ELSEVIER

Geoderma 103 (2001) 149–160

GEODERMA

www.elsevier.com/locate/geoderma

A comparison of prediction methods for the creation of field-extent soil property maps

T.F.A. Bishop *, A.B. McBratney

Department of Agricultural Chemistry and Soil Science, University of Sydney, Ross St., Building A03, Sydney, NSW 2006, Australia

Received 12 February 2000; received in revised form 8 September 2000; accepted 9 February 2001

Abstract

We compare various prediction methods for mapping of soil cation exchange capacity using different combinations of secondary information. The prediction methods used are statistical analysis (generalised additive model, regression tree, multiple linear regression), geostatistical interpolation (ordinary kriging) and the hybrid techniques (regression-kriging and kriging with external drift). The secondary spatial information used are terrain attributes, bare soil colour aerial photograph, bare soil LANDSAT TM imagery, crop yield data and soil apparent electrical conductivity (EC_a). A modification of jackknifing was used as the validation method. This involved 100 jackknife partitions to examine the stability of the validation indices with different realisations of the data set. Root-mean-square error (RMSE) was used as the validation index, with the mean RMSE used to judge the prediction quality. The best prediction methods were kriging with external drift, multiple linear regression and generalised additive models. They were best in combination with soil EC_a or the bare soil colour aerial photograph.

Keywords: Jackknife; Soil prediction; Regression; Geostatistical; Precision agriculture

1. Introduction

Due to high cost and time-consuming nature of soil sampling, research in developing methods for the creation of soil maps from sparse soil data is becoming increasingly important. In the past 20 years, the development of

* Corresponding author. Fax: +61-2351-5108.
E-mail address: t.bishop@agec.usyd.edu.au (T.F.A. Bishop).

0016-7061/01/$ - see front matter
PII: S0016-7061(01)00074-X

prediction methods that use cheap secondary information to spatially extend sparse and expensive soil measurements has been a sharpening focus of research (e.g. Odeh et al., 1994; Gessler et al., 1995). Existing prediction methods vary widely from simple linear regression (Moore et al., 1993) to advanced nonlinear regression methods, e.g. regression trees (McKenzie and Ryan, 1997), and geostatistical methods such as co-kriging (Odeh et al., 1995) or the hybrid methods (McBratney et al., 2000) such as kriging with external drift (Goovaerts, 1998). Wherever secondary attributes are used, they are most commonly sourced from the digital elevation models and remote sensing.

A limitation to adopting most of the methods above at the within-field extent has been due to the lack of high-resolution secondary information and a lack of interest in quantifying within-field variability. Precision agriculture (PA) has provoked an increasing interest in the management of within-field soil variability, and by nature, involves the collection of high-resolution secondary information. Examples of secondary information include remote sensing imagery, elevation data and crop yield data. Furthermore, a number of proximal soil sensors are becoming more available; examples are the Soil Doctor (Colburn, 1999) and the VERIS conductivity cart (Lund et al., 1999). Consequently, the potential for using the secondary information to aid soil mapping at the within-field extent is greater than ever before.

The aim of this paper is to present a modification of the jackknife validation method and use it to:

- compare different prediction methods for soil prediction,
- compares the usefulness for soil prediction of different types of secondary information,
- choose the best combination of prediction method and secondary information.

2. Methods

2.1. The data set

The study site is a 74-ha field called ‘East Creek’ (29°42’36.3 S, 150°10’36.1 E) on a dryland grain cropping farm, located 25 km east of Moree in northern New South Wales, Australia. As a result of the ongoing research activities in PA, the following types of secondary information have been obtained:

- bare soil colour aerial photograph taken on the 29th April 1997
  - raw reflectance in the red, blue and green portion of the electromagnetic spectrum at 1-m grain size;

- bare soil LANDSAT TM image taken on the 31st May 1997
  - raw reflectance in the blue$^{450-520\ \mu m}$, green$^{520-600\ \mu m}$, red$^{630-690\ \mu m}$, NIR$^{760-900\ \mu m}$ and MIR$^{1550-1750\ and\ 2080-2350\ \mu m}$ portion of the electromagnetic spectrum at 25-m grain size;
- crop yield data (1996 sorghum and 1997 wheat) using an Ag Leader yield monitor
  - point measurements at a sampling density of 1407 observations/ha for the 1996 sorghum crop and 517 observations/ha for the 1997 wheat crop;
- soil apparent electrical conductivity (EC_a) data collected on March 1st 1999 using a VERIS 3100 conductivity cart (Lund et al., 1999)
  - point measurements at three depths: 0–30, 0–90 and 30–90 cm, at a sampling density of 54 observations/ha;
- elevation data collected on December 17th 1999 using an Ashtech GG-12 GPS receiver
  - point measurements at a sampling density of 316 observations/ha.

