

DIGITAL ELEVATION MAPS FOR SPATIAL MODELLING OF SOIL SERVICES

Carolyn Hedley, Pierre Roudier, Leo Valette

Landcare Research, Palmerston North
hedleyc@landcareresearch.co.nz

Abstract

Valuable soil information can be obtained from digital elevation maps because the derived terrain attributes improve the spatial prediction of key soil attributes (e.g. topsoil depth and organic carbon) that impact on soil function and the services soils provide. Zones of higher soil organic carbon delineated on maps suggest areas of enhanced water storage, structure, cation exchange capacity, nitrogen stores, and dissolved organic carbon, which in turn attenuates nitrate leaching to freshwaters.

While the publically available national 25-m resolution digital elevation map is suited to catchment-scale studies, recent technological developments provide digital elevation maps at much higher resolutions ($\leq 10\text{m}$).

Digital elevation maps derived from electromagnetic and LiDAR (Light Detection And Ranging) surveys are at the appropriate scale for paddock- and farm-scale spatial prediction of topography-controlled soil properties. Soil electromagnetic surveys can provide digital elevation maps at an approximate 10-m resolution, and airborne LiDAR surveys provide digital elevation maps at $\leq 1\text{-m}$ resolution.

Primary terrain attributes, such as slope and aspect, and secondary terrain attributes, such as flow accumulation pathways, can be quantitatively derived from digital elevation maps. For example, a multiple flow direction algorithm assesses upslope area for any one pixel and allows this accumulated upslope flow to be distributed among all downslope directions.

The derived datalayers can be used to select sampling positions (stratified sampling) and for digital soil mapping of selected soil attributes. These methods are applied at two case study sites:

An electromagnetic survey-derived digital elevation map is used to investigate topsoil depth on 75 ha of irrigated rolling downlands, in North Otago. Topsoil depths ranged between 0 and 76 cm, tending to increase downslope, and a random forest model was used to predict the topsoil depth (R^2 0.5; RMSE 9 cm).

A LiDAR-derived digital elevation map is used to produce a preliminary soil organic carbon map for Massey University Tuapaka hill country farm (476 ha). Soil organic carbon stocks ranged between 42 [at lowest elevations] and 194 [at highest elevations] T/ha to 30 cm soil. Model predictions for the whole farm (calibration set [$n=35$]: r^2 0.91; RMSE 10.6 t/ha; validation set [$n=15$]: (r^2 : 0.16; RMSE 21.7 t/ha) are preliminary because, to date, only one subcatchment of the farm has been sampled.

Introduction

New forms of spatial data are increasingly becoming available to assist high resolution (i.e. <10-m resolution; “paddock-scale”) spatial modelling of soil attributes. This refines previous methods that used lower resolution maps, for example, the national 25-m resolution digital elevation map (available from <https://lris.scinfo.org.nz/>).

Key soil attributes such as topsoil depth and organic matter content affect the services the soil provides. For example, a paddock-scale soil organic carbon (SOC) map indicates areas of higher buffering capacity (high SOC) and lower buffering capacity (low SOC). High SOC zones are more able to interact with high nitrogen doses, e.g. from dairy cow urine, slowing the transformation of organic nitrogen through useful N sources for plant and microbial activity (ammonium) to soluble nitrates which potentially contaminate freshwater resources. The filtering and buffering properties of soils are linked to their position in the landscape, and a high resolution soil attribute map therefore indicates how soil services relate to topographic features. The ability to understand the pattern of soil variability and relate it to nitrate leaching risk is needed to inform and improve management of grazed pastoral soils (e.g. Lilburne & Webb, 2002). A map of topsoil depth can be similarly interpreted, because topsoils are defined as the soil horizon where organic matter visibly accumulates.

Examples of the new forms of data available for this spatial modelling include those derived from (1) ground-based electromagnetic (EM) sensors (Adamchuk, 2004; Hedley et al., 2013) and (2) airborne LiDAR surveys (Mulder et al., 2013).

EM sensors measure the apparent electrical conductivity (EC) of a soil, responding to soil textural and moisture differences. When mobilised with accurate Global Positioning System (GPS) survey equipment the data can be processed into soil EC maps, which quantify soil variability (Hedley et al., 2004; Doolittle & Brevik, 2014). The EM data points are georeferenced and so the elevation data, collected simultaneously, can be processed into a digital elevation map.

LiDAR is a laser-based device analogous to Radar, and the range (accurate distance) to an object is determined by measuring the time delay between transmission of a pulse and detection of the reflected signal. Airborne LiDAR units can survey several hundreds of hectares in one day. Ground-based devices also exist and are used for smaller scale applications, such as monitoring of unstable hillslopes. Typical positional accuracy of these devices is +/- 0.25 m and vertical accuracy is +/-0.15 m.

The most difficult part of digital terrain analysis is the creation of the digital elevation model (DEM) (Quinn et al., 1991). Poor datasets with irregular point data can cause artefacts and mismatches with other data sources, for example sink features, formed when data are too coarse and valley bottoms are not recorded, as well as dam features that occur with data are too coarse and small channel features missed, in some sections along their length. However LiDAR data provides very dense data on a regular grid (typically one pulse per square metre), which minimises the problems associated with DEM models.

