected unit within the study area. This generates a new candidate scheme for which the objective function can be evaluated with Equation (9), and compared with that of the previous scheme. The new candidate scheme is accepted if it has a smaller value of the objective function than the previous one. If the new scheme has a larger value of the objective function then it is accepted or rejected at random; the probability of acceptance is given by (Wadoux, Brus, Rico-Ramirez, & Heuvelink, 2017):

$$P(\mathrm{accept}) = \exp\left(\frac{\bar{\sigma}_{\mathrm{P}}^2(\mathrm{old}) - \bar{\sigma}_{\mathrm{P}}^2(\mathrm{new})}{\alpha}\right), \quad (10)$$

where the control parameter α is a temperature parameter. The temperature is kept constant during a set of perturbations, called a chain, after which it is decreased to a value of $\beta \times \alpha$ for $\beta < 1$. In this way, the risk of the optimizer becoming trapped in a local but not a global minimum is reduced. We used the implementation provided by the R package spsann (Samuel-Rosa, 2017) through the optimUSER function. The initial temperature α was set to 3 with a cooling parameter β of 0.9. These were chosen so that P(accept) is close to 1 in the first chain and generally zero at the final chain. The maximum number of chains is set to 200, so that the total number of iterations is $N \times 200$. The process stops if the determined number of iterations ($N \times 200$) is reached or if the criterion remains constant for ten chains. The candidate locations are the centre of cells of a square grid.

2.3 | Scenario 1

The first scenario considers the case where the variogram is known. We characterize the spatial correlation ρ by the second parametrization of the isotropic Matérn model (Matérn, 1986) given by Stein (2006, p. 31):

$$\rho(h) = \frac{1}{2^{\nu-1}\Gamma(\nu)}\left(\frac{2\nu^{\frac{1}{2}}h}{a}\right)^{\nu}\mathcal{K}_\nu\left(\frac{2\nu^{\frac{1}{2}}h}{a}\right), \quad (11)$$

where h is the separation distance, $\mathcal{K}_\nu$ is the modified Bessel function of the second kind of order ν (see Abramowitz & Stegun, 1972, pp. 374-379) and Γ is the gamma function. The correlation function $\rho(h)$ has parameters a and ν. Parameter a is the distance parameter, which indicates how fast the correlation decays with increasing h and ν is the smoothness parameter. Stein (2006) noted that ν is the critical parameter in the Matérn correlation model. The larger is ν, the smoother is Z. We chose a Matérn model for its flexibility in modelling the spatial covariance with a small number of parameters (Minasny & McBratney, 2005).

For the first scenario, we generated a square area of 100 m × 100 m. Spatial coverage schemes of size $N = 60$, 61, … , 200 are derived by discretization of the area into N geographical strata using the stratify method from the R package spcosa (Walvoort et al., 2010). The spatial

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

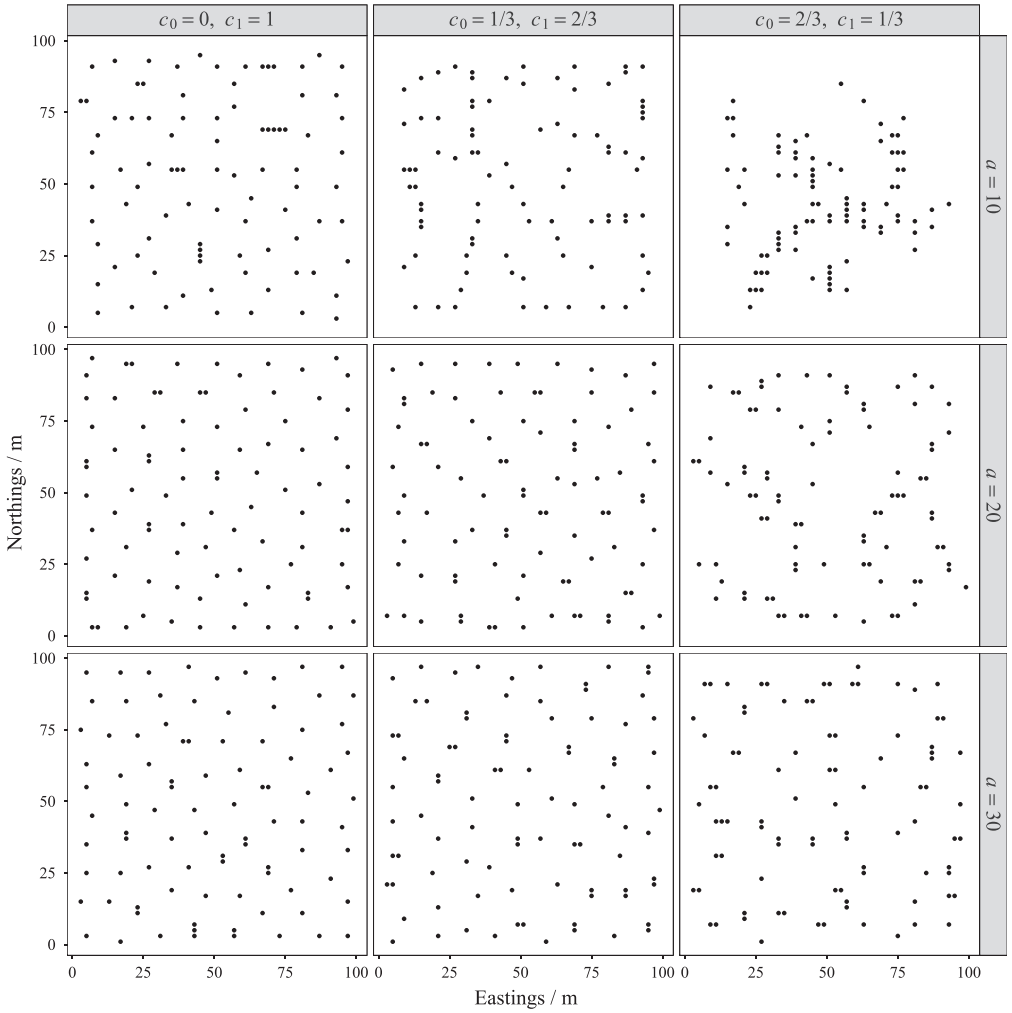


FIGURE 1 Optimized 90-unit schemes for different variogram parameters and $\nu = 0.2$

coverage units are taken in the centroid of the strata, which is equivalent to minimizing the mean squared shortest distance between a location in the region and the nearest sampling location. In addition, we also generated samples of size $N = 60, 61, \ldots, 200$ in which the sampling locations were distributed according to a spatial coverage scheme, with a subset of 10% of units positioned at an arbitrary distance that was short relative to the spacing between neighbouring points in the basic spatial coverage survey. This arbitrary short distance was set to 2 m because of a mean spacing between neighbouring locations in the sc scheme of 6.6 m for $N = 60$ and 12.5 m for $N = 200$. These close-pair units were selected by simple random sampling without replacement in a randomly chosen direction from 0–360 degrees. We repeated the selection of close-pairs several times to determine the sampling variation in total variance. Since the latter was small, we did not pursue this any further because this confirmed the very tight confidence intervals in the Lark and Marchant (2018) study. We considered different sets of variogram parameter values, all of which had a total sill variance of one. Four values of ν were tested: $\nu = 0.5$ (equivalent to the exponential variogram), $\nu = 0.2$ (rougher than the exponential), $\nu = 1.1$ and $\nu = 2$ (smoother than the exponential). These four ν values were combined with each of three distance parameter: $a = 10$, 20 and 30, and three ratios of the nugget (c_0) to total sill variance ($c_0 + c_1 = 1$) for strong ($c_0 = 0$), moderate ($c_0 = 1/3$) and weak ($c_0 = 2/3$) spatial dependence. Note that we use the nugget to sill ratio to characterize the spatial dependence of a model with known parameters, but this should not be done when comparing empirical variograms because the magnitude of the nugget variance is likely to depend in part on the sampling scheme. Each of the $4 \times 3 \times 3 = 36$ scenarios were optimized for a fixed sample size $N = 90$ in the way described in the previous section. To speed up computations the criterion was evaluated at 34×34 locations on a regular square grid of spacing 3 m.

