Journal of Applied Sciences

ISSN 1812-5654

Journal of Applied Sciences 9 (8): 1562-1567, 2009
ISSN 1812-5654

Predictive Infiltration Rate Mapping with Improved Soil and Terrain Predictors

H.R. Motaghian and J. Mohammadi
Department of Soil Science, Agriculture College, Shahrekord University, Shahrekord, Iran

Abstract: This study addresses the issue of incorporating soil and terrain covariates into predictive mapping of infiltration rate (IR) values in a semi-arid region in Iran. Besides, multiple linear regression of IR values against some soil and terrain variables, three geostatistical models including ordinary kriging, ordinary cokriging and simple kriging with varying local means were used. A 10 fold validation approach was used with the mean MAE and RMSE as validation indices to judge the prediction quality. The best prediction model was ordinary cokriging followed by simple kriging with varying local means. These methods were best in combination with some soil and terrain covariables.

Key words: Geostatistics, infiltration rate, kriging, predictive modeling

INTRODUCTION

In the context of a growing demand of high-resolution spatial information for environmental protection and management, efficient and accurate prediction models are needed. Many studies have shown that useful predictive relation exist between environmental variables and soil properties (McBratney *et al*., 2003; Scull *et al*., 2003). Environmental correlation is a method of digital soil mapping (or predictive soil mapping) that exploits these relations (Mckenzie and Ryan, 1999; Mckenzie and Gallant, 2007). In the past 20 years, a number of prediction models that use cheap and easily measured ancillary information has been developed and tested. These models vary widely from simple linear regression (Moore *et al*., 1993a) to advanced nonlinear regression methods such as regression trees (Mckenzie and Ryan, 1999) and geostatistical models such as cokriging (Ersahin, 2003), simple kriging with varying local means (Meul and Van Meirvenne, 2003; Goovaerts, 1999) or the hybrid models (McBratney *et al*., 2000) such as kriging with external drift (Hudson and Wackernagel, 1994; Bourennane *et al*., 1996; Goovaerts, 1999; Bourennane and King, 2003) and regression-kriging (Odeh *et al*., 1994; Knotters *et al*., 1995; Hengel *et al*., 2004). Wherever auxiliary variables are used; they are most commonly sourced from digital elevation models (McBratney *et al*., 2000).

Soil hydraulic properties like infiltration rate (IR) and their inter-dependencies constitute an important part of data and information needed in many *pedo-hydrological* predictions. They are the main variables controlling the key processes such as water and chemicals movement and transport in the soil profiles. Among the various soil hydraulic properties, infiltration rate (IR) is reported to have the highest variability at different spatial scales (Mallants *et al*., 1996; Tsegaye and Hill, 1998; Ersahin, 2003). Measurements of soil hydraulic properties are relatively cost and time-consuming and become impractical when hydrologic predictions are needed for large areas. Therefore, predicting soil hydraulic properties such as IR from auxiliary data becomes an alternative to measurements in many applications (Pachepsky *et al*., 2006). A close spatial relationship between soil hydraulic properties and other easily measured soil and terrain attributes, as auxiliary variables, can be exploited to predict these hydraulic characteristics, as target variables, with a reasonable accuracy at unobserved locations.

In this study, two geostatistical models (ordinary cokriging, simple kriging with varying local means) were used to incorporate secondary information (soil properties and terrain attributes derived from digital elevation model) into the spatial prediction of infiltration rate values at a catchment scale. A 10 fold jackknifing procedure was used to compare the prediction performances of these two predictions models with the straightforward ordinary kriging and multiple regression models.

MATERIALS AND METHODS

Study site, soil sampling and laboratory analysis: The study area is located in central Zagros region of the ChaharMahal Va Bakhtiari Province, Iran, about 65 km South-west of the city of Shahrekord (Fig. 1). The main

Corresponding Author: J. Mohammadi, Department of Soil Science, Agriculture College, Shahrekord University, Shahrekord, Iran
Tel/Fax: +98-381-4424228

