

Peter M. Atkinson and Christopher D. Lloyd

## Contents

74.1	Introduction .....	1462
74.2	Characterizing Spatial Variation .....	1463
74.2.1	The Experimental Variogram .....	1463
74.2.2	Variogram Model Fitting .....	1464
74.2.3	Example .....	1466
74.3	Spatial Interpolation with Kriging .....	1468
74.3.1	Simple and Ordinary Kriging .....	1468
74.3.2	Simulation .....	1470
74.4	The Change of Support Problem .....	1470
74.4.1	Regularization .....	1470
74.4.2	Variogram Deconvolution .....	1471
74.4.3	Area-to-Point Kriging .....	1471
74.5	Geostatistics and Regional Science .....	1473
74.6	Conclusions .....	1474
	References .....	1474

---

## Abstract

Characterizing the spatial structure of variables in the regional sciences is important for several reasons. Firstly, the spatial structure may itself be of interest. The structure of a population variable tells us something about how the population is configured spatially. For example, is the population clustered by some properties, but not others? Secondly, mapping variables from sparse

---

P.M. Atkinson (✉)

Geography and Environment, University of Southampton, Southampton, UK  
e-mail: [P.M.Atkinson@soton.ac.uk](mailto:P.M.Atkinson@soton.ac.uk)

C.D. Lloyd

School of Environmental Sciences, University of Liverpool, Liverpool, UK  
e-mail: [C.D.Lloyd@liverpool.ac.uk](mailto:C.D.Lloyd@liverpool.ac.uk)

sample observations or transferring values between areal units requires knowledge of how the property of interest varies spatially. Thirdly, we require knowledge of spatial variation in order to design sampling strategies which make the most of the effort, time, and money expended in sampling. Geostatistics comprises a set of principles and tools which can be applied to characterize or model spatial variation and use that model to optimize the mapping, simulation, and sampling of spatial properties. This chapter provides an introduction to some key ideas in geostatistics, with a particular focus on the kinds of applications which may be of interest for regional scientists.

---

## 74.1 Introduction

Geostatistics provides a body of techniques which can be used to characterize spatial variation, interpolate and simulate spatial variables, and design optimal spatial sampling strategies (Journel and Huijbregts 1978; Goovaerts 1997; Haining et al. 2010). To date, most applications of geostatistics have been in the physical sciences. However, there is an increasing diversity of research in the social sciences which makes use of geostatistical methods. This chapter reviews some of the key principles underlying geostatistics and considers the contributions these approaches may make in a regional science context.

Underlying classical geostatistics is the random function (RF) model (Journel and Huijbregts 1978; Chilès and Delfiner 1999). A RF is a set of random variables (RVs) which vary as a function of location  $\mathbf{x}$ . A RV is a stochastic process, a simple discrete example of which is the rolling of a die with the outcome being a value between one and six. Each outcome of this process is termed a *realization*. In most practical applications of geostatistics, variables are continuous. As an example, in a geostatistical framework, population density  $z$  at location  $\mathbf{x}$  is considered a RV, and the set of RVs at locations  $z(\mathbf{x}_i) i = 1, \dots, n$  is a RF, or more precisely, since  $n$  is finite, a random *vector*. Observations of population densities are termed regionalized variables (ReVs), and they are treated as stochastic realizations of an underlying RF.

In geostatistics, it is common to fit a stationary RF model (Journel and Huijbregts 1978; Chilès and Delfiner 1999). First, a spatially stationary (i.e., constant) mean parameter is usually defined. However, various alternatives are common in which the mean is allowed to vary across space (Goovaerts 1997) including those RF models that include a spatially varying “trend” (e.g., a two-dimensional polynomial or some spatially varying function of covariates; Hudson and Wackernagel 1994). A second parameter (more strictly a model comprising several parameters) that is usually defined is the stationary spatial covariance function (which implies second-order stationarity; see Journel and Huijbregts 1978) or variogram (defined further below, which implies intrinsic stationarity), either of which can be used to represent the “character” of spatial variation. Since the variogram function is more widely applicable than the spatial covariance function, we will refer to it here and in the remainder of the text. The mean and

variogram are, thus, the parameters that define the RF, and, thus, it is these parameters that need to be estimated to provide a useful model for geostatistical inference (e.g., prediction of the value of the property of interest at some unobserved location).

One might be wondering what utility the variogram brings and why stationarity is required as a modeling decision. We offer the following lay person explanation. In geostatistics, the central modeling decision is that the character of spatial variation can be treated as “stationary” (i.e., independent of location). This decision to fit a RF model that is parameterized by a stationary function representing the character of spatial variation is crucial to geostatistical inference. It means that it is possible to (i) lump together the observed spatial variation from different pairs of locations separated by specific distances, and possibly directions (this is essentially what the variogram does); (ii) use that information to infer the character of spatial variation at new pairs of locations; and (iii) crucially, given an observation representing one of a new pair of locations, predict at the other unobserved location based on the imputed character of spatial variation and, thus, how related we expect the two locations to be.

