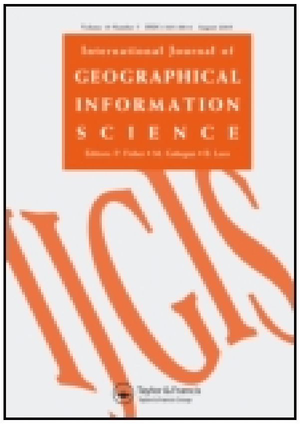

International journal of geographical information systems
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tgis19

Soil-landscape modelling and spatial prediction of soil attributes

P. E. GESSLER [a] [b] , I. D. MOORE (deceased) [c] , N. J. McKENZIE [d] & P. J. RYAN [e]

[a] CSIRO Division of Soils, GPO Box 639, Canberra, ACT 2601

[b] Australia and Centre for Resource & Environmental Studies, Australian National University, Canberra, Australia E-mail: paulg@cbr.soils.csiro.au

[c] Centre for Resource & Environmental Studies, Australian National University, Canberra, Australia

[d] CSIRO Division of Soils, GPO Box 639, Canberra, ACT 2601, Australia

[e] CSIRO Division of Forestry, Canberra, Australia

Available online: 05 Feb 2007

To cite this article: P. E. GESSLER, I. D. MOORE (deceased), N. J. McKENZIE & P. J. RYAN (1995): Soil-landscape modelling and spatial prediction of soil attributes, International journal of geographical information systems, 9:4, 421-432

To link to this article: http://dx.doi.org/10.1080/02693799508902047

INT. J. GEOGRAPHICAL INFORMATION SYSTEMS, 1995, VOL. 9, NO. 4, 421–432

Research Article

Soil-landscape modelling and spatial prediction of soil attributes

P. E. GESSLER
CSIRO Division of Soils, GPO Box 639, Canberra, ACT 2601, Australia and Centre for Resource & Environmental Studies, Australian National University, Canberra, Australia
email: paulg@cbr.soils.csiro.au

I. D. MOORE (deceased)
Centre for Resource & Environmental Studies, Australian National University, Canberra, Australia

N. J. McKENZIE
CSIRO Division of Soils, GPO Box 639, Canberra, ACT 2601, Australia

and P. J. RYAN
CSIRO Division of Forestry, Canberra, Australia

Abstract. Explicit and quantitative models for the spatial prediction of soil and landscape attributes are required for environmental modelling and management. In this study, advances in the spatial representation of hydrological and geomorphological processes using terrain analysis techniques are integrated with the development of a field sampling and soil-landscape model building strategy. Statistical models are developed using relationships between terrain attributes (plan curvature, compound topographic index, upslope mean plan curvature) and soil attributes (A horizon depth, Solum depth, E horizon presence/absence) in an area with uniform geology and geomorphic history. These techniques seem to provide appropriate methodologies for spatial prediction and understanding soil landscape processes.

1. Introduction

Environmental models require spatial representation of soils because they modify material and energy fluxes at the earth's surface. Ideally, spatial predictions of soil layers, individual soil attributes and, eventually, soil-landscape processes, are needed at a scale appropriate for environmental management (Moore *et al.* 1993). The challenge is to develop explicit, quantitative, and spatially realistic models of the soil-landscape continuum useful for a variety of purposes beyond taxonomic classification (McSweeney *et al.* 1994). A promising development is the potential for correlating soil attributes with terrain and environmental attributes that are simple to measure and have physical meaning (Moore *et al.* 1993, McKenzie and Austin 1993). The underlying hypothesis of Moore *et al.* (1993) was that the development of soil toposequences often occurs in response to the way water moves through and over, the landscape. Water movement is in turn controlled by the geometry of the land surface and underlying materials. The geometry of the land surface, therefore, can be used as a first approximation for predicting the movement of water and related material (Moore *et al.* 1991).

There has been a trend in recent work (Bouma 1989, Gessler *et al.* 1989, Baize and Girard 1992, FitzPatrick 1993) towards using soil layers rather than soil profiles or pedons as the basic object for study. Soil layers may have a pedogenic (soil horizon) or geomorphic (stratigraphic unit) origin. Regardless of origin, they form a logical building block for spatial modelling and interpretation of how sequences of layers behave. The soil layers at any location are a result of integrated pedo-geomorphic and hydrological processes (Simonson 1959, Butler 1964). As such, a description of the arrangement, dimension and nature of the soil layers at locations in the landscape may be used as a link or pointer to the spatial distribution of processes and vice-versa.

