Spatial Sampling

70

Eric M. Delmelle

Contents

Abstract

Spatial sampling is the process of collecting observations in a two-dimensional framework. Careful attention is paid to (1) the quantity of the samples, dictated by the budget at hand, and (2) the location of the samples. A sampling scheme is generally designed to maximize the probability of capturing the spatial variation of the variable under study. Once initial samples have been collected and its variation documented, additional measurements can be taken at other locations. This approach is known as second-phase sampling, and various optimization criteria have recently been proposed to determine the optimal location of these

E.M. Delmelle
Department of Geography and Earth Sciences, University of North Carolina at Charlotte, Charlotte, NC, USA
e-mail: Eric.Delmelle@uncc.edu

M.M. Fischer, P. Nijkamp (eds.), *Handbook of Regional Science*,
DOI 10.1007/978-3-642-23430-9_73,

new observations. In this chapter, we review fundamentals of spatial sampling and second-phase designs. Their characteristics and merits under different situations are discussed, while a numerical example illustrates a modeling strategy to use covariate information in guiding the location of new samples. The chapter ends with a discussion on heuristic methods to accelerate the search procedure.

70.1 Introduction

70.1.1 Context

According to Haining (2003), spatial or two-dimensional sampling has been applied to many disciplines such as mining, soil studies, telecommunications, ecology, geology, and geography, to cite a few (Akella et al. 2011). Scientists may be constrained by available budget and time to acquire a certain number of samples instead of trying to obtain information everywhere (Müller 1998; Thompson 2002; Delmelle 2009). It is generally desirable to find samples that are as representative as possible from the real data. Not only the cost of a complete census is prohibitive, it is time-consuming (Haining 1990) and it may result in redundant data when those are spatially autocorrelated (Griffith 2005). The autocorrelation function is defined as the similarity of the values of the variable of interest as a function of their separating distance (Gatrell 1979). This similarity decreases as the distance among sample points increases. Positive autocorrelation occurs when nearby observations are more alike than samples collected further away. Sparse sampling is less costly, but the variability of the variable of interest may go unnoticed. Consequently, not only the quantity of the samples is important but also their locations.

70.1.2 One-Dimensional Sampling

Pioneering research on sampling was devoted to one-dimensional problems (see, e.g., Cochran 1946; Madow 1946, 1953; Madow and Madow 1949). Cochran documented the efficiency associated with random sampling, systematic sampling, and stratified sampling. A *random sampling* scheme (Fig. 70.1a) allocates n sample points randomly within a population of interest. Each location is equally likely selected. In a *systematic random sampling* (Fig. 70.1b), the population is partitioned into a prespecified number of intervals. For each interval, a number of samples are collected, and the total of all samples is of size n. In a *systematic sampling* scheme (Fig. 70.1c), the population of interest is divided into n intervals of similar size. The first element is chosen within the first interval, starting at the origin, and the remaining $n-1$ elements are aligned according to the same, fixed interval. A discussion of these configurations to the field of natural resources can be found in Stevens and Olsen (2004).

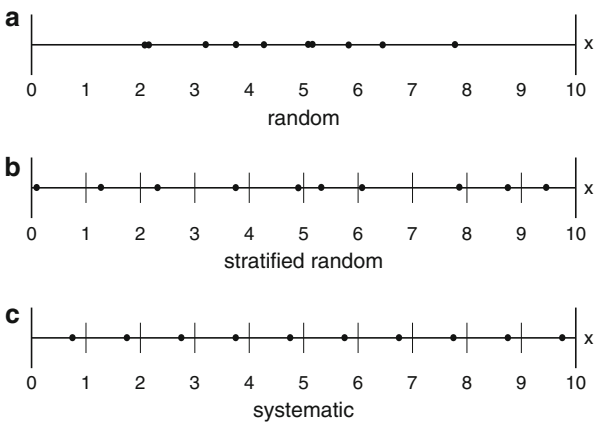

Fig. 70.1 One-dimensional sampling schemes for $n = 10$. The x-axis is partitioned in 10 intervals for cases (**b**) and (**c**). The random sampling locations have been generated using MATLAB rand function