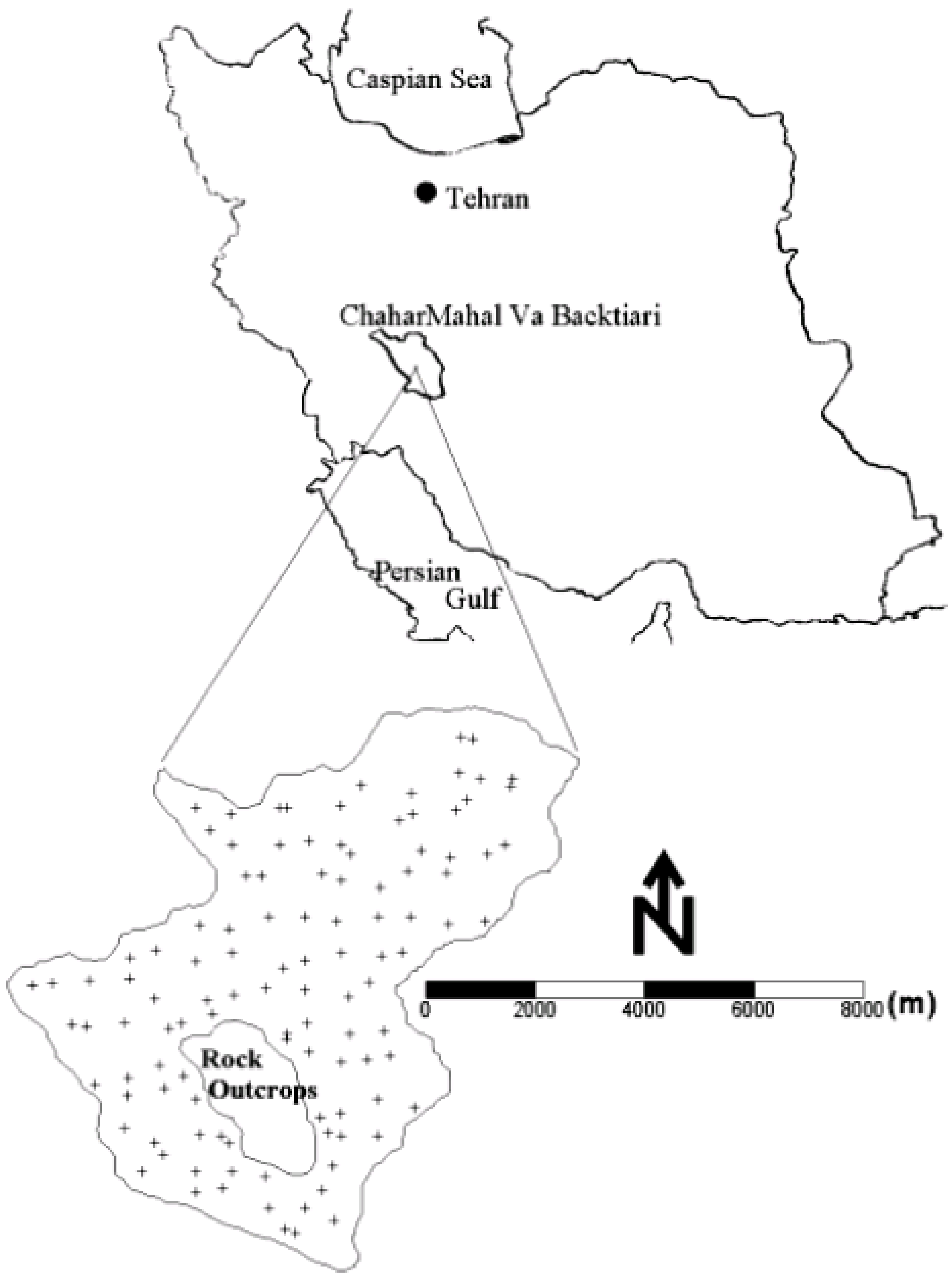

Fig. 1: Study area with sampling locations

landform features within this 92 km^2 catchment include plateaus, upland terraces, gently rolling hills and alluvial plains surrounded by mountains. Elevation ranges from 2393 to 2944 m above sea level with slope gradients approximately 5 to 30%. The catchment is predominately underlain by quaternary deposits composed of limestone, sandstone and conglomerate. The primary land uses within the catchment include degraded rangelands and dry farming. Irrigated farming is conducted in a very small portion of the catchment.