A powerful feature of DEMs is being able to derive a wide range of terrain attribute layers. The most basic and commonly used primary terrain attributes include surface derivatives such as slope, aspect, and curvature. Secondary terrain attributes are calculated from a combination of two or more primary terrain attributes to model spatial variation of processes across a landscape; the most commonly used being the ‘topographic wetness index’ (TWI), which is defined as the natural logarithm of a specific catchment area divided by the tangent

of the slope, and used for hydrological modelling (Moore et al., 1991). Similarly, the SAGA Wetness Index (SWI), based on a modified catchment area, provides a more sensitive prediction for sub-catchments, and is often preferred for paddock scale analyses (e.g. Bohner & Selige, 2006). The wetness indices (WI) are calculated using some form of the basic equation (1):

$$WI = \ln\left(\frac{A_s}{\tan(S)}\right) \quad (1)$$

where A_s is specific catchment area (m^2m^{-1}), and S is slope (degree).

The specific catchment area is the upslope contributing area per unit contour length. The number of cells that drain through any point is estimated using a single flow or multiple flow direction algorithm (Wolock & McCabe, 1995). The multiple flow algorithm is preferable because it distributes water flow to all neighbouring cells with lower elevation, using slope towards neighbouring cells as a weighting factor for proportional distribution (Holmgren, 1994).

Algorithms are also available to derive other terrain attributes, such as landform elements (Schmidt & Hewitt, 2004) and global solar irradiation. Global solar irradiation represents the annual cumulative received solar energy, summing daily direct, diffuse and reflected solar irradiation, which are computed from the DEM and the slope and aspect maps (Hofierka & Suri, 2002).

The DEM-derived terrain attribute layers can then be utilised for selection of the sampling positions (Minasny & McBratney, 2006; Roudier et al., 2012; Roudier & Hedley, 2013) and for soil attribute modelling. One type of spatial modelling for soil attributes is the regression tree classification of which random forests (Breiman et al., 1984) is one example. The tree structure is generated by partitioning the data recursively into a number of groups, each time maximising some measure of difference in the response variable in the resulting two groups. This method deals very well with nonlinearity. Using a decision algorithm, the tree model decides automatically the splitting variables, splitting points, and tree topology. These regression tree models have been widely adopted for predicting soil attributes, being relatively easy to interpret (McBratney et al., 2003).

This paper reports the development of a method for high resolution soil attribute mapping, using DEMs derived from an EM survey and airborne LiDAR data, and discusses application of the derived information.

Case study 1: 75 ha irrigated field in North Otago

The study area is an arable field, irrigated by a 490-m centre pivot, near Oamaru. The landscape is rolling hills, and the soils are Waiareka clay loam soils, derived from basalt rock, classified as Typic Vertic Melanic soils (Hewitt, 2011).

An EM survey provided the DEM (Fig. 1), and terrain attributes were extracted in GRASS GIS (Grass Development Team 2012). The terrain attributes (elevation, slope, flow accumulation pathways) and the SAGA wetness index were used with EM (to 0.5 m) and EM (to 1.0 m) for stratified sampling (Fig. 2) to assess available water-holding capacity (AWC) and topsoil depth. Nine sampling positions were derived from these covariate datalayers, following the procedure described by Roudier et al. (2012).

