GEOSCIENCE AUSTRALIA

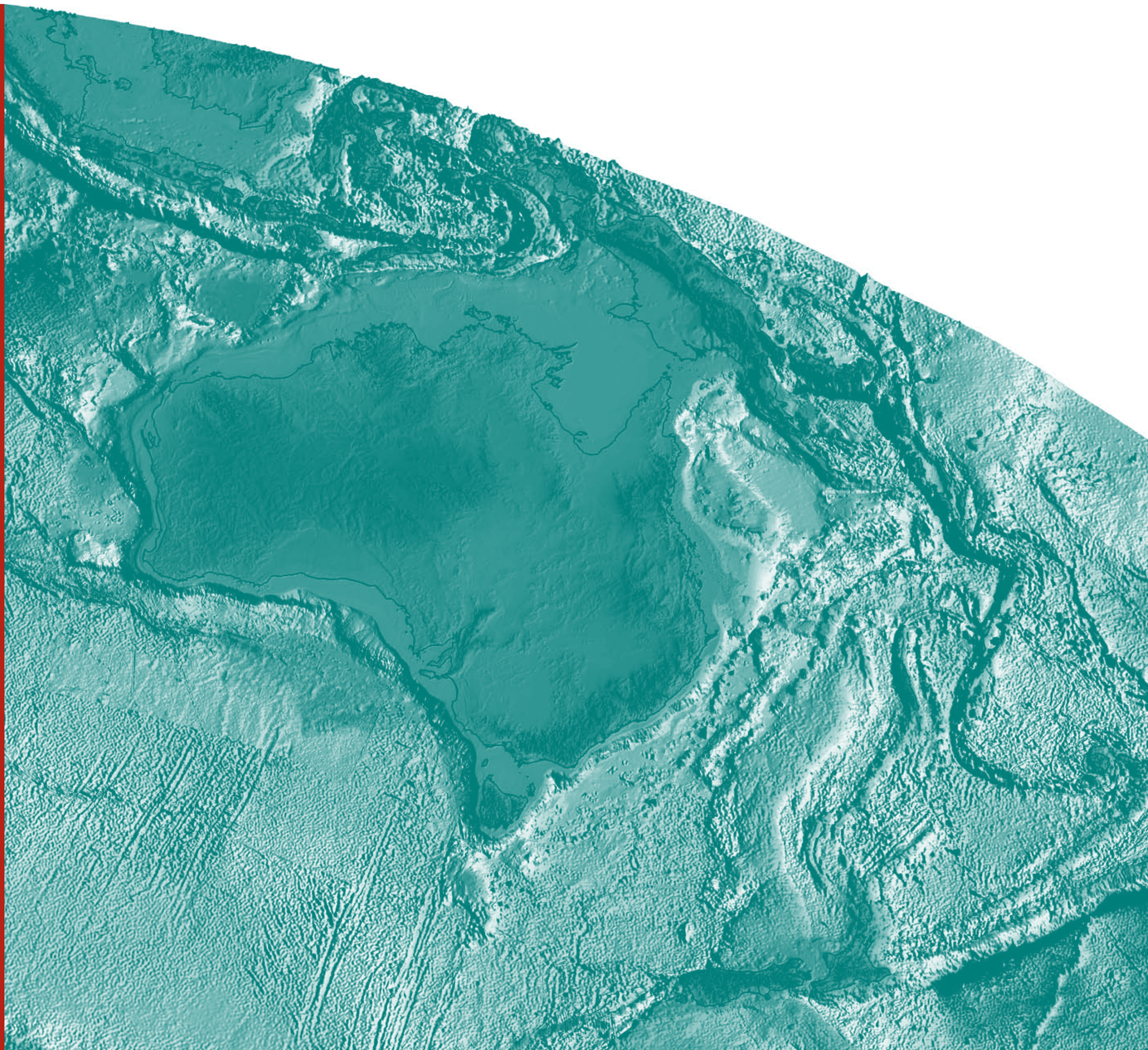
Australian Government
Geoscience Australia

A Review of Spatial Interpolation Methods for Environmental Scientists

Jin Li and Andrew D. Heap

Record
2008/23

GeoCat #
68229



APPLYING GEOSCIENCE TO AUSTRALIA'S MOST IMPORTANT CHALLENGES

Geoscience Australia Record 2008/23

A Review of Spatial Interpolation Methods for Environmental Scientists

Jin Li and Andrew D. Heap

Geoscience Australia, GPO Box 378, Canberra, ACT 2601, Australia

Australian Government
Geoscience Australia

Department of Resources, Energy and Tourism
Minister for Resources and Energy: The Hon. Martin Ferguson, AM MP
Secretary: Dr Peter Boxall, AO

Geoscience Australia
Chief Executive Officer: Dr Neil Williams, PSM

ISSN 1448-2177
ISBN 978 1 921498 30 5 (hardcopy)
ISBN 978 1 921498 29 9 (CD-ROM)
ISBN 978 1 921498 28 2 (webcopy)

GeoCat # 68229

Bibliographic reference: Li, J. and Heap, A.D., 2008. *A Review of Spatial Interpolation Methods for Environmental Scientists*. Geoscience Australia, Record 2008/23, 137 pp.

Correspondence for feedback:

Sales Centre
Geoscience Australia
GPO Box 378
Canberra
ACT 2601
Australia

sales@ga.gov.au

Abstract

Spatial continuous data (spatial continuous surfaces) play a significant role in planning, risk assessment and decision making in environmental management. They are, however, usually not readily available and often difficult and expensive to acquire, especially for mountainous and deep marine regions. As geographic information systems (GIS) and modelling techniques are becoming powerful tools in natural resource management and biological conservation, spatial continuous data of environmental variables are increasingly required. Environmental managers often require spatial continuous data over a region of interest to make effective and informed decisions and scientists need accurate data which are well-distributed across a region to make justified interpretations. However, spatial distribution data of natural phenomena are often collected from point sources. The marine environment in Australia is a typical case, where: seabed mapping, habitat classification, and prediction of marine biodiversity, essential for marine biodiversity conservation, need reliable spatial continuous data of the marine environment. In most of the Australian marine region, such data are not available, and only sparsely and unevenly scattered point samples have been collected. Therefore, spatial interpolation techniques are essential for estimating biophysical variables for the unsampled locations.

The spatial interpolation methods, including geostatistics, have been developed for and applied to various disciplines. They are data-specific or even variable-specific. Many factors including sample size, sampling design and data properties affect the estimations of the methods. There are no consistent findings about how these factors affect the performance of the spatial interpolators. Therefore, it is difficult to select an appropriate spatial interpolation method for a given input dataset. This review aims to provide some guidelines and suggestions in relation to the application of the spatial interpolation methods to environmental data by comparing the features of the commonly applied spatial interpolators.

Over 40 spatial interpolation methods are briefly described. They fall into three categories, namely: non-geostatistical interpolators (12), geostatistical interpolators (22) and combined methods (8). Many sub-methods or types are also included, for example, 8 stratified methods and 6 types of regression kriging. A list of the methods that are not commonly used is also provided for those readers interested.

The spatial interpolation methods were developed either for specific disciplines or even for specific variables based on the data properties modelled. Each method has its specific assumptions and features. These features, such as global versus local, exact versus inexact, deterministic versus stochastic, and gradual versus abrupt, are discussed. In total, 26 spatial interpolation methods are compared theoretically and their features are summarised.

Commonly used assessment measures are summarised in relation to: 1) the performance of variogram models, 2) the performance of the spatial interpolation methods, and 3) the performance of a spatial interpolation method for datasets with different sample sizes. The criteria used to judge each measurement are also discussed. Two new measurements are proposed and a procedure is developed to compare the results of the performance of the spatial interpolation methods for different variables and from various disciplines.

The spatial interpolation methods have been applied to many disciplines. The focus of this review is mainly on comparative studies in environmental sciences. For each of the 51 comparative studies considered, the following information, including the methods compared, sampling design, sample size, area of region studied and the results, is summarised. The performance of 62 methods and sub-methods in the 51 comparative studies is compared. Four types of method groups are identified based on the frequency of their application.

Several factors that affect the performance of the spatial interpolation methods are discussed, including sampling design, sample spatial distribution, data quality, correlation between the primary and secondary variables, and interaction among various factors. The impacts of sample density, variation in the data, sampling design and stratification on the estimations of the spatial interpolation methods are quantified using data from 77 cases in 17 reviewed comparative studies. The results show that variation within the data is a dominant factor and has tremendous impacts on the performance of the spatial interpolators. As the variation increases, the accuracy of all methods decreases. Irregular-spaced sampling design and stratification would improve the accuracy of estimation. There is no evidence of the effects of sampling density on the performance of the spatial interpolation methods in this comprehensive comparative study.

A total of 26 commonly used spatial interpolation methods are then classified based on their features to provide an overview of relationships among these methods. These features are quantified and a cluster analysis is conducted to show similarities among these spatial interpolators. They are classified into 10 groups. To provide guidelines for potential users, a decision tree for selecting an appropriate method from the 26 spatial interpolation methods is developed according to the availability and nature of data.

Finally, a list of available software packages for spatial interpolation is provided. Some important factors for spatial interpolation in marine environmental science are discussed, and recommendations are made for applying spatial interpolation methods to marine environmental data in Geoscience Australia.

Abbreviations

AIC: Akaike information criteria
AK: Akima's interpolator
ANOVA: analysis of variance
ASE: averaged standard error
BIC: Bayesian information criteria
BK: block kriging
CART: regression tree
CCK: colocated cokriging
CK: cokriging
Cl: classification
CV: coefficient of variation
DEM: digital elevation model
DK: disjunctive kriging
DuK: dual kriging
EBLUP: empirical best linear unbiased predictor
EF: model efficiency
FK: factorial kriging
FS: Fourier series
GAM: generalised additive model
GCV: generalised cross validation
GIDS: gradient plus inverse distance squared
GIS: geographic information systems
GLM: generalised linear model
GM: global mean
GRNN: general regression neural network
GWR: geographically weighted regression
ICK: indicator cokriging
IDS: inverse distance squared
IDW: inverse distance weighting
IK: indicator kriging
IKED: IK with external drift
KED: kriging with an external drift
KT: kriging with a trend
KWS: kriging within strata
LM: linear regression model
LMM: linear mixed model

LR: lapse rate
LTS: local trend surfaces
MAE: mean absolute error
MBE: mean bias error
MBK: model-based kriging
MCMC: Markov chain Monte Carlo
ME: mean error
MFK: multivariate factorial kriging
MG: multiGaussian
MIK: median indicator kriging
MSE: mean squared error
MSE2: mean standardised error
MSRE: mean square reduced error
MWRCK: moving window regression residual cokriging
NaN: natural neighbours
NN: nearest neighbours
OCCK: ordinary colocated cokriging
OCK: ordinary CK
OCKWS: ordinary cokriging within strata
OIDW: optimal IDW
OIK: ordinary IK
OK: ordinary kriging
OKWS: ordinary KWS
PCA: principal component analysis
PCK: principal component kriging
PK: probability kriging
RBFN: radial basis function network
REML: residual maximum likelihood method
REML-EBLUP: residual maximum likelihood-empirical best linear unbiased predictor
RK: regression kriging
RMAE: relative mean absolute error
RMSE: root mean squared error
RMSSE: root mean square standardised error
RND: relative nugget difference
RRMSE: relative root mean square error
RVar: ratio of the variance of estimated values to the variance of the observed values
SCCK: simple colocated cokriging

SCK: simple cokriging
SCKWS: simple cokriging within strata
SIK: simple IK
SK: simple kriging
SKlm: SK with varying local means
SKWS: simple KWS
SMSE: standardised mean square error
SOCK: standardised OCK
StGM: stratified GM
StIDS: stratified IDS
StIDW: stratified IDW
StIDW-0: stratified moving average
StNN: stratified NN
StOCK: stratified OCK
StOK: stratified OK
StSK: stratified SK
StTPS: stratified TPS
TIN: triangular irregular network
TPS: thin plate splines or Laplacian smoothing splines
TSA: trend surface analysis
UK: universal kriging
UK-LD: UK with a linear drift
UK-QD: UK with quadratic drift

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Spatial continuous data (or spatial continuous surfaces) play a significant role in planning, risk assessment, and decision making in environmental management. They are, however, usually not always readily available and often difficult and expensive to acquire, especially for mountainous or deep marine regions. Environmental data collected on field surveys are typically from point sources. However, environmental managers often require spatial continuous data over the region of interest to make effective and confident decisions, and scientists need accurate spatial continuous data across a region to make justified interpretations.

As geographic information systems (GIS) and modelling techniques are becoming powerful tools in natural resource management and biological conservation, spatial continuous data of environmental variables are increasingly required (Collins and Bolstad, 1996; Hartkamp *et al*., 1999). Thus, the values of an attribute at unsampled points need to be estimated, meaning that spatial interpolation from point data to spatial continuous data is necessary. It is also necessary when 1) the discretized surface has a different level of resolution, cell size or orientation from that required; 2) a continuous surface is represented by a data model that is different from that required; and 3) the data we have do not cover the domain of interest completely (Burrough and McDonnell, 1998). In such instances, spatial interpolation methods provide tools to fulfil such task by estimating the values of an environmental variable at unsampled sites using data from point observations within the same region. Predicting the values of a variable at points outside the region covered by existing observations is called extrapolation (Burrough and McDonnell, 1998). In this study, extrapolation is regarded as part of interpolation because all spatial interpolation methods can be used to generate an extrapolation.

In Australia, point biophysical data from the marine environment are collected for seabed mapping and habitat classification purposes, where biophysical data are sparsely and unevenly distributed, principally because of the high costs and difficulties associated with collecting samples from many regions of the marine environment. Spatial interpolation and extrapolation of the point data are required for such purposes and also for the prediction of marine biodiversity, biological conservation and ecosystem management. To support conservation and management of Australia's marine biodiversity as part of the United Nations Convention on Biological Diversity (United Nations, 1993), Geoscience Australia has undertaken studies that also require the spatial interpolation of the point biophysical data.

The spatial interpolation methods, including geostatistics, have been developed for and applied in various disciplines (Zhou *et al*., 2007). They are proposed for specific data types or a specific variable. Many factors including sample size, sampling design, and the nature of the data affect the estimation of a spatial interpolator. There are no consistent findings about how these factors affect the performance of the spatial interpolators. Therefore, it is difficult to address the key issue in spatial interpolation that is how to select an appropriate spatial interpolation method for a given input dataset (Burrough and McDonnell, 1998).

This review aims to provide guidelines and suggestions useful to environmental scientists, especially in marine sciences, on the spatial interpolation of biophysical data by comparing the features of commonly applied spatial interpolators. This review covers several aspects of spatial interpolation, which are presented in eight chapters. Following this introduction, Chapter 2 contains brief descriptions of the commonly used spatial interpolation methods. Features of 26 spatial interpolation methods are discussed and theoretically compared in Chapter 3. Several measurements that are usually used to assess the performance of variogram models and the spatial interpolation methods are presented in Chapter 4; and also two new measurements are proposed to assess the performance of the spatial interpolation methods using the results from different disciplines and for different variables. Applications of the spatial interpolators in environmental sciences are briefly described and then the results from 51 comparative studies are compared in Chapter 5. Factors that affect the performance of the spatial interpolators are discussed and examined in Chapter 6. The 26 methods discussed in Chapter 3 are then classified and a decision tree for selecting an appropriate interpolator according to the nature of input dataset is developed in Chapter 7. Finally, in Chapter 8, a list of software packages for spatial interpolation is provided, several important issues in applying the spatial interpolators are discussed in the context of marine environmental sciences and some recommendations are made for spatial interpolation in marine environmental sciences.

Because this review is for environmental science researchers, jargon and mathematical and statistical formulas are avoided whenever possible. However, some mathematical and statistical nomenclature is provided to maintain a rigorous discussion of the methods. Explanations and/or definitions are provided for statistical terms, and equations are presented in a simplified, concise and easy-to-follow version. Relevant literature is also provided for further reference.

Three significant challenges were encountered in this review, namely: 1) sometimes the same method is presented with different names in different references; 2) mathematical symbols often change with references although they represent the same concept; and 3) methods are not described clearly in some references. Efforts have been made to match different names and symbols with the right methods and concepts and to assign the correct names to the methods used in various studies. In some cases, when it was impossible to find information on the method used in a reference, the reference is either discarded or a note is made to avoid confusion.

Chapter 2: Spatial Interpolation Methods

Numerous methods have been developed for spatial interpolation in various disciplines and there are a number of different terms used to distinguish them, including: "interpolating" and "non-interpolating" methods or "interpolators" and "non-interpolators" (Laslett *et al.*, 1987). In this review, all methods are referred to as spatial interpolation methods or spatial interpolators. The spatial interpolation methods covered in this review are only those commonly used or cited in environmental studies. As such, the list of the methods in this review is not an exhaustive one.

In this chapter, a total of 42 spatial interpolation methods are briefly described. They fall into three categories: 1) non-geostatistical methods, 2) geostatistical methods, and 3) combined methods (Table 2.1). In geostatistics, the methods that are capable of using secondary information (see section 2.1.5 for definition) are often referred to as "multivariate", while the methods that do not use the secondary information are called "univariate" methods. Here, it must be noted that multivariate usually refers to more than one response variable, despite of the fact that in some references it also refers to more than one explanatory variable (usually referred to as multiple variables). A brief introduction to geostatistics is provided prior to the descriptions of the geostatistical methods. The level of description of each method depends on the nature of the method. If it is relatively simple and straightforward, a full description of the method is provided. If it is relatively complex, then a brief description is provided and relevant publications for further reading are cited.

Estimations of nearly all spatial interpolation methods can be represented as weighted averages of sampled data. They all share the same general estimation formula, as follows:

$$\hat{z}(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i) \qquad (1)$$

where $\hat{z}$ is the estimated value of an attribute at the point of interest x_0, z is the observed value at the sampled point x_i, λ_i is the weight assigned to the sampled point, and n represents the number of sampled points used for the estimation (Webster and Oliver, 2001). The attribute is usually called the primary variable, especially in geostatistics.

Table 2.1. The spatial interpolation methods considered in this review.

Non-geostatistical	Geostatistical		Combined method
	Univariate	Multivariate	
Nearest neighbours	Simple kriging	Universal kriging	Classification combined other interpolation methods
Triangular irregular network related interpolations	Ordinary kriging	SK with varying local means	Trend surface analysis combined with kriging
Natural neighbours	Block kriging	Kriging with an external drift	Lapse rate combined with kriging
Inverse distance weighting	Factorial kriging	Simple cokriging	Linear mixed model
Regression models	Dual kriging	Ordinary cokriging	Regression trees combined with kriging
Trend surface analysis	Indicator kriging	Standardised OCK	Residual maximum likelihood-empirical best linear unbiased predictor
Splines and local trend surfaces	Disjunctive kriging	Principal component kriging	Regression kriging
Thin plate splines	Model-based kriging	Colocated cokriging	Gradient plus inverse distance squared
Classification	Simulation	Kriging within strata	
Regression tree		Multivariate factorial kriging	
Fourier series		Indicator kriging	
Lapse rate		Indicator cokriging	
		Probability kriging	
		Simulation	

2.1. Non-Geostatistical Interpolators

A total of 12 non-geostatistical interpolation methods are briefly described.

2.1.1. Nearest Neighbours

The nearest neighbours (NN) method predicts the value of an attribute at an unsampled point based on the value of the nearest sample by drawing perpendicular bisectors between sampled points (n), forming such as Thiessen (or Dirichlet/Voronoi) polygons (V_i, i=1,2,…, n). This produces one polygon per sample and the sample is located in the centre of the polygon, such that in each polygon all points are nearer to its enclosed sample point than to any other sample points (Isaaks and Srivastava, 1989; Ripley, 1981; Webster and Oliver, 2001). The estimations of the attribute at unsampled points within polygon V_i are the measured value at the nearest single sampled data point x_i that is $\hat{z}(x_0) = z(x_i)$. The weights are:

$$\lambda_i = \begin{cases} 1 & if \quad x_i \in V_i, \\ 0 & otherwise. \end{cases} \tag{2}$$

All points (or locations) within each polygon are assigned the same value (Ripley, 1981; Webster and Oliver, 2001). A number of algorithms exist to generate the polygons (Gold and Condal, 1995), including pycnophylactic interpolation (Burrough and McDonnell, 1998).

2.1.2. Triangular Irregular Network

The triangular irregular network (TIN) was developed by Peuker and co-workers (1978) for digital elevation modelling that avoids the redundancies of the altitude matrix in the grid system (Burrough and McDonnell, 1998). In TIN, all sampled points are joined into a series of triangles based on a Delauney's triangulation. Each triangle is empty so it does not contain any of the sampled points. TIN forms a different basis for making estimates in comparison with those used in NN. The value of a point within a triangle is estimated by linear or cubic polynomial interpolation (Ripley, 1981; R Development Core Team, 2007; Webster and Oliver, 2001). The advantages and disadvantages of TIN are discussed in Burrough and McDonnell (1998).

2.1.3. Natural Neighbours

The natural neighbours (NaN) method was introduced by Sibson (1981). It combines the best features of NN and TIN (Webster and Oliver, 2001). The first step is a triangulation of the data by Delauney's method, in which the apices of the triangles are the sample points in adjacent Thiessen polygons. This triangulation is unique except where the data are on a regular rectangular grid. To estimate the value of a point, it is inserted into the tessellation and then its value is determined by sample points within its bounding polygons. For each neighbour, the area of the portion of its original polygon that became incorporated in the tile of the new point is calculated. These areas are scaled to sum to 1 and are used as weights for the corresponding samples (Webster and Oliver, 2001).

2.1.4. Inverse Distance Weighting

The inverse distance weighting or inverse distance weighted (IDW) method estimates the values of an attribute at unsampled points using a linear combination of values at sampled points weighted by an inverse function of the distance from the point of interest to the sampled points. The assumption is that sampled points closer to the unsampled point are more similar to it than those further away in their values. The weights can be expressed as:

$$\lambda_i = \frac{1/d_i^p}{\sum_{i=1}^{n} 1/d_i^p} \tag{3}$$

where d_i is the distance between x_0 and x_i, p is a power parameter, and n represents the number of sampled points used for the estimation. The main factor affecting the accuracy of IDW is the value of the power parameter (Isaaks and Srivastava, 1989). Weights diminish as the distance increases, especially when the value of the power parameter increases, so nearby samples have a heavier weight and have more influence on the estimation, and the resultant spatial interpolation is local (see section 3.1.1 for definition) (Isaaks and Srivastava, 1989).

The choice of power parameter and neighbourhood size is arbitrary (Webster and Oliver, 2001). The most popular choice of p is 2 and the resulting method is often called inverse square distance or inverse distance squared (IDS). The power parameter can also be chosen on the basis of error measurement (*e.g.*, minimum mean absolute error, resulting the optimal IDW) (Collins and Bolstad, 1996). The smoothness of the estimated surface increases as the power parameter increases, and it was found that the estimated results become less satisfactory when p is 1 and 2 compared with p is 4 (Ripley, 1981). IDW is referred to as “moving average” when p is zero (Brus *et al.*, 1996; Hosseini *et al.*, 1993; Laslett *et al.*, 1987), “linear interpolation” when p is 1 and “weighted moving average” when p is not equal to 1 (Burrough and McDonnell, 1998).

2.1.5. Regression Models

This method is essentially a linear regression model (LM) and assumes that the data are independent of each other, normally distributed and homogeneous in variance. Regression methods explore a possible functional relationship between the primary variable and explanatory variables (*e.g.*, geographical coordinates, elevation) that are easy to measure (Burrough and McDonnell, 1998). These explanatory variables are usually referred to as secondary variables, auxiliary variables or ancillary variables. The information provided by these variables is called secondary information. The final model can be selected by a thorough understanding of the relationships between the primary variable and secondary variables and/or by using Akaike information criteria (AIC) or Bayesian information criteria (BIC).

2.1.6. Trend Surface Analysis

The trend surface analysis (TSA) is a special case of LM, which only uses geographical coordinates to predict the values of the primary variable. TSA separates the data into regional trends and local variations (Collins and Bolstad, 1996). TSA shares the same assumption as LM, and always contains all variables. It has also been extended to include other variables (Collins and Bolstad, 1996), in which case, it should be classified as LM.

2.1.7. Splines and Local Trend Surfaces

The splines consist of polynomials with each polynomial of degree p being local rather than global. The polynomials describe pieces of a line or surface (*i.e.*, they are fitted to a small number of data points exactly) and are fitted together so that they join smoothly (Burrough and McDonnell, 1998; Webster and Oliver, 2001). The places where the pieces join are called knots. The choice of knots is arbitrary and may have a dramatic impact on the estimation (Burrough and McDonnell, 1998). For degree p = 1, 2, or 3, a spline is called linear, quadratic or cubic respectively. Typically the splines are of degree 3 and they are cubic splines (Webster and Oliver, 2001).

The local trend surfaces (LTS) fit a polynomial surface for each predicted point using the nearby samples (Venables and Ripley, 2002). There are two approaches in LTS. The first is a local polynomial regression fitting that is detailed by Cleveland *et al*. (in: Chambers and Hastie, 1992) and Cleveland and Devlin (1988). The second is a bilinear or bicubic spline that was developed to implement bivariate interpolation onto a grid for irregularly spaced point data (Akima, 1978; Akima, 1996). This method is also known as Akima's interpolator (AK). Both approaches are unable to choose the smoothness (Venables and Ripley, 2002).

2.1.8. Thin Plate Splines

Thin plate splines (TPS), formally known as "laplacian smoothing splines", were developed principally by Wahba and Wendelberger (1980) for climatic data. The smoothing parameter is calculated by minimising the generalised cross validation function (GCV). This method is relatively robust because the minimisation of GCV directly addresses the predictive accuracy and is less dependent on the veracity of the underlying statistical model (Hutchinson, 1995). TPS provides a measure of spatial accuracy (Hutchinson, 1995; Wahba and Wendelberger, 1980).

2.1.9. Classification

The classification method (Cl) uses easily accessible soft information (*e.g.*, soil types, vegetation types, or administrative areas) to divide the region of interest into sub-regions that can be characterised by the moments (*i.e.*, mean, variance) of the attribute measured at locations within the region of interest (Burrough and McDonnell, 1998). The model for classification is:

$$\hat{z}(x_0) = \mu + \alpha_k + \varepsilon \qquad (4)$$

where $\hat{z}$ is the estimated value of the attribute at location x_0, μ is the general mean of the attribute over the region of interest, α_k is the deviation between μ and the mean of unit (type) k, and ε is the residual (pooled within-unit) error (Burrough and McDonnell, 1998). Cl can be computed using the analysis of variance (ANOVA) method or LM by specifying the attribute as a response variable and the soft information as an explanatory factor with k classes. This method shares the same assumptions as LM.

2.1.10. Regression Tree

The regression tree (CART), also known as binary decision trees, uses binary recursive partitioning whereby the data of the primary variable are successively split along the gradient of the explanatory variables into two descendent subsets or nodes. These splits occur so that at any node the split is selected to maximise the difference between two split groups or branches (Breiman *et al*., 1984). The mean value of the primary variable in each terminal node can then be used to map the variable across the region of interest (Balk and Elder, 2000).

2.1.11. Fourier Series

The Fourier series (FS) method is used to estimate the values of an attribute by interpolating the samples using a linear combination of sine and cosine waves in two-dimensional space (Davis, 1973), as follows:

$$\hat{Z}_{ij} = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\alpha_{nm}\cos\frac{2n\pi X_i}{\lambda_x}\cos\frac{2n\pi Y_j}{\lambda_y} + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\beta_{nm}\cos\frac{2n\pi X_i}{\lambda_x}\sin\frac{2n\pi Y_j}{\lambda_y} + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\gamma_{nm}\sin\frac{2n\pi X_i}{\lambda_x}\cos\frac{2n\pi Y_j}{\lambda_y} + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\delta_{nm}\sin\frac{2n\pi X_i}{\lambda_x}\sin\frac{2n\pi Y_j}{\lambda_y} \quad (5)$$

where $\hat{Z}_{ij}$ is the value estimated at a point with coordinate X_i and Y_j, α, β, γ and δ are coefficients, and λ_x and λ_y are the wavelength along x and y axis. The FS method has been proved useful in sedimentary geology in relation to periodic features such as spatial distribution of sand dunes, ripple marks and gilgai (Burrough, 1991; Davis, 1973). This method is only applicable to strict periodic variables. Given its rare application, this method will not be considered any further.

2.1.12. Lapse Rate

The lapse rate (LR) was developed to estimate air temperature in relation to elevation/altitude. It uses the temperature value of the nearest weather station and the difference in elevation to estimate air temperature at the unsampled point on the basis of the relationship between air temperature and elevation for a region. It is also termed smart interpolation (Vicente-Serrano *et al.*, 2003; Willmott and Matsuura, 1995). It makes the assumption that the lapse rate is constant across the study region (Collins and Bolstad, 1996). Several variants of LR have been proposed for air temperature (Stahl *et al.*, 2006). Given it is limited to only predicting temperature using elevation, this method will not be discussed any further.

2.2. Geostatistics

A brief introduction to geostatistics is initially provided for reference before the description of geostatistical interpolators. Most of the information about geostatistics in this section and in the next section is from Goovaerts (1997), and other relevant references are also cited when necessary.

2.2.1. Introduction of Geostatistics

Geostatistics is usually believed to have originated from the work in geology and mining by Krige (1951), but it can be traced back to the early 1910s in agronomy and 1930s in meteorology (Webster and Oliver, 2001). It was developed by Matheron (1963) with his theory of regionalised variables (Wackernagel, 2003). "*A mineralized phenomenon can be characterized by the spatial distribution of a certain number of* *measurable quantities called regionalized variables* (page 10)"; and this concept is termed regionalisation (Journel and Huijbregts, 1978). Other key concepts of geostatistics include: "*When a variable is distributed in space, it is said to be regionalized* (page 27)" and "*geostatistical theory is based on the observation that the variabilities of all regionalized variables have a particular structure* (page 10)" (Journel and Huijbregts, 1978). Geostatistics includes several methods that use kriging algorithms for estimating continuous attributes. Kriging is a generic name for a family of generalised least-squares regression algorithms, used in recognition of the pioneering work of Danie Krige (1951).

2.2.2. Semivariance and Variogram

Semivariance (γ) of Z between two data points is an important concept in geostatistics and is defined as:

$$\gamma(x_i, x_0) = \gamma(h) = \frac{1}{2}\mathrm{var}[Z(x_i) - Z(x_0)] \tag{6}$$

where h is the distance between point x_i and x_0 and $\gamma(h)$ is the semivariogram (commonly referred to as variogram) (Webster and Oliver, 2001).

A plot of $\hat{\gamma}(h)$ against h is known as the experimental variogram (Figure 2.1), which displays several important features (Burrough and McDonnell, 1998). The first is the "nugget", a positive value of $\hat{\gamma}(h)$ at h close to 0, which is the residual reflecting the variance of sampling errors and the spatial variance at shorter distance than the minimum sample spacing. The "range" is a value of distance at which the "sill" is reached. Samples separated by a distance larger than the range are spatially independent because the estimated semivariance of differences will be invariant with sample separation distance. If the ratio of sill to nugget is close to 1, then most of the variability is non-spatial (Hartkamp *et al.*, 1999). The range provides information about the size of a search window used in the spatial interpolation methods (Burrough and McDonnell, 1998).

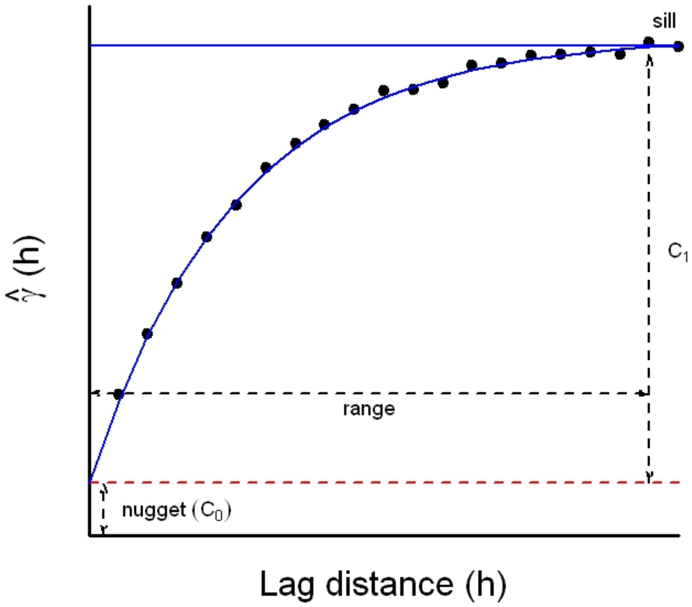


Figure 2.1. An example of a semivariogram as illustrated by an exponential model, with range, nugget (C_0) and sill (C_0+C_1).

The semivariance can be estimated from the data, as follows:

$$\hat{\gamma}(h) = \frac{1}{2n}\sum_{i=1}^{n}(z(x_i) - z(x_i + h))^2 \qquad (7)$$

where n is the number of pairs of sample points separated by distance h (Burrough and McDonnell, 1998). Variogram modelling and estimation is extremely important for structural analysis and spatial interpolation (Burrough and McDonnell, 1998). The variogram models may consist of simple models, including: Nugget, Exponential, Spherical, Gaussian, Linear, and Power model or the nested sum of one or more simple models (Burrough and McDonnell, 1998; Pebesma, 2004; Webster and Oliver, 2001). Four commonly used variogram models are illustrated in Figure 2.2 based on equations in Burrough and McDonnell (1998).

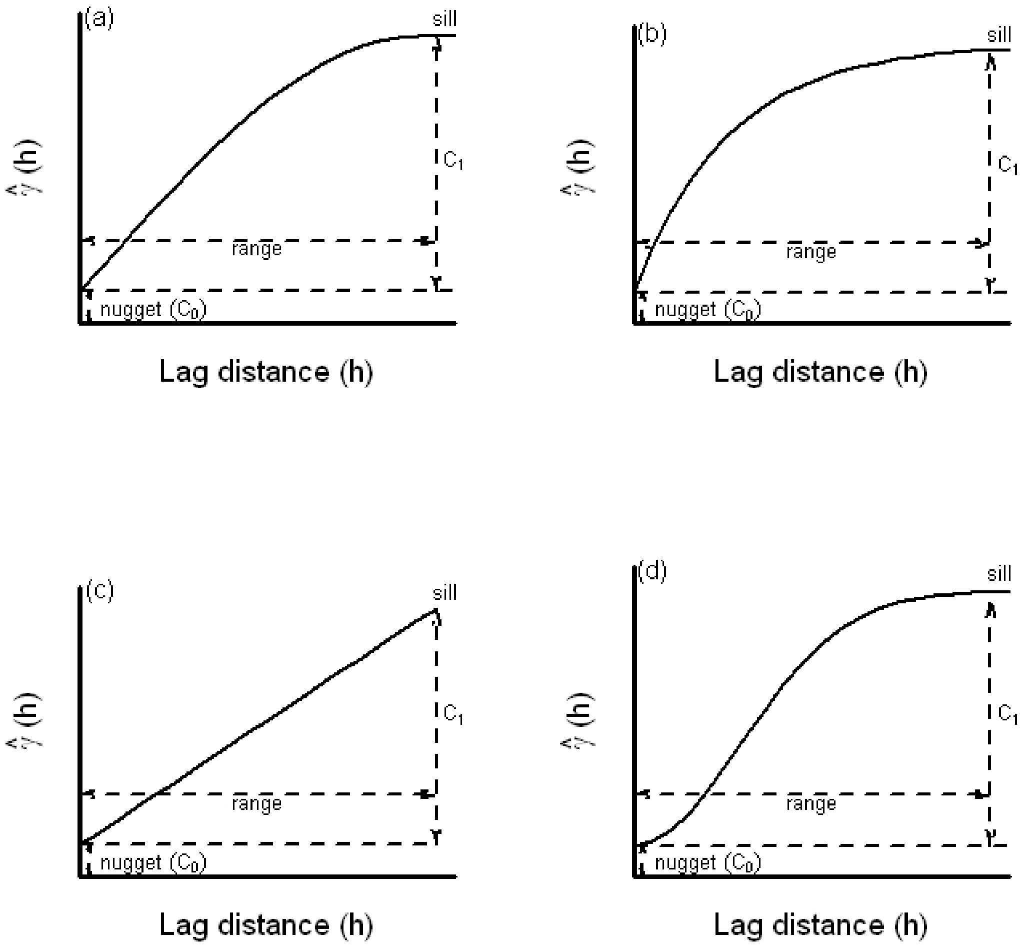


Figure 2.2. Examples of four commonly used variogram models: (a) spherical; (b) exponential; (c) linear; and (d) Gaussian.

2.2.3. Kriging Estimator

All kriging estimators are variants of the basic equation (8), which is a slight modification of equation (1), as follows:

$$\hat{Z}(x_0) - \mu = \sum_{i=1}^{n} \lambda_i \left[Z(x_i) - \mu(x_0) \right] \tag{8}$$

where μ is a known stationary mean, assumed to be constant over the whole domain and calculated as the average of the data (Wackernagel, 2003). The parameter λ_i is kriging weight; n is the number of sampled points used to make the estimation and depends on the size of the search window; and $\mu(x_0)$ is the mean of samples within the search window.

The kriging weights are estimated by minimising the variance, as follows:

$$\begin{aligned}&\mathrm{var}[\hat{Z}(x_0)] = E[\{\hat{Z}(x_0) - Z(x_0)\}^2]\\ &= E\left[(\hat{Z}(x_0))^2 + (Z(x_0))^2 - 2\hat{Z}(x_0)Z(x_0)\right]\\ &= \sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j C(x_i - x_j) + C(x_0 - x_0) - 2\sum_{i=1}^{n}\lambda_i C(x_i - x_0)\end{aligned} \tag{9}$$

where $Z(x_0)$ is the true value expected at point x_0, n represents the number of observations to be included in the estimation, and $C(x_i$-$x_j)$ = Cov[Z(x_i), Z(x_j)] (Isaaks and Srivastava, 1989).

The step by step procedures for finding equation (9) and linking it to γ are given by Clark and Harper (2001). The assumptions of kriging are stationarity of difference between x and $x+h$ and variance of differences, which define the requirements for the intrinsic hypothesis (Burrough and McDonnell, 1998; Journel and Huijbregts, 1978). This means that semivariance does not depend on the location of samples and only depends on the distance between samples, thus the semivariance is isotropic.

2.3. Geostatistical Interpolators

In this section 22 geostatistical interpolators are briefly described. Geostatistical approaches are used to: 1) describe spatial patterns and interpolate the values of the primary variable at unsampled locations; and 2) model the uncertainty or error of the estimated surface.

2.3.1. Simple Kriging

The estimation of simple kriging (SK) is based on equation (7) and a slightly modified equation (8), leading to equation (10) as:

$$\hat{Z}(x_0) = \sum_{i=1}^{n}\lambda_i Z(x_i) + \left[1 - \sum_{i=1}^{n}\lambda_i\right]\mu \tag{10}$$

where μ is a known stationary mean. The parameter μ is assumed constant over the whole domain and calculated as the average of the data (Wackernagel, 2003). SK is used to estimate residuals from this reference value μ given a priori and is therefore sometimes referred to as “kriging with known mean” (Wackernagel, 2003). The parameter $\mu(x_0)$ in equation (8) is replaced by the stationary mean μ in equation (10). The number of sampled points used to make the estimation in equation (10) is determined by the range of influence of the semivariogram (Burrough and McDonnell, 1998). Because SK does not have a non-bias condition, $1-\sum_{i=1}^{n}\lambda_i$ is not necessarily 0; the greater the value of $1-\sum_{i=1}^{n}\lambda_i$, the more the estimator will be drawn toward the mean; and in general the value of $1-\sum_{i=1}^{n}\lambda_i$ increases in relative poorly sampled regions (Boufassa and Armstrong, 1989). SK assumes second-order stationary that is constant mean, variance and covariance over the domain or the region of interest (Wackernagel, 2003; Webster and Oliver, 2001). Because such an assumption is often too restrictive, ordinary kriging (no a priori mean) is most often used (Burrough and McDonnell, 1998).

2.3.2. Ordinary Kriging

The ordinary kriging (OK) is similar to SK and the only difference is that OK estimates the value of the attribute using equations (7) and (8) by replacing μ with a local mean $\mu(x_0)$ that is the mean of samples within the search window, and forcing $\left[1-\sum_{i=1}^{n}\lambda_i\right]=0$, that is $\sum_{i=1}^{n}\lambda_i=1$, which is achieved by plugging it into equation (8) (Clark and Harper, 2001; Goovaerts, 1997). Therefore, OK essentially uses equations (7) and (1) to make the estimation. OK estimates the local constant mean, then performs SK on the corresponding residuals, and only requires the stationary mean of the local search window (Goovaerts, 1997).

2.3.3. Kriging with a Trend

The kriging with a trend (KT) is normally called universal kriging (UK) that was proposed by Matheron (1969). It is an extension of OK by incorporating the local trend within the neighbourhood search widow as a smoothly varying function of the coordinates. UK estimates the trend components within each search neighbourhood window and then performs SK on the corresponding residuals.

2.3.4. Block Kriging

The block kriging (BK) is a generic name for estimation of average values of the primary variable over a segment, a surface, or a volume of any size or shape (Goovaerts, 1997). It is an extension of OK and estimates a block value instead of a point value by replacing the point-to-point covariance with the point-to-block covariance (Wackernagel, 2003). Essentially, BK is block OK and OK is "point" OK.