In this scenario we compared for each variogram the size of the sc and sc_+ samples required to attain the same value of the objective function as the optimized scheme of 90 units.

We also compared the average total prediction variance that resulted from the geostatistical survey of each random function with the prediction variance that would result from using an estimate of the simple random sample of the field mean as a predictor of the value at points. If the design-

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

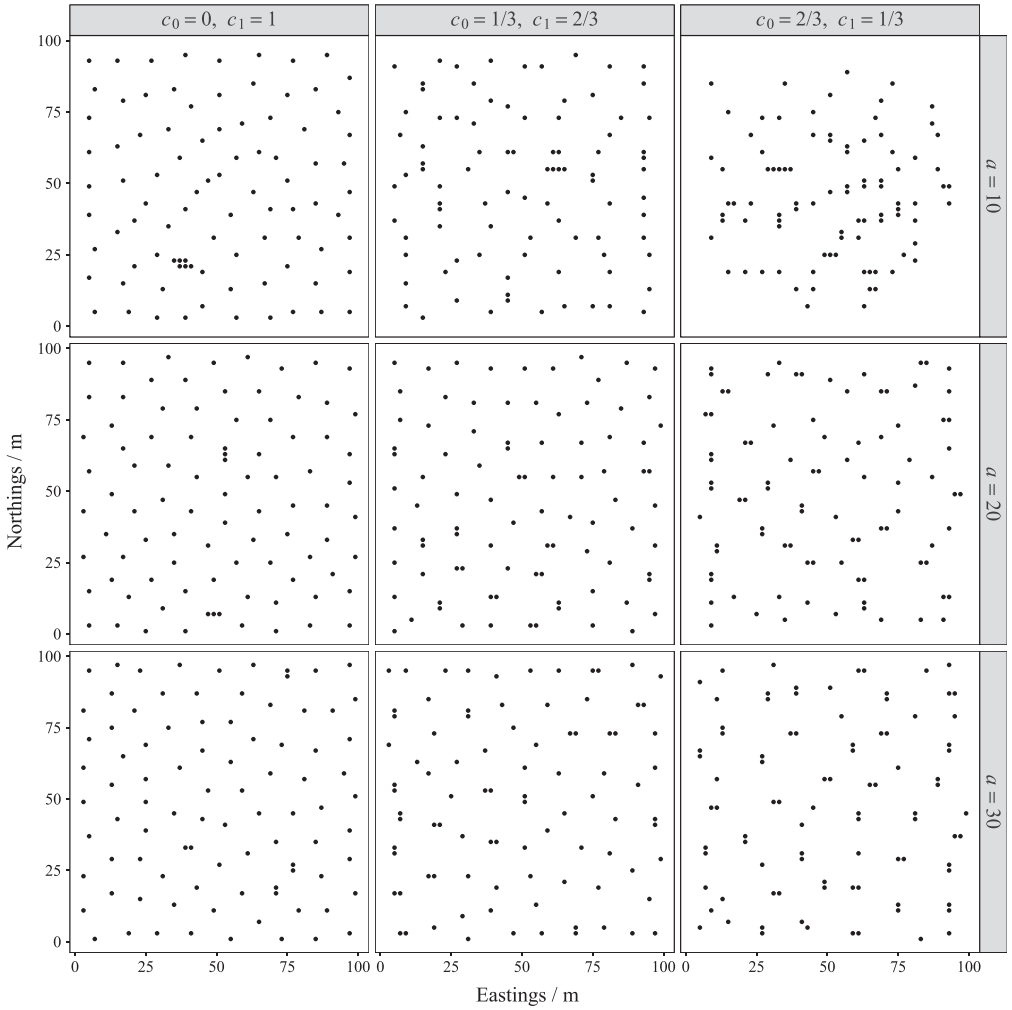


FIGURE 2 Optimized 90-unit schemes for different variogram parameters and $\nu = 0.5$

based survey consists of N locations selected by simple random sampling this prediction variance is equal to (Brus, De Gruijter, & Breeuwsma (1992, Eq.7)):

$$\bar{\sigma}^2_{\mathrm{DB}} = \sigma^2\left(1 + \frac{1}{N}\right), \tag{12}$$

where σ^2 is the dispersion variance (the variance of the variable within the study area) and the σ^2/N term reflects the uncertainty in estimating the field-scale mean of the property of interest with simple random sampling (Brus & De Gruijter, 1993). Instead of the spatial variance (dispersion variance) for a single realization, we used the model expectation of the dispersion variance in Equation (12), so that the model expectation of the spatial mean of the design-based estimation error variance at points was also obtained. For each set of variogram parameters, we determined the smallest sample size of a geostatistical survey which led to the average total prediction variance being less than this design-based prediction variance. We determined the dispersion variance for each random function from the average variance of 1000 lower-upper (LU)-simulations of the function at 2000 random locations across the study area.

2.4 | Scenario 2

The second scenario considered a survey of soil clay content where no field-specific information about the variogram was available. In such a circumstance, McBratney and Pringle (1999) suggested that the average of previously published soil clay variograms should provide useful information for assessing soil sampling schemes.

Here we used data from a published study on field-scale variability of soil variograms. We used a compilation of soil clay variograms, provided by Paterson et al. (2018). They were gathered from the existing literature, based on untransformed data and physical measurements. We converted the exponential, spherical and linear clay variograms to a Matérn model (Equation 11) by re-estimating their parameters using a least squares approach. In this way, we compared surveys using variograms with the same number of estimated parameters. From the set of Matérn clay variograms, we derived an average experimental variogram as in McBratney and Pringle (1999). Each variogram for soil clay was evaluated at a set of closely-spaced lag intervals. Each value of semivariance was transformed to its fourth root. The average value of the fourth root of the variogram