70.1.3 Two-Dimensional Sampling

Necessary and common to both spatial and nonspatial sampling strategies are (i) the size of the sampling set, which is dictated by the budget (or time) at hand; (ii) the configuration of the sampling design; (iii) an estimator to characterize the population; and (iv) an estimation of the sampling variance to compute confidence intervals. Das (1950) has documented the variation of the sampling variance of two-dimensional designs. A *simple random sampling* design (Fig. 70.2a) randomly selects m sample points in a study region, generally denoted $\mathfrak{D}$, where each location has an equal opportunity to be sampled. In a *systematic sampling* design, (illustrations given in Fig. 70.2b–d), the study region is discretized into m intervals of equal size $\triangle$. The first element is randomly or purposively chosen within the first interval, and so are other points in the remaining regions. If the first sample is chosen at random, the resulting scheme is called *systematic random sampling*. When the first sample point is not chosen at random, the resulting configuration is called *regular systematic sampling*. A *centric systematic sampling* occurs when the first point is chosen in the center of the first interval, resulting in a checkerboard configuration. The most common regular geometric configurations are the equilateral triangular grid, the rectangular (square) grid, and the hexagonal one (Cressie 1991). The benefits of a systematic approach reside in a good spreading of observations across $\mathfrak{D}$, guaranteeing a maximized sampling coverage, and preventing sampling clustering and redundancy. This design however presents two inconveniences:

(a) The distribution of separating distances in $\mathfrak{D}$ is not represented well because many pairs of points are separated by the same distancesss,
(b) If the spatial process shows evidence of recurrence, periodicities, there is a risk that the variation of the variable will remain uncaptured, because the systematic design coincides in frequency with a regular pattern in the landscape (Overton and Stehman 1993).

A *systematic random method* addresses the second concern since it combines both systematic and random procedures (Dalton et al. 1975). One observation is

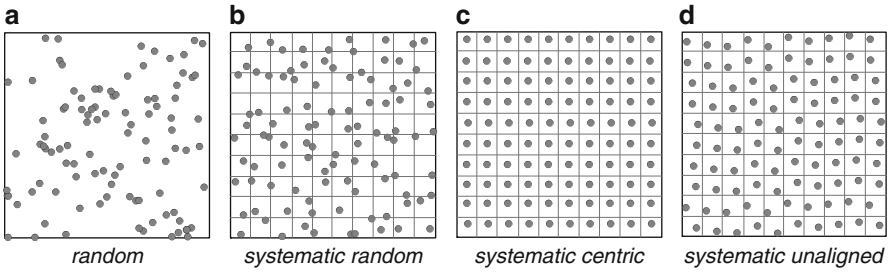

Fig. 70.2 Two-dimensional sampling schemes for $n = 100$. In figures (**b**), (**c**), and (**d**), both x- and y-axis have been divided into 10 intervals. Points were randomly generated using MATLAB rand function

randomly selected within each cell. However, sample density needs to be high enough in order to document the strength of the spatial relationship (e.g., variogram) among observations. From Fig. 70.2b, some patches of $\mathfrak{D}$ remain undersampled, while others regions show evidence of clustered observations. A *systematic unaligned* scheme prevents this problem from occurring by imposing a stronger restriction on the random allocation of observations (King 1969).

In *stratified sampling* (Delmelle 2009), the population (or $\mathfrak{D}$) is partitioned into nonoverlapping strata. A set of samples is collected for each stratum, where the sum of the samples over all strata must equal m (strata may be of different size, for instance, a census tract). The knowledge of the underlying process is a determining factor in defining the shape and size of each stratum. Smaller strata are preferred in nonhomogeneous subregions.

Evaluation of Sampling Strategies. Following Quenouille's approach of a linear autocorrelation model, stratified random sampling is generally considered to yield a smaller variance than a systematic design. However, if the autocorrelation function is not linear (for instance, exponential), systematic sampling is the most efficient technique, followed by stratified random sampling and random sampling. Overton and Stehman (1993) presented some numerical results illustrating the magnitude of the differences of the three aforementioned designs under various population models. When sampling a phenomenon characterized by a regular pattern in the landscape, a systematic unaligned configuration is generally preferred (Delmelle 2009).

70.2 Geostatistical Sampling

An essential commonality of many natural phenomena is its spatial continuity in the geographical space. The field of geostatistics (Matheron 1963) provides a set of regression techniques to mathematically summarize the spatial variation of the phenomenon and use this information to predict the phenomenon under study at unsampled locations. Central to geostatistics is kriging, an interpolation technique that uses the semivariogram, a function which reflects the dissimilarity of pairs of points at different distance lags. The strength of this correlation determines the

weighting scheme used to create a prediction surface at unsampled locations, while minimizing the estimation error. As the distance separating two sample points increases, their similarity decreases and the influence on the weighting scheme diminishes. Beyond a specific distance called the range where autocorrelation is very small, the semivariogram flattens out (see, e.g., Ripley (1981) and Cressie (1991) for various summaries).

