Digital soil mapping of key soil properties over New South Wales

Version 2.0

Department of Planning and Environment

NSW GOVERNMENT

Cover photo: grazing lands of North West of New South Wales. Simone Cottrell/DPE

Preferred citation:

Gray JM (2023) *Digital soil mapping of key soil properties over New South Wales,* version 2.0, Technical Report, NSW Department of Planning and Environment, Parramatta.

Published by:

Environment and Heritage
Department of Planning and Environment
Locked Bag 5022, Parramatta NSW 2124
Phone: +61 2 9995 5000 (switchboard)
Phone: 1300 361 967 (Environment and Heritage enquiries)
TTY users: phone 133 677, then ask for 1300 361 967
Speak and listen users: phone 1300 555 727, then ask for 1300 361 967
Email: info@environment.nsw.gov.au
Website: www.environment.nsw.gov.au

Report pollution and environmental incidents
Environment Line: 131 555 (NSW only) or info@environment.nsw.gov.au
See also www.environment.nsw.gov.au

ISBN 978-1-923018-48-8
EHG 2023/0163
April 2023

Find out more about your environment at:

www.environment.nsw.gov.au

Summary

Digital soil maps (DSMs) have been prepared for a suite of key soil properties over NSW, updating those presented in OEH (2018). DSMs are maps derived through quantitative modelling techniques that are based on relationships between soil attributes and the environment. They are presented for soil organic carbon (SOC), pH, cation exchange capacity (CEC), sum-of-bases, available phosphorus (bray), bulk density, sand, silt and clay. The maps are at 100 m spatial resolution and cover multiple soil depth intervals down to 2 m, consistent with major Australian and international systems.

Random forest decision tree modelling techniques were applied. Validation results for the maps indicate overall moderate to strong but variable performance, with Lin's concordance values over 0.8 for some properties and depth intervals, but less for other property/depth combinations.

The maps provide at least a useful first approximation of these soil properties across the state. They are available for download as geotiff files through the NSW Government Sharing and Enabling Environmental Data (SEED) portal and may also be viewed directly through the Department of Planning and Environment's soil and landscape spatial viewer eSPADE.

These DSMs represent an alternative and complementary product to the Australia-wide DSMs presented within the Soil and Landscape Grid of Australia, being based on NSW data alone rather than nation-wide data. These new maps also complement existing conventional soil landscape products available for much of NSW.

Together with the existing maps, the new DSMs help inform on vital soil conditions across NSW and assist in the ongoing sustainable management and protection of our soil resources. They also provide valuable input data for a range of other natural resource and environmental modelling systems throughout the state.

Contents

List of tables

List of figures

Shortened forms

SCARP	Australian Soil Carbon Research Program
CEC	cation exchange capacity
DEM	digital elevation model
department, the	NSW Department of Planning and Environment
DSM	digital soil map, digital soil mapping
GIS	geographical information system
LCCC	Lin's concordance correlation coefficient
LLS	Local Land Services
P	phosphorus
MAE	mean absolute error
ME	mean error
MedAE	median absolute error
MER	monitoring, evaluation and reporting
MLR	multiple linear regression
MSE	mean square error
RF	random forest (decision-tree modelling technique)
RMSE	root mean square error
SALIS	Soil and Land Information System (NSW)
SEED	Sharing and Enabling Environmental Data (NSW environmental data portal)
SLGA	Soil and Landscape Grid of Australia
SOC	soil organic carbon

1. Introduction

Digital soil mapping (DSM) has become established over the last 2 decades as an important avenue for acquiring and presenting important soil information. DSMs are prepared through quantitative modelling techniques based on relationships between soil properties or classes and the environment. The underlying models have their roots in the fundamental soil equation of Dokuchaev (1899) and Jenny (1941): *s = f(cl, o, r, p, t, ...),* which states that soils are a function of climate, organisms, relief, parent material and time. More recently, this conceptual model has been further advanced with the '*s, c, o, r, p, a, n*' approach of McBratney et al. (2003), which has the additional factors of *s* (a soil attribute predictor) and *n* (a geographic position predictor), and also incorporates modelling of residual errors.

The DSM approach uses available environmental data to represent each of these soil-influencing factors, developing the relationships over known soil data points then extrapolating these relationships over broad regions using continuous environmental data grids (e.g. climate grids or digital elevation model grids).

DSMs have the potential to be a valuable complement to existing, conventional soil landscape mapping products over NSW. They can:

- provide estimates of specific soil properties or classes down to fine grid size (e.g. 100 m pixels), rather than broad polygons
- provide coverage for soil data for areas of the state with nil or only very broad-scale existing soil data
- provide estimates for soil properties that may not have been included in original soil surveys and laboratory analyses
- provide data in a spatial format more readily applied by quantitative environmental modellers, e.g. ecological, hydrological and climate change modellers.

However, conventional soil maps do retain many advantages over DSMs, including their more holistic approach to describing soil character, and the 2 mapping forms are best applied together. Further discussion on the relationship of DSM to conventional soil mapping is given in the 'Discussion' chapter (see Section 4.4).

1.1 Existing digital soil mapping over NSW

Regional and broader-scale DSMs have been undertaken over NSW in a number of projects during the past 2 decades. A suite of statewide DSMs at 100 m resolution were presented in OEH (2018). These are now superseded by the maps presented here, which used a more sophisticated modelling process involving multiple iterations of each map, allowing improved estimates of uncertainty, and which cover a slightly different suite of soil properties. The updated maps are available from the NSW Government's Sharing and Enabling Environmental Data (SEED) portal and via the Department of Planning and Environment (the department) eSPADE soil and landscape spatial viewer.

An earlier major national project resulted in the release of the Soil and Landscape Grid of Australia (SLGA) (Grundy et al. 2015; Viscarra Rossel et al. 2015), and more recently an updated SLGA. This includes DSMs for a wide range of key soil properties over the entire Australian continent at 3 arc second grid (approximately 90 m), with 6 depth intervals down to 2 m. Associated maps presenting upper and lower 95% confidence level predictions are also provided. This product represents Australia's contribution to a proposed global DSM: GlobalSoilMap.net (Sanchez et al. 2009; Arrouays et al. 2014).

Prior to this, digital mapping over Australia's agricultural zone had been carried out for a range of soil properties including pH, soil organic carbon (SOC), total phosphorus (P) and clay over topsoils and subsoils by Henderson et al. (2005) at 250 m resolution. Soil carbon was further digitally mapped across most or all of the country by Bui et al. (2009), Viscarra Rossel et al. (2014) and Gray et al. (2015a). Local catchment and field-scale DSMs have been carried out within NSW for various soil properties, for example by Minasny et al. (2006); Malone et al. (2009); Triantafilis et al. (2009) and Karunaratne et al. (2014).

1.2 Aims

The broad-scale Australia-wide DSM products of the SLGA were developed using soil data from all over Australia. In this current project, DSMs are prepared using only NSW soil data and environmental covariate layers, thus the models are more specific to this state. These new maps should provide further modelled evidence of soil properties across NSW that can be compared to and complement the existing continental-scale DSMs and also the NSW conventional soil maps.

This project aims to:

- use random forest (RF) decision-tree modelling techniques to prepare DSMs over NSW, covering the key soil properties of SOC, pH, cation exchange capacity (CEC), sum-of-bases, available P (bray), bulk density (BD), clay, sand and silt
- use 100 m resolution raster over 6 depth intervals down to 2 m, consistent with the SLGA and GlobalSoilMap.net, plus 0–10 cm, 10–30 cm depths and others for SOC
- provide validation results for the models and maps
- discuss interpretation and use of the maps.

The subject soil properties are essential for effective agricultural, hydrological, climatic, ecological and other scientific studies. The properties of SOC, pH_{ca}, CEC, sum-of-bases and available P are indicators of a soil's chemical condition, its nutrient status and potential to retain nutrients. The clay, silt and sand content and BD of a soil control its texture and physical behaviour, including water holding capacity and permeability, and also influence many chemical characteristics. The storage of carbon in soil is considered a potential vital avenue in addressing global climate change (Baldock et al. 2012; Gray et al. 2022).

2. Methods

2.1 Overview

The DSMs of the 9 soil properties were prepared at 8 or more depth intervals down to 2 m. They were based on soil survey data available over NSW. These data were randomly divided into training and validation subsets, at an approximate 80:20 ratio. Environmental covariate data representing the main soil-forming factors were applied in the initial training models and final maps production. These were derived from field survey data and various environmental data grids covering the entire state.

The modelling used an RF decision-tree method, however multiple linear regression (MLR) modelling assisted in the selection of environmental variables and in the understanding of influencing environmental factors for each soil property. Validation of the initial models and then the final DSM was carried out using the validation dataset. Figure 1 presents an overview of the DSM process applied in this project.

2.2 Soil data

Soil data for most soil properties was derived from the NSW Soil and Land Information System (SALIS). Exceptions were SOC and BD over upper 30 cm layers, which used data from the 2008 NSW monitoring, evaluation and reporting (MER) program (Chapman et al. 2011; OEH 2014), the Australian Soil Carbon Research Program (SCARP) (Sanderman et al. 2011; Baldock et al. 2013) and recent data from NSW Department of Primary Industries (DPI) (see Gray et al. 2022).

Final profile numbers, plus the laboratory analytical method, are listed in Table 1. The soil dataset was randomly apportioned 80% as training data and 20% as validation data for each soil property/depth interval.

In addition to SOC concentration (%), SOC mass (kg/m^3) and SOC stocks (Mg/ha) were derived by applying BD data. BD was available together with LECO SOC in the datasets down to 30 cm, but approximate estimates were derived from SLGA spatial maps deeper than 30 cm. The laboratory values derived from the Walkley–Black method (used for the 30–100 cm interval) are reported to underestimate values (Skjemstad et al. 2000) but no correction factor was applied for this.

To avoid reporting 2 separate pH test results, pH_w values were converted into pH_{ca} values using the correlation tables of Henderson and Bui (2002). The latter method of calculating pH is preferred in Australia as it more closely represents the ionic soil solutions typically found in the field, and thus gives more consistent results.

The soil datasets were organised into 6 standard depth intervals down to 2 m; that is, 0–5, 5–15, 15–30, 30–60, 60–100 and 100–200 cm, consistent with the SLGA and the global DSM project GlobalSoil.Net. Additional depth intervals of 0–10 cm and 10–30 cm were also applied, plus the 0–30 cm and 0–100 cm interval for SOC. Soil property values reported for the original depth interval of each soil horizon were converted into these standard depth intervals using the equal area splining process of Bishop et al. (1999) and Malone et al. (2009).

Table 1 Soil properties: laboratory methods and profile numbers

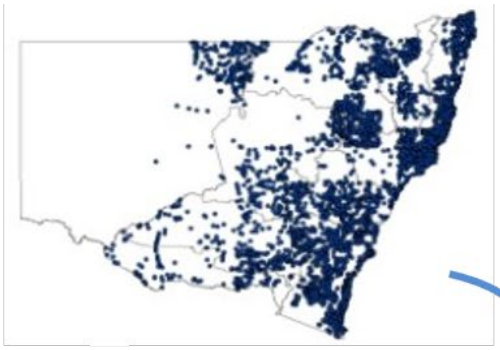
For laboratory methods see OEH (2017)

Soil property	Units	Laboratory method	No. of profiles
SOC	%, kg/m^3 Mg/ha %	6B2 (LECO combustion) (for 0–30 cm) 6A1 (Walkley–Black wet oxidation) and 6B2 (LECO) (for 30–100 cm)	2,160[1] 6,462[2]
pH$_{(CaCl_2)}$	pH units	4B1 (pH of 1:5 soil/0.01M calcium chloride extract). Includes conversions from 4A1 (pH 1:5 soil/water suspension)	11,610
CEC	cmol$_c$/kg	15F1 (silver thiourea)	6,064
Sum-of-bases[3]	cmol$_c$/kg	15F1 (silver thiourea) and 15D1 (pretreatment soluble salts)	6,929
P(bray)	mg/kg (ppm)	9E1 (fluoride-extractable P)	5,968
Bulk density	Mg/m^3	503 (core methods – mass of known vol.)	1,665
Sand (fine & total)	%	517 (sieves and hydrometer)	7,224
Silt	%	as above	7,224
Clay	%	as above	7,224

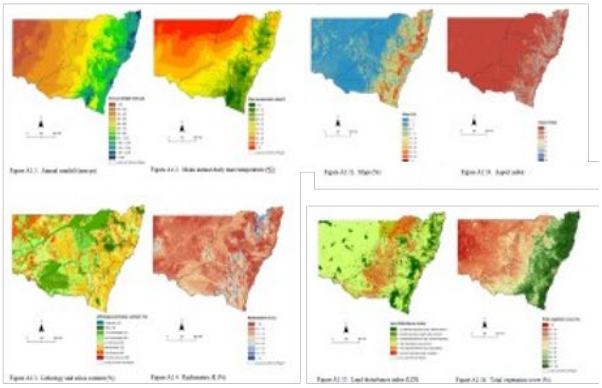
[1] SOC from MER, SCARP and DPI research programs

[2] SOC from SALIS (below 30 cm)

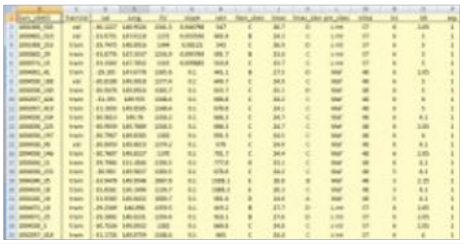
[3] Calcium, magnesium, sodium and potassium



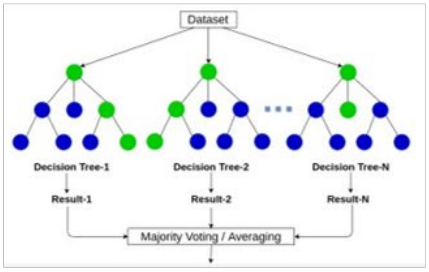
1. Acquire soil data points across NSW with associated laboratory values



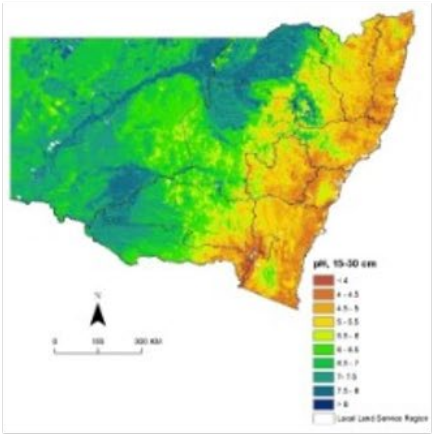
2. Acquire environmental data as grids across NSW



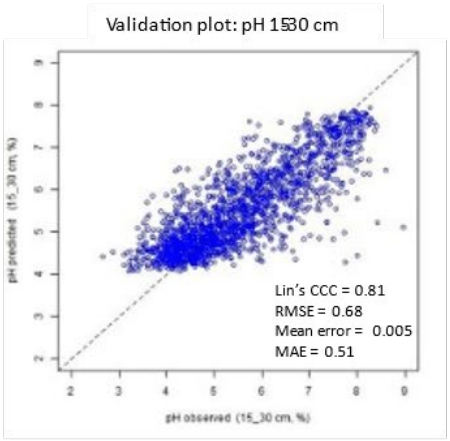
3. Prepare matrix of soil laboratory and environmental data, separate into training data (80%) and validation data (20%)



4. Run random forest decision trees linking soil properties to environmental data using training data



5. Use models with environmental grids to prepare soil property maps across NSW



6. Assess accuracy of the final maps using validation data

Figure 1 Overview of the DSM process

2.3 Covariates

Covariates were selected to effectively represent each of the key soil-forming factors of climate, parent material, relief, biota and age, as outlined below. Further detail is provided in the cited references.

Climate

- **Mean annual rainfall** (mm p.a., *rain*) derived from 2.5 km Australia-wide climate grids from the Australian Bureau of Meteorology with resampling of cell values down to a 100 m grid. For most soil properties, the grids cover the 1980–2005 period, which overlaps or slightly predates the period when most of the soil profiles were collected (see Appendix A, Figure A.1). However, for SOC, the climate grids covered the period approximately 20 years prior to the date of most of the soil carbon sampling, i.e. 1990–2010.
- **Mean annual daily maximum temperature** (°C, *Tmax*) – as above (see Appendix A, Figure A.2).
- **Mean annual daily minimum temperature** (°C, *Tmin*) – as above (see Appendix A, Figure A.3).

Parent material

- **Lithology** – the basis of this covariate was the lithology of the parent material, and more specifically the silica content (%), which is applied as a silica index. Silica content provides a simple but meaningful quantitative estimate of the chemical composition of most parent materials. It generally has a direct relationship to quartz content and an inverse relationship with basic cation content (Gray et al. 2016). For example, granite is moderately siliceous with approximately 73% silica, while basalt is mafic material with only approximately 48% silica. Higher silica content parent materials typically give rise to soils with more quartzose sandier textures with lower chemical fertility.

  For model development, the description of parent material or geologic unit recorded at each site by the soil surveyor was used. For the final map preparation, lithological classes and silica index values were applied manually to each geological formation as identified in the seamless digital geology map of the Geological Survey of NSW, a multi-scaled product between 1:25,000 and 1:500,000 scales. (Colquhoun et al. 2022). For poorly defined Cainozoic unconsolidated material, such as unqualified ‘alluvium’ or ‘colluvium’ for which their broad composition is unknown, lithological classes were allocated following reference to existing soil type maps. This exploited clear soil type to parent material relationships, such as black vertosols (Isbell and NCST 2021) with the mafic class and highly sandy arenosols with the upper siliceous class. Note that calcareous and some other lithology types could not be characterised in any meaningful way by their silica content. These occupy small areas across NSW and generally occur as minor components of broader mixed geology units (see Appendix A, Figure A.4).
- **Gamma radiometrics** – radiometric potassium (*rad_K*), uranium (*rad_U*) and thorium (*rad_Th*); 90 m grids developed by and sourced from Geoscience Australia (see Appendix A, Figures A.5–7).
- **NIR clay components** – the relative proportions of kaolin, illite and smectite clays (0–20 cm depth) derived from DSM techniques based on laboratory near infra-red (NIR) spectroscopy (Viscarra Rossel 2011); 90 m grids sourced through the CSIRO Data Access Portal via the SLGA (see Appendix A, Figures A.8–10).

Relief

- **Topographic wetness index** (*TWI*) – a widely used index that represents potential hydrological conditions based on slope and catchment area, as derived from digital elevation models (DEMs) (Gallant and Austin 2015); sourced through the CSIRO Data Access Portal via the SLGA (see Appendix A, Figure A.11).
- **Slope** – slope gradient in percent as derived from a 100 m DEM (see Appendix A, Figure A.12).
- **Aspect index** (*Asp*) – an index to represent the amount of solar radiation received by sites, ranging from 1 for flat areas and gentle north or north-west facing slopes (high radiation in southern hemisphere) to 10 for steep south and south-east facing slopes (low radiation) (Gray et al. 2015b) (see Appendix A, Figure A.13).

Biota

- **Land disturbance index** (*LDI*) – an index that reflects the intensity of disturbance associated with the land use (Gray et al. 2015b), where 1 denotes natural ecosystems and 6 denotes intensive cropping, based on 1:25,000 scale land-use mapping for 2007 (for training data where site data was lacking) and 2017 (for final grids) (DPE 2021) (see Appendix A, Figure A.14)
- **Total vegetation cover** (*vegtot)* – combined photosynthetic and non-photosynthetic vegetation (in %) derived from MODIS fractional vegetation data, 2000–2017 (Guerschman and Hill 2018) (see Appendix A, Figure A.15).

Age

- **Weathering index** (*W_I*) – an index to represent the degree of weathering of parent materials, regolith and soil, based on gamma radiometric data (Wilford 2012); 90 m grids were sourced from Geoscience Australia. The index is considered reflective of the age factor in the *clorpt* and *scorpan* frameworks (see Appendix A, Figure A.16).

2.4 Modelling approach

All analyses were carried out using R statistical software (R Core Team 2022). The soil datasets were apportioned 80% as training data and 20% as validation data using a simple random data splitting approach with modelling by RF decision-tree models (randomForest package, Liaw and Wiener 2018). Variable importance plots and examination of MLR models allowed the identification of the most influential environmental variables to apply in the final RF models. Final maps were prepared with the RF models, using 10 bootstrap samples and stacking the resulting outputs (using customised code with the abovementioned package). The 10 bootstrap samples together with the 200 trees applied within each bootstrap, which gave 2,000 iterations for each soil property/depth interval, were considered sufficient for the purpose of this study. No significant differences were observed in the mean values, standard deviations and prediction limits in trials with 10 and 100 bootstraps (see Appendix B).

A natural log transformation was applied to the SOC, CEC, sum-of-bases and P(bray) values to achieve normality. Upper 95% and lower 5% prediction limit maps were derived using results from the multiple RF iterations. The difference between these prediction limits gave the 90% prediction interval, as presented in the associated maps.

The kriging of residual errors was trialled in order to incorporate any consistent spatial patterns of modelling error (Odeh et al. 1995), but was not adopted in the final maps due to persistent anomalies.

The variable importance plots demonstrate the relative influence of each environmental variable for each soil property. The metric ‘increase in mean square error (MSE)’ was

adopted in these plots, which denotes the increase in error, or decline in performance of the model, due to the removal of the subject variable. The direction of influence of the subject variable on each variable was derived from MLR models performed simultaneously with the same dataset.

2.5 Map production

The maps generated by R in the above process were further formatted using ESRI ArcGIS software. Maps at finer scales covering Local Land Services (LLS) or other areas of particular interest may also be prepared in GIS environments, with enhancement with towns, roads and waterways, such as those presented in Figures 2 and 3. The maps may also be viewed at different scales with underlying infrastructure, terrain or satellite image layers through eSPADE.

2.6 Validation

The final DSMs were validated using the validation datasets (20% of original data points). Lin's concordance correlation coefficient (LCCC) was used to measure the level of agreement of predicted values with observed values relative to the 1:1 line (Lin 1989). The root mean square error (RMSE), mean error (ME) and mean/median absolute error (MAE/MedAE) of results was also determined. These statistics, together with the prediction interval maps, provide an indication of uncertainty levels sufficient to assess the performance of these maps.

The adoption of more sophisticated techniques such as the cross-validation techniques demonstrated by Malone et al. (2014) and Kidd et al. (2015) might further improve the reliability of the validation results. The collection of independent validation data using a design-based sampling approach would also improve the assessment of accuracy and quality of the maps (Brus et al. 2011).

The extent to which the summation of all particle sizes (sand, silt and clay, but not coarse fragments) approximates 100% provides an indication of the reliability of these maps.

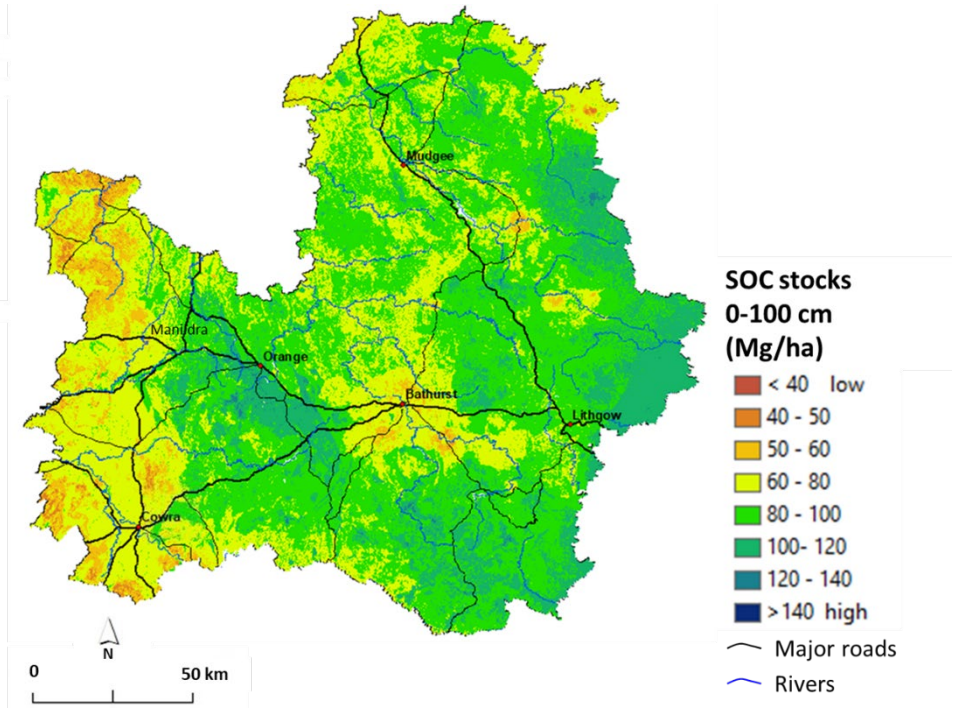


Figure 2 **Moderate-scale map of SOC stocks (to 100 cm depth) over Central Tablelands LLS**

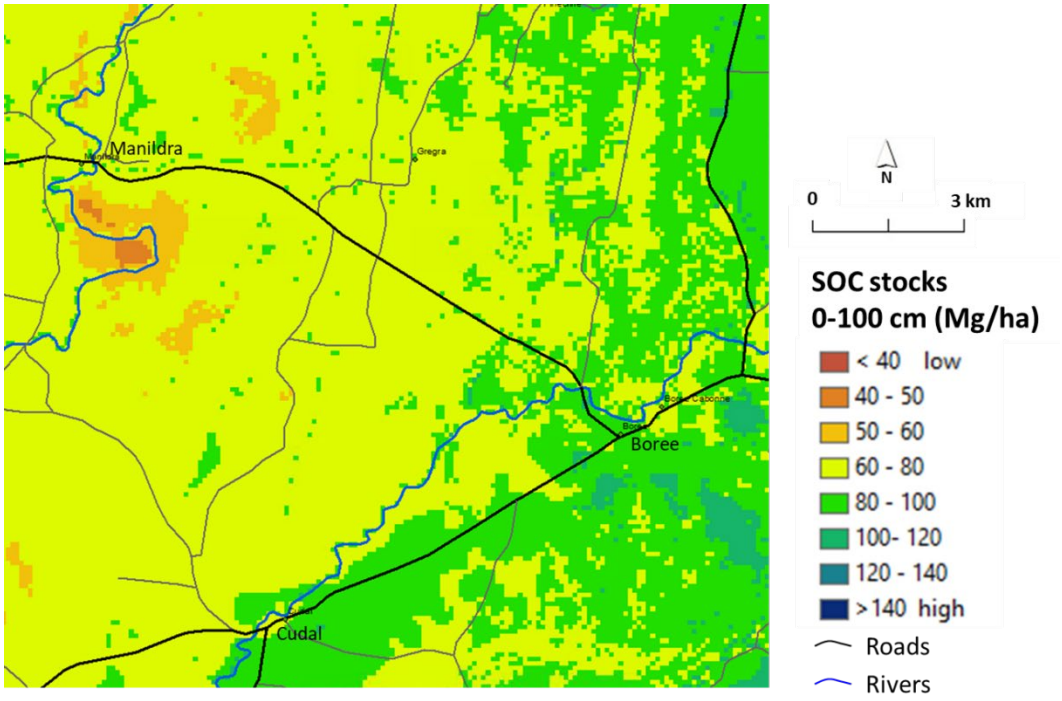


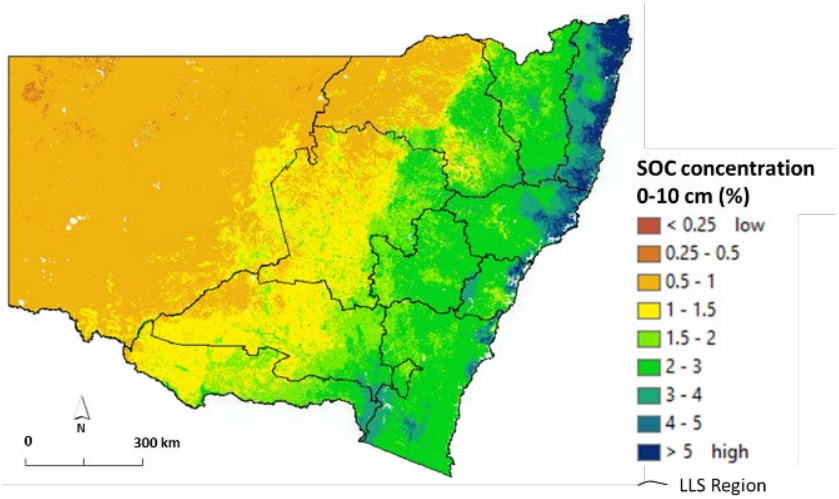
Figure 3 **Fine-scale map of SOC stocks (to 100 cm depth) over Manildra region, central NSW**

3. The digital soil maps

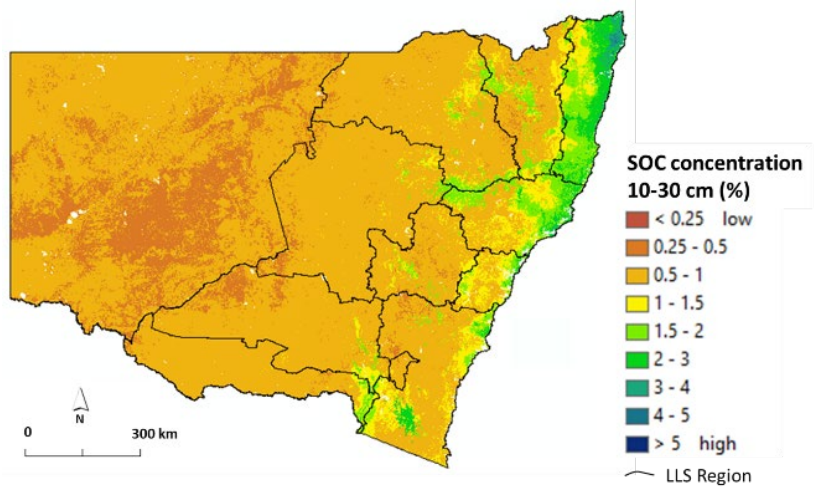
Map images for 3 selected depth intervals (0–10, 10–30 and 30–60 cm) for each soil property are presented in Figures 4 to 15. Full digital maps (in geotiff format) for all depth intervals are available for download through the NSW Government environmental data portal SEED and may also be viewed directly through the department's soil and landscape spatial viewer eSPADE. The following images are presented in terms of classes (e.g. <1%, 1–2%, etc.). Variable importance plots (with direction of influence of each variable), validation statistics of the selected depth intervals, and a brief summary of the results are also presented. Validation statistics for all depth intervals are also presented in Appendix C. The 90% prediction interval maps for the selected depth intervals are presented in Appendix D.