2.3.5. Factorial Kriging

The factorial kriging (FK) is designed to determine the origins of the value of a continuous attribute (Goovaerts, 1997). It models the experimental semivariogram as a linear combination of a few basic structure models to represent the different factors operating at different scales (*e.g.*, local and regional scales). FK can decompose the kriging estimates into different components such as nugget, short-range, long-range, and trend, and such components could be filtered in mapping if considered as noise. For example, the nugget component at sampled points could be filtered to remove discontinuities (peaks) at the sampled points, while the long-range component could be filtered to enhance the short-range variability of the attribute. FK assumes that noise and the underlying signal are additive and that the noise is homoscedastic. Given the nature of this method, it will not be further considered in this review.

2.3.6. Dual Kriging

The dual kriging (DuK) estimates the covariance values instead of data values to elucidate the filtering properties of kriging (Goovaerts, 1997). It also reduces the computational cost of kriging when used with a global search neighbourhood. It includes dual SK, dual OK, and dual FK. Given its restricted application, it will not be given further consideration in this review.

2.3.7. Simple Kriging with Varying Local Means

The SK with varying local means (SKlm) is an extension of SK by replacing the stationary mean with known varying means at each point that depend on the secondary information (Goovaerts, 1997). If the secondary variable is categorical, the primary local mean is the mean of the primary variable within a specific category of the secondary variable. If it is continuous, the primary local mean is a function of the secondary variable or can be acquired by discretising it into classes. SK is then used to produce the weights and estimates.

2.3.8. Kriging with an External Drift

The kriging with an external drift (KED) is similar to UK but incorporates the local trend within the neighbourhood search window as a linear function of a smoothly varying secondary variable instead of as a function of the spatial coordinates (Goovaerts, 1997). The trend of the primary variable must be linearly related to that of the secondary variable. This secondary variable should vary smoothly in space and is measured at all primary data points and at all points being estimated. KED is also called UK or external drift kriging in Pebesma (2004). KED could be extended to include both secondary variables and coordinate information if gstat is used (personal communication with Edzer Pebesma, at useR! 2008, Dortmund, Germany, August 13, 2008).

2.3.9. Cokriging

Unlike SK within strata (see section 2.3.15), SKlm and KED that require the availability of information of secondary variables at all points being estimated, cokriging (CK) is proposed to use non-exhaustive secondary information and to explicitly account for the spatial cross correlation between the primary and secondary variables (Goovaerts, 1997). Equation (8) can be extended to incorporate the additional information to derive equation (11), as follows:

$$\hat{Z}_1(x_0) - \mu_1 = \sum_{i_1=1}^{n_1} \lambda_{i_1} \left[Z_1(x_{i_1}) - \mu_1(x_{i_1}) \right] + \sum_{j=2}^{n_v} \sum_{i_j=1}^{n_j} \lambda_{i_j} [Z_j(x_{i_j}) - \mu_j(x_{i_j})] \tag{11}$$

where μ_1 is a known stationary mean of the primary variable, $Z_1(x_{i_1})$ is the data of the primary variable at point i_1, $\mu_1(x_{i_1})$ is the mean of samples within the search window, n_1 is the number of sampled points within the search window for point x_0 used to make the estimation, (λ_{i_1}) is the weight selected to minimise the estimation variance of the primary variable, n_v is the number of secondary variables, n_j is the number of jth secondary variable within the search widow, λ_{i_j} is the weight assigned to i_j^{th} point of j^{th} secondary variable, $Z_j(x_{i_j})$ is the data at i_j^{th} point of j^{th} secondary variable, and $\mu_j(x_{i_j})$ is the mean of samples of j^{th} secondary variable within the search widow.

The cross-semivariance (or cross-variogram) can be estimated from data using the following equation:

$$\hat{\gamma}_{12}(h) = \frac{1}{2n} \sum_{i=1}^{n} [z_1(x_i) - z_1(x_i + h)][z_2(x_i) - z_2(x_i + h)] \tag{12}$$

where n is the number of pairs of sample points of variable z_1 and z_2 at point x_i, x_1+h separated by distance h (Burrough and McDonnell, 1998). Cross-semivariances can increase or decrease with h depending on the correlation between the two variables and the Cauchy-Schwartz relation must be checked to ensure a positive CK estimation variance in all circumstances (Burrough and McDonnell, 1998).

2.3.10. Simple Cokriging

Replacing $\mu_1(x_{i_1})$ with the stationary mean (μ_1) of the primary variable, and replacing $\mu_j(x_{i_j})$ with the stationary mean μ_j of the secondary variables in equation (11) will give the linear estimator of simple cokriging (SCK) (Goovaerts, 1997) as:

$$\begin{aligned} \hat{Z}_1(x_0) &= \sum_{i_1=1}^{n_1} \lambda_{i_1} \left[Z_1(x_{i_1}) - \mu_1 \right] + \mu_1 + \sum_{j=2}^{n_v} \sum_{i_j=1}^{n_j} \lambda_{i_j} [Z_j(x_{i_j}) - \mu_j] \\ &= \sum_{i_1=1}^{n_1} \lambda_{i_1} \left[Z_1(x_{i_1}) \right] + \sum_{j=2}^{n_v} \sum_{i_j=1}^{n_j} \lambda_{i_j} Z_j(x_{i_j}) + (1 - \sum_{i_1=1}^{n_1} \lambda_{i_1}) \mu_1 - \sum_{j=2}^{n_v} \sum_{i_j=1}^{n_j} \lambda_{i_j} \mu_j \end{aligned} \tag{13}$$

If the primary and secondary variables are not correlated, the SCK estimator reverts to the SK estimator (Goovaerts, 1997). The weights generally decrease as the corresponding data points get farther away from the point of interest. When the point of interest is beyond the correlation range of both the primary and secondary data, the SCK estimator then reverts to the stationary mean of the primary variable. If all secondary variables are recorded at every sampled point, it is referred to as "equally sampled" or "isotopic". If the primary variable is undersampled relative to the secondary variables, it is referred to as "undersampled" or "heterotopic". When the secondary variables are linearly dependent, one should be kept and other correlated variables discarded, and multivariate analysis such as principal component analysis (PCA) may be used to eliminate such dependency. The sill of the cross semivariogram model is the correlation coefficient between the primary and secondary variables.

2.3.11. Ordinary Cokriging

The ordinary cokriging (OCK) is similar to SCK (Goovaerts, 1997). The only difference is that OCK estimates the value of the primary variable using equation (13) by replacing μ_1 and μ_j with a local mean $\mu_1(x_0)$ and $\mu_j(x_0)$ (*i.e.*, the mean of samples within the search window), and forcing $\sum_{i_1=1}^{n_1}\lambda_{i_1}=1$, and $\sum_{i_j=1}^{n_j}\lambda_{i_j}=0$. These two constraints may result in negative and/or small weights. To reduce the occurrence of negative weights, these two constraints are combined to form the single constraint: $\sum_{i_1=1}^{n_1}\lambda_{i_1}+\sum_{i_j=1}^{n_j}\lambda_{i_j}=1$.

OCK amounts to estimating the local primary and secondary means and applying the SCK estimator (equation 13) with these estimates of the means rather than the stationary means (Goovaerts, 1997).

Related primary information, such as constraint intervals (indicating the intervals of the primary variable) or categorical information (indicating occurrence of a particular facies), is referred to as soft information rather than secondary information because they relate directly to the primary variable (Goovaerts, 1997). OCK can also be used for interval-type soft information by replacing $z_j(x_{i_j})$ with indicator data. However, all soft information is treated as secondary information in this review.

2.3.12. Standardised Ordinary Cokriging

The OCK has two drawbacks by calling for the secondary data weights to sum to zero (Goovaerts, 1997). The first is that some of the weights are negative, thus increasing the risk of getting unacceptable estimates. The second is that most of the weights tend to be small, thus reducing the influence of the secondary data. To overcome these drawbacks, the standardised OCK (SOCK) estimator was introduced, which calls for knowledge of the stationary means of both the primary and secondary variables. These means can be estimated from the sample means. SOCK still accounts for the local departures from the overall means as OCK.

2.3.13. Principal Component Kriging

The principal component kriging (PCK) applies PCA to a few (*nv*) secondary variables to generate *nv* orthogonal or uncorrelated PCA components (Goovaerts, 1997). OK is then applied to each of the components to get principal component estimates. The final estimate is then generated as a linear combination of the principal component estimates weighted by their loadings and plus the local attribute mean.

2.3.14. Colocated Cokriging

The colocated cokriging (CCK) is a variant of CK (Goovaerts, 1997). It only uses the single secondary datum of any given type closest to the point being estimated. Like CK, CCK can also have several variants like simple colocated cokriging (SCCK), and ordinary colocated cokriging (OCCK). CCK is proposed to overcome problems, such as screening effects of samples of the secondary variables close to or colocated with the point of interest. This situation arises when the sample densities of the secondary variables are much higher than that of the primary variable. OCCK is also the preferred method for categorical soft information.

2.3.15. Kriging within Strata

The kriging within strata (KWS) is characterised by 1) stratifying the study region based on secondary information, 2) calculating experimental semivariogram within each stratum, and 3) estimating the primary variable within each specific stratum using the semivariogram model and the closest primary data samples within the stratum (Goovaerts, 1997). It may include simple KWS (SKWS), ordinary KWS (OKWS), simple cokriging within strata (SCKWS) and ordinary cokriging within strata (OCKWS).

2.3.16. Multivariate Factorial Kriging

The multivariate factorial kriging (MFK), also called factorial kriging analysis, analyses the relationships between variables at the spatial scales detected and modelled from experimental semivariograms (Goovaerts, 1997). Given that, it will not be further considered in this review.

2.3.17. Indicator Kriging

The indicator kriging (IK) and all its extensions below are proposed mainly to assess the uncertainty about the unknown rather than to estimate the unknown (Goovaerts, 1997), so these methods will only be briefly described here. IK may also be applied to the spatial generalisation for categorical data (Jerosch *et al*., 2006). Besides IK, the multiGaussian (MG) approach and local uncertainty models (*e.g.*, conditional variance, local entropy, and interquartile range) are also proposed for this purpose, and are not described in this review.

In IK, the original data are transformed from a continuous to a new scale and different indicator coding types are proposed depending on the nature of the local information available (Goovaerts, 1997). For hard data (a precise measurement of the primary variable) a binary indicator scale is resulted by scoring it as 1 if it is less than or equal to a specified threshold, and 0 otherwise. This is the most commonly used indicator coding type. Indicator coding is also used for other data types like constraint intervals, soft categorical data, soft continuous data and colocated sources of information. IK may include simple IK (SIK) and ordinary IK (OIK). They are similar to SK and OK, but vary by only replacing the values of continuous attribute with the indicator values. IK is exact as SK and OK. Like OK, IK can also be extended to kriging over a block, and like SK, SIK can be extended to SIK with local prior means. IK can also be combined with KED by applying the KED framework for each indicator, resulting in IK with external drift (IKED) (Haberlandt, 2007).

2.3.18. Indicator Cokriging

SIK and OIK ignore indicator data at the thresholds that are different from that being estimated (Goovaerts, 1997). Information from all thresholds can be accounted for using the cokriging method introduced in section 2.3.9. The primary and secondary variables in CK are replaced by indicator values at each threshold, resulting in indicator cokriging (ICK). In short, ICK is the sum of IK over all thresholds.

IK and ICK estimators become identical when all indicator data are intrinsically correlated (*i.e.*, all indicator direct- and cross- semivariogram models are proportional to a common semivariogram, and all vectors of hard indicator data are equally sampled; Goovaerts, 1997). The resultant estimator is called median indicator kriging (MIK) because the common model is usually inferred from the indicator semivariogram at the median threshold value.

When hard and soft indicator data are used as the primary and secondary variables in ICK, the resultant estimator is called soft cokriging, in which the soft information needs no longer to be exhaustive (Goovaerts, 1997). The demanding computation of soft cokriging can be alleviated by a Markov-Bayes algorithm. When soft data are much more numerous than hard data, only the datum closest to the point being estimated (*e.g.*, the colocated soft indicator datum) is retained in the ICK, resulting colocated ICK.

2.3.19. Probability Kriging

The probability kriging (PK) is the cokriging of the indicator data using the rank-order transform as a secondary variable (Goovaerts, 1997). The indicator data are values of 0 or 1. The rank-order transform is the standardised ranks that are the rank-order of each datum of the primary variable divided by sample size. Replacing the values of the primary variable in CK by indicator data and using the rank-order transform as the secondary variable in CK would result in a PK estimator.

2.3.20. Disjunctive Kriging

The disjunctive kriging (DK) is used for the primary variable that the conventional transformations (*e.g.*, logarithm or square-rooting) cannot yield a near-normal distribution. In DK, the primary variable is transformed into Hermite polynomials, which are a series of normally distributed sub-variables that are kriged separately. The resultant estimates are summed to give the DK estimator (Gaus *et al*., 2003). DK also provides an estimate of the conditional probability that a random variable located at a point, or averaged over a block in two-dimensional space, exceeds certain thresholds. DK produces a nonlinear unbiased, distribution-dependent estimator with the characteristics minimum variance of errors (Burrough and McDonnell, 1998; Yates *et al*., 1986). The theory of disjunctive kriging and examples of its practical application are described by Armstrong and Matheron (1986a; 1986b), Rendu (1980) and Oliver *et al*. (1996).

Methods for uncertainty assessment at unsampled points will not be discussed any further in this review. For those interested in application of these approaches, further information could be found in the following studies. Lark and Ferguson (2004) compared IK and DK for mapping risk of soil nutrient deficiency or excess. Emery (2006) used ordinary multiGaussian kriging, ordinary DK, ordinary IK and conditional expectation for assessing the risk of deficiency or excess of a soil property at unsampled locations. Carr and Deng (1987) compared DK and IK using earthquake ground motion data.

2.3.21. Model-based Kriging

Model-based kriging (MBK) was developed by Diggle *et al*. (1998). This method embeds the linear kriging methodology within a more general distributional framework that is characteristically similar to the structure of a generalized linear model. A Bayesian approach is adopted and implemented via the Markov chain

Monte Carlo (MCMC) methodology, to predict arbitrary functionals of an unobserved latent process whilst making a proper allowance for the uncertainty in the estimation of any model parameters (Moyeed and Papritz, 2002). This method was further illustrated in Diggle and Ribeiro Jr. (2007). Given its heavy computational demanding (Moyeed and Papritz, 2002), this method will not be considered further in this review as it is not applicable to large dataset such as those for the Australian marine region.

2.3.22. Simulation

The MG and indicator-based algorithms provide only models of local uncertainty in that each conditional cumulative distribution function is specific to one single location (Goovaerts, 1997). Stochastic simulation models the spatial uncertainty about attribute values at several locations taken together by generating multiple realisations of the joint distribution of attribute values in space. It is heavy computationally demanding and its interpolated surface should be very similar to OK interpolation if enough realisations are used (>500) (Burrough and McDonnell, 1998). Given the nature of this approach, it will not be described further and detailed information of stochastic simulation can be found in Goovaerts (1997). There is also brief description of procedures of simulation in Burrough and McDonnell (1998).

2.4. Combined Procedures

A range of combined procedures could be developed to generate the estimation based on the above spatial interpolation methods and other statistical approaches. Here, eight methods identified in the literature are listed. Within each method there may be several sub-methods or types.

2.4.1. Classification Combined with Other Spatial Interpolation Methods

A significant explanatory secondary categorical variable, such as vegetation classes or geomorphic units, can be used as prior information to stratify the data into subsets for each category. Other spatial interpolation methods are then applied within each stratum, resulting in various combinations, such as:

1) “SK within classes” (Stein *et al.*, 1988; Voltz and Webster, 1990) or “stratified SK” (StSK) (Brus *et al.*, 1996);

2) “Stratified OK” (OK with classes, *i.e.*, StOK) (Hernandez-Stefanoni and Ponce-Hernandez, 2006; Stein *et al.*, 1988);

3) “Stratified OCK” (StOCK) (Hernandez-Stefanoni and Ponce-Hernandez, 2006);

4) “Stratified TPS” (StTPS) (Brus *et al*., 1996);

5) “Stratified NN” (StNN) (Brus *et al*., 1996);

6) “Stratified moving average” (a special case of IDW with limited number of neighbourhoods and a zero distance power) (StIDW-0) (Brus *et al*., 1996);

7) “Stratified global mean (GM)” (StGM) (Brus *et al*., 1996); and

8) “Stratified IDW” (StIDW) (Hernandez-Stefanoni and Ponce-Hernandez, 2006) and “stratified IDS” (StIDS) (Brus *et al*., 1996).

Separate variograms could be derived for each class if there are enough samples (Stein *et al*., 1988), otherwise a single pooled variogram within classes can be used if separate variograms require too many samples (Voltz and Webster, 1990) or, in other words, if the sample size is too small. The first three methods in the list above are variants of KWS.

2.4.2. Trend Surface Analysis Combined with Kriging

The TSA is fitted to the data, which describes the large scale (global) spatial variability; the residuals from the TSA are then modelled using OK and OCK; and the final estimates are the sum of the kriged residuals and the estimated trend surface (Wang *et al*., 2005).

2.4.3. Lapse Rate Combined with Kriging

Wang *et al*. (1996), cited by (Li *et al*., 2005), proposed LR combined with kriging. The data are corrected using LR to the same altitude. OK is then applied to the corrected data that no longer contain any altitude information. Lastly, the estimations are the sum of the kriged values and the corrections of altitude effect by using LR of air temperature.

2.4.4. Linear Mixed Model

An approach, similar to TSA combined with kriging, was proposed in the context of the linear mixed model (LMM) (Gilmour *et al*., 2004; Welham *et al*., 2004). The global trend that accounts for changes with coordinates was fitted as the fixed effects; local trend that is the difference between the observed values was estimated through a covariance structure; and both global and local trends were estimated using the residual maximum likelihood method (REML) in a single analysis. The resultant surface is the sum of global trend and local trend.

2.4.5. Regression Tree Combined with Kriging

Regression tree (CART) is fitted to the data to produce a tree with optimal tree size; the residuals produced by the tree regression analysis are then analysed using OK (Balk and Elder, 2000; Bishop and McBratney, 2001; Erxleben *et al*., 2002) or OCK (Balk and Elder, 2000; Erxleben *et al*., 2002). The final estimations are the sum of the CART estimates and the kriged residuals.

2.4.6. Residual Maximum Likelihood-empirical Best Linear Unbiased Predictor

Residual maximum likelihood-empirical best linear unbiased predictor (REML-EBLUP) is implemented in the context of the linear mixed model that comprises fixed effects (such as a trend model), random effects (*i.e.*, spatially dependent random variables) and an independent random variable (Lark *et al*., 2006). It consists of these three computations to generate the EBLUP estimate for an unsampled point. The first is the estimation of a variance structure (*e.g.*, a variogram) for some specified linear mixed model using REML. This is then used to obtain estimates of the model coefficients for the fixed and random effects that we need to form the EBLUP. The estimated variance model is then used to compute the EBLUP estimates or predictions at unsampled points. The predicted value consists of two components, namely: 1) a prediction based on the polynomial trend model or external drift variables and 2) a kriging estimate of the spatially dependent random effect.

2.4.7. Regression Kriging

Several types of regression kriging (RK) have been proposed (Hengl, 2007; Minasny and McBratney, 2007; Odeh *et al*., 1994; Odeh *et al*., 1995):

1) RK-A, type A (Odeh *et al*., 1994), called “kriging combined with (linear) regression” (Ahmed and De Marsily, 1987; Knotters *et al*., 1995), performs regression and then kriges the regressed values;

2) RK-B, type B (Odeh *et al*., 1994), called “kriging with a guess field” (Ahmed and De Marsily, 1987), involves regression, then kriges the predicted values and the residuals separately, and sums both values to generate the final prediction;

3) RK-C, type C (Odeh *et al*., 1995), variously called “regression with residual simple kriging” (Asli and Marcotte, 1995), “detrended kriging” (Jef *et al*., 2006; Nalder and Wein, 1998), “modified residual kriging and cokriging” (Erxleben *et al*., 2002; Martínez-Cob, 1996) or “residual kriging” (Mardikis *et al*., 2005), is similar to type B, but only kriges the residuals and then sums the predicted values

and the kriged values to obtain the final prediction;

4) RK-D, is similar to RK-C, but uses generalised least squares (GLS) instead. GLS is fitted to the data, the GLS residuals are interpolated using SK, and then the final estimates are the sum of the kriged residuals and the GLS surface (Hengl, 2007);

5) RK-E is a mixture of RK-C and RK-A. The key feature is that points with either measured or predicted values (values estimated using RK-C, *i.e.*, the sum of the predicted values by regression and the kriged values from residuals) are treated as equivalent in the interpolation process (Li *et al*., 2007); and

6) RK-F, a further type of RK, is similar to RK-C, but uses generalised additive model (GAM) instead. GAM is fitted to the data, the GAM residuals are interpolated using OK, and then the final estimates are the sum of the kriged residuals and the GAM surface (Bishop and McBratney, 2001).

In RK, the regression can have any form, such as generalised linear models (GLM) (Gotway and Stroup, 1997; Pebesma *et al.*, 2005) or non-linear models, which provides a possibility to include more ancillary variables (Li *et al*., 2007). A generic framework for RK was proposed by Hengl *et al*. (2004)

2.4.8. Gradient Plus Inverse Distance Squared

Gradient plus inverse distance squared (GIDS) is proposed by Nalder and Wein (1998). It performs LM using spatial coordinates and elevation data as secondary information and then the developed model is used to predict the primary variable. The predicted values are then used to derive the estimation for the unsampled points using IDS.

Several other methods have been proposed by various authors and applied to environmental data in addition to those described above. These methods are not described in this review because they are either relatively straightforward or have been rarely used. These methods include:

1) LM combined with IDW (Jarvis and Stuart, 2001; Jef *et al*., 2006; Vicente-Serrano *et al*., 2003);

2) LM combined with partial TPS (Jarvis and Stuart, 2001) and with splines (Vicente-Serrano *et al*., 2003);

3) Kriging combined with Q-mode factor analysis (Juang and Lee, 1998);

4) IK combined with PCK, resulting in "indicator principal component kriging" (Suro-Perez and Journel, 1991);

5) Density estimation (Silverman, 1981);

6) A "moving window regression residual cokriging" (MWRCK) (Sun, 1998);

7) LSZ or "Bayesian multivariate interpolation with data missing-by-design" (Sun, 1998);

8) A radial basis function network (RBFN, a variant of neural network) and an improved RBFN by Lin and Chen (2004);

9) General regression neural network (GRNN) (Kanevski *et al.*, 2008); and

10) "A consistently well behaved method of interpolation" (Stineman, 1980).

Geographically weighted regression (GWR) (Fotheringham *et al.*, 2002) has immense potential for the spatial interpolation of the environmental data. The advantages of this method are that it: 1) is based on the traditional regression framework, and 2) incorporates local spatial relationships into the framework in an intuitive and explicit manner (Fotheringham *et al.*, 2002). Although no reference on the application of this method has been acquired in this review, many unpublished applications are available online.

The field of geostatistics reached its peak around 1996-1998 in terms of the annual total citation of articles, but the total number of annually published articles is still growing (Zhou *et al.*, 2007). The development of hybrid methods is certainly not over and the methods will continue to evolve both from theoretical and practical aspects (Hengl, 2007). Five developments are anticipated in the near future in geostatistics according to Hengl (2007): 1) design of more sophisticated prediction models, 2) derivation of local regression-kriging, 3) development of user-friendly sampling optimisation packages, 4) intelligent data analysis reports generation, and 5) introduction of multi-temporal, multi-variate prediction models.

In short, the spatial interpolation methods are developed for specific types of environmental data or even specific environmental variable. They each have their specific assumptions, data requirements and properties and will be compared in the next chapter.

Chapter 3: Features and Theoretical Comparison of Spatial Interpolation Methods

The spatial interpolation methods described in Chapter 2 were developed either for specific disciplines or even for specific variables according to the nature of the data. Therefore, each of the methods has its specific assumptions and features. In this chapter, a few important features, such as global versus local, exact versus inexact, deterministic versus stochastic, and gradual versus abrupt, are discussed. The features of each of the individual methods are then summarised and compared.

3.1. Features of Spatial Interpolation Methods

3.1.1. Global versus Local

The spatial interpolation methods can be classified as either global or local methods. Global methods use all available data of the region of interest to derive the estimation and capture the general trend. Local methods operate within a small area around the point being estimated (*i.e.*, use samples within a search window) and capture the local or short-range variation (Burrough and McDonnell, 1998).

3.1.2. Exactness

The spatial interpolation methods can be either "exact' or "inexact". A method that generates an estimate that is the same as the observed value at a sampled point is called an exact method. All other methods are inexact, which means that their predicted value at the point differs from its known value (Burrough and McDonnell, 1998).

3.1.3. Deterministic versus Stochastic

Stochastic methods incorporate the concept of randomness and provide both estimations (*i.e.*, deterministic part) and associated errors (stochastic part, *i.e.*, uncertainties represented as estimated variances). All other methods are deterministic because they do not incorporate such errors and only produce the estimations. In other words, deterministic methods have no assessment of errors with the predicted values, while stochastic methods provide an assessment of the errors associated with the predicted values.

3.1.4. Gradual versus Abrupt

Some methods (*e.g.*, NN) produce a discrete and abrupt surface, while some other methods (*e.g.*, distance-based weighted averages) produce a smooth and gradual surface. The smoothness depends on the criteria used in the selection of the weight values in relation to the distance. Criteria include simple distance relations (*e.g.*, IDW), minimisation of variance (*e.g.*, SK, OK and CK), and minimisation of curvature and enforcement of smoothness (*e.g.*, splines).

3.1.5. Linear Kriging versus Nonlinear Kriging

Kriging methods are often classified as linear and nonlinear (Moyeed and Papritz, 2002; Papritz and Moyeed, 1999). There are no formal definitions for linear and nonlinear kriging. Linear kriging can be defined as kriging methods that derive the estimation using observed values by assuming a normal distribution of the samples. Non-linear kriging are those methods that derive predictions based on the transformed values of the observed data. Linear kriging may include SK, OK and UK. Nonlinear kriging methods consist of DK, IK, multiGaussian kriging, lognormal OK and MBK. Nonlinear kriging methods have two major advantages over linear kriging, namely: 1) they were developed to model the conditional distribution of the primary variable (*i.e.*, to give an estimate of its probability distribution conditional on the available information); and 2) their estimations should theoretically be more precise when a Gaussian random process is inappropriate to model the observations.

3.1.6. Univariate versus Multivariate

The spatial interpolation methods that only use samples of the primary variable in deriving the estimation are termed “univariate” methods. Methods that also use the secondary variables are often referred to as “multivariate” methods. In geostatistics, the methods accounting for a single variable, such as SK, OK, BK, are univariate, and methods accounting for secondary information, like SCK, OCK, KED, are multivariate. Although UK is classified as a method accounting for a single variable by Goovaerts (1997), it is considered as a multivariate method in this review because it uses coordinate information. TPS can also be extended to include a multivariate spline function (Burrough and McDonnell, 1998; Hutchinson, 1995; Mitasova *et al.*, 1995).

It should be noted that in the gstat package in R, multivariate kriging means that there are several variables that can be either primary variables with or without secondary variables, while multiple kriging implies that several primary variables are kriged separately at the same time (personal communication with Edzer Pebesma, 25 and 26 August 2008).

3.1.7. Irregular versus Regular System

The estimation can be interpolated based on either the regular grid system or irregular (*e.g.*, triangular) network. The advantages and disadvantages of an irregular network are discussed in Burrough and McDonnell (1998). Most of the methods in this review are based on the point data in a regular grid system that has several advantages (Burrough and McDonnell, 1998).

3.2. Comparison of the Features

The features of the spatial interpolation methods vary from one method to another, so it is difficult to summarise the features of all methods in one table. In this section, the non-geostatistical methods are first compared with the common features of geostatistical methods. The features of geostatistical methods are then summarised and compared amongst themselves.

3.2.1. Non-geostatistical Methods and Kriging Methods

The features of non-geostatistical methods and geostatistics are summarised in Table 3.1, which is mainly based on the Burrough *et al*. (1998). For geostatistical methods, their common features are summarised for a comparison with non-geostatistical methods.

The NN is a special case of IDW with p being zero and n equal to 1 (Brus *et al*., 1996; Laslett *et al*., 1987). NN is best for qualitative data when other spatial interpolation methods are not applicable (Burrough and McDonnell, 1998; Hartkamp *et al*., 1999). The disadvantages are that the estimated values at each point are based on just one sample point, there is no error estimate, and other nearby sampled points are ignored (Webster and Oliver, 2001).

The TIN is a simple, local and deterministic method (Webster and Oliver, 2001). It is better than NN, although each estimate still depends on only three samples. The estimated surface is continuous but with abrupt changes in gradient at the margins of the triangles (Webster and Oliver, 2001).

The NaN is local and deterministic (Webster and Oliver, 2001). It is somewhat better than NN and TIN because its estimated surface is continuous and smooth except at the data points where its derivative is discontinuous. However, such abrupt changes can be smoothed (Sibson, 1981). At local maxima and minima in such data it can generate an artefact known as “Prussian helmets” (Sibson, 1981; Webster and Oliver, 2001).

The IDW works well with regularly spaced data, but it is unable to account for clustering (Isaaks and Srivastava, 1989).

The TSA is considered to be a stochastic method (Collins and Bolstad, 1996). However, in other publications it is described as a deterministic method with a local stochastic component (Burrough and McDonnell, 1998).

The classification method assumes that all spatial changes take place at boundaries, which are sharp instead of gradual (Burrough and McDonnell, 1998).

Table 3.1. Comparison of non-geostatistical spatial interpolation methods and kriging as a generic model for geostatistical methods (mainly modified from Burrough and McDonnell (1998).

Method	Assumption	Univariable / Multivariable	Deterministic/ stochastic	Local/ global	Exact/ inexact	Abrupt/ gradual	Limitation of the procedure	Computing load	Output data structure	Suitability
Nearest neighbours (NN)	Best local predictor is nearest data point	Univariable	Deterministic	Local	Exact	Abrupt	No error assessment, only one data point per polygon. Tessellation pattern depends on distribution of data.	Small	Polygons or gridded surface	Nominal data from point observations
Triangulation (TIN)	Best local predictor is data points on the surrounding triangle	Univariable	Deterministic	Local	Exact	Abrupt	No error assessment. TIN pattern depends on distribution of data and there a few ways to form triangulation and no one is better than any other.	Small	Triangles or Gridded surface	Quick interpolation from sparse data on regular or irregularly spaced samples.
Natural neighbours (NaN)	Best local predictor is data points in the surrounding polygons	Univariable	Deterministic	Local	Exact	Gradual or abrupt	No error assessment	Small	Gridded surface	Quick interpolation from sparse data on regular or irregularly spaced samples.
Inverse distance weighting (IDW)	underlying surface is smooth	Univariable	Deterministic	Local	Inexact (but can be forced to be exact)	Gradual	No error assessment. Results depend on size of search window and choice of weighting parameter. Poor choice of window can give artefacts when used with high data densities such as digitised contours.	Small	Gridded surface, contours	Quick interpolation from sparse data on regular grid or irregularly spaced samples.
Regression models (LM)	Samples are independent, normal and homogeneous in variance	Univariable/ Multivariable	Stochastic	Global	Inexact	Gradual if inputs have gradual variation	Results depend on the fit of the regression model and the quality and detail of the input data surfaces. Error assessment possible if input errors are known.	Small	Polygons or continuous, gridded surface	Simple numerical modelling of expensive data when better methods are not available or budgets are limited
Trend surface analysis (TSA)	Phenomenological explanation of trend, normally distributed data	Multivariable	Stochastic	Global	Inexact	Gradual	Physical meaning of trend may be unclear. Outliers and edge effects may distort surface. Error assessment limited to goodness of fit.	Small	continuous, gridded surface	Quick assessments and removal of spatial trend

Method	Assumption	Univariable / Multivariable	Deterministic/ stochastic	Local/ global	Exact/ inexact	Abrupt/ gradual	Limitation of the procedure	Computing load	Output data structure	Suitability
Splines & Local trend surfaces (LTS)	Best local predictor is the nearest data point and data normality	Multivariable	Stochastic	Local	Inexact	Gradual	Results depend on span parameter and detail of the input data surfaces.	Moderate	continuous, gridded surface	Quick interpolation from sparse data on regular grid or irregularly spaced samples.
Classification (Cl)	Homogeneity within boundaries	Univariable	Deterministic "soft" information	Global	Inexact	Abrupt	Delineation of areas and classes may be subjective. Error assessment limited to within-class standard deviations.	Small	Classified polygons	Quick assessments when data are sparse Removing systematic differences before continuous interpolation from data points
Regression tree (CART)	Phenomenological explanation of variance	Multivariable	Stochastic	Global	Inexact	?	?	Small	Gridded surface	
Thin plat splines (TPS)	Underlying surface is smooth everywhere	Univariable/ Multivariable	Deterministic	Local	Exact	Gradual	Goodness of fit possible, but within the assumptions that the fitted surface is perfectly smooth.	Small	Gridded surface, contours	Quick interpolation (univariate or multivariate) of digital elevation data and related attributes to create digital elevation models (DEM) from moderately detailed data
Kriging*	Interpolated surface is smooth. Statistical stationarity and the intrinsic hypothesis.	Univariable/ Multivariable	Stochastic	Local	Exact	Gradual	Error assessment depends on variogram and distribution of data points and size of interpolated blocks. Requires care when modelling spatial correlation structures.	Moderate	Gridded surface	When data are sufficient to compute variograms, kriging provides a good interpolator for sparse data. Binary and nominal data can be interpolated with Indicator kriging. Soft information can also be incorporated as trends or stratification. Multivariate data can be interpolated with co-kriging.

* Of the kriging methods, BK is an inexact interpolator.

Splines are deterministic with locally stochastic properties. Splines are piece-wise functions using a few points at a time. The interpolation predictions can be quickly calculated and predictions are very close to the values being interpolated, providing the measurement errors associated with the data are small (Burrough and McDonnell, 1998; Mitasova *et al*., 1995). Splines retain small-scale features, but there are no direct estimates of the errors (Burrough and McDonnell, 1998). The application of splines and other nonparametric regression models to data on a grid is sometimes questionable because the dataset does not have the direct information needed for reliable prediction and the dataset yields no direct information on residual variance (Laslett, 1994).

Exact splines may produce local artefacts of excessively high or low values. These artefacts can be removed using TPS, where an exact spline surface is replaced by a locally smoothed average (Burrough and McDonnell, 1998). TPS can also be extended to include multivariate spline function (Burrough and McDonnell, 1998; Hutchinson, 1995; Mitasova *et al*., 1995). TPS may provide a view of reality that is unrealistically smooth and thus generate misleading results (Burrough and McDonnell, 1998).

Kriging has few advantages and also some drawbacks (Nalder and Wein, 1998). It provides the best linear unbiased estimate. It also provides a measure of the error or uncertainty at the unsampled points; and it is an exact method with an exception of BK. It does not produce edge-effects resulting from trying to force a polynomial to fit the data as with TSA (Collins and Bolstad, 1996). However, it assumes stationarity of data, which is usually not true, although this assumption can be relaxed with specific forms of kriging. Definition of the required variogram models is time consuming and somewhat subjective; and definition of neighbourhoods is also required which is difficult to do objectively. It also assumes that the data are isotropic. Data transformation may be needed for non-stationary data and anisotropy needs to be considered.

Kriging also requires a large number of samples, at least 100, to produce a reliable estimation of variogram (Webster and Oliver, 1992). This requirement could be overcome by using the REML variogram because predictions based on REML variograms were generally more accurate than those from the conventional moment variograms with fewer than 100 samples (Kerry and Oliver, 2007). In such cases, a sample size of 50 appears adequate (Kerry and Oliver, 2007).

Kriging variance is independent of the data values. Hence, it does not reflect the uncertainty expected at a specific point. It is a ranking index of data geometry (and size) and is not a measure of the local spread of errors (Goovaerts, 1997). The error variance provided by kriging algorithms is also poorly correlated with actual estimation error. Therefore, in general, the kriging variance cannot be used alone as a measure of local uncertainty (Goovaerts, 1997).

3.2.2. Geostatistical Methods

The features of 16 geostatistical methods are summarised in Table 3.2 according to the findings in Goovaerts (1997) and other previous studies cited in Chapter 2. Some of the methods reviewed in Chapter 2 are excluded because either they are for uncertainty assessment, for categorical variables, or rarely used in practice. However, a few frequently used variants or sub-methods are included. Prior to the comparison and discussion of geostatistical methods, issues relevant to kriging weights and search neighbourhood windows are discussed.

Table 3.2. A comparison of geostatistical spatial interpolation methods.

Geostatistical method	Univariable/ Multivariable	Stationary/ local mean	Local trend	Information of coordinates	Secondary variable	Point/ block estimation	Exhaustive secondary information	Stratification	Orthogonalisation of secondary information	Single or multiple samples in the search window
Simple kriging (SK)	Univariate	stationary	no	no	no	point	na	no	na	multiple
Ordinary kriging (OK)	Univariate	local	no	no	no	point	na	no	na	multiple
Block kriging (BK)	Univariate	local	no	no	no	block	na	no	na	multiple
Universal kriging (UK)	Multivariate	local	yes	yes	no	point	yes	no	no	multiple
SK with varying local means (SKlm)	Multivariate	local	no/yes*	no	yes	point	yes	no	no	multiple
Kriging with an external drift (KED)	Multivariate	local	yes	yes	yes	point	yes	no	no	multiple
Simple cokriging (SCK)	Multivariate	stationary	no	no	yes	point	no	no	no	multiple
Ordinary cokriging (OCK)	Multivariate	local	no	no	yes	point	no	no	no	multiple
Standardised OCK (SOCK)	Multivariate	both#	no	no	yes	point	no	no	no	multiple
Principal component kriging (PCK)	Multivariate	local	no	no	yes	point	no	no	yes	multiple
Simple colocated cokriging (SCCK)	Multivariate	stationary	no	no	yes	point	no	no	no	single
Ordinary colocated cokriging (OCCK)	Multivariate	local	no	no	yes	point	no	no	no	single
Simple kriging within strata (SKWS)	Multivariate	within strata stationary	no	no	no	point	na	yes	na	multiple
Ordinary kriging within strata (OKWS)	Multivariate	local	no	no	no	point	na	yes	na	multiple
Simple cokriging within strata (SCKWS)	Multivariate	within strata stationary	no	no	yes	point	no	yes	no	multiple
Ordinary cokriging within strata (OCKWS)	Multivariate	local	no	no	yes	point	no	yes	no	multiple

* The local trend is "no" if the secondary variable is categorical and "yes" if it is continuous.

Need the stationary means of both the primary and secondary variables.

Kriging Weights

A number of factors affect kriging weights (Webster and Oliver, 2001). The value of the weight assigned to a sample increases as the distance to the estimated point or block decreases. The larger the nugget, the smaller the value of the weight of a sample that is nearest to the point or block estimated. The value of the weight of the nearest sample decreases but the weights of the more distant samples increase as the block size increases. Clustered samples carry less weight individually than isolated samples at the same distance. Samples can be screened by those lying between them and the target point.

Goovaerts (1997) raised several issues in relation to kriging weights. Kriging weights depend only on the shape of the semivariogram (*i.e.*, the relative nugget effect and anisotropy, correlation range), not on its global sill. Samples outside the correlation range may be assigned a non-zero weight, due to their contribution to the estimation of the trend component at points of interest. Samples may get negative weights when they are screened by a closer sample. Negative weights allow the kriging estimate to take values outside the range of the data, which is referred to as the non-convexity of the estimator and may yield unacceptable results such as negative estimates or estimated proportions larger than one. A larger nugget effect reduces the impact of distance of sample points to the point of interest and also reduces the screening effect. In the presence of a pure nugget effect, all weights are equal to $1/n$ and the kriging estimate then reverts to the arithmetic average of the data retained (*i.e.*, moving average).

There are a few ways to deal with the non-convexity problems (*i.e.*, the kriging estimate takes values outside the range of the data due to negative weights; Goovaerts 1997). The first is to force all weights to be positive. The next is to add to all the weights a constant (*i.e.*, the absolute value of the largest negative weight), and then reset the weights to sum to one. The third is to reset any faulty estimate to the nearest bound (*e.g.*, 0 if negative values are not acceptable, or 1 for excessive proportions). The last method is to impose constraints on the kriging estimates through the use of indicator constraint intervals.

Search Neighbourhood Window

Several aspects should be considered in selecting the search neighbourhood window (Goovaerts, 1997). The shape of the window is typically taken as a circle centred on the point being estimated. When the variation of the data is anisotropic, an ellipse with its major axis oriented along the direction of the maximum continuity should be used. If samples are clustered, the window should be split into equal angle sectors, but quadrants should be avoided when data are grided and the estimation grid is aligned with the sampling grid.

Although there are reasons for restricting the size of the window, one should avoid limiting a priori the maximum search distance to the correlation range of data (Goovaerts, 1997). In a sub-region where sample density is low, the search distance should be increased to retain enough data or simply retain *n* closest samples regardless their distance. The semivariogram distance $\gamma(x_0\text{-}x_i)$, instead of Euclidian distance $|x_0\text{-}x_i|$, should be used for data selection so that data are preferentially selected along the direction of maximum continuity.

In addition to size and orientation of the window, the minimum and maximum number of samples need to be specified for estimation (Goovaerts, 1997). The minimum needs to be equal at least to the number of constraints on kriging weights and should be larger where data are clustered. The maximum should be limited to depict local features of the attribute and should be larger to depict long-range structures. Isaaks and Srivastava (1989) have also discussed the search strategy.