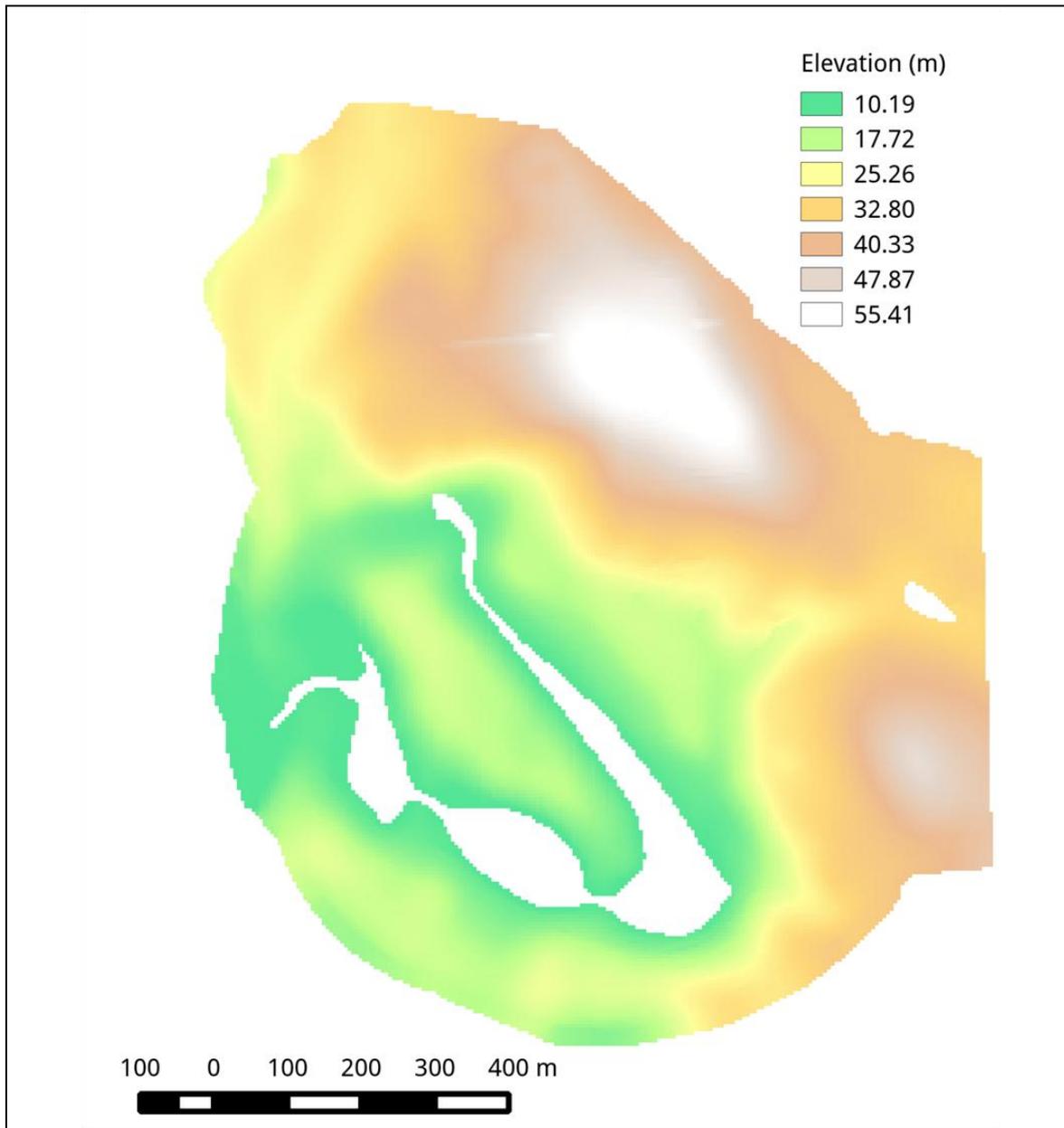


Figure 1 DEM derived from the EM survey for an irrigated field in North Otago rolling downlands.

Additional topsoil depth data ($n=50$) was collected by soil augering selected downslope transects. The spatial prediction model was derived using a random forest data mining method with the covariate datalayers and collected measured data points (Fig. 3).

Ninety percent of the measurements are used in a training set and ten percent for an independent test set. The independent test sets were randomly removed through an iterative process. This 10-fold cross validation procedure was repeated 30 times. The random forest data mining method is a machine learning algorithm which infers rules between the point estimates (in this case topsoil depth) and the values of the covariates observed at the sampling positions using boosted regression techniques.

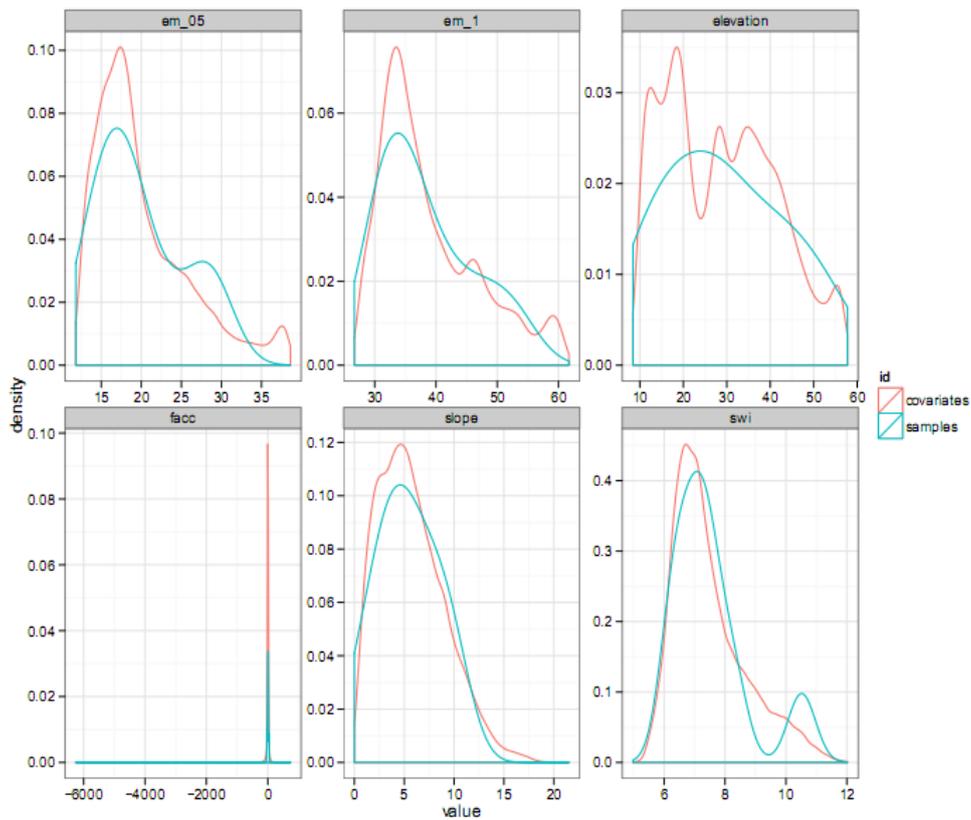


Figure 2 Population distribution of the covariates (red line) and of the subset of sampling positions (blue lines) obtained using a Latin hypercube method to select sampling (see e.g. Roudier et al. (2012) for explanation).

