Review

# Spatial interpolation methods applied in the environmental sciences: A review

CrossMark

Jin Li*, Andrew D. Heap

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

A B S T R A C T

Spatially continuous data of environmental variables are often required for environmental sciences and management. However, information for environmental variables is usually collected by point sampling, particularly for the mountainous region and deep ocean area. Thus, methods generating such spatially continuous data by using point samples become essential tools. Spatial interpolation methods (SIMs) are, however, often data-specific or even variable-specific. Many factors affect the predictive performance of the methods and previous studies have shown that their effects are not consistent. Hence it is difficult to select an appropriate method for a given dataset. This review aims to provide guidelines and suggestions regarding application of SIMs to environmental data by comparing the features of the commonly applied methods which fall into three categories, namely: non-geostatistical interpolation methods, geostatistical interpolation methods and combined methods. Factors affecting the performance, including sampling design, sample spatial distribution, data quality, correlation between primary and secondary variables, and interaction among factors, are discussed. A total of 25 commonly applied methods are then classified based on their features to provide an overview of the relationships among them. These features are quantified and then clustered to show similarities among these 25 methods. An easy to use decision tree for selecting an appropriate method from these 25 methods is developed based on data availability, data nature, expected estimation, and features of the method. Finally, a list of software packages for spatial interpolation is provided.

## 1. Introduction

Spatially continuous data (or GIS layers) of biophysical variables are increasingly required in environmental sciences and management. They are usually not readily available and they are difficult and expensive to acquire, especially in areas that are difficult to access (e.g., mountainous or deep marine regions). Spatial distribution data of biophysical variables are often collected from point sources. However, environmental managers often require spatially continuous data over the region of interest to make effective and confident decisions and scientists need accurate spatially continuous data across a region to make justified interpretations. Therefore, SIMs become essential for estimating biophysical variables at the unsampled locations. Spatial interpolation is defined as predicting the values of a primary variable at points within the same region of sampled locations, while predicting the values at points outside the region covered by existing observations is called extrapolation (Burrough and McDonnell, 1998). In this review, we mainly concentrate on the interpolation and also briefly discuss the limitations for extrapolation in relevant sections.

SIMs, including geostatistics, have been developed for and applied in various disciplines, such as mining engineering (Journel and Huijbregts, 1978) and environmental sciences (Burrough and McDonnell, 1998; Goovaerts, 1997; Webster and Oliver, 2001). 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). 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). The top 10 fields which 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 (Zhou et al., 2007).

Different SIMs are developed for specific data types. Many factors affect the predictions of SIMs (Li and Heap, 2011; Li et al., 2011b). There are no consistent findings about how these factors

* Corresponding author. Tel.: +61 (02) 6249 9899; fax: +61 (02) 6249 9956.
E-mail address: Jin.Li@ga.gov.au (J. Li).

affect the performance of SIMs. It is difficult to address the key issue in the application of the methods to environmental data; namely how to select an appropriate method for a given input dataset (Burrough and McDonnell, 1998). Therefore, a thorough review of commonly applied SIMs is necessary.

This review aims to provide guidelines and suggestions for environmental scientists, including marine scientists, regarding application of SIMs to biophysical data. It covers: 1) the commonly used SIMs in environmental sciences; 2) the features and theoretical comparison of 25 SIMs; 3) factors affecting the performance of SIMs; 4) classification of these 25 methods to illustrate their relationship structure; 5) an easy to use decision tree for selecting an appropriate method from these methods based on data nature, expected estimation and features of the method; and 6) a list of software packages for spatial interpolation.

## 2. Spatial interpolation methods

In this section, terms used for SIMs are clarified, and SIMs are then introduced and classified. The trend of spatial interpolation field is depicted; and methods newly introduced and novel hybrid methods developed for spatial interpolation are briefly introduced and discussed. Finally, potential methods for spatial interpolation in environmental sciences are discussed.

A number of methods have been developed for spatial interpolation and many terms have been used to distinguish them, including: 'deterministic' and 'stochastic' methods (Myers, 1994), or "interpolating" and "non-interpolating" methods, or "interpolators" and "non-interpolators" (Laslett et al., 1987). In this review, all methods are referred to as SIMs. The SIMs 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.

A total of 38 methods for spatial interpolation, which are briefly described by Li and Heap (2008), are listed in Table 1. They fall into three categories: 1) non-geostatistical methods, 2) geostatistical methods and 3) combined methods. Estimations by nearly all SIMs can be represented as weighted averages of sampled data (Webster and Oliver, 2001). Because this is a review for environmental science researchers, jargon and mathematical and statistical formulae are avoided whenever possible, but relevant references can be sourced in Li and Heap (2008, 2011) for further reference. Nearly all these methods share the same general estimation formula, as follows:

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

where $\widehat{z}$ is the estimated value of the primary variable at the point of interest $x_0$, $z$ is the observed value at the sampled point $x_i$, $\lambda_i$ is the weight assigned to the sampled point, and $n$ represents the number of sampled points used for the estimation (Webster and Oliver, 2001). Taking two most commonly compared method, inverse distance weighting (IDW) and ordinary kriging (OK) (Li and Heap, 2011) as an example, the weight for IDW is: $\lambda_i = 1/d_i^p / \sum_{i=1}^{n} 1/d_i^p$, where $d_i$ is the distance between $x_0$ and $x_i$, and $p$ is an exponent; while the weight for OK is estimated by minimising the variance of the prediction errors (Isaaks and Srivastava, 1989). Basic concepts (e.g., kriging and semivariogram) and assumptions of kriging and some other SIMs are briefly reviewed by Li and Heap (2008) and detailed in Cressie (1993), Goovaerts (1997), Isaaks and Srivastava (1989), Journel and Huijbregts (1978), Wackernagel (2003) and Webster and Oliver (2001).

The field of geostatistics reached its peak around 1996–1998 in terms of the annual total citation of articles and the total number of annually published articles is still growing (Zhou et al., 2007), but Hengl et al. (2009) found that the field of geostatistics has stabilised since 1999 in terms of average citation. The development of combined methods is certainly ongoing 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.

A number of methods from the machine learning field have been recently introduced to the spatial interpolation of environmental variables and several novel hybrid methods have been developed. These methods include:

1) support vector machine (SVM) (Gilardi, 2002; Li et al., 2011c, 2010),
2) random forest (RF) (Bustamante et al., 2011; Li et al., 2011a, 2011b, 2012),
3) neural network (Kanevski et al., 2008; Li et al., 2010; Lin and Chen, 2004),
4) neuro-fuzzy network (Friedel and Iwashita, 2013; Klesk, 2008; Özkan, 2006; Salski, 2003; Strebel et al., 2013),

**Table 1**
The SIMs considered in this review.

| Non-geostatistical | Geostatistical (kriging) | | Combined methods |
|---|---|---|---|
| | Univariate | Multivariate | |
| Nearest neighbours (NN) | Simple kriging (SK) | Universal kriging (UK) | Classification combined with other interpolation methods |
| Triangular irregular network related interpolations (TIN) | Ordinary kriging (OK) | SK with varying local means (SKlm) | Trend surface analysis combined with kriging |
| Natural neighbours (NaN) | Factorial kriging (FK) | Kriging with an external drift (KED) | Lapse rate combined with kriging |
| Inverse distance weighting (IDW) | Dual kriging (DuK) | Simple cokriging (SCK) | Linear mixed model (LMM) |
| Regression models (LM) | Indicator kriging (IK) | Ordinary cokriging (OCK) | Regression trees combined with kriging |
| Trend surface analysis (TSA) | Disjunctive kriging (DK) | Standardised OCK (SOCK) | Residual maximum likelihood-empirical best linear unbiased predictor (REML-EBLUP) |
| Splines and local trend surfaces (LTS) | Model-based kriging (MBK) | Principal component kriging (PCK) | Regression kriging (RK) |
| Thin plate splines (TPS) | | Colocated cokriging (CCK) | Gradient plus inverse distance squared (GIDS) |
| Classification (Cl) | | Kriging within strata (KWS) | |
| Regression tree (RT) | | Multivariate factorial kriging (MFK) | |
| | | IK with an external drift (IKED) | |
| | | Indicator cokriging (ICK) | |
| | | Probability kriging (PK) | |

5) boosted decision tree (BDT) (Li et al., 2012),
6) the combination of SVM with IDW or ordinary kriging (OK) (Li et al., 2011a, 2011b),
7) the combination of RF with IDW or OK (RFIDW, RFOK) (Li et al., 2011a, 2010, 2012),
8) general regression neural network (GRNN) (Li et al., 2011c),
9) the combination of GRNN with IDW or OK (Li et al., 2012), and
10) the combination of BDT with IDW or OK (Li et al., 2012).

These methods, especially the newly developed hybrid methods RFIDW and RFOK, may play a significant role in spatial interpolation in the future because of their high prediction accuracy (Li, 2013; Li et al., 2011b, 2012; Sanabria et al., 2013). However, their applications are still rare, and so they are not further discussed in this review.

Geographically weighted regression (GWR) (Fotheringham et al., 2002) has potential for the spatial interpolation of the environmental data. The advantages of this method are that it is based on the traditional regression framework and incorporates local relationships into the framework in an intuitive and explicit manner (Fotheringham et al., 2002). The application of this method to spatial interpolation would be worthwhile (Wheeler and Páez, 2010); but its applications to environmental sciences are still rare (Harris et al., 2011) and should be cautiously exercised particularly when sample size is small (Páez et al., 2011).

