Ecological Informatics 6 (2011) 228–241

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors

Jin Li *, Andrew D. Heap

Marine & Coastal Environment, PMD, Geoscience Australia, GPO Box 378, Canberra, ACT 2601, Australia

ARTICLE INFO

Article history:
Received 25 August 2009
Received in revised form 22 November 2010
Accepted 5 December 2010
Available online 16 December 2010

Keywords:
Spatial interpolator
Geostatistics
Kriging
Data variation
Sample density

ABSTRACT

Spatial interpolation methods have been applied to many disciplines. Many factors affect the performance of the methods, but there are no consistent findings about their effects. In this study, we use comparative studies in environmental sciences to assess the performance and to quantify the impacts of data properties on the performance. Two new measures are proposed to compare the performance of the methods applied to variables with different units/scales. A total of 53 comparative studies were assessed and the performance of 72 methods/sub-methods compared is analysed. The impacts of sample density, data variation and sampling design on the estimations of 32 methods are quantified using data derived from their application to 80 variables. Inverse distance weighting (IDW), ordinary kriging (OK), and ordinary co-kriging (OCK) are the most frequently used methods. Data variation is a dominant impact factor and has significant effects on the performance of the methods. As the variation increases, the accuracy of all methods decreases and the magnitude of decrease is method dependent. Irregular-spaced sampling design might improve the accuracy of estimation. The effect of sampling density on the performance of the methods is found not to be significant. The implications of these findings are discussed.

Contents

1. Introduction

Spatially continuous data (or spatially continuous surfaces) play a significant role in environmental sciences and management. Environmental managers usually require spatially continuous data over the

* Corresponding author. Tel.: +61 2 6249 9899; fax: +61 2 6249 9956.
E-mail address: Jin.Li@ga.gov.au (J. Li).

1574-9541/$ – see front matter. Crown Copyright © 2010 Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.ecoinf.2010.12.003

region of interest to make effective and confident decisions, and scientists need accurate spatially continuous data across a region to make justified interpretations. Such data are, however, usually not always readily available and often difficult and expensive to acquire, especially for mountainous or deep marine regions. Moreover, environmental data collected from field surveys are often from point sources. Thus, the values of an attribute at unsampled points need to be estimated in order to generate spatially continuous data. In such instances, spatial interpolation methods provide a tool for estimating the values of an environmental variable at unsampled sites using data from point observations.

Spatial interpolation methods are developed for specific data types or a specific variable. The key features of the common used methods have been compared by Li and Heap (2008). They have been applied to various disciplines, such as mining engineering (Journel and Huijbregts, 1978) and environmental sciences (Goovaerts, 1997; Burrough and McDonnell, 1998; Webster and Oliver, 2001). A bibliographic research (Zhou et al., 2007) found that the top 10 fields that employ geostatistics are: 1) geosciences, 2) water resources, 3) environmental sciences, 4) agriculture or soil sciences, 5) mathematics, 6) statistics and probability, 7) ecology, 8) civil engineering, 9) petroleum engineering, and 10) limnology.

Many factors affect the performance of spatial interpolation methods. These include: sampling density (Isaaks and Srivastava, 1989; Englund et al., 1992; Burrough and McDonnell, 1998; Dirks et al., 1998; Hartkamp et al., 1999; Stahl et al., 2006), sample spatial distribution (Collins and Bolstad, 1996), sample clustering (Isaaks and Srivastava, 1989; Laslett, 1994; Zimmerman et al., 1999), surface type (MacEachren and Davidson, 1987; Stein et al., 1988; Voltz and Webster, 1990; Zimmerman et al., 1999), data variance (Collins and Bolstad, 1996; Martínez-Cob, 1996; Schloeder et al., 2001), data normality (Rossi et al., 1992; Weber and Englund, 1992; Cressie, 1993; Wu et al., 2006), quality of secondary information (Ahmed and De Marsily, 1987; Collins and Bolstad, 1996; Martínez-Cob, 1996; Goovaerts, 1997; Juang and Lee, 1998; Goovaerts, 2000; Bishop and McBratney, 2001; Wang et al., 2005; Hernandez-Stefanoni and Ponce-Hernandez, 2006; Hengl, 2007), stratification (Brus et al., 1996; Voltz and Webster, 1990), and grid size or resolution (Hengl, 2007). Interactions among different factors may also exist (Zimmerman et al., 1999). 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). However, there are no consistent findings about how these factors affect the performance of the spatial interpolators.

In this review, we assess the performance of spatial interpolation methods using the information derived from the comparative studies of the methods in environmental sciences that include meteorology and water resources, ecology, agriculture and soil science and marine environmental science. We also analyse the impacts of sample density (km^2/sample), data variation in terms of coefficient of variation (CV), and sampling design based on the information from the comparative studies.

2. Assessment measures

With wide and increasing applications of spatial interpolation methods, there is a growing concern about their accuracy and precision (Hartkamp et al., 1999). The statistics of the differences (absolute and squared) between measured and predicted values at sampled points are often used as an indicator of the performance of an inexact method (Burrough and McDonnell, 1998). Several error measures have been proposed. Measures for assessing the performance of the spatial interpolation methods have been briefly reviewed by Li and Heap (2008) as in (Table 1). Commonly used error measures include: mean error (ME), mean absolute error (MAE), mean squared error (MSE) and root mean squared error (RMSE). ME is used for determining the degree of bias in estimates and is often referred to as "bias" (Isaaks and Srivastava, 1989), but it should be used cautiously as an indicator of accuracy because negative and positive estimates counteract each other and the resultant ME tends to be lower than actual error (Nalder and Wein, 1998). In this table, we also include the mean of percent error (MPE) indicating a better model when its value approaching zero and the ratio of the standard deviation of predicted to observed values (RSD) indicating a better model as its value approaching 1. MPE is a mean of ME and suffers the shortage of ME, and RSD is equivalent to RVar. RMSE provides a measure of error size, but it is sensitive to outliers as it places a lot of weight on large errors (Hernandez-Stefanoni and Ponce-Hernandez, 2006). MSE suffers the same drawbacks as RMSE, whereas MAE is less sensitive to extreme values (Vicente-Serrano et al., 2003) and indicates the extent to which the estimate can be in error (Nalder and Wein, 1998). MAE and RMSE are similar measures because they give estimates of the average error, but they do not provide information about the relative size of the average difference and the nature of differences comprising them (Willmott, 1982). However, MAE and RMSE are among the best overall measures of model performance as they summarise the mean difference in the units of observed and predicted values (Willmott, 1982).

All of the measures listed in Table 1 assess the performance of spatial interpolation methods for individual primary variables. The magnitude of these measures depends on the unit/scale of the primary variable. For example, the magnitude of MAE is usually lower than 1% for soil organic matter, but it can be easily beyond 100 mm for rainfall or even over 1000 mm for tropical regions. In this review, we need to compare the performance of spatial interpolation methods among different studies, in which the primary variables are usually in different units/scales. It is not

Table 1
Measurements used to assess the performance of spatial interpolation methods (Ahmed and De Marsily, 1987; Burrough and McDonnell, 1998; Hu et al., 2004; Isaaks and Srivastava, 1989; Willmott, 1981; Vicente-Serrano et al., 2003). The measurements considered are those commonly used, so the list of the measurements is intentionally non-exhaustive.

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

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

possible to use these unit-/scale-dependent measures for such comparisons, so new measures are needed.

On the basis of MAE and RMSE, we propose two new measures that remove the effect of unit/scale and are then not sensitive to the changes in unit/scale. The first is relative mean absolute error (RMAE) that is given as:

$$RMAE = \frac{1}{n}\sum_{i=1}^{n} |(p_i - o_i)/o_i| 100 \quad (1)$$

And the second is relative root mean square error (RRMSE), as follows:

$$RRMSE = [\frac{1}{n}\sum_{i=1}^{n} ([p_i - o_i]/o_i)^2]^{1/2} 100 \quad (2)$$

where n is number of observations or samples, o is observed value, p is predicted or estimated values.

However, information required to calculate RMAE and RRMSE is not always available. To overcome this problem, RMAE and RRMSE are modified by using MAE/mean (*i.e.*, mean of the validation dataset) and RMSE/mean instead. The latter is also called standardised RMSE (Haberlandt, 2007). The mean of the validation dataset is, however, often not reported either. They are further modified by using the mean of the dataset for estimation that is more frequently available. These two new measures and their modified versions can be interpreted as measures of the relative errors in estimations. They are used in this study to compare the performance of spatial interpolation methods applied to variables with different units/scales from various studies in different disciplines.

3. Comparison of spatial interpolation methods applied to various disciplines

In this review, we select 53 comparative studies that compared the performance of spatial interpolation methods in environmental sciences. Three significant challenges were encountered in this review: 1) sometimes the same method is presented with different names; 2) different mathematical symbols are often used although they represent the same concept; and 3) methods are not described clearly in some studies. Efforts have been made to match different names and symbols with the right methods and concepts and to assign the correct names to the methods used in various studies. When it was impossible to find information on the method used in a study, the study is either discarded or a note is made to avoid confusion.

3.1. Comparative studies

The spatial interpolation methods compared are listed in Table 2 and the results are briefly stated for the 53 comparative studies in Table 3. The comparative studies are also summarised in the Supplementary online material (Supplementary 1), including the methods compared, sampling design, sample size, area of the region studied, and the associated results. The frequency and recommendation times of each spatial interpolation method compared in these 53 studies are summarised in Table 4. The recommendations of a spatial interpolation method (Table 4) should be assessed in together with the methods compared in Table 3 as they depend on the methods compared.