Mathematical Expression for Kriging. A variable of interest Y is collected at m supports within a study region $\mathfrak{D}$. Using data values of the primary variable, an empirical semivariogram $\widehat{\gamma}(h)$ summarizes the variance of values separated by a particular distance lag (h):

$$\widehat{\gamma}(h) = \frac{1}{2d(h)} \sum_{|\mathbf{s}_i - \mathbf{s}_j| = h} (y(\mathbf{s}_i) - y(\mathbf{s}_j))^2 \tag{70.1}$$

where $d(h)$ is the number of pairs of points for a given lag value, and $y(\mathbf{s}_i)$ the observation value at location $\mathbf{s}_i$. The semivariogram is characterized by a nugget effect a and a sill σ^2 where $\widehat{\gamma}(h)$ levels out. The nugget effect reflects the spatial dependence at microscales, caused by measurement errors at distances smaller than sampling distances (Cressie 1991). Once the lag distance exceeds the range r, there is no spatial dependence between the sample sites anymore. The semivariogram function $\widehat{\gamma}(h)$ becomes constant at a value called the sill σ^2. A model $\gamma(h)$ is fitted to the experimental variogram, for instance, an exponential model:

$$\gamma(h) = \sigma^2 \left(1 - e^{\frac{-3h}{r}}\right) \tag{70.2}$$

In the presence of a nugget effect a, Eq. (70.2) becomes

$$\gamma(h) = a + (\sigma^2 - a)\left(1 - e^{\frac{-3h}{r}}\right) \tag{70.3}$$

Equation (70.4) denotes the corresponding covariogram $C(h)$ that summarizes the covariance between any two points:

$$C(h) = C(0) - \gamma(h) = \sigma^2 - \gamma(h) \tag{70.4}$$

The interpolated, kriged value at a location $\mathbf{s}$ in space is given as a weighted mean of surrounding values, where each value is weighted according to the variogram model:

$$\widehat{y}(\mathbf{s}) = \sum_{i=1}^{W} w_i(\mathbf{s}) y(\mathbf{s}_i) \tag{70.5}$$

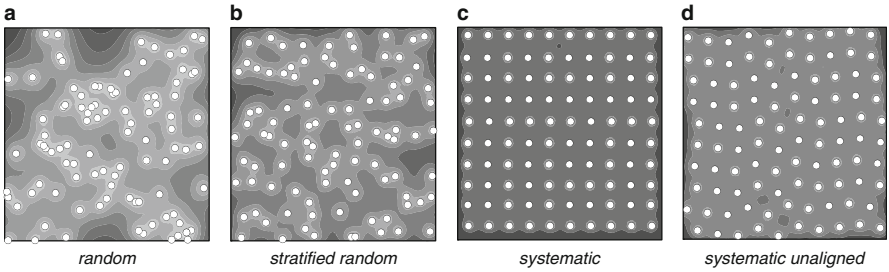

Fig. 70.3 Kriging variance associated with the two-dimensional sampling schemes of Fig. 70.3

where W is the set of neighboring points that are used to estimate the interpolated value at location $\mathbf{s}$, and $w_i(\mathbf{s})$ is the weight associated with each surrounding point, which is a function of the semivariogram function. The weight of each sample can be determined by an exponential function (Eq. (70.2)). For computational purposes, kriging is performed on a set of grid nodes $\mathbf{s}_g$ $(g = 1, 2, \ldots, G)$. Kriging yields an associated variance that measures the prediction uncertainty. The kriging variance at a location $\mathbf{s}$ is given by

$$\sigma_k^2(\mathbf{s}) = \sigma^2 - \mathbf{c}^T(\mathbf{s})\mathbf{C}^{-1}\mathbf{c}(\mathbf{s}) \tag{70.6}$$

where $\mathbf{c}^T$ is the transpose of the covariance matrix $\mathbf{C}$ based on the covariogram function and $\mathbf{C}^{-1}$ its inverse. The overall kriging variance (σ_k^2) is obtained by integrating Eq. (70.6) over the region $\mathfrak{D}$. Computationally, it is easier to perform a spatial of $\mathfrak{D}$ and sum the kriging variance over all grid points $\mathbf{s}_g$:

$$\int_{\mathfrak{D}} \sigma_k^2(\mathbf{s}_g) \approx \frac{1}{G} \sum_{g \varepsilon G} \sigma_k^2(\mathbf{s}_g) \tag{70.7}$$

The kriging variance can be calculated with an estimated variogram and the known location of existing sampling points. The kriging variance solely depends on the spatial dependence and configuration of the observations (Cressie 1991). Figure 70.3 summarizes the variation in the kriging variance estimate for the four designs of Fig. 70.2.

Van Groenigen et al. (1998, 1999) suggest that initial sampling schemes should be optimized for a reliable estimation of the variogram function, which can either be used for the prediction of the variable under study or to help designing additional sampling phase(s). For the former, two strategies have been suggested in the literature:

(a) A geometric coverage of sample points over the study region is generally desirable to guarantee enough pairs of points at different distances.
(b) Points need to be distributed in the multivariate field to capture as much variation as possible.