A few further developments in spatial interpolation are worth attention, although their applications to environmental sciences are rare. The first is its extension to spatio-temporal data. Since the publication of the review of geostatistical space-time models (Kyriakidis and Journel, 1999), there are rapid developments in statistics for spatio-temporal data (Banerjee et al., 2004; Cressie and Wikle, 2011). The second is the Bayesian kriging using prior information (Handcock and Stein, 1993; Pilz et al., 2012; Strebel et al., 2013). The third is the ability to account for the spatial support in kriging (Goovaerts, 2008; Gotway and Young, 2002, 2007). The fourth is the hierarchical Bayesian spatio-temporal modelling (Bakar and Sahu, 2012; Cressie and Wikle, 2011). As more and more spatial data and prior information become available over time in environmental sciences, statistical methods for such data shall get increasingly important and their applications would increase. And the last is new developments in the combined methods such as clustering assisted regression method (Tang et al., 2012).

## 3. Features and theoretical comparison of SIMs

### 3.1. Important features

#### 3.1.1. Global versus local

Global methods use all available data in the region of interest to derive the estimation and capture the general trend(s). Local methods operate within a small area around the point being estimated (*i.e.*, using samples in a search window) and capture the local or short-range variation(s) (Burrough and McDonnell, 1998).

#### 3.1.2. Exactness

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 estimations.

#### 3.1.4. Gradual versus abrupt

Some methods (e.g., nearest neighbours (NN)) produce a discrete and abrupt surface, while 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 between the point of interest and sample points. Criteria include simple distance relations (*e.g.*, IDW), minimisation of variance (*e.g.*, kriging methods) and minimisation of curvature and enforcement of smoothness (*e.g.*, splines).

#### 3.1.5. Convex versus non-convex

Convex methods (e.g., IDW) yield estimates that are always valued between the minimum and maximum of the observed values, whereas non-convex methods (e.g., kriging methods) can yield estimates outside of the range of the observed values (Goovaerts, 1997; Li and Heap, 2008). Convexity of the estimators may depend on factors such as weights assigned for samples and difference between the ranges of predictors for the samples and for the unobserved locations. Possible negative weights resulting from screen effect (Journel and Huijbregts, 1978; Webster and Oliver, 2001) for kriging methods could lead to non-convex estimates (Goovaerts, 1997), while weights for IDW are always positive thus result in convex estimates. When the ranges of the predictors for samples are shorter than those for the unobserved locations, non-convex predictions would be produced by regression related methods (Li et al., 2010, 2012).

#### 3.1.6. Univariate versus multivariate

Spatial interpolation methods which only use samples of the primary variable in deriving the estimation are termed "univariate" methods. Methods which also use explanatory variables (*i.e.*, the secondary variables) are often referred to as "multivariate" methods in geostatistics (see Wackernagel, 2003), which is however often termed as "multiple" (e.g., multiple regression). The "multivariate" usually refers to multiple responses/dependent variables in classic statistics (e.g., multivariate analysis, multivariate ANOVA), despite the fact that in some references it also refers to more than one explanatory variable (that is usually referred to as multiple variables); and univariate refers to only one response variable in classic statistics. In geostatistics, methods accounting for a single variable, such as simple kriging (SK) and OK, are univariate; and methods accounting for secondary information, like simple cokriging (SCK), ordinary cokriging (OCK), kriging with an external drift (KED), are multivariate. Although universal kriging (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. Thin plate splines (TPS) can also be extended to include a multivariate spline function (Burrough and McDonnell, 1998; Hutchinson, 1995; Mitasova et al., 1995; Wahba, 1990).

It should be noted that in the gstat package in R, multivariate kriging means that there are several variables which can be either primary variables or both primary and secondary variables, while multiple kriging implies that several primary variables are kriged separately at the same time (pers. comm. with E. Pebesma, August 2008).

#### 3.1.7. Linear versus nonlinear

Geostatistical methods are often classified as linear or nonlinear (Moyeed and Papritz, 2002; Papritz and Moyeed, 1999). Linear

kriging can be defined as kriging methods which derive the estimation using observed values by assuming a normal distribution of samples, such as SK, OK and UK. Non-linear kriging are those methods which derive predictions based on the transformed values of the observed data, such as disjunctive kriging (DK), indicator kriging (IK), lognormal OK and model-based kriging (MBK). Nonlinear kriging methods have two major advantages over linear kriging, namely they give an estimate of its probability distribution conditional on the available information and their estimations should theoretically be more precise when a Gaussian random process is inappropriate to model the observations.

### 3.2. Comparison of SIMs

SIMs, both non-geostatistical and geostatistical methods, are compared in terms of their features in Table 2. Features of SIMs vary from one method to another, so it is difficult to summarise the features of all methods in one table. The rest of this section intends to explain Table 2 and also provides further supplementary information. Non-geostatistical methods are first compared with geostatistical methods. Geostatistical methods are then compared amongst themselves.

#### 3.2.1. Non-geostatistical methods and kriging methods

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 SIMs are not applicable (Burrough and McDonnell, 1998; Hartkamp et al., 1999). The disadvantages are only one sample point considered and other nearby sampled points ignored for estimated values, with no error estimate (Webster and Oliver, 2001).

Triangular irregular network (TIN) is more accurate than NN, although each estimate still depends on only three samples (Webster and Oliver, 2001). The estimated surface is continuous but with abrupt changes in gradient at the margins of the triangles (Webster and Oliver, 2001).

Natural neighbours (NaN) 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 (Webster and Oliver, 2001). 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).

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

Trend surface analysis (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).

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

Splines retain small-scale features, but there are no direct estimates of the errors (Burrough and McDonnell, 1998). The predictions are very close to the values being interpolated, providing the measurement errors are small (Burrough and McDonnell, 1998; Mitasova et al., 1995). 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; Wahba, 1990; Wood, 2006). TPS can also be extended to include a multivariate spline function (Burrough and McDonnell, 1998; Hutchinson, 1995; Mitasova et al., 1995; Wahba, 1990). TPS can be viewed as an extension of linear regression where the parametric model is replaced by a smooth non-parametric function, and TPS allows an estimate of prediction uncertainty (Haylock et al., 2008). In contrast to local regression methods, one strength of TPS is its dependence on all samples (McKenney et al., 2011). TPS may produce unrealistically smooth results and thus be misleading (Burrough and McDonnell, 1998).

Kriging provides the best linear unbiased estimate, and a measure of the error or uncertainty (Burrough and McDonnell, 1998; Goovaerts, 1997; Journel and Huijbregts, 1978; Nalder and Wein, 1998). It does not produce edge-effects resulting from forcing a polynomial to fit data as with TSA (Collins and Bolstad, 1996). However, its assumption of data stationarity is usually not true, although this assumption can be relaxed with specific forms of kriging. The exactness of kriging (Burrough and McDonnell, 1998; Goovaerts, 1997) is also debatable due to the measurement error component in the nugget effect (Cressie, 1993). Definition of the required variogram models is time consuming and subjective. Data transformation may be needed for non-stationary data and anisotropy needs to be considered.

Kriging variance is independent of data values. Hence, it does not reflect the uncertainty expected at a specific point (Goovaerts, 1997). 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). However, this independence is expected because of data stationarity assumption of kriging methods (Webster and Oliver, 2001; pers. comm. with Noel Cressie in December 2010).

#### 3.2.2. Geostatistical methods

The features of 15 geostatistical methods are summarised in Table 3 according to the findings in Goovaerts (1997) and other studies cited in Li and Heap (2008). Some of the methods listed in Table 1 are excluded because either they are for uncertainty assessment or for categorical variables; and a few frequently used variants or sub-methods are included.

*Simple kriging versus ordinary kriging*: OK is usually preferred to SK because 1) OK 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; 3) OK estimates better follow the data fluctuations than SK estimates; and 4) OK estimates change 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 primary variable (Goovaerts, 1997). However, OK requires a stationary mean of the local search window. If the mean of the primary variable is not constant (or non-stationary) within the search window, regression kriging (RK) (Knotters et al., 1995), universal cokriging (Stein et al., 1988), KED and UK (Verfaillie et al., 2006) could be used to deal with such data.

*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 primary variable 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*: 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

**Table 2**
Comparison of non-geostatistical SIMs and kriging as a generic model for geostatistical methods.

| Method | Assumption | Univariable/ multivariable | Deterministic/ stochastic | Local/global | Exact/ inexact | Abrupt/ gradual | Convex/ non-convex | 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 | Convex | 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 predictors are data points on the surrounding triangle | Univariable | Deterministic | Local | Exact | Abrupt | Convex | 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 predictors are data points in the surrounding polygons | Univariable | Deterministic | Local | Exact | Gradual or abrupt | Convex | 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 | Convex | 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 | Abrupt/ gradual | Can be non-convex | 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 | Can be non-convex | 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 |
| Splines & local trend surfaces (LTS) | Best local predictor is the nearest data point and data normality | Multivariable | Stochastic | Local | Inexact | Gradual | Convex | 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 | Convex | 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 (RT) | Phenomenological explanation of variance | Multivariable | Stochastic | Global | Inexact | ? | Convex | ? | Small | Gridded surface | |

**Table 2** (*continued* )

| Method | Assumption | Univariable/ multivariable | Deterministic/ stochastic | Local/global | Exact/ inexact | Abrupt/ gradual | Convex/ non-convex | Limitation of the procedure | Computing load | Output data structure | Suitability |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Thin plate splines (TPS) | Underlying surface is smooth everywhere | Univariable/ multivariable | Stochastic | Global | Exact/ inexact | Abrupt/ gradual | Can be non-convex | Error assessment of spatial predictions is 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[a] | Interpolated surface is smooth. Statistical stationarity and the intrinsic hypothesis. | Univariable/ multivariable | Stochastic | Local | Exact | Gradual | Can be non-convex | 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. |

Mainly modified from Burrough and McDonnell (1998).

