Whelan, B.M., McBratney, A.B. & Minasny, B. (2002). Vesper 1.5 – spatial prediction software for precision agriculture. In P.C. Robert, R.H. Rust & W.E. Larson (eds) Precision Agriculture, Proceedings of the 6th International Conference on Precision Agriculture, ASA/CSSA/SSSA, Madison, Wisconsin, 14p.

VESPER 1.5 – SPATIAL PREDICTION SOFTWARE FOR PRECISION AGRICULTURE

B.M. Whelan, A.B. McBratney and B. Minasny

Australian Centre for Precision Agriculture (ACPA),
University of Sydney, Sydney, Australia.

ABSTRACT

VESPER 1.5 is a shareware software program, written to provide rigourous spatial prediction techniques for the precision agriculture industry. It offers a range of options to deal with data sets of varying data density, spatial distribution, and observation uncertainty. Such data sets are now gathered from a range of real-time yield, soil and crop sensors and through manual sampling regimes. Specifically, the program provides the flexibility to calculate global and local variogram models, undertake global and local kriging in either punctual or block form and output the parameters and estimates in an ASCII text format. The program provides control of the semivariogram calculation and choice of models that may be fit to the input data. A boundary and prediction grid may be generated in the software or supplied as an external file. VEPSER 1.5 allows user defined neighbourhood and prediction-block sizes, along with a number of more advanced controls. It provides a real-time graphical display of the semivariogram modeling and a progress (and final) map of the kriged estimates. The value of the local variogram/kriging process in dealing with data sets generated for precision agriculture operations is shown here with a statistical comparison of the standard prediction techniques over a 100ha field. A comparison using a small portion (~1ha) of another field is also provided to illustrate both the visual impact of each technique and introduce the benefits block kriging of estimates brings to many of these data sets. Having the ability to tailor the prediction process to individual data sets is essential for Precision Agriculture (PA) where data quantity, density and measurement quality varies.

Keywords: spatial prediction, local variograms, block kriging, digital maps.

INTRODUCTION

Precision Agriculture (PA) tools, in particular crop yield monitoring, soil electrical conductivity measurement and intensive soil sampling have provided spatially dense data sets for use in crop management. And the desire to extract

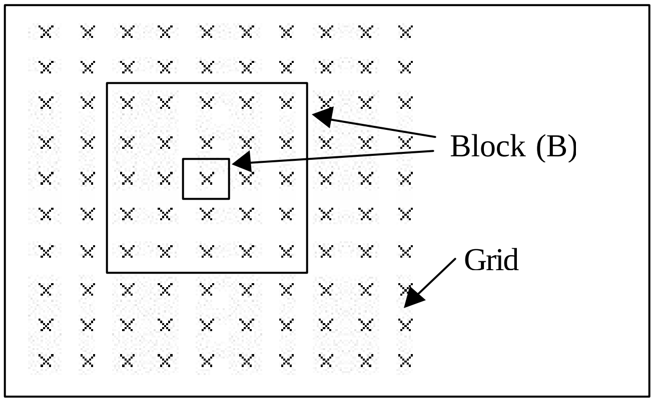


Figure1. Generalised map model.

valuable information from these data sets has also brought the process of digital map construction into wider use. All digital maps are based on some form of map model and usually require a spatial prediction procedure to produce a continuous surface map. The particular map model and the spatial prediction procedure chosen will have an impact on the predictions and the final map.

Map Model Description

Digital maps are constructed using a map model (Figure 1) whereby values are represented as a set of blocks (*B*) the centres of which are located on a grid (*G*). These models may take a number of general forms. According to Goodchild (1992) the blocks may have sides equal to the grid spacing (a raster model), the blocks may be points on a regular grid (a grid model) or they may be points and the grid irregular, or infinitely fine, with missing values or values equal to zero (a point model).

Spatial Prediction Techniques

Any form of spatial prediction is based on the premise that observations made in close proximity to each other are more likely to be similar than observations separated by larger distances. This is the concept of spatial dependence. The process of spatial prediction requires that a model of the spatial variability (spatial dependence) in a data set be constructed or assumed so that estimates at the unsampled locations (prediction points) may be made on the basis of their location in space relative to actual observation points. It is the form of these models, and the assumptions underlying the choice of the same, which generally distinguish the major spatial prediction methods.

Global methods use all the data to determine a general model for spatial dependence. This model is then applied, in association with the whole data set, in the prediction process at every prediction point. Local prediction methods use only points 'neighbouring' the prediction point in the prediction operation. In the case of local predictors, a singular form of the spatial variance model may be constructed for the entire data set and applied in each neighbourhood, or an individual model may be constructed, and used exclusively for, each neighbourhood. Local methods may therefore be the preferred option, especially on large data sets, and where a single variance model may be inappropriate.