Simple Kriging versus Ordinary Kriging

OK is usually preferred to SK because: 1) it requires neither knowledge nor stationarity of the mean over the region of interest; 2) OK allows one to account for local variation of the mean by limiting the domain of the stationarity of the mean to the local neighbourhood centred on the point being estimated, but SK requires the stationary mean of the whole region of interest; 3) OK estimates better follow the data fluctuations than SK estimates; and 4) OK estimates changes proportionally with the local data means (Goovaerts, 1997). Hence the OK with local search neighbourhood already accounts for trends (varying mean) in the values of the attribute (Goovaerts, 1997). However, OK requires a stationary mean of the local search window. This requirement of stationarity of OK and OCK has been relaxed to allow them to krige non-stationary data by using RK (Knotters *et al*., 1995), universal CK (Stein *et al*., 1988), and KED and UK (Verfaillie *et al*., 2006) if the statistical properties of the primary variable are not constant within the search window.

Ordinary Kriging versus Universal Kriging

OK and UK yield similar interpolating estimates, but quite different extrapolating estimates, depending on the trend fitted to the last few data values (Goovaerts, 1997). OK with local search neighbourhoods is preferred in interpolations because it provides results similar to UK estimates, but is easier to implement. In extrapolations, UK should be used whenever the attribute suggests a particular function form for extrapolating a trend fitted from the sampled data. UK may yield aberrant extrapolation estimates (*e.g.*, negative estimates depending on the trend fitted to the last few values; Goovaerts, 1997).

Kriging with External Drift versus Simple Kriging with Varying Local Means

Like UK, KED amounts to evaluating the regression coefficients from samples within each search window, estimating the trend component at all primary data points and at the point being estimated, and then performing SK on the corresponding residuals (Goovaerts, 1997). The KED estimator is similar to SKlm. KED and SKlm differ in their definition of the trend component. The trend coefficients are derived once and independently of the kriging system in SKlm, but are implicitly estimated through the kriging system within each search window in KED (Goovaerts, 1997).

Simple Kriging versus Simple Cokriging

The SK and SCK are compared according to Goovaerts (1997). SCK is theoretically better than SK because its error variance is always smaller than or equal to the error variance of SK. However, SCK is much more demanding than SK because of the additional modelling and computational requirements. SCK and SK produce identical estimates when: 1) the primary and secondary variables are uncorrelated; or 2) the primary and secondary variables are measured at the same locations and the cross covariance is proportional to the primary autocovariance. SK and SCK estimates are essentially the same in the isotopic case and the difference between estimates increases as the samples of secondary variables become more numerous than those of the primary variable. Cokriging improves over kriging only when the secondary variables are better sampled than the primary variable, or more accurately reflect the real world. CK is most effective when the covariate is highly correlated with the primary variable (Hartkamp *et al*., 1999).

Goovaerts (1997) discussed further the effects of a secondary variable. The contribution of the secondary variable to the SCK estimate should depend on: 1) correlation between the primary and secondary variables, 2) its pattern of spatial continuity, 3) the spatial configuration of the primary and secondary sample points, and 4) the sample density of each variable. The relative influence of the secondary variable can be measured by the ratio of the sums of absolute values of the secondary and primary data weights. The ratio increases exponentially as: 1) the correlation coefficient increases, 2) the relative nugget effect on the primary semivariogram model increases, 3) the secondary data points get further away from the primary data point and get closer to the point of interest, and 4) the sample size of the secondary variable increases. The secondary variable may screen the influence of the colocated primary data when both the primary and secondary variables are highly correlated and the secondary variable varies more continuously in space than the primary variable.

Ordinary Kriging versus Ordinary Cokriging

If both the primary and secondary variables are all measured at the same points then OCK will not produce estimates that are different from OK (Burrough and McDonnell, 1998).

Colocated Cokriging versus Cokriging

The CCK is valuable alternative to CK when the sample density is high for the secondary variables (Goovaerts, 1997). It avoids instability caused by highly redundant secondary data and is computationally fast. However, it requires: 1) the samples of secondary variables at all points being estimated; and 2) knowledge of the stationary means of the primary and secondary variables. CK and the computationally fast CCK give similar results.

Colocated Cokriging versus Kriging with External Drift

The CCK and KED use exhaustively sampled secondary information, but they differ in many aspects (Goovaerts, 1997). In CCK the colocated datum directly influences the primary cokriging estimate and CCK accounts for the global linear correlation between primary and secondary variables as captured by the semivariogram. In KED the secondary information provides information only about the primary trend of the point of interest and tends to influence strongly the estimate especially when the estimated slope of the local trend model is large. The influence of the residual covariance required by KED is not straightforward. Modelling direct and cross semivariograms in CCK is straightforward although computationally demanding.

Block Kriging versus Ordinary Kriging

The BK estimates vary more smoothly in space than OK estimates; and the smoothness increases with increasing size of the block (Goovaerts, 1997). BK smoothes out short-range variation of the attribute and can erase the artefact discontinuities near sample points. As such, BK is not an exact method. If the objective is to map large-scale features of an attribute, BK is preferred to point kriging (Goovaerts, 1997).

Limitations of Principal Component Kriging

Although PCK remains computationally less demanding than cokriging, it suffers from three major limitations (Goovaerts, 1997). The first is that only those data points where all variables are jointly measured can be considered. The next is that the cross correlation between principal components at $h \neq 0$ may not be negligible. The third is that the modelling of the principal components' semivariogram cannot capitalise on available ancillary information about the original variables.

Features of Regression Kriging

The RK is mathematically equivalent to the spatial interpolation method variously called UK and KED (Hengl *et al*., 2007). RK combines a regression of the primary variable on secondary variables with SK of the regression residuals, but UK and KED use secondary variables directly to solve the kriging weights. The advantage of RK is its ability to extend the method to a broader range of regression techniques such as GAM, and CART (Bishop and McBratney, 2001), and GLM (Gotway and Stroup, 1997; Pebesma, 2005) and to allow separate interpretation of the two interpolated components (Hengl *et al*., 2007).

Knotters *et al*. (1995) discussed the advantages and disadvantages of RK in comparison with CK. Three advantages were identified: 1) in RK, the relationship between the primary variable and the secondary variable can have any form and is physically interpretable, but CK does not use physically interpretable relationships and assumes a linear relationship; 2) for RK, only a model of the spatial correlation of the primary variable is required, but in CK with *m* variables, *m* generalised covariance functions and $m(m-1)/2$ generalised cross-covariance functions need to be estimated; and 3) RK is thus less computationally demanding and therefore is more efficient than CK. RK also considers the local trend within the search window by kriging non-stationary data. A disadvantage of RK is that assumptions are required about the errors of the values predicted by the regression model, namely: they are unsystematic, not autocorrelated and not correlated with the variable. Knotters *et al*. (1995) suggested that by adding variables to the regression model, thereby explaining a greater part of the variance, the assumption of the absence of autocorrelation of the errors will be satisfied more. We would also argue that if a GLM is used, such assumptions could be avoided. Several limitations of RK were also discussed by Hengl (2007), including data quality, sample size, reliable estimation of the covariance/correlation structure, extrapolation outside the sampled feature space, secondary variables with uneven relationships to the primary variable, and intermediate-scale modelling.

Splines versus Regression Kriging-D

It was claimed that regularised splines with tension and RK-D would yield very similar results; and the major difference is that the splines require a user specified smoothing parameter, while the smoothing is determined objectively in the kriging (Hengl, 2007).

Residual Maximum Likelihood-empirical Best Linear Unbiased Predictor versus Regression Kriging

There is a bias at long lags when the variogram of the residuals were estimated using RK-C (Lark *et al*., 2006). Such bias may be reduced but not removed when using RK-D. However, the bias in the variance for REML is very small and negligible by comparison with the bias for RK-C (Lark *et al*., 2006). Such bias will have two consequences: 1) underestimation of the overall variation of the random variable and 2) incorrect estimation of spatial structure (Lark *et al*., 2006). Therefore, REML-EBLUP was recommended over various RK types unless datasets are very large because REML-EBLUP is applicable only when the sample size is small (<200) (Minasny and McBratney, 2007).

Model-Based Kriging versus Other Kriging Methods

OK, lognormal OK, DK, and IK ignore the additional uncertainty caused by inferring the covariance structure from the data when they make predictions; but MBK does not suffer from this obvious disadvantage because it incorporates parameter uncertainty in a natural way into the predictions (Moyeed and Papritz, 2002). However, MBK is computationally demanding and not suitable for large datasets (> 300 samples) (Moyeed and Papritz, 2002). This limitation may be overcome as the computing power increases.

Simulation

Conditional bias (*i.e.*, an underestimation of large values and an overestimation of small values) dramatically affects the evaluation of the extent of an attribute. Simulated maps could correct for the conditional bias of estimated maps if they are applied (Goovaerts, 1997).

Chapter 4: Assessment Measures

In this chapter, performance assessment measures are reviewed. These measurements are for assessing the performance of: 1) the variogram models, 2) the spatial interpolation methods, and 3) the spatial interpolation methods for datasets with different sample sizes. Two new measurements are proposed for assessing the performance of the spatial interpolation methods for different variables. The measurements reviewed are those commonly used, so the list of the measurements is intentionally non-exhaustive.

4.1. Performance of Variogram Models

As discussed in Chapter 2, there are a number of variogram models that could be employed; and different variogram models may lead to different interpolations. Thus selecting an appropriate model to capture the features of the data is critical. The ratio of the square sum of deviance to the total sum of squares provides information on which model best fits the semivariance. If the model fits the semivariogram well, the ratio will be small, otherwise the ratio will approach 1 (Hartkamp *et al*., 1999). Cross-validation techniques could be used to choose the best semivariogram model from the candidates (such as spherical, exponential, and Gaussian). Cross-validation techniques could also be used to select an optimal search radius which minimises the kriging variance. The best-fit variogram model could also be determined by selecting the model having the lowest AIC (Erxleben *et al*., 2002).

The spatial structure of the data affects the performance of geostatistical interpolators. To test for anisotropy, the semivariogram needs to be determined in different directions. To ensure isotropy, the semivariogram model should be unaffected by the direction in which h is taken (Hartkamp *et al*., 1999). The variogram may show directional changes in different spatial scales, different semivariances, and different forms (Webster and Oliver, 2001).

The structural variance, which determines the variance due to spatial dependence explained by the variogram model, is calculated as the difference of total variance and nugget variance divided by the total variance (Hernandez-Stefanoni and Ponce-Hernandez, 2006). The relative "noisy" nature of the spatial variability is represented by the values of the nugget variance (Hernandez-Stefanoni and Ponce-Hernandez, 2006). The nugget variance may result from 1) the sampling error, and 2) the spatial dependence that may exist at finer scales than the minimum separation distance between samples (Hernandez-Stefanoni and Ponce-Hernandez, 2006).

The ratio of nugget to sill reflects the spatial heterogeneity of the data (Robertson *et* *al*., 1997; Wang *et al*., 2005). If the ratio is big, the spatial variation is mainly resulted from the random process and the measurement error is high, and if it is small, the variation is mainly due to the spatial structure.

4.2. Performance of Spatial Interpolation Methods

With the wide and increasing applications of the spatial interpolation methods, there is also a growing concern about their accuracy and precision (Hartkamp *et al*., 1999). As any other statistical modelling techniques, the spatial interpolation methods also produce a certain degree of errors associated with the estimation.

The statistics of the differences (absolute and squared) between the measured and predicted values at sampled points are often used as an indicator of the performance of an inexact method (for definition of exactness see section 3.1.2) (Burrough and McDonnell, 1998). Several error measurements have been proposed (Table 4.1). Commonly used error measurements include: mean error (ME), mean absolute error (MAE), mean squared error (MSE) and root mean squared error (RMSE). ME is used for determining the degree of bias in the estimates, often referred to as the bias (Isaaks and Srivastava, 1989) but it should be used cautiously as an indicator of accuracy because negative and positive estimates counteract each other and resultant ME tends to be lower than actual error (Nalder and Wein, 1998). RMSE provides a measure of the error size, but it is sensitive to outliers as it places a lot of weight on large errors (Hernandez-Stefanoni and Ponce-Hernandez, 2006). MSE suffers the same drawbacks as RMSE. Whereas MAE is less sensitive to extreme values (Willmott, 1982; Vicente-Serrano *et al*., 2003) and indicates the extent to which the estimate can be in error (Nalder and Wein, 1998). MAE and RMSE are argued to be similar measures, and they give estimates of the average error, but they do not provide information about the relative size of the average difference and the nature of differences comprising them (Willmott, 1982). Of course, we can also use cross-validation in together with these measurements to assess the performance of both exact and inexact methods.

Table 4.1. Measurements used to assess the performance of the spatial interpolation methods (Ahmed and De Marsily, 1987; Burrough and McDonnell, 1998; Hu *et al.*, 2004; Isaaks and Srivastava, 1989; Vicente-Serrano *et al.*, 2003).

Measurement	Definition*
Mean error (ME) or mean bias error (MBE)	$ME = \frac{1}{n}\sum_{i=1}^{n}(p_i - o_i)$
Mean absolute error (MAE)	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert p_i - o_i \rvert$
Mean square error (MSE)	$MSE = \frac{1}{n}\sum_{i=1}^{n}(p_i - o_i)^2$
Root mean square error (RMSE)	$RMSE = [\frac{1}{n}\sum_{i=1}^{n}(p_i - o_i)^2]^{1/2}$
Mean square reduced error (MSRE)	$MSRE = \frac{1}{n}\sum_{i=1}^{n}(p_i - o_i)^2 / s^2$
Mean standardised error (MSE2)	$MSE2 = \frac{1}{n}\sum_{i=1}^{n}(p_{si} - o_{si})$
Root mean square standardised error (RMSSE)	$RMSSE = [\frac{1}{n}\sum_{i=1}^{n}(p_{si} - o_{si})^2]^{1/2}$
Averaged standard error (ASE)	$ASE = [\frac{1}{n}\sum_{i=1}^{n}(p_i - (\sum_{i=1}^{n} p_i)/n)^2]^{1/2}$
Willmott's D	$D = 1 - \frac{\sum_{i=1}^{n}(p_i - o_i)^2}{\sum_{i=1}^{n}(\lvert p'_i \rvert + \lvert o'_i \rvert)^2}$
Ratio of the variance of estimated values to the variance of the observed values (RVar)	$RVar = \frac{Var[p]}{Var[o]}$
Model efficiency (EF)	$EF = 1 - \frac{\sum_{i=1}^{n}(p_i - o_i)^2}{\sum_{i=1}^{n}(\overline{o} + o_i)^2}$

* n: number of observations or samples; o: observed value; p: predicted or estimated values; o_s: standardised observed value; p_s: standardised predicted value; s: standard deviation of the estimation error; $\overline{o}$: mean of observed values; o'_i: o_i- $\overline{o}$; and p'_i: p_i- $\overline{o}$.

Hu *et al*. discussed several criteria for using error measurements to judge the performance of the spatial interpolation methods (Hu *et al*., 2004). If ME, MSE, and MSE2 are closer to zero, and RMSE is smaller, the better the model. ASE and RSME should be the same or close. If ASE>RSME, then the method overestimates the primary variable. If ASE<RSME, then the method underestimates the primary variable. RMSSE should be close to 1. If RMSSE>1, the method underestimates the primary variable, and if RMSSE<1, it overestimates the primary variable.

MSRE, also called studentised residuals for regression diagnostics in statistics (Venables and Ripley 2002) or standardised mean square error (SMSE; Martínez, 1996) , should approach one (Ahmed and De Marsily, 1987).

The index of agreement, or Willmott's D, scales with the magnitude of the variable, retains mean information, and does not amplify outliers (Willmott, 1982). If D is closer 1, the more accurate the method is considered to be (Vicente-Serrano *et al*., 2003).

The closer RVar is to 1, the better the ability of a spatial interpolation method to preserve the observed variance (Haberlandt, 2007).

Greenwood, Neeteson and Draycott (1985) cited by Vicente-Serrano *et al*. (2003) proposed an accuracy measurement known as model efficiency (EF). The closer EF to 1, the better the method. If EF is close to zero, it indicates that the mean value of the observations is more reliable than the estimations and the model has significant limitations (Vicente-Serrano *et al*., 2003).

RMSE and MAE are argued to be among the best overall measures of model performance as they summarise the mean difference in the units of observed and predicted values (Willmott, 1982).

The correlation between the observed values and predicted values, usually described by Pearson's product-moment correlation coefficient or coefficient of determination, is also a commonly used performance measurement. However, it is argued that it should not be used as a model performance measure because it is insufficient and often misleading (Willmott, 1981; Willmott, 1982).

4.3. Performance of Spatial Interpolation Method for Datasets with Different Sample Sizes

The performance of a spatial interpolation method for two different sizes of datasets was compared in terms of the nugget effects of the variogram (Hartkamp *et al*., 1999). As an indication of measurement accuracy, if the nugget of the large dataset is larger than the nugget of the small dataset, then the large dataset is probably less accurate, providing that 1) for each variogram, the number of lags and the lag distance are kept constant; and 2) the model type fitted through the variogram is also the same for each dataset. Thus the nugget difference is independent of model, number of lags, and lag distance. The relative nugget difference (RND) can be calculated as:

$$RND = (\frac{Nugget_{large-dataset} - Nugget_{small-dataset}}{Nugget_{large-dataset}}) * 100\% \quad (14)$$

4.4. Performance of Spatial Interpolation Methods for Different Variables

All of the measures listed in section 4.2 have been developed to assess the performance of the spatial interpolation methods for individual primary variables. The magnitude of these measurements depends on the unit of the primary variable. In some cases, it is necessary to compare the performance of the spatial interpolation methods among different studies, in which the primary variables are in different measurement units or scales. It is impossible to use the scale-dependent measurements for such comparisons, so new types of measurements are needed to compare results for variables with different measurement units.

Here we propose two new measurements that remove the effect of measurement units and they are not sensitive to the changes in measurement unit or scale. The first is relative mean absolute error (RMAE) that is given as**:**

$$RMAE = \frac{1}{n}\sum_{i=1}^{n} |(p_i - o_i)/o_i| \quad (15)$$

RMAE can be understood as a relative error in predictions. And the second is relative root mean square error (RRMSE), as follows**:**

$$RRMSE = [\frac{1}{n}\sum_{i=1}^{n}([p_i - o_i]/o_i)^2]^{1/2} \quad (16)$$

However, in most published studies, the information required to calculate RMAE and RRMSE is not available. To overcome this problem, RMAE and RRMSE are modified by using MAE/mean (*i.e.*, mean of the validation dataset) and RMSE/mean instead. The second one is also called standardised RMSE (Haberlandt, 2007). The mean of the validation dataset is, however, often not reported in publications. They are further modified by using the mean of the dataset for estimation that is more frequently available in publications. These new measurements provide effective tools to compare the results of various variables from different studies and from various disciplines.

Chapter 5: Comparison of Spatial Interpolation Methods Applied to Various Disciplines

The spatial interpolation methods have been applied to many disciplines such as mining engineering (Journel and Huijbregts, 1978) and environmental sciences (Burrough and McDonnell, 1998; Goovaerts, 1997; Webster and Oliver, 2001). On the basis of a bibliographic research (Zhou *et al*., 2007), it was found that the top 10 fields that employ geostatistics are: 1) geosciences, 2) water resources, 3) environmental sciences, 4) agriculture or soil sciences, 5) mathematics, 6) statistics and probability, 7) ecology, 8) civil engineering, 9) petroleum engineering and 10) limnology. Our focus in this review is on environmental sciences, including: water resources, meteorology, ecology, agriculture and soil science, and marine environmental science. Examples of the application of different spatial interpolation methods in each of these disciplines are provided in this chapter. The focus is further narrowed down on comparative studies that compared the performance of the spatial interpolation methods.

Sample density, sample size and spatial distribution of samples as discussed in the next chapter are important in assessing the performance of the spatial interpolation methods. However, such information is often not clearly stated or unavailable in publications. In this review, such information is provided whenever it is available in the references, and information on the area of region studied, resolution, sampling strategy or experimental design is also provided. The spatial interpolation methods compared, sampling design, sample size, area of region interested, and results are summarised for each of 51 comparative studies, namely: 16 studies in meteorology and water resources, one study in ecology, 25 studies in agriculture and soil sciences, four studies in marine environmental science and five in other disciplines (Appendix A).

5.1. Comparison by Studies

The 51 comparative studies reviewed above are summarised in Table 5.1. The spatial interpolation methods compared are listed and the results are briefly discussed. The frequency of each spatial interpolation method compared in these 51 studies is summarised in Table 5.2. Given that the times of recommendation of a spatial interpolation method in Table 5.2 depend on the methods compared, it should be assessed in together with all spatial interpolation methods compared (Table 5.1).

These 51 comparative studies illustrate the following major characteristics:

1. The spatial interpolation methods have been applied widely in environmental

sciences, with about 62 various methods including combined methods employed;

2. Different studies have compared a suite of different methods, which makes it difficult to draw general conclusions. However, by numbers, OK, IDW including IDS and OCK are the most commonly compared methods (Table 5.2);
3. In general, kriging methods perform better than non-geostatistical methods, with only a few exceptions;
4. RK, KED and OCK frequently performed better than other methods, and IDS TPS and LM occasionally outperformed other methods;
5. GIDS and other highlighted methods in Table 5.2 are worthy of attention because of their good performance; and
6. Stratification may improve the estimation.

Table 5.1. Summary of the 51 reviewed comparative studies.

No	Reference	Discipline	Methods compared	Result
1	Hartkamp *et al.*, 1999	Meteorology and Water resources	IDW, TPS & OCK	No difference, but TPS preferred.
2	Erxleben *et al.*, 2002		IDW, OK, RK-C, OCK, CART with OK & CART with OCK	CART with OK and CART with OCK more accurate.
3	Martínez-Cob, 1996		OK, OCK & RK-C	OCK more accurate.
4	Vicente-Serrano *et al.*, 2003		TSA, LM, NN, IDW, splines, SK, OK, BK, OK, UK, OCK, LM with IDS & splines with LM	Kriging and LM more accurate.
5	Haberlandt, 2007		NN, IDS, OK, OIK, KED & IKED	KED the best.
6	Collins and Bolstad, 1996		IDS, OIDW, splines, LM, TSA, LR, kriging & CK	LM the best.
7	Jarvis and Stuart, 2001		LM-IDW, TSA, RK-C & partial TPS with secondary variables	Partial TPS with secondary variables the best.
8	Jef *et al.*, 2006		RK-C, IDW (with distance power 4) & LM with IDW	RK-C the best.
9	Goovaerts, 2000		SKlm, KED, OCCK, LM, NN, IDS & OK	SKlm the best.
10	Nalder and Wein, 1998		GIDS, IDS, NN, CK, OK, RK-C & UK	GIDS preferred.
11	Mardikis *et al.*, 2005		GIDS, IDS, OK & RK-C	GIDS the best.
12	Naoum and Tsanis, 2004		Splines, IDW, NN, LM & kriging	Kriging preferred.
13	(Lin and Chen, 2004		RBFN, improved RBFN & OK	Improved RBFN the best.
14	Sun, 1998		MWRCK, CK & LSZ	LSZ the best.
15	Li *et al.*, 2005		IDS, OK, OCK & OK combined with LR	OK combined with LR the best.
16	Hosseini *et al.*, 1993		OK, UK, TSA, IDW & AK	OK preferred.
17	Hernandez-Stefanoni and Ponce-Hernandez, 2006	Ecology	OK, OCK, IDS, StOK, StOCK, StIDW & Cl	StOK the best.
18	Schloeder *et al.*, 2001	Agriculture and soil science	OK, IDW & TPS	OK and IDW better.
19	Wang *et al.*, 2005		TSA-OK & TSA-OCK	TSA-OCK better.
20	Voltz and Webster, 1990		SK, StSK, Cl & cubic spline	StSK the best.
21	Brus *et al.*, 1996		Cl, GM, IDS, OK, NN, IDW-0, TPS & their combination with soil strata	StOK the best.
22	Van Kuilenburg *et al.*, 1982		NN, IDS & OK	OK preferred.
23	Goovaerts, 1997		OCK, SCK, SOCK & OCCK	OCK and SOCK better.
24	Goovaerts, 1997		KED & SKlm	Similar
25	Goovaerts, 1997		OIK & OICK	Similar
26	Hu *et al.*, 2004		SK, OK, lognormal kriging, UK, DK & IDW	UK the best.

No	Reference	Discipline	Methods compared	Result
27	Moyeed and Papritz, 2002	Agriculture and soil science	OK, lognormal OK, DK, IK & MBK	Similar
28	Laslett *et al*., 1987		TPS, OK, global means and medians, NN, IDW-0, IDS, AK, NaN & TSA	TPS and OK better.
29	Laslett and McBratney, 1990		NN, TPS, AK, SK? & REML UK	REML UK the best.
30	Laslett, 1994		Cubic splines & SK	SK better.
31	Knotters *et al*., 1995		OK, OCK & RK-A	RK-A the best.
32	Bishop and McBratney, 2001		GAM, LM, CART, OK, KED, RK-F & RK-C	KED the best.
33	Odeh *et al*., 1994		LM, OK, UK, OCK, RK-A, and RK-B	RK the best.
34	Odeh *et al*., 1995		LM, OK, UK, OCK, RK-A, RK-B & RK-C	RK-C the best.
35	Meul and Van Meirvenne, 2003		OK, UK, SKlm & OCK	UK + OCK the best.
36	Minasny and McBratney, 2007		REML-EBLUP, OK & RK-C	RK-C recommended.
37	Li *et al*., 2007		OK, OCK and RK-E	RK-E better.
38	Bourennane *et al*., 2000		KED & LM	KED better.
39	Ahmed and De Marsily, 1987		OCK, KED, RK-A & RK-B	OCK and RK-A preferred.
40	Wu *et al*., 2006		OK & OCK	OCK better.
41	Gotway *et al*., 1996		OK & IDW	OK better.
42	Kravchenko and Bullock, 1999		OK, lognormal OK & IDW	Lognormal OK better.
43	ICES, 2005	Marine environmental science	OK & KED	KED better.
44	Verfaillie *et al*., 2006		OK, KED & LM	KED the best.
45	Rivoirard and Wieland, 2001		KED & OK	KED better.
46	Ruddick, 2006		OK, OCK, IDW, NN & T2R	Similar.
47	Boufassa and Armstrong, 1989	Other fields	OK, lognormal OK, SK, lognormal SK, disjunctive OK & disjunctive SK	SK and OK recommended.
48	Isaaks and Srivastava, 1989		OK, IDS, TIN & NN	OK the best.
49	Zimmerman *et al*., 1999		OK, UK & IDS	OK preferred.
50	Weber and Englund, 1992		OK, SK, lognormal OK, rank OK, global mean, IDW, TSA & Projected Slope	IDS better.
51	Puente and Bras, 1986		UK, DK & local mean estimator	UK better.

Table 5.2. Frequency of the spatial interpolation methods compared and the number of times the method was recommended in the 51 reviewed comparative studies. Methods with 100% rate of recommendation are highlighted.

Method	Frequency	Recommendations	Method	Frequency	Recommendations
OK	37	8	disjunctive OK	1	0
IDW	18	1	disjunctive SK	1	0
OCK	14	4	GAM	1	0
LM	13	2	IK	1	0
IDS	12	1	IKED	1	0
NN	11	0	**improved RBFN**	1	1
KED	9	6	local mean	1	0
RK-C	9	3	lognormal SK	1	0
UK	8	2	LR	1	0
SK	7	2	**LSZ**	1	1
TSA	7	0	MWRCK	1	0
Splines	6	0	NaN	1	0
TPS	6	3	OICK	1	0
CK	4	0	**OK combined with LR**	1	1
lognormal OK	4	1	Projected Slope	1	0
RK-A	4	2	RBFN	1	0
AK	3	0	**REML UK**	1	1
Cl	3	0	REML-EBLUP	1	0
DK	3	0	**RK-E**	1	1
GM	3	0	RK-F	1	0
kriging	3	2	**SOCK**	1	1
RK-B	3	0	StIDS	1	0
SKlm	3	1	StIDW	1	0
BK	2	0	StOCK-	1	0
GIDS	2	2	**StSK**	1	1
OCCK	2	0	StTPS	1	0
OIK	2	0	T2R	1	0
StOK	2	2	TIN	1	0
CART with OCK	1	1	**TSA-OCK**	1	1
CART	1	0	**UK+OCK**	1	1
CART with OK	1	1			

5.2. Comparison by Variables

Of the 51 comparative studies from the various disciplines contained in this review, 17 were selected for a comparative analysis. The criteria for the selection are that the following information should be reported: 1) the mean and CV of the primary variable for either the estimation dataset or validation dataset, 2) the sample size for the estimation and validation datasets, 3) the area of the region studied, and 4) appropriate accuracy measurements of the spatial interpolation methods (*i.e.*, MAE and/or RMSE or MSE). Of course, the spatial interpolation methods need to be named properly, appropriately referenced, or clearly described. The information has been summarised in Appendix B. In the 17 studies, there were 33 methods and their variants and 77 cases (*i.e.*, variables). For some methods, the method and its variants have to be grouped into one method. Taking IDW as example, some studies clearly stated the power of distance, but in others no such information was provided, so IDW and its variants are treated as a single method. However, information on their variants is provided in the appendix for those interested.

In this section, the frequency and accuracy of the spatial interpolation methods are discussed. The performance of various methods is then further compared in relation to sampling density, variation in the data, and sampling design in the next chapter, based on the information provided by the 17 selected studies.

5.2.1. Frequency of the Spatial Interpolation Methods Compared

The frequency with which the individual spatial interpolation method was compared varies considerably between methods in the 17 comparative studies (Fig. 5.1). The spatial interpolation methods can be divided into four groups in terms of their frequency. The first group contains the most frequently compared methods with a frequency > 30 which are OK, IDW, IDS and TPS. The high frequency of IDW and TPS was mainly because their variants were all counted individually, thus increased the number of occurrences of these two methods (see Appendix B). The next group contains frequently compared methods with a frequency between 20 and 30 which are two methods, OCK and RK-C. The third group includes GIDS, IDW-0, LM, TSA and UK that were less frequently compared (with a frequency between 8 and 15). The last group contains the remaining methods that were occasionally compared (with a frequency < 8).

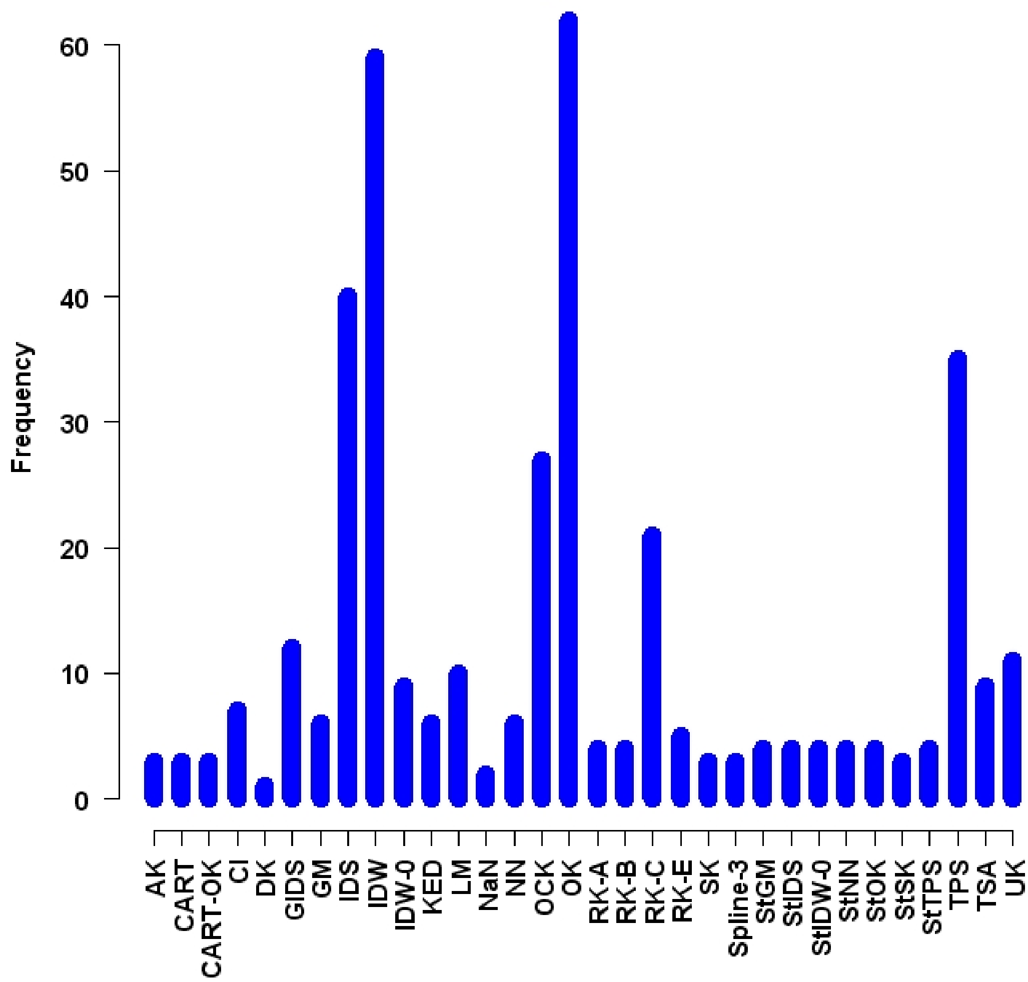


Figure 5.1. The frequency of 33 spatial interpolation methods compared in the 17 reviewed comparative studies.

5.2.2. Performance of the Spatial Interpolation Methods Compared

The performance of the spatial interpolation methods compared exhibits dramatic variation in terms of RMAE and RRMSE (Figs. 5.2 and 5.3). The RMAE values of some methods such as DK, KED, LM, are missing because the studies reviewed did not report this information. Of the four most frequently compared methods, OK is the most accurate. In the second group, RK-C performed better than OCK and both of these methods are more accurate than the four most frequently compared methods in group 1. Likewise, GIDS performed better than other less frequently compared methods in group 3 and is also more accurate than all methods in group 1 and 2. For the occasionally used methods (group 4), the results are not reliable due the small number of times of application. In general, RK-C, OCK, KED and GIDS are the best performing (*i.e.* the most accurate) methods.

However, these conclusions are only based on the results from 77 cases in the 17 comparative studies. Other methods might display similar features, but were unfortunately not compared in this review due to the lack of relevant information for appropriate comparison between different variables. Moreover, some comparative studies may have been missed in this review because only 51 comparative studies are assessed but there are 2,866 publications identified by the Institute for Scientific Information between 1967 to 2005 in geostatistics (Zhou *et al*., 2007). Nonetheless, the most influential comparative studies are believed to have been included in this review.

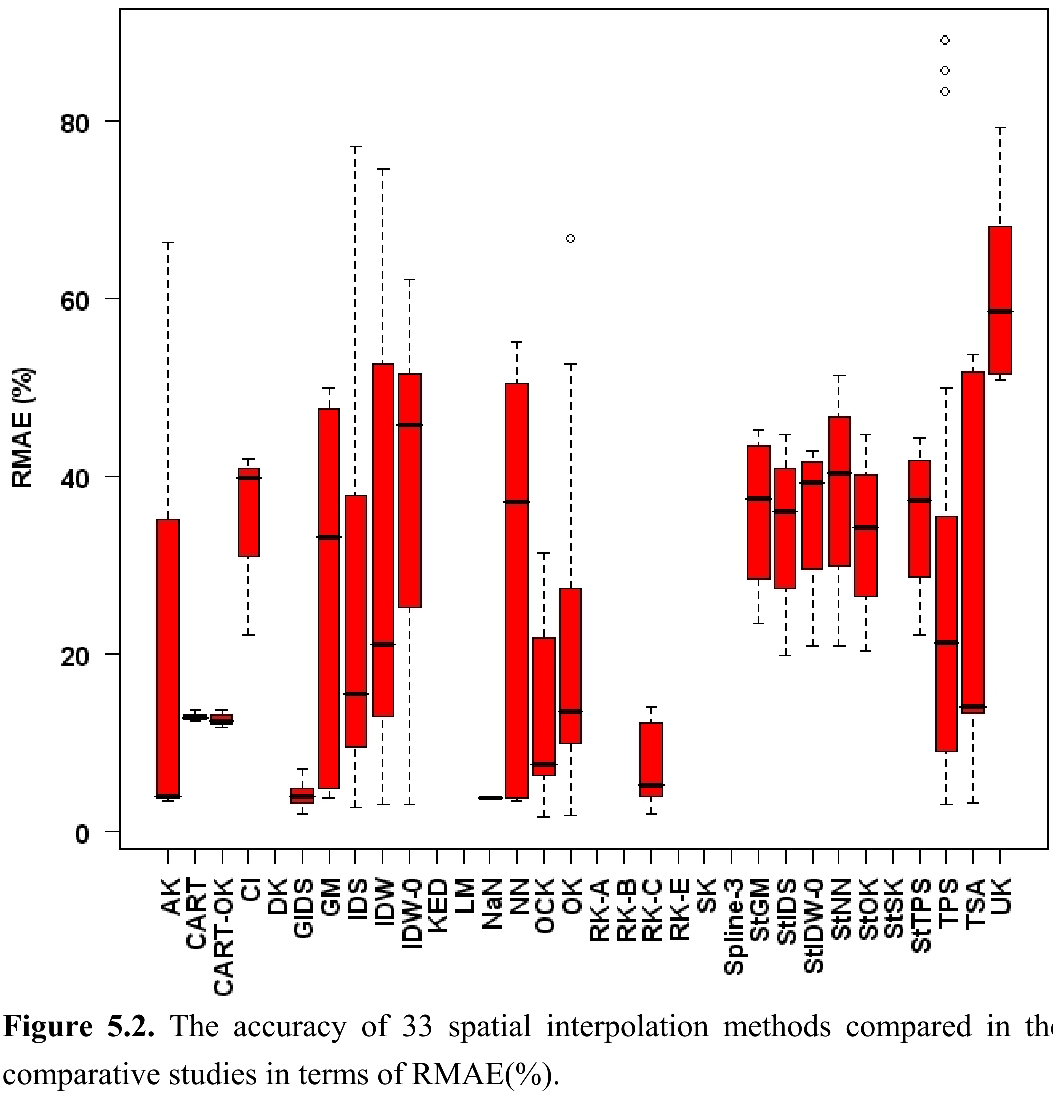


Figure 5.2. The accuracy of 33 spatial interpolation methods compared in the 17 comparative studies in terms of RMAE(%).

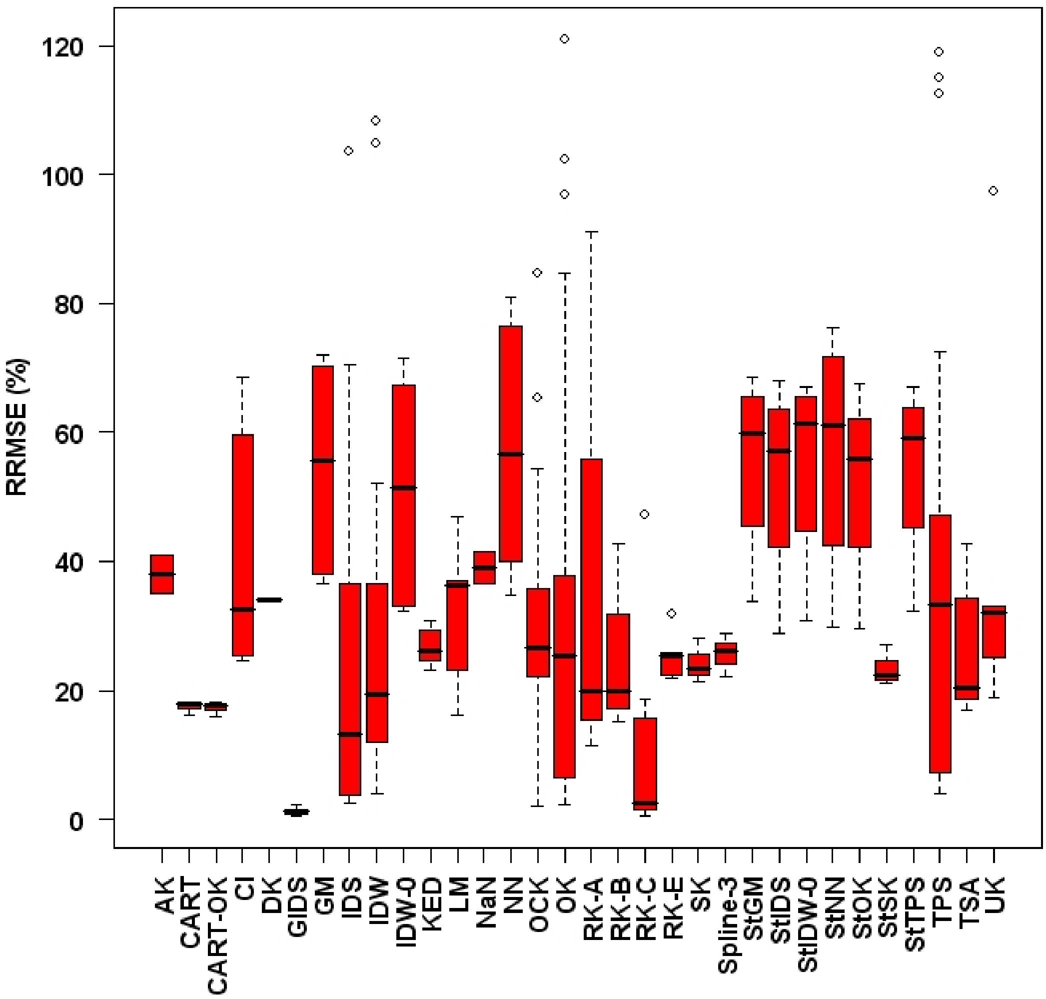


Figure 5.3. The accuracy of 33 spatial interpolation methods compared in the 17 comparative studies in terms of RRMSE(%).

