

What is the spatial representation of digital soil maps? An issue of the spatial entity

Brendan Malone^A, Alex McBratney^A and Budiman Minasny^A

^AFaculty of Agriculture Food & Natural Resources, The University of Sydney, NSW 2006, Australia, Email b.malone@usyd.edu.au, a.mcbratney@usyd.edu.au, b.minasny@usyd.edu.au.

Abstract

The spatial entity used within a digital soil mapping framework can have profound implications for data reliability and interpretability. We compare the quality of results and maps of the variability of available water capacity (AWC) generated by both the classical point prediction approach and an innovative method that involves a block estimator. Rather than a single prediction point in the centre of a pixel, the block estimator method takes the weighted average of the soil attribute across the entire extent of a pixel through a block kriging approach. The results indicate AWC varies continuously across the landscape. However, maps produced by our block estimator method result in predictions and maps which are more fluid and smoothed reflecting a more realistic representation of the variability.

Key Words

Digital soil mapping, block kriging, available water capacity.

Introduction

The pixel model is an efficient method in which to display the spatial variability of soil properties and classes in a map format (Grunwald 2006). It is generally assumed that the value predicted for a soil property or class at a pixel is the same at every point within the extent of that pixel, no matter the resolution. In actuality, the predicted soil value is a single point located at the centre of the pixel (Figure 1). Thus the short range soil variability that may occur within a pixel will not be accurately addressed. A more appropriate method in which to assign a predicted soil value to a pixel is by way of a bulked mean within the entire extent of the pixel.

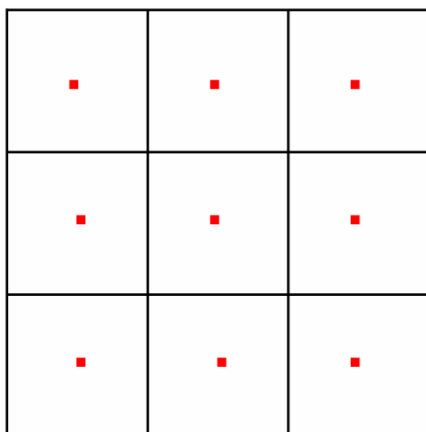


Figure 1. Representation of the pixel model and associated point prediction in the centre of the pixel.

For this study we demonstrate a method in which a mean prediction of soil properties within a pixel can be evaluated by using an existing point predicted digital soil map. This method involves the use of block kriging, where the kriged value represents a statistically weighted average of the entire extent of the pixel (Whelan *et al.* 2001). Using available water capacity (AWC) as an exemplar soil property, we compare and contrast maps produced by both methods i.e. point predictions vs. bulked mean predictions.

Methods

Study area

The study site (1500km²) is situated near Narrabri (30.32S 149.78E), 500km NNW of Sydney, NSW, Australia. Agricultural enterprises such as cropping and pastoral farming are predominant in the area (Figure2). The soil dataset consists of 341 soil profiles (Figure 2). The dataset describes and quantifies

various soil morphological, physical and chemical attributes at depth intervals of 0–0.1, 0.1–0.2, 0.3–0.4, 0.7–0.8, 1.2–1.3 and 2.5–2.6m (McGarry *et al.* 1989). The focus of this study was the prediction of AWC in the top 10cm of the soil profile; therefore we were only concerned with measurements in the 0–10cm depth increment. Using sand, clay and organic matter as inputs (McGarry *et al.* 1989) a pedo-transfer function generated estimates of AWC at each sample point (Minasny *et al.* 2006).

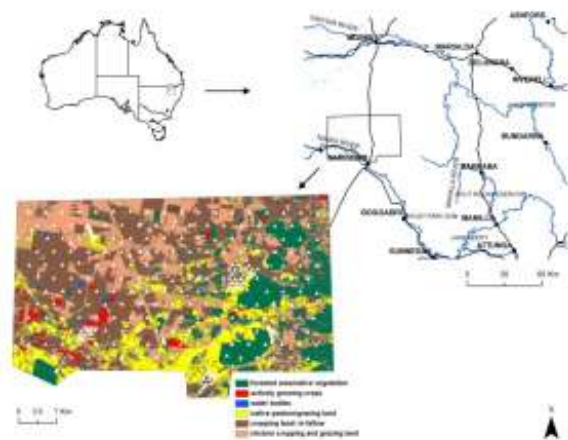


Figure 2. The Edgeroi study area.

Environmental data

A number of environmental covariates were sourced and interpolated onto a common grid of 90m resolution. These included:

- 3 arc-second (90m) digital elevation model (DEM). First and second derivatives, namely: slope, terrain wetness index (TWI), flow path length, altitude above channel network (AOCN) and multi-resolution index of valley bottom flatness (MRVBF) were determined.
- Landsat 7 ETM+ images from 2003. The Landsat bands were used for the approximation of land cover and land use. Vegetation cover and type was estimated using the Normalised Difference Vegetation Index (NDVI). Furthermore, the band ratios or more commonly, soil enhancement ratios of b3/b2, b3/b7 and b5/b7 were also derived.
- Gamma-radiometric survey data which provides a measure of the spatial distribution of three radioactive elements (potassium-K, thorium-Th and uranium-U) in the top 30–45 cm of the earth's crust. This data was used to approximate the distribution of various parent materials over the landscape.