Spatial prediction methods whose principle requires the prediction to exactly reproduce the data values at sites where data is available are said to act as interpolators. There is a variety of prediction techniques which may be applied to mapping continuous surfaces. The most widely known include: global means and medians; local moving means; inverse distance squared interpolation; Akima's interpolation; natural neighbour interpolation; quadratic trend; Laplacian smoothing splines; and various forms of kriging.

The prediction technique of choice for map production in precision agriculture will depend on the expected use of the map. However, real-time sensors that intensively sample variables such as crop yield, produce large data sets containing a wealth of information on small-scale spatial variability. By definition, precision agricultural techniques should aim to identify the quality of the data and preserve the appropriate degree of detail.

VESPER

VESPER 1.5 (Variogram Estimation and Spatial Prediction plus Error) is a PC-Windows software program developed by the ACPA that allows the geostatistical spatial prediction procedures of punctual and block kriging to be applied to data sets gathered for PA management. The program also offers the further options of global or local kriging, using global or local semivariograms.

Input and ouput files are controlled through the ‘File’ panel (Figure 2a). Input data with associated Cartesian coordinate locations is required to enable spatial analysis. The output files record the specific session setup details, variogram model parameters and the prediction locations, values and associated prediction variance.

The ‘Variogram’ panel provides the choice of global or local semivariogram estimation and provides access to a choice of models (Figure 2b) which may be fit to the semivariogram using 3 possible weighting procedures (Figure 2c). Nonlinear least-squares estimation is used in the model fitting process. The model may be chosen from a comprehensive range of options. Provision is made for comparison of the ‘goodness of fit’ of the numerous models through the Akaike Information Criteria (Akaike, 1973) and sum of squared error (SSE). If a global semivariogram is required, the ‘Fit Variogram’ button provides access to an interactive calculation and modeling panel (Figure 2d) from which the final model parameters are extracted for use in the subsequent kriging procedures. The global modelling panel now also provides for subjective model fitting through interactive parameter control bars. This is useful in small data sets and applications where emphasis needs to be placed on particular regions of the sampling separation distance.