3.1 Soil organic carbon

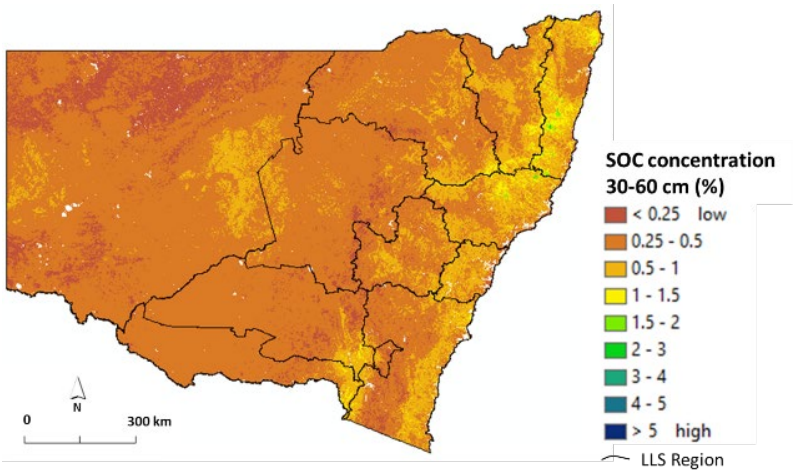
3.1.1 SOC %



a: SOC % 0–10 cm, mean[1]



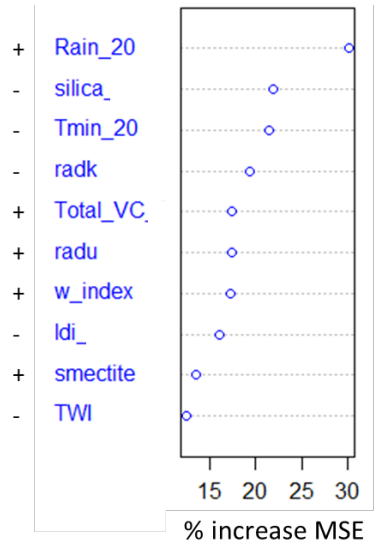
b: SOC % 10–30 cm, mean[1]



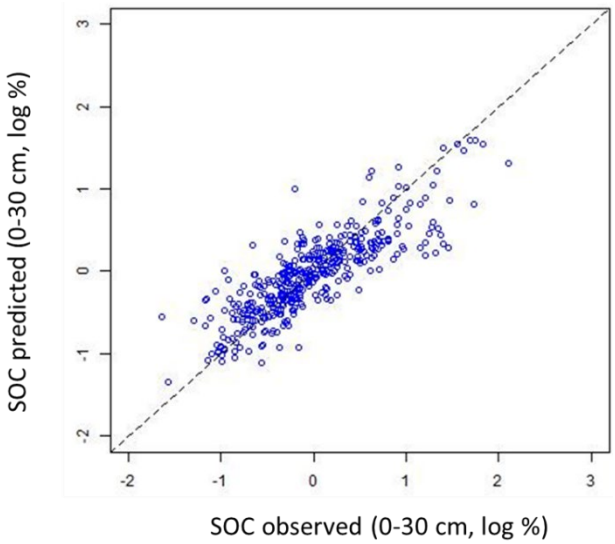
c: SOC % 30–60 cm, mean[1]

Figure 4 Selection of maps for SOC concentration

[1] 90% prediction interval maps are provided in Appendix D.1.1



a: SOC % 0–30 cm – variable importance plot (with direction of influence)



b: SOC % 0–30 cm – map validation plot

Figure 5 Selection of plots for SOC concentration

Table 2 Map validation statistics for selection of layers: SOC %

(see Appendix C for all layers to 100 cm)

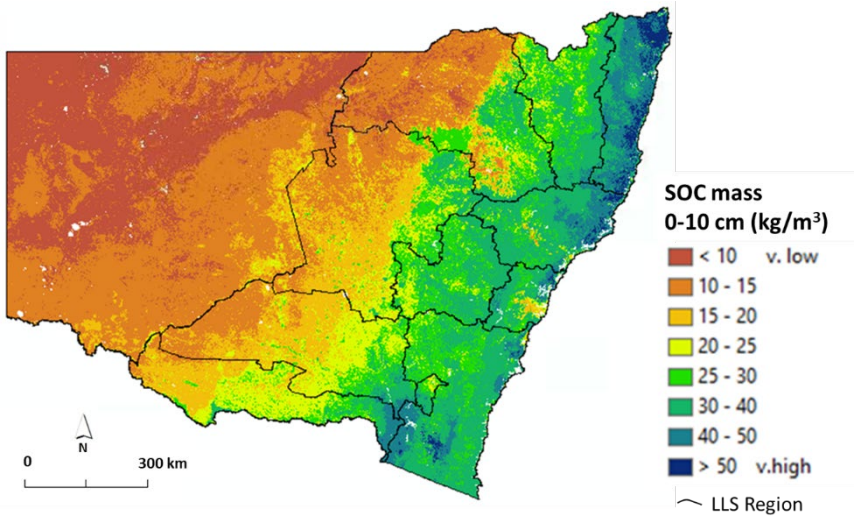
Depth (cm)	N	LCCC	RMSE (log units)	ME (log units)	MAE (log units)
0–10	415	0.82	0.36	–0.02	0.27
10–30	415	0.73	0.41	–0.004	0.30
0–30	417	0.82	0.34	–0.01	0.25
30–60	1177	0.32	0.97	–0.02	0.65

N: validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

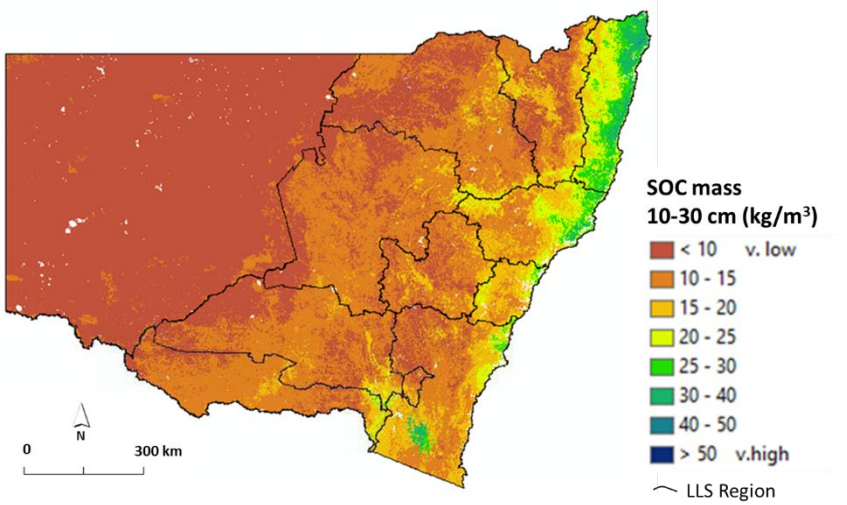
Summary

The maps are of high statistical strength in the upper 30 cm with average LCCC values over 0.8. However, they are weaker in the lower depths, with LCCC values dropping below 0.3 (see Appendix C). The maps reveal highest SOC concentrations in the upper surface layers, over 5% in the north-east coastal region in the 0–10 cm interval, but rarely exceeding 1.5% in the 30–60 cm depth interval. Values increase from west to east, primarily reflecting the increasing rainfall and lower temperatures. Values also increase with less siliceous (more clay rich) soil/parent materials and higher vegetation cover as demonstrated by the variable importance plot in Figure 5a.

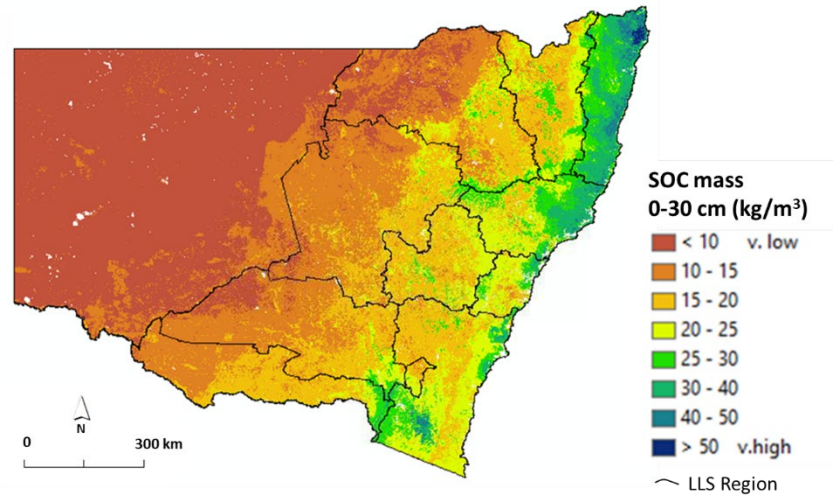
3.1.2 SOC mass (kg/m^3)



a: SOC mass (kg/m^3) 0–10 cm, mean[1]



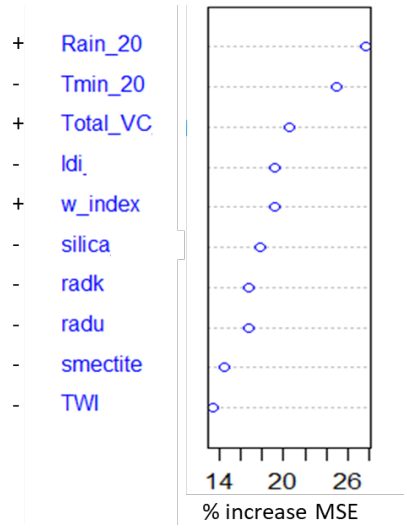
b: SOC mass (kg/m^3) 10–30 cm, mean[1]



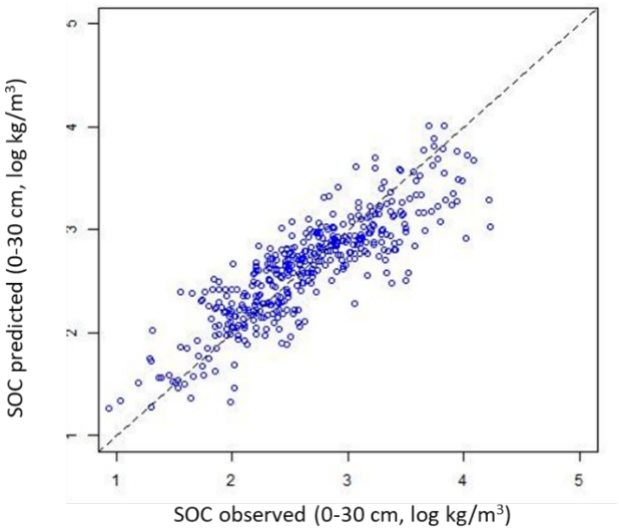
c: SOC mass (kg/m^3) 0–30 cm, mean[1]

Figure 6 Selection of maps for SOC mass

[1] 90% prediction interval maps are provided in Appendix D.1.2



a: SOC mass (kg/m^3) 0–30 cm – variable importance plot (with direction of influence)



b: SOC mass (kg/m^3) 0–30 cm – map validation plot

Figure 7 Selection of plots for SOC mass

Table 3 Map validation statistics for selection of layers: SOC mass

(see Appendix C for other layers to 30 cm)

Depth (cm)	N	LCCC	RMSE (log units)	ME (log units)	MAE (log units)
0–10	417	0.83	0.33	–0.02	0.25
10–30	418	0.77	0.36	0.0045	0.27
0–30	418	0.84	0.32	–0.02	0.24

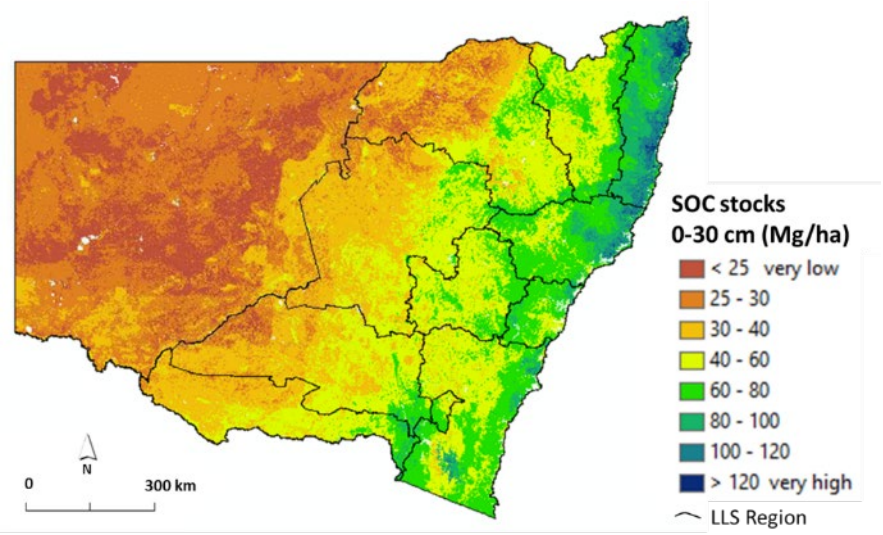
N: validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

Summary

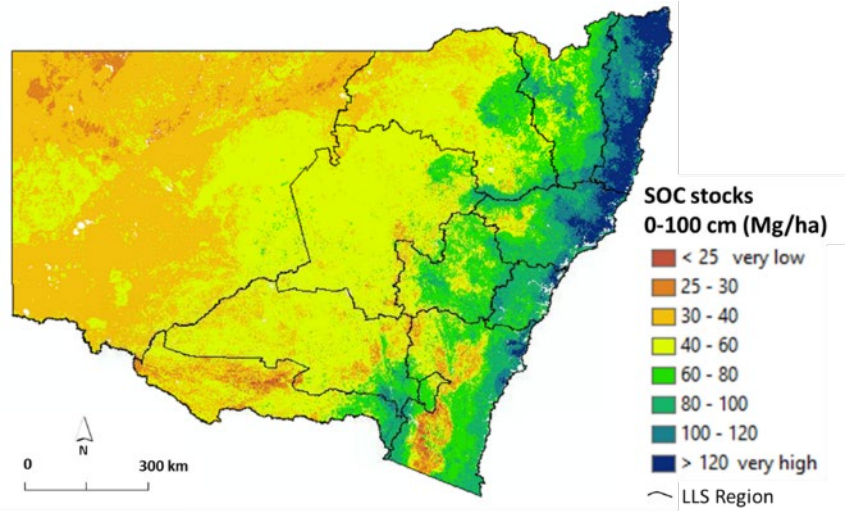
As for SOC concentration, the maps for SOC mass (in kg/m^3) are of high strength in the upper 30 cm with average LCCC values over 0.8. There was insufficient laboratory derived BD data over NSW to allow mass calculations below 30 cm.

The maps reveal highest SOC mass in the upper surface layers, over 50 kg/m^3 in the north-east coastal region in the 0–10 cm interval, but rarely exceeding 30 kg/m^3 in the 10–30 cm depth interval. As for SOC concentration, values increase from west to east, primarily reflecting the increasing rainfall and lower temperatures. Values also increase with higher vegetation cover, less intensive land use (indicated by negative LDI variable), and less siliceous (more clay rich) soil/parent materials and as demonstrated by the variable importance plot in Figure 7a.

3.1.3 SOC stocks (Mg/ha)



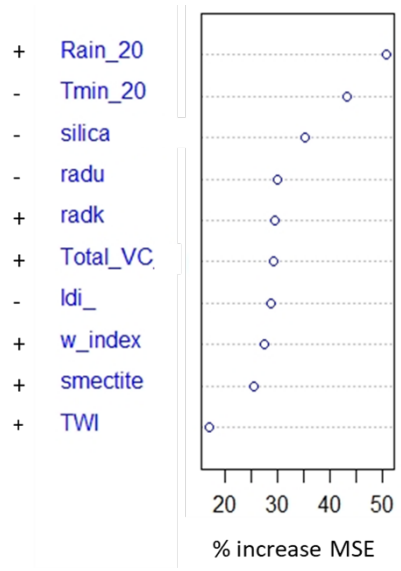
a: SOC stocks (Mg/ha) 0–30 cm, mean[1]



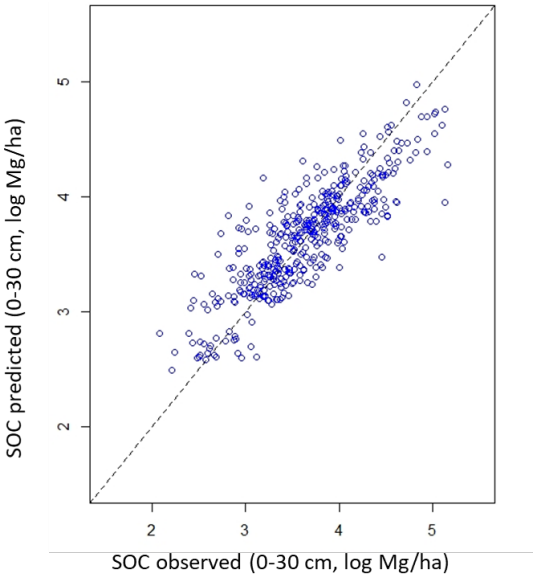
b: SOC stocks (Mg/ha) 0–100 cm, mean[1]

Figure 8 Selection of maps for SOC stocks

[1] 90% prediction interval maps are provided in Appendix D.1.3



a: SOC stocks (Mg/ha) 0–30 cm – variable importance plot (with direction of influence)



b: SOC stocks (Mg/ha) 0–30 cm – map validation plot

Figure 9 Selection of plots for SOC stocks

Table 4 Map validation statistics for selection of layers: SOC stocks

(see Appendix C for other layers to 100 cm)

Depth (cm)	N	LCCC	RMSE (log units)	ME (log units)	MAE (log units)
0–10	418	0.84	0.33	–0.01	0.24
0–30	430	0.80	0.33	0.005	0.26
0–100	431	0.78	0.23	–0.003	0.16

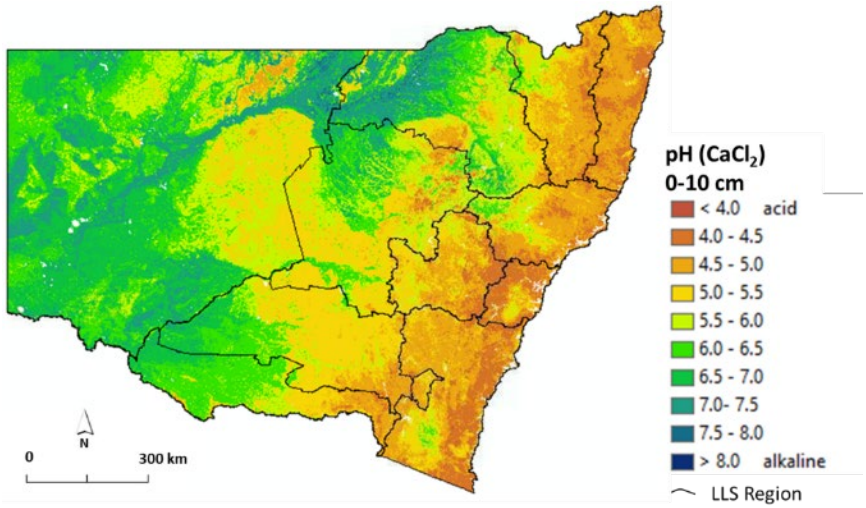
N: validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

Summary

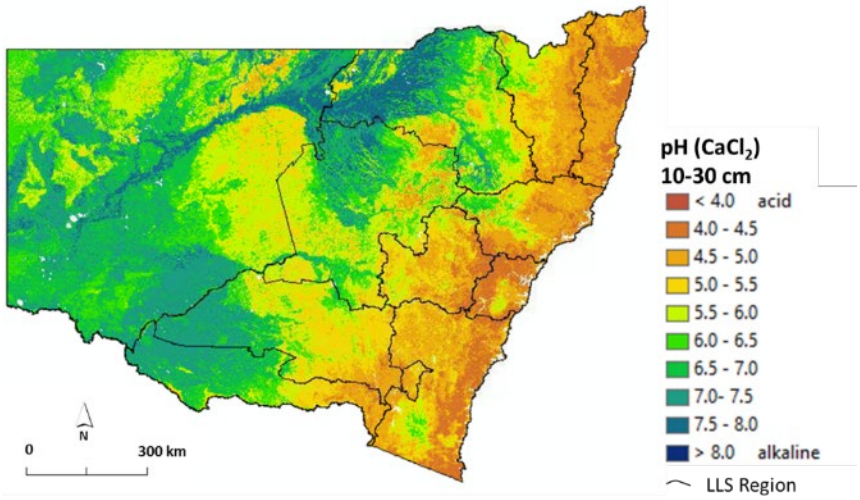
As for SOC concentration and mass, the maps for SOC stocks (in Mg/ha) are of high strength in the upper 30 cm with an LCCC value of 0.84 for the 0–10 cm interval. The validation results for 0–100 cm also demonstrated strong performance.

The SOC stocks are highest in the 0–100 cm interval as expected due to the greater volume of soil per unit area over this depth. Lower SOC concentrations and mass at the deeper depths are compensated for by the greater volume of soil. As for SOC concentration and mass, values increase from west to east, primarily reflecting the increasing rainfall and lower temperatures. Values also increase with less siliceous (more clay rich) soil/parent materials, higher vegetation cover and less intensive land use (LDI) as demonstrated by the variable importance plot in Figure 9a.

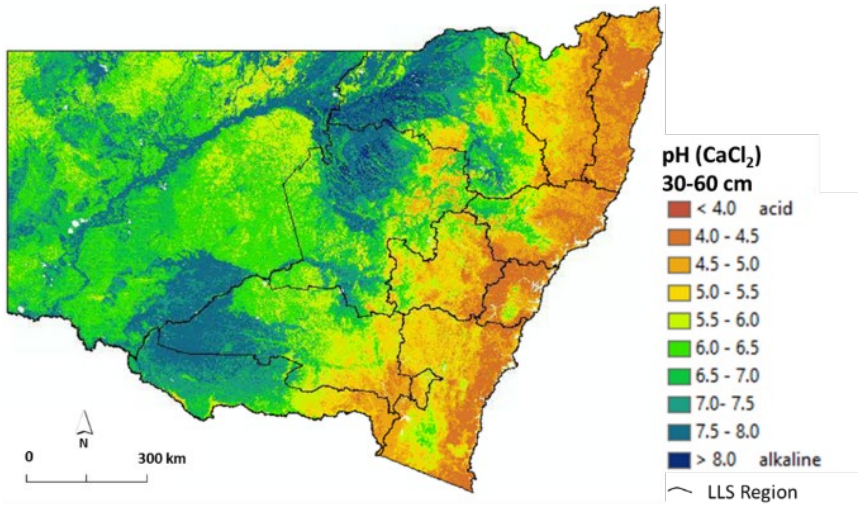
3.2 pH($CaCl_2$) (pH units)



a: pH($CaCl_2$) 0–10 cm, mean[1]



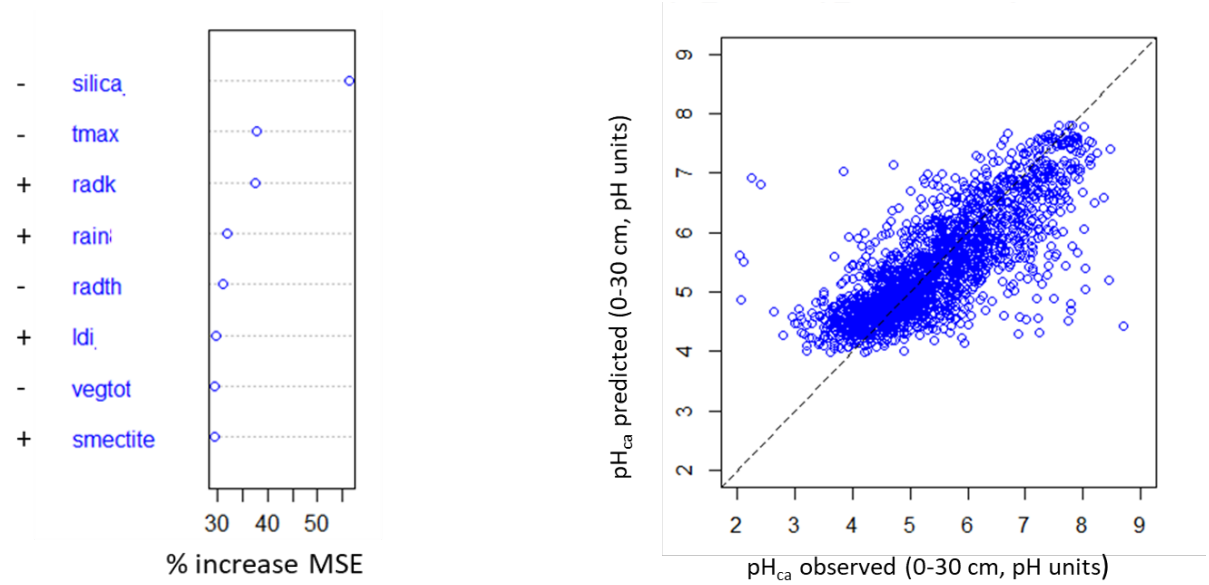
b: pH($CaCl_2$) 10–30 cm, mean[1]



c: pH($CaCl_2$) 30–60 cm, mean[1]

Figure 10 Selection of maps for pH

[1] 90% prediction interval maps are provided in Appendix D.2



a: pH($CaCl_2$) 0–30 cm – variable importance plot (with direction of influence)

b: pH($CaCl_2$) 0–30 cm – map validation plot

Figure 11 Selection of plots for pH

Table 5 Map validation statistics for selection of layers: pH

(see Appendix C for all layers to 200 cm)

Depth (cm)	N	LCCC	RMSE	ME	MAE
0–10	2098	0.74	0.67	0.02	0.50
10–30	2025	0.80	0.66	–0.002	0.50
0–30	2031	0.76	0.70	0.005	0.51
30–60	1710	0.82	0.74	–0.02	0.57

N: Validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

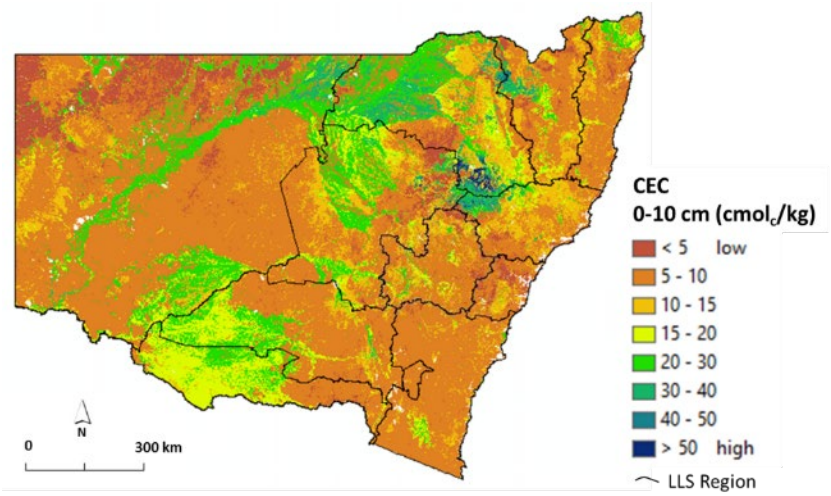
Summary

The pH maps are of high statistical strength over all depth intervals. LCCC values increase slightly with increasing depth, e.g. rising to 0.84 over the 100–200 cm interval (see Appendix C). However, other statistical indicators such as RMSE and MAE suggest slightly weaker performance with depth.

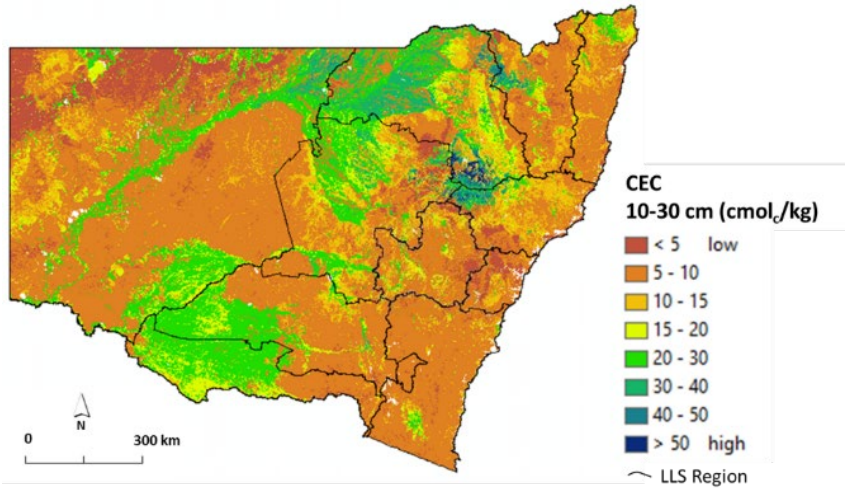
The maps reveal lowest pH values (i.e. more acidic) in the upper surface layers, becoming gradually higher (more alkaline) with depth. Soils become more acidic (lower pH) with increasing siliceous (more sandy) soil/parent materials, with the silica variable being the dominant influence and displaying a negative trend in the variable importance plot in Figure 11a. Soils also become more acidic from west to east, reflecting the increasing rainfall and declining temperatures with the associated higher leaching potential in the east of the state. Increasing land disturbance, i.e. more agricultural use, is associated with more alkaline soils, which are typically more fertile and suitable for agriculture.

