Assessing the suitability of lower Murrumbidgee valley soils for irrigated cotton production

Jonathon James Moore
BScAgr (Hons I)

Supervisor: Professor Stephen Cattle
Associate Supervisor: Dr. Patrick Filippi

A thesis submitted to fulfil requirements for the degree of

Doctor of Philosophy

2024

School of Life and Environmental Science

Faculty of Science

The University of Sydney

New South Wales

Australia

The research reported in this thesis was supported by the award of a Research Training scholarship to the PhD Candidate.

Statement of Originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree of other purpose.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Jonathon James Moore

30 June 2024

Thesis Summary

Soil is a finite and irreplaceable resource that underpins agricultural systems, regulates water cycles and presents opportunities for climate change mitigation. Consequently, the sustainable management of soil is essential for guarding the security of agricultural production and ecosystems. For this to occur, however, an accurate understanding of inherent soil properties and how they vary spatially, is required. Within the lower Murrumbidgee valley of southern NSW, Australia, there is a paucity of available data and land managers require more information on both the chemical and physical properties of their soils and the presence of potential constraints to agricultural production.

While the development of novel approaches and 'blue-sky' research is important, the targeted application of developed methods is essential in filling knowledge gaps and allowing benefits to reach the end user, in this instance, farmers and land managers. Within the lower Murrumbidgee valley farmers are a key stakeholder of the soil resource. Thus, they should be a key consideration when designing, undertaking and communicating projects to ensure their outcomes are tangible. Therefore, rather than seeking to develop novel approaches to digital soil mapping (DSM) and digital soil assessment (DSA), this thesis utilises various methods recognised in literature to develop outputs targeted to the cotton industry in southern NSW. Using a collected dataset of 153 soil cores to 1 m depth from across the region, soil profiles are morphologically described, soil properties are modelled and mapped at the within-field and regional scales, before a regionally specific classification is developed.

In this thesis the review of literature, Chapter 2, synthesises knowledge from various disciplines to tell the story of how soils of the lower Murrumbidgee valley have evolved. Considerations of the regional geomorphological characteristics explain how the soils have developed, the extent to which they vary and how, despite a rich history of soils research, there are limitations to the publicly available soils information and DSM products. To prevent unnecessary repetition in different research chapters, Chapter 3 outlines the datasets used in this study: the 'full dataset', the 'valley-wide subset' and the 'within-field subset', as well how soil samples were obtained. Along with outlining the laboratory methods used to obtain the data used in the ensuing chapters, it also describes how visible near infrared (VisNIR) spectroscopy was used to predict sand and clay content. A description on how spatial covariates were acquired is included also.

Chapter 4 is the first research chapter of this thesis and entails the morphological description of soil cores from the 'full dataset' as outlined in Chapter 3. It was determined that previous soil groupings did not appropriately account for the 153 described soil profiles. Consequently, a new classification was developed, with ten soil classes presented alongside a bifurcating classification key. These classes were placed in four broader groups: Red-Brown clays, Grey clays, Earths and Deep sands. Clays, of both colour groups, were the most prolific soil types across the region. This is shown spatially through the mapping of soil groups across the region to a 90 m resolution using a random forest classification model. Higher density sampling within two fields allowed for transects to be developed. At one site a streambed sand, importantly distinguished from dune sands, was identified and the distinct morphological variability seen in soils to a depth of 1 m at the within-field scale is highlighted.

Chapter 5 examines how the morphological variability identified in Chapter 4 relates to irrigated cotton production at the within-field level. It identifies the complex relationship between cotton lint yields and inherent soil properties within two paddocks on two farms that have significant yield variability. It is demonstrated that on both sites clay and sand content are the most positively and negatively correlated properties to yield, respectively. This associates more clayey textured soils with higher cotton yields as a result of soil hydrological attributes such as water holding capacity, a trend demonstrated through the derivation of production zones using available yield data. Bespoke digital soil maps are produced to visualise the spatial variability in pH, electrical conductivity of the extract (ECe), exchangeable sodium percentage (ESP), cation exchange capacity (CEC), sand and clay content at a field scale using varied sample sizes. This demonstrated the potential of simple linear models to predict soil properties at a moderate-to-good quality. Considerations are then given as to how management zones can be used to improve irrigation and overall production efficiencies.

Chapter 6 uses digital soil mapping (DSM) as a tool to understand the spatial variability of pH, ECe, ESP, CEC, sand and clay content across the lower Murrumbidgee valley. Extreme Gradient Boosting (XGBoost) models produced moderate model quality statistics when assessed using a leave-one-site-out-cross-validation (LOSOCV). The inclusion of layer mid-depth as a predictor variable allowed for the concurrent modelling of the sampled depths, resulting in interpretable and dynamic models. Statistical methods combined with an expert assessment of landscape trends

allowed for the most parsimonious, but also pedologically sound, model for each soil property to be developed. To undertake this process the understanding of the region's geomorphology acquired in Chapter 2 was required. SHapley Additive exPlanation (SHAP) values were effective in interpreting the impact of predictor variables on model outputs. These values showed that the role of spatial covariates varied depending on the soil property being modelled. An external point support validation of available state, national and global DSM products for subsoil clay content showed that the model statistics reported in literature for each product were poorer than when validated using the data from this study. This indicates that these broader scale DSMs were unable to identify local variability in subsoil clay content at the point support and suggests that region-specific models are the best approach moving forward.

Chapter 7 utilises the DSM products developed in Chapter 6 to undertake a DSA of the region for irrigated cotton production. DSAs are the application of DSMs into decision making aids capable of answering specific questions. Multiple DSA products are created including depth to constraint maps and the overall suitability classification which consisted of four classes: well suited, suited, marginally suited and unsuited. Land was placed in these classes based on regionally specific rulesets for the soil-landscape parameters of slope, plant available water capacity (PAWC) and soil chemical constraints. The latter was determined based on the depth to which a plant limiting threshold for pH or exchangeable sodium percentage (ESP) was reached. PAWC was predicted using pedotransfer functions incorporating cation exchange capacity, sand and clay content. In total, 16.7% of the land in the lower Murrumbidgee valley was classed as well suited to irrigated cotton production and 18.4% was classed as unsuited with the remainder comprised of marginally suited (34%) or suited (31%) land. A large proportion of the land classed as well suited has already been developed for broadacre irrigated cropping. Understandings of the landscape gained through the preceding chapters of this thesis, including the literature review, informed the targeted selection of parameters and development of specific rulesets. It was apparent that, in the case of ESP, previously used plant growth thresholds are not always appropriate as, if these had been applied, land known to be capable of successfully growing cotton would have been classed as unsuitable. The classification showed that areas most suited to irrigated cotton production that have not been developed are in the south and central north of the lower Murrumbidgee valley.

Overall, the use of different approaches, from morphological classifications to bespoke paddock and regional DSMs are able to identify the extent of soil variability within the lower Murrumbidgee valley of southern NSW. Through considerations of the landscape's geomorphological characteristics, this variability is linked and discussed in relation to pathways of palaeochannel systems. It is demonstrated that significant variability occurs at both a regional and paddock scale. Sharp contrasts are present in soil morphological features, with deep sands and heavy clays occurring directly adjacent to each other with no topographical variability. It is made clear that when this occurs the sandier soil texture has a significant, negative impact on cotton lint yield, reducing the efficacy of presently used irrigation systems. At a regional scale it is apparent that, like the within-field scale, the suitability for irrigated cotton production varies because of largely inherent soil characteristics. Finally, conclusions will be drawn with the major findings synthesised into two discussion points aimed at the stakeholders that this research was targeted towards: farmers and agronomists.

Acknowledgements

The last few years have without a doubt been some of the most challenging, but also rewarding, of my life. I have learnt an immeasurable amount and for this I am grateful. I would first like to thank the Cotton Research and Development Corporation (CRDC) for funding this project. The experiences within the industry, from conferences to grower days have been incredible and have allowed me to learn about and grow within the industry.

The journey to completing this PhD has not been easy. Beginning a PhD in the month that the world went into lockdown was not ideal, nor were the lockdowns that ensued over the next two years, especially considering this project hinged on fieldwork almost 500 km away. Then came the flooding and inability to access sample sites, Stephen, I think you may have been right when you called the project cursed. This necessitated the need to be adaptable, and alter the project as required, and this would not have been possible without my incredible supervisors.

Firstly, Stephen, it has been a privilege to complete this project under your tutelage. I could not have completed this PhD without your support and guidance. I am truly appreciative of everything you have taught me and sincerely thank you for always leaving your door open for me to ask countless questions. You have undoubtedly improved my scientific writing ability and your immense passion for soil has rubbed off on me. Patrick, you too were instrumental in allowing this project to proceed through challenges and for this I am grateful. You have advanced my data and mapping skills immeasurably and have generally brightened up each day I am in the office. It has been a magnificent experience learning from you and I appreciate your patience when I bombarded you with potentially silly questions.

The field sampling campaigns for this project were…. eventful. Firstly, James, what great times we had driving to Hay, staying a night, and then driving back because it started raining again. On a serious note, it was great to spend time in the field and at conferences with you and I appreciate your support. I would also like to thank all of the technical staff, particularly Iona, Hero, Adriana and Sami for your support throughout the project, I couldn't have done it without you. To the remaining original work 'clumpers', Mikaela and Si Yang, you have both been wonderful to work alongside and I appreciate the support you have both given me throughout the past few years. To all the other PhD candidates and post-docs at ATP it has been a pleasure to spend time with you over the last few years. Not only has it been fun, but you have all taught me a lot. Thank you to all current and former staff, particularly; Alex, Balwant, Budi, Liana, Damien, Brett, Tom and Floris. It has been incredible to see how you undertake research and your dedication to advancing agriculture is inspiring.

To all of my friends, thank you for the good times and support over the last few years, I couldn't have done it without you. Lastly, thank you to my family. I could not have gotten through this without your support and inspiration, so, Mum, Dad and Solange, please know how much I appreciate it.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABS	Australian Bureau of Statistics
AoA	Area of Applicability
ASC	Australian Soil Classification
BE	Barest earth
BOM	Bureau of Meteorology
CEC	Cation exchange capacity
CIA	Coleambally Irrigation Area
CRDC	Cotton Research and Development Corporation
DEM	Digital elevation model
DSA	Digital soil assessment
DSM	Digital soil mapping
ECe	Electrical conductivity of the extract
ESP	Exchangeable sodium percentage
ha	Hectares
ka	Kilo-annum (1000 years)
LCCC	Lin’s Concordance Correlation Coefficient
LOSOCV	Leave-one-site-out-cross-validation
MDB	Murray Darling Basin
MIA	Murrumbidgee Irrigation Area
ML	Machine Learning
NSW	New South Wales
OSL	Optically stimulated luminescence
PAWC	Plant available water capacity
Rad	Radiometric
RF	Random Forest

RMSE	Root Mean Squared Error
SHAP	Shapley Additive exPlanation
TA	Terrain attribute
TL	Thermoluminescence
USDA	United States Department of Agriculture
VisNIR	Visible near infrared
XGBoost	Extreme Gradient Boosting

Chapter 1

1. General introduction

An understanding of soil properties and how they vary spatially is essential for economically and environmentally sustainable agricultural management decision making. The highly heterogeneous nature of soils, however, can make this challenging. Historically, spatial understandings of soil variability have relied upon traditional survey methods. Within the lower Murrumbidgee region this consisted of samples being taken at various distances along spaced transects, with interpolations made based on surface inspections, landscape knowledge and recurring patterns of vegetation (van Dijk and Talsma, 1964). In Australia, however, there has been a steady decline in the number of traditional soil surveys being undertaken (Kidd et al., 2020). Since the turn of the century, improved computational capacity and the availability of spatial data has seen an increase in the application of digital soil mapping (DSM), however, the accuracy of these products is reliant on observations of soil as input data. The sparseness of this data in some regions of Australia, including the lower Murrumbidgee valley in southern NSW, has seen previous research identify that models may not be applicable to this region (Pozza et al., 2022) or are unable to identify soil variability across smaller spatial extents, such as the farm scale (Han et al., 2022). This negates the potential of these resources to be used as decision-support tools by farmers.

The lower Murrumbidgee valley in southern NSW falls centrally within the Murray Darling Basin (MDB), within southern NSW, Australia (Figure 1.1). The MDB in its entirety contributes over AUD $30 billion to the economy per annum (Murray Darling Basin Authority [MDBA], 2024). Within the lower Murrumbidgee valley are the Murrumbidgee (MIA) and Coleambally (CIA) Irrigation Areas, developed in the 1910s and 1960s, respectively. The longstanding agricultural productivity in the MIA and CIA is underpinned by knowledge gained from soil surveys undertaken prior to, or during, the development of these irrigation areas. An understanding of the complex patterns of soil distribution was able to guide infrastructure investment and allow for appropriate crops to be grown in specific areas. These traditionally ranged from horticulture on lighter, sandier soil types to rice, or other broadacre crops, on heavier clay soils (Figure 1.1). The diversity of agriculture within the region is therefore tied to the diversity of the soil.

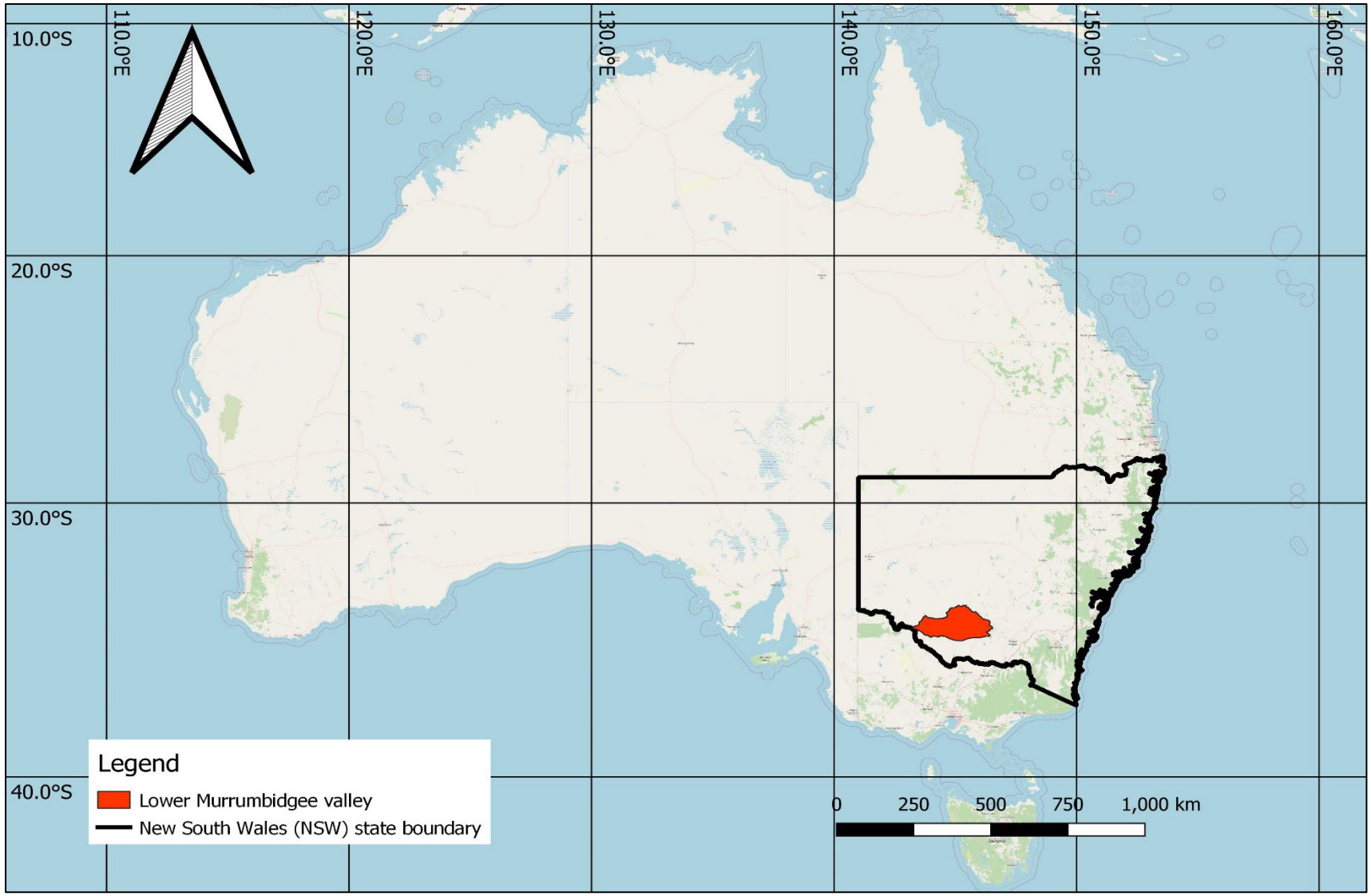


Figure 1.1. A map of Australia showing the location of the lower Murrumbidgee valley and state of New South Wales.

While the Murrumbidgee River is presently the major landscape feature of the region, it has had little role in the deposition of the plain's alluvial sediments. Instead, the surficial sediments of the lower Murrumbidgee valley have resulted from four periods of temporally distinct palaeochannel activity over the last 100,000 years. As a consequence of the changing watercourses, there are numerous points on the landscape where these depositional systems are intertwined, with highly variable soils occurring as a result. While an understanding of these geomorphic features informed surveys in the east of the region, beyond the bounds of the MIA and CIA there is a paucity of available soil data, and information more broadly. Even within the surveyed areas, anthropogenic influences are likely to have significantly altered soils, with land reformation a common practice to achieve consistent gradients for irrigation. Further, much of the legacy data for the region is not geolocated, reducing its efficacy in either understanding spatial-temporal variability or being incorporated into contemporary DSM exercises.

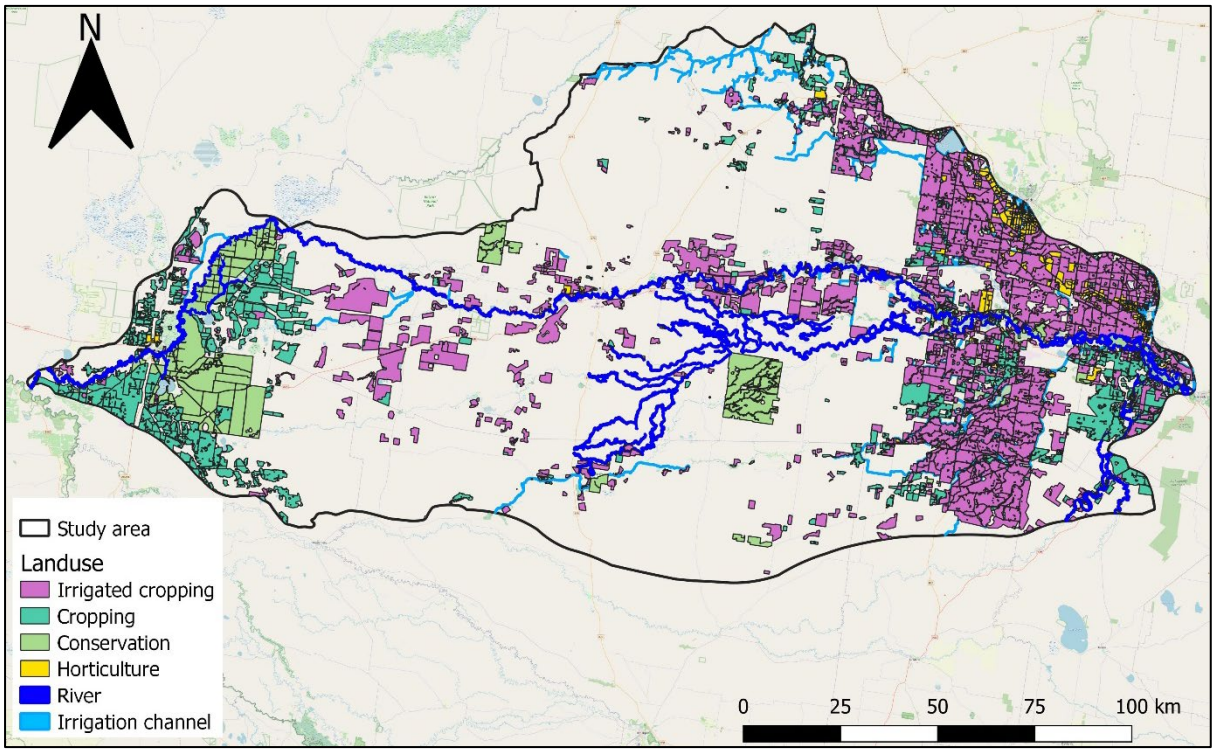


Figure 1.2. The distribution of landuses in the lower Murrumbidgee region, including irrigated broadacre cropping, dryland cropping, conservation and horticulture (NSW Department of Climate Change, Energy, the Environment and Water [DCCEEW], 2023). Also shown are the Murrumbidgee River and its distributary systems as well as irrigation supply channels. The majority of the uncoloured area is used for grazing.

The lack of accurately mapped soil types in the area, as identified by Holland and Eastwood (2014), is largely due to a dearth of available soil data within the region. This can be visualised through the density of sample sites contained within the National Soil Site Collation (Figure

1.2). This issue is not unique to the lower Murrumbidgee valley, with patches of both dense and sparse soil data spread throughout Australia (Figure 1.2). This can result in spatial autocorrelation, where DSM model quality statistics are inflated due to clusters of datapoints. When these models make predictions into an area where data is sparse, such as the lower Murrumbidgee valley, the model may inaccurately represent the landscape due to the representation of data in dense as opposed to sparse areas, diminishing the value of the product. An example of this was seen in a study mapping sodicity across the MDB using collated soil surveys and publicly available data (Pozza et al., 2022). An area of applicability analysis showed that the model was not applicable to almost the entirety of the lower Murrumbidgee valley (Pozza et al., 2022). The impact of this paucity of information was specifically noted following the expansion of the cotton industry from traditional growing regions in northern NSW, including the Gwydir, Namoi and Macquarie valleys, into the south of the state (Holland and Eastwood, 2014).

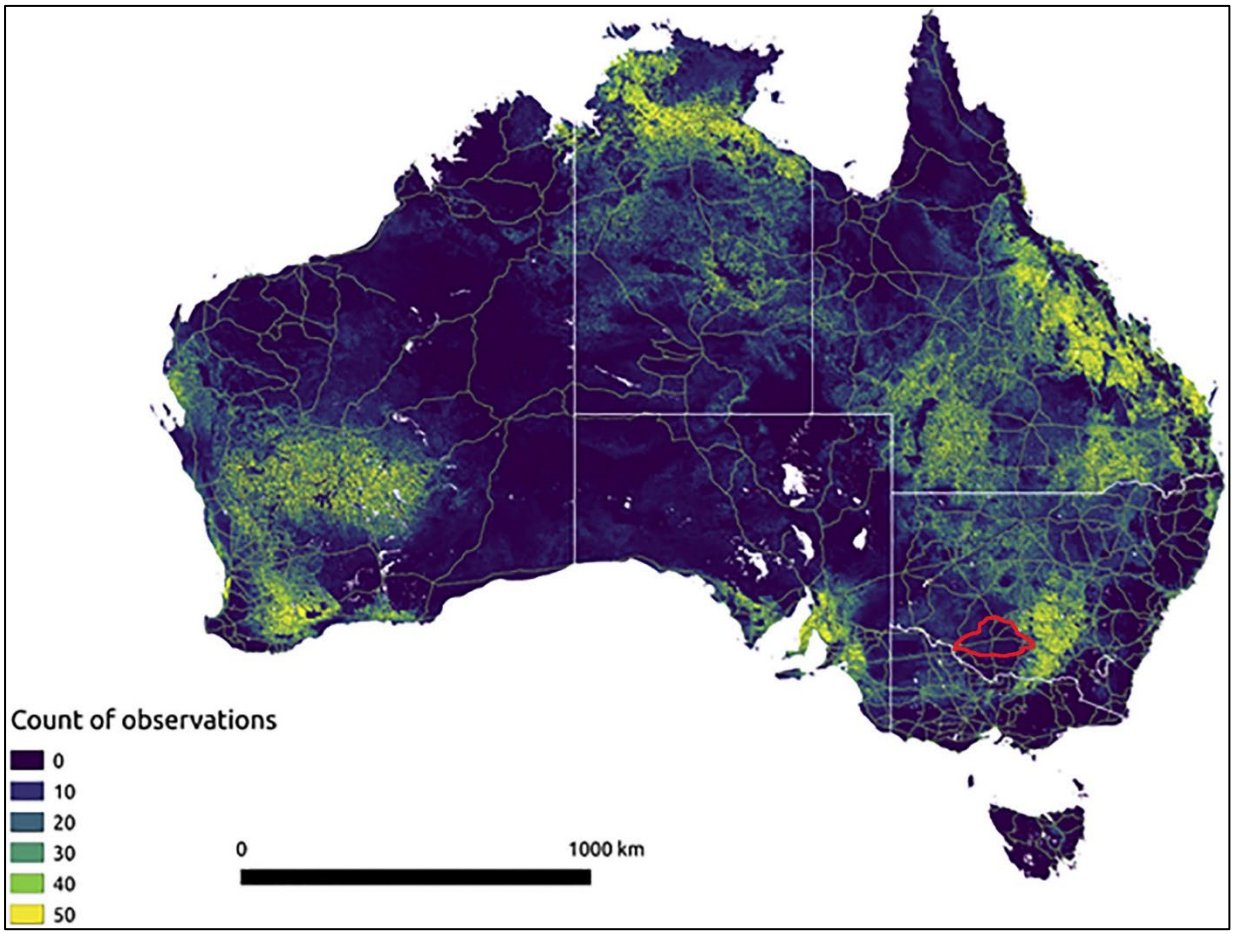


Figure 1.3. A map showing the relationship between soil observations and the covariate feature space that these samples are able to represent. Yellow areas have a greater density of samples in the covariate feature space that is associated with the sample locations (Searle et al. 2021, Fig. 1). The approximate bounds of the lower Murrumbidgee valley are shown in red.

Australia is one of the top 10 producers, and the third largest exporter, of cotton globally (Khan et al., 2020) with a production forecast of 5 million bales in 2023/24 (USDA, 2024). Cotton production in Australia first came to prominence in the 1960's with little expansion into southern NSW in the following decades despite a despite a research centre in Griffith (Conaty et al., 2022). Over the last decade, however, the industry has made a resurgence as a profitable and attractive crop to farmers within the region. Gupta and Hughes (2018) note an increase in cotton plantings from less than 5,000 hectares (ha) before 2010-11 to roughly 35,500 ha in the 2015-16 season. As a primarily irrigated crop, the area sown to cotton annually varies in accordance with water allocations (Conaty et al., 2022). The USDA (2019) note that in the drought affected 2018/19 season, southern growing regions accounted for 25% of Australia's total cotton crop with an area of 43,639 ha sown (ABS, 2020). In the recent 2023-24 season approximately 60,000 ha of cotton was planted across the lower Murrumbidgee valley (S. O'Rafferty, personal communication, June 28, 2024).

Along with being grown in the MIA and CIA, significant areas of land have been developed for irrigated cotton production in the west of the region (Figure 1.1). In these reaches, much of the prior soils research to have occurred in the east of the region is not applicable due to differences in the soils present. A lack of effectively mapped soil, and the associated lack of understanding of inherent properties, underpins knowledge gaps identified by Holland and Eastwood (2014) in their soil scoping study for cotton in southern regions. Their study evaluated the current understanding of cotton-growing soils and soil management, reaching strong conclusions regarding the lack of presently available soil data and information for stakeholders. As the cotton industry has evolved primarily in northern NSW, this is where much of the soils-focused research has occurred (Cattle & Field, 2013). While there are opportunities for this research to be translated to new areas, an understanding of the inherent condition and capacity of the soil is first required. Holland and Eastwood (2014) noted that "cotton growers need to know more about their soils so that they are able to refine soil-related management practices". The key takeaway from this report is a need for southern NSW best management practices to be developed; however, before this can occur, the distribution of soils and potential constraints such as sodicity and alkalinity need to be understood (Holland and Eastwood, 2014). The same issue has been investigated near the town of Hillston in the neighbouring Lachlan valley, with several papers published on the topic providing an improved understanding of how soils vary between growing regions and the impact of irrigation on temporal change (Filippi, 2017; Filippi et al, 2018a, 2018b; Onus et al., 2003).

This thesis has evolved from the findings of the scoping study undertaken by Holland and Eastwood (2014), alongside discussions with cotton industry stakeholders within the lower Murrumbidgee valley. The chapters therein explore the known variability of the region's soils while developing outputs to fill identified knowledge gaps.

The literature review of this thesis aims to provide context to the region, specifically focusing on the sources of soil variability and patterns of soil distribution. A review of current knowledge surrounding geology, geomorphology and soil is essential before considerations are made on the future enhancement of the region's soils knowledge. Due to the influence of ancient geomorphic features on this landscape, what appears to be a flat plain is actually a complex network of paleochannel sediments both at and below the surface. This thesis utilises a dataset of 153 soil cores extracted to a depth of 1 m from across the lower Murrumbidgee valley. A theme throughout is the application of proven scientific methods to understand the characteristics of soils and how they are distributed, creating information that will be of benefit to the end-user. Specifically, the aims of this thesis with their research questions are:

1. To morphologically classify lower Murrumbidgee valley soils and discuss their occurrence in relation to palaeochannel systems at the regional and within-field scale
    - **a.** What are the typical morphological characteristics of lower Murrumbidgee valley soils?
    - **b.** Can soil cores extracted from across the entirety of the region be appropriately classified using previously-developed soil groups?
    - **c.** How is the distribution of soils related to palaeochannel systems?
    - **d.** What is the extent of soil morphological variability at the within-field scale?

2. To assess how variability of inherent soil properties at the within-field scale impacts irrigated cotton production
    - **a.** What is the extent of soil variability between yield-derived production zones within paddocks used to grow cotton?
    - **b.** How are inherent soil properties correlated to cotton lint yields?
    - **c.** Can paddock scale digital soil maps accurately represent soil variability spatially?

3. To model and map multiple soil properties across the lower Murrumbidgee valley

   a. Can digital soil mapping be used as a tool to accurately model and then map soils at multiple depths concurrently?
   b. How do presently available digital soil mapping products perform when validated against external datasets from the lower Murrumbidgee valley at the point support?

4. To identify the most limiting factor to, and classify the overall suitability for, irrigated cotton production in the lower Murrumbidgee valley
   a. How does the depth to which soil chemical constraints are reached vary across the region?
   b. What are the soil-landscape factors most limiting to irrigated cotton production?
   c. What proportion of the lower Murrumbidgee is considered well suited to irrigated cotton production?

In addressing these aims and answering these research questions, this body of work does not intend to create novel approaches to DSM. Instead, it seeks to use DSM as a tool to develop information for stakeholders surrounding the distribution of fundamental soil properties and how they relate to irrigated cotton production in the lower Murrumbidgee valley. The outcomes of this thesis will improve our foundational understanding of soils and their spatial variability within the lower Murrumbidgee valley. This information can then be used to make informed decisions on where cotton should be grown within the region, guide management decisions and provide a platform to explore more applied research questions investigating farming systems. We cannot manage what we do not know, and this study seeks to provide the basis for the informed management of soils in irrigated cotton systems within the lower Murrumbidgee valley.

Chapter 2

2. Understanding the development and distribution of soils within the lower Murrumbidgee valley, NSW, Australia.

Abstract

To understand the soils of the lower Murrumbidgee valley in southern NSW, one must first look backwards at their evolution before looking forwards at how they can be sustainably managed. While the area consists almost entirely of fluviatile sediments, the central feature of this region, the Murrumbidgee River, has had little influence on the landscape. Instead, the river incises itself within sediments that were deposited by numerous, temporally distinct palaeochannel systems. Complicating this spaghetti like mosaic of sediments are the aeolian elements of parna and remobilised sand dunes. Understanding the features of each geomorphic process is essential in understanding soil distribution, characteristics and classifications. This review identifies the breadth of studies, and differing hypotheses, within the field of geomorphology as well as how they have informed understandings of soil distribution. Not described in literature is how the geomorphic-soil relationship becomes more complex in the west of the region, when numerous paleochannel anabranches result in fine scale variability. A consequence of this is that sandy and clayey soils occur adjacent to each other with no topographical change. An examination of literature from the past 70 years also highlights the importance of appropriately cataloguing data and making resources available. While it is acknowledged that the region has a rich history of landscape and soil studies, many of the latter are presently redundant as they have either not been digitised or are not publicly available. This has flow on effects, with a lack of data reducing the efficacy of new digital soil mapping (DSM) methods, the limitations and opportunities of which are considered. The availability of this data could also allow studies to be undertaken assessing temporal change as a result of management. This narrative review seeks to journey from geology to technology, synthesising knowledge from various disciplines to highlight how sequences of sediment deposition have resulted in the surficial soils we study today. In doing this it will discuss what has been done and what is known, identify limitations to previous studies and our current understanding of the soils more broadly and briefly present opportunities to address these. Ultimately, it highlights how synthesising legacy knowledge from various disciplines can help to understand and tell the story of the soils we study today.

2.1. Introduction

The lower Murrumbidgee valley of southern NSW falls centrally within the Murray Darling Basin (MDB) of southeastern Australia. The MDB covers an area of 1,061,469 km^2 across New South Wales, Victoria, the Australian Capital Territory, Queensland and South Australia. The Basin is a major food bowl, contributing over AUD $30 billion to the economy annually with the Murrumbidgee Irrigation Area (MIA), a 3789 km^2 subregion of the lower Murrumbidgee valley, contributing AUD $5 billion (MI, 2010; MDBA, 2024). Recently, changing climatic conditions have placed pressure on agricultural production systems with prolonged periods of drought limiting water availability for irrigation (Wheeler et al., 2019). While favourable soils in the region have contributed to strong agricultural productivity, a recent scoping study identified significant soils-related knowledge gaps within the region (Holland and Eastwood, 2014).

While the MDB includes 77,000 km of river systems, these 'current' watercourses have not significantly influenced the development of the landscape and its soils. One such major river is the Murrumbidgee, which is presently the central feature of the lower Murrumbidgee valley. While many may consider this subregion of the MDB as a relatively lifeless plain, it is rich in soil diversity and has been described as containing some of the most characteristic landforms in Australia (Bowler, 1978).

The lower Murrumbidgee is located centrally within the fluvio-lacustrine area of the Murray Basin, the geology of which is described by Brown (1989). Three geologically distinct sedimentary units; the Renmark Group, the Shepparton Formation and the Calivil Formation, are presented by Brown and Stephenson (1991). The Shepparton Formation was further differentiated into the 'upper Shepparton' and 'lower Shepparton' formations by Prathapar et al. (1997) with a younger 'Coonambidgal Formation' incising itself within the former (Pucillo, 2005).

These youngest units, the upper Shepparton and Coonambidgal Formations, are complex and the distribution of the soils therein is tied to geomorphology. Numerous studies, such as Butler (1950;1958), Langford-Smith (1960), Butler et al. (1973), and Page and Nanson (1996) have described the geomorphological features of the lower Murrumbidgee, including the dominant fluviatile and subdominant aeolian elements. Given that patterns of sediment and soil distribution result from these processes, there is value in reviewing this work. Further complexity is added to understanding the landscape due to varying nomenclature, for example,

the fluviatile systems were initially described as prior streams (Butler, 1950) but will be referred to in this review as ‘palaeochannels’ owing to more recent studies (Page et al., 1996).

Over the last 100,000 years fluctuating climatic conditions determined cycles between sediment deposition and soil development, resulting in the intersection of alluvial fans, mixed with layered aeolian deposits of varying depths. To understand these interactions, numerous stratigraphic models have been developed to explain landscape degradation, sediment aggradation and the resulting soil distribution at a conceptual level (Butler, 1958; Page & Nanson, 1996; Pels, 1964, 1971; Pucillo, 2005; Schumm, 1968). In these too, however, complexities arise with units in these models not centrally named, instead, given geographically specific titles. Identifying trends in this work allows for a general understanding of how sediment deposition varies perpendicular to paleochannel systems.

The success of the Murrumbidgee and Coleambally Irrigation Areas, within the regions east, has been underpinned by soil surveys undertaken between the 1930s and 1970s (Butler, 1979; Taylor & Hooper, 1938; Stannard, 1968, 1970; van Dijk, 1958, 1964; van Dijk & Talsma, 1964). These surveys, combined with geomorphological studies, allowed for soil features such as subplasticity and waterlogging susceptibility to be understood. Over the last 80 years extensive work has also examined the soil attributes with an agronomic focus (Lang and Hicks, 1975). Much of this work, however, has been confined to the bounds of the irrigation areas, beyond which there is a lack of publicly available data. It is in these less understood areas that interactions between the lower reaches of differently aged palaeochannel systems become more complex, resulting in ‘mosaic’ like variability (Murphy et al., 2000).

More recently, digital soil mapping (DSM) resources have become available, however, their efficacy in translating information spatially is reduced due to low data quantity and quality. Further, soils may have been anthropogenically altered since being initially examined. Despite this, technological advancements provide opportunities to harness legacy maps (Malone et al., 2017) and expert knowledge (Bui et al., 2020) in the development of digital products.

Previous reviews of the lower Murrumbidgee valley have examined geological and geomorphological processes (Pucillo, 2005), collated historical soil data (Hornbuckle and Christen, 1998; Hornbuckle et al., 2008a, 2008b; Thacker et al., 2008) and identified limitations in understandings of soil as they relate to cotton production (Holland & Eastwood, 2014).

This narrative review seeks to synthesise these branches of study, from the geology to the technology, through discussions on the ancient landscape processes and how these relate to past

soil studies and patterns of soil distribution. Lastly, the review will discuss limitations with presently available resources while also presenting opportunities to enhance our understanding of the landscape.

2.2. Study area

The area examined in this review (Figure 2.1), termed the 'lower Murrumbidgee valley', is akin to the classification of the Murrumbidgee subregion within the Riverina bioregion by Morgan and Terrey (1992). As determined by the Interim Biogeographic Regionalisation for Australia (IBRA7), the Riverina bioregion is one of 89 within Australia accounting for 8.9% of New South Wales's landmass covering approximately 9,580,000 ha. The Riverina has also been described and termed the Riverine Plain by Pain at el. (2011) in their derivation of Australia's physiographic regions. The lower Murrumbidgee valley totals 2,577,100 ha (25,771 km^2) encompassing towns including Griffith, Leeton, Hay, Balranald and Coleambally. The region extends from 145.34°, -33.87° in the north to 145.21°, -35.18° in the south, and, 146.53°, -34.73° in the east to 143.21°, -34.71° at the westmost point (Figure 2.1).

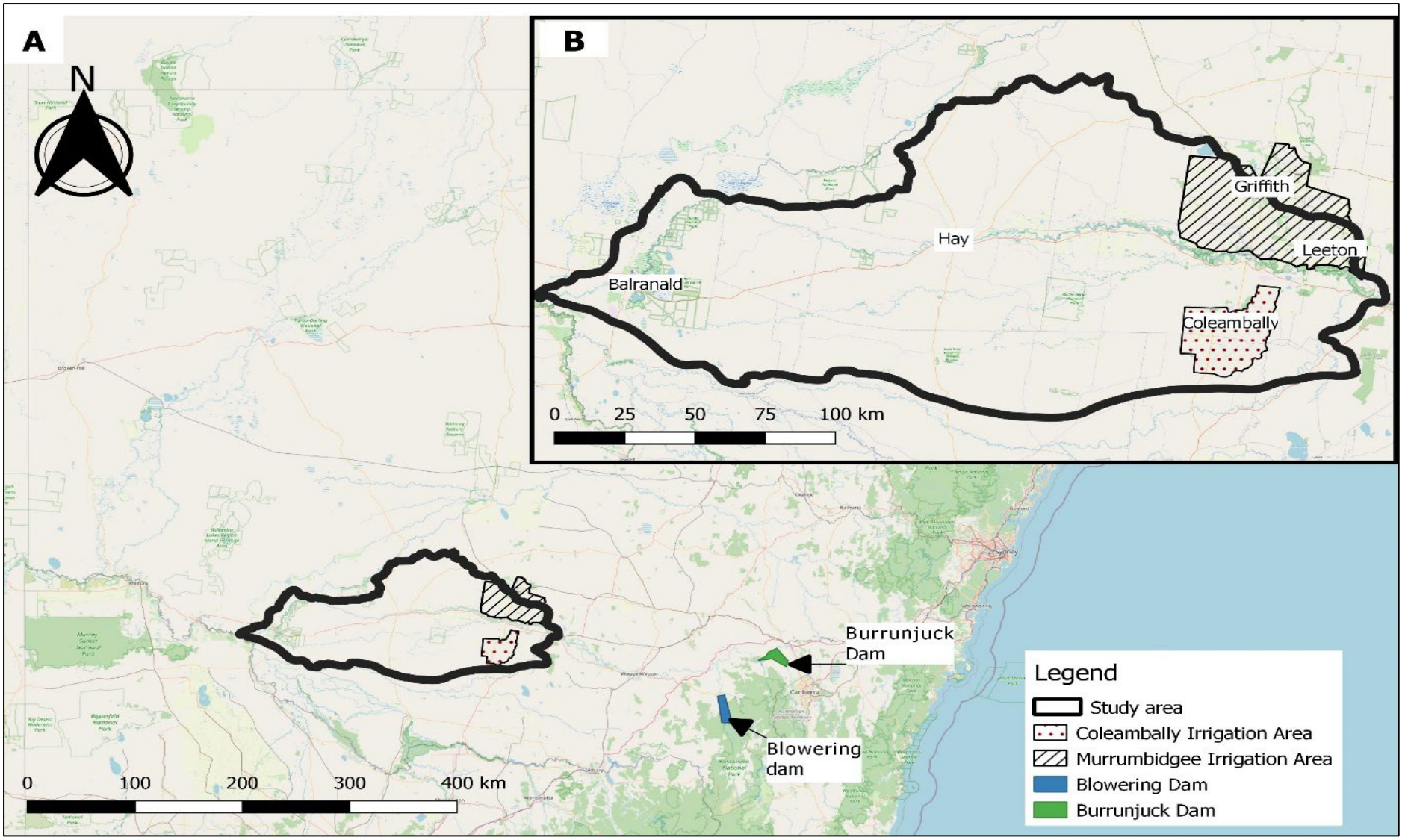


Figure 2.1. A) The location of the lower Murrumbidgee valley within New South Wales. Shown also are the locations of Burrunjuck and Blowering Dams. B) The study area in relation to the Murrumbidgee and Coleambally Irrigation areas.

The Murrumbidgee River is the defining feature of the modern landscape and has been central to the region's economic development. Along the river, weirs are used to manage water flows while allowing for diversions away from the river into irrigation supply channels. There are no major tributaries, however, a distributary of the Murrumbidgee River is Yanco Creek which diverts from the river south of Leeton in a south-westerly direction. The creek continues as the Billabong and Colombo Creek before meeting the Edward River, itself an anabranch of the Murray River. In the rest of the region there are numerous channels that flood seasonally, filling lakes; including Yanga Lake, and swamps; such as Cumbung Swamp.

In contrast to the flat, alluvial plains of the lower Murrumbidgee, the broader catchment and Murrumbidgee River rise in the Great Dividing Range at elevations of over 2000 m. It is in the upper catchment where Burrunjuck (1,028,000 ML capacity) and Blowering (1,628,000 ML capacity) Dams are located (Figure 2.1). The completion of Burrunjuck Dam (1912) allowed for irrigation within the Murrumbidgee Irrigation Area (MIA). Similarly, the construction of Blowering Dam (1968) provided the opportunity for development of the Coleambally Irrigation Area (CIA) while storing water from the Hydro-Electric Scheme. As a source of water and sediment on the alluvial lower Murrumbidgee, land management and mineral weathering in the upper sub-catchments must be considered as a source of salinity which may impact the condition of soil downstream (Conyers et al., 2008). Olley and Scott (2002) provide a thorough report on the state of the river and its changes over the 180 years since European settlement.

2.2.1. Climate

The current climate of the broader Riverina region is classified by Stern et al. (2000) as dry and semi-arid with cool winters and hot summers. Rainfall decreases along an east-west gradient from 432 mm (Leeton) to less than 350 mm (Balranald) (Bureau of Meteorology [BOM], 2023). At a localised scale, Zhou et al. (2009) applied spatial analysis techniques utilising 30 years' climate data to produce two schemes of climatic zoning across the eastern MIA. The impact of past climate conditions is discussed in relation to geomorphology and soil development within this review.

The climatic conditions within the upper Murrumbidgee catchment are the dominant factor affecting river flows and the availability of water for irrigated agricultural production. Green et al. (2011) provided a report detailing the water resources and management of the broader

catchment. In some areas above the primary water storages, average annual rainfall exceeds 1500 mm (BOM, 2023).

2.2.2. Agricultural production

Access to water for irrigation through either extraction from the Murrumbidgee River or ground water aquifers allows for diverse agricultural production. Variability in productivity is driven by seasonal water availability and variations in soil type. The development of irrigation infrastructure in the MIA during the 1910s allowed for the establishment of highly successful horticultural industries on the region's lighter soils, with rice common on heavier soil types. Similarly, the development of the CIA in the 1960s allowed for the expansion of irrigated agriculture into areas with heavier soils more conducive to broadacre cropping. Substantial changes have occurred to the dominant industries of the region over the last century with notable expansions of cotton (annual) and tree nut (perennial) crops. Presently, broadacre irrigated summer crops including rice, maize and cotton, in association with winter cereal crops, are common on heavier soil types. Permanent plantings of citrus and vineyards are common, along with tree nut crops on lighter soils. Areas not developed for irrigation, particularly in western parts, are used for grazing. The increase in perennial plantations has the potential to place pressure on water allocations. While essential in the economic development of the region, the allocation of water for agriculture is a divisive issue in some communities. Kandasamy et al. (2014) used the Murrumbidgee basin as an example to examine the changing dynamics of communities in balancing environmental sustainability and agricultural production.

2.3. Landscape characteristics, geology and geomorphology

2.3.1. Geological evolution

Excluding minor occurrences of bedrock outcrops in the far east, the study area is comprised entirely of Cenozoic era fluvio-lacustrine sediments which fill the Murray Basin. Tectonic activity in the early Cenozoic era (65 mya–present) resulted in the formation of the Murray Basin, as described by Brown and Stephenson (1991), through the upwarping of Proterozoic (2500–541 mya) and Palaeozoic (541–252 mya) rocks which are presently observed on the eastern boundaries of the now fluviolacustrine basin (Page et al., 1996). The depth to bedrock underlying these early Cenozoic sediments generally increases from east-to-west. Pucillo

(2005) stated a maximum depth of 600 m in the central west of the Plain, while Pels (1968) noted depths to bedrock of 153 m near Coleambally, 465–156 m in the MIA and over 300 m near Hay. Brown (1989) presented a schematic cross section categorizing the sediments into three geologically distinct units; the Renmark Group and the Shepparton and Calivil Formations of the Wynghu Group (Figure 2.2). Detailed discussions on these geological units are provided by Brown (1989) and Brown and Stephenson (1991).

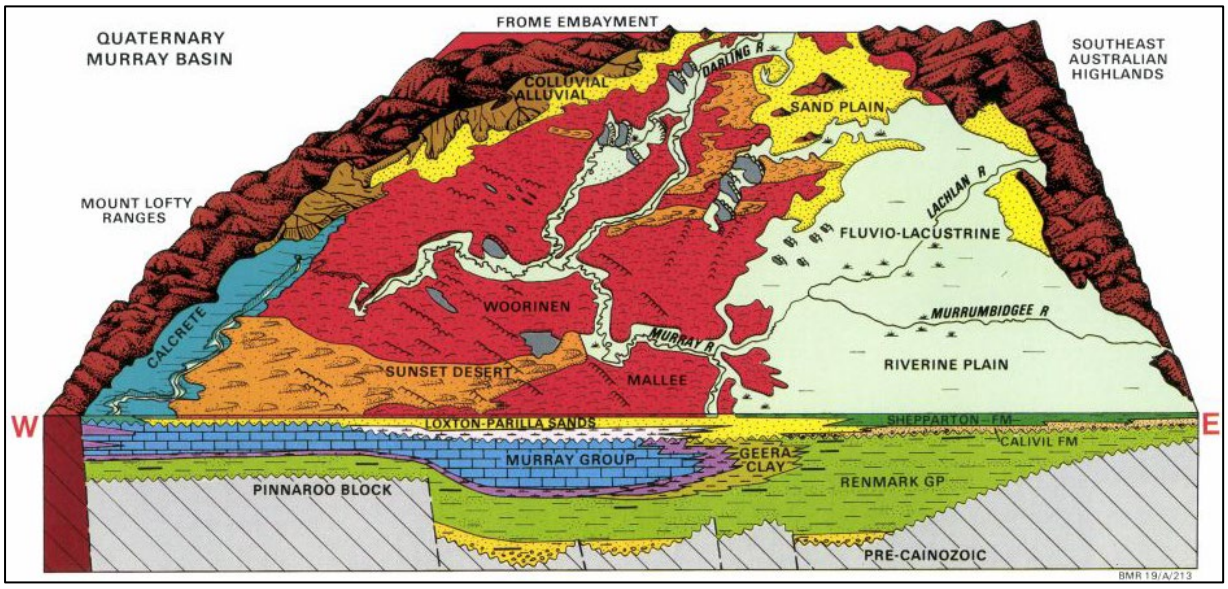


Figure 2.2. A cross sectional diagram of the Murray basin by Brown (1989, Figure 5) illustrating the presence of fluvio-lacustrine material in the central Plain with the three major stratigraphic units to a depth of approximately 600 m.

A brief description on various phases of geological evolution is provided by Page et al. (1996). Of relevance to this study is the onset of semi-arid conditions from 0.5-0.4 mya when, alongside the formation of aeolian landscapes in the neighbouring Mallee region (Wasson, 1989), fluvial deposits of the Shepparton Formation were laid down, forming the modern Riverine Plain (Page et al., 1996). Prathapar et al. (1997) observed significant differences in hydrological properties at increasing depth at sites surrounding Coleambally leading to the splitting of the Shepparton Formation into two; an upper and lower formation. Within the CIA, Pels (1968) discussed the general geology of unconsolidated sediments and outcrops of consolidated sediments through the examination of 750 boreholes, providing data to understand groundwater movement and salinity (Pels, 1968).

A new, younger sedimentary unit associated with the active river and recently active palaeochannel systems, the Coonambidgal Formation, is also present, incising itself within the upper Shepparton Formation sediments. It is these upper Shepparton and Coonambidgal

sediments that have been the focus of geomorphic studies and are most relevant to the 'modern' agricultural soils of the plain.

2.3.2. Geomorphology

The lower Murrumbidgee valley, and broader Riverine Plain, which may appear monotonous and bland, is in fact rich in geomorphic complexity even within geological units. Studies of these geomorphological characteristics have been essential in the stratigraphic categorizing of sediments and understanding soil distribution. Since the pioneering work of Butler (1950), several studies have focused on understanding relationships between the dominant fluviatile and secondary aeolian elements. Despite conjecture regarding the correct nomenclature in early work, the relic fluviatile elements will be termed 'palaeochannels' in this review. The relationship between these two geomorphic features is important. Bullard and McTainsh (2003) reviewed interactions between aeolian and fluvial elements within landscapes to highlight the importance of examining interactions concurrently as opposed to the traditional approach of examining each domain exclusively. For example, the activity of source bordering dunes is directly influenced by the movement of fluviatile sediments which, in turn, are a product of the climate. Locally, Fried (1993) postulated that increased sinuosity in younger palaeochannels was a result of increased sediment loads caused by increased parna deposition which was then remobilisation in the upper catchment. Butler et al. (1973) synthesised prior geomorphological research to produce a map of the distribution of geomorphic features across the entire Riverine Plain.