The study area was sampled on a pseudo-regular initial grid spacing of 1 km during June and July 2007. Sampling was performed at different spatial scales resulting in the minimum distance between samples of about 150 m. In total, 111 sample points were considered for different soil analysis. All 111 soil samples were analyzed for sand, silt and clay contents by the hydrometer method (Gee and Bauder, 1986). In addition, at each sampling site, undisturbed soil samples were obtained from the topsoil using 200 cm^3 steel cores to determine soil bulk density. Infiltration tests were conducted on all 111 sample sites using double-ring variable-water level infiltrometers. The internal diameter was 30 cm for inner and 45 cm for outer ring.

Slope gradient, slope aspect, contour curvature and profile curvature were derived from a 100 m resolution digital elevation model in accordance with the procedure of Moore *et al*. (1993b).

Prediction models: A brief description of the prediction models used is given below:

Statistical linear regression: In classical least-squares regression technique, regression model was created by relating the target variable (IR) to the soil and terrain covariates. This model then used to predict the target variables to testing sample locations where all variables and co-variables have been determined.

Ordinary kriging: Ordinary kriging (OK) is one of the most popular and basic univariate kriging methods with extensive applications in soil science. The basic idea is to estimate the target variable at unsampled location as a linear combination of neighboring observations. It relies on a weighting scheme where closer observations have greater impact on the final prediction. The weighting scheme is dictated by the variogram. The weights are determined such as to minimize the estimation variance, while ensuring the unbiasedness of the predictor (Goovaerts, 1997).

Ordinary cokriging: Cokriging (COK) is a multivariate extension of OK in which the predictor is calculated by using simultaneously the interrelationships (auto-correlation) between the target data values and spatial co-dependencies (cross-correlation) between target and co-variables. In the early cokriging studies (Goulard and Voltz, 1992), the ancillary variables were other soil variables, indicating that other soil variables are themselves useful predictors of the target variable. Later, this geostatistical multivariate algorithm was performed with detailed ancillary information of environmental variables derived from digital elevation models and satellite images (Odeh *et al*., 1994; Goovaerts, 1999).

Simple kriging with varying local means: In simple kriging with varying local means (SKLM), the unknown local stationary mean from the OK prediction is replaced by known varying local means derived from ancillary information (Goovaerts, 1997). Local means can be derived from the secondary information using a linear first order relationship. If the local means are derived using a linear statistical regression model, the SKLM estimate can be considered as the sum of the regression estimate and the simple kriging estimate of the residual value at unsampled location (Goovaerts, 1999).

In the case of hybrid models, ordinary cokriging and simple kriging with varying local means were performed using covariates that had significant correlation coefficient (r) with IR ($p<0.05$).

Validation of prediction performance: In comparing the performance of different methods, a modified jackknifing method was used (Bishop and McBratney, 2001). The procedure involves randomly splitting the data into two, the prediction and validation subsets. The prediction set is used to create a model, which is then used to predict onto the validation sites, thus providing an independent assessment of the prediction performance. Various validation indices can be used as a measure of prediction quality, the most common of which are root-mean-square error (RMSE) and mean absolute error (MAE). The latter yields a more balanced perspective of the goodness of fit at moderate IR, whereas the RMSE measures the goodness of fit relevant to high IR predicted values (Kisi and Ozturk, 2007). The RMSE and MAE can be used together to diagnose the variation in the errors in a set of predictions. The RMSE will always be larger or equal to the MAE. The greater difference between them indicates the greater variance in the individual errors in the sample. If the RMSE equals MAE, then all the errors are of the same magnitude.

To overcome the problems arising with single jackknifing method when researchers faced with a small sample size, a multiple jackknifing approach (10 fold jackknifing) was adopted (Bishop and McBratney, 2001). This involved random selection of both the prediction and the validation sets a number of times; in this case 10 times. Each time, the sample size is 86 for the prediction set and 25 for the validation test. For each of the 10 random selections, all of the different predictions models were created which were used to predict onto the validation set. Therefore, 10 realisations of RMSE and MAE were obtained for each of the prediction models. Finally, the mean RMSE and MAE values for each of the prediction models were presented.