3.3 Cation exchange capacity ($cmol_c/kg$)

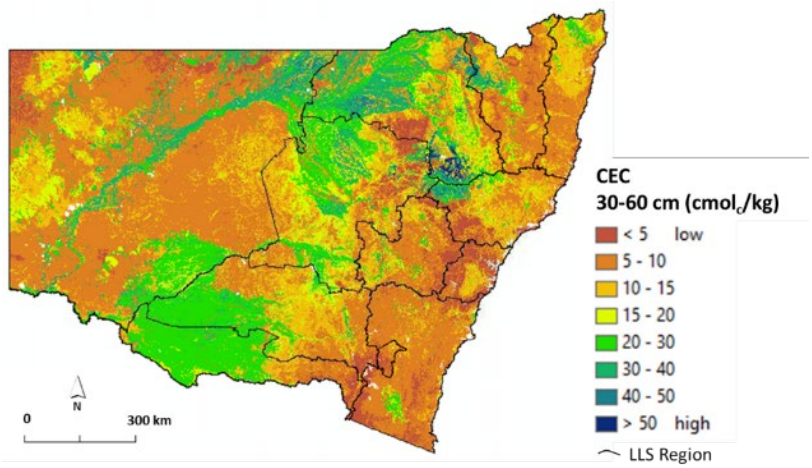
(Total exchange sites in soil; including sites occupied by basic and acidic exchangeable cations)



a: CEC ($cmol_c/kg$) 0–10 cm, mean[1]



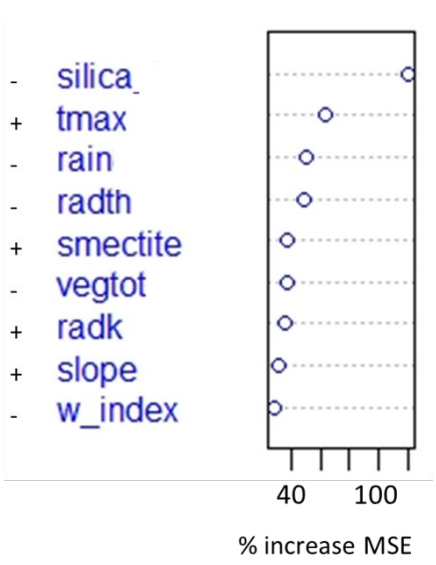
b: CEC ($cmol_c/kg$) 10–30 cm, mean[1]



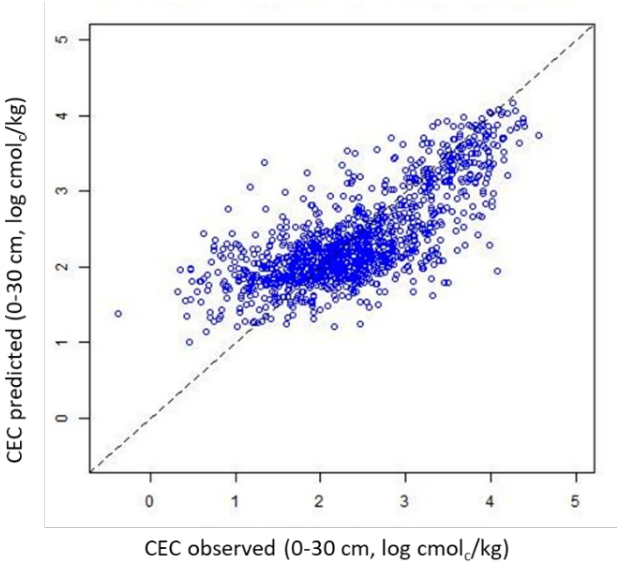
c: CEC ($cmol_c/kg$) 30–60 cm, mean[1]

Figure 12 Selection of maps for CEC

[1] 90% prediction interval maps are provided in Appendix D.3



a: CEC ($cmol_c/kg$) 0–30 cm – variable importance plot (with direction of influence)



b: CEC ($cmol_c/kg$) 0–30 cm – map validation plot

Figure 13 Selection of plots for CEC

Table 6 Map validation statistics for selection of layers: CEC

(see Appendix C for all layers to 200 cm)

Depth (cm)	N	LCCC	RMSE (log units)	ME (log units)	MAE (log units)
0–10	1113	0.57	0.68	0.002	0.47
10–30	1059	0.69	0.59	–0.005	0.44
0–30	1247	0.70	0.57	–0.04	0.43
30–60	1120	0.67	0.67	0.05	0.48

N: Validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

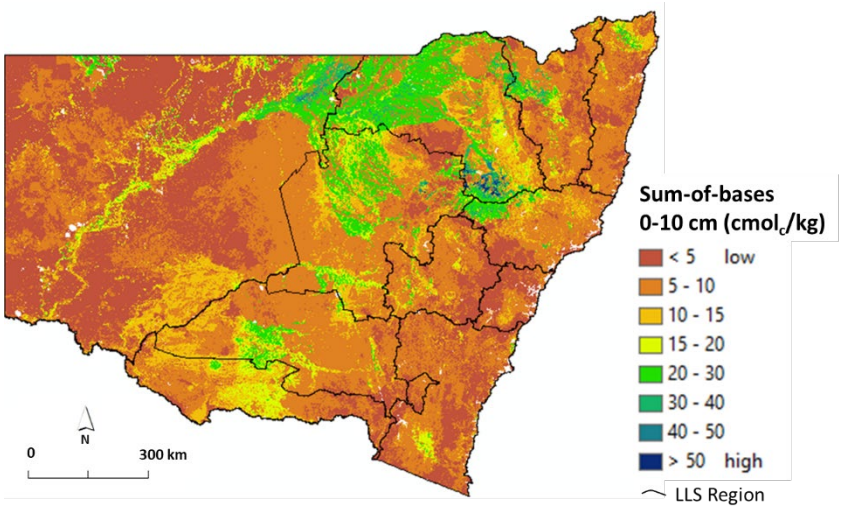
Summary

The CEC maps are of moderate to high strength over all depth intervals. LCCC values increase slightly with increasing depth, e.g. rising from 0.57 for the 0–10 cm interval to 0.70 over the 60–100 cm interval, an improving trend also repeated with other statistics, such as RMSE and MAE (see Appendix C).

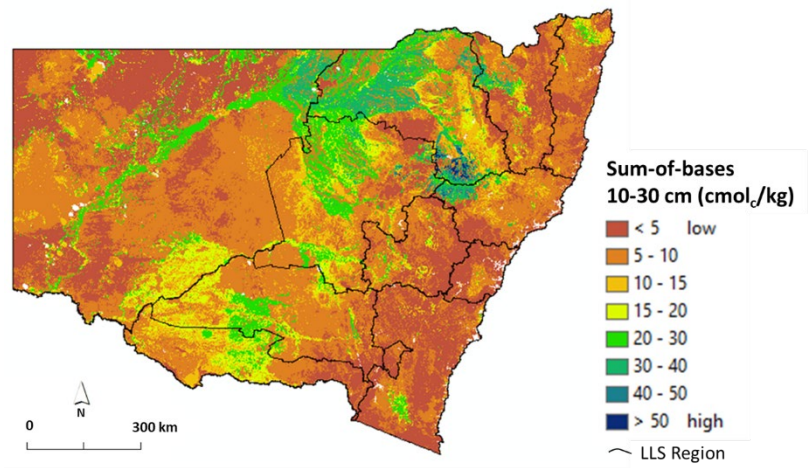
The maps reveal lowest CEC (i.e. ability to store cations /macro-nutrients) in the upper surface soil layers, becoming gradually higher in CEC with depth, with increasing ability to store macro-nutrients and thus becoming more fertile. Soils achieve higher CEC levels with increasing mafic character of soil/parent materials, with the silica variable being the dominant influence and displaying a negative trend in the variable importance plot in Figure 13a. The CEC of soils also increases with drier climatic conditions, i.e. lower rainfall and higher temperatures. Increasing land disturbance, i.e. more agricultural use, is associated with more fertile soils able to retain macro-nutrients that are more suitable for agriculture.

3.4 Sum-of-bases ($cmol_c/kg$)

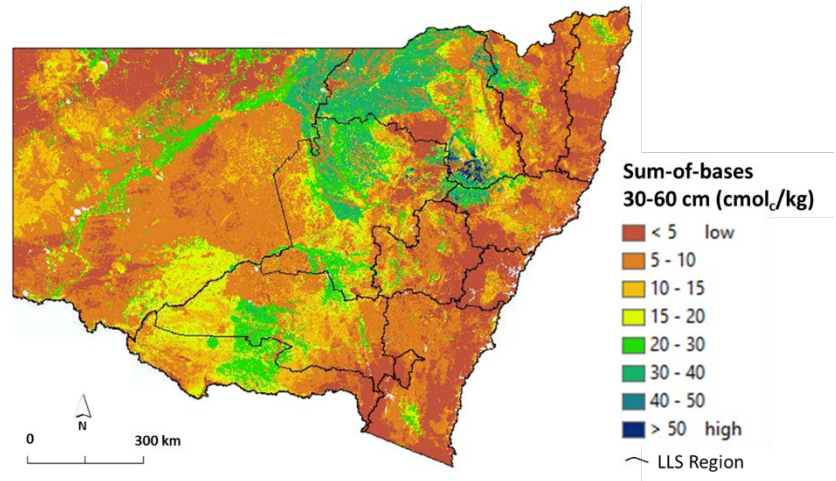
(Sum of exchangeable calcium, magnesium, sodium and potassium cations)



a: Sum-of-bases ($cmol_c/kg$) 0–10 cm, mean[1]



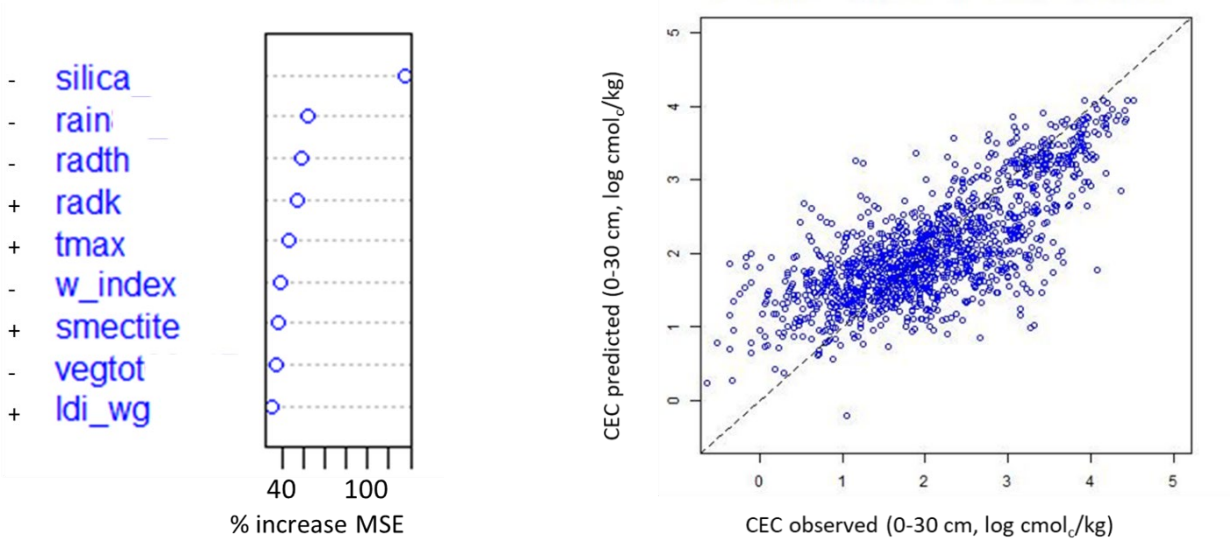
b: Sum-of-bases ($cmol_c/kg$) 10–30 cm, mean[1]



c: Sum-of-bases ($cmol_c/kg$) 30–60 cm, mean[1]

Figure 14 Selection of maps for sum-of-bases

[1] 90% prediction interval maps are provided in Appendix D.4



a: Sum-of-bases (cmol$_c$/kg) 0–30 cm – variable importance plot (with direction of influence)

b: Sum-of-bases (cmol$_c$/kg) 0–30 cm – map validation plot

Figure 15 Selection of plots for sum-of-bases

Table 7 Map validation statistics for selection of layers: sum-of-bases

(see Appendix C for all layers to 200 cm)

Depth (cm)	N	LCCC	RMSE (log units)	ME (log units)	MAE (log units)
0–10	1323	0.69	0.68	–0.004	0.51
10–30	1289	0.74	0.68	0.03	0.52
0–30	1324	0.72	0.67	–0.008	0.52
30–60	1189	0.76	0.71	–0.007	0.53

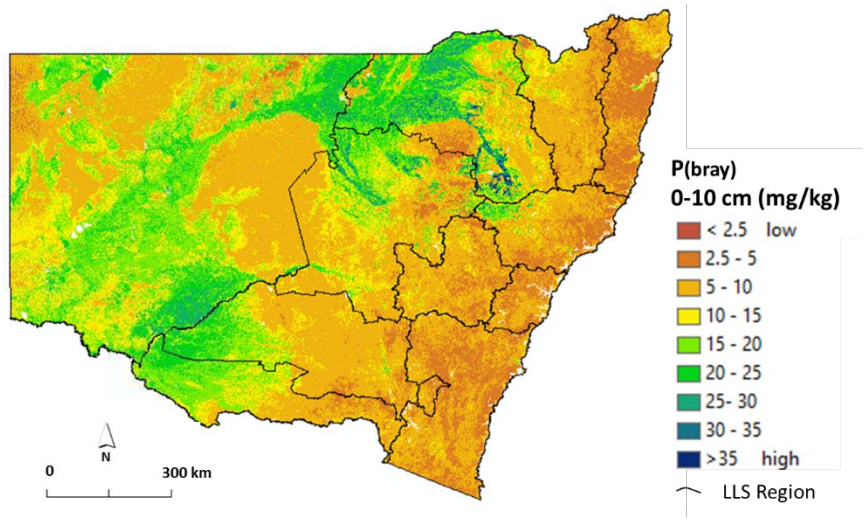
N: Validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

Summary

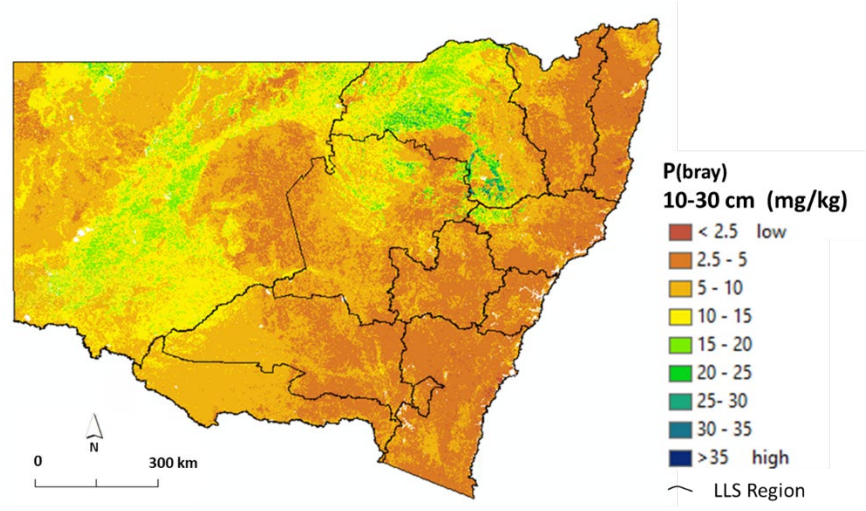
The sum-of-bases maps are of moderate to high statistical strength over all depth intervals. LCCC values slightly increase with increasing depth, e.g. rising from 0.69 for the 0–10 cm interval to 0.76 or more below 30 cm, an improving trend also repeated with other statistics such as RMSE and MAE (see Appendix C).

The maps reveal lowest sum-of-bases values (i.e. lower cations or macro-nutrients) in the upper surface layers, becoming gradually more macro-nutrient rich with depth. Soils become more macro-nutrient rich and fertile with increasing mafic character of soil/parent materials, with the silica variable being the dominant influence and displaying a negative trend in the variable importance plot in Figure 15a. Sum-of-bases in soils also increases with drier climatic conditions, i.e. lower rainfall and higher temperatures, with the associated lower leaching of cations. Increasing land disturbance, i.e. more agricultural use, is associated with more macro-nutrient rich and fertile soils that are more suitable for agriculture.

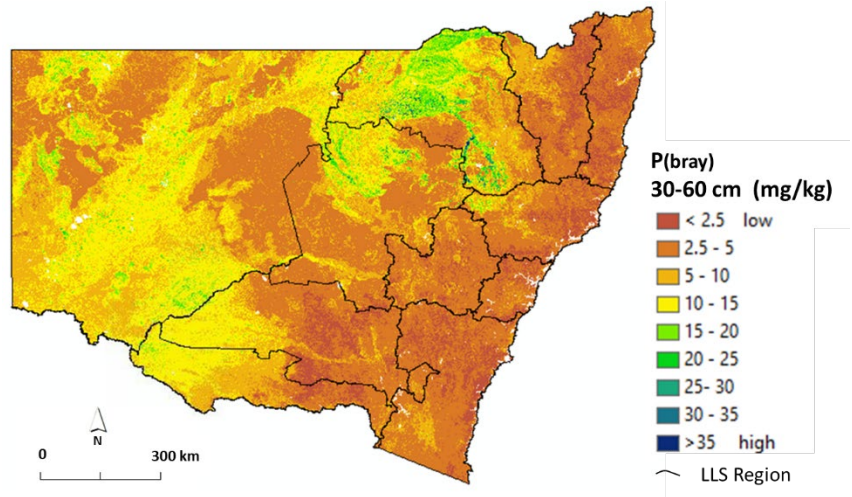
3.5 Available phosphorus (P(bray), mg/kg)



a: P(bray) (mg/kg) 0–10 cm, mean[1]



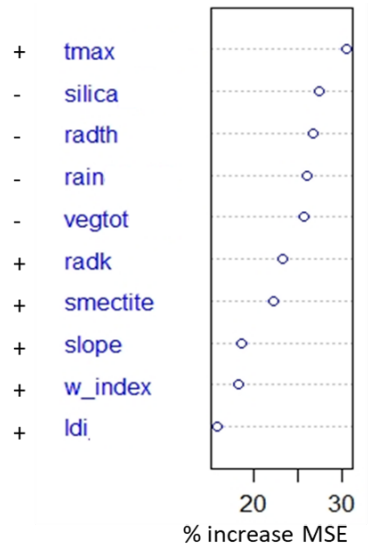
b: P(bray) (mg/kg) 10–30 cm, mean[1]



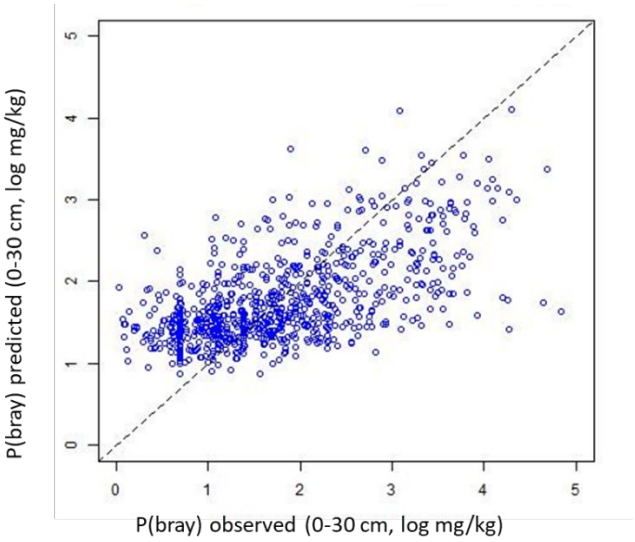
c: P(bray) (mg/kg) 30–60 cm, mean[1]

Figure 16 Selection of maps for available P(bray)

[1] 90% prediction interval maps are provided in Appendix D.5



a: P(bray) (mg/kg) 0–30 cm – variable importance plot (with direction of influence)



b: P(bray) (mg/kg) 0–30 cm – map validation plot

Figure 17 Selection of plots for available P(bray)

Table 8 Map validation statistics for selection of layers: available P(bray)

(see Appendix C for all layers to 200 cm)

Depth (cm)	N	LCCC	RMSE (log units)	ME (log units)	MAE (log units)
0–10	845	0.42	0.85	–0.004	0.68
10–30	808	0.49	0.81	–0.05	0.63
0–30	801	0.52	0.76	–0.007	0.60
30–60	560	0.48	0.88	0.09	0.69

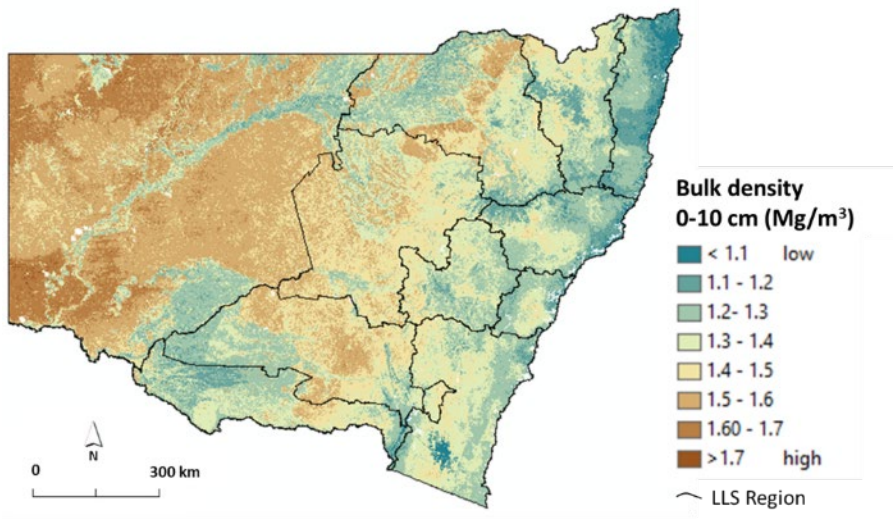
N: Validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

Summary

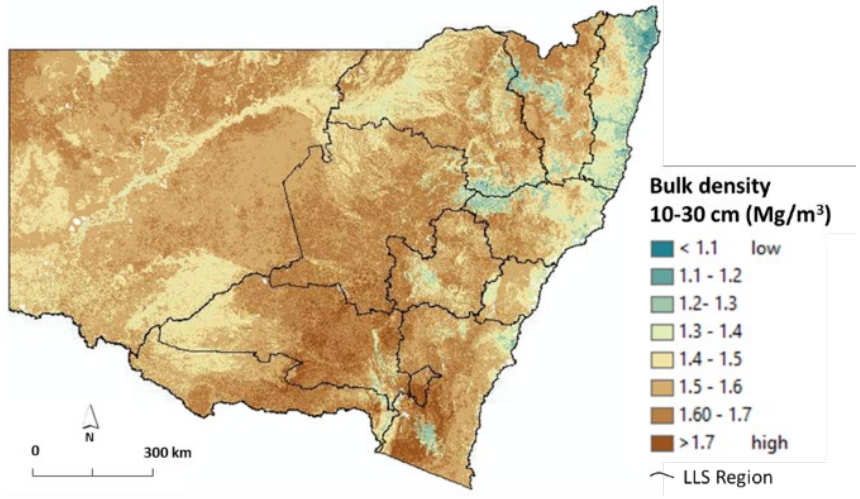
The available P(bray) maps are of only low to moderate statistical strength over all depth intervals. LCCC values increase slightly with increasing depth, e.g. rising from 0.42 for the 0–10 cm interval to 0.55 over the 100–200 cm interval, an improving trend, but other statistics such as RMSE and MAE remain constant or decline slightly in performance with depth (see Appendix C).

The maps reveal highest available phosphorus (an important plant nutrient) in the upper surface layers, declining gradually with depth. Soils increase in available P, and become more fertile for agricultural plants, with increasing mafic character of soil/parent materials, with the silica variable being the dominant influence (with negative trend) in the variable importance plot in Figure 17a. Available P in soils is also shown to increase with drier climatic conditions, i.e. lower rainfall and higher temperatures, with the associated lower leaching of cations. Increasing land disturbance, i.e. more agricultural use, is associated with more available P in soils that are more suitable for agriculture.

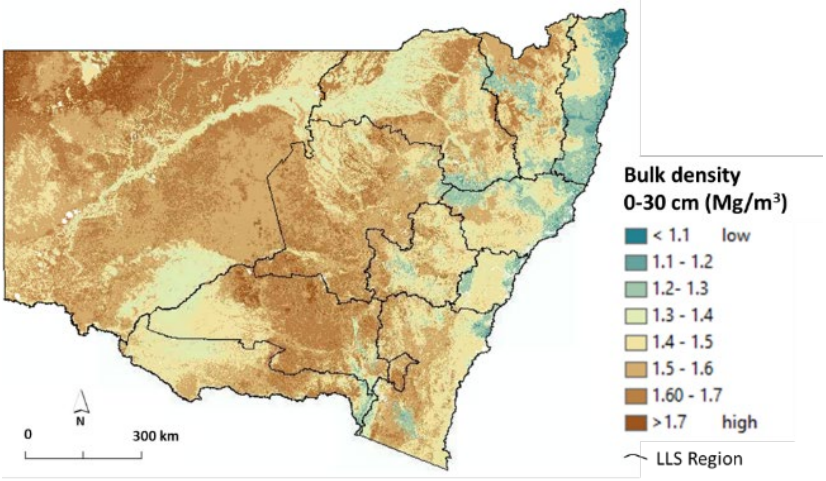
3.6 Bulk density (Mg/m^3)



a: BD (Mg/m^3) 0–10 cm, mean[1]



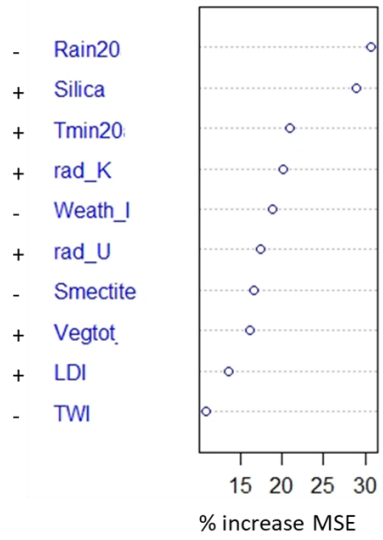
b: BD (Mg/m^3) 10–30 cm, mean[1]



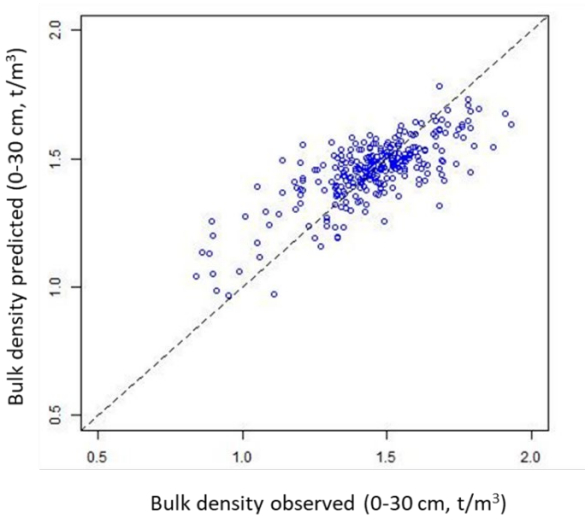
c: BD (Mg/m^3) 0–30 cm, mean[1]

Figure 18 Selection of maps for BD

[1] 90% prediction interval maps are provided in Appendix D.6



a: BD (Mg/m^3) 0–30 cm – variable importance plot (with direction of influence)



b: BD (Mg/m^3) 0–30 cm – map validation plot

Figure 19 Selection of plots for BD

Table 9 Map validation statistics for selection of layers: BD

(see Appendix C for other layers to 30 cm)

Depth (cm)	N	LCCC	RMSE (Mg/m^3)	ME (Mg/m^3)	MAE (Mg/m^3)
0–10	303	0.51	0.18	0.01	0.14
10–30	294	0.56	0.17	0.001	0.13
0–30	296	0.70	0.13	0.0008	0.094

N: Validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

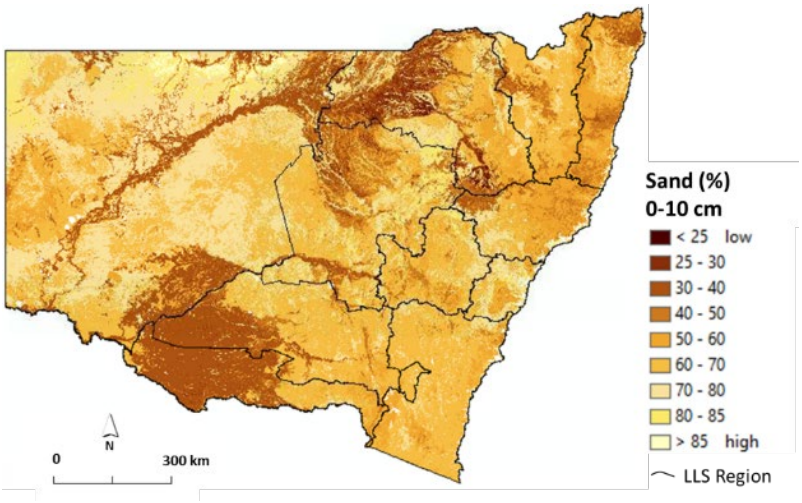
Summary

The BD maps were only prepared down to 30 cm, the depth to which there was reliable data for NSW. The maps are of moderate statistical strength, with LCCC values varying between 0.51 and 0.70.

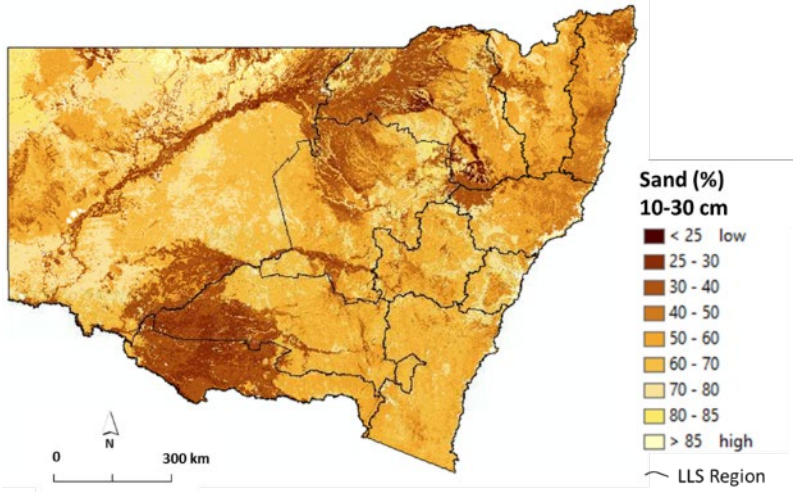
The maps reveal lowest BD in the upper surface layers, increasing gradually with depth. Soils are revealed to increase in BD with drier climatic conditions, i.e. lower rainfall and higher temperatures, as revealed by the variable importance plot in Figure 19a. Values increase with increasing siliceous character of the parent material, as revealed by the positive trend with silica in the variable importance plot. Similarly, values increase with increasing disturbance of the soil (e.g. increased cultivation), as shown by the positive trend with LDI.