However, soil-landscape processes operate across a range of spatial and temporal scales (Allen and Starr 1982, Kachanowski 1988) and it is clear that imprinting of past climates, truncation by over-riding processes, and process synergisms occur (Malanson *et al.* 1990, Allison 1991). Consequently, soil attributes exhibit different and complex scales of variation (Butler 1964, Beckett and Webster 1971, Burrough 1993). Thus, our expectations for deciphering the relationship between pattern and process should vary within and between physiographic domains. This reinforces the need to develop environmental correlations using exploratory data analysis (Tukey 1977) followed by explicit definition based on physically interpretable statistical models (Chambers and Hastie 1992, McKenzie and Austin 1993).

The general form of such statistical models being:

$$S_i = f_i(\text{slope, catchment position, solar radiation, gamma radiometrics, } \ldots)$$

where:

- S is individual soil attribute (e.g. soil depth, pH, etc.);
- f is a function of one or several environmental attributes;
- i is the physiographic domain characterized by common climate, parent material, geomorphic history, vegetation, etc.

In this approach, the statistical model is developed using data from measurements of soil attributes (response variable) made in the field at locations where measurements of environmental attributes (explanatory variable(s)) are available. Spatial prediction is then achieved using environmental variables, such as slope, that may be generated using digital terrain methods or other techniques. The environmental variables must be easier to obtain than soil variables and be available for the complete study area, otherwise, intensive sampling of soil variables in association with an interpolation or surface fitting procedure would be a more efficient method for spatial prediction. The definition of the physiographic domain where a developed model applies depends on the scale and purpose of the work (McSweeney *et al.* 1994). It could be for broadly defined regions (e.g. river basins, land systems) or more local areas defined by hillslopes within a given geomorphic unit. With this approach, primary data can be re-analysed with different combinations of response and explanatory variables, and statistical methods can be varied as suggested by exploratory data analysis and general field observation.

Australia contains vast areas with scant land resource information. The resources for collecting basic data sets to understand environmental function and management are limited (McKenzie 1991). This paper presents initial results on the testing of a method for developing explicit soil-landscape models using pedological knowledge, spatial analysis, field sampling, exploratory data analysis and statistical modelling techniques. The broad aims of this work were to develop:

(*a*) Procedures for the quantitative characterization of landform because of its importance as a local scale predictor of soil attributes.
(*b*) A rational and efficient soil sampling strategy.
(*c*) Robust statistical models for the spatial prediction of soil attributes in an area with uniform geology and geomorphic history; and
(*d*) Quantitative methods for comparing and understanding soil-landscape processes.

Developed models and predictions may then be used to parameterize other models for environmental management (e.g. estimation of erosion hazard, crop growth, water quality, nutrient cycling) and for simulation of impacts due to changes in land use.

2. 2-D spatial characterization of processes: terrain analysis

Moore *et al.* (1991) review terrain analysis and its application in the earth sciences. Primary and secondary (or compound) topographic attributes are recognized and they present a table summarizing the significance of these attributes for characterizing the spatial distribution of landscape processes. Many of the attributes have potential use as spatial predictors of soil attributes. Primary attributes are directly calculated from elevation data and include areal measures such as specific catchment area and point measures including the first and second derivatives such as slope, aspect, plan and profile curvature. Secondary attributes involve combinations of the primary attributes that quantify the contextual nature of points or characterize the spatial variability of specific processes occurring in the landscape or both. Methods of computation are presented by Moore *et al.* (1991, 1993).

Digital topographic attributes are scale dependent and if these effects are not considered, computed attributes may be meaningless or the processes of interest may be masked (Moore *et al.* 1991, 1994). Moore *et al.* (1994) report critical differences in the computation methods of primary and secondary topographic attributes and, for example, advise against the use of the D8 method of flow direction computation. This method does not allow flow dispersion and produces unrealistic flow patterns. This significantly influences the computation of flow accumulation which is critical to the computation of many spatial hydrological and soil-landscape attributes such as catchment and dispersal areas. Differences in environmental attribute correlations and model development will occur due to physiographic setting, scale of analysis, computation methods, and others (data structure, quality and error). It is essential for a modelling framework to have explicit definition of decisions relating to the particular combination of methods applied (McSweeney *et al.* 1994, Wagenet *et al.* 1994) so that other workers can evaluate, repeat or improve the model.