Different combinations of the wide range of the secondary information were used in the comparison of the prediction methods. The point elevation data were rasterised to a 5-m pixel digital elevation model (DEM) using the TOPOGRID tool in ArcInfo, an earlier version of ANUDEM (Hutchinson, 1989). The DEM was then used to derive slope, plan curvature, profile curvature, upslope area and compound topographic index using the GRID module of ArcInfo (ESRI, 1997). The EC_a and yield data were interpolated onto the same grid using kriging with local variograms (Minasny et al., 1999). The reflectance values for the LANDSAT TM and aerial photography were found at each grid node by spatial coincidence. Thus, all of the secondary information was available on a common 5-m grid.

Additionally in March 1999, 113 soil cores, each to a depth of 90 cm, were taken from the field. For the purposes of this study, the CEC of the 0–15-cm layer was used as the target variable. To find secondary information values at each sampling location, the values at the nearest node of the 5-m grid were used.

2.2. Prediction methods

A brief description of the prediction methods used is given below.

2.2.1. Statistical methods

Multiple linear regression (MLR): the classical least-squares regression technique.

Generalised additive model (GAM): a regression method where non-linearity is introduced by splitting the range of variables and fitting piece-wise constant functions (Chambers and Hastie, 1992). In the case presented here, cubic smoothing splines were used.

Regression tree (RT): a hierarchical method that splits the data into nodes in a binary fashion until the node is too homogenous or there are too few observations. Splitting is optimised by minimising residual deviance (Venables and Ripley, 1994). An example of its application for mapping soil is presented by McKenzie and Ryan (1997).

2.2.2. Geostatistical method

Ordinary kriging (OK): a univariate interpolation method with extensive applications in soil science (e.g. Burgess and Webster, 1980). Relies on a weighting scheme where closer sample locations have greater impact on the final prediction. The weighting scheme is dictated by the variogram.

2.2.3. Hybrid methods

Regression-kriging: this method involves the ordinary kriging of the residuals of the generic regression techniques (MLR-k, GAM-k, RTR-k). The kriged residual values are then added to the regression predictions to provide an improved prediction. An example where this method has been applied is by Odeh et al. (1995).

Kriging with external drift (KED): an interpolation method that incorporates secondary information into the kriging system (Goovaerts, 1998). It uses the secondary information to find the local means of the target variable and performs simple kriging on the residuals (Goovearts, 2000). A linear relationship is assumed to exist between the secondary variable and the target variable.

It is not the aim of this paper to carry out an exhaustive comparison of the prediction methods. Rather, the aim is to present a comparison of a number of older techniques, e.g. MLR, OK, with the more recent innovations, e.g. KED, GAM, RT.

2.3. Prediction models and their validation

For the purpose of comparison, each prediction method was used in combination with each class of secondary information. The classes were crop yield data, EC_a data, terrain attributes, aerial photo bands and LANDSAT TM bands. The Akaike Information Criterion (AIC) (Akaike, 1973) was used to decide upon the most parsimonious number of predictor variables for the MLR and GAM prediction methods. The cross-validation method described by Venables and Ripley (1994) was used to optimise the size of the regression tree. In the case of the hybrid techniques, regression-kriging was only performed when spatial autocorrelation was evident in the residuals. KED was performed using the secondary variable from each class of secondary information (Table 1 and Fig. 1) that had the highest correlation coefficient (r) with CEC. They were: EC_a: 0–30 cm ($r = 0.74$), crop yield: wheat 1997 ($r = 0.21$), digital terrain data:

Table 1
Correlations of secondary information with CEC (0–15 cm)