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

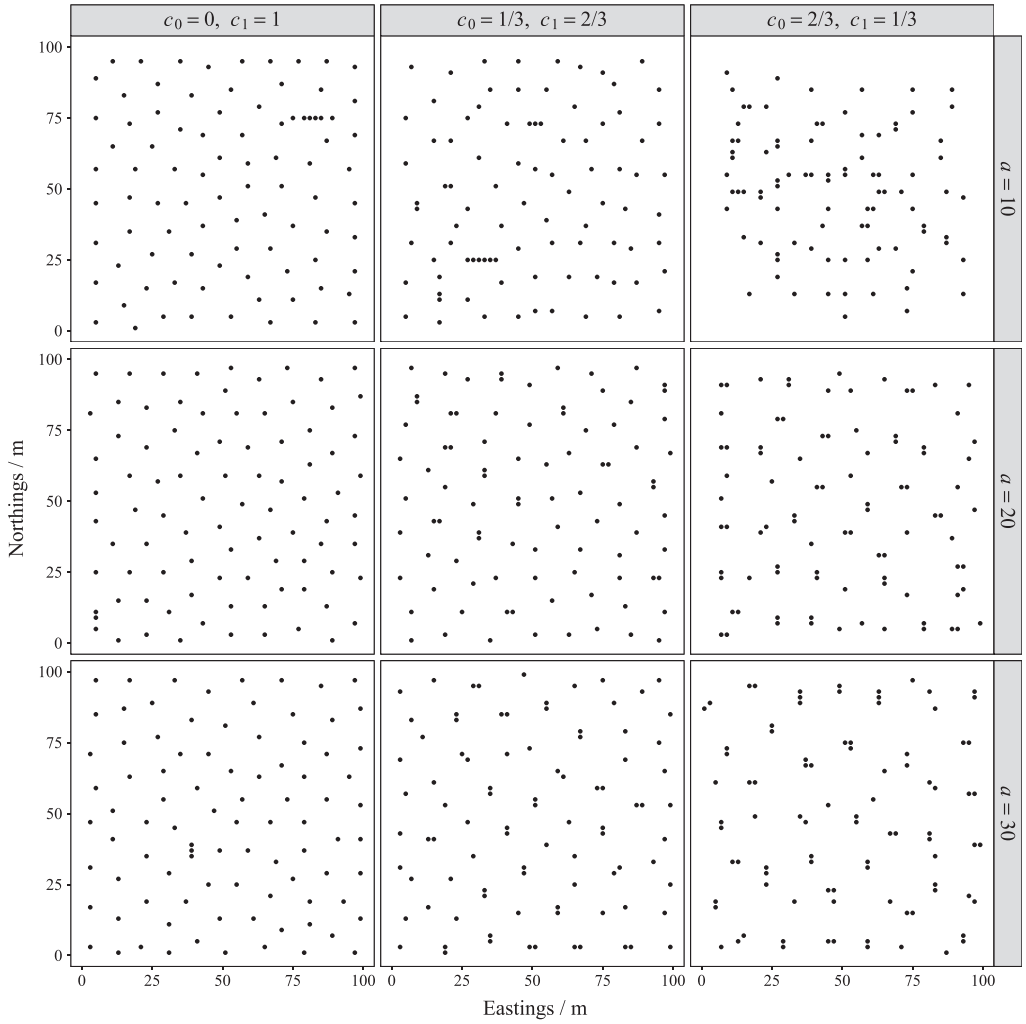


FIGURE 3 Optimized 90-unit schemes for different variogram parameters and $\nu = 1.1$

was computed at each lag interval over all the clay variograms and the resulting values were back-transformed to their fourth power. The fourth root is used to give a normally distributed variable even when the underlying variable includes extreme values (Cressie & Hawkins, 1980). Finally, a Matérn correlation function (Equation 11) was fitted by non-linear least squares to the average experimental variogram. The estimated Matérn correlation function was similar to an exponential variogram ($\nu = 0.5$) with a nugget variance $c_0 = 2.6$, a partial sill $c_1 = 8.0$ and a distance parameter $a = 44.1$ m (effective range is about 85 m). We then optimized the distribution of 90 sample units within a 500 m × 500 m region, using the mean total prediction error variance, Equation (9), as the objective function specifying the parameters of the average variogram. The objective function was evaluated at a centred square grid of 25 × 25 points with a spacing of 20 m. We then found, for the random function with parameters estimated for each clay variogram, the value of the objective function achieved by optimizing sample schemes of size $N = 60, 61, \ldots, 200$, and the corresponding number of observations in an sc and an sc_+ scheme required to match the value of the objective function achievable by optimization with the average clay variogram.

3 | RESULTS

3.1 | Scenario 1

Figures 1–4 show 90-unit sampling schemes optimized to minimize the expected total error with different values of the nugget to sill ratio, different distance parameters a and smoothness parameters of 0.2, 0.5, 1.1 and 2, respectively. In all schemes, the sampling locations are generally evenly dispersed over the area with some close-pair units. When the nugget to sill ratio increases (larger c_0), the number of close-pairs tends to increase substantially. The pattern for larger values of the distance parameter a is reversed. The larger is a, the smaller are the transects of close-pairs. When $c_0 = 2/3$, $c_1 = 1/3$ and $a = 10$ the sample size seems insufficient to cover the whole area. This might indicate that for this variogram and study area, 90 units were insufficient to both estimate the variogram and predict the soil property

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

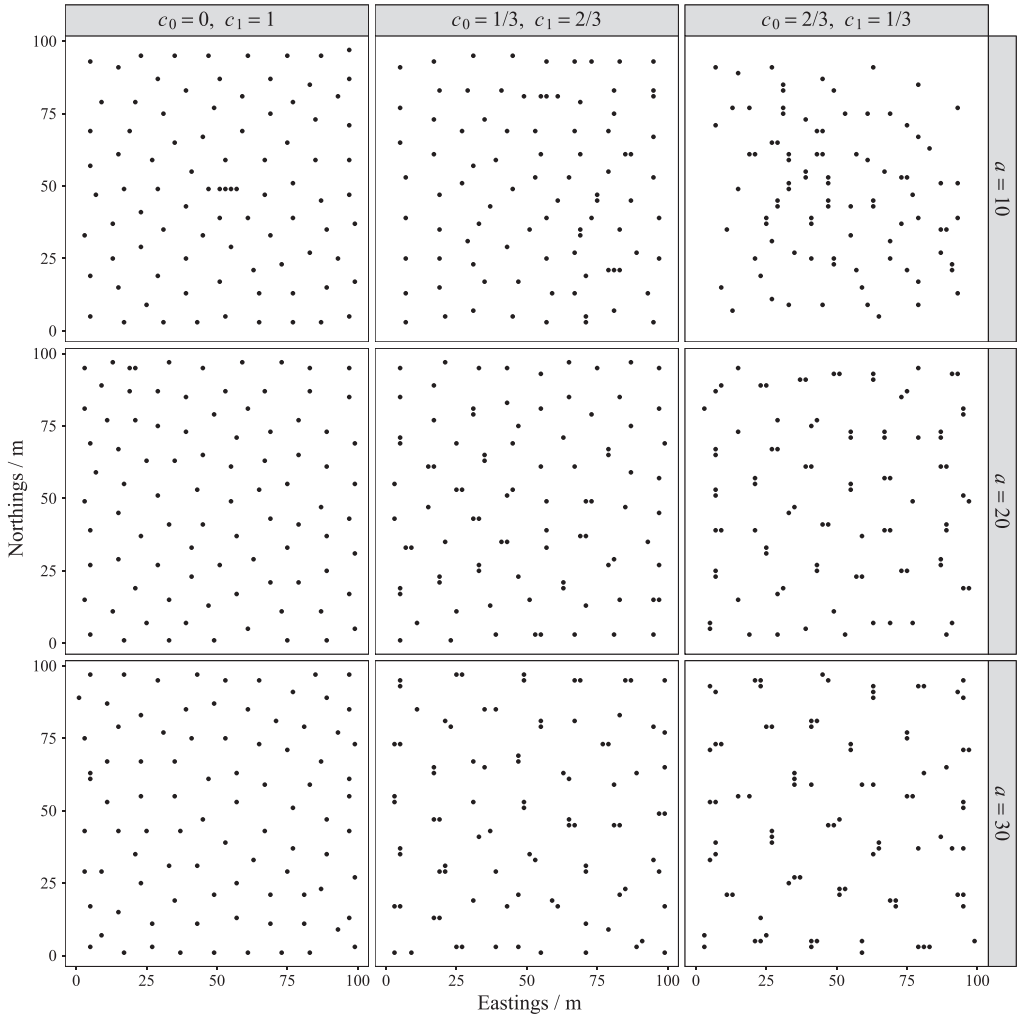


FIGURE 4 Optimized 90-unit schemes for different variogram parameters and $\nu = 2$

across the region. All values of ν tested had comparable patterns for the optimized schemes.

Figures 5–8 show the values of the objective function for each variogram type for the sc or sc_+ schemes compared to the values of the objective function from the optimized 90-unit sample scheme. The values of objective function for the sc schemes show a rougher pattern than those of the objective function for the sc_+ schemes. The sc schemes performed poorly in most cases. The poor performance was less pronounced for large values of a when ν was 1.1 or 2. In such cases, sc schemes were only slightly worse than the optimized schemes. The sc_+ schemes always performed slightly worse than the optimized schemes. With increasing nugget to sill ratio, the sc_+ schemes needed an increasing number of additional units to reach the same value for the objective function as the optimized sample scheme. This was valid for all values of ν tested.