3.7 Sand

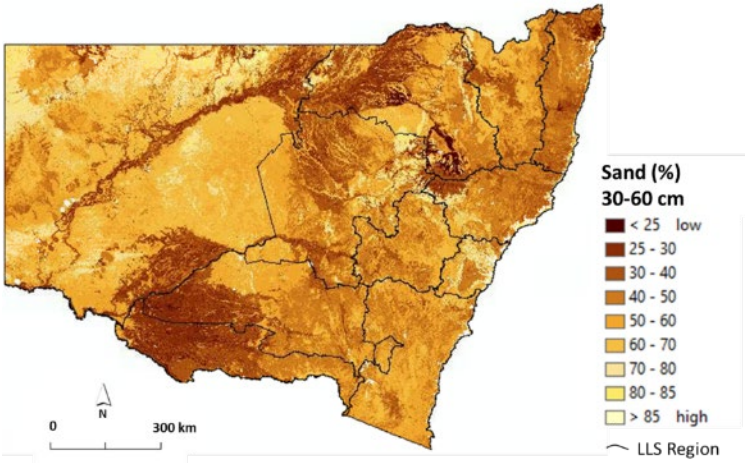
3.7.4 Total sand (%)



a: Total sand (%) 0–10 cm, mean[1]



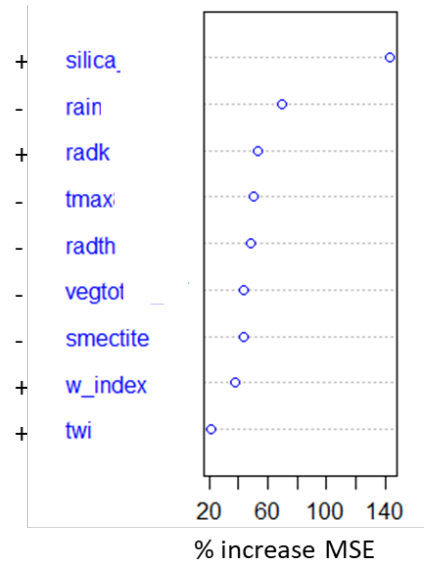
b: Total sand (%) 10–30 cm, mean[1]



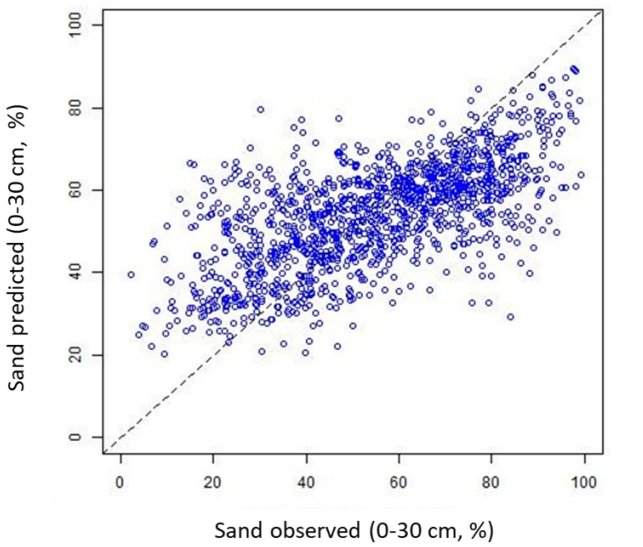
c: Total sand (%) 30–60 cm, mean[1]

Figure 20 Selection of maps for total sand %

[1] 90% prediction interval maps are provided in Appendix D.7.1



a: Total sand (%) 0–30 cm – variable importance plot (with direction of influence)



b: Total sand (%) 0–30 cm – map validation plot

Figure 21 Selection of plots for total sand (%)

Table 10 Map validation statistics for selection of layers: total sand %

(see Appendix C for all layers to 200 cm)

Depth (cm)	N	LCCC (%)	RMSE (%)	ME (%)	MAE (%)
0–10	1317	0.63	14.7	-1.0	11.5
10–30	1285	0.64	14.9	0.20	11.9
0–30	1317	0.57	16.1	–0.60	12.7
30–60	1154	0.53	17.2	0.9	13.6

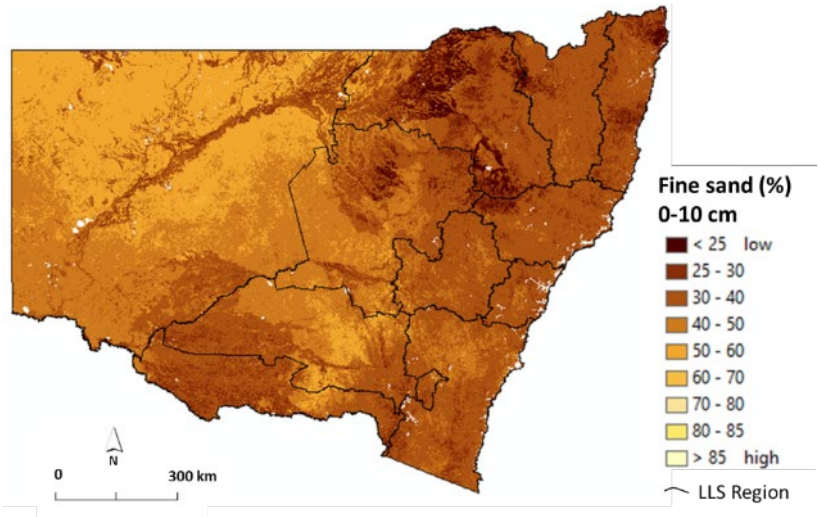
N: Validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

Summary

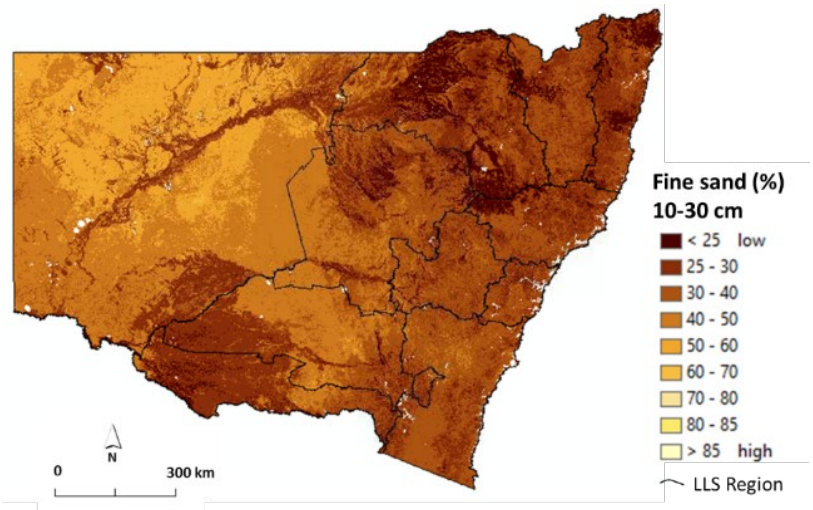
The maps for total sand (%) are of moderate statistical strength over all depth intervals, with LCCC values declining slightly with increasing depth, e.g. declining from 0.63 over the 0–10 cm interval to 0.52 for the 100–200 cm interval (see Appendix C). The other statistical indicators such as RMSE and MAE also suggest weaker performance with depth. Appendix E presents the summation of all particle sizes (excluding coarse fragments), and reveals moderate approximation to 100%, providing a degree of confidence in the reliability of these maps.

The maps reveal that total sand % remain essentially constant with depth. Soils become more sandy with increasingly siliceous parent materials, with the silica variable being the clearly dominant influence and displaying a positive trend in the variable importance plot in Figure 21a. The negative trend with rainfall suggests less sandy soils in high rainfall zones.

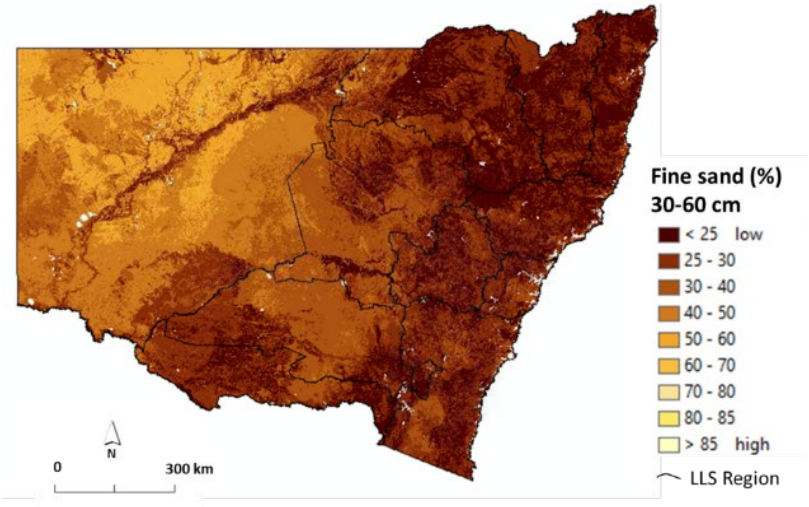
3.7.5 Fine sand (%)



a: Fine sand (%) 0–10 cm, mean[1]



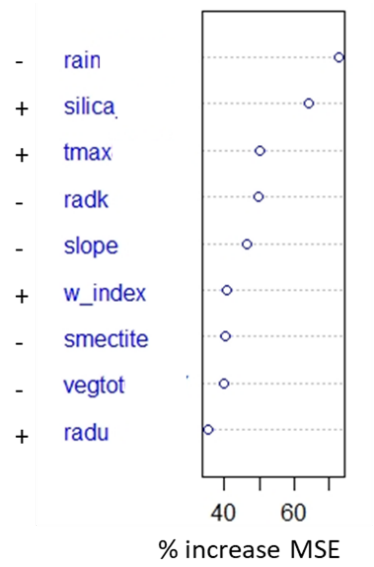
b: Fine sand (%) 10–30 cm, mean[1]



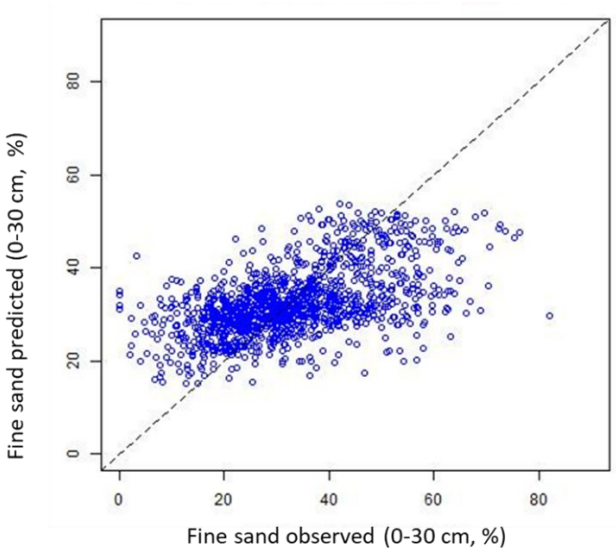
c: Fine sand (%) 30–60 cm, mean[1]

Figure 22 Selection of maps for fine sand %

[1]90% prediction interval maps are provided in Appendix D.7.2



a: Fine sand (%) 0–30 cm – variable importance plot (with direction of influence)



b: Fine sand (%) 0–30 cm – map validation plot

Figure 23 Selection of plots for fine sand %

Table 11 Map validation statistics for selection of layers: fine sand %

(see Appendix C for all layers to 200 cm)

Depth (cm)	N	LCCC (%)	RMSE (%)	ME (%)	MAE (%)
0–10	1320	0.45	11.6	0.5	9.0
10–30	1286	0.46	10.8	–0.4	8.4
0–30	1319	0.44	11.4	–0.3	8.7
30–60	1159	0.37	11.6	–0.4	8.8

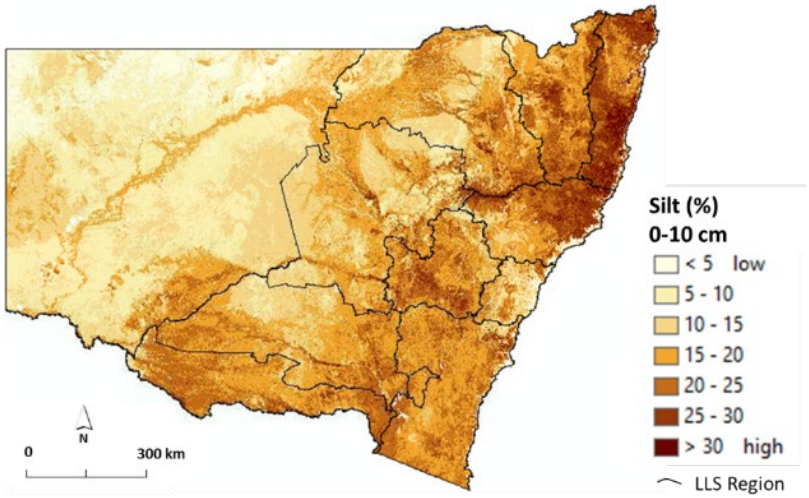
N: Validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

Summary

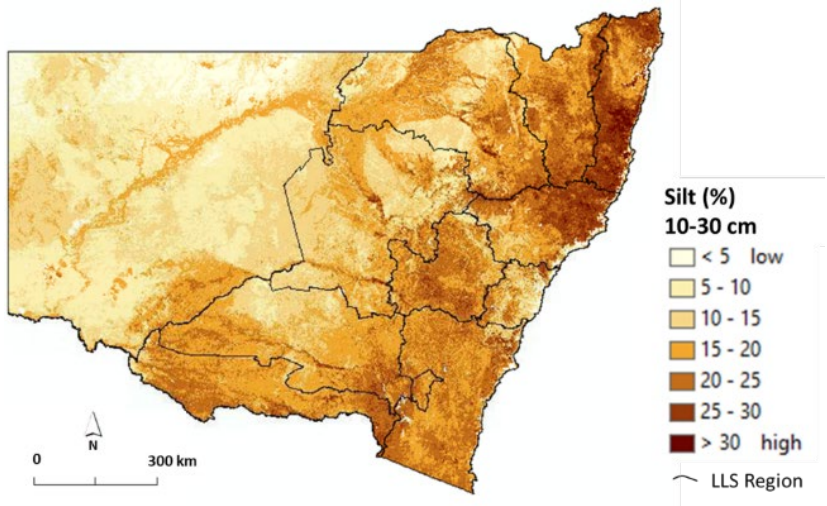
The maps for fine sand (%, 0.062–0.25 mm) are of only low statistical strength, with LCCC values ranging from approximately 0.45 in surface layers to as low as 0.34 in the deeper layers (see Appendix C) The other statistical indicators do not vary substantially with depth.

The maps reveal that fine sand % gradually increases with depth. As for total sand, soils increase in fine sand content with increasingly siliceous parent materials, with the silica variable being the clearly dominant influence and displaying a positive trend in the variable importance plot in Figure 23a. Similarly, the negative trend with rainfall suggests less fine sands in soils in high rainfall zones.

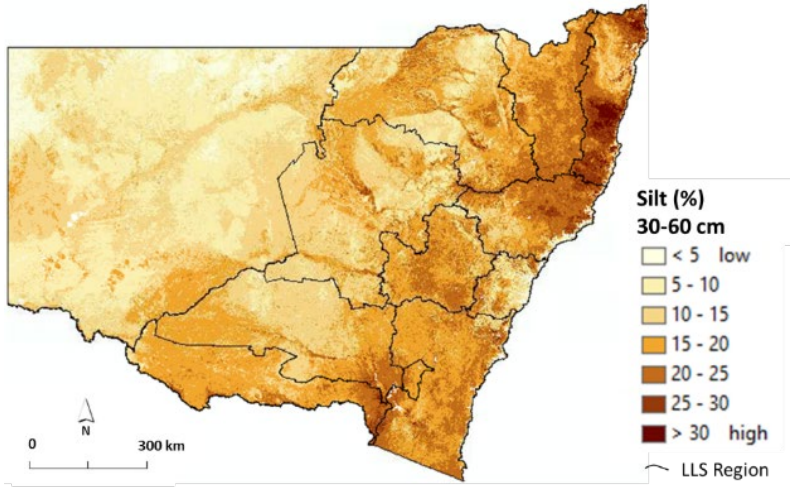
3.8 Silt (%)



a: Silt (%) 0–10 cm, mean[1]



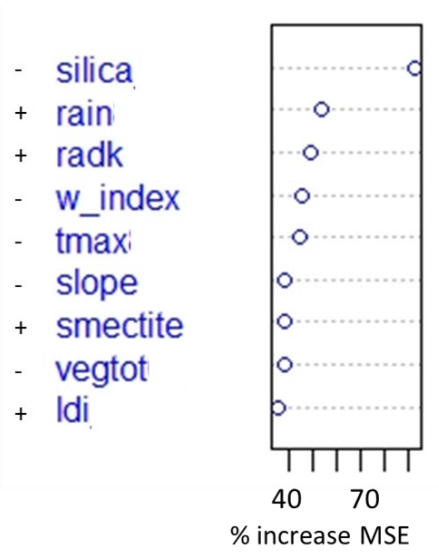
b: Silt (%) 10–30 cm, mean[1]



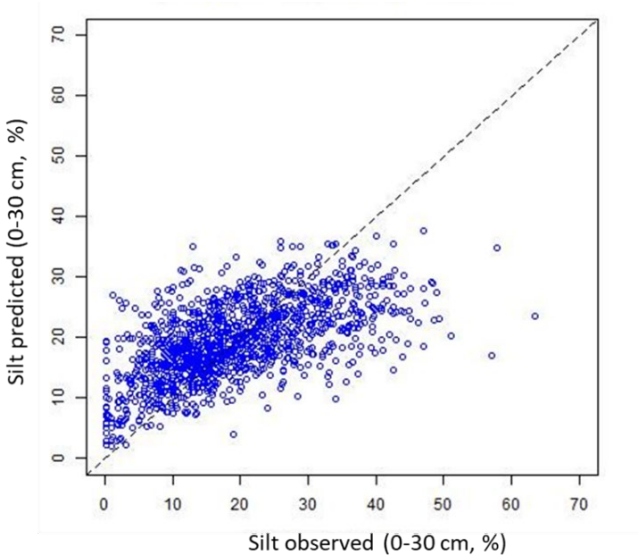
c: Silt (%) 30–60 cm, mean[1]

Figure 24 Selection of maps for silt %

[1] 90% prediction interval maps are provided in Appendix D.8



a: Silt (%) 0–30 cm – variable importance plot (with direction of influence)



b: Silt (%) 0–30 cm – map validation plot

Figure 25 Selection of plots for silt %

Table 12 Map validation statistics for selection of layers: silt %

(see Appendix C for all layers to 200 cm)

Depth (cm)	N	LCCC (%)	RMSE (%)	ME (%)	MAE (%)
0–10	1320	0.55	8.4	0.5	6.5
10–30	1286	0.55	8.5	0.4	6.4
0–30	1319	0.52	8.5	0.2	6.5
30–60	1159	0.49	8.4	0.1	6.4

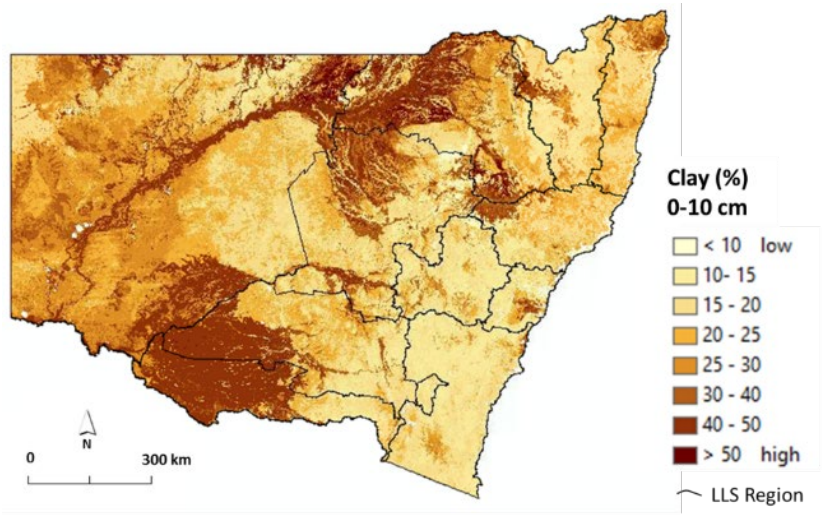
N: Validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

Summary

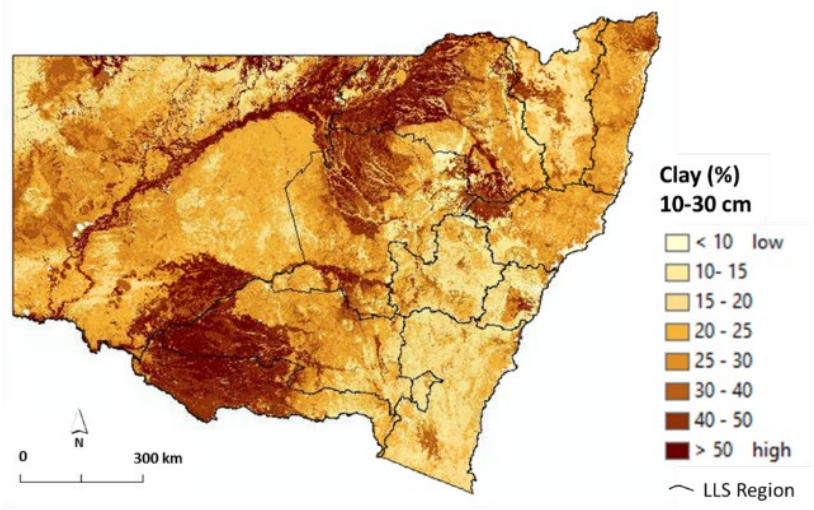
The maps for silt (%) are generally of moderate statistical strength down to 60 cm (LCCC 0.50–0.55) but drop to low strength (0.41 or less) in the layers below 60 cm (see Appendix C). A similar pattern is revealed by other statistical indicators such as RMSE and MAE.

The maps reveal that silt % generally decreases with depth. Soils become less silty with increasingly siliceous parent materials, with the silica variable being the clearly dominant influence and displaying a negative trend in the variable importance plot in Figure 25a. The positive trend with rainfall and negative trend with temperatures suggests more silty soils under higher rainfall and cooler temperature conditions.

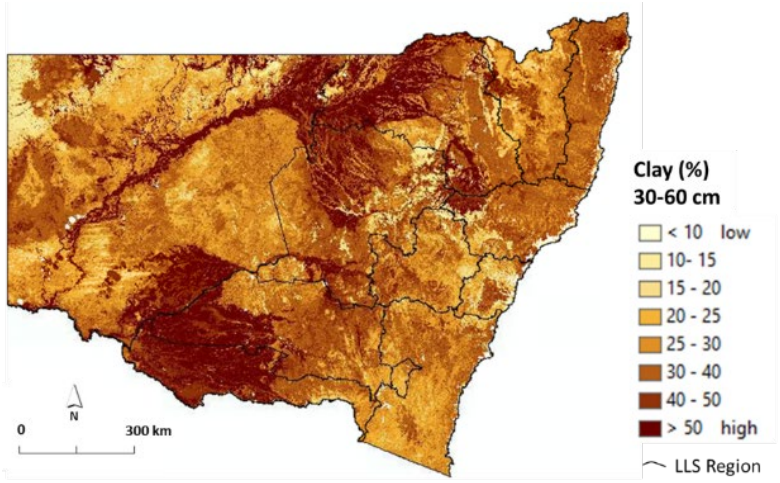
3.9 Clay (%)



a: Clay (%) 0–10 cm, mean[1]



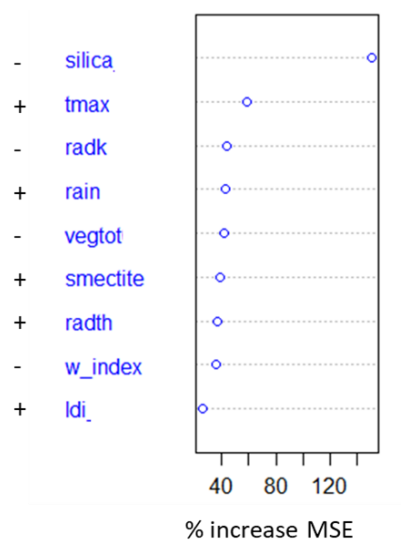
b: Clay (%) 10–30 cm, mean[1]



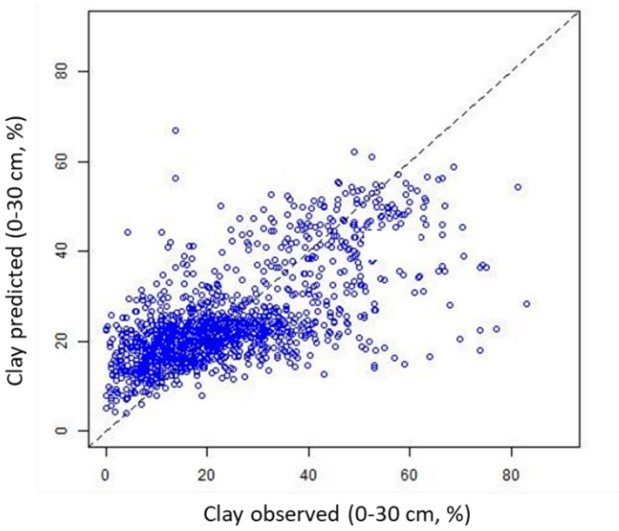
c: Clay (%) 30–60 cm, mean[1]

Figure 26 Selection of maps for clay %

[1] 90% prediction interval maps are provided in Appendix D.9



a: Clay (%) 0–30 cm – variable importance plot (with direction of influence)



b: Clay (%) 0–30 cm – map validation plot

Figure 27 **Selection of plots for clay %**

Table 13 **Map validation statistics for selection of layers: clay %**

(see Appendix C for all layers to 200 cm)

Depth (cm)	N	LCCC (%)	RMSE (%)	ME (%)	MAE (%)
0–10	1320	0.63	11.6	0.2	8.1
10–30	1285	0.62	12.1	–0.2	9.2
0–30	1319	0.63	11.7	0.1	8.8
30–60	1159	0.53	14.5	–0.2	11.5

N: Validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error

Summary

The maps for clay (%) are generally of moderate statistical strength down to 30 cm (LCCC 0.61–0.66) but are of only low to moderate strength below this (see Appendix C). A similar pattern is revealed by other statistical indicators such as RMSE and MAE.

The maps reveal that clay % generally increases with depth. Soils become less clay rich with increasingly siliceous parent materials, with the silica variable being the clearly dominant influence and displaying a negative trend in the variable importance plot in Figure 27a. The positive trend with both rainfall and temperatures suggests more clay rich soils under higher rainfall and warmer temperature conditions, possibly reflecting higher degrees of weathering of feldspathic minerals to clay in these zones.

4. Discussion

4.1 Map validation

Map validation results of the DSMs, using the randomly withdrawn validation datasets for each soil property/depth interval, are presented for selected depth intervals for each soil property in sections 3.1 to 3.9, and in Appendix C for all depth intervals.

The statistical indicators of predictive performance of the maps vary between the different soil properties and depth intervals. Moderate to strong values of the LCCC (0.7 to 0.85) are achieved for SOC (upper 30 cm), pH and sum-of-bases. Moderate concordance values (0.55 to 0.7) are typically achieved for CEC, BD, total sand and clay, while low to moderate concordance values (0.4 to 0.55) are evident for P(bray), fine sand and silt. Low concordance values (<0.4) are evident in deeper layers (below 60 cm) for SOC, fine sand and silt. This indication of performance is supported by the other statistical metrics of RMSE, MAE and MedAE.

The statistical values demonstrate varying patterns of strength with depth for different soil properties. Statistical performance clearly decreases with depth for SOC, total sand, silt and clay, but remains broadly constant for the other soil properties.

A further indication of reliability of the particle size maps, total sand, silt and clay, is the extent to which these 3 maps sum to 100% (excluding the coarse fragment component). The maps and table presented in Appendix E reveal a moderately high proportion of pixels had summed values close to 100%, with over 99% being within 10 percentage points of this target, suggesting moderately strong performance in this respect.

4.2 Strategy for use

The DSMs presented here provide useful first approximations of key soil properties at different depth intervals across NSW, down to 100 m spatial resolutions.

The maps are best viewed with GIS software after downloading from SEED, in either continuous or categorical format, together with other infrastructure or terrain layers to reliably locate images (see Figure 2 and Figure 3). Categorical map legends can be applied with use of ‘layer files’ as also provided in the SEED packages. The precise estimate for any particular pixel can be viewed using the GIS ‘identify’ tool. The maps may also be viewed through eSPADE, where they may be simultaneously viewed with underlying infrastructure, terrain or satellite image layers. Images of selected depth intervals are presented in sections 3.1 to 3.9.

The validation results are generally indicative of at least moderate map reliability, except at deeper levels for some soil properties. The 90% prediction interval maps as presented in Appendix D, together with the 5% and 95% prediction limit maps (available on SEED) also provide an indication of the uncertainties associated with the maps. More sophisticated techniques, such as those applied by Malone et al. (2014) and Kidd et al. (2015), may be adopted in the future to derive more certain prediction limits.

By applying the maps in conjunction with the variable importance plots (as presented in section 3), an understanding can be gained of the influencing factors behind the spatial distribution of the soil properties. For example, rising pH levels can be seen to be associated with drier conditions (lower rainfall and higher temperatures), less siliceous parent materials and more agriculturally intensive land uses.