3. Material and methods

3.1. *Study region*

The study region is the Wagga Wagga 1 : 100 000 topographic map sheet located on the western slopes of the Great Dividing Range in southeastern Australia (147°E, 35°S; 147°E, 35°30′S; 147°30′E, 35°30′S; 147°30′E, 35°S). This region was chosen because it has a diverse range of geological units, landforms, soils and land uses typical of the broader Murray-Darling River Basin. Areas with distinct combinations of geology and landform (physiographic domains) have been delineated and later work will develop soil-landscape models in each area for testing and comparison. This paper focuses more specifically on initial methodology development in a 100 km^2 pilot study area (centered on 147°27′E, 35°24′S) dominated by gently-rolling erosional landforms on Ordovician metasediments. The dominant land use is pastoral grazing.

3.2. *Soil-landscape model development*

Two methods have recently been proposed for development of explicit and quantitative soil-landscape models (McKenzie and Austin 1993; McSweeney *et al.* 1994). Both methods are similar in approach and require a definition of purpose, scale of application and stratification of the physiographic domain for field sampling. The work reported here is aimed at developing spatial models of soil layer patterns within the Ordovician meta-sediment physiographic domain. The models of soil layer patterns are viewed as critical if they are to lead to eventual spatial prediction of individual soil attributes and soil-landscape processes in three and four (time) dimensions. The scale of application is the hillslope within small catchments and is intended to provide information at the local land management level in the study area. The soil layer is used as the basic object of study and the catchment is the boundary of the system, due to its significance for spatially related hydrological and erosional processes.

Stratification into distinct physiographic domains for soil-landscape model development is a critical initial step. The quality of stratification depends on the availability of prior information on soils, geology, vegetation, landform and surficial materials. At the onset of this work a 1 : 100 000 geology map (Raymond 1992) was generated and initial stratification into physiographic domains was performed using these data. Additional data layers (soils, landform, stratigraphy, vegetation, climate) are being generated as part of a collaborative project, and subsequent work will look at stratification using these integrated data more specifically. The focus here is on the methods of explicit soil-landscape model development within one physiographic domain.

Digital contours (10 m contour interval), streamlines and spot heights registered to the Australian Map Grid (AMG-UTM) were obtained from the New South Wales Land Information Centre in digital form. A base-line 20 m × 20 m grid digital elevation model (DEM) for the 100 km^2 study area was generated using the program ANUDEM (Hutchinson 1989). Scaling parameters, fractal and error properties of this surface are reported elsewhere (Moore *et al.* 1994). Seventeen catchments were delineated and a full range of primary and secondary topographic attributes were generated for each catchment using the methods of Moore *et al.* (1993, 1994). Flow or area accumulation (i.e. specific catchment area) was calculated using the FRho8 flow dispersion algorithm (Moore *et al.* 1994), and a 100 cell channel initiation threshold. When this threshold is reached, flow accumulation switches from dispersive to channelized flow using the D8 method.

As part of this work, a new algorithm was developed for computing upslope statistical moments (i.e., upslope mean slope, upslope mean plan curvature etc.,) to provide quantitative information about the upslope catchment area feeding into each grid cell. The S-Plus language (Statistical Sciences 1992) was used to develop a function to create graphical displays of the probability density function and listing of the statistical moments and Moran statistics (Goodchild 1986) for each terrain attribute on a catchment basis. This function enables the rapid characterization of a catchment and quantitative comparison of the overall differences between catchments or specific zones within a catchment. A contiguous five catchment subarea (20 868 cells or 834·72 hectares) was selected for soil-landscape model development because it encompassed a range of the topographic variability (including aspect) characteristic of the physiographic domain as a whole.

3.3. *Development of an explicit and quantitative sampling strategy*

An iterative sampling strategy using four criteria was used to select field sample sites. First, the sampling plan aimed to reflect the provisional predictive pedologic model by sampling evenly along the predictive variable(s) in attribute space. Secondly, randomization was used to achieve an unbiased sample. Thirdly, sampling inefficiencies due to spatial dependence in soil attributes were minimized. Fourthly, locational error between the digital terrain model and the real world was minimized.

3.3.1. *Provisional predictive pedologic model & randomization*

When a soil surveyor initiates a survey in un-mapped territory, he or she often begins with an implied model or mental construct and begins testing hypotheses with sample points. This provisional predictive pedologic model (McKenzie and Austin 1993) evolves as points are sampled. But much of this information about continuous soil-landscape variation is lost or subsumed when map unit lines are drawn. A common provisional model is the catena (Latin = a chain) soil-landscape (Milne 1935) that implies a concordance of soil pattern with landform as one traverses from hilltop to valley bottom along toposequences. The compound topographic index (CTI), often referred to as the steady-state wetness index, is a quantification of catenary landscape position. It is defined as:

$$\mathrm{CTI} = \ln\,(A_s/\tan\beta) \tag{1}$$

where A_s is the specific catchment area (area (m^2) per unit width orthogonal to the flow direction) and β is the slope angle. Moore *et al.* (1993) showed that the CTI is correlated with several soil attributes such as silt percentage ($r = 0{\cdot}61$), organic matter content ($r = 0{\cdot}57$), phosphorus ($r = 0{\cdot}53$) and A horizon depth ($r = 0{\cdot}55$) in the soil surface of a small toposequence. The CTI was used in this work as an explicit and quantitative provisional predictive pedologic model. To develop a robust statistical model for testing hypothesized correlations, it is sensible to sample evenly in CTI attribute space. Thus, the CTI was divided evenly into five equal percentile classes (figure 1 (*a*)). The goal of this work was to develop a soil-landscape model applicable to the broader Ordovician metasediment physiographic domain. Therefore, the percentile break-points were computed using all the grid cells falling on this bedrock type in the 100 km^2 study area. Figure 1 (*b*) shows a spatial display of the percentile classes for the study catchments. The percentile classes also provide convenient strata or patches that can be used for randomization to meet the second sampling criterion.

3.3.2. *Spatial dependence*

Soil attributes show varying degrees of spatial dependence and this reduces the efficiency of random sampling (McBratney and Webster 1983). Spacing sample sites using information about the spatial dependence structure increases the information content of samples. No *a priori* information on the spatial dependence structure of the soil attributes of interest was available. Instead we postulated that the spatial dependence structure of the CTI related in a general way to the spatial dependence structure of the soil attributes of interest. Moran's I coefficient (Goodchild 1986), which characterizes the overall strength of spatial dependence, is 0·70 for the CTI cells on the Ordovician meta-sediments in the 100 km^2 study area. This indicates strong

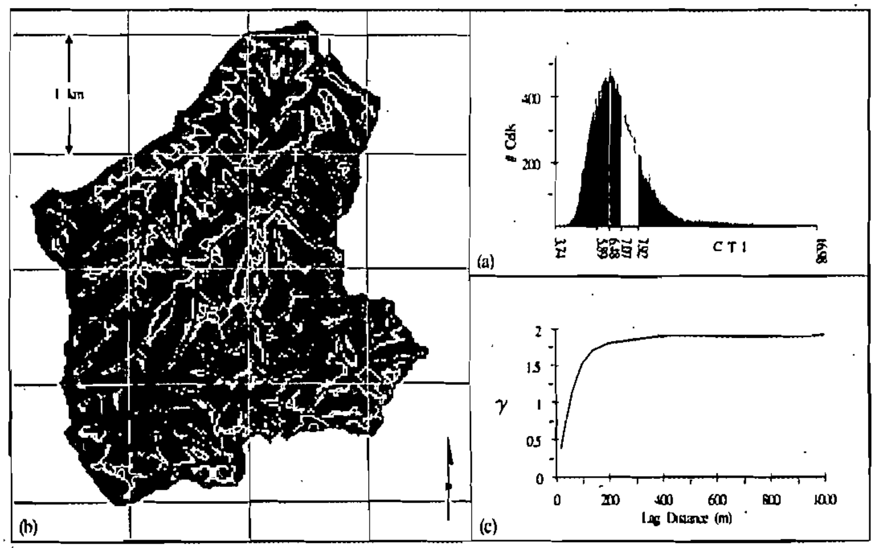

Figure 1. (*a*) Twenty percentile histogram of CTI, (*b*) spatial display of CTI for study catchments and (*c*) CTI variogram.