5.3. Complicating and Confounding Factors

Several complicating and confounding factors were encountered in this review that may have some bearing on the outcomes. Information about the study region, experimental design and primary variable were missing at times. The measurements of the performance of the spatial interpolation methods varied between studies. Occasionally the methods used for interpolation were not clearly or adequately described or referenced. For example, some studies mentioned the use of kriging or cokriging. This is not sufficient because there are many different kriging methods and more than one cokriging method. All these factors make it difficult to compare the performance of the spatial interpolation methods using results from the published studies, consequently preventing any possible generalisation of the observed patterns. Only 5 out of 16 studies in meteorology and water resources, and 12 of 25 studies in soil science provided appropriate information for possible comparison between different variables and studies. All of the studies reviewed in the other disciplines failed to report relevant statistics for further comparative research between different variables.

Here we would recommend that future studies should report relevant information clearly in their publications, including: the area of region studied; experimental and sampling design, particularly the sample size of datasets for estimation and validation; summary statistics of the primary variable for both estimation and validation datasets; and appropriate references or descriptions of the spatial interpolation methods used. The measurements should include at least MAE or MSE for comparing the results of different variables. Correlation coefficient has been used in many studies as a measurement of the performance of the spatial interpolation methods. However, as discussed in section 4.2, it is often misleading and it should be either avoided or extreme care should be taken in using it.

In summary, the spatial interpolation methods have been applied in many disciplines. Although some methods perform better than others, there is no consistent pattern in the performance observed and thus no definite conclusion could be drawn on which method is the best or most appropriate. It is clear that some methods are only applicable to a certain types of data. The performance of a spatial interpolation method depends not only on the features of the method itself, but also on other factors such as the nature and quality of the data. These factors are discussed in the next chapter. It was even argued that improvements in prediction do not rely on more sophisticated methods, but rather on gathering more useful and high quality data (Minasny and McBratney, 2007).

Chapter 6: Factors Affecting the Performance of Spatial Interpolation Methods

In this chapter, several factors that affect the performance of the spatial interpolation methods are discussed. The impacts of sampling density, variation in the data, sampling design and stratification on the estimation of the spatial interpolation methods are quantified using data from 77 cases in the 17 comparative studies (Appendix B). The variability, sampling density, detectability and other properties of some common environmental variables have been summarised by Hengl (2007).

6.1. Sampling Design and Sample Spatial Distribution

6.1.1. Data Density

Data density plays a significant role in the performance of the spatial interpolation methods. The following sections discuss the effects of data density on the performance of the spatial interpolation methods.

High Density

When data density is high, most methods produce similar results (Burrough and McDonnell, 1998). It was found that kriging does not show significantly greater improvement in prediction than simpler methods, such as IDS and NN for high-density networks (*i.e.*, 13 rain gauges over a 35 km^2 region) (Dirks *et al*., 1998). Bregt (1992), cited in (Brus *et al*., 1996), compared local mean, global mean, IDW and kriging at several grid densities ranging from 8 to 200 samples per km^2 for the depth to the pyritic layer and found no statistically significant differences between these methods at any density. Little difference was also found in the performance of OK, UK, UK with a linear drift, IDS and TSA for a intensively sampled region, however the interpolated surfaces were very different, resulting a preference for OK (Hosseini *et al*., 1993).

Using datasets of regularly spaced and high density samples, Gotway *et al*. (1996) found that the use of wider sample spacings greatly reduced the information in the resultant maps, although the sample density was still relatively high.

Low Density

When data are sparse, the underlying assumptions about the variation among samples may differ and the choice of a spatial interpolation method and parameters may become critical (Burrough and McDonnell, 1998; Hartkamp *et al*., 1999). The performance of the spatial interpolation methods is better when the sample density is greater (Englund *et al*., 1992; Isaaks and Srivastava, 1989; Stahl *et al*., 2006). However, it is claimed that the accuracy of regression modelling is not really dependent on the sampling density, but rather on how well the data are sampled and how significant the correlation is between the primary variable and secondary variable(s) (Hengl, 2007).

Sample size also affects the predicted error. It was found that with small samples, both UK and DK may dramatically over- or under-predict the predicted estimation error (Puente and Bras, 1986). This suggests that such predicted errors should not be used in an absolute sense, but as a relative measure of spatial estimation accuracy (Puente and Bras, 1986). In addition, it is found that the smoothing of the estimations (or map) increased at lower sample densities (Goovaerts, 1997). Issues relating to sample size are further discussed below.

6.1.2. Sample Spatial Distribution

Sample spatial distribution may affect the performance of the spatial interpolation methods. Splines performed much better when dense, regularly-spaced data were available, but not for irregular-spaced data (Collins and Bolstad, 1996). For irregularly-spaced data, the interpolated map was more variable where sample density was high than where it was low, which may result in structures that are pure artefacts of the data configuration; and one potential solution is to use simulation algorithms instead of kriging algorithms (Goovaerts, 1997). In contrast, sample patterns (*i.e.*, random, cellular stratified, and regular grid) were found not to be significant in determining the performance of OK (Englund *et al*., 1992).

Sample clustering affects the accuracy of the estimations and the effects may also depend on the spatial interpolation methods. High clustering reduced the correlation coefficient between the observed and estimated values for all four methods studied, OK, IDS, TIN and NN; reduced the MAE for OK, TIN and NN and increased the MAE for IDS; reduced the MSE for OK, NN, increased the MSE for IDS, while had little influence on TIN (Isaaks and Srivastava, 1989). SK outperformed cubic splines if the sample points were highly clustered (Laslett, 1994). In addition, sample clustering reduced the accuracy of all methods tested (*i.e.*, OK, UK, IDS) (Zimmerman *et al*., 1999).

While spatial scale, relative spatial density and distribution of samples can be determinant factors on the performance of the spatial interpolation methods (Collins and Bolstad, 1996), other relevant factors may also be important. For example, altitudinal and seasonal changes in data have been shown to play a significant role in predictions (Stahl *et al*., 2006). Where temporal scales are short, preliminary data analyses are especially important to determine the suitability of a particular spatial interpolation method (Collins and Bolstad, 1996).

6.1.3. Surface Type

The surface type may play a significant role in the performance of the spatial interpolation methods. The variability in the surface tremendously increases the estimation error of the spatial interpolation methods; and estimation error consistently increases with an increasing rate as sample size decreases (MacEachren and Davidson, 1987). It has also been found that the performance of the spatial interpolation methods decreased with increasing variability of the surface (Zimmerman *et al*., 1999). Distinct and sharp spatial changes, like changing soil types across a region, may also cause problems with the estimations (Stein *et al*., 1988; Voltz and Webster, 1990).

6.1.4. Sample Size, Sampling Design and Variogram

Sample size and sampling configuration or design affect the reliability of the variogram. Generally, if the sample size is <50, the variograms derived are often erratic with little or no evident spatial structure (Webster and Oliver, 2001). The larger the sample size from which the variogram is computed, the more precisely is it estimated, although the precision is unknown in most instances (Brus and de Gruijter, 1994; Webster and Oliver, 2001). If the sample size is too small, a noisy variogram would be generated (Burrough and McDonnell, 1998).

Sample spacing must relate to the scale or scales of variation in a region, otherwise samples might be too sparsely spaced to identify correlation and could result in a pure nugget (Webster and Oliver, 2001). In such cases, the accuracy of the estimation could be reduced, as evidenced by the findings in Gotway *et al*. (1996) that the use of wider sampling spacings greatly reduced the information in the resultant maps. In addition, the smoothness of the estimations (or map) may increase with the relative nugget effect (Goovaerts, 1997).

The spatial structure of the data may also affect the sample size and variogram. For data with a short range of variograms, intensive sampling with a large proportion of clustered points is required; and conversely for variables with a long range, fewer and more evenly spaced samples are required (Marchant and Lark, 2006). Variogram is also sensitive to sample clustering, particularly when it is combined with a proportional effect that is a form of heteroscedasticity where the local mean and local variance of data are related (Goovaerts, 1997).

The number of pairs of samples at each lag is an important factor that needs to be considered in modelling the variogram. A rule of thumb, as suggested by Burrough and McDonnell (1998) is that at least 50-100 samples are necessary to achieve a stable variogram. Alternatively, 30-50 pairs of samples with the lag distance less than half of the dimension of sampled region are required to achieve the same result (Journel and Huijbregts, 1978). For REML variograms, 50 samples may be adequate (Kerry and Oliver, 2007). Even a sample size of 28 has been suggested for kriging and CK in a case study (Chang *et al*., 1998). Another rule of thumb is that the product of the lag interval distance and the number of lags should not exceed half of the largest dimension of the region of interest (Verfaillie *et al*., 2006). In addition, Burrough and McDonnell (1998) discussed some issues regarding how to use variograms to optimise the sampling so as to improve the overall estimations.

6.1.5. Sample Size and Spatial Interpolation Methods

The impacts of sample size on the estimation depend on the spatial interpolation methods. On the basis of the comparison of SK (incorrectly termed OK in the study, but in fact it is SK as it used global mean, see page 14 in Hengl 2007) and RK-D on two datasets with sample sizes of 222 and 2251 respectively, it was found that RK-D performed better than SK in terms of the level of detail and accuracy, and RK-D (222) even performed better than SK (2251) (Hengl, 2007). It was suggested that future studies should focus more on the quality of sampling and on quality of auxiliary environmental predictors, rather than on making more observations (Hengl, 2007). Findings in this research imply that the effects of sample size on the estimations also depend on the spatial interpolation methods. In practice, we believe there is a threshold beyond which any increase in sample size does not improve much the accuracy of the estimations; otherwise sample size is still a critical factor. Other factors like variance inherited in the data also play a significant role (as discussed below). Care should be taken in applying this suggestion in future research.

OK, OCK and RK-E were compared for several sample sizes that are 40, 70, 100, 130 and 160 (Li *et al*., 2007). The results showed that as sample size increased, the performance of all three methods increased, with exceptions that OK and OCK were more accurate when sample size was 70 than when sample size was 100, RK-E was less accurate when sample was 160 than when sample size was 130, 100 and 70 in terms of RMSE. A similar result was observed by Wang *et al*. (2005) for TSA-OK and TSA-OCK.

KED and LM were applied for sample densities of 40, 50, 75, 100, 125 and 150 (Bourennane *et al*., 2000). The results revealed that despite a couple of anomalies, generally KED performed better when the sample size increased. The performance of

LM remained largely stable across all the sample sizes, which implies that 40 samples provided sufficient information or there is no useful information contained in the extra samples for linear model.

The exceptions found in these studies imply that factors other than sample size may play a major role in determining the performance of a spatial interpolator. It is likely that in these cases, other properties of the data, such as spatial distribution and spatial structure, also influenced the performance. Notwithstanding these additional factors, the results of these studies into the effects of sample size suggest that its effect on the performance of the spatial interpolators depends largely on the features of the spatial interpolators themselves. Therefore, there is no definite conclusion on the relation of sample size and the performance of the spatial interpolators.

6.2. Data Quality

Five major factors relevant to the quality of the data are discussed in this section: distribution, isotropism and anisotropism, variance and range, accuracy, spatial correlation and other factors, and secondary variables. The sources of errors in spatial continuous data and factors affecting the reliability of spatial continuous data have been discussed in Burrough and McDonnell (1998).

6.2.1. Distribution

Data normality can influence the estimation of certain spatial interpolation methods that assume that the input data are distributed normally about their mean. Data normality can be tested using such as the Kolgorov-Smirnov test. If this assumption is not met, log transformation is commonly applied, thus resulting in lognormal methods (*e.g.*, lognormal kriging; Cressie, 1993). The predictions are then transformed back to the original scale by a marginally unbiased back transformation proposed by Cressie (1993). However, back-transforming the estimated values can be problematic because exponentiation tends to exaggerate any interpolation-related error (Goovaerts, 1997). Other transformation functions may also be used to achieve the normality, resulting in trans-Gaussian kriging and multi-Gaussian kriging (Cressie, 1993). Rank and normal score transformation could also be applied prior to kriging (Rossi *et al*., 1992; Weber and Englund, 1992; Wu *et al*., 2006). In addition, the prediction error may also be used to determine whether the data should be transformed (Nalder and Wein, 1998).

Wu *et al*. (2006) found that data transformation of highly skewed data generally improved the estimations by OK and OCK, especially for low concentrations of zinc, but the differences among normal score, log-normal and rank-order transformations were relatively small for OCK. Kravchenko and Bullock (1999) also found a similar result that log-transformation generally improved the performance of OK. OK failed to model the conditional distribution of the marginally skewed data, while the nonlinear methods modelled the conditional distribution with similar success (Moyeed and Papritz, 2002). In contrast, log transformation was found to have little effect on the performance of OK (Moyeed and Papritz, 2002), or even reduce the accuracy of OK prediction (Weber and Englund, 1992).

6.2.2. Isotropism and Anisotropism

Isotropism of data is assumed for kriging methods. Data may display evidence of anisotropism, which should be considered in the modelling; otherwise biased estimation may result. However, in some cases, the anisotropism could be ignored to simplify model fitting and to maintain some consistency between the semivariograms in the multivariate model (Martínez-Cob, 1996). Conditions that allow for this are: 1) anisotropism is not evident with the specified search distance; 2) the secondary variable and primary variable are colocated, thus the influence of surrounding values would be small, so anisotropy would make little difference; and 3) the directions of maximum and minimum spatial variability for the different variables did not coincide. A similar result was also observed by Haberlandt (2007). It was found that the impact of the semivariogram on interpolation performance was not great because no significant differences could be found in prediction performance between isotropic and anisotropic variograms, although anisotropy was clearly present in the data. The best estimations were obtained using an automatic fitting procedure with isotropic variograms.

6.2.3. Variance and Range

The variance of the data affects the performance of the spatial interpolators and the resultant predictions. The performance of the spatial interpolation methods decreased rapidly when the coefficient of variation (CV) increased (Martínez-Cob, 1996; Schloeder *et al*., 2001).

It was also found that the variance and range of the data can influence the performance and choice of a spatial interpolation technique after comparing eight spatial interpolators across two regions for two temperature variables (maximum and minimum) at three temporal scales (Collins and Bolstad, 1996). When temperature variances were large, the performance of all spatial interpolation techniques suffered, which means that increasing temperature variance negatively affected the performance of the spatial interpolators. As the temperature range increased, MAE values across all spatial interpolators also increased significantly. Temporal scale also affected the choice of a spatial interpolator as temperature range, temperature variance, and temperature correlation with elevation, all changed with temporal scale (Collins and Bolstad, 1996).

The chosen sampling scheme also affects the performance of the spatial interpolation methods through the variation in the data. Data should be collected at a range of separations to capture changes in the scales of the variation (Laslett, 1994).

6.2.4. Accuracy

Data accuracy is an important factor influencing the estimations of the spatial interpolation methods. Where the data are not representative of the surface being modelled, it may result in interpolation biases (Collins and Bolstad, 1996). Where sample elevations are not representative of regional elevations, care must be taken in comparing observed and interpolated data. Data noise can negatively affect the performance of the spatial interpolation methods (*i.e.*, OK, UK and IDS in Zimmerman *et al*., 1999; NaN in Webster and Oliver, 2001). When data are too noisy, a pure nugget effect is produced in the variogram and the resultant interpolation is not sensible (Burrough and McDonnell, 1998). In contrast, sampling precision (*i.e.*, zero error and high-level normally distributed error with a relative standard deviation of 32% of the true value) was found not to be significant in determining the performance of a spatial interpolator, being OK in this case (Englund *et al*., 1992)

Outliers affect the performance of the spatial interpolation methods and interact with sampling schemes. The variogram is sensitive to outliers and to extreme values (Webster and Oliver, 2001). Exceptionally large or small values will distort the average as evident from its definition (*i.e.*, Equation 6 and 7). This effect depends on the location of the data point in the region and also on the spatial pattern of data (Webster and Oliver, 2001). All outliers must be regarded with suspicion and investigated. Outliers should be removed if they are believed to not belong to the population and strongly skewed distributions need to be transformed to approximate normal before conducting geostatistical analyses (Webster and Oliver, 2001). For example, removing outliers resulted in considerable improvement in the performance of the spatial interpolation methods, particularly when additional samples were included to allow estimation of short-range variation (Laslett and McBratney, 1990).

6.2.5. Spatial Correlation and Other Factors

Spatial correlation in samples is also essential for reliable estimation. The performance of OK, UK and IDS was negatively affected when the spatial correlation between samples decreased (Zimmerman *et al*., 1999).

The performance of different spatial interpolation methods changes with the variable estimated. It was found that the best method varied as a function of the region and the spatial scale required for estimation (Vicente-Serrano *et al*., 2003). Accuracy was lower in regions of great topographic complexity and regions with contrasting atmospheric or oceanic influences than in flatter regions or regions with constant atmospheric patterns. Even the performance of the same spatial interpolation method differed considerably with different variables, which resulted from the fact that data quality (in this case variance and range) changed with different variables.

6.2.6. Secondary Variables

The quality of secondary information is important for methods using auxiliary information. In these methods, secondary variables are assumed to be well and accurately sampled at a large number of locations in space and to give a good image of the underlying structure of the primary variable that means they need to be strongly correlated with the primary variables (Ahmed and De Marsily, 1987).

6.3. Correlation between Primary and Secondary Variables

Correlation between the primary and secondary variables is critical for the spatial interpolation methods that use auxiliary information. A number of studies have shown that the strength of the correlation between the primary and secondary variables can considerably affect the performance of CK and OCK (Ahmed and De Marsily, 1987; Goovaerts, 1997; Hernandez-Stefanoni and Ponce-Hernandez, 2006; Juang and Lee, 1998). In addition, the performance of GAM, LM, CART, OK, KED, RK-C and RK-F depended on the choice of secondary information (Bishop and McBratney, 2001). Conversely, Optimal IDW (OIDW) was found to be superior over kriging when data were isotropic and the primary variable was not correlated with secondary variable (Collins and Bolstad, 1996). Non-related or non-significant variables can be eliminated using a stepwise procedure in regression models.

As the correlation increases, the information brought from the secondary variable on to the primary value increases (Goovaerts, 1997). The accuracy of regression modelling depends on how well the data are sampled and how significant the correlation is between the primary variable and secondary variable (Hengl, 2007). It has been shown that stronger correlations would result in more accurate estimations by CK and OCK (Goovaerts, 1997), by OCK over OK and RK-C (Martínez-Cob, 1996) and by SKlm, KED and OCCK (Goovaerts, 2000). In another study, it was found that as correlations between elevation and temperature increased, MAE values dropped significantly for those spatial interpolation methods which used elevation as ancillary information (Collins and Bolstad, 1996). For a correlation >0.4, SCK and

OCK performed better than other methods (SK, OK, LM), and KED was almost as accurate as CK (Asli and Marcotte, 1995). When the correlation increased from 0.77 to 0.99, the RMSE for CK was reduced by 48.3% (Wang *et al*., 2005).

Erxleben *et al*.(2002) have suggested that Moran's I should be used to test whether the primary variable and the secondary variable are spatially independent in terms of cross-correlation statistics. They concluded that only variables that were spatially cross-correlated with the primary variable should be included in OCK models.

6.4. Other Issues

The choice of semivariogram models may play a significant role in the resultant estimation. It was found that the first-order trend OK performed better with Gaussian semi-variogram model than with spherical and exponential models (Hu *et al*., 2004). Some practical guidelines are provided for selecting an appropriate variogram model by Hartkamp *et al* (1999), Goovaerts (2000) and Cressie (1993). Some relevant issues are also discussed in sections 4.1 and 6.1.

The scale (*i.e.*, grid size or resolution) can also affect the accuracy of the estimations. As the grid becomes coarser, the overall information content will progressively decrease (Hengl, 2007). The accuracy may increase as the grid size decreases, but computing time will also increase (Hengl, 2007).

6.5. Interaction among Factors

Interactions among different factors may also exist and should be considered in evaluating the performance of the spatial interpolation methods. All two-way interactions of method, surface type, sampling pattern, noise, and correlation and three way interactions of method-surface type-sampling patterns, method-surface type-noise, and surface type-sampling pattern-noise, were found to significantly affect the performance of the spatial interpolation methods (Zimmerman *et al*., 1999).

6.6. Impacts of Data Quality

The performance of various methods is analysed in relation to sampling density, variation in the data, sampling design and stratification based on information in the 17 reviewed comparative studies in Appendix B. Two measurements proposed in section 4.3 are used to assess the performance of the spatial interpolation methods for different variables.

6.6.1. Sampling Density

The sampling density may play a role in the performance of the spatial interpolation methods as discussed above. Considering all 17 reviewed comparative studies, there is no apparent pattern to the performance of the spatial interpolation methods in relation to sampling density (*i.e.*, area per sample) in terms of RMAE and RRMSE (Figs. 6.1 and 6.2). It is often argued that if the sample size is big enough, then the effects of sample size would disappear, which means that a threshold exists. Apparently this assertion is not true as seen in this review because, in intensely sampled cases, there is still a clear pattern where as sample size increases the performance of the spatial interpolation methods continues to improve (Figs. 6.3 and 6.4). However, this result is probably misleading because the pattern is confounded by the effects of variation in the data as illustrated in Figs. 6.12 to 6.16 in 6.6.2.

For each method, there is little relation between the performance and the sampling density in terms of RMAE and RRMSE (Figs. 6.5 and 6.6). The difference in observations for RMAE and RRMSE is due to the fact that in some studies, both MAE and RMSE were reported, but in some only one of them was presented. The effects of sampling density are dominated by the other factors such variation in the data.

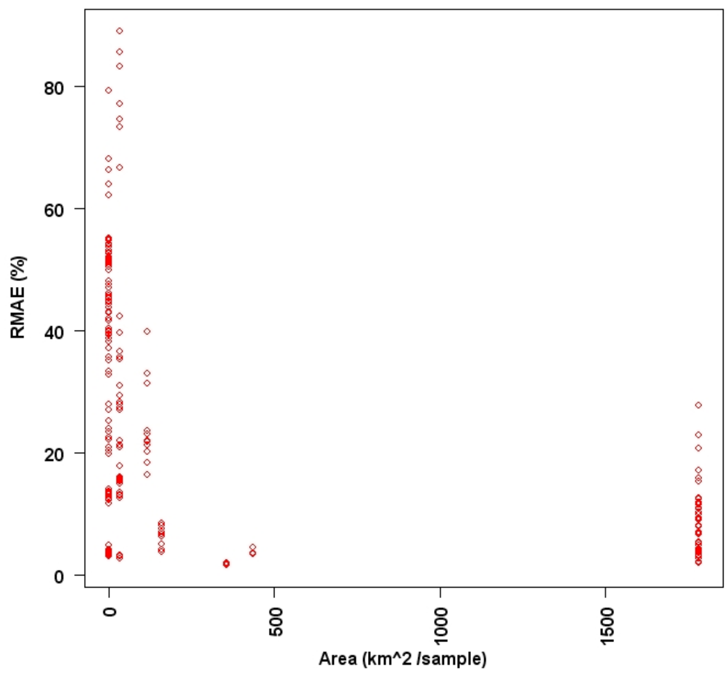


Figure 6.1. Effects of sampling density on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in terms of RMAE(%).

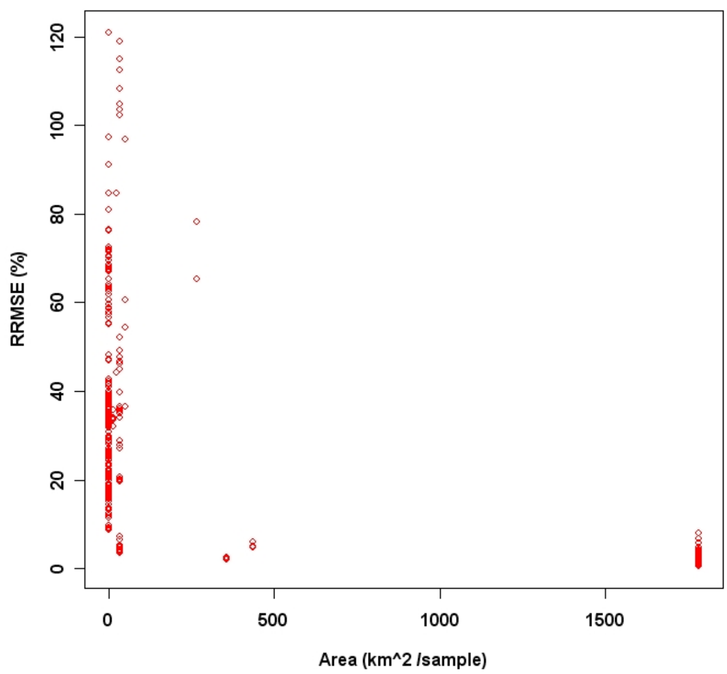


Figure 6.2. Effects of sampling density on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in terms of RRMSE(%).

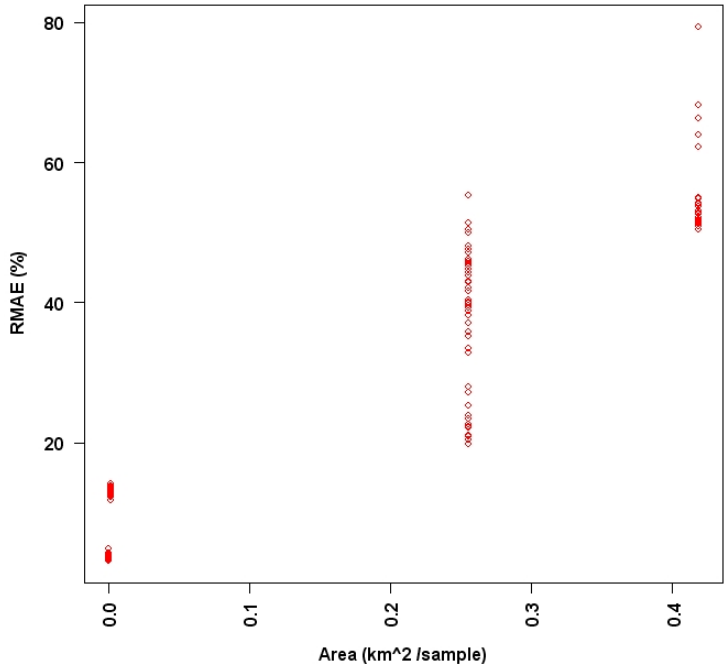


Figure 6.3. Effects of sampling density on the accuracy of the spatial interpolation methods compared in the 17 comparative studies with intensely sampled cases in terms of RMAE(%).

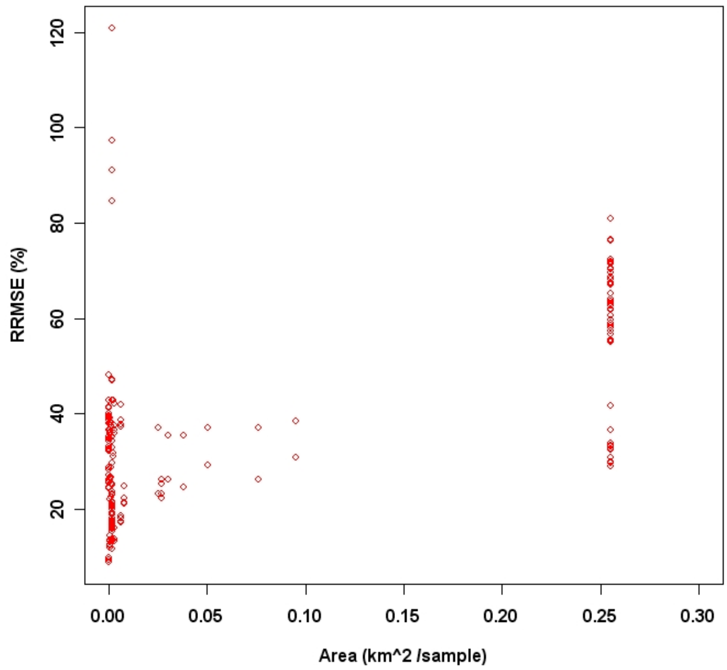


Figure 6.4. Effects of sampling density on the accuracy of the spatial interpolation methods compared in the 17 comparative studies with intensely sampled cases in terms of RRMSE(%).

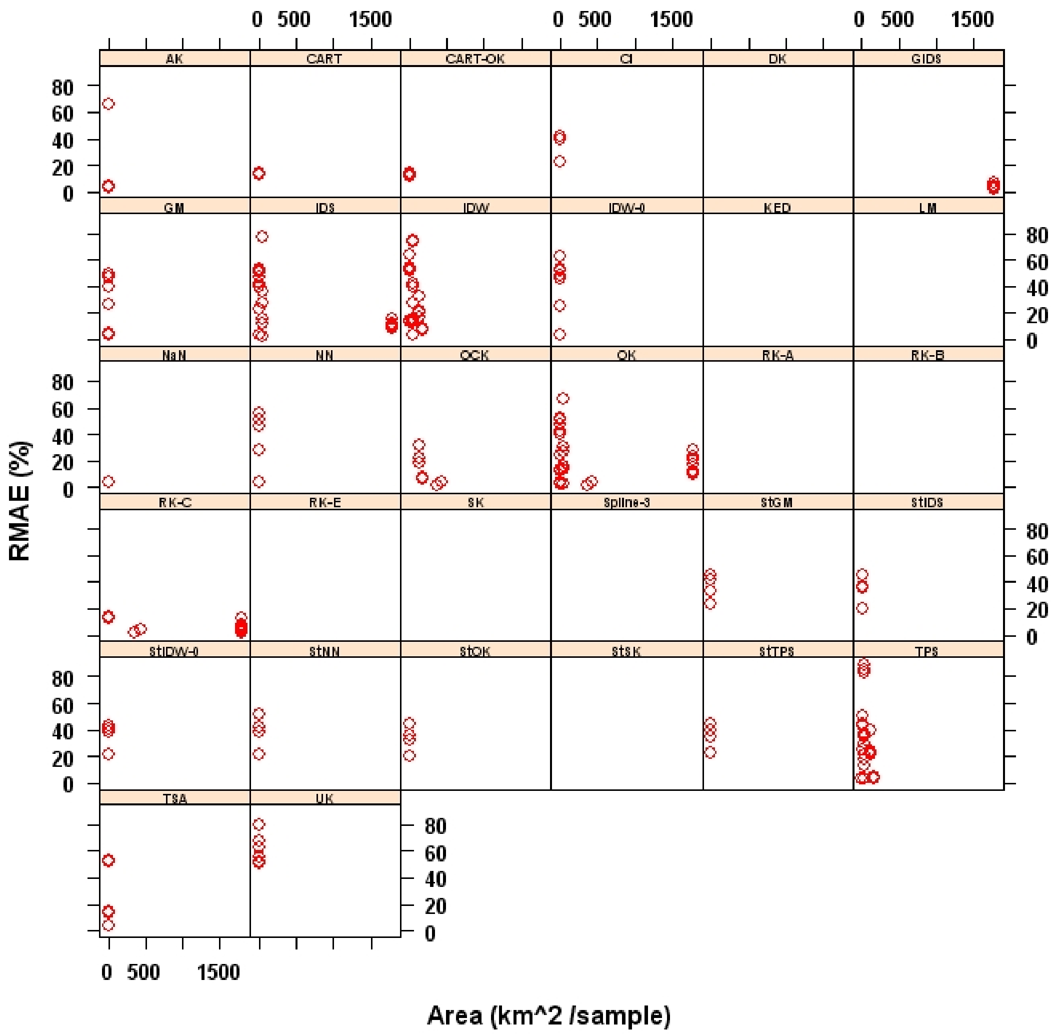


Figure 6.5. Effects of sampling density on the accuracy of each spatial interpolation method compared in the 17 comparative studies in terms of RMAE(%).

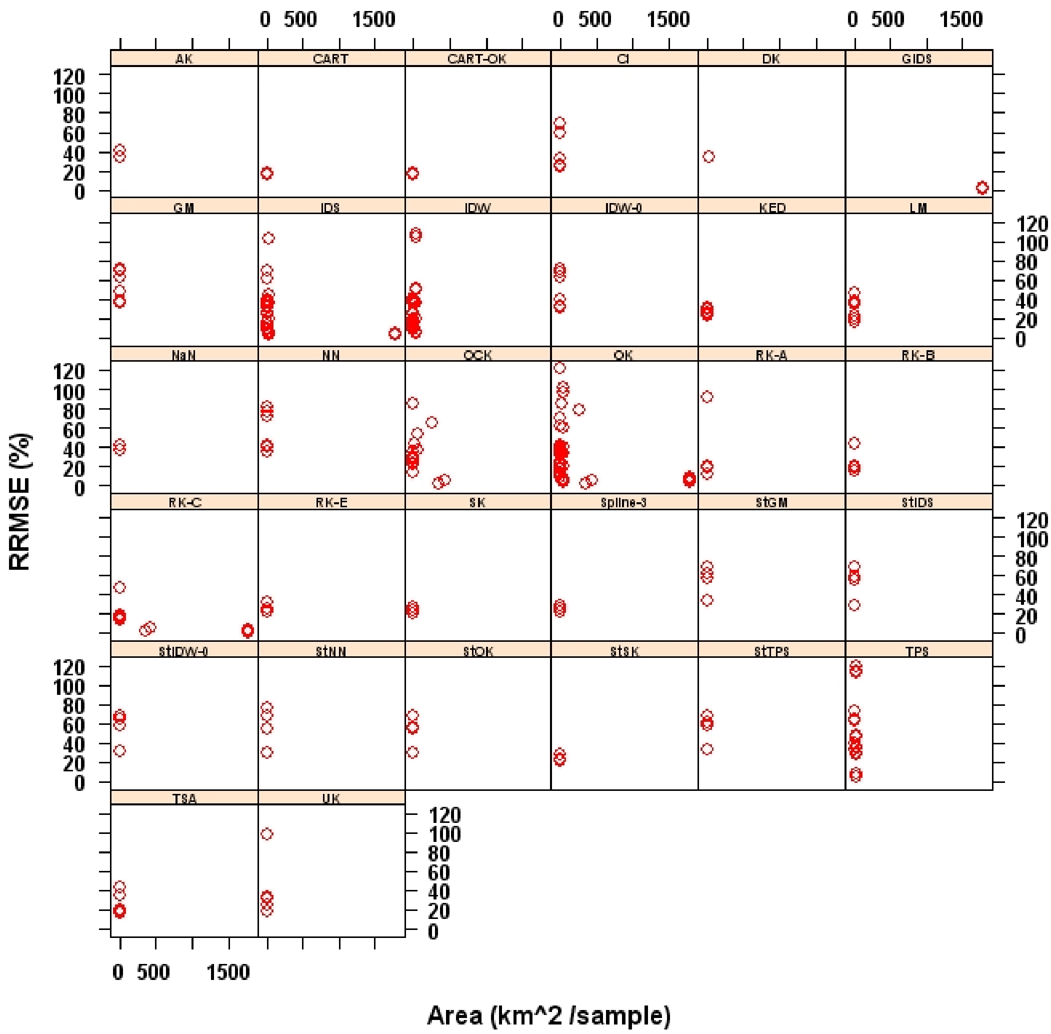


Figure 6.6. Effects of sampling density on the accuracy of each spatial interpolation method compared in the 17 comparative studies in terms of RRMSE(%).

6.6.2. Data Variation

Data variation significantly affects the performance of the spatial interpolation methods. There is a strong pattern of the performance of the spatial interpolation methods in relation to the variation in the data in terms of RMAE and RRMSE (Figs. 6.7 and 6.8). As the variation increases, the performance declines, which is consistent with previous findings (Collins and Bolstad, 1996; Martínez-Cob, 1996; Schloeder *et al*., 2001).

This relationship maintains when the sampling density changes in terms of RMAE and RRMSE (Figs. 6.9 and 6.10). The pattern is further illustrated for datasets with high sample densities where the area per sample is <50 km^2 (Figs. 6.11 and 6.12) and even <0.3 km^2 (Figs. 6.13, 6.14, 6.15 and 6.16). This relationship also persists for data with relatively low sample densities where the area per sample is >1500 km^2, although in this case the results were only available from one study and the area per sample is 1783 km^2 (Figs. 6.17 and 6.18). These results suggest that the effects of variation in data on the performance of the spatial interpolation methods are independent of sampling density.

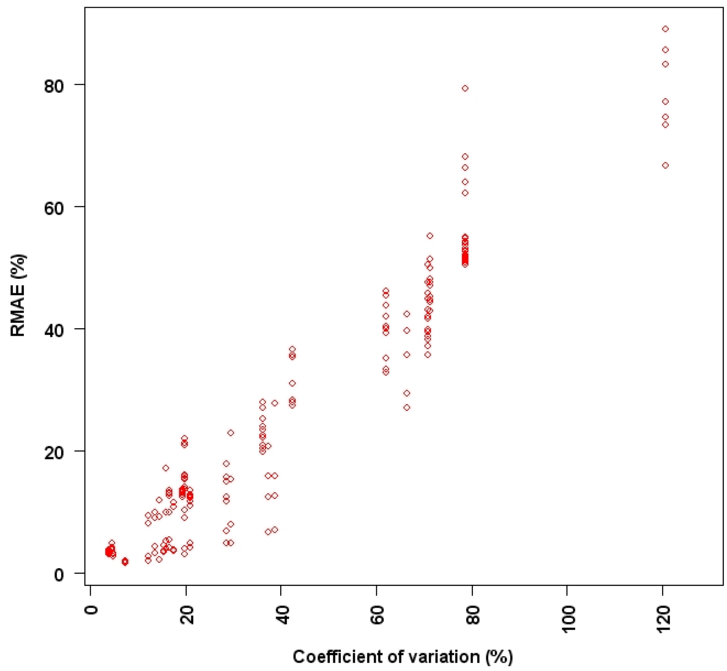


Figure 6.7. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in terms of RMAE(%).

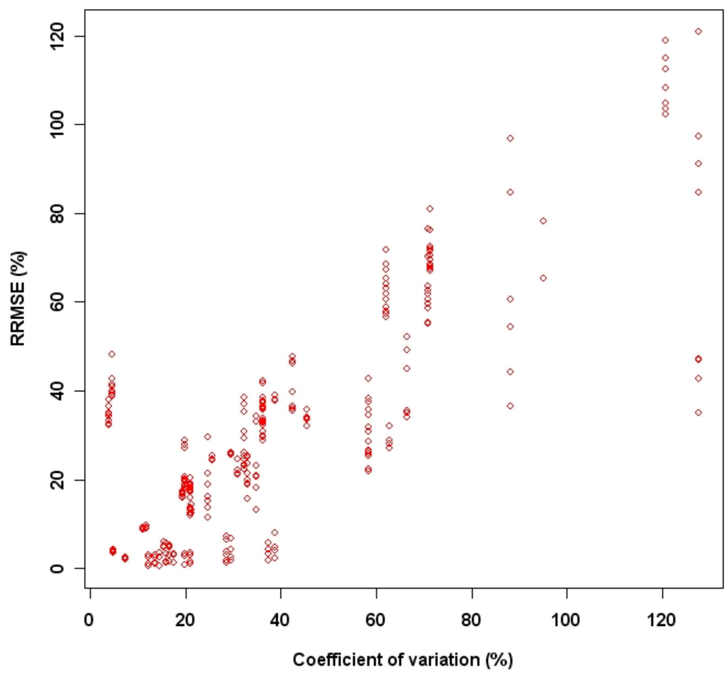


Figure 6.8. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in terms of RRMSE(%).

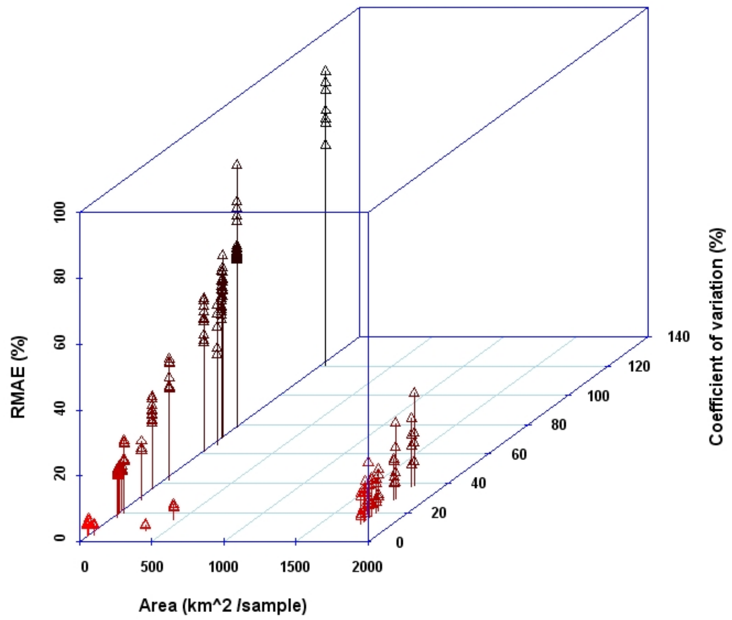


Figure 6.9. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in relation to the sample density in terms of RMAE(%).

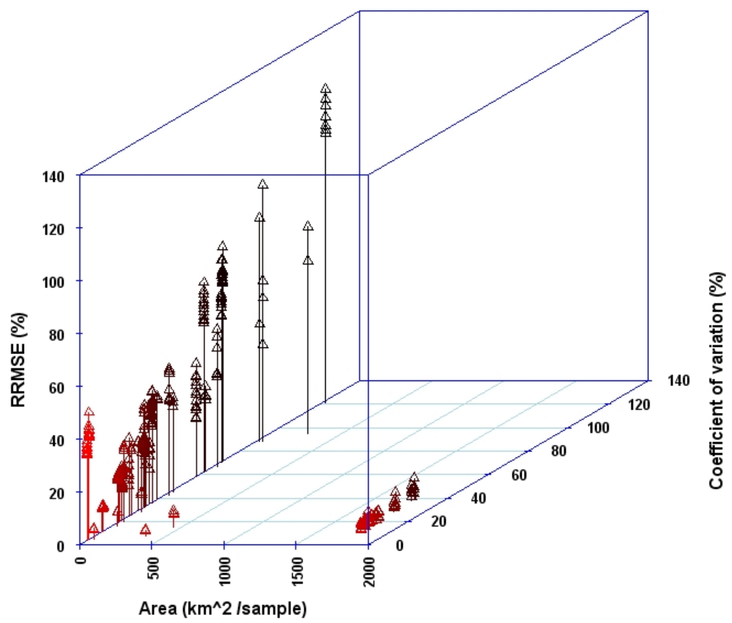


Figure 6.10. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in relation to the sample density in terms of RRMSE(%).