[a] Kriging methods can be either point kriging or block kriging (BK). In this study, all kriging methods refer to point kriging methods unless otherwise specified. BK methods are extensions of point kriging methods and are inexact and the most commonly used BK is block OK (Burrough and McDonnell, 1998; Goovaerts, 1997; Isaaks and Srivastava, 1989).

**Table 3**
A comparison of geostatistical SIMs.

| Geostatistical method | Univariable/ multivariable | Stationary/ local mean | Local trend | Information of coordinates | Secondary variable | Exhaustive secondary information | Stratification | Orthogonalisation of secondary information | Single or multiple samples in the search window |
|---|---|---|---|---|---|---|---|---|---|
| Simple kriging (SK) | Univariate | Stationary | No | No | No | na | No | na | Multiple |
| Ordinary kriging (OK) | UNIVARIATE | Local | No | No | No | na | No | na | Multiple |
| Universal kriging (UK) | Multivariate | Local | Yes | Yes | No | Yes | No | No | Multiple |
| SK with varying local means (SKlm) | Multivariate | Local | No/yes[a] | No | Yes | Yes | No | No | Multiple |
| Kriging with an external drift (KED) | Multivariate | Local | Yes | Yes | Yes | Yes | No | No | Multiple |
| Simple cokriging (SCK) | Multivariate | Stationary | No | No | Yes | No | No | No | Multiple |
| Ordinary cokriging (OCK) | Multivariate | Local | No | No | Yes | No | No | No | Multiple |
| Standardised OCK (SOCK) | Multivariate | Both[b] | No | No | Yes | No | No | No | Multiple |
| Principal component kriging (PCK) | Multivariate | Local | No | No | Yes | No | No | Yes | Multiple |
| Simple colocated cokriging (SCCK) | Multivariate | Stationary | No | No | Yes | No | No | No | Single |
| Ordinary colocated cokriging (OCCK) | Multivariate | Local | No | No | Yes | No | No | No | Single |
| Simple kriging within strata (SKWS) | Multivariate | Within strata stationary | No | No | No | na | Yes | na | Multiple |
| Ordinary kriging within strata (OKWS) | Multivariate | Local | No | No | No | na | Yes | na | Multiple |
| Simple cokriging within strata (SCKWS) | Multivariate | Within strata stationary | No | No | Yes | No | Yes | No | Multiple |
| Ordinary cokriging within strata (OCKWS) | Multivariate | Local | No | No | Yes | No | Yes | No | Multiple |

[a] The local trend is "no" if the secondary variable is categorical and "yes" if it is continuous.
[b] Need the stationary means of both the primary and secondary variables.

estimated, and then performing SK on the corresponding residuals (Goovaerts, 1997). 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, but SCK needs additional modelling and computational requirements. Both methods produce identical estimates when either the primary and secondary variables are uncorrelated or the primary and secondary variables are measured at the same locations and the cross covariance is proportional to the primary autocovariance. Their estimates are essentially the same in the isotropic case and the difference between estimates increases as the samples of secondary variables become more numerous than those of the primary variable. Cokriging (CK) improves over kriging only when the secondary variables are better sampled than the primary variable, or more accurately reflect the real world. 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 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 which are different from OK (Burrough and McDonnell, 1998).

*Colocated cokriging versus cokriging*: Colocated cokriging (CCK) is a 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 it is computationally fast. However, it requires the samples of the secondary variables at all

points being estimated and 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*: 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 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 point kriging*: Block kriging (BK) smooths out short-range variation of the primary variable and can erase the artefact discontinuities near sample points as often observed in the predictions of point kriging. BK estimates vary more smoothly in space than point kriging estimates such as block OK estimates vs. OK estimates (Burrough and McDonnell, 1998; Isaaks and Srivastava, 1989); and the smoothness increases with increasing size of the block (Goovaerts, 1997). As such, predictions of BK are not exact. BK is preferred to point kriging for mapping large-scale features (Goovaerts, 1997).

*Regression kriging versus other related methods*: RK is mathematically equivalent to UK and KED, 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 (Hengl et al., 2007). However, KED and UK differ from RK as the formers conduct a local assessment of the relationship between the primary variable and secondary variables when a local search window is used. Six types of RK (i.e. RK-A to RK-F) have been identified by Li and Heap (2008). The advantage of RK is its ability to extend the method to a broader range of regression techniques such as generalised additive models (GAM) and regression trees (RT) (Bishop and McBratney, 2001), and

generalised linear models (GLM) (Gotway and Stroup, 1997), 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: 1) in RK, the relationship between the primary 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) RK is less computationally demanding and is more efficient than CK; and 3) RK also considers the local trend within the search window by kriging non-stationary data. A disadvantage of RK is that the errors of the predictions are assumed to be unsystematic, not auto-correlated and not correlated with the variable. However, the kriging component of RK does assume the errors to be spatially correlated. 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 better satisfied. 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.

## 4. Factors affecting the performance of SIMs

The predictive performance of SIMs that is usually measured using prediction errors (Bennett et al., 2013; Li and Heap, 2011) is affected by many factors. In this review, we focus on those factors which were commonly encountered in previous studies.

### 4.1. Sampling design and spatial distribution of samples

#### 4.1.1. Data density
*High density*: When data density is high, most methods produce similar results (Burrough and McDonnell, 1998). In analysing rainfall data, Dirks et al. (1998) found that kriging does not show significantly greater improvement in prediction than simpler methods, such as inverse distance squared (IDS) and NN for high-density networks (*i.e.*, 13 rain gauges over a 35 km$^2$ region). In processing soil data, Bregt (1992) 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 an intensively sampled region, however the interpolated surfaces were very different, resulting in 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 more widely spaced samples 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 SIM and parameters may become critical (Burrough and McDonnell, 1998; Hartkamp et al., 1999). The performance of SIMs 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 dependent on the sampling density, but rather on how well the data are sampled and how strong the correlation is between the primary variable and secondary variable(s) (Hengl, 2007). Sample density also affects the predicted error. It was found that with low sample density, both UK and DK may dramatically over- or under-predict the prediction error (Puente and Bras, 1986). This suggests that

such prediction errors should not be used in an absolute sense, but as a relative measure of spatial estimation accuracy (Puente and Bras, 1986). However, the effects of sample density on the prediction accuracy are found to be marginal in a review of 32 methods applied in 80 application cases (Li and Heap, 2011). In addition, it was found that the smoothness of the estimations increased at lower sample densities (Goovaerts, 1997). Issues relating to sample size are further discussed below.

#### 4.1.2. Spatial distribution of samples
Sample spatial distribution may affect the performance of SIMs. Splines perform better when dense, regularly-spaced data are used, than for irregular-spaced data (Collins and Bolstad, 1996). However, application of splines and other nonparametric regression models to data on a grid is sometimes questionable, because the data does not have the direct information needed for reliable prediction and yields no direct information on residual variance (Laslett, 1994). For irregularly-spaced data, the interpolated surface is more variable where sample density is high than where it is low, which may result in structures which are pure artefacts of data configuration. 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).

Spatial clustering of samples affects the accuracy of the estimations and the effects may depend on SIMs. High clustering: 1) reduced the correlation coefficient between the observed and estimated values for all four methods studied, OK, IDS, TIN and NN; 2) reduced the mean absolute error (MAE) for OK, TIN and NN and increased the MAE for IDS; 3) reduced the mean squared error (MSE) for OK and NN; 4) increased the MSE for IDS and 5) had little influence on the estimations from 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 OK, UK and IDS (Zimmerman et al., 1999).

The chosen sampling scheme also affects the performance of SIMs 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).

#### 4.1.3. Temporal variation
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 SIM (Collins and Bolstad, 1996).

#### 4.1.4. Surface type and landscapes
The variation in the surface substantially increases the estimation error of SIMs; and the estimation error consistently increases with an increasing rate as sample size decreases (MacEachren and Davidson, 1987). The performance of SIMs decreases with increasing variation in the surface (Zimmerman et al., 1999). Distinct and sharp spatial changes, like changing soil types across a region, may also cause problems (Stein et al., 1988; Voltz and Webster, 1990). The existence of physical barriers and the difference in landscapes may also affect the performance of SIMs and the accuracy of spatial predictions (Jensen et al., 2006; Little et al., 1997; López-Quílez and Muñoz, 2009; Zhu and Lin, 2010).

#### 4.1.5. Sample size, sampling design and variogram
Sample size and sampling 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 is generated (Burrough and McDonnell, 1998).