RESULTS AND DISCUSSION

Statistical characterization of soil properties: The global mean for IR was 4.31 cm h^{-1} and the median was 3.40 cm h^{-1}. IR values ranged from 0.37 to 14.12 cm h^{-1}, with a CV of 72% (Table 1). Comparing with earlier study, variation in IR at the regional scale is large. Ersahin (2003) found a mean value of 5.11 cm h^{-1} and a CV 37% for 50 field-measured IRs.

Table 1: Descriptive statistics of infiltration rate (cm h^{-1}) for the surface soil

Number of sample points	111
Mean	4.31
Median	3.40
Minimum	0.37
Maximum	14.12
Standard deviation (SD)	3.12
Coefficient of variation (CV)%	72.30

Table 2: Simple pearson correlation coefficients (r) between field-measured infiltration rate (IR) values and soil and terrain attributes

Parameters	IR	Bulk density	Clay	Silt	Sand	Elevation	Slope gradient	Slope aspect	Profile curvature	Contour curvature
IR	1.00	0.23	-0.24	-0.10	0.22	0.21	-0.07	0.09	0.17	0.06
Bulk density		1.00	-0.42	-0.31	0.51	-0.18	-0.14	0.29	0.05	-0.01
Clay			1.00	-0.01	-0.62	0.11	0.25	-0.22	0.01	-0.14
Silt				1.00	-0.79	0.01	0.17	-0.18	-0.06	-0.07
Sand					1.00	-0.10	-0.29	0.27	0.13	0.05
Elevation						1.00	0.15	-0.10	0.31	0.26
Slope gradient							1.00	-0.36	-0.28	-0.12
Slope aspect								1.00	-0.04	-0.17
Profile curvature									1.00	0.63
Contour curvature										1.00

Statistical relationships between infiltration rate values and soil and terrain attributes: IR showed some correlation with clay, sand and bulk density ($p<0.05$). These correlations, though not particularly strong, may be used to enhance the prediction performance of different models. No significant relationships ($p<0.05$) between IR values and terrain attributes, except elevation and profile curvature, were found (Table 2). Spatial resolution of the digital elevation model can significantly affect the terrain attributes. Thus, the correlation between soil and terrain depends on the resolution of interest. For resolution of <100 m, as described by McBratney *et al.* (2003), local terrain attributes such as slope gradient, slope aspect and curvature, are found to be good predictors of soil variability.

Comparison of model performance: OK, Which uses only primary data (IR values), was considered as a reference to assess the actual gain of accounting for environmental covariates (Table 3). Overall, the best prediction methods were COK followed closely by SKLM. Multiple linear regression performed worst of all. Goovaerts (1999) compared linear regression and ordinary kriging with simple kriging with varying local means and cokriging for rainfall erosivity mapping purposes, favouring the latter. Bourennane *et al*. (1996) showed that prediction of horizon thickness is more accurate with the use of a slope map as external drift.

Whereas, the SKLM approach use covariables only to inform on the local mean of infiltration rate, cokriging incorporates the secondary variables directly into the computation of the estimate (Goovaerts, 1997). Cokriging may seem more demanding in that several variograms must be inferred and jointly modeled. However, the results presented here show that this complexity pays off for this data set. Maps of IR created by models indicate that all kriging- based maps clearly reflect the similar global and local spatial distribution of IR values over the study area. However, comparing these maps with ones generated by statistical linear regression indicates that more details of spatial variability of IR values were revealed by maps created by kriging- based models (Fig. 2).

Table 3: Mean MAE and RMSE of different prediction models

Prediction models	Mean MAE	Mean RMSE
REG	2.45	4.23
OK	1.83	3.34
COK	0.83	1.78
SKLM	1.25	2.87