Secondary information	Correlation coefficient
EC_a 0–30 cm	0.74
EC_a 0–90 cm	0.51
EC_a 30–90 cm	0.38
Sorghum yield 1996	0.15
Wheat yield 1997	0.21
Aerial photo blue band	−0.22
Aerial photo red band	−0.61
Aerial photo green band	−0.47
LANDSAT blue band	−0.33
LANDSAT green band	−0.38
LANDSAT red band	−0.40
LANDSAT NIR band	−0.32
LANDSAT MIR1 band	−0.43
LANDSAT MIR2 band	−0.45
Elevation	0.38
Slope	−0.15
Profile curvature	0.01
Plan curvature	0.00
Compound topographic index	−0.08
Upslope area	−0.09

elevation ($r = 0.37$), aerial photograph: red band ($r = -0.61$), LANDSAT TM: mid IR ($r = -0.45$).

A modified jackknifing method was used for the validation of the methods used here. 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 quality or uncertainty. Various validation indices can be used as a measure of prediction quality, the most common of which are the root-mean-square error (RMSE) and mean error. Some obvious problems arise with the jackknifing method when researchers are faced with a very small sample size. In such situations, not enough data are available for accurate variogram estimation, which is necessary for geostatistical or hybrid methods. Webster and Oliver (1992) indicated that 100 points is a bare minimum required for variogram estimation. Similarly, the sample size may not be large enough for independent use in the prediction and validation sets. In validation sets with a small sample size, outliers can easily influence the value of validation indices. An example is a situation where the validation locations are in clusters or are located near the boundary of the area of interest. Usually in such a situation, a single jackknifing does not give a true indication of the prediction.