Sampling design has significant influence on the prediction accuracy with irregular designs preferred to regular ones (Li and Heap, 2011). Sample spacing must relate to the scale or scales of the variation of the primary variable 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 prediction accuracy might be reduced, as demonstrated by Gotway et al. (1996) which show that wider sampling spacings greatly reduce the information in the resultant maps. In addition, the smoothness of the estimations (or map) will increase with the relative nugget effect (Goovaerts, 1997).

The spatial structure of data may also affect 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). The variogram is also sensitive to sample clustering, particularly when it is combined with a proportional effect that is a form of heteroscedasticity (i.e., the local mean and local variance of data are related) (Goovaerts, 1997). Sample size is important in variogram modelling because the number of pairs of samples at each lag is an important factor for generating reliable variogram parameters. 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; or at least 100, to produce a reliable estimation of the variogram (Webster and Oliver, 1992). 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). This requirement could be overcome by using the residual maximum likelihood method (REML) variogram (Kerry and Oliver, 2007). Predictions based on REML variograms were generally more accurate than those from the conventional moment variograms with fewer than 100 samples, and a sample size of 50 appears adequate for REML variograms (Kerry and Oliver, 2007). Even a sample size of as low as 28 has been suggested for kriging and CK (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, there is a bias at long lags when the variogram of the residuals were estimated using RK-C (Lark et al., 2006; Li and Heap, 2008). Such bias may be reduced but not removed when using RK-D (Lark et al., 2006; Li and Heap, 2008). 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, residual maximum likelihood-empirical best linear unbiased predictor (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).

#### 4.1.6. Sample size and SIMs

The impacts of sample size on the estimation depend on SIMs. RK-D performed better than SK in terms of the level of detail and accuracy, and RK-D (with 222 samples) even performed better than SK (with 2251 samples), leading to a suggestion 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). In practice, we believe there is a threshold beyond which any increase in sample size would not significantly improve 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). Therefore, care should be taken in applying this suggestion.

OK, OCK and RK-E (Li and Heap, 2008) were compared for sample sizes of 40, 70, 100, 130 and 160 (Li et al., 2007). The results showed that as sample size increases, the performance of all three methods improves, with exceptions of that: 1) OK and OCK are more accurate when sample size is 70 than when sample size is 100 and 2) RK-E is less accurate at a sample size of 160 compared with a sample size of 130, 100 and 70. A similar result is observed by Wang et al. (2005) for combined methods TSA-OK and TSA-OCK.

KED and linear regression model (LM) were applied for sample sizes of 40, 50, 75, 100, 125 and 150 (Bourennane et al., 2000). The results revealed that despite a couple of anomalies, generally KED performs better when the sample size increases. The performance of LM remains largely stable across all the sample sizes, which implies that 40 samples provide sufficient information or there is no useful information contained in the extra samples for the 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 SIM. It is likely that in these cases, other data properties, 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 SIMs depends largely on the features of the methods themselves.

### 4.2. Data nature and quality

Five major factors relevant to data quality are discussed in relation to the performance of SIMs: distribution, isotropy and anisotropy, variance and range, accuracy, and spatial correlation and relevant factors. The sources of errors in spatially continuous data and factors affecting the reliability of spatially continuous data have been discussed by Burrough and McDonnell (1998).

#### 4.2.1. Distribution

Data normality can influence the estimation of certain SIMs which assume that the input data are distributed normally about their mean. If this assumption is not met, log transformation is commonly applied, thus resulting in lognormal methods (e.g., lognormal kriging; Cressie, 1993). 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).

In interpolating soil zinc data, Wu et al. (2006) found that 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 that log-transformation generally improved the performance of OK. OK failed to model the marginally skewed data (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 could reduce the accuracy of OK prediction (Weber and Englund, 1992). These findings suggest

that the effects of data transformation on the performance of kriging methods vary with studies and should be examined in the future individual studies.

#### 4.2.2. Isotropy and anisotropy

Isotropy of data is assumed for kriging methods. Data may display evidence of anisotropy which should be considered, otherwise biased estimations may result. However, in some cases, the anisotropy could be ignored to simplify model fitting and to maintain some consistency between the semivariograms in the multivariate model (Martínez-Cob, 1996). Conditions which allow for this are: 1) anisotropy is not evident within the specified search window; 2) the secondary 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 do not coincide. Haberlandt (2007) also found no significant differences in prediction performance between isotropic and anisotropic variograms, although anisotropy was clearly present in the data.

#### 4.2.3. Variance and range

The variance of data negatively affects the performance of SIMs and the resultant predictions. Their performance decreases rapidly when the coefficient of variation (CV) increases as evidenced in previous studies (Martínez-Cob, 1996; Schloeder et al., 2001) and as further examined and confirmed based on 80 published application cases by Li and Heap (2011). As the data range increased, the performance of all methods compared improved significantly (Collins and Bolstad, 1996).

#### 4.2.4. Accuracy

Data noise can negatively affect the performance of SIMs, including OK, UK and IDS (Zimmerman et al., 1999), and NaN (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 was found not to be significant in determining the performance of OK in certain circumstances (Englund et al., 1992). When data are not representative of the surface being modelled, interpolation biases may result (Collins and Bolstad, 1996).

Outliers affect the performance of SIMs and also 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 of semivariance as evident from its definition. The effects depend on the location of data points in the region and also on the spatial pattern of data (Webster and Oliver, 2001). Outliers should be removed if they are believed to not belong to the population and strongly skewed distributions need to be transformed to approximately normal before conducting geostatistical analyses (Webster and Oliver, 2001). Removal of outliers can result in considerable improvement in the performance of SIMs, particularly when additional samples are included to allow estimation of short-range variation (Laslett and McBratney, 1990).

#### 4.2.5. Spatial correlation and relevant factors

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

The performance of different SIMs changes with the variable estimated. The best methods vary as a function of the region and the spatial scale required for estimation (Vicente-Serrano et al., 2003). In their study on interpolating annual precipitation and temperature, the accuracy is 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 SIM differs considerably with different variables due to difference in the data variance and range.

#### 4.3. Correlation between primary and secondary variables

Correlation between primary and secondary variables is critical for SIMs which use secondary information. In these methods, secondary variables are assumed to be well and accurately sampled at a large number of locations in space or at some resolution (e.g., optical remote sensing data or acoustic backscatter data) and to provide an accurate representation of the underlying structure of the primary variable. This means they need to be strongly correlated with the primary variable(s) (Ahmed and De Marsily, 1987). 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; Li et al., 2011c, 2010). In addition, the performance of GAM, LM, RT, OK, KED, RK-C and RK-F depended on the choice of secondary information (Bishop and McBratney, 2001; Li et al., 2011c, 2010). 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).

As the correlation increases, information provided by the secondary variable to the primary variable increases (Goovaerts, 1997). Stronger correlations result in more accurate estimations of CK and OCK (Goovaerts, 1997), and SKlm, KED and ordinary colocated cokriging (OCCK) (Goovaerts, 2000). In one particular study, as correlation between elevation and temperature increases, the performance of SIMs using elevation as ancillary information improves significantly (Collins and Bolstad, 1996). For a correlation >0.4, SCK and OCK perform better than other methods (SK, OK and LM), and KED is almost as accurate as CK (Asli and Marcotte, 1995). When the correlation increased from 0.77 to 0.99, the root mean squared error (RMSE) for CK was reduced by 48.3% (Wang et al., 2005).

Erxleben et al. (2002) suggest that the bivariate 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 conclude that only variables which are spatially cross-correlated with the primary variable should be included in OCK models.

#### 4.4. Interaction among factors

Interactions among different factors may also exist and should be considered in evaluating the performance of SIMs. 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 SIMs (Zimmerman et al., 1999). Sample density, data variation and sampling design were also found to have significantly interactive effects on the prediction accuracy of SIMs in a review of 80 application cases in environmental sciences (Li and Heap, 2011).

#### 4.5. Other issues

The choice of semivariogram models may play a significant role in the resultant estimation. First-order trend OK performs better with a Gaussian semi-variogram model than with spherical and

exponential models (Hu et al., 2004). Some practical guidelines for selecting an appropriate variogram model are provided by Hartkamp et al. (1999), Goovaerts (2000) and Cressie (1993).

Grid size (resolution) can also affect the accuracy of 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).

## 5. Classification of SIMs

SIMs are classified based on their features to provide an overview of their differences and relationships. These features are then quantified and clustered to show similarities and relationships among these SIMs.

### 5.1. Classification tree of SIMs

Classification of SIMs has not been addressed before apart from a study by Lam (1983) who proposed a simple classification of four types of SIMs. In this review, we adopt an approach used in taxonomy to classify the 25 SIMs according to their features (Fig. 1). In this figure, SIMs are classified based on their features summarised in Tables 2 and 3 and their comparisons in Section 3. This classification tree illustrates the relationship structure among these methods and forms a basis for a decision tree to select an appropriate method in practice in Section 6.