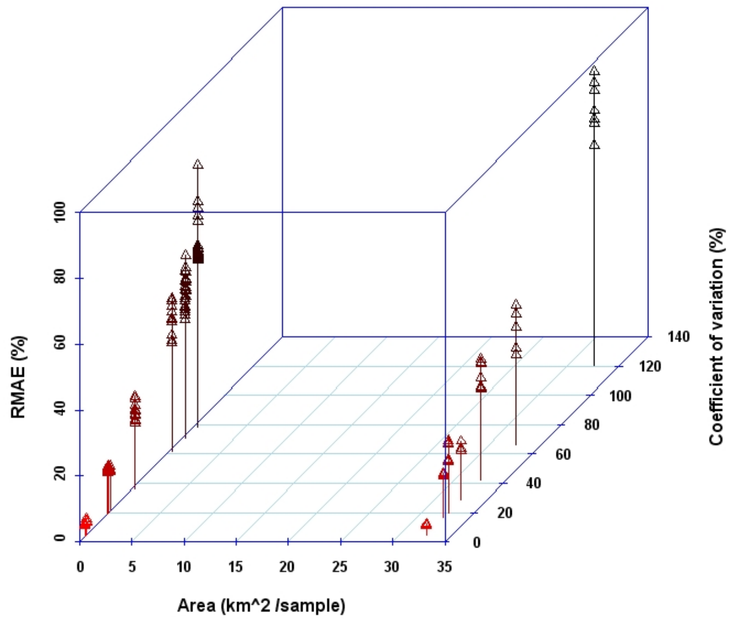


Figure 6.11. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in relation to high sample density in terms of RMAE(%).

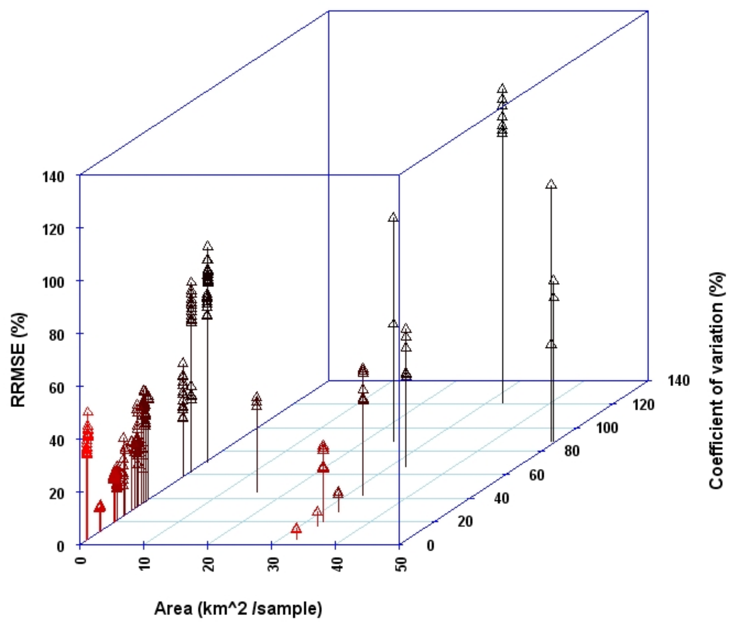


Figure 6.12. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in relation to high sample density in terms of RRMSE(%).

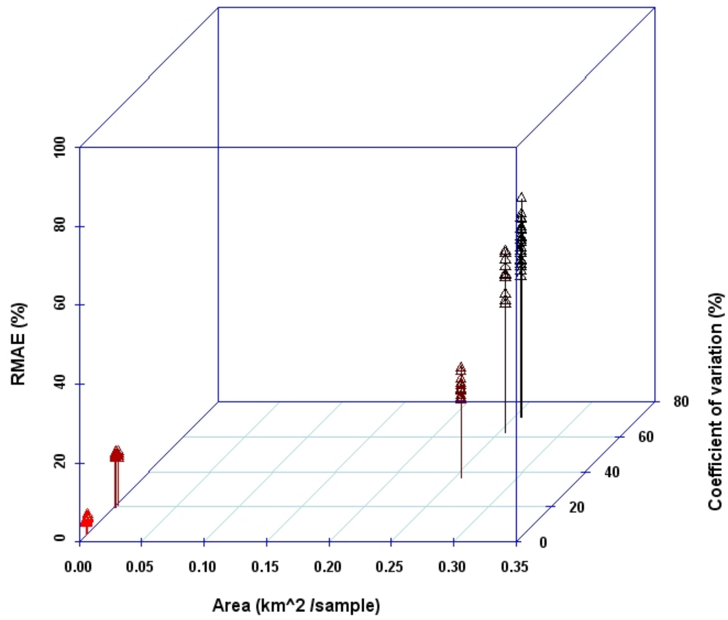


Figure 6.13. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in relation to very high sample density in terms of RMAE(%).

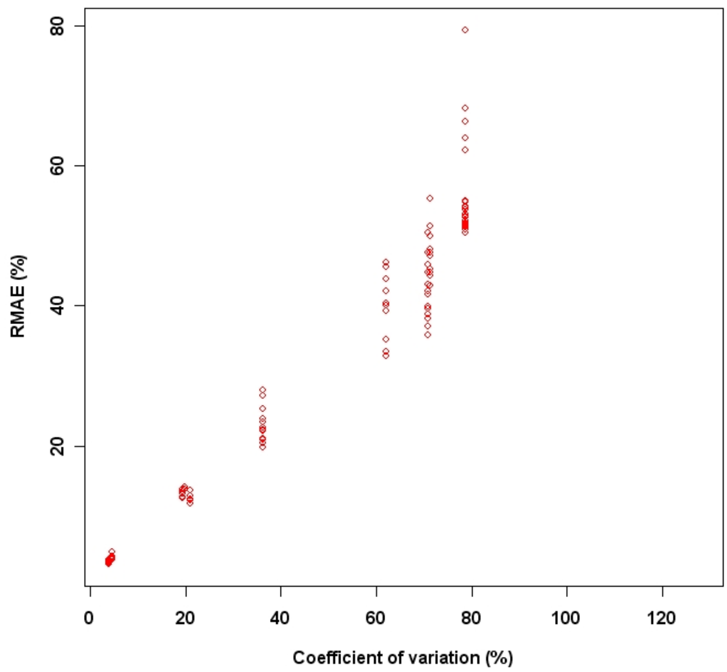


Figure 6.14. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies with very high sample density (<0.3 km^2 per sample) in terms of RMAE(%).

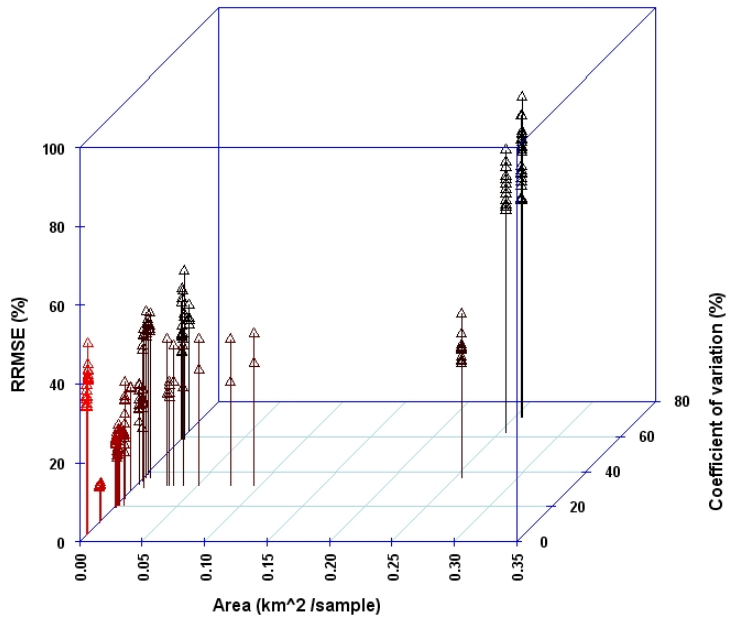


Figure 6.15. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in relation to very high sample density in terms of RRMSE(%).

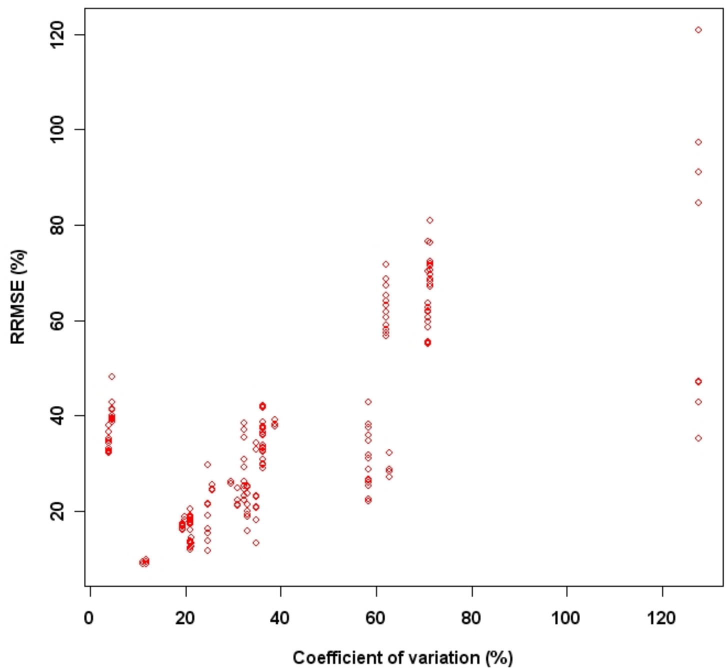


Figure 6.16. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies with very high sample density (<0.3 km^2 per sample) in terms of RRMSE(%).

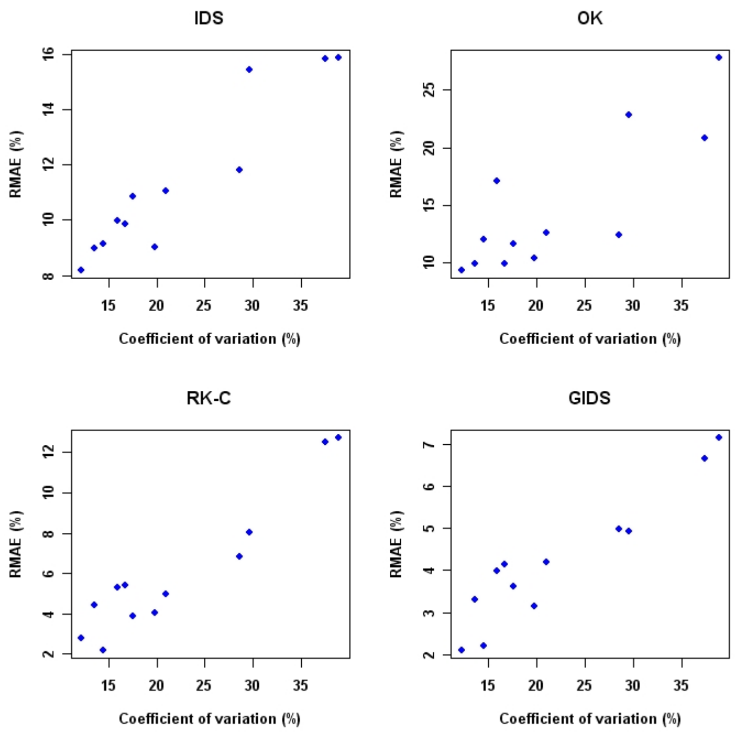


Figure 6.17. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in relation to low sample density in terms of RMAE(%).

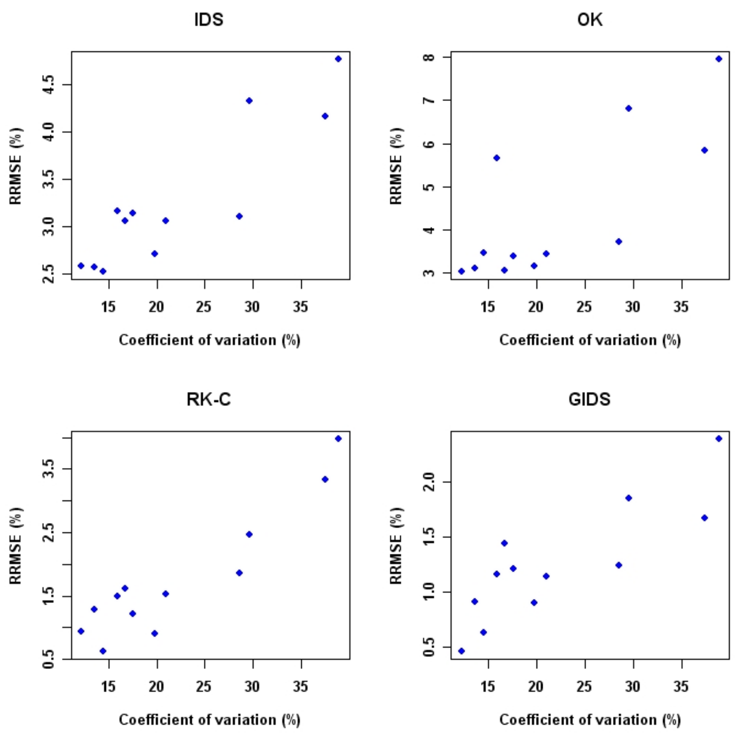


Figure 6.18. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in relation to low sample density in terms of RRMSE(%).

The relationship between the performance of the spatial interpolation methods and the variation in the data is largely maintained for each individual method in terms of RMAE and RRMSE (Figs. 6.19 and 6.20; see also Figs. 6.17 and 6.18). The results show that the performance of all frequently used methods is affected by the variation in the data, but the overall impact is method-dependent. GIDS and RK-C are less sensitive to the variation in the data than OK, OCK, IDS, IDW and TPS, and OCK is less sensitive to the variation in the data than OK in terms of RRMSE. Such method dependency was also observed in the improved RBFN, which performed well especially when the variance of the reference surface was large in comparison with RBFN and OK (Lin and Chen, 2004).

The results are consistent with findings by Gotway *et al*. (1996) who found that the performance of a spatial interpolator (IDW in this case) may be affected by variation of the dataset in terms of CV. Although they claimed that the performance of OK was generally unaffected by variation in the data, but it turned out to be that the performance of OK, like IDW, also declines as the variation increases (as illustrated in Figs A.1 and A.2). This again supports the results observed in this review.

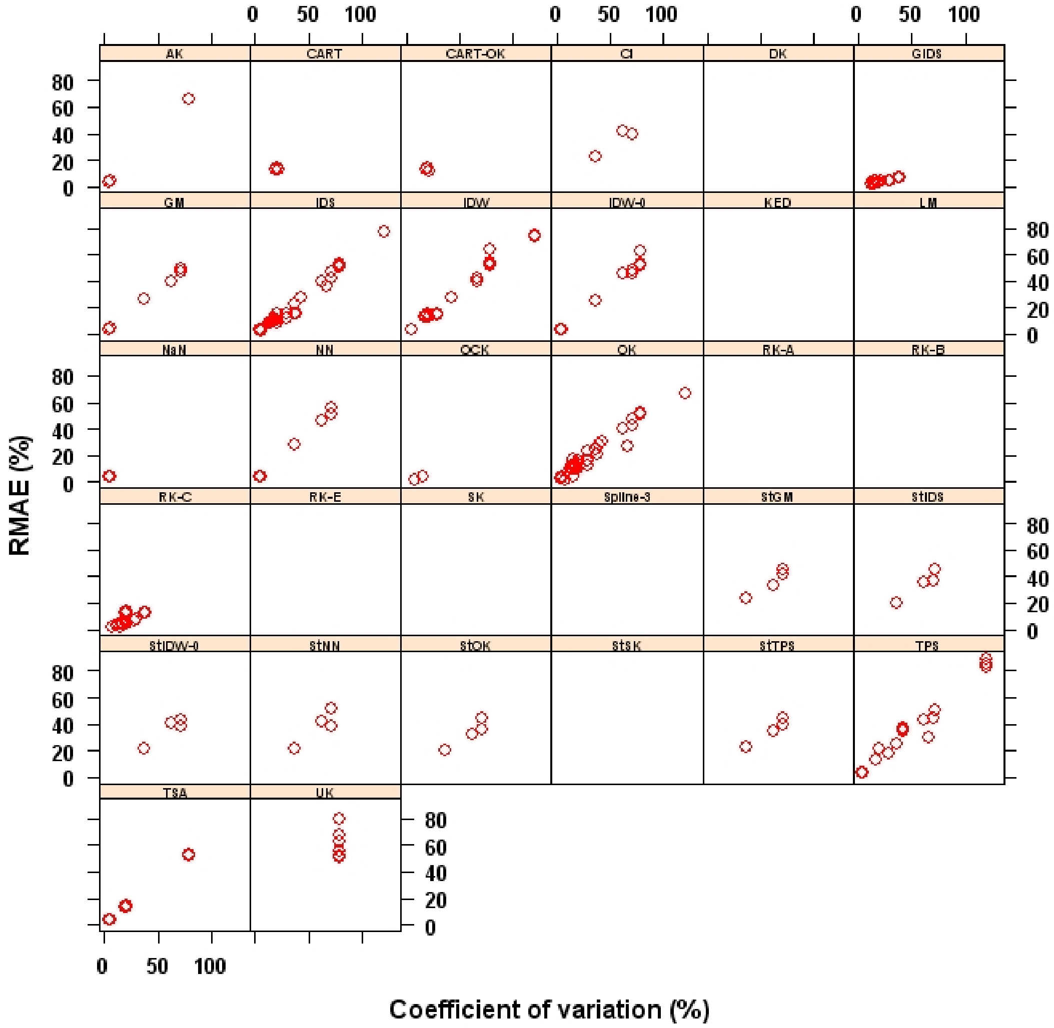


Figure 6.19. Effects of the variation in the data on the accuracy of each spatial interpolation method compared in the 17 comparative studies in terms of RMAE(%).

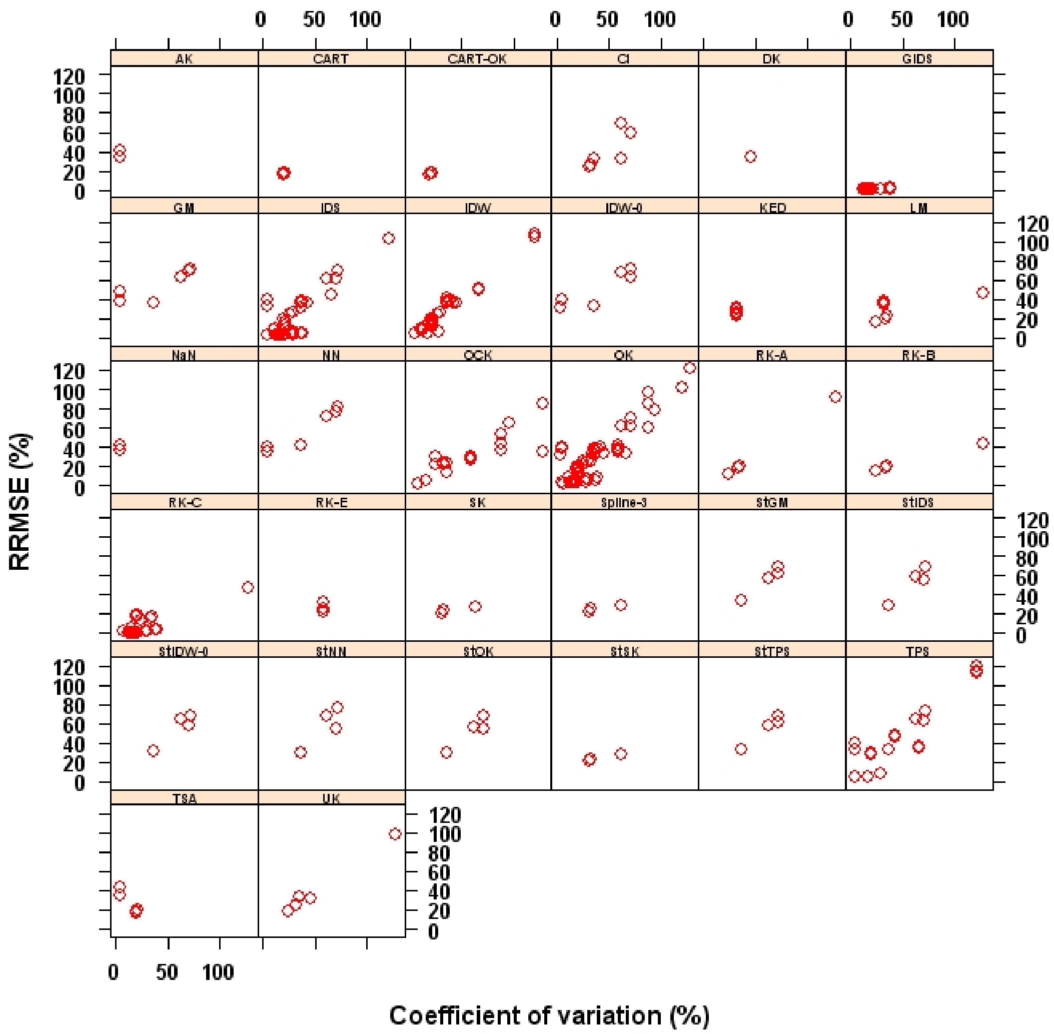


Figure 6.20. Effects of the variation in the data on the accuracy of each spatial interpolation method compared in the 17 comparative studies in terms of RRMSE(%).

6.6.3. Sampling Design

Difference in sampling design affects the performance of the spatial interpolation methods. In the 17 comparative studies, samples collected from point locations that are irregularly distributed in space lead to a higher accuracy of the estimations of the spatial interpolators than samples colleted from regularly distributed points (Fig. 6.21 and 6.22). Therefore, the irregular sampling design could improve the performance of the spatial interpolation methods. However, splines performed poor for irregular spaced data (Collins and Bolstad, 1996); and sample patterns were found not to be significant in determining the performance of the spatial interpolator (*i.e.*, OK) (Englund *et al*., 1992).

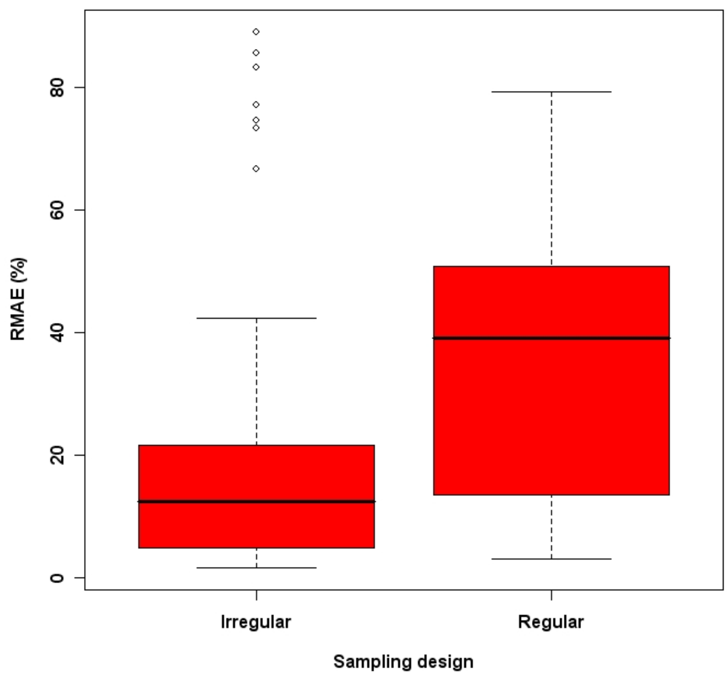


Figure 6.21. Effects of sampling design on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in relation to the sampling design in terms of RMAE(%).

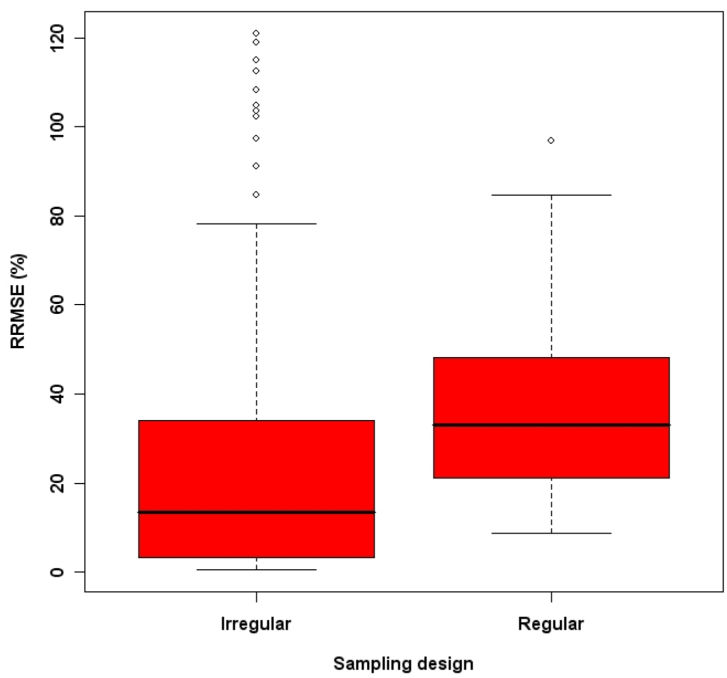


Figure 6.22. Effects of sampling design on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in relation to the sampling design in terms of RRMSE(%).

It is obvious that although the average value of the RMAE and RRMSE is relatively lower for samples collected from irregularly distributed points than from regularly sampled ones, the variation in the RMAE and RRMSE is much higher for samples colleted from irregularly distributed points (Figs. 6.21 and 6.22). This high variation in the RMAE and RRMSE is mainly due to the relatively high variations in the datasets for the irregularly spaced samples (Figs. 6.23 and 6.24). The high CV comes from three studies that used irregularly sampling method (Odeh *et al*., 1994; Odeh *et al*., 1995; Schloeder *et al*., 2001). However, this does not necessarily mean that the irregularly sampling method would always lead to high variation in the collected data, and further investigation is warranted.

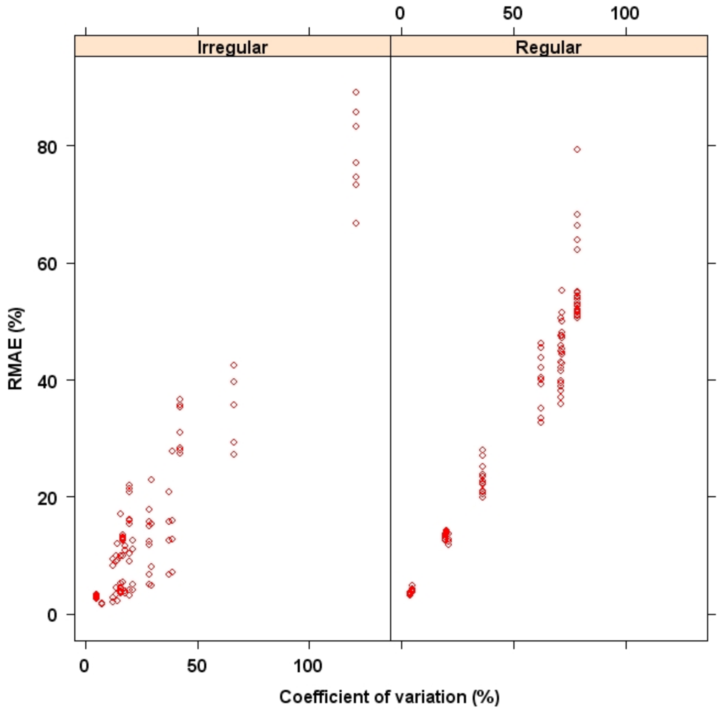


Figure 6.23. Effects of sampling design and the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in terms of RMAE(%).

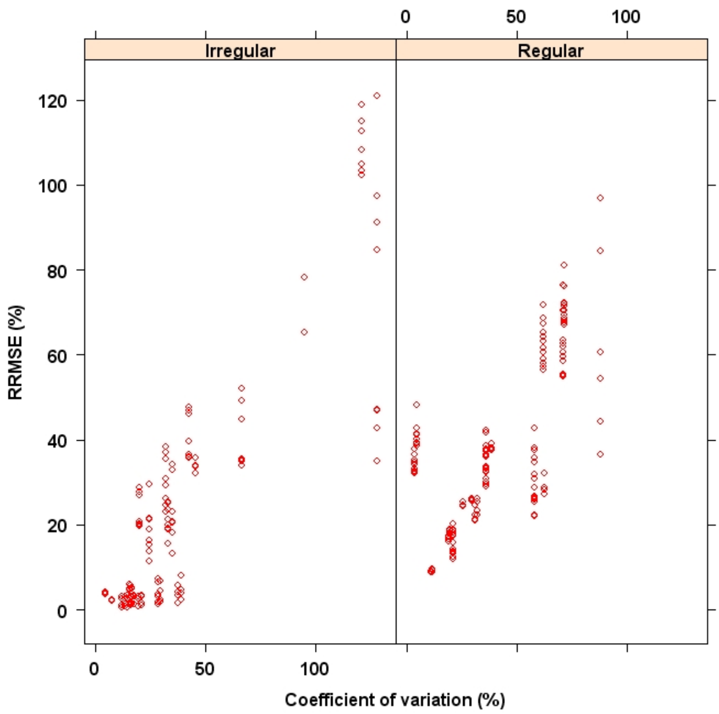


Figure 6.24. Effects of sampling design and the variation in the data on the accuracy of the spatial interpolation methods compared in the 17 comparative studies in terms of RRMSE(%).

6.6.4 Stratification

Stratification pertains to the method of spatial interpolation rather than features of the data. Although it was claimed that there was no statistically significant stratification (Brus *et al*., 1996), the trend of improvement in the performance of the spatial interpolation methods is apparent for all methods compared in terms of RMAE and RRMSE (Figs. 6.25 and 6.26). It was found that stratification improved the performance of SK by Voltz and Webster (1990), because it avoids the effects of non-stationary. However, stratification has two major limitations: 1) it may dramatically reduce the number of samples in the kriging neighbourhood, and 2) it depends on the goodness of the classification (Voltz and Webster, 1990). Further research into the effects of stratification is warranted.

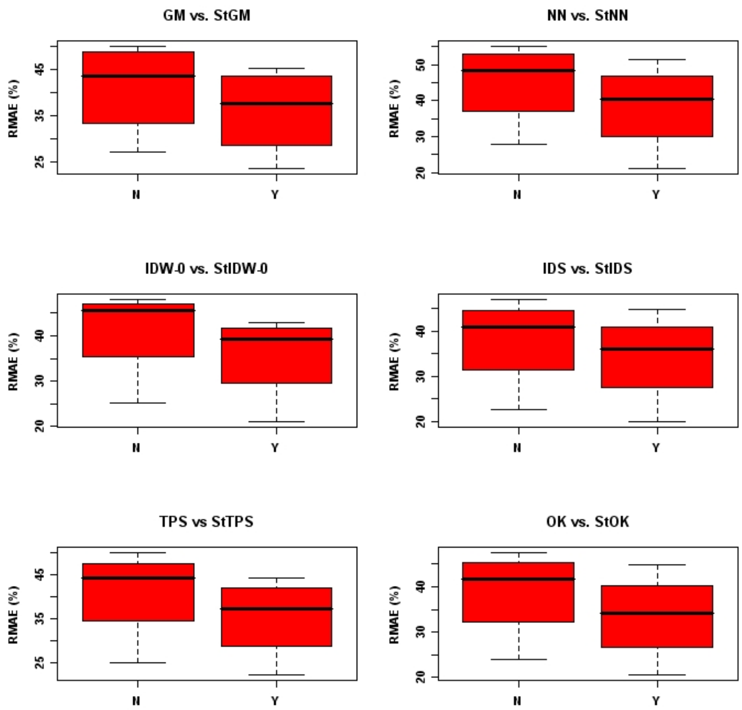


Figure 6.25. Effects of stratification on the accuracy of the spatial interpolation methods compared by Brus *et al*. (1996) in terms of RMAE(%).

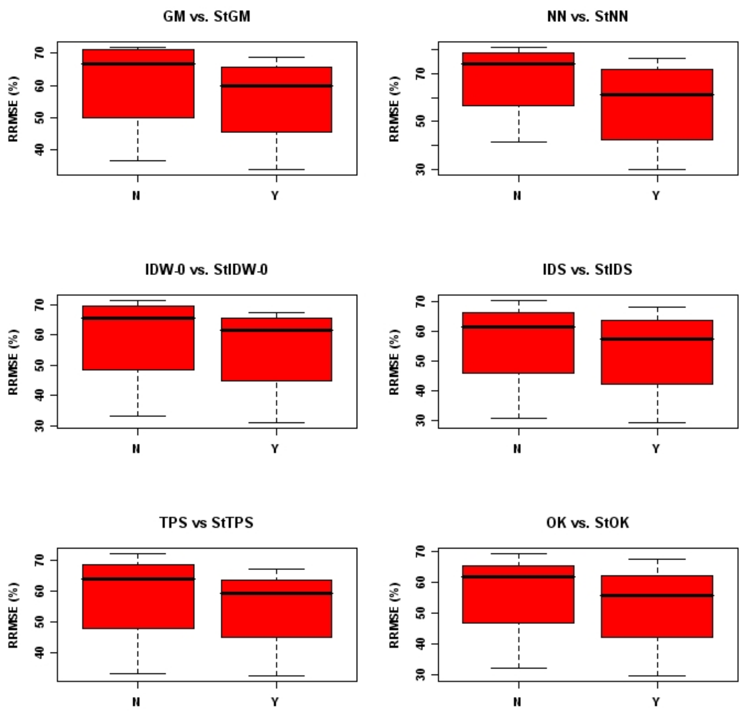


Figure 6.26. Effects of stratification on the accuracy of the spatial interpolation methods compared by Brus *et al*. (1996) in terms of RRMSE(%).

Chapter 7: Classification and Selection of the Methods

In this chapter, the spatial interpolation methods are classified based on their features to provide an overview of the differences and relationships among the various spatial interpolation methods. These features are then quantified and a cluster analysis is conducted to show similarities and relationships among these spatial interpolators. Lastly, a decision tree is developed for selecting an appropriate method according to the nature and availability of data to provide guidelines for potential users.

7.1. Classification of Spatial Interpolation Methods

The spatial interpolation methods are classified based on the comparisons summarised in Tables 3.1 and 3.2. This classification assists in further understanding these methods and provides a base for developing a decision tree for selecting an appropriate method in practice.

The classification of the spatial interpolation methods has not been addressed before apart from the study of Lam (1983) who proposed a simple classification of four types of the spatial interpolation methods. In this review, an approach used in taxonomy is adopted to classify the 26 spatial interpolation methods according to their features, as follows:

1 Non-geostatistical, no error assessment
  2 Deterministic
    3 Global ……………………………………………..…................................**Cl**
    3* Local
      4 Exact
        5 Abrupt
          6 Tessellation and using one sample…...**NN**
          6* Using more than one sample
            7 Triangulation and using three samples...**TIN**
            7* Combination of triangulation & tessellation ....**NaN**
        5* Gradual
          8 Univariate………..……......…..**NaN**
          8* Univariate/multivariate, exact within smoothing limit.......…………. **TPS**
      4* Inexact ………………………..**IDW**
  2* Stochastic
    9 Global
      10 Abrupt………………………..……………………………………..…...**CART**
      10*Gradual
        11 Coordinates only……………………………….………………………**TSA**
        11* Coordinates and other secondary variables……………...……………..**LM**
    9*Local……………………………………………………………..**Splines & LTS**
1* Geostatistical, with error assessment
  12 Univariate
    13 Stationary mean……………………………………………………………**SK**

13* Local means
14 Point estimate……………………………………………………..……OK
14* Block estimate…………………………………………………………BK
12* Multivariate
15 Stationary mean
16 Non-stratification
17 Search window with multiple samples………...…………………….…SCK
17* Search window with single sample……………..…………….………SCCK
16* Stratification
18 Non-continuous secondary information…………..…………………SKWS
18* With continuous secondary information……….…………….…...SCKWS
15*Local means
19 Exhaustive secondary information and/or local trend
20 Coordinates only…………………….……………………………….…..UK
20* Non-coordinate secondary variable
21 A secondary variable and search window with multiple samples
22 Regression coefficients estimated within each search window...….KED
22* Regression coefficients estimated once……...….………………….SKlm
21* One or more secondary variable and search window with single sample……………………………………………………………....OCCK
19* Non-exhaustive secondary information and no local trend
23 Stratification
24 No secondary information……………………………...….….……OKWS
24* Secondary information…………………………………………OCKWS
23* Non-stratification
25 Orthogonalisation of secondary information………...……………….……PCK
25* Non-orthogonalisation of secondary information
26 No information of the stationary means of the primary and secondary variables……………………………………………...……...……OCK
26* With information of the stationary means of both the primary and secondary variables...……………………………..…...……..SOCK

7.2. Similarity of Spatial Interpolation Methods

The similarity between 26 spatial interpolation methods is analysed in this section. A total of 21 features extracted from Table 3.1 and 3.2 and from those used for the classification are converted into qualitative variables with factor levels ranging from 0 to 1 or not applicable (na) (Table 7.1). Information of each feature for each of the 26 spatial interpolation methods is summarised in Table 7.2.

Table 7.1. Conversion between feature status and factor levels.

No	Feature	Level		
		0	1	na*
1	Univariate	no	yes	
2	Multivariate	no	yes	
3	Deterministic/stochastic	deterministic	stochastic	
4	Local/global	global	local	
5	Exact/inexact	exact	inexact	
6	Abrupt transition	no	yes	
7	Gradual transition	no	yes	
8	Output: polygons	no	yes	
9	Output: triangular	no	yes	
10	Output: grids	no	yes	
11	Stationary/ local mean	stationary	local	na
12	Stationary mean of secondary variable	no	yes	na
13	Local trend-constant	no	yes	na
14	Local trend-non-constant	no	yes	na
15	Info of coordinates	no	yes	
16	Secondary variables	no	yes	
17	Point/ block	point	block	
18	Exhaustive secondary information	no	yes	na
19	Stratification	no	yes	
20	Orthogonalisation of secondary information	no	yes	na
21	Single or multiple samples in the search window	single	multiple	

*na: not applicable.

Table 7.2. The qualified data of the 21 features of 26 spatial interpolation methods. For the feature corresponding to each number please see Table 7.1. The methods are arranged in an order according to the results from Figure 7.1. The bold values highlight the key differences among the methods within each non-single-method group.

Method	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
UK	0	1	1	1	0	0	1	0	0	1	1	0	**0**	1	**1**	**0**	0	1	0	0	1
SKlm	0	1	1	1	0	0	1	0	0	1	1	0	**1**	1	**0**	**1**	0	1	0	0	1
KED	0	1	1	1	0	0	1	0	0	1	1	0	**0**	1	**1**	**1**	0	1	0	0	1
SKWS	0	1	1	1	0	0	1	0	0	1	**0**	0	1	0	0	0	0	na	1	na	1
OKWS	0	1	1	1	0	0	1	0	0	1	**1**	0	1	0	0	0	0	na	1	na	1
SCK	0	1	1	1	0	0	1	0	0	1	**0**	**0**	1	0	0	1	0	0	**0**	**0**	**1**
OCK	0	1	1	1	0	0	1	0	0	1	**1**	**0**	1	0	0	1	0	0	**0**	**0**	**1**
SOCK	0	1	1	1	0	0	1	0	0	1	**1**	**1**	1	0	0	1	0	0	**0**	**0**	**1**
PCK	0	1	1	1	0	0	1	0	0	1	**1**	**0**	1	0	0	1	0	0	**0**	**1**	**1**
SCCK	0	1	1	1	0	0	1	0	0	1	**0**	**0**	1	0	0	1	0	0	**0**	**0**	**0**
OCCK	0	1	1	1	0	0	1	0	0	1	**1**	**0**	1	0	0	1	0	0	**0**	**0**	**0**
SCKWS	0	1	1	1	0	0	1	0	0	1	**0**	**0**	1	0	0	1	0	0	**1**	**0**	**1**
OCKWS	0	1	1	1	0	0	1	0	0	1	**1**	**0**	1	0	0	1	0	0	**1**	**0**	**1**
LM	1	1	1	0	1	0	1	1	0	1	na	0	1	1	1	1	0	1	0	0	1
CART	1	1	1	0	1	1	0	0	0	1	na	0	na	na	1	1	0	1	0	0	1
TSA	0	1	1	**0**	1	0	1	0	0	1	na	0	na	na	1	0	0	na	0	na	1
LTS	0	1	1	**1**	1	0	1	0	0	1	na	0	na	na	1	0	0	na	0	na	1
Cl	1	0	0	0	1	1	0	1	0	0	na	na	na	na	1	1	1	1	0	na	0
SK	1	0	1	1	0	0	1	0	0	1	**0**	na	1	0	0	0	**0**	na	0	na	1
OK	1	0	1	1	0	0	1	0	0	1	**1**	na	1	0	0	0	**0**	na	0	na	1
BK	1	0	1	1	0	0	1	0	0	1	**1**	na	1	0	0	0	**1**	na	0	na	1
IDW	1	**0**	0	1	**1**	0	1	0	0	1	na	na	na	na	0	0	0	na	0	na	1
TPS	1	**1**	0	1	**0**	0	1	0	0	1	na	na	na	na	0	0	0	na	0	na	1
NN	1	0	0	1	0	1	**0**	**1**	**0**	1	na	na	na	na	0	0	1	na	0	na	**0**
TIN	1	0	0	1	0	1	**0**	**0**	**1**	1	na	na	na	na	0	0	1	na	0	na	**1**
NaN	1	0	0	1	0	1	**1**	**0**	**0**	1	na	na	na	na	0	0	1	na	0	na	**1**

The data in Table 7.2 were first analysed using hierarchical cluster analysis on Gower's distance in R 2.6.2 (R Development Core Team, 2007). In a further cluster analysis, NA was replaced by 0, because it was treated as missing when calculating the distance in the first analysis. However, both analyses produced the same classification results. The results from the first analysis are presented in Figure 7.1. If a threshold line is added at 0.2 in Figure 7.1, these methods could be classified into 10 groups. The results show that:

1) LM, CART and Cl each forms a single method group, and they are group 4, 5 and 7 respectively, which indicate that they are different for each other and also from all the other methods.