Moreover, optimal sampling strategies exist to reduce the kriging variance associated with the interpolation process. The next paragraphs illustrate three common objectives in spatial sampling: variogram estimation, minimization of the kriging variance, and sampling in a multivariate field.

70.2.1 Designs for Variogram Estimation

Traditional ways to evaluate the goodness of a sampling scheme do not incorporate the spatial structure of the variable. The increasing use of geostatistics as a least-squares interpolation technique, however, has fostered research on optimizing sampling configurations to maximize the amount of information obtained during a first sampling phase. Matérn (1960) and Yfantis et al. (1987) have suggested that the use of an equilateral triangular sampling grid (Fig. 70.4) can yield to a very reliable estimation of the variogram and predict the mean over a study region, assuming radially symmetric, decreasing covariances.

Systematic designs (Fig. 70.2c, d) offer the advantage of good coverage of observations, capturing the main features of the variogram (Van Groenigen et al. 1999). It may be necessary to strategically design a scheme where a subset of the observations are evenly spread across the study area and the remaining points clustered together to capture the autocorrelation function at very small distances (Delmelle 2009).

The reliability of the variogram function depends on the number of pairs of points within each distance band. Russo (1984) and Warrick and Myers (1987) have proposed some strategies to reproduce an a priori defined ideal distribution of pairs of points, based on a given variogram function. The procedure allows to account for the variation in distance and direction (anisotropy[1]). Corsten and Stein (1994) use a nested sampling design for a better estimation of the nugget effect. A nested sampling design consists of taking observations according to a hierarchical scheme, with decreasing distances between observations. This type of sampling scheme distributes a high number of observations in some parts of the area and a low observation density in other regions. This in turn generates only a few distances for which variogram values are available. Taking into account a prior information of the spatial structure of the variable and assuming a stationary variable, Van Groenigen and Stein (1998) have combined two different objectives to allocate samples during an initial phase. The first objective called the Warrick/Myers criterion ensures optimal estimation of the covariogram and aims at redistributing pairs of points over the distance and direction lags according to a prespecified distribution. The second criterion, called *minimization of the mean of the shortest distances* (MMSD), requires all sampling points spread evenly to ensure that

[1]Anisotropy is a property of a natural process, where the autocorrelation among points changes with the distance and direction between two locations. We talk about an isotropic process however when there is no effect of direction in the spatial autocorrelation of the primary variable.

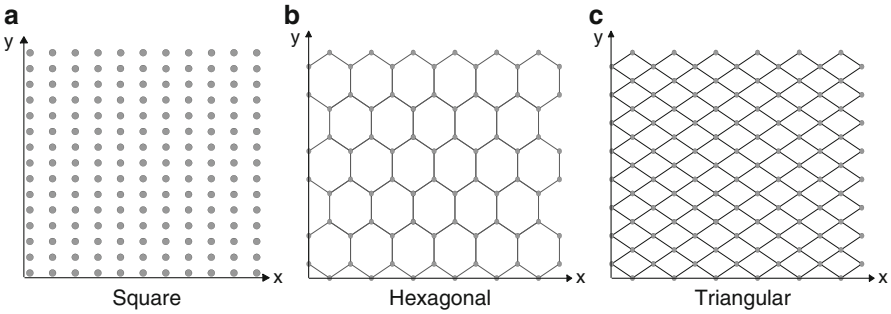

Fig. 70.4 Three common geometric sampling schemes

unsampled locations are never far from a sampling point. The second criterion suggested by the authors is of deterministic nature, resulting an even spreading pairs of points across the study area, which is similar in nature to a systematic pattern.

70.2.2 Optimal Designs to Minimize the Kriging Variance

The kriging procedure generates a minimum-error estimate of the variable of interest. This uncertainty is minimal – or zero when there is no nugget effect – at existing sampling points and increases with the distance to the nearest samples. One approach suggested in the literature is to design a sampling configuration to minimize this uncertainty. Since continuous sampling is not feasible, seeking the best sampling procedure must be carried out on a discretized grid. Using an a priori variogram model (Eq. (70.4)), it is possible to design an initial sampling scheme S to minimize the overall kriging variance (Eq. (70.8)) or the maximum kriging variance (Eq. (70.9)).

$$\underbrace{\text{MINIMIZE}}_{\{s_1,\ldots,s_m\}} J(S) = \frac{1}{G}\sum_{g\varepsilon G}\sigma_k^2(s_g;S) \tag{70.8}$$

$$\underbrace{\text{MINIMIZE}}_{\{s_1,\ldots,s_m\}} J(S) = \frac{1}{G}\underbrace{\sup}_{g\in G}\left\{\sigma_k^2(s_g;S)\right\} \tag{70.9}$$