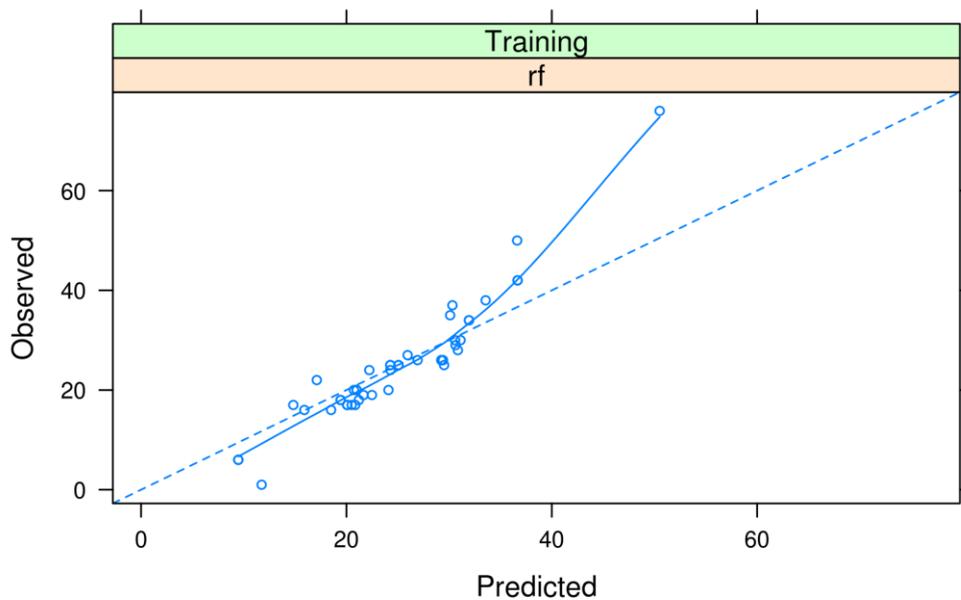


Figure 3 Relationship between the observed and predicted values of topsoil depth (cm), derived from data mining modelling process.

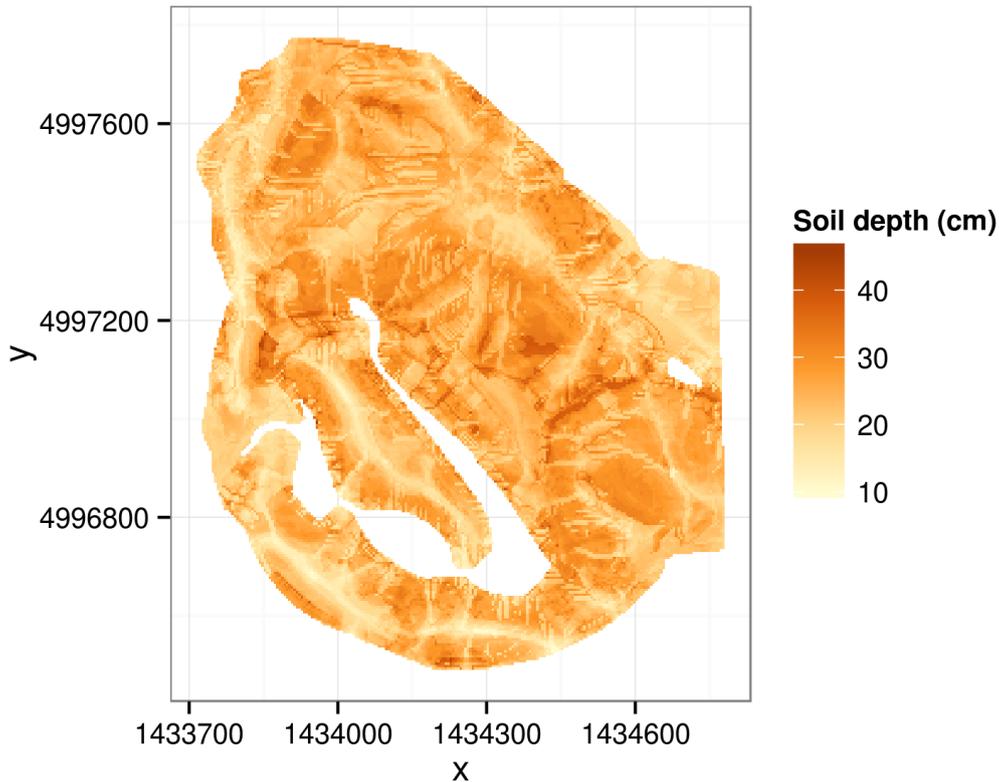


Figure 4 Map of topsoil depth for a North Otago 75 ha field, irrigated by a variable rate centre pivot.

The resulting map (Fig. 4) shows that topsoil depths typically varied between 10 and 50 cm, and confirmed and delineated field observations of topsoil depth increase down slopes. Actual observation of topsoils were that topsoil was absent in some summit positions and as deep as 76 cm in some valley bottoms (mean = 25 cm).