2) Group 1: SKlm, UK and KED are alike and they all use secondary information and/or coordinate information in making their estimations and they are different in local trend and utilisation of coordinate and secondary information.

3) Group 2: SKWS and OKWS form a group of kriging methods without using secondary information but with stratification; and they differ due to the stationary mean for SKWS and local means for OKWS.

4) Group 3: all cokriging methods are grouped together. Within this group there are two subgroups as distinguished by stratification. Methods with stratification differ in the choice of mean; and methods without stratification differ in the choice of mean, using secondary information, the number of samples in search window and orthogonalisation in producing their estimations.

5) Group 6: TSA and LTS form a group that uses coordinate information in deriving the estimations; and they are different in estimation in that the TSA is a global approach and LTS is a local one.

6) Group 8: BK, SK and OK are most similar and they do not use secondary information; and their differences are from the choice of point or block estimation and the choice of the stationary mean or local means.

7) Group 9: IDW and TPS are in the same group and their features are similar; and they differ in exactness of estimation and TPS could be multivariate.

8) Group 10: NN, TIN and NaN are similar for most of features considered except the output and the smoothness.

The relationship among these 10 groups can be further explored if a threshold line is added at 0.4 in Figure 7.1, and these groups can be merged into four major groups. Group 1, 2 and 3 are similar because of their common features like multivariate, stochastic, local, inexact, gradual, grid-output and with point estimation. Group 4, 5 and 6 share similar features that are multivariate, stochastic, inexact, utilisation of coordinate information, point-estimation, stratification, multiple samples in search widow and with grid-output. Group 7 is unique due to the combination of the requirement of exhaustive secondary information and block estimation, and polygon-output. Group 8, 9 and 10 are all featured with univariate, local, and non-utilisation of coordinate and secondary information.

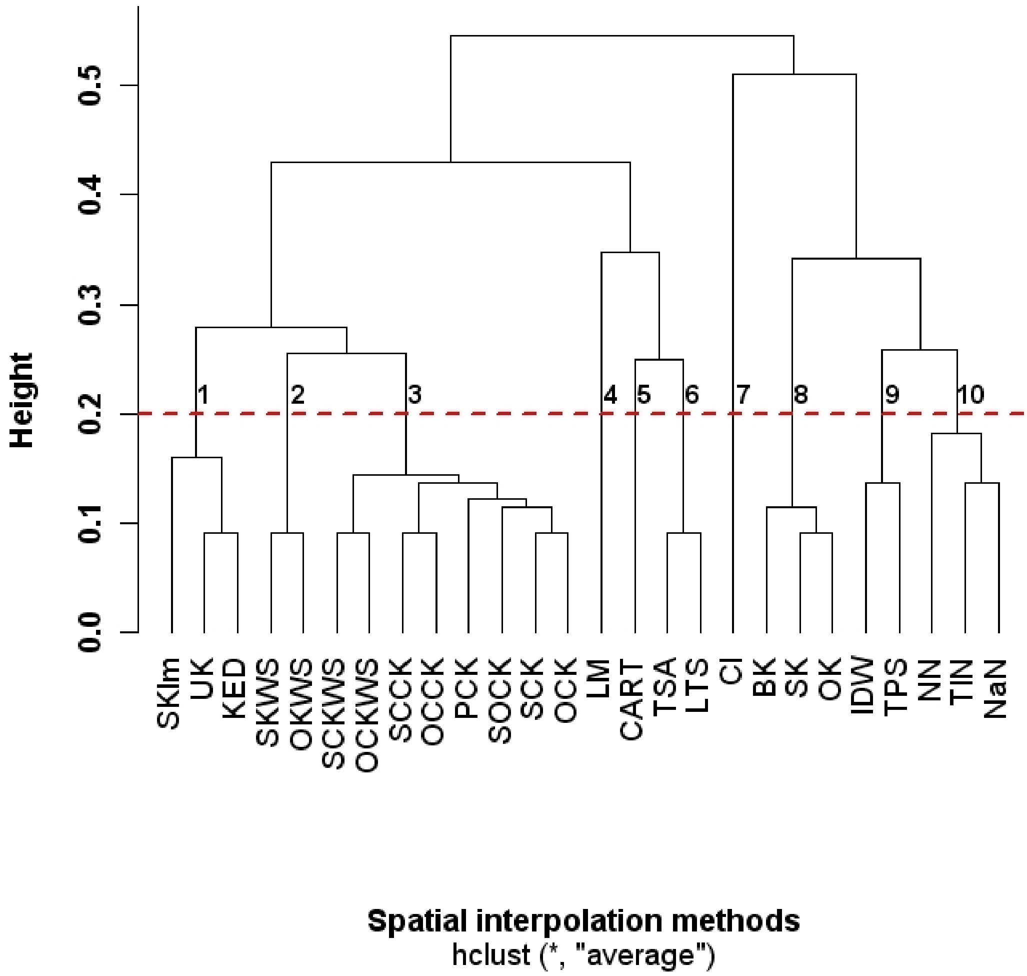


Figure 7.1. Classification of the spatial interpolation methods based on the 21 binary features in Table 7.2.

7.3. Selection of Spatial Interpolation Methods

Selection of an appropriate spatial interpolation method for the data at hand is critical, but it is not an easy task. The performance of the spatial interpolators depends on many factors including: the variable under study, the spatial configuration of the data, and the underlying assumptions of the spatial interpolation methods. It seems that there is no simple answer regarding the choice of an appropriate spatial interpolator, because a method is "best" only for specific situations (Isaaks and Srivastava, 1989).

There are a number of factors that should be considered in making an appropriate selection, given that there is no one best spatial interpolation method. The choice of method may depend on the assumption and properties of each method, nature and spatial structure of the data for the primary variable, sample size or sample density and distribution, the availability of secondary information, and the factors discussed in Chapter 6. They can be used prior to interpolation to eliminate some inappropriate methods. The availability of software may also be an important issue. The computational demands are also crucial depending on the sample size, the power of the computer, and the efficiency of software.

In this section, a decision tree for selecting an appropriate spatial interpolation method is developed according to the availability and nature of the data and the expected estimation in combination with the features of each spatial interpolator. All 26 spatial interpolation methods listed in Table 7.2 are considered. Again this decision tree is represented in an easy following taxonomic fashion.

1 Data or residuals show spatial structure
  2 Estimation of continuous variable
    3 No information of secondary variables available
        4 Global mean known ……………………………..…...............................**SK**
        4* Global mean unknown and using local means
          5 Point estimation……..……………………..**OK**
          5* Block estimation……………………………………………………..**BK**
    3* Information of secondary variables available
      6 Global mean known
        7 Secondary variable is only categorical
          8 Stratification……………………..………………………………...**SKWS**
          8* Non-stratification………………………………………..…..……….**SKlm**
        7* Secondary variable is not only categorical
          9 Stratification……………………..…………………………………**SCKWS**
          9* Non-stratification
            10 Sparse samples of secondary variable and multiple samples in search window……….………………………………………………...…**SCK**
            10* Dense samples of secondary variable and single sample in search window.………………………………..…..……………………**SCCK**
      6* Global mean unknown and using local means
        11 Secondary information available for each point being estimated
          12 Spatial trend is apparent and only coordinates available ……….....……**UK**
          12* Other secondary variable available
            13 An apparent global relation with the secondary variable…………**SKlm**
            13* The relation is not so apparent…………………………….……….**KED**
        11* Secondary information not available for each point being estimated
          14 Secondary variables including a categorical variable
            15 Only a categorical variable available
              16 Multiple samples in search window……………..…………..….…**OKWS**
              16* Dense samples of secondary variable and single sample in search

window…………………………………...…………………….**OCCK**
15* Other secondary information available……………………….**OCKWS**
14* Secondary variables without categorical variable
17 Sparse samples of secondary variable and multiple samples in search window
18 Many secondary variables and PCA needed..…………………….**PCK**
18* PCA not needed to reduce the number of secondary variables
19 Avoid negative weights and artificially limiting the effect of secondary variable…………………………………………………**SOCK**
19* Accept above two drawbacks………………………………**OCK**
17* Dense samples of secondary variable and single sample in search window………………………………………………………**OCCK**
2* Estimation of categorical variable or uncertainty assessment….**IK & its variants**
1* Data or residuals show no spatial structure
20 No secondary variables available
21 Abrupt estimation acceptable
22 Using single sample for estimation………………………………..…….**NN**
22* Using multiple samples for estimation
23 Using three samples for estimation…………………………………...**TIN**
23* Using more than three natural neighbour samples for estimation……..**NaN**
21* Abrupt estimation unacceptable
24 Using more than three natural neighbour samples weighted by area...…..**NaN**
24* Using nearest several samples weighted by distance…………...……….**IDW**
20* Secondary variable available
25 Using information of coordinates
26 Only coordinates information used with inexact estimation
27 Using nearby samples……..……………………………...**Splines & LTS**
27* Using all samples…………………………………………...……**TSA**
26* May use other variables with exact estimation……………...…………..**TPS**
25* Not using information of coordinates
28 Only categorical secondary information available
29 Only one variable available……………………...……………………..**Cl**
29* Multiple variables available………………………………………**CART**
28* Continuous secondary information available
30 Univariate or multiple secondary information…………………….….**LM**
30* Require multiple secondary information……………………………**CART**

The above decision tree only provides a guideline for selecting an appropriate spatial interpolator according to the nature and availability of the data and the expected outcomes. There are also many other factors as discussed in previous chapters that could influence the choice of the spatial interpolation methods. For example, one might use a spatial interpolator that does not incorporate secondary information even if such information is available if it is considered a reasonable approach. Joint application of two spatial interpolation methods might produce additional benefits such as the combined procedures in section 2.4. If sharp spatial changes, such as those caused by soil and rock types, vegetation classes, and habitat types, are expected, stratified spatial interpolation methods may be used to improve the estimation

(Hernandez-Stefanoni and Ponce-Hernandez, 2006; Stein *et al.*, 1988; Voltz and Webster, 1990).

For kriging methods, a number of factors including sample size, isotropy and anisotropy of the data, need to be considered for selecting appropriate variogram model. Data transformation may need to be considered when the data are skewed and anisotropic. Three methods of data transformation (logarithms, standardised rank order, and normal scores) can be employed to reduce the skewness (Wu *et al*., 2006). Non-stationary methods like KED should be used in cases with a general anisotropy or trend (*i.e.* drift) (Verfaillie *et al*., 2006). If different types of nonstationarity exist inside a study region, application of different spatial interpolation methods to each type may be a good practice because the estimation resulted from the combination of the results from different methods can be more precise than when only a single method is used (Meul and Van Meirvenne, 2003).

It is recommended that one should try several search strategies on a test subregion before running any kriging over an entire region (Goovaerts, 1997). Cross-validation can be used to evaluate the effects of different search parameters on the estimations, but it should be noted that the search strategy that generates the best cross-validated results may not necessarily produce the best estimations at unsampled locations.

When datasets consist of relatively few samples, it is recommended that least square error and ranking procedures should be used rather than Delfiner's methodology for estimating the generalised covariance function (Puente and Bras, 1986).

Guidelines have been also proposed in previous studies for selecting a spatial interpolator from subsets of the methods listed above. For instance, a decision tree for selecting a suitable spatial prediction method from RK, OK, IDW and LM was developed by Hengl (2007). Pebesma (2004) proposed a decision tree for IDW, TSA and a few kriging methods available in gstat package in R. There are also some guidelines for choosing between DK and IK according to the nature and structure of the data (Lark and Ferguson, 2004). In addition, several steps have been provided for using kriging methods by Burrough and McDonnell (1998) who also provide guidelines for selecting an appropriate spatial interpolation method.

This review is the first to provide guidelines in the form of a decision tree for selecting an appropriate spatial interpolator from 26 spatial interpolation methods according to data properties. It is the most comprehensive comparative study that has been published for environmental scientists. The next chapter lists software packages for the application of the spatial interpolation methods to the environmental data.

Chapter 8: Software Packages and Recommendation for Marine Environmental Scientists

8.1. Software Packages

There are many software packages that contain functions to interpolate spatial point data to spatial continuous data (Table 8.1). The list of software packages and the spatial interpolation methods in each package is acquired from various sources and is not exhaustive.

Several packages in *R* perform spatial interpolation, including: *akima*, *deldir*, *fields*, *geoR*, *GeoRglm*, *GRASS*, *gstat*, *spatial*, *sgeostat*, *RandomFields*, and *tripack*. Large parts of the *geoR* and *GeoRglm* packages address the uncertainty of estimated covariance parameters in Bayesian framework (also known as MBK; Diggle and Ribeiro Jr., 2007) (Pebesma, 2004). Due to heavy computational requirements, MBK seems to be only relevant to datasets of small sample sizes (Moyeed and Papritz, 2002). Given that the power of the computer has been increased dramatically recently and some improved algorithms have been adopted in geo-statistics that could easily handle sample size of 10,000 (personal communication with Edzer Pebesma, 9 July 2008), the previous statements made about the computation requirements of the spatial interpolation methods may no longer valid. This should be taken into account in the future studies.

Computer programs available for the methods of surface pattern analysis are listed and briefly described by Legendre and Legendre (1998). They include *Geo-EAS*, *GEOSTAT*, *GSLib*, *ISATIS*, *Kellogg's*, *MACGRIDZO* and *UNIMAP*, which also include some spatial interpolation methods. The computing capacities of some popular statistical and GIS packages were compared by Hengl (2007). Some spatial interpolation methods are also available in *GS+* (Robertson, 2000).

Two types of the spatial interpolation methods, OK and UK, are provided in the *S+SpatialStats* module in *S-PLUS* (Kaluzny *et al*., 1998).

Table 8.1. Availability of the spatial interpolation methods in several commonly used software packages.

Method/ package	ArcGIS/ ArcView GIS	GS+	R												S-PLUS	SURFER
			stats	akima	deldir	fields	geoR	geoRglm	GRASS	gstat	spatial	sgeostat	RandomFields	tripack		
NN	yes				yes									yes		yes
TIN	yes				yes									yes	yes	yes
NaN	yes															yes
Cl		yes	yes												yes	
TSA	yes										yes				yes	
IDW	yes	yes								yes						yes
LM	yes		yes												yes	yes
TPS	yes		yes	yes		yes									yes	
SK	yes					yes?	yes			yes		yes?	yes			
OK	yes	yes					yes		yes	yes			yes		yes	yes
UK	yes					yes	yes			yes	yes				yes	yes
SCK	yes									yes						
OCK	yes	yes								yes						
Universal CK	yes									yes						
BK	yes									yes					yes	
SCCK										yes						
KED							yes			yes						
StOK/StSK		yes								yes						
IK	yes									yes						
MBK							yes	yes								
Simulation										yes						

8.2. Important Factors and Recommendation

8.2.1. Important Factors

There are many potential factors that should be considered as secondary information for the spatial interpolation of marine environmental data. For instance, in marine science, combined flow bed shear stress (*i.e.*, a combination of the effects of surface ocean waves, tidal, wind and density driven ocean currents) strongly influence benthic habitats on the continental shelf by mobilising sediments or directly influencing organisms. It was suggested that both magnitude and frequency of combined-flow bed shear stresses must be considered when characterising the benthic environment (Hemer, 2006).

Bathymetry has been used to improve the performance of the spatial interpolators (Verfaillie *et al*., 2006). The relation between the bathymetry and grain-size depends on the morphology, topography, and the substrate type (Verfaillie *et al*., 2006), so inclusion of such information would probably further improve the estimations. Distance to coastline may be important in improving the prediction of geospatial data. Other factors like those used in Whiteway *et al*. (2007) may also provide useful information to improve the prediction.

As in statistical analysis, understanding the mechanisms underpinning the observed phenomena and incorporation of professional knowledge in the estimation could improve the performance of the spatial interpolation methods. For example, inclusion of in-water distance instead of Euclidean distance and trend component improved the prediction accuracy by a reduction of 10-30% of the prediction error variance in a study for predicting contaminant and water quality variables in an estuary (Little *et al*., 1997).

There are also many other factors affecting the performance of the spatial interpolation methods as discussed in previous chapters. These factors should be considered in data collection, field survey design and selection of the spatial interpolation methods. Irregular sampling design may be preferred over regular one, but not for splines.

8.2.2. Recommendation for Marine Environmental Scientists

A number of the spatial interpolation methods show their strength in practical application. For instance, RK-C and GIDS are less sensitive to the variation in the data and more accurate than other methods. KED and OCK have also proven to obtain high accuracy when appropriate high quality secondary information is available. RK-D and those methods highlighted in Table 5.2 are also worth a further investigation. OK and UK could be good candidates when the correlation between the primary variable and secondary variables are weak. TPS should be considered if no spatial structure is detected. IDW, although performs poorly in most cases, should provide a good control as it is a standard spatial interpolation tool used for geospatial data in Geoscience Australia. In addition, GWR is worth attention in the future.

Stratification could improve the estimation of the spatial interpolators when relevant information is available by reducing the variance of the data (Stein *et al*., 1988; Voltz and Webster, 1990). The geomorphic features of Australian continental margin (Harris *et al*., 2005; Heap and Harris, 2008) would provide valuable information for employing stratification method together with the spatial interpolation techniques. If such information stratifies relevant Australian marine environmental variables so that the variance within each geomorphic feature is reduced, then the accuracy of spatial interpolation of the environmental variables is expected to be improved.

LR and its variants (Stahl *et al*., 2006) are developed specifically for air temperature in relation to elevation. Given this, they are probably not as applicable to other environmental science disciplines. However, if similar relationships can be found between seabed physical variables and bathymetry, LR could be possibly adopted to spatial interpolation of marine environmental variables.

Acknowledgements

We would like to thank Hideyasu Shimadzu and Brendan Brooke for their valuable comments and suggestions. We also thank Edzer Pebesma for his helpful explanation about computational aspect of geo-statistical methods in gstat package. We are grateful to Maggie Tran for her careful proofreading. Tara Anderson, Shoaib Burq, David Ryan and Frederic Saint-Cast are acknowledged for providing references relevant to marine environmental science. This record is published with permission of the Chief Executive Officer, Geoscience Australia

References:

Ahmed, S. and De Marsily, G., 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resources Research, 23(9): 1717-1737.

Akima, H., 1978. A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points. ACM Transactions on Mathematical Software, 4(2): 148-159.

Akima, H., 1996. Algorithm 761: scattered-data surface fitting that has the accuracy of a cubic polynomial. ACM Transactions on Mathematical Software, 22: 362-371.

Armstrong, M. and Matheron, G., 1986a. Disjunctive kriging revisited: Part I. Mathematical Geology, 18(8): 711-728.

Armstrong, M. and Matheron, G., 1986b. Disjunctive kriging revisited: Part II. Mathematical Geology, 18(8): 729-741.

Asli, M. and Marcotte, D., 1995. Comparison of approaches to spatial estimation in a bivariate context. Mathematical Geology, 27: 641-658.

Balk, B. and Elder, K., 2000. Combining binary decision tree and geostatistical methods to estimate snow distribution in a mountain watershed. Water Resources research, 36(1): 13-26.

Bishop, T.F.A. and McBratney, A.B., 2001. A comparison of prediction methods for the creation of field-extent soil property maps. Geoderma, 103: 149-160.

Boufassa, A. and Armstrong, M., 1989. Comparison between different kriging estimators. Mathematical Geology, 21(3): 331-345.

Bourennane, H., King, D. and Couturier, A., 2000. Comparison of kriging with external drift and simple linear regression for predicting soil horizon thickness with different sample densities. Geoderma, 97: 255-271.

Bregt, A.K., 1992. Processing of soil survey data, Agricultural University of Wageningen, The Netherlands.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J., 1984. Classification and regression trees. Wadsworth, Belmont.

Brus, D.J. and de Gruijter, J.J., 1994. Estimating of non-ergodic variograms and their sampling variance by design-based sampling strategies. Mathematical Geology, 26(4): 437-454.

Brus, D.J. *et al*., 1996. The performance of spatial interpolation methods and choropleth maps to estimate properties at points: a soil survey case study. Environmetrics, 7: 1-16.

Burrough, P.A., 1991. Principles of Geographical Information Systems for Land Resources Assessment. Monographs on Soil and Resources Survey, 12. Oxford Press, Oxford, 194 pp.

Burrough, P.A. and McDonnell, R.A., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford, 333 pp.

Carr, J.R. and Deng, E.D., 1987. Comparison of two techniques for applying disjunctive kriging: The Gaussian anamorphosis model versus the direct statistical inference of the bivariate distributions. Mathematical Geology, 19(1): 57-68.

Chambers, J.M. and Hastie, T.J., 1992. Statistical Models in S. Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, CA, 608 pp.

Chang, Y.H., Scrimshaw, M.D., Emmerson, R.H.C. and Lester, J.N., 1998. Geostatistical analysis of sampling uncertainty at the Tollesbury Managed

Retreat site in Blackwater Estuary, Essex, UK: Kriging and cokriging approach to minimise sampling density. The Science of the Total Environment, 221: 43-57.

Clark, I. and Harper, W.V., 2001. Practical Geostatistics 2000. Geostokos (Ecosse) Limited, 342 pp.

Cleveland, W.S. and Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting. Journal of the American Statistical Association, 83(403): 596-610.

Collins, F.C. and Bolstad, P.V., 1996. A comparison of spatial interpolation techniques in temperature estimation, Proceedings, Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, NM. Santa Barbara, CA: National Center for Geographic Information and Analysis, Santa Barbara.

Cressie, N.A., 1993. Statistics for Spatial Data. John Wiley & Sons, Inc., New York, 900 pp.

Danielsson, A., Carman, R., Rahm, L. and Aigars, J., 1998. Spatial estimation of nutrient distributions in the Gulf of Riga sediments using cokriging. Estuarine, Coastal and Shelf Science, 46: 713-722.

Davis, J.C., 1973. Statistics and Data Analysis in Geology. John Wiley & Sons, Inc., New York, 550 pp.

Diggle, P.J. and Ribeiro Jr., P.J., 2007. Model-based Geostatistics. Springer, New York, 228 pp.

Diggle, P.J., Tawn, J.A. and Moyeed, R.A., 1998. Model-based geostatistics. Applied Statistics, 47(3): 299-350.

Dirks, K.N., Hay, J.E., Stow, C.D. and Harris, D., 1998. High-resolution studies of rainfall on Norfolk Island Part II: interpolation of rainfall data. Journal of Hydrology, 208: 187-193.

Emery, X., 2006. Ordinary multigaussian kriging for mapping conditional probabilities of soil properties. Geoderma, 132: 75-88.

Englund, E., Weber, D. and Leviant, N., 1992. The effects of sampling design parameters on block selection. Mathematical Geology, 24(3): 329-343.

Erxleben, J., Elder, K. and Davis, R., 2002. Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains. Hydrological Processes, 16: 3627-3649.

Fotheringham, A.S., Brunsdon, C. and Charlton, M., 2002. Geographically Weighted Regression: the analysis of spatially varying relationships. John Wiley & Sons, Ltd, Chichester, 269 pp.

Gaus, I., Kinniburgh, D.G., Talbot, J.C. and Webster, R., 2003. Geostatistical analysis of arsenic concentration in groundwater in Bangladesh using disjunctive kriging. Environmental Geology, 44(8): 939-948.

Gilmour, A., Cullis, B., Welham, S., Gogel, B. and Thompson, R., 2004. An efficient computing strategy for prediction in mixed linear models. Computational Statistics & Data Analysis, 44: 571-586.

Goff, J.A., Jenkins, C.J. and Williams, S.J., 2008. Seabed mapping and characterization of sediment variability using the usSEABED data base. Continental Shelf Research, 28: 614-633.

Gold, C.M. and Condal, A.R., 1995. A spatial data structure integrating GIS and simulation in a marine environment. Marine Geodesy, 18: 213-228.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York, 483 pp.

Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology, 228: 113-129.

Gotway, C.A., Ferguson, R.B., Hergert, G.W. and Peterson, T.A., 1996. Comparison of kriging and inverse-distance methods for mapping parameters. Soil Science Society of American Journal, 60: 1237-1247.

Gotway, C.A. and Stroup, W.W., 1997. A generalized linear model approach to spatial data analysis and prediction. Journal of Agricultural, Biological, and Environmental Statistics, 2(2): 157-178.

Greenwood, D.J., Neeteson, J.J. and Draycott, A., 1985. Response of potatoes to N fertilizer: dynamic model. Plant Soil, 85: 185-203.

Guarini, J.-M. *et al*., 1998. Dynamics of spatial patterns of microphytobenthic biomass: inferences from a geostatistical analysis of two comprehensive surveys in Marennes-Oleron Bay (France). Marine Ecology Progress Series, 166: 131-141.

Haberlandt, U., 2007. Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. Journal of Hydrology, 332: 144-157.

Harris, P. *et al*., 2005. Geomorphic Features of the Continental Margin of Australia, Geoscience Australia.

Hartkamp, A.D., De Beurs, K., Stein, A. and White, J.W., 1999. Interpolation Techniques for Climate Variables, CIMMYT, Mexico, D.F.

Heap, A.D. and Harris, P.T., 2008. Geomorphology of the Australian margin and adjacent seafloor. Australian Journal of Earth Sciences, 55: 555-585.

Hemer, M.A., 2006. The magnitude and frequency of combined flow bed shear stress as a measure of exposure on the Australian continental shelf. Continental Shelf Research, 26: 1258-1280.

Hengl, T., 2007. A Practical Guide to Geostatistical Mapping of Environmental Variables, Luxembourg: Office for Official Publication of the European Communities.

Hengl, T., Heuvelink, G.B.M. and Rossiter, D.G., 2007. About regression-kriging: From equations to case studies. Computer & Geosciences, 33: 1301-1315.

Hengl, T., Heuvelink, G.B.M. and Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma, 120: 75-93.

Hernandez-Stefanoni, J.L. and Ponce-Hernandez, R., 2006. Mapping the spatial variability of plant diversity in a tropical forest: comparison of spatial interpolation methods. Environmental Monitoring and Assessment, 117: 307-334.

Hosseini, E., Gallichand, J. and Caron, J., 1993. Comparison of several interpolators for smoothing hydraulic conductivity data in South West Iran. American Society of Agricultural Engineers, 36(6): 1687-1693.

Hu, K., Li, B., Lu, Y. and Zhang, F., 2004. Comparison of various spatial interpolation methods for non-stationary regional soil mercury content. Environmental Science, 25(3): 132-137.

Hutchinson, M.F., 1995. Interpolating mean rainfall using thin plate smoothing splines. International Journal of Geographical Information Systems, 9(4): 385-403.

Hutchinson, M.F. and Gessler, P.E., 1994. Splines - more than just a smooth interpolator. Geoderma, 62: 45-67.

ICES, 2005. Report of the Working Group on Marine Habitat Mapping (WGMHM), ICES CM 2005/E:05, Bremerhaven, Germany.

Isaaks, E.H. and Srivastava, R.M., 1989. Applied Geostatistics. Oxford University Press, New York, 561 pp.

Jarvis, C.H. and Stuart, N., 2001. A comparison among strategies for interpolating maximum and minimum daily air temperature. Part II: the interaction between number of guiding variables and the type of interpolation method. Journal of Applied Meteorology, 40: 1075-1084.

Jef, H., Clements, M., Gerwin, D. and Frans, F., 2006. Spatial interpolation of ambient ozone concentrations from sparse monitoring points in Belgium. Journal of Environmental Monitoring, 8: 1129-1135.

Jerosch, K., Schlüter, M. and Roland, P., 2006. Spatial analysis of marine categorical information using indicator kriging applied to georeferenced video mosaics of the deep-sea Håkon Mosby Mud Volcano. Ecological Informatics, 1: 391-406.

Journel, A.G. and Huijbregts, C.J., 1978. Mining Geostatistics. Academic Press, London, 600 pp.

Juang, K.W. and Lee, D.Y., 1998. A comparison of three kriging methods using auxiliary variables in heavy-metal contaminated soils. Journal of Environmental Quality, 27: 355-363.

Kanevski, M., Timonin, V. and Pozdnoukhov, A., 2008. Automatic decision-oriented mapping of pollution data. ICCSA, 1: 678-691.

Kerry, R. and Oliver, M.A., 2007. Comparing sampling needs for variograms of soil properties computed by the method of moments and residual maximum likelihood. Geoderma, 140(4): 383-396.

Knotters, M., Brus, D.J. and Oude Voshaar, J.H., 1995. A comparison of kriging, co-kriging and kriging combined with regression for spatial interpolation of horizon depth with censored observations. Geoderma, 67: 227-246.

Kravchenko, A.K. and Bullock, D.G., 1999. A comparison study of interpolation methods for mapping soil properties. Agronomy Journal, 91: 393-400.

Krige, D.G., 1951. A statistical approach to some mine valuations problems at the Witwatersrand. Journal of the Chemical, Metallurgical and Mining Society of South Africa, 52: 119-139.

Kaluzny, S.P., Vega, S.C., Cardoso, T.P. and Shelly, A.A., 1998. S+SpatialStats: User's Manual for Windows and UNIX. Springer, New York.

Lam, N.S.-N., 1983. Spatial interpolation methods: a review. The American Cartographer, 10(2): 129-139.

Lark, R.M., Cullis, B.R. and Welham, S.J., 2006. On spatial prediction of soil properties in the presence of a spatial trend: the empirical best linear unbiased predictor (E-BLUP) with REML. European Journal of Soil Science, 57(6): 787-799.

Lark, R.M. and Ferguson, R.B., 2004. Mapping risk of soil nutrient deficiency or excess by disjunctive and indicator kriging. Geoderma, 118: 39-53.

Laslett, G.M., 1994. Kriging and splines: an empirical comparison of their predictive performance in some applications. Journal of the American Statistical Association, 89(426): 391-400.

Laslett, G.M. and McBratney, A.B., 1990. Further comparison of spatial methods for predicting soil pH. Soil Science Society of America Journal, 54: 1553-1558.

Laslett, G.M., McBratney, A.B., Pahl, P.J. and Hutchinson, M.F., 1987. Comparison of several spatial prediction methods for soil pH. Journal of Soil Science, 38: 325-341.

Legendre, P. and L., L., 1998. Numerical Ecology. ELSEVIER, Amsterdam, 853 pp.
Lehmann, A., 1997. A GIS approach of aquatic plant spatial heterogeneity in relation to sediment and depth gradients, Lake Geneva, Switzerland. Aquatic Botany, 58: 347-361.
Li, X., Cheng, G. and Lu, L., 2005. Spatial analysis of air temperature in the Qinghai-Tibet Plateau. Arctic, Antarctic, and Alpine Research, 37(2): 246-252.
Li, Y., Shi, Z., Wu, C., Li, H. and Li, F., 2007. Improved prediction and reduction of sampling density for soil salinity by different geostatistical methods. Agricultural Science in China, 6(7): 832-841.
Lin, G. and Chen, L., 2004. A spatial interpolation method based on radial basis function networks incorporating a semivariogram model. Journal of Hydrology, 288: 288-298.
Little, L.S., Edwards, D. and Porter, D.E., 1997. Kriging in estuaries: as the cow flies, or as the fish swims? Journal of Experimental Marine Biology and Ecology, 213: 1-11.
Lucio, P.S., Bodevan, E.C., Dupont, H.S. and Ribeiro, L.V., 2006. Directional kriging: a proposal to determine sediment transport. Journal of Coastal Research, 22(6): 1340-1348.
MacEachren, A.M. and Davidson, J.V., 1987. Sampling and isometric mapping of continuous geographic surfaces. The American Cartographer, 14(4): 299-320.
Marchant, B.P. and Lark, R.M., 2006. Adaptive sampling and reconnaissance surveys for geostatistical mapping of the soil. European Journal of Soil Science, 57: 831-845.
Mardikis, M.G., Kalivas, D.P. and Kollias, V.J., 2005. Comparison of interpolation methods for the prediction of reference evapotranspiration - an application in Greece. Water Resources Management, 19: 251-278.
Martínez-Cob, A., 1996. Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. Journal of Hydrology, 174: 19-35.
Matheron, G., 1963. Principles of geostatistics. Economic Geology, 58: 1246-1266.
Matheron, G., 1969. Le Krigeage universel. Cahiers du Centre de Morphologie Mathematique, Ecole des Mines de Paris, Fontainebleau.
Maynou, F.X., Sardà, F. and Conan, G.Y., 1998. Assessment of the spatial structure and biomass evaluation of *Nephrops norvegicus* (L.) populations in the northwestern Mediterranean by geostatistics. ICES Journal of Marine Science, 55: 102-120.
Meul, M. and Van Meirvenne, M., 2003. Kriging soil texture under different types of nonstationarity. Geoderma, 112: 217-233.
Minasny, B. and McBratney, A.B., 2007. Spatial prediction of soil properties using EBLUP with the Matern covariance function. Geoderma, 140: 324-336.
Mitasova, H. *et al*., 1995. Modelling spatially and temporally distributed phenomena: new methods and tools for GRASS GIS. International Journal of Geographical Information Systems, 9(4): 433-446.
Moyeed, R.A. and Papritz, A., 2002. An empirical comparison of kriging methods for nonlinear spatial point prediction. Mathematical Geology, 34(4): 365-386.
Nalder, I.A. and Wein, R.W., 1998. Spatial interpolation of climatic Normals: test of a new method in the Canadian boreal forest. Agricultural and Forest Meteorology, 92: 211-225.
Naoum, S. and Tsanis, I.K., 2004. Ranking spatial interpolation techniques using a GIS-based DSS. Global Nest: the International Journal, 6(1): 1-20.
Odeh, I.O.A., McBratney, A.B. and Chittleborough, D.J., 1994. Spatial prediction of

soil properties from landform attributes derived from a digital elevation model. Geoderma, 63: 197-214.

Odeh, I.O.A., McBratney, A.B. and Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma, 67: 215-226.

Oliver, M.A., Webster, R. and McGrath, S.P., 1996. Disjunctive kriging for environmental management. Environmetrics, 7: 333-358.

Papritz, A. and Moyeed, R.A., 1999. Linear and non-linear kriging methods: tools for monitoring soil pollution. In: V. Barnett, A. Stein and K.F. Turkman (Editors), Statistical Aspects of Health and the Environment. Statistics for the Environment. John Wiley & Sons, Chichester, pp. 303-336.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Computer & Geosciences, 30: 683-691.

Pebesma, E.J., Duin, R.N.M. and Burrough, P.A., 2005. Mapping sea bird densities over the North Sea: spatially aggregated estimates and temporal changes. Environmetrics, 16(6): 573-587.

Petitgas, P., 1993. Use of a disjunctive kriging to model areas of high pelagic fish density in acoustic fisheries surveys. Aquatic Living Resources, 6: 201-209.

Porter, D.E., Edwards, D., Scott, G., Jones, B. and Street, W.S., 1997. Assessing the impacts of anthropogenic and physiographic influences on grass shrimp in localized salt-marsh estuaries Aquatic Botany, 58: 289-306.

Puente, C.E. and Bras, R.L., 1986. Disjunctive kriging, universal kriging, or no kriging: Small sample results with simulated fields. Mathematical Geology, 18(3): 287-305.

R Development Core Team, 2007. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Rendu, J.-M., 1980. Disjunctive kriging: Comparison of theory with actual results. Mathematical Geology, 12(4): 305-320.

Ripley, B.D., 1981. Spatial Statistics. John Wiley & Sons, New York, 252 pp.

Rivoirard, J. and Wieland, K., 2001. Correcting for the effect of daylight in abundance estimation of juvenile haddock (*Melanogrammus aeglefinus*) in the North sea: application of kriging with external drift. ICES Journal of Marine Science, 58: 1272-1285.

Robertson, G.P., 2000. GS^+: Geostatistics for the Environmental Science. Gamma Design Software, Plainwell, Michigan USA.

Robertson, G.P. *et al.*, 1997. Soil resources, microbial activity, and primary production across an agricultural ecosystem. Ecological Applications, 7: 158-170.

Rossi, R.E., Mulla, D.J., Journel, A.G. and Franz, E.H., 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecological Monographs, 62(2): 277-314.

Ruddick, R., 2006? Data interpolation methods in the Geoscience Australia seascape maps, Geoscience Australia, Canberra.

Schloeder, C.A., Zimmerman, N.E. and Jacobs, M.J., 2001. Comparison of methods for interpolating soil properties using limited data. Soil Science Society of American Journal, 65: 470-479.

Sibson, R., 1981. A brief description of natural neighbour interpolation. In: V. Barnett (Editor), Interpreting Multivariate Data. John Wiley and Sons, Chichester, pp. 21-36.

Silverman, B.W., 1981. Density estimation for univariate and bivariate data. In: V.

Barnett (Editor), Interpreting Multivariate Data. John Wiley and Sons, Chichester, pp. 37-53.
Stahl, K., Moore, R.D., Floyer, J.A., Asplin, M.G. and McKendry, I.G., 2006. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. Agricultural and Forest Meteorology, 139: 224-236.
Stein, A., Hoogerwerf, M. and Bouma, J., 1988. Use of soil map delineations to improve (co-)kriging of point data on moisture deficits. Geoderma, 43: 163-177.
Stineman, R.W., 1980. A consistently well behaved method of interpolation. Creative Computing, 6(7): 54-57.
Sun, W., 1998. Comparison of a cokriging method with a Bayesian alternative. Environmetrics, 9: 445-457.
Suro-Pérez, V. and Journel, A.G., 1991. Indicator principal component kriging. Mathematical Geology, 23(5): 759-788.
United Nations, 1993. Convention on Biological Diversity. Treaty Series: No. 30619. Concluded at Rio de Janeiro on 5 June 1992.
Van Kuilenburg, J., De Gruijter, J.J., Marsman, B.A. and Bouma, J., 1982. Accuracy of spatial interpolation between point data on soil moisture supply capacity, compared with estimates from mapping units. Geoderma, 27: 311-325.
Venables, W.N. and Ripley, B.D., 2002. Modern Applied Statistics with S-Plus. Springer-Verlag, New York, 495 pp.
Verfaillie, E., van Lancker, V. and van Meirvenne, M., 2006. Multivariate geostatistics for the predictive modelling of the surficial sand distribution in shelf seas. Continental Shelf Research, 26: 2454-2468.
Vicente-Serrano, S.M., Saz-Sánchez, M.A. and Cuadrat, J.M., 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. Climate Research, 24: 161-180.
Voltz, M. and Webster, R., 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. Journal of Soil Science, 41: 473-490.
Wackernagel, H., 2003. Multivariate Geostatistics: An Introduction with Applications. Springer, Berlin, 387 pp.
Wahba, G. and Wendelberger, J., 1980. Some new mathematical methods for variational objective analysis using splines and cross-validation. Monthly Weather Review, 108: 1122-1145.
Wang, H., Liu, G. and Gong, P., 2005. Use of cokriging to improve estimates of soil salt solute spatial distribution in the Yellow River delta. Acta Geographica Sinica, 60(3): 511-518.
Wang, J., Li, Q., Fischer, M.M., Chen, J. and Chen, H., 1996. Adaptive structure model of earthquake trend regionalization. Earthquake Research in China, 12 (supplement): 78-88.
Weber, D. and Englund, E., 1992. Evaluation and comparison of spatial interpolators. Mathematical Geology, 24(4): 381-391.
Webster, R. and Oliver, M., 2001. Geostatistics for Environmental Scientists. John Wiley & Sons, Ltd, Chichester, 271 pp.
Webster, R. and Oliver, M.A., 1992. Sample adequately to estimate variograms of soil properties. Journal of Soil Science, 43: 177-192.
Welham, S., Cullis, B.R., Gogel, B., Gilmour, A. and Thompson, R., 2004. Prediction

in linear mixed models. Australian and New Zealand Journal of Statistics, 46: 325-347.
Whiteway, T. *et al*., 2007. Seascapes of the Australian Margin and Adjacent Sea Floor: Methodology and Results, Geoscience Australia.
Willmott, C.J., 1981. On the validation of models. Physical Geography, 2: 184-194.
Willmott, C.J., 1982. Some comments on the evaluation of model performance. Bulletin American Meteorological Society, 63(11): 1309-1313.
Willmott, C.J. and Matsuura, K., 1995. Smart interpolation of annually averaged air temperature in the United States. Journal of Applied Meteorology, 34: 2577-2586.
Wu, J., Norvell, W.A. and Welch, R.M., 2006. Kriging on highly skewed data for DTPA-extractable soil Zn with auxiliary information for pH and organic carbon. Geoderma, 134: 187-199.
Yates, S.R., Warrick, A.W. and Myers, D.E., 1986. Disjunctive kriging: 1. Overview of estimation and conditional probability Water Resources Research, 22(5): 615-621.
Zhou, F., Guo, H.-C., Ho, Y.-S. and Wu, C.-Z., 2007. Scientometric analysis of geostatistics using multivariate methods. Scientometrics, 73(3): 265-279.
Zimmerman, D., Pavlik, C., Ruggles, A. and Armstrong, M.P., 1999. An experimental comparison of ordinary and universal kriging and inverse distance weighting. Mathematical Geology, 31(4): 375-390.