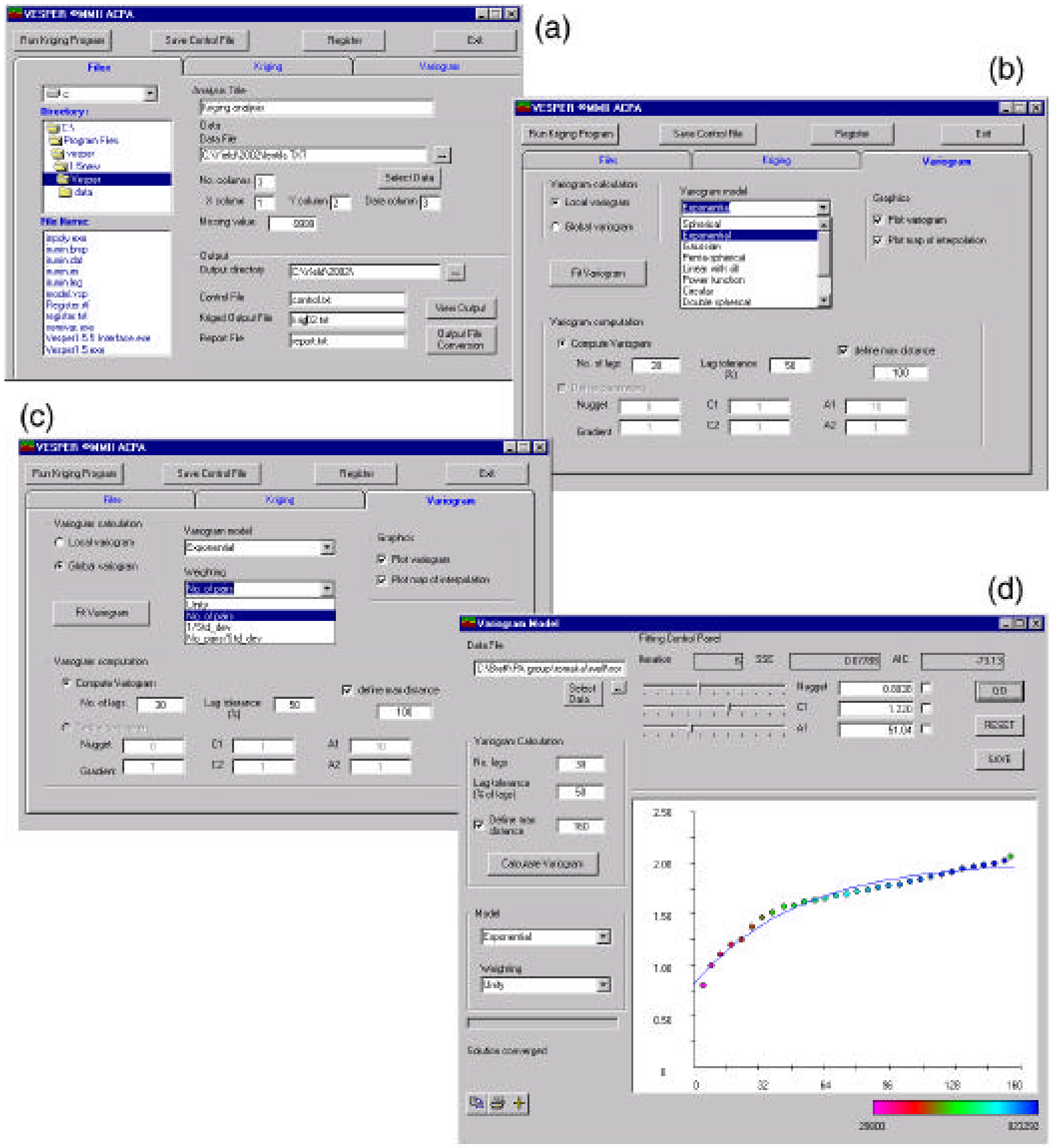


Figure 2. Operational panels – file input/output control panel (a), variogram panel showing available models (b), variogram panel showing weighting options for model fitting (c), global variogram operation window (d).

The 'Kriging' panel (Figure 3a) provides kriging type (ordinary or simple) and method (punctual or block) options. Here it is also possible to define the block size (if relevant), set neighbourhood limits based on radial distance or number of data points and manipulate the kriging region. For most PA applications, the field boundary will provide the limits of the kriging region. VESPER 1.5 provides the option of importing an existing boundary file or describing the field boundary using an interactive drawing tool (Figure 3b). The prediction grid (at user-defined distances) may then be produced with the software (Figure 3c) or a previous grid file imported. These features are important for the continuity of prediction sites through time within a field.

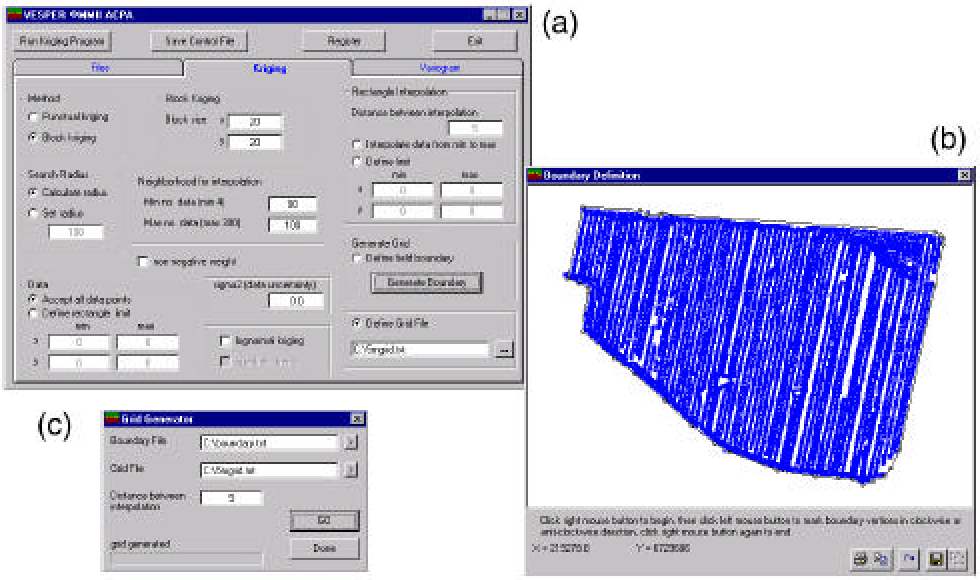


Figure 3. Kriging panel (a) interactive boundary construction window (b) and prediction grid setup window (c).

In operation, VESPER 1.5 provides a window displaying the operational progress (Figure 4). For all forms of kriging a prediction progess map is produced along with a count of visited versus total prediction sites. The graphical progress facilities can be disengaged to increase the speed of the prediction process.

Local semivariograms are calculated for each neighbourhood during the local kriging process, but the maximum distance and number of lags required for estimating the local semivariograms is set through the 'Variogram' panel. Kriging with local variograms involves searching for the data points within the defined neighbourhood surrounding each prediction site, estimating the variogram cloud for the data points and fitting a model, then predicting a value (and its uncertainty) for the attribute under question at each prediction site. Note in Figure 4 (a) and (b) that this local method allows changes in local variability to be reflected in the variogram parameters for each prediction.