For each set of variogram parameters, Table 1 reports the number of additional samples necessary when using the sc_+ scheme to reach the objective function of the optimized 90-unit scheme. Overall, the sc_+ scheme needs at least 8 and a maximum of 59 additional units to achieve the objective function of the optimized 90-unit scheme. As mentioned previously, there is a clear trend associated with the nugget to sill ratio. The larger is the ratio, the larger is the number of additional units in the sc_+ scheme. This effect was slightly diminished for increasing values of a.

Figure 9 shows the objective function for the sc and sc_+ schemes for the case where the smoothness parameter $\nu = 0.5$ was either fixed (known) or estimated (with uncertainty) with parameters $c_0 = 0$, $c_1 = 1$ and $a = 20$. When the smoothness was estimated there was a marked difference between the total error variance for the sc and sc_+ schemes when there were fewer than about 200 sample points in total. With larger sample sizes (above 220) the difference became negligible. When the smoothness is known (equivalent to assuming an exponential variogram), there were still minor differences between the sc and sc_+ scheme objective functions but they rapidly converged to the same values (from about 120 units).

Table 2 shows the minimum sample size required for the expected total variance to be smaller than the estimation variance of the target property that would result from a design-based survey of the same size. The sc_+ schemes needed on

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

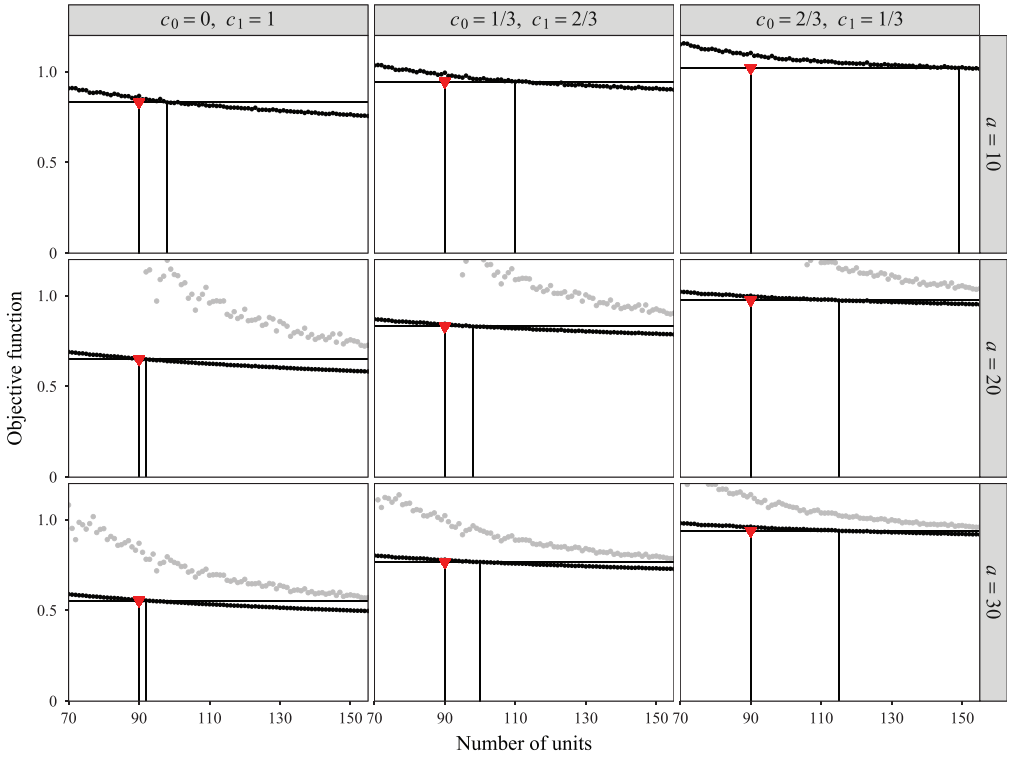


FIGURE 5 Value of objective function for sc_+ (black dots), sc (grey dots) and optimized (red triangle) schemes. The spacing between the two vertical lines indicates the number of extra units required for sc_+ to achieve an optimized objective function value for $\nu = 0.2$

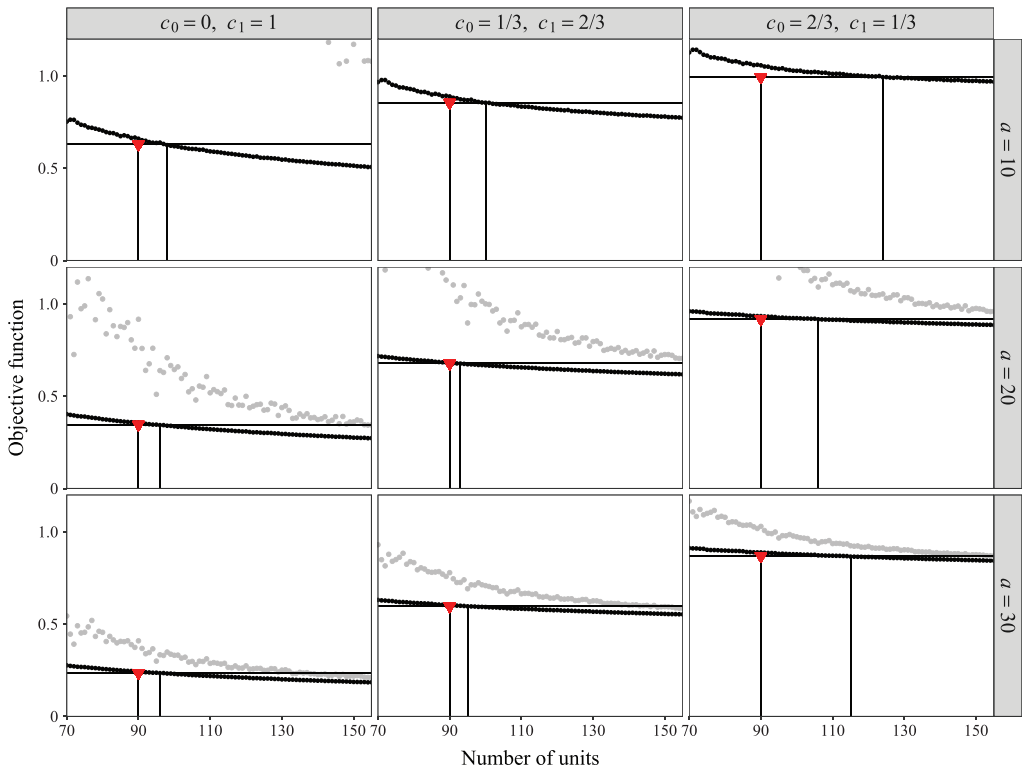


FIGURE 6 Value of objective function for sc_+ (black dots), sc (grey dots) and optimized (red triangle) schemes. The spacing between the two vertical lines indicates the number of extra units required for sc_+ to achieve an optimized objective function value for $\nu = 0.5$