Appendices

Appendix A. Applications of Spatial interpolation Methods in Various Disciplines

A.1. Meteorology and Water Resources

A total of 16 studies that compared the performance of the spatial interpolation methods in meteorology and water resources have been reviewed in this section.

Study1: OK, UK with a linear drift (UK-LD), UK with quadratic drift (UK-QD), TSA (with first, second, third and fourth order), IDW (with distance powers 0, 1, 2, 3 and 4) and AK were employed for smoothing hydraulic conductivity data using 383 samples, 500 m apart over an area of 16,000 ha (Hosseini *et al*., 1993). Leave-one-out cross-validation was used for validating the performance of these spatial interpolators. The best methods were OK, UK with a linear drift and IDS in terms of MAE. Although overall, the poor performance of all methods due to the very high variation (coefficient of variation (CV) = 78.68%) in the data, OK was considered the most appropriate method because of its precision and the smoothness of its interpolated surface. The similarity in the precision of these methods might be due to the high sampling density, as discussed in the Chapter 6.

Study 2: IDS, Optimal IDW (OIDW, the power parameter is chosen based on the minimum MAE), cubic splines, LM, TSA, LR, kriging, and CK were compared in temperature estimation (*i.e.*, minimum and maximum) for three temporal scales (10 year average, seasonal and daily) in two regions (Collins and Bolstad, 1996). The results show that:

1) IDS performed consistent poorly across all temporal and regional scales. Where data were sparse, the results were implausible. It also suffered from discontinuities at station locations resulting in temperature peaks (*i.e.*, "birds eye" patterns in the interpolated surface). OIDW was recommended over IDS because OIDW would always yield equal or better MAE results than IDS since the power parameter was chosen on the basis of minimum MAE. When the data were isotropic, IDS had lower MAE values than kriging. OIDW, however, was not always as visually plausible as kriging. When the data were not correlated and isotropic, the most preferred method turned out to be OIDW.
2) TSA was not recommended for temperature interpolation because: a) it was not representative of the original data range as its interpolated temperature range was typically narrower than the original data range; b) although it tended to capture broad regional trends, due to bias introduced by multicollinearity, these trends

were suspect; and c) where station distribution resulted in extrapolation beyond the convex data hull, the estimated temperatures were well beyond the original data range.

3) Splines generated poor visual and cross validation results. Where data variances were high, splines tended to have interpolated values well outside the observed data range. Splines also produced high MAE values across all temporal scales, across both regions, for both minimum and maximum temperatures. Splines performed much better when dense, regularly-spaced data were available. A similar conclusion was made by Hutchinson and Gessler (1994) that cubic splines was generally not recommended for interpolation of irregularly-spaced data. Splines are useful for quickly obtaining a clear map showing the main features of the variable, but they are not an accurate spatial interpolator. Generally speaking, splines produced more outliers than kriging.
4) LM was clearly superior to all other methods, with the lowest MAE value of the eight methods considered. The performance of LM did not appear to be affected by data range. However, care must be taken with LM to ensure the results are representative of the original data range. Where station elevations are not representative of regional elevations, care must be taken in comparing observed and interpolated data.
5) While LR performed poorly in terms of MAE, its results were more plausible than methods that did not use elevation as ancillary information. When elevation and temperature were not correlated, LR degraded into a NN where estimated values simply took on the value of the nearest station point. LR resulted in some banding effects and "island-like" isothermal tessellations around certain influential stations. Outlier stations were less noticeable with LM than with LR. In Stahl *et al.*'s study (2006), it was found that methods that compute local lapse rates performed better for datasets with a greater number of higher-elevation stations.
6) Kriging produced better results than OIDW when the data were anisotropic. Kriging appeared similar to splines, but had lower MAE values than splines for every case tested. Perhaps the greatest advantage of kriging is that the geostatistical process provides the users greater information about the spatial variability of the regionalised variable of interest via the semivariogram and variogram surfaces.
7) CK produced visually implausible results and when elevation and temperature were not correlated. Its overall performance beared a strong resemblance to kriging. However, further evaluation is not possible because the methods of kriging and cokriging method used in this study were not stated.
8) Stronger correlations between elevation and temperature favoured LM and LR.

LR, CK and LM were found to be inappropriate when correlations between temperature and elevation were below 0.72. Inverse distance squared, optimal inverse distance, and kriging showed similar robustness to a priori data range, correlation (between elevation and temperature), and variance. Of all the methods assessed, splines seemed to be most sensitive to a priori data characteristics. Kriging was favoured over optimal inverse distance when data were anisotropic. When data were isotropic, OIDW was favoured. When data variance was high or correlation between temperature and elevation were low, CK produced specking or "birds eye" effects around station locations.

Study 3: OK, OCK and RK-C were used to interpolate long-term mean total annual reference evapotranspiration and long-term mean total annual precipitation for a 47,000 km^2 complex topographic region (Martínez-Cob, 1996). A total of 108 and 132 samples were used for estimation for evapotranspiration and precipitation respectively, and 50 samples for validation. Elevation was used as secondary information. It was found that OCK was more accurate thane OK and RK-C, and that RK-C did not improve OK results. The good performance of OCK was probably due to a good correlation between the primary and secondary variables.

Study 4: GIDS, IDS, NN, CK, OK, RK-C, and UK were compared for the spatial interpolation of monthly temperature and monthly precipitation using 32 samples with elevation as the secondary information (Nalder and Wein, 1998). Leave-one-out method was used for validation. GIDS produced the lowest MAE and RMSE while providing low MEs for both temperature and precipitation. It also produced a more consistent performance than any other method from month to month. However, it was not significantly better than RK-C for temperature or better than OK, CK, NN and IDS for precipitation. Nevertheless, GIDS was considered preferable because it is robust and relatively simple to apply.

Study 5: MWRCK, CK and LSZ (a Bayesian alternative) were used to predict the NO_3 in rainfall using 48 monthly samples for 35 datasets with SO_4 as secondary information and leave-one-out cross validation was used (Sun, 1998). It was found that LSZ yielded a slightly smaller MSE than MWRCK. LSZ also produced an almost correct coverage probability, but MWRCK yielded a low coverage probability.

Study 6: IDW, TPS and OCK were evaluated using a DEM to a resolution of 1 km^2 over a 20,000 km^2 square region (Hartkamp *et al.*, 1999). A total of 169 samples of monthly precipitation were used for predictions and 25 samples for validation. For monthly mean maximum temperature 125 samples were used for estimation and 15 samples for validation. The validation showed no difference among the three methods for predicting precipitation. For maximum temperature, splines performed best. The rigid prerequisites of cokriging regarding the statistical properties of the data used (*e.g.*, normal distribution, non-stationarity), along with its computational demands, may put this approach at a disadvantage. Taking into account error prediction, data assumptions, and computational simplicity, TPS was recommended for interpolating climate variables.

Study 7: SKlm, KED, OCCK, LM, NN, IDS, and OK were compared using 36 samples of annual and monthly rainfall data in a 5000 km^2 region (Goovaerts, 2000). Leave-one-out cross validation was used for validating the performance of these spatial interpolators. This study revealed that: 1) three multivariate geostatistical methods (*i.e.*, SKlm, KED and OCCK) using elevation as a secondary variable outperformed the other spatial interpolators in terms of MSE; 2) SKlm yielded the best predictions; and it also provided an easier way to incorporate several secondary variables than the other two multivariate methods; 3) OK was better than LM when the correlation coefficient between rainfall and elevation was smaller than 0.75; 4) the three multivariate methods generally reduced the OK prediction error as long as the correlation coefficient was larger than 0.75; and the benefit of the multivariate techniques was marginal when the correlation was too small; and 5) when the correlation ranged between 0.4 and 0.7 and the secondary information (*i.e.*, elevation data) exhibited a much smaller relative nugget effect than rainfall data, inclusion of secondary information still improved the prediction, in particular when the nugget effect of their cross semivariogram was small.

Study 8: LM with IDW, TSA extended to include secondary information, RK-C (detrended OK) and partial TPS incorporating secondary information were applied to interpolating the maximum and minimum daily air temperatures to a resolution of 1 km using 174 samples in England and Wales region (Jarvis and Stuart, 2001). Several variables including elevation and land cover classes were used as secondary information or “guiding variables”. Prior to the inclusion of secondary variables, all methods produced similar estimates for both minimum and maximum temperatures in terms of RMSE, with the exception of TSA which was less accurate than the others. After incorporating the covariates, the performance of all methods improved considerably, and the differences in estimation accuracy among partial PTS, OK and IDW results were not significant although the performance of TSA was poorer. Best accuracies were achieved using partial TPS.

Study 9: IDW, OK, RK-C and OCK, and CART with OK and OCK were used to estimate snow depth using three datasets of 550 samples collected from an area of one km^2 within each of three study regions (Erxleben *et al.*, 2002). Elevation, slope, aspect, net solar radiation and vegetation were used as the secondary information. Leave-one-out cross-validation was used to assess the performance of these spatial interpolators. It was found that the tree-based models provided the most accurate estimate, explaining up to about 30% of the observed variability. Kriging of the regression tree residuals did not substantially improve the models. The poor performance of all models was largely due to the lack of spatial structure. The value for the power parameter for IDW was not provided in this study. However, it was found previously that combining regression tree and OK or OCK could explain 6-20% more of the variance and thus improve the estimation of snow depth over 6.9 km^2 in the same region (Balk and Elder, 2000).

Study 10: IDW (with power parameters 1, 2 and 3), LM, NN, splines (with a tension parameter of 400 and 500, and a smoothing parameter of 0 and 400), TSA (with first, second, third, fourth and fifth order), SK, OK, BK, OK with anisotropy, UK with linear drift, UK with quadratic drift, CK (perhaps OCK), LM plus IDS, and splines (with tension of 400) of the residuals of LM were applied to annual precipitation and temperature for a mountainous region of over 20,000 km^2 (Vicente-Serrano *et al.*, 2003). Several variables including longitude, latitude, distance to Mediterranean Sea, elevation, radiation and their two-way interaction were used as secondary information. These variables were highly correlated with the primary variable. A total of 99 samples were used for precipitation and 61 for temperature, of which 70% were used for estimation and the remaining 30% for validation. It was found that the geostatistical methods and a regression model generated the best estimations for precipitation and the regression-based method produced the most accurate results for temperature.

Study 11: RBFN, improved RBFN and OK were compared using 20 simulated datasets consisting of 25 samples randomly sampled from 100 evenly spacing grid points in a square area of 81 grid cells and the remaining 75 data points were used for validation (Lin and Chen, 2004). The data for each grid point of each dataset was generated using an exponential semivariogram model. These methods were also applied to 64 datasets of hourly rainfall records of 55 samples in a region of 2726 km^2 and leave-one-out cross validation was used. The results showed that improved RBFN was the best, then OK and the least is RBFN regardless of the arrangement of sample points in terms of RMSE. A similar pattern of performance was observed for the rainfall data. The improved RBFN had a higher computation speed for larger dataset as compared to OK.

Study 12: Two spline methods (regularized and tension), IDW, NN, LM and kriging with different variogram models were compared using daily rainfall records of sample size 187, 280 and 374 in a region of 41,284 km^2 with a resolution of 500 m, 1,000 m, 5,000 m, and 10,000 m (Naoum and Tsanis, 2004). The remaining 280, 187 and 93 samples out of the total 467 samples were probably employed for validation respectively. Kriging with exponential and universal_1 models showed consistent performance and provided reliable estimates in terms of MAE and standard deviation regardless of sample size and resolution. However, it is not clear which specific kriging method was used in this research and this study was conducted in ArcView GIS 3.2.

Study 13: IDS, OK, OCK and OK combined with LR were applied for spatial analysis of monthly mean air temperature using 90 samples in the Qinghai-Tibet Plateau with a resolution of 0.5° (Li *et al*., 2005). The results revealed that the performance of these methods was as follows: OK combined with LR performed the best, followed by COK and OK, and then IDS in terms of subjective analysis. The limited availability of the secondary information affected the performance of COK.

Study 14: GIDS, IDS, OK, and RK-C were compared for the prediction of reference evapotranspiration using 74 samples for estimation and 19 samples for validation in a region of 131,944 km^2 with a resolution of 7.62 km (Mardikis *et al*., 2005). GIDS was found to be the most accurate method in terms of MAE and RMSE. GIDS also performed better than NN, LM and several LR related methods (Stahl *et al*., 2006),.

Study 15: RK-C, IDW (with distance power 4) and LM with IDW were used for the spatial interpolation of ambient ozone concentrations to a resolution of 5 km from sparse monitoring points (sample size = 38) in Belgium (land area is 30,278 km^2). Population density data were used as auxiliary data (Jef *et al*., 2006). Leave-one-out method was used for validation. RK-C was the best spatial interpolator in terms of RMSE and it significantly improved the estimation in comparison with IDW.

Study 16: NN, IDS, OK, OIK, KED and indicator KED (IKED) were compared for predicting hourly precipitation (21 stations with 64 time steps) with secondary information from daily precipitation (281 stations), elevation and radar in a region of 25,000 km^2 with 14,436 grid cells (Haberlandt, 2007). It was found that KED and IKED clearly outperformed the other methods in terms of standardised RMSE. The best performance was achieved when all additional information were used simultaneously with KED. IKED produced, in some cases, smaller RMSEs than the compared methods, which used the original data, but at the expense of a significant loss of variance.

A.2. Ecology

Only one study was found in ecology that compared the performance of the spatial interpolation methods. StOK was found to be the best in terms of the accuracy of the estimations (*i.e.*, MAE, RMSE and correlation coefficient) in comparison with six other methods that are OK, OCK, IDS, StOCK, StIDW with p=1, and Cl using 141 samples of plant diversity data collected from a tropical landscape mosaic of 64 km^2 in GS+ (Hernandez-Stefanoni and Ponce-Hernandez, 2006). Vegetation indices were used as secondary information and leave-one-out method was used for validation. The relative poor performance of OCK and StOCK was due to the poor correlation between the primary variable and secondary variables.

Geostatistical tools have been applied for modelling and interpreting ecological spatial dependence to many aspects of ecology and their application in this discipline dates back to 1960 (Rossi *et al.*, 1992). LM was used to predict the spatial distribution of biomass of three plant species using aquatic environmental variables including sediment and depth (Lehmann, 1997). OK was used to assess the impacts of disturbances on the distribution of grass shrimp in estuaries (Porter *et al.*, 1997).

A.3. Agriculture and Soil Science

A total of 25 studies have been reviewed in this section, which compared the performance of the spatial interpolation methods in agriculture and soil science. McBratney *et al*. (2003) summarised some application of the spatial interpolation methods in soil science.

Study 1: NN, IDS and OK were applied for predicting the soil moisture using 530 sample for estimation and 661 samples for validation in an area of 359 ha (Van Kuilenburg *et al.*, 1982). OK was the most accurate of these methods, but it was only marginally better than IDS in terms of RMSE.

Study 2: OCK, KED, RK-A and RK-B were applied for estimating transmissivity at a resolution of 3 km using 72 samples of transmissivity measurements and 235 samples of specific capacity data as secondary information in a region of 80 by 40 km (Ahmed and De Marsily, 1987). These methods were also applied to 15 simulated datasets. The following conclusions were drawn: 1) OCK was found to be the most rigorous method and should be used if the residuals of the regression of one variable on the other are spatially correlated and if the correlation coefficients between the primary and secondary variables are high, but it requires all variables have a significant number of common data points for a reasonable estimation of the cross variogram; 2) RK-A could be used only if the residuals of the regression are spatially uncorrelated or if the correlation coefficients between the primary and secondary variables are high, and it requires all variables to have a significant number of common data points to fit a regression; 3) KED could be used for an unlimited number of variables, like OCK and it does not require any common data points between the variables; and 4) RK-B did not show any advantage.

Study 3: TPS and OK were found to be the most accurate methods for estimating soil pH data using 121 samples collected in an area of 10 by 10 grid of 10 m spacing in comparison with global means and medians, NN, IDW-0 (averages of three nearest samples), IDS, AK, NaN, and quadratic TSA in terms of MSE based on 64 validation samples (Laslett *et al*., 1987).

Study 4: Laslett and McBratney (1990) further compared NN, TPS, AK, global kriging using generalised covariances (SK?) and REML UK (a global UK fitted by REML) using regularly spaced 121 soil samples as in study 3 to predict soil pH, with and without 80 additional samples to account for spatial variation at short distances. Two datasets of 121 and 64 samples were used for validation. The results showed that REML UK trend was consistently the best performing method. The inclusion of close data pairs usually improved the predictions of the methods, but no dramatic improvement occurred with REML UK as the best overall method for the dataset.

Study 5: SK, StSK, Cl, and a cubic spline were evaluated and compared for clay content data (Voltz and Webster, 1990). One dataset consists of 321 samples collected along a transect 3.2 km long at 10 m intervals; of which 107 samples were used for estimation and 212 for validation. Two other datasets were collected from point locations with 100 m spacing in an area of 92 ha. The first dataset consisted of 34 samples for estimation and 143 for validation and the second one with 114 samples for prediction and 63 for validation. StSK and SK were found to perform better than Cl and spline in terms of MSE, and StSK was more precise than SK.

Study 6: Cubic splines and SK were compared using two sampled surface datasets (Laslett, 1994). One dataset consisted of 1150 samples of heights measured at 1 micron intervals along the drum of a roller and the other dataset was collected along a transect at 365 sampling locations of 4 m spacing. The results revealed that SK sometimes outperformed the splines by considerable margins, particularly if the samples are highly clustered; and it never performed worse than the splines in terms of MSE.

Study 7: LM, OK, UK, OCK, RK-A, and RK-B were applied to 161 samples of four soil properties data in an area of 400 by 700 m and 71 samples were used for validation (Odeh *et al*., 1994). RK-A produced the best results for the depth of solum and subsoil clay, while RK-B gave the best estimation for the depth of bedrock and topsoil gravel in terms of RMSE. Generally speaking RK methods outperformed all the other methods. The poor performance of OK was largely due to the strong trend in the data as evidenced by the performance of LM.

Study 8: LM, OK, UK, isotopic OCK, heterotopic OCK, RK-A, RK-B and RK-C were further compared using the datasets in Study 7 by Odeh *et al*.(1995). In this study, both RK-C and heterotopic OCK performed well in terms of RMSE, and RK-C generally performed the best and was more flexible than heterotopic OCK.

Study 9: OK, OCK and RK-A (kriging combined with regression) were compared for 539, 141, 55 and 33 samples of non-stationary data (with a drift of degree 0, 1, and 2) for soil layer depth in an area of about 97 ha (Knotters *et al*., 1995). A total of 117 samples were used for validation. Soil electrical conductivity was used as auxiliary variable. RK-A was found to be the most accurate method in terms of RMSE.

Study 10: Cl, GM, IDS, OK, NN, IDW-0, and TPS and their combination with soil strata (*i.e.*, stratified using soil units) were used to estimate soil properties (thickness of A1 horizon, maximum areic mass of phosphate adsorbed by soil, mean highest water table, and mean lowest water table) using 188 samples (a square grid of 12 x 16 points, spaced 500 m apart) in a region of 6 x 8 km (Brus *et al*., 1996). A dataset consisting of 96 samples was used for validation. Soil units were used to stratify the data. It was found that: 1) IDS and OK were more reliable, although differences between methods were small and not statistically significant, 2) the stratification slightly improved the estimation but the effect was not statistically significant, and 3) the combined effect of methods and stratification was usually not significant. However, OK performed better for values near data points. Combined with soil map stratification, OK was a more reliable estimator in the sense that it estimated all soil properties well.

Study 11: OK and IDW (with distance power parameters 1, 2 and 4) were compared using two datasets of regularly spaced and high density samples (Gotway *et al*., 1996). The first dataset consisted of 255 soil N and organic matter samples collected from an area of 90 x 518 m with a sampling space of 6 x 15 m plus 60 extra samples at closer spacings of 0.76, 1.5, 2.3, 3.7, and 5.3 m. Samples were divided into subsets for the prediction (195, 119) and validation (119 and 136) respectively. The second dataset consisted of about 1388 soil N and organic matter samples collected from an area of 53 ha. Samples were again divided into subsets for the prediction (731, 657) and validation (657, 709) respectively, and two further subsets of 192 samples and 88 samples were sampled on a regular grid of size 48 x 48 m and 72 x 72 m respectively for prediction, and 657 samples were used for validation. The results revealed that 1) the accuracy of IDW tended to increase with the power of distance for datasets with a CV of <25%; 2) for datasets with greater variation (with a CV of >25%), IDW using high distance powers (2 or 4) can produce poor estimations; 3) OK was slightly better than IDW for all sampling scenarios in terms of MSE; 4) "the accuracy of predictions from kriging was generally unaffected by the coefficient of variation", but this statement is not correct because a further analysis showed that OK had a similar response to changes in CV for both soil nitrate (Figure A.1) and soil organic matter (Figure A.2); and 5) the use of wider sampling spacings greatly reduced the information in the resultant maps.

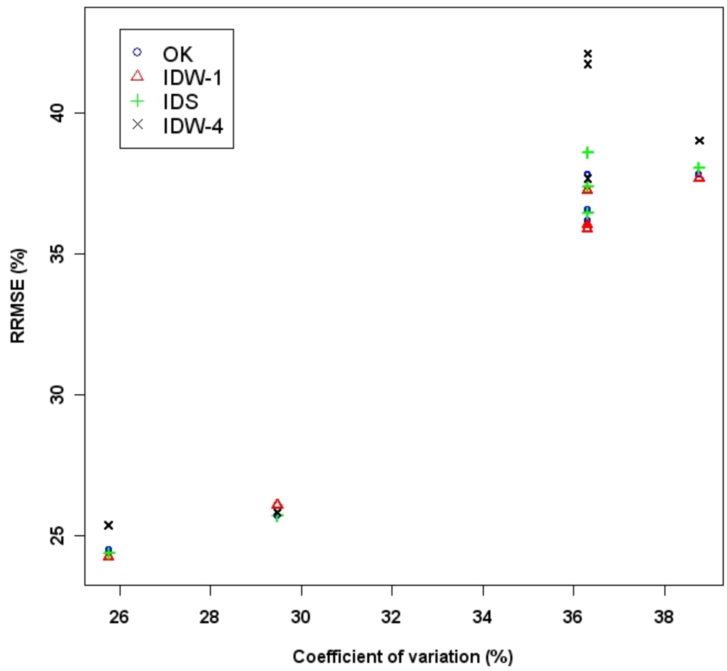


Figure A.1. Impacts of the variation in the dataset of soil nitrate on the performance of four spatial interpolation methods based on the results of Gotway *et al*. (1996).

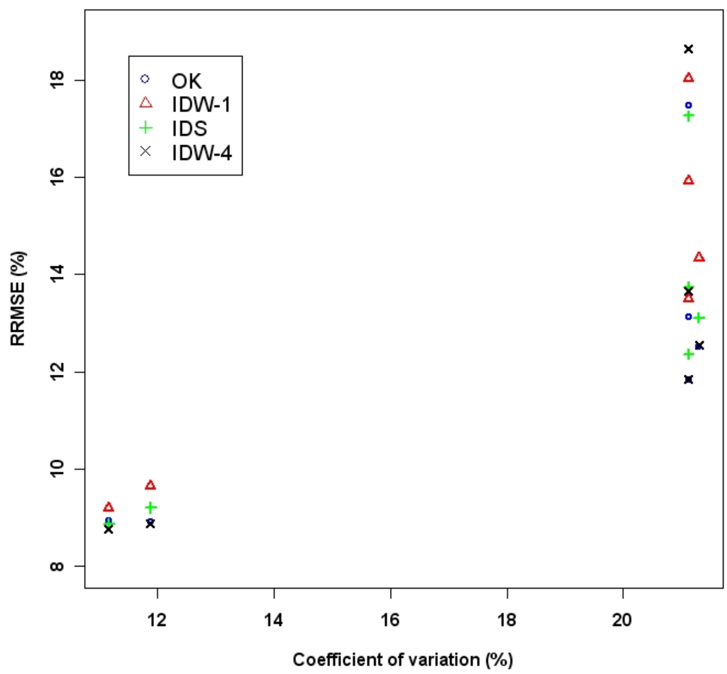


Figure A.2. Impacts of the variation in the dataset of soil organic matter on the performance of four spatial interpolation methods based on the results of Gotway *et al*. (1996).

Study 12: Goovaerts (1997) compared several geostatistical methods using the Jura dataset collected at Lausanne. The results showed that OCK was compared with SCK, SOCK and OCCK. OCK estimates better followed the data fluctuations than SCK estimates. OCK and SOCK performed equally. OCCK only slightly reduced the accuracy of estimation in comparison with OCK.

Study 13: KED and SKlm estimates displayed similar long-range features, but KED yielded more local details and might produce unacceptable negative estimations for soil mineral concentration using the Jura dataset (Goovaerts, 1997).

Study 14: OIK and OICK with a single unbiasedness constraint produced similar estimations except that OICK resulted in more variable estimations beyond data range and OICK might produce estimations outside the data range such as predicted probabilities being negative or >1 (Goovaerts, 1997). PK, as a special case of OICK, shared the same features of OICK as discussed above.

Study 15: OK, lognormal OK and IDW (with distance powers 1, 2, 3 and 4) were compared for predicting soil P and K using 30 datasets of sample size ranging from 36 to 1752, with sample spacing from 25 to 100 m (Kravchenko and Bullock, 1999). The results showed that OK with the optimal number of the neighbouring samples, a careful selected variogram model and appropriate log-transformation generally performed better than IDW in terms of ME and MAE.

Study 16: KED and LM were applied for predicting the thickness of a silty-clay-loam horizon with different sample densities (40, 50, 75, 100, 125 and 150) in an area of 380 ha and with a resolution of 20 m (Bourennane *et al*., 2000). A total of 69 samples were used for validation. The results showed that irrespective of sample size, KED estimates were on average more accurate than LM in terms of RMSE. KED performed better when the sample size increased, but the performance of LM remained unchanged over all sample sizes. Moreover, KED performed even better at a sample size of 40 than LM at whatever the sample size. KED can improve estimations, resulting in a considerable reduction of sampling intensity and while maintaining high prediction accuracy.

Study 17: GAM, LM, CART, OK, KED, RK-F and RK-C were compared for soil cation exchange capacity with a number of secondary variables in an area of 74 ha (Bishop and McBratney, 2001). A total of 113 samples were used for estimation and leave-one-out method was used for validation. It was found that the better prediction methods were KED, RK-C and RK-F in terms of RMSE, of which KED was the best.

Study 18: OK, IDW (with distance powers 0.5, 1 and 2, and two search radii 12 and 22 km), and TPS with tension method (with weights 0.01, 0.1, and 0.5, and two variations of the number of points parameter = 8 and 16) were applied to predict several soil properties (clay content, pH, Na, Ca, Mg, total available P, and organic matter) using 44 samples in a region of 20 x 70 km (Schloeder *et al.*, 2001). Leave-one-out method was used for validation. OK and IDW produced similar accuracies and were similarly effective, and TPS with tension performed poorly by comparison.

Study 19: OK, lognormal OK, DK, IK and MBK were compared using a calibration dataset consisting of 500 samples of the trace element (Co and Cu) concentrations and a validation dataset consisting of 2149 samples over a region of 3500 km^2 (Moyeed and Papritz, 2002). No method was found to be superior to the others when the data were marginally skewed. OK failed to model the conditional distribution of the marginally skewed data. Between them, the nonlinear methods modelled the conditional distribution with similar success.

Study 20: OK, UK, SKlm and OCK were compared for soil silt content for 96 samples with elevation as secondary information in an area of 8 x 18 km (Meul and Van Meirvenne, 2003). A total of 164 samples were used for validation. OCK best accounted for the global trend while UK was best for accounting for the local nonstationarity in terms of MSE. Estimations from combining the results of these two methods (*i.e.*, UK + OCK) were more precise than when any single method was used over the entire study region.

Study 21: SK, OK, lognormal kriging, UK, DK and IDW and their combinations with linear and quadratic trend were applied to 70 samples of soil surface Hg content data collected over a region of 1039 km^2 (Hu *et al.*, 2004). Leave-one-out method was used for validation. It was found that the methods with a trend effect were better than those without a trend effect; and first-order trend UK method was the best, while the IDW the worst. However, the results are slightly contentious as the value of the power parameter used for IDW was not stated in the study.

Study 22: TSA-OK and TSA-OCK (perhaps OCK) were compared for predicting soil Cl^- concentration using 119, 120 and 239 regularly spaced samples collected over a region of 5862 km^2 (Wang *et al.*, 2005). Soil total salt was used as secondary information and leave-one-out cross validation was used to assess the performance of these spatial interpolation methods. The results revealed that OCK was more accurate than OK in terms of RMSE.

Study 23: OK and OCK was applied for predicting soil Zn using 293 samples of highly skewed data with a testing set of 294 samples in the 18 counties of Northern North Dakota (approximately 78,000 km^2) (Wu *et al*., 2006). Soil organic carbon and pH were used as auxiliary variables for OCK. Three methods of data transformation (logarithms, standardised rank order, and normal scores) were carried out to reduce the skewness. OCK, using soil organic carbon or pH as secondary variable, was consistently more accurate than OK in terms of RMSE. OCK with soil organic carbon and pH together provided additional benefit. Data transformation generally improved the estimations, especially for low Zn concentrations. The differences in the performance between normal score OCK, log-normal OCK and rank-ordered OCK were relatively small.

Study 24: OK, OCK and RK-E were compared using 160 samples of soil bulk electrical conductivity over an area of 10.5 ha (Li *et al*., 2007). A total of 80 samples were used for validation. The results showed that irrespective of the sample size of the primary variable (*i.e.*, 40, 70, 100, 130 and 160), RK-E produced, on average, more accurate predictions than OK and COK in terms of RMSE. RK-E was more accurate at a sample size of 70 compared to OCK at any of the sample size. And RK-E was more accurate at a sample size of 40 than OK at any of the sample size. The study concluded that RK-E showed promise for improving predictions with considerable reduction of sampling intensity while maintaining high prediction accuracy.

Study 25: REML-EBLUP, OK and RK-C were compared for four topsoil properties datasets, namely: 1) 155 Zn concentration samples in an area of 10 km^2, 2) 399 samples of soil pH with 200 m spacing between sites, 3) 341 samples of soil clay content with approximately 2.8 km sampling spacing, and 4) 248 samples of soil nitrogen in an area of 83 ha (Minasny and McBratney, 2007). The results revealed that although REML-EBLUP generally improved the prediction in terms of RMSE, the improvement was small compared with RK-C. Therefore, RK-C is a robust method for practical application, while REML-EBLUP is useful when the spatial trend is strong and the number of observations is small (<200).

A.4. Marine Environmental Science

A total of four studies that compared the performance of the spatial interpolation methods in marine environmental science have been reviewed in this section.

Study 1: KED and OK was used for predicting haddock of ages 1, 2 and 3 years using 200-300 samples in each year from 1983 to 1997 in a region of about 369,154 km^2 (Rivoirard and Wieland, 2001). Day/night indicator and time of day were used as external drift. The results from cross-validation indicated that KED with day/night indicator and with time of day performed better than OK in terms of MSE.

Study 2: OK and KED with bathymetry as external drift were compared for grain size data (ICES, 2005). The results revealed that the estimations from KED were 15.7% more accurate than those generated by OK in terms of MSE. This research was further reported in detail by Verfaillie *et al*.(2006), as discussed next.

Study 3: OK, KED and LM were applied for predicting the surficial sand distribution on continental shelf over a region of 3600 km^2 using about 6,000 samples, of which 70% of the samples were used for estimation and the remaining for validation (Verfaillie *et al*., 2006). Bathymetry, with a resolution of 80 m, was used as a secondary variable. KED proved to be the most accurate method in terms of ME, MAE, MSE and RMSE; and the resulting map was more realistic than that from the other methods and separated clearly the sediment distribution over the sandbanks from the swales.

Study 4: OK, OCK, IDW with three distance power parameter (1, 1.5, and 2) and two search radii, NN and Topo to Raster (T2R, hydrological splining) were compared using sand data in the northern Australian marine region (Ruddick, 2007). Sample size for prediction and validation, summary statistics of input data and area of study region were not reported in this study. However, the results showed that IDW, OK, NN and T2R performed similarly in terms of RMSE, although T2R produced slightly poor estimations. The failure of OCK was attributed to the weak correlation between sand concentrations and bathymetry.

The comparison studies of the spatial interpolation methods in marine environmental science are scarce. However, there are some applications of the spatial interpolation methods for estimating marine environmental variables. DK has been used to model regions of high fish density (Petitgas, 1993). NN has been used to interpolate bathymetry, sea temperature and seismic data (Gold and Condal, 1995). OK has been used to predict the spatial patterns of microphytobenthic biomass (Guarini *et al*., 1998) and the spatial distribution of nutrients in the surficial sediments (Danielsson *et al*., 1998). OK and DK have been used for the assessment of the spatial structure and biomass evaluation in Mediterranean Sea (Maynou *et al*., 1998). IK has been applied to mapping of the spatial distribution of benthic communities following a categorical classification scheme (Jerosch *et al*., 2006). Kriging has been used to determine sediment transport (Lucio *et al*., 2006). Finally, OK has been used for seabed mapping and characterisation of sediment variability (Goff *et al*., 2008).

A.5. Other Disciplines

A total of five studies that compared the performance of the spatial interpolation methods in other disciplines have been reviewed in this section.

Study 1: UK, DK and a local mean estimator were compared using 36 datasets chosen randomly on simulated stationary and nonstationary fields (Puente and Bras, 1986). The fields were generated on a rectangular region of 30, 000 km^2, with side on proportion 2 to 1, on a regular 21 x 11 grid. Half of the datasets consist of about 50 points and the other half of datasets about 30 points. The results revealed that in all cases, UK and DK performed better than the local mean estimator, with UK either performed better than or as good as DK in terms of MSE. UK performed particularly well with nonstationary fields, but often underestimated the predicted estimation variance.

Study 2: OK, lognormal OK, SK, lognormal SK, disjunctive OK and disjunctive SK were compared using 122 samples in geology (Boufassa and Armstrong, 1989). The results showed that lognormal kriging produced comparable results to the corresponding type of disjunctive kriging. Results produced by linear kriging (OK and SK) were similar to those produced by the corresponding nonlinear methods. OK, SK, disjunctive OK and disjunctive SK can produce negative estimates because of the presence of negative weights, but lognormal OK and lognormal SK never generate negative estimates since the estimator is an exponential. It was recommended that SK should be used when the mean of the distribution is known, otherwise OK should be used. The similar results from linear and nonlinear methods were argued to be due to the CV was only 1.53. However, this CV is very high and the observed similarity might actually result from the high sampling density instead.

Study 3: It was found that the accuracy of four spatial interpolation methods was as follows: OK > IDS > TIN > NN in terms of MAE when the samples were least clustered; and the order became OK > TIN > IDS > NN in terms of MAE and MSE when the samples were most clustered (Isaaks and Srivastava, 1989). The datasets used in this study are probably simulated data.

Study 4: OK, SK, lognormal OK, rank OK, global mean, IDW (with distance power 1), IDS, TSA and Projected Slope were compared using 54 subsets of data drawn from an exhaustive set of 19,800 data points in a 110 x 180 array, and 198 block (2 x 2 array) estimates were made with each method for each subset (Weber and Englund, 1992). The subsets were drawn independently according to a factorial design with three sample sizes (104, 198 and 308), three sample patterns (random, cellular stratified, and regular grid), and two sampling precisions (zero error and high level normally distributed error with a relative standard deviation of 32% of the true value). It was found that the accuracy of these methods in terms of MSE was as follows: IDS > IDW > OK > SK > the other methods compared. However, it was suggested that the results from this study should not be interpreted to mean the IDW and IDS are superior to kriging methods in all cases.

Study 5: Four spatial interpolation methods (OK, UK, IDS-6 using the nearest six observations, and IDS-12 using the nearest 12 observations) were compared using a factorial computational experiment that included three simulated surface types (plane, sombrero and Morrison's surface), four sampling patterns (hexagonal, inhibited, random and clustered), two variances and two correlation strength parameters (Zimmerman *et al*., 1999). The results revealed that the two kriging methods performed substantially better than the two IDS methods over all levels of surface type, sampling pattern, noise and correlation. Moreover, there was little difference between two kriging methods, although OK performed marginally better when the data were inhibited, clustered, high of noise or less correlated or when the data were for a sombrero surface.

Appendix B. Summary statistics of the information from the 17 reviewed comparative studies.