Burgess et al. (1981) estimate kriging variances for different scenarios of sampling densities, nugget effects, and size of study regions. The strategy attempts to identify the minimum number of samples necessary to reach a certain level of kriging variance. General findings are the increase of the prediction error as the nugget effect increases or when the sampling density is reduced. An equilateral

triangular configuration of sampling points (Fig. 70.4c) is best under isotropic conditions, but a square grid at the same density is nearly as good, and is preferred for data collection convenience. An equilateral triangle design will keep the variance to a minimum, because it reduces the farthest distance from initial sample points to nonsample points. A square grid performs well, especially in case of isotropy (McBratney and Webster 1981; McBratney et al. 1981). When directional discontinuities are present, a square grid pattern is preferred (Olea 1984; Yfantis et al. 1987).

70.2.3 Sampling in a Multivariate Context

McBratney and Webster (1983) have discussed the importance of spatial sampling to multivariate fields. Sample data can be very difficult to collect, and very expensive, especially when monitoring air or soil pollution, for instance (Haining 1990). Secondary data can be a valuable asset if they are available continuously over a study area and combined within the primary variable (Hengl et al. 2003). Secondary spatial data sources can include maps, digital elevation models, and national, socioeconomic, and demographic census data. Cross-variograms express the spatial relationships among those variables. In turn, this information is capitalized to calibrate the parameters of the kriging equations. When the variogram of the primary variable and the cross-variograms are known a priori, an improved sampling configuration can be obtained. A rule of thumb consists of locating the observations of the main variable where covariates exhibit substantial spatial variation (Delmelle and Goovaerts 2009). Secondary variables should be used to reduce the sampling effort in areas where their local contribution in predicting the primary variable is maximum (Delmelle 2009). If a set of covariates predicts accurately the data value where no initial sample has been collected yet, there is little incentive to perform sampling at that location. On the other hand, when covariates perform poorly in estimating the primary variable, additional samples are necessary.

70.3 Second-Phase Sampling

Second-phase spatial sampling is defined as the addition of new observations to improve the overall prediction of the variable of interest. A set M of m initial measurements has been collected, and a variogram that summarizes the spatial structure of the variable of interest will help to determine the location and size for an additional set and location of their samples. It is generally agreed in the literature that the objective function aims to collect new samples to reduce the prediction error (kriging variance) by as much as possible.

Mathematical Expression for Minimizing of the Kriging Variance in a Second Phase. We add a set of n new sample points to the initial sample set of size m.

Using the variogram function from the first sample set, the change in kriging variance $\triangle\sigma_k^2$ is over all grid points $\mathbf{s}_g$:

$$\triangle\sigma_k^2 = \frac{1}{G}\left[\sum_{g\varepsilon G}\sigma_k^{\text{old}}(\mathbf{s}_g) - \sum_{g\varepsilon G}\sigma_k^{\text{new}}(\mathbf{s}_g)\right] \tag{70.10}$$

where σ_k^{old} is the mean kriging variance calculated with the set of $[m]$ initial sample points and σ_k^{new} is the mean kriging variance with the $[m+n]$ additional set of points. From Eq. (70.10)

$$\sigma_k^{\text{old}}(\mathbf{s}_g) = \sigma^2 - \underbrace{\mathbf{c}(\mathbf{s}_g)}_{[1,m]} \times \underbrace{\mathbf{C}^{-1}}_{[m]} \times \underbrace{\mathbf{c}^T(\mathbf{s}_g)}_{[m,1]} \tag{70.11}$$

$$\sigma_k^{\text{new}}(\mathbf{s}_g) = \sigma^2 - \underbrace{\mathbf{c}(\mathbf{s}_g)}_{[1,m+n]} \times \underbrace{\mathbf{C}^{-1}}_{[m+n]} \times \underbrace{\mathbf{c}^T(\mathbf{s}_g)}_{[m+n,1]} \tag{70.12}$$

The objective function helps to locate the set of additional n points that will maximize this change in kriging variance (Christakos and Olea 1992; Van Groenigen et al. 1999; Rogerson et al. 2004). The n additional points are to be chosen from a set of size $(N-m)$, that is, all possible sample sites in $\mathfrak{D}$ except the m ones selected during the first sampling phase. In that case, there are $\binom{N-m}{n}$ possible combinations and it is almost impossible to find the optimal using. The objective function is formulated as follows:

$$\underbrace{\text{MAXIMIZE}}_{\{s_{m+1},\ldots,s_{m+n}\}} J(S) = \frac{1}{G}\sum_{g\varepsilon G}\triangle\sigma_k^2(s_g;S) \tag{70.13}$$

Incorporating Secondary Information in a Second Sampling Phase. New samples can be collected in areas where secondary variables do not provide good estimates of the primary variable. Consider the situation where the primary data is supplemented by k additional secondary variables X_i $(\forall i = 1,\ldots,k)$ available at G grid nodes $\mathbf{s}_g$ $(g = 1,2,\ldots,G)$. Local regression techniques such as geographically weighted regression (Brunsdon et al. 1996) provide locally linear regression estimates at every point i, using distance weighted samples. Our goal is to sample in those areas characterized by low local r^2, since it is in those areas that covariates are not performing well in predicting the outcome of the primary variable. A local r^2 can be conceived as how well covariates predict the main variable locally, for instance, from a GWR model.