4.3 Limitations

There are a number of limitations and sources of uncertainty associated with these DSMs, including:

- **inadequate representativeness of training data**: the data may not be entirely representative of all environments, i.e. combinations of covariate space. Some environments may not be sufficiently well sampled, as there was no formal sampling design during collection of the data points over the soil survey program covering several decades. A sampling design such as a Latin hypercube approach (Minasny and McBratney 2006) can help address this problem. The tendency of sampling points to be in easily accessible locations rather than being randomly selected also introduces modelling weaknesses
- **weaknesses in categorical training data**: there may be considerable variation within many of the classes used in model and map development. Each lithology class (or silica index) actually represents a range of compositions. Additionally, some parent materials can be difficult to reliably classify, e.g. all shales were assumed to be upper-intermediate class (silica index 62%), whereas in reality they may vary between lower-intermediate class to mid-siliceous (silica index 57–73%). The LDI is a very coarse indicator of land use, land management and biotic conditions at a site
- **weakness in covariate data grids for map production**: errors will occur due to the scale (i.e. pixel or polygon size), particularly with coarse-scale polygonal datasets. For example, there may be considerable lithological variation within individual geological units and the associated map polygons. Likewise, land use may vary from that identified in the land-use grid. The Bureau of Meteorology climate grids required downscaling from the original 2.5 km to 100 m resolution
- **remote sensed data errors**: data such as the gamma radiometrics and hyperspectral derived clay composition have uncertainties arising from the complexity of geophysical–soil relationships, the fact signals usually only relate to soil properties of the top few millimetres, and high noise-to-signal ratios (due to issues such as coarse resolutions and interference by water and vegetation) (McBratney et al. 2003; Mulder et al. 2011)
- **time issues**: the LDI does not consider the period of time a new land-use regime has been in operation, thus soil property imprints from a previous land use may still be evident in the analysed soil. The current recorded land use and groundcover may differ from that at the time of profile collection. An issue arises when soils are old enough to have been influenced by previous climatic conditions and they therefore carry an imprint from those conditions, rather than being entirely influenced by current climate conditions. Soil carbon, in particular, is influenced even by recent climate conditions, such as the previous few years
- **laboratory analysis errors**: including sample collection and handling errors, differences due to different laboratory techniques used and, for SOC, possible under-estimation of values by the Walkley–Black method (Skjemstad et al. 2000).

Inherent weaknesses and uncertainties in DSM are discussed more broadly in Nelson et al. (2011), Bishop et al. (2015) and Robinson et al. (2015).

Despite these weaknesses, the validation results are promising and suggest the models and resulting maps are useful in providing first approximations of a range of soil properties across the landscape in NSW.

4.4 Relationship to existing DSM and conventional soil survey products

The DSMs for NSW presented here represent an alternative and complementary product to the Australia-wide maps presented within the updated SLGA (Grundy et al. 2015; Viscarra Rossel et al. 2015). Both products are at similar scale and have the same depth intervals; however, the SLGA used slightly different modelling techniques and applied data from across all of Australia, while the NSW maps presented here are based on NSW data only. The SLGA deals with a slightly wider range of soil properties than this NSW project, with additional maps for total nitrogen, available water-holding capacity, liquid limits, coarse fragments, soil depth and depth to hard rock. The updated SLGA also includes nationwide maps for soil carbon fractions, which compare with NSW maps of these fractions, also available on SEED (Gray et al. 2019).

It has been contended by the SLGA Working Group that the bringing together of several DSM products derived through different methods and data sources can lead to an ultimately more reliable product (pers. comm., SLGA Working Group meeting, Canberra, 28 October 2014). This is exemplified by the combination of national and state-specific products in the primary product presented in the SLGA (Grundy et al. 2015). These NSW maps appear to compare well with the national SLGA maps, certainly in their broad trends of spatial distribution of soil properties. These new maps provide a further line of evidence for DSM results for the various soil properties.

The maps also complement the conventional polygon-based soil landscape maps available over NSW; for example, through eSPADE. Both products have their advantages and disadvantages. On the one hand, the DSMs provide detail of variation in soil properties within polygons. They also cover a range of soil properties that are not consistently available across the whole state from conventional mapping. The DSMs are particularly useful where no detailed conventional soil mapping has been completed. The raster products can be more easily incorporated into many other environmental modelling programs, such as spatial biodiversity modelling.

On the other hand, the soil landscape maps provide a more holistic overview of soil character than the individual soil properties of DSMs. They provide vital details on key soil character that are generally less well covered by these DSMs or by DSM projects more generally. For example, they provide detail on soil type (Australian Soil Classification [ASC], Great Soil Group [GSG] or other system), profile type (uniform, gradational or texture contrast soil forms), profile depths, colour, soil structure, coarse fragments and various other macro- and micro-morphological features. Importantly, they combine the different soil properties together into a single holistic soil description, which can be more useful and powerful than dealing with separate individual soil properties.

The polygonal format of the conventional maps typically lends itself better to applying land-use and management recommendations on the ground. Many land managers and regional planners find it useful to have discrete polygons with similar land management requirements identified for them on a map or other spatial layer, rather than a set of purely raster-based products that may require additional interpretation.

4.5 Conclusion

DSMs have been produced for a range of key soil properties over NSW, with coverage over multiple soil depth intervals down to 2 m. Validation results of the maps indicate generally moderate to high effectiveness. They are useful in providing at least a first approximation of these properties. The maps may be viewed directly through the department's spatial viewer eSPADE and are also available for download as GIS files through the SEED environmental data portal.

Further development of the DSM techniques applied in this project, including use of other emerging remote-sensed datasets, and application of more sophisticated measures of map uncertainty, may improve the quality and effectiveness of these maps.

The DSMs are a potentially useful complementary product to the existing DSMs and conventional soil landscape products available for much of NSW. Both the digital and conventional soil products can be used together to help inform on soil conditions throughout the state. They can thus assist in the ongoing sustainable management and protection of this vital natural resource. They also provide valuable input data for a range of other natural resource, environmental and climate change modelling systems, critical for effective environmental management across NSW.

Appendix A: Covariate layers

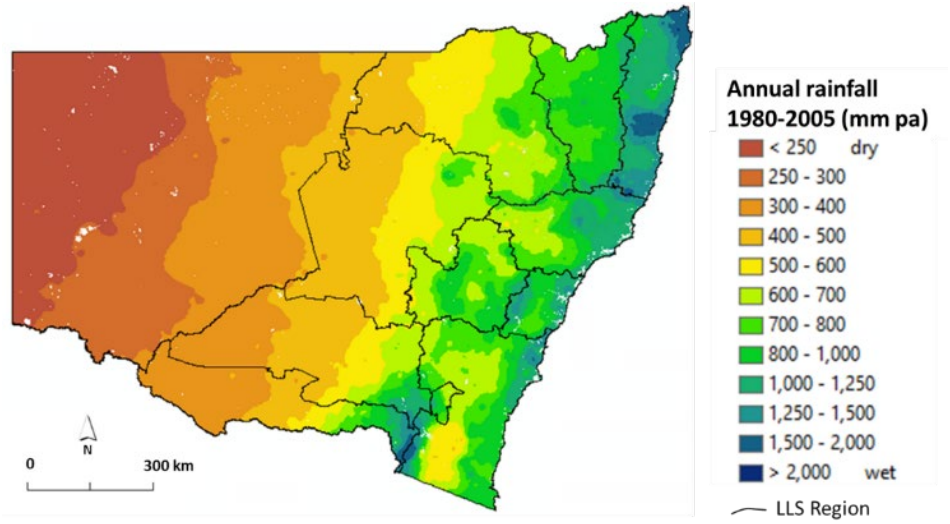


Figure A.1 Annual rainfall, 1980–2005 (mm p.a.)

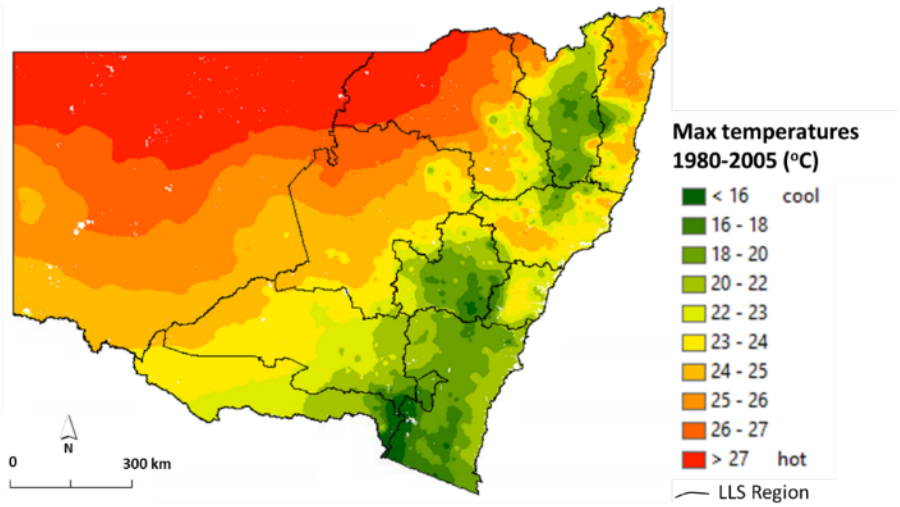


Figure A.2 Mean annual daily maximum temperature, 1980–2005 (°C)

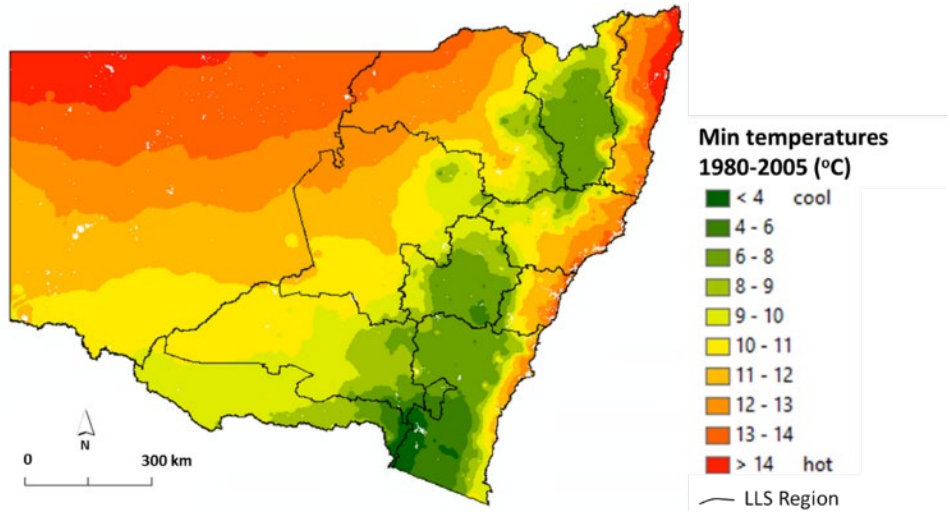


Figure A.3 Mean annual daily minimum temperature, 1980–2005 (°C)

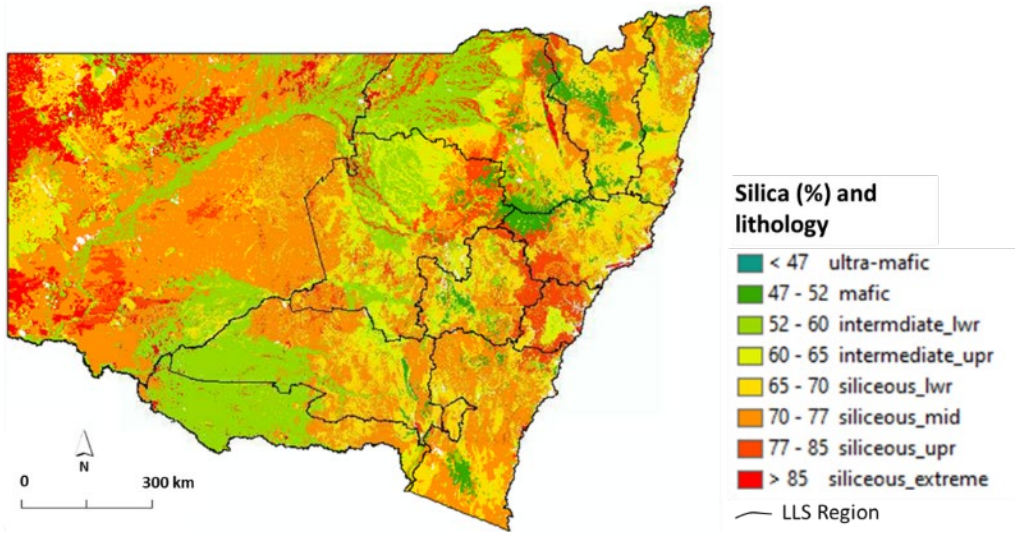


Figure A.4 Silica % and lithology

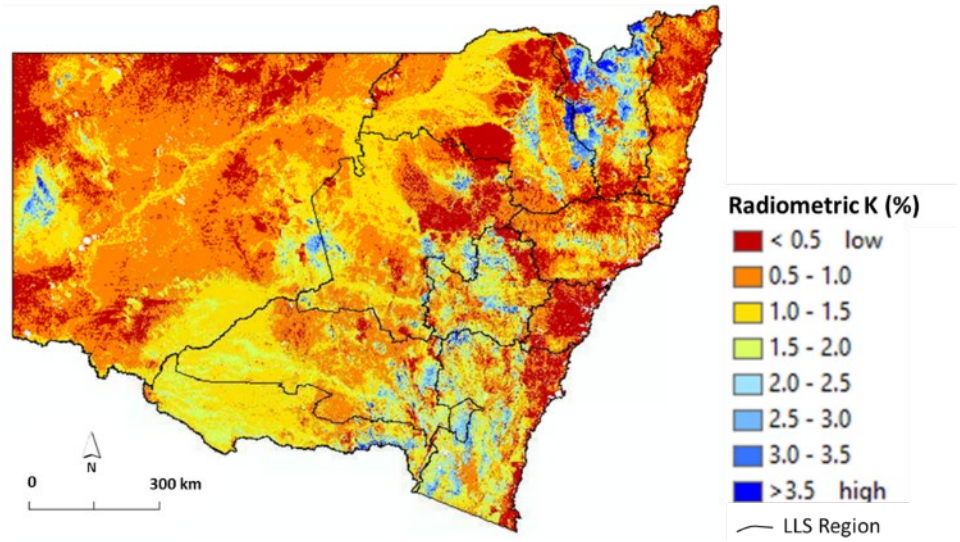


Figure A.5 Radiometric K (%)

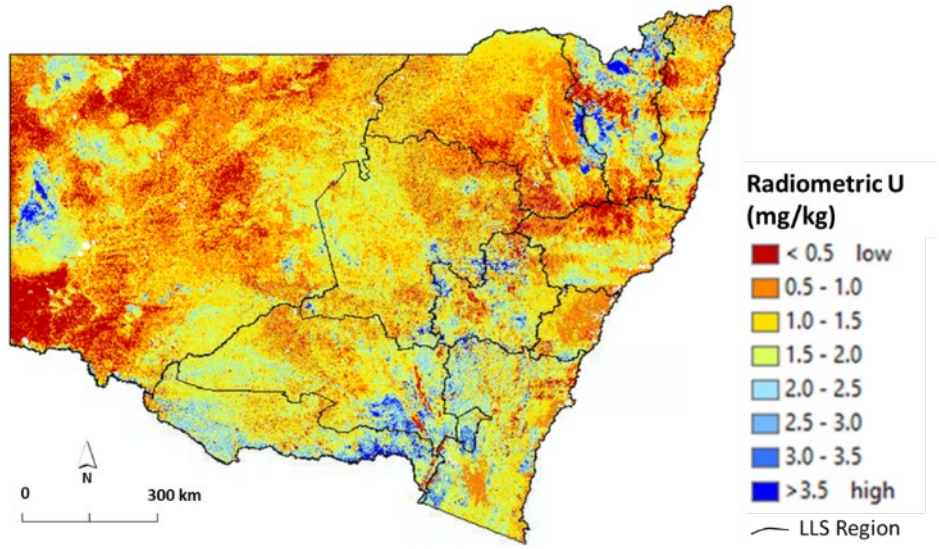


Figure A.6 Radiometric U (mg/kg)

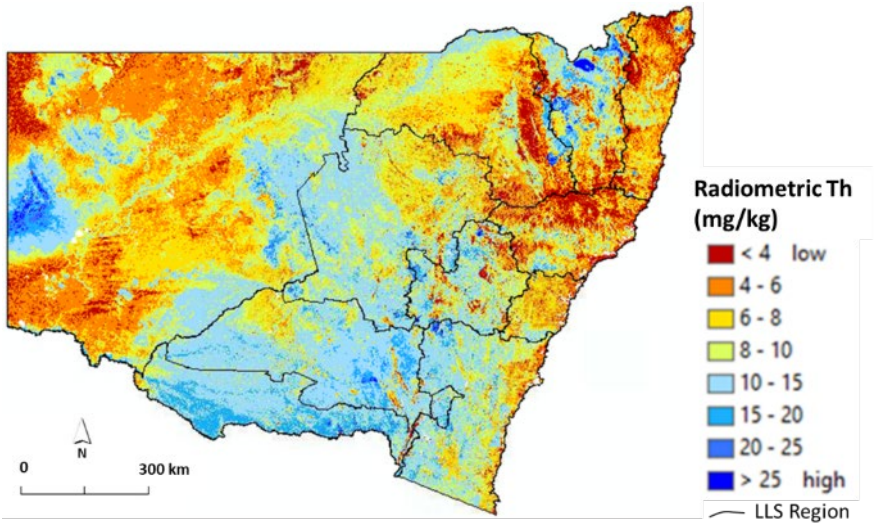


Figure A.7 Radiometric Th (mg/kg)

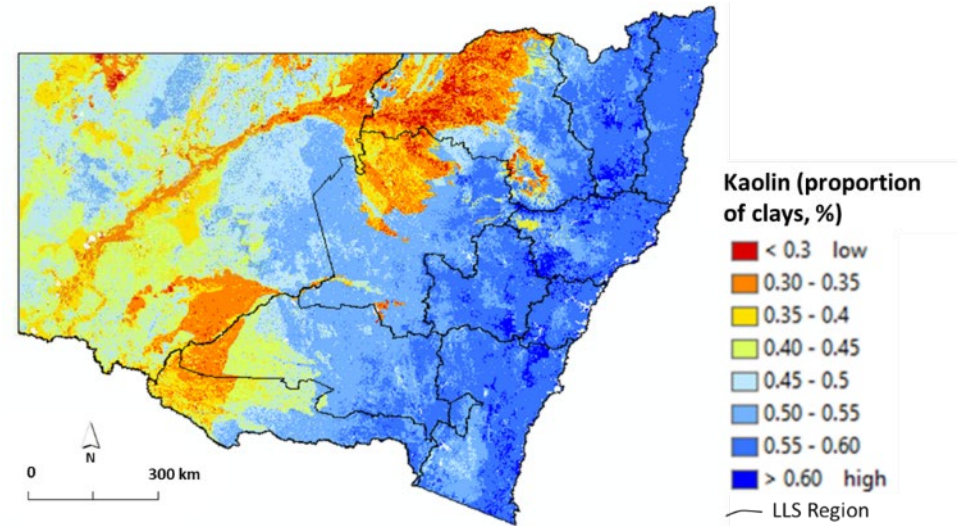


Figure A.8 Kaolin (component fraction)

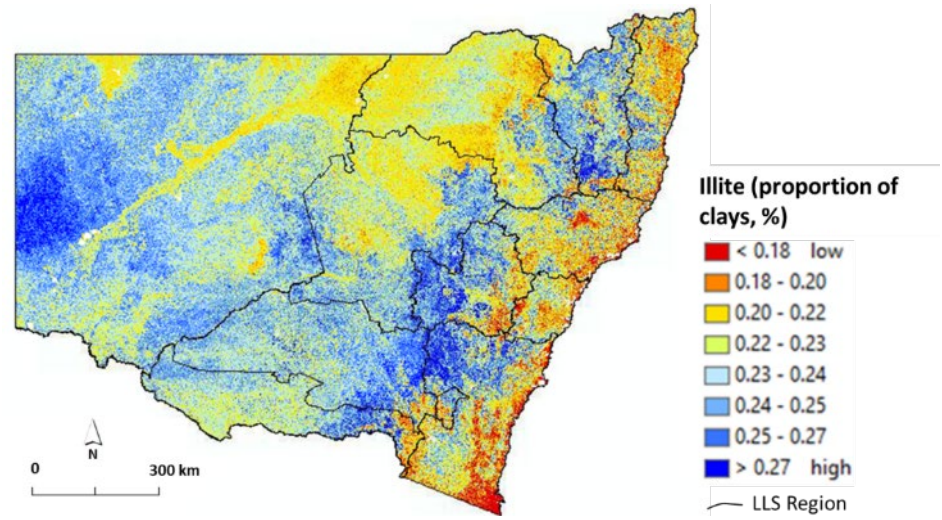


Figure A.9 Illite (component fraction)

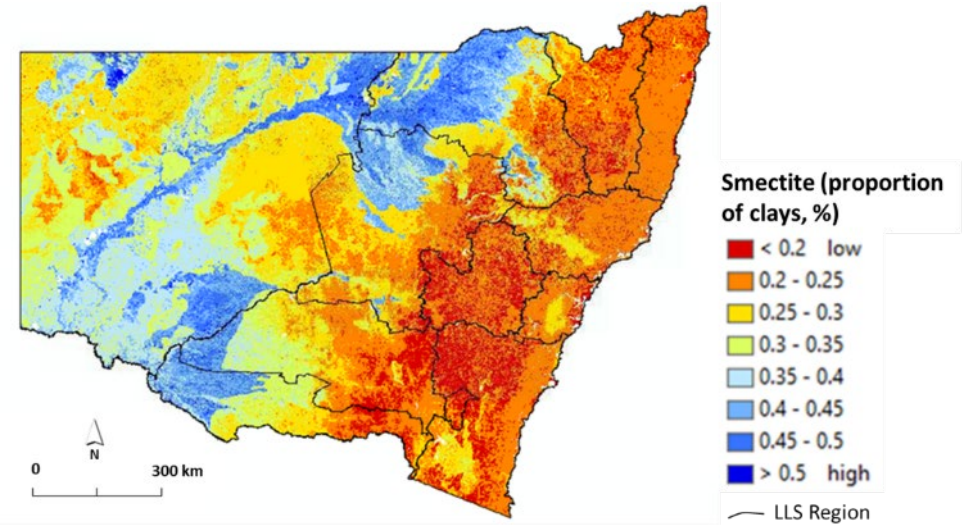


Figure A.10 Smectite (component fraction)

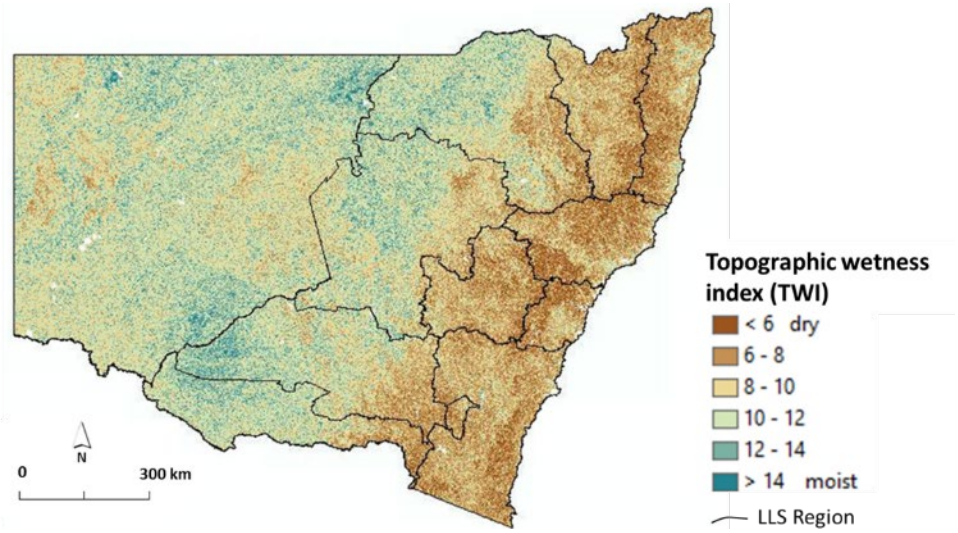


Figure A.11 Topographic wetness index

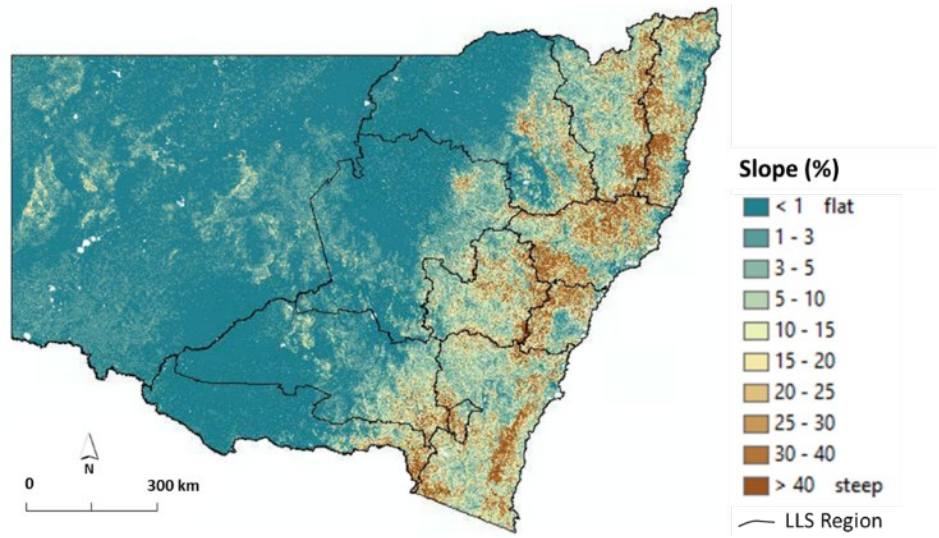


Figure A.12 Slope (%)

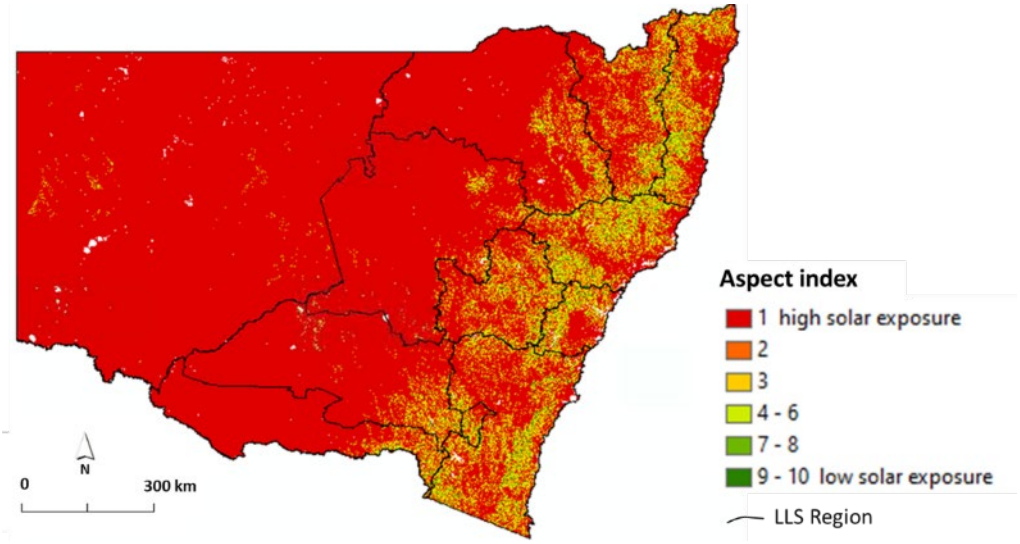


Figure A.13 Aspect index

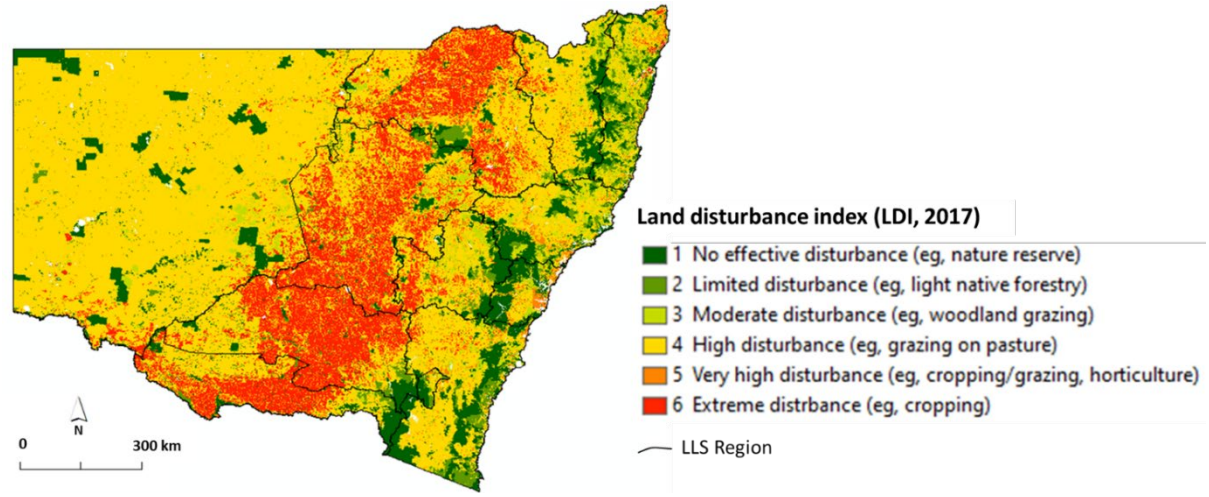


Figure A.14 Land disturbance index, 2017 (LDI)

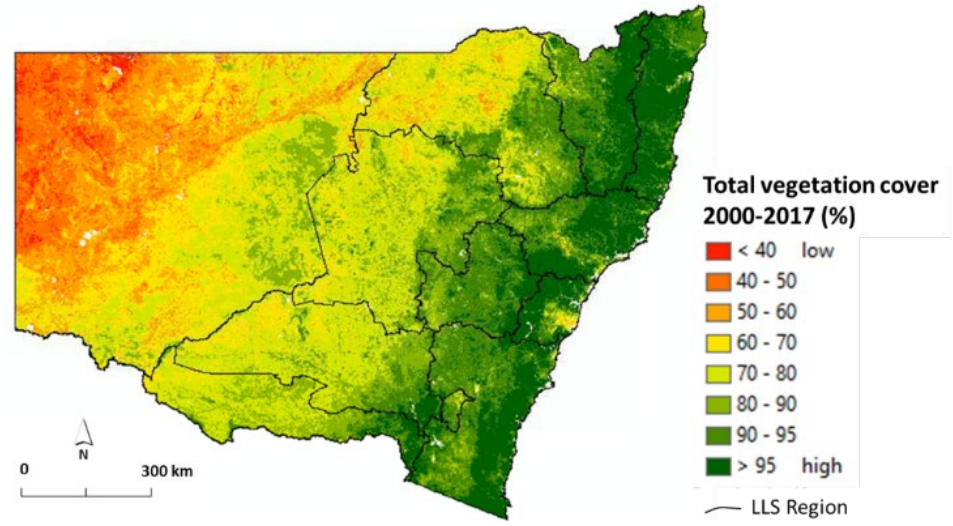


Figure A.15 Total vegetation cover, 2000–2017 (%)

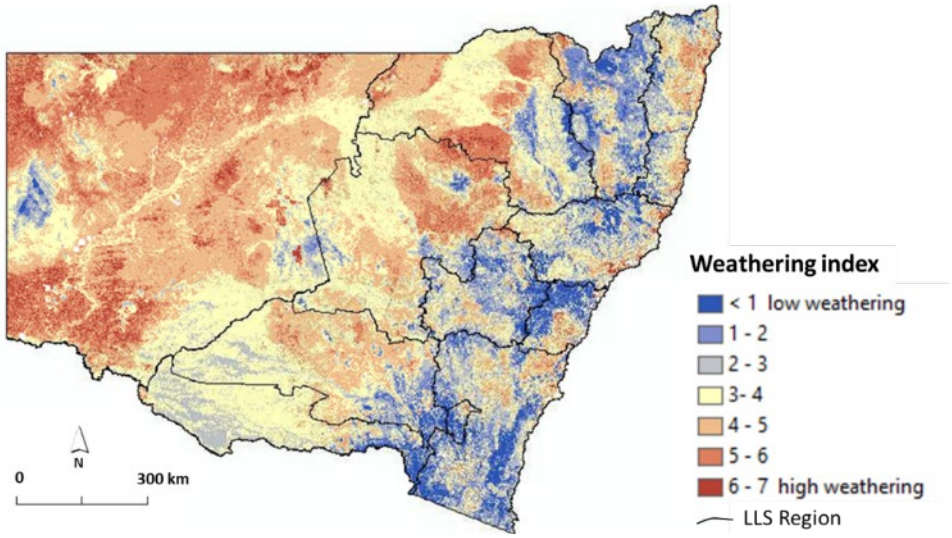


Figure A.16 Weathering index

Appendix B: Influence of bootstrap numbers

Influence of bootstrap sample number (n) on prediction limit statistics in RF models

The following table presents predictions derived from 3 validation pixels for SOC concentration (% 0–30 cm) using differing bootstrap numbers (n). Each bootstrap run comprised 200 trees (iterations) in the RF modelling, thus 10 bootstraps equate to 2,000 iterations.