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
1	(Hartkamp *et al*., 1999)	Meteorology	Precipitation-April	20000	Irregular	169	118.3432	IDW	IDW	5.9		1.9500		33.05	
1				20000	Irregular	169	118.3432	OCK	OCK	5.9		1.8500		31.36	
1				20000	Irregular	169	118.3432	TPS	TPS	5.9		2.3500		39.83	
2	(Hartkamp *et al*., 1999)	Meteorology	Precipitation-May	20000	Irregular	169	118.3432	IDW	IDW	27.2		6.0000		22.06	
2				20000	Irregular	169	118.3432	OCK	OCK	27.2		5.9500		21.88	
2				20000	Irregular	169	118.3432	TPS	TPS	27.2		5.8000		21.32	
3	(Hartkamp *et al*., 1999)	Meteorology	Precipitation-August	20000	Irregular	169	118.3432	IDW	IDW	197.6		32.5500		16.47	
3				20000	Irregular	169	118.3432	OCK	OCK	197.6		36.3500		18.40	
3				20000	Irregular	169	118.3432	TPS	TPS	197.6		45.5000		23.03	
4	(Hartkamp *et al*., 1999)	Meteorology	Precipitation-September	20000	Irregular	169	118.3432	IDW	IDW	166.4		33.8000		20.31	
4				20000	Irregular	169	118.3432	OCK	OCK	166.4		36.7000		22.06	
4				20000	Irregular	169	118.3432	TPS	TPS	166.4		39.3000		23.62	
5	(Hartkamp *et al*., 1999)	Meteorology	Temperature-April	20000	Irregular	125	160	IDW	IDW	31.9		2.7000		8.46	
5				20000	Irregular	125	160	OCK	OCK	31.9		2.6000		8.15	
5				20000	Irregular	125	160	TPS	TPS	31.9		1.6000		5.02	
6	(Hartkamp *et al*., 1999)	Meteorology	Temperature-May	20000	Irregular	125	160	IDW	IDW	32.9		2.5000		7.60	
6				20000	Irregular	125	160	OCK	OCK	32.9		2.3000		6.99	
6				20000	Irregular	125	160	TPS	TPS	32.9		1.4000		4.26	
7	(Hartkamp *et al*., 1999)	Meteorology	Temperature-August	20000	Irregular	125	160	IDW	IDW	28.3		2.0000		7.07	
7				20000	Irregular	125	160	OCK	OCK	28.3		1.8000		6.36	
7				20000	Irregular	125	160	TPS	TPS	28.3		1.2000		4.24	
8	(Hartkamp *et al*., 1999)	Meteorology	Temperature-September	20000	Irregular	125	160	IDW	IDW	28.2		1.9000		6.74	
8				20000	Irregular	125	160	OCK	OCK	28.2		1.9000		6.74	
8				20000	Irregular	125	160	TPS	TPS	28.2		1.1000		3.90	
9	(Erxleben *et al*., 2002)	Meteorology	Snow depth1	1	Regular	549	0.001821	IDW	IDW	0.58	19.83	0.0822	0.1090	14.17	18.79
9				1	Regular	549	0.001821	OK	OK	0.58	19.83	0.0820	0.1092	14.14	18.83
9				1	Regular	549	0.001821	TSA	TSA	0.58	19.83	0.0817	0.1088	14.09	18.76
9				1	Regular	549	0.001821	RK-C	RK-C	0.58	19.83	0.0817	0.1087	14.09	18.74
9				1	Regular	549	0.001821	CART	CART	0.58	19.83	0.0800	0.1043	13.79	17.98
9				1	Regular	549	0.001821	CART-OK	CART-OK	0.58	19.83	0.0797	0.1048	13.74	18.07

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
10	(Erxleben *et al.*, 2002)	Meteorology	Snow depth2	1	Regular	549	0.001821	IDW	IDW	1.09	19.36	0.1503	0.1877	13.79	17.22
10				1	Regular	549	0.001821	OK	OK	1.09	19.36	0.1479	0.1904	13.57	17.47
10				1	Regular	549	0.001821	TSA	TSA	1.09	19.36	0.1468	0.1851	13.47	16.98
10				1	Regular	549	0.001821	RK-C	RK-C	1.09	19.36	0.1429	0.1814	13.11	16.64
10				1	Regular	549	0.001821	CART	CART	1.09	19.36	0.1391	0.1764	12.76	16.18
10				1	Regular	549	0.001821	CART-OK	CART-OK	1.09	19.36	0.1367	0.1748	12.54	16.04
11	(Erxleben *et al.*, 2002)	Meteorology	Snow depth3	1	Regular	549	0.001821	IDW	IDW	1.77	21.07	0.2268	0.3368	12.81	19.03
11				1	Regular	549	0.001821	OK	OK	1.77	21.07	0.2200	0.3351	12.43	18.93
11				1	Regular	549	0.001821	TSA	TSA	1.77	21.07	0.2409	0.3595	13.61	20.31
11				1	Regular	549	0.001821	RK-C	RK-C	1.77	21.07	0.2180	0.3322	12.32	18.77
11				1	Regular	549	0.001821	CART	CART	1.77	21.07	0.2196	0.3200	12.41	18.08
11				1	Regular	549	0.001821	CART-OK	CART-OK	1.77	21.07	0.2094	0.3134	11.83	17.71
12	Martínez-Cob, 1996	Meteorology	Evapotranspiration	47000	Irregular	108	435.1852	OK	OK	1087	15.6	40.3000	51.9711	3.71	4.78
12				47000	Irregular	108	435.1852	OCK	OCK	1087	15.6	38.6000	53.3198	3.55	4.91
12				47000	Irregular	108	435.1852	RK-C	RK-C	1087	15.6	49.3000	64.7765	4.54	5.96
13	Martínez-Cob, 1996	Meteorology	Precipitation	47000	Irregular	132	356.0606	OK	OK	2770.6	7.4	48.9000	62.3699	1.77	2.25
13				47000	Irregular	132	356.0606	OCK	OCK	2770.6	7.4	44.8000	59.4390	1.62	2.15
13				47000	Irregular	132	356.0606	RK-C	RK-C	2770.6	7.4	53.8000	67.3795	1.94	2.43
14	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-Jan	131944	Irregular	74	1783.027	IDS	IDS	1.2	37.5	0.1900	0.0500	15.83	4.17
14				131944	Irregular	74	1783.027	GIDS	GIDS	1.2	37.5	0.0800	0.0200	6.67	1.67
14				131944	Irregular	74	1783.027	OK	OK	1.2	37.5	0.2500	0.0700	20.83	5.83
14				131944	Irregular	74	1783.027	RK-C	RK-C	1.2	37.5	0.1500	0.0400	12.50	3.33
15	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-Feb	131944	Irregular	74	1783.027	IDS	IDS	1.61	28.6	0.1900	0.0500	11.80	3.11
15				131944	Irregular	74	1783.027	GIDS	GIDS	1.61	28.6	0.0800	0.0200	4.97	1.24
15				131944	Irregular	74	1783.027	OK	OK	1.61	28.6	0.2000	0.0600	12.42	3.73
15				131944	Irregular	74	1783.027	RK-C	RK-C	1.61	28.6	0.1100	0.0300	6.83	1.86
16	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-March	131944	Irregular	74	1783.027	IDS	IDS	2.22	19.8	0.2000	0.0600	9.01	2.70
16				131944	Irregular	74	1783.027	GIDS	GIDS	2.22	19.8	0.0700	0.0200	3.15	0.90
16				131944	Irregular	74	1783.027	OK	OK	2.22	19.8	0.2300	0.0700	10.36	3.15
16				131944	Irregular	74	1783.027	RK-C	RK-C	2.22	19.8	0.0900	0.0200	4.05	0.90
17	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-April	131944	Irregular	74	1783.027	IDS	IDS	3.17	14.5	0.2900	0.0800	9.15	2.52
17				131944	Irregular	74	1783.027	GIDS	GIDS	3.17	14.5	0.0700	0.0200	2.21	0.63
17				131944	Irregular	74	1783.027	OK	OK	3.17	14.5	0.3800	0.1100	11.99	3.47

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
17				131944	Irregular	74	1783.027	RK-C	RK-C	3.17	14.5	0.0700	0.0200	2.21	0.63
18	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-May	131944	Irregular	74	1783.027	IDS	IDS	4.27	12.2	0.3500	0.1100	8.20	2.58
18				131944	Irregular	74	1783.027	GIDS	GIDS	4.27	12.2	0.0900	0.0200	2.11	0.47
18				131944	Irregular	74	1783.027	OK	OK	4.27	12.2	0.4000	0.1300	9.37	3.04
18				131944	Irregular	74	1783.027	RK-C	RK-C	4.27	12.2	0.1200	0.0400	2.81	0.94
19	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-June	131944	Irregular	74	1783.027	IDS	IDS	5.45	13.6	0.4900	0.1400	8.99	2.57
19				131944	Irregular	74	1783.027	GIDS	GIDS	5.45	13.6	0.1800	0.0500	3.30	0.92
19				131944	Irregular	74	1783.027	OK	OK	5.45	13.6	0.5400	0.1700	9.91	3.12
19				131944	Irregular	74	1783.027	RK-C	RK-C	5.45	13.6	0.2400	0.0700	4.40	1.28
20	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-July	131944	Irregular	74	1783.027	IDS	IDS	6.02	15.9	0.6000	0.1900	9.97	3.16
20				131944	Irregular	74	1783.027	GIDS	GIDS	6.02	15.9	0.2400	0.0700	3.99	1.16
20				131944	Irregular	74	1783.027	OK	OK	6.02	15.9	1.0300	0.3400	17.11	5.65
20				131944	Irregular	74	1783.027	RK-C	RK-C	6.02	15.9	0.3200	0.0900	5.32	1.50
21	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-Aug	131944	Irregular	74	1783.027	IDS	IDS	5.57	16.7	0.5500	0.1700	9.87	3.05
21				131944	Irregular	74	1783.027	GIDS	GIDS	5.57	16.7	0.2300	0.0800	4.13	1.44
21				131944	Irregular	74	1783.027	OK	OK	5.57	16.7	0.5500	0.1700	9.87	3.05
21				131944	Irregular	74	1783.027	RK-C	RK-C	5.57	16.7	0.3000	0.0900	5.39	1.62
22	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-Sept	131944	Irregular	74	1783.027	IDS	IDS	4.14	17.6	0.4500	0.1300	10.87	3.14
22				131944	Irregular	74	1783.027	GIDS	GIDS	4.14	17.6	0.1500	0.0500	3.62	1.21
22				131944	Irregular	74	1783.027	OK	OK	4.14	17.6	0.4800	0.1400	11.59	3.38
22				131944	Irregular	74	1783.027	RK-C	RK-C	4.14	17.6	0.1600	0.0500	3.86	1.21
23	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-Oct	131944	Irregular	74	1783.027	IDS	IDS	2.62	21	0.2900	0.0800	11.07	3.05
23				131944	Irregular	74	1783.027	GIDS	GIDS	2.62	21	0.1100	0.0300	4.20	1.15
23				131944	Irregular	74	1783.027	OK	OK	2.62	21	0.3300	0.0900	12.60	3.44
23				131944	Irregular	74	1783.027	RK-C	RK-C	2.62	21	0.1300	0.0400	4.96	1.53
24	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-Nov	131944	Irregular	74	1783.027	IDS	IDS	1.62	29.6	0.2500	0.0700	15.43	4.32
24				131944	Irregular	74	1783.027	GIDS	GIDS	1.62	29.6	0.0800	0.0300	4.94	1.85
24				131944	Irregular	74	1783.027	OK	OK	1.62	29.6	0.3700	0.1100	22.84	6.79
24				131944	Irregular	74	1783.027	RK-C	RK-C	1.62	29.6	0.1300	0.0400	8.02	2.47
25	Mardikis *et al.*, 2005	Meteorology	Evapotranspiration-Dec	131944	Irregular	74	1783.027	IDS	IDS	1.26	38.9	0.2000	0.0600	15.87	4.76
25				131944	Irregular	74	1783.027	GIDS	GIDS	1.26	38.9	0.0900	0.0300	7.14	2.38
25				131944	Irregular	74	1783.027	OK	OK	1.26	38.9	0.3500	0.1000	27.78	7.94
25				131944	Irregular	74	1783.027	RK-C	RK-C	1.26	38.9	0.1600	0.0500	12.70	3.97
26	Hosseini *et al.*, 1993	Water resources	Hydraulic conductivity-6	160	Regular	382	0.418848	OK	OK	2.72	78.68	1.4320		52.65	

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
26				160	Regular	382	0.418848	IDW-0	IDW-0	2.72	78.68	1.4320		52.65	
26				160	Regular	382	0.418848	IDW-1	IDW	2.72	78.68	1.4350		52.76	
26				160	Regular	382	0.418848	IDW-2	IDS	2.72	78.68	1.4460		53.16	
26				160	Regular	382	0.418848	IDW-3	IDW	2.72	78.68	1.4610		53.71	
26				160	Regular	382	0.418848	IDW-4	IDW	2.72	78.68	1.4920		54.85	
26				160	Regular	382	0.418848	UK-LD	UK	2.72	78.68	1.4970		55.04	
26				160	Regular	382	0.418848	UK-QD	UK	2.72	78.68	2.1590		79.38	
27	Hosseini *et al.*, 1993	Water resources	Hydraulic conductivity-12	160	Regular	382	0.418848	OK	OK	2.72	78.68	1.3990		51.43	
27				160	Regular	382	0.418848	IDW-0	IDW-0	2.72	78.68	1.4050		51.65	
27				160	Regular	382	0.418848	IDW-1	IDW	2.72	78.68	1.4020		51.54	
27				160	Regular	382	0.418848	IDW-2	IDS	2.72	78.68	1.4110		51.88	
27				160	Regular	382	0.418848	IDW-3	IDW	2.72	78.68	1.4300		52.57	
27				160	Regular	382	0.418848	IDW-4	IDW	2.72	78.68	1.4750		54.23	
27				160	Regular	382	0.418848	UK-LD	UK	2.72	78.68	1.4020		51.54	
27				160	Regular	382	0.418848	UK-QD	UK	2.72	78.68	1.8550		68.20	
28	Hosseini *et al.*, 1993	Water resources	Hydraulic conductivity-18	160	Regular	382	0.418848	OK	OK	2.72	78.68	1.3730		50.48	
28				160	Regular	382	0.418848	IDW-0	IDW-0	2.72	78.68	1.6930		62.24	
28				160	Regular	382	0.418848	IDW-1	IDW	2.72	78.68	1.7390		63.93	
28				160	Regular	382	0.418848	IDW-2	IDS	2.72	78.68	1.3920		51.18	
28				160	Regular	382	0.418848	IDW-3	IDW	2.72	78.68	1.4170		52.10	
28				160	Regular	382	0.418848	IDW-4	IDW	2.72	78.68	1.4710		54.08	
28				160	Regular	382	0.418848	UK-LD	UK	2.72	78.68	1.3830		50.85	
28				160	Regular	382	0.418848	UK-QD	UK	2.72	78.68	1.6930		62.24	
29	Hosseini *et al.*, 1993	Water resources	Hydraulic conductivity	160	Regular	382	0.418848	AK	AK	2.72	78.68	1.8050		66.36	
29				160	Regular	382	0.418848	TSA-1	TSA	2.72	78.68	1.4620		53.75	
29				160	Regular	382	0.418848	TSA-2	TSA	2.72	78.68	1.4460		53.16	
29				160	Regular	382	0.418848	TSA-3	TSA	2.72	78.68	1.4060		51.69	
29				160	Regular	382	0.418848	TSA-4	TSA	2.72	78.68	1.4080		51.76	
30	Schloeder *et al.*, 2001	Soil science	Clay	1400	Irregular	43	32.55814	OK	OK	631	16.64	83.0000	32.2800	13.15	5.12
30				1400	Irregular	43	32.55814	IDW-0.5	IDW	631	16.64	81.8000	31.3688	12.96	4.97
30				1400	Irregular	43	32.55814	IDW-1	IDW	631	16.64	80.1000	30.7083	12.69	4.87
30				1400	Irregular	43	32.55814	IDW-2	IDS	631	16.64	79.5000	30.6105	12.60	4.85
30				1400	Irregular	43	32.55814	TPS-0.01	TPS	631	16.64	82.7000	31.4006	13.11	4.98
30				1400	Irregular	43	32.55814	TPS-0.1	TPS	631	16.64	83.3000	31.8748	13.20	5.05

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
30				1400	Irregular	43	32.55814	TPS-0.5	TPS	631	16.64	85.0000	32.6497	13.47	5.17
31	Schloeder *et al.*, 2001	Soil science	pH	1400	Irregular	43	32.55814	OK	OK	7.82	4.86	0.2100	0.2828	2.69	3.62
31				1400	Irregular	43	32.55814	IDW-0.5	IDW	7.82	4.86	0.2600	0.3317	3.32	4.24
31				1400	Irregular	43	32.55814	IDW-1	IDW	7.82	4.86	0.2500	0.3162	3.20	4.04
31				1400	Irregular	43	32.55814	IDW-2	IDS	7.82	4.86	0.2200	0.3000	2.81	3.84
31				1400	Irregular	43	32.55814	TPS-0.01	TPS	7.82	4.86	0.2400	0.3162	3.07	4.04
31				1400	Irregular	43	32.55814	TPS-0.1	TPS	7.82	4.86	0.2400	0.3317	3.07	4.24
31				1400	Irregular	43	32.55814	TPS-0.5	TPS	7.82	4.86	0.2500	0.3317	3.20	4.24
32	Schloeder *et al.*, 2001	Soil science	Na	1400	Irregular	43	32.55814	OK	OK	3.61	66.48	0.9800	1.2247	27.15	33.93
32				1400	Irregular	43	32.55814	IDW-0.5	IDW	3.61	66.48	1.5300	1.8815	42.38	52.12
32				1400	Irregular	43	32.55814	IDW-1	IDW	3.61	66.48	1.4300	1.7776	39.61	49.24
32				1400	Irregular	43	32.55814	IDW-2	IDS	3.61	66.48	1.2900	1.6186	35.73	44.84
32				1400	Irregular	43	32.55814	TPS-0.01	TPS	3.61	66.48	1.0600	1.2806	29.36	35.47
32				1400	Irregular	43	32.55814	TPS-0.1	TPS	3.61	66.48	1.0600	1.2689	29.36	35.15
32				1400	Irregular	43	32.55814	TPS-0.5	TPS	3.61	66.48	1.0600	1.2649	29.36	35.04
33	Schloeder *et al.*, 2001	Soil science	Ca	1400	Irregular	43	32.55814	OK	OK	27.56	19.92	4.4400	5.7166	16.11	20.74
33				1400	Irregular	43	32.55814	IDW-0.5	IDW	27.56	19.92	4.2600	5.4827	15.46	19.89
33				1400	Irregular	43	32.55814	IDW-1	IDW	27.56	19.92	4.2500	5.4360	15.42	19.72
33				1400	Irregular	43	32.55814	IDW-2	IDS	27.56	19.92	4.3900	5.5498	15.93	20.14
33				1400	Irregular	43	32.55814	TPS-0.01	TPS	27.56	19.92	5.7600	7.4559	20.90	27.05
33				1400	Irregular	43	32.55814	TPS-0.1	TPS	27.56	19.92	5.8800	7.6440	21.34	27.74
33				1400	Irregular	43	32.55814	TPS-0.5	TPS	27.56	19.92	6.0600	7.9265	21.99	28.76
34	Schloeder *et al.*, 2001	Soil science	Mg	1400	Irregular	43	32.55814	OK	OK	7.64	42.41	2.3700	3.0332	31.02	39.70
34				1400	Irregular	43	32.55814	IDW-0.5	IDW	7.64	42.41	2.1600	2.7964	28.27	36.60
34				1400	Irregular	43	32.55814	IDW-1	IDW	7.64	42.41	2.1400	2.7477	28.01	35.97
34				1400	Irregular	43	32.55814	IDW-2	IDS	7.64	42.41	2.1000	2.7221	27.49	35.63
34				1400	Irregular	43	32.55814	TPS-0.01	TPS	7.64	42.41	2.7000	3.5256	35.34	46.15
34				1400	Irregular	43	32.55814	TPS-0.1	TPS	7.64	42.41	2.7300	3.5749	35.73	46.79
34				1400	Irregular	43	32.55814	TPS-0.5	TPS	7.64	42.41	2.8000	3.6401	36.65	47.64
35	Schloeder *et al.*, 2001	Soil science	P	1400	Irregular	43	32.55814	OK	OK	4.68	120.73	3.1200	4.7812	66.67	102.16
35				1400	Irregular	43	32.55814	IDW-0.5	IDW	4.68	120.73	3.4900	5.0666	74.57	108.26
35				1400	Irregular	43	32.55814	IDW-1	IDW	4.68	120.73	3.4300	4.9071	73.29	104.85
35				1400	Irregular	43	32.55814	IDW-2	IDS	4.68	120.73	3.6100	4.8415	77.14	103.45
35				1400	Irregular	43	32.55814	TPS-0.01	TPS	4.68	120.73	3.9000	5.2659	83.33	112.52

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
35				1400	Irregular	43	32.55814	TPS-0.1	TPS	4.68	120.73	4.0100	5.3796	85.68	114.95
35				1400	Irregular	43	32.55814	TPS-0.5	TPS	4.68	120.73	4.1700	5.5624	89.10	118.85
36	Schloeder *et al.*, 2001	Soil science	Soil organic matter	1400	Irregular	43	32.55814	OK	OK	14	28.57	2.2000	0.8944	15.71	6.39
36				1400	Irregular	43	32.55814	IDW-0.5	IDW	14	28.57	2.1000	0.8944	15.00	6.39
36				1400	Irregular	43	32.55814	IDW-1	IDW	14	28.57	2.2000	0.8944	15.71	6.39
36				1400	Irregular	43	32.55814	IDW-2	IDS	14	28.57	2.2000	0.8944	15.71	6.39
36				1400	Irregular	43	32.55814	TPS-0.01	TPS	14	28.57	2.5000	1.0000	17.86	7.14
36				1400	Irregular	43	32.55814	TPS-0.1	TPS	14	28.57	2.5000	1.0000	17.86	7.14
36				1400	Irregular	43	32.55814	TPS-0.5	TPS	14	28.57	2.5000	1.0000	17.86	7.14
37	Wang *et al.*, 2005	Soil science	Cl	5862	Regular	238	24.63025	OK	OK	0.23	88.1		0.1902		84.53
37				5862	Regular	238	24.63025	OCK	OCK	0.23	88.1		0.0997		44.31
38	Wang *et al.*, 2005		Cl	5862	Regular	119	49.2605	OK	OK	0.23	88.1		0.2179		96.84
38				5862	Regular	119	49.2605	OCK	OCK	0.23	88.1		0.0824		36.62
39	Wang *et al.*, 2005		Cl	5862	Regular	118	49.67797	OK	OK	0.23	88.1		0.1364		60.62
39				5862	Regular	118	49.67797	OCK	OCK	0.23	88.1		0.1222		54.31
40	Voltz and Webster, 1990	Soil science	Clay	0.032	Regular	107	0.000299	Classification	Cl	25.1	62.82		8.0747		32.17
40				0.032	Regular	107	0.000299	SK	SK	25.1	62.82		7.0640		28.14
40				0.032	Regular	107	0.000299	StSK	StSK	25.1	62.82		6.8044		27.11
40				0.032	Regular	107	0.000299	Cubic Spline	Spline-3	25.1	62.82		7.2042		28.70
41	Voltz and Webster, 1990	Soil science	Clay	0.92	Regular	34	0.027059	Classification	Cl	20.4	32.33		5.1575		25.28
41				0.92	Regular	34	0.027059	SK	SK	20.4	32.33		4.7434		23.25
41				0.92	Regular	34	0.027059	StSK	StSK	20.4	32.33		4.5387		22.25
41				0.92	Regular	34	0.027059	Cubic Spline	Spline-3	20.4	32.33		5.3292		26.12
42	Voltz and Webster, 1990	Soil science	Clay	0.92	Regular	114	0.00807	Classification	Cl	21.6	31.06		5.3292		24.67
42				0.92	Regular	114	0.00807	SK	SK	21.6	31.06		4.6043		21.32
42				0.92	Regular	114	0.00807	StSK	StSK	21.6	31.06		4.5607		21.11
42				0.92	Regular	114	0.00807	Cubic Spline	Spline-3	21.6	31.06		4.7958		22.20
43	Brus *et al.*, 1996	Soil science	Thickness of A1 horizon	48	Regular	188	0.255319	Classification	Cl	29	62.07	12.2000	19.9000	42.07	68.62
43				48	Regular	188	0.255319	GM	GM	29	62.07	11.4000	18.3000	39.31	63.10
43				48	Regular	188	0.255319	StGM	StGM	29	62.07	9.7000	16.6000	33.45	57.24
43				48	Regular	188	0.255319	IDW-0	IDW-0	29	62.07	13.2000	19.5000	45.52	67.24
43				48	Regular	188	0.255319	StIDW-0	StIDW-0	29	62.07	11.7000	18.6000	40.34	64.14
43				48	Regular	188	0.255319	NN	NN	29	62.07	13.4000	20.8000	46.21	71.72
43				48	Regular	188	0.255319	StNN	StNN	29	62.07	12.2000	19.5000	42.07	67.24

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
43				48	Regular	188	0.255319	IDS	IDS	29	62.07	11.6000	17.6000	40.00	60.69
43				48	Regular	188	0.255319	StIDS	StIDS	29	62.07	10.2000	17.1000	35.17	58.97
43				48	Regular	188	0.255319	TPS	TPS	29	62.07	12.7000	18.9000	43.79	65.17
43				48	Regular	188	0.255319	StTPS	StTPS	29	62.07	10.2000	16.8000	35.17	57.93
43				48	Regular	188	0.255319	OK	OK	29	62.07	11.7000	17.9000	40.34	61.72
43				48	Regular	188	0.255319	StOK	StOK	29	62.07	9.5000	16.4000	32.76	56.55
44	Brus *et al*., 1996	Soil science	P	48	Regular	188	0.255319	Classification	Cl	2.1	71.43				
44				48	Regular	188	0.255319	GM	GM	2.1	71.43	1.0500	1.5100	50.00	71.90
44				48	Regular	188	0.255319	StGM	StGM	2.1	71.43	0.9500	1.4400	45.24	68.57
44				48	Regular	188	0.255319	IDW-0	IDW-0	2.1	71.43	1.0100	1.5000	48.10	71.43
44				48	Regular	188	0.255319	StIDW-0	StIDW-0	2.1	71.43	0.9000	1.4100	42.86	67.14
44				48	Regular	188	0.255319	NN	NN	2.1	71.43	1.1600	1.7000	55.24	80.95
44				48	Regular	188	0.255319	StNN	StNN	2.1	71.43	1.0800	1.6000	51.43	76.19
44				48	Regular	188	0.255319	IDS	IDS	2.1	71.43	0.9900	1.4800	47.14	70.48
44				48	Regular	188	0.255319	StIDS	StIDS	2.1	71.43	0.9400	1.4300	44.76	68.10
44				48	Regular	188	0.255319	TPS	TPS	2.1	71.43	1.0500	1.5200	50.00	72.38
44				48	Regular	188	0.255319	StTPS	StTPS	2.1	71.43	0.9300	1.4100	44.29	67.14
44				48	Regular	188	0.255319	OK	OK	2.1	71.43	1.0000	1.4600	47.62	69.52
44				48	Regular	188	0.255319	StOK	StOK	2.1	71.43	0.9400	1.4200	44.76	67.62
45	Brus *et al*., 1996	Soil science	Mean lowest water table	48	Regular	188	0.255319	Classification	Cl	62	70.97	24.7000	37.0000	39.84	59.68
45				48	Regular	188	0.255319	GM	GM	62	70.97	29.5000	43.6000	47.58	70.32
45				48	Regular	188	0.255319	StGM	StGM	62	70.97	25.8000	38.8000	41.61	62.58
45				48	Regular	188	0.255319	IDW-0	IDW-0	62	70.97	28.4000	39.4000	45.81	63.55
45				48	Regular	188	0.255319	StIDW-0	StIDW-0	62	70.97	23.7000	36.3000	38.23	58.55
45				48	Regular	188	0.255319	NN	NN	62	70.97	31.3000	47.4000	50.48	76.45
45				48	Regular	188	0.255319	StNN	StNN	62	70.97	24.1000	34.2000	38.87	55.16
45				48	Regular	188	0.255319	IDS	IDS	62	70.97	26.1000	38.4000	42.10	61.94
45				48	Regular	188	0.255319	StIDS	StIDS	62	70.97	23.0000	34.4000	37.10	55.48
45				48	Regular	188	0.255319	TPS	TPS	62	70.97	27.8000	38.8000	44.84	62.58
45				48	Regular	188	0.255319	StTPS	StTPS	62	70.97	24.5000	37.5000	39.52	60.48
45				48	Regular	188	0.255319	OK	OK	62	70.97	26.7000	38.3000	43.06	61.77
45				48	Regular	188	0.255319	StOK	StOK	62	70.97	22.2000	34.1000	35.81	55.00
46	Brus *et al*., 1996	Soil science	Mean highest water table	48	Regular	188	0.255319	Classification	Cl	132	36.36	29.3000	43.2000	22.20	32.73
46				48	Regular	188	0.255319	GM	GM	132	36.36	35.8000	48.2000	27.12	36.52

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
46				48	Regular	188	0.255319	StGM	StGM	132	36.36	30.9000	44.6000	23.41	33.79
46				48	Regular	188	0.255319	IDW-0	IDW-0	132	36.36	33.3000	43.6000	25.23	33.03
46				48	Regular	188	0.255319	StIDW-0	StIDW-0	132	36.36	27.6000	40.7000	20.91	30.83
46				48	Regular	188	0.255319	NN	NN	132	36.36	36.9000	55.0000	27.95	41.67
46				48	Regular	188	0.255319	StNN	StNN	132	36.36	27.7000	39.5000	20.98	29.92
46				48	Regular	188	0.255319	IDS	IDS	132	36.36	29.9000	40.6000	22.65	30.76
46				48	Regular	188	0.255319	StIDS	StIDS	132	36.36	26.2000	38.2000	19.85	28.94
46				48	Regular	188	0.255319	TPS	TPS	132	36.36	33.3000	44.0000	25.23	33.33
46				48	Regular	188	0.255319	StTPS	StTPS	132	36.36	29.4000	42.7000	22.27	32.35
46				48	Regular	188	0.255319	OK	OK	132	36.36	31.6000	42.7000	23.94	32.35
46				48	Regular	188	0.255319	StOK	StOK	132	36.36	27.0000	39.0000	20.45	29.55
47	Hu *et al*., 2004	Soil science	Hg	1039	Irregular	69	15.05797	OK	OK	0.08	45.57		0.0267		33.77
47				1039	Irregular	69	15.05797	Lognormal OK	OK	0.08	45.57		0.0266		33.66
47				1039	Irregular	69	15.05797	DK	DK	0.08	45.57		0.0269		34.00
47				1039	Irregular	69	15.05797	UK with a linear trend	UK	0.08	45.57		0.0254		32.14
47				1039	Irregular	69	15.05797	IDW	IDW	0.08	45.57		0.0282		35.70
48	Laslett *et al*., 1987	Soil science	pH in water	0.01	Regular	121	8.26E-05	NaN	NaN	5.26	3.9	0.1930	1.9209	3.67	36.52
48				0.01	Regular	121	8.26E-05	AK	AK	5.26	3.9	0.1867	1.8466	3.55	35.11
48				0.01	Regular	121	8.26E-05	IDS	IDS	5.26	3.9	0.1700	1.7464	3.23	33.20
48				0.01	Regular	121	8.26E-05	GM	GM	5.26	3.9	0.1975	1.9925	3.75	37.88
48				0.01	Regular	121	8.26E-05	NN	NN	5.26	3.9	0.1831	1.8303	3.48	34.80
48				0.01	Regular	121	8.26E-05	IDW-0	IDW-0	5.26	3.9	0.1636	1.6941	3.11	32.21
48				0.01	Regular	121	8.26E-05	TSA	TSA	5.26	3.9	0.1747	1.7972	3.32	34.17
48				0.01	Regular	121	8.26E-05	TPS	TPS	5.26	3.9	0.1650	1.7088	3.14	32.49
48				0.01	Regular	121	8.26E-05	OK isotropic	OK	5.26	3.9	0.1656	1.7146	3.15	32.60
48				0.01	Regular	121	8.26E-05	OK anisotropic	OK	5.26	3.9	0.1622	1.7059	3.08	32.43
49	Laslett *et al*., 1987	Soil science	pH in CaCl2	0.01	Regular	121	8.26E-05	NaN	NaN	4.49	4.76	0.1809	1.8601	4.03	41.43
49				0.01	Regular	121	8.26E-05	AK	AK	4.49	4.76	0.1834	1.8439	4.09	41.07
49				0.01	Regular	121	8.26E-05	IDS	IDS	4.49	4.76	0.1733	1.7550	3.86	39.09
49				0.01	Regular	121	8.26E-05	GM	GM	4.49	4.76	0.2167	2.1587	4.83	48.08
49				0.01	Regular	121	8.26E-05	NN	NN	4.49	4.76	0.1745	1.7972	3.89	40.03
49				0.01	Regular	121	8.26E-05	IDW-0	IDW-0	4.49	4.76	0.1723	1.7692	3.84	39.40
49				0.01	Regular	121	8.26E-05	TSA	TSA	4.49	4.76	0.1892	1.9183	4.21	42.72

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
49				0.01	Regular	121	8.26E-05	TPS	TPS	4.49	4.76	0.1708	1.7378	3.80	38.70
49				0.01	Regular	121	8.26E-05	OK isotropic	OK	4.49	4.76	0.1714	1.7521	3.82	39.02
49				0.01	Regular	121	8.26E-05	OK anisotropic	OK	4.49	4.76	0.1723	1.7720	3.84	39.47
50	Odeh *et al*., 1994, 1995	Soil science	Depth of solum	0.28	Irregular	161	0.001739	LM	LM	73.5	24.66		11.9200		16.22
50				0.28	Irregular	161	0.001739	OK	OK	73.5	24.66		15.7600		21.44
50				0.28	Irregular	161	0.001739	UK	UK	73.5	24.66		13.8900		18.90
50				0.28	Irregular	161	0.001739	OCK	OCK	73.5	24.66		15.7000		21.36
50				0.28	Irregular	161	0.001739	RK-A	RK-A	73.5	24.66		8.4500		11.50
50				0.28	Irregular	161	0.001739	RK-B	RK-B	73.5	24.66		11.2000		15.24
50				0.28	Irregular	161	0.001739	OCK heterotropic	OCK	73.5	24.66		21.7400		29.58
50				0.28	Irregular	161	0.001739	RK-C	RK-C	73.5	24.66		10.0100		13.62
51	Odeh *et al*., 1994, 1995	Soil science	Depth of bedrock	0.28	Irregular	161	0.001739	LM	LM	105.5	32.96		21.0400		19.94
51				0.28	Irregular	161	0.001739	OK	OK	105.5	32.96		26.7100		25.32
51				0.28	Irregular	161	0.001739	UK	UK	105.5	32.96		26.4300		25.05
51				0.28	Irregular	161	0.001739	OCK	OCK	105.5	32.96		24.8600		23.56
51				0.28	Irregular	161	0.001739	RK-A	RK-A	105.5	32.96		20.2200		19.17
51				0.28	Irregular	161	0.001739	RK-B	RK-B	105.5	32.96		19.8900		18.85
51				0.28	Irregular	161	0.001739	OCK heterotropic	OCK	105.5	32.96		22.4500		21.28
51				0.28	Irregular	161	0.001739	RK-C	RK-C	105.5	32.96		16.5100		15.65
52	Odeh *et al*., 1994, 1995	Soil science	Topsoil gravel	0.28	Irregular	161	0.001739	LM	LM	10.6	127.69		4.9700		46.89
52				0.28	Irregular	161	0.001739	OK	OK	10.6	127.69		12.8200		120.94
52				0.28	Irregular	161	0.001739	UK	UK	10.6	127.69		10.3100		97.26
52				0.28	Irregular	161	0.001739	OCK	OCK	10.6	127.69		8.9800		84.72
52				0.28	Irregular	161	0.001739	RK-A	RK-A	10.6	127.69		9.6500		91.04
52				0.28	Irregular	161	0.001739	RK-B	RK-B	10.6	127.69		4.5400		42.83
52				0.28	Irregular	161	0.001739	OCK heterotropic	OCK	10.6	127.69		3.7200		35.09
52				0.28	Irregular	161	0.001739	RK-C	RK-C	10.6	127.69		5.0100		47.26
53	Odeh *et al*., 1994, 1995	Soil science	Subsoil clay	0.28	Irregular	161	0.001739	LM	LM	44.4	34.99		10.2200		23.02
53				0.28	Irregular	161	0.001739	OK	OK	44.4	34.99		15.2000		34.23
53				0.28	Irregular	161	0.001739	UK	UK	44.4	34.99		14.6300		32.95
53				0.28	Irregular	161	0.001739	OCK	OCK	44.4	34.99		10.2400		23.06
53				0.28	Irregular	161	0.001739	RK-A	RK-A	44.4	34.99		9.1100		20.52

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
53				0.28	Irregular	161	0.001739	RK-B	RK-B	44.4	34.99		9.2600		20.86
53				0.28	Irregular	161	0.001739	OCK heterotopic	OCK	44.4	34.99		5.8900		13.27
53				0.28	Irregular	161	0.001739	RK-C	RK-C	44.4	34.99		8.0400		18.11
54	Li *et al.*, 2007	Soil science	Electrical conductivity	0.105	Regular	160	0.000656	OK	OK	123.8	58.32		42.8500		34.61
54				0.105	Regular	160	0.000656	OCK	OCK	123.8	58.32		32.8600		26.54
54				0.105	Regular	160	0.000656	RK-E	RK-E	123.8	58.32		32.1500		25.97
55	Li *et al.*, 2007	Soil science	Electrical conductivity	0.105	Regular	130	0.000808	OK	OK	123.8	58.32		44.4100		35.87
55				0.105	Regular	130	0.000808	OCK	OCK	123.8	58.32		32.5400		26.28
55				0.105	Regular	130	0.000808	RK-E	RK-E	123.8	58.32		27.2200		21.99
56	Li *et al.*, 2007	Soil science	Electrical conductivity	0.105	Regular	100	0.00105	OK	OK	123.8	58.32		47.2800		38.19
56				0.105	Regular	100	0.00105	OCK	OCK	123.8	58.32		35.5100		28.68
56				0.105	Regular	100	0.00105	RK-E	RK-E	123.8	58.32		27.6800		22.36
57	Li *et al.*, 2007	Soil science	Electrical conductivity	0.105	Regular	70	0.0015	OK	OK	123.8	58.32		46.5400		37.59
57				0.105	Regular	70	0.0015	OCK	OCK	123.8	58.32		33.0600		26.70
57				0.105	Regular	70	0.0015	RK-E	RK-E	123.8	58.32		31.3000		25.28
58	Li *et al.*, 2007	Soil science	Electrical conductivity	0.105	Regular	40	0.002625	OK	OK	123.8	58.32		52.8500		42.69
58				0.105	Regular	40	0.002625	OCK	OCK	123.8	58.32		38.2900		30.93
58				0.105	Regular	40	0.002625	RK-E	RK-E	123.8	58.32		39.3800		31.81
59	Bourennane *et al.*, 2000	Soil science	Thickness of a soil horizon	3.8	Irregular	40	0.095	LM	LM	0.65	32.31		0.2500		38.46
59				3.8	Irregular	40	0.095	KED	KED	0.65	32.31		0.2000		30.77
60	Bourennane *et al.*, 2000	Soil science	Thickness of a soil horizon	3.8	Irregular	50	0.076	LM	LM	0.65	32.31		0.2400		36.92
60				3.8	Irregular	50	0.076	KED	KED	0.65	32.31		0.1700		26.15
61	Bourennane *et al.*, 2000	Soil science	Thickness of a soil horizon	3.8	Irregular	75	0.050667	LM	LM	0.65	32.31		0.2400		36.92
61				3.8	Irregular	75	0.050667	KED	KED	0.65	32.31		0.1900		29.23
62	Bourennane *et al.*, 2000	Soil science	Thickness of a soil horizon	3.8	Irregular	100	0.038	LM	LM	0.65	32.31		0.2300		35.38
62				3.8	Irregular	100	0.038	KED	KED	0.65	32.31		0.1600		24.62
63	Bourennane *et al.*, 2000	Soil science	Thickness of a soil horizon	3.8	Irregular	125	0.0304	LM	LM	0.65	32.31		0.2300		35.38
63				3.8	Irregular	125	0.0304	KED	KED	0.65	32.31		0.1700		26.15
64	Bourennane *et al.*, 2000	Soil science	Thickness of a soil horizon	3.8	Irregular	150	0.025333	LM	LM	0.65	32.31		0.2400		36.92
64				3.8	Irregular	150	0.025333	KED	KED	0.65	32.31		0.1500		23.08
65	Wu *et al.*, 2006	Soil science	Soil zinc	78000	Irregular	293	266.2116	OK	OK	1.01	95.05		0.7900		78.22
65				78000	Irregular	293	266.2116	OCK	OCK	1.01	95.05		0.6600		65.35

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
66	Gotway *et al.*, 1996	Soil science	Soil nitrate	0.047	Regular	195	0.000241	OK	OK	3.97	29.47		1.0198		25.69
66				0.047	Regular	195	0.000241	IDW-1	IDW	3.97	29.47		1.0344		26.06
66				0.047	Regular	195	0.000241	IDS	IDS	3.97	29.47		1.0198		25.69
66				0.047	Regular	195	0.000241	IDW-4	IDW	3.97	29.47		1.0247		25.81
67	Gotway *et al.*, 1996	Soil science	Soil nitrate	0.047	Regular	119	0.000395	OK	OK	4	25.75		0.9798		24.49
67				0.047	Regular	119	0.000395	IDW-1	IDW	4	25.75		0.9695		24.24
67				0.047	Regular	119	0.000395	IDS	IDS	4	25.75		0.9747		24.37
67				0.047	Regular	119	0.000395	IDW-4	IDW	4	25.75		1.0149		25.37
68	Gotway *et al.*, 1996	Soil science	Soil organic matter	0.047	Regular	195	0.000241	OK	OK	27.1	11.88		2.4166		8.92
68				0.047	Regular	195	0.000241	IDW-1	IDW	27.1	11.88		2.6134		9.64
68				0.047	Regular	195	0.000241	IDS	IDS	27.1	11.88		2.4960		9.21
68				0.047	Regular	195	0.000241	IDW-4	IDW	27.1	11.88		2.4083		8.89
69	Gotway *et al.*, 1996	Soil science	Soil organic matter	0.047	Regular	119	0.000395	OK	OK	26.7	11.16		2.3854		8.93
69				0.047	Regular	119	0.000395	IDW-1	IDW	26.7	11.16		2.4556		9.20
69				0.047	Regular	119	0.000395	IDS	IDS	26.7	11.16		2.3685		8.87
69				0.047	Regular	119	0.000395	IDW-4	IDW	26.7	11.16		2.3409		8.77
70	Gotway *et al.*, 1996	Soil science	Soil nitrate	0.53	Regular	731	0.000725	OK	OK	3.14	36.31		1.1358		36.17
70				0.53	Regular	731	0.000725	IDW-1	IDW	3.14	36.31		1.1314		36.03
70				0.53	Regular	731	0.000725	IDS	IDS	3.14	36.31		1.1446		36.45
70				0.53	Regular	731	0.000725	IDW-4	IDW	3.14	36.31		1.1832		37.68
71	Gotway *et al.*, 1996	Soil science	Soil nitrate	0.53	Regular	657	0.000807	OK	OK	3.25	38.77		1.2288		37.81
71				0.53	Regular	657	0.000807	IDW-1	IDW	3.25	38.77		1.2247		37.68
71				0.53	Regular	657	0.000807	IDS	IDS	3.25	38.77		1.2369		38.06
71				0.53	Regular	657	0.000807	IDW-4	IDW	3.25	38.77		1.2689		39.04
72	Gotway *et al.*, 1996	Soil science	Soil nitrate	0.53	Regular	192	0.00276	OK	OK	3.14	36.31		1.1489		36.59
72				0.53	Regular	192	0.00276	IDW-1	IDW	3.14	36.31		1.1269		35.89
72				0.53	Regular	192	0.00276	IDS	IDS	3.14	36.31		1.1747		37.41
72				0.53	Regular	192	0.00276	IDW-4	IDW	3.14	36.31		1.3229		42.13
73	Gotway *et al.*, 1996	Soil science	Soil nitrate	0.53	Regular	88	0.006023	OK	OK	3.14	36.31		1.1874		37.82
73				0.53	Regular	88	0.006023	IDW-1	IDW	3.14	36.31		1.1705		37.28
73				0.53	Regular	88	0.006023	IDS	IDS	3.14	36.31		1.2124		38.61
73				0.53	Regular	88	0.006023	IDW-4	IDW	3.14	36.31		1.3115		41.77
74	Gotway *et al.*, 1996	Soil science	Soil organic matter	0.53	Regular	731	0.000725	OK	OK	22.4	21.12		2.6495		11.83
74				0.53	Regular	731	0.000725	IDW-1	IDW	22.4	21.12		3.0265		13.51

Case study	Reference	Discipline	Variable	Area (km^2)	Sampling design	Sample size	Area per sample	method	Method	Mean	CV (%)	MAE	RMSE	RMAE (%)	RRMSE (%)
74				0.53	Regular	731	0.000725	IDS	IDS	22.4	21.12		2.7695		12.36
74				0.53	Regular	731	0.000725	IDW-4	IDW	22.4	21.12		2.6552		11.85
75	Gotway *et al*., 1996	Soil science	Soil organic matter	0.53	Regular	657	0.000807	OK	OK	22.3	21.3		2.7893		12.51
75				0.53	Regular	657	0.000807	IDW-1	IDW	22.3	21.3		3.2000		14.35
75				0.53	Regular	657	0.000807	IDS	IDS	22.3	21.3		2.9240		13.11
75				0.53	Regular	657	0.000807	IDW-4	IDW	22.3	21.3		2.8000		12.56
76	Gotway *et al*., 1996	Soil science	Soil organic matter	0.53	Regular	192	0.00276	OK	OK	22.4	21.12		2.9394		13.12
76				0.53	Regular	192	0.00276	IDW-1	IDW	22.4	21.12		3.5693		15.93
76				0.53	Regular	192	0.00276	IDS	IDS	22.4	21.12		3.0773		13.74
76				0.53	Regular	192	0.00276	IDW-4	IDW	22.4	21.12		3.0594		13.66
77	Gotway *et al*., 1996	Soil science	Soil organic matter	0.53	Regular	88	0.006023	OK	OK	22.4	21.12		3.9128		17.47
77				0.53	Regular	88	0.006023	IDW-1	IDW	22.4	21.12		4.0410		18.04
77				0.53	Regular	88	0.006023	IDS	IDS	22.4	21.12		3.8665		17.26
77				0.53	Regular	88	0.006023	IDW-4	IDW	22.4	21.12		4.1749		18.64