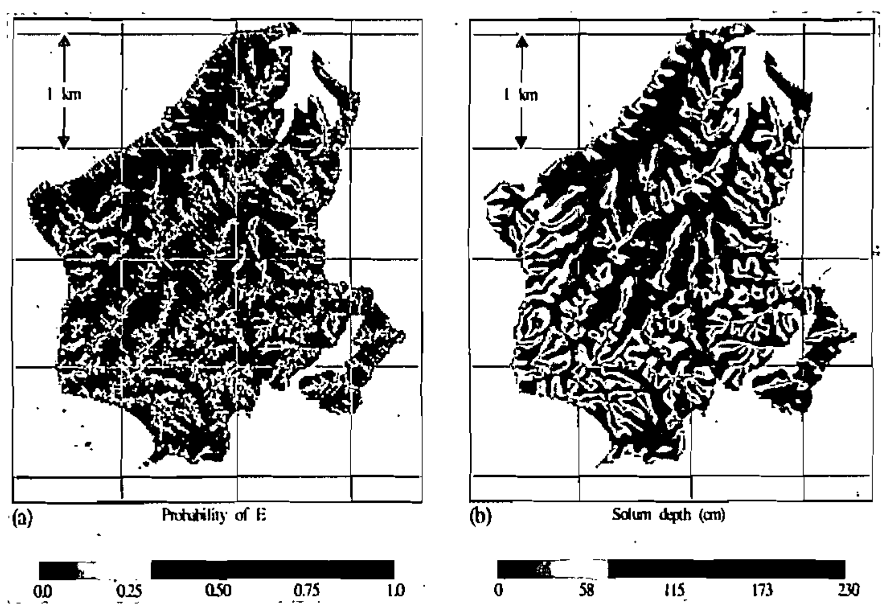

Figure 3. (*a*) Probability of E horizon and (*b*) predicted solum depth.

spatial dependence in the CTI. The variogram (Webster and Oliver 1990) is a common method of quantifying the spatial dependence structure of a regionalized variable (Matheron 1971). Figure 1 (*c*) shows the computed variogram for the CTI cells on the Ordovician metasediments. This variogram shows a range (distance within which spatial dependence occurs) of approximately 500 m. This suggests that statistical independence can best be maintained by spacing samples 500 m or more apart. An assumption is that the spatial dependence is stationary across the landscape. Subsequent sampling may be useful at nested scales within this distance to develop a useful understanding of short-range variation for individual soil attributes. Short range variation was not of primary interest and will not be discussed in this paper.

3.3.3. *Location of sample sites*

Accurate location of field sample points allocated using a geographical information system (GIS) is critical to the development of robust statistical models. To minimize locational errors, samples were located only in attribute patches with a minimum size of 3 × 3 grid cells (0·36 ha). This was accomplished by using a two cell erosion and dilation procedure to eliminate thin areas and small patches.

3.4. *Sample site allocation and data collection*

Sites were allocated in two batches of 30 samples. Six samples were distributed in each CTI percentile class according to the following iterative scheme. The patches for each class were numbered from 1 to *n* (total number of patches for percentile class). A random number generator was used to produce a random number vector of length *n*. Sites were selected sequentially from randomly selected patches and Australian Map Grid coordinates produced for each site. Sites within 500 m of previously selected sites were discarded and the next random patch selected until six sites were allocated for each class. Each site was located in the field using a global positioning satellite (GPS) receiver. The slope, aspect, elevation and specific catchment area attributes for each site were output from the GIS and used in the field to refine site placement and ensure consistency. At each site a 71 mm diameter core was taken to a maximum depth of 2·3 m. The cores were described according to McDonald *et al.* (1990).

Diagnostic morphological attributes that characterize the soil layers were used for model development. These attributes were: A horizon depth, E horizon presence/absence, E horizon depth, mottle presence/absence, depth to mottles, A horizon clay percentage, B horizon clay percentage and solum depth (A + E + B horizon depths). Results pertaining only to A horizon depth, solum depth and the probability of encountering an E horizon are presented to demonstrate the methodology. The A horizon depth is a general guide to nutrient status of soils in the study area and also an indicator of surface stability to erosional and depositional processes. An E horizon is indicative of downward or lateral percolation and leaching processes and periods of water logging. This has an impact on biological productivity and trafficability. Solum depth provides an indication of the available water capacity, and also exerts a major control on biological productivity.

3.5. *Exploratory data analysis and statistical model development*

A matrix of scatter plots (Cleveland 1993) was developed to identify patterns or structures within the data and to provide an indication of soil and terrain attribute correlations. Figure 2 (*a*) shows a scatter plot of solum depth versus CTI and figure 2 (*b*) a box plot of upslope mean plan curvature versus E horizon presence (1) or