2.3.2.1. Palaeochannels

Palaeochannels refer to previous river systems which crossed landscapes, in this instance the lower Murrumbidgee valley, depositing vast amounts of sediment across the plain. This section of the review draws heavily on the work of Pucillo (2005) who provided a detailed review on the development of theories relating to palaeochannels in the Coleambally region. The present Murrumbidgee River system contains a small sedimentary load compared to the palaeochannel systems which previously crossed the landscape (Schumm, 1968). Estimates of the streamload characteristics of palaeochannels is vital in understanding sediment distribution which impacts soil characteristics across the landscape. Discussion on these influences focuses on two phases of palaeochannel activity: 'aggradation', where sediments are deposited resulting in the infilling of channel beds, and, 'degradation', where sediments from the streambed are eroded, carving out new or deeper palaeochannel pathways.

Discussions of palaeochannels within the region have evolved since the theory of 'prior streams' as a basis of soil formation across the Riverine Plain was first proposed by Butler (1950). It was postulated that the region's soils are reflections of the patterns of low sinuosity streams extending westward across the area consisting of aggraded bed-load channels with levees (Butler, 1950; Pucillo, 2005). In developing a model for stratigraphy, Butler (1958) proposed separate phases of prior stream activity, where deposition occurred during arid phases.

Langford-Smith (1960) questioned the periodicity and climatic assumptions of Butler's theory, instead suggesting a system of a single riverine phase, akin to modern deltaic river systems with deposition occurring under more humid conditions (Figure 2.3). This theory was later modified to be more similar to those of Butler (1950, 1958), whereby channel incision occurs under established pluvial periods while deposition occurs as these phases wane (Pucillo, 2005).

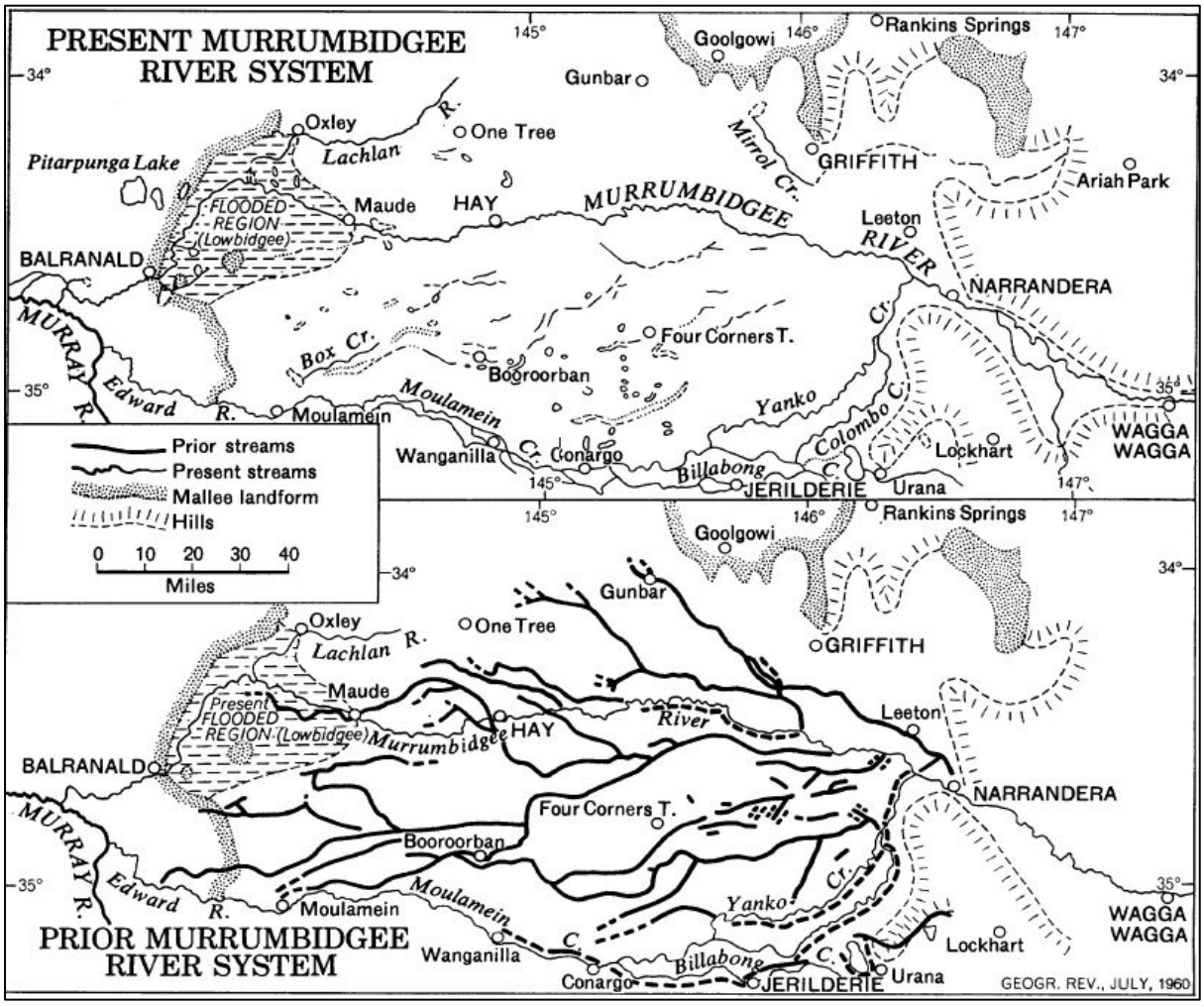


Figure 2.3. A map produced by Langford-Smith (1960, Figure 4) showing deltaic like pathways of prior streams across the lower Murrumbidgee valley (bottom). The western region surrounding Balranald provides a nexus of lower floodplain systems alongside the eastern Mallee landforms. The pathways of present streams are also shown (top).

In undertaking work around the Cadell fault in the neighbouring Murray valley, Pels (1966) identified that the youngest prior stream systems defined by Butler (1958) were more complex than had been previously considered. Pels (1966) instead presented a new categorisation of 'ancestral rivers' which post-dated the prior stream systems but pre-dated water courses of the present-day. He theorised that these ancestral rivers incised themselves within the sediments of prior stream systems and while prior streams dissipated on the plain, the ancestral rivers continued beyond, exhibiting a higher suspended streamload and greater sinuosity (Pels, 1966).

Schumm (1968) discusses the importance of climate, and impacts of its changes, to streamloads of palaeochannel systems, suggesting that the ancestral rivers developed during a humid phase with higher runoff and peak-discharges than both the modern river systems and prior streams. Findings are presented on how different alterations to the hydrological properties of the systems impact their form. For example, a decrease in sand load and increase in discharge between prior streams and ancestral rivers resulted in substantial morphological changes, with channel depth and sinuosity increasing but gradient, meander wavelength and width decreasing (Schumm, 1968). These, and other changes, have significant impacts on presently observed landscape characteristics. Fried (1993) questioned that changes to channel morphology could not be a result of catchment changes alone (Bullard & McTainsh, 2003). It was instead proposed that increased upland aeolian deposition during the Last Glacial Maximum resulted in increased sinuosity as a result of higher streamloads when these sediments were remobilised as conditions changed (Fried, 1993). Consequently, the deposition of these materials on banks in low-flow periods may have stabilised the river system resulting in the observed development of the more sinuous, younger, ancestral rivers (Fried, 1993).

Bowler (1978) identified limitations to the terminology representing the previous ancestral rivers and prior streams. The difficulty in classifying systems into these two types is shown where attributes relating to both ancestral rivers and prior streams are observed within the same channel system (Bowler, 1978). To overcome this Bowler (1978), in the neighbouring Murray valley, utilised the term "complexes" and the application of geographic names to identified channel units.

Prior to the advent of thermoluminescence (TL) dating techniques researchers were limited by the capacity of radiocarbon dating methods to identify sediment ages. In undertaking TL dating on sediments from the extensive borehole network within the Murrumbidgee, Page and Nanson (1996) and Page et al. (1996) showed the terminology of prior streams (Butler, 1950) and ancestral rivers (Pels, 1971) to have been incorrectly applied due to false assumptions on their

chronology and mode of formation. Instead, they adopted the term palaeochannels, identifying four periods of activity, terming them, the Coleambally phase (occurring from 105–80 ka), the Kerarbury phase (occurring from 55-35 ka), the Gum Creek phase (occurring from 35-25 ka), and the Yanco phase (occurring from 20-13 ka). The modern flow regime began to be established approximately 12 ka. More recently there has been conjecture surrounding these sediment ages. Mueller et al. (2018) employed optically stimulated luminescence (OSL) alongside TL dating methods to revisit sediment chronology of the Gum Creek and Yanco palaeochannel systems. For samples examined using both methods it was shown that TL ages are between 54 and 73% older than the corresponding OSL ages (Mueller et al., 2018). Consequently, it was proposed that the periods of enhanced palaeochannel activity for the Gum Creek and Yanco systems are revised to 41-29 and 29-18 ka, respectively (Mueller et al., 2018).

Differences in characteristics have been observed between the four palaeochannel phases. Channels of the older, Coleambally phase, are described as being “mixed-load, laterally migrating sinuous palaeochannels with occasional transitions to a straighter bedload-dominated mode” concluding with a “bedload dominated episode resulting in aggradational palaeochannels on the surface of the Plain” (Page & Nanson, 1996). The Gum Creek phase exhibited the same characteristics, however, vertically aggrading bedloads are only observed on the downstream reaches (Page & Nanson, 1996). The younger Yanco system exhibits no terminating aggradational episode, being “characterised entirely by large mixed-load sinuous migrational palaeochannels” (Page & Nanson, 1996). The Manning equation was used on reconstructed cross sections at different locations for each palaeochannel phase to estimate the bankfull discharge and width of the systems, showing each to be significantly greater than the modern Murrumbidgee River, concurring with prior hypotheses (Table 2.1).

Table 2.1. Estimated bankfull discharges based on reconstructed cross-sections of palaeochannel systems at locations within the lower-Murrumbidgee, adapted from Page and Nanson (1996, Table 2). Palaeodischarge ratio was determined by dividing the estimated palaeochannel bankfull discharge by measured present bankfull discharges at Darlington Point (313 m3 s-1) for the Coleambally, Kerarbury and Gum Creek systems and at Hay (278 m3 s-1) for the Yanco system (from Page and Nanson, 1996). Also shown are statistics for the modern Murrumbidgee river at Hay, combining data from Page and Nanson (1996) and Speer (2018)* who averaged data from 1873-1983 from the town gauge.

System	Reach	Slope	Mean Depth (m)	Width (m)	Mean velocity (m s^{-1})	Bankfull discharge (m^3 s^{-1})	Palaeodischarge Ratio
Coleambally	Bundure	0.00025	6.5	165	1.59	1740	5.6
Coleambally	Yamma	0.00027	5.5	185	1.47	1500	4.8
Kerarbury	Waddi	0.00026	7.0	220	1.7	2160	8.3
Gum Creek	Tombullen	0.00018	5.0	215	1.13	1220	3.9
Yanco	Rhyola	0.00010	5.5	250	0.90	1240	4.5
Murrumbidgee River	Hay	<0.0002	2.6*	115*	-	278	N/A

A map of the palaeochannel systems was produced, with accompanying descriptions of their characteristics (Figure 2.4). While similar to earlier maps, including the geomorphic map of Butler et al. (1973), accurate sediment ages and improved insights into streamload and channel morphology add significant value in understanding the landscape influences of each system.

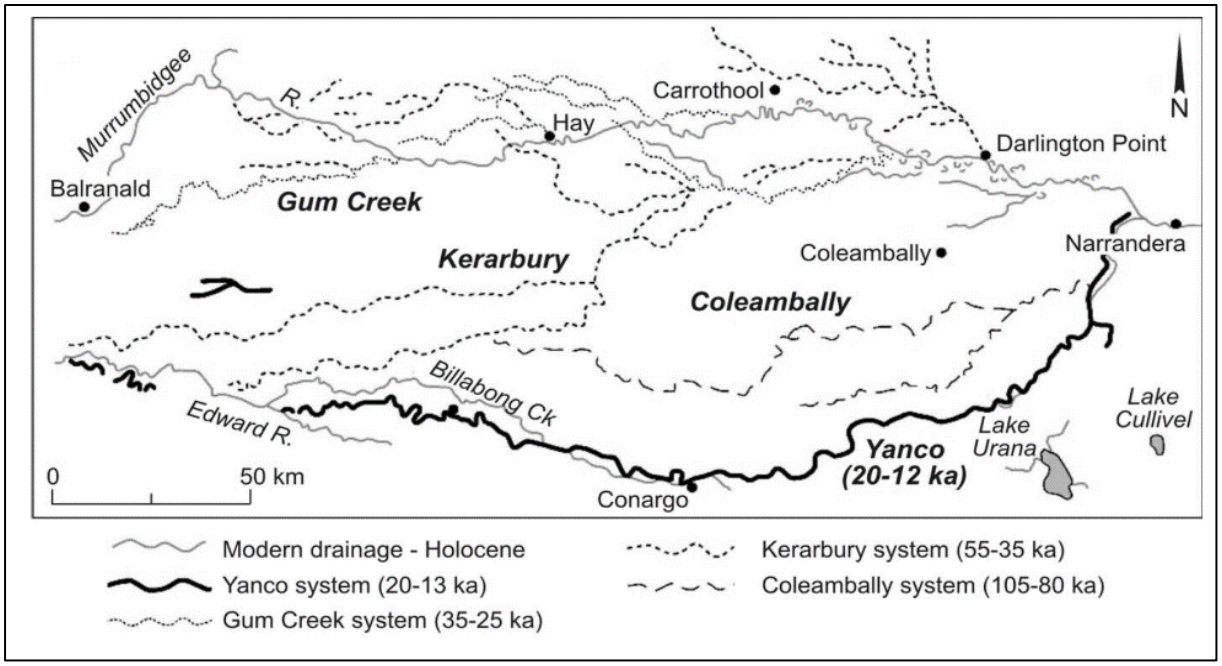


Figure 2.4. A map of the Late Quaternary palaeochannel systems of the lower Murrumbidgee valley presented by Page et al. (2009, Figure 2).

Technological advancements and the utilisation of remote and proximally sensed data have allowed palaeochannels to be studied in greater detail. Alongside discussing the development

of theories relating to the region's palaeochannels, Pucillo (2005) examined four periods of palaeochannel activity within the Coleambally area, referring to distinct systems as 'Types'. Type 1 units consist of already identified palaeochannel sediments within the Upper Shepparton and Coonambidgal geological units, however, the older Type 2, 3 and 4 facies were previously undifferentiated (Pucillo, 2005). These channels occurred at depths of 12-25 m and >25 m, described as being laterally extensive and belonging to the "coarse channel fluvial group" (Pucillo, 2005).

Fuentes et al. (2020) identified the potential of Natural Language Processing (NLP) modelling, machine learning and spatial interpolation techniques to produce 3D lithological maps of the Coleambally region. This approach has the potential to automate 3D lithological mapping from text inputs allowing for the qualitative, interpreted data to be utilised quantitatively (Fuentes et al., 2020). The importance of accurately mapping and characterising palaeochannel systems extends beyond accounting for patterns of soil distribution. Buried palaeochannel systems strongly influence groundwater hydrology, impacting the suitability for, and efficiency of, irrigated agricultural systems. Similarly, appropriate identification can present the opportunity for managed aquifer recharge processes to occur (Harvey et al., 2024; Page et al., 2023).

2.3.2.2. Aeolian elements

Sediments of aeolian origin are a widespread component of the Murrumbidgee landscape and include sand dunes and more clayey deposits. Evidence of aeolian deposition in the Riverina was noted by Butler (1950) through the observation of a uniform sheet covering the region with little variation in character across different conditions, naming it 'parna' (Butler, 1956). Butler (1956) and van Dijk (1958) identified the observed surface sheet to be the most recent parna deposit with more 'parna sheets' resulting from different phases of aeolian deposition buried below the surface. Previously, parna had been discussed as 'loess', an aeolian sediment reported around the world, as an influence on southeastern Australia (Crocker, 1946; Hills, 1939). Butler (1956) noted, however, the characteristics of this material to be different from loess with respect to its particle size grading and being more clayey, occurring as earthy, homogenous, calcareous, uniform sheets (Butler, 1956). The material is accepted to have been transported as silt and fine sand-sized pellets of calcareous clay (Cattle & Smith, 2018). As parna occurs through the deposition of clay aggregates eroded from previously weathered materials and developed soil, Butler (1956) termed these as 'old' soils.

The key property of soils derived from parna is subplasticity, whereby the texture grade becomes more clayey following the mechanical working of soil (Cattle & Smith, 2018). After

first being termed by Butler (1956), there was a focus on researching subplasticity, with a special edition of the Australian Journal of Soil Research published specifically examining the soil property (Blackmore, 1976; Brewer & Blackmore, 1976; Butler, 1976; McIntyre, 1976; Norrish & Tiller, 1976; Walker & Hutka, 1976). Van Dijk (1958) presented three degrees of subplasticity determined by the change in texture following a period of working. These degrees of subplasticity underpin subgroupings of soil types within the 'hillslope' class discussed later in this review.

In creating their geomorphic map, Butler et al. (1973) noted difficulties in mapping the distribution of parna due to its lack of independent geomorphic form. When parna is deposited, the reworking of materials and different weathering regimes resulting from landscape position mean that not all parna derived soils exhibit the same characteristics (Greene et al., 2009). Despite this, there is a developed understanding of the distribution of various parna phases at a regional (Figure 2.5*)* and more local (van Dijk, 1958) level. The pioneering work of Butler (1956) and Butler and Hutton (1956) described a 'parnaless zone' approximately 120 km in width surrounding Hay in the west of the study area (Figure 2.5).

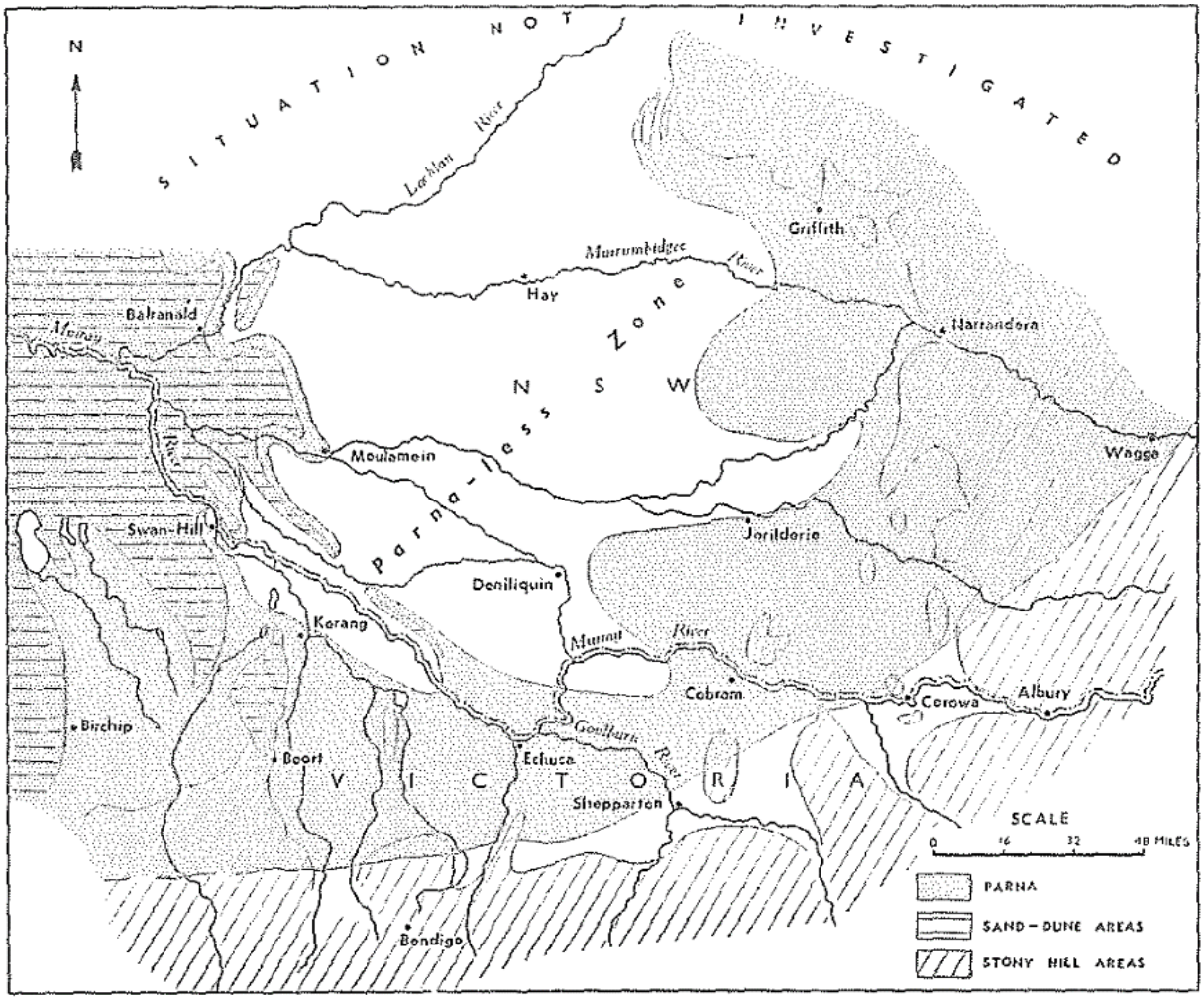


Figure 2.5. The theorised distribution of parna, including the parneless zone, across the lower Riverine Plain including the lower Murrumbidgee valley as developed by Butler (1956, Figure 1).

This zone occurs in an area less researched than those in the east of the region. Based on the conceptual model for the pathways of aeolian dust presented by Cattle et al. (2009) (Figure 2.6) it is likely that even within the parnaless zone, localised aeolian deposition has occurred although is not clearly discernible as a 'sheet' as is the case in the east. The sediments of this zone will therefore contain parna as a secondary material associated with the deposits from palaeochannels rather than as a distinguished stratigraphic unit. As a result, over the plain more broadly it is likely that there is parna mixed with alluvial sediments (Greene et al., 2009**)**. It is also likely that parna deposits in the upper catchment were eroded and redeposited across the plain during periods of palaeochannel activity (Greene et al., 2009) (Figure 2.6). This has already been discussed as a reason for changes in palaeochannel morphology, as postulated by Fried (1993).

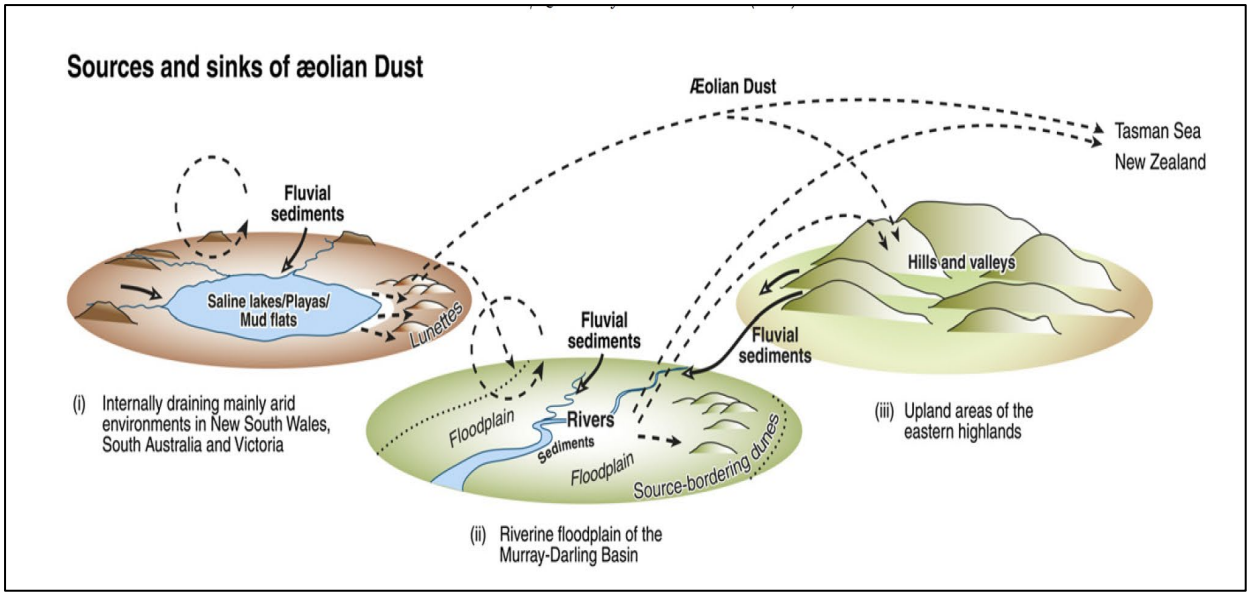


Figure 2.6. A conceptual diagram developed by Cattle et al. (2009, Figure 1) highlighting the pathways of aeolian dust within south-eastern Australia. Within this schematic, it is possible for the lower Murrumbidgee to act as both a source and sink for aeolian dust.

The first parna sheet to be characterised was Widgelli parna, the upper and most extensive sheet across the region (Butler, 1956; Butler & Hutton, 1956; Butler, 1958). This sheet varies in thickness; Butler and Hutton (1956) estimate a thickness of 3–5 ft (0.9-1.5 m) east of the parnaless zone, while in the Griffith-Yenda region van Dijk (1958) estimates a thickness of 1–4 ft (0.3–1.2 m).

Butler and Hutton (1956) and Butler (1958) emphasised the significant variability observed within Widgelli parna soils. This is a result of varying drainage conditions, emphasising the importance of understanding the relationship between the surface and older, buried materials. Observed parna sheets are also better preserved on slopes than on the lower floodplain. Van Dijk (1958) differentiates the Widgelli parna into two sheets, the most recent Tabbita parna which covers the entire area to a depth of 1- 4 ft (0.3–1.2 m) and the largely buried Bingar parna. Buried at depths often over 4.5 metres is the Cocoparra parna layer. The development of soil profiles within buried layers suggests periods of high humidity following parna deposition, in contrast to the aridity occurring prior (van Dijk, 1958).

Alongside the deposition of parna, secondary aeolian actions have occurred across the area. Van Dijk and Talsma (1964) noted this through the observed modification of dunes bordering palaeochannel stream beds through blown-out depressions, lunettes and, on occasion, sand sheets extending across the landscape within the Coleambally region. Pucillo (2005) identified and described the activity of source-bordering dunes at three sites on different palaeochannel belt sequences within the Coleambally area. As well as associating sediment ages with periods

of palaeochannel activity, a model for dune development is presented (Figure 2.7). This theorises that sediments are replenished in channels during fluvial periods and are windswept across banks as water subsides (Pucillo, 2005). The presence of vegetation stabilised these sediments, allowing dunes to develop, before they become remobilised as vegetation decreased during periods of aridity, such as the Last Glacial Maximum (Pucillo, 2005).

Page et al. (2001) also examined source bordering dunes near Wagga Wagga, approximately 100 km east of the study area, denoting three distinct stratigraphic units. Thermoluminescence dating of these sediments correlates well to periods of palaeochannel activity identified by Page et al. (1996) on the lower floodplain. The identified Clarendon Unit (15-25 ka) aligns with activity of the Yanco system, the Glenfield Unit (35–60 ka) aligns with the Kerarbury system and the Yarrangundry Unit (80-120 ka) with the Coleambally system. While consisting primarily of locally derived sands, calcareous clay fractions emanating from western areas were also observed (Page et al., 2001).

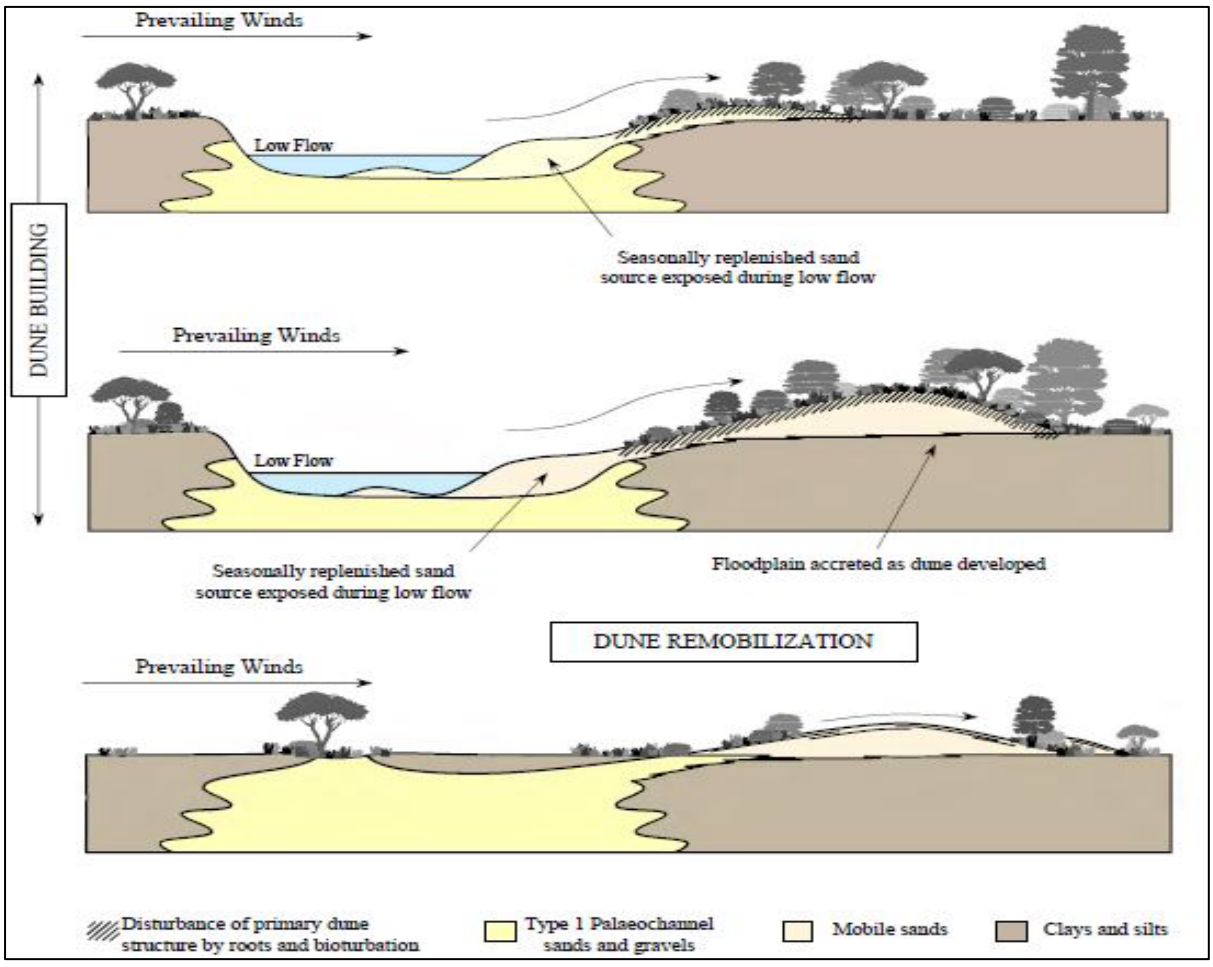


Figure 2.7. A diagram series from Pucillo (2005, figure 7.15) illustrating the evolution of source-bordering dunes on the Coleambally and Kerarbury Palaeochannel systems. Seasonal replenishment of the source sands occurred during flood discharges and dune remobilisation resulted from a reduction in stabilising vegetation throughout the Last Glacial Maximum.

2.3.3. Sediment stratigraphy

Following the theorising of prior streams (Butler, 1950) and parna deposition (Butler, 1956) new models of sedimentary deposition were required to explain observed soil patterns. These landscape representations differentiated sedimentary units within what would later be termed the broader 'Coonambidgal' and 'Shepparton' geological units by Brown and Stephenson (1991). Butler (1958) presented the first stratigraphic model for the region consisting of five depositional units to have influenced the upper Riverine Plain through fluvial and aeolian action (Figure 2.8). Of the five depositional units, determined from soil associations, there are three fluvial units; Coonambidgal, Mayrung and Quiamong; two aeolian units, Colongulac and Widgelli; and a mixed fluvial-aeolian unit, Katandra (Butler, 1958). The Widgelli parna unit, a layer of which mantles the plain, was proposed by Butler (1950; 1958) to indicate periods of aridity, acting as a stratigraphic marker. This schematic highlights the complex interactions of differently aged sediments (Figure 2.8). A layer of Widgelli Parna initially mantled previous Katandra and Quiamong units, however, at points it has been degraded by a Mayrung 'prior stream' (Figure 2.8). This stream has then aggraded, with a soils resulting from the two depositional processes meeting when the outer bounds of this stream's sediments meet the uneroded Widgelli Parna (Figure 2.8).

Accompanying the stratigraphic model (Figure 2.8) is a representation of sediment ages over time with respect to changes in the climate (Figure 2.9). Using Figure 2.8 as a companion to Figure 2.9 allows the visualisation of the complex layering of differently aged sediments where oscillating climatic conditions were a driver of changes between soil development and sediment deposition. The model (Figure 2.9) was developed following pedological examinations of sediments from each stratigraphic unit (Figure 2.8). Where little or no soil development was observed there was believed to be little time between depositional phases (Butler, 1958). For example, minimal soil development was observed in the Quiamong phase, suggesting it was followed by an arid climate and covered by deposits of the Widgelli parna unit (Butler, 1958). Conversely, substantial soil development within the Mayrung unit is considered to result from longer periods of more humid conditions before this was mantled by the Coonambidgal unit (Butler, 1958).

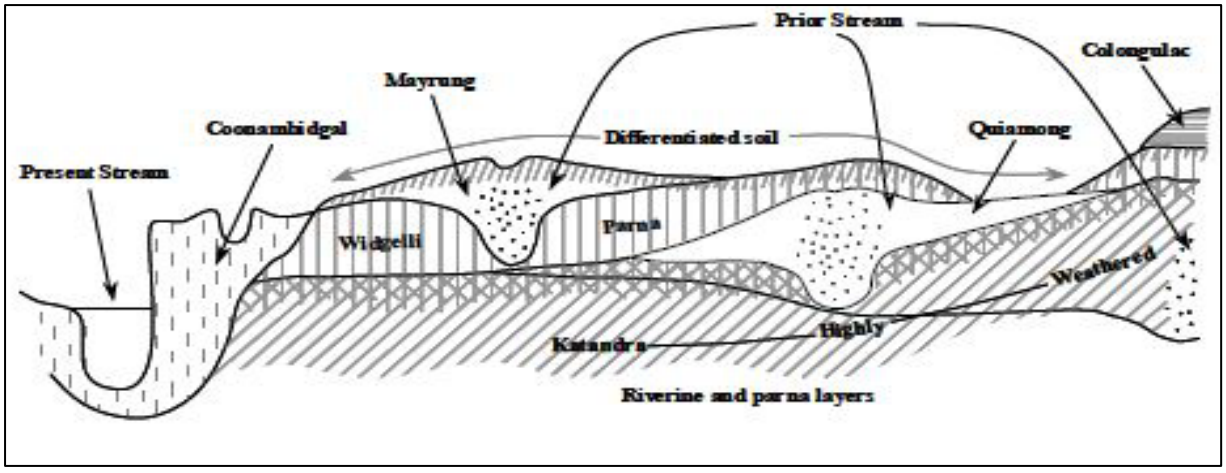


Figure 2.8. The stratigraphic model of Butler (1958, Figure 2) representing the interactions of different sediment groups across the landscape.

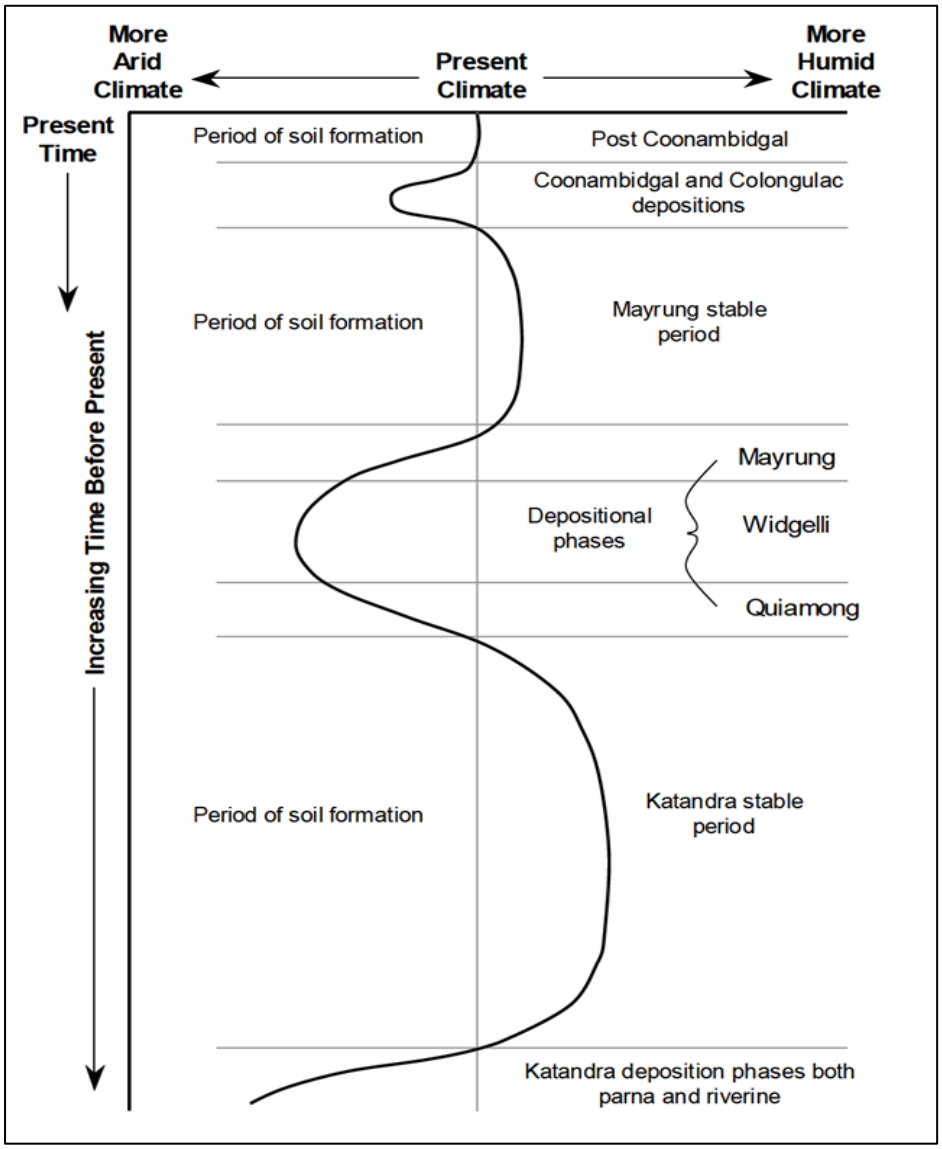


Figure 2.9. A diagrammatic representation of the depositional systems of Butler (1958) in association with climatic conditions and soil forming intervals made to accompany the stratigraphic model (Figure *2.8*) (from Pucillo, 2005, Figure 3.8).

In applying the work of Butler to the neighbouring Murray Valley, Pels (1966, 1971) developed a chronological model for three distinct periods of river activity, each comprising of aggradation and degradation (Figure 2.10). All periods of ancestral river activity began with the degradation of the landscape through river incision while a climatic shift to aridity resulted in sediment aggradation. Radiocarbon dating estimates were used to determine sediment ages. Pels (1971) estimated that the older Coonambidgal I channels incised themselves within the landscape prior to the deposition of sediments between 40,000 and 30,000 years ago. The Coonambidgal II channels incised between 30,000 and 26,000 years ago before sediments aggraded between 26,000 and 13,400 years ago, while the most recent ancestral river developed between 13,400 and 10,000 years ago, with sediments aggrading from 10,000 to 4,200 years ago (Figure 2.10).

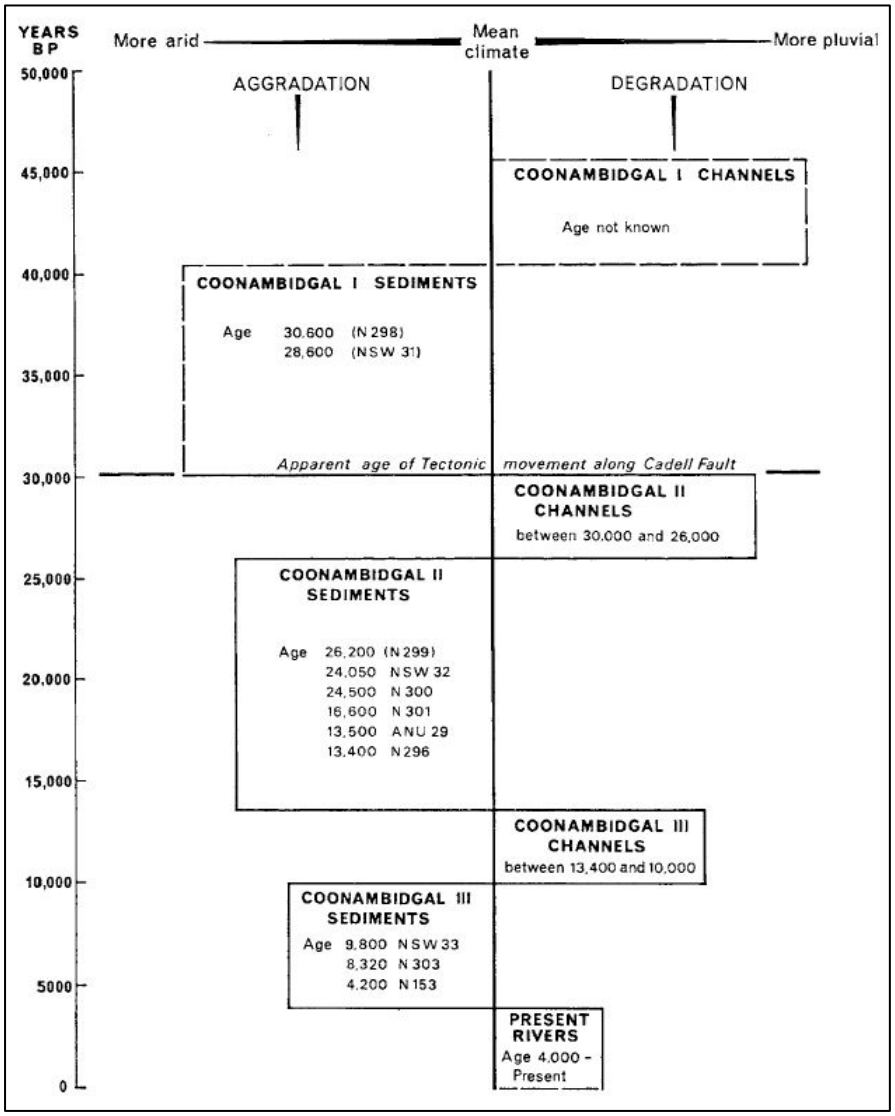


Figure 2.10. The chronological model of Pels (1971, Figure 4.4) showing the aggradation (deposition of sediments within streambeds) and degradation (incision of watercourses within the landscape) of Coonambidgal channels in association with climatic conditions dated based on radiocarbon estimates.

Schumm (1968) presented a new 'hybrid' stratigraphic model by synthesising previously used terminology and adapting the stratigraphic model developed by Butler (1958) (Figure 2.11). Schumm (1968) suggested that under humid conditions channels exhibited higher discharge rates and sinuosity, more like the current Murrumbidgee River, compared to arid climates where channels were wider, depositing sediments within riverbeds.

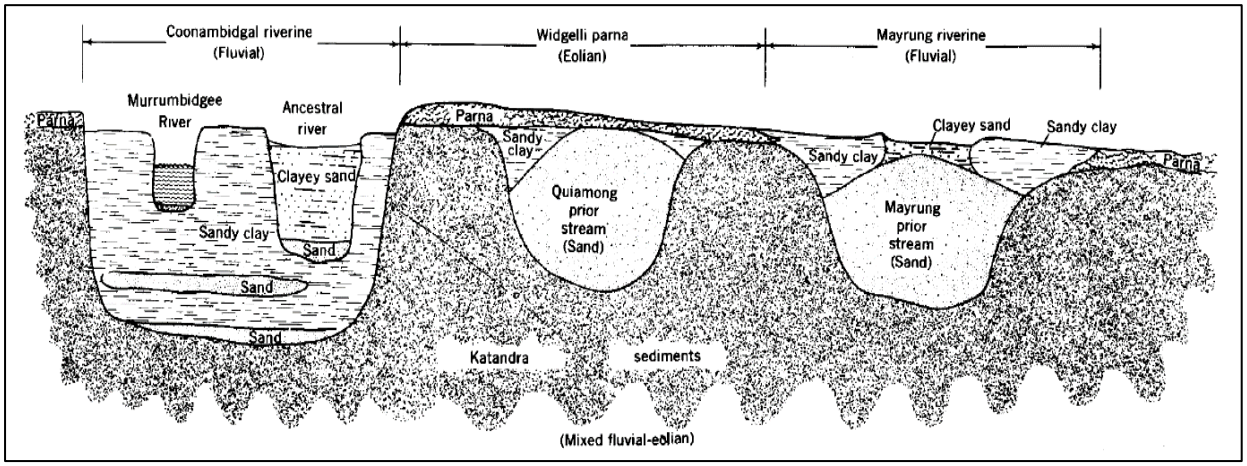


Figure 2.11. Surficial stratigraphic model of Schumm (1968, Figure 2) incorporating previous work of Butler (1958) and Pels (1964).

The thermoluminescence techniques employed by Page et al. (1996) and Page and Nanson (1996) allowed the development of a new, conceptual, model estimating the chronological development of morphologically different palaeochannel systems alongside aeolian features with typical sediment textures shown (Figure 2.12).

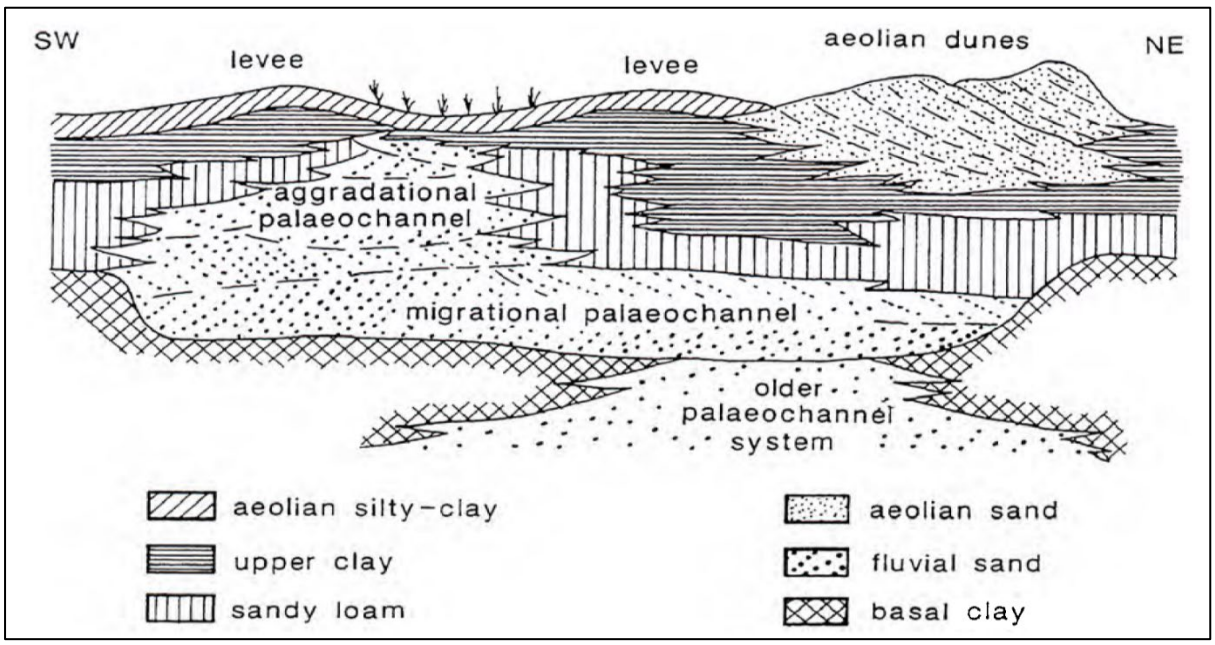


Figure 2.12. A revised stratigraphic model for palaeochannels within the lower Murrumbidgee by Page and Nanson (1996, Figure 9) showing the development of migrational and aggradational palaeochannel alongside surface features.

Analysis of extensive borehole datasets allowed Pucillo (2005) to provide estimates of sediment texture to a depth of 180 m (Figure 2.13). This encompassed the geologically distinct Calivil Formation and Renmark Group. Further, Pucillo (2005) presented a stratigraphic model to a depth of 35 meters, with four distinct palaeochannel sequences identified. The Type 1 unit incorporates palaeochannel sequences encompassed within already discussed stratigraphic models. The Type 2, 3 and 4 units, however, had not previously been differentiated from the broader Shepparton Formation. Previously, stratigraphic models were limited to the Coonambidgal and Upper Shepparton Formations as defined by Brown and Stephenson (1991) and Prathapar et al. (1997).