We now show (i) how the RF model can be estimated through calculation of the empirical or experimental variogram and the fitting of mathematical models with parameters and (ii) how the parameterized RF model can be used for spatial prediction through a technique known as Kriging and in other geostatistical operations.

---

## 74.2 Characterizing Spatial Variation

Most geostatistical analyses proceed with a summary of the properties of the data, and this is generally regarded as good practice. As in any statistical analysis, summary statistics and the histogram may suggest that the data should be transformed in some way (e.g., taking a logarithm so that the transformed distribution is approximately normal).

The spatial structure of a variable can be characterized with several structure functions, most commonly the variogram (outlined below).

### 74.2.1 The Experimental Variogram

The variogram cloud relates (half the) the squared differences (i.e., the semivariances) between paired data values to the distance (and possibly direction) by which they are separated, and it, thus, provides a summary of how different observations are as a function of distance.

The variogram (or, more properly, semivariogram) collapses the information contained in the variogram cloud; the semivariances within distance bands are averaged, and so, for example, there is an average semivariance for data pairs separated by 0–10 m, 10–20 m, and so on. The plot of average semivariance against

distance band indicates how spatially dependent the variable is. In practice, many properties tend to become more dissimilar as the separation distance increases.

The experimental variogram,  $\hat{\gamma}(\mathbf{h})$ , is estimated from  $p(\mathbf{h})$  paired observations,  $z(\mathbf{x}_\alpha), z(\mathbf{x}_\alpha + \mathbf{h}), \alpha = 1, 2, \dots, p(\mathbf{h})$  with:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2p(\mathbf{h})} \sum_{\alpha=1}^{p(\mathbf{h})} \{z(\mathbf{x}_\alpha) - z(\mathbf{x}_\alpha + \mathbf{h})\}^2 \quad (74.1)$$

Semivariances can be computed within particular directional tolerances. For example, only data pairs which are aligned approximately north–south could be included in Eq. (74.1). In this way, it is possible to assess how spatial variation differs as a function of direction.

Where the variable of interest is a rate calculated from a numerator and a population-based denominator (e.g., as for disease rate), the standard experimental variogram can be modified easily to account for spatially varying underlying populations. That is, greater weight can be given to zones with large populations, and thus, it is possible to account for spatially varying uncertainties in rates (see Goovaerts 2005; Goovaerts et al. 2005).

## 74.2.2 Variogram Model Fitting

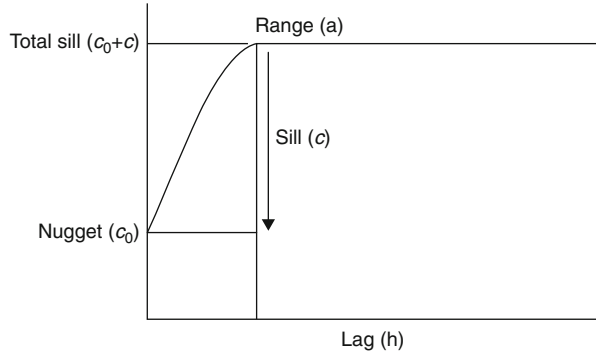
The variogram may be used as a means in itself to analyze the spatial structure of a variable (e.g., Berberoglu et al. 2000; Lloyd et al. 2004). More commonly, it is the first stage in Kriging-based spatial interpolation, as described below. A model can be fitted to the variogram, for example, by weighted least squares (Cressie 1985; McBratney and Webster 1986). Then the coefficients of the fitted model can be used as an input to the Kriging process. In principle, the variogram model provides information on how much weight is given to observations as a function of their distance from prediction locations. This use of information on spatial structure makes Kriging distinct from other commonly applied methods of interpolation such as inverse distance weighting, whereby weights are a simple function of distance and no account is taken of the specific spatial structure of the data being analyzed. The most commonly used models are taken from a set of “permissible” or “authorized” models.

Some of the most commonly used authorized models are detailed below. The nugget effect model, which indicates measurement error and microscale variation (variation at a distance smaller than the sample spacing), is given by:

$$\gamma(h) = \begin{cases} 0 & \text{for } h = 0 \\ c_0 & \text{for } |h| > 0 \end{cases} \quad (74.2)$$

Three of the most frequently used bounded models are the spherical model, the exponential model, and the Gaussian model, and these are each defined below. Each

**Fig. 74.1** Bounded variogram model: nugget effect and spherical model



of these models is “bounded” in that each has a maximum value of semivariance defined by a “sill” parameter. The exponential model is given by:

$$\gamma(h) = c \left[ 1 - \exp\left(-\frac{h}{d}\right) \right] \tag{74.3}$$

where  $c$  is the sill of the exponential model and  $d$  is the non-linear distance parameter. The exponential model reaches the sill asymptotically, and the practical range is  $3d$  (i.e., the separation at which approximately 95 % of the sill is reached).

The spherical model is a very widely used variogram model, and its form corresponds well with what is often observed in real-world studies, with an almost linear growth in semivariance to a particular separation and then stabilization (Armstrong 1998). It is given by:

$$\gamma(h) = \begin{cases} c \left[ 1.5 \frac{h}{a} - 0.5 \left(\frac{h}{a}\right)^3 \right] & \text{if } h \leq a \\ c & \text{if } h > a \end{cases} \tag{74.4}$$

where  $a$  is the non-linear parameter, known as the range.