### 5.2. Similarity between SIMs

The similarity between 25 SIMs is analysed. A total of 23 features extracted from Tables 2 and 3 and from those used for the classification (Fig. 1) are converted into variables with factor levels being either 0, 1 or not applicable (na) (Table 4). Information of each feature for each of the 25 methods is summarised in Table 5. These data were analysed using hierarchical cluster analysis on Gower's distance in R 2.6.2 (R Development Core Team, 2007) and the results are shown in Fig. 2. If a threshold line is placed at Height = 0.2 in Fig. 2, these methods could be classified into 11 groups:

- Group 3: NN, TIN and NaN are similar for most of the features considered except the output and the smoothness.
- Group 4: SK and OK are most similar and they do not use secondary information. Their difference is the choice of the stationary mean or local means.
- Group 5: Simple kriging within strata (SKWS) and ordinary kriging within strata (OKWS) form a group of kriging methods which do not use secondary information but apply stratification. They differ due to the stationary mean for SKWS and local means for OKWS.
- Group 6: All cokriging methods are grouped together. In this group there are two subgroups distinguished by the application of stratification. Methods with stratification (simple cokriging within strata (SCKWS) and ordinary cokriging within strata (OCKWS)) differ in the choice of mean. Methods without stratification (simple colocated cokriging (SCCK), ordinary colocated cokriging (OCCK), PCK, standardised OCK (SOCK), SCK and OCK) differ in the choice of mean, use of secondary information, the number of samples in a search window, and orthogonalisation in producing their estimations.
- Group 7: SKlm, UK and KED are similar and use secondary information and/or coordinate information in making their estimations. They are different in local trend and utilisation of coordinate and secondary information.

- Group 8: TSA and local trend surfaces (LTS) form a group that uses coordinate information in deriving the estimations. TSA is a global approach and LTS is a local one.
- CI, IDW, TPS, LM, and RT form single method group and they are groups 1, 2, 9, 10 and 11 respectively, which indicate that they are different from each other and also from all the other methods.

The relationship among these 11 groups can be further explored if a threshold line is placed at Height = 0.4 in Fig. 2 and these groups can be merged into three major groups. The first major group, containing groups 1, 2 and 3, is non-geostatistical, deterministic, univariate, local and non-utilisation of coordinate and secondary information. The second one consists of groups 4, 5, 6 and 7 and is geostatistical, stochastic, local, exact and gradual. The last one formed by groups 8, 9, 10 and 11 is non-geostatistical, multivariate, stochastic and inexact.

## 6. Selection of SIMs

Selection of an appropriate SIM for the data at hand is critical, but not easy. The performance of SIMs depends on many factors. There is no simple answer regarding the choice of an appropriate SIM, because a method is "best" only for specific situations (Isaaks and Srivastava, 1989). A number of factors should be considered in making an appropriate selection. The choice of method may depend on the assumption and properties of each method, the nature and spatial structure of data for the primary variable, sample size and distribution, the availability of secondary information and many other factors as discussed in Section 4. They can be used to eliminate some inappropriate methods. The availability of software may also be an important issue. The computational demands are also crucial depending on sample size, the power of computer and the efficiency of software.

Guidelines have been proposed in previous studies for selecting a SIM from subsets of the methods listed above. For instance, a decision tree for selecting a suitable spatial prediction method from four methods (RK, OK, IDW and LM) was developed by Hengl (2007). Pebesma (2004) and Bivand et al. (2008) proposed a decision tree for IDW, TSA and three kriging methods in gstat package in R. There are also guidelines for choosing between DK and IK according to the nature and structure of data (Lark and Ferguson, 2004). A general method has been developed for selecting among SIMs based on unbiased estimation of MSE by Huang and Chen (2007) and they applied this method to OK, UK and TPS. In addition, several steps have been provided for using kriging methods by Burrough and McDonnell (1998).

In this review, we develop a decision tree according to the availability and nature of data and the expected estimation in combination with the features of each SIM (Fig. 3). All 25 methods in Table 5 are considered. The decision tree is illustrated in a taxonomic fashion. It provides guidelines for selecting an appropriate SIM from these methods. It is easy to use for selecting a method based on the information available. For example, for a *continuous primary variable* (2) with *spatial structure* (*with a nonlinear variogram*) (1) to make predictions using *local means* (4*) and *without secondary variable* (3), the candidate interpolation method is OK in Fig. 3.

Many other factors as discussed in previous sections could influence the choice of a SIM. For example, one might use a SIM that does not incorporate secondary information even if such information is available if it is considered as a reasonable approach. Joint application of two SIMs might produce additional benefits such as the combined procedures in Li and Heap (2008). If distinct spatial changes, such as those in soil and rock types, vegetation

classes and habitat types, are expected, stratified SIMs may be used to improve the estimation (Hernandez-Stefanoni and Ponce-Hernandez, 2006; Stein et al., 1988; Voltz and Webster, 1990). Moreover, if more than one method can be applied, cross-validation in combination with the measures of predictive errors (Bennett et al., 2013; Gneiting et al., 2007; Li, 2013; Li and Heap, 2008; Li et al., 2011b) can be used to select a method that minimises the predictive error.

For kriging methods, a number of factors, including sample size, and isotropy and anisotropy of data, need to be considered for selecting appropriate variogram models. 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 SIMs 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).

```
1 Non-geostatistical, no error assessment (except TPS)
    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…... ………………………………………………………………………………………………NaN
            4* Inexact ………………………………………………………………………………………………………IDW
    2* Stochastic
        8 Global
            9 Coordinates only……………………………………………………………………………………………TSA
            9* Coordinates and other secondary variables
                10 Abrupt………………………………………………………………………………………………RT
                10*Gradual and can be abrupt if categorical secondary variables used
                    11 No error assessment ………………………………………………………………………………LM
                    11* With error assessment ……………………………………...………… TPS
        8*Local…………………………………………………  ………………………………………Splines & LTS
1* Geostatistical, with error assessment
    12 Univariate
        13 Stationary mean…………………………………………………………………………………………SK
        13* Local means……………………………………………………………...…………………………OK
    12* Multivariate
        14 Stationary mean
            15 Non-stratification
                16 Search window with multiple samples…………...……………………………………………SCK
                16* Search window with single sample…………………………………………………………SCCK
            15* Stratification
                17 Non-continuous secondary information…………………………………………………………SKWS
                17* With continuous secondary information……………………………………………………SCKWS
        14*Local means
            18 Exhaustive secondary information and/or local trend
                19 Coordinates only………………………………………………………………………………UK
                19* Non-coordinate secondary variable
                    20 A secondary variable and search window with multiple samples
                        21 Regression coefficients estimated within each search window…...……………………KED
                        21* Regression coefficients estimated once…….....………………………………………SKlm
                    20* One or more secondary variable and search window with single
                    sample…………………………………………………………….....………………………OCCK
            18* Non-exhaustive secondary information and no local trend
                22 Stratification
                    23 No secondary information…………………………………………………………………OKWS
                    23* Secondary information…………………………………………………………………OCKWS
                22* Non-stratification
                    24 Orthogonalisation of secondary information…………...………………………………………PCK
                    24* Non-orthogonalisation of secondary information
                        25 No information of the stationary means of the primary and secondary
                        variables…………………………………………………………………………………OCK
                        25* With information of the stationary means of both the primary and secondary
                        variables...………………………………………….....………………  ……………………SOCK
```

**Fig. 1.** Classification tree of 25 SIMs based on their features.

**Table 4**
Conversion between feature status and factor levels.

| No | Feature | Level | | |
|----|---------|-------|---|---|
| | | 0 | 1 | na[a] |
| 1 | Univariate | No | Yes | |
| 2 | Multivariate | No | Yes | |
| 3 | Deterministic/stochastic | Deterministic | Stochastic | |
| 4 | Local/global | Global | Local | |
| 5 | Exact | No | Yes | |
| 6 | Inexact | No | Yes | |
| 7 | Abrupt transition | No | Yes | |
| 8 | Gradual transition | No | Yes | |
| 9 | Convex | No | Yes | |
| 10 | Output: polygons | No | Yes | |
| 11 | Output: triangular | No | Yes | |
| 12 | Output: grids | No | Yes | |
| 13 | Stationary/local mean | Stationary | Local | na |
| 14 | Stationary mean of secondary variable | No | Yes | na |
| 15 | Local trend-constant | No | Yes | na |
| 16 | Local trend-non-constant | No | Yes | na |
| 17 | Info of coordinates | No | Yes | |
| 18 | Secondary variables | No | Yes | |
| 19 | Point/block | Point | Block | |
| 20 | Exhaustive secondary information | No | Yes | na |
| 21 | Stratification | No | Yes | |
| 22 | Orthogonalisation of secondary information | No | Yes | na |
| 23 | Single or multiple samples in the search window | Single | Multiple | |

[a] na: not applicable.
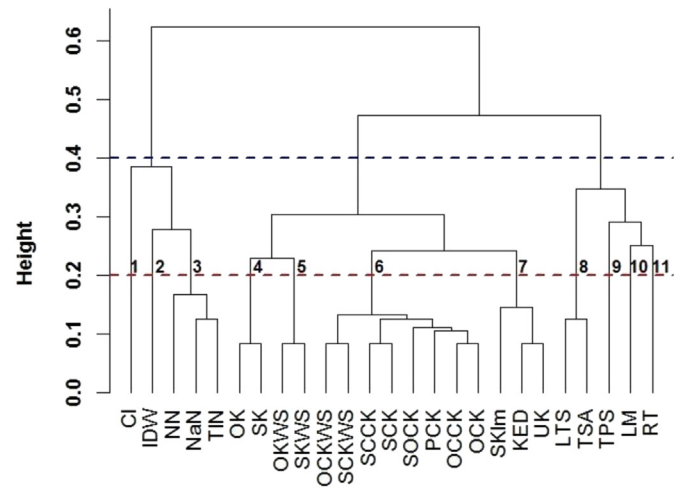
**Fig. 2.** Classification of 25 SIMs based on the 23 binary features in Table 5 using hierarchical cluster analysis.