The 53 comparative studies reveal the following features:

1. the spatial interpolation methods have been applied widely in environmental sciences, with 72 methods/sub-methods employed including combined methods;
2. different studies have compared a suite of different methods, which makes it difficult to draw general conclusions. However, ordinary kriging (OK), inverse distance weighting (IDW) including inverse distance squared (IDS) and ordinary co-kriging (OCK) are the most frequently compared methods (Table 4);
3. in general, kriging methods perform better than non-geostatistical methods; of which kriging with an external drift (KED) is the most highly recommended method; and
4. gradient plus inverse distance squared (GIDS) and other highlighted methods in Table 4 have low frequency, but are worthy of attention because of their perfect rate of recommendation.

Table 2
Full name for the abbreviations of spatial interpolation methods in this review.

Abbreviation	Full name	Abbreviation	Full name
AK	Akima's interpolator	OIK	Ordinary indicator kriging
BK	Block kriging	OK	Ordinary kriging
CART	Regression tree	RBFN	Radial basis function network
CK	Cokriging	REML	Residual maximum likelihood method
Cl	Classification	REML-EBLUP	Residual maximum likelihood-empirical best linear unbiased predictor
DK	Disjunctive kriging	RK*	Regression kriging
GAM	Generalised additive model	SCK	Simple CK
GIDS	Gradient plus inverse distance squared	SK	Simple kriging
GM	Global mean	SKlm	SK with varying local means
IDS	Inverse distance squared	SOCK	Standardised OCK
IDW	Inverse distance weighting	StIDS	Stratified IDS
IK	Indicator kriging	StIDW	Stratified IDW
IKED	IK with external drift	StGM	Stratified GM
KED	Kriging with an external drift	StNN	Stratified NN
LM	Linear regression model	StOCK	Stratified OCK
LR	Lapse rate	StOK	Stratified OK
LSZ	Bayesian MWRCK	StSK	Stratified SK
MBK	Model-based kriging	StTPS	Stratified TPS
MWRCK	Moving window regression residual cokriging	T2R	Topo to Raster
NaN	Natural neighbours	TIN	Triangular irregular network
NN	Nearest neighbours	TPS	Thin plate splines
OCCK	Ordinary colocated cokriging	TSA	Trend surface analysis
OCK	Ordinary CK	UK	Universal kriging
OICK	Ordinary indicator cokriging		

* There are a few types of RK and their definitions are given by Li and Heap (2008).

3.2. Spatial interpolation methods

To compare the performance of spatial interpolation methods, the following information is considered essential: 1) sampling design; 2) the mean and CV of the primary variable for either the estimation dataset or validation dataset; 3) the sample size for the estimation and validation datasets; 4) the area of the region studied; and 5) appropriate accuracy measurements of the spatial interpolation methods (*i.e.*, MAE and/or RMSE or MSE). Of course, the spatial interpolation methods need to be properly named, appropriately referenced, or clearly described. Only 18 of the 53 comparative studies reported all of the essential information and are selected for a quantitative comparative analysis. In these 18 studies, 32 methods/sub-methods were applied to 80 environmental variables (*i.e.*, cases) (Supplementary 2). For some methods, the method and its variants have to be grouped into one method. Taking IDW as an example, some studies clearly stated the power of distance, but in others such information was not provided, so IDW and its variants are treated as a single

Table 3
Summary of the 53 comparative studies reviewed. The full name of the abbreviation of each spatial interpolation method is listed in Table 1 and these methods are briefly described by Li and Heap, 2008.

No	Reference	Discipline	Methods compared	Result
1	Hosseini et al. (1993)	Meteorology and Water resources	OK, UK, TSA, IDW and AK	OK preferred.
2	Collins and Bolstad (1996)		IDS, IDW, splines, LM, TSA, LR, kriging and CK	LM the best.
3	Martínez-Cob (1996)		OK, OCK and RK-C	OCK more accurate.
4	Nalder and Wein (1998)		GIDS, IDS, NN, CK, OK, RK-C and UK	GIDS preferred.
5	Sun (1998)		MWRCK, CK and LSZ	LSZ the best.
6	Hartkamp et al. (1999)		IDW, TPS and OCK	No difference, but TPS preferred.
7	Goovaerts (2000)		SKlm, KED, OCCK, LM, NN, IDS and OK	SKlm the best.
8	Javis and Stuart (2001)		LM-IDW, TSA, RK-C and TPS	TPS the best.
9	Erxleben et al. (2002)		IDW, OK, RK-C, OCK, CART with OK and CART with OCK	CART with OK and CART with OCK more accurate.
10	Vicente-Serrano et al. (2003)		TSA, LM, NN, IDW, splines, SK, OK, BK, UK, OCK, LM with IDS and splines with LM	Kriging and LM more accurate.
11	Lin and Chen (2004)		RBFN, improved RBFN and OK	Improved RBFN the best.
12	Naoum and Tsanis (2004)		Splines, IDW, NN, LM and kriging	Kriging preferred.
13	Li et al. (2005)		IDS, OK, OCK and OK combined with LR	OK combined with LR the best.
14	Mardikis et al. (2005)		GIDS, IDS, OK and RK-C	GIDS the best.
15	Jef et al. (2006)		RK-C, IDW (with distance power 4) and LM with IDW	RK-C the best.
16	Haberlandt (2007)		NN, IDS, OK, OIK, KED and IKED	KED the best.
17	Hernandez-Stefanoni and Ponce-Hernandez (2006)	Ecology	OK, OCK, IDS, StOK, StOCK, StIDW and Cl	StOK the best.
18	Van Kuilenburg et al. (1982)	Agriculture and soil science	NN, IDS and OK	OK preferred.
19	Ahmed and De Marsily (1987)		OCK, KED, RK-A and RK-B	OCK and RK-A preferred.
20	Laslett et al. (1987)		TPS, OK, GM and global medians, NN, IDW-0, IDS, AK, NaN and TSA	TPS and OK better.
21	Laslett and McBratney (1990)		NN, TPS, AK, SK? and REML UK	REML UK the best.
22	Voltz and Webster (1990)		SK, StSK, Cl and cubic splines	StSK the best.
23	Laslett (1994)		Cubic splines and SK	SK better.
24	Odeh et al. (1994)		LM, OK, UK, OCK, RK-A, and RK-B	RK the best.
25	Knotters et al. (1995)		OK, OCK and RK-A	RK-A the best.
26	Odeh et al. (1995)		LM, OK, UK, OCK, RK-A, RK-B and RK-C	RK-C the best.
27	Brus et al. (1996)	Agriculture and soil science	Cl, GM, IDS, OK, NN, IDW-0, TPS and their combination with soil strata	StOK the best.
28	Gotway et al. (1996)		OK and IDW	OK better.
29	Goovaerts (1997)		OCK, SCK, SOCK and OCCK	OCK and SOCK better.
30	Goovaerts (1997)		KED and SKlm	Similar.
31	Goovaerts (1997)		OIK and OICK	Similar.
32	Kravchenko and Bullock (1999)		OK, lognormal OK and IDW	Lognormal OK better.
33	Bourennane et al. (2000)		KED and LM	KED better.
34	Bishop and McBratney (2001)		GAM, LM, CART, OK, KED, RK-F and RK-C	KED the best.
35	Schloeder et al. (2001)		OK, IDW and TPS	OK and IDW better.
36	Moyeed and Papritz (2002)		OK, lognormal OK, DK, IK and MBK	Similar.
37	Meul and Van Meirvenne (2003)		OK, UK, SKlm, OCK and UK + OCK	UK + OCK the best.
38	Hengl et al. (2004)		Cl, OK, LM and RK-C	RK-C preferred.
39	Hu et al. (2004)		SK, OK, lognormal kriging, UK, DK and IDW	UK the best.
40	Wang et al. (2005)		TSA-OK and TSA-OCK	TSA-OCK better.
41	Wu et al. (2006)		OK and OCK	OCK better.
42	Li et al. (2007)		OK, OCK and RK-E	RK-E better.
43	Minasny and McBratney (2007)		REML-EBLUP, OK and RK-C	RK-C recommended.
44	Rivoirard and Wieland (2001)	Marine environmental science	KED and OK	KED better.
45	ICES (2005)		OK and KED	KED better.
46	Verfaillie et al. (2006)		OK, KED and LM	KED the best.
47	Bello-Pineda and Hernández-Stefanoni (2007)		BK, IDS and IDW	BK the best.
48	Ruddick (2007)		OK, OCK, IDW, NN and T2R	Similar.
49	Puente and Bras (1986)	Other fields	UK, DK and local mean estimator	UK better.
50	Boufassa and Armstrong (1989)		OK, lognormal OK, SK, lognormal SK, disjunctive OK and disjunctive SK	SK and OK recommended.
51	Isaaks and Srivastava (1989)		OK, IDS, TIN and NN	OK the best.
52	Weber and Englund (1992)		OK, SK, lognormal OK, rank OK, GM, IDS, IDW, TSA and Projected Slope	IDS better.
53	Zimmerman et al. (1999)		OK, UK and IDS	OK preferred.

method. IDS is separated from other variants of IDW if the information of the power is available.