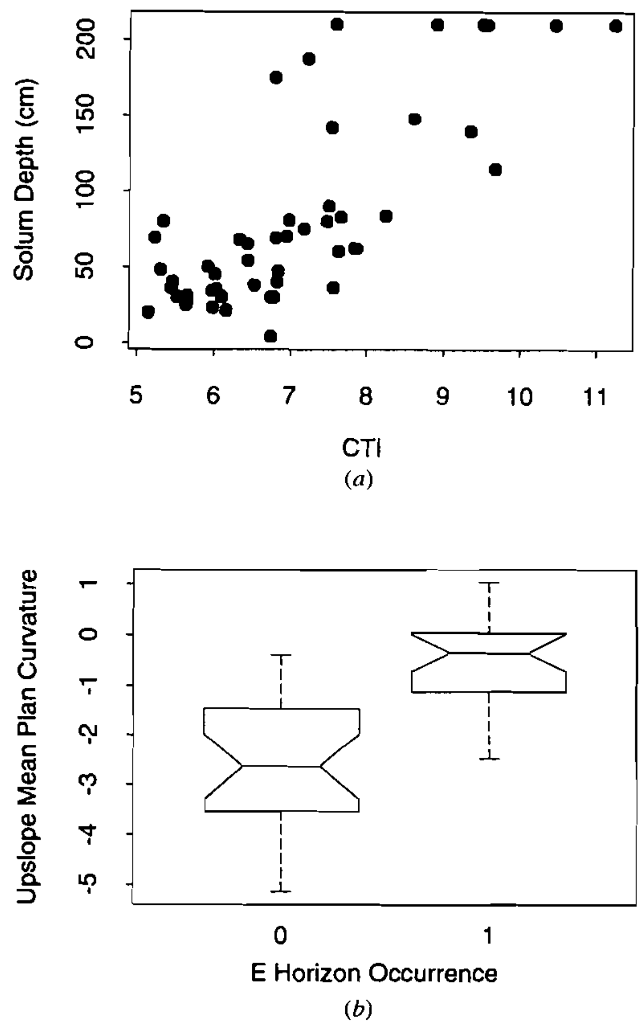

Figure 2. (*a*) scatterplot of solum depth versus CTI, (*b*) boxplot of upslope mean plan curvature versus absence (0) or presence (1) of an E horizon.

absence (0). This illustrates a simple visualization of relationships between soil (response) and terrain (explanatory) attributes that provided the first indication of predictive potential. This was followed by a stepwise exhaustive search technique (Statistical Sciences 1992) that considers possible subsets of explanatory variables based on the residual sum of squares. Statistical modelling was then performed using generalized linear models (McCullagh and Nelder 1989). Diagnostic methods of identifying outliers, influential observations and violations of model assumptions were used routinely (Cooke and Weisberg 1982).

Two types of generalized linear model were used. The first was a multiple regression with an identity link function and poisson error function. It is similar to a classical least squares multiple regression, except the poisson error function is specified, in this instance, because the variance increases with the fitted mean. The second type of model was used for predicting a binary response variable, in this instance, the probability of encountering an E or bleached horizon in the upper part of the soil profile. This generalized linear model uses a logistic link function and binomial errors and is often referred to as a logistic regression model. The proportion of variation accounted for by a logistic model cannot be expressed using a statistic analogous to R^2. Model adequacy is assessed in terms of the prediction errors and the reduction in residual deviance which is distributed approximately like χ^2 (McCullagh and Nelder 1989).

4. Results and discussion

The ubiquitous and substantial short range variation of soil attributes places a fundamental limit on the quality of spatial prediction. This issue has been avoided in traditional soil surveys by the delineation of somewhat qualitative and subjective map unit lines based on morphological soil types (McSweeney *et al.* 1994). Webster (1977) concluded that the variation accounted for by a typical general purpose soil survey would range from about half the total variance for soil physical attributes to less than one tenth for some soil chemical attributes. This provides an informal measure for judging the success of a statistical model. Statistical models that predict soil attributes using topographic attributes are presented in table 1. The percent reduction in deviance provides an indication of the proportion of response variability explained by the fitted model and is similar to the R^2 for multiple regression. The results in table 1 are encouraging because of the large reductions in deviance accounted for by the fitted models.

As expected, CTI was a useful predictor because it combines contextual and site information via the upslope catchment area and slope, respectively. Plan curvature was not expected to have a strong predictive power because it does not include contextual information. However, it was significant in predicting the A horizon and solum depth in combination with CTI. This suggests that local scale pedogenic as well as hillslope scale processes are influencing soil profile development. Upslope mean plan curvature

Table 1. Regression equations for prediction of soil attributes (Standard errors are shown in parentheses)*.

Regression models	Reduction in deviance (%)
A horizon depth = 0·92 + 5·67 plancrv + 4·88 CTI	63%
SE (14·1) (1·4) (1·9)	
Solum depth = − 57·95 + 12·83 plancrv + 21·46 CTI	68%
SE (39·4) (3·9) (5·2)	

Logistic regression model

$\ln(p/(1-p)) = 2{\cdot}52 + 1{\cdot}68$ umplacrv

re-arranging gives:

p(E horizon + $= \exp(2{\cdot}52 + 1{\cdot}68\ \text{umplancrv})/(1 + \exp(2{\cdot}52 + 1{\cdot}68\ \text{umplancrv}))$

Analysis of deviance

Model	Deviance	Residual deviance	Df	Pr(Chi)
Null		69·31	49	
umplancrv	29·43	39·88	48	<0·001

* CTI = compound topographic index.
plancrv = plan curvature.
umplancrv = upslope mean plan curvature.
p = probability that an E horizon is present.

provided the best logistic model fit for the probability of an E horizon occurrence. This indicates that the overall upslope convergent and divergent flow processes may control E horizon development. The next best logistic fit was provided by CTI, which in part, measures some of the same types of landscape processes as upslope mean plan curvature. Figure 3 displays the spatial extension of the logistic model for E horizon presence/absence (figure 3 (*a*)) and the regression model for solum depth (figure 3 (*b*)) for the study catchments.