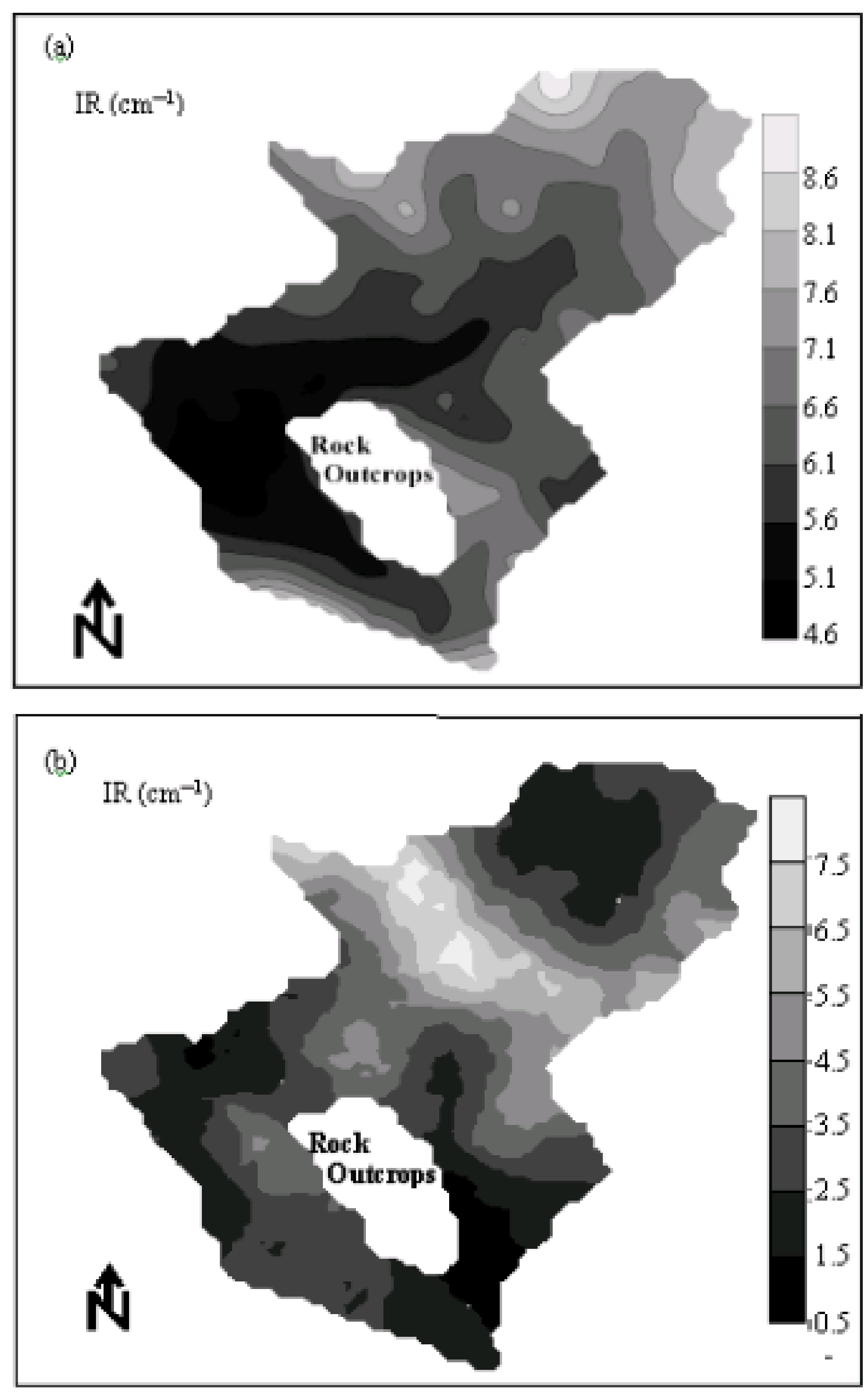

Fig. 2: Map of infiltration rate (cm h^{-1}) created by (a) statistical linear regression model and (b) ordinary cokriging model

CONCLUSIONS

Due to high cost and time-consuming nature of soil hydraulic properties, research in developing models for the creation of digital map of soil properties from sparse soil data using secondary variables is becoming increasingly important. Recently, sources of ancillary information are increasingly available from digital terrain modeling parameters to various air-and space-born remote sensing images. Under these circumstances, environmental correlation is now feasible for predictive soil mapping across large areas. Hence, the classical statistical models and generic geostatistical techniques such as ordinary kriging are likely to be replaced with hybrid models. The results here indicate that even when only the poorly correlated ancillary variables are available, the hybrid models may still perform better than the classical regression models and generic geostatistical models, such as ordinary kriging that use only the target variable.