The output for all kriging operations is a five column ASCII text file containing the prediction point ID, location coordinates, the predicted value and the kriging variance. Access is provided at the end of the kriging process to manage the data delimeter and include/exclude the point ID and header descriptions in the final output file. This allows input formats to be tailored to GIS and map display programs. An input file detailing the exact settings for each prediction session is also saved along with a report file logging global variogram parameters or the parameters of each local variogram depending on the operation. Other details of the data and the kriging session are also recorded in this file for future reference. A surface map of the estimates and the prediction variance (Figure 5) can also be accessed at the completion of the kriging procedure.

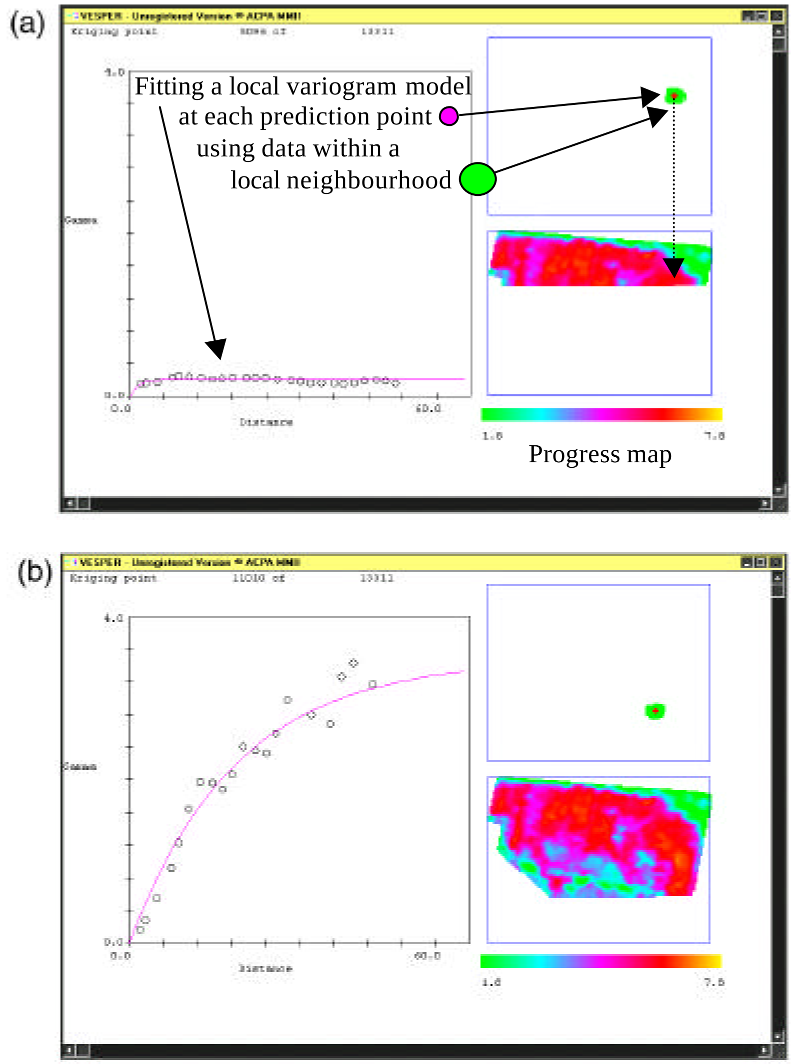


Figure 4. Local variogram, data neighbourhood and prediction point display for and area with low variability (a) and higher variability (b).

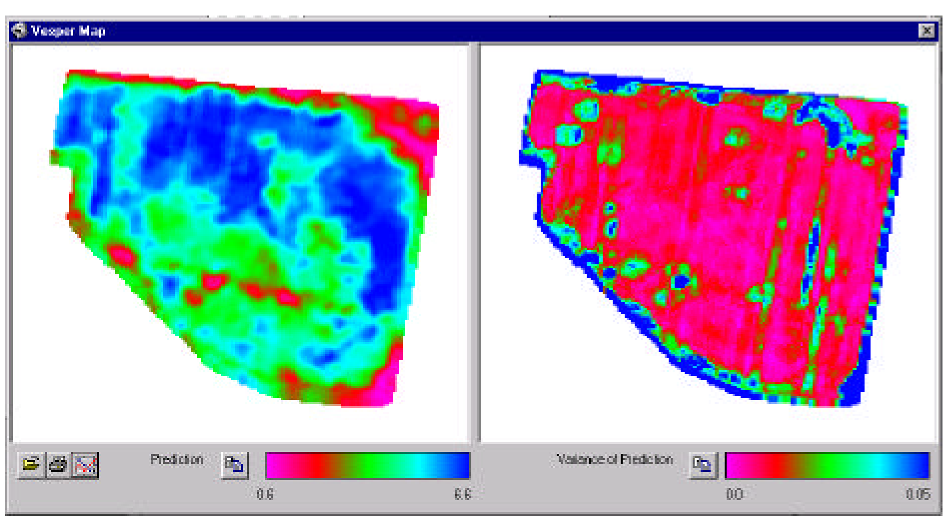


Figure 5. Output maps for prediction estimates and prediction variance.