The Gaussian model is given by:

$$\gamma(h) = c \left[ 1 - \exp\left(-\frac{h^2}{d^2}\right) \right] \tag{74.5}$$

As for the exponential model, the Gaussian model does not reach a sill at a finite distance, but rather approaches a defined sill asymptotically. The practical range is  $a \times \sqrt{3}$  (Journel and Huijbregts 1978). Variograms with parabolic behavior at the origin, as represented by the Gaussian model here, are indicative of very regular spatial variation (Journel and Huijbregts 1978). Authorized models can be used in positive linear combination where a single model is insufficient to represent well the form of the variogram (Webster and McBratney 1989). Figure 74.1 depicts a variogram model comprising the combination of a nugget effect and a spherical component. In practice, the combination of the nugget effect and one or more permissible models is common.

Where the experimental variogram does not reach a maximum, it may be desirable to fit a trend model to the data and model the residuals with a bounded model. In some circumstances, it may be desirable to select an unbounded variogram model. The most widely used unbounded model is the power model:

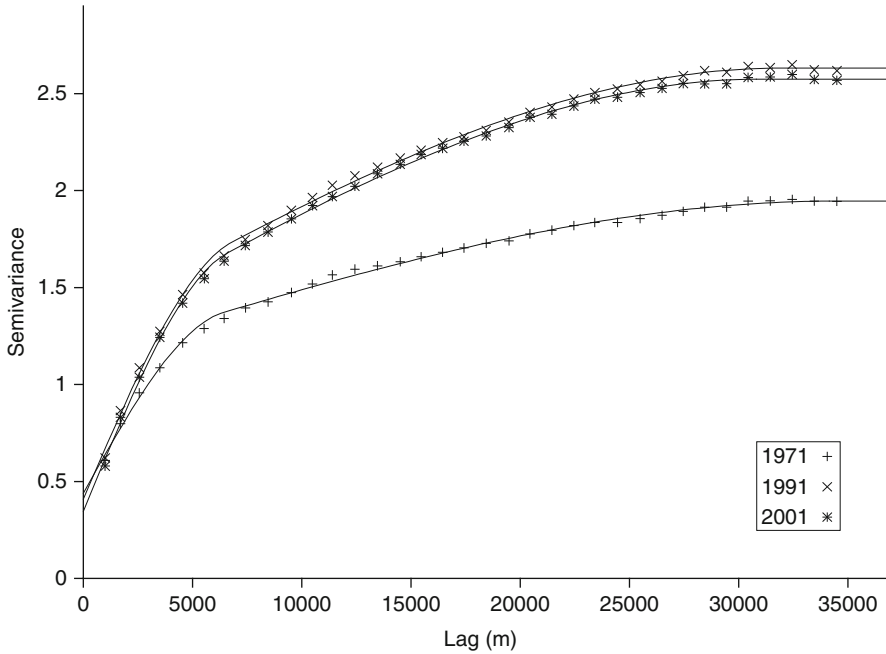
$$\gamma(h) = mh^\omega \quad (74.6)$$

where  $\omega$  is a power  $0 < \omega < 2$  with a positive slope,  $m$  (Deutsch and Journel 1998); the linear model is a special case of the power model.