Measuring soil hydraulic properties like saturated hydraulic conductivity in the field is expensive and time consuming. Therefore, prediction of these properties with a reasonable accuracy is crucial. The results presented here show potential to improve model predictions by using hybrid models such as ordinary cokriging and simple kriging with varying local means.

Finally, this study showed that multiple jackknifing would be more suitable validation approach for situations where the sample data set is not sufficiently large. In this study, only 10 realisations of random selection of both prediction and validation set, with ratio of 3.4, were created. Further research is needed to find the optimum number of realisations and ratio for robust estimation of validation indices.

REFERENCES

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

Bourennane, H., D. King, P. Chery and A. Bruand, 1996. Improving the kriging of a soil variable using slope gradient as external drift. Eur. J. Soil Sci., 47: 473-483.

Bourennane, H. and D. King, 2003. Using multiple externals drifts to estimate a soil variable. Geoderma, 114: 1-18.

Ersahin, S., 2003. Comparing ordinary kriging and cokriging to estimate infiltration rate. Soil Sci. Soc. Am. J., 67: 1848-1855.

Gee, G.W. and J.W. Bauder, 1986. Particle Size Analysis Methods of Soil Analysis. ASA and SSSA, Madison, WI, pp: 404-407.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. 1st Edn., Oxford University Press, London.

Goovaerts, P., 1999. Using elevation to aid the geostatistical mapping of rainfall erosivity. Catena, 34: 227-242.

Goulard, M. and M. Voltz, 1992. Linear coregionalization model: Tools for estimation and choice of cross-variogram matrix. Math. Geol., 24: 269-286.

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

Hudson, G. and H. Wackernagel, 1994. Mapping temperature using kriging with external drift: Theory and an example from Scotland. Int. J. Climatol., 14: 77-91.

Kisi, O. and O. Ozturk, 2007. Adaptive neurofuzzy computing technique for evapotranspiration estimation. J. Irrigation Drain Eng., 133: 368-379.

Knotters, M., D.J. Brus and J.H. Oude Voshaar, 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.

Mallants, D., P.M. Binayak, J. Diederik and J. Feyen, 1996. Spatial variability of hydraulic properties in a multi-layered soil profile. Soil Sci., 161: 167-181.

McBratney, A.B., I.O.A. Odeh, T.F.A. Bishop, M.S. Dunbar and T.M. Shatar, 2000. An overview of pedometric techniques for use in soil survey. Geoderma, 97: 293-327.

McBratney, A.B., M.L. Mendonça-Santos and B. Minasny, 2003. On digital soil mapping. Geoderma, 117: 3-52.

Mckenzie, N.J. and P.J. Ryan, 1999. Spatial prediction of soil properties using environmental correlation. Geoderma, 89: 67-94.

Mckenzie, N.J. and J.C. Gallant, 2007. Digital Soil Mapping: An Introductory Perspective. In: Developments in Soil Science, Lagacherie, P., A.B. McBratney and M. Voltz, (Eds.). Vol. 31. Elsevier B.V., The Netherlands, pp: 327-349.

Meul, M. and M. van Meirvenne, 2003. Kriging soil texture under different types on nonstationarity. Geoderma, 112: 217-233.

Moore, I.D., P.E. Gessler, G.A. Nielsen and G.A. Peterson, 1993a. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J., 57: 443-452.

Moore, I.D., A. Lewis and J.C. Gallant, 1993b. Terrain Attributes: Estimation Methods and Scale Effects. In: Modeling Change in Environmental Systems, Jakeman, A.J. (Ed.). John Wiley and Sons, New York.

Odeh, I.O.A., A.B. McBratney and D.J. Chittleborough, 1994. Spatial prediction of soil properties from landform attributes derived from a digital elevation model. Geoderma, 63: 197-215.

Pachepsky, Y.A., W.J. Rawels and H.S. Lin, 2006. Hydropedology and pedotransfer functions. Geoderma, 131: 308-316.

Scull, P., J. Franklin, O.A. Chadwick and D. McArthur, 2003. Predictive soil mapping: A review. Prog. Phys. Geogr., 27: 171-197.

Tsegaye, T. and R.L. Hill, 1998. Intensive tillage effects on spatial variability of soil physical properties. Soil Sci., 16: 143-154.