The advantage of this form of mapping over conventional methods is that individual soil attributes rather than soil types are predicted with a specified accuracy and precision. Assumptions of high covariance between soil attributes, implicit in the mapping of traditional soil types, are avoided. The sampling procedure used here also enables the exploration and identification patterns in the data that may relate to process thresholds in the landscape. Subsequent quantitative delineation of process zones (e.g. zones of net erosion) can be used for land management planning.

5. Conclusions

We began with a provisional pedologic model where CTI was hypothesized to be a strong controlling variable and designed our sampling plan accordingly. The field data supported this assertion and provided evidence of other useful explanatory variables. The identification of plan curvature and upslope mean plan curvature as useful predictors demonstrates a key feature of our methodology. Models are proposed and then tested. During the testing phase, new hypotheses of landscape processes controlling soil distribution are formulated and these may be tested to further improve our capacity for spatial prediction. In conventional surveys, this process is undertaken in the minds of surveyors as they traverse a region and develop mental and sometimes verbal models for spatial prediction.

Our long-term goal is to develop a quantitative and statistical analogue to the conventional method that is explicit, consistent and repeatable. Evidence is not confused with interpretation and models can be communicated in an objective way. At present, a large body of knowledge is trapped within the minds of soil surveyors and is eventually lost. Our procedure meets with Hewitt's (1993) demands for a scientific rather than subjective procedure for developing explicit and quantitative soil-landscape models for spatial prediction. These methods provide a basis for understanding soil-landscape processes and may be integrated with other spatial interpolation techniques such as kriging and splines (Hutchinson and Gessler 1994). Information about scale (Moore *et al.* 1994) and error (Burrough in press) must also be incorporated in an explicit fashion.

Acknowledgments

The authors thank Linda Ashton, John Hutka and Chris Moran for assistance. This study was funded in part by Grant No. NRMS–M218 from the Murray-Darling Basin Commission and by the Water Research Foundation of Australia. Professor Ian Moore passed away before the completion of this paper. He was a friend, mentor and colleague who will be dearly missed. We will endeavour to carry on his spirit and work.

References

Allen, F., and Starr, T. B., 1982, *Hierarchy: Perspectives for Ecological Complexity.* (Chicago: University of Chicago Press).

Allison, R. J., 1991, Slopes and slope processes. *Progress in Physical Geography*, **4**, 423–437.

BAIZE, D., and GIRARD, M. C., 1992, *Pedological reference base: main soils of Europe, Referentiel pedologique, Principaux sols d'Europe* (Versailles, France: INRA).

BECKETT, P. H. T., and WEBSTER, R., 1971, Soil variability: a review. *Soils and Fertilizers*, **34**, 1–15.

BOUMA, J., 1989, Land qualities in space and time. In *Land qualities in space and time, Proceedings of a symposium organized by the International Society of Soil Science, Wageningen, The Netherlands* (Wageningen: Pudoc), pp. 3–13.

BURROUGH, P. A., 1993, Soil variability: a late 20th century view. *Soil and Fertilizers*, **5**, 529–562.

BURROUGH, P. A., in press, Spatial data quality and error analysis issues: GIS functions and environmental modelling. *Proceedings of the 2nd International Conference on Integrating Environmental Modelling and GIS* (Boulder: GIS World).

BUTLER, B. E., 1964, Can pedology be rationalized: a review of the general study of soils. Presidential address. Publication 3, Australian Society Soil Science, CSIRO Division of Soils, Canberra, Australia.

CHAMBERS, J. M., and HASTIE, T. J., 1992, *Statistical Models in S.* (California: Wadsworth & Brooks).

CLEVELAND, W. S., 1993, *Visualizing Data.* (New Jersey: Hobart Press).

COOKE, R. D., and WEISBERG, S., 1982, *Residuals and Influence in Regression* (London: Chapman & Hall).

FITZPATRICK, E. A., 1993, Introduction: soil horizons. *Catena*, **20**, 361–362.