3.2.1. Frequency of the spatial interpolation methods compared

The frequency of individual spatial interpolation method applied to the 80 cases varies considerably among methods (Fig. 1). The spatial interpolation methods can be divided into four groups in terms of their frequency. The first group contains OK, IDW and IDS which were the most frequently compared methods with a frequency of >30. The next group contains OCK, regression residual kriging (RK-C) and thin plate splines (TPS) which were frequently compared methods with a frequency between 20 and 30. The third group includes Cl,

Table 4
Frequency of the spatial interpolation methods compared and the number of times the method was recommended in the 53 comparative studies. Methods with 100% rate of recommendation are highlighted.

No.	Method	Frequency	Rate of recommendation (%)	No.	Method	Frequency	Rate of recommendation (%)
1	OK	36	22	37	IK	1	0
2	IDW	16	6	38	IKED	1	0
3	OCK	15	27	39	Improved RBFN	1	100
4	LM	10	20	40	LM with IDS	1	0
5	IDS	14	7	41	LM with IDW	1	0
6	NN	11	0	42	local mean	1	0
7	RK-C	10	40	43	lognormal kriging	1	0
8	KED	9	67	44	lognormal SK	1	0
9	UK	9	22	45	LR	1	0
10	SK	7	29	46	**LSZ**	1	100
11	TPS	6	50	47	MBK	1	0
12	TSA	6	0	48	MWRCK	1	0
13	Splines	5	0	49	NaN	1	0
14	CI	4	0	50	OICK	1	0
15	lognormal OK	4	25	51	**OK combined with LR**	1	100
16	RK-A	4	75	52	Projected Slope	1	0
17	AK	3	0	53	RBFN	1	0
18	BK	2	50	54	**REML UK**	1	100
19	CK	2	0	55	REML-EBLUP	1	0
20	DK	3	0	56	**RK-E**	1	100
21	GM	3	0	57	RK-F	1	0
22	kriging	2	50	58	rank OK	1	0
23	RK-B	3	33	59	SCK	1	0
24	SKlm	3	33	60	**SOCK**	1	100
25	**GIDS**	2	100	61	Splines with LM	1	0
26	OCCK	2	0	62	StGM	1	0
27	OIK	2	0	63	StIDS	1	0
28	**StIDW**	2	0	64	StNN	1	0
29	**StOK**	2	100	65	StOCK	1	0
30	**CART with OCK**	1	100	66	**StSK**	1	100
31	CART	1	0	67	StTPS	1	0
32	**CART with OK**	1	100	68	T2R	1	0
33	disjunctive OK	1	0	69	TIN	1	0
34	disjunctive SK	1	0	70	**TSA-OCK**	1	100
35	GAM			71	**TSA-OK**	1	0
36	Global medians	1	0	72	**UK + OCK**	1	100

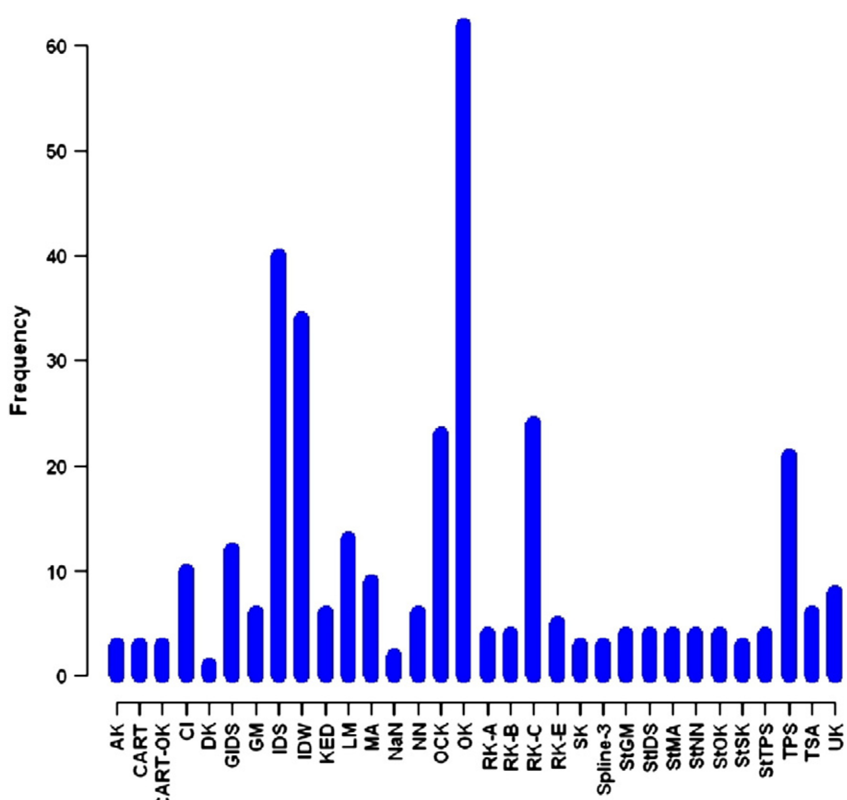

Fig. 1. The frequency of 32 spatial interpolation methods compared using 80 cases in the 18 comparative studies.

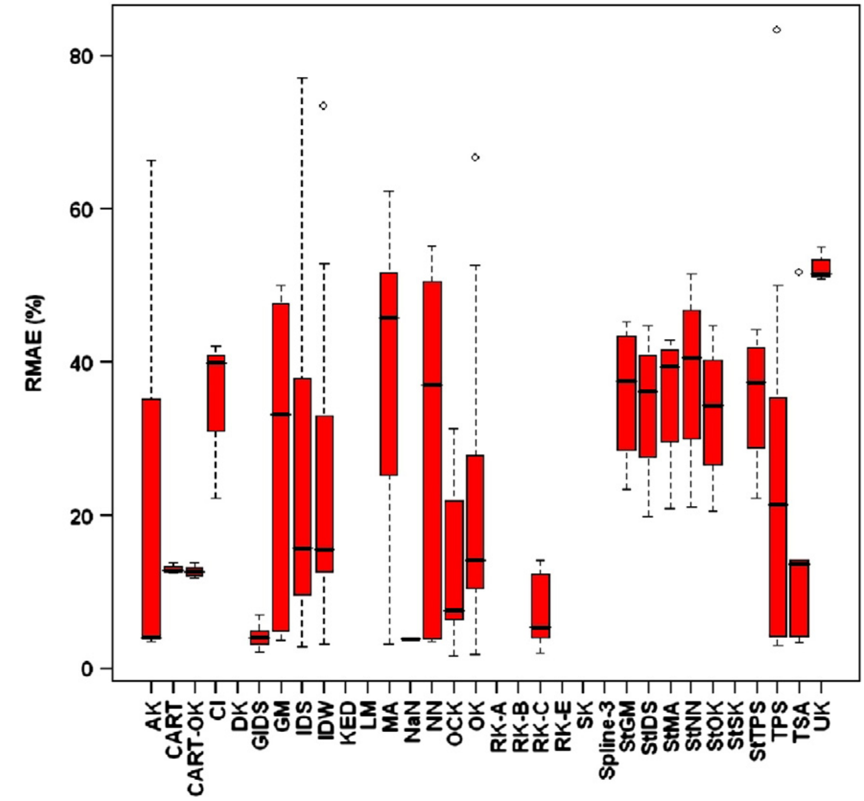

Fig. 2. The accuracy of 32 spatial interpolation methods compared using 80 cases in the 18 comparative studies in terms of RMAE (%).

GIDS, and LM which were less frequently compared, with a frequency between 10 and 20. The last group contains the remaining methods that were only occasionally compared with a frequency of <10.

3.2.2. Performance of the spatial interpolation methods compared

The performance of the spatial interpolation methods compared in the 80 cases exhibits dramatic variation (Figs. 2 and 3) and the

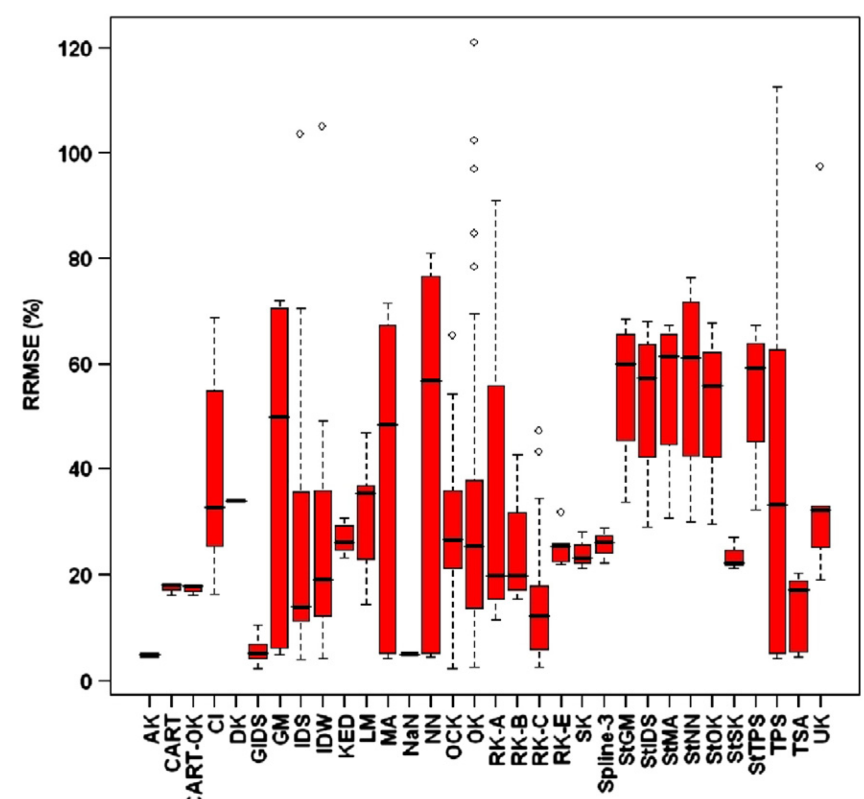

Fig. 3. The accuracy of 32 spatial interpolation methods compared using 80 cases in the 18 comparative studies in terms of RRMSE (%).