Case study 2: 470 ha Massey University Tuapaka Farm

The study area is Massey University Tuapaka Farm, situated 15 km north-west of Palmerston North. The 470-ha farm is characterised by relatively flat terraces at lowest elevations (50–100 m), and rolling to very steep slopes, rising to the summit of the Tararua Ranges (360 m elevation). The highest elevation of the farm is 360 m. The soils are mapped as Tokomaru soils on the terrace soils, Halcombe hill and steepland soils, Makara Steepland soils, Shannon and Tuapaka soils, Korokoro soils on the sloping land, with Ramiha soils on the summit (Pollok & McLaughlin, 1986).

An airborne LiDAR survey was commissioned for Tuapaka Farm. A 5 m DEM was derived from the resultant point cloud from the LiDAR survey, using SPDlib software (Bunting et al., 2013). Figure 5 is a hillshade visualisation of the digital elevation map.

Terrain attributes were derived from the DEM (Fig. 6). These include:

- Slope
- Aspect

- SWI
- Landform elements: valley pit, flat, peak, ridge, shoulder, spur, slope, hollow, and footslope
- Global solar irradiation, (using slope and aspect).

Additional covariate datalayers used were parent material (four classes from an existing soil map) (Pollok & McLaughlin, 1986), rainfall, and a mean water balance. The mean water balance was calculated for different altitudes (from Land Environments of New Zealand (LENZ); Leathwick et al., 2002) using mean daily temperature, mean daily solar radiation, and mean rainfall, with a 100-m resolution.

The ten covariate datalayers (four examples provided in Fig. 6) were then used to select fifty sampling positions, from lowest to highest elevation, for one sub-catchment on the farm, in which the main farm track is located. Positions were derived using an operationally constrained conditioned Latin hypercube method (cLH) (Roudier et al., 2012; Roudier et al., 2013). The operational constraint adds a weighting index for ease of access, and the cLH statistically selects a subset of positions with the same population distribution as the total population.

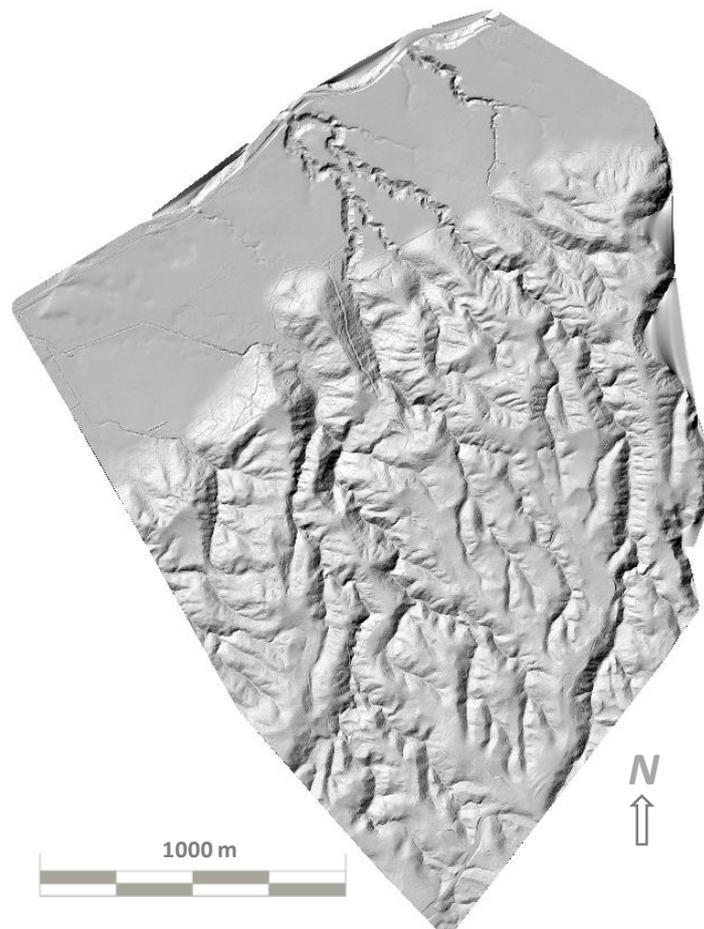


Figure 5 Hillshade visualisation of Massey University hill country farm Tuapaka, created from the LidAR data derived digital elevation model (5-m resolution) showing a terrace surface bordering the Manawatu River floodplains in the north, and the dissected hill slopes of the Tararua Ranges in the south.