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

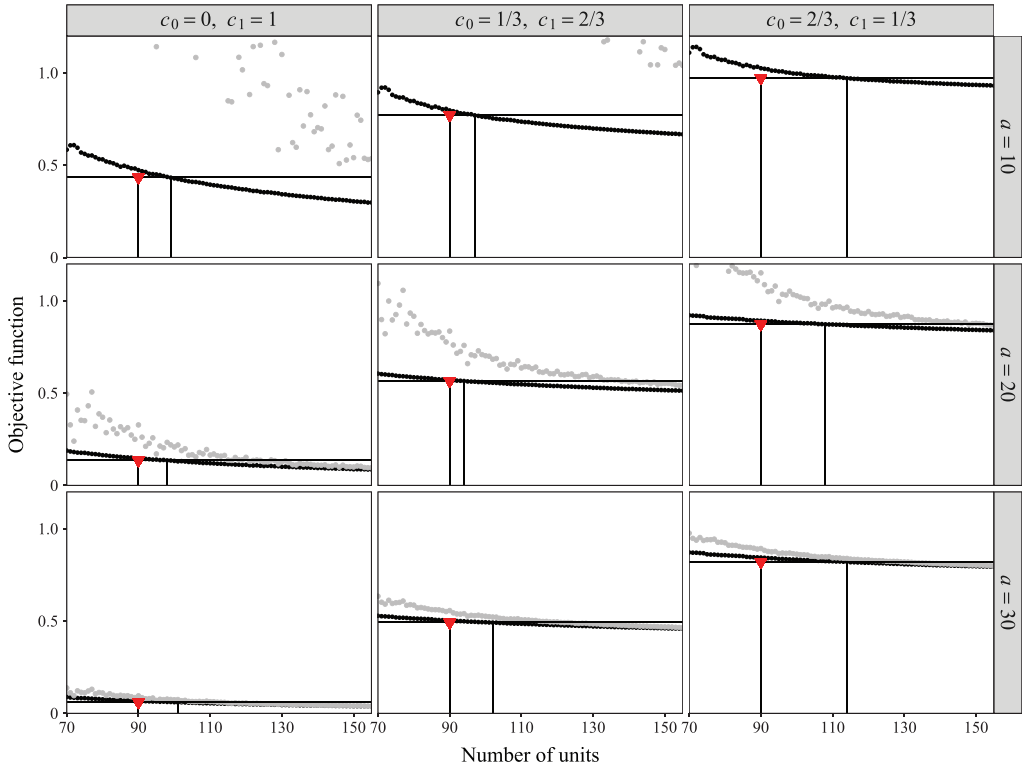


FIGURE 7 Value of objective function for sc_+ (black dots), sc (grey dots) and optimized (red triangle) schemes. The spacing between the two vertical lines indicates the number of extra units required for sc_+ to achieve an optimized objective function value for $\nu = 1.1$

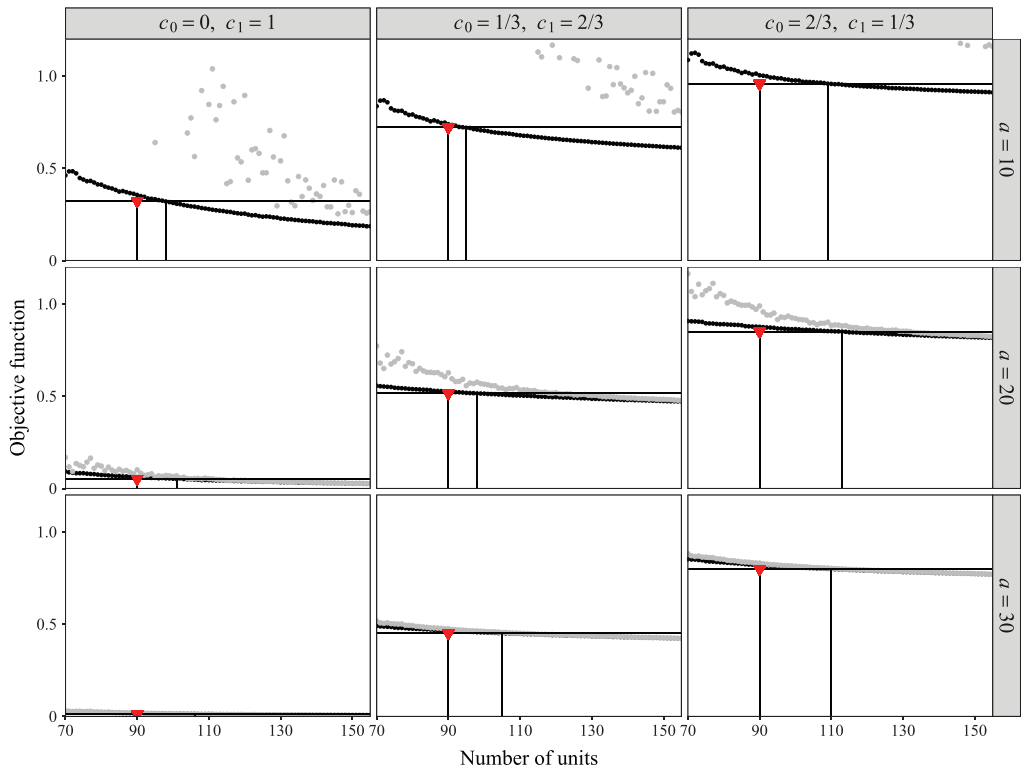


FIGURE 8 Value of objective function for sc_+ (black dots), sc (grey dots) and optimized (red triangle) schemes. The spacing between the two vertical lines indicates the number of extra units required for sc_+ to achieve an optimized objective function value for $\nu = 2$

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Additional number of sampling units in the spatial coverage supplemented with close-pairs of points schemes (sc_+) required to achieve the same objective function as that of the optimized 90-unit survey

c_0	c_1	a	$\nu = 0.2$	$\nu = 0.5$	$\nu = 1.1$	$\nu = 2$
0	1	10	8	8	9	8
1/3	2/3	10	20	10	7	5
2/3	1/3	10	59	34	24	19
0	1	20	2	6	8	11
1/3	2/3	20	8	3	4	8
2/3	1/3	20	25	16	18	23
0	1	30	2	6	11	16
1/3	2/3	30	10	5	11	15
2/3	1/3	30	25	25	11	20

average fewer units than the sc schemes. There is a clear association between an increase in the required sample size, increase in the nugget to sill ratio and decrease in the smoothness and distance parameters. When compared to the effective range of the target property (i.e. the distance at which the spatially correlated portion of the variogram attains 95% of the sill), the minimum number of units increased with decreasing values of the effective range. The dispersion variance (denoted σ^2 in Table 2) increased with larger values of the nugget to sill ratio and larger values of the distance parameter.

3.2 | Scenario 2

Figure 10 shows an example of sc and sc_+, as well as the optimized 90-unit scheme obtained by minimization of the expected total error using the average soil clay variogram. The optimized scheme had sampling units dispersed evenly over the area with a number of close-pair units. The number of close-pair units seems slightly larger than that of the sc_+ scheme. While the sc_+ and optimized scheme share some similarity in the pattern of sampling locations, the sc scheme is very different from the optimized scheme.

This is confirmed by Figure 11 which shows values of the objective function for sc_+, sc and optimized schemes using the average clay variogram. The sc scheme performed poorly until about 200 units. In contrast, the sc_+ had objective function values closer to that of the optimized scheme. Twenty-two additional locations were required for the sc_+ scheme to reach the objective function of the optimized scheme, which was achieved with a total of 11 close-pairs in the sc_+ scheme (out of 112).

Figure 12 shows the standardized soil clay variograms and the average variogram. First, the average variogram was used to compute the optimized scheme. Second, we found the sample size for the sc_+ scheme for each separate clay variogram to achieve the total variance of the optimized scheme. Overall, the optimized scheme was fairly robust with contrasting standardized soil clay variograms because it gave about the same total variance for most of the individual variograms as for the average variogram. Figure 12 shows that a large number of additional units were needed (>100) when large sill values of the variogram were reached in a short distance. In addition, fewer units were needed (< −5) when the total sill was reached at large distances. For similar values of the distance parameter, more units were needed for larger values of the nugget variance, e.g. weaker spatial dependence at short distances (see for example the two clay variograms with similar distance parameters but different nugget values).