Table 5
Effects of methods, CV, sample density and sampling design on the performance of spatial interpolation methods in terms of RMAE (%). Due to unbalanced design, the interactions between methods and other variables were not considered in this analysis. The data were extracted from Supplementary 2. The data were analysed using a linear model with arcsine transformation of RMAE (%) in R (R Development Core Team, 2007).

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Methods	22	3.5960	0.1635	72.8664	0.0000
CV	1	6.4592	6.4592	2879.4499	0.0000
Sample density	1	0.0030	0.0030	1.3469	0.2476
Sampling design	1	0.0126	0.0126	5.6207	0.0190
CV: Sample density	1	0.0147	0.0147	6.5712	0.0113
CV:Sampling design	1	0.0260	0.0260	11.5727	0.0009
Sample density: Sampling design	1	0.0023	0.0023	1.0306	0.3116
CV: Sample density: Sampling design	1	0.0257	0.0257	11.4725	0.0009
Residuals	156	0.3499	0.0022		

difference in the performance between the methods is significantly different in terms of RMAE and RRMSE (Tables 5 and 6). The RMAE values of some methods, such as KED, linear regression model (LM), are missing because the studies reviewed did not report such information. Of the 32 compared methods, GIDS, natural neighbours (NaN), RK-C and OCK appeared to be more accurate than the rest of the methods in terms of RMAE; and GIDS, NaN and RK-C were more accurate than the rest of the methods in terms of RRMSE. Since NaN was only an occasionally used method (Fig. 1), the result is not reliable due to the small number of times it was applied. The performance of the methods also depended on CV (Figs. 4 and 5). GIDS and RK-C were the most accurate methods, but they were only applied to datasets with small CV. GIDS, OCK and RK-C were less sensitive to CV.

These conclusions are only based on the results from 80 cases in the 18 comparative studies. Other methods may display similar features, but were not compared in this review due to the lack of relevant information for appropriate comparison between different variables. Moreover, some comparative studies may have been missed because only 53 comparative studies are assessed even though there are 2866 publications in geostatistics between 1967 and 2005 (Zhou et al., 2007). Nonetheless, the most influential comparative studies are believed to have been included in this review.

3.3. Complicating and confounding factors

Several complicating and confounding factors were encountered in this review that may have some bearing on the results. Information about the study region, experimental design and primary variable were missing at times. The measures for the performance of the spatial interpolation methods varied between studies. Occasionally the methods used for interpolation were not clearly or adequately described or referenced. For example, some studies mentioned the use of kriging or cokriging. This is not sufficient because there are many different kriging methods and more than one cokriging method. All of these factors make it difficult to compare the performance of the spatial interpolation methods using results from the published studies, consequently preventing any possible generalisation of the observed patterns. Only 5 out of 16 studies in meteorology and water resources, and 13 of 26 studies in soil science provided sufficient information for possible comparison between different variables and studies. All of the studies reviewed in the other disciplines failed to report relevant statistics for further comparative research between different variables.

Here we would recommend that future studies should report relevant information clearly in their publications, including: the area of region studied; experimental and sampling design, particularly the sample size of datasets for estimation and validation; summary statistics of the primary variable for both estimation and validation datasets; and appropriate references or descriptions of the spatial interpolation methods used. The measures should include at least MAE or MSE for comparing the results of different variables. Correlation coefficient has been used in many studies as a measure of the performance of the spatial interpolation methods. However, as discussed by Li and Heap (2008), it is often misleading and it should be either avoided or extreme care should be taken in using it.

Table 6
Effects of methods, CV, sample density and sampling design on the performance of spatial interpolation methods in terms of RRMSE (%). The data were extracted from Supplementary 2. Due to the unbalanced design, the interactions between methods and other variables were not considered in this analysis. The data were analysed using a linear model with square root transformation of RRMSE (%) in R (R Development Core Team, 2007).

	Df	Sum sq	Mean sq	F value	Pr(>F)
Methods	31	374.4800	12.0800	19.6441	0.0000
CV	1	715.4300	715.4300	1163.4079	0.0000
Sample density	1	2.5700	2.5700	4.1823	0.0419
Sampling.design	1	8.0000	8.0000	13.0124	0.0004
CV: Sample density	1	0.0013	0.0013	0.0021	0.9632
CV:Sampling.design	1	23.3500	23.3500	37.9769	0.0000
Sample density: Sampling.design	1	4.0700	4.0700	6.6117	0.0107
CV: Sample density: Sampling.design	1	5.5900	5.5900	9.0858	0.0028
Residuals	257	158.0400	0.6100		

4. Factors affecting the performance of spatial interpolation methods

4.1. Results

The effects of impact factors that are sampling density (km^2/sample), CV, and sampling design (regular vs. irregular) on the performance of various methods was quantified in terms of RMAE and RRMSE using the data from 80 cases in the 18 comparative studies (Tables 5 and 6). CV and sampling design significantly affected the performance of the methods in terms of both RMAE and RRMSE; sampling density had little impact on the performance in term of RRMSE; and there were strong interactions among these three factors in terms of both RMAE and RRMSE. Studies with regular design had high sampling density (Figs. 6a and 7a), whereas the sampling density for studies with irregular design varied (Figs. 6b and 7b). For RMAE, it is apparent that as CV increased, RMAE also increased but it increased more for studies with regular sampling design than for those with irregular sampling design (Fig. 6). Interactions among these three factors are shown by a decrease of RMAE at the intermediate sampling density for studies with irregular design. In terms of RRMSE, the patterns observed (Fig. 7) were similar to those for RMAE. The relationship between CV and RMAE (Fig. 6b) and RRMSE (Fig. 7b) at low sampling density was weak. This was further illustrated in Figs. 8 and 9, which showed that there was an apparent pattern between the performance of the methods and CV for each method, although the magnitude changed with methods.

4.2. Discussion

4.2.1. Sampling density

It is often argued that if the sample density is big enough, then its effects would disappear, which means that a threshold number of sample density exists where addition of more samples does not improve the performance of a spatial interpolation method (Li and Heap, 2008). This assertion is not supported by the findings in this review. Our results show that sampling density had little impact on the performance of the methods (Figs. 6 and 7), although sampling

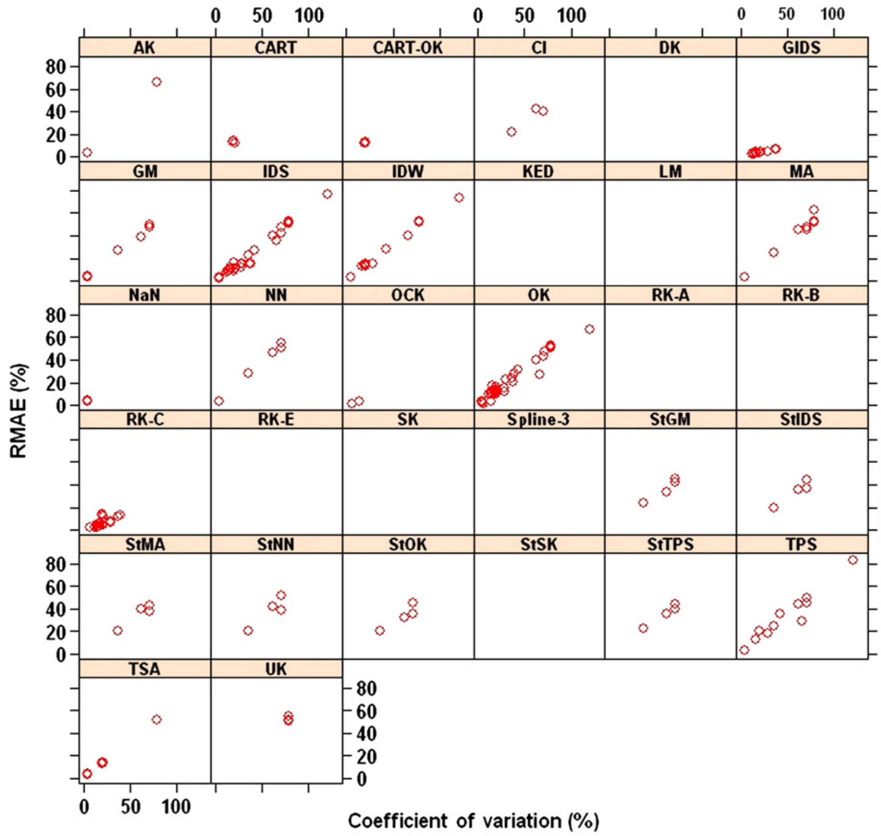

Fig. 4. Effects of the variation in the data on the accuracy of each spatial interpolation method compared using 80 cases in the 18 comparative studies in terms of RMAE (%).