Formulating the Second-Phase Sampling Problem. This approach is proposed by Cressie and has been applied by Rogerson et al. (2004) and Delmelle and

Goovaerts (2009) to weight the kriging variance, where the importance of a location to be sampled is represented by a weight $w(\mathbf{s})$, which is location specific.

$$\underbrace{\text{MAXIMIZE}}_{\{s_{m+1},\ldots,s_{m+n}\}} J(S) = \frac{1}{G}\sum_{g\varepsilon G} w(s_g)\triangle\sigma_k^2(s_g;S) \tag{70.14}$$

The weight should reflect the importance provided locally by covariates, but could also account for the rapid change in spatial structure of the primary variable at $\mathbf{s}_g$ (Delmelle and Goovaerts 2009).

70.4 Numerical Example

A numerical example is provided to gain insight into the structure of the sampling problem. The goal is to maximize the change in the weighted kriging variance. As a hypothetical example, we simulated a synthetic snowfall data in a 10×10 km bounding box.

Minimizing the Kriging Variance. Figure 70.5 displays the initial set of 50 measurements and the associated interpolated map, based on an exponential semivariogram with a range of 3,000 m, a nugget effect $C(0) = 0$, and sill $a = 0.025$. The amount of snowfall is simulated to be minimal in the upper northwestern corner and increases steadily southeastwards. Figure 70.6 on the left is an interpolated contour map of the prediction error. The variance increases away from existing data points, to reach maximum values in the corner of the study area. The right figure displays the discretized study area, generating a set $P = 51 \times 51$ potential points. If the goal is to maximize the change in kriging variance only (Eq. (70.13)), the location of the new points would be far away from existing ones.

Weights to Reflect Sampling Priorities. An example of sampling weights is given in Fig. 70.3 on the left and is multiplied by the kriging variance on the right. As a result of this multiplication, some locations exhibiting high weight for second-phase sampling (where we observe a stronger variation in the spatial structure of the primary variable) may not be recommended for further sampling since they are located in the close vicinity of an existing initial sampling point. For instance, on the left figure, the location $[6,705,000;\ 4,718,500]$ is characterized by a high sampling weight. However, the region has already been sampled and consequently the likelihood for second-phase sampling decreases. Consider that we have the intention to add one sample point that would optimize Eq. (70.14). To find the point that would maximize the change in kriging variance, summed over all grid points, an iterative procedure is necessary which evaluates the score of each candidate sample points to the objective function. Such an enumeration can be very time-consuming, and for computational purposes, it is less demanding to select a point where the weighted kriging variance value is maximum. From Fig. 70.7, the location of the point exhibiting the maximum weighted kriging variance was

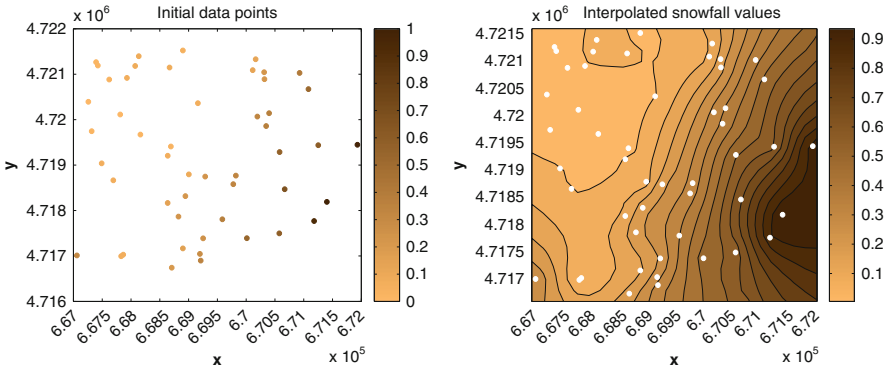

Fig. 70.5 *Dark colors* denote regions characterized by heavier amounts of snow. Units are in feet