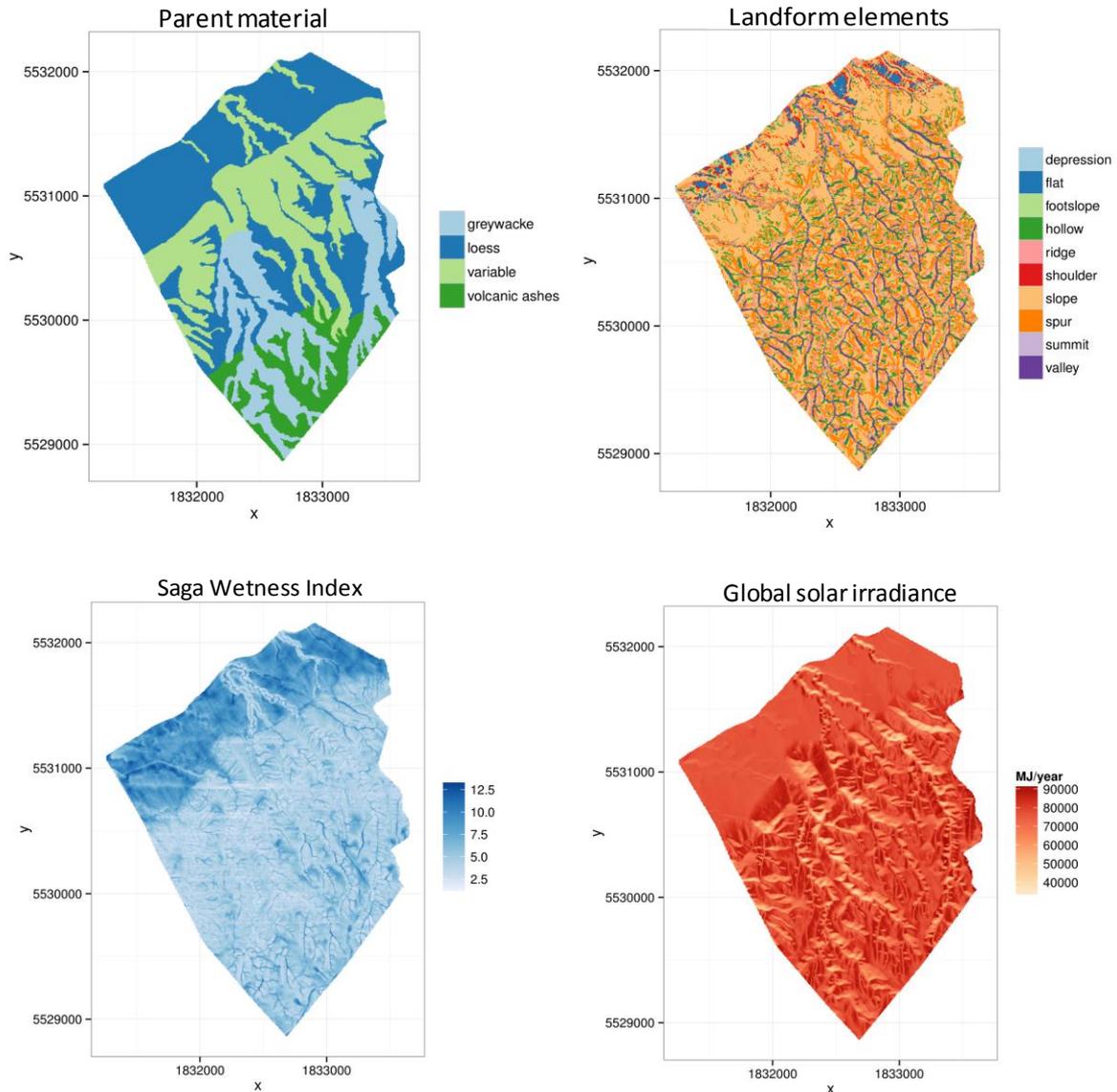


Figure 6 Examples of four of the covariate datalayers extracted from the DEM: parent material, landform elements, Saga Wetness Index and global solar irradiance.

Soil samples were collected from these fifty positions to measure volumetric SOC stocks to 30 cm depth from the soil surface. This dataset was then used to calibrate a spatial model using a random forest regression tree algorithm. The algorithm infers rules between SOC measurements and the values of the covariates observed at the sampling positions. The relative importance of the selected covariates was rainfall, elevation, water balance, parent material, TWI, solar irradiation, slope, aspect, SWI, and landform elements.

The fifty measurements indicate that SOC stocks range from 40 t/ha (to 30 cm soil depth) at lowest elevations, to 190 t/ha (to 30 cm soil depth) at the highest elevations where Ramiha soils occur, and the spatial prediction for the farm is shown in Figure 7. Model predictions for the whole farm (calibration set [$n=35$]: r^2 0.91; RMSE 10.6 t/ha; validation set [$n=15$]: (r^2 : 0.16; RMSE 21.7 t/ha) are preliminary because, to date, only one subcatchment of the farm has been sampled.