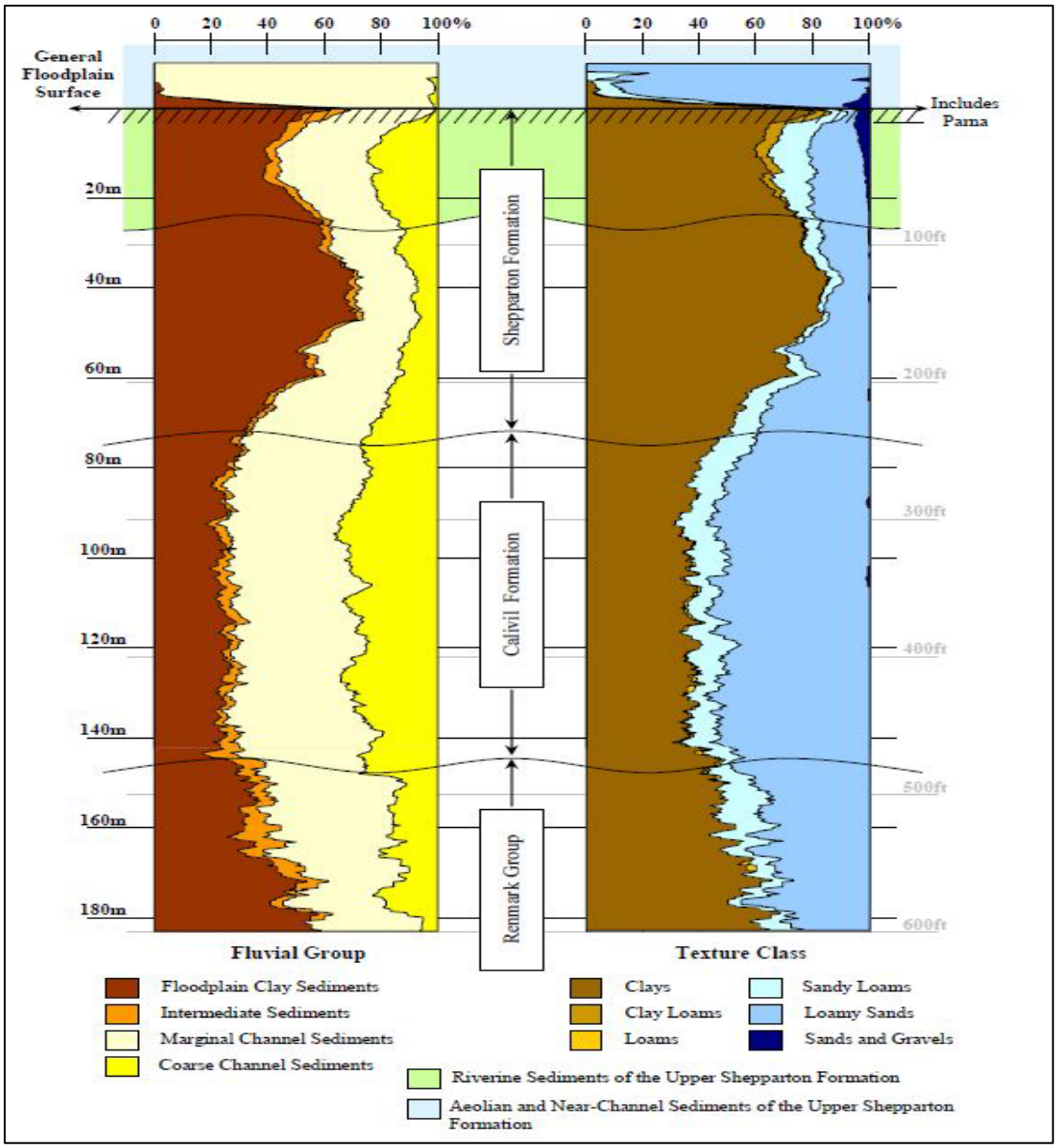


Figure 2.13. Sediment depth fractions from analysis of the entire borehole dataset within the Coleambally region for texture class and fluvial group classification (Pucillo, 2005, Figure 5.4). Also included is the geological group at each depth.

Many of the sedimentary or stratigraphic units are given localised geographic names relating to where the specific studies occurred. Consequently, a comparison of these units reveals conjecture between nomenclature while also identifying similarities in the ages of separately identified geological or stratigraphic units. (Table 2.2). For example, Brown and Stephenson (1991), in their geological study of the landscape, categorise Coonambidgal I sediments (Pels, 1971) within the Shepparton (geological) formation while the Coonambidgal II and III (Pels, 1971) are akin to the Coonambidgal of Butler (1958) (Pucillo, 2005).

Table 2.2. A synthesis of geological and stratigraphic units, listed under the identifying author, estimating the age of when different systems were active in depositing sediments over the last 135 ka years. This highlights that despite differing nomenclature between units identified in different sub-regions there are strong similarities in the ages that each system was active.

Age (ka)	Geological unit	Stratigraphic unit			
	Brown & Stephenson (1991)	*Page et al. (1996)*	*Pucillo (2005)*	*Pels (1971)*	*Butler (1958)*
0	Coonambidgal	Modern (Holocene) River systems			
5				Coonambidgal III	Coonambidgal
10		Yanco			
15				Coonambidgal II	
20	***Debated***				Mayrung
25		Gum creek			Widgelli
30					Quiamong
35		Kerarbury	Type 1		
40				Coonambidgal I	
45					
50					
55					
60					
65					
70					
75					
80		Coleambally			
85					
90					
95					
100	Shepparton				Unspecified Katandra sediments
105					
110					
115					
120					
125					
130					
135					
…			Type 2*		
…					
			Type 3*		
…					
…			Type 4*		
…					
…					

2.4. Soil distribution in the lower Murrumbidgee valley:

The development of sustainable agricultural industries relies upon accurate understandings of soil condition and variability. Historical, targeted, institutional soil surveys were essential in the development of the MIA and CIA between the 1930s and 1970s. Considering these survey outputs alongside geomorphological studies allows for features such as sub-plasticity as well as patterns of salinity and waterlogging susceptibility to be understood. Across the last century, significant secondary research examining soil in relation to agronomic practices has also occurred. More recently there has been an increase in 'grey' data resulting from privately undertaken soils research. This has enhanced stakeholders' anecdotal understanding of soil variability at the farm scale. Despite the extent of this work over an extended period of time there are significant issues with data availability and temporal changes to the soil.

2.4.1. Soil surveys

The establishment of the MIA and CIA in the 1910s and 1960s, respectively, necessitated the need for accurate soil data. Within these subregions, maps resulting from soil surveys dictated the establishment of specific production systems in accordance with the identified soil type. Presently across the lower Murrumbidgee valley, only these two irrigation areas have been publicly surveyed since the expansion of agriculture in the region over a century ago, with 94 soil types identified, described, and mapped across multiple surveys and reclassifications. These surveyed areas, however, equate to only approximately 5.6%, or, 1450 km^2 of the 25,771km^2 area described in this review.

2.4.1.1. Murrumbidgee Irrigation Area (MIA)

Taylor and Hooper (1938) published the first soil survey within the region, comprising 13,439 ha of land focused on horticultural soils in the areas surrounding Leeton and Griffith. Fifty-five identified and described soil types were mapped, with observations provided on soil conditions and the likely success of horticultural production on different soil types. At this time, alluvial plains were considered a result of modern river systems, resulting in difficulties accurately describing soil property distributions and patterns of salinity. Taylor and Hooper (1938) also noted that some soils exhibited a uniqueness in behaviour, terming 'light and heavy' clays soils which would later be redefined as subplastic.

Advancements in understandings of the region's geomorphology, coupled with demand for soils knowledge, necessitated that Butler (1979) release a revised edition of the survey. Soil

type descriptions were reframed within the context of prior streams (referred to as palaeochannels in this review) and parna. By considering the many soil types as soil sequences resulting from now understood geomorphic features, the complexity arising from the large number of soil units was reduced (Butler, 1979). The distribution of soil properties, such as salinity, was also explained. Following the updating or removal of outdated sections, soil type descriptions and maps were presented largely unchanged, with slight adjustments to the composition of soil groups (Table 2.3). The report also included soil series across the landscape and a key to named soil types with particular characteristics.

Table 2.3. Soil type groupings from the revised survey edition of Butler (1979, Table 4) where superscripts; [A] denotes a deep subsoil going heavy and grey and [B] denotes a deep subsoil becoming sandy.

Group	Soil types
Subplastic soils of the hillslopes	Ballingall loam, Lakeview loam, Merungle loam, Tharbogang loam, Wyangan loam, Types 1, 2, 3, 4, 5, 6, 7, 10, 11, 15, 16 and 17.
Subplastic soils of the lower slopes and plains	Bilbul[A] loam and clay loam, Griffith[A] loam and clay loam, Hanwood[B] sandy loam and loam, Jondaryan[B] loam and clay loam, Stanbridge[B] sandy loam and loam, Yenda[A] sandy loam and loam and Type 9.
Plastic soils of the plains	(a) Lighter subsoil sub-group Fivebough[B] sandy loam, Mirrool loam, Willimbong loam, Yoogali loam. (b) Normal sub-group Beelangera clay loam, Camarooka sandy loam, Leeton clay loam, Types 8, 12 and 13.
Sandy soils	Banna sand, Hyandra sandy loam, Tenningerie sand and sandy loam, Wamoon sand and sandy loam, Yambil sandy loam, Yandera sandy loam and loam, Types 7, 10 and 14
Mallee soils	A, B, C

Targeted surveying work was also undertaken by van Dijk (1958;1961) in the Griffith – Yenda region focusing on parna (1958) and the southern portion of the MIA (1961). In the latter work, van Dijk (1961) identified 21 soil types which were categorised within 14 soil associations, where each association has one or two dominant soil types as well as subdominant and minor

soil types, where relevant (Table 2.4). The categorisation of each association was informed by topography and linked to one of six soil landscape units. These units were mapped over an area of approximately 100,000 acres (40,500 ha). Comments on soil properties and suitability for agricultural systems were also made.

As an example, the Thulabin soil association, one of six associations within the Whitton clay plain landscape unit, is included below (Table 2.4).

Table 2.4. The Thulabin soil association, tabulated, adapted from van Dijk (1961, Table 4). This is an example of one of the 14 identified soil associations.

Soil association	Type location	Dominant soil types and their topography		Subdominant soil types and their topography		Minor soil types and their topography	
Thulabin	On all plains	Thulabin clay loam Thulabin loam Thulabin sandy loam	Well-drained, frequently slightly domed plains	Thulabin sand	Slight rises and domed plains	Tenningerie sand Sandmount sand	Small low dunes
						Tuppal clay loam and Wunnamura Clay	Swamp depression
						Birganbigal clay loam	Level, well-drained plain section

2.4.1.2. Coleambally Irrigation Area (CIA)

Soils within the CIA were first examined by van Dijk and Talsma (1964) prior to the development of the irrigation scheme. Their work resulted in two maps being produced; a detailed reconnaissance map of part of the Coleambally area and a soil landscape map of the Coleambally area (Figure 2.14). In the former, more detailed study, auger holes were examined every 400-800 m along traverses spread between 1.2-1.6 km apart. Surface inspections and examinations of aerial photographs were used to validate areas between traverses and assisted in developing conclusive boundaries between soil units. The second, broader study utilised the same methodologies with traverses at 8 km intervals. The methods of categorisation into soil landscape and association units are akin to those of van Dijk (1961). All soil types identified had already been published and were categorised within greater soil groups, with basic descriptions of the surface, subsoil and topography provided.

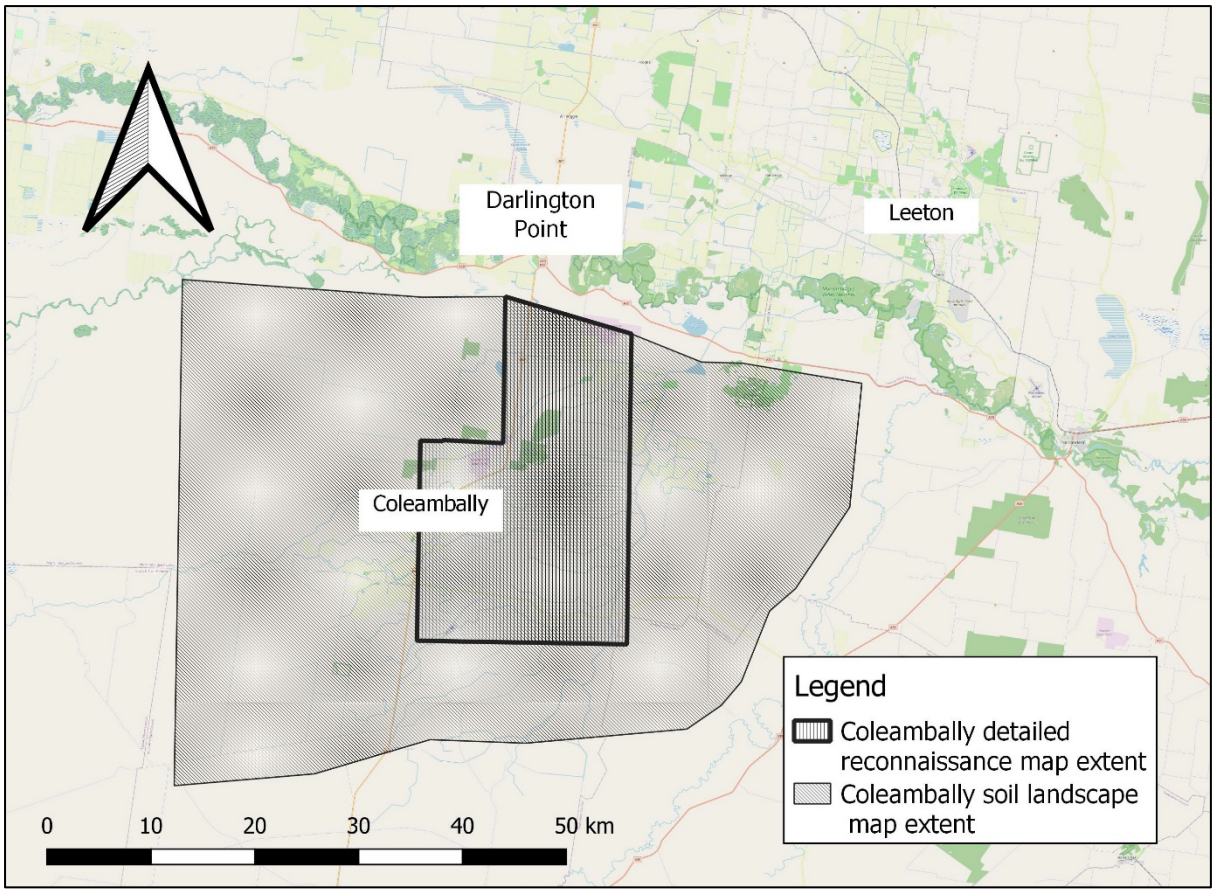


Figure 2.14. The approximate bounds of the detailed reconnaissance and soil landscape surveys of the Coleambally area undertaken by van Dijk and Talsma (1964)

Utilising landscape characteristics, recurring sequences of soil patterns were used to develop ten soil associations (Table 2.5), where each association is defined by one or two dominant soil types alongside subdominant and minor types. For example, the Biranbigil association is

located on flat, well drained plains with dominant and subdominant soil types of the Birganbigal clay loam and Wilbriggie clay loam, respectively (Table 2.5). These associations, which are generalisations of the detailed study of soil type distribution, were mapped over 150,000 acres (61,000 ha) to produce the detailed reconnaissance map.

The broader soil landscape map utilised the ten soil associations as well as geomorphic and drainage conditions, factors significant in soil development, to identify 16 soil landscapes (van Dijk & Talsma, 1964). Each landscape contains a dominant soil association with potentially multiple subdominant associations (Table 2.6). The Carabury plain, for example, is one of four landscapes associated with the gilgaied clay plains of the region. The dominant and subdominant soil associations for this landscape unit are the Yooroobla and Wilbriggie soil associations, respectively, and the Great Soil Group (Stephens, 1953) is ‘Grey and brown soils of heavy texture’, equating to Vertosols in the Australian Soil Classification (ASC). This map covers an additional 250,000 acres (101,000 ha) with the distribution of each association closely tied to geomorphological features.

Table 2.5. The 10 soil associations identified in the Coleambally detailed reconnaissance survey adapted from van Dijk and Talsma (1964, Table 4). In some instances, associations contained up to six subdominant or minor soil types. To simplify, only the first two listed subdominant or minor types are shown.

Soil association	Physiography	Soil types		
		Dominant	*Subdominant*	*Minor*
Birganbigal	Flat, well drained plains	Birganbigil clay loam	Willbriggie clay loam	Thulabin loam and clay loam
				Mundiwa clay loam
Cobram	Narrow belts with curves and slightly winding elongated sandy rises on river ridge tracts	Cobram sandy loam	Finley loam	Tenningerie sand
			Moira loam	Sandmount sand
Danberry	Low, flat sections and low slopes on the flanks of the river ridge tracts	Danberry loamy sand	Danberry sand	Sandmount sand
			Danberry sandy loam	Tenningerie sand
Sandmount	High broken relief on river ridge tracts	Sandmount sand	Wamoon sand	Thulabin sandy loam
			Thulabin sand	Danberry sand
Tenningerie	Low broken relief on river ridge tracts	Tenningerie sand	Sandmount sand	Danberry clay loam
			Wammon sand	Tuppal clay loam
Thulabin	Well-drained slightly undulating sections and flanks of river ridge tracts; occasional domes and lunettes on the plains	Thulabin loam	Thulabin sandy loam	Tenningerie sand
			Thulabin clay loam	Tuppal clay loam
Tuppal	Swampy depressions	Tuppal clay loam	Wandook clay	Gogeldrie clay
			Wunnamurra clay	
Willbriggie	Moderately well-drained plains with slightly lobate surface relief, swampy depressions and occasional domes	Willbriggie clay loam	Yooroobla clay	Morago clay loam
			Coree clay loam	Wunnamurra clay
Wunnamurra	Poorly drained plains subject to frequent flooding	Wunnamurra clay	Yooroobla clay	Mundiwa clay
			Coree clay loam	Tuppal clay loam
Yooroobla	Moderately well to poorly drained plains subject to flooding	Yooroobla clay	Mundiwa clay loam	Coree clay loam
			Morago clay loam	Tuppal clay loam

Table 2.6. The 16 soil landscapes identified in the broader Coleambally landscape survey. Included are the physiographic unit, soil landscape, dominant and subdominant associations (Table *2.5*) and generalised Great Soil Group (van Dijk & Talsma, 1964, Table 6)

Physiography	**Soil Landscape**	**Soil Associations** ***Dominant (italics)*** **and subdominant**	**Characteristic Great Soil Group (Stephens, 1953)**
Continuous river ridge	Waddi river ridge Yamma river ridge	*Cobram,* Sandmount, Tenningerie, Thulabin	Deep sands Red-brown earths Solodized solonetz
Broken river ridges	Cararbury river ridge Tubbo river ridge Boona river ridge	*Sandmount,* Tenningerie, Danberry, Thulabin	
	Banandra river ridge Gidgell river ridge	*Danberry, Thulabin*	
Apex plain	Ourendumbee plain	*Birganbigil*	Red-brown earths
Gilgaied clay plains	Jurambula plain	*Yooroobla,* Wunnamurra	Grey and brown soils of heavy texture
	Cararbury plain	*Yooroobla,* Willbriggie	
	Gidgell plain	*Yooroobla,* Willbriggie, Wunnamurra	
	Mourndah plain	*Wunnamurra,* Yooroobla	
Brown plains	Boobalbundi plain Argoon plain	*Willbriggie*	Transitional red-brown earths
	Goolgumbla plain Yamma plain	*Willbriggie,* Willbriggie light subsoil phase	

Accompanying descriptions on soil distribution are comments on land suitability relating to specific crops including citrus, pasture species, cereals and rice. Each individual soil type is placed within one of five landuse classes:

- C: Suited to citrus and most horticultural crops,
- Y: Restricted horticultural use due to poor drainage, successfully used for lucerne, pastures and cereals,
- P: Well suited for quality pasture with potential for cropping in rotations,
- G: Dense slaking surfaces provide suitability for rice growing, and,
- O: Commonly comprised of gilgai soils provide a lasting favourable tilth well suited to pasture (van Dijk & Talsma, 1964).

Similar to the soil landscapes (van Dijk & Talsma, 1964), Stannard (1968) described and mapped (unpublished) seven physiographic units resulting from micro-topographical differences in the landscape. While van Dijk and Talsma (1964) utilised only previously identified soil types in their survey, Stannard (1970) observed differences sufficient to create 14 new soil types ranging from self-mulching clays to deep sands. In the accompanying handbook, Stannard (1970) provided morphological descriptions of all soil types with an accompanying map issued to landholders, however, this was not published.

2.4.1.3. Murray Valley

The Murray valley, located adjacent to the Murrumbidgee region within the Riverine Plain, exhibits similar landscape characteristics and consequently shares similar soil types. Soil surveys have occurred more widely in the Murray valley, focusing primarily on land suitability and classification for irrigation during periods of land development following the Second World War. A number of these surveys were relied upon in understanding the soil types surrounding Coleambally. Surveys of note include:

- Soils of the Berriquin Irrigation District (Smith, 1945),
- The soil and land-use survey of the Wakool Irrigation District (Smith et al., 1943),
- Soils of the Deniboota Irrigation District and their classification for irrigation (Johnston, 1952),
- Jernargo extension of the Berriquin Irrigation District, NSW (Churchward & Flint, 1956),
- The soils of the East Murrakool district, New South Wales, and their relation to land use under irrigation (Churchwood, 1956), and,
- The soils and land use of the Denimein Irrigation District, New South Wales (Churchward, 1958).

2.4.2. Reclassifications of soil types

The diversity and variability of the lower Murrumbidgee valley's soils creates difficulties in effectively summarising and communicating research to stakeholders, especially when differently named soil or landscape series are presented for each surveying exercise. The generalisation of soil groups and their spatial occurrence is reliant on landscape position, highlighting the importance of understanding the region's past geomorphological history. Hornbuckle and Christen (1999) addressed this in their report on soil physical properties believing that, based on recurrent patterns of association and morphological similarities, the 94 identified soil types fitted within six groups: Self-mulching clays, Hard setting clays, Red Brown Earths, Transitional red brown earths, Sands over clays and Deep sands, with subgroups where necessary (Table 2.7). This work was expanded to include the Murray Valley and republished as a three-report series of: (a) Accessing soil map data via Google Earth software (Hornbuckle et al., 2008a), (b) Physical soil properties (Hornbuckle et al., 2008b) and (c) Chemical soil properties (Thacker et al., 2008).

These reports are effective in simplifying, synthesising and effectively disseminating data from difficult-to-access research. Data was accessed through searches of unpublished databases containing 691 references to soil studies within the region, alongside systemic key word searches of public sources (Hornbuckle et al., 2008b). Data was reviewed within each soil group (Table 2.7), with soil properties presented for four major characteristics: soil physical composition, water movement, water retention/moisture characteristics and soil salinity/sodicity (Hornbuckle & Christen, 1999).

Table 2.7. General features of the six soil groups first presented by Hornbuckle and Christen (1999, Table 3-1). The ASC classifications presented may no longer be an accurate representation of these groups, for example, an additional soil order means that the Deep sands would now likely be classed as Arenosols.

	Soil group					
Feature	Self-mulching clays	Hard-setting clays	Red-brown earths	Transitional red-brown earths	Sands over Clay	Deep sands
Topsoil	Clay	Loam-clay	Loam-sandy loam	Loam-clay loam	Sand-loam, cemented	Sand
Depth (m)	0.05-0.15	≤ 0.05	0.1-0.25	≤0.1	0.25-1	>1
Subsoil	Heavy clay, lime	Heavy clay	Heavy clay	Heavy clay	Mottled medium clay	Mottled clay sand – light clay
Deep subsoil	Medium clay, concretionary lime	Medium clay, crystalline gypsum common	Sandy clay, often micaceous	Medium clay, crystalline gypsum common	Medium clay, can be sandy at depth	Light clay
ASC classification	Vertosol	Sodosol	Chromosol	Chromosol	Chromosol	Rudosol

2.4.3. Secondary, production focused, soil studies

Historical soil studies have the capacity to provide a baseline resource for understanding changes to soil characteristics following the development of irrigation systems. Government research centres established in Griffith, Yanco and Deniliquin facilitated substantial soils focused work in the region. Much of this work examined specific production systems for different crops, with comments on the soil condition or basic soil data noted. Lang and Hicks (1975) presented a bibliography of soils research to have occurred within the Riverina. This catalogued all prior work relating to soils within the region based on specific study areas. Data from many of these projects was incorporated within the work of Hornbuckle et al. (2008b) and Thacker et al. (2008) where it was used to summarise differences in soil properties between the six re-classified soil groups (Table 2.7).

A broader study undertaken by Talsma (1968) examined and mapped salinity from east of Coleambally to south of Maude, in the west of the region, utilising over 350 datapoints. In

averaging measurements at 2-4 ft (0.6–0.9 m) depths, a gradual trend of increasing salinity from east to west is shown (Figure 2.15). Analytical data showed a distinct difference in the salinity variability of heavy grey and brown soils in the west, with a marked salinity increase below a depth of 1 ft (0.3 m), compared to the heavy grey and brown soils in the east. Utilising a smaller dataset, Exchangeable Sodium Percentage (ESP) was generally low (<7%) in the east, however, increased to 7-22% at the surface and 15-30% in subsoils west of the K transect (Figure 2.15).

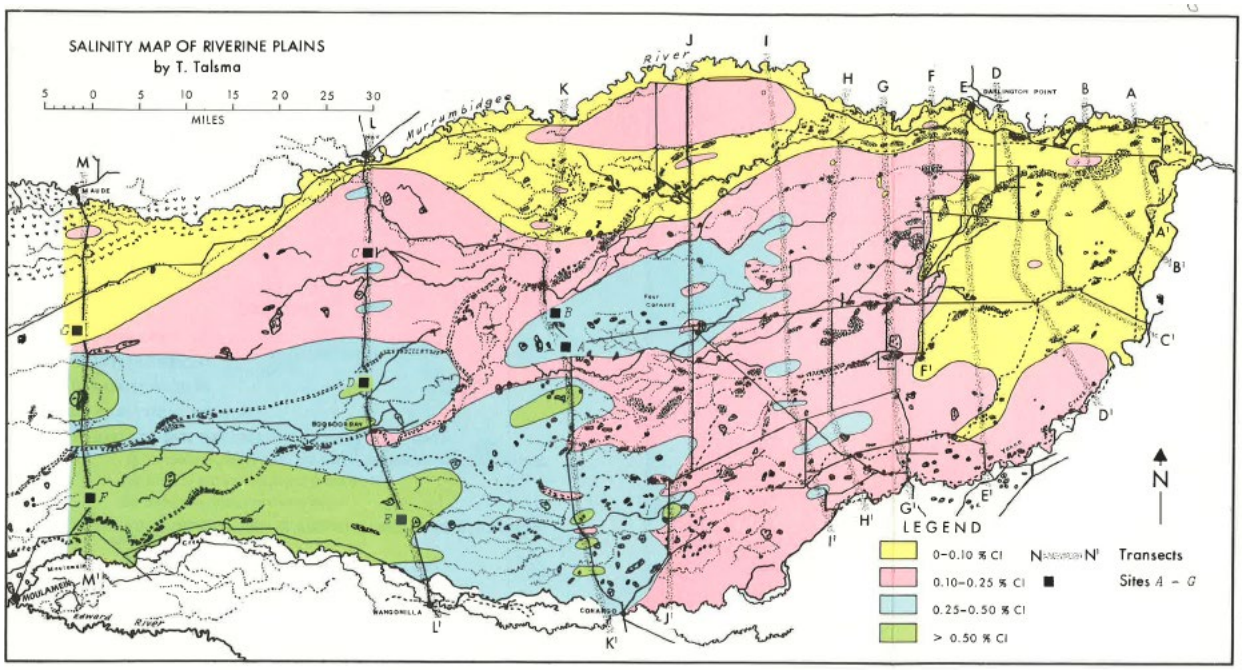


Figure 2.15. The salinity map produced by Talsma (1968, Figure 8) for the Riverine Plain south of the Murrumbidgee River between the Yanco Creek anabranch and the town of Maude.

2.4.4. Soil series in relation to palaeochannel systems

Patterns of soil property distribution are strongly correlated with geomorphologic units and features. In the first theory on prior streams, Butler (1950) theorised that channel beds contain coarse channel sediments with adjacent levees, or dunes, comprised of sandy clay loams and sandy clays. Soils become more clayey and contain more salts moving laterally away from these streams with salinity greatest in the downstream (westerly) reaches of the palaeochannel systems (Butler, 1950). This model was slightly modified by Butler (1958) and Langford-Smith (1960) (Figure 2.16). Both models showcased clear increases in the clayeyness of soil moving away from channel systems (Figure 2.16). Butler (1958) theorised that as clay content increases, salt content also increases while lime content decreases. The occurrence of occasional sand dunes, as further explained by Pucillo (2005) is also included in the schematic (Figure 2.16).

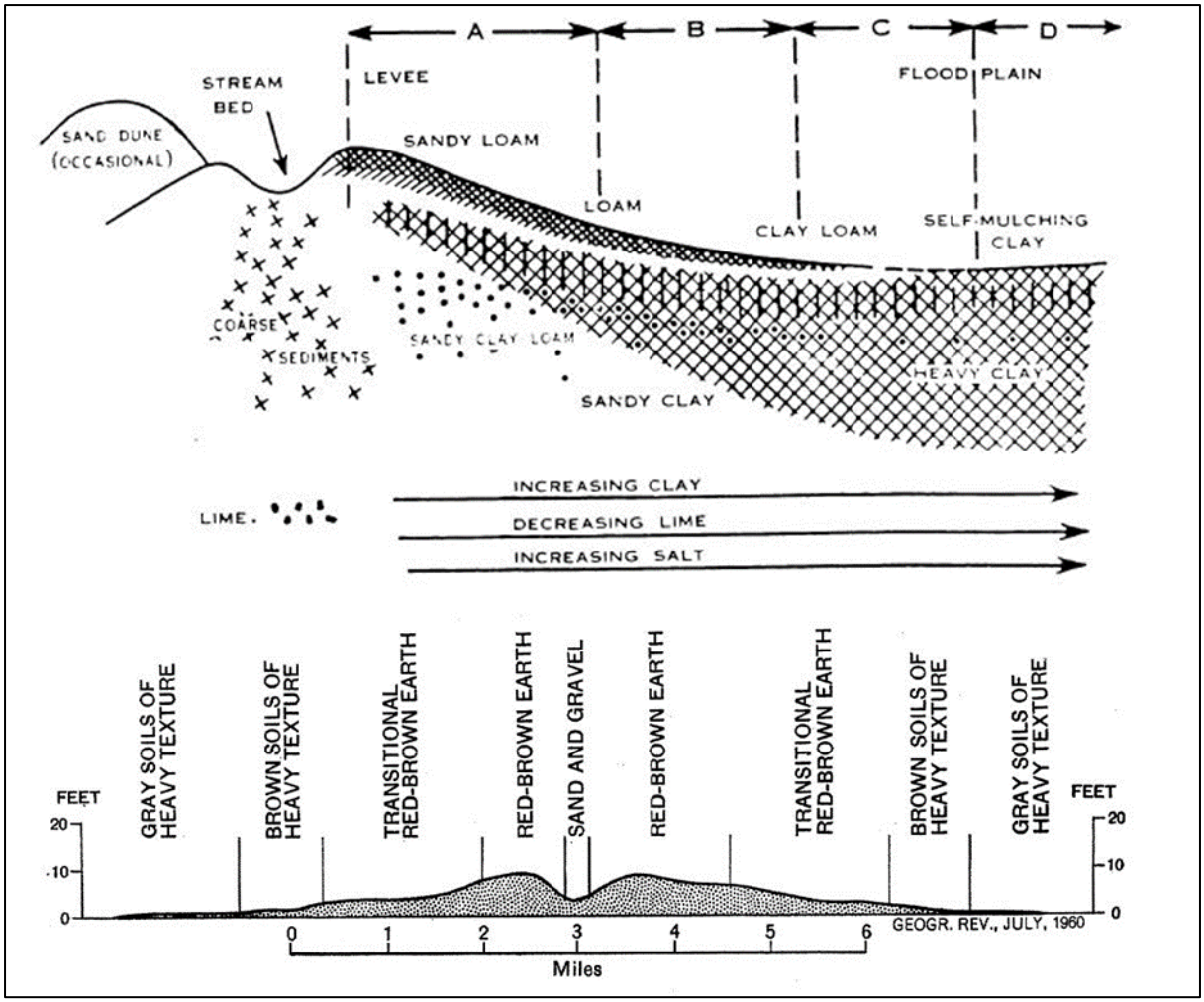


Figure 2.16. Two 'characteristic' soil sequences along a perpendicular traverse to palaeochannels presented by Butler (1958, Figure 5) (top) and Langford-Smith (1960, Figure 10) (bottom).

Murphy et al. (2000) aptly describes the west of the study area, between Hay and Balranald, as "a mosaic of Grey Vertosols and Red Chromosols" with the almost level plain traversed by prior streams. Murphy et al. (2000) classifed shallow streambeds as Hypocalcic Calcarosols with scalded margins and levees (Murphy et al., 2000). Where not subject to secondary actions, preserved source bordering dunes are also present. A recent change to the Australian Soil Classification (ASC) may result in these palaeochannel streambeds and dunes being classified as Arenosols instead of Calcarosols (Isbell & National Committee on Soil and Terrain [NCST], 2021).

In the area described by Murphy et al. (2000) there are also more complex relationships. The patterns of soil distribution are clearer in upstream areas or where palaeochannel systems with higher streamloads fully traversed the plain. Not described in literature is how the relationship becomes less clear on distributary, or terminal channels where streamflows decreased, and sinuosity increased, as these systems terminated. Such distributary channels can be observed

on the plain north of Hay and towards Balranald. Here, streambeds of coarse sediments are sharply incised within grey clays, with no graduated change in texture across a traverse. Levees, as described by Butler (1958) and Langford-Smith (1960), are less common in these instances.

2.4.5. Digital soil mapping products

This review has so far focused on traditional survey methods in association with studies of landscape formation. While effective, soil survey processes are both cost and time intensive. Kidd et al. (2020) briefly discussed the decline in soil surveys within Australia since the 1990s, due to decreased funding and availability of expertise, before presenting how DSM transitioned from a research topic to operationally filling this void. Over this time, improvements in computational capacity and the availability of spatial data have seen significant interest and advancements in the field of DSM globally (Minasny & McBratney, 2016). Where traditional soil surveys required a high level of local knowledge, DSM instead utilises statistical relationships between soil and environmental data to make predictions of soil properties at specific locations (Arrouays et al., 2014; Han et al., 2022). Ma et al. (2019) also presented how DSM can work to support traditional pedological methods to gain increased insights into soil development and processes.

In noting commonalities among recently created digital soil maps, McBratney et al. (2003) presented the 'scorpan' model as a framework for DSM. Where the longstanding equation of Jenny (1941) explained factors which account for the formation of soils, the scorpan approach is instead an empirical quantitative description of soil and spatial factors which can be used as predictors (McBratney et al., 2003).

$$S_c \text{ or } S_a = f(s, c, o, r, p, a, n) + e$$

The equation presents soil, as either a specific soil property (S_a) or a soil class (S_c), as a function of soil (s), climate (c), organisms (o), relief (r), parent material (p), age (a) and spatial position (n), with spatially correlated residuals also incorporated (McBratney et al., 2003). Since this time there have been a diverse range of reviews focusing on advancements and applications of DSM (Arrouays et al., 2014; Minasny & McBratney, 2016; Malone et al., 2017; Searle et al., 2021; Chen et al., 2022; Adeniyi et al., 2024).

One key benefit of DSMs is their accessibility. Whereas historical, analogue data is at risk of being lost, public digital resources are readily available and simple to interpret. Currently, there

are three widely utilised digital soil maps covering New South Wales, including the entirety of this study area. These are:

- The state NSW Digital Soil Map (Gray, 2023; NSW OEH, 2017),
- The national Soil and Landscape grid of Australia (SLGA) (Grundy et al., 2015; Viscarra Rossel et al., 2015), and,
- The global SoilGrids250m global soil map (Hengl et al., 2017a; Poggio et al., 2021).

Utilising the SLGA as an example, it presently contains 20 nationally mapped soil properties at depth intervals of; 0-0.05, 0.05-0.15, 0.15-0.3, 0.3-0.6, 0.6-1 and 1-2 m. These layers have the potential to be utilised as important covariates in secondary environmental (Andres et al., 2021; Rifai et al., 2024) and agricultural (Al-Shammari et al., 2021; He & Wang, 2019; Lawes et al., 2023) studies. Of relevance to this study area was a model of exchangeable sodium percentage (ESP) developed by Pozza et al. (2022) across the Murray Darling basin. This study utilised publicly available covariate data alongside soil point datasets to predict ESP to a depth of 1 metre, and to identify the depth where sodicity constraints of 6, 10 and 15% ESP were exceeded (Pozza et al., 2022). The produced models and maps show that at a depth of 0.2-0.3 m, 82% of this study area exceeds an ESP of 10% (Hazleton & Murphy, 2007) while at 0.9-1 m depth this increases to 99% of the study area (Pozza et al., 2022).

2.5. Current limitations to available soils information

Despite the breadth of work undertaken over the last century, there are knowledge gaps regarding the region's soils. A report by Holland and Eastwood (2014) addressed these in a scoping study evaluating the understanding of the region's soils in relation to cotton production. While this report focused on the cotton industry, it identified limitations in presently available soil information relevant to other agricultural sectors and general land management. In areas of land without accurate soil data, or those where an understanding of inherent soil properties is limited, it is difficult to know the locations most at risk of soil degradation and temporal change. As a result, information regarding both threats to the soil, through anthropogenic degradation, and opportunities, such as increased carbon sequestration, cannot be communicated.

2.5.1. Data and resource availability

Presently, a major limitation regarding information on the region's soils is the availability of geolocated soil data, which underpins the accuracy of DSM activities. There are also

shortcomings in stakeholders' ability to access legacy soil resources including soil surveys, landscape maps and secondary work focusing on soil through an agronomic lens. Soil surveys and landscape maps are publicly available for the MIA and CIA, however, beyond these areas such resources do not exist. This is noted by Holland and Eastwood (2014) with respect to land surrounding Hay, where there has been a significant expansion of the cotton industry. While the NSW state government generated a significant number of soil map sheets across the state from the 1960s to the early 2000s, they are not available for much of this study area.

This is not to say that the examining of specific soil properties has not occurred. In fact, the bibliography of soils research within the Riverina presented by Lang and Hicks (1975) identifies extensive works to have occurred over the prior 40 years. Much of this work, however, was never digitised. Consequently, it is not available through searches of public databases and risks being lost. Some of these private, or unpublished, databases were included in the report of Murrumbidgee soil types presented by Hornbuckle and Christen (1999). Despite this, for many of the 94 individual soil types identified within the MIA and CIA, there was limited or no soil site data available. Where data is available, difficulties may arise geolocating sample sites, undermining its value.

Over recent years there has been an expansion in companies offering privately contracted soils 'research'. This has seen an increase in private soil surveys as opposed to the institutional, public surveys that have been discussed earlier. In many instances, this work focuses on soil through an agronomic lens, primarily concerned with plant nutrients and the topsoil. While this is an important aspect of the farming system, it neglects fundamental soil properties, such as the mineral suite and texture, which are important determinants of susceptibility to degradation. While these works ultimately have the potential to build stakeholder capacity and inform regional understandings of soil properties and their distribution, the findings are not publicly available.

2.5.2. Temporal variability

Changes to the soil over time also have the potential to negatively impact the usefulness of available data and resources. While soil change is a natural process, it can be exacerbated by anthropogenic activities related to farming systems. Consequently, where significant changes to the soil have occurred the value of historical resources, including soil surveys, can be diminished unless data has been appropriately catalogued and can be used to study temporal change.

The most significant anthropogenic changes have resulted from land reformation. To allow for appropriate water flow, land under broadacre flood/furrow irrigated cropping systems commonly requires a gradient of 0.05-0.5% (Esfandiari & Maheshwari, 2001; Devkota et al., 2021). To facilitate this, areas of higher elevation are 'cut', with the material placed in lower elevation, 'fill' areas. Consequently, there is the potential for upper soil horizons to be completely removed in 'cut' regions with lower B horizons now at the surface. Similarly, in 'fill' areas, the previous topsoil can now become the subsoil. While this does not classify the soil as an anthroposol, per the ASC (Isbell & NCST, 2021), it significantly alters the soil condition. Cay and Cattle (2005) presented a study examining how this process impacts soil properties in cotton production systems in the neighbouring Lachlan Valley. This work noted that cotton growers had observed yields in 'cut' areas to be up to 5 bales/ha lower than in 'fill' areas (Cay & Cattle, 2005). Of the soil profiles examined, the oldest, most pedologically differentiated soil exhibited the greatest change between native, 'cut' and 'filled' profiles (Cay & Cattle, 2005). Jessop et al. (1985) identified that, on texture contrast soils in the Macquarie Valley, there were significantly fewer cotton bolls, an indicator of potential yield, in 'cut' compared to 'filled' areas. In a study in southern Queensland, Jones et al. (2022) identified that cut and fill was the variable most negatively contributing to irrigated cotton yield in 26% of the study area. The change in soil properties identified by Cay and Cattle (2005), and broader negative influences on yield, is likely to be present in the Murrumbidgee Valley, with soils significantly altered since first being surveyed. These changes complicate utilising historical data to understand the influence of soil on crop yields.

Exacerbated soil change also occurs in instances where land has not been physically reformed. Filippi et al. (2016) provided a detailed discussion on global soil degradation issues influenced by agriculture, including acidification, erosion, salinisation and soil fertility. A variety of global examples are given to showcase how these processes impact specific soil properties over time (Filippi et al., 2016). Across the cotton growing region near Hillston, adjacent to this study area, Filippi et al. (2018b) identified a mild acidification trend between 2002 and 2015. Across the same area and time-period, Filippi et al. (2018a) examined changes to soil electrical conductivity (EC) and exchangeable sodium percentage (ESP). ESP values tended to decrease over time, most significantly in cotton growing areas, while EC both increased, primarily in horticultural soils, and decreased, in soils used to grow cotton, with statistically significant change observed in isolated areas only (Filippi et al., 2018a). These changes relate to the

proportion of salts being added to the soil either through saline irrigation water or specific fertilisers, in association excess water application or heavy rainfall events.

Physical deterioration is also noted as a major soil degradation issue by Filippi et al. (2016). Physical change to agricultural soils often focuses on compaction which, measured through soil bulk density, increases as a result of heavy machinery and livestock trafficking (Hamza & Anderson, 2005). Soils of the region are also known to be susceptible to coalescence (Cockroft & Olsson, 2000; Daniells, 2012) and hardsetting (Mullins et al., 1990). While these processes occur naturally through wetting and drying cycles, they have the potential to be accelerated by irrigation cycles. This may impact the long-term sustainability of crops reliant upon irrigation water by limiting the effective rooting depth of plants and/or increasing waterlogging susceptibility.

2.5.3. Soil sequences

The presence of levees and/or dunes in association with palaeochannel systems is a key component of the soil sequences discussed earlier in this review. Where these are absent, difficulties may arise in applying a generalised soil sequence to palaeochannel systems. The absence of levees may be a result of palaeochannel form or due to secondary aeolian action as already discussed. A background understanding of the region is required to prevent incorrect conclusions regarding soil distribution being drawn based on the presence and/or characteristics of levee systems. A potential secondary consequence of dunes remobilising as sheets across the plain (van Dijk & Talsma, 1964) is the masking of other surface soil features. This, however, has not been reported.

The influence of ‘cut’ and ‘fill’ practices has already been discussed. This too has the potential to diminish the value of soil sequences by moving a part of the soil profile from one location on the sequence to another, thus, altering the sequence. In areas where parna sheets are present near the surface and development for irrigation has occurred, these sheets may have been mixed with other layers of the soil profile.

2.5.4. Spatial supports and uncertainty of digital soil resources

The ability for digital soil maps to accurately model soil properties is reliant upon the quality and abundance of input data. When considering the value of the available DSMs, both the farm and regional scale are important. While much of this work, like earlier geomorphological and

soil landscape studies, occurs at a broader scale in most instances the key stakeholders are private landowners. Consequently, the ability to downscale resources to support decisions at the farm scale is important. Searle et al. (2021) show soil data to be sparsely available in this region compared to other areas of the state. Consequently, it is not surprising that accuracy of model predictions decreases. In the lower Murrumbidgee, the NSW product (Gray, 2023; NSW OEH, 2017) predicts to a 'very low' confidence class, noting "very limited soil and landscape data is available at the regional scale". When examining this map (Figure 2.17), the soil classification for more than 70% of the region is "Vertosols". While localised knowledge supports a high presence of vertosols across the plains, directly intertwined within these are sands associated with palaeochannel streambeds, levees or source bordering dunes (Murphy et al., 2000). While this map identifies more significant source bordering dunes and levees as rudosols, it fails to acknowledge the presence of these finer scale palaeochannel deposits (Figure 2.17).

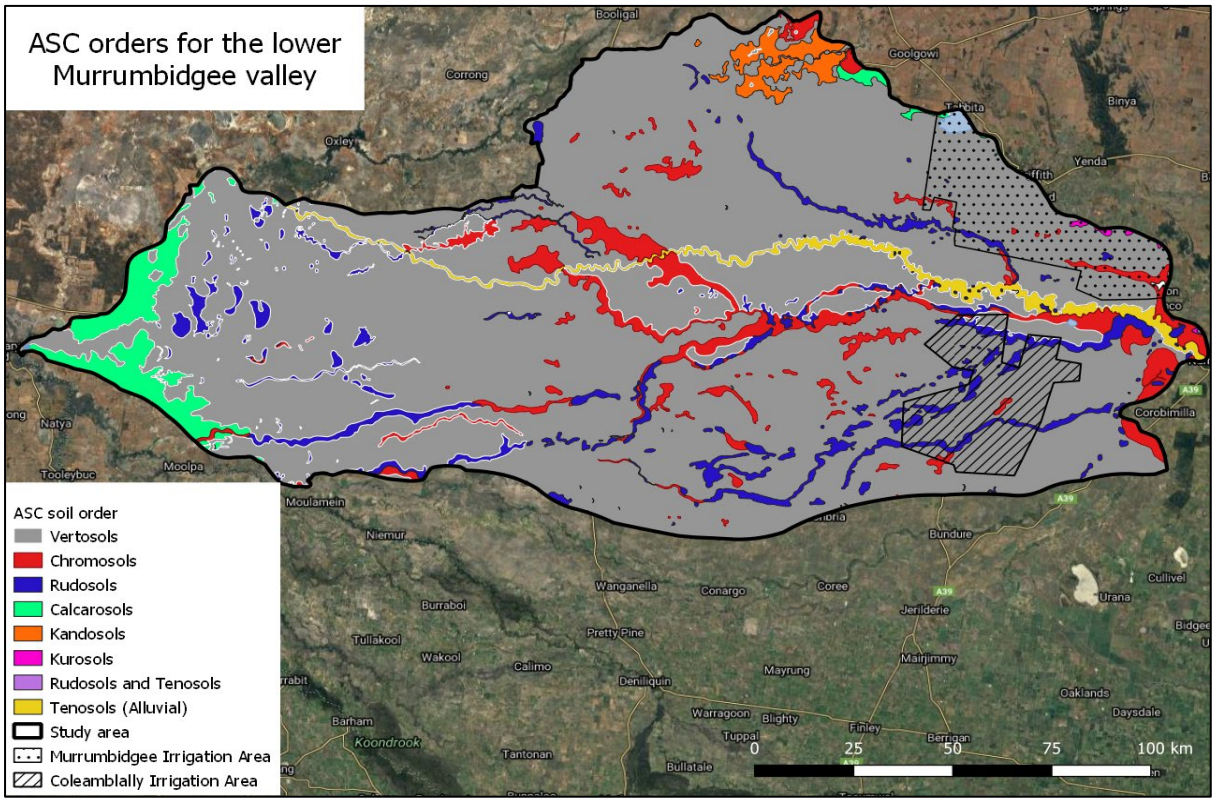


Figure 2.17. The Australian Soil Classification (ASC) orders, predicted at a very low confidence class, for the study area. The extent of the Murrumbidgee and Coleambally Irrigation Areas is also shown (NSW OEH, 2017).

The state (NSW DSM), national (SLGA) and global (SoilGrids250m) DSM products typically use a single model to make predictions across the entirety of the state/nation (Gray, 2023; Grundy et al., 2015; Hengl et al., 2017a; NSW OEH, 2017; Poggio et al., 2021; Viscarra Rossel

et al., 2014). Consequently, while the variability in this study area is known to be significant, it is at a much more localised scale than what these models, over larger areas, are built on and these variations are not detected. This is described in detail by Han et al. (2022) who examined the predictive capacity of the three different DSM products using external datasets at the point, strata (field management zone) and farm scale across 14 locations from central Queensland to southern NSW. For 0-0.3 m clay content reported validation statistics (LCCC) were 0.53 and 0.56 for the NSW and SLGA DSMs, respectively. When assessed against the external validation dataset at the different support levels, however, model validation statistics were below what was reported. At the point scale external validation LCCCs of 0.22 (NSW) and 0.33 (SLGA) were produced for clay content at 0.3-0.6 m. For the SLGA product the LCCC increased at the strata spatial support (LCCC=0.58) and remained similar at the broader farm scale (LCCC=0.51). External validation statistics, however, decreased for the NSW product at the strata support (LCCC=0.20) and again at the farm scale (LCCC=0.07). Soil organic carbon was validated at the same depth, with Han et al. (2022) concluding that when assessed against independent datasets, DSM products perform worse than their published model values and are unable to represent within field variability well. Similarly, Bishop et al. (2015) identified that prediction quality was much greater at a support scale of soil-land use complexes compared to point and 48 m block supports. Searle et al. (2021) also identified issues with the area of applicability of DSMs resulting from some areas nationally having dense numbers of datapoints in contrast to others, such as the lower Murrumbidgee valley, where data is sparse.

Practically, this suggests that while the pixel size of the DSM resources is theoretically fine enough to predict variability at the field scale, the products in their current form do not identify this variability. Han et al. (2022) did not find that national models, built with a larger soil observation dataset, outperformed the state model. This suggests that the distribution and density of soil observation datapoints is vital. Therefore, while the current digital soil mapping resources do identify broad landscape trends across the lower Murrumbidgee valley, they are not able to identify localised variability. This undermines their potential for use in decision support at the farm or paddock scale.

2.6. Opportunities to enhance understandings of the region's soils

Increasing available data underpins the potential knowledge enhancement of lower Murrumbidgee valley soils. This is central to the development of robust digital soil maps and will allow for a greater understanding of how the region's soils have changed, or may change,

temporally. Farmers, along with consultants and agronomists as intermediaries, are a primary stakeholder of the soil and, consequently, it is important to consider their requirements as the end-users of research while understanding opportunities and potential drawbacks in sharing soil data.

2.6.1. Publicly available data

The ability to manage and effectively deliver soil data is currently a major priority in Australia. The Federal Government National Soil Strategy (CSIRO, n.d.; Department of Agriculture, Fisheries and Forestry [DAFF], 2024) provided $15 million in the 2021-2022 budget to assist the Australian National Soil Information System (ANSIS) in improving this data storage and delivery. While often focused on surface soil samples, farmers regularly maintain a large database of soil measurements. In 2022 the Australian Government also implemented the Historical Soil Data Capture Payments program to increase data availability (DAFF, 2023). The 12-month program resulted in 553 data owners participating, with 291,203 results submitted (DAFF, 2023). Programs sharing data have the potential to be effective in supplementing soil survey activities, which are both time and cost intensive. Within the lower Murrumbidgee valley there is the potential for specific industries to work with stakeholders, encouraging them to share data to advance the localised industry. Essential to this is effective communication of the benefits that the stakeholders will receive and certainty surrounding data protection and sharing of data through appropriate terms of reference and use.

2.6.2. Legacy soil data

Where legacy soil data is available there is value in undertaking studies to determine how these soils have changed over time. At known geolocated sample points, if the temporal changes can be quantified there is the capacity to extend these findings over larger areas. Further, such quantification of change can then be harnessed to inform how soils might change into the future, including which soil types are most, or least, at risk of specific soil degradation processes.

There is also the potential to utilise geomorphological studies in soil mapping through a geo-pedological approach. Zinck et al., (2016) provided detail on integrating the two disciplines with numerous methodological examples in their book ‘Geopedology’. In Argentina, Leizica et al. (2022) utilised geomorphological maps alongside satellite NDVI and slope to identify

soil sampling locations, creating soil management zones by extrapolating soil type and capability using the geomorphic map.