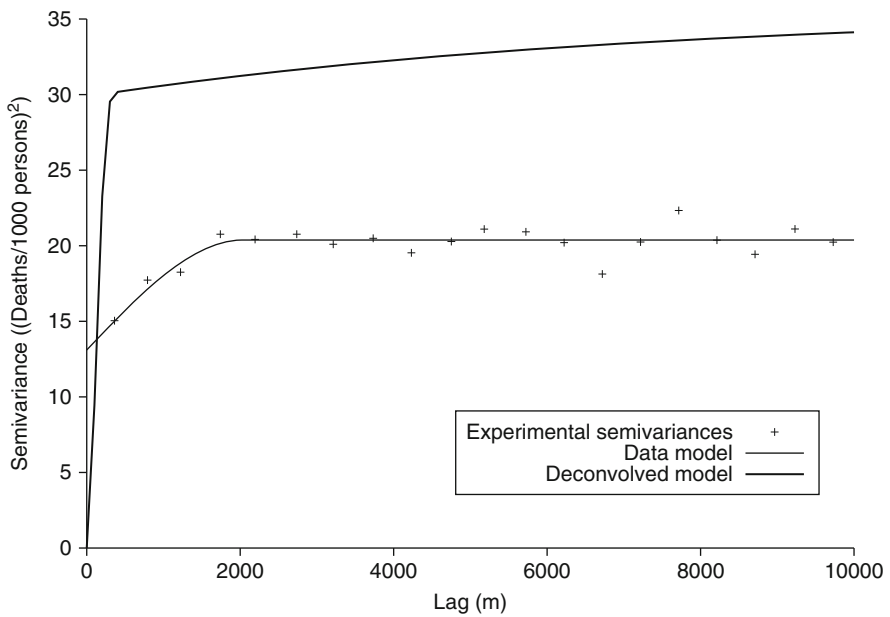
### 74.2.3 Example

The variogram characterizes the spatial structure in a variable. Where the variable represents some characteristic of a population, the variogram captures information on the dominant spatial scales of variation in that population property. Figure 74.3 shows variograms estimated from log-ratio transformed data on the percentage of Catholics in Northern Ireland in 1971, 1991, and 2001 for 1-km-square cells. The data are from the Census of Population, and the log ratios are computed with  $1/\sqrt{2} \times \ln(\text{Catholics by religion (\%)/non-Catholics by religion (\%)})$ . The percentages were computed with  $n_1 + 1$  and  $n_2 + 1$ , where  $n_1$  is the number of Catholics and  $n_2$  is the number of non-Catholics for each 1-km cell; this avoids logging zeros and reflects uncertainties in small counts. The data and transformation rationale are described in Lloyd (2010). Briefly, where the variables used sum to a constant (e.g., percentages sum to 100), use of raw values may be problematic, and Aitchison (1986) suggests that use of log ratios is a suitable approach; most applications of such approaches are with respect to analysis of compositions with multiple parts. The present analysis is based on only one variable (religion or community background) with only two groups (Catholic or non-Catholic), which can, thus, be expressed as a single ratio. However, Filzmoser et al. (2009) show that univariate statistical methods should not be applied directly to (raw) compositional data. Furthermore, Pawlowsky and Burger (1992) argue that structural analysis should not be conducted using raw proportions or percentages as the constant-sum constraint may lead to a distorted picture of the spatial covariance structure and possibly erroneous interpretations.

The variograms in Fig. 74.2 were fitted with a nugget effect and two spherical components. This combination of models reflects nested spatial structures. That is, the variogram models exhibit two breaks corresponding to distances of approximately 5 and 30 km, and these spatial scales are represented by the two model ranges. The ranges and the nugget effects are similar for each of the 3 census years, while the total sill, which represents the magnitude of variation, for 1971 is much smaller than the sills for 1991 and 2001. These results suggest that the major scales of variation did not change between 1971 and 2001 – the areas which were dominated by Catholics or Protestants were broadly similar. What did change was the magnitude of differences between places. In support of previous studies



**Fig. 74.2** Variogram of religion log ratios for 1971, 1991, and 2001 in Northern Ireland



**Fig. 74.3** Variogram of deaths/1,000 persons for SOAs with fitted model and deconvolved model

of residential segregation in Northern Ireland, the variograms suggest that the concentrations of Catholics and Protestants increased between 1971 and 1991, but there was little change (at the Northern Ireland (NI) scale at least) between 1991 and 2001. In other words, in 1991 Catholic areas were more Catholic than they had been in 1971, while Protestant areas had also become more Protestant. Lloyd (2012) discussed these results in more depth and also estimates variograms locally to enable assessment of how population spatial structures vary across NI. The derivation of local variograms is discussed by Lloyd (2011).

## 74.3 Spatial Interpolation with Kriging

### 74.3.1 Simple and Ordinary Kriging

Ordinary Kriging (OK) predictions are weighted averages of the  $n$  locally available data. The OK weights define the best linear unbiased predictor (BLUP). The OK prediction,  $\hat{z}_{OK}(\mathbf{x}_0)$ , is defined as:

$$\hat{z}_{OK}(\mathbf{x}_0) = \sum_{\alpha=1}^n \lambda_{\alpha}^{OK} z(\mathbf{x}_{\alpha}) \quad (74.7)$$

with the constraint that the weights,  $\lambda_{\alpha}^{OK}$ , sum to one to ensure an unbiased prediction:

$$\sum_{\alpha=1}^n \lambda_{\alpha}^{OK} = 1 \quad (74.8)$$

Using the Kriging system, appropriate weights can be determined, and these are multiplied by the available observations before summing the products to obtain the predicted value. These weights are derived given the coefficients of a model fitted to the variogram (or another function such as the covariance function) (Oliver 2010).

The Kriging prediction error must have an expected value of zero:

$$E\{\hat{Z}_{OK}(\mathbf{x}_0) - Z(\mathbf{x}_0)\} = 0 \quad (74.9)$$

The Kriging (or prediction) variance,  $\sigma_{OK}^2$ , is expressed as:

$$\begin{aligned} \hat{\sigma}_{OK}^2(\mathbf{x}_0) &= E\{[\hat{Z}_{OK}(\mathbf{x}_0) - Z(\mathbf{x}_0)]^2\} \\ &= 2 \sum_{\alpha=1}^n \lambda_{\alpha}^{OK} \gamma(\mathbf{x}_{\alpha} - \mathbf{x}_0) - \sum_{\alpha=1}^n \sum_{\beta=1}^n \lambda_{\alpha}^{OK} \lambda_{\beta}^{OK} \gamma(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}) \end{aligned} \quad (74.10)$$

That is, we seek the values of  $\lambda_1, \dots, \lambda_n$  (the weights) that minimize this expression with the constraint that the weights sum to one [Eq. (74.8)].