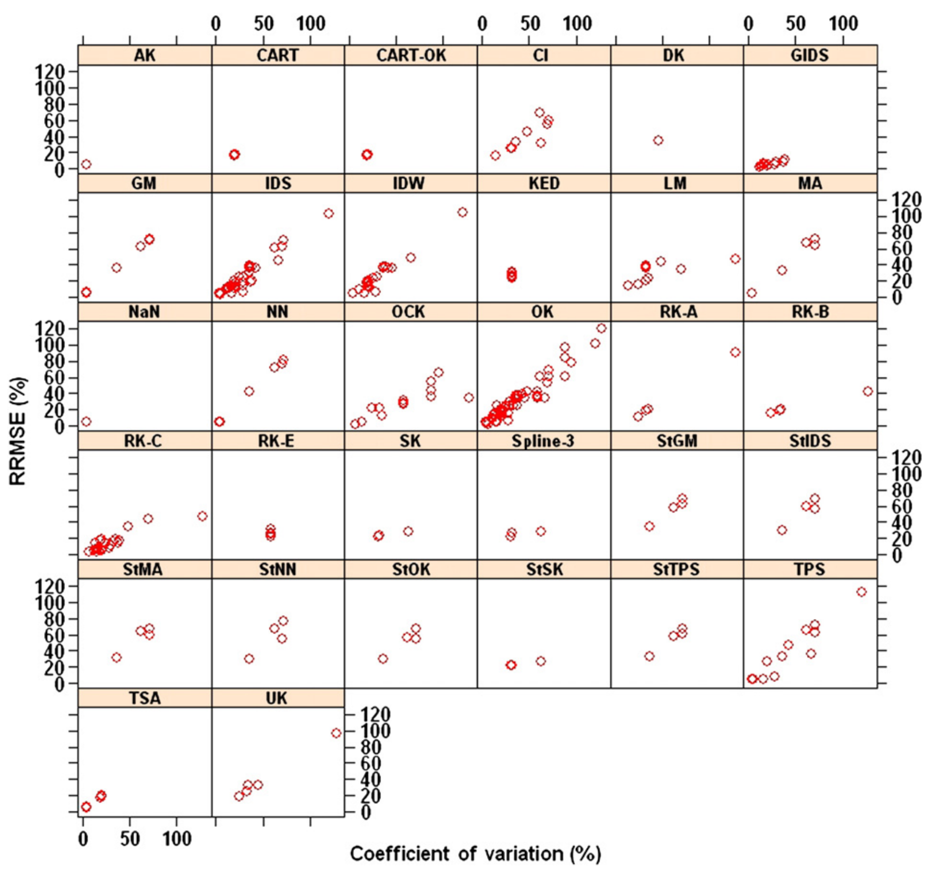

Fig. 5. Effects of the variation in the data on the accuracy of each spatial interpolation method compared using 80 cases in the 18 comparative studies in terms of RRMSE (%).

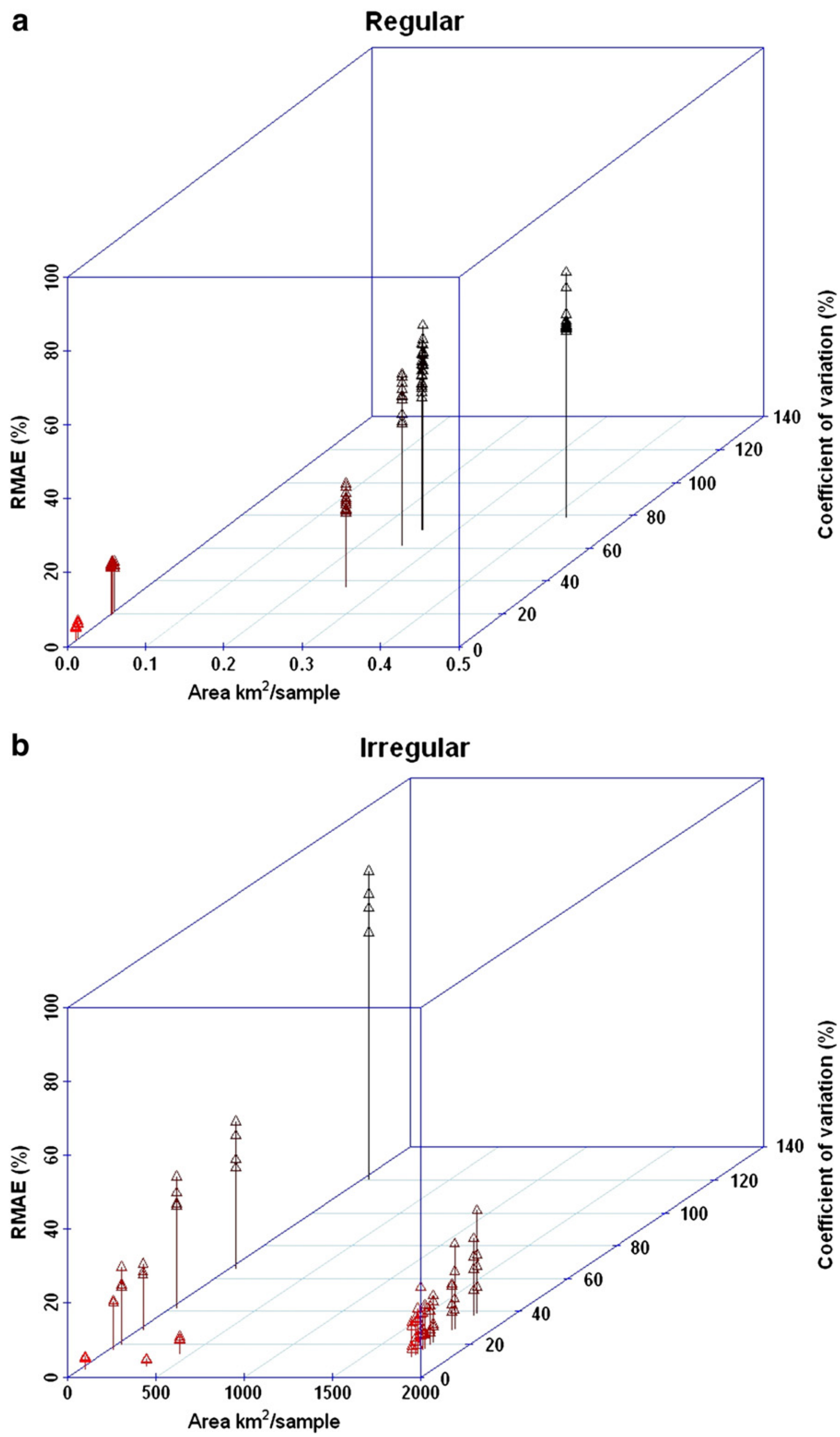

Fig. 6. Effects of sampling density (km^2/sample) and variation in the data on the accuracy of the spatial interpolation methods compared using 80 cases in the 18 comparative studies in terms of RMAE (%): (a) regular sampling and (b) irregular sampling.

density significantly interacted with sample density and sampling design in terms of RMAE (Table 5) and RRMSE (Table 6). For each method, there is little relationship between the performance and the sampling density in terms of RMAE and RRMSE. The difference in observations for RMAE and RRMSE is due to the fact that in some studies, both MAE and RMSE were reported, but in some only one of them was presented. This non-significant effect of sampling density is probably because different studies focused on phenomena that are on different scales. To achieve certain accuracy in spatial interpolation, different phenomenon may have different sample density even if the sample size is the same. For example, phenomena at small scale (e.g., patterns of ripples on a dune) would result in relatively high sample density due to relatively small survey area, while sampling phenomena at larger scale (e.g., patterns of dunes) the same sample size as for