IMPACT OF PREDICTION TECHNIQUES ON DIGITAL MAPS

Comparison by Distribution and Performance Rankings

Individual wheat yield values, collected at a frequency of 1 Hz from a 100 ha field in NSW, Australia, were randomly allocated into one of two equal-size datasets. One data set was used as input values for the prediction processes, the other provided the prediction locations and test values for a comparison of point (punctual) prediction techniques. Local mean, inverse-distance squared, local kriging with a global variogram are compared along with the less common technique of local kriging with a local variogram (Haas, 1990). A search neighbourhood of 100 data points was used as standard.

Table 1 shows the resulting frequency distributions and performance rankings of the prediction techniques in comparison to the observed values at the locations. The rankings (1:4 – where the closest prediction value to the observation value gets a ranking = 1) are calculated at each point and then summed for each technique. The final performance rank is allocated from the lowest to the highest sum of ranks.

Here the estimates from the kriging procedures most closely match the original observation values and thereby maintain more of the original frequency distribution. Local kriging with a local semivariogram has performed the best. Inverse distance-squared, while performing third overall, has registered the smallest frequency of number one ranks.

Comparison by Spatial Representation

To visually demonstrate the results of the different prediction methods on crop yield data, a small portion (~1ha) of another field and crop has been chosen. Sorghum yield data, acquired using a real-time yield monitor in 7 metre wide harvest runs, was predicted onto a regular 1 metre grid using the point prediction

Table 1. Wheat yield frequency distribution and performance rankings for spatial prediction techniques on a 100ha field in NSW, Australia. (26337 observations)

Technique	Max. (t/ha)	Min. (t/ha)	Mean (t/ha)	Sum of ranks	Median rank	No. of ranks = 1	Final Rank
Test data	6.26	0.92	3.71				
Local kriging w/ local variogram	5.99	1.01	3.71	59152	2	9150	1
Local kriging w/ global variogram	5.88	1.11	3.71	60688	2	7421	2
Inverse distance-squared	5.71	1.01	3.72	63382	3	4480	3
Local mean	5.01	1.87	3.72	80168	4	5284	4

methods of local inverse distance-squared, local kriging with a global semivariogram and local punctual kriging with a local semivariogram. In addition, local block kriging with a local semivariogram has been undertaken.

Block kriging has rarely been used since Burgess & Webster (1980) introduced geostatistical spatial prediction techniques into soil science, and software for performing it is rather scarce. Block kriging attempts to predict the weighted average of a variable over some block of length (*dx*) and width (*dy*) centred about some prediction point (*x0, y0*). It should be noted that the locations (*x0, y0* - the prediction grid or raster) can be closer together than the block length or width. This in fact gives an aesthetically pleasing, smooth map. The major advantage of using block kriging is that the estimate of the block mean, not surprisingly, improves as the block dimensions increase.

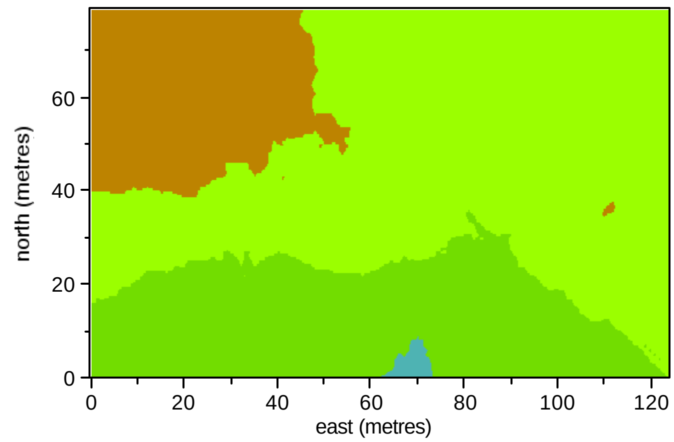
In Figure 6a shows the map produced by the simple process of local moving mean. The map is smoothed by the moving window operation and the fact that all data points receive equal weight in the prediction process. Figure 6b, the inverse distance method, places a lot of varibility in the map by virtue of honouring the very high and low peaks in the harvest data. It is easy to distinguish the harvest operation lines that run NW/SE in the surface map. Because the inverse distance squared model is fixed, and its radius of influence is small, the map takes on the characteristic "spottiness" of maps made using this technique.

Local kriging with a global semivariogram (Figure 6c) has smoothed out the map to a degree and the harvest operation lines are not evident because the variogram has captured a longer spatial dependence in the data set than the fixed inverse distance model. Data points from further out in the neighbourhood have been given some influence on the prediction at each point.