Data analysis

We used a regression kriging approach to predict AWC at each sample point and across the study area. A neural network was constructed on a training data set of 261 points (leaving 80 for validation). The input variables were the various layers of environmental covariates. The target variable was the measured value of AWC at each training point. After the modeling process, residuals were evaluated for each point. A semi-variogram was used to assess the spatial distribution of residuals. We used an exponential model to krig the residuals onto the common 90m grid of the Edgeroi. For model validation, the profile formulae were applied to the 80 withheld data points. Residuals, estimated from the semi-variogram model of the training step residuals were added to the prediction resulting in a final prediction.

For mapping point estimates of AWC, profile formulae from the neural network were applied to the common 90m grid geo-database where only information relating to the environmental data existed to make predictions of AWC. A final prediction was determined by adding the kriged residual to the prediction at each point. A point predicted map of AWC was the resulting product of this procedure.

The final predictions of the point estimates of AWC were used as inputs for the block mean estimation. As a first step to the block estimation method is the realization uncertainty of the prediction at each point or pixel. The residual estimate is a good indication of this. As we could quantify this element of uncertainty, we incorporated it within the block kriging procedure by defining it as the σ^2 parameter. The σ^2 parameter was calculated by evaluating the variance of all the residuals across the study area. Secondly, the common grid was off-set by 1m to the left from the original grid. With the incorporation of the σ^2

parameter, block kriging was performed onto the offset grid. The block size used had equal-side dimensions of 90m. For each prediction via block, local exponential variograms were used. A block estimated predicted map of AWC was the resulting product of this procedure.

Results

For neural network training we used a 3 hidden node network. Lin's concordance correlation coefficient (CCC) between the observed and predicted AWC values was 0.61, indicating a substantial agreement along the 45° line (Lin 2000). Validation results were less impressive where a fair agreement (CCC = 0.27) was observed between the observed and predicted (Figure 3a). There was very little spatial autocorrelation of residuals beyond a separation distance of 83m (Figure 3b) which resulted in only a minor improvement in the final prediction where the CCC was 0.34 indicating still some significant deviations from the 45° line (Figure 3c).

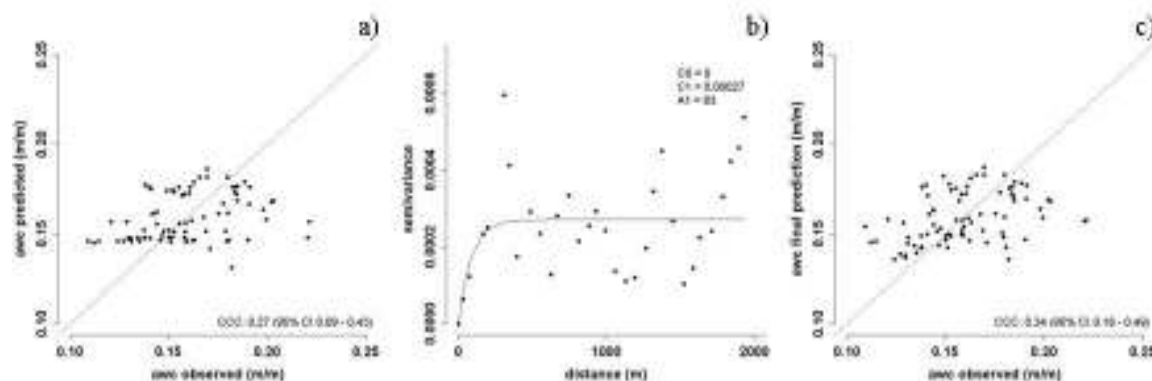


Figure 3. Validation results. (a) Observed vs. predicted plot (b) semi-variogram of residuals (c) Observed vs. final prediction plot.

The model results indicate there was some correlation between the target variable AWC and the available environmental data. This is shown in Figure 4a where AWC varies significantly across the study area. Some notable patterns exist such as a generally higher to lower gradient of AWC in the east to west direction. This coincides with an increasing proportion of land used for cropping as one moves in a westerly direction across the study area. Additionally the topography itself also becomes more open and flatter in the western area in comparison to the more undulating land features in the east.