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

**Table 5**
The quantified data of the 23 features of 25 SIMs. For the feature corresponding to each number please see Table 4. The methods are arranged in an order according to the results from Fig. 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 | 22 | 23 |
|--------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Cl | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | na | na | na | na | 1 | 1 | 1 | 1 | 0 | na | 0 |
| IDW | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | na | na | na | na | 0 | 0 | 0 | na | 0 | na | 1 |
| NN | 1 | 0 | 0 | 1 | 1 | 0 | 1 | **0** | 1 | **1** | **0** | 1 | na | na | na | na | 0 | 0 | 1 | na | 0 | na | **0** |
| NaN | 1 | 0 | 0 | 1 | 1 | 0 | 1 | **1** | 1 | **0** | **0** | 1 | na | na | na | na | 0 | 0 | 1 | na | 0 | na | 1 |
| TIN | 1 | 0 | 0 | 1 | 1 | 0 | 1 | **0** | 1 | **0** | **1** | 1 | na | na | na | na | 0 | 0 | 1 | na | 0 | na | 1 |
| OK | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **1** | na | 1 | 0 | 0 | 0 | 0 | na | 0 | na | 1 |
| SK | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **0** | na | 1 | 0 | 0 | 0 | 0 | na | 0 | na | 1 |
| OKWS | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **1** | 0 | 1 | 0 | 0 | 0 | 0 | na | 1 | na | 1 |
| SKWS | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **0** | 0 | 1 | 0 | 0 | 0 | 0 | na | 1 | na | 1 |
| OCKWS | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **1** | **0** | 1 | 0 | 0 | 1 | 0 | 0 | **1** | **0** | 1 |
| SCKWS | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **0** | **0** | 1 | 0 | 0 | 1 | 0 | 0 | **1** | **0** | 1 |
| SCCK | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **0** | **0** | 1 | 0 | 0 | 1 | 0 | 0 | **0** | **0** | 0 |
| SCK | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **0** | **0** | 1 | 0 | 0 | 1 | 0 | 0 | **0** | **0** | 1 |
| SOCK | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **1** | **1** | 1 | 0 | 0 | 1 | 0 | 0 | **0** | **0** | 1 |
| PCK | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **1** | **0** | 1 | 0 | 0 | 1 | 0 | 0 | **0** | **1** | 1 |
| OCCK | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **1** | **0** | 1 | 0 | 0 | 1 | 0 | 0 | **0** | **0** | 0 |
| OCK | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | **1** | **0** | 1 | 0 | 0 | 1 | 0 | 0 | **0** | **0** | 1 |
| SKlm | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | **1** | 1 | **0** | **1** | 0 | 1 | 0 | 0 | 1 |
| KED | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | **0** | 1 | **1** | **1** | 0 | 1 | 0 | 0 | 1 |
| UK | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | **0** | 1 | **1** | **0** | 0 | 1 | 0 | 0 | 1 |
| LTS | 0 | 1 | 1 | **1** | 0 | 1 | 0 | 1 | **1** | 0 | 0 | 1 | na | 0 | na | na | 1 | 0 | 0 | na | 0 | na | 1 |
| TSA | 0 | 1 | 1 | **0** | 0 | 1 | 0 | 1 | **0** | 0 | 0 | 1 | na | 0 | na | na | 1 | 0 | 0 | na | 0 | na | 1 |
| TPS | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | na | na | 0 | 1 | 1 | 1 | 0 | 1 | 0 | na | 1 |
| LM | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | na | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 |
| RT | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | na | 0 | na | na | 1 | 1 | 0 | 1 | 0 | 0 | 1 |

1 Data or residuals show spatial structure or a non-linear variogram
　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……………………………………………………………………………**OK**
　　3* Information of secondary variables available
　　　5 Global mean known
　　　　6 Secondary variable is only categorical
　　　　　7 Stratification……………………………………………………………………………………………………**SKWS**
　　　　　7* Non-stratification………………………………………………………………………………………………**SKlm**
　　　　6* Secondary variable is not only categorical
　　　　　8 Stratification……………………………………………………………………………………………………**SCKWS**
　　　　　8* Non-stratification
　　　　　　9 Sparse samples of secondary variable and multiple samples in search window…………………………………………………………………………………………………………**SCK**
　　　　　　9* Dense samples of secondary variable and single sample in search window…………………………………………………………………… ……………………………………………**SCCK**
　　　5* Global mean unknown and using local means
　　　　10 Secondary information available for each point being estimated
　　　　　11 Spatial trend is apparent and only coordinates available …………..………………………………………………**UK**
　　　　　11* Other secondary variable available
　　　　　　12 An apparent global relation with the secondary variable…………..……………………………………………**SKlm**
　　　　　　12* The relation is not so apparent…………..………………………………………………………………**KED**
　　　　10* Secondary information not available for each point being estimated
　　　　　13 Secondary variables including a categorical variable
　　　　　　14 Only a categorical variable available
　　　　　　　15 Multiple samples in search window………………………………………………………………………**OKWS**
　　　　　　　15* Dense samples of secondary variable and single sample in search window……………………………………………..………………………………………………………**OCCK**
　　　　　　14* Other secondary information available………………………………………………………………………**OCKWS**
　　　　　13* Secondary variables without categorical variable
　　　　　　16 Sparse samples of secondary variable and multiple samples in search window
　　　　　　　17 Many secondary variables and PCA needed..………………………… ……………………………………**PCK**
　　　　　　　17* PCA not needed to reduce the number of secondary variables
　　　　　　　　18 Avoid negative weights and artificially limiting the effect of secondary variable……………………………………………………………………………………………………**SOCK**
　　　　　　　　18* Accept above two drawbacks……………………………………………………………………**OCK**
　　　　　　16* 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 or linear variogram or sample size is too small to derive a reliable variogram
　19 No secondary variables available
　　20 Abrupt estimation acceptable
　　　21 Using single sample for estimation……………………………………………………………………………...**NN**
　　　21* Using multiple samples for estimation
　　　　22 Using three samples for estimation…………………………………………………………………………**TIN**
　　　　22* Using more than three natural neighbour samples for estimation………..………………………………………**NaN**
　　20* Abrupt estimation unacceptable
　　　23 Using more than three natural neighbour samples weighted by area..…………………………………………………**NaN**
　　　23* Using nearest several samples weighted by distance…………..………………………………………………**IDW**
　19* Secondary variable available
　　24 Using information of coordinates
　　　25 Only coordinates information used with inexact estimation
　　　　26 Using nearby samples…….…..………………………………………………………………………**Splines & LTS**
　　　　26* Using all samples………………………………..…..……………………………………………………**TSA**
　　　25* May use other variables with exact or inexact estimation…..…………………………………………………**TPS**
　　24* Not using information of coordinates
　　　27 Only categorical secondary information available
　　　　28 Only one variable available…………………………..……………………………………………………**Cl**
　　　　28* Multiple variables available………………………..…..…………………………………………………**RT**
　　　27* Continuous secondary information available
　　　　29 Univariate or multiple secondary information……………………..……………………………………………**LM**
　　　　29* Require multiple secondary information…………………..…..…………………………………………………**RT**

**Fig. 3.** A decision tree of the SIMs.

**Table 6**
Availability of the SIMs in several commonly used software packages.

| Method/package | ArcGIS[a] | GS+ | R | | | | | | | | | | | | S-PLUS[b] | SURFER |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | stats | akima | deldir | fields | geoR | geoRglm | GRASS | Gstat | spatial | Sgeostat | RandomFields | Tripack | | |
| NN | Yes | | | | 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 | | | | | Yes | Yes |
| TPS | Yes | | 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 | | | | | | |
| SCCK | | | | | | | | | | Yes | | | | | | |
| KED | | | | | | | Yes | | | Yes | | | | | | |
| StOK/StSK | | Yes | | | | | | | | Yes | | | | | | |
| IK | Yes | | | | | | | | | Yes | | | | | | |
| MBK | | | | | | | Yes | Yes | | | | | | | | |

[a] Including ArcGIS Geostatistical Analyst, an extension to ArcGIS.
[b] Including the S+SpatialStats module (Kaluzny et al., 1998).

estimations at unsampled locations when samples may not be representative of the study area such as samples being scarce and preferentially located (Goovaerts, 1997). Visual examination of the predictions is an essential and further step to validate the predictions (Li et al., 2011a, 2011b, 2011c).

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).

## 7. Software packages

Many software packages contain functions to interpolate spatial point data to spatially continuous data (Table 6). The list of software packages and SIMs in each package is acquired from various sources and is not exhaustive. However, for each method there is at least one software package provided.