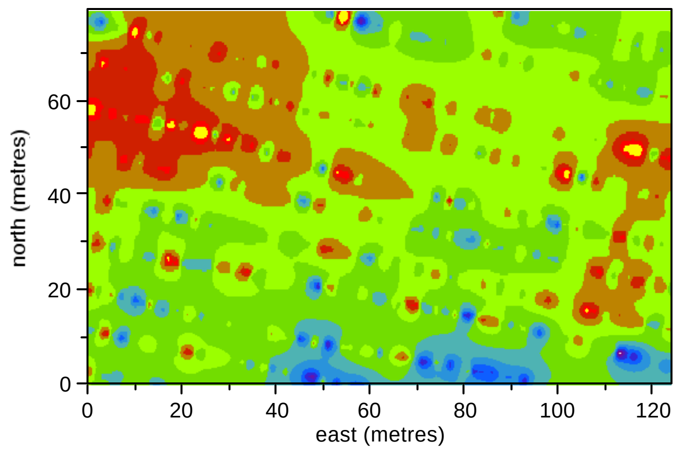
Local Kriging with local variograms (Figure 7a) restore some of the local variability because the changes in spatial dependence between the local neighbourhoods is included. Changing the map model from point estimates to estimates representing the weighted average yield in a 20 metre block around each prediction point (Figure 7b) removes some of this variability from the estimates.

That the form of spatial prediction chosen for map construction may be significantly influential on the final prediction surface is not a new concept. A number of studies (e.g. Laslett et al. (1987), Wollenhaupt et al. (1994), Weber & Englund (1994), Whelan et al. (1996), Gotway et al. (1996)) show that in general inverse distance techniques are sensitive to the degree of inherent variability in a data set, the neighbourhood population used in each prediction and the power of distance used in the weighting calculation. Alternatively, the accuracy of ordinary kriging generally displays little sensitivity to the variability in the data sets and the accuracy of the estimates improves with increasing neighbourhood populations.

(a)



(b)



(c)

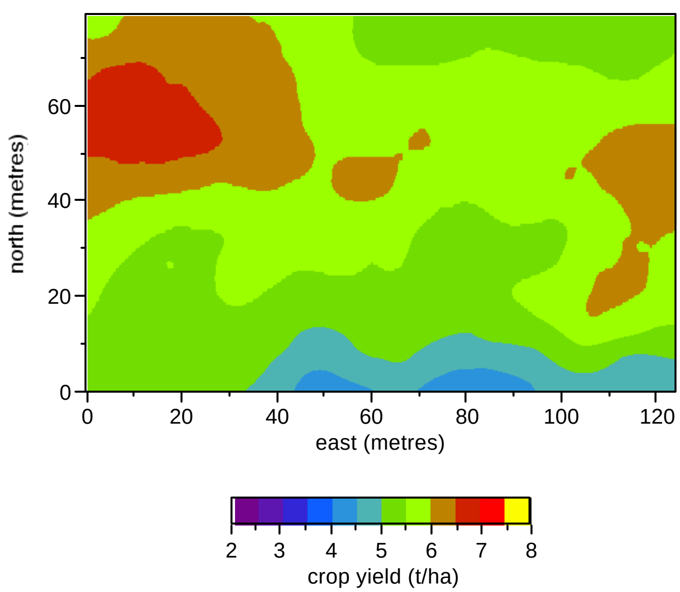
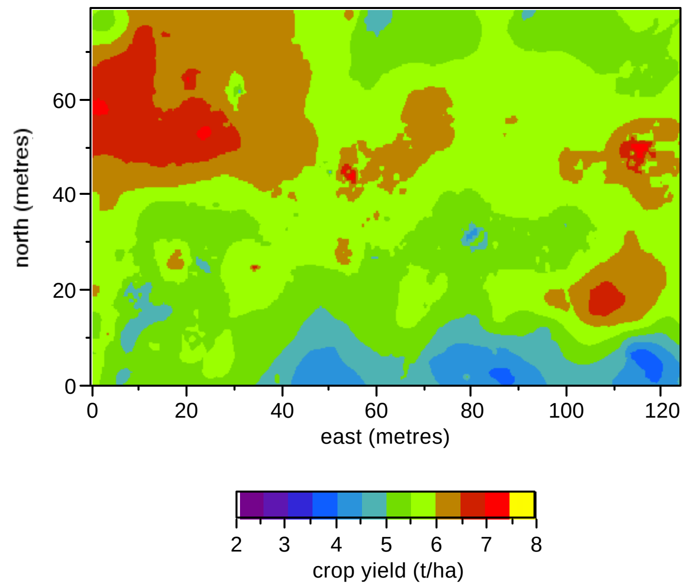


Figure 6. Crop yield maps constructed using different prediction procedures. (a) local mean (b) inverse distance-squared (c) local punctual kriging with a global variogram.