2.6.3. Digital soil mapping approaches

In addition to the discussed current limitations of DSM products in the lower Murrumbidgee valley, there are significant opportunities. When considering lessons from the operationalisation of DSM in Australia (Kidd et al., 2020) there are clear pathways to improve resources both spatially and temporally (Searle et al., 2021). A key benefit of DSMs is their accessibility. There is the potential to generate readily available, easily interpretable soil maps allowing stakeholders to understand variability of specific soil properties more accurately at a farm and regional scale.

Regionally, the lower Murrumbidgee valley will benefit greatly from increased data accessibility. Once this data is available, there is a greater capacity to develop a more refined, regionally specific model. It is likely that such a model will have greater sensitivity to localised variability, with the potential to outperform larger state, national or global DSMs (Han et al., 2022). In creating these resources there are opportunities to use higher resolution covariates (Kidd et al., 2020) and frameworks to downscale data to the desired resolution (Malone et al., 2012; Malone et al., 2017). Legacy soil maps and surveyor knowledge can also be incorporated through various methodologies (Sulaeman et al., 2013; Bui et al., 2020; Ellili-Bargaoui et al., 2020). There is the capacity to develop DSMs at the farm and field scale utilising publicly available (remotely sensed) and privately collected (proximally sensed) data as model covariates. The increase in bespoke model quality from incorporating both covariate types is discussed by Filippi et al. (2024).

2.7. Conclusions

While the lower Murrumbidgee may appear as a relatively lifeless, flat plain, it in fact has a rich geomorphological history and diverse, highly variable soils. From a stratigraphic perspective it can be considered as ‘paleochannel spaghetti’ in ‘3D’ where differently aged sediments with different characteristics are intertwined at, and below, the surface. At the surface, mosaics of differently coloured sediments weave their way across the landscape. Below the surface, layers of older sediments continue to be intertwined also, with strings of sands among more clayey textured sediments impacting hydrology. As well as necessitating considerations of deep drainage, this hydrological variability influences the pedological development of surficial sediments, continuing to impact the soils that are farmed today. Atop of this are varyingly thick layers of aeolian deposits comprised of parna or remobilised dunes. While these interactions and the resulting patterns of soil distribution are complex, legacy studies highlight that they can be understood. At present, however, the efficacy of digital resources for the lower Murrumbidgee valley is reduced due to the sparseness of available data. Similarly, while this review draws on a wide range of studies from across over half a century, from a soils perspective this has almost solely occurred in the MIA and CIA. In the west, however, there is a paucity of available information regarding soil types and their characteristics. While geomorphological maps extend through the western reaches, there are no considerations of how palaeochannel anabranches impact soil variability. It is in these areas where variability in patterns of sediment distribution become most complex and future research should be targeted. New DSM methodologies present opportunities to address gaps in spatial understandings, however, more geolocated data is required to produce models that are accurate at a local level, a necessity given the fine scale at which variability occurs.

Chapter 3

3. Materials and methods – study area, soil sampling and soil property datasets

3.1. Introduction

While each chapter of this thesis has a different objective, there is a base dataset, or a subset thereof, that each chapter uses. To avoid unnecessary repetition in each of the research chapters the relevant background information on the obtaining of these datasets will be described here. This first step in this consists of differentiating between the 'valley-wide' and 'within-field' data subsets and describing the methodology used to identify sampling sites. Combined, these subsets form the 'full-dataset'. This full dataset is utilised in Chapters 4, 6 and 7 of this thesis. The within-field subset is focused on soils at the farm scale and is utilised in Chapters 4 and 5. The laboratory methods used to collect data will be described, as will the use of visible near infrared (VisNIR) spectroscopy to make predictions of particle size distribution. Finally, while different modelling approaches and combinations of predictor variables are used in each of the chapters, some background information on digital soil mapping (DSM) is also provided.

3.2. Study area, population centres and agricultural production

The area examined in this study, referred to as the "lower Murrumbidgee", was defined spatially by extracting the area within the bounds of the Murrumbidgee River valley catchment and the Riverine Plain bioregion (Figure 2.1). While the lower Murrumbidgee's alluvial soils have developed as a result of palaeochannel system activity, as described in the review of literature, the modern Murrumbidgee River is the defining landscape feature. While this Holocene system has played only a minor role in landscape evolution and soil development it is essential to the region, providing water for human consumption and irrigation.

The area is defined as dry and semi-arid (Stern et al., 2000) with rainfall decreasing along an east-west gradient from Leeton (432 mm rainfall annually) to Balranald (350 mm rainfall annually) (Table 3.1). Despite this, higher rainfall in the upstream areas of the Murrumbidgee catchment allows for access to water for irrigation which is the primary driver of the region's economy. The region includes two purposely designed irrigation areas, the development of which were tied to the building of dams in the upper, high rainfall regions. The older Murrumbidgee Irrigation Area (MIA) was developed during the 1910s following the construction of Burrunjuck Dam (1,028,000 ML capacity). The more recent Coleambally Irrigation Area (CIA) was developed in the 1960s alongside the construction of Blowering Dam (1,628,000 ML capacity).

Table 3.1. Summary climatic statistics from population centres across the region showing the summer and winter mean maximum and minimum temperatures as well as the average number (no.) of rainfall days and mean annual rainfall. Data collated from BOM (2023).

Location	Summer mean maximum temperature (ºC)	Summer mean minimum temperature (ºC)	Winter mean maximum temperature (ºC)	Winter mean minimum temperature (ºC)	Mean no. rainfall days (>-1mm)	Mean annual rainfall (mm)
Narrandera	32.3	16.8	15.2	3.6	59	433
Yanco	32.5	17.9	15.3	5.2	53	394
Griffith	32.2	16.7	15.3	3.9	48	397
Hay	32.3	16.0	16.1	3.9	50	367
Balranald	32.2	16.0	16.5	4.2	46	323

Beyond these named irrigation areas, direct extraction from the Murrumbidgee River or groundwater aquifers allows for irrigation to occur. The selection of crops grown relies primarily on water availability and soil type. Recently there has been an expansion of cotton production on heavier soil types, with a need identified to increase the information relating to soils that is available to these landowners (Holland and Eastwood, 2014). Prior to the recent expansion of the cotton industry within the region, a productivity report submission found the MIA alone was estimated to contribute over AUD $5 billion to the economy (MI, 2010).

3.3. Soil dataset collection

A soil sampling plan was formulated to account for the main soil types under cotton production while also representing variability in soils not used for irrigated agriculture. A total of 153 soil cores with a diameter of 40 mm were extracted to a depth of 1 m (Figure 3.1). The soil coring rig used to extract cores was manufactured by Christie Engineering (Sydney, Australia) to specific design specifications in 2016 and utilises a hydraulic hammer as opposed to a rotating drill mechanism. The rig is mounted on a Toyota Landcruiser ute and is powered by a 4 stroke Honda GX270 motor. At each sampling site notes were made on the surface condition of the soil and surrounding landscape. Once extracted, soil cores were placed in half-cut polyvinyl chloride (PVC) pipes, wrapped in plastic sheeting and returned to the laboratory. The sample sites and obtained data are described in two subsets that, combined, form the ‘full dataset’

(Table 3.2). These subsets will, henceforth, be referred to as the ‘valley-wide subset’ and the ‘within-field subset’.

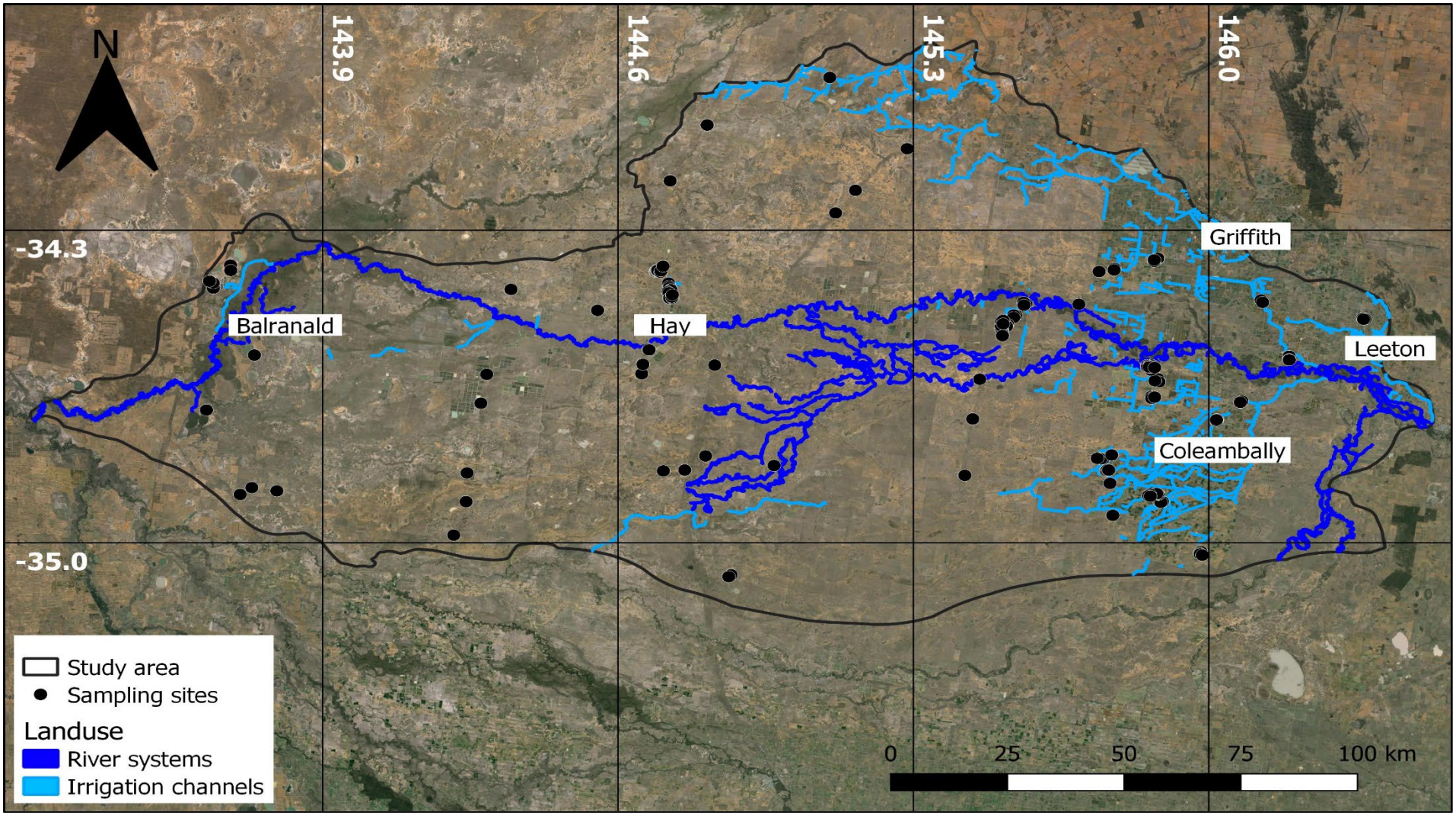


Figure 3.1. The location of the 153 sample sites from across the study area where soil cores were extracted to a depth of 1 metre. Shown also are the major population centres, pathways of the Murrumbidgee River and its distributary systems as well as Irrigation supply channels (NSW DCCEEW, 2023).

Table 3.2. The breakdown of sample sites within the two subsets showing the farm, landuse and number of sites.

Subset	Farm	Landuse	Number of sites
Valley-wide		Irrigated cotton	73
		Natural condition	43
		Dryland cropping	2
		Total valley wide	118
Within-field	*Farm A*	Irrigated cotton	8
		Natural condition	1
	Farm B	Irrigated cotton	8
	Farm C	Irrigated cotton	18
		Total within-field	35
Full dataset total			153

3.3.1. Valley wide subset

The first subset is the valley-wide subset, consisting of 118 soil cores taken from various landuses across the entirety of the region (Table 3.2). The selection of sampling sites was undertaken in two parts. Firstly, as this project is aimed towards cotton production, a purposive sampling plan was formulated with local stakeholders to identify 17 cotton farms from across the region that were representative of soil variability. Across these farms, 63 soil cores were taken, with multiple cores taken per farm where soil variability was present. This results in the observed clusters of sample points (Figure 3.1). Secondly, a covariate stack of remotely sensed, publicly available data was used to group the study area into clusters on the basis of similarity (Table 3.3). A stratified random sampling plan was then developed and executed within these clusters.

3.3.1.1. K-means clustering and stratified random sampling plan

In total, 55 locations were selected for sampling following K-means clustering of remotely sensed covariate data and the implementation of a stratified random sampling plan. This methodology intended to identify sites to compliment the purposive sampling plan. In undertaking this, a 30 × 30 grid was created for the study area before each spatial covariate was extracted on the grid (Table 3.3).

Table 3.3. The full suite of covariate data utilised in the development of sampling clusters in this study

Data category	Data description	Data source	Resolution
Terrain	Elevation (m)	Gallant et al. 2011	30 m
	Slope (%)	Gallant and Austin (2012a)	30 m
	Slope (degree)	Gallant and Austin (2012a)	30 m
	Slope (relief)	Gallant and Austin (2012b)	30 m
	TWI	Gallant and Austin (2012c)	30 m
	MrVBF	Gallant et al. (2012)	30 m
	MrRTF	Gallant et al. (2013)	30 m
	Aspect	Gallant and Austin (2012d)	30 m
	TPI (class)	Gallant and Austin (2012e)	30 m
	TPI (mask)	Gallant and Austin (2012e)	30 m
Radiometric	Dose rate	Minty et al. (2009)	100 m
	Uranium (ppm)	Minty et al. (2009)	100 m
	Potassium (%)	Minty et al. (2009)	100 m
	Thorium (ppm)	Minty et al. (2009)	100 m
	Uranium: Thorium (ratio)	Minty et al. (2009)	100 m
	$Uranium^2$:Thorium (ratio)	Minty et al. (2009)	100 m
	Uranium: Potassium (ratio)	Minty et al. (2009)	100 m
	Potassium: Thorium (ratio)	Minty et al. (2009)	100 m
Geological	Silica content (%)	Gray et al. (2016)	1:250,000
Satellite	Landsat 7 NDVI 5^{th} percentile	Gorelick et al. (2017)	30 m
	Landsat 7 NDVI 50^{th} percentile	Gorelick et al. (2017)	30 m
	Landsat 7 NDVI 95^{th} percentile	Gorelick et al. (2017)	30 m
	Landsat 7 Band 3 (Red) 5^{th} percentile	Gorelick et al. (2017)	30 m
	Landsat 7 Band 3 (Red) 50^{th} percentile	Gorelick et al. (2017)	30 m
	Landsat 7 Band 3 (Red) 95^{th} percentile	Gorelick et al. (2017)	30 m

Following visual and statistical analysis of different covariate combinations, it was decided that NDVI 50^{th} percentile, radiometric thorium, silica and slope (%) would be used to formulate the clusters. These variables were selected based on their ability to represent known landscape variability, including the influence of previously mapped paleochannel systems (Page and Nanson, 1996), various patterns of differently described plains (Butler et al., 1973) and previously published digital soil resources (NSW OEH, 2017). DEM was not included in the clustering analysis, due to the unrealistic east-to-west banding effect it caused while additional covariates complicated the model resulting in speckled, non-continuous, clusters. A visual

examination of the 'elbow plot' (Figure 3.2), and considerations of the appropriate weighting of 55 sample sites among clusters, resulted in six clusters being chosen.

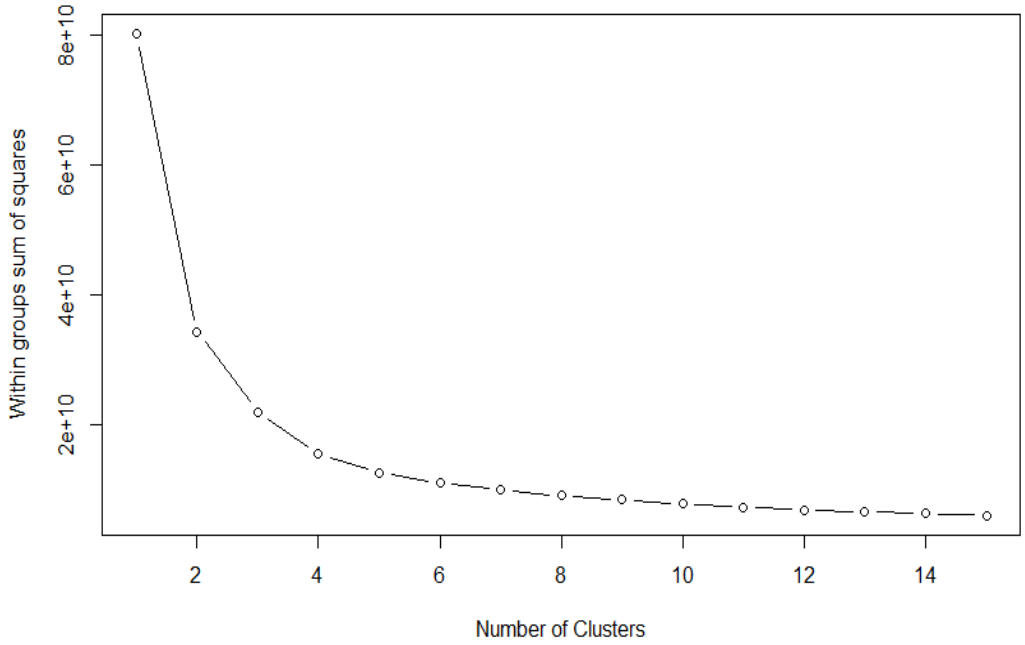


Figure 3.2. The 'elbow plot' showcasing a slowing rate in the decrease of the within groups sum of squares as the number of clusters for the region extends beyond six. This was chosen to balance the statistically optimal number of clusters without limiting how the 55 soil cores would be distributed between them.

The covariate stack containing NDVI 50th percentile, radiometric thorium, silica and slope was then partitioned into the six clusters with parameters set to have 20 different initialisations and a maximum of 100 iterations for each initialisation. Once each cluster was defined, a stratified random sampling plan was employed whereby the number of samples assigned per cluster was weighted depending on the overall size of that cluster (Figure 3.3).

Over the course of 2020, 2021 and 2022 numerous alterations to this sampling plan were necessary. Initially, the outlined weighting toward cotton farms was to be paired with a larger number of non-cotton soils, however, an inability to travel to the region due to multiple pandemic enforced lockdowns meant that the overall number of cores taken needed to be reduced. Further, La Niña weather events resulted in significant flooding in the region during 2021 and 2022 making some sampling points inaccessible due to road closures. In these instances, sampling sites were moved to the nearest accessible point within the predefined cluster.

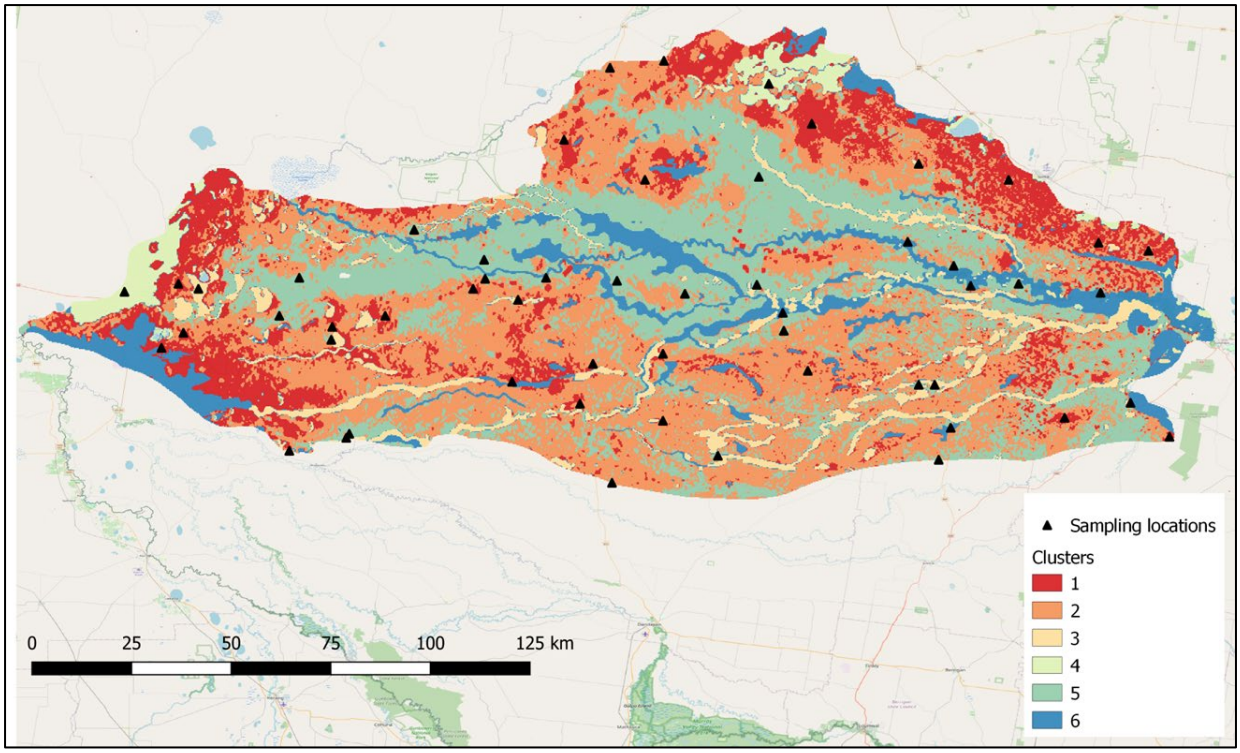


Figure 3.3. The finalised clusters and identified sampling locations. Of note are the spatial patterns of each cluster which accurately represent previous work by following the paths of known palaeochannel systems and their associated drainage plains.

3.3.2. Within-field subsets

The second, within-field, subset encompassed 35 soil cores (Table 3.2), including one native vegetation site, extracted at a higher density from three paddocks across the study area (Figure 3.4). The names of the farms on which these paddocks were located have been anonymised to Farm A, Farm B and Farm C. Data obtained from Farm A and Farm C is the basis of Chapter 5, examining the influence of paleochannels on within-field soil variability and how this impacts crop production. It is also referenced in Chapter 4 in discussions on localised soil morphological variability. These sites were selected due to their proximity to known palaeochannels of different characteristics and the available proximally sensed soil and crop yield data. The selection of sample sites occurred through clustering yield data and considerations of proximally sensed data, with more detail surrounding this provided in Chapter 5.

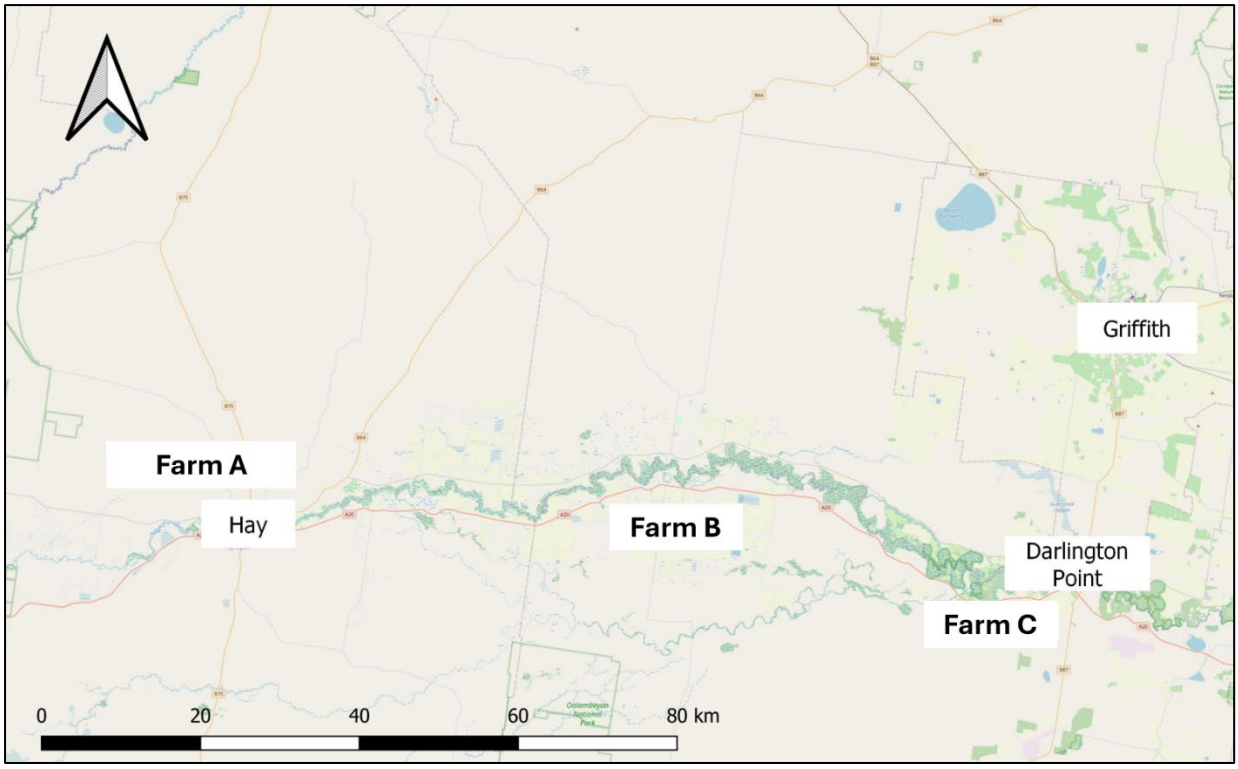


Figure 3.4. The location of the farms where the within-field subsets were sampled in this study shown in relation to the population centres of Griffith, Hay and Darlington point.

3.3.2.1. Farm A subset (n=9)

The Farm A subset involved the extraction of 8 soil cores from an 82 hectare paddock and one paired native vegetation site on the property 'Farm A', located 18 km north-west of Hay (Figure 3.4). The site at Farm A was chosen due to a sand seam that is incised in the heavy clay landscape. This is reflected in both crop yield and proximally sensed soil data. This stream can be a considered a terminal, distributary channel associated with the Gum Creek palaeochannel system which is estimated to have been active between 35 to 25 ka (Page and Nanson, 1996).

3.3.2.2. Farm B subset (n=8)

The Farm B subset involved the extraction of eight soil cores from a 240 ha paddock on the property 'Farm B', located centrally within the region near the town of Carrathool, approximately 50 km east of Hay (Figure 3.4). The paddock is located 'midstream', on plains approximately 25 km north of the Kerarbury Palaeochannel system (Page and Nanson, 1996).

3.3.2.3. Farm C subset (n=18)

The Farm C subset relates to 18 soil cores which were extracted from a 94 hectare paddock on the property 'Farm C', located approximately 12 km southwest of the township of Darlington Point (Figure 3.4). The paddock was chosen due to its close proximity to the Waddi River Ridge (Waddi Reach) (van Dijk and Talsma, 1964), an arm of the Kerabury system which is estimated

to have been active between 55-35 ka (Page and Nanson, 1996). Unlike the terminal nature of the sand seam at Farm A, at a site nearby the width of the Waddi Reach was estimated to be 220 m with a bankfull discharge of 2610 m^3 s^{-3}, 8.3 times greater than the modern river. Consequently, it has had a more significant effect on the landscape with a preserved source bordering dune still present. For the Farm C subset, 18 cores were extracted from a 94 hectare paddock.

3.4. Soil property analysis

Once extracted and returned to the laboratory, all cores (n=153) were vertically halved to reveal a more 'natural face'. Each core was then horizonated and morphologically described. The occurrence and observed depth of lime, gypsum, coarse fragments (>2 mm) and subdominant colours were noted. This data is utilised in Chapter 4 of this thesis.

It was decided that each soil core would be examined at five depths, these being 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m. A shallower surface sample of 0-0.05 m was considered, however, given many of soils within this study have been heavily, and in some instances recently, cultivated it was decided that a surface sample of 0-0.1 m would be more appropriate.

After morphological examination, depth increment samples were air dried (40ºC), ground and passed through a 2 mm sieve. In total, this resulted in 764 samples for further processing and analysis as one core was only extracted to 0.8 m.

All dried and ground soil samples were analysed utilising traditional laboratory methods to determine pH, electrical conductivity and base cations. Due to the large number of samples and the time-consuming nature of determining soil particle size distribution using the pipette method of particle size analysis (Glendon & Or, 2002), not all samples underwent particle size analysis, with samples from the valley-wide and within-field subsets selected for analysis separately. All samples were then hand ground to 200 μm before being scanned with visible near infrared spectroscopy (VisNIR).

3.4.1. Soil pH

Soil pH was analysed for all samples (n=764) in a 1:5 soil to water suspension using a SevenCompactTM s220 pH/Ion meter. The meter was calibrated utilising solutions with a known pH of 4, 7 and 10 after every 20 samples.

3.4.2. Soil electrical conductivity

As a measure of salinity, soil electrical conductivity (EC) was measured for all samples in the full dataset (n=764) in a 1:5 soil to water extract using a S230 Conductivity meter. The meter was calibrated using a 1.413 (dS/m) solution every 20 samples. The measured EC was converted to soil electrical conductivity of the extract (ECe) through a conversion factor based on the soil texture grade (Table 3.4).

Table 3.4. Conversion factors to estimate ECe (dS/m) from EC1:5 (dS/m) on the basis of soil texture as adapted from Hazleton and Murphy (2007)

Soil texture	**Conversion factor**
Sand (S), Loamy sand (LS), Clayey sand (CS)	23
Sandy loam (SL)	14
Loam (L) Silty loam (ZL), Sandy clay loam (SCL)	9.5
Clay loam (CL), Silty clay loam (ZCL), Clay loam, sandy (CLS)	8.6
Light clay (LC), Light medium clay (LMC)	8.6
Medium clay (MC)	7.5
Heavy clay (HC)	5.8

3.4.3. Exchangeable base cations

For all samples (n=764), 35 g of soil was sent for external analysis of exchangeable base cations (Ca^{2+}, Mg^{2+}, K^{+} and Na^{+}) as well as the overall cation exchange capacity (CEC). Due to known alkalinity of samples the 15C1 method of Rayment and Lyons (2011) was chosen. This entails a pretreatment of 60% aqueous ethanol and 20% aqueous glycerol to remove soluble salts. The extraction of cations by leaching then occurred using 1 M ammonium chloride (NH_4Cl) at pH 8.5. The exchangeable sodium percentage (ESP), an indicator of soil sodicity, was determined for each sample by dividing the exchangeable Na^{+} cations (cmol(+)/kg) by the overall CEC (cmol(+)/kg) and multiplying this value by 100.

3.4.4. Particle size analysis

In total, 205 samples from the full dataset underwent particle size analysis using the pipette method (Glendon & Or, 2002). This occurred in two stages due to the prolonged period over which fieldwork occurred. Samples were selected for analysis separately from the within-field subset (105 samples analysed) and the valley-wide subset (100 samples analysed).

For each of 35 soil cores in the within-field subset, three depths underwent particle size analysis. The 0-0.1 m sample for each core was analysed with the depths then alternating between 0.1-0.3 and 0.6-0.8 m or 0.3-0.6 and 0.8-1 m between cores to understand variability down the profile while balancing the time taken to undertake analysis. This produced 105 datapoints across 35 soil cores for sand, silt and clay content. This was undertaken to allow initial analysis for Chapter 5 to proceed.

The remaining (n=100) samples were selected for particle size analysis through spectroscopic clustering, which is discussed below. Categorisations of soil texture for use in converting soil EC to soil ECe (Table 3.4) were made using the triangular texture diagram and field texture grade table (CSIRO Publishing and the National Committee on Soil and Terrain [NCST], 2009).

3.5. Using visible near infrared (VisNIR) spectroscopy to predict particle size distribution

While pH and EC can be measured rapidly in the laboratory and, in this instance, funding enabled the external analysis of exchangeable base cations, undertaking particle size analysis is both time and labour intensive. Consequently, VisNIR spectroscopic measurements were used to identify an appropriate subset of samples for analysis before being used to model sand and clay content with machine learning techniques. This approach is well established, and it is known that soil particle size can be accurately predicted using this method (Bricklemyer and Brown, 2010; Ackerson et al., 2015; Chen et al., 2021; Gozukara et al., 2022). Once finely ground, all samples (n=764) were scanned using a contact probe with inbuilt halogen lamp on an ASF Agrispec spectrophotometer (Malvern Panalytical, Boulder, Co, USA). The instrument was calibrated using a Spectralon white tile with re-calibration occurring after every 10 scans.

3.5.1. Treatment of spectroscopic data

Spectroscopic data was imported into the software R (R Core Team, 2021) and analysed using functions from the ‘soilspec’ package, following methods detailed by Wadoux et al. (2021). Firstly, spectra were coerced into a single dataframe before the reflectance was converted to absorbance (absorbance = log(1/reflectance)). Spectral trimming then occurred from 500-2451 nm to remove noisy portions before a splice correction was performed at 1000 and 1800 nm. Spectra were smoothed using a Savitzky–Golay filter with a window size of 11 and fitting a second order polynomial before a standard normal variate transformation was performed.

Finally, to reduce the overall data size, spectral reduction occurred by resampling spectra at 8 nm intervals.

3.5.2. Exploratory analysis of spectral data

A principal component analysis (PCA) was performed to reduce the dimensionality of data and assist in the detection of outliers. It was shown that the first three principal components cumulatively explained 95.2% of the total variance. Utilising these scores, the Mahalanobis distance between the spectra and centre of the principal component (PC) scores was calculated, with scores exceeding a level of 3 identified as potential outliers. Two samples were identified as potential outliers (Figure 3.5). To confirm if these measurements were outliers, the Mahalanobis distance values were plotted while selected samples were then also plotted in association with the top three PC scores and in an absorbance wavelength plot (Figure 3.5). These outliers were removed before proceeding with analysis.

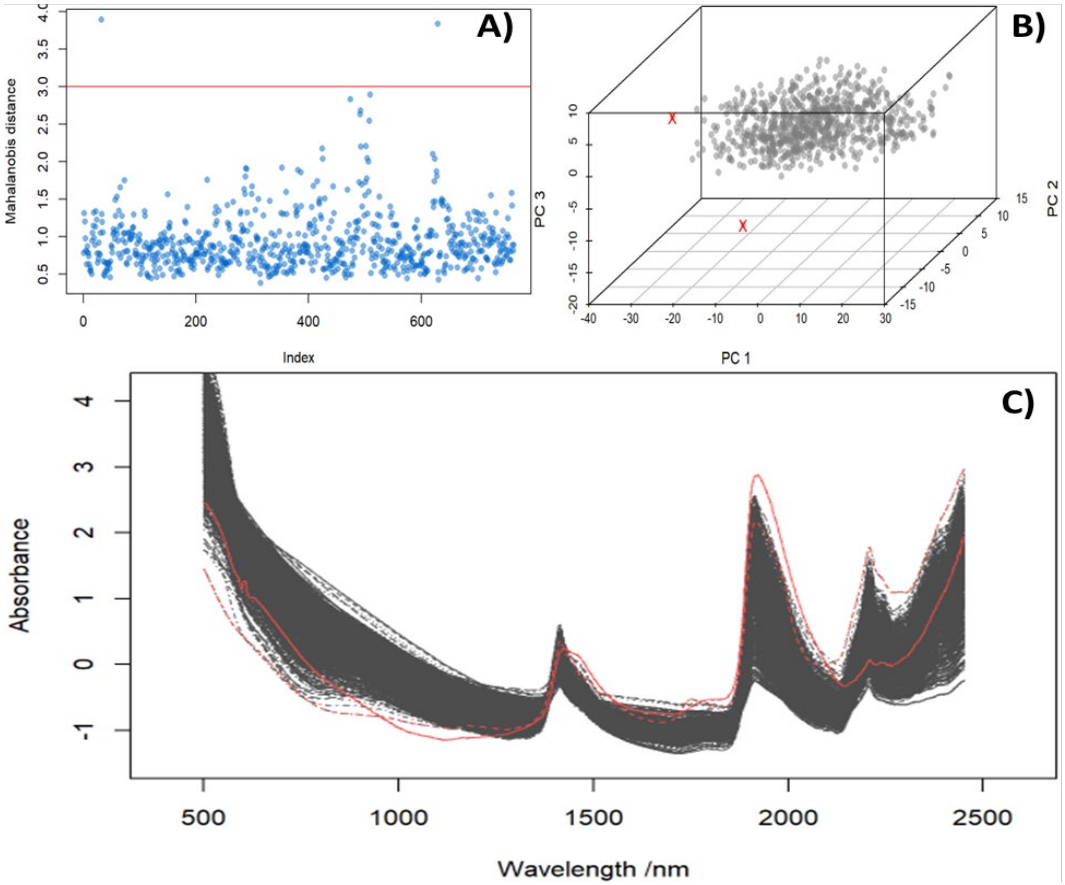


Figure 3.5. Output figures highlighting the detection and validation of outlier spectra where; **A)** indicates two spectra with a Mahalanobis score >3 out of the entire sample set (n=764), **B)** shows these two spectra, represented with red X marks, in relation to the top three PCs, and, **C)** shows the normalised absorbance of these two spectra, highlighted in red, against the remaining dataset across the wavelengths examined.

3.5.3. Selection of samples for particle size analysis testing

Following the deletion of outliers, eight scaled PCs, accounting for 99.9% of variance, were retained as the dataset dimensionality had already been reduced, meaning that the dataset size was not an issue. These PCs were then used to select samples representative of the valley-wide subset (n=589). This excluded all samples for cores from the within-field subset (n=175) that had already been particle size analysed through a separate selection process. A cluster stability score was then generated for sampling sizes between 30 and 360 in increments of 30 to compare three sampling methods; k-means clustering, conditioned latin hypercube sampling and Kennard-stone sampling. For each sample size (30-360) the three methods were compared on the basis of their mean square distance (msd) and standard deviation of mean square distance (sd_msd). The conditioned latin hypercube sampling and k-means clustering produced similar results with little decrease in the measured msd beyond a sample size of 100.

K-means clustering, on the basis that it is robust and widely used, was chosen to randomly select 100 samples, whereby each sample is that located closest to the centre of 100 created clusters. The decision to select 100 samples was a balance between lowering the msd and the time taken to undertake analysis. The selected samples then underwent particle size analysis using the pipette method of Glendon and Or (2002) as previously described.

3.5.4. Model predictions of sand and clay content

Spectroscopy was used also in a purely operational capacity to estimate the percentage sand and clay content of samples. Utilising the particle size distribution data obtained from both the within-field (n=105) and valley-wide (n=100) subsets, two cubist models were developed to predict sand and clay content. The first iteration of these models utilised only spectroscopic data as a predictor variable. Given, however, cation exchange capacity (CEC) measurements were available for all samples, the second iteration of this model included CEC as a predictor alongside the spectra.

Model validation was undertaken using a leave-one-sample-out-cross-validation (LOSOCV). Model quality was assessed primarily using the Lin's Concordance Correlation Coefficient (LCCC), a unit-less statistic showing the agreement between the predicted and observed measurement considering precision and accuracy whereby a value closer to 1 indicates a higher quality model (Lin, 1989). The root mean square error (RMSE) and bias of each model are also reported.

Given the strong relationship between CEC and soil texture, whereby more clayey soils generally have a higher CEC, it is not surprising that the models containing CEC outperformed those without. In this instance, the model predicting clay content with CEC produced an LCCC in the LOSOCV of 0.89 while the model with spectra as the sole predictor produced an LCCC of only 0.70, also in the LOSOCV (Figure 3.6). Similarly, the model containing CEC and spectra to predict sand content produced an LCCC in the LOSOCV of 0.91 as opposed to the model with spectra alone, which produced an LCCC of 0.75, also through the LOSOCV (Figure 3.7). Consequently, predictions for sand and clay were made using the models containing CEC and spectra.

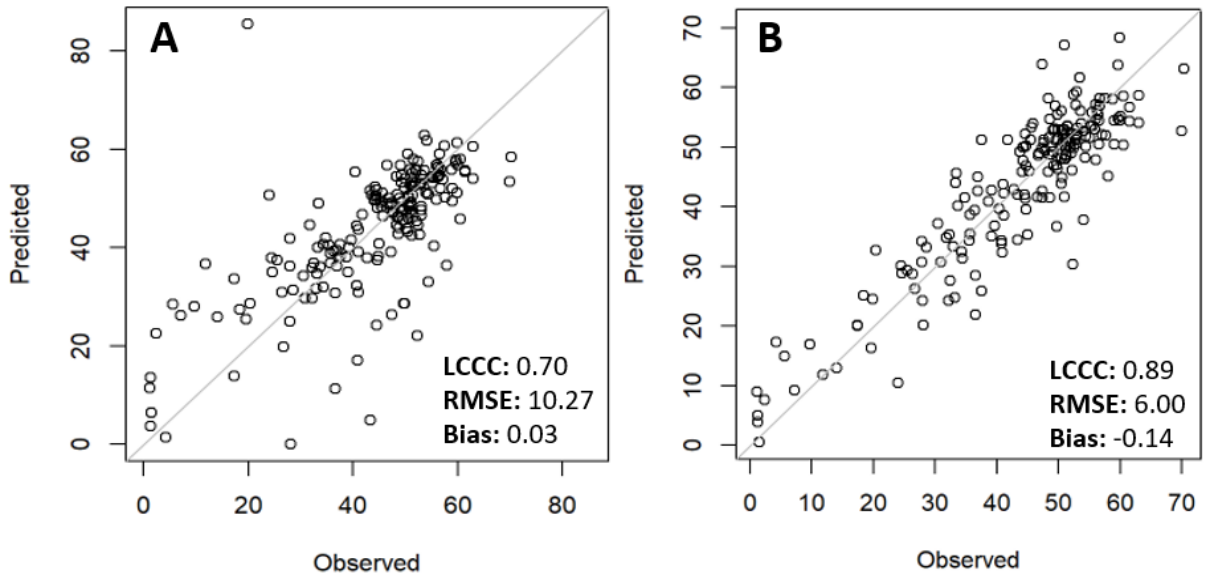


Figure 3.6. Plotted predicted vs observed values for models predicting clay content containing only spectral values (A), and spectral and CEC values (B), with relevant model statistics reported.

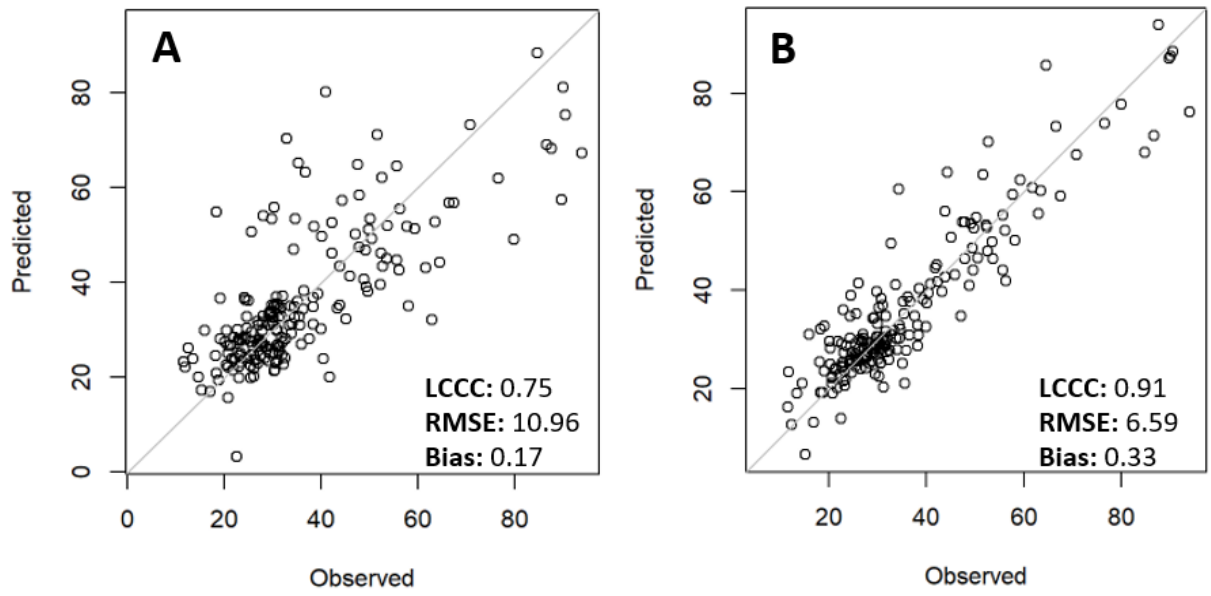


Figure 3.7. Plotted predicted vs observed values for models predicting sand content containing only spectral values (A), and spectral and CEC values (B), with relevant model statistics reported.

3.6. General digital soil modelling and mapping approach

All data analysis within this thesis occurred using various packages within the program 'R' (R Core Team, 2023). Produced maps were either directly exported from R or were formatted using 'QGIS' (QGIS.org, 2023).

While each chapter of this thesis employs different modelling approaches to predict categorical and continuous variables at different scales, the theory behind these approaches is consistent. These methods rely on an adaptation of the now commonly applied 'scorpan' approach to DSM which was presented by McBratney et al. (2003). Whereas the traditional equation of Jenny (1943) sought to understand factors affecting soil formation, the 'scorpan' approach seeks to allow for the prediction of soil classes, or properties, through statistical relationships between predictor variables, whereby;

S_c *or* $S_a = f(s, c, o, r, p, a, n) + e.$

In this instance a soil property (S_a) or a soil class (S_c) is predicted as a function of soil (s), climate (c), organisms (o), relief (r), parent material (p), age (a) and spatial position (n), with spatially correlated residuals (e) also incorporated (McBratney et al., 2003).

This thesis is not designed to develop novel approaches for DSM, nor will it provide detailed reviews of the field with substantial resources already available in this area (McBratney et al., 2003; Lagacherie, 2008; Zhang et al., 2017; Malone et al., 2017). Instead, it focuses on applying

robust approaches of DSM as tools for quantifying different aspects of the soil resource, at different scales, with different outcomes showcased in each chapter. The regional and farm scale approach to DSM allows for the incorporation of local knowledge to inform the selection of specific covariate, or predictor, variables within the 'scorpan' function. This thesis will incorporate publicly available spatial covariates from four categories: terrain, radiometric, geological and satellite (Table 3.3).

Climate data was not incorporated into models as rainfall follows the same trend and rate of change along the east-west axis of the region as elevation, whereby, as elevation decreases in a westerly direction rainfall also decreases. There is also little variation in average summer, or winter, temperatures across the region (Table 3.1).

Throughout each of the following chapters different combinations of spatial covariates were chosen as model predictors representative of specific factors. These covariates will consistently be considered as terrain, satellite, geological and radiometric attributes (Table 3.3)

Terrain attributes include DEM; slope percentage, degree and relief; aspect; topographic position index (TPI) class and mask; topographic wetness index (TWI); multi-resolution ridge top flatness (MrRTF); and multi-resolution valley bottom flatness (MrVBF) (Table 3.3). These were derived from the Shuttle Radar Topography Mission (SRTM) and accessed through both the CSIRO Data Access Portal or by using the dataharvester package within R (Haan et al., 2023).

Satellite attributes, excluding the Landsat 7 imagery utilised to develop clusters in this chapter, refer to the barest earth dataset developed by Roberts et al. (2019) and was obtained through the dataharvester package (Haan et al., 2023). The barest earth dataset was developed by extraction noise reduced, cloud free estimates of barest state land from a combined 30 year period of Landsat datasets (Roberts et al., 2019). The dataset includes the blue, green, red, NIR, SWIR and SWIR2 wavelengths (Table 3.3).

Radiometric attributes refer to potassium (% K), thorium (ppm eTh), Uranium (ppm eU), ratios of these and the total dose as presented by Minty et al. (2009) in their radiometric map of Australia (Table 3.3). These products are the results of the Australia-Wide Airborne Geophysical Survey which was back-calibrated and thus effectively levelled to the Atomic Energy Agency's radioelement datum (Minty et al., 2009). This data was obtained through the dataharvester package also (Haan et al., 2023).

The geological attribute incorporated throughout this thesis is the silica index developed by Gray et al. (2016) (Table 3.3). To assist in selecting sample sites on private properties, and for modelling in Chapter 5, proximally sensed, privately held, radiometric and electromagnetic induction covariates were also used.

3.7. Conclusion

The research chapters of this thesis will utilise the full dataset of this study, alongside the within-field subset, in different ways to examine how soil properties vary at both the within-field and valley-wide scale in the lower Murrumbidgee valley of southern NSW. External factors throughout the duration of this project necessitated that plans for obtaining soil data were robust and adaptable. Consequently, a purposive plan was implemented to account for the main soil types under cotton production; a stratified random sampling plan was used to identify sampling locations across the valley as a whole; and yield and proximally sensed data was used to identify sampling sites within-field to examine the influence of soil properties on cotton yield variability. The soil properties that are consistently referred to are texture (sand and clay content), cation exchange capacity (CEC), electrical conductivity of the extract (ECe), pH and exchangeable sodium percentage (ESP). While the applications of this data vary, the methods through which it was obtained are consistent. These consisted of both traditional laboratory and visible near infrared (VisNIR) spectroscopic methods. The latter were utilised to make predictions of sand content and clay content due to the large sample size and time-consuming nature of particle size analysis methods. Utilising spectra alongside cation exchange capacity (CEC) cubist models predicting clay and sand content produced Lin's Concordance Correlation Coefficients (LCCCs) of 0.89 and 0.91, respectively, following a leave-one-site-out-cross-validation (LOSOCV).

This study aims to showcase how morphological and analytical methods can be applied to increase understandings of inherent soil properties and their distribution within the lower Murrumbidgee valley. The specific methods employed in the subsequent research chapters reflects the desire of stakeholders to have increased information on the soil resource within this region.

Chapter 4

4. Morphological classifications of lower Murrumbidgee valley soils

4.1. Abstract

The lower Murrumbidgee valley in southern NSW exhibits diverse and highly variable soils The distribution of these soils is linked to palaeochannel systems. Previous classifications of soil types within localised groups have been focused on earthy soils in the region's east. Within the west of the region, however, there is a prolific occurrence of red and grey clays in association with these earths, as well as sands. This study morphologically described 153 soil cores from across the lower Murrumbidgee valley to a depth of one metre. These descriptions, alongside analytical data, were used to consider if the current soil groupings were representative of the whole region or if a new classification was required. Based on the dominance of clay soils it was decided to create a new classification. Soils were differentiated based on colour, texture and the presence, or absence, of lime and gypsum. This classification placed soil profiles in ten classes: Red-brown clays (n=16), Red-brown calcic clays (n=31), Red-brown gypsic clays (n=10), Grey clays (n=17), Grey calcic clays (n=26), Grey gypsic clays (n=7), Transitional earth soils (n=9), Texture contrast soils (n=6), Deep sands (n=3) and exceptions, which did not fall into the prior nine classes but were still described (n=28). The characteristics and distribution of each soil class is discussed at a regional level in relation to palaeochannel systems. The distinctness of morphological variability at the within-field scale is shown through transects of sites sampled on, or adjacent to, two palaeochannel systems. All soil classes, including the exceptions, were placed in four broader groups: Red-brown clays, Grey clays, Earths and Deep sands that were modelled and mapped across the lower Murrumbidgee valley using a random forest classification model. While this model appropriately identified the prolific occurrence of clays it underpredicted the sandier soil groups. There was significant variability in measured soil electrical conductivity of the extract (ECe) and exchangeable sodium percentage (ESP) within classes. Future research could, therefore, increase the dataset size and seek to differentiate classes further on account of these agriculturally important properties.