Validation point 1

Vector name	Sample (bootstrap) size	Mean	SD	LPL (90% conf)	UPL (90% conf)
Bst10	10	0.68	0.036	0.61	0.75
Bst25	25	0.67	0.029	0.62	0.72
Bst50	50	0.67	0.031	0.62	0.72
Bst100	100	0.67	0.037	0.61	0.73

Validation point 2

Vector name	Sample (bootstrap) size	Mean	SD	LPL (90% conf)	UPL (90% conf)
Bst10	10	2.15	0.139	1.89	2.42
Bst25	25	2.06	0.151	1.81	2.33
Bst50	50	2.06	0.147	1.81	2.31
Bst100	100	2.09	0.144	1.85	2.33

Validation point 3

Vector name	Sample (bootstrap) size	Mean	SD	LPL (90% conf)	UPL (90% conf)
Bst10	10	0.99	0.105	0.79	1.20
Bst25	25	1.00	0.110	0.81	1.19
Bst50	50	1.00	0.105	0.82	1.18
Bst100	100	1.018	0.098	0.85	1.18

It is evident there is no significant difference between statistics using different n values between 10 and 100, nor are there consistent trends depending on n.

Figure B.1 presents outputs from 2 random points on maps of SOC% 0–30 cm over ACT, using 10, 25, 50 and 100 bootstrap runs. The points show:

- lower 5% prediction limit (PL)
- mean
- upper 95% prediction limit.

The outputs reveal no significant difference in values between the different bootstrap values. Note each bootstrap run comprised 200 trees (iterations).

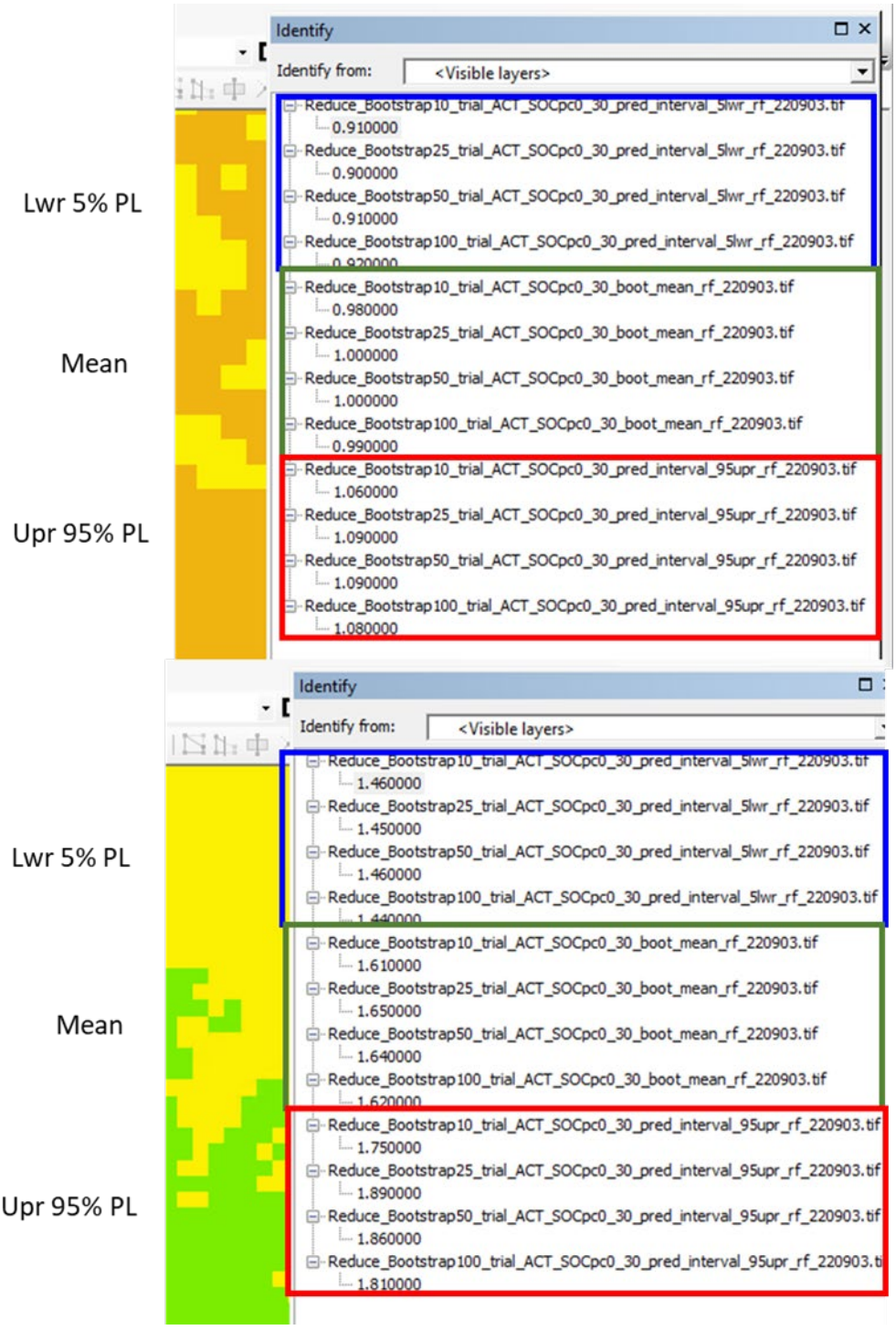


Figure B.1 Mean and prediction limits from 2 random points over ACT SOC% (0–30 cm) maps with different bootstrap numbers

Figure B.2 presents the reduction in prediction error that occurs with increasing number of trees in RF models for SOC% 0–30 cm. This study used 200 trees in each bootstrap run, after which the reduction in error was insignificant.

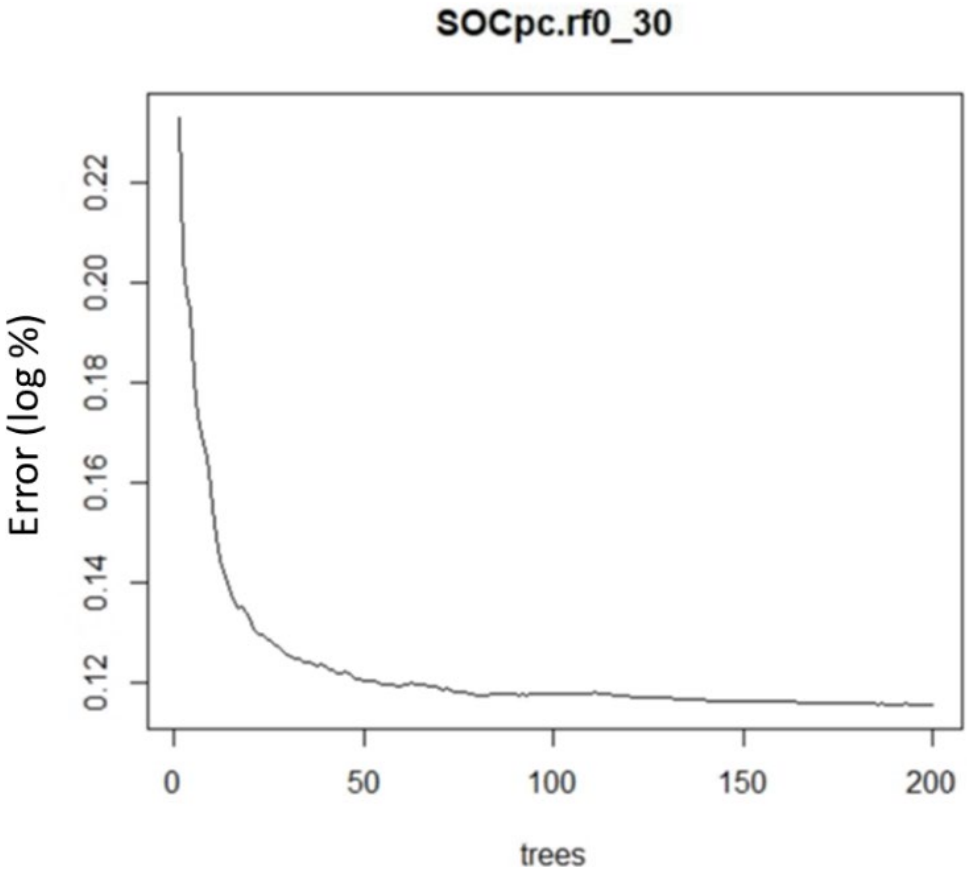


Figure B.2 **Reduction in prediction error with number of trees in each RF run: NSW SOC% 0–30 cm**

Appendix C: Map validation results (all depths)

Property	Depth (cm)	N	LCCC	RMSE	ME	MAE	MedAE
SOC concent'n (log units)	0–10	415	0.82	0.36	–0.02	0.27	0.20
	10–30	415	0.73	0.41	–0.004	0.30	0.23
	0–30	417	0.82	0.34	–0.01	0.25	0.18
	0–5	416	0.81	0.40	–0.004	0.30	0.24
	5–15	415	0.82	0.33	–0.01	0.25	0.19
	15–30	415	0.76	0.38	0.008	0.28	0.20
	30–60	1,177	0.32	0.97	–0.02	0.65	0.44
	60–100	909	0.18	1.11	–0.01	0.78	0.52
	0–100	417	0.81	0.22	0.003	0.16	0.12
SOC mass (log units)	0–10	417	0.83	0.33	–0.02	0.25	0.19
	10–30	418	0.77	0.36	0.0045	0.27	0.21
	0–30	418	0.84	0.32	–0.02	0.24	0.19
SOC stock (log units)	0–10	418	0.84	0.33	–0.01	0.24	0.18
	0–30	430	0.80	0.33	0.005	0.26	0.19
	0–100	431	0.78	0.23	–0.003	0.16	0.11
pH (pH units $CaCl_2$)	0–10	2,098	0.74	0.67	0.02	0.50	0.39
	10–30	2,025	0.80	0.66	–0.002	0.50	0.39
	0–30	2,031	0.76	0.70	0.005	0.51	0.37
	0–5	2,098	0.70	0.70	0.02	0.53	0.41
	5–15	2,089	0.76	0.67	0.002	0.50	0.38
	15–30	1,997	0.81	0.68	0.005	0.51	0.40
	30–60	1,710	0.82	0.74	–0.02	0.57	0.44
	60–100	1,330	0.82	0.84	0.01	0.63	0.47
	100–200	549	0.84	0.83	–0.08	0.63	0.50
CEC (log units)	0–10	1,113	0.57	0.68	0.002	0.47	0.34
	10–30	1,059	0.69	0.59	–0.005	0.44	0.35
	0–30	1,247	0.70	0.57	–0.04	0.43	0.35
	0–5	1,247	0.65	0.62	–0.004	0.45	0.35
	5–15	1,241	0.67	0.60	–0.008	0.45	0.45
	15–30	1,190	0.69	0.60	0.007	0.46	0.38
	30–60	1,120	0.67	0.67	0.05	0.48	0.36
	60–100	885	070	0.65	–0.02	0.47	0.36
	100–200	342	0.65	0.75	0.09	0.49	0.32

Property	Depth (cm)	N	LCCC	RMSE	ME	MAE	MedAE
Sum-of-bases (log units)	0–10	1,323	0.69	0.68	−0.004	0.51	0.40
	10–30	1,289	0.74	0.68	0.03	0.52	0.40
	0–30	1,324	0.72	0.67	−0.008	0.52	0.42
	0–5	1,323	0.71	0.68	0.02	0.52	0.41
	5–15	1,318	0.73	0.65	−0.02	0.50	0.41
	15–30	1,266	0.75	0.69	−0.01	0.53	0.43
	30–60	1,189	0.76	0.71	−0.007	0.53	0.40
	60–100	946	0.76	0.74	0.004	0.54	0.41
	100–200	369	0.77	0.75	0.03	0.56	0.43
P(bray) (mg/kg)	0–10	845	0.42	0.85	−0.004	0.68	0.58
	10–30	808	0.49	0.81	−0.05	0.63	0.51
	0–30	801	0.52	0.76	−0.007	0.60	0.52
	0–5	854	0.47	0.82	0.02	0.66	0.56
	5–15	858	0.48	0.80	0.04	0.64	0.55
	15–30	723	0.52	0.80	0.04	0.63	0.53
	30–60	560	0.48	0.88	0.09	0.69	0.57
	60–100	337	0.50	0.94	0.15	0.72	0.57
	100–200	184	0.55	1.00	0.03	0.81	0.70
Bulk density (Mg/m^3)	0–10	303	0.51	0.18	0.01	0.14	0.11
	10–30	294	0.56	0.17	0.001	0.13	0.10
	0–30 cm	296	0.70	0.13	0.001	0.09	0.07
Sand total (%)	0–10	1,317	0.63	14.7	−1.0	11.5	9.3
	10–30	1,285	0.64	14.9	0.20	11.9	9.9
	0–30	1,317	0.57	16.1	−0.60	12.7	10.8
	0–5	1,315	0.58	15.0	−0.5	11.7	9.9
	5–15	1,311	0.60	15.6	0.1	12.3	10.1
	15–30	1,261	0.60	15.6	−0.5	12.4	10.3
	30–60	1,154	0.53	17.2	0.9	13.6	11.1
	60–100	874	0.49	18.6	0.2	14.6	11.7
	100–200	388	0.52	19.1	0.7	15.2	12.6
Sand fine (%)	0–10	1,320	0.45	11.6	0.5	9.0	7.4
	10–30	1,286	0.46	10.8	−0.4	8.4	6.8
	0–30	1,319	0.44	11.4	−0.3	8.7	6.8
	0–5	1,320	0.41	12.3	0.05	9.6	7.9
	5–15	1,313	0.47	11.4	−0.2	8.8	7.0
	15–30	1,264	0.41	11.4	0.3	8.8	7.2
	30–60	1,159	0.37	11.6	−0.4	8.8	7.3
	60–100	875	0.34	11.8	0.5	9.4	7.7
	100–200	393	0.38	11.0	1.0	7.2	8.6

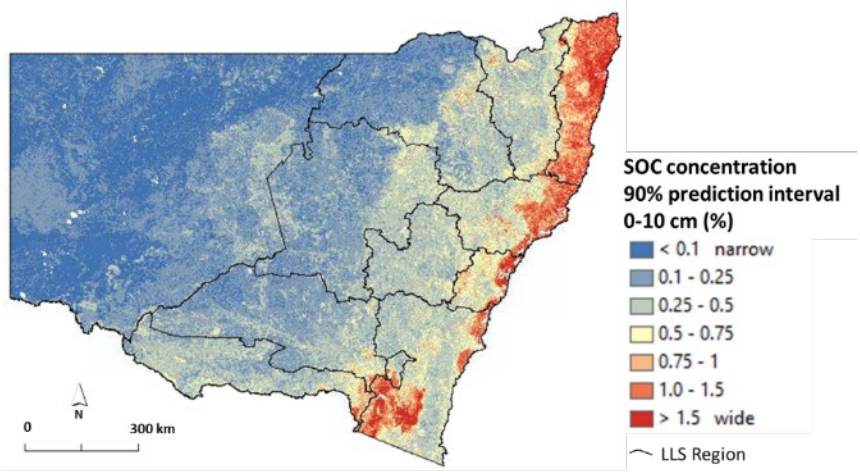
Property	Depth (cm)	N	LCCC	RMSE	ME	MAE	MedAE
Silt (%)	0–10	1,320	0.55	8.4	0.5	6.5	5.2
	10–30	1,286	0.55	8.5	0.4	6.4	5.0
	0–30	1,319	0.52	8.5	0.2	6.5	5.3
	0–5	1,320	0.54	8.6	0.33	6..6	5..4
	5–15	1,313	0.54	8.8	0.37	6.7	5.3
	15–30	1,264	0.51	9.0	0.06	6.8	5.4
	30–60	1,159	0.49	8.4	0.1	6.4	5.0
	60–100	875	0.41	8.7	−0.004	6.5	5.0
	100–200	393	0.32	9.5	0.4	7.0	5.5
Clay (%)	0–10	1,320	0.63	11.6	0.2	8.1	6.0
	10–30	1,285	0.62	12.1	−0.2	9.2	7.2
	0–30	1,319	0.63	11.7	0.1	8.8	7.1
	0–5	1,320	0.61	11.2	0.4	8.2	6.2
	5–15	1,313	0.66	10.9	0.003	8.2	6.6
	15–30	1,264	0.61	13.1	0.4	10.1	8.2
	30–60	1,159	0.53	14.5	−0.2	11.5	9.9
	60–100	875	0.43	17.2	1.0	13.7	11.4
	100–200	393	0.51	16.2	0.50	12.8	10.0

N: validation sample number; LCCC: Lin's concordance correlation coefficient; RMSE: root mean square error; ME: mean error (positive means predictions overestimate); MAE: mean absolute error; MedAE: median absolute error

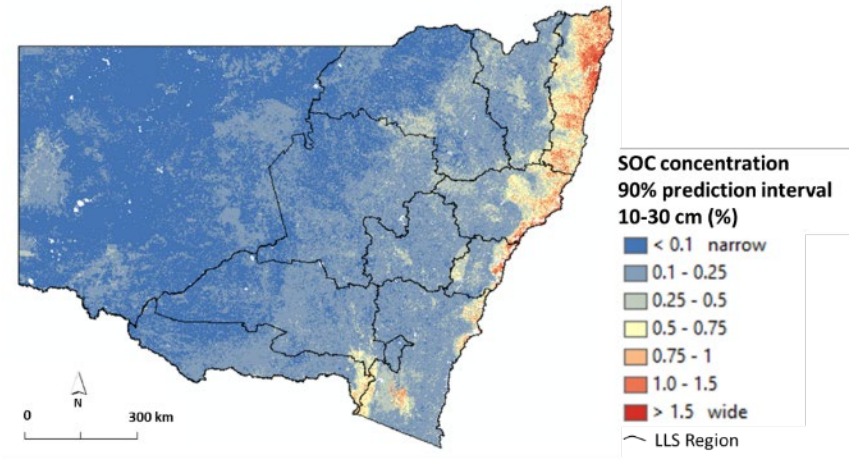
Appendix D: 90% prediction intervals

D.1 Soil organic carbon

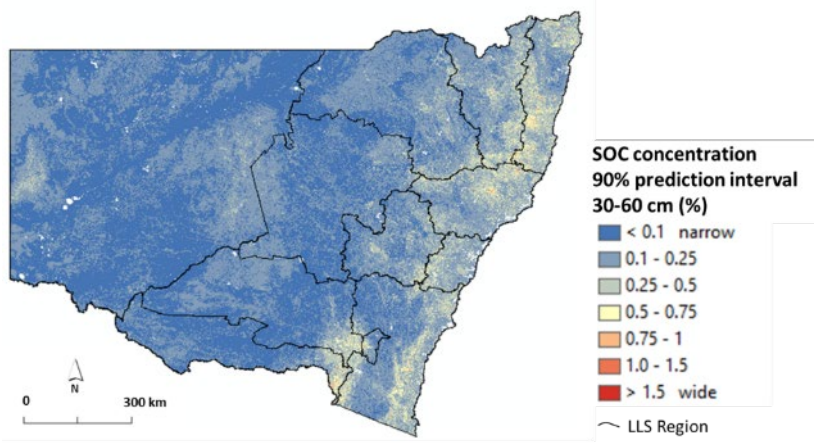
D.1.1 SOC %



a: SOC % 0–10 cm

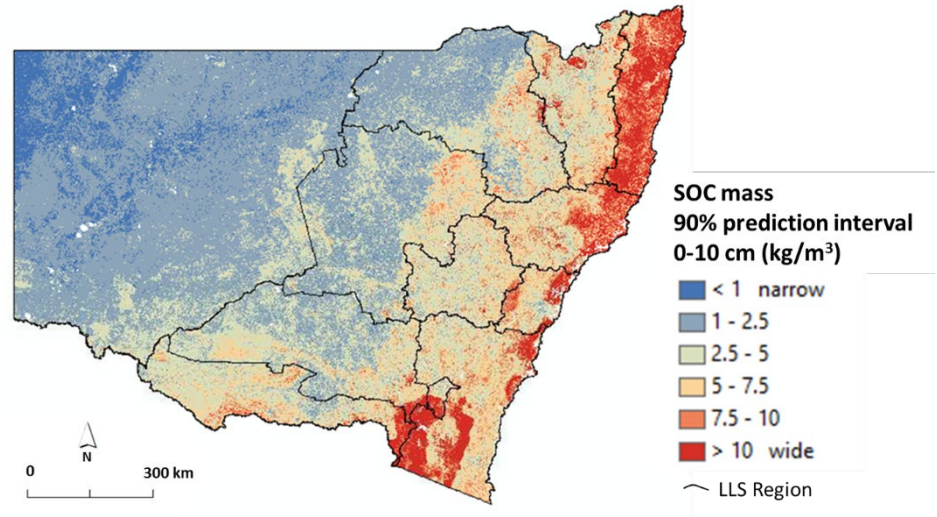


b: SOC % 10–30 cm

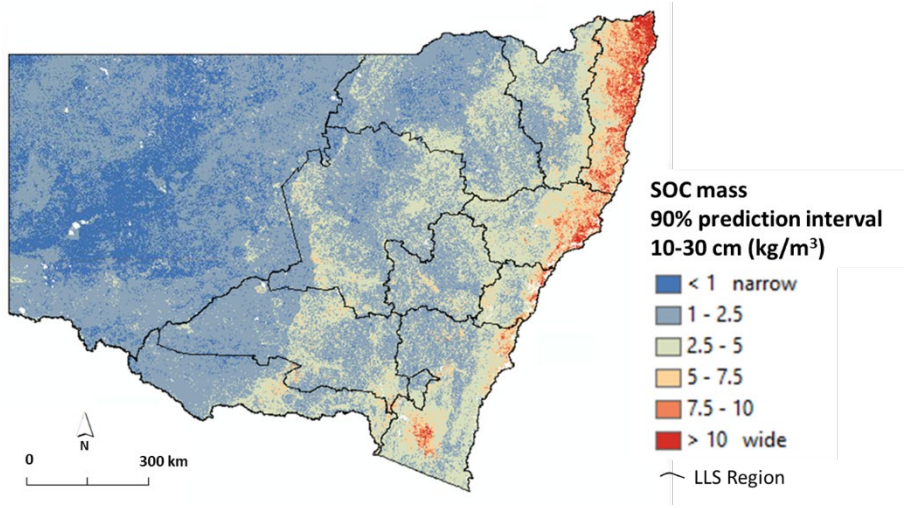


c. SOC % 30–60 cm

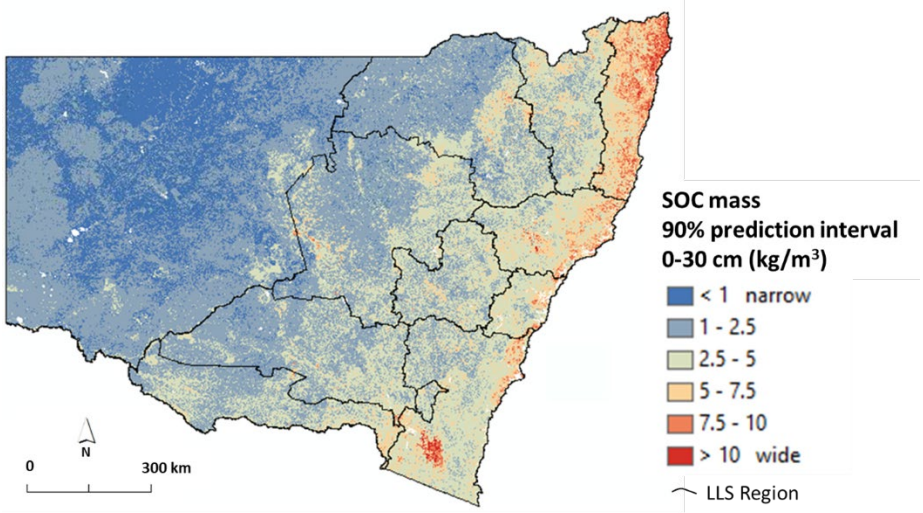
D.1.2 SOC mass (kg/m^3)



a: SOC mass (kg/m^3) 0–10 cm

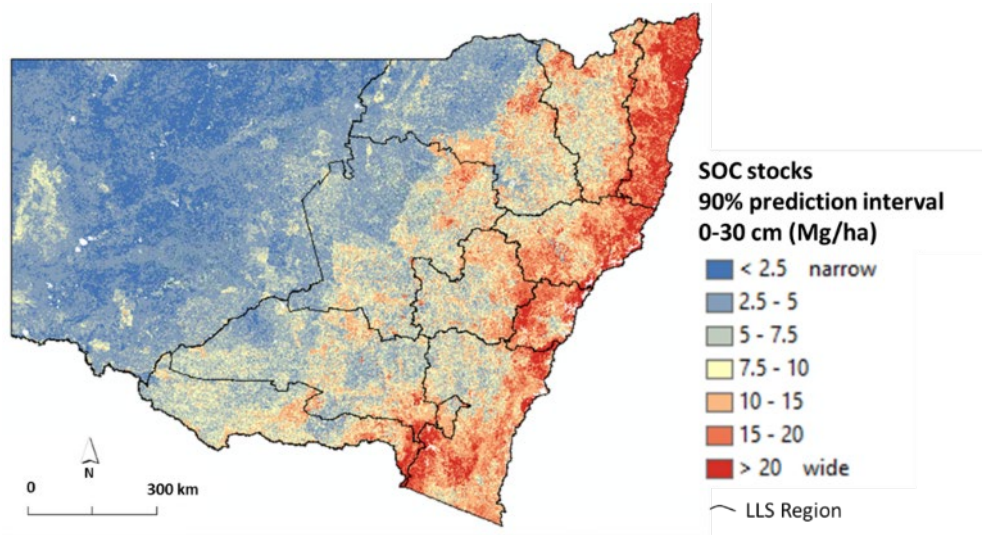


b: SOC mass (kg/m^3) 10–30 cm

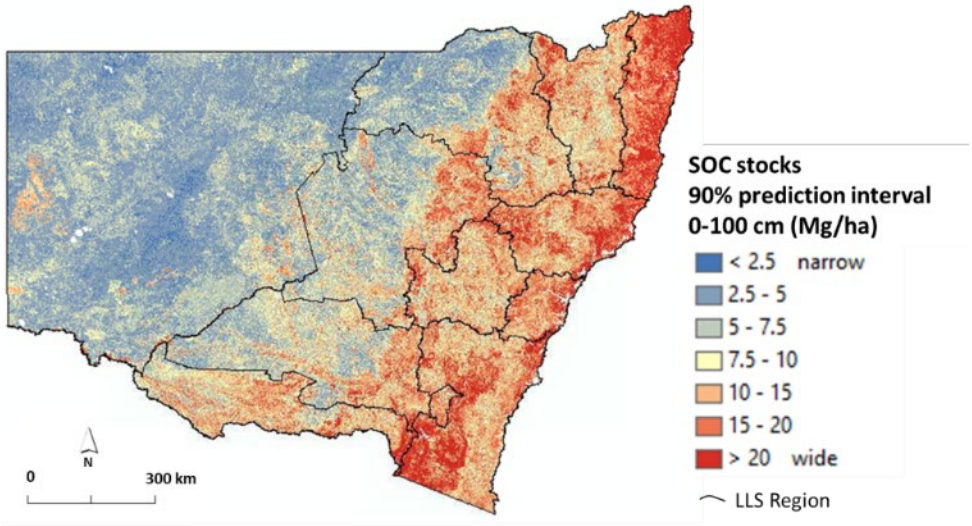


c: SOC mass (kg/m^3) 0–30 cm

D.1.3 SOC stocks (Mg/ha)