(a)



(b)

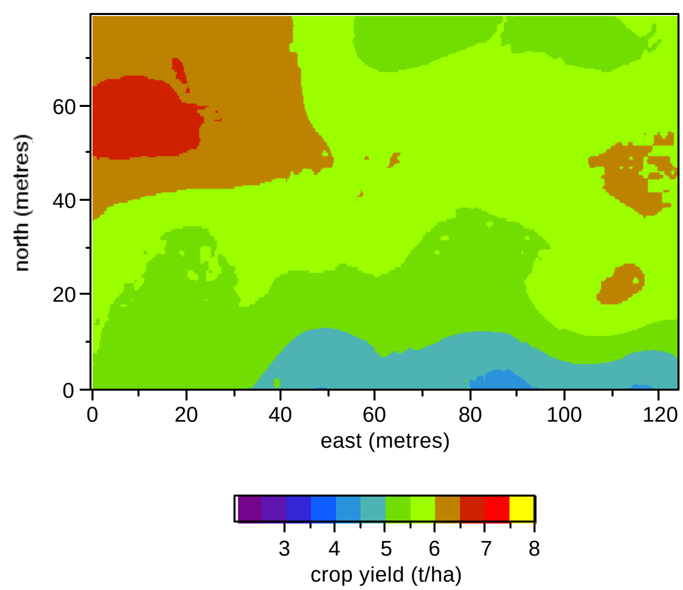


Figure 7. Crop yield maps constructed using different prediction procedures. (a) local punctual kriging with a local variogram (b) local block kriging with a local variogram.

The observed inefficiencies of the inverse distance squared prediction technique can be attributed to two main problems. Firstly, the spatial variability in a data set is not used to determine the spatial dependence model for use in the prediction process. Secondly, the method is an exact interpolator that passes through the data points, and this may not be sensible if there is uncertainty in the observations. Kriging only operates as an interpolator if the semivariogram nugget value (C0) equals zero. With any positive C0 value, close range uncertainty in the observations will be reflected in the kriged surface. Such uncertainty may arise in either the value of the observed attribute or its spatial location.

This point is often overlooked in assessing the suitability of prediction techniques but should be a given a high priority in PA owing to the potential (and real) errors associated with real-time sensors and GPS receivers (Lark et al. (1997); Whelan and McBratney (2002); Arslan and Colvin (2002)). In such cases, block kriging estimates for an area should prove extremely useful in reducing the carryover of errors into the final maps. Block kriging also offers a robust method of estimating values for an area that represents the smallest differentially manageable land unit in a farming operation (usually governed by implement width and operational dynamics).

Block kriging may be undertaken using a global semivariogram but once the number of data points rises above 500 it seems wasteful to assume a single semivariogram within the field. A global semivariogram may prove too restrictive in its representation of local spatial structure whereas local semivariogram estimation and kriging offers the ability to preserve the true local spatial variability in the predictions. If the chosen neighbourhood is reasonably small, the use of local semivariograms should also negate the possible requirement for trend analysis and removal prior to semivariogram estimation and kriging.

A further advantage in the use of kriging techniques lies in the provision of a prediction variance estimate (Laslett et al., 1987; Brus et al., 1996) which may be used to produce confidence limits on the predicted values. The reporting of such limits should be mandatory for digital maps as they will have important ramifications on the extrapolation of management information (Whelan and McBratney (1999); Cuppit and Whelan (2001). The uncertainty may also be used to determine the most suitable mapping class delineations in digital maps. For example, if the 95% confidence interval in crop yield estimates is +/- 1.0 t/ha, classifying a field using classes less than 1.0 t/ha would be misleading. A classification system based on the uncertainty in the yield data may prove useful in the future.

On the other hand, criticisms that have been levelled at the kriging techniques' complexity and related computational expense (e.g. Murphy et al., 1995). Astoundingly, this one line of criticism has apparently overridden all the advantages discussed above, and led to the general acceptance of the inverse distance method as the prediction method of choice in the emerging mapping packages for Precision Agriculture. While there may be some instances where a prediction map is required quickly (e.g. soil attribute maps for interpolation to fertiliser application maps), at present the author believes this is not a rational reason for discarding the advantages incumbent with kriging techniques. Certainly for crop yield maps, the computational time would be far outweighed by the single fact that the map represents a great deal of time, effort and expense taken to grow a crop. Ultimately, it is the integration of an entire seasons crop growth information.

Where the computational expense may become important (and indeed the choice of prediction technique possibly unimportant) is when the observation sampling scheme is inadequate in terms of sample size, sample strategy, or both. Sample size is probably considered the most crucial parameter (Englund et al. 1992) with an increasing number of observations generally offering greater prediction accuracy. Numerous studies on the effect of sample strategy for regionalised variables have been reported since the early theoretical work of McBratney et al. (1981). The general axiom to emerge is that sampling schemes which fail to produce a sample set representative of the actual spatial variability in the attribute of interest will hinder accurate prediction by any method. Data sets from calibrated real-time sensors should not fall into this category, but traditional soil sampling operations may produce such data.