This minimization is achieved through Lagrange multipliers. The conditions for the minimization are given by the OK system comprising  $n + 1$  equations and  $n + 1$  unknowns:

$$\begin{cases} \sum_{\beta=1}^n \lambda_{\beta}^{\text{OK}} \gamma(\mathbf{x}_{\alpha} - \mathbf{x}_{\beta}) + \psi_{\text{OK}} = \gamma(\mathbf{x}_{\alpha} - \mathbf{x}_0) \alpha = 1, \dots, n \\ \sum_{\beta=1}^n \lambda_{\beta}^{\text{OK}} = 1 \end{cases} \quad (74.11)$$

where  $\psi_{\text{OK}}$  is a Lagrange multiplier. Given  $\psi_{\text{OK}}$ , the prediction variance of OK can be derived with:

$$\hat{\sigma}_{\text{OK}}^2 = \sum_{\alpha=1}^n \lambda_{\alpha}^{\text{OK}} \gamma(\mathbf{x}_{\alpha} - \mathbf{x}_0) + \psi_{\text{OK}} \quad (74.12)$$

The Kriging variance estimates the prediction variance (i.e., the square of the prediction error) based on the linear algebra given above. It is, thus, a measure of confidence in predictions and is a function of the form of the variogram, the sample configuration, and the sample support (Journel and Huijbregts 1978). The Kriging variance is, however, not conditional on the data values locally, and this has led some researchers to use alternative approaches such as conditional simulation (discussed in the next section) to build models of “spatial” uncertainty (Goovaerts 1997).

There are two standard varieties of OK: punctual OK and block OK. With punctual OK the predictions cover the same area or volume (the support,  $\mathbf{v}$ ; see Atkinson and Tate 2000) as the observations. For example, if observations of some property of a population are made through a grid-based census with a grid cell size of 1 km, as in NI, then predictions made from those data will also have a support of (i.e., represent) 1-km cells. In block OK, the predictions are made to a larger support than the observations. With punctual OK the original observation data are honored. That is, they are retained in the output map. Block OK predictions are averages over areas (i.e., the support has increased). Thus, at  $\mathbf{x}_0$  the prediction is not the same as an observation and does not need to honor it. In regional science contexts, data are rarely available on point supports, and often a concern may be how to predict from areas to points, a topic discussed below.

The choice of variogram model affects the Kriging weights and, therefore, the predictions. However, if the form of two models is similar at the origin of the variogram, then the two sets of results may be similar (Armstrong 1998). The choice of nugget effect may have marked implications for both the predictions and the Kriging variance. As the nugget effect is increased, the predictions become closer to the global average (Isaaks and Srivastava 1989).

Poisson Kriging provides an appropriate means to interpolate rare events such as mortality. The Poisson Kriging system incorporates the population-weighted mean of the rates (Goovaerts 2005) which accounts for the variability due to population

size, with larger weights where the population size is larger and, therefore, where the data may be considered more reliable.

Other forms of Kriging include Kriging with a trend model (where large-scale trends in the data are explicitly taken into account), cokriging (where information on secondary variables is used in the prediction process) (Wackernagel 2003), and indicator Kriging (where the data are transformed into a set of thresholds and the probability of exceeding thresholds is estimated) (Goovaerts, 1997).

### 74.3.2 Simulation

Kriging predictions are weighted moving averages, and thus, Kriging is a smoothing interpolator. In block Kriging, some amount of smoothing happens naturally through averaging over the larger support, but some occur due to the interpolation process itself, which effectively extends the support out to include the observations. The latter smoothing is an unwanted artifact of the interpolation process. A means of overcoming this unwanted smoothing is conditional simulation (Journel 1996; Goovaerts 1997; Dungan 1999). With conditional simulation, predictions are drawn from equally probable joint realizations of the RVs which comprise a RF model (Deutsch and Journel 1998). The simulated values are not the expected values (as in Kriging), but rather they are drawn from the conditional cumulative distribution function (ccdf), as a function of the available data, including both the observations and previously simulated data, and the (modeled) spatial variation. Since the approach is stochastic, multiple simulated realizations can be obtained. The range of values obtained represents the “spatial” uncertainty.

---

## 74.4 The Change of Support Problem

In regional science contexts, data are often available for zones rather than points. While individual or household level data are available in some contexts, these are usually provided without detailed spatial information, and spatially aggregated data usually offer the only means of exploring detailed spatial patterns. The data support,  $v$ , is defined as the geometrical size, shape, and orientation of the units associated with the measurements (Atkinson and Tate 2000). Thus, making predictions from areas to points corresponds to a change of support. Geostatistics offers the means to (i) explore how the spatial structure of a variable changes with change of support and (ii) change the support by interpolation to an alternative zonal system or to a quasi-point support (Schabenberger and Gotway 2005).

### 74.4.1 Regularization

For many applications, the variogram defined on a point support is not available. Indeed, it is physically impossible to measure on a point support given that all

measurements are integrals over a positive finite support. Thus, only values over a positive support (area) may be available. The variogram of such aggregated data is termed the regularized or areal variogram (Goovaerts 2008).