4 | DISCUSSION

For all optimized schemes, there was a number of close-pair units. This shows that sampling units at short distances had a critical effect on decreasing the total expected error (which encompasses uncertainty in the variogram parameters and kriging variance). The number of close-pair units increased according to the nugget to sill ratio and to a lesser extent relative to the distance parameter. This was an expected result, because a random variable with a small spatial correlation distance and large nugget to sill ratio had to be sampled at a large number of short-distance locations to ensure minimization of uncertainty in both variogram parameters and prediction error variances (Marchant & Lark, 2007). This explains why sc schemes performed poorly in all cases. The sc schemes lacked close-pair units to estimate the spatial correlation over short distances which have a large effect on total expected error. Sampling schemes containing a subset of 10% as close-pairs (suggested by Lark and Marchant (2018)) provide a robust strategy to ensure a reasonably small total expected error. In the case of a small distance parameter or large nugget to sill ratio, 10% of close-pairs does not provide sufficient information and it is better to either increase the ratio of units taken at short distances or to use an optimized scheme.

The test presented in Figure 9 suggests that the importance of close pairs is reduced if the smoothness parameter is assumed to be known. In practice, however, this was not the case. Assuming a particular smoothness value (e.g. 0.5 for the exponential variogram) for a regularly sampled soil property led to a substantial proportion of the uncertainty being disregarded. This choice was somewhat subjective because it was related to the decision of the modeller and the range of possibilities we allowed in our model. We point out that close pairs are not only important when the nugget to sill ratio is large (Table 1) and the range of spatial correlation ($3 \times a$ if $\nu = 0.5$) is small relative to the size of the study area (Table 2), but also when one needs to estimate the additional Matérn model parameter ν (Figure 9).

Results from our second scenario showed that the optimized scheme based on the average variogram was fairly robust for contrasting soil clay variograms. For several variograms, an sc_+ scheme outperformed the optimized scheme. This was an unexpected result at first sight. The reason is

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

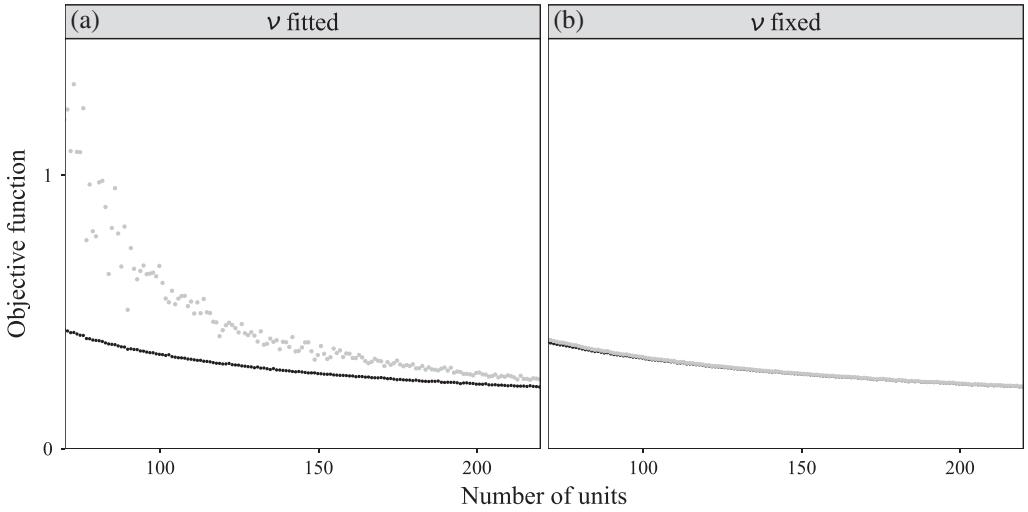


FIGURE 9 Value of objective function for sc_+ scheme (black dots) and sc scheme (grey dots) for $c_0 = 0$, $c_1 = 1$, $a = 20$ and $\nu = 0.5$. In (a) the smoothness parameter is to be estimated while in (b) it is assumed to be known

that the sampling scheme was optimized for the average variogram, and can therefore be suboptimal for an individual variogram. The results of the second scenario suggested that databases of variogram parameters (e.g. the one of Paterson et al. (2018)) can be used to derive an average variogram, and that the latter can be used to guide sampling (McBratney & Pringle, 1999) or to predict a soil property from fewer units than usually required for estimating variogram parameters (Kerry & Oliver, 2004). An average variogram could also provide prior information for expert or Bayesian elicitation of the variogram (Cui, Stein, & Myers, 1995).

In our two scenarios, close-pair units were taken at a fixed distance from one of the spatial coverage units. There might be room for further research on how these close-pair units should be selected. For example, in several optimized schemes, transects of several units can be seen. Further tests on our scenario 1 (not shown) suggested that selecting close-pairs in a cluster might reduce substantially the number of additional units needed with an increasing nugget to sill ratio. Such a scheme would, however, rely heavily on the assumption of stationarity (i.e. that the short-scale variation in the cluster indicates the short-scale variation across the study area). Our results here were for ordinary kriging in which the local mean of the variable was assumed to be constant. Sampling to support universal kriging (to model a non-stationary mean) or to support kriging with external drift (to model dependence of the mean on covariates) introduces other considerations, specifically estimation of the fixed effects parameters in the model. This requires further work. We speculate that the supplemented spatial coverage schemes that we have shown to be efficient for ordinary kriging would also be efficient for universal kriging, in that the spatial coverage points would ensure reliable estimation of trend parameters, and the close-pairs would similarly ensure that the variance parameters are estimated precisely.

For the optimized schemes in scenario 1, derived from a variogram with a small distance parameter and large nugget to sill ratio, the sample size seems too small compared to the size of the study area. Table 2 shows that this is indeed the case for several variogram types, and especially if the units are based on a spatial coverage scheme. This can lead to

TABLE 2 Minimum number of units required for the expected total prediction error variance to be smaller than the estimation variance of the target property that would result from a design-based survey of the same size. The dispersion variance is derived by averaging the variance of 1000 simulations using the lower-upper (LU) decomposition (Davis, 1987). The simulations are realized using 2000 units, selected by simple random sampling. In addition, the effective ranges of the different variogram types, denoted r, are reported

			$\nu = 0.2$				$\nu = 0.5$				$\nu = 1.1$				$\nu = 2$			
c_0	c_1	a	σ^2	r	sc	sc_+	σ^2	r	sc	sc_+	σ^2	r	sc	sc_+	σ^2	r	sc	sc_+
0	1	10	0.97	22	>200	75	0.98	21	164	61	0.97	20	104	54	0.98	19	95	49
1/3	2/3	10	0.98	22	>200	79	0.99	21	>200	52	0.98	20	128	67	0.98	19	109	72
2/3	1/3	10	0.99	22	>200	>200	0.99	21	>200	83	0.99	20	>200	79	0.99	19	158	72
0	1	20	0.93	45	95	28	0.93	42	84	20	0.93	40	54	20	0.92	38	42	24
1/3	2/3	20	0.95	45	>200	77	0.95	42	95	24	0.94	40	62	24	0.94	38	48	20
2/3	1/3	20	0.98	45	195	145	0.98	42	124	66	0.97	40	163	65	0.97	38	163	>200
0	1	30	0.88	67	104	20	0.87	64	54	22	0.83	60	31	11	0.85	57	24	13
1/3	2/3	30	0.91	67	77	23	0.90	64	62	16	0.91	60	48	16	0.87	57	45	15
2/3	1/3	30	0.96	67	>200	136	0.95	64	92	147	0.94	60	72	27	0.93	57	73	46