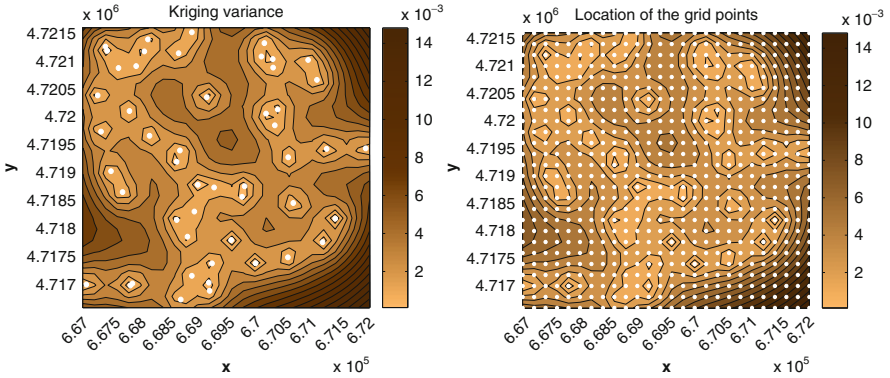

Fig. 70.6 The prediction error on the left and the location of grid nodes $\mathbf{s}_g$

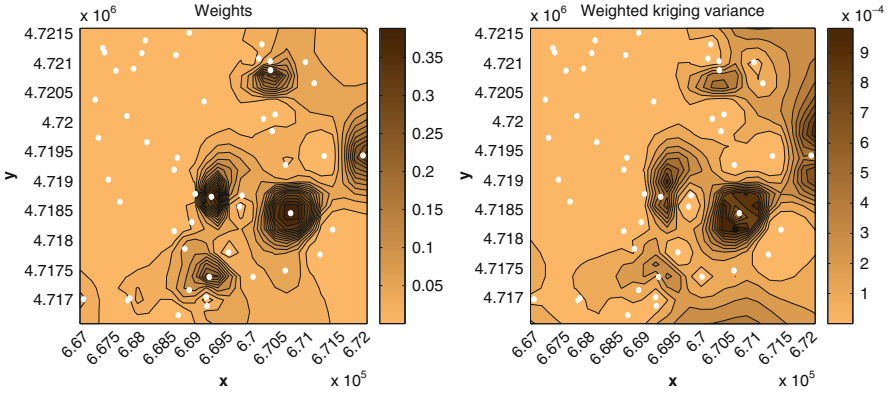

Fig. 70.7 Combined objectives 2 and 3 (on the *left*), multiplied with the kriging variance (on the *right*)

$[6,70,600;\ 4,718,200]$. Once an optimal (or near optimal) point has been added, it is possible to recompute the objective function. It is also desirable to adapt the constraints as the iteration continues.

70.5 Search Strategies

The set of candidate sampling locations may be large, and it is desirable to rely on heuristic techniques to return an acceptable solution in a limited time frame. A heuristic guides the search towards a sample set S that is optimal (or near optimal) to a predefined objective function, for instance, the set S^* is optimal to the objective function J defined in Eq. (70.14). The efficiency of a heuristic depends on its capacity to give as often as possible a solution close to S^*. In second-phase sampling, a heuristic that would select m points at random would not return a very good value for J. Examples of search techniques include the *greedy* algorithm, *simulated annealing*, *tabu search*, and *genetic algorithms*, among others (Christakos and Olea 1992; Delmelle and Goovaerts 2009). The *greedy* algorithm builds a solution sequentially only accepting improving moves, while the other methods improve the value of the objective function by iterations starting from an initial solution s_0, but also accepting non-improving moves. Specifically, the three first methods usually remain stuck in a local optimum while the last three – also called *metaheuristics* – may find the optimal solution s^*. Greedy leads to a unique solution S_0^+ and does not explore the entire set of candidate samples. *Simulated annealing* authorizes to occasionally decrease the objective function (in our case to be maximized) in order to continue exploring for better solutions. Note that the simulated annealing algorithm does not always converge. Both *tabu search* and *genetic algorithm* techniques have not been applied to sampling optimization. They both lead to an optimal solution, yet the tabu search algorithm tends to cycle.

70.6 Conclusion

Accurate and effective spatial sampling strategies are very important when researchers are limited by their available budget (or time). A careful design is crucial to identify the main features of the phenomenon under study and avoid that its spatial characteristics remain unnoticed. For instance, incorporating some randomness in a systematic sampling design may be useful to document patterns with periodicities. Once initial samplings have been collected, a variogram can be built, which ultimately helps designing a second-phase sampling survey (away from existing samples and where the variation is maximum). When the set of candidate locations is large and the objective nonlinear, heuristic methods may be necessary to find a set optimal to some sampling criteria. The methods illustrated in this chapter may easily be extended to areal data (for instance, census tracts or socioeconomic strata). Some areas may be deemed more important for sampling, and the proposed objectives are flexible to reflect sampling priorities.