### 74.4.2 Variogram Deconvolution

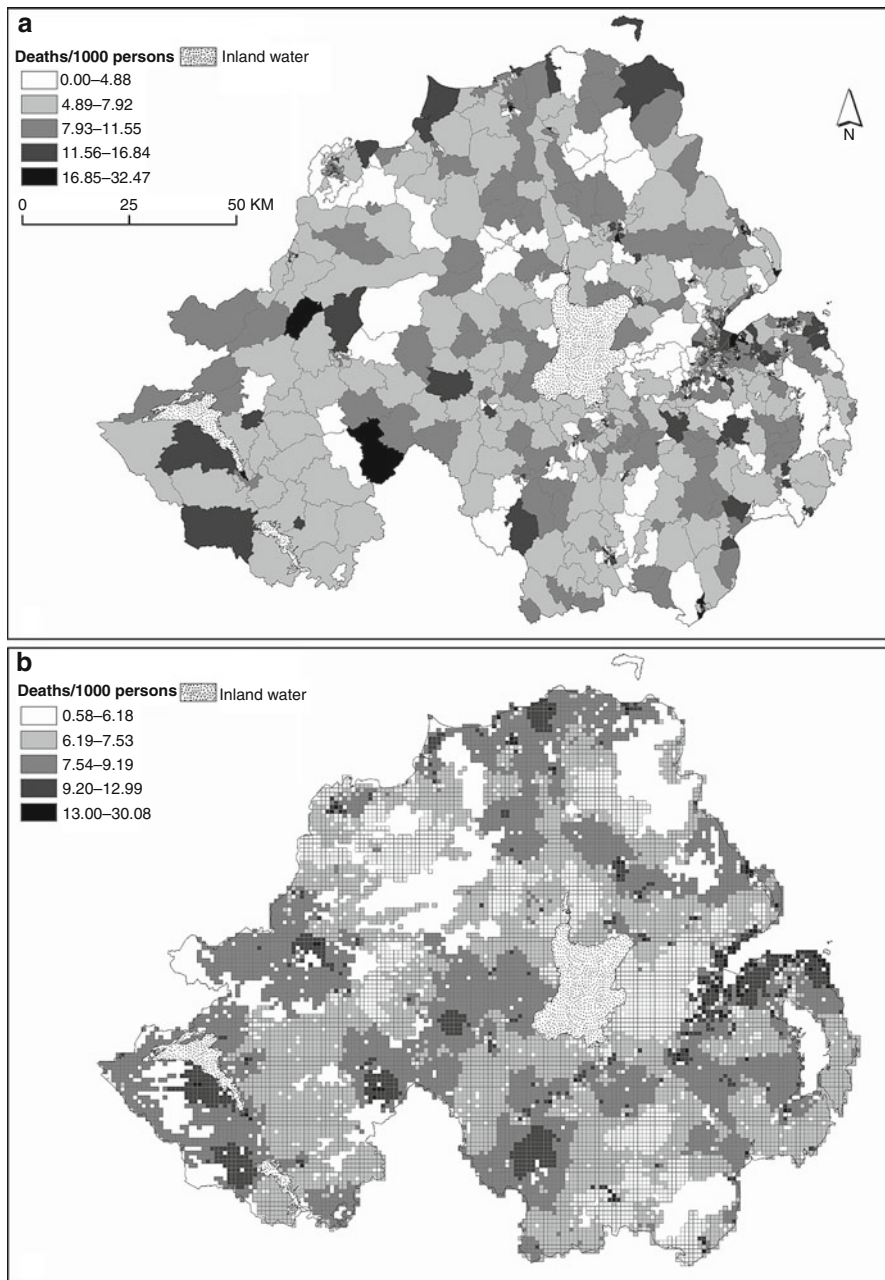
Theoretically, given the point support variogram, it is possible to estimate the variogram for any support. Through variogram deconvolution, the point support variogram can be estimated from the variogram estimated from areal data. Atkinson and Curran (1995) derived the point support variogram from the areal support variogram, with the areas defined as regular grids, following an iterative procedure. Of greater relevance for many regional science applications is variogram deconvolution for irregular supports, that is, those cases where the data are available on irregular zones (such as census or administrative zones) rather than regular cells. Goovaerts (2008) presents a procedure for variogram deconvolution given irregular zones, and this method is implemented in the STIS software (see <http://www.biomedware.com/>).

The above deconvolution method is illustrated here through an example. The case study makes use of data on deaths per 1,000 of the population in Northern Ireland (NI). The death data are counts for 2008, and the denominator is given as the midyear estimates of total population in 2008 for super output areas (SOAs). Figure 74.3 gives the variogram estimated from deaths/1,000 persons over SOAs. The deconvolved model, derived using the method of Goovaerts (2008), is also given. The application of the deconvolved model for Kriging is illustrated below.

### 74.4.3 Area-to-Point Kriging

Given the deconvolution procedures outlined above, it is possible to make predictions at point locations given data defined on areal supports. Kyriakidis (2004) and Goovaerts (2008) show how the Kriging system is adapted in the case of areal data supports and point prediction locations.