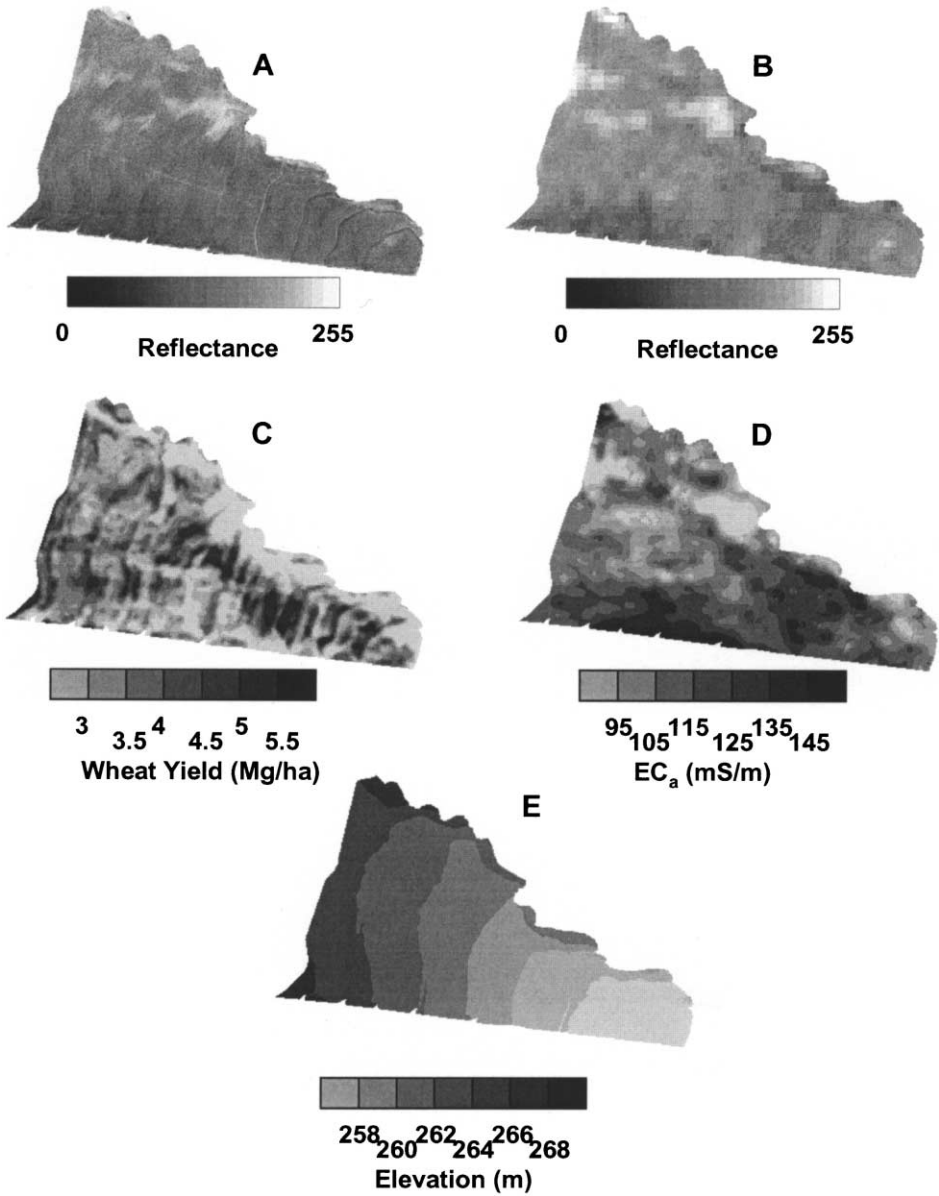

Fig. 1. Maps of information layer with highest correlation coefficient for each class of secondary information: (A) red band of colour aerial photograph; (B) mid IR band of LANDSAT TM; (C) 1997 wheat yield; (D) Veris 0–30 cm; (E) elevation.

In this study the sample size is only 113. To overcome the problems as described above, a multiple jackknifing approach was adopted. This involved random selection of both the prediction and the validation sets a number of times; in this case 100 times. The sample size each time is 95 for the prediction

set and 18 for the validation set. For each of the 100 samplings, all of the different prediction models were created, including variogram estimation, fitting regression models, etc, which were used to predict onto the validation sites. Using this procedure, RMSE was calculated for each validation set.

$$\mathrm{RMSE} = \sqrt{\left[\frac{1}{n}\sum_{i=1}^{n}\left(z(x_i) - z^*(x_i)\right)2\right]} \quad (1)$$

where $z(x_i)$ = actual CEC content, and $z^*(x_i)$ = predicted CEC content.

Therefore, 100 realisations of RMSE were obtained for each of the prediction models.

3. Results

3.1. Robustness of validation methods

The histograms of RMSE are shown in Fig. 2. Generally, the histograms exhibited normal distributions. The histograms also indicate the dispersion of RMSE around the mean, which is true for all the prediction models. This is indicative of instability of the validation indices, and thus proves that a single jackknifing may not provide the right answer in determining the best prediction

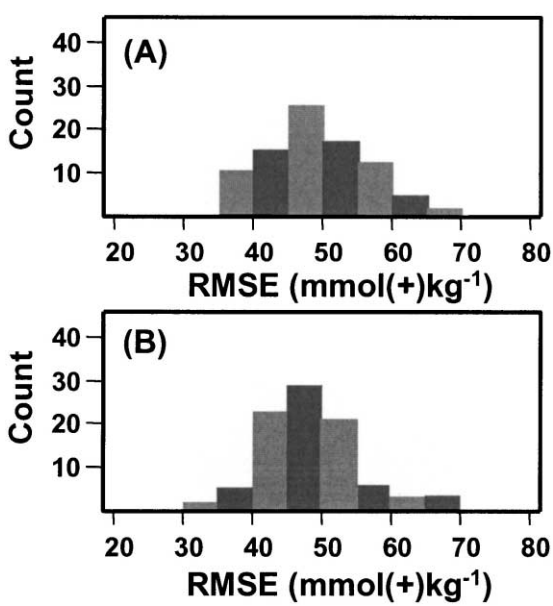

Fig. 2. Histograms of RMSE for (A) MLR–Veris and (B) GAM–Veris combinations.

method, especially in situations where the validation set involves a low number of sample size. Therefore, the use of the multiple jackknifing seemed appropriate.

In deciding the best prediction model, we used the mean value of RMSE as it gives an overall indication of the prediction quality. The mean RMSE values for each of the prediction models used here are presented in Table 2.

3.2. The best prediction methods

The comparative performance of the prediction models is gauged by using the RMSE of OK as the standard, as OK does not take into account the secondary information. Overall, the best prediction methods were KED followed closely by MLR-k. Both of these methods outperformed OK when used with any of the classes of secondary information. All of the regression methods (without kriging of residuals) were inferior to ordinary kriging except when the highly correlated EC_a variable was used. Regression trees performed worst of all, while GAMs were only slightly inferior to MLR and KED. The performance of all of the regression methods was improved when model residuals were incorporated by regression-kriging. The fact that MLR outperformed the GAMs is contrary to previously published work (Odeh et al., 1997) as it would be expected that more advanced regression methods would outperform simpler methods such as MLR. Reasons for this unusual result could be study-specific or over fitting of the GAM even though stepwise procedures were used to achieve parsimony.