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 a Bayesian framework (also known as MBK; Diggle and Ribeiro, 2007; Pebesma, 2004). Due to heavy computational requirements, MBK seems to be only relevant to datasets of small sample sizes (Moyeed and Papritz, 2002). However, this and previous relevant statements about the computation requirements may no longer be valid because of the increase in computing power since then and adoption of some improved algorithms. For example, a sample size of 10,000 can be easily handled in geo-statistics now (pers. comm. with E. Pebesma, 9 July 2008). Some other R packages that are not included in this study can be found at CRAN task views: analysis of spatial data (http://cran.r-project.org/web/views/Spatial.html) and handling and analysing spatio-temporal data (http://cran.r-project.org/web/views/SpatioTemporal.html).

A few SIMs are also available in *GS+* (Robertson, 2000), ArcGIS Geostatistical Analyst (an extension to ArcGIS™) and SURFER. Two types of SIMs, OK and UK, are provided in the *S+SpatialStats* module in *S-PLUS* (Kaluzny et al., 1998). TPS is also implemented in the ANUSPIN software (Hutchinson, 1995, 2011; Xu and Hutchinson, 2013).

A list of geostatistical software (including e.g., SGeMS, Space-time routines, WinGslib) with some kriging methods are compared and discussed by Goovaerts (2010). Computer programs available for 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 SIMs. The computing capacities of some popular statistical and GIS packages were compared by Hengl (2007).

In addition, besides above packages for SIMs, software available for optimal spatial sampling design to improve spatial interpolation was reviewed by Spöck (2012). Moreover, a MATLAB and Octave toolbox was developed for spatial sampling design optimisation (Spöck, 2012).

## References

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

Asli, M., Marcotte, D., 1995. Comparison of approaches to spatial estimation in a bivariate context. Math. Geol. 27, 641–658.

Bakar, K.S., Sahu, S.K., 2012. spTimer: Spatio-temporal Bayesian Modelling Using R. University of Southampton, UK.

Banerjee, S., Carlin, B.P., Gelfand, A.E., 2004. Hierarchical Modeling and Analysis for Spatial Data. Chapman & Hall/CRC, Boca Raton.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013.

Characterising performance of environmental models. Environ. Model. Softw. 40, 1—20.

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

Bivand, R.S., Pebesma, E.J., Gómez-Rubio, V., 2008. Applied Spatial Data Analysis with R. Springer, New York.

Bourrennane, H., King, D., 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.

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

Brus, D.J., de Gruijter, J.J., Marsman, B.A., Visschers, B.A., Bregt, A.K., Breeuwsma, A., 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., McDonnell, R.A., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford.

Bustamante, R.H., Dichmont, C.M., Ellis, N., Griffiths, S., Rochester, W.A., Burford, M.A., Rothlisberg, P.C., Dell, Q., Tonks, M., Lozano-Montes, H., Deng, R., Wassenberg, T., Okey, T.A., Revill, A., van der Velde, T., Moeseneder, C., Cheers, S., Donovan, A., Taranto, T., Salini, G., Fry, G., Tickell, S., Pascual, R., Smith, F., Morello, E., 2011. Effects of Trawling on the Benthos and Biodiversity: Development and Delivery of a Spatially-explicit Management Framework for the Northern Prawn Fishery. Final report to the project FRDC 2005/050. CSIRO Marine and Atmospheric Research, Cleveland, p. 382.

Chang, Y.H., Scrimshaw, M.D., Emmerson, R.H.C., 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. Sci. Total Environ. 221, 43—57.

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

Cressie, N.A., 1993. Statistics for Spatial Data (revised edition). John Wiley & Sons, Inc., New York.

Cressie, N., Wikle, C.K., 2011. Statistics for Spatio-temporal Data. John Wiley & Sons, New Jersey.

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

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

Englund, E., Weber, D., Leviant, N., 1992. The effects of sampling design parameters on block selection. Math. Geol. 24 (3), 329—343.

Erxleben, J., Elder, K., Davis, R., 2002. Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains. Hydrol. Process. 16, 3627—3649.

Fotheringham, A.S., Brunsdon, C., Charlton, M., 2002. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. John Wiley & Sons, Ltd, Chichester.

Friedel, M.J., Iwashita, F., 2013. Hybrid modelling of spatial continuity for application to numerical inverse problems. Environ. Model. Softw. 43, 60—79.

Gilardi, N., 2002. Machine Learning for Spatial Data Analysis, p. 73.

Gneiting, T., Balabdaoui, F., Raftery, A.E., 2007. Probabilistic forecasts, calibration and sharpness. J.R. Stat. Soc. Ser. B 69, 243—268.

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

Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J. Hydrol. 228, 113—129.

Goovaerts, P., 2008. Kriging and semivariogram deconvolution in presence of irregular geographical units. Math. Geosci. 40, 101—128.

Goovaerts, P., 2010. Geostatistical software. In: Fischer, M.M., Getis, A. (Eds.), Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer-Verlag, Berlin, pp. 125—134.

Gotway, C.A., Ferguson, R.B., Hergert, G.W., Peterson, T.A., 1996. Comparison of kriging and inverse-distance methods for mapping parameters. Soil Sci. Soc. Am. J. 60, 1237—1247.

Gotway, C.A., Stroup, W.W., 1997. A generalized linear model approach to spatial data analysis and prediction. J. Agric. Biol. Environ. Stat. 2 (2), 157—178.

Gotway, C.A., Young, L.J., 2002. Combining incompatible spatial data. J. Am. Stat. Assoc. 97, 632—648.

Gotway, C.A., Young, L.J., 2007. A geostatistical approach to linking geographically aggregated data from different sources. J. Comput. Graph. Stat. 16, 115—135.

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

Handcock, M.S., Stein, M.L., 1993. A Bayesian analysis of kriging. Technometrics 35, 403—410.

Harris, P., Brunsdon, C., Fotheringham, A.S., 2011. Links, comparisons and extensions of the geographically weighted regression model when used as a spatial predictor. Stoch. Environ. Res. Risk Assess. 25, 123—138.

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

Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D., New, M., 2008. A European daily high-resolution gridded data set of surfaces temperature and precipitation for 1950—2006. J. Geophys. Res. 113, D20119.

Hengl, T., 2007. A Practical Guide to Geostatistical Mapping of Environmental Variables. Office for Official Publication of the European Communities, Luxembourg, p. 143.

Hengl, T., Heuvelink, G.B.M., Rossiter, D.G., 2007. About regression-kriging: from equations to case studies. Comput. Geosci. 33, 1301—1315.

Hengl, T., Minasny, B., Gould, M., 2009. A geostatistical analysis of geostatistics. Scientometrics 80 (2), 491—514.

Hernandez-Stefanoni, J.L., Ponce-Hernandez, R., 2006. Mapping the spatial variability of plant diversity in a tropical forest: comparison of spatial interpolation methods. Environ. Monit. Assess. 117, 307—334.

Hosseini, E., Gallichand, J., Caron, J., 1993. Comparison of several interpolators for smoothing hydraulic conductivity data in South West Iran. Am. Soc. Agric. Eng. 36 (6), 1687—1693.

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

Huang, H., Chen, C., 2007. Optimal geostatistical model selection. J. Am. Stat. Assoc. 102 (479), 1009—1024.

Hutchinson, M.F., 1995. Interpolating mean rainfall using thin plate smoothing splines. Int. J. Geogr. Inform. Syst. 9 (4), 385—403.

Hutchinson, M.F., 2011. ANUSPLIN Version 4.3. Available on-line at. http://fennerschool.anu.edu.au/publications/software/anusplin.php.

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

Jensen, O.P., Christman, M.C., Miller, T.J., 2006. Landscape-based geostatistics: a case study of the distribution of blue crab in Chesapeake Bay. Environmetrics 17, 605—621.

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

Juang, K.W., Lee, D.Y., 1998. A comparison of three kriging methods using auxiliary variables in heavy-metal contaminated soils. J. Environ. Qual. 27, 355—363.

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

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

Kerry, R., 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.

Klesk, P., 2008. Construction of neurofuzzy network capable of extrapolating (and interpolating) with respect to the convex hull of a set of input samples in $R^n$. IEEE Trans. Fuzzy Syst. 16 (5), 1161—1179.

Knotters, M., Brus, D.J., 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., Bullock, D.G., 1999. A comparison study of interpolation methods for mapping soil properties. Agron J. 91, 393—400.

Krige, D.G., 1951. A statistical approach to some mine valuations problems at the Witwatersrand. J. Chem. Metall. Min. Soc. S. Afr. 52, 119—139.

Kyriakidis, P.C., Journel, A.G., 1999. Geostatistics space-time models: a review. Math. Geol. 31 (6), 651—684.

Lam, N.S.-N., 1983. Spatial interpolation methods: a review. Am. Cartogr. 10 (2), 129—139.

Lark, R.M., Cullis, B.R., 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. Eur. J. Soil Sci. 57 (6), 787—799.

Lark, R.M., 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. J. Am. Stat. Assoc. 89 (426), 391—400.

Laslett, G.M., McBratney, A.B., 1990. Further comparison of spatial methods for predicting soil pH. Soil Sci. Soc. Am. J. 54, 1553—1558.

Laslett, G.M., McBratney, A.B., Pahl, P.J., Hutchinson, M.F., 1987. Comparison of several spatial prediction methods for soil pH. J. Soil Sci. 38, 325—341.