Area-to-point Kriging is illustrated given the death rate data and the deconvolved variogram detailed in the previous section. Ordinary Kriging with Poisson population (2008 midyear estimates for each SOA) adjustment was applied with a population denominator of 1,000 (i.e., the deaths are rates per 1,000 of the population). The discretization geography was 1-km cells populated in 2001 (from 2001 Census of Population), with numbers of persons per cell (2001 Census counts) as weights. The destination geography (locations where estimates are required) was 1-km cells (as for the discretization geography). Figure 74.4 shows deaths/1,000 persons (A) for SOAs and (B) derived using area-to-point Kriging at 1-km cells. In areas covered by large SOAs (with lower density populations), the increased detail is particularly apparent.



**Fig. 74.4** Deaths/1,000 persons (a) for SOAs and (b) derived using area-to-point Kriging at 1-km cells (only populated cells shown). Source: 2001 Census, Output Area Boundaries. Crown copyright 2003

## 74.5 Geostatistics and Regional Science

Geostatistics originated in the mining industry through the work of Danie Krige in the 1950s and particularly through Georges Matheron who developed the Theory of Regionalized Variables in the 1960s and 1970s. From the original mining geology application, domain geostatistics found further application in the dominant field of petroleum geology, particularly through the Stanford Center for Reservoir Forecasting (SCRF) led by Andre Journel, but also in hydrogeology led by researchers such as Jaime Gómez-Hernández, in soil science led by the work of Richard Webster (see Webster and Oliver 2000) and Alex McBratney and subsequently Pierre Goovaerts, and in remote sensing and environmental science more generally. Thus, most applications of geostatistics can be divided into mining, petroleum, and “environmental”, and indeed, many conferences such as GeoENV organized themselves in this way for many years. However, the potential of geostatistics in the social sciences has been highlighted in recent research.

Much interest has been generated in the use of geostatistics to study a range of population-based properties, including rates based on census attributes as exemplified in this chapter and disease outcomes. This has been enabled by two key developments: (i) the pioneering work of Pascal Monestiez and Pierre Goovaerts on Poisson Kriging (Goovaerts 2005; Monestiez 2006), which enables the prediction of counts (e.g., mortality or morbidity) given an underlying population (e.g., at-risk population), and (ii) the parallel development in statistics of what has become known as model-based geostatistics. For example, in relation to the former, population-weighted variograms enable robust estimation in cases where the population underlying the observations is variable in magnitude. Moreover, Poisson Kriging allows the spatial interpolation of variables representing “rare” events, such as mortality.

Model-based geostatistics deserves special mention here as it represents an alternative Bayesian approach to geostatistics that has been made popular by statisticians such as Peter Diggle and is favored by the epidemiological community. Model-based geostatistics differs from the classical approach presented in this chapter in that the parameters of the RF model are regarded as uncertain, whereas in the classical approach these parameters are estimated through variogram model fitting and then regarded as “fixed” for the purposes of Kriging. As a result, model-based geostatistics conveys naturally the uncertainty in parameter estimation through to the prediction, resulting in a more complete description of prediction uncertainty. The general approach advocated by Diggle and others is two staged: (i) regression is used to predict the variable of interest from covariates, and (ii) geostatistics is used to predict the residuals at unobserved locations based on spatial dependence in those residuals. Since model-based geostatistics naturally allows for the adoption of a range of regression link functions (including the Poisson and logistic functions), the framework is general and can be applied readily to the population-based prediction problems described above. Thus, the development of new approaches, or adaptations of existing methods, has made the potential range of applications of geostatistical methods in regional science much wider.

The utility of area-to-point Kriging is much greater than may at first seem the case. Censuses are the primary means of enumerating the populations of the world's nations. However, the majority of them use arbitrary areal units to convey the original data, which were measured at the individual level, to the public (e.g., for reasons of preventing disclosure of personal information). Consequently, key variables relating to the world's population are obscured by (i) the aggregation effect which is effectively the convolution discussed above and (ii) the zonation effect which means that alternative realizations of the set of arbitrary boundaries may lead to very different realizations. This is the well-known modifiable areal unit problem (MAUP) which has dogged census analysis for decades. Area-to-point Kriging tackles the MAUP head on allowing researchers to bypass the sampling frame imposed by a particular set of areal units and re-present the census data on any support. While there are obvious limits to what can be achieved (it is not possible to generate more information than one started with), the results are visually stunning and do represent an optimal linear solution to the problem (Kyriakidis 2004; Goovaerts 2008). Additionally, applications that require common supports for multiple time periods (e.g., for which zonal systems may differ) may also benefit from recent developments in variogram deconvolution and area-to-point Kriging.

---

## 74.6 Conclusions

This chapter has introduced classical geostatistics and emphasized the advances that have been made in recent years that have opened up the possibility of applying geostatistics in social science and regional science. The ease of access to software which implements modern geostatistical methods means that researchers are increasingly likely to be held back only by their knowledge, rather than appropriate tools. It is hoped that this chapter will provide a useful starting point for those who wish to explore geostatistical methods in the context of regional science.