Table 2
Mean RMSE of different prediction models*

Secondary information	Number of secondary variables	MLR	RT[a]	GAM	KED	Mean[b]
Air photo	3	5.71/5.20	5.90	5.79/5.43	5.31	5.56
EC_a	3	4.90/4.66	5.10	4.85/4.75	4.62	4.81
LANDSAT TM	6	6.48/5.53	7.27	6.94/6.18	5.41	6.31
Terrain data	6	6.49/5.60	7.09	6.86/6.13	5.64	6.32
Yield data	2	6.97/5.60	7.48	7.23/5.89	5.49	6.45
Mean[c]		6.11/5.32	6.57	6.32/5.65	5.29	
Ordinary kriging[d]		5.64				

* Numbers to the right of the forward slash indicate the RMSE for regression-kriging if performed. MLR = multiple linear regression; RT = regression tree; GAM = generalised additive model; KED = kriging with external drift.

[a] No spatial autocorrelation was evident in the residuals of the RT models, so regression-kriging was not performed.

[b] Mean of each type of secondary information across all prediction methods.

[c] Mean of each different prediction method across all types of secondary information.

[d] Ordinary kriging results were not include in calculations of means (see footnotes b and c).

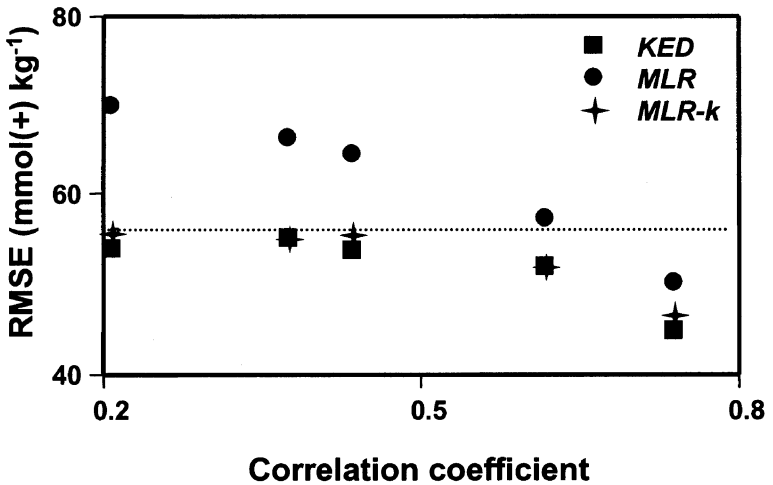

Fig. 3. The relationship between RMSE and correlation coefficient.[1]

3.3. The best combinations of secondary information for predicting topsoil CEC

When considering the quality of prediction provided by each of the classes of secondary information, it can be seen that the larger the correlation of the ancillary attributes with CEC (Table 1), the smaller the RMSE (Table 2). This is illustrated in Fig. 3 where the mean RMSE is plotted against correlation coefficient for KED, MLR and MLR-k in combination with (from left to right) 1997 wheat yield, elevation, LANDSAT MIR $^{2.08-2.35\ nm}$ band, red band of the colour aerial photograph and EC_a 0–30 cm. As can be seen, the greater the correlation between CEC and the secondary variable, the smaller the RMSE.

In general terms, the EC_a measurements clearly provide the best predictions followed by the aerial photographic bands (Table 2). The LANDSAT TM bands produced worse predictions in comparison to the aerial photograph bands due to their inferior spatial resolution (1 vs. 25 m) (Fig. 1B). The elevation data was reasonably well correlated with CEC but the other terrain attributes exhibit poor correlation with CEC. Yield data, when used as covariates in the hybrid methods (e.g. KED and MLR-k), give results with the largest RMSE compared with the results produced using the other secondary information. However, the yield data results are still superior to the predicted CEC by ordinary kriging (Table 2).