4.2. Introduction

Understandings of inherent soil properties are vital in developing and maintaining sustainable agricultural systems. The lower Murrumbidgee valley, in southern NSW, is known for its agricultural productivity contributing over AUD $5 billion to the economy annually (MI, 2010). Soil surveys are valuable in evaluating land for agricultural production (Bouma, 1986), and have underpinned the success of agriculture in the Murrumbidgee and Coleambally Irrigation Areas, developed in the 1910s and 1960s, respectively. These surveys and soil type descriptions, including the work of Taylor and Hooper (1938), van Dijk (1968, 1961), van Dijk and Talsma (1964), Stannard (1970) and Butler (1979) identified 94 localised soil types that were categorised into landscape associations and soil series. Beyond the bounds of these irrigation areas, however, there is limited publicly available soil information. High variability presents challenges in appropriately translating research, with landholders requiring further information on their soils (Holland & Eastwood, 2014). Soil classifications may form a part of the solution to this, with standardised descriptions reducing complexity and allowing for the effective extension of knowledge.

The lower Murrumbidgee Valley is located centrally within the Riverine Plain and Murray basin, an area described by Brown and Stephenson (1991) as fluviolacustrine. Four distinct geological units of Cenozoic era sediments fill this basin to depths of up to 600 m (Brown, 1989; Pucillo, 2005). Diverse geomorphological studies of the surficial units, the Upper Shepparton and Coonambidgal Formations, have examined the characteristics of, and relationships between, the dominant fluviatile and secondary aeolian landscape influences, with a geomorphic map produced by Butler et al. (1973).

The fluviatile sediments that fill the basin, and determine patterns of soil distribution, result from palaeochannel systems. These systems have been widely studied since first theorised by Butler (1950) as an origin of soil formation (Bowler, 1978; Butler, 1958; Langford-Smith, 1960; Pels, 1964, 1966, 1971; Schumm, 1968). More recently, thermoluminescence sediment dating and stream reconstructions allowed the delineation and description of four main palaeochannel systems within the last 100,000 years (Mueller et al., 2018; Page & Nanson, 1996; Page et al., 1996; Page et al., 2009). The primary reaches of these systems were mapped, with each palaeochannel following unique watercourses and exhibiting estimated streamloads up to eight times greater than the Murrumbidgee River (Page & Nanson, 1996; Page et al., 2009). During periods of activity, younger systems would incise within sediments of previous systems leading to new sediments being deposited atop the plain, before low-flow periods led

to the aggradation of sands and coarse sediments within the streambed, before the cycle repeated.

Excluding the aeolian influenced hillslopes (Butler, 1979) and western mallee landscapes (Bowler & Magee, 1978; Mitchell, 2002; Wasson, 1989), the streamloads of palaeochannel systems determined the distribution of alluvial sediments across of the plain. Clay particles (>0.002 mm), being much finer than sands (0.02-2 mm), remain in suspension longer, thus, being carried further from river courses before being deposited. This formed the basis of stratigraphic models and toposequences, such as those presented by Butler (1958) and Langford-Smith (1960), whereby sandy streambeds containing coarse fragments are bounded by deep, sandy levees, transitioning to texture contrast soils, transitional loams and lastly, heavy clays, with this sequence occurring over 7-8 miles (11-13 km). Secondary aeolian action through the remobilisation of streambed sands onto source bordering dunes and then across the plain can also occur (van Dijk Talsma, 1946; Pucillo, 2005).

While theoretical toposequences and geomorphic maps have formed the basis of understanding soil distribution, the relationships become more complex in the west of the region which Murphy et al. (2000) described as a mosaic. This is the result of deltaic-like, often terminal, palaeochannel anabranch fans incising within clays without altering landscape features. As a result, these are difficult to detect. A lack of soil surveys in the region's west adds to difficulty in understanding this variability. The main aeolian feature is, 'parna', an earthy, homogenous, calcareous sediment that mantles the landscape in the east but is absent as a discernible sheet in the west (Butler, 1958). A characteristic of these sediments is 'subplasticity', whereby the texture becomes more clayey when worked (Cattle & Smith, 2018). Where this mantle is present, variability of buried sediments may be masked.

The current understanding of the region's soils is primarily limited to within developed irrigation areas. This is despite the region having a rich history of research extending for over a century. Along with the identification of 94 local soil types, significant agronomically focused soils research has occurred. A bibliography on soils of the region collated by Lang and Hicks (1975) includes references to over 500 reports relating to the region's soils over the preceding 40 years. Despite this breadth of work, much of it has not been digitised and is not publicly available. As a result, there is a paucity of reliable data beyond basic soil descriptions. Hornbuckle and Christen (1999) reviewed historical research from the region through accessing 691 references relating to soils contained within published and unpublished databases. The 94 soil types identified within the east of the area were placed into six soil

groups, with further subgroups where required, primarily based on soil hydrological properties and texture. This work was expanded upon through two further editions (Hornbuckle et al., 2008b; Thacker et al., 2008) incorporating the neighbouring Murray Valley which shares similar characteristics to the south and west of the lower Murrumbidgee valley. Despite this, there are subgroups presented where no soil data was available (Hornbuckle et al., 2008b; Thacker et al., 2008).

This study sought to morphologically describe and then classify 153 soil cores extracted to a depth of 1 m from across the lower Murrumbidgee valley. In doing this, an assessment will be made on if the presently developed regional soil classes are appropriate for soils from across the entirety of the region. The distribution of these classed soils, at both a regional and localised scale will be discussed through the lens of the more widely researched geomorphological features.

4.3. Materials and methods

4.3.1. Soil sampling

This study utilised the 153 soil cores described as the 'full-dataset' in Chapter 3 of this thesis. In brief, at each sampling point soil cores were extracted to a depth of one metre using a vehicle mounted hydraulic corer. Once collected, soil cores were packaged whole and transported to the laboratory for analysis. At each sample site, notes were recorded on the surface condition and surrounding landscape. The coordinates of each core location were recorded using a Garmin GPSMAP 66s handheld GPS.

4.3.2. Morphological description of soils

Once in the laboratory, soil cores were unpacked and morphologically described prior to being dried and ground. This process involved splitting soil cores to reveal a 'more natural', unsmeared face. Each soil core was then horizonated based primarily on soil colour, texture and, where possible, structure. The soil colour was determined on 'moist' soil using the Munsell Color book (Munsell Color, 2010), with the dominant and subdominant (where present) colour recorded for each horizon. Observations were made on the presence and depth of coarse fragments (particles >2 mm), pedogenic segregations such as lime and gypsum and observable signs of waterlogging through mottling. Descriptions of these features and sub-dominant soil

colours followed the guidelines of the 'Australian Soil and Land Survey Field Handbook' (NCST, 2009).

4.3.3. Laboratory analysis of soil samples

The laboratory processes undertaken in this study, including predictions of soil texture using spectroscopy data, are described in chapter 3 of this thesis. Consequently, only a summary is provided here.

Once morphologically described, soil cores were subsampled into five depth increments (0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m) which were air-dried and ground to pass through a 2 mm sieve. Soil Electrical Conductivity (EC) and pH for all samples were measured in a 1:5 soil to water solution using a S230 Conductivity meter and SevenCompactTM s220 pH/Ion meter, respectively. A conversion factor was used to convert soil EC to soil ECe (electrical conductivity of the extract) based on the texture grade of each sample. The exchangeable calcium, magnesium, potassium, sodium and cation exchange capacity (CEC) was determined through the 15C1 method of Rayment and Lyons (2011) for all samples (n=764). Collected data allowed for the determination of the exchangeable sodium percentage (ESP) and calcium to magnesium ration (Ca:Mg).

All samples (n=764) were scanned using a contact probe on an ASF Agrispec spectrophotometer (Malvern Panalytical, Boulder, Co, USA). Data was then processed and K-means clustering used to identify representative subsamples to undertake particle size analysis on. In undertaking this analysis, the pipette method (Glendon & Or, 2002) was used to determine the particle size analysis of 205 samples. Utilising this data alongside CEC data and VisNIR spectra, a cubist model was built to predict sand and clay content for all samples.

4.3.4. Classification of soils within morphological groups

Once morphological and analytical data was obtained, it was assessed to consider how soils could be classed. Upon considering the previous soil groupings of Hornbuckle and Christen (1999) (Table 4.1) alongside the NSW Digital soil map, which classifies approximately 70% of the region as Vertosols (Gray, 2023), it was decided to develop a new, regional classification scheme that focused on clay soils. This was primarily because this study included areas in the west of the region where heavy clays are prolific. In contrast, previous work including surveys (van Dijk & Talsma, 1964; Butler, 1979), focused on the east of the study area, where hillslope

soils or Red-brown earths are dominant. A bifurcating classification key incorporating the findings of this study was created to assist in this process. Each of the soil classes was then classified to the Subgroup level of the Australian Soil Classification (ASC) using the average data for each class (Isbell & NCST, 2021).

Table 4.1. The six broad soil groups first categorised by Hornbuckle and Christen (1999) with descriptions of the topsoil, subsoil and deep subsoil as well as the generalised ASC classification.

	Soil group					
Feature	Self-mulching clays	Hard-setting clays	Red-brown earths	Transitional red-brown earths	Sands over Clay	Deep sands
Topsoil	Clay	Loam-clay	Loam-sandy loam	Loam-clay loam	Sand-loam, cemented	Sand
Depth (m)	0.05-0.15	≤ 0.05	0.1-0.25	≤0.1	0.25-1	>1
Subsoil	Heavy clay, lime	Heavy clay	Heavy clay	Heavy clay	Mottled medium clay	Mottled clay sand – light clay
Deep subsoil	Medium clay, concretionary lime	Medium clay, crystalline gypsum common	Sandy clay, often micaceous	Medium clay, crystalline gypsum common	Medium clay, can be sandy at depth	Light clay
ASC classification	Vertosol	Sodosol	Chromosol	Chromosol	Chromosol	Rudosol

4.3.5. Modelling and mapping of soil groups

The modelling and mapping of soils in this chapter is used only as a tool to visualise the distribution of soil types across the area. A stack of spatial covariates previously extracted on to a 90 m grid through resampling using the 'bilinear' method was assembled (Table 4.2). Covariates were then removed due to collinearity (dose rate, barest earth green and barest earth red), local landscape knowledge (MrVBF) and having an overall negative model impact (slope (%).

Table 4.2. Spatial covariates obtained for use in this analysis prior to the removal of slope (%), MrVBF dose rate, barest earth green and barest earth SWIR 2. Covariates fall within different data categories with the original resolution prior to resampling also shown. Note: 1 arc second corresponds to approximately 30 m.

Data category	Data layer	Data source	Original resolution
Terrain	DEM (m)	Gallant et al. (2011)	1 arc second
	Slope (%)	Gallant & Austin (2012a)	3 arc seconds
	MrRTF	Gallant et al. (2013)	3 arc seconds
	MrVBF	Gallant et al. (2012)	3 arc seconds
	TWI	Gallant & Austin (2012c)	3 arc seconds
Radiometric	Dose rate	Minty et al. (2009)	100 m
	Uranium (ppm)	Minty et al. (2009)	100 m
	Potassium (%)	Minty et al. (2009)	100 m
	Thorium (ppm)	Minty et al. (2009)	100 m
Geological	Silica (%)	Gray et al. (2016)	1:250,000
Satellite	Barest earth blue	Roberts et al., 2019	25
	Barest earth red	Roberts et al., 2019	25
	Barest earth green	Roberts et al., 2019	25
	Barest earth NIR	Roberts et al., 2019	25
	Barest earth SWIR	Roberts et al., 2019	25
	Barest earth SWIR 2	Roberts et al., 2019	25

Utilising the remaining spatial covariates, two random forest classification models were built using the 'randomForest' package within R (Liaw & Wiener, 2002; R Core Team, 2023). The first modelled the nine individual soil classes (exceptions class excluded) while the second placed each of the 10 soil classes (exceptions class included) into four broader groups: Red-brown clays, Grey clays, Earths and Deep sands. Model validation occurred through a leave-one-site-out-cross-validation (LOSOCV) whereby the model is built on n-1 sites before being predicted at the omitted site. The statistics for both models were assessed. The higher quality model was predicted onto a 90 m grid of the study area producing a soil class map.

4.3.6. Localised distribution of soil classes

Utilising available geomorphological and landscape resources from Butler (1958), Langford-Smith (1960), van Dijk and Talsma (1964), Butler et al. (1973), Page and Nanson (1996) and Page et al. (2009) the influence of palaeochannel systems on soil distribution was examined. This involved the identification of two sites located on, or adjacent to, differently characterised

palaeochannel systems. These sites are two of those described in the 'within-field subsets' section of Chapter 3.

In summary, the first site, 'Farm A', is located approximately 18 km northwest of Hay. A paddock was identified which contained a palaeochannel streambed incised into heavy clays. These terminal palaeochannel beds are believed to be associated with the Gum Creek system which is estimated to have been active between 35-25 ka (Page & Nanson, 1996). This streambed is defined as an 'undefined stream trace' with 'broad deposition patterns visible' within a 'scalded plain' by Butler et al. (1973). Eight soil cores were extracted to identify the extent of morphological variability within this 82 ha field.

At the second site, 'Farm C', located 12 km southwest of Darlington Point and 39 km southwest of Griffith, a cluster of soil cores was taken adjacent to the Waddi River Ridge (van Dijk & Talsma, 1964). This ridge is located upstream on the Waddi Reach of the Kerarbury paleochannel system (Page et al., 2009; van Dijk & Talsma, 1964) and can be considered as a source bordering dune developed on the palaeochannel system levee with an elevation approximately 2 m above the surrounding landscape. At a site nearby it was estimated that this reach had a width of 220 m and bankfull discharge (streamload) 8 times greater than the present Murrumbidgee River (Page & Nanson, 1996). At this site, 18 soil cores were extracted from a 94 ha paddock with a further five cores taken on or south of the adjacent Waddi River Ridge.

Utilising this data, schematic transect diagrams to a depth of 1 metre were developed to showcase trends in soil texture and pedogenic features at the within-field scale.

4.4. Results

4.4.1. Creation of soil classes, their characteristics and the classification key

A morphological approach using soil texture and profile descriptions was employed to describe each soil core. Upon an assessment of this data, a new classification with ten soil classes was developed. Soil profiles were initially grouped based on texture before being further differentiated. Once the variety of soils was understood, a classification key was developed to assist with this process, where each profile is placed in first class that it meets the qualifying criteria for (Figure 4.1). This key, and the classes therein, was developed on account of the prolific nature of clay soils within the region. The number of soil classes was decided upon to appropriately recognise morphological variability while not overclassifying the relatively small sample size (n=153).

As well as texture, soil colour and the presence of lime and/or gypsum was used to differentiate classes further. Consequently, each of the examined soil cores was placed within one of 10 classes:

- *Red-Brown clays*; soils with >35% clay throughout the first metre of the soil profile where the dominant colour is red or brown and no lime or gypsum is present,
- *Red-Brown calcic clays*; soils with >35% clay throughout the first metre of the soil profile where the dominant colour is red or brown, lime is present at any depth and gypsum is absent,
- *Red-Brown gypsic clays*; soils with >35% clay throughout the first metre of the soil profile where the dominant colour is red or brown and gypsum is present. Lime may also be observed,
- *Grey clays*; soils with >35% clay throughout the first metre of the soil profile where the dominant colour is grey and no lime or gypsum is present,
- *Grey calcic clays*; soils with >35% clay throughout the first metre of the soil profile where the dominant colour is grey, lime is present at any depth and gypsum is absent,
- *Grey gypsic clays*; soils with >35% clay throughout the first metre of the soil profile where the dominant colour is grey and gypsum is present. Lime may also be observed,
- *Texture contrast soils*; Soils with a sandy or loamy textured surface horizon that is greater than 0.1 m thick showcasing a strong texture contrast with a clear, sharp or abrupt boundary to the underlying clay subsoil,
- *Transitional earth soils*; Soils with a sandy or loamy surface horizon less than 0.1 m deep exhibiting a strong texture contrast with a clear, sharp or abrupt boundary to the underlying clay horizons,
- *Deep sands*; soils with <10% clay throughout the first metre of the profile with the exclusion of the surface soil, and,
- *Exceptions*; these are soils which do not fall within any of the above classes.

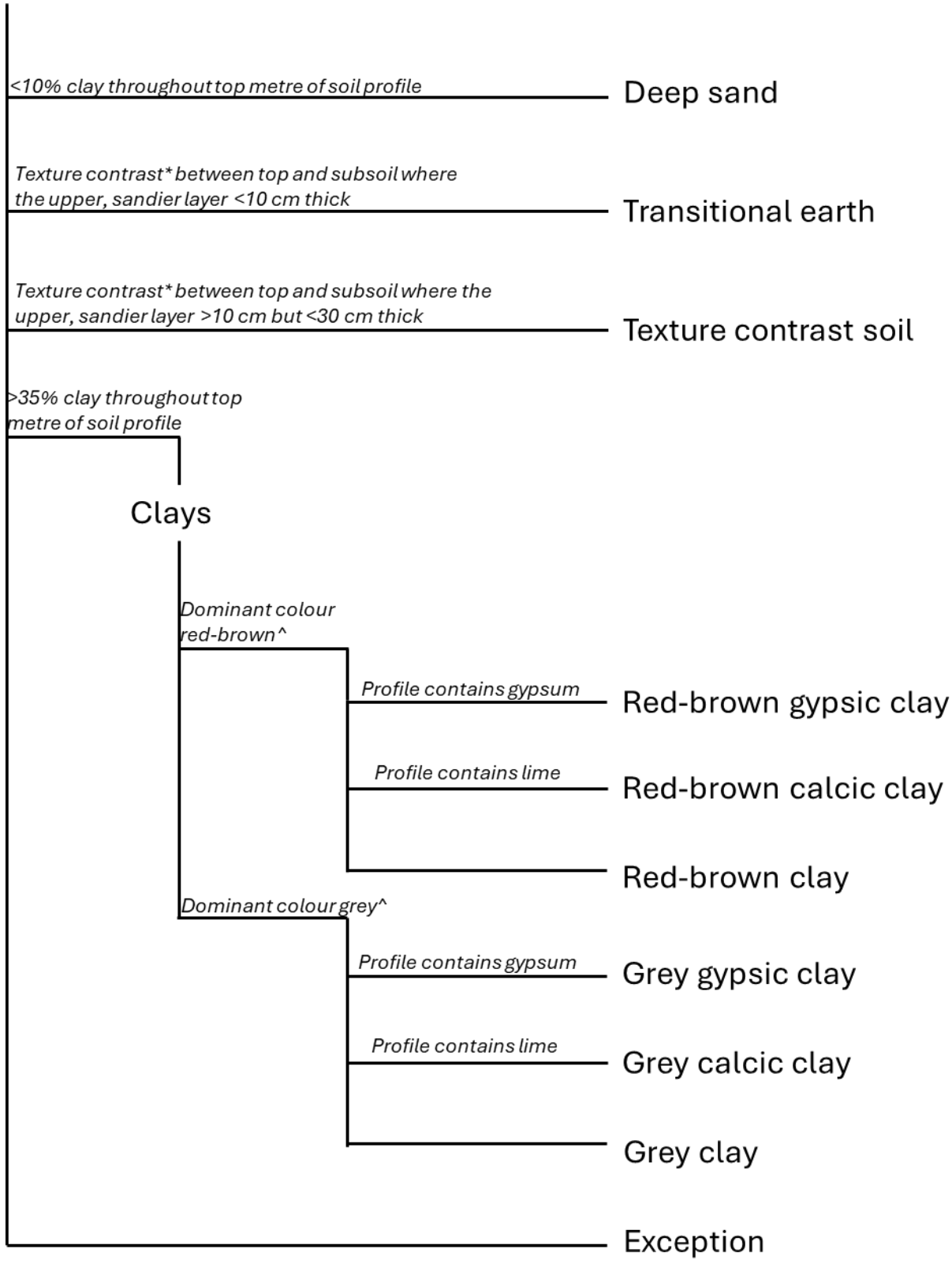


Figure 4.1. The classification key employed to class soil profiles in this study. A soil profile is placed in the first class, e.g. Deep sand, where the criteria is met. *The texture contrast criteria are explained in-text. ^Soil colour was determined using the moist Munsell colour with soils categorised as red, brown or grey based on the Australian Soil Classification (Isbell & NCST, 2021).

4.4.1.1. Deep sands

Deep sand soils were categorised based on containing <10% clay throughout the observed profile. Soils within this group are considered as Arenosols within the ASC (Isbell & NCST, 2021).

4.4.1.2. Transitional earth soils

Soils within the Transitional earth class contain an overlaying sandy or loamy textured horizon <0.1 m thick with a clear, sharp or abrupt boundary to a more clayey underlying horizon (Hornbuckle et al., 2008b; Isbell & NCST, 2021). If the overlaying horizon contains less than 20% clay, the underlying horizon must have a minimum clay content of 20% and be twice as high as the overlying horizon (Isbell & NCST, 2021). Alternatively, if the overlying horizon has a clay content between 20% and 35% the underlying horizon must show an increase of at least 20% clay (Isbell & NCST, 2021). There are no colour requirements for this class.

4.4.1.3. Texture contrast soils

Soils within the Texture contrast class exhibit the same characteristic as the previously described transitional earths, however, for texture contrast soils the overlaying sandy or loamy horizon is between 0.1 m and 0.3 m thick.

4.4.1.1. Clays

Soils were categorised into six classes within the broader clay group as opposed to the two presented by Hornbuckle et al. (2008b). These are; Red-brown clays; Red-brown calcic clays; Red-brown gypsic clays; Grey clays; Grey calcic clays, and Grey gypsic clays. Clay soils within each of these groups were first differentiated from other classes on the basis of their clay content exceeding 35% throughout the first metre of the profile, with the exclusion of the top 50 mm of soil (Isbell & NCST, 2021). The hardsetting clays of Hornbuckle et al. (2008b) may fall within these classes, however, have not been explicitly categorised. Once placed in this broader group, soils were differentiated based on the dominant colour class being grey or red/brown (Isbell & NCST, 2021). Finally, based on the occurrence of lime or gypsum at any point within the one metre soil core, samples were classified as calcic or gypsic (either red-brown or grey), respectively. Where profiles included both lime and gypsum they were classed as gypsic.

4.4.1.2. Exceptions

Soil cores examined which did not fall within the previously described categories have been noted as exceptions in this study. While there are trends observed within these profiles, they do

not occur widely within the region and, thus, do not necessitate a further differentiation from a classification perspective. They have, however, still been described. Soils within this classification may include uniformly textured soils with >10% but <35% clay content throughout the profile, buried soils or soils that become sandier at depth.

4.4.2. Classification of soils within classes and common morphological features

With the exclusion of 28 samples, identified as 'exceptions', all soil profiles were placed within one of the nine pre-defined classes within four broader soil groups (Table 4.3). The majority (n=107) of the 153 soil cores examined were classified as clays with the remainder classed as Texture contrast soils (n=6), Transitional earths (n=9) and Deep sands (n=3). In many instances soils within the exceptions category exhibited similarities to soils within the 'Earths' group, however, there is not sufficient texture differentiation to fall within either the Transitional or Texture contrast classes. With the exclusion of the exceptions, exemplar soil cores for each of the soil classes are shown to a depth of 1 m (Figure 4.5). Each soil class was also classified using the Australian Soil Classification (ASC) to the Subgroup level (Table 4.3) (Isbell & NCST, 2021).

Table 4.3. The nine soil classes, and exceptions, placed within four broader soil groups. Shown is the number (n) of profiles within each class. The ASC classification to Subgroup level is also provided utilising average measurements for each class (Isbell & NCST, 2021).

Soil Group	Class	n	Australian soil classification (ASC) equivalent
Red-brown clays	Brown clays	16	Episodic Epipedal Grey Vertosol
	Red-brown calcic clays	31	Endocalcareous Epipedal Grey Vertosol
	Red-brown gypsic clays	10	Gypsic Epipedal Grey Vertosol
Grey clays	Grey clays	17	Episodic Epipedal Grey Vertosol
	Grey calcic clays	26	Epihypersodic-Endocalcerous Epipedal Grey Vertosol
	Grey gypsic clays	7	Episodic-Gypsic Epipedal Grey Vertosol
Earths	Transitional soils	9	Sodic Dystrophic Brown Chromosol
	Texture contrast soils	6	Dystrophic Subnatric Brown Sodosol
Sands	Deep sands	3	Basic Palic Brown Arenosol
	Exceptions	28	N/A

Within each description of these soil classes the categorisation of attribute values for pH, salinity (ECe), sodicity (ESP) and CEC follow those of Hazleton and Murphy (2007). When describing trends with depth, the 'surface horizon' relates to 0-0.15 m, the 'upper subsoil' 0.15-0.6 m and the 'lower subsoil' 0.6-1 m.

In describing the morphology of each soil core there are commonalities regarding the presence and characteristics of lime, gypsum and subdominant colours.

Lime

Pedogenic lime, that being lime formed in-situ, was observed in three main forms within the examined soil profiles (Figure 4.2). Firstly, lime may be present as veins running vertically through the soil profile along bio- or macro- pores with sharp boundaries against the adjacent soil particles. These veins are most common in the heavier clay soils below 0.5 m. The second observed form of lime was nodular (Figure 4.2), with individual nodules observed up to a width of 15 mm. Where lime is present higher in the soil profile (<0.6 m) nodules are the most common pedogenic form, with smaller nodules generally occurring at shallower depths. The third form of lime noted in this study was diffuse pedogenic segregations. These segregations were observed to only occur in depths >0.7m and are softer than the nodular or vein forms. In many instances soil profiles would contain multiple forms of lime, for instance, smaller nodules higher in the profile (0.5-0.7 m) with diffuse pedogenic, potentially crystalline, carbonates observed below 0.9 m.

Gypsum

Gypsum in clay soils occurred at depths >0.8 m and was crystalline in nature (Figure 4.3). Gypsum was not observed in non-clay soil classes. Individually these lath-like crystals were medium in size, up to 5 mm, with one grouping of these crystals occupying approximately 0.1 m of the profile below 0.9 m in one core. Where both gypsum and lime were observed in the soil profile, the presence of gypsum was generally at a greater depth than lime. There are, however, instances where crystalline gypsum occurs adjacent to diffuse pedogenic segregations of lime as observed in the Red-brown calcic clay profile (Figure 4.5). In this instance, nodular lime was also observed within the soil profile below 0.35 m but above the depth where gypsum was present.

Subdominant soil colour

Two primary forms of subdominant soil colours were observed that were unrelated to mottling (Figure 4.4). In both instances the contrast between the dominant and subdominant colours is prominent with clear or sharp boundaries. The first form of subdominant colours occurs vertically through the soil in connected pore spaces or root channels. The presence of these root channels and their interaction with the surrounding soil particles results in the oxidation of iron, resulting in a reddish colour, along the channel (Figure 4.4). An alternate theory for this

occurrence is that the red colouration is a result of rubified sand grains filling previously open soil channels and pore spaces.

The second instance is where a subdominant colour is observed as a transition between two horizons of different colours (Figure 4.4). Such changes are common in the region, frequently occurring in clay soils at depths >0.7 m. Where soil colour changes from grey to red in the subsoil it suggests the deposition of younger alluvial material atop an older, buried profile. Observed in Figure 4.4 is such a change, whereby the upper horizons are grey in colour with a reddish hue becoming dominant from 0.75 m. More analysis, however, is required to examine the pedological features of the buried soil below 1 m to comment on this further.

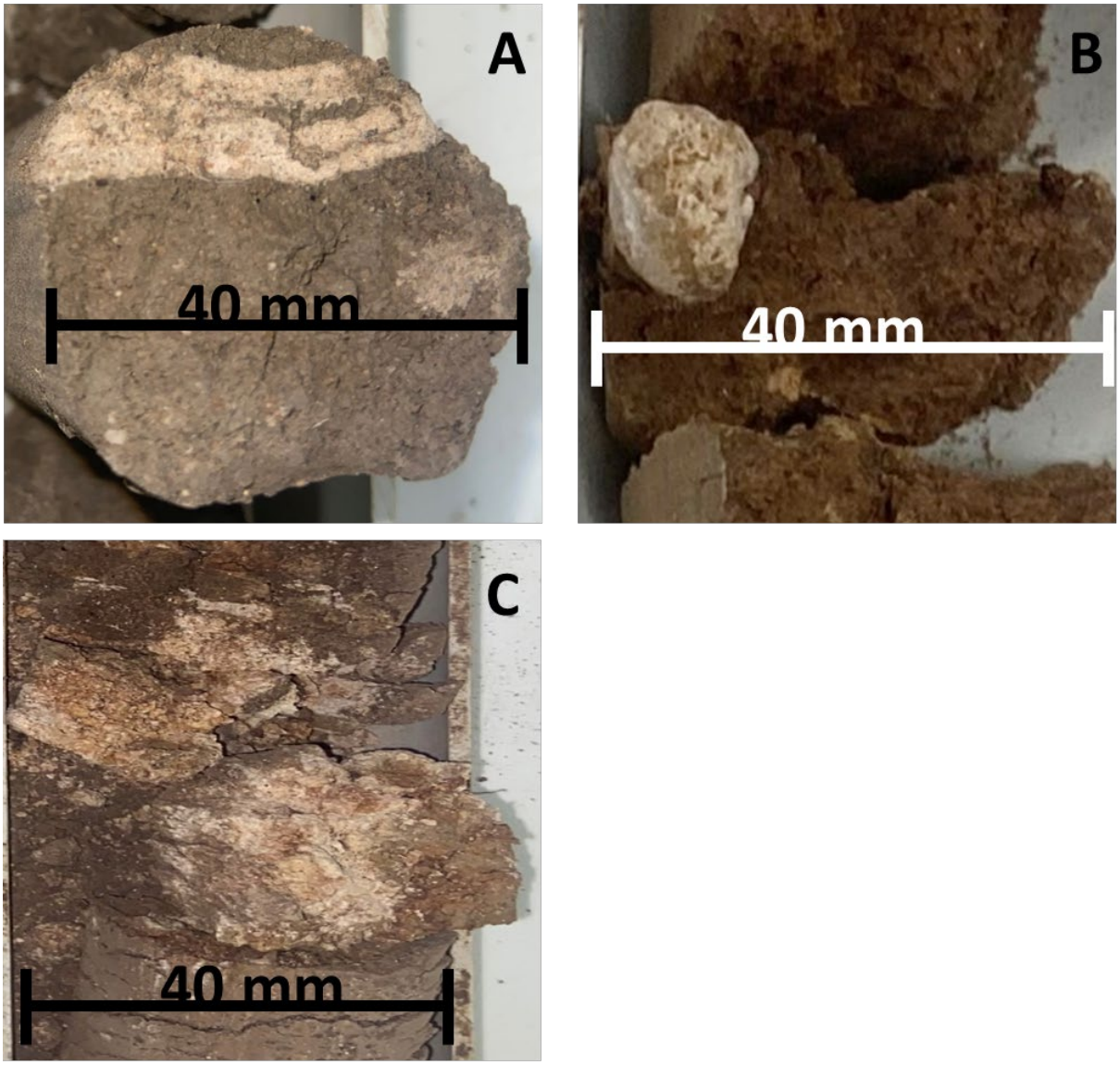


Figure 4.2. Three morphologically distinct occurrences of lime within the soils described in this study. A) Veins of lime throughout the profile in bio- or macro-pores, B) Nodules of lime varying in sizes up to a width of 0.015 m, as pictured in this instance, and, C) Diffuse pedogenic segregations of lime most commonly occurring below 0.7 m.

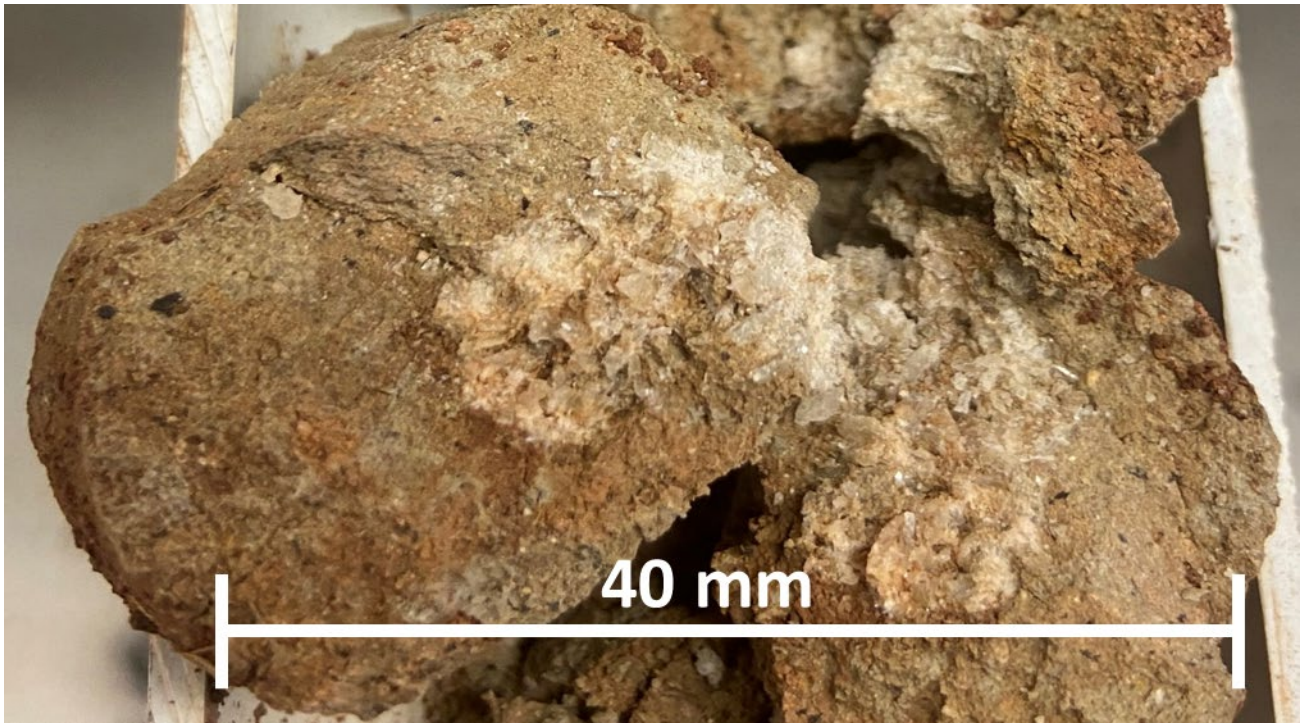


Figure 4.3. An example of crystalline gypsum occurring at a depth of 0.8 m in a Grey clay with a subdominant greyish brown colour also present at depth.

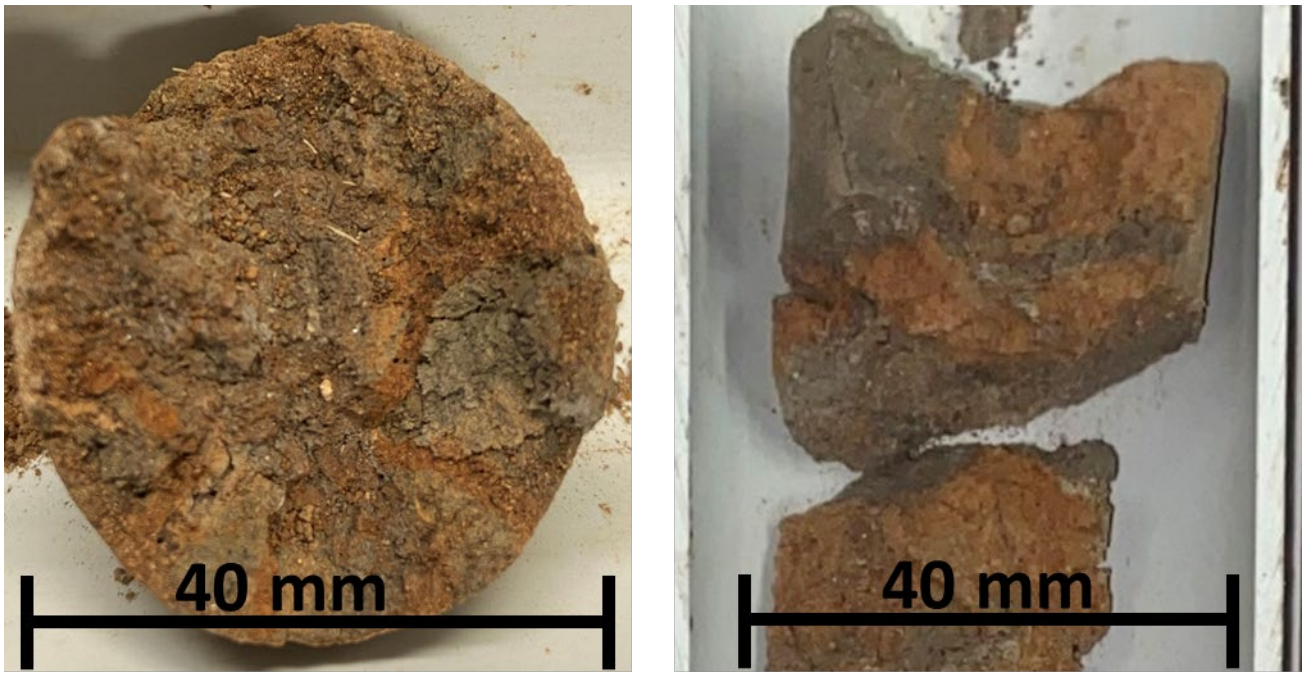


Figure 4.4. Two examples of instances of subdominant soil colours unrelated to waterlogging. Left) a redder hue is present in bio- or macro-pores. Right) There is a sharp contrast between colours proceeding an eventual change. In this instance, the dominant colour of the upper subsoil soil, grey, becomes subdominant.

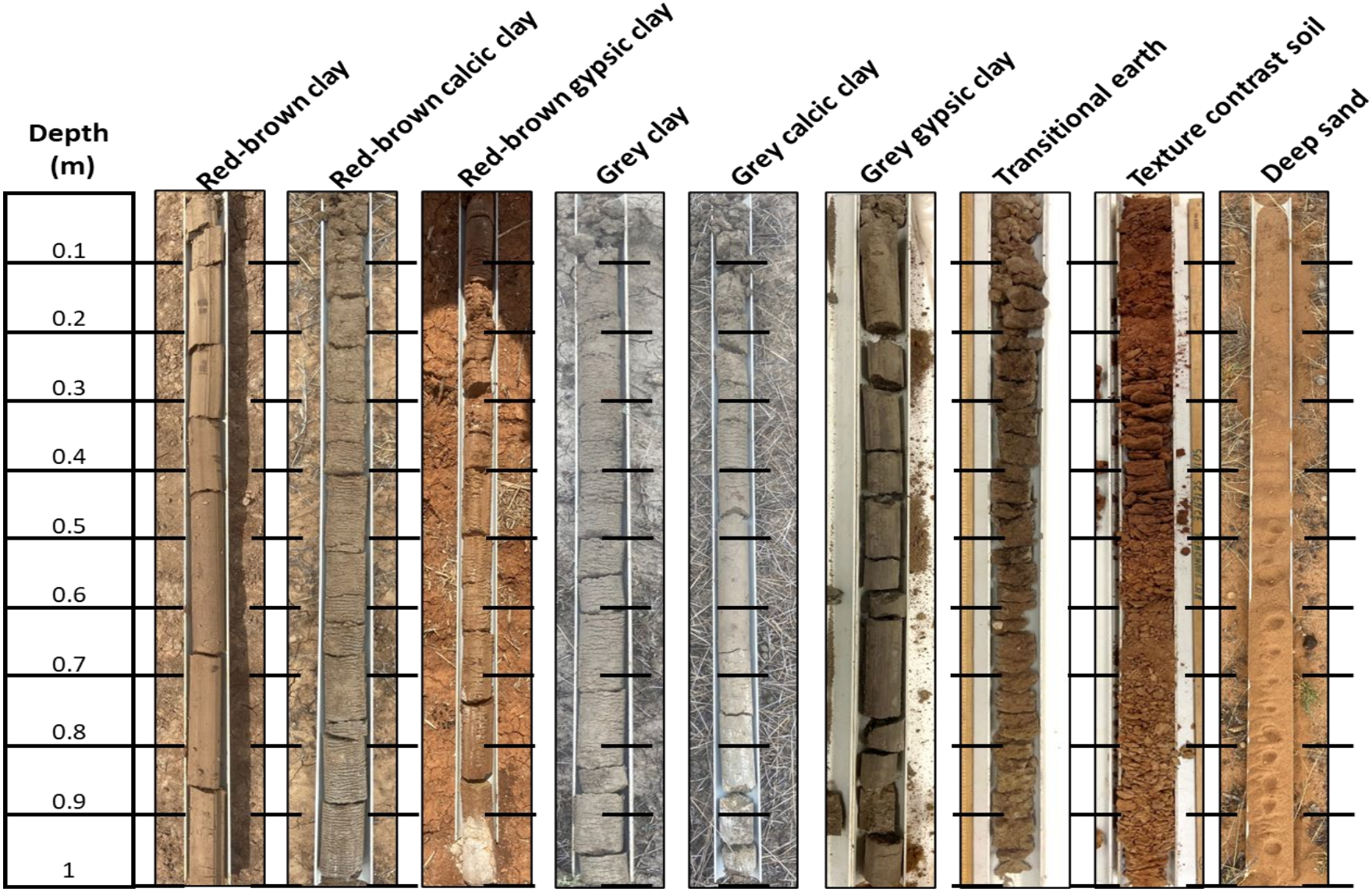


Figure 4.5. Exemplar soil cores to a depth of 1 metre for each of the nine classified soil classes

4.4.3. Clays

Of the 153 profiles examined within this study 107 are classed as clays. Within this grouping, 57 are calcic, either Red-brown or Grey. Clays of both colours occur throughout the region, however, a higher proportion of clays in the west-southwest are grey (Figure 4.6). In contrast, the majority of clays in the region's north are ged-brown (Figure 4.6). There is no clear east-west trend for calcic clays of either colour (Figure 4.6). For all six clay groups, soils are light, medium or heavy clay in texture throughout the profile with the exclusion of the surface 50 mm. There is limited capacity to observe structure within soil cores with a diameter of 40 mm. Consequently, when considering the categorisation of these soils within the Australian Soil Classification (Isbell & NCST, 2021) it is possible some profiles would be considered Dermosols instead of Vertosol. Within each clay class there is significant variability in measured salinity (ECe) and sodicity (ESP).

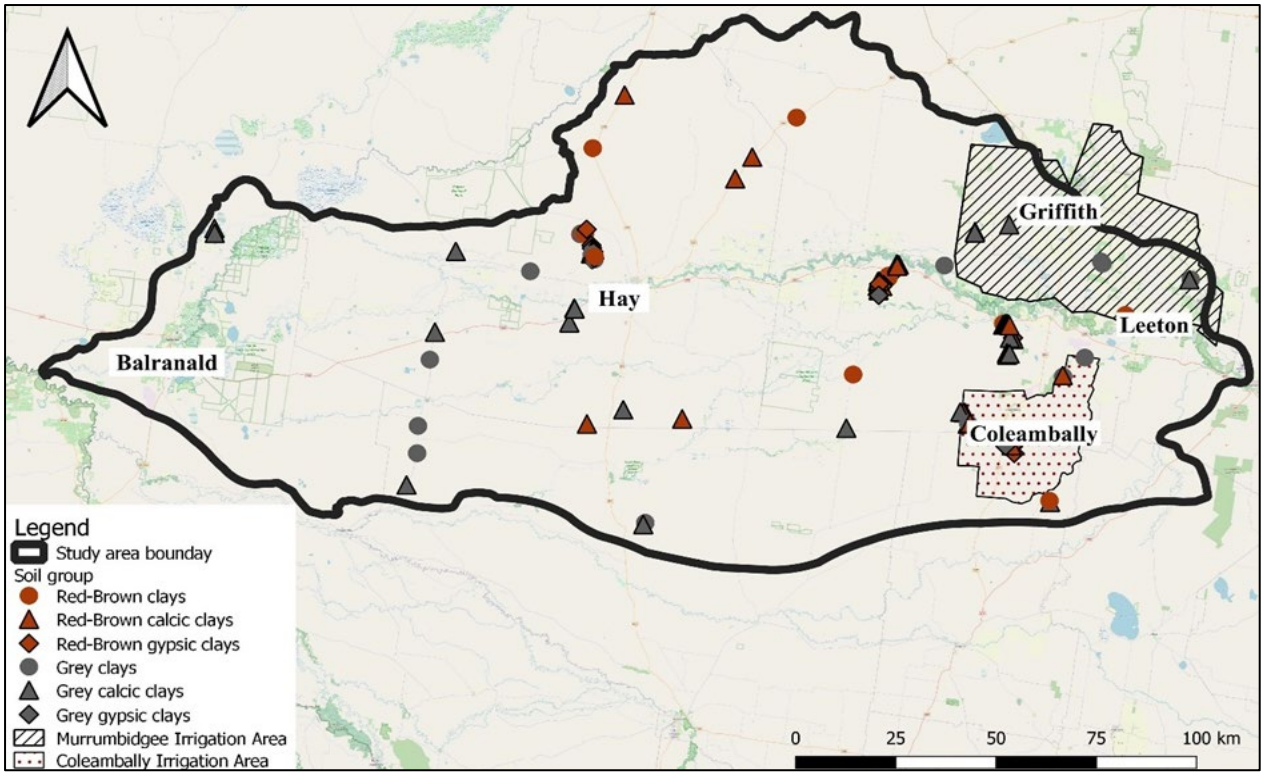


Figure 4.6. The location soil cores classed as clays throughout the study area. Soils are more commonly grey in the west-southwest of the area while north of Hay red-brown clays are most prevalent.

4.4.3.1. Red-brown clays

Sixteen soil profiles were classified as Red-brown clays. While no soils examined in this study were distinctly red, the class has not been reduced to 'Brown clays' on the basis that red coloured clays are likely present more broadly and red horizons were observed. Unlike the

calcic or gypsic Red-brown clays, lime or gypsum is not present in the top 1 m of these soils. Soils within this class occur across the plain, (Figure 4.6) and may exhibit signs of waterlogging (n=7), subdominant colours in the lower subsoil (n=11) and/or coarse fragments (n=5). The observation of coarse fragments is more common in the region's west. Soils in this class are most commonly brown, strong brown or yellowish brown. Red or reddish-brown horizons were also present. In the east, soils may exhibit a greyish-brown surface horizon (<0.1 m). Profiles may exhibit a colour change, becoming greyish brown, olive brown or grey in the lower subsoil (>0.8 m). In these instances, the dominant colour within the upper horizons, brown-strong brown or yellowish brown, is still present as a subdominant colour.

The clay content increases slightly from the surface (mean=47%) to lower subsoil (mean=51%), as does CEC, which ranges from low (6-12 cmol(+)/kg) to high (25-40 cmol(+)/kg). These soils are alkaline, trending from neutral or slightly basic at the surface to strongly basic in the lower subsoil.

There is significant variability in the measured ECe and ESP of the Red-brown clays. High (ECe 8-16) or extreme (ECe>16) salinity is seen in the lower subsoil. Soils may also be strongly sodic (ESP>14) in the upper subsoil as well as the lower subsoil. The occurrence of high or extreme salinity, and strong sodicity, is seen only in Red-brown clays in the west of the region.

Table 4.4. Summary statistics for 16 profiles classified as Red-brown clays. Displayed are the minimum, mean and maximum values for clay content, sand content, pH, ECe, CEC, ESP and the calcium to magnesium ratio (Ca:Mg).

Depth *(m)*		Clay *(%)*	Sand *(%)*	pH	ECe *(dS/m)*	CEC *(cmol(+)/kg)*	ESP *(%)*	Ca:Mg
0-0.1	Min	28	19	6.3	0.70	10	0.8	0.59
	Mean	47	32	7.5	1.93	19	7.1	1.27
	Max	58	57	8.5	5.38	23	26.3	2.02
0.1-0.3	Min	34	22	6.9	0.45	15	1.2	0.63
	Mean	48	29	7.9	4.99	20	11.9	1.15
	Max	58	42	9.0	24.9	26	38.1	2.16
0.3-0.6	Min	34	15	6.6	0.04	16	2.26	0.62
	Mean	51	28	8.3	5.90	23	15.2	1.14
	Max	59	43	8.9	37.2	28	35.6	1.94
0.6-0.8	Min	39	17	7.2	0.04	20	2.28	0.63
	Mean	52	26	8.4	7.65	25	15.8	1.24
	Max	62	40	9.1	49.1	30	35.5	2.41
0.8-1	Min	38	14	7.3	0.02	18	2.36	0.75
	Mean	51	27	8.3	7.76	25	14.9	1.40
	Max	60	38	9.0	37.0	35	32.5	2.38

4.4.3.2. Red-brown calcic clays

Thirty-one soils are classified as Red-brown calcic clays, differentiated from the Red-brown clays due to the presence of lime within the first metre of the soil profile. Where lime occurs at depths <0.6 m, it appears as hard nodules whereas below this depth it may occur in veins, as segregations or as larger nodules. These soils are distributed across the plain, often occurring in close proximity to Red-brown clays, Red-brown gypsic clays and Transitional earths classes (Figure 4.6, Figure 4.8). Soils in this class may exhibit waterlogging (n=6), mottling or subdominant soil colours in the lower subsoil (n=25) and/or coarse fragments (n=5).

The dominant colour of Red-brown calcic clays ranges from brown or dark brown to red or reddish-brown, with variability occurring between horizons throughout the profile. In limited instances, some observed horizons are dark yellowish brown. In the lower subsoil the dominant soil colour may change to become greyish brown or grey. The surface soil may also be greyish brown (n=4) overlaying brown or dark brown horizons.

Red-Brown calcic clays are more basic than the other classes within the Red-brown clay group, with slight alkalinity at the surface that increases with depth (mean = 8.7). Trends in CEC and clay content mirror those described for the Red-brown clay group. Some Red-brown calcic

clays exhibit extreme salinity (ECe>16) from the upper subsoil, becoming more common in the lower subsoil, and strong sodicity (ESP>14) throughout the first metre of the soil profile. Consequently, 'average' soils in this class can be considered marginally sodic to sodic.

Table 4.5. Summary statistics for 31 profiles classified as Red-brown calcic clays. Displayed are the minimum, mean and maximum values for clay content, sand content, pH, ECe, CEC, ESP and the calcium to magnesium ratio (Ca:Mg).

Depth *(m)*		Clay *(%)*	Sand *(%)*	pH	ECe *(dS/m)*	CEC *(cmol(+)/kg)*	ESP *(%)*	Ca:Mg
0-0.1	Min	32	22	6.3	0.72	10	0.6	0.58
	Mean	45	35	7.7	1.40	19	5.9	1.34
	Max	56	52	8.7	3.28	29	17.1	2.25
0.1-0.3	Min	36	24	5.5	0.85	11	0.9	0.58
	Mean	49	33	8.1	2.93	20	9	1.27
	Max	39	48	9.1	16.3	30	31.8	1.98
0.3-0.6	Min	36	15	7.7	0.82	11	0.9	0.50
	Mean	51	29	8.6	4.22	24	12	1.15
	Max	60	49	9.4	24.3	30	37.9	1.87
0.6-0.8	Min	37	18	7.6	0.84	11	0.6	0.75
	Mean	51	28	8.7	8.08	25	12.1	1.38
	Max	61	49	9.7	39.2	39	30	5.66
0.8-1	Min	35	19	7.5	0.83	9.0	0.2	0.67
	Mean	49	29	8.7	8.37	23	12.2	1.38
	Max	57	50	9.8	39.6	32	28.6	2.27

4.4.3.3. Red-brown gypsic clays

Ten examined soil profiles are classified as Red-brown gypsic clays. As well as exhibiting gypsum, eight of these profiles contain lime in the subsoil. Where present, lime most commonly occurs as nodules rather than pedogenic segregations and is at a shallower depths than the gypsum. The lath-like crystals of gypsum were observed individually and in grouped segregations, the latter more common in the lower subsoil. Soils within this class are distributed across the plain and often occur in association with Red-brown calcic clays (Figure 4.6). Within this class, profiles can also contain coarse fragments at any depth (n=4) while mottling or subdominant colours in the lower subsoil are common, the former suggesting susceptibility to waterlogging.