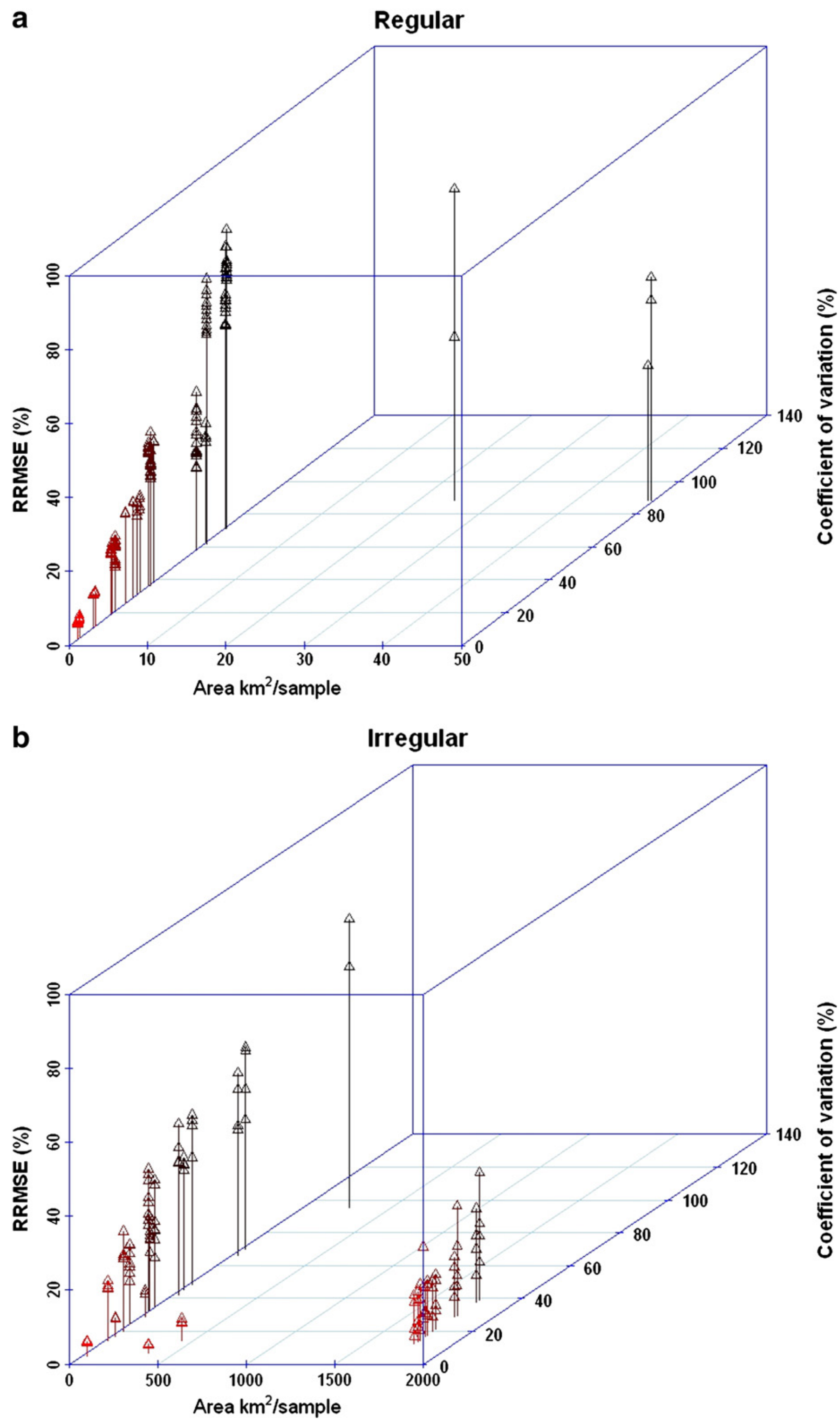

Fig. 7. Effects of sampling density (km^2/sample) and variation in the data on the accuracy of the spatial interpolation methods compared using 80 cases in the 18 comparative studies in terms of RRMSE (%): (a) regular sampling and (b) irregular sampling.

ripple would be sufficient, but resultant sample density would be relatively lower because of larger survey area. It is apparent that size of study area plays an important role. Therefore, comparison of the findings from studies of phenomena on different scales may not be able to detect the effects of sample density. Thus this non-significant effect of sampling density observed does not necessarily suggest that we should not increase the sample density in any future study.

Sampling density has long been argued to play a significant role in the performance of the spatial interpolation methods. When data density is high, most methods produce similar results (Burrough and McDonnell, 1998), which is supported by previous findings (Hosseini et al., 1993; Dirks et al., 1998).

When sample density is relatively low, the performance of the spatial interpolation methods is better when the sample density increases (Isaaks

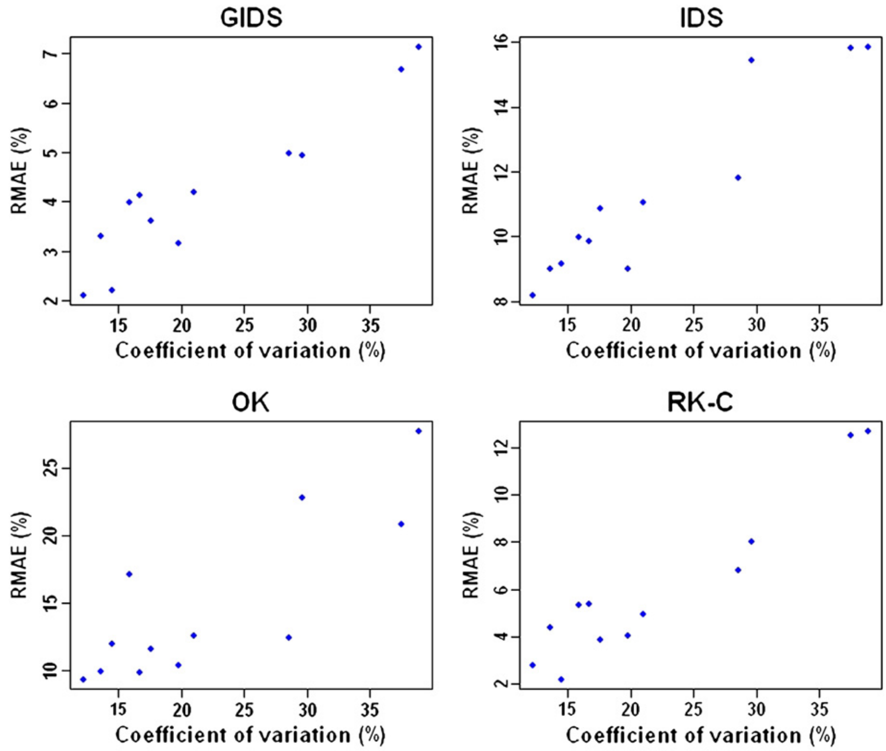

Fig. 8. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in relation to low sample density (1783 km^2 per sample) in terms of RMAE (%).

and Srivastava, 1989; Englund et al., 1992; Stahl et al., 2006). Gotway et al. (1996) found that the use of wider sample spacings greatly reduced the information in the resultant maps, although the sample density was still relatively high. In contrast, it is claimed that the accuracy of regression modelling is not really dependent on the sampling density, but rather on how well the data are sampled and how significant the correlation is between the primary variable and secondary variable(s) (Hengl, 2007). This is largely supported by our findings as discussed below.

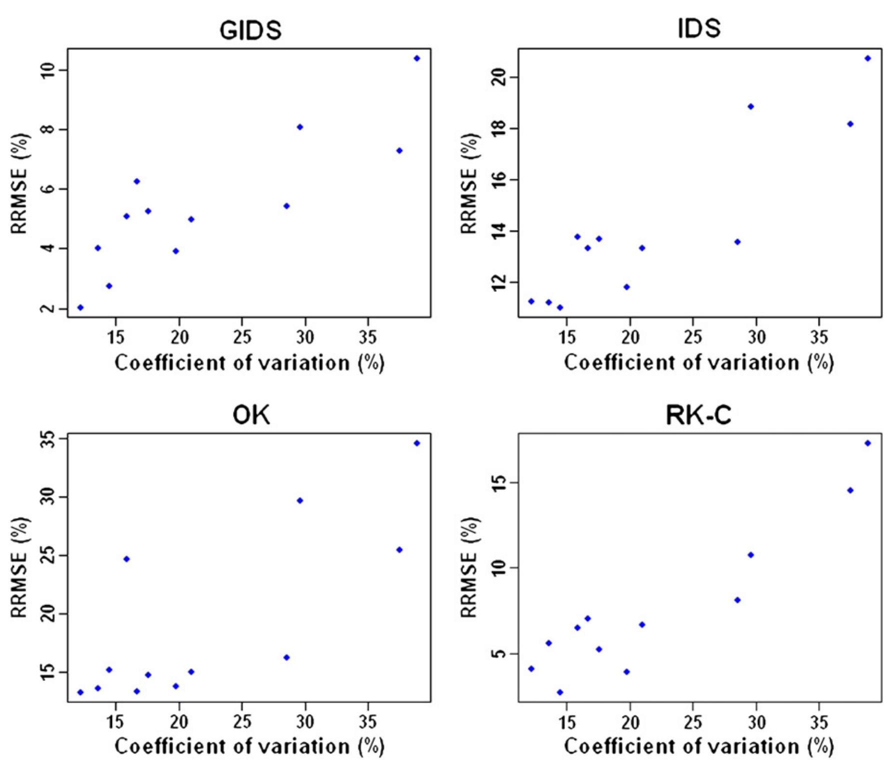

Fig. 9. Effects of the variation in the data on the accuracy of the spatial interpolation methods compared in relation to low sample density (1783 km^2 per sample) in terms of RRMSE (%).

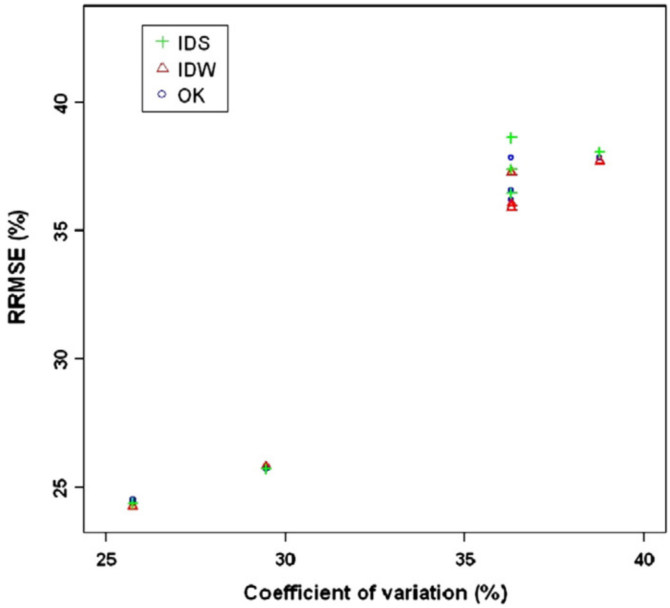

Fig. 10. Impacts of the variation in the dataset of soil nitrate on the performance of four spatial interpolation methods based on the results of Gotway et al. (1996).