EC_a is a function of moisture, structure, mineralogy and texture (Rhoades, 1992). Thus, EC_a as measured, provides a good indication of the variability in these properties. Furthermore, the VERIS measures within the profile, giving an

[1] The dashed line represents the mean RMSE of ordinary kriging.

indication of variability with depth. CEC is commonly related to soil texture and mineralogy, and any secondary information that depicts this variation will be a good predictor of CEC, i.e. EC_a. The field in this study had contrasting soil types in terms of colour and texture: coarse-textured Red Chromosols (Isbell, 1996) and heavier Grey and Brown Vertosols (Isbell, 1996). Therefore, the bare soil remote sensed imagery also gives a reasonable indication of CEC variability due to textural differences in the soil types (Fig. 1A and B). The imagery suffered since it only provided an indication of the top few millimeters of soil, which in some cases do not truly reflect the top 15 cm from which the soil samples were taken. The LANDSAT TM imagery, while providing a greater range of bands, was not a good predictor possibly due to its much coarser spatial resolution (25 m) (Fig. 1B).

Variation in elevation and slope reflects a reasonable variability in soil due to erosion along the north boundary of the field, as the elevation decreases sharply (Fig. 1E) and as the slope increases. The eroded section is characterized by coarse-textured soil, which is picked up by the changes in elevation and slope. There was little relationship between other terrain attributes and CEC variability due to the flatness of the majority of the field. Overall, the best prediction model was a combination of KED with EC_a (0–30 cm) as the secondary variable (Fig. 4).

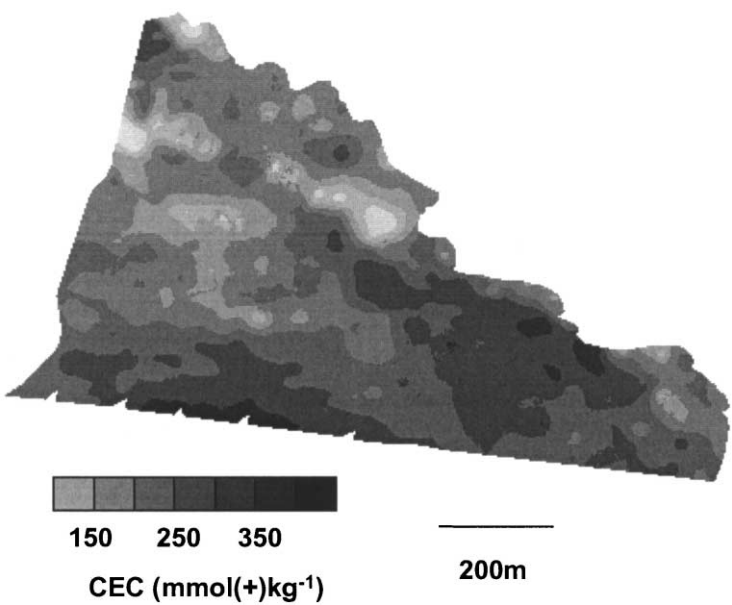

Fig. 4. Map of CEC created using kriging with external drift, with EC_a 0–30 cm as the secondary variable.

4. Conclusions

The results here indicate that when secondary information is available, it should be used to model the deterministic trend in the variation of a soil attribute, using the hybrid methods such as regression with kriging of the residuals or kriging with external drift. The generic geostatistical techniques, such as ordinary kriging that use only the target soil variable, cannot match the prediction performance of the methods that incorporate densely measured, correlated secondary information. The generic geostatistical methods are not designed to do so. Even when only the poorly correlated secondary attributes are available, e.g. crop yield, the hybrid methods may still perform better than the generic geostatistical method. The growing prominence of PA in Australia, Europe and USA has resulted in an increasing availability of yield data, which as shown by the results in this paper, can be used to improve the quality of soil maps in comparison with maps produced by univariate geostatistical methods alone. The other types of secondary information used in this paper, while being less common, are vital data inputs for a PA management programme, and in the near future will be more commonly available. Therefore, increasingly, a variety of secondary information will be available to aid in the production of within-field soil maps. The better prediction methods were KED, MLR-k and GAM-k, any of which have been proved to be useful in conjunction with EC_a as the secondary variable, to more accurately predict CEC than would be possible with univariate methods such as ordinary kriging.