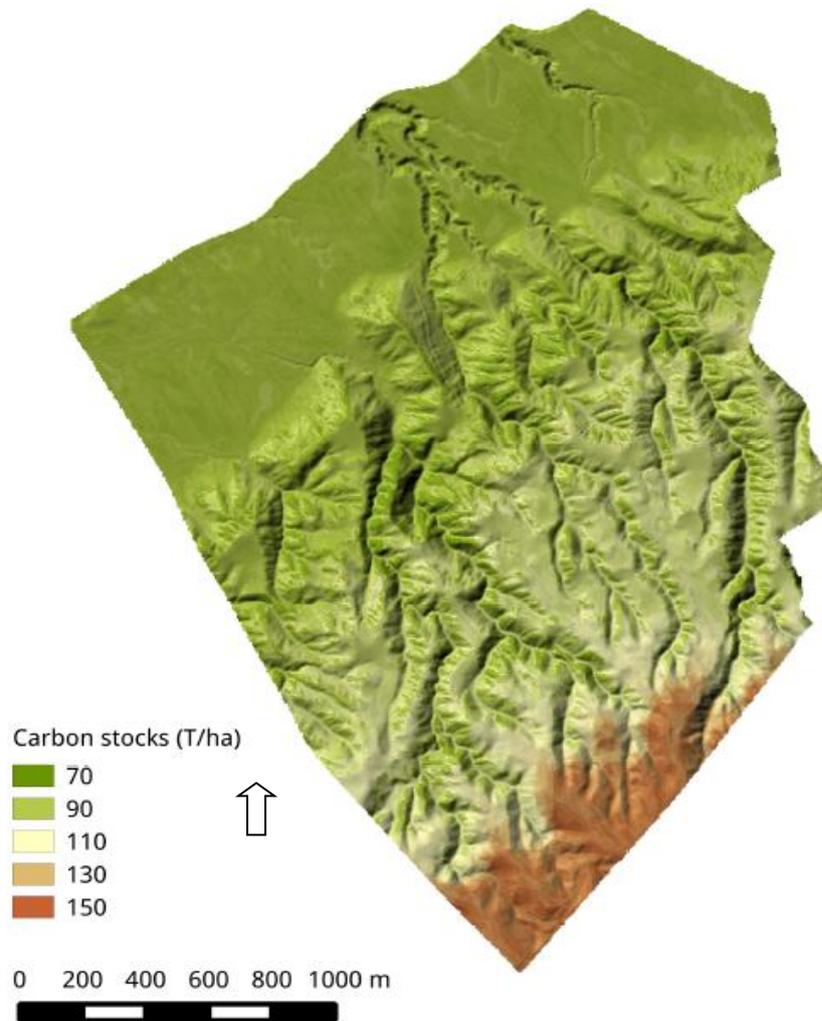


Figure 7 Digital map of soil organic carbon stocks blended into the hill shade representation of the digital elevation map (Valette, 2013).

Discussion

The topsoil depth map derived for the North Otago site, where a variable rate centre pivot irrigation system has been installed, can be used for further modelling exercises and as a management tool. The rolling topography drives the need for variable rate irrigation at this site, and topsoil depth has been found to generally increase with decreasing elevation. Topsoil depth is an easier soil attribute to measure in the field than other attributes, such as AWC, and is a major driver of AWC differences in the Waiareka clay loam soils, found on the site. Future work will use the topsoil predictions to refine spatial predictions of AWC, using the sparser dataset of measured AWC values ($n=9$).

The covariate datalayers can be used to derive management zones, and to determine positions for soil moisture monitoring. Real-time soil moisture monitoring in zones with known AWC can then be used spatially to predict daily soil moisture status (Hedley et al., 2013). This information can be used advantageously to control the precision irrigation system. In the first year of our trials at this site, irrigation was reduced to south-facing slopes and excluded from swampy areas in the valley floor area. Water savings were about 27%, equating to a financial

saving of \$216/ha (Swallow, 2013). This exemplifies the valuable application of this spatial modelling method to improve management of irrigated soils and landscapes, aiming for more efficient use of allocated freshwaters; and reduced drainage, to minimise the risk of contaminated waterways.

The SOC map derived for Tuapaka farm is also a useful management tool, being one indicator of soil quality. The effect of soil organic matter on soil properties is to increase water retention and availability and provide cation exchange capacity (CEC). CEC controls the ability of a soil to retain nutrients within the root zone. It is also a measure of the buffering capacity of a soil against pH change, the ability to chelate and complex ions, and to contribute to soil structure and form stable aggregates. Stable aggregates in soil sustain biological activity and biodiversity by providing food and habitat for soil animals and microorganisms (Sparling et al., 2006). Sparling et al. (2006) estimated that the extra organic matter in high C soils in their dairy farm study was worth NZ\$27–150 ha/yr in terms of increased milk solids production. This supports the case for the development of a refined method to assess SOC stocks in preparation for any future farm-scale greenhouse gas and carbon accounting system.

The SOC map can be used to derive a refined baseline carbon level for greenhouse gas emissions trading schemes. It helps localise the variables controlling soil carbon stocks, as well as to identify potential project locations for soil-based carbon sequestration opportunities. In addition it serves as an input into mechanistic simulation models (Minasny et al., 2012).

As the samples at the Tuapaka farm were only collected from one sub-catchment, the estimate is preliminary, and further field sampling is planned to refine the spatial model for the whole farm.

Our future research aims to use this map to derive one mean value of SOC stocks for the whole farm, with reduced uncertainty compared with traditional aspatial approaches. Traditional techniques for assessing SOC stocks at the farm-scale are very time-consuming and costly. Stratified sampling based on mathematical analysis of high resolution datalayers provides a robust quantitative method to assess SOC stocks for an area, with known uncertainty. Because the uncertainty is known, and a mathematical method exists to quantify the uncertainty, the spatial modelling method can be refined with added measurements, to obtain a required level of certainty.

Conclusions

This paper has reported an innovative method to map soil attributes at high resolution. Covariate datalayers were assembled, and these were mainly derived from DEMs. A statistical analysis of the covariate datalayers was used to select sampling positions, termed ‘stratified sampling’. The datalayers were used again to develop a spatial model to interpolate the measured values at fixed positions across the whole study area.

Digital elevation maps are fundamental to most forms of digital soil mapping. Many covariate datalayers can be extracted from DEMs to improve spatial models. High resolution DEMs derived from LiDAR are particularly powerful for farm-scale and paddock-scale studies of the spatial variability of soil attributes.