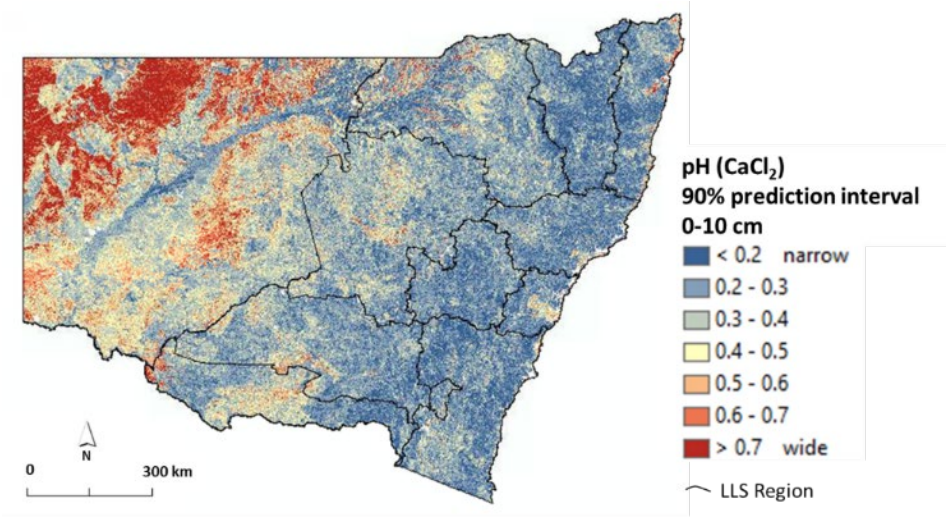


a: SOC stocks (Mg/ha) 0–30 cm

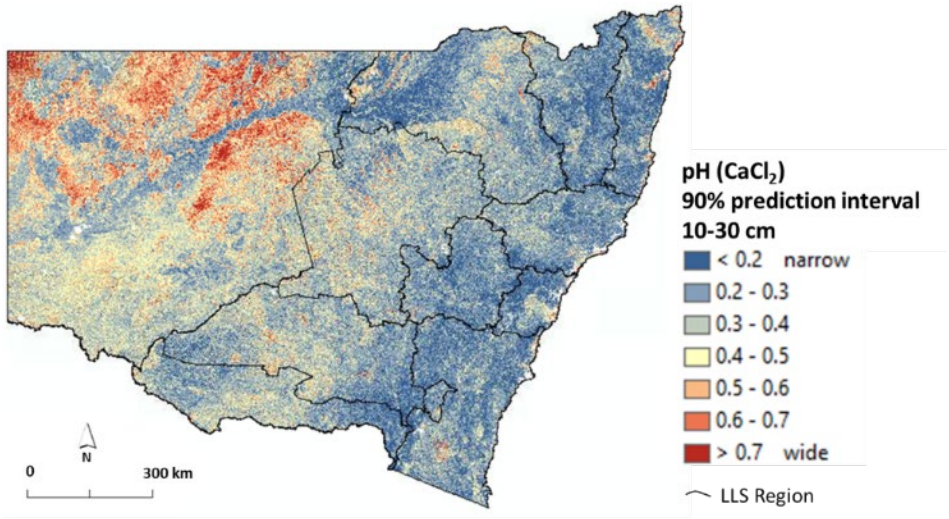


b: SOC stocks (Mg/ha) 0–100 cm

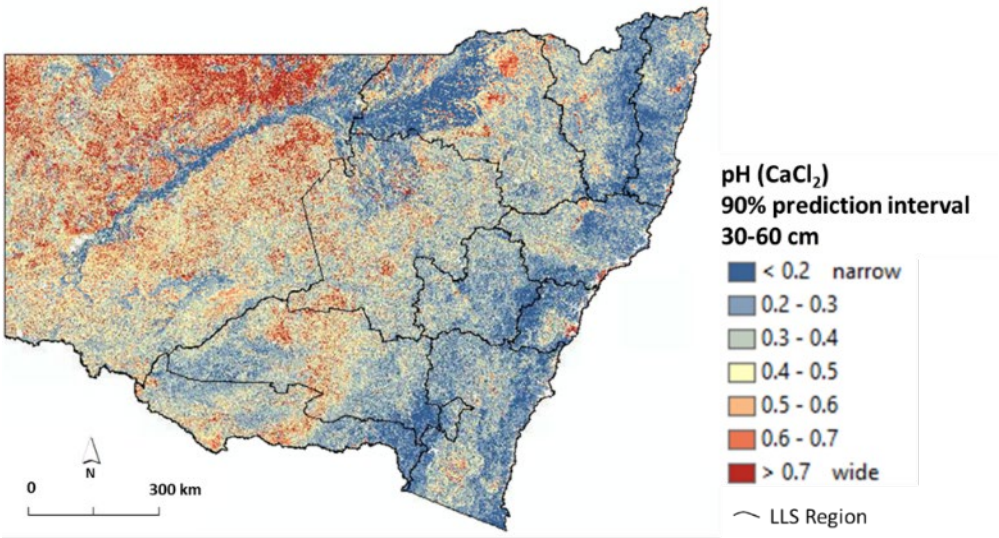
D.2 pH$_{(CaCl_2)}$ (pH units)



a: pH$_{(CaCl_2)}$ 0–10 cm



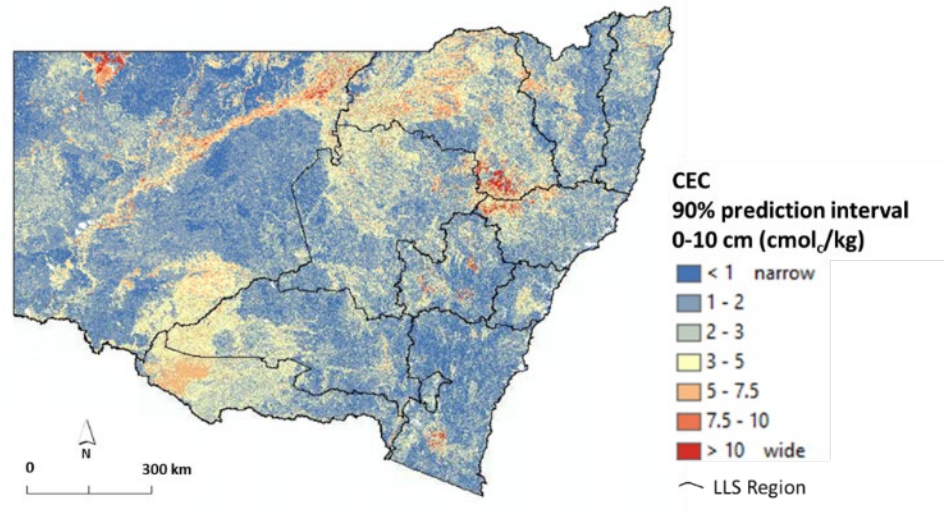
b: pH$_{(CaCl_2)}$ 10–30 cm



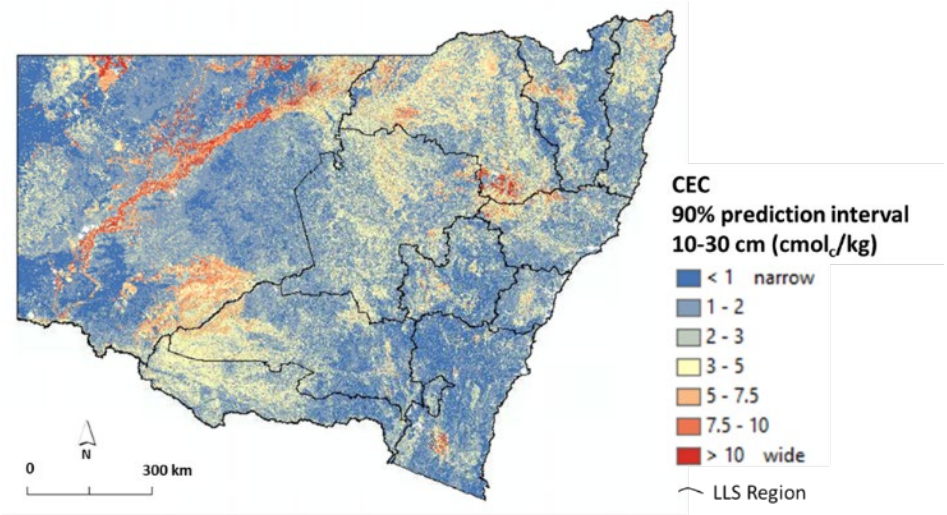
c: pH$_{(CaCl_2)}$ 30–60 cm

D.3 Cation exchange capacity ($cmol_c/kg$)

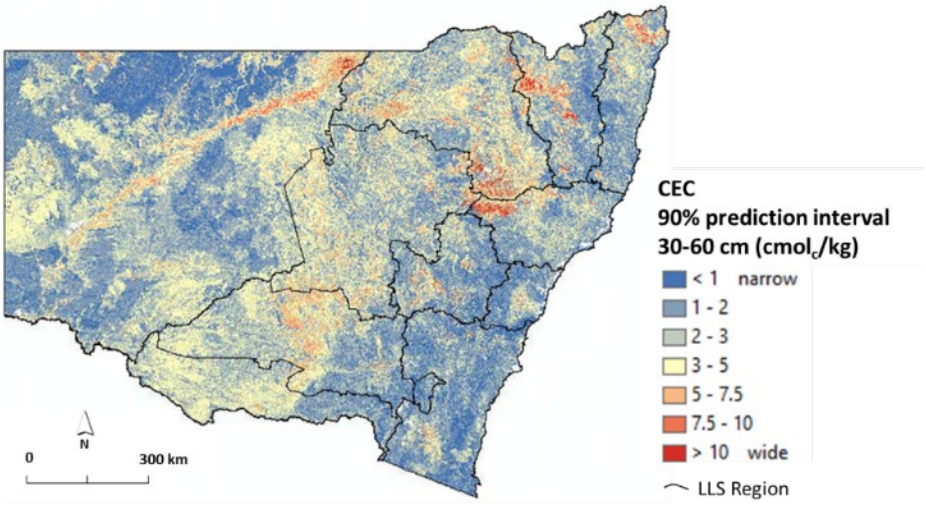
(Total exchange sites in soil; including sites occupied by basic and acidic exchangeable cations)



a: CEC ($cmol_c/kg$) 0–10 cm



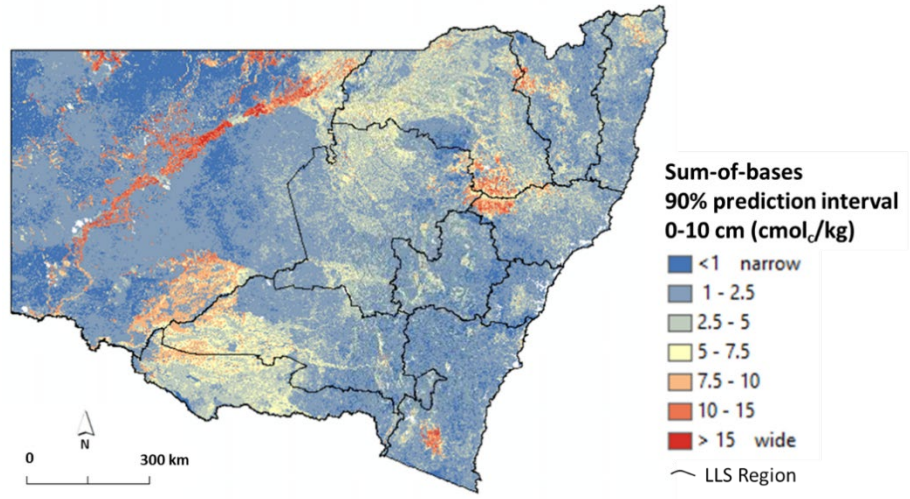
b: CEC ($cmol_c/kg$) 10–30 cm



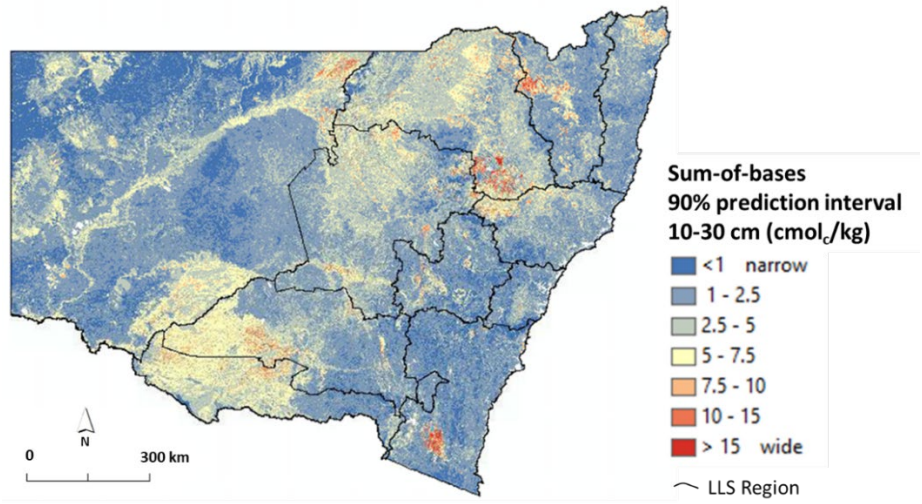
c: CEC ($cmol_c/kg$) 30–60 cm

D.4 Sum-of-bases ($cmol_c/kg$)

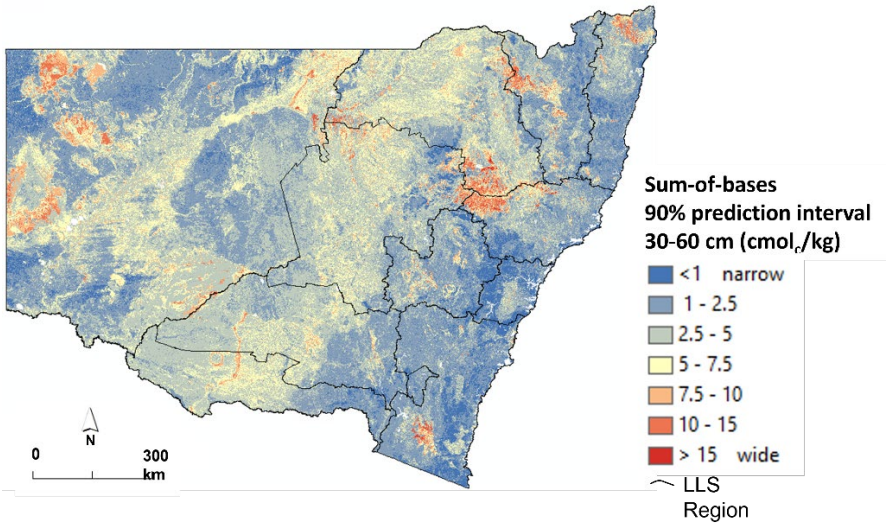
(Sum of exchangeable calcium, magnesium, sodium and potassium cations)



a: Sum-of-bases ($cmol_c/kg$) 0–10 cm, mean

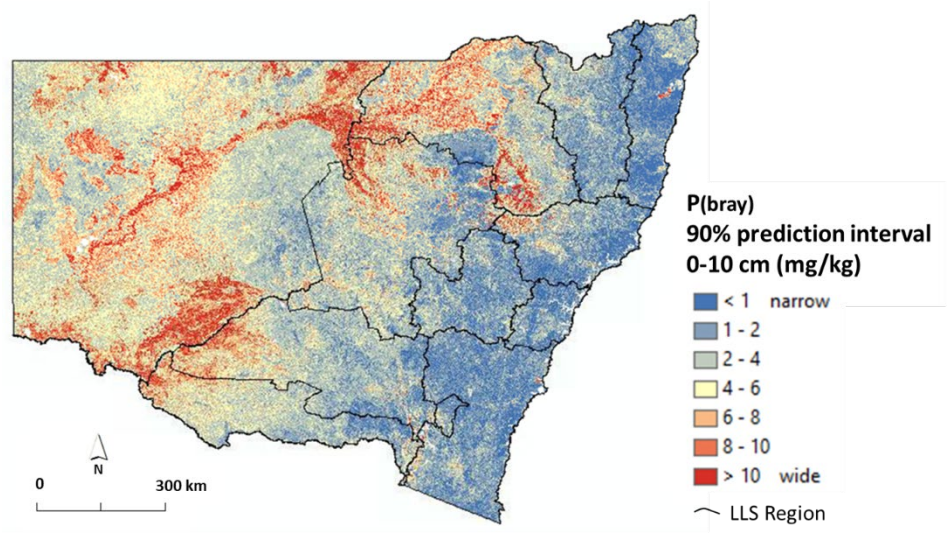


b: Sum-of-bases ($cmol_c/kg$) 10–30 cm, mean

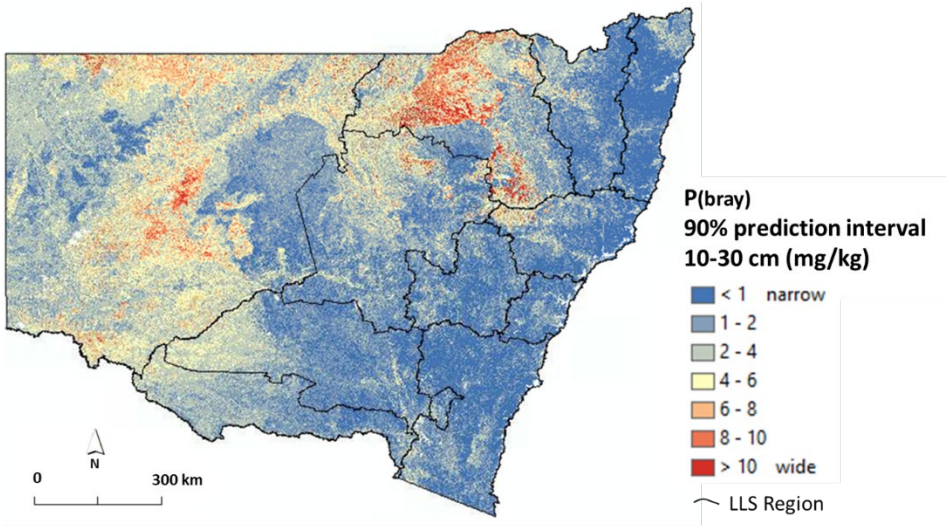


c: Sum-of-bases ($cmol_c/kg$) 30–60 cm

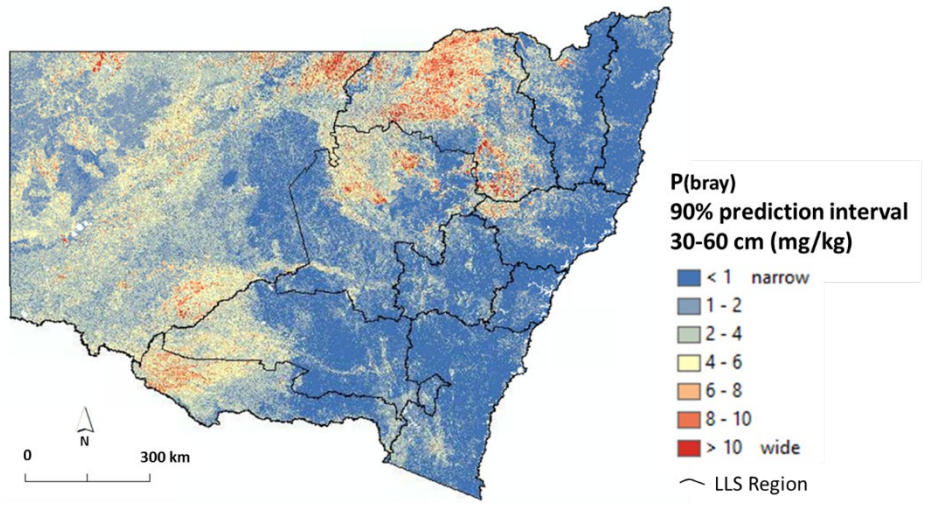
D.5 Available phosphorus (P(bray), mg/kg)



a: P(bray) (mg/kg) 0–10 cm

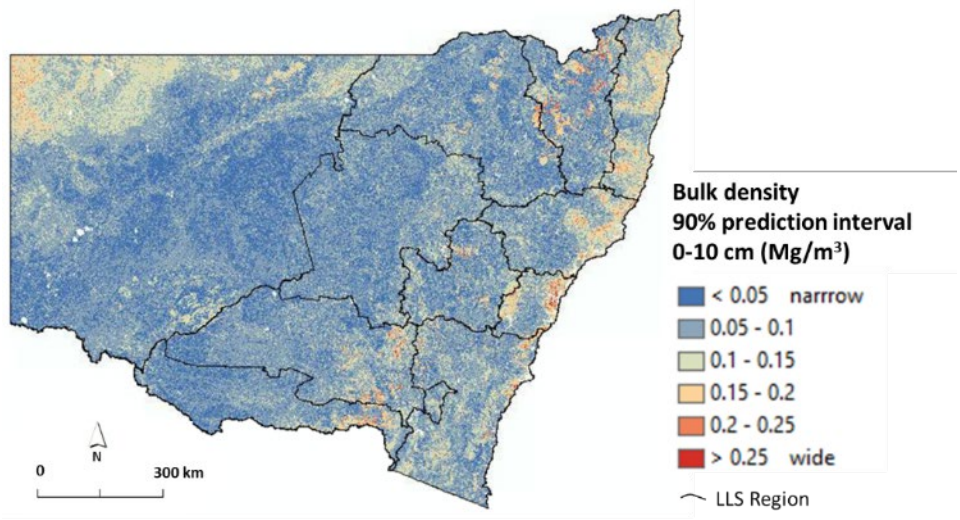


b: P(bray) (mg/kg) 10–30 cm

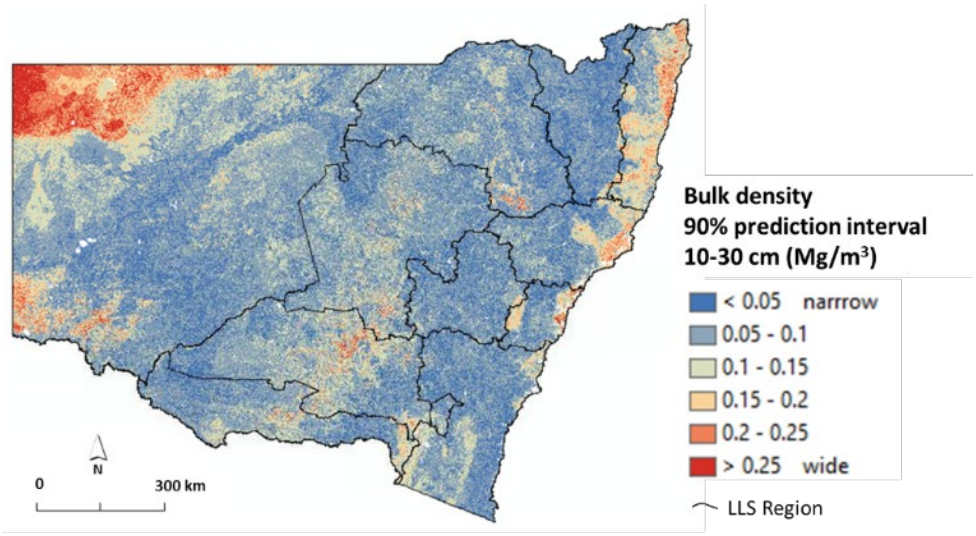


c: P(bray) (mg/kg) 30–60 cm

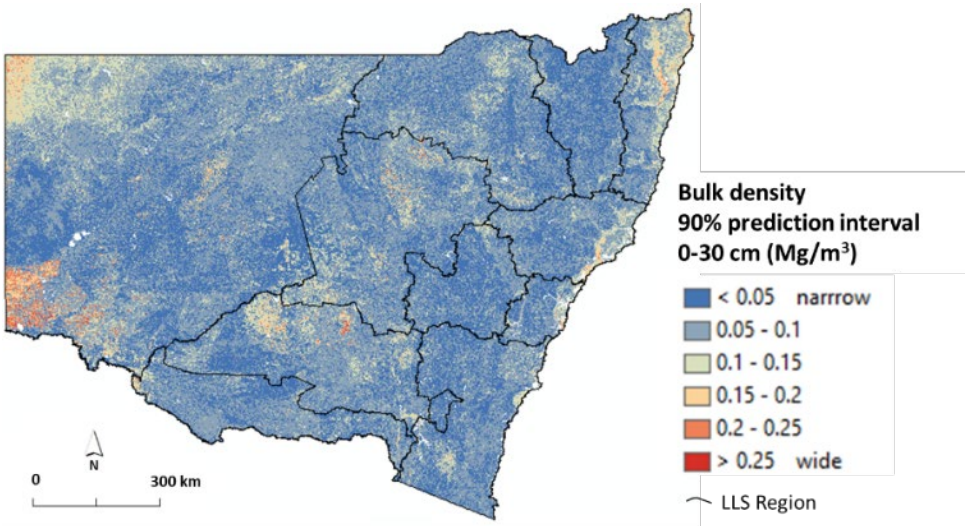
D.6 Bulk density (Mg/m^3)



a: BD (Mg/m^3) 0–10 cm



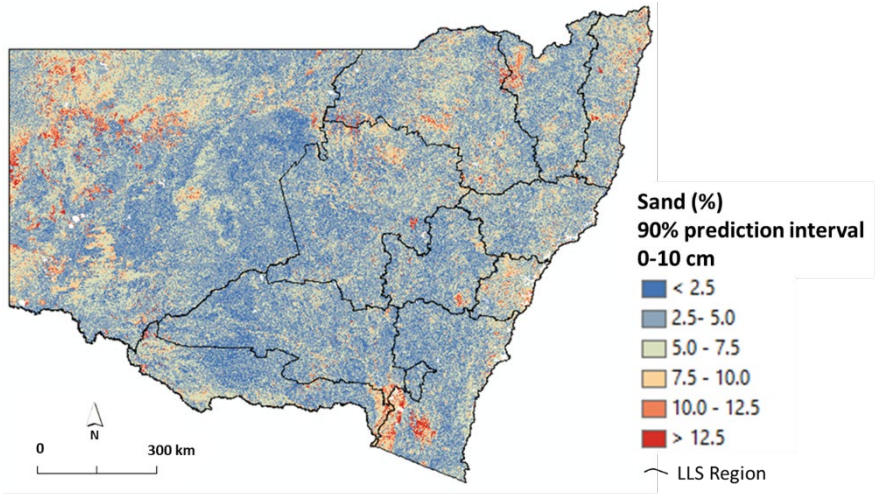
b: BD (Mg/m^3) 10–30 cm



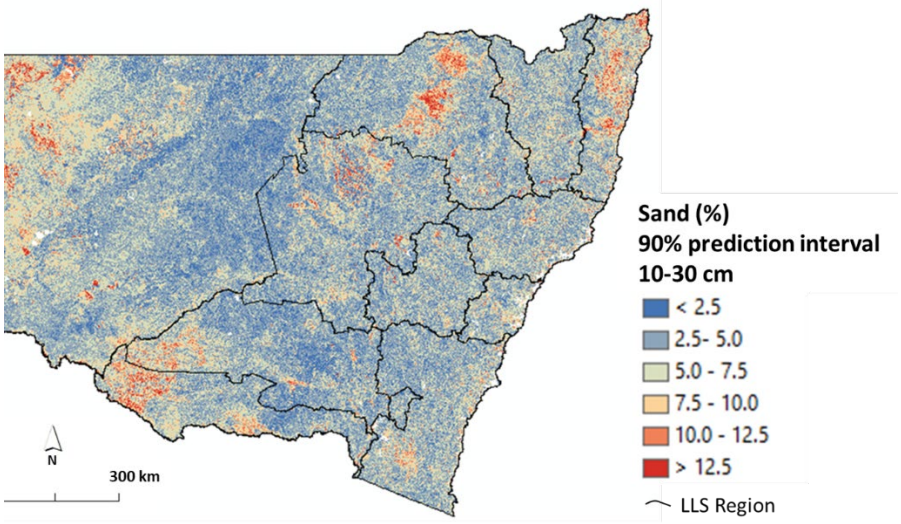
c: BD (Mg/m^3) 0–30 cm

D.7 Sand

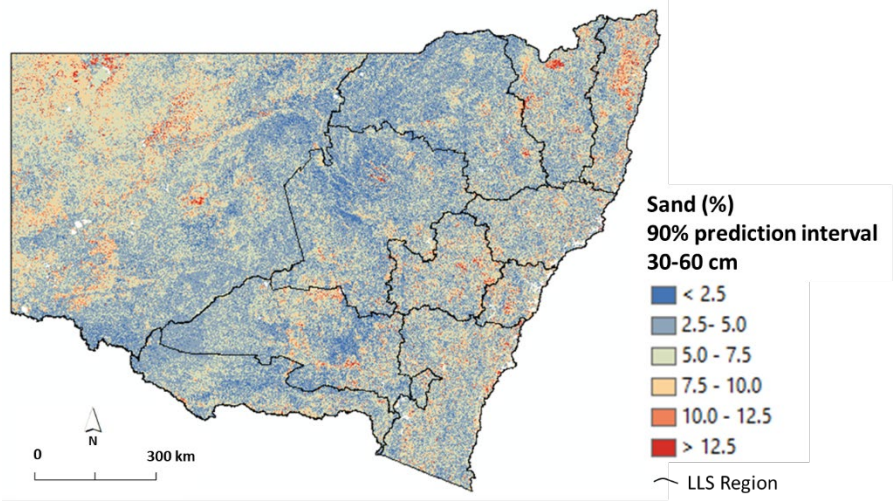
D.7.1 Total sand (%)