GESSLER, P. E., MCSWEENEY, K., KIEFER, R., and MORRISON, L., 1989, Analysis of contemporary and historical soil/vegetation/land use patterns in southwest Wisconsin utilizing GIS and remote sensing technologies. *Proceedings of the ASPRS/ACSM Annual Convention* (Falls Church, VA: ASPRS/ACSM), pp. 85–92.

GOODCHILD, M. F., 1986, Spatial autocorrelation. *CATMOG—Concepts and techniques in modern geography* (Norwich: Geo Abstracts).

HEWITT, A. E., 1993, Predictive modelling in soil survey. *Soils and Fertilizers*, **3**, 305–314.

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.

HUTCHINSON, M. F., and GESSLER, P. E., 1994, Splines: more than just a smooth interpolator. *Geoderma*, **62**, 45–67.

KACHANOWSKI, R. G., 1988, Processes in soils—from pedon to landscape. In *Scales and Global Change* edited by T. Rosswall (New York: Wiley and Sons), 153–177.

MALANSON, G. P., BUTLER, D. R., and WALSH, S. J., 1990, Chaos theory in physical geography. *Physical Geography*, **11**, 293–304.

MATHERON, G., 1971, *The theory of regionalized variables and its applications. Les Cahiers du centre de morphologie mathematique de Fontainebleu. Ecole Nationale Superieur des Mines de Paris.*

MCBRATNEY, A. B., and WEBSTER, R., 1983, Optimal interpolation and isarithmic mapping of soil properties. V. Co-regionalization and multiple sampling strategy. *Journal of Soil Science*, **34**, 137–162.

MCCULLAGH, P., and NELDER, J. A., 1989, Generalized linear models. *Monographs on Statistics and Applied Probability* No. 37. (London: Chapman & Hall).

MCDONALD, R. C., ISBELL, R. F., SPEIGHT, J. G., WALKER, J., and HOPKINS, M. S., 1990, *Australian Soil and Land Survey Field Handbook.* Second Edition. (Sydney: Inkata Press).

MCKENZIE, N. J., 1991, *A strategy for coordinating soil survey and land evaluation in Australia.* Divisional Report 114, CSIRO Division of Soils, Canberra.

MCKENZIE, N. J., and AUSTIN, M. P., 1993, A quantitative Australian approach to medium and small scale surveys based on soil stratigraphy and environmental correlation. *Geoderma*, **57**, 329–355.

MCSWEENEY, K. M., GESSLER, P. E., SLATER, B., HAMMER, R. D., BELL, J., and PETERSON, G. W., 1994, Towards a new framework for modelling the soil-landscape continuum. In *Factors of Soil Formation: A Fiftieth Anniversary Retrospective*, Special Publication 33 (Madison, WI: Soil Science Society of America), 127–145.

MILNE, G., 1935, Some suggested units of classification and mapping particularly for East African soils. *Soils Research*, **4**, 3.

MOORE, I. D., GRAYSON, R. B., and LADSON, A. R., 1991, Digital terrain modelling: review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, **5**, 3–30.

MOORE, I. D., GESSLER, P. E., NIELSEN, G. A., and PETERSEN, G. A., 1993, Soil attribute prediction using terrain analysis. *Soil Science America, Journal*, **57**, 443–452.

MOORE, I. D., LEWIS, A., and GALLANT, J. C., 1994, Terrain attributes: estimation methods and scale effects. In *Modelling Change in Environmental Systems*, edited by A. J. Jakeman, M. B. Beck and M. McAleer (London: Wiley), pp. 189–214.

RAYMOND, O. L., 1992, *Geology of the Wagga Wagga sheet, Sheet 8327, Australian Geological Survey Organization, Canberra, Australia.*

SIMONSON, R. W., 1959, Outline of a generalized theory of soil genesis. *Soil Science Society of America Proceedings*, **23**, 152–156.

STATISTICAL SCIENCES, 1992, *S-Plus Programmer's Manual.* Statistical Sciences Corporation, Seattle, Washington.

TUKEY, J. W., 1977, *Exploratory Data Analysis* (Reading: Addison-Wesley).

WAGENET, R. J., BOUMA, J., and HUTSON, J. L., 1994, Modelling water and chemical fluxes as driving forces of pedogenesis. In *Quantitative Modelling of Soil Forming Processes*, Special Publication 39 (Madison, WI: Soil Science Society of America), pp. 17–36.

WEBSTER, R., 1977, *Quantitative and Numerical Methods in Soil Classification and Survey* (Oxford: Clarendon Press).

WEBSTER, R., and OLIVER, M. A., 1990, *Statistical Methods in Soil and Land Resource Survey* (Oxford: Oxford University).