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

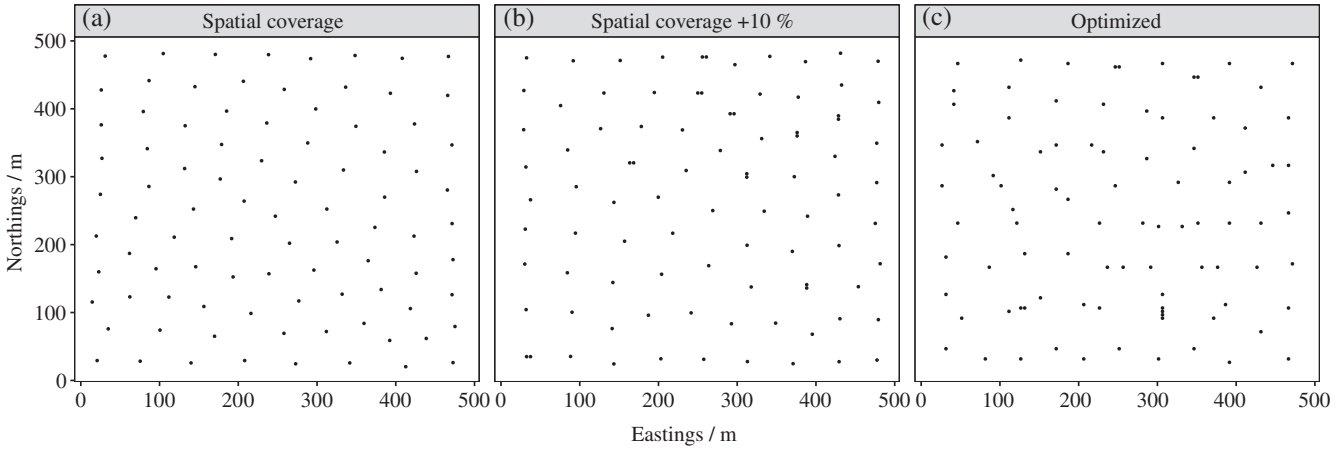


FIGURE 10 Example of 90-unit sc scheme (a), sc_+ scheme (b) and optimized scheme for the average soil clay variogram (c)

situations where the expected total variance is larger than the total sill variance. In such circumstances adding close-paired observations might resolve the problem of parameter estimation, but the overall sampling scheme remains inadequate for the task of spatial mapping because the spacing between neighbouring observations in the spatial coverage scheme is not sufficiently small relative to the range of spatial dependence. If one kriges from a grid with spacing larger than the range, then the prediction error variance is equal to the sill variance plus the Lagrange parameter, which is equivalent to the second term for the prediction variance of the spatial mean as a point predictor in Equation (12). This points us to the fact that, in these circumstances, where we cannot afford a grid with spacing that is small relative to the range, spatial prediction by kriging is not an option. In these circumstances point prediction might, in the worst case, be the regional mean of the variable, estimated by design-based sampling and with a prediction error variance computed from Equation (12). It might be possible to do better by estimating mean values within subregions of the area of interest such as soil map units (Webster & Beckett, 1968), again by design-based sampling, or by undertaking design-based sampling to estimate parameters of a predictive relation between the soil property of interest and covariates such as data from remote sensors. We hope that this clarifies why we refer to design-based estimation in the paper. It is not the case that design-based sampling does not provide a basis for spatial prediction. Design-based simply refers to the sampling scheme (probability sampling) and the basis for estimation from the data. The resulting design-based mean (for a region or subregion) may then be treated as a spatial prediction, as discussed by Webster and Lark (2013).

Table 2 also shows that for large value of smoothness ($\nu = 1.1$ or 2) and small nugget to sill ratios, the minimum number of units needed to make geostatistical analysis more accurate than a design-based estimate, on average, is surprisingly small. This can be explained by the relatively large values of the effective range ($r = 60$ and $r = 57$) for the

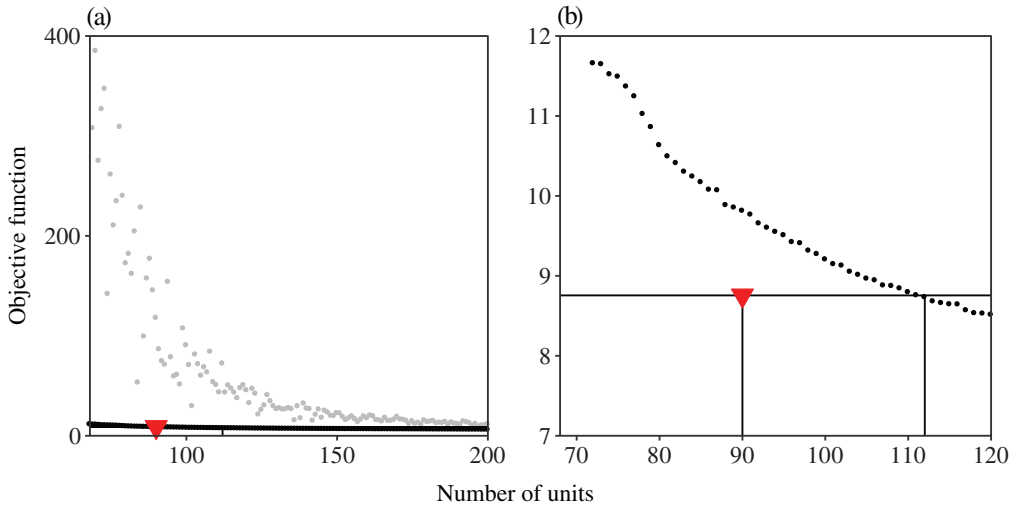


FIGURE 11 Values of objective function for sc_+ scheme (black dots) and sc scheme (grey dots) and optimized scheme (the red triangle). The spacing between the two vertical lines in (b) indicates the extra units required for sc_+ to achieve an objective function value from the optimized scheme for the soil clay average variogram

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

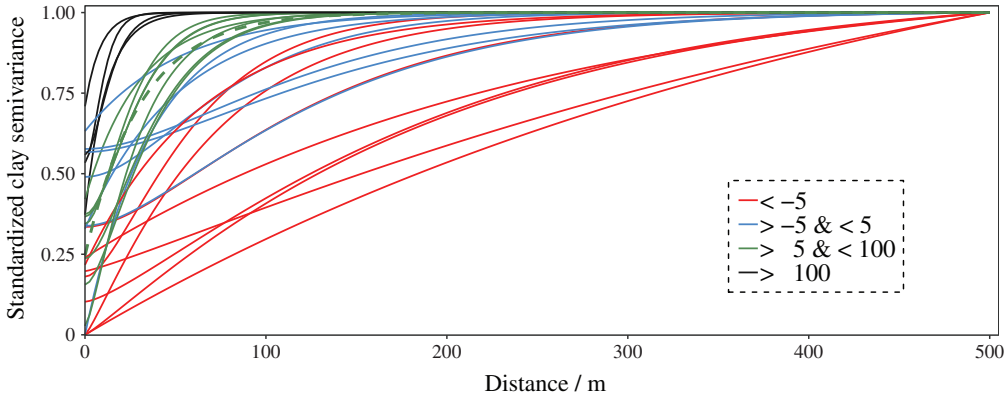


FIGURE 12 Standardized variograms of soil clay. The dashed line represents the average variogram. The color defines the number of additional sc_+ units required to reach the same value of the objective function for each variogram as the scheme optimized with the average variogram

case study (square of 100 m × 100 m); most sampling units were within the range of spatial correlation. However, for random functions with larger nugget to sill ratios, the design-based survey was more accurate even when the survey consisted of more than 200 units. Thus, the number of units required to estimate the variogram from a geostatistical survey depended on the degree of spatial correlation of the target property. We acknowledge that these total prediction variances are based upon a Taylor series approximation to the true variances.