a: Total sand (%) 0–10 cm

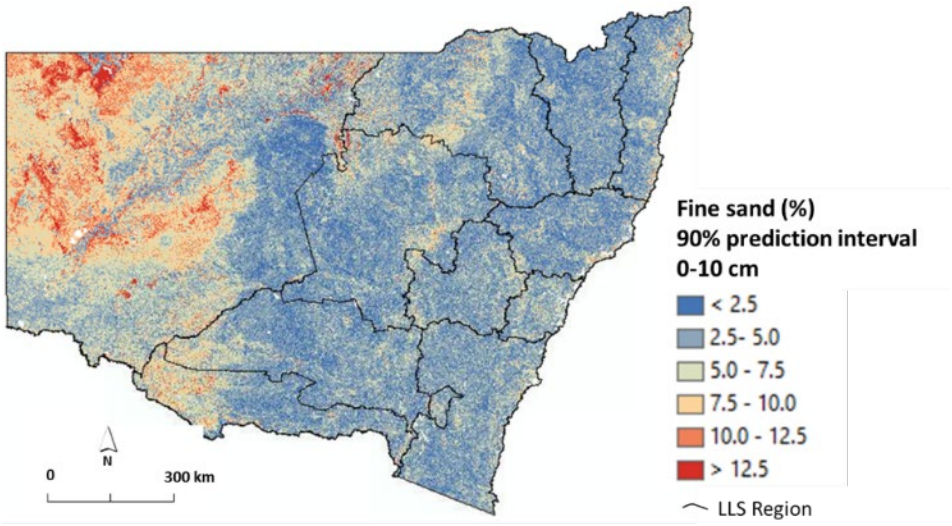


b: Total sand (%) 10–30 cm

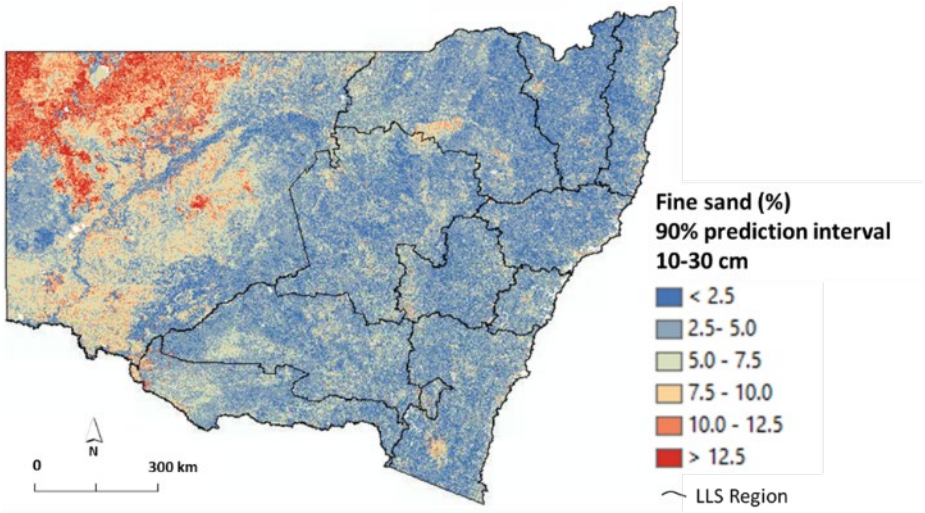


c: Total sand (%) 30–60 cm

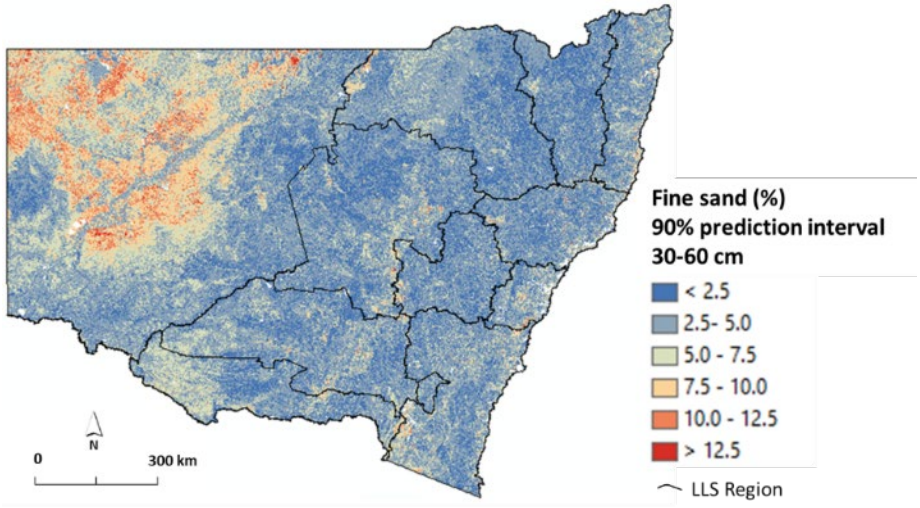
D.7.2 Fine sand (%)



a: Fine sand (%) 0–10 cm

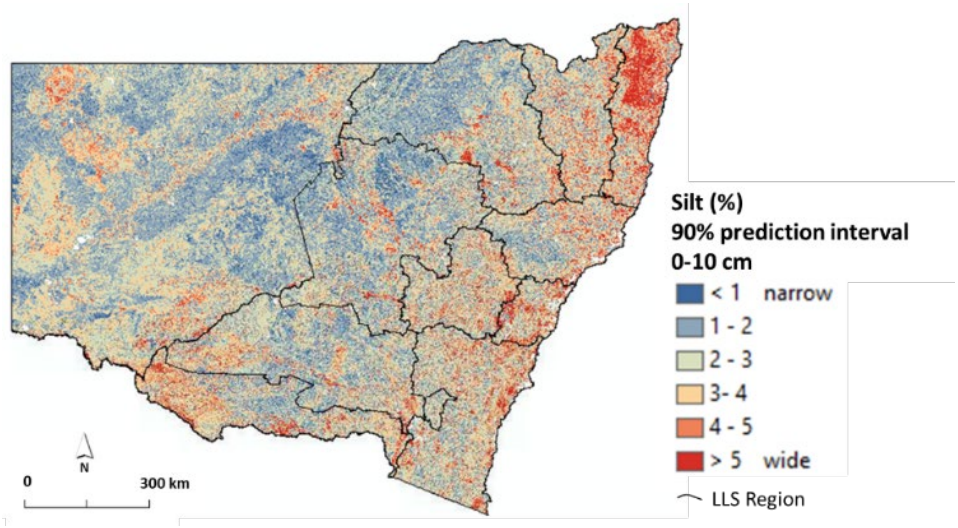


b: Fine sand (%) 10–30 cm

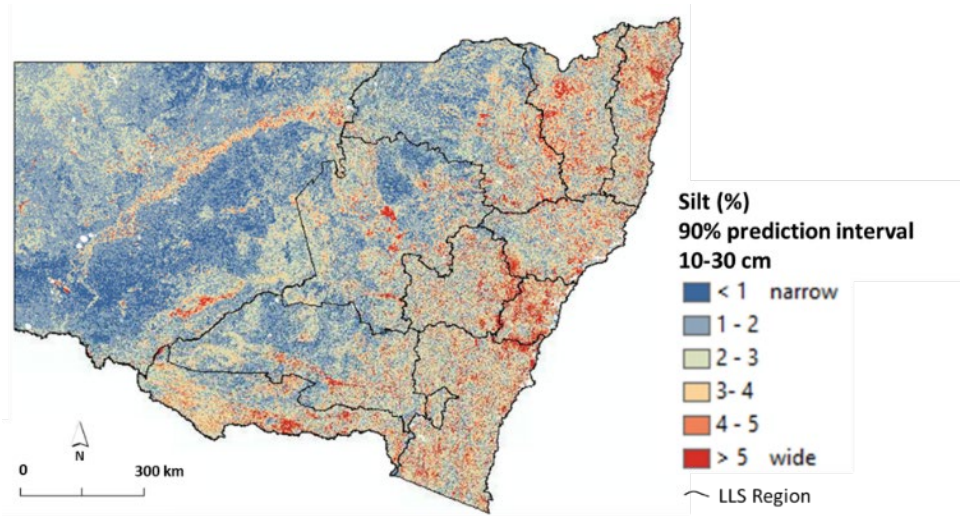


c. Fine sand (%) 30–60 cm

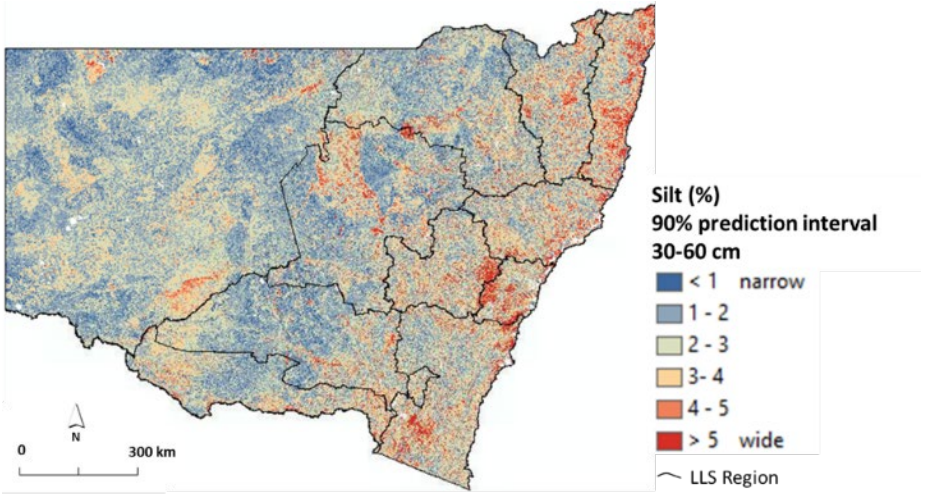
D.8 Silt (%)



a: Silt (%) 0–10 cm

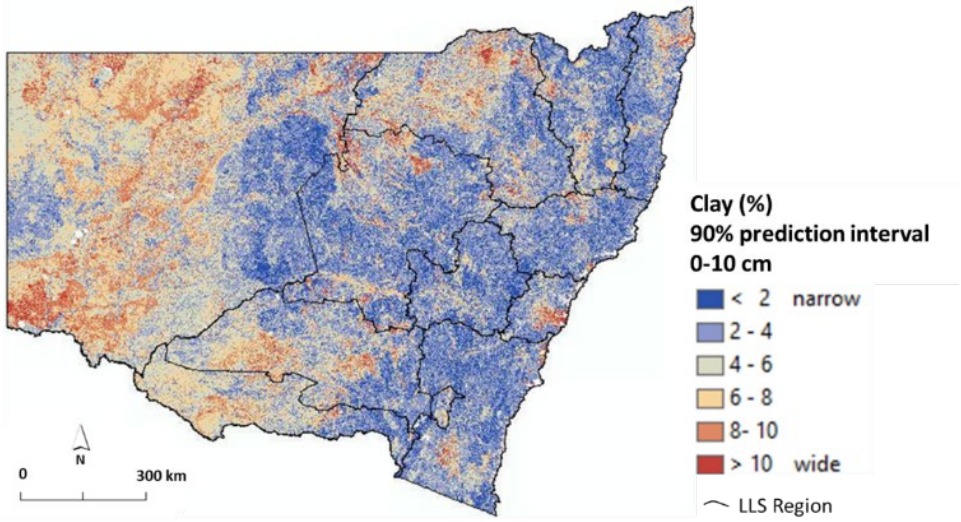


b: Silt (%) 10–30 cm

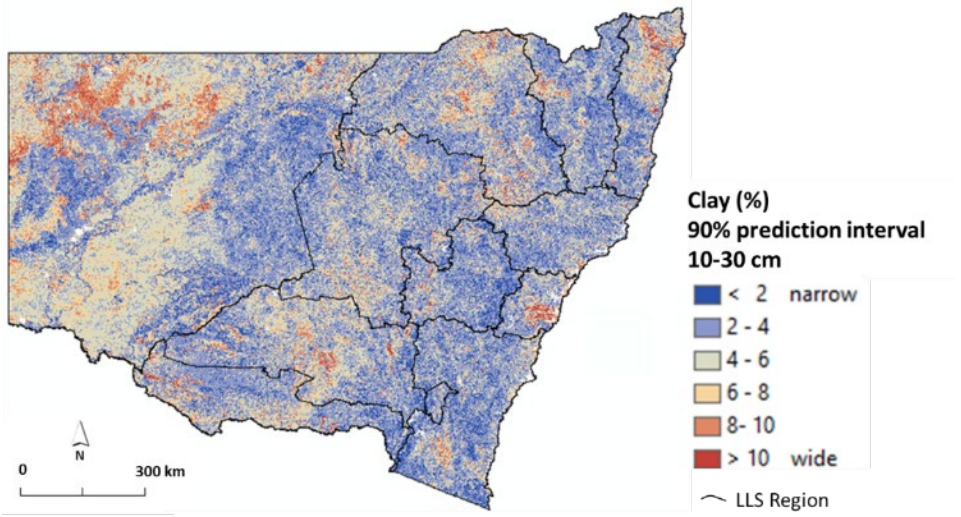


c: Silt (%) 30–60 cm

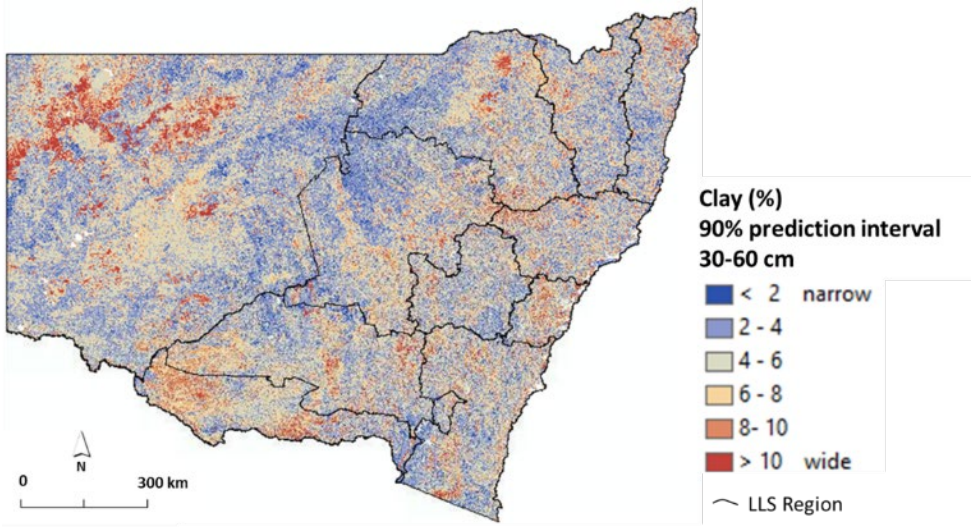
D.9 Clay (%)



a: Clay (%) 0–10 cm

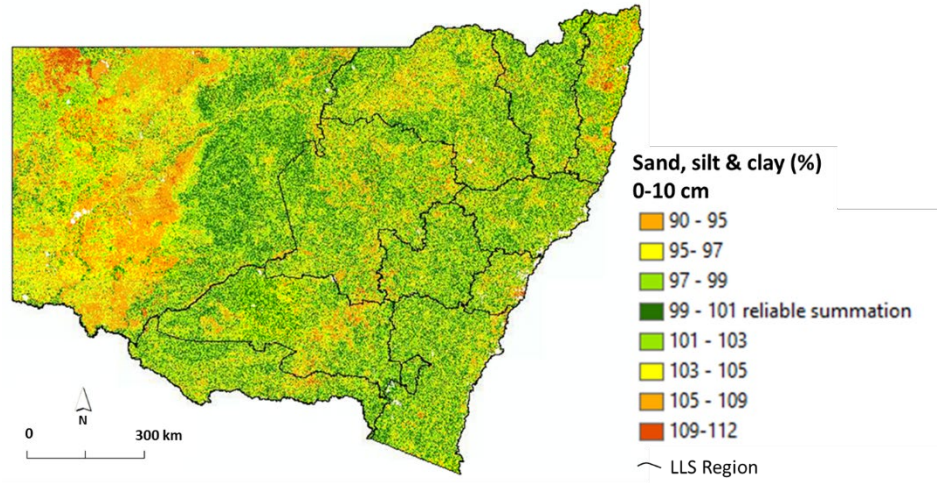


b: Clay (%) 10–30 cm

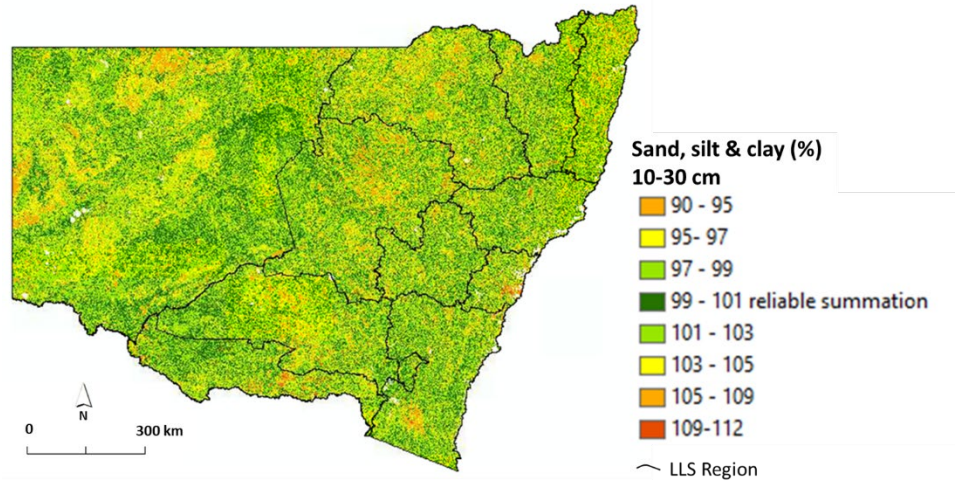


c: Clay (%) 30–60 cm

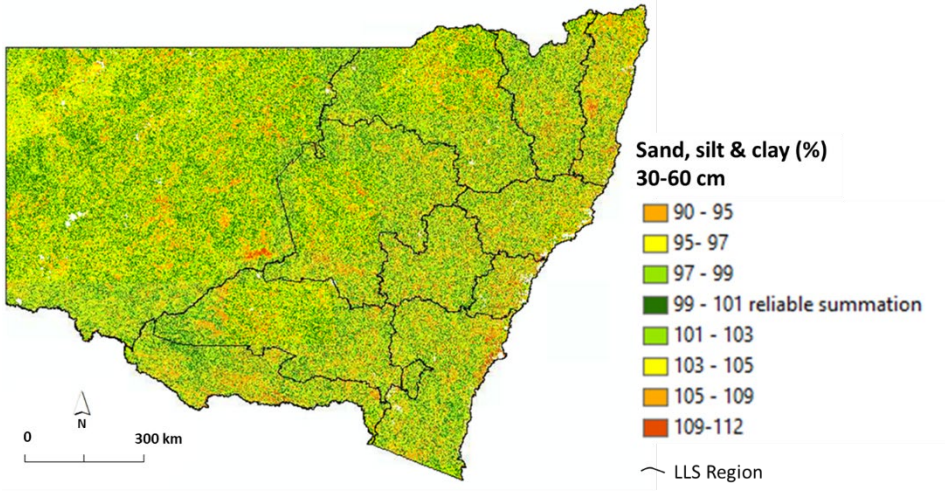
Appendix E: Summation of particle sizes (sand, silt and clay)



a: Sand, silt and clay (%) 0–10 cm



b: Sand, silt and clay (%) 10–30 cm



c: Sand, silt and clay (%) 30–60 cm

Table E.1 **Proportion of summed particle sizes (sand, silt and clay) falling in different reliability ranges**

Depth (cm)	99–101% range	95–105% range	90–110% range
0–10	22.1	82.0	99.3
10–30	21.0	81.7	99.7
30–60	26.8	91.2	99.0

References

Arrouays D, McKenzie N, Hempel J, Richer de Forges A and McBratney A (eds) (2014) *GlobalSoilMap: Basis of the global spatial soil information system*, CRC Press/Balkema, The Netherlands.

Baldock JA, Sanderman J, Macdonald L, Allen D, Cowie A, Dalal R, Davy M, Doyle R, Herrmann T, Murphy D and Robertson F (2013) *Australian Soil Carbon Research Program*, v1, data collection, CSIRO, available at: https://data.csiro.au/dap/landingpage?pid=csiro%3A5883

Baldock JA, Wheeler I, McKenzie N and McBratney A (2012) 'Soils and climate change: potential impacts on carbon stocks and greenhouse gas emissions, and future research for Australian agriculture', *Crop and Pasture Science*, 63:269–283.

Bishop TFA, Horta A and Karunaratne SB (2015) 'Validation of digital soil maps at different spatial supports', *Geoderma*, 241–242:238–249.

Bishop TFA, McBratney AB and Laslett GM (1999) 'Modelling soil attribute depth functions with equal-area quadratic smoothing splines', *Geoderma*, 91:27–45.

Brus DJ, Kempen B and Heuvelink GBM (2011) 'Sampling for validation of digital soil maps', *European Journal of Soil Science*, 62:394–407.

Bui E, Henderson B and Viergever K (2009) 'Using knowledge discovery with data mining from the Australian Soil Resource Information System database to inform soil carbon mapping in Australia', *Global Biogeochemical Cycles*, 23:1–15.

Chapman GA, Gray JM, Murphy BW, Atkinson G, Leys JF, Muller R, Peasely B, Wilson BR, Bowman, G, McInnes-Clarke SK, Tulau MJ, Morand DT and Yang X (2011) *Monitoring, Evaluation and Reporting of Soil Condition in New South Wales: 2008 Program*, NSW Department of Environment, Climate Change and Water, Sydney.

Colquhoun GP, Hughes KS, Deyssing L, Ballard JC, Folkes CB, Phillips G, Troedson AL and Fitzherbert JA (2022) *New South Wales Seamless Geology dataset*, version 2.2, Geological Survey of New South Wales, Department of Regional NSW, Maitland, https://search.geoscience.nsw.gov.au/product/9232

Dokuchaev VV (1899) *On the Theory of Natural Zones*, St. Petersburg, Russia.

DPE (Department of Planning and Environment) (2021) *NSW Landuse 2007 and 2017*, NSW Department of Planning and Environment, Parramatta, accessed 13 September 2022, https://datasets.seed.nsw.gov.au/dataset/nsw-landuseac11c (for 2007); https://datasets.seed.nsw.gov.au/dataset/nsw-landuse-2017-v1p2-f0ed

Gallant JC and Austin JM (2015) 'Derivation of terrain covariates for digital soil mapping in Australia', *Soil Research*, 53:895–90.

Gray JM, Bishop TFA and Wilford JR (2016) 'Lithology and soil relationships for soil modelling and mapping', *Catena*, 147:429–440, doi: 10.1016/j.catena.2016.07.045.

Gray JM, Bishop TFA and Wilson BR (2015a) 'Factors controlling soil organic carbon stocks with depth in eastern Australia', *Soil Science Society of America Journal*, 79:1741–1751, doi: 10.2136/sssaj2015.06.0224

Gray JM, Bishop TFA and Yang X (2015b) 'Pragmatic models for the prediction and digital mapping of soil properties in eastern Australia', *Soil Research*, 53:24–42, doi: 10.1071/SR13306.

Gray JM, Karunaratne SB, Bishop TFA, Wilson BR and Veeragathipillai M (2019) 'Driving factors of soil organic carbon fractions over New South Wales, Australia', *Geoderma*, 353:213-226, doi: 10.1016/j.geoderma.2019.06.032, data available at: https://datasets.seed.nsw.gov.au/dataset/soil-organic-carbon-fractions-over-nsw

Gray JM, Wang BA, Waters CM, Orgill SE, Cowie AL and Ng EL (2022) 'Digital mapping of soil carbon sequestration potential with enhanced vegetation cover over New South Wales, Australia', *Soil Use and Management*, 38:229–247, doi: 10.1111/sum.12766, data available at: https://datasets.seed.nsw.gov.au/dataset/soil-carbon-sequestration-potential-with-enhanced-vegetation-cover-over-nsw

Grundy MJ, Viscarra Rossel RA, Searle RD, Wilson PL, Chen C and Gregory LJ (2015) 'Soil and Landscape Grid of Australia', *Soil Research*, 53:835–844, updated SLGA website.

Guerschman JP and Hill MJ (2018) 'Calibration and validation of the Australian fractional cover product for MODIS collection 6', *Remote Sensing Letters*, 9(7):696–705, doi: 10.1080/2150704X.2018.1465611, data accessed 13 September 2022 at: https://eo-data.csiro.au/remotesensing/v310/australia/monthly/cover/

Henderson BL and Bui EN (2002) 'An improved calibration curve between soil pH assessed in water and $CaCl_2$', *Australian Journal of Soil Research*, 40:1399–1405.

Henderson BL, Bui EN, Moran CJ and Simon DAP (2005) 'Australia-wide predictions of soil properties using decision trees', *Geoderma*, 124:383–398.

Isbell RF and NCST (2021) *The Australian Soil Classification*, third edition, with National Committee on Soil and Terrain, CSIRO Publishing, Melbourne.

Jenny H (1941) *Factors of soil formation*, McGraw-Hill Book Company, New York.

Kidd D, Webb M, Malone B, Minasny B and McBratney AB (2015) 'Digital soil assessment of agricultural suitability, versatility and capital in Tasmania, Australia', *Geoderma Regional*, 6:7–21.

Karunaratne SB, Bishop TFA, Baldock JA and Odeh IOA (2014) 'Catchment scale mapping of measurable soil organic carbon fractions', *Geoderma*, 219–220:14–23.

Liaw A and Wiener M (2018) *Package 'randomForest' Breiman and Cutler's Random Forests for Classification and Regression*, CRAN Repository, https://cran.r-project.org/web/packages/randomForest/randomForest.pdf

Lin LI (1989) 'A concordance correlation coefficient to evaluate reproducibility', *Biometrics*, 45:255–268.

Malone BP, McBratney AB, Minasny B and Laslett GM (2009) 'Mapping continuous depth functions of soil carbon storage and available water capacity', *Geoderma*, 154:137–152.

Malone BP, Minasny B, Odgers NP and McBratney AB (2014) 'Using model averaging to combine soil property rasters from legacy soil maps and from point data', *Geoderma*, 232–234:34–44.

McBratney AB, Mendonça Santos ML and Minasny B (2003) 'On digital soil mapping', *Geoderma*, 117:3–52.

Minasny B and McBratney AB (2006) 'A conditioned Latin hypercube method for sampling in the presence of ancillary information', *Computer Geoscience*, 32:1378–1388.

Minasny B, McBratney AB, Mendonça-Santos ML, Odeh IOA and Guyon B (2006) 'Prediction and digital mapping of soil carbon storage in the Lower Namoi Valley', *Australian Journal of Soil Research*, 44:233–244.

Mulder VL, De Bruin S, Schaepman ME and Mayr TR (2011) 'The use of remote sensing in soil and terrain mapping – a review', *Geoderma*, 162:1–19.

Nelson MA, Bishop TFA, Odeh IOA and Triantafilis J (2011) 'An error budget for different sources of error in digital soil mapping', *European Journal of Soil Science*, 62:417–430.

Odeh IOA, McBratney AB and Chittleborough DI (1995) 'Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging', *Geoderma*, 67:215–226.

OEH (Office of Environment and Heritage) (2014) *Soil condition and land management in NSW: Final results from the 2008–09 monitoring evaluation and reporting program*, Technical Report, NSW Office of Environment and Heritage, Sydney, accessed 6 May 2017, www.environment.nsw.gov.au/research-and-publications/publications-search/soil-condition-and-land-management-in-new-south-wales

OEH (2017) *SALIS laboratory soil test methods, soil physical and chemical analysis*, NSW Office of Environment and Heritage, Sydney.

OEH (2018) *Digital soil mapping of key soil properties over NSW, version 1.2*, Technical Report, NSW Office of Environment and Heritage, Sydney.

R Core Team (2022) *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria, accessed 18 August 2022, www.R-project.org/

Robinson NJ, Benke KK and Norng S (2015) 'Identification and interpretation of sources of uncertainty in soils change in a global systems-based modelling process', *Soil Research*, 53:592–604.

Sanchez PA, Ahamed S, Carré F, Hartemink A, Hempel J, Huising J, Lagacherie P, McBratney A, McKenzie N, de Lourdes Mendonça-Santos M, Minasny B, Montanarella L, Okoth P, Palm C, Sachs J, Shepherd K, Vågen T, Vanlauwe B, Walsh M, Winowiecki L and Zhang G (2009) 'Digital soil map of the world', *Science*, 325:680–681.

Sanderman J, Baldock J, Hawke B, Macdonald L, Massis-Puccini A and Szarvas S (2011) *National Soil Carbon Research Programme: Field and Laboratory Methodologies*, CSIRO Land and Water, Urrbrae, South Australia, accessed 17 October 2016, https://csiropedia.csiro.au/wp-content/uploads/2016/06/SAF-SCaRP-methods.pdf

Skjemstad JO, Spouncer LR and Beech A (2000) *Carbon conversion factors for historical soil carbon data*, Technical Report No. 15, National Carbon Accounting System, Australian Greenhouse Office, Canberra, https://publications.csiro.au/rpr/download?pid=procite:bd31d914-db08-4ae3-a77b-e0d23070581d&dsid=DS1.

Triantafilis J, Lesch SM, Lau KL and Buchanan SM (2009) 'Field level digital soil mapping of cation exchange capacity using electromagnetic induction and a hierarchical spatial regression model', *Australian Journal of Soil Research*, 47:651–663.

Viscarra Rossel RA (2011) 'Fine-resolution multiscale mapping of clay minerals in Australian soils measured with near infrared spectra', *Journal of Geophysical Research*, 116:F04023, 15pp.

Viscarra Rossel RA, Chen C, Grundy M, Searle R, Clifford D and Campbell PH (2015) 'The Australian three-dimensional soil grid: Australia's contribution to the *GlobalSoilMap* project', *Soil Research*, 53:845–864.

Viscarra Rossel RA, Webster R, Bui EN and Baldock JA (2014) 'Baseline map of organic carbon in Australian soil to support national carbon accounting and monitoring under climate change', *Global Change Biology*, 20:2953–2970.

Wilford J (2012) 'A weathering intensity index for the Australian continent using airborne gamma-ray spectrometry and digital terrain analysis', *Geoderma*, 183:124–142.

More information

- Australian Soil Carbon Research Program (SCARP)
- eSPADE soil and landscape spatial viewer
- GlobalSoilMap.net project
- SEED data portal
- SALIS (Soil and Land Information System)
- Soil and Landscape Grid of Australia (SLGA)

Acknowledgements

This report was prepared by Jonathan Gray, with assistance from other staff from the Science Strategy and Impact Branch and Environmental Protection Science Branch, of the Science, Economics and Insights Division, NSW Department of Planning and Environment. Senani Karunaratne (CSIRO) and Mark Young (Department of Planning and Environment) provided particularly useful advice and assistance.