Legendre, P., Legendre, L., 1998. Numerical Ecology, second ed. Elsevier, Amsterdam.

Li, J., 2013. Predictive modelling using random forest and its hybrid methods with geostatistical techniques in marine environmental geosciences. In: Christen, P., Kennedy, P., Liu, L., Ong, K.-L., Stranieri, A., Zhao, Y. (Eds.), The Proceedings of the Eleventh Australasian Data Mining Conference (AusDM 2013), Canberra, Australia, 13—15 November 2013, Conferences in Research and Practice in Information Technology, vol. 146.

Li, J., Heap, A., 2008. A Review of Spatial Interpolation Methods for Environmental Scientists. Geoscience Australia, p. 137. Record 2008/23.

Li, J., Heap, A., 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: performance and impact factors. Ecol. Inform. 6, 228—241.

Li, J., Heap, A., Potter, A., Daniell, J.J., 2011a. Predicting Seabed Mud Content across the Australian Margin II: Performance of Machine Learning Methods and Their Combination with Ordinary Kriging and Inverse Distance Squared. Geoscience Australia, p. 69. Record 2011/07.

Li, J., Heap, A.D., Potter, A., Daniell, J., 2011b. Application of machine learning methods to spatial interpolation of environmental variables. Environ. Model. Softw. 26, 1647—1659.

Li, J., Heap, A.D., Potter, A., Huang, Z., Daniell, J., 2011c. Can we improve the spatial predictions of seabed sediments? A case study of spatial interpolation of mud content across the southwest Australian margin. Cont. Shelf Res. 31, 1365—1376.

Li, J., Potter, A., Huang, Z., Daniell, J.J., Heap, A., 2010. Predicting Seabed Mud Content across the Australian Margin: Comparison of Statistical and Mathematical Techniques Using a Simulation Experiment. Geoscience Australia, p. 146, 2010/11.

Li, J., Potter, A., Huang, Z., Heap, A., 2012. Predicting Seabed Sand Content across the Australian Margin Using Machine Learning and Geostatistical Methods. Geoscience Australia, p. 115. Record 2012/48.

Li, Y., Shi, Z., Wu, C., Li, H., Li, F., 2007. Improved prediction and reduction of sampling density for soil salinity by different geostatistical methods. Agric. Sci. China 6 (7), 832—841.

Lin, G., Chen, L., 2004. A spatial interpolation method based on radial basis function networks incorporating a semivariogram model. J. Hydrol. 288, 288—298.

Little, L.S., Edwards, D., Porter, D.E., 1997. Kriging in estuaries: as the crow flies, or as the fish swims? J. Exp. Mar. Biol. Ecol. 213, 1—11.

López-Quílez, A., Muñoz, F., 2009. Geostatistical computing of acoustic maps in the presence of barriers. Math. Comput. Model. 50, 929—938.

MacEachren, A.M., Davidson, J.V., 1987. Sampling and isometric mapping of continuous geographic surfaces. Am. Cartogr. 14 (4), 299—320.

Marchant, B.P., Lark, R.M., 2006. Adaptive sampling and reconnaissance surveys for geostatistical mapping of the soil. Eur. J. Soil Sci. 57, 831—845.

Martínez-Cob, A., 1996. Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. J. Hydrol. 174, 19—35.

Matheron, G., 1963. Principles of geostatistics. Econ. Geol. 58, 1246—1266.

McKenney, D.W., Hutchinson, M.F., Papadopol, P., Lawrence, K., Pedlar, J., Campbell, K., Milewska, E., Hopkinson, R.F., Price, D., Owen, T., December 2011. Customized spatial climate models for North America. Bull. Am. Meteorol. Soc., 1611—1622.

Meul, M., Van Meirvenne, M., 2003. Kriging soil texture under different types of nonstationarity. Geoderma 112, 217—233.

Minasny, B., McBratney, A.B., 2007. Spatial prediction of soil properties using EBLUP with the Matern covariance function. Geoderma 140, 324—336.

Mitasova, H., Mitas, L., Brown, W.M., Gerdes, D.P., Kosinovsky, I., Baker, T., 1995. Modelling spatially and temporally distributed phenomena: new methods and tools for GRASS GIS. Int. J. Geogr. Inf. Syst. 9 (4), 433—446.

Moyeed, R.A., Papritz, A., 2002. An empirical comparison of kriging methods for nonlinear spatial point prediction. Math. Geol. 34 (4), 365—386.

Myers, D.E., 1994. Spatial interpolation: an overview. Geoderma 62, 17—28.

Nalder, I.A., Wein, R.W., 1998. Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. Agric. For. Meteorol. 92, 211—225.

Özkan, C., 2006. Surface interpolation by adaptive neuro-fuzzy inference system based local ordinary kriging. In: Narayanan, P.J., Narayanan, P.J., Narayanan, P.J. (Eds.), ACCV 2006, LNCS, vol. 3851. Springer-Verlag, Berlin, pp. 196—205.

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

Páez, A., Farber, S., Wheeler, D.C., 2011. A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships. Environ. Plan. A 43 (12), 2992—3010.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Comput. Geosci. 30, 683—691.

Pilz, J., Kazianka, H., Spöck, G., 2012. Some advances in Bayesian spatial prediction and sampling design. Spat. Stat. 1, 65—81.

Puente, C.E., Bras, R.L., 1986. Disjunctive kriging, universal kriging, or no kriging: small sample results with simulated fields. Math. Geol. 18 (3), 287—305.

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

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

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

Salski, A., 2003. Ecological applications of fuzzy logic. In: Recknagel, F. (Ed.), Ecological Informatics: Understanding Ecology by Biologically-inspired Computation. Springer, Berlin, pp. 3—14.

Sanabria, L.A., Qin, X., Li, J., Cechet, R.P., Lucas, C., 2013. Spatial interpolation of McArthur's forest fire danger index across Australia: observational study. Environ. Model. Softw. 50, 37—50.

Schloeder, C.A., Zimmerman, N.E., Jacobs, M.J., 2001. Comparison of methods for interpolating soil properties using limited data. Soil Sci. Soc. Am. J. 65, 470—479.

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

Spöck, G., 2012. Spatial sampling design based on spectral approximations to the random field. Environ. Model. Softw. 33, 48—60.

Stahl, K., Moore, R.D., Floyer, J.A., Asplin, M.G., 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. Agric. For. Meteorol. 139, 224—236.

Stein, A., Hoogerwerf, M., Bouma, J., 1988. Use of soil map delineations to improve (co-)kriging of point data on moisture deficits. Geoderma 43, 163—177.

Strebel, K., Espinosa, G., Giralt, F., Kindler, A., Rallo, R., Richter, M., Schlink, U., 2013. Modelling airborne benzene in space and time with self-organizing maps and Bayesian techniques. Environ. Model. Softw. 41, 151—162.

Tang, L., Su, X., Shao, G., Zhang, H., Zhao, J., 2012. A clustering-assisted regression (CAR) approach for developing spatial climate data sets in China. Environ. Model. Softw. 38, 122—128.

Verfaillie, E., van Lancker, V., van Meirvenne, M., 2006. Multivariate geostatistics for the predictive modelling of the surficial sand distribution in shelf seas. Cont. Shelf Res. 26, 2454—2468.

Vicente-Serrano, S.M., Saz-Sánchez, M.A., Cuadrat, J.M., 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. Clim. Res. 24, 161—180.

Voltz, M., Webster, R., 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. J. Soil Sci. 41, 473—490.

Wahba, G., 1990. Spline Models for Observational Data. Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania.

Wackernagel, H., 2003. Multivariate Geostatistics: An Introduction with Applications, third ed. Springer, Berlin.

Wang, H., Liu, G., Gong, P., 2005. Use of cokriging to improve estimates of soil salt solute spatial distribution in the Yellow River delta. Acta Geogr. Sin. 60 (3), 511—518.

Weber, D., Englund, E., 1992. Evaluation and comparison of spatial interpolators. Math. Geol. 24 (4), 381—391.

Webster, R., Oliver, M., 2001. Geostatistics for Environmental Scientists. John Wiley & Sons, Ltd, Chichester.

Webster, R., Oliver, M.A., 1992. Sample adequately to estimate variograms of soil properties. J. Soil Sci. 43, 177—192.

Wheeler, D.C., Páez, A., 2010. Geographically weighted regression. In: Fischer, M.M., Getis, A. (Eds.), Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer, Heidelberg, pp. 461—486.

Wood, S.N., 2006. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC, Boca Raton.

Wu, J., Norvell, W.A., 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.

Xu, T., Hutchinson, M.F., 2013. New developments and applications in the ANUCLIM spatial climatic and bioclimatic modelling package. Environ. Model. Softw. 40, 267—279.

Zhou, F., Guo, H.-C., Ho, Y.-S., Wu, C.-Z., 2007. Scientometric analysis of geostatistics using multivariate methods. Scientometrics 73 (3), 265—279.

Zhu, Q., Lin, H.S., 2010. Comparing ordinary kriging and regression kriging for soil properties in contrasting landscapes. Pedosphere 20 (5), 594—606.

Zimmerman, D., Pavlik, C., Ruggles, A., Armstrong, M.P., 1999. An experimental comparison of ordinary and universal kriging and inverse distance weighting. Math. Geol. 31 (4), 375—390.