Finally, this study shows that the multiple jacknife method is more suitable for situations where the sample size is not sufficiently large to provide for validation. In this study, the ratio of the prediction sample size to validation sample size is 5.3. Further research is needed to:

(i) examine the robustness of single jackknife divisions with different ratios,
(ii) develop guidelines for the minimum ratio necessary for the robust estimation of validation indices from single jackknife divisions,
(iii) examine the possibility of using other validation indices for the robust measure of prediction quality using single or multiple jackknifing.

References

Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle. In: Petrov, B.N., Csaki, F. (Eds.), Second International Symposium on Information Theory. Akademia Kiado, Budapest, Hungary, pp. 267–281.

Burgess, T.M., Webster, R., 1980. Optimal interpolation and isarithmic mapping of soil properties: I. The semivariogram and punctual kriging. Journal of Soil Science 31, 315–331.

Chambers, J.M., Hastie, T.J., 1992. Statistical Models. S. Wadsworth and Brooks, CA, p. 608.
Colburn, J.W., 1999. Soil Doctor multi-parameter, real time soil sensor and concurrent input control system. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), Proceedings of the Fourth International Conference on Precision Agriculture. ASA-CSSA-SSSA, Madison, WI, pp. 1011–1022.
ESRI, 1997. Understand GIS: The Arc/Info Method. Environmental Systems Research Institute, Redlands, CA, p. 302.
Gessler, P.E., Moore, I.D., McKenzie, N.J., Ryan, P.J., 1995. Soil landscape modelling and spatial prediction of soil attributes. International Journal of Geographical Information Systems 9, 421–432.
Goovaerts, P., 1998. Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties. Biology and Fertility of Soils 27, 315–334.
Goovearts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology 228, 113–129.
Hutchinson, M.F., 1989. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. Journal of Hydrology 106, 211–232.
Isbell, R.F., 1996. The Australian Soil Classification. CSIRO, Melbourne, Australia, p. 205.
Lund, E.D., Colin, P.E., Christy, D., Drummond, P.E., 1999. Applying soil conductivity technology to Precision Agriculture. In: Robert, P.C., Rust, R.H., Larson, W.E. (Eds.), Proceedings of the Fourth International Conference on Precision Agriculture. ASA-CSSA-SSSA, Madison, WI, pp. 1089–1100.
McBratney, A.B., Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S., Shatar, M., 2000. An overview of pedometric techniques for use in soil survey. Geoderma 97, 293–328.
McKenzie, N.J., Ryan, P.J., 1997. Spatial prediction of soil properties using environmental correlation. Geoderma 89, 67–94.
Minasny, B., McBratney, A., Whelan, B., 1999. VESPER—program for automatic variogram estimation and spatial prediction with error. Proceedings of the 3rd Conference of the Working Group on Pedometrics of the International Union of Soil Science (IUSS), September 27–29, The University of Sydney, pp. 67–68.
Moore, I.D., Gessler, P.E., Nielsen, G.A., Peterson, G.A., 1993. Soil attribute prediction using terrain analysis. Soil Science Society of America Journal 57, 443–452.
Odeh, I.O.A., Gessler, P.E., McKenzie, N.J., McBratney, A.B., 1994. Using attributes derived from digital elevation models for spatial prediction of soil properties. Resource Technology '94, Melbourne, Australia, pp. 451–463.
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, 215–225.
Odeh, I.O.A., McBratney, A.B., Slater, B.K., 1997. Predicting soil properties from ancillary information: non-spatial models compared with geostatistical and combined methods. In: Baafi, E.Y., Schofield, N.A. (Eds.), Geostatistics Wollongong '96, vol. 2, Kluwer Academic Publishers, Netherlands, pp. 1008–1019.
Rhoades, J.D., 1992. Instrumental field methods of salinity appraisal. In: Topp, G.C., Reynolds, W.D., Green, R.E. (Eds.), Advances in the Measurement of Soil Physical Properties: Bringing Theory into Practise. Soil Science Society of America, Madison, WI, USA, pp. 299–336.
Venables, W.N., Ripley, B.D., 1994. Modern Applied Statistics with S-Plus. Verlag, New York, p. 462.
Webster, R., Oliver, M.A., 1992. Sampling adequately to estimate variograms of soil properties. Journal of Soil Science 43, 177–192.