Red-brown gypsic clays are brown, dark brown, strong brown, reddish-brown or red throughout the profile. The dominant soil colour frequently changes between horizons. In the

lower subsoil light yellowish brown, grey or greyish brown mottles, or subdominant colours, are common.

Soils profiles within this class are neutral-to-slightly alkaline at the surface and are most basic in the upper subsoil (mean=8.7). Similar to the previously described Red-brown clay classes, profiles may be strongly sodic (ESP>14) and are most sodic in the upper subsoil (mean =15.6) or from 0.6-0.8 m (mean=16.3). One profile, located in the west of the region, exhibits extreme salinity (ECe>16) in the lower subsoil. Both clay content and CEC increase from the surface to depth, with three profiles exhibiting a very high CEC (>40 cmol(+)/kg) in the lower subsoil.

Table 4.6. Summary statistics for 10 profiles classified as Red-brown gypsic clays. Displayed are the minimum, mean and maximum values for clay content, sand content, pH, ECe, CEC, ESP and the calcium to magnesium ratio (Ca:Mg).

Depth *(m)*		Clay *(%)*	Sand *(%)*	pH	ECe *(dS/m)*	CEC *(cmol(+)/kg)*	ESP *(%)*	Ca:Mg
0-0.1	Min	35	27	6.5	0.62	12	1.0	0.50
	Mean	47	32	7.4	1.30	20	4.5	1.16
	Max	56	50	8.4	2.04	26	7.8	1.69
0.1-0.3	Min	47	21	7.2	0.74	20	3.0	0.54
	Mean	52	26	8.3	2.10	24	10.2	1.10
	Max	56	37	9.0	5.97	27	18.5	1.83
0.3-0.6	Min	45	12	7.1	0.78	18	7.2	0.58
	Mean	51	25	8.7	4.19	24	15.6	1.00
	Max	56	34	9.3	15.0	28	26.9	1.34
0.6-0.8	Min	44	16	7.1	1.05	21	8.3	0.55
	Mean	53	25	8.5	6.80	25	16.3	1.18
	Max	59	33	9.3	28.4	31	31.8	1.97
0.8-1	Min	42	1	6.2	1.12	21	3.4	0.93
	Mean	54	20	8.0	12.5	33	11.7	2.22
	Max	63	27	9.4	34.5	53	22.9	5.12

4.4.3.4. Grey clays

Seventeen soil profiles are classified as Grey clays on account of their grey colour and the absence of lime or gypsum at any point within the first metre of the soil profile. Soils in this class occur across the region, often in close association with grey calcic or gypsic clays as well as in a mosaic with all classes of Red-brown clays (Figure 4.6). Coarse fragments were observed in one profile, while mottling or subdominant colours occurred in four profiles.

The most common dominant colours were grey, dark grey or dark greyish brown with these colours often occurring in association with each other. Two profiles exhibited a change to light olive brown or light yellowish brown below 90 cm in the lower subsoil, while one profile had a greyish brown surface soil.

Grey clays showcase minimal variability in clay content from the surface (mean=49%) to lower subsoil (mean=53%) while mean CEC (cmol(+)/kg) increased from 19 to 26 across these depths (Table 4.7). There is variability in the pH of these profiles, with a slight trend from neutral-to-mildly alkaline at the surface (mean=7.4) to moderately alkaline in the lower subsoil (mean=8.5) (Table 4.7). Two profiles, however, are moderately acidic (5.6-6.0) to neutral (6.6-7.3) throughout the profile. Strong sodicity (ESP>14) is common in grey clays, with a maximum ESP of 48.3% observed in the lower subsoil of a profile in the region's west (Table 4.7). Grey clays may be extreme saline (ECe>16) in the lower subsoil.

Table 4.7. Summary statistics for 17 profiles classified as Grey clays. Displayed are the minimum, mean and maximum values for clay content, sand content, pH, ECe, CEC, ESP and the calcium to magnesium ratio (Ca:Mg).

Depth *(m)*		Clay *(%)*	Sand *(%)*	pH	ECe *(dS/m)*	CEC *(cmol(+)/kg)*	ESP *(%)*	Ca:Mg
0-0.1	Min	32	23	6.0	0.30	12	0.9	0.44
	Mean	49	32	7.4	1.83	19	6.6	1.40
	Max	61	51	9.0	9.88	25	23.8	2.63
0.1-0.3	Min	45	17	6.8	0.02	15	0.9	0.39
	Mean	54	27	7.8	3.02	23	12.5	1.22
	Max	64	43	9.7	20.0	30	34.6	2.03
0.3-0.6	Min	43	12	6.3	0.26	13	1.0	0.36
	Mean	54	26	8.1	7.42	24	17.5	1.15
	Max	63	38	9.9	37.1	31	40.0	2.07
0.6-0.8	Min	39	17	6.1	0.28	13	0.9	0.32
	Mean	54	26	8.2	8.82	25	21.4	1.02
	Max	64	44	9.5	37.7	31	45.5	1.98
0.8-1	Min	42	18	6.8	0.71	13	2.2	0.34
	Mean	53	25	8.1	12.2	26	21.0	1.37
	Max	63	38	9.5	35.0	43	48.3	5.00

4.4.3.5. Grey calcic clays

A total of 26 profiles are classified as Grey calcic clays, differentiated from the Grey clay class based on the presence of lime, and from Grey gypsic clays on the absence of gypsum in the top

metre of the profile. Lime can be present in the upper subsoil but is more prevalent in the lower subsoil. When occurring higher in the profile, the most common form of lime is as hardened nodules. As depth increases, lime occurring in veins or as pedogenic segregations is more common. Grey calcic clays may contain coarse fragments (n=10), subdominant soil colours in the lower subsoil (n=16) and/or signs of waterlogging (n=7).

The most common colour throughout the soil profile is grey, with combinations of greyish brown, dark grey or dark greyish brown noted. One profile exhibited a greyish brown surface soil while two profiles exhibited a light olive brown colour in the lower subsoil.

A similar trend of moderate increases in clay content and CEC, as has already been described for Grey clays, was observed. One soil profile located on undisturbed land near Balranald, however, contains a significantly higher CEC in the lower subsoil (mean = 68.5 cmol(+)/kg) (Table 4.8). Soils within this class trend from being neutral or slightly acidic at the surface (mean pH=7.4) to moderately or strongly alkaline in the lower subsoil (mean pH=8.5) (Table 4.8). Strong sodicity (ESP>14) is common within and below the upper subsoil (Table 4.8), with a greater concentration of sodic soils in the region's west. While average salinity (ECe) is below the extreme threshold (ECe>16) at all depths, five profiles exceed this limit at some point. Each of these profiles are found in locations that have never been cleared or irrigated.

Table 4.8. Summary statistics for 26 profiles classified as Grey calcic clays. Displayed are the minimum, mean and maximum values for clay content, sand content, pH, ECe, CEC, ESP and the calcium to magnesium ratio (Ca:Mg).

Depth *(m)*		Clay *(%)*	Sand *(%)*	pH	ECe *(dS/m)*	CEC *(cmol(+)/kg)*	ESP *(%)*	Ca:Mg
0-0.1	Min	19	22	6.4	0.55	8.0	0.6	0.33
	Mean	47	34	7.4	2.05	18	5.8	1.34
	Max	60	59	8.8	13.1	24	19.1	3.33
0.1-0.3	Min	36	18	6.1	0.55	10	0.9	0.33
	Mean	50	31	7.8	3.22	20	10.6	1.30
	Max	63	63	9.0	25.1	29	34.5	4.23
0.3-0.6	Min	33	18	6.9	0.77	11	0.8	0.29
	Mean	52	28	8.3	5.60	22	15.6	1.20
	Max	62	47	9.5	29.4	31	45.2	3.19
0.6-0.8	Min	37	17	7.0	1.12	16	1.0	0.37
	Mean	53	27	8.5	7.75	25	17.2	1.41
	Max	63	41	9.9	37.1	73	43.6	7.32
0.8-1	Min	38	10	7.1	0.66	14	1.1	0.76
	Mean	52	27	8.5	10.5	27	15.3	1.82
	Max	62	52	10	42.1	64	35.9	8.73

4.4.3.6. Grey gypsic clays

Seven profiles were classified as Grey gypsic clays due to the presence of gypsum in the first metre of the profile, with six of these also containing lime. Gypsum was present as either individual lath-like crystals, more common above 0.8 m depth, or as segregated groupings of crystals below 0.8 m. Observed lime was nodular, with nodule size commonly increasing with depth. Within this class coarse fragments (n=4), subdominant colours (n=4) and mottling suggesting waterlogging (n=2) were observed. Five of the seven soils in this class are located among a cluster of sites approximately 15 km northwest of Hay (Figure 4.6).

All profiles have a dominant colour of grey or dark grey, with greyish brown and dark greyish brown potentially occurring as subdominant colours. Two profiles exhibit brown horizons in the lower subsoil (>0.8 m) where grey or dark grey transitions to become the subdominant colour.

Excluding a slight increase from the surface (mean=52%) to upper subsoil (mean=58%) there is minimal change in clay content throughout the profile (Table 4.9). Despite this, there is an increase in the CEC from the surface (mean=22 cmol(+)/kg) to the lower subsoil (mean=37 cmol(+)/kg). Soils show, on average, a neutral (6.6-7.3) to mildly alkaline (7.4-7.8) pH

throughout the profile (Table 4.9). All profiles exhibit extreme salinity (ECe>16) in the lower subsoil (Table 4.9). Five profiles also surpass the threshold for strong sodicity (ESP>14) with sodicity being greatest in the upper subsoil (mean=19.5).

Table 4.9. Summary statistics for seven profiles classified as Grey gypsic clays. Displayed are the minimum, mean and maximum values for clay content (%), sand content (%), pH, ECe (dS/m), CEC (cmol(+)/kg), ESP (%) and the calcium to magnesium ratio (Ca:Mg).

Depth *(m)*		Clay *(%)*	Sand *(%)*	pH	ECe *(dS/m)*	CEC *(cmol(+)/kg)*	ESP *(%)*	Ca:Mg
0-0.1	Min	44	20	7.1	0.85	18	2.7	0.61
	Mean	52	29	7.3	1.58	22	7.6	0.92
	Max	65	34	7.5	2.23	25	13.9	1.10
0.1-0.3	Min	50	19	7.3	0.49	21	4.8	0.54
	Mean	58	25	7.9	3.20	25	15.7	0.93
	Max	70	32	8.8	5.73	28	23.8	1.43
0.3-0.6	Min	52	21	6.7	1.53	22	7.9	0.55
	Mean	57	25	7.8	10.7	26	19.5	1.15
	Max	62	29	8.8	25.7	34	32.5	1.76
0.6-0.8	Min	54	16	6.3	3.35	26	6.9	0.52
	Mean	58	21	7.2	20.00	33	14.0	2.14
	Max	64	27	7.9	24.3	41	34.5	4.13
0.8-1	Min	50	15	6.1	19.6	25	4.8	0.91
	Mean	58	19	7	24.7	37	12.2	2.72
	Max	70	30	7.8	29.9	48	25.9	5.47

4.4.4. Texture contrast soils

The six soils within this class exhibit a clear, sharp or abrupt contrast between the lighter textured surface and more clayey underlying subsoil horizons (Table 4.10). These soils appear primarily in the east of the region (n=5) with one profile located southwest of Hay (Figure 4.7). Texture contrast soils may contain nodular lime in the lower subsoil (n=4) and exhibit subdominant colours or mottling (n=4). No samples contained observable coarse fragments or showed clear signs of waterlogging.

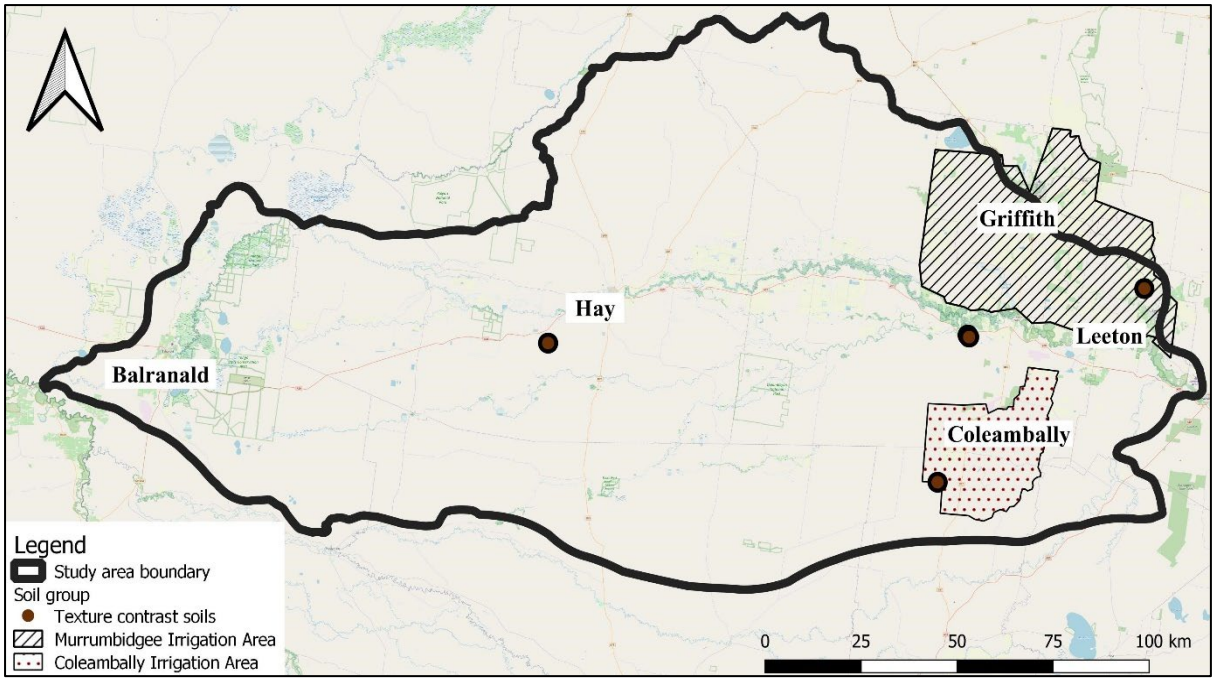


Figure 4.7. The location of the six cores classified as Texture contrast soils in this study.

There is variability in the colour of the texture contrast soils. Three profiles are brown, yellowish dark red or yellowish red throughout the profile. Two soils contain a grey, lighter surface soil overlaying the heavier, brown subsoil while one profile is grey throughout.

Clay content increases from an average of 22% (0-0.3 m) to an average of 39% (0.3-0.6 m), 46% (0.6-0.8 m) and 42% (0.8-1 m) (Table 4.10). Soil pH and CEC values have a similar trend, with averages for both soil properties increasing from 5.6 and 7 cmol(+)/kg at the surface to 8.9 and 18 cmol(+)/kg in the lower subsoil, respectively (Table 4.10). One profile is sodic below 30 cm (ESP>14), however, did not exhibit a columnar structure as required for it to be differentiated as a sand over clay soil per Hornbuckle et al. (2008b).

Table 4.10. Summary statistics for six profiles classified as Texture contrast soils. Displayed are the minimum, mean and maximum values for clay content, sand content, pH, ECe, CEC, ESP and the calcium to magnesium ratio (Ca:Mg).

Depth *(m)*		Clay *(%)*	Sand *(%)*	pH	ECe *(dS/m)*	CEC *(cmol(+)/kg)*	ESP *(%)*	Ca:Mg
0-0.1	Min	6	44	6.2	0.83	2	2.1	1.35
	Mean	22	62	5.6	1.09	7	3.8	2.56
	Max	31	87	6.9	1.35	10	5.4	6.30
0.1-0.3	Min	12	48	5.9	0.35	2	1.9	0.94
	Mean	22	64	6.8	1.02	6	5.2	1.97
	Max	30	89	7.5	1.69	11	7.0	3.30
0.3-0.6	Min	21	31	7.8	0.46	4	5.1	0.71
	Mean	39	45	8.2	1.62	15	6.4	1.43
	Max	51	67	9.0	2.76	24	18.4	2.57
0.6-0.8	Min	43	25	8.0	0.86	14	5.2	0.79
	Mean	46	35	8.7	2.63	20	11.5	1.40
	Max	50	43	9.2	5.44	31	22.7	2.83
0.8-1	Min	35	30	8.2	1.67	10	2.1	0.70
	Mean	42	41	8.9	2.73	18	9.9	1.62
	Max	52	52	9.3	4.29	27	23.2	3.27

4.4.5. Transitional earth soils

Nine soils are classified as Transitional earths, observed primarily in the east of the region, with one soil located north of Hay (Figure 4.8). Transitional earths may show evidence of mottling or subdominant soil colours (n=8), nodular lime in the subsoil (n=5) and/or coarse fragments (n=3). Soils within this class are most commonly dark brown to yellowish red throughout the profile. In some instances, a dark greyish brown horizon may overlay a dark brown or yellowish red subsoil.

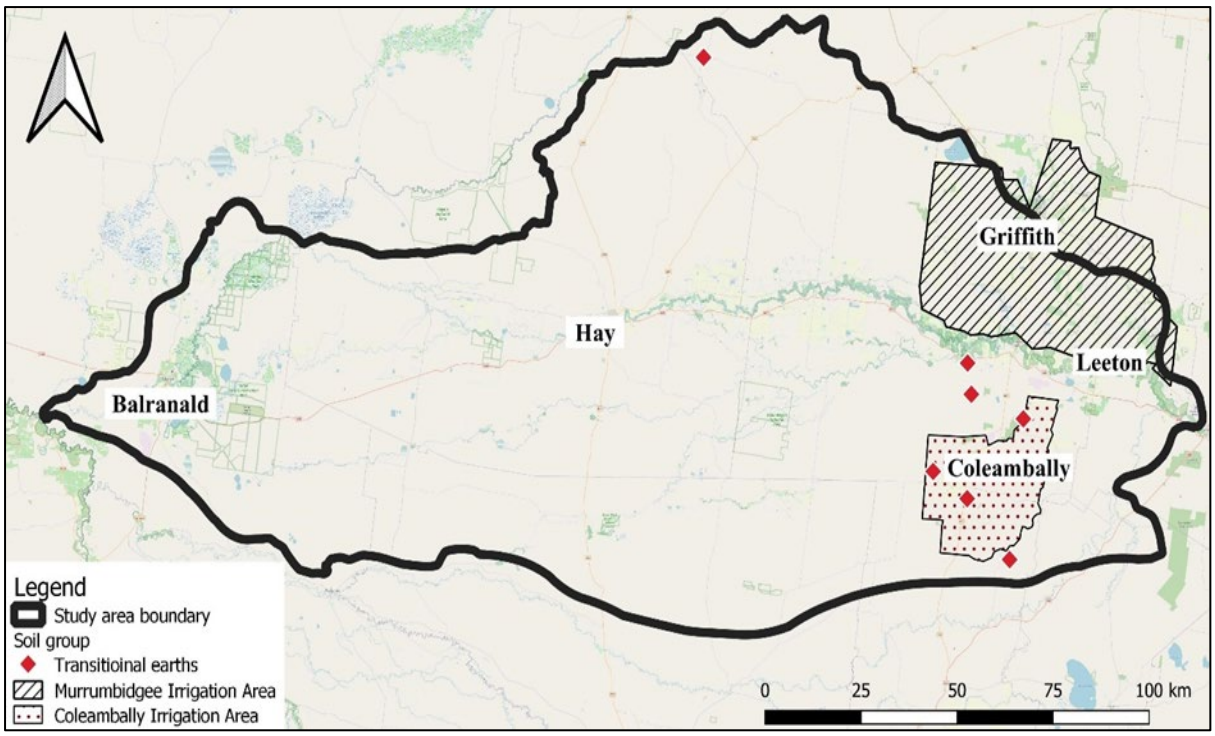


Figure 4.8. The locations of the nine sites where the soil profiles were classified as transitional earths.

There is an increase in clay content from the surface (mean = 26%) to 0.1-0.3 m (mean = 48%) which is mirrored by the CEC which increases from 8 cmol(+)/kg (0-0.1 m) to 17 cmol(+)/kg (0.1-0.3 m) (Table 4.11). The pH of soils increases from neutral at the surface (mean=6.7) to strongly alkaline in the lower subsoil (mean=8.6) (Table 4.11). The Transitional earth located north of Hay is strongly sodic (ESP>14) and extremely saline (ECe>16) from the upper subsoil to depth (Table 4.11). 'Average' Transitional earth soils, however, are not considered to be extremely saline or strongly sodic.

Table 4.11. Summary statistics for seven profiles classified as Transitional earths. Displayed are the minimum, mean and maximum values for clay content, sand content, pH, ECe, CEC, ESP and the calcium to magnesium ratio (Ca:Mg).

Depth *(m)*		Clay *(%)*	Sand *(%)*	pH	ECe *(dS/m)*	CEC *(cmol(+)/kg)*	ESP *(%)*	Ca:Mg
0-0.1	Min	17	44	5.9	0.29	3	0.6	0.76
	Mean	26	54	6.7	1.03	8	2.4	1.54
	Max	34	67	8	1.52	13	6.4	2.32
0.1-0.3	Min	37	24	6.6	0.26	13	0.9	0.69
	Mean	48	34	7.6	1.27	17	5.7	1.26
	Max	57	47	9.6	3.42	23	21.1	2.33
0.3-0.6	Min	35	21	7.7	0.80	13	1.6	0.68
	Mean	49	31	8.5	3.17	21	8.0	1.15
	Max	57	43	9.3	14.9	26	33.5	2.42
0.6-0.8	Min	40	20	8.2	0.60	12	1.7	0.79
	Mean	52	28	8.6	5.83	23	7.5	1.33
	Max	59	45	8.9	35.4	28	18.3	3.57
0.8-1	Min	40	23	7.9	0.56	13	2.1	0.79
	Mean	51	31	8.6	7.64	23	9.0	1.26
	Max	57	41	8.9	36.2	33	26.0	2.04

4.4.6. Deep sands

Three soil profiles were classified as Deep sands, containing less than 10% clay content throughout the first metre of the soil profile (Table 4.12). The sand content of these profiles increases slightly with depth while the pH is neutral to slightly basic (Table 3.4). All profiles have a very low CEC (<3.6 cmol(+)/kg) and were a brown-dark yellowish brown throughout the profile (Table 4.12).

Table 4.12. Summary statistics for three profiles classified as Deep sands. Displayed are the minimum, mean and maximum values for clay content, sand content, pH, ECe, CEC, ESP and the calcium to magnesium ratio (Ca:Mg).

Depth *(m)*		Clay *(%)*	Sand *(%)*	pH	ECe *(dS/m)*	CEC *(cmol(+)/kg)*	ESP *(%)*	Ca:Mg
0-0.1	Min	4.3	76	6.8	1.16	1.8	1.1	1.43
	Mean	7.3	84	7.1	1.56	2.6	7.2	1.55
	Max	9.7	90	7.4	1.56	3.6	12.2	1.73
0.1-0.3	Min	1.2	76	6.9	0.93	1.5	6.3	1.08
	Mean	4.7	86	7.2	1.06	2.0	8.1	1.65
	Max	9.2	94	7.5	1.29	3.0	9.3	2.36
0.3-0.6	Min	1.3	90	7.3	0.66	0.9	2.3	1.33
	Mean	3.3	90	7.4	0.87	1.2	9.6	1.56
	Max	7.3	90	7.6	1.19	1.6	13.8	1.90
0.6-0.8	Min	3.4	90	7.6	0.48	0.8	2.9	1.24
	Mean	4.9	90	7.6	0.73	1.0	4.7	2.01
	Max	6.9	90	7.7	0.86	1.5	6.7	3.17
0.8-1	Min	1.5	88	7.0	0.66	0.9	2.0	1.25
	Mean	3.4	89	7.3	0.74	1.1	6.6	1.64
	Max	6.4	91	7.6	0.86	1.3	15.4	2.10

These soils are associated with palaeochannel systems, occurring on source-bordering sand dunes or within relic streambeds. Of the three cores examined within this class, two are considered streambed sands, located northwest of Hay, while one is a mid-slope dune soil located south of Hay (Figure 4.9). While coarse fragments were present throughout the profile of the streambed sands, they were not observed in the dune sand. The dune sand also contained a higher ESP than the two streambed sands.

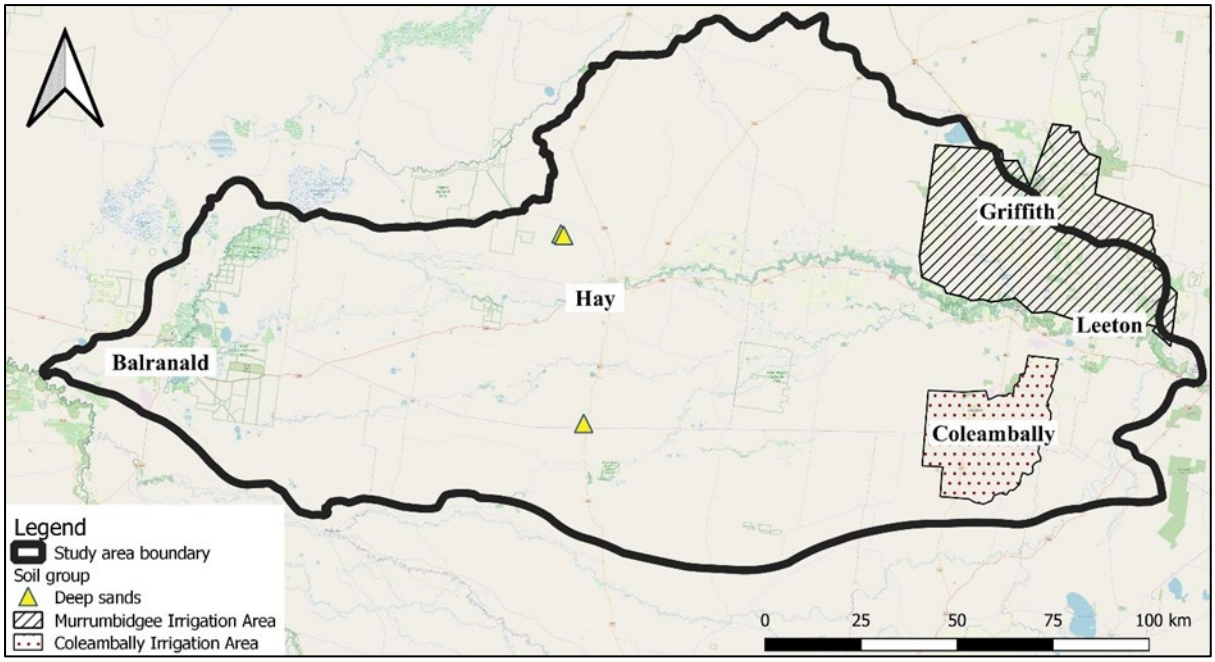


Figure 4.9. The location of the three sites where soils were classified as deep sands.

4.4.7. Exceptions

The 28 soil profiles that did not fall within one of the previous nine classes can be differentiated into a further four categories which do not occur widely over the region. The first of these are 'lakebed adjacent' soils. Four soil cores fit within this category. All profiles are a loamy texture to a depth of 1 metre, vary in colour, from brown to grey, and are highly alkaline (pH>8.8) at all depths. These soils were observed north-west of Balranald and are located adjacent to seasonally flooded lakes (Figure 4.10).

The second category of soils are 'Dunefield soils' located south of Balranald (Figure 4.10). This region is the interface of the lower Murrumbidgee plains and the Mallee dunefields. These soils are loamy, calcareous earths, however, texture varies at depth with the soil most clayey at the surface. This indicates the potential presence of buried horizons. Of the three profiles observed, one exhibited a sand lens at 0.38 m (Figure *4.11*).

The third category are 'Loam' or 'Gradational loam' soils. These soils are similar to those within the Texture contrast and Transitional earth classes, however, do not make the criteria due to gradual changes in texture. Consequently, these soils are observed across the plain in similar locations to the Texture contrast (Figure 4.7) and Transitional earth (Figure 4.8) classes. They are most commonly brown-to-brownish grey in colour with a loamy texture that may become clayey at depth. Lime is common in the lower subsoil.

Soils that become sandier at depth fall within the fourth category which is referred to as 'unclassified'. These soils are clayey at the surface, becoming sandier to a depth of 1 metre or are loamy at the surface, becoming clayey in the upper subsoil before becoming sandier again in the lower subsoil.

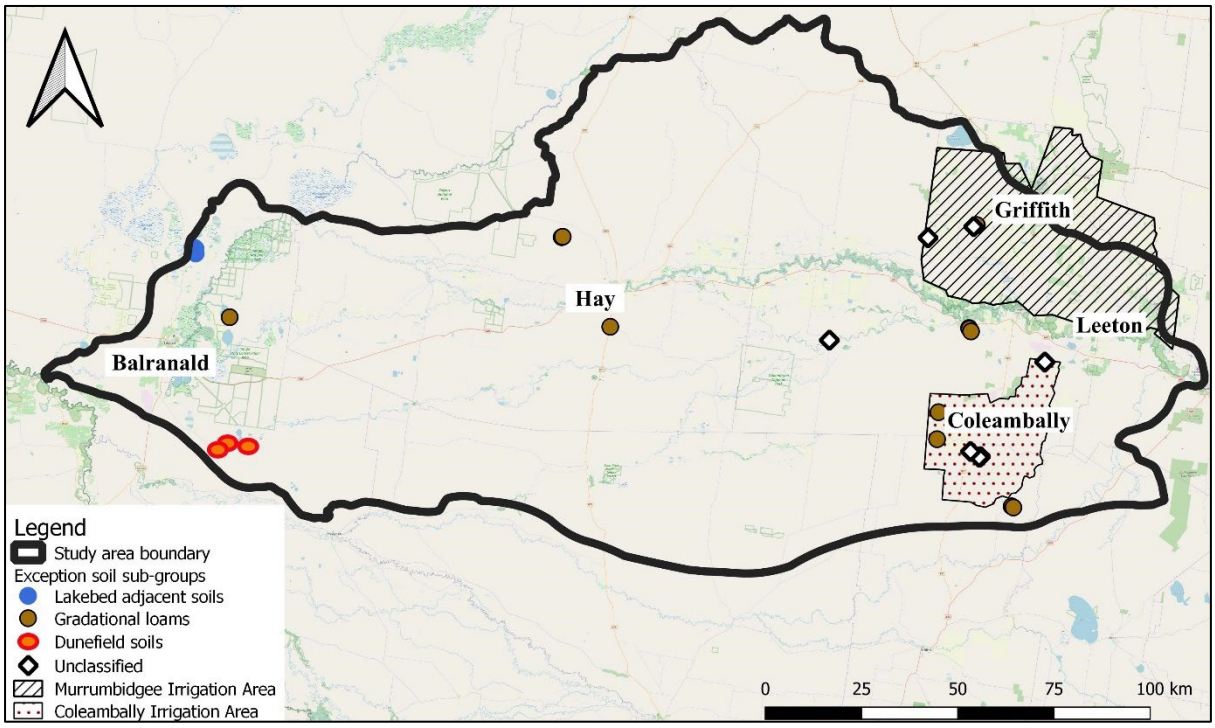


Figure 4.10. The location of the 28 sample sites that did not fall within one of the nine soil classes and, are thus, considered as exceptions. Shown are the four sub-categories of lakebed adjacent soils, gradational loams, dunefield soils and unclassified,

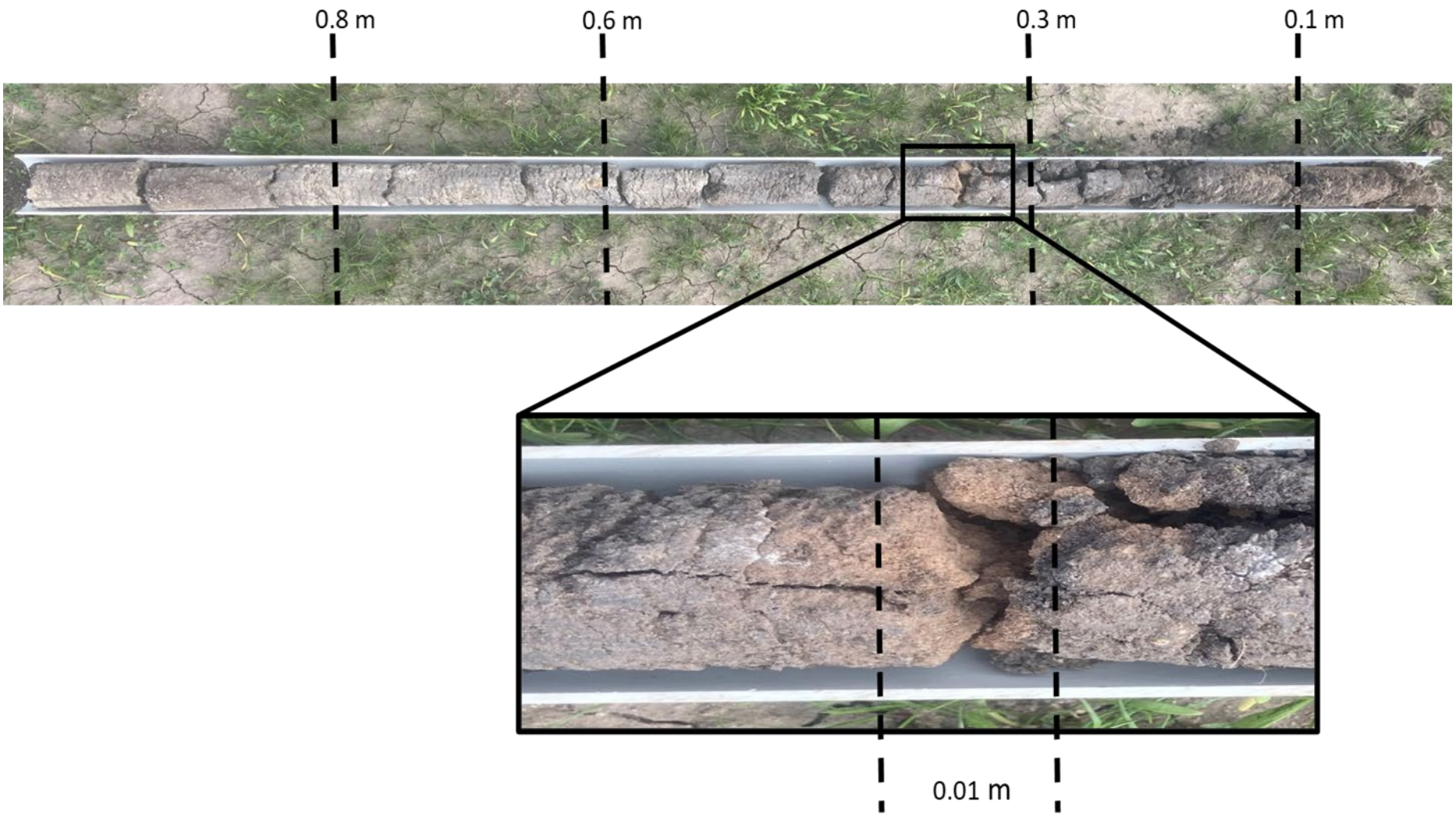


Figure 4.11. An exemplar profile of the calcareous ‘dunefield’ soils which occur at the interface of the lower Murrumbidgee region and Mallee regions. In these soils lime is common below 0.2 m depth. Present in this core is a lens of sand approximately 10 mm thick at a depth of 0.38 m.

4.4.8. Soil class models and maps

4.4.8.1. Model selection

Through a LOSOCV it was shown that the generalised model, with four broader groups, outperformed the more specific soil class model with out of bag estimates of error rates of 39.5% and 64.83%, respectively. In this model the highest error rate was for the Sands group, with the three profiles predicted as earths (Table 4.13). Red-brown and Grey clays produced the lowest class error with rates of 0.33 and 0.36, respectively (Table 4.13).

Table 4.13. The breakdown of predictions of the four broader soil groups modelled in this study. In this model texture contrast and transitional earth soils, as well as appropriately selected 'exceptions' were placed within the earths group.

	Red-brown clays	Earths	Grey clays	Sands	Class error
Red-brown clays	38	6	13	0	0.33
Earths	10	17	6	1	0.50
Grey clays	13	5	32	0	0.36
Sands	0	3	0	0	1.00

4.4.8.2. Valley-wide soil group map

Clays are predicted to be dominant across the region, with Red-brown clays more prevalent in the east and Grey clays more prolific in the west (Figure 4.12). Where the plain ends (143.5°), and the landscape intersects with the neighbouring Mallee region, the map accurately predicts the presence of the Earths group (Figure 4.12). The Earths group is also predicted to occur in the east and northeast of the region as well as along palaeochannel tracts (Figure 4.12).

The developed model showcased an inability to accurately class soils as Sands. Consequently, they are sparsely projected in the map (Figure 4.12). Where sands are present, they fall adjacent to Earth soils in the centre of the study and in small, isolated pockets, among Grey clays in the central-north of the region (-35.375°, 144.5°) (Figure 4.12). It is likely that where areas are classed as Earths, and follow known palaeochannel systems, Deep sands will be present either within the streambed or as source bordering dunes.

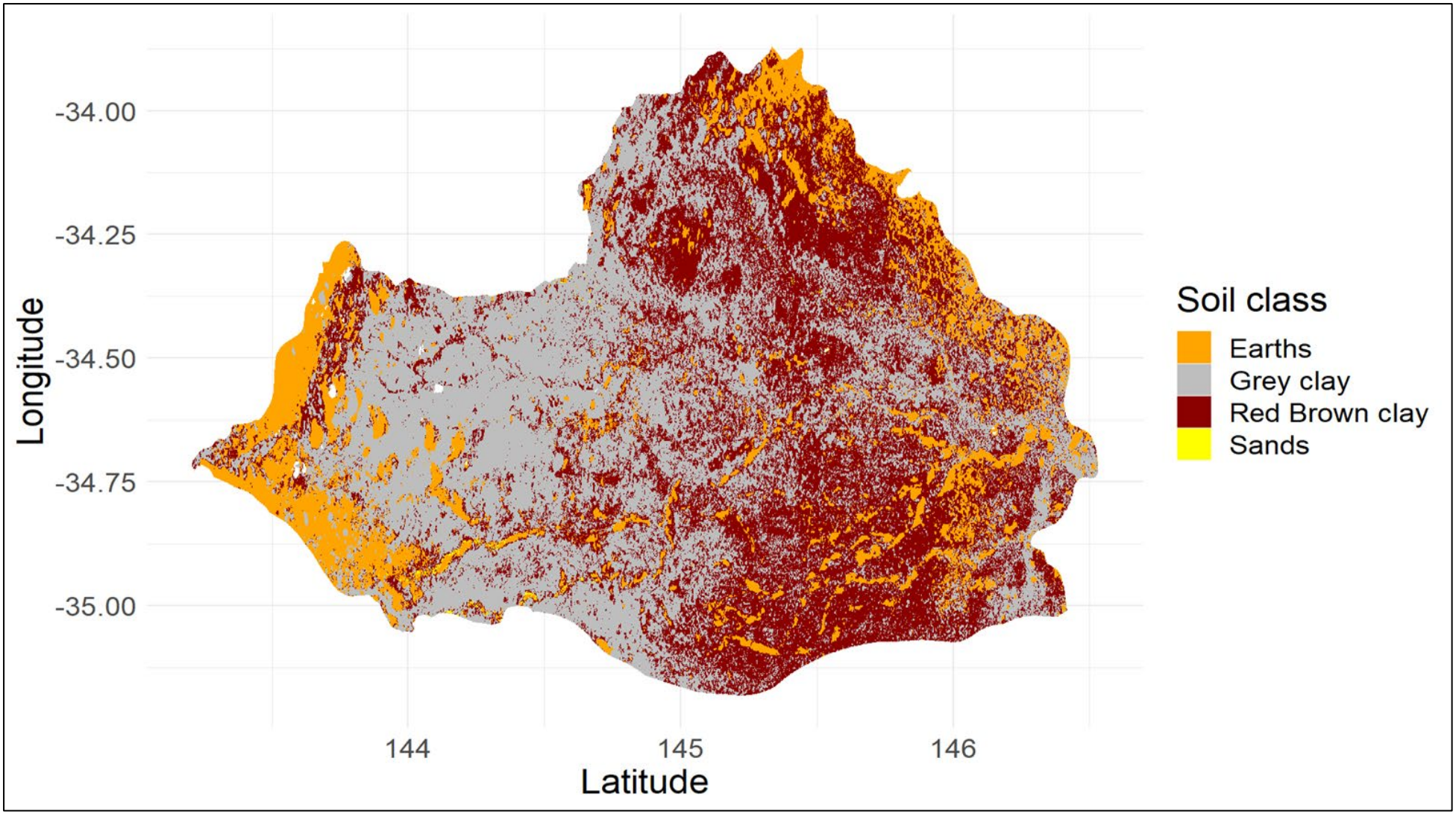


Figure 4.12. A map showcasing the predicted distribution of four soil groups: Grey clays, Red-Brown clays, Earths and Sands across the region at a 90 m resolution. In the west of the region, predictions are not made where lakes are present

4.4.9. Localised distribution of soil classes

4.4.9.1. Farm A

There is significant variability in the morphology of soil cores extracted from the 82 ha paddock on Farm A northwest of Hay (Figure 4.13). The defining features of this paddock and the surrounding landscape are terminal palaeochannel streams. Unlike major palaeochannel systems, no levees are observed adjacent to this streambed in the cultivated or uncultivated areas. Instead, a narrow, sandy, palaeochannel streambed is incised into the heavier clays of the plain (Figure 4.13). These more clayey soils do not contain lime or gypsum and can be either red-brown or grey, with no clear trend in colour as distance from the streambed increases. Interestingly, sandier areas continue beyond the approximate bounds of the stream, with a gradational loam occurring between a red-brown and grey clay in the west of the field (Figure 4.13). In this loam, lime is observed in the lower subsoil. The gradational loam located centrally within the streambed is narrowly excluded from being a deep sand due to a clay content of 19% in the lower subsoil with some structural development. Excluding this profile, and the two deep sands, the remaining five cores are strongly sodic (ESP>14).

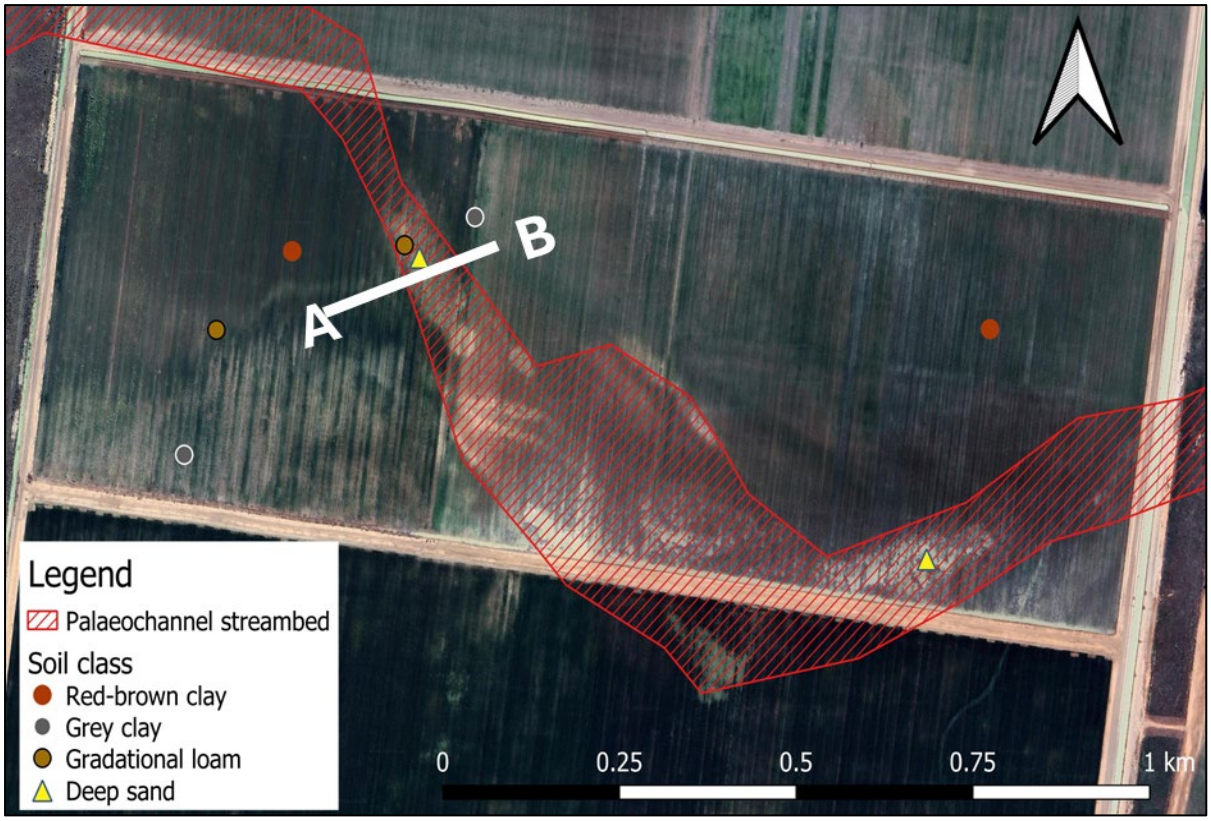


Figure 4.13. The paddock examined utilising the Farm A subset data from this study. The site, located northwest of Hay, contains a sand streambed incising itself within the heavy clay plain. Located within this paddock are Red-brown clays, grey clays, gradational loams and deep sands.

Utilising aggregated soil profile data, the developed soil transect shows that coarse fragments are limited to the sandy palaeochannel streambed and the upper horizons of the grey clay, to the northeast, that the streambed is sharply incised within (Figure 4.14). To the northeast, salinity (ECe) increases as the distance from the streambed increases. There is a slightly more gradual increase in clay content perpendicular to the streambed in the southwest (Figure 4.14). In this instance there are no coarse fragments or trends in salinity with the dominant soil colour being brown instead of grey.

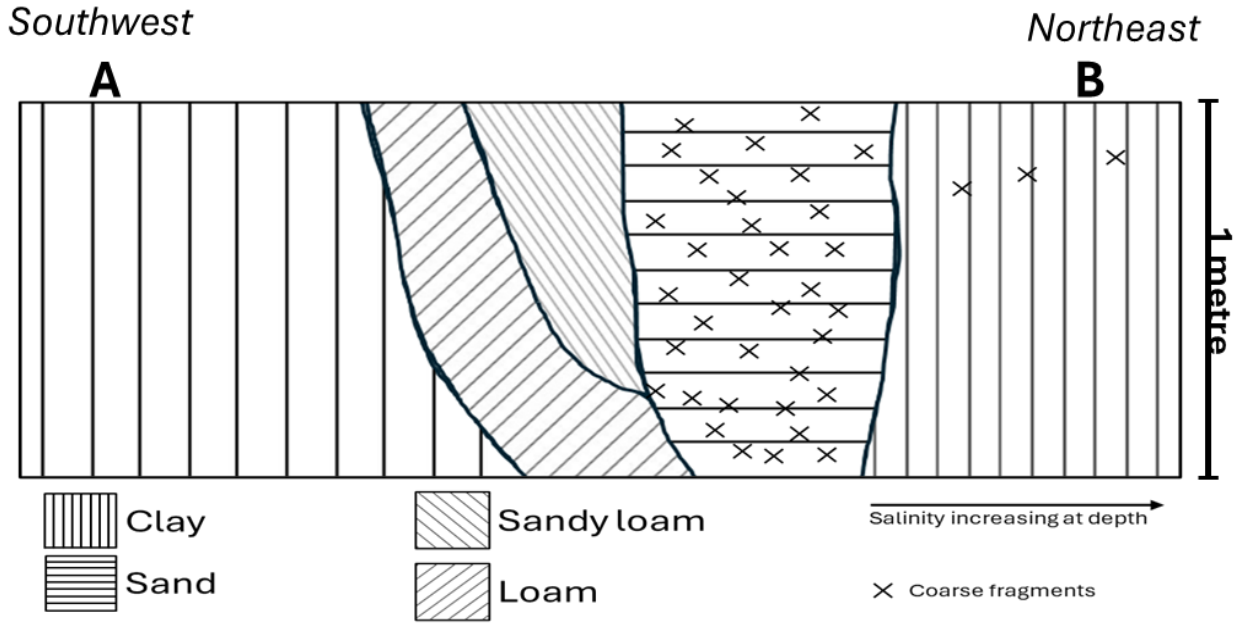


Figure 4.14. Cross-sectional soil diagram along the A-B transect (Figure 4.13) within the field at Farm A.

4.4.9.2. Farm C

There is greater uniformity in the soil cores examined from adjacent to the upper Waddi River Ridge (Figure 4.12) than those described at Farm A. The loam profile, located within the approximate confines of the river ridge, is a uniformly textured sandy loam with a neutral pH. A previously mapped extension of the river ridge (van Dijk and Talsma, 1964), which has likely been eroded or anthropogenically altered, contains two gradational loam soils. These are a loamy texture at the surface, gradually becoming more clayey to the lower subsoil where lime is present (Figure 4.16). Within this area the brown transitional soil has a loam surface overlaying a heavy clay texture in the upper and lower subsoil.

Located equidistant to the north of the primary river ridge are Texture contrast soils (n=3) and Red-brown clays (n=12). Except for one soil core sampled from the northwest of the paddock, and furthest from the ridge, all Red-Brown clays contained lime in the deep subsoil (Figure

4.15). Across the paddock alkalinity increases with depth. There is no clear spatial trend regarding sodicity at depth within the Red-brown clay profiles. While some profiles exhibit strong sodicity (ESP>14) others remain classified as non-sodic (ESP<6).