References

- Adamchuk, V.I., Hummel, J.W., Morgan, M.T., Upadhyaya, S.K. 2004. On-the-go soil sensors for precision agriculture. *Computers and Electronics in Agriculture* 48: 272-294.
- Breiman, L. 2001. Random forests. *Machine Learning* 45(1): 5-32.
- Bohner, J., Selige, T. 2006. Spatial prediction of soil attributes using terrain analysis and climate regionalisation. In: *SAGA analyses and modelling applications* (Eds. J. Bohner, K.R. McCloy and J. Dtrobl). Gottinger Geographische Abhandlungen 115: 13-28.
- Bunting, P., Armston, J., Clewey, D., Lucas, R. 2013. Sorted pulse data (spd) library. Part 1: A generic framework for lidar data from pulsed laser systems in terrestrial environments. *Computers and Geosciences* 56: 197-206.
- Doolittle, J.A., Brevik, E.C. 2014. The use of electromagnetic induction techniques in soils studies. *Geoderma* 223–225: 33-45.
- GRASS Development Team. 2012 *Geographic Resources Analysis Support Systems (GRASS 7) Software*. Open Source Geospatial Foundation.
- Hedley, C.B., Yule, I.J., Eastwood, C.R., Shepherd, T.G., Arnold, G. 2004. Rapid identification of soil textural and management zones using electromagnetic induction sensing of soils. *Australian Journal of Soil Research* 42(4): 389-400.
- Hedley, C.B., Roudier, P., Yule, I.J., Ekanayake, J., Bradbury, S. 2013. Soil water status and water table modelling using electromagnetic surveys for precision irrigation scheduling. *Geoderma* 199: 22-29.
- Hewitt, A.H. 2011. *New Zealand Soil Classification*. 3rd edn. Landcare Research Science Series No. 1. Lincoln, Manaaki Whenua Press. 136 p.
- Holmgren, P. 1994. Multiple flow direction algorithms for runoff modelling in grid based elevation models: an empirical approach. *Hydrological Processes* 8(4): 327-334.
- Hofierka, J., Suri, M. 2002. The solar radiation model for open source gis: implementation and applications. *Proceedings of the Open Source GIS-GRASS users conference*. Pp. 1-19.
- Leathwick, J., Morgan, F., Wilson, G., Rutledge, D., McLeod, M., Johnston, K. 2002. *Land environments of New Zealand: technical guide*. Hamilton, Waikato Regional council. 237 p.
- Lilburne, L., Webb, T. 2002. Effect of soil variability, within and between soil taxonomic units, on simulated nitrate leaching under arable farming, New Zealand. *Australian Journal of Soil Research* 40: 1187-1199.
- McBratney, A.B., Santos, M.M., Minasny, B. 2003. On digital soil mapping. *Geoderma* 117(1-2): 3-52.
- Mulder, V.L.; Bruin, S. de; Schaepman, M.E.; Mayr, T. 2013. The use of remote sensing in soil and terrain mapping: review. *Geoderma* 162 (1-2): 1-19. ISSN 0016-7061
- Minasny, B., McBratney, A.B., Malone, B.P., Wheeler, I. 2012. Digital mapping of soil carbon. *Advances in Agronomy* 118: 1-47.
- Minasny, B., McBratney, A.B. 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers and Geosciences* 32: 1378-1388.

- Moore, I.D., Grayson, R.B., Ladson, A.R. 1991. Digital terrain modelling: a review of hydrological, geomorphological and biological applications, *Hydrological Processes* 5: 3-30.
- Pollok, J., McLaughlin, B. 1986. *A user-friendly guide to the soils of Taupaka Farm*. Palmerston North, Massey University Tuapaka Farm Series Publication No. 3. 56 p.
- Quinn, P., Beven, K., Chevallier, P., Planchin, O. 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes* 5: 59-79.
- Roudier, P, Hedley, C.B. 2013. Smart sampling to assist on-farm nutrient management. In: *Accurate and efficient use of nutrients on farms* (Eds L.D. Currie and C L. Christensen). <http://flrc.massey.ac.nz/publications.html>. Occasional Report No. 26. Fertilizer and Lime Research Centre, Massey University, Palmerston North, New Zealand.
- Roudier, P., Beaudette, D.E., Hewitt, A.E. 2012. A conditioned Latin hypercube sampling algorithm incorporating operational constraints. In: *Digital soil assessments and beyond* (Eds. B. Minasny, B.P. Malone and A.B. McBratney). Boca Raton, FL, CRC Press, Pp. 227-231.
- Sparling, G.P., Wheeler, D., Vesely, E.T., Schipper, L.A. 2006. What is soil organic matter worth? *Journal of Environmental Quality* 35: 548–557.
- Schmidt, J., Hewitt, A., 2004. Fuzzy land element classification from DTMs based on geometry and terrain position. *Geoderma* 121: 243-256.
- Swallow, A. 2013. Simple steps to zone management. *Rural News*, 3 December 2013..
- Valette, L. 2013. Farm-scale mapping of soil organic carbon on a hill country farm using visible near infrared spectroscopy. Masters thesis submitted in partial fulfilment of the M.Sc. specialisation: AgroTic, Montpellier, SupAgro, France, September 2013. 52 p.
- Wolock, D.M., McCabe, G.J. 1995. Comparison of single and multiple flow direction algorithms for computing topographic parameters in TOPMODEL. *Water Resources Research* 31(5): 1315-1324.