5 | CONCLUSIONS

From the results and discussion we draw the following conclusions.

- The sc schemes performed poorly in almost all cases because of the lack of information at short distance to estimate the variogram parameters.
- Uncertainty of the sc scheme was mainly characterized by uncertainty of the smoothness parameter. Performance of the sc scheme can therefore be greatly improved by assuming that the smoothness is known, for example with an exponential variogram. However, in practice we have no justification for making such an assumption.
- The benefit of using an optimized scheme over an sc_+ scheme was clear but still generally modest. In addition, the optimization required the variogram parameters to be known.
- The benefit of using an optimized scheme over an sc_+ scheme became more important with an increasing nugget to sill ratio (weaker spatial dependence). In this case, geostatistical survey was unlikely to be effective.
- For a random variable with zero nugget and a large range of spatial correlation fewer than 15 observations were required to obtain average total prediction variances that were smaller than the prediction variance of the design-based estimate of the regional mean, treated as a point prediction at each location. However, 200 observations of a random variable with a substantial nugget effect were insufficient to meet the same criterion.
- When the scale of spatial variation of the soil property was not known, using an average variogram for optimizing the sampling scheme is a robust strategy.
- Overall, the tests conducted showed that there was little evidence of large benefits from optimizing sampling schemes. Therefore, it is better in most cases to use a spatial coverage scheme supplemented by a subset of close-pair units unless prior knowledge of the variogram is available (e.g. reconnaissance survey).

ACKNOWLEDGEMENTS

This project received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 607000. Ben Marchant's contribution is published with the permission of the Executive Director of the British Geological Survey (Natural Environment Research Council). We thank the anonymous reviewers and the editor for their constructive comments, which helped to improve this manuscript.

ORCID

Alexandre M.J.-C. Wadoux https://orcid.org/0000-0001-7325-9716
Richard M. Lark https://orcid.org/0000-0003-2571-8521

REFERENCES

Abramowitz, M., & Stegun, I. A. (1972). *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables* (Vol. 55). New York, NY: Dover Publications.

Brus, D. J., & De Gruijter, J. J. (1993). Design-based versus model-based estimates of spatial means: Theory and application in environmental soil science. *Environmetrics*, *4*, 123–152.

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Brus, D. J., De Gruijter, J. J., & Breeuwsma, A. (1992). Strategies for updating soil survey information: A case study to estimate phosphate sorption characteristics. *Journal of Soil Science*, *43*, 567–581.

Brus, D. J., De Gruijter, J. J., & Van Groenigen, J. W. (2007). Designing spatial coverage samples using the k-means clustering algorithm. *Developments in Soil Science*, *31*, 183–192.

Brus, D. J., Spätjens, L. E. E. M., & De Gruijter, J. J. (1999). A sampling scheme for estimating the mean extractable phosphorus concentration of fields for environmental regulation. *Geoderma*, *89*, 129–148.

Cressie, N., & Hawkins, D. M. (1980). Robust estimation of the variogram: I. *Journal of the International Association for Mathematical Geology*, *12*, 115–125.

Cui, H., Stein, A., & Myers, D. E. (1995). Extension of spatial information, bayesian kriging and updating of prior variogram parameters. *Environmetrics*, *6*, 373–384.

Davis, M. W. (1987). Production of conditional simulations via the LU triangular decomposition of the covariance matrix. *Mathematical Geology*, *19*, 91–98.

De Gruijter, J. J., Brus, D. J., Bierkens, M. F. P., & Knotters, M. (2006). *Sampling for Natural Resource Monitoring* (pp. 166-168). Dordrecht: Springer Science & Business Media.

Kerry, R., & Oliver, M. (2004). Average variograms to guide soil sampling. *International Journal of Applied Earth Observation and Geoinformation*, *5*, 307–325.

Kerry, R., & Oliver, M. (2007). Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. *Geoderma*, *140*, 383–396.

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, *220*, 671–680.

Kitanidis, P. K. (1987). Parametric estimation of covariances of regionalized variables. *JAWRA Journal of the American Water Resources Association*, *23*, 557–567.

Lark, R., & Marchant, B. (2018). How should a spatial-coverage sample design for a geostatistical soil survey be supplemented to support estimation of spatial covariance parameters? *Geoderma*, *319*, 89–99.

Marchant, B., & Lark, R. (2007). Optimized sample schemes for geostatistical surveys. *Mathematical Geology*, *39*, 113–134.

Matérn, B. (1986). *Spatial Variation*. Berlin: Springer.

McBratney, A. B., & Pringle, M. J. (1999). Estimating average and proportional variograms of soil properties and their potential use in precision agriculture. *Precision Agriculture*, *1*, 125–152.

Minasny, B., & McBratney, A. B. (2005). The Matérn function as a general model for soil variograms. *Geoderma*, *128*, 192–207.

Paterson, S., McBratney, A. B., Minasny, B., & Pringle, M. J. (2018). Chapter 21: Variograms of soil properties for agricultural and environmental applications. In *Pedometrics* (pp. 623–667). Berlin: Springer.

Royle, J. A., & Nychka, D. (1998). An algorithm for the construction of spatial coverage designs with implementation in SPLUS. *Computers & Geosciences*, *24*, 479–488.

Samuel-Rosa, A. 2017. *Spsann: Optimization of Sample Configurations using Spatial Simulated Annealing*. Retrieved from https://CRAN.R-project.org/package=spsann R package version 2.1–0.

Starks, T. H. (1986). Determination of support in soil sampling. *Mathematical Geology*, *18*, 529–537.

Stein, M. L. (2006). *Interpolation of Spatial Data: Some Theory for Kriging*. Dordrecht: Springer Science & Business Media.

Van Groenigen, J. W., Siderius, W., & Stein, A. (1999). Constrained optimisation of soil sampling for minimisation of the kriging variance. *Geoderma*, *87*, 239–259.

Wadoux, A. M. J.-C., Brus, D. J., Rico-Ramirez, M. A., & Heuvelink, G. B. M. (2017). Sampling design optimisation for rainfall prediction using a non-stationary geostatistical model. *Advances in Water Resources*, *107*, 126–138.

Walvoort, D. J. J., Brus, D. J., & De Gruijter, J. J. (2010). An R package for spatial coverage sampling and random sampling from compact geographical strata by k-means. *Computers & Geosciences*, *36*, 1261–1267.

Webster, R., & Beckett, P. (1968). Quality and usefulness of soil maps. *Nature*, *219*, 680.

Webster, R., & Lark, R. M. (2013). *Field Sampling for Environmental Science and Management*. London: Routledge.

Webster, R., & Oliver, M. A. (1992). Sample adequately to estimate variograms of soil properties. *Journal of Soil Science*, *43*, 177–192.

Webster, R., & Oliver, M. A. (2007). *Geostatistics for Environmental Scientists*. Chichester: John Wiley & Sons.

Yfantis, E. A., Flatman, G. T., & Behar, J. V. (1987). Efficiency of kriging estimation for square, triangular, and hexagonal grids. *Mathematical Geology*, *19*, 183–205.

Zhu, Z., & Stein, M. L. (2006). Spatial sampling design for prediction with estimated parameters. *Journal of Agricultural, Biological, and Environmental Statistics*, *11*, 24.

How to cite this article: Wadoux AM, Marchant BP, Lark RM. Efficient sampling for geostatistical surveys. *Eur J Soil Sci*. 2019;70:975–989. https://doi.org/10.1111/ejss.12797

13652389, 2019, 5, Downloaded from https://bsssjournals.onlinelibrary.wiley.com/doi/10.1111/ejss.12797 by James Cook University, Wiley Online Library on [21/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License