**Acknowledgments** The authors thank the anonymous referees for their comments and thank the editors for their patience while this chapter was finalized. The Northern Ireland Statistics and Research Agency are thanked for access to data. Census output is Crown copyright and is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland. Northern Ireland Statistics and Research Agency, 2001 Census: Standard Area Statistics (Northern Ireland) [computer file]. ESRC/JISC Census Programme, Census Dissemination Unit, Mimas (University of Manchester).

---

## References

- Aitchison J (1986) *The statistical analysis of compositional data*. Chapman and Hall, London
- Armstrong M (1998) *Basic linear geostatistics*. Springer, Berlin
- Atkinson PM, Curran PJ (1995) Defining an optimal size of support for remote sensing investigations. *IEEE Trans Geosci Remote Sens* 33(3):768–776

- Atkinson PM, Tate NJ (2000) Spatial scale problems and geostatistical solutions: a review. *Prof Geogr* 52(4):607–623
- Berberoglu S, Lloyd CD, Atkinson PM, Curran PJ (2000) The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean. *Comput Geosci* 26(4):385–396
- Chilès J-P, Delfiner P (1999) *Geostatistics: modeling uncertainty*. Wiley, New York
- Cressie NAC (1985) Fitting variogram models by weighted least squares. *Math Geol* 17(5):563–586
- Deutsch CV, Journel AG (1998) *GSLIB: geostatistical software and user's guide*, 2nd edn. Oxford University Press, New York
- Dungan JL (1999) Conditional simulation. In: Stein A, van der Meer F, Gorte B (eds) *Spatial statistics for remote sensing*. Kluwer, Dordrecht, pp 135–152
- Filzmoser P, Hron K, Reimann C (2009) Univariate statistical analysis of environmental (compositional) data: problems and possibilities. *Sci Total Environ* 407(23):6100–6108
- Goovaerts P (1997) *Geostatistics for natural resources evaluation*. Oxford University Press, New York
- Goovaerts P (2005) Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging. *Int J Health Geogr* 4:31pp
- Goovaerts P (2008) Kriging and semivariogram deconvolution in the presence of irregular geographical units. *Math Geosci* 40(1):101–128
- Goovaerts P, Jacquez GM, Greiling D (2005) Exploring scale-dependent correlations between cancer mortality rates using factorial kriging and population-weighted semivariograms. *Geogr Anal* 37(2):152–182
- Haining RP, Kerry R, Oliver MA (2010) Geography, spatial data analysis, and geostatistics: an overview. *Geogr Anal* 42(1):7–31
- Hudson G, Wackernagel H (1994) Mapping temperature using kriging with external drift: theory and an example from Scotland. *Int J Climatol* 14(1):77–91
- Isaaks EH, Srivastava RM (1989) *An introduction to applied geostatistics*. Oxford University Press, New York
- Journel AG (1996) Modelling uncertainty and spatial dependence: stochastic imaging. *Int J Geogr Inf Syst* 10(5):517–522
- Journel AG, Huijbregts CJ (1978) *Mining geostatistics*. Academic, London
- Kyriakidis PC (2004) A geostatistical framework for area-to-point spatial interpolation. *Geogr Anal* 36(3):259–289
- Lloyd CD (2010) Exploring population spatial concentrations in Northern Ireland by community background and other characteristics: an application of geographically weighted spatial statistics. *Int J Geogr Inf Sci* 24(8):1193–1221
- Lloyd CD (2011) *Local models for spatial analysis*, 2nd edn. CRC Press, Boca Raton
- Lloyd CD (2012) Analysing the spatial scale of population concentrations by religion in Northern Ireland using global and local variograms. *Int J Geogr Inf Sci* 26(1):57–73
- Lloyd CD, Berberoglu S, Curran PJ, Atkinson PM (2004) A comparison of texture measures for the per-field classification of Mediterranean land cover. *Int J Remote Sens* 25(19):3943–3965
- McBratney AB, Webster R (1986) Choosing functions for semi-variograms of soil properties and fitting them to sampling estimates. *J Soil Sci* 37(4):617–639
- Monestiez P, Dubroca L, Bonnin E, Durbec J-P, Guinet C (2006) Geostatistical modelling of spatial distribution of *Balaenoptera physalus* in the Northwestern Mediterranean Sea from sparse count data and heterogeneous observation efforts. *Ecol Model* 193(3–4):615–628
- Oliver MA (2010) The variogram and kriging. In: Fischer MM, Getis A (eds) *Handbook of applied spatial analysis. Software tools, methods and applications*. Springer, Berlin/Heidelberg/New York, pp 319–352
- Pawlowsky V, Burger H (1992) Spatial structure analysis of regionalized compositions. *Math Geol* 24(6):675–691

- 
- Schabenberger O, Gotway CA (2005) *Statistical methods for spatial data analysis*. Chapman and Hall/CRC, Boca Raton
- Wackernagel H (2003) *Multivariate geostatistics. An introduction with applications*, 3rd edn. Springer, Berlin
- Webster R, McBratney AB (1989) On the Akaike information criterion for choosing models for variograms of soil properties. *J Soil Sci* 40(3):493–496
- Webster R, Oliver MA (2000) *Geostatistics for environmental scientists*. Wiley, Chichester