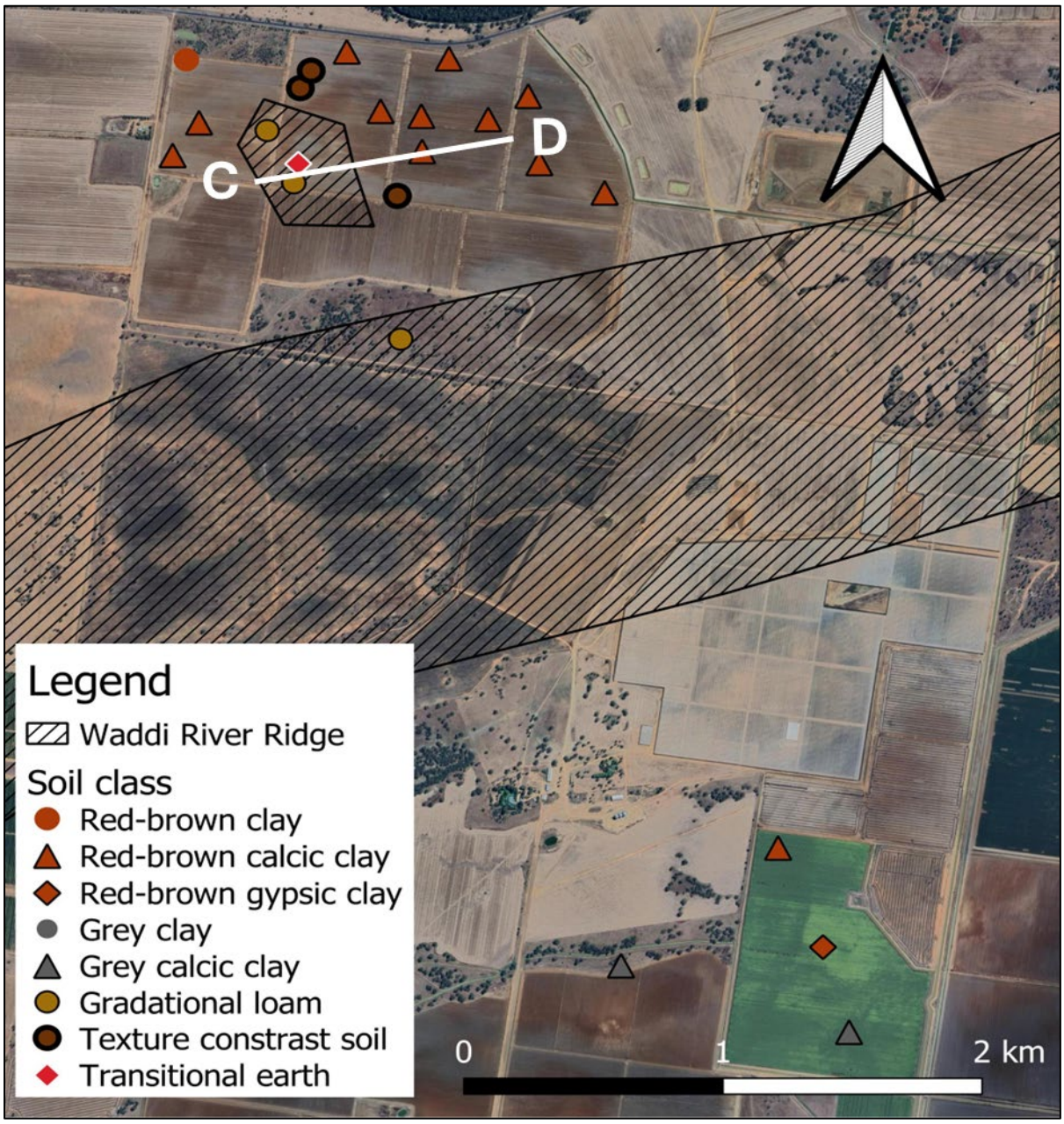


Figure 4.15. Soil cores taken from Farm C on, or adjacent to, the Waddi River Ridge. Within a distance of four kilometres seven different soil classes were observed, Shown north of the ridge which cross the area is an extension of it.

The transect formulated between points C-D in this paddock shows that coarse fragments are present at depth (0.5-1 m) in the Texture contrast and Red-brown calcic clay soils toward the east. Lime is most prevalent as nodules (0.5-0.9 m) or softer, diffuse pedogenic segregations (>0.9 m) in the Red-brown calcic clays, however, is also observed in the Gradational loam and Texture contrast soils (Figure 4.16). The paddock is uniformly clayey, however, within the area

defined as an extension of the ridge, loam upper horizons overlay this clay to varying depths (Figure 4.16). It is difficult to understand this trend due to landscape alterations through land reformation.

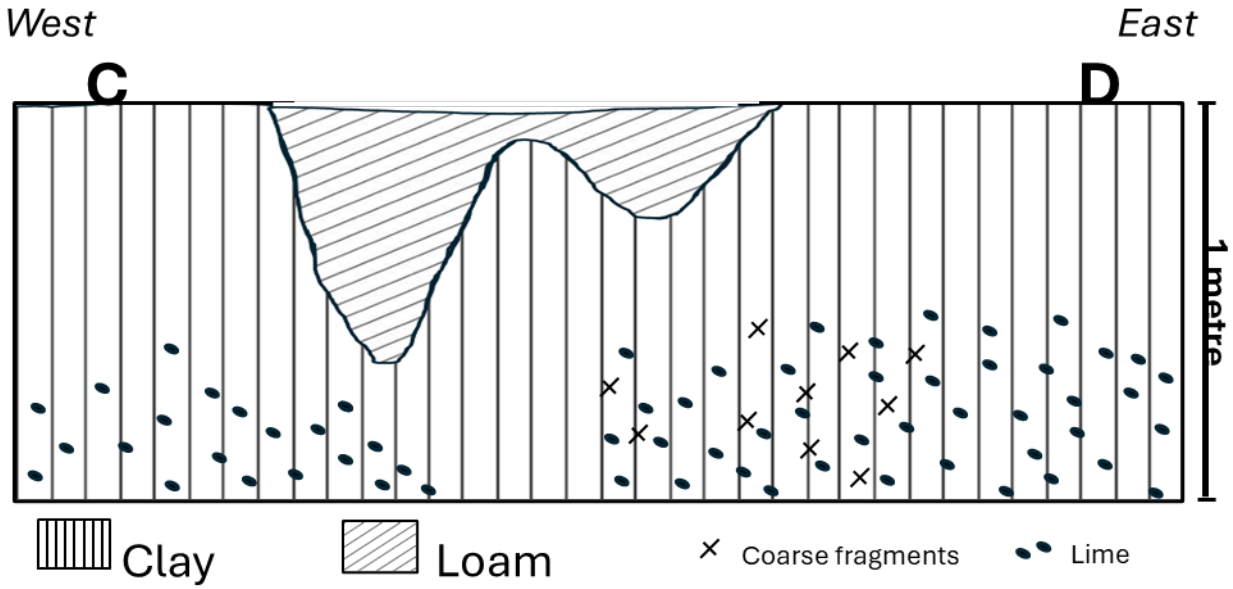


Figure 4.16. Cross-sectional soil diagram developed using morphological observations from soil cores along the C-D transect (Figure 4.15) within the field at Farm A.

Profiles examined south of the river ridge include a Red-brown gypsic clay, a Red-brown calcic clay and two Grey calcic clays (Figure 4.15). The two Grey gypsic clays occur at a greater distance from the ridge than any soils observed to the north (Figure 4.15). The westerly grey core was extracted from a native vegetation site. The core extracted from the irrigated paddock is more acidic throughout the profile, with a surface pH of 6.4, compared to 7.5 in the unaltered soil, and a pH in the lower subsoil of 8.8, compared to 9.0 in the unaltered soil. In the lower subsoil the native site is more sodic, with an ESP of 17% compared to the irrigated soil which exhibited an ESP of 10.6%. Excluding the presence of gypsum in one of the Red-brown clays, both are akin to the sites located north of the ridge.

4.5. Discussion

4.5.1. Development of soil classes

As a result of sampling sites from across the entirety of the lower Murrumbidgee and not just within the eastern irrigation areas, a high proportion of the profiles were classified as clays. Consequently, an emphasis was placed on differentiating between these clays without using localised geographic terminology meaning the developed soil classes differ from those

presented in past soil surveys in the lower MIA (van Dijk, 1961) and CIA (van Dijk & Talsma, 1964) or reclassifications (Hornbuckle & Christen, 1999; Hornbuckle et al., 2008b; Thacker et al., 2008). In the former instances, soil types (units) were categorised as dominant, subdominant and minor within a geographical association largely informed by landscape physiography. While the incorporation of localised landscape knowledge adds information, it reduces the efficacy in translating knowledge to areas where these associations are not mapped. This study categorises soils in a manner that does not require knowledge of isolated physiographic units, instead, presenting classes applicable to the region in its entirety.

As the basis of this classification was primarily soil texture and colour, there is the capacity for classes to be further sub-divided. For example, within the Red-brown clays significant variability is seen in the ECe and ESP measurements suggesting sub-divisions of Red-brown salic or sodic clays would be possible. Not observed in this study are the 'Sands over clays'. These soils are described by Hornbuckle and Christen (1999) and later Hornbuckle et al. (2008b) as texture contrast soils with dense, coarsely structured subsoils that can crack into columns with a rounded top under dry conditions. Within the ASC these soils would be classified as Sodosols. These soils are known to occur within the region, as identified during surveys by Taylor and Hooper (1938) and Stannard (1970). It is possible that some profiles classified as a Texture contrast or Transitional earth soils are in fact Sand over clays given measured sodicity at depth. A columnar structure in the subsoil, however, was a requirement for this class and was unable to be clearly discerned from a 40 mm diameter soil core.

4.5.2. Spatial distribution of soil classes

Due to limited data availability for some of the nine individual soil classes, the broader soil group model was chosen to map soil distribution across the lower Murrumbidgee valley. Even in this instance the low number of data points for the Deep sand group leads to a 100% class error prediction. Sharififar et al. (2019) have previously highlighted the impact that data imbalances can have on model uncertainty.

Despite drawbacks the resulting map appropriately displays known landscape trends including the dominance of clay soils across the region. The prevalence of Red-brown clays in the southeast and north of the region suggests these soils to be older than the grey clays to the west, with the redder colouring suggesting extended exposure to weathering processes and subsequent rubification (Sánchez-Marañón, 2011). Rubification relates to yellowing or reddening of soils in warm climates due to the liberation of iron, or pedogenic hematite

formation, as minerals undergo weathering (Kemp, 1985). This was described briefly by Cay and Cattle (2005) in relation to differently aged sediments in the neighbouring Lachlan valley. It is appropriate then to consider that, theoretically, grey soils are more likely to be younger and associated with more recently active palaeochannels.

These patterns are consistent with maps of palaeochannel systems, where the area in the southeast is associated with the older Coleambally system (active 85-105 ka) while, further west, the younger Kerarbury system (35-55 ka) was more recently active in sediment deposition. Patterns in the region's north are also consistent with this theory, with Langford-Smith (1960) mapping prior streams in this area that would predate the paleochannels mapped by Page and Nanson (1996) and Page et al. (2009). Continuous passages of grey soils, hypothesised to be associated with the Kerarbury system (active 35-55 ka) are likely to have been deposited atop older sediments that, when not buried, are shown as Red-brown clays in the map. This interaction of differently aged sediments was observed where a grey top and upper subsoil overlayed an older, red-brown lower subsoil. This pedological development in the lower subsoil suggests an extended period of sediment exposure, supporting the theories of Butler (1958).

The Earth soils detected in the western fringes of the region occur primarily as Calcarosols where the Mallee landscape meets the Murrumbidgee. In the region's east, these Earths have been widely described as Texture contrast soils with subplastic properties resulting from the influence of parna deposits (Butler, 1979; Cattle and Smith, 2018; Hornbuckle and Christen, 1999; Hornbuckle et al., 2008b; van Dijk, 1958). In the south of the region, mapped tracts of these soils are associated with known pathways of the older Coleambally palaeochannel system (Page & Nanson, 1996).

There are, however, limitations to the soil group map. For example, Murphy et al. (2000) described the region's west (144º-144.5º) as a mosaic of Grey Vertosols and hard Red Chromosols, Calcarosols and Rudosols. Observations of the sand seam at Farm A, described below, show finer scale variability to be present, with an examination of satellite imagery highlighting that this variability is not constrained to this site. Within the region's west, neither the streambed sands or texture contrast soils (Murphy et al., 2000) are represented to the appropriate extent in this model and resulting map.

Soil variability at the farm, or within-field scale, does not support previous theoretical toposequences (Butler, 1958; Langford-Smith, 1960). At Farm C, for example, within the

examined distances of the Waddi River Ridge the dominant soils should consist of loams with clay horizons only occurring at depth. This was, however, only observed in six of the 23 classed soil profiles. This site supported Butler's (1958) theory that lime is greatest, and salinity lowest, closer to palaeochannels with the inverse occurring as distance from the system increases. The results at Farm C are consistent with the survey findings of van Dijk and Talsma (1964) who identified the areas adjacent to the Waddi River Ridge as falling within the locally characterised Cararbury soil association. Within this soil association 'local' soil types, including the Yooroobla and Wunnamurra clays, showcase similarities to the soils described in this study (van Dijk & Talsma, 1964).

The site at Farm A tells a more interesting morphological story, highlighting that where palaeochannel streams are terminal the 'mosaic' becomes more complex. At these sites it is shown that deep sandy soils occur directly adjacent to heavy clays. Further, there is no clear differentiation between Red-brown and Grey clays. While colour differences may result from microtopographical differences, these are not observable, with loam soils also present. This is of interest as it is expected that the grey sediment would be associated with the younger system and should be more uniformly distributed. Instead, grey and red-brown soils occur directly adjacent to each other with no clear differentiation. Along with soil colour, CEC can indicate the extent of soil weathering, whereby a lower CEC suggests an extended period exposure to weathering, suggesting an older soil. In comparing the CECs between the red-brown and grey clays in this field, however, the red-brown clays have a higher CEC at each of the five depths, contradicting the theory that the red-brown soils are older.

It is hypothesised that the soils at Farm A are located on terminal anabranches of the Gum Creek system, active from 25-35 ka, which have incised into older clay sediments. By incorporating knowledge presented in literature, three hypotheses are presented regarding this streambed sand and the absence of characteristic levees. Firstly, despite a low streamload small levees were produced, however, these have since been remobilised across the landscape. Secondly, where small levees were present, instead of being remobilised across the landscape they infilled shallow streambeds resulting in a sand like seam. And, thirdly, levees were never present.

Prior streams, predating the mapped palaeochannel systems of Page et al. (2009), were hypothesised by Langford-Smith (1960) to have been present both north and south of this property (Figure 2.3). As well as identifying the Gum Creek system, Page et al. (2009) mapped the occurrence of the Kerarbury system, which is theorised to be older than the Gum Creek

system but younger than the prior stream that Langford-Smith (1960) mapped to the north of the paddock. It could, therefore, be the case that this is the intersection of these two clay fans, the younger grey soils, associated with the Kerarbury system, and the older red-brown soils, associated with the prior stream of Langford-Smith (1960). At the intersection a younger palaeochannel anabranch has incised, resulting in the deep sand. This aligns with layering of sediments proposed by Butler (1958), however, in this instance instead of the Mayrung and Widgelli Parna units, both sediments are likely of alluvial origin due to their westerly position. Instances of colour patches in the southwest of the paddock may be the result of previous microtopographical differences which have been removed through land reformation.

Ultimately this highlights that Arenosols and Vertosols can occur adjacent to each other, with no topographical difference. This is different to previous studies where the sandy soils are associated with dunes rather than streambeds. The distribution of grey and brown soils does not follow any clear pattern. Consequently, these smaller terminal streams incising themselves within heavy clays produces a pattern akin to spaghetti that strings across the western plain. This creates difficulty in mapping spatial variability at broader scales, as was witnessed in the valley scale soil group map where only minimal pockets of Earths and Deep sands were noted in this part of the region. This is likely why soil associations containing dominant and subdominant soil types were favoured by early researchers.

4.5.3. Clays

Despite the alluvial sediments of the plain being unrelated to the modern system, surface sediments from recent palaeochannel systems are still relatively young. Consequently, profiles are not highly pedologically differentiated and, thus, clays were prevalent throughout the study area. While initial theoretical transects of the region (Butler, 1958) suggested lime decreases as clay content increases, however, calcic soils were the most common sub-class of both Red-brown and Grey clays in this study. The findings of this study are consistent with summaries of clay soils in the region's east by Thacker et al. (2008) and Cay and Cattle (2005), the later noting lime in abundance within clay soils of the neighbouring Lachlan valley. Where gypsum and lime are both present within the first metre of soil, the occurrence of lime above gypsum is expected, given gypsum is more soluble (Anderson et al., 2021).

The clays identified in this study have been classified as Vertosols within the ASC (Isbell & NCST, 2021). There is, however, difficulty in making this classification from a 40 mm soil core. Consequently, it is possible that some clays in any of the six classes could be classed as

Dermosols (Isbell & NCST, 2021). Each class was placed within the Epipedal Group, however, as this was done on the average profile for each class, some profiles would be considered massive. Similarly, as the Subgroup classifications are based on chemical characteristics, and the averages for each class came from highly variable data, some clay profiles with a low ESP would be placed in different subgroups. For example, non-sodic Brown clay profiles would be classed within the Haplic Subgroup.

Sodic soils occur extensively throughout Australia's agricultural regions (Jayawardane & Chan, 1994) and were frequently observed in this study. This is reflected with the Brown, Grey and Grey gypsic clay classes being classified within an Episodic or Episodic-Gypsic subgroup as well as the Grey calcic clays being classified as Epihypersodic-Endocalcerous. A structural collapse resulting from the high proportion of sodium ions was observed in some cores within this study that exhibited extreme (>14% ESP) sodicity (Rengasamy & Olsson, 1991; Dodd et al., 2013a). A further derivation of structurally collapsing sodic soils into sub-categories may be of value to allow more accurate identification of, and advisement regarding, these soils.

The chosen differentiation between clay soils was on the basis that soil colour can be a valuable indicator of the past pedological history related to soil formation and a potential indicator of soil organic carbon, drainage conditions, aeration, mineralogy and iron content (Ibáñez-Asensio et al., 2013; Hartemink & Minasny, 2014). As already discussed in relation to soil distribution, red-brown or yellow soils occur following the pedogenic release of iron from primary minerals suggesting they are more developed, and thus older, than grey soils. Within the clay classes it is common to see the layering of sediments, with grey clays overlaying a red-brown subsoil. At a finer scale, however, it was difficult to identify a clear spatial trend of differently aged soils. As soil colour can also indicate mineralogy (Hartemink & Minasny, 2014), it is possible that varying colour patterns are the result of different parent materials for differently aged sediment deposits. Alternatively, younger deposits may be red-brown if the source of sediment was older, more pedologically developed upstream sediments. This could include upstream sediments that were initially of aeolian origin, in accordance with the pathways of Cattle et al. (2009) and Greene et al. (2009).

A comparison of soil CEC has already been undertaken at Farm A, however, an analysis of one Grey calcic clay and two Red-brown calcic clays 50 km south of Hay shows the grey soil to have a higher CEC at all depths. Similarly, comparing three Red-brown calcic clays, located north of Hay, to three Grey calcic clays, located south or southwest of Hay, showed the grey soils to have a higher average CEC at depths of 0-0.1, 0.1-0.3 and 0.8-1 m. This supports the

notion that the grey clays are younger. These sites are exposed the same environmental conditions but a higher CEC in the grey clays indicates that there has been less mineral weathering. It may then be the case that the described soils at Farm A are an outlier.

Measured pH, alongside variability in soil salinity (ECe) and sodicity (ESP) to 0.3 m, are similar to findings for the clay groups of Thacker et al. (2008) in their review of available regional soil data. This study, however, did not differentiate between self-mulching and hardsetting clays. Summary pH statistics for work previously undertaken on hardsetting clays shows them to be slightly more acidic than both those in this study and the previously differentiated self-mulching group of Thacker et al. (2008). While similar variability is observed in ECe and ESP data, there is an increase in extremely sodic measurements at 0.3 m in the hardsetting clays. This comparison, however, may not be valid as data contained within previous reviews is limited to the Murrumbidgee and Coleambally Irrigation Areas with both salinity and sodicity known to increase along the east-west axis (Talsma, 1968). This trend was observed in this study, with increased pedogenic lime also seen, an expected occurrence given that aridity also increases along this gradient.

4.5.4. Texture contrast soils

While the average texture contrast soil in this study was considered a Dystrophic Subnatric Brown Sodosol, some individual profiles within this class could be considered as Haplic or Sodic Dystrophic Brown Chromosols on account of lower ESP. Similarly, varied coloured soils, such as the grey profile, would be placed in different Suborders. It is important to note that some soils within the region contain subplastic properties. If subplasticity is present soils are excluded from being classed as Sodosols due to landuse impacts and would therefore be classified as Chromosols and likely placed within the Subplastic Great Group (Isbell & NCST, 2021).

Texture contrast soils have previously been described as Red-brown earths (Hornbuckle & Christen, 1998) which are differentiated into three subclasses based on their subplasticity (van Dijk, 1958). Subplasticity refers to the soil texture grade becoming more clayey following mechanical working (Cattle & Smith, 2018) and has been widely discussed in literature (Blackmore, 1976; Brewer & Blackmore, 1976; Butler, 1976; McIntyre, 1976; Norrish & Tiller, 1976; Walker & Hutka, 1976). These soils are also referred to as ‘hillslope soils’ (Butler, 1979) and are prevalent in the east-northeast of the region. The Red-brown Earth classification has been broadened to Texture contrast soils as profiles in the region’s west fit the texture contrast

classification but contain grey horizons. Further, while the physical properties of these soils are well understood, Thacker et al. (2008) note that there is limited available chemical data.

While texture contrast soils are included in the 'mosaic' description of the landscape presented by Murphy et al. (2000), these soils are pedologically developed as indicated by their yellow or red colour. In contrast, the Texture contrast soil observed west of Hay in this study was grey throughout. Given this, it is unlikely that the texture contrast results from the illuviation of clay particles over an extended period, nor has it been anthropogenically altered, as it is a native site. Consequently this site, located near the Murrumbidgee River, is likely the result of coarser sediments from a Holocene flood burying an alluvial clay associated with a palaeochannel system depositional event. Excluding this grey site, the presence of a texture contrast and red, or yellow, colours indicate that these soils have undergone a greater degree of pedological development than the previously described clay soils. Another possibility for these soils is that a sandy clay loam has buried a light medium clay which had previously been exposed for an extended period resulting in the red colouration.

4.5.5. Transitional earths

Soils classed as Transitional earths have previously been considered Transitional Red-brown Earths (Hornbuckle & Christen, 1998) however, as with the Texture contrast soils, this grouping has also been broadened on the basis of colour. Based on the average data obtained in this study, Transitional earth soils were classed as Sodic Dystrophic Brown Chromosols. As described above, however, profiles may be placed in separate Suborders on account of their colour in the B2t horizon. A sodic B2t horizon would also result in the profile being classified as a Sodosol (Isbell & NCST, 2021). The Transitional earths of this study are more alkaline to a depth of 1 m than the summary statistics presented by Thacker et al. (2008). To a depth of 0.3 m there is similar variability in ESP measurements, with the summary statistics of Thacker et al. (2008) also exhibiting variability from <1% to >20%.

These soils have been used successfully for broadacre irrigated agriculture, with crops grown including sorghum and ponded rice (Beecher, 1991; Cook et al., 1989). Consequently, much research on these soil types has focused on hydrological properties and water infiltration related to management of these systems. Khan et al. (2002) summarised that, from 0.2-0.6 m depth, hydraulic conductivity is lower than in texture contrast soils, ranging from 0.026 to 10 mm/day. Cook et al. (1989) found that steady state infiltration in these soils increased following slotted

treatments of gypsum. These soils are well suited to ponded rice production, however, are often unsuited for furrow irrigation (North & Schultz, 2017).

The Transitional earth observed in the region's north is primarily surrounded by Red-brown clays, some of which are calcic. It is possible that this is a more pedologically developed version of these soils or, like the texture contrast profile described above, it is the result of 0-0.1 m of coarser sediment burying an alluvial clay. It has been previously stated that the Transitional Red-Brown Earths commonly contain lime and gypsum in the subsoil (Thacker et al. 2008). While pedogenic lime was observed in this study, gypsum was not seen within the first metre of the soil profile.

4.5.6. Deep sands

Two of the three Deep sands in this study differ significantly from those previously described. In summarising previous research, Hornbuckle and Christen (1999) noted that deep sands are primarily of aeolian origin occurring as dunes. Stannard (1970) described 13 local sand soil types surrounding Coleambally, 10 of which are associated with hillslopes or rises. Consequently, while not observed at sites within this study, Deep Sands are known to be present in the region's east. These sands, like the site observed south of Hay, have developed because of secondary aeolian action as described by van Dijk and Talsma (1964). The three non-sandhill soil types (Stannard, 1970) become clayey within the first metre of the soil profile. Summary statistics for these sandhill soils are like those in this study, with soils exhibiting a low electrical conductivity and neutral to moderately alkaline pH (Thacker et al. 2008).

The two sand profiles described at Farm A, however, are morphologically differentiated from these previously described soils due to the presence of coarse fragments and absence of clay within the first metre of the profile. Thus, they have been termed streambed sands as they are not of aeolian origin. While Hornbuckle et al. (2008b) note a streambed sand, termed 'Stream bed soil type A', no data or source for this classification is provided. A further factor behind the differentiation of these sand types is the measured ESP. The dune sand south of Hay was marginally-to-strongly sodic at all depths while the streambed sands were non-sodic.

Based on satellite imagery, these soils, while only in isolated pockets linked to physiography, occur throughout the region. They have previously been classified within the Siliceous Sands Great Soil Group (Stace et al., 1969). Where these sands are sharply incised within heavy clays, resulting in abrupt texture changes, there are potentially significant impacts on irrigated

agricultural systems. Murphy et al. (2000) noted beds of Hypocalcic Calcarosols and, while descriptions of these soils may exhibit some similar characteristics to the sandy soils described here, the Deep sands class in this study is not calcic. Thus, it is proposed that these streambed sands be differentiated from other sand classes in future. Further, given their impact on potential agricultural management it is advisable that further work examining their distribution and characteristics beyond a depth of 1 m is undertaken.

4.5.7. Exceptions

The 'Dunefield soils', identified from three soil cores taken south of Balranald are located on what Mitchell (2002) terms the 'Mallee Cliffs Sandplains'. Located on the interface of the Mallee and Murrumbidgee regions, this landscape produces interesting toposequences as observed in this study, with calcareous loamy sands, associated with aeolian dunes, becoming more clayey down an eastward slope, ending in heavier brown or grey clays associated with Murrumbidgee palaeochannel systems (Mitchell, 2002). The upper soils in this sequence, compromising two profiles in this study, exhibit rubification, through reddened coatings on sand particles. Rubification in this instance refers specifically to the coating of quartz particles in a thin reddish film because of iron being released from minerals (Ben-Dor et al., 2006). As there is no clear gradational reddening, it is likely that this rubification is the result of a clay-rich cutan coating on quartz grains which has then oxidised and can be maintained on the grain even if it is remobilised (Bowler & Magee, 1978).

Bowler and Magee (1978) described the area surrounding the three core locations as the boundary between 'linear dunes' and the 'alluviated riverine plain'. Pell et al. (2001) described these lunettes as falling within the Woorinen Formation, one of three sands within the Mallee dunefields. It is suggested that these linear dunes have resulted from multiple periods of dune remobilisation, with the last episode to have occurred between 16,000 and 15,000 years ago (Bowler & Magee, 1978; Bowler et al., 1976).

Soils considered loams or gradational loams, with the exclusion of one profile described at Farm A, are most closely associated with Transitional earths or Texture contrast classes. Where these soils have been anthropogenically altered through agriculturally systems it is possible that under a natural condition they would have been classed as either Transitional or Texture contrast soils.

Soils considered as unclassified are differentiated based on unusual trends of texture throughout the profile whereby clay content either decreases consistently from the surface or increases to a maximum at 0.1-0.6 m before decreasing. Soils of this nature have been previously identified by Stannard (1970) within the Coleambally region. In these soils, such as the 'Crommelin Clay', the lightening of texture does not occur until the deep subsoil below 1.06 m (Stannard, 1970). Consequently, the soils within this category under natural conditions are likely the result of a younger, surface layer, overlaying an older buried profile. If these profiles are used for agricultural production, they may have been altered by anthropogenic activities such as the cutting and filling of sediments to create level gradients for irrigation, the impact of which have been discussed by Cay and Cattle (2005).

4.6. Conclusion

This study highlights the variability in soil types at both a regional and within-field scale within the lower Murrumbidgee valley. Morphological descriptions of 153 soil cores from across the study area highlighted the prolific occurrence of clay soils, particularly within the previously unsurveyed west of the region. This necessitated the development of a new regional classification. Based on recurring trends in the collected samples, a bifurcating key was created to place soils within ten classes: Red-brown clays (n=16), Red-brown calcic clays (n=31), Red-brown gypsic clays (n=10), Grey clays (n=17), Grey calcic clays (n=26), Grey gypsic clays (n=7), Transitional earth soils (n=9), Texture contrast soils (n=6) and Deep sands (n=3). The remaining cores were classed as exceptions (n=28). Lime was frequently observed in the clay classes as veins, nodules or diffuse pedogenic segregations while subdominant colours were also common. The distribution of soil classes, visualised through a soil group map produced using a random forest classification model, is clearly related to the pathways of palaeochannel systems. Sharp colour changes from a grey to red-brown colour within the soil profile suggest that older, pedologically developed, soil profiles have been buried by younger sediments. The development of two field-scale transects highlighted the sharp changes in soil morphology that occur. At Farm A, streambed sands were identified and it is proposed that these soils be differentiated from other dune sands based on their position within the landscape. There were instances where anthropogenic activities were believed to have altered the soil profile. Despite the development of new soil classes, there is still significant variability in soil chemical properties such as ECe and ESP. As these properties are of importance to soil management considerations, further research should look to expand on this work through the further development of subclasses.

Chapter 5

5. Examining the influence of lower Murrumbidgee valley palaeochannel systems on soil properties and assessing their impact on within-field cotton yield variability

Abstract

Distributary patterns of soils within the lower Murrumbidgee valley of southern NSW are related to the pathways of relic river systems defined as palaeochannels. Differing pathways, ages and characteristics of these deltaic like systems result in highly variable soils at the within-field scale. It has previously been stated that an improved understanding of this variability, and how it relates to cotton production, is required to improve management practices. This study examined the within-field variability and relationships to cotton lint yield of soil properties including pH, soil electrical conductivity of the extract (ECe), exchangeable sodium percentage (ESP), cation exchange capacity (CEC), clay content and sand content. This occurred in two paddocks on different farms; Farm A and Farm C, at five depth increments; 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8, 0.8-1 m. At Farm A 2018 cotton lint yield varied from 4-11 bales/ha in an 82 ha paddock while at Farm C, cotton yield ranged from 6-17 bales/ha (2018) and 6-18 bales/ha (2019) in a 94 ha paddock. A point correlation (r) analysis revealed clay and sand content to be the soil properties most positively and negatively correlated to crop yield, respectively. At all depths at Farm A the strength of this relationship was >0.89 for clay and <-0.87 for sand. Also of note was the positive relationship between ESP and cotton lint yield at all depths, on both farms (0.06-0.75). This analysis revealed that the strongest correlations were in the subsoil. Simple linear models are demonstrated as being moderate-to-good at predicting soil properties within-field when assessed using the Lin's Concordance Correlation Coefficient (LCCC) through a leave-one-site-out-cross-validation (LOSOCV). At Farm C, linear models outperformed random forest models, with LCCCs obtained through the LOSOCV ranging from 0.62-0.73. At Farm A, only linear models were developed with LOSOCV obtained LCCCs varying from 0.45-0.65. For each paddock, two production zones were delineated reflecting differences in crop yields and associated soil properties. The identification of these zones can allow for variable rate management to increase production efficiency and profitability. This research shows that variability in soil texture within-field significantly impacts irrigated cotton production, with considerations given on the implementation of different irrigation systems in the future.

5.1. Introduction

The distribution of soils within the lower Murrumbidgee valley in southern NSW occurs primarily as a result of palaeochannel systems. Within the region four, temporally distinct periods of palaeochannel activity have occurred over the last 100,000 years (Page & Nanson, 1996). Because of this, differently aged sediments are intertwined across the landscape resulting in sharp variations in soil properties. An understanding of this variability at the within-field scale is required for sustainable irrigated cotton production.

While Australia exhibits the highest average cotton lint yield per unit of land globally, there is still significant yield variability even at a within-field scale (Constable & Bange, 2015; Nachimuthu et al., 2024). Historically, cotton has been grown primarily on highly fertile Vertosols in northern NSW and central-southern Queensland (Conaty et al., 2022; Isbell & NCST, 2021). Consequently, much of the research relating to production systems and their soils has occurred in these regions (Cattle & Field, 2013). Over the last two decades, however, there has been a significant expansion of the cotton industry in southern NSW (Gupta & Hughes, 2018). A review of the information available regarding the region's soils in relation to cotton production by Holland and Eastwood (2014) found that growers required more information on soil attributes including sodicity, salinity and soil pH.

Soil sodicity results from excess sodium ions on charged clay particles and organic matter (Cattle and Field, 2013; Filippi et al., 2020). The consequence of this imbalance is the repelling of soil particles, resulting in the dispersion of soils and their structural decline (Cattle and Field, 2013; Rengasamy and Olsson, 1991). This collapse in soil structure can reduce the depth that soil plants roots are able to explore, also affecting soil porosity and hydrology, negatively impacting production. Soil pH is another key determinant of crop growth and development with different crops having different optimal pH ranges, most commonly from 5.5 to 9.0 (Hazleton and Murphy, 2007), values exceeding which will limit plant growth through toxicity and/or the inability to access essential nutrients. In excess, soil salinity can interfere with the plant uptake of water and essential nutrients; impacting growth, development and ultimately crop yields (Butcher et al., 2016; Munns, 2002; Sharif et al., 2019).

The texture of soils refers to the proportion of differently sized mineral particles, also impacting the production potential of soils (Wang et al., 2021). Soils with a higher clay content exhibit a higher water holding capacity, with the ability to hold water more tightly between soil particles and store water between irrigation cycles or across fallow periods (Hake & Grimes, 2010).

Clay soils also have a lower saturated hydraulic conductivity, with Hazleton and Murphy (2007) estimating a typical range of 2.5-10 mm/hour in moderately pedal light clay compared to over 120 mm/hour in a weakly pedal sandy loam. In the lower Murrumbidgee valley, all cotton is grown under irrigation (ABS, 2022), with at least 80% of Australia's cotton crop produced using gravity-surface-irrigation systems (Roth et al., 2013). These systems, however, are best suited to soils with a high water holding capacity and low-medium infiltration rate (North & Schultz, 2017).

Globally, studies have identified the complex relationships between cotton yield and the suite of soil attributes such as texture, electrical conductivity, exchangeable cations, slope and pH (Constable & Bange, 2015; Elms et al., 2006; Ping et al., 2004). Nachimuthu et al. (2024) state that investigating individual soil properties in high compared to low yielding areas could provide further insights into yield variability, with few studies in this field. In their study, they examined five farms, two of which were in the Riverina, identifying soil sodicity to be negatively related to cotton lint yield (Nachimuthu et al., 2024). Their Riverina soil sites, however, were both heavy clay soils and textural variability is known to occur more broadly across the region (Nachimuthu et al., 2024). Understanding the influences of soil properties is essential prior to the development of management zones or the implementation of variable rate irrigation. Opportunities for this may increase given a recent, albeit slight, uptake in lateral move or pivot systems in the cotton industry (Neupane & Guo, 2019; Roth et al., 2013).

Given the heterogeneous nature of soils, an understanding of spatial variability is critical. While state, national and global digital soil mapping (DSM) products are available for the lower Murrumbidgee valley, research suggests that these products are poor at identifying variability at the point, field or farm supports (Han et al., 2022). Filippi et al. (2024) present bespoke farm or paddock scale DSMs developed using proximally and/or remotely sensed covariates as an opportunity to overcome this. Kidd et al. (2020) stated that at the farm level DSM development has been "historically driven by the private market" while many paddock scale DSMs are also only available for the topsoil and are generated using simple interpolation methods between grid sampling points (Filippi et al., 2024).

Management or production zones also present an opportunity to assess the extent of spatial soil variability in relation to crop yields while providing the basis for site specific crop management (Whelan & McBratney, 2000). The derivation of fields into homogeneous zones allows for more informed decisions, targeted towards specific crop and soil types, to be made at spatially specific locations (Whelan & McBratney, 2000). There are various data sources available for

deriving management zones including yield, soil characteristics and remote sensing, which can be combined or used individually (Ali et al., 2022; Javadi et al., 2022; Leo, 2022; Leo et al., 2023; Whelan & McBratney, 2003). K-means clustering is one of the most commonly used algorithms for delineating clusters once data layers have been collated (Leo, 2022; Zeraatpisheh et al., 2022). Once derived clusters, also referred to as zones, can be managed appropriately on an individual basis to improve economic and environmental outcomes (Leo et al., 2023; Zhang et al., 2002).

To examine the influence of soil variability on crop yields, two farms on differently characterised paleochannel systems were selected within the lower Murrumbidgee valley. Available cotton lint yield monitor data from a harvester-mounted sensor and collected soil data was used to identify point correlations between cotton lint yields and a suite of important soil properties (pH, soil electrical conductivity of the extract (ECe), exchangeable sodium percentage (ESP), exchangeable calcium (Ca^{2+}), exchangeable magnesium (Mg^{2+}), exchangeable potassium (K^{+}), cation exchange capacity (CEC) and sand and clay content). This was done at five depth increments: 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8, 0.8-1 m. Production zones were derived using yield data to assess how these soil properties varied between high and low yielding zones. Lastly, proximal and remotely sensed covariates were used to developed linear and random forest models to predict and map soil properties at all depths concurrently.

5.2. Materials and methods

5.2.1. Study sites

The paddocks examined are two of those described as ‘within-field subsets’ in Chapter 3 as well as in relation to the ‘localised distribution of soil classes’ in the preceding chapter of this thesis. In the latter, a detailed discussion on the morphological characteristics of each paddocks’ soils is also provided. Consequently, only a summary is provided below.

5.2.1.1. Farm A

The first paddock examined in this study is located on Farm A, approximately 18 km north-west of the township of Hay. The 82 ha paddock is used for broadacre irrigated cropping, with rotations including cotton, corn, wheat and barley. The paddock is irrigated using siphons over banks in a furrow system where land reformation has occurred to create a consistent gradient. Incising itself within the paddock, resulting in variability observable from satellite imagery

(Figure 5.1), is a terminal palaeochannel streambed believed to be associated with the Gum Creek system (Page et al., 2009). Soils within the paddock vary from deep sands to heavy brown or grey clays.

Figure 5.1. Satellite imagery of the 82 ha paddock at Farm A.

5.2.1.2. Farm C

The second paddock examined is located on Farm C, approximately 12 km southwest of the township of Darlington Point and 3 km south of the Murrumbidgee River. The 94 ha paddock is used to grow crops including wheat, canola, barley, corn and cotton which are irrigated using a pontoon system. The paddock is located approximately 2 km north of the Waddi River Ridge (van Dijk & Talsma, 1964), an upper arm of the Kerarbury system with a streamflow estimated to be 8.3 times greater than the modern river (Page & Nanson, 1996). Like the paddock at Farm A, variability is noticeable from satellite imagery (Figure 5.2). Soils in this paddock range from texture contrast and transitional earths to heavy brown clays.

Figure 5.2. Satellite imagery of the 94 ha paddock at Farm C.

5.2.2. Crop yield data

All data preparation and analysis in this study was undertaken using R Version 4.3.2 (R Core Team, 2023). Available yield data for summer and winter crops was obtained for both fields. At Farm A this consisted of 2018 cotton and 2020 and 2021 wheat yields. At Farm C, yield data was obtained for 2018 and 2019 cotton as well as 2019 barley and 2020 wheat crops. The raw yield data for each crop in both fields was assessed before being predicted onto a 5 m grid using an Inverse Distance Weighting function contained within the 'gstat' package in 'R' (Pebesma, 2004). The minimum and maximum number of points to be considered was set to 10 and 100, respectively (Figure 5.3; Figure 5.4).

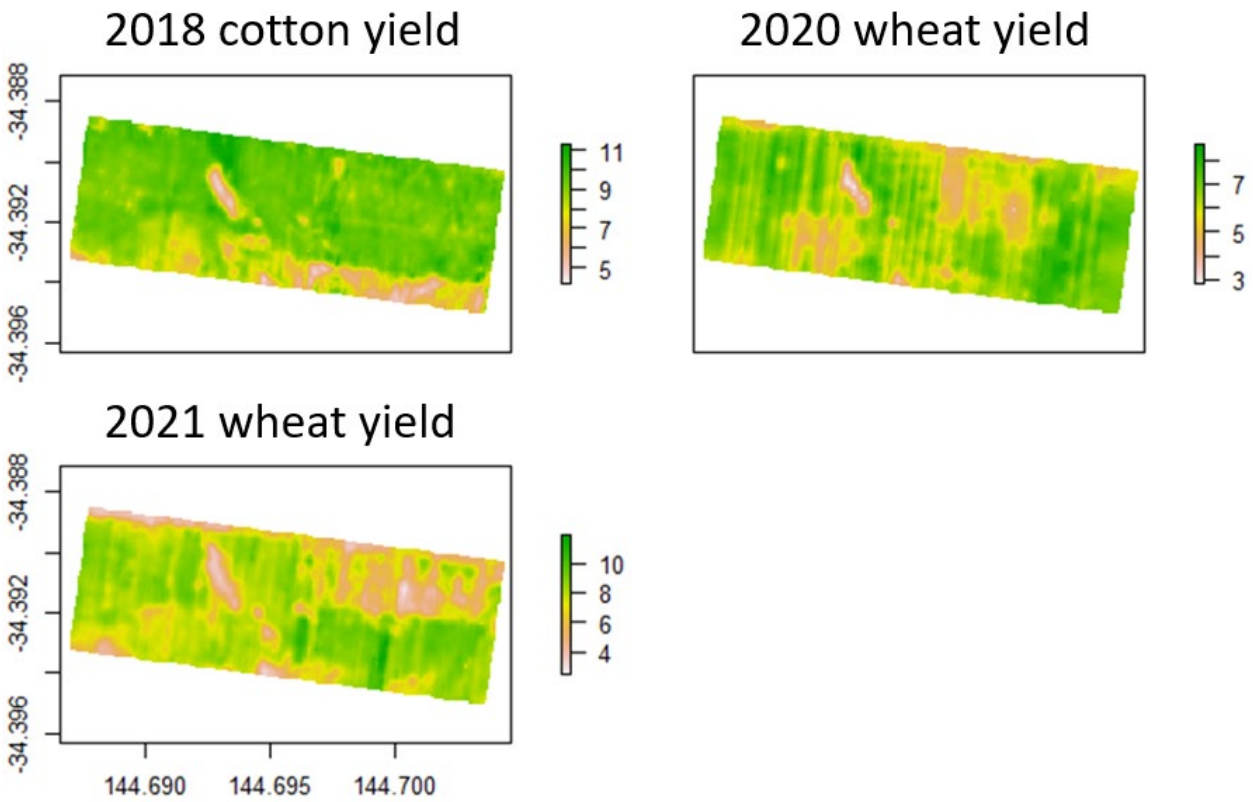


Figure 5.3. Yield data at a five metre resolution for the 2018 cotton (bales/ha) and 2020 and 2021 wheat (tonnes/ha) crops in the examined field at Farm A.

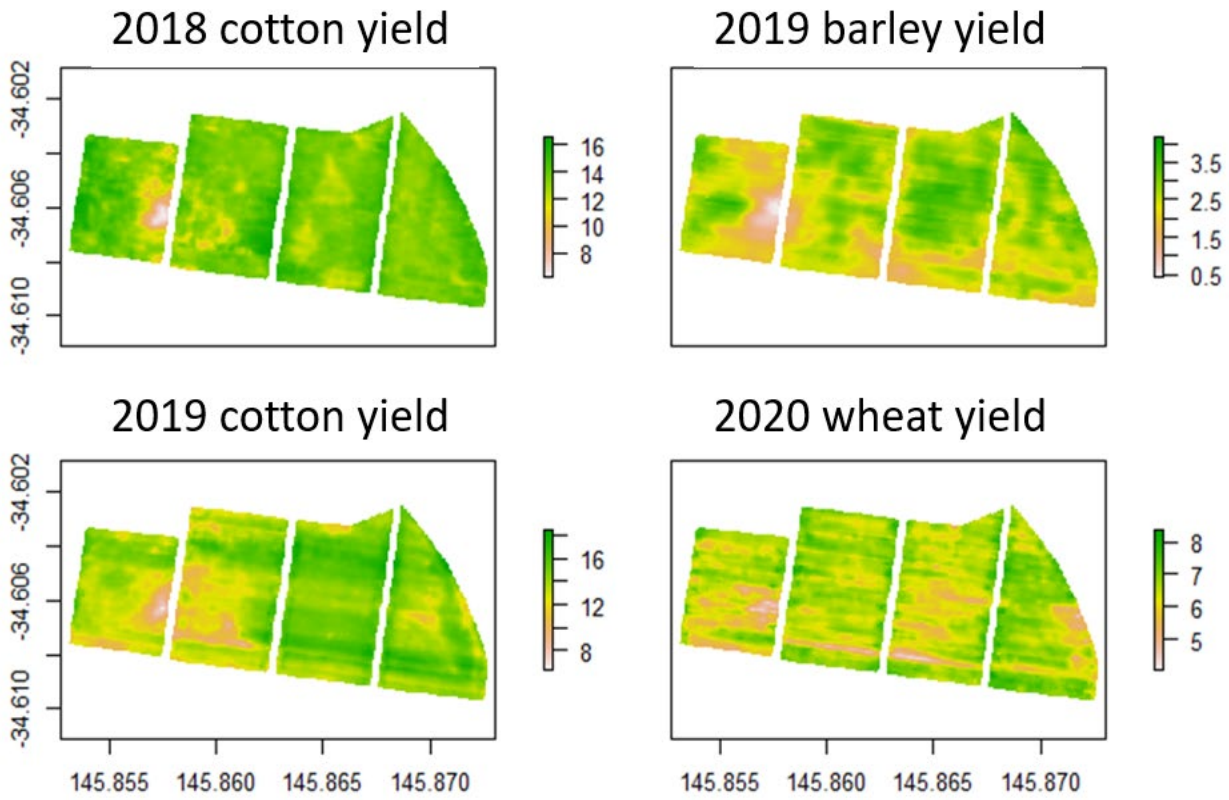


Figure 5.4. Yield data at a five metre resolution for the 2018 and 2019 cotton (bales/ha), 2019 barley (tonnes/ha) and 2020 wheat (tonnes/ha) crops in the examined field at Farm C.

5.2.3. Development of production zones

Utilising the yield data for both fields the Pearsons correlation coefficient (r) was used to assess the strength and direction of relationships between crops over different years. Yield data was discussed with each landholder to confirm that observed variability was consistent with their understanding of the paddock and not due to unrelated management effects, such as blocked irrigation tail drains or areas being under-sown. As this study is primarily focused on cotton yield variability, the anecdotal spatial relationship of barley and wheat yields with cotton yield, including for years where data had not been maintained, was also discussed. This local knowledge alongside the strength of correlations was used to decide what yield data would be used to develop production zones.

Once satisfied with the quality of data, selected yield data was scaled before undergoing k-means clustering. While the statistically optimum number of clusters is commonly determined through the elbow method, in both fields two production zones were generated. This was done to suit the practicality and purpose of this study which focuses on the application of methods to meet the needs of producers within the region. As there were distinctly variable zones observed in the raw yield data it was decided that two clusters would appropriately differentiate variability and allow for a zonal sampling plan.

5.2.4. Soil sampling

The production zones were discussed with landholders alongside proximally sensed soil data prior to the selection of sampling sites. At Farm A, a total of eight sample sites were chosen, with a clearer delineation between the two clusters within the paddock. Within the two clusters sample sites were randomly selected with the number of sites per cluster determined by the cluster's size. In the larger Cluster one, five cores were extracted while in the smaller, Cluster two, three cores were taken to allow for appropriate replicability. At Farm C, 18 sites were sampled, with 13 cores extracted from the larger, Cluster one, and five from Cluster two.

At each sample site, soil cores were extracted to a depth of one metre using a vehicle mounted hydraulic soil corer. The location of each core was recorded using a handheld GPS unit. Once collected, cores were packaged in their natural condition and returned to the laboratory for further analysis.

5.2.5. Soil property datasets

As the methods employed to obtain soil property datasets are described in chapter 3, only a brief description will be provided here. On return to the laboratory and being morphologically described, soil cores were subsampled at five depth increments (0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m) before being air dried (40ºC) and ground to pass through a 2 mm sieve.

Following completion of drying and grinding, soil pH and Electrical Conductivity (EC) were measured for all samples in a 1:5 soil to water suspension using a SevenCompactTM s220 pH/Ion meter and S230 Conductivity meter, respectively (Rayment & Lyons, 2011). A texture grade conversion factor was used to convert soil EC to soil ECe (electrical conductivity of the extract). Subsamples were sent for external analysis of Ca^{2+}, Mg^{2+}, K^{+} and Na^{+} as well as cation exchange capacity (CEC) and exchangeable sodium percentage (ESP) using the 15C1 method of Rayment and Lyons (2011).

A mixed approach was taken to determine soil particle size distribution, using the pipette method of particle size analysis (Glendon & Or, 2002) alongside VisNIR spectroscopy to predict soil texture. Initially, however, the spectral library could not be curated due to enforced sampling delays encountered during this project. To allow analysis to proceed, the pipette method of particle size analysis was used to determine the particle size distribution of a subset of samples from this study. In order to balance this labour-intensive process with an understanding of soil texture variability, samples from the 0-0.1 m depth were analysed for all cores, with subsamples from each core then alternating between the depths of 0.1-0.3 and 0.6-0.8 m, or, 0.3-0.6 and 0.8-1 m. Once further sampling was possible, and all cores from the broader ‘valley-wide’ study area were obtained, samples were finely ground, by hand, before being scanned on an ASF Agrispec spectrophotometer using a contact probe (Malvern Panalytical, Boulder, Co, USA). Utilising spectral (n=754) and laboratory data from the ‘full-dataset’ (n=205), cubist models were developed to predict clay (%) and sand (%) for all samples.

5.2.6. Proximal and remotely sensed covariates

A combination of proximal and remotely sensed data was used as soil property model covariates (Table 5.1). Privately held, proximally sensed electromagnetic induction (EMI) and gamma radiometric sensor data were obtained from landholders. Proximal surveying was conducted on a 24 m swath, with the position recorded with an EMLID Real-Time Kinematic

(RTK) GPS unit. Electromagnetic induction, used to measure soil apparent electrical conductivity (ECa) which can be used to estimate various soil properties and how they vary spatially, was determined using a DUALEM-21S instrument with measurements recorded at multiple depths (Dealem Inc., Milton, ON, Canada). A RSX-1 gamma radiometric detector with a 4 L sodium iodine crystal was used to record gamma radiometric data (Radiation Solutions Inc., Mississauga, ON, Canada). To produce continuous layers across a 5 m grid from this raw data, Inverse Distance Weighting (IDW), with a minimum and maximum number of points to consider for interpolation of 10 and 100 respectively, was used.

Eight publicly available, remotely sensed covariates were downloaded through the dataharvester package in R (Table 5.1) (Haan et al., 2023; R Core Team, 2023). For both fields the covariates were extracted onto a 5 m grid as to align with yield and proximal data using the point-in-polygon interpolation.

Table 5.1. Proximally and remotely sensed covariate data obtained for use in modelling for this study. The data category, description, source and initial resolution are shown.