The impacts of sample density on the estimation depend on the spatial interpolation methods (Bourennane et al., 2000; Wang et al., 2005; Hengl, 2007; Li et al., 2007), which is consistent with the findings in this review. This is because the underlying assumptions about the variation among samples may differ and the choice of a spatial interpolation method may become critical (Burrough and McDonnell, 1998; Hartkamp et al., 1999). Hence there is no definite conclusion on the relation of sample density and the performance of the spatial interpolators. The effects of sampling density are dominated by data variation as discussed below.

4.2.2. Data variation

There is a strong pattern of the performance of the spatial interpolation methods in relation to CV in terms of RMAE and RRMSE (Figs. 6 and 7). The observed patterns suggest that the effects of CV on the performance of the spatial interpolation methods are independent of sampling density. As CV increases, the performance declines, which is consistent with the previous findings (Collins and Bolstad, 1996; Martínez-Cob, 1996; Schläpfer and Schmid, 1999). This result is also consistent with the results in Gotway et al. (1996). Although they claimed that the performance of OK was generally unaffected by CV, it turned out that the performance of OK, like IDW, also declines as CV increases (Figs. 10 and 11). Large data variation suggests a high heterogeneity in the surface to be estimated or in the primary variable. Therefore, when the data variation is high, sample density must be increased to capture the spatial variation of the primary variable.

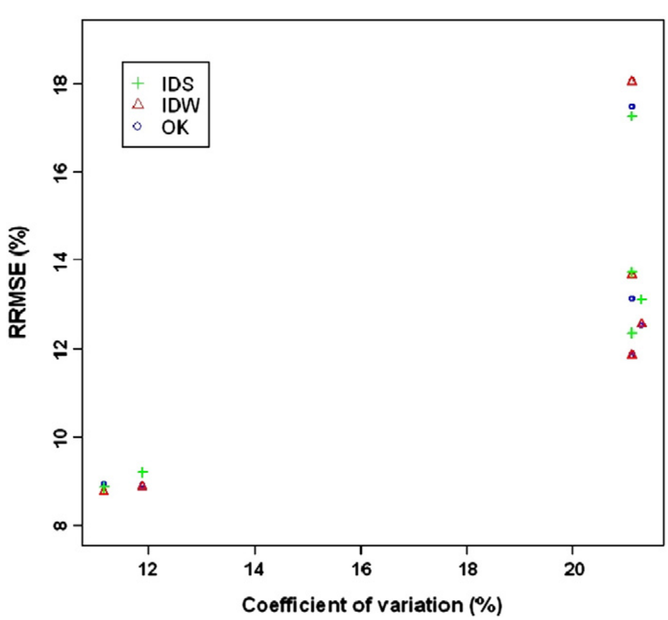

Fig. 11. Impacts of the variation in the dataset of soil organic matter on the performance of four spatial interpolation methods based on the results of Gotway et al. (1996).

The relationship between the performance of the methods and CV is largely maintained for each individual method in terms of RMAE and RRMSE (Figs. 4 and 5; see also Figs. 8 and 9). The results show that the performance of all frequently used methods is affected by CV, but the overall impact is method-dependent. GIDS and RK-C are less sensitive to CV than OK, OCK, IDS, IDW and TPS. OCK is less sensitive to CV than OK in terms of RRMSE. Such method dependency was also observed in a previous study (Lin and Chen, 2004).

The significant impact of CV observed in this review supports the view that improvements in prediction do not rely on more sophisticated methods, but rather on gathering more useful and high quality data (Minasny and McBratney, 2007). This finding also supports the argument that future studies should focus more on the quality of sampling and auxiliary environmental predictors (Hengl, 2007).

4.2.3. Sampling design

Samples collected from point locations that are irregularly distributed in space lead to a higher accuracy of the estimations of the spatial interpolators than samples collected from the regularly distributed points (Figs. 6 and 7). However, it should be noted that studies with regular design usually have higher sample density than those with irregular design, which may affect the reliability of any derived inference. Moreover, splines perform poorly for irregularly spaced data (Collins and Bolstad, 1996) and sample patterns are found not to be significant in determining the performance of the spatial interpolator (*i.e.*, OK) (Englund et al., 1992). For irregularly spaced data, the interpolated map is more variable where sample density is high than where it is low, which may result in patterns that are pure artefacts of the data configuration, and one potential solution is to use simulation algorithms instead of kriging algorithms (Goovaerts, 1997). Therefore, the irregular sampling design could improve the performance of the spatial interpolation methods, but care needs to taken in applying this finding.

A few more issues related to sampling design need further discussions. Sample clustering was found to affect the accuracy of the estimations and the effects may also depend on the spatial interpolation methods (Isaaks and Srivastava, 1989; Laslett, 1994; Zimmerman et al., 1999). Surface type may also influence the performance of the methods due to the variability in the surface (MacEachren and Davidson, 1987; Zimmerman et al., 1999). Distinct and sharp spatial changes, like changing soil types across a region, may also cause problems with the estimations (Stein et al., 1988; Voltz and Webster, 1990). The chosen sampling scheme also affects the performance of the methods through data variation; and data should be collected at a range of separations to capture changes in the scales of the variation (Laslett, 1994). Sample stratification may (Voltz and Webster, 1990) or may not (Brus et al., 1996) improve the performance of the methods. Sample stratification may avoid the effects of non-stationary, but it has two major limitations: 1) it may dramatically reduce the number of samples in the kriging neighbourhood, and 2) it depends on the goodness of the classification (Voltz and Webster, 1990). These issues need to be considered in sampling design and further research into the effects of stratification is warranted.

5. Summary

Over 70 spatial interpolation methods/sub-methods have been applied in several disciplines in environmental science. OK, IDW including IDS and OCK are the most frequently compared methods. The performance of 32 methods in 18 comparative studies has been analysed using two newly proposed measures and a procedure has been developed to compare the performance of the methods applied to multiple variables with different measuring units/scales. The impacts of sample density, CV, and sampling design on the estimations of the methods have been quantified. The performance of a spatial interpolation method depends not only on the features of the method itself, but also on other factors such as data variation and sampling design. The sensitivity of the methods to CV is method-dependent. In general, kriging methods perform better than non-geostatistical methods; of which KED is the most highly recommended method, and GIDS, OCK and RK-C were less sensitive to CV. Therefore, methods that are less sensitive to data noise should be identified and employed in application. When the data variation is high, sample density should be increased to capture the spatial changes, thus to improve the performance of the spatial interpolation methods.

Supplementary data to this article can be found online at doi: 10.1016/j.ecoinf.2010.12.003.

Acknowledgements

We would like to thank Brendan Brooke, Peter-Harris, Tomislav Hengl, Augusto Sanabria, Hideyasu Shimadzu and Xiufu Zhang for their valuable comments and suggestions. We are grateful to Maggie Tran for her careful proofreading. Tara Anderson, Shoaib Burq, David Ryan and Frederic Saint-Cast are acknowledged for providing references relevant to marine environmental science. This paper is published with permission of the Chief Executive Officer, Geoscience Australia.