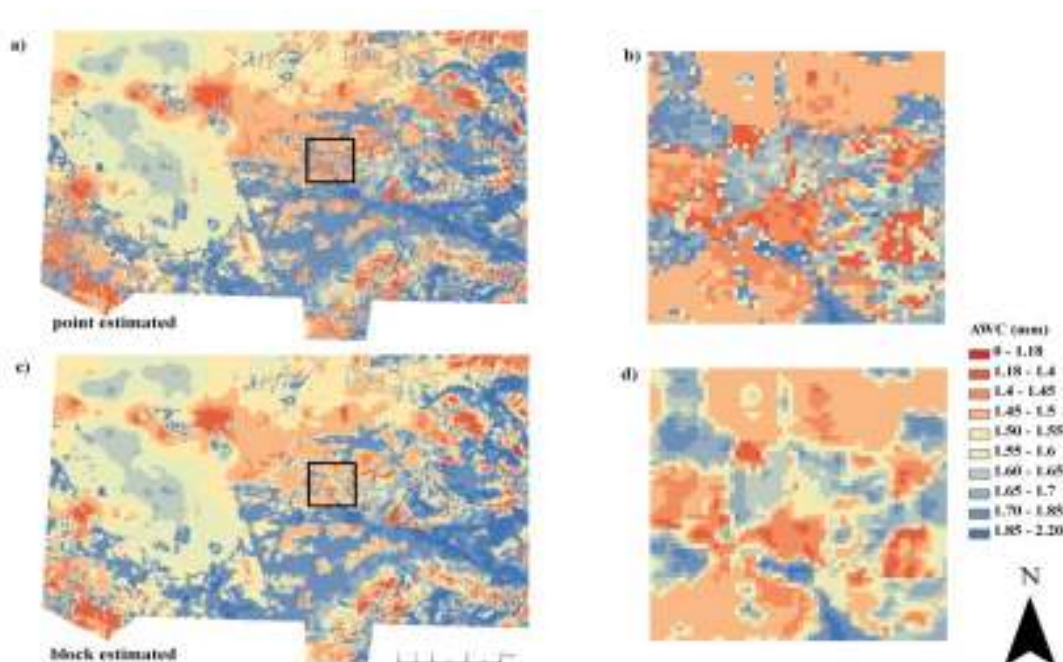


Figure 3. Comparisons of maps of AWC variability in the 0–10cm depth range generated from point predictions (a) and block estimated predictions (c). Zoomed in zone of 5km² of predictions made from points (b) or blocks (d).

The result of the block estimation method was overall very positive due to the fact that the visual appearance of the map is a lot smoother and clearer for interpretation than that of the point estimated map (Figure 3c). Broadly, the spatial variation of AWC changes very little when comparing both maps. The key difference is that in areas where there is significant local variation, the point predictions appear quite noisy in comparison to the block estimates where there is a more gradual transition in predicted values across an area. This phenomenon is illustrated where we selected a 5km² zone within the study area that displayed what appeared to be a high degree of spatial variability. Overall, it is difficult to separate both methods in terms of the general variation of AWC in the zone. However, for the point estimated predictions (Figure 3b), the zone is clearly pixilated or noisy. On the other hand, for the block estimations (Figure 3d), interpretation of AWC variability is easier to define as the noisiness has been smoothed resulting in a clearer map.

Discussion and conclusions

Block estimations of AWC across the Edgeroi significantly improved the visual quality and interpretability of the generated map. The procedure had the effect of smoothing out some of the noisiness that results from point predictions resulting in a more realistic depiction of AWC across the Edgeroi area. In a general context, the spatial entity of a soil map has significant implications for the interpretability of soil information. Essentially a pixel value of a soil property is a single observation located at the centre of that pixel. It is problematic to assume this value is the same for the entire extent of the pixel. Thus a more appropriate method for making predictions within a pixel is the incorporation of a block estimator in which makes a prediction that represents a statistically weighted average across the entire extent of a pixel. The result is that collectively, predictions are more continuous from one pixel to the next. This improves the quality of predictions, enhances the interpretability of the resulting soil map and realistically illustrates the spatial variation of soil properties across a defined area. As this method is relatively simple to implement, this method could easily be appended to the methodologies of other DSM projects.

References

- Grunwald S (2006) What do we really know about the space-time continuum of soil-landscapes. In 'Environmental Soil-Landscape Modeling'. (Ed. S Grunwald) pp. 3–36. (CRC Press: Boca Raton, FL).
- McGarry D, Ward WT, McBratney AB (1989) 'Soil studies in the Lower Namoi Valley: methods and data. The Edgeroi Dataset.' (CSIRO Division of Soils: Adelaide).
- Minasny B, McBratney AB, Mendonca-Santos ML, Odeh IOA, Guyon B (2006) Prediction and digital mapping of soil carbon storage in the Lower Namoi Valley. *Australian Journal of Soil Research* **44**, 233–244.
- Whelan BM, McBratney AB, Minasny B (2001) Vesper – Spatial prediction software for precision agriculture. In 'ECPA 2001. Third European Conference on Precision Agriculture.'. Agro Montpellier, Ecole Nationale Agronomique de Montpellier. (Eds G Grenier, S Blackmore) pp. 139–144.