CONCLUDING REMARKS

Spatial prediction methods used in PA should accurately represent the spatial variability of sampled field attributes and maintain the principle of minimum information loss. However, data used in any spatial prediction procedure should be of known precision and that precision used to guide the choice of spatial predictor. Due to imprecision in crop yield measurement and within-field location, interpolators (exact spatial predictors) are generally not optimal.

The results presented show that the form of spatial prediction chosen for mapping yield has a significant influence on the final prediction surface. Local kriging using a local variogram appears well suited as a spatial prediction method for dense data-sets. In particular, local block kriging reduces the estimate uncertainty when compared with punctual kriging and may be an optimal mapping technique for the current generation real-time yield and soil sensors.

Ultimately, any software devised for spatial prediction in precision agriculture applications should include options that will optimally support the management decisions that will be formulated upon the prediction results.

VESPER is available as shareware from the ACPA at www.usyd.edu.au/su/agric/acpa

REFERENCES

Akaike, H. (1973). Information theory and an extension of maximum likelihood principle. *In* B.N. Petrov & F. Csáki (ed.), Second International Symposium on Information Theory. Akadémia Kiadó, Budapest. pp 267-281.

Arslan, S., Colvin, T.S. (2002). Grain yield mapping: yield sensing, yield reconstruction and errors. Precision Agriculture 3 : 135-154.

Brus, D.J., de Gruijter, J.J., Marsman, B.A., Visschers, R., Bregt, A.K., Breeuwsma, A., Bouma, J. (1996). The performance of spatial interpolation methods and choropleth maps to estimate soil properties at points: a soil survey case study. Environmetrics 7: 1-16.

Burgess, T.M., Webster, R. (1980). Optimal interpolation and isarithmic mapping of soil properties. II. Block kriging. J. Soil Sci. 31 : 333-341.

Cupitt, J., Whelan, B.M. (2001). Determining potential within-field crop management zones. *In* G. Grenier and S. Blackmore (ed.), ECPA 2001, Proceedings of the 3rd European Conference on Precision Agriculture, Montpellier, France, agro-Montpellier, Montpellier, France pp 7-12.

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

Goodchild, M.F. (1992). Geographical data modelling. Computers and Geosciences 18: 401-408.

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

Haas, T.C. (1990). Kriging and automated semivariogram modelling within a moving window. Atmospheric Environment 24A: 1759-1769.

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

Lark, R.M., Stafford, J.V., Bolam, H.C. (1997). Limitations on the spatial resolution of yield mapping for combinable crops. J. Agric. Engng Res. 66: 183-193.

McBratney, A.B., Webster, R., Burgess, T.M. (1981). The design of optimal sampling schemes for local estimation and mapping of regional variables – I : theory and method. Computers and Geosciences 7: 331-334.

Murphy, D.P., Schnug, E., Haneklaus, S. (1995). Yield mapping – a guide to improved techniques and strategies. *In* P.C. Robert, R.H. Rust & W.E. Larson (ed.), Site-Specific Management for Agricultural Systems: Proceedings of the 2nd International Conference on Precision Agriculture, ASA/CSSA/SSS, Madison. pp. 32-47.

Weber, D.D., Englund, E.J. (1994). Evaluation and comparison of spatial interpolators II. Math. Geol. 26: 589-603.

Whelan, B.M., McBratney, A.B. (1997). Sorghum grain-flow convolution within a conventional combine harvester. *In* J.V. Stafford (ed.), Precision Agriculture 1997, Bios, Oxford, England. pp 759-766.

Whelan, B.M., McBratney, A.B. (1999). Prediction uncertainty and implications for digital map resolution. *In* P.C. Robert, R.H. Rust & W.E. Larson (ed.), Precision Agriculture, Proceedings of the 4th International Conference on Precision Agriculture, ASA/CSSA/SSSA, Madison, Wisconsin. pp 1185-1196.

Whelan, B.M., McBratney, A.B. (2002). A parametric transfer function for grain-flow within a conventional combine harvester. Precision Agriculture 3 : 123-134.

Whelan, B.M., McBratney, A.B., Viscarra Rossel, R.A. (1996). Spatial prediction for Precision Agriculture. *In* P.C. Robert, R.H. Rust & W.E. Larson (ed.), Precision Agriculture: Proceedings of the 3rd International Conference on Precision Agriculture, ASA/CSSA/SSS, Madison. pp. 331-342.

Wollenhaupt, N.C., Wolkowski, R.P., Clayton, M.K. (1994). Mapping soil test phosphorus and potassium for variable-rate fertilizer application. J. Prod. Agric. 7 : 441-448.