References

Ahmed, S., De Marsily, G., 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resources Research 23, 1717–1737.
Bello-Pineda, J., Hernández-Stefanoni, J.L., 2007. Comparing the performance of two spatial interpolation methods for creating a digital bathymetric model of the Yucatan submerged platform. Pan-American Journal of Aquatic Sciences 2, 247–254.
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.
Boufassa, A., Armstrong, M., 1989. Comparison between different kriging estimators. Mathematical Geology 21, 331–345.
Bourennane, 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.
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, p. 333.
Collins, F.C., Bolstad, P.V., 1996. A comparison of spatial interpolation techniques in temperature estimation. Proceedings, Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, NM. National Center for Geographic Information and Analysis, Santa Barbara, Santa Barbara, CA.
Cressie, N.A., 1993. Statistics for Spatial Data. John Wiley & Sons, Inc., New York, p. 900.
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. Journal of Hydrology 208, 187–193.
Englund, E., Weber, D., Leviant, N., 1992. The effects of sampling design parameters on block selection. Mathematical Geology 24, 329–343.
Erxleben, J., Elder, K., Davis, R., 2002. Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky Mountains. Hydrological Processes 16, 3627–3649.
Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York, p. 483.
Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology 228, 113–129.
Gotway, C.A., Ferguson, R.B., Hergert, G.W., Peterson, T.A., 1996. Comparison of kriging and inverse-distance methods for mapping parameters. Soil Science Society of American Journal 60, 1237–1247.
Haberlandt, U., 2007. Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. Journal of Hydrology 332, 144–157.
Hartkamp, A.D., De Beurs, K., Stein, A., White, J.W., 1999. Interpolation Techniques for Climate Variables. CIMMYT, Mexico, D.F.
Hengl, T., 2007. A Practical Guide to Geostatistical Mapping of Environmental Variables. JRC Scientific and Technichal Reports. Office for Official Publication of the European Communities, Luxembourg.
Hengl, T., Heuvelink, G.B.M., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120, 75–93.
Hernandez-Stefanoni, J.L., Ponce-Hernandez, R., 2006. Mapping the spatial variability of plant diversity in a tropical forest: comparison of spatial interpolation methods. Environmental Monitoring and Assessment 117, 307–334.
Hosseini, E., Gallichand, J., Caron, J., 1993. Comparison of several interpolators for smoothing hydraulic conductivity data in South West Iran. American Society of Agricultural Engineers 36, 1687–1693.
Hu, K., Li, B., Lu, Y., Zhang, F., 2004. Comparison of various spatial interpolation methods for non-stationary regional soil mercury content. Environmental Science 25, 132–137.
ICES, 2005. Report of the Working Group on Marine Habitat Mapping (WGMHM), ICES CM 2005/E:05. Bremerhaven, Germany.
Isaaks, E.H., Srivastava, R.M., 1989. Applied Geostatistics. Oxford University Press, New York, p. 561.
Javis, C.H., Stuart, N., 2001. A comparison among strategies for interpolating maximum and minimum daily air temperature. Part II: the interaction between number of guiding variables and the type of interpolation method. Journal of Applied Meteorology 40, 1075–1084.
Jef, H., Clements, M., Gerwin, D., Frans, F., 2006. Spatial interpolation of ambient ozone concentrations from sparse monitoring points in Belgium. Journal of Environmental Monitoring 8, 1129–1135.
Journel, A.G., Huijbregts, C.J., 1978. Mining Geostatistics. Academic Press, London, p. 600.
Juang, K.W., Lee, D.Y., 1998. A comparison of three kriging methods using auxiliary variables in heavy-metal contaminated soils. Journal of Environmental Quality 27, 355–363.
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. Agronomy Journal 91, 393–400.
Laslett, G.M., 1994. Kriging and splines: an empirical comparison of their predictive performance in some applications. Journal of the American Statistical Association 89, 391–400.
Laslett, G.M., McBratney, A.B., 1990. Further comparison of spatial methods for predicting soil pH. Soil Science Society of America Journal 54, 1553–1558.
Laslett, G.M., McBratney, A.B., Pahl, P.J., Hutchinson, M.F., 1987. Comparison of several spatial prediction methods for soil pH. Journal of Soil Science 38, 325–341.
Li, J., Heap, A., 2008. A Review of Spatial Interpolation Methods for Environmental Scientists No. Record 2008/23. Geoscience Australia, Canberra.
Li, X., Cheng, G., Lu, L., 2005. Spatial analysis of air temperature in the Qinghai–Tibet Plateau. Arctic, Antarctic, and Alpine Research 37, 246–252.
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. Agricultural Science in China 6, 832–841.
Lin, G., Chen, L., 2004. A spatial interpolation method based on radial basis function networks incorporating a semivariogram model. Journal of Hydrology 288, 288–298.
MacEachren, A.M., Davidson, J.V., 1987. Sampling and isometric mapping of continuous geographic surfaces. The American Cartographer 14, 299–320.
Mardikis, M.G., Kalivas, D.P., Kollias, V.J., 2005. Comparison of interpolation methods for the prediction of reference evapotranspiration — an application in Greece. Water Resources Management 19, 251–278.
Martínez-Cob, A., 1996. Multivariate geostatistical analysis of evapotranspiration and precipitation in mountainous terrain. Journal of Hydrology 174, 19–35.
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.
Moyeed, R.A., Papritz, A., 2002. An empirical comparison of kriging methods for nonlinear spatial point prediction. Mathematical Geology 34, 365–386.
Nalder, I.A., Wein, R.W., 1998. Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. Agricultural and Forest Meteorology 92, 211–225.
Naoum, S., Tsanis, I.K., 2004. Ranking spatial interpolation techniques using a GIS-based DSS. Global Nest: the International Journal 6, 1–20.
Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma 63, 197–214.
Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma 67, 215–226.
Puente, C.E., Bras, R.L., 1986. Disjunctive kriging, universal kriging, or no kriging: small sample results with simulated fields. Mathematical Geology 18, 287–305.
R Development Core Team, 2007. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
Rivoirard, J., Wieland, K., 2001. Correcting for the effect of daylight in abundance estimation of juvenile haddock (*Melanogrammus aeglefinus*) in the North sea:

application of kriging with external drift. ICES Journal of Marine Science 58, 1272–1285.
Rossi, R.E., Mulla, D.J., Journel, A.G., Franz, E.H., 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecological Monographs 62, 277–314.
Ruddick, R., 2007. Data Interpolation Methods in the Geoscience Australia Seascape Maps. Geoscience Australia, Canberra.
Schläpfer, F., Schmid, B., 1999. Ecosystem effects of biodiversity: a classification of hypotheses and exploration of empirical results. Ecological Applications 9, 893–912.
Schloeder, C.A., Zimmerman, N.E., Jacobs, M.J., 2001. Comparison of methods for interpolating soil properties using limited data. Soil Science Society of American Journal 65, 470–479.
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. Agricultural and Forest Meteorology 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.
Sun, W., 1998. Comparison of a cokriging method with a Bayesian alternative. Environmetrics 9, 445–457.
Van Kuilenburg, J., De Gruijter, J.J., Marsman, B.A., Bouma, J., 1982. Accuracy of spatial interpolation between point data on soil moisture supply capacity, compared with estimates from mapping units. Geoderma 27, 311–325.
Verfaillie, E., van Lancker, V., van Meirvenne, M., 2006. Multivariate geostatistics for the predictive modelling of the surficial sand distribution in shelf seas. Continental Shelf Research 26, 2454–2468.
Vicente-Serrano, S.M., Saz-Sánchez, M.A., Cuadrat, J.M., 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. Climate Research 24, 161–180.
Voltz, M., Webster, R., 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. Journal of Soil Science 41, 473–490.
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 Geographica Sinica 60, 511–518.
Weber, D., Englund, E., 1992. Evaluation and comparison of spatial interpolators. Mathematical Geology 24, 381–391.
Webster, R., Oliver, M., 2001. Geostatistics for Environmental Scientists. John Wiley & Sons, Ltd, Chichester, p. 271.
Willmott, C.J., 1982. Some comments on the evaluation of model performance. Bulletin American Meteorological Society 63, 1309–1313.
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.
Zhou, F., Guo, H.-C., Ho, Y.-S., Wu, C.-Z., 2007. Scientometric analysis of geostatistics using multivariate methods. Scientometrics 73, 265–279.
Zimmerman, D., Pavlik, C., Ruggles, A., Armstrong, M.P., 1999. An experimental comparison of ordinary and universal kriging and inverse distance weighting. Mathematical Geology 31, 375–390.

FURTHER REFERENCES FOR THE SUPPLEMENTARY MATERIALS

Balk, B., K., Elder, 2000. Combining binary decision tree and geostatistical methods to estimate snow distribution in a mountain watershed. Water Resources research 36, 13–26.
Danielsson, A., Carman, R., Rahm, L., Aigars, J., 1998. Spatial estimation of nutrient distributions in the Gulf of Riga sediments using cokriging. Estuarine, Coastal and Shelf Science 46, 713–722.
Goff, J.A., Jenkins, C.J., Williams, S.J., 2008. Seabed mapping and characterization of sediment variability using the usSEABED data base. Continental Shelf Research 28, 614–633.
Gold, C.M., Condal, A.R., 1995. A spatial data structure integrating GIS and simulation in a marine environment. Marine Geodesy 18, 213–228.
Guarini, J.-M., Blanchard, G.F., Bacher, C., Gros, P., Riera, P., Richard, P., Gouleau, D., Galois, R., Prou, J., Sauriau, P.-G., 1998. Dynamics of spatial patterns of microphytobenthic biomass: inferences from a geostatistical analysis of two comprehensive surveys in Marennes–Oleron Bay (France). Marine Ecology Progress Series 166, 131–141.
Hutchinson, M.F., Gessler, P.E., 1994. Splines — more than just a smooth interpolator. Geoderma 62, 45–67.
Jerosch, K., Schlüter, M., Roland, P., 2006. Spatial analysis of marine categorical information using indicator kriging applied to georeferenced video mosaics of the deep-sea Håkon Mosby Mud Volcano. Ecological Informatics 1, 391–406.
Lehmann, A., 1997. A GIS approach of aquatic plant spatial heterogeneity in relation to sediment and depth gradients, Lake Geneva, Switzerland. Aquatic Botany 58, 347–361.
Lucio, P.S., Bodevan, E.C., Dupont, H.S., Ribeiro, L.V., 2006. Directional kriging: a proposal to determine sediment transport. Journal of Coastal Research 22, 1340–1348.
Maynou, F.X., Sardà, F., Conan, G.Y., 1998. Assessment of the spatial structure and biomass evaluation of *Nephrops norvegicus* (L.) populations in the northwestern Mediterranean by geostatistics. ICES Journal of Marine Science 55, 102–120.
McBratney, A.B., Mendonca Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52.
Petitgas, P., 1993. Use of a disjunctive kriging to model areas of high pelagic fish density in acoustic fisheries surveys. Aquatic Living Resources 6, 201–209.
Porter, D.E., Edwards, D., Scott, G., Jones, B., Street, W.S., 1997. Assessing the impacts of anthropogenic and physiographic influences on grass shrimp in localized salt-marsh estuaries. Aquatic Botany 58, 289–306.
Willmott, C.J., 1981. On the validation of models. Physical Geography 2, 184–194.