References

Akella MR, Delmette EM, Batta R, Rogerson P, Blatt A (2011) Adaptive cell tower location using geostatistics. Geogr Anal 42(3):227–244

Brunsdon C, Fotheringham AS, Charlton ME (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr Anal 28(4):281–298

Burgess TM, Webster R, McBratney AB (1981) Optimal interpolation and isarithmic mapping of soil properties: IV. Sampling strategy. J Soil Sci 32(4):643–659

Christakos G, Olea RA (1992) Sampling design for spatially distributed hydrogeologic and environmental processes. Adv Water Res 15(4):219–237

Cochran WG (1946) Relative accuracy of systematic and stratified random samples for a certain class of populations. Ann Math Stat 17(2):164–177

Corsten LCA, Stein A (1994) Nested sampling for estimating spatial semivariograms compared to other designs. Appl Stoch Models Data Anal 10(2):103–122

Cressie N (1991) Statistics for spatial data. Wiley, New York

Dalton R, Garlick J, Minshull R, Robinson A (1975) Sampling techniques in geography. Goerges Philip and Son, London

Das AC (1950) Two-dimensional systematic sampling and the associated stratified and random sampling. Sankhyā 10(1–2):95–108

Delmelle E (2009) Spatial sampling. In: Rogerson P, Fotheringham S (eds) The SAGE handbook of spatial analysis. SAGE London

Delmelle E, Goovaerts P (2009) Second-phase spatial sampling designs for non-stationary spatial variables. Geoderma 153(1–2):205–216

Gatrell AC (1979) Autocorrelation in spaces. Environ Plan A 11(5):507–516

Griffith DA (2005) Effective geographic sample size in the presence of spatial autocorrelation. Ann Assoc Am Geogr 95(4):740–760

Haining RP (1990) Spatial data analysis in the social and environmental sciences. Cambridge University Press, Cambridge, UK

Hengl T, Rossiter DG, Stein A (2003) Soil sampling strategies for spatial prediction by correlation with auxiliary maps. Aust J Soil Res 41(8):1403–1422

King LJ (1969) Statistical analysis in geography. Prentice-Hall, Englewood Cliffs, pp 217–222

Madow LH (1946) Systematic sampling and its relation to other sampling designs. J Am Stat Assoc 41(234):204–217

Madow WG (1953) On the theory of systematic sampling. III. Comparison of centered and random start systematic sampling. Ann Math Stat 24(1):101–106

Madow WG, Madow LH (1949) On the theory of systematic sampling. I. Ann Math Stat 15(1):1–24

Matérn B (1960) Spatial variation. Springer, Berlin/Heidelberg/New York, p 151

Matheron G (1963) Principles of geostatistics. Econ Geol 58(8):1246–1266

McBratney AB, Webster R (1981) The design of optimal sampling schemes for local estimation and mapping of regionalized variables: II. Program and examples. Comput Geosci 7(4):331–334

McBratney AB, Webster R (1983) Optimal interpolation and isarithmic mapping of soil properties: V. Co-regionalization and multiple sampling strategy. J Soil Sci 34(1):137–162

McBratney AB, Webster R, Burgess TM (1981) The design of optimal sampling schemes for local estimation and mapping of regionalized variables: I. Theory and method. Comput Geosci 7(4):335–365

Müller W (1998) Collecting spatial data: optimal design of experiments for random fields. Physica, Heidelberg

Olea RA (1984) Sampling design optimization for spatial functions. Math Geol 16(4):369–392

Overton WS, Stehman SV (1993) Properties of designs for sampling continuous spatial resources from a triangular grid. Commun Stat Theory Methods 22(9):2641–2660

Ripley BD (1981) Spatial statistics. Wiley, New York, p 252

Rogerson PA, Delmelle EM, Batta R, Akella MR, Blatt A, Wilson G (2004) Optimal sampling design for variables with varying spatial importance. Geogr Anal 36(2):177–194
Russo D (1984) Design of an optimal sampling network for estimating the variogram. Soil Sci Soc Am J 48(4):708–716
Stevens D, Olsen A (2004) Spatially balanced sampling of natural resources. J Am Stat Assoc 99(465):262–278
Thompson SK (2002) Sampling, 2nd edn. Wiley, New York, p 367
Van Groenigen JW, Stein A (1998) Constrained optimization of spatial sampling using continuous simulated annealing. J Environ Qual 27(5):1078–1086
Van Groenigen JW, Siderius W, Stein A (1999) Constrained optimisation of soil sampling for minimisation of the kriging variance. Geoderma 87(3–4):239–259
Warrick AW, Myers DE (1987) Optimization of sampling locations for variogram calculations. Water Resour Res 23(3):496–500
Yfantis EA, Flatman GT, Behar JV (1987) Efficiency of kriging estimation for square, triangular and hexagonal grids. Math Geol 19(3):183–205