Data type	Category	Data description	Access source	Resolution
Proximally sensed data	Electromagnetic induction	ECa 50 cm	Privately held	5 m
		ECa 150 cm	Privately held	5 m
		ECa 300 cm	Privately held	5 m
	Gamma radiometrics	Total dose	Privately held	5 m
		Thorium	Privately held	5 m
		Potassium	Privately held	5 m
		Uranium	Privately held	5 m
	Elevation	Elevation	Privately held	5 m
Remotely sensed data	Bare earth	TWI	Gallant and Austin (2012c)	30 m
		Blue band	Roberts et al. (2019)	30 m
		Red band	Roberts et al. (2019)	30 m
		Green band	Roberts et al. (2019)	30 m
		NIR band	Roberts et al. (2019)	30 m
		SWIR 1610	Roberts et al. (2019)	30 m
		SWIR 2200	Roberts et al. (2019)	30 m
Geological data		Silica	Gray et al. (2016)	1:250 000

5.2.7. Data analysis, modelling and mapping

Across both fields, data frames containing measured pH, ECe, CEC, ESP, Ca:Mg, sand content and clay content were created. Boxplots were used to visualise variability in measured data between production zones within each field. The Pearson correlation coefficient was used to understand relationships between point soil data and crop yields at each sample depth with a correlation matrix produced for each paddock.

5.2.7.1. Development of models for soil properties

Two approaches were used to model the target soil properties of pH, ECe, CEC, ESP, clay (%) and sand (%). At Farm A (n=8) linear models were used on account of the low sample size. At Farm C (n=18), linear models were compared with random forest (RF) models. For all models, the mid-depth of each layer being assessed (0.05, 0.2, 0.45, 0.7, 0.9 m) was included as a predictor variable, allowing for the modelling of each soil layer concurrently. Implementing

this '3D' modelling approach increases the cohesivity and interpretability of models and their outputs while also increasing the number of samples used to build the model (Roudier et al., 2020). Consequently, the number of samples used to build models increased to 40 (Farm A) and 90 (Farm C).

For each soil property, graphical and statistical summaries were used to identify potential outliers and determine if data transformations were required. If outliers were visually present, and returned a z-score > 3, they were removed. On this basis, one high value outlier was removed from the ECe dataset at Farm C. The overall skewness of the dataset was also measured along with the relationship to predictor variables to consider if data transformations were required.

Following the preparation of each soil property dataset, all remote and proximally sensed covariates were included in the initial model. Model assumptions were assessed before a Variance Inflation Factor (VIF) function was used from within the 'car' package in R to remove variables on the basis of collinearity (Fox & Weisberg, 2019). In this instance, predictor variables containing a VIF >10 were removed from the model. A simultaneous backwards and forwards elimination was then undertaken using the stepAIC function within the R 'MASS' package to produce the most parsimonious model (Venables & Ripley, 2002). As the models were used to predict all depths concurrently, depth was forced into each model at the beginning of the process and was not allowed to be eliminated. This was repeated for each soil property. Due to the higher number of samples at Farm C, RF models were also built for each soil property to allow for a comparison of modelling approaches. Each of the RF models included the same covariates as the corresponding linear model.

5.2.7.2. Model validation

Model performance was assessed using a leave-one-site-out-cross-validation (LOSOCV). This process trains the dataset on n-1 sites, with prediction occurring at all depths at the removed site. This is repeated for all sites. It is considered beneficial for smaller datasets as it does not rely on splitting the data into calibration and validation subsets. Model quality was assessed using the Lin's Concordance Correlation Coefficient (LCCC). As this statistic is unitless it allows for comparisons between models (Lin, 1989). A LCCC value closer to 1 indicates a higher quality model. In addition to this, the Root Mean Square Error (RMSE) and bias are also reported.

5.2.7.3. Map production

For each soil property the selected model was predicted onto a 5 m grid of each paddock utilising the spatial covariates included in the final model. As linear models can predict beyond the extent of the training dataset, for each soil property a minimum and maximum prediction was set. The minimum value was set to 5% below the minimum measured value, unless this was negative, and the maximum 5% above the maximum measured value. This ensured minimal overextrapolation of predicted values. If data was transformed, an appropriate back transformation took place. Data was then subset into the five sample depths, converted to a spatial dataframe and projected as a map.

5.3. Results

5.3.1. Production zone outputs

There is variability in the strength of relationships between crop yields across different years at both Farm A and Farm C (Figure 5.5). With respect to cotton production, the strongest correlations are observed at Farm C between the 2018 and 2019 cotton yields (0.43) and 2019 cotton and barley (0.36) (Figure 5.5). Farm A presents only a very weak correlation between 2020 wheat yield and 2018 cotton yield (0.18), however, a moderately positive correlation is observed between the 2020 and 2021 wheat yields (0.41) (Figure 5.5).

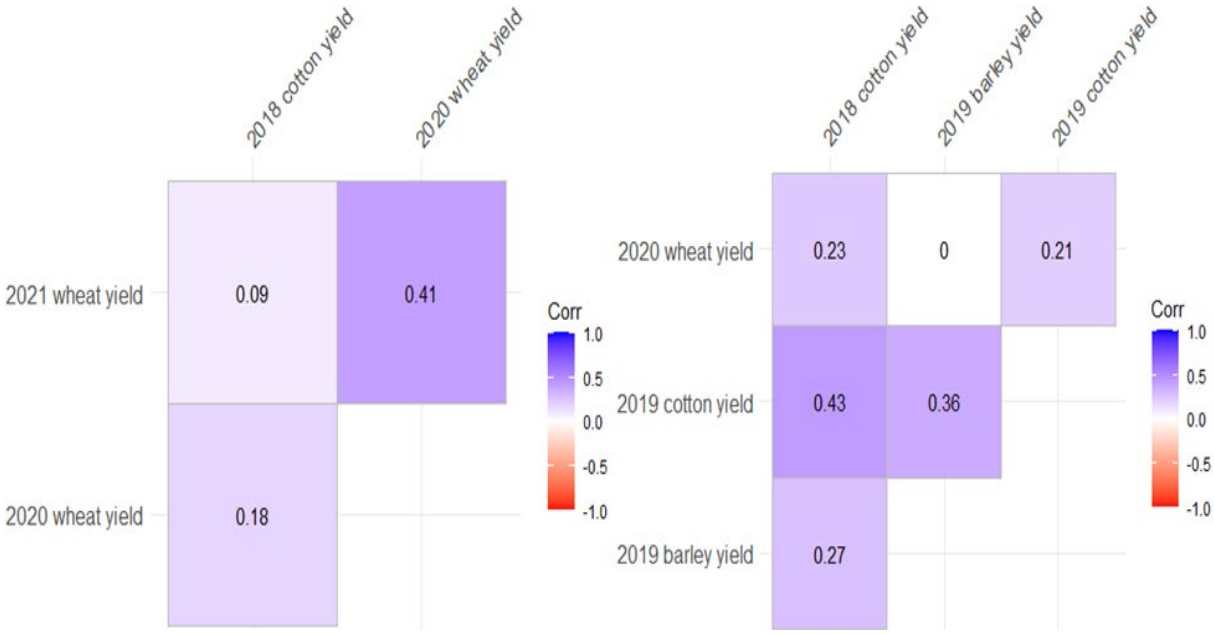


Figure 5.5. Correlation matrices between cotton (2018) and wheat (2020, 2021) yields in the examined field at Farm A (left) and between cotton (2018, 2019), barley (2019) and wheat (2020) in the field at Farm C (right). A correlation (r) closer to 1 indicates a strong, positive relationship while a number closer to -1 a stronger negative relationship.

Based on these correlations and discussions with stakeholders it was decided that 2020 wheat and 2018 cotton would be used to develop production zones at Farm A. At Farm C 2018 and 2019 cotton as well as 2019 barley were used. At both Farm A (Figure 5.6) and Farm C (Figure 5.7) Cluster 1 represents the higher, and Cluster 2 the lower, yielding zone as shown by extended boxplot whiskers and lower median yields (Figure 5.3; Figure 5.4). These clusters align with spatially observed extremities in the raw yield data.

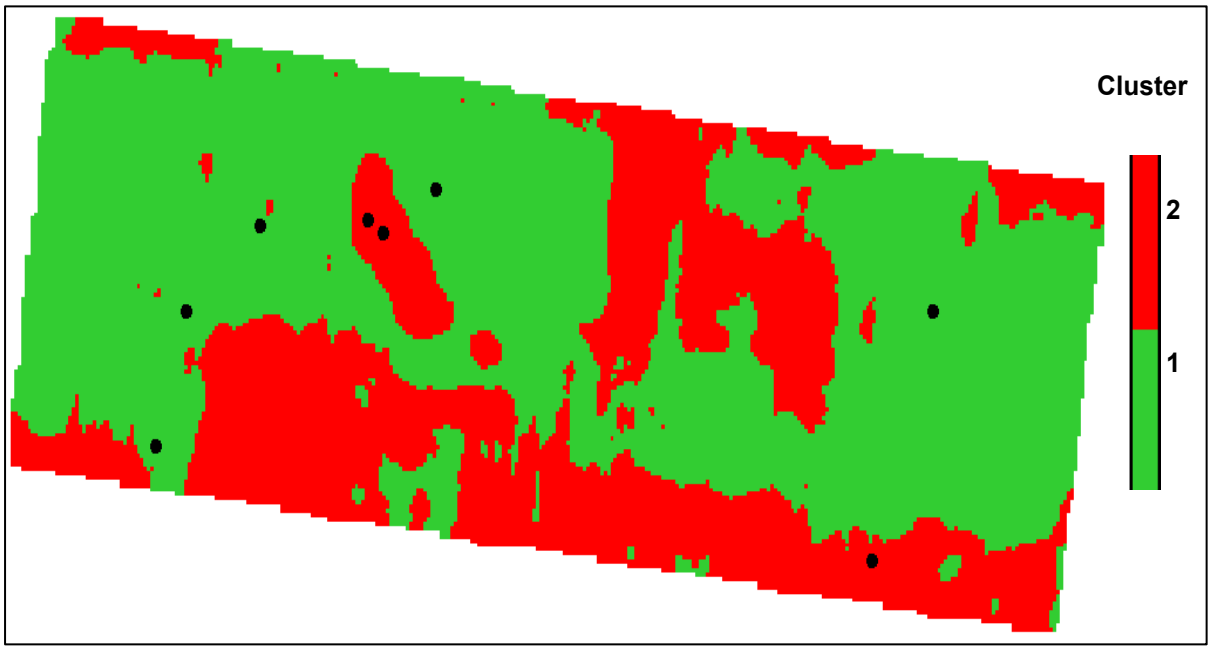


Figure 5.6. Production zones at Farm A developed through k-means clustering incorporating 2018 cotton and 2020 wheat yields. Selected sampling sites are shown as black dots. At each site, soil cores were extracted to a depth of one metre.

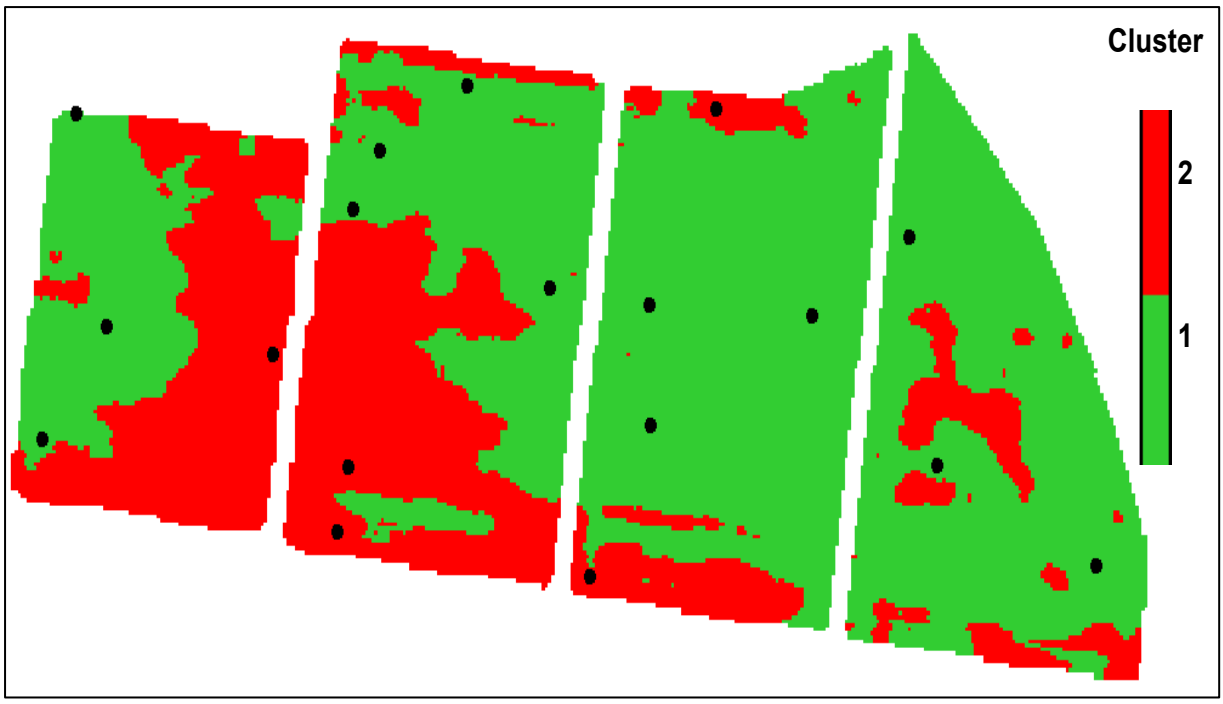


Figure 5.7. Production zones at Farm C developed through k-means clustering incorporating 2018 cotton, 2019 cotton and 2019 barley yields. Sites where soil cores were extracted to a depth of one metre are shown as black dots.

The 2018 cotton crop at Farm A has a minimum yield of 4.2 bales/ha in Cluster 2 as opposed to 6.7 bales/ha in Cluster 1 (Figure 5.8). The 2018 and 2019 cotton crops at Farm C have minimum yields of 6.3 and 6.4 bales/ha, respectively, in Cluster 2 compared to minimum corresponding yields of 11.4 and 11.5 bales/ha in the higher yielding Cluster 1 (Figure 5.9).

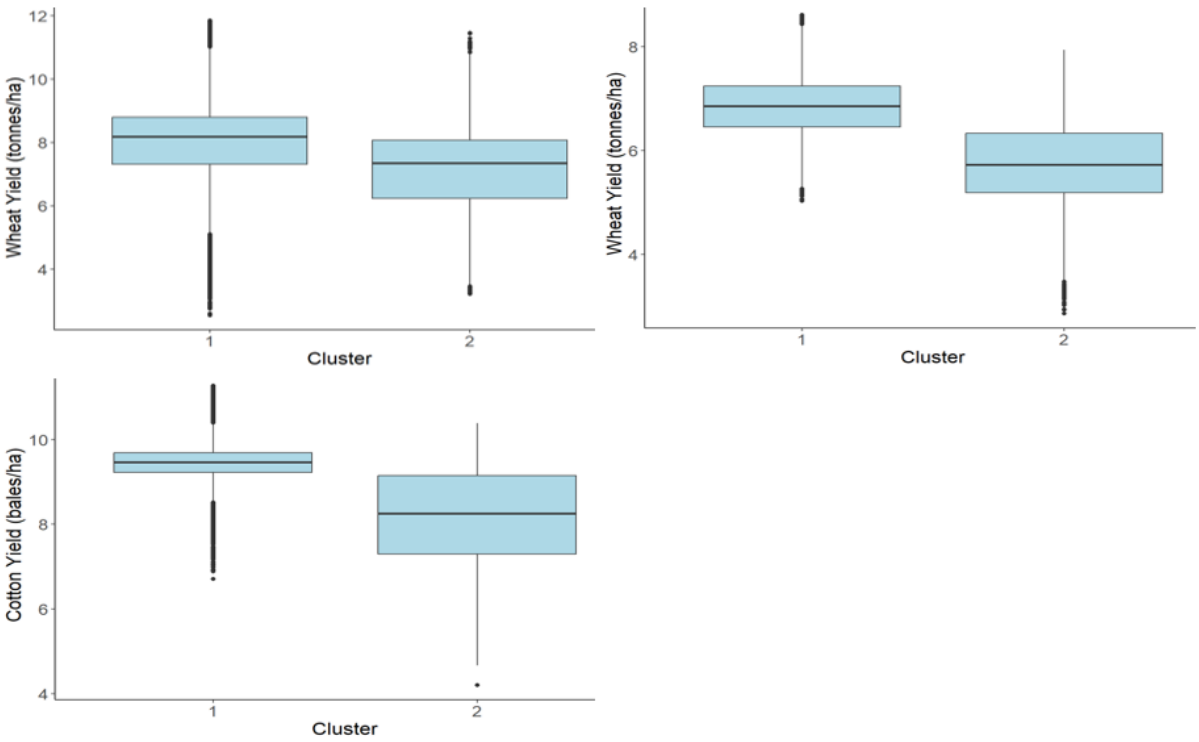


Figure 5.8. Variability in crop yield for the two production zones (clusters) for 2020 (top right) and 2021 (top left) wheat (tonnes/ha) and 2018 cotton (bales/ha) (bottom left) crops at Farm A.

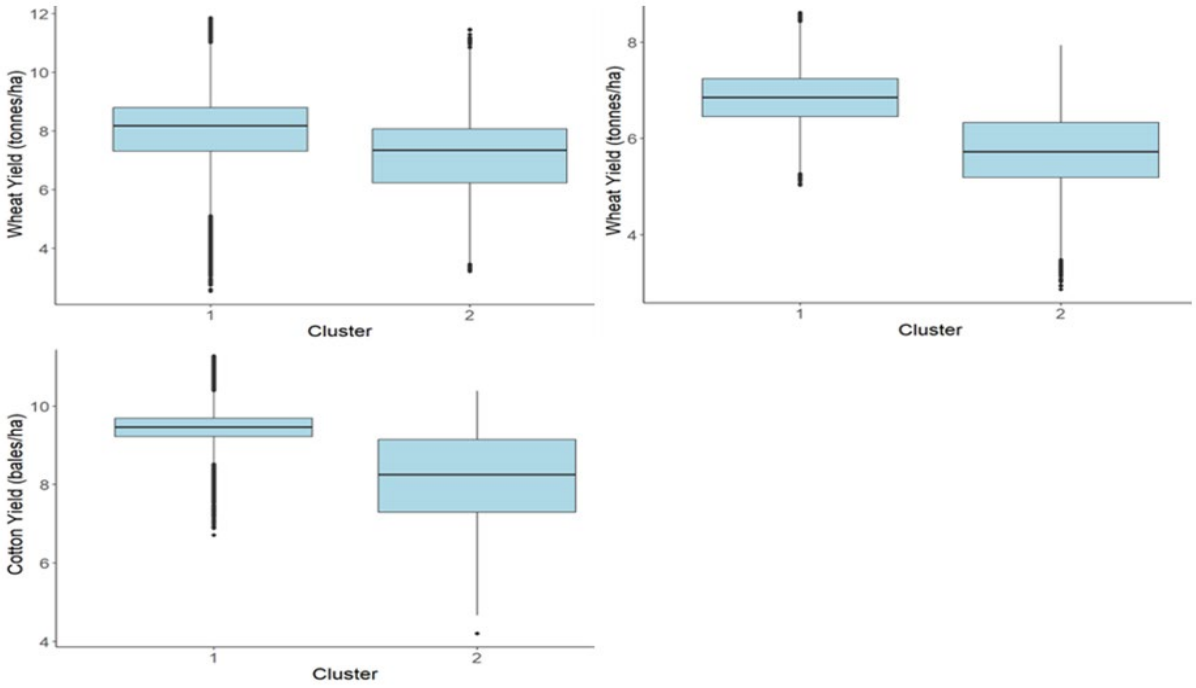


Figure 5.9. Variability in crop yield between the two production zones (clusters) for 2019 barley (tonnes/ha), 2020 wheat (tonnes/ha), and 2018 (bottom right) and 2019 (bottom left) cotton (bales/ha) crops at Farm C.

5.3.2. Soil variability between production zones

5.3.2.1. pH

There is a general trend of increasing alkalinity with depth across all clusters at both farms (Figure 5.10). At Farm A, both clusters exhibit the highest median pH at 0.6-0.8 m, with a decrease at 0.8-1 m. There is significant variability within clusters at depths of 0-0.1 and 0.1-0.3 m at Farm C and at all depths in Cluster 1 at Farm A. Despite the within cluster variability, the higher yielding Cluster 1 at Farm A exhibits greater alkalinity than Cluster 2. The differentiation between clusters at Farm C is less clear. At depths of 0.3-0.6, 0.6-0.8 and 0.8-1 m the pH of clusters 1 and 2 approaches and on occasion surpasses 9, a point where alkalinity is likely to impede plant growth (Figure 5.10).

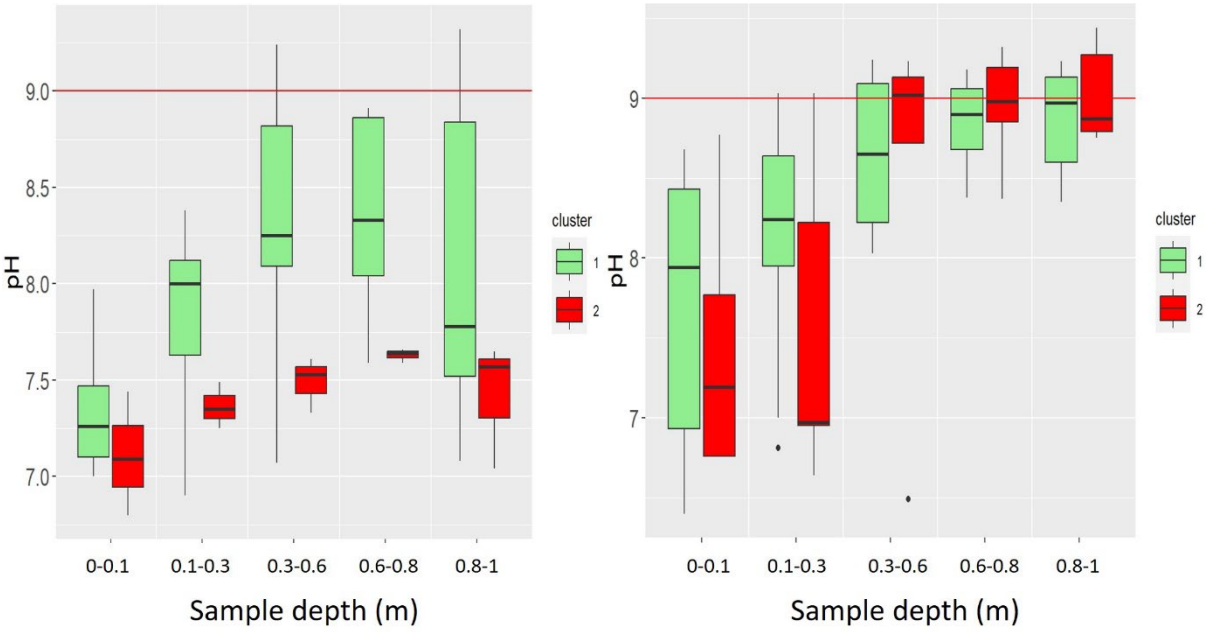


Figure 5.10. The pH variability between the two production zones (clusters) at five sample depths for the field at Farm A (left) and Farm C (right). The solid red line represents the point (pH 9) where alkalinity is likely to significantly impact plant growth (Hazleton & Murphy, 2007).

5.3.2.2. Electrical conductivity of the extract (ECe)

ECe increases with depth in both clusters at Farm C as well as in Cluster 1 at Farm A (Figure 5.11). At Farm A there is a clear difference in ECe between clusters at depth, where some samples in Cluster 1 increase to exceed the salinity tolerance of wheat (6.0 dS/m) and cotton (7.7 dS/m) at 0.8-1 m. Conversely, the measured ECe in Cluster 2 is below these levels, decreasing slightly with depth. At Farm C, the median ECe of Cluster 1 is greater than Cluster 2 at all depths excluding the surface. Despite increasing ECe at depth, no measured values at Farm C cross the salinity threshold for wheat or cotton (Figure 5.11).

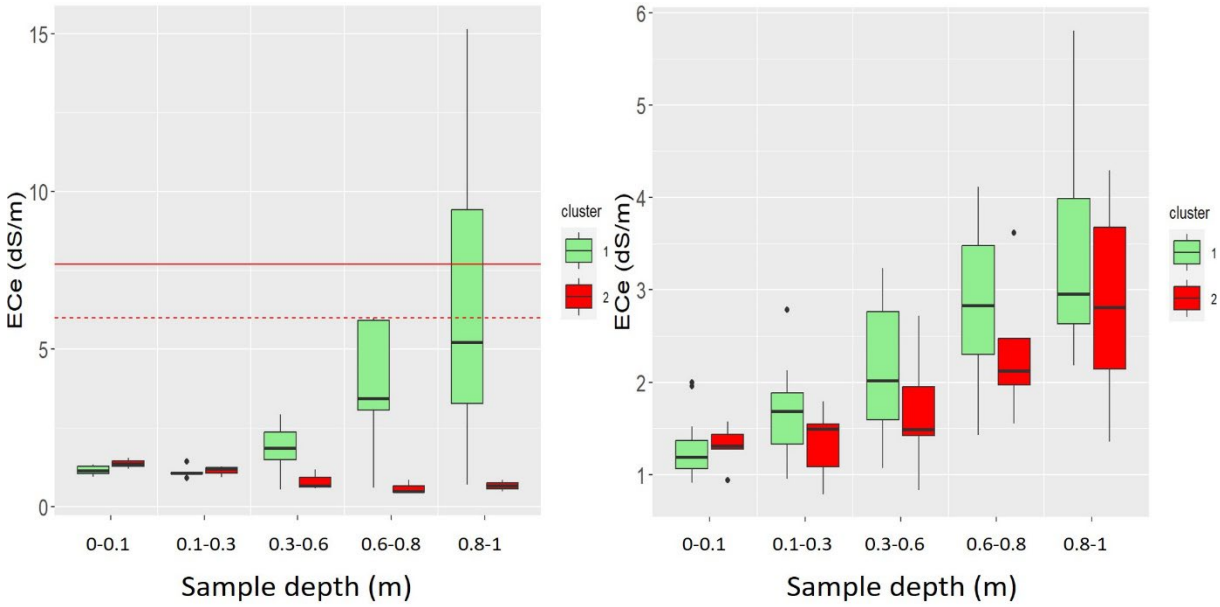


Figure 5.11. Variability at the five sampled depths for ECe (ds/m) between the two production zones (clusters) at Farm A (left) and Farm C (right). The dashed and solid red lines indicate where salinity will significantly impact wheat (6) and cotton (7.7) growth, respectively (Hazleton & Murphy, 2007).

5.3.2.3. Cation exchange capacity (CEC)

There is a clear difference in CEC between clusters at Farm A (Figure 5.12). Cluster 1 exhibits a significantly higher CEC at all depths, with the highest median values at 0.3-0.6 and 0.6-0.8 m before a slight decrease at 0.8-1 m. The inverse is seen in the lower yielding Cluster 2. A similar, although less distinct, trend is observed at Farm C, with a higher median CEC in Cluster 1 at all depths. In Cluster 1 there is higher variability at depths of 0-0.1 and 0.1-0.3 m while Cluster 2 exhibits greater within cluster variability below 0.3 m (Figure 5.12).

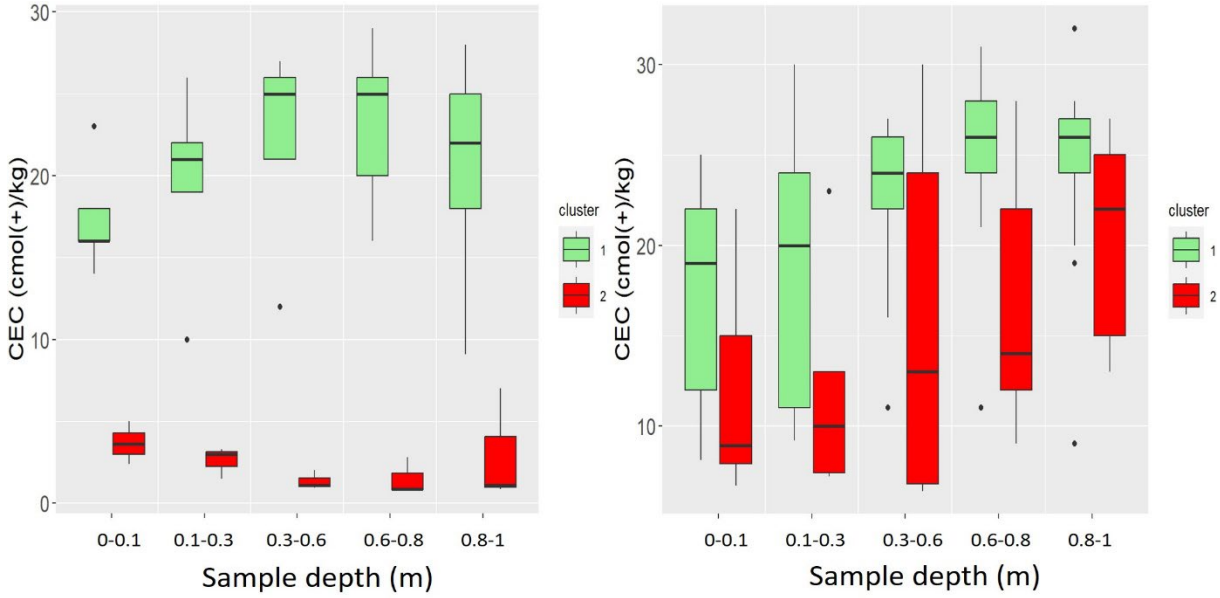


Figure 5.12. Cation Exchange Capacity (CEC) (cmol(+)/kg) variability between the two production zones (clusters) at five sample depths for the field at Farm A (left) and Farm C (right).

5.3.2.4. *Exchangeable sodium percentage (ESP)*

Clusters in both fields exhibit similar ESP values at depths of 0-0.1 and 0.1-0.3 m (Figure 5.13). At Farm A, there is a clear delineation of ESP between clusters at depths of 0.3-0.6, 0.6-0.8 and 0.8-1 m where ESP increases towards 30% in Cluster 1, exceeding potential sodicity thresholds. Conversely, ESP slightly decreases at these depths in Cluster 2. A similar trend is observed between clusters at Farm C. While the median ESP decreases in Cluster 2 below 0.3 m, there is within-cluster variability and some samples exceed 10%. Median values for Cluster 1 at Farm C increase with depth. In this cluster some samples exceed 10% ESP at 0.3-0.6 and 0.6-0.8 m, and 15% ESP at 0.8-1 m (Figure 5.13).

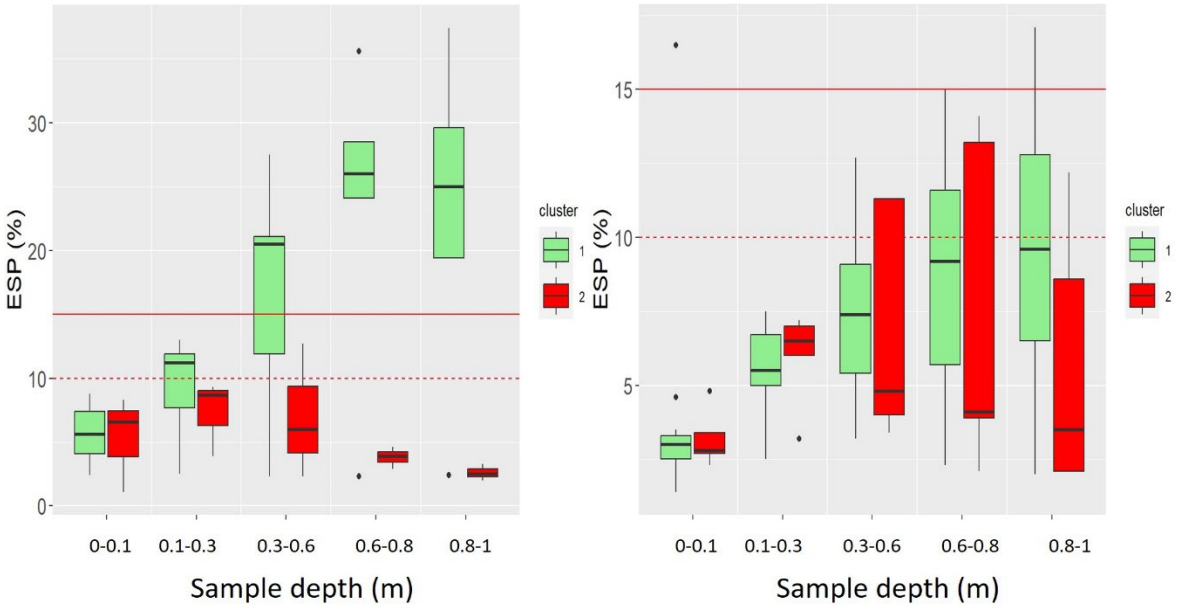


Figure 5.13. Variability in Exchangeable Sodium Percentage (ESP) (%) between production zones (clusters) at Farm A (left) and Farm C (right). The solid red line (15%) and dashed red line (10%) indicate two potential sodicity thresholds (Hazleton and Murphy, 2007; Pozza et al., 2022).

5.3.2.5. *Calcium to magnesium ratio (Ca:Mg)*

There is variability in the Ca:Mg between clusters at Farm A, with Cluster 2 having a higher median ratio at all depths (Figure 5.14). There are measured values below a level of 1 at all depths for Cluster 2, however, the median remains above this level. Both clusters, in both fields, show a decreasing trend in Ca:Mg from the surface to the 0.3-0.6 m before the ratio increases beyond this depth. At Farm C there is less distinction in the Ca:Mg between clusters. The median ratio is highest for both clusters at the surface (0-0.1 m). The ratio falls below a level of 1 at one site, in Cluster 1 at 0.1-0.3 m, and multiple sites in both clusters at depths of 0.3-0.6, 0.6-0.8 and 0.8-1 m. Within cluster variability also increases with depth (Figure 5.14).

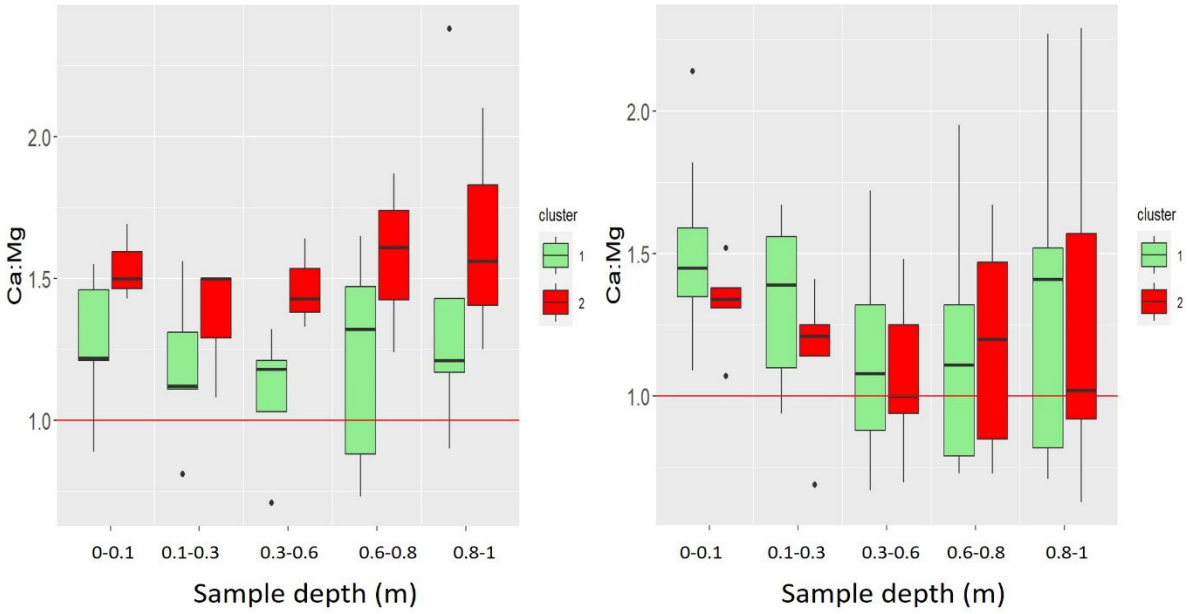


Figure 5.14. Variability of the Calcium to Magnesium ratio (Ca:Mg) between the two production zones (clusters) at the five sample depths for the field at Farm A (left) Farm C (right). The solid red line represents the point where calcium becomes deficient. A ratio lying between 1-4 suggest low calcium and <1 suggests a calcium deficiency.

5.3.2.6. Soil texture

Across both fields where sand content increases, clay content decreases. Due to this relationship, where previous figures displayed plots for each farm by soil property, this section will instead display sand and clay content adjacent to each other for each farm. There is a clear distinction in soil texture between clusters at Farm A (Figure 5.15). Median values of sand content exceed 75% at all depths in Cluster 2 while in Cluster 1 the median is below 37% at all depths (Figure 5.15). The inverse is seen in clay content, with values significantly greater in Cluster 1. There is no clear trend of texture changing at depth in either cluster (Figure 5.15).

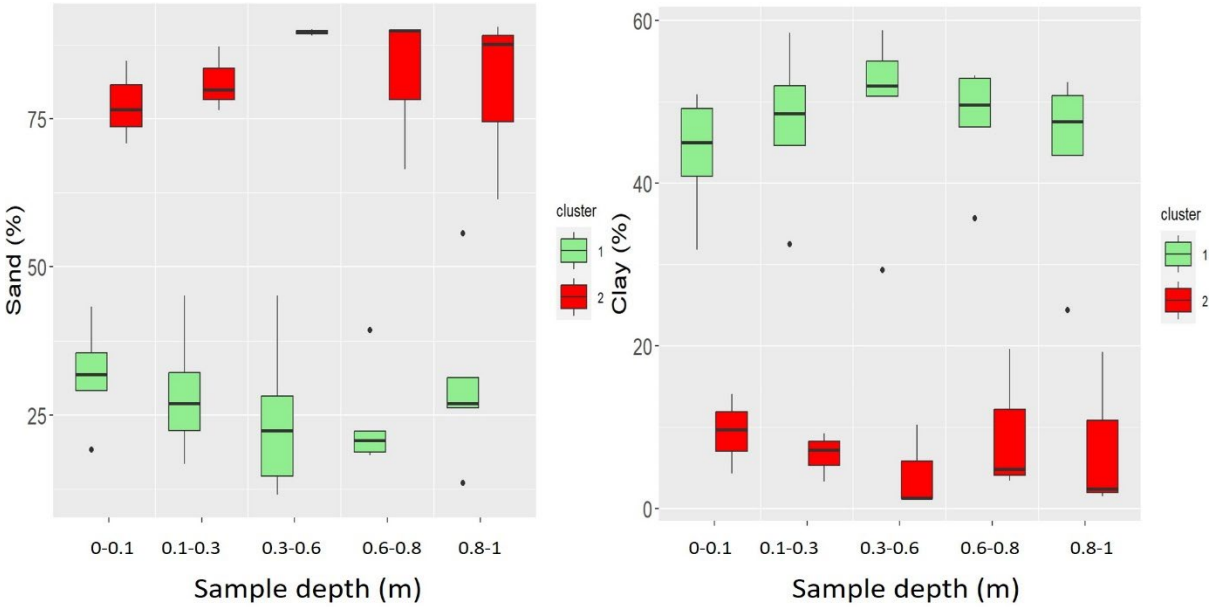


Figure 5.15. Variability of the sand (left) and clay (right) fractions as a percentage of soil between the two production zones (clusters) at the five measured soil depths in the field at Farm A.

At Farm C there is a trending change of soil texture at depth, with soils becoming more clayey down the profile (Figure 5.16). At all depths there is a higher median clay and lower median sand content in Cluster 1. There is greater within-cluster variability in Cluster 2 (Figure 5.16).

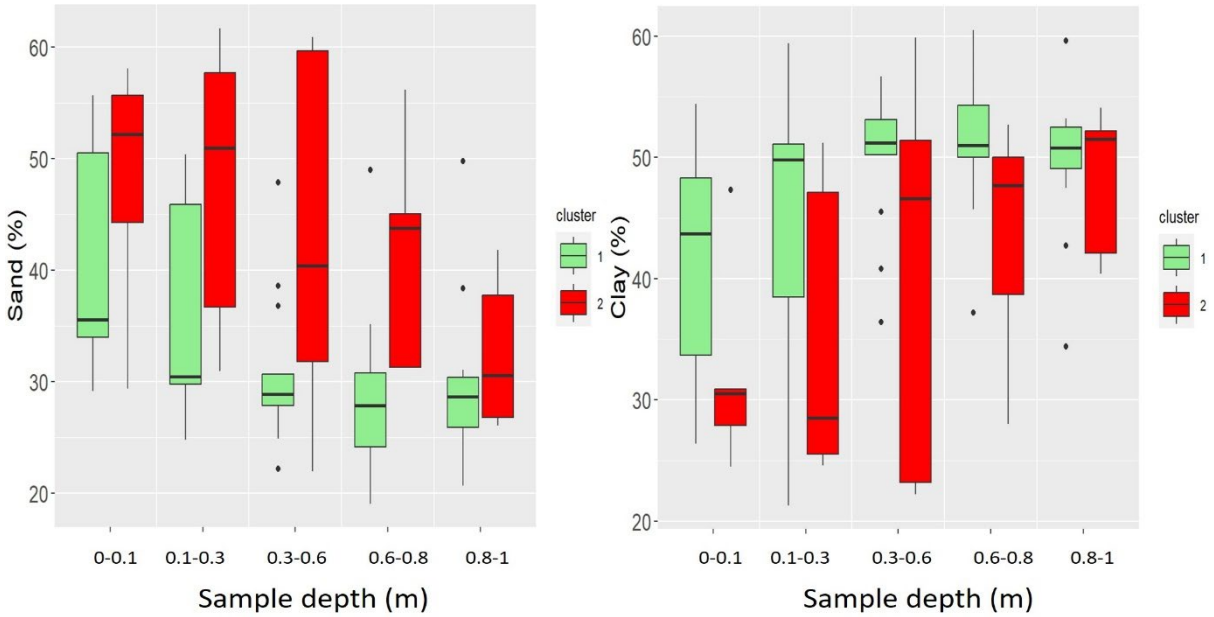


Figure 5.16. Variability of the sand (left) and clay (right) fractions as a percentage of soil between the two production zones (clusters) at the five measured soil depths in the field at Farm C.

5.3.3. Soil point data correlations to crop yield

For this analysis the strength of correlations will be defined as strong (0.7-1), moderate (0.5-0.7), weak (0.3-0.5) and very weak (0-0.3). Where the correlation statistic (r) is >0 the variable shows a positive relationship while values <0 indicate a negative relationship. A value of 0

indicates that the variables have no relationship while values of 1 or -1 indicate a perfect positive or negative relationship, respectively.

5.3.3.1. Farm A

The soil properties most strongly correlated to crop yields are sand content and clay content, representing soil texture. These correlations are strongest at 0.6-0.8 m where sand and clay share strong correlations of -0.96 and 0.95 to cotton yield, respectively (Figure 5.17). At all depths, correlations are >0.89 (clay) and <-0.87 (sand) (Figure 5.17).

As individual cations (K^+, Mg^{2+} and Ca^{2+}) and CEC are strongly positively correlated to clay content, they are also moderately-to-strongly positively correlated to cotton yield at all depths, with values >0.72 (Figure 5.17). ESP is moderately positively correlated to cotton yield at 0.3-0.6 and 0.6-0.8 m (0.75), however, exhibits a very weak, but still positive correlation at the surface (0.12).

The influence of ECe on cotton yield varies with depth, exhibiting a moderately negative correlation at the surface (-0.54) compared to a moderately strong correlation (0.68) at 0.6-0.8 m (Figure 5.17). The relationship between pH and cotton yield increases from very weak (0.27) at the surface to moderate (0.71) at 0.6-0.8 m. Similar correlation trends are observed for 2020 and 2021 wheat yields, however, the strongest correlations are with the 2021 wheat crop yields and ECe (0.76), pH (0.89) and ESP (0.92) at 0.6-0.8 m (Figure 5.17).

There are also strong correlations between soil properties, with these being most significant at 0.6-0.8 m (Figure 5.17). Sand exhibits a moderate negative correlation to ESP (-0.62) and ECe (-0.67) and strong negative correlations to all other soil properties. Correlations between clay content and soil properties are the inverse of sand. A weak negative correlation is seen between pH and ECe the surface (-0.41) compared to a moderately positive correlation of 0.58 at 0.6-0.8 m while the strength of the correlation between pH and ESP increases from 0.18 to 0.93 between these depths (Figure 5.17).

5.3.3.2. Farm C

At Farm C the strongest correlations between cotton yield and soil attributes are with sand content and clay content at 0.3-0.6 and 0.6-0.8 m (Figure 5.18). At 0.3-0.6 m the relationship between sand content and 2018 cotton (-0.64) is more strongly negative than 2019 cotton (-0.53). The relationship becomes more similar at 0.6-0.8 m with values of -0.58 (2018 cotton) and -0.60 (2019 cotton). The inverse is true for clay content, which is more strongly correlated to 2018 cotton at 0.6-0.8 m (0.66) than 0.3-0.6 m (0.59) (Figure 5.18). Despite this variability

the relationship between subsoil texture and cotton lint yields is clear, with more clayey soils positively correlated to crop yield. Generally, the depths where the strongest correlations occur between soil attributes and cotton yield varies more at Farm C than at Farm A.

As CEC is moderately-to-strongly correlated to clay content it shares similar correlations with cotton yields (Figure 5.18). These are strongest at 0.6-0.8 m, with moderate positive correlations to both 2018 (0.52) and 2019 (0.58) cotton yield. ESP is most strongly correlated to 2018 cotton yield at 0.3-0.6 m (0.49) and 2019 cotton at 0.6-0.8 m (0.44).

A very weak, negative correlation is seen between pH and 2019 cotton yield at the surface (-0.11) and 0.8-1 m (-0.28) with a moderate correlation to 2018 yield at this depth (-0.35). Soil ECe is weakly negatively correlated to 2019 cotton yield (-0.11) at the surface compared to a moderately positive correlation (0.51) at 0.6-0.8 m (Figure 5.18).

Correlations are weaker between soil properties and non-cotton crop yields (Figure 5.18). Sand content (-0.71) and clay content (0.68), however, exhibit strong or moderate correlations with barley at 0.6-0.8 m. At 0-0.1 and 0.1-0.3 m pH shows moderate negative correlations with wheat of -0.58 and -0.49, respectively. At these depths there is only a very weak correlation between pH and cotton yield. At 0-0.1 and 0.1-0.3 m ECe showcases correlations of 0.25 and -0.4 with wheat yield, respectively (Figure 5.18).

The strongest correlations between soil properties are seen between texture (sand and clay) and exchangeable cations at 0.3-0.6 m (Figure 5.18). Unlike Farm A, pH has a moderate correlation to ECe at all depths, with the relationship strongest at 0.1-0.3 m (0.55). Soil ESP shares a negative correlation to pH (-0.06) and ECe (-0.37) at 0.1-0.3 m compared to moderately positive correlations of 0.55 and 0.51, respectively, at 0.6-0.8 m. CEC is moderately positively correlated to pH at 0-0.1 m (0.67) compared to a negative correlation (-0.31) at 0.6-0.8 m (Figure 5.18).

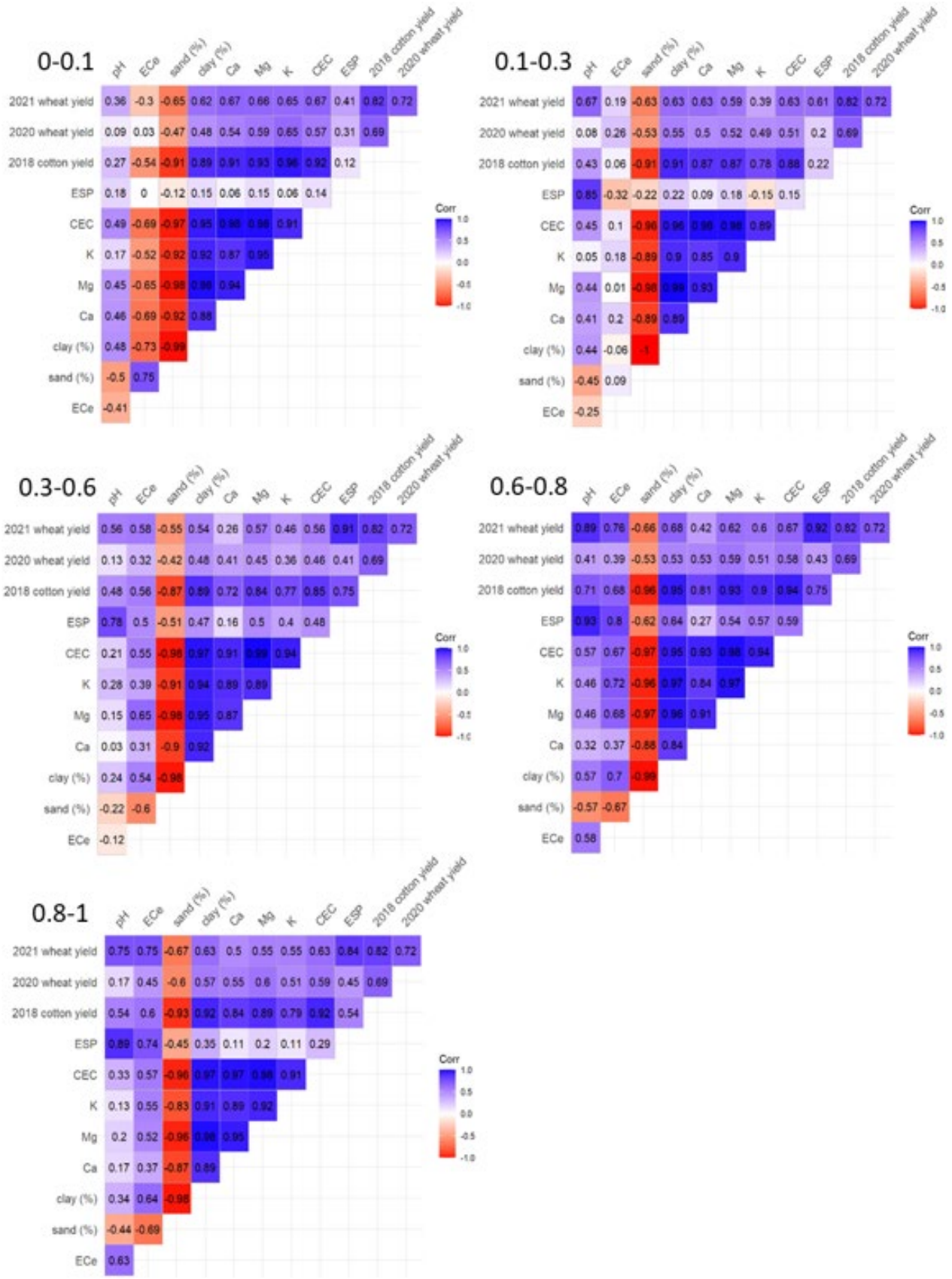


Figure 5.17. Point data correlation matrices between measured soil properties and crop yields at the five sample depths (0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1) for sample points at Farm A. Values approaching 1, dark blue, show a strong positive correlation while values approaching -1, dark red, indicate a strong negative correlation. A value of 0 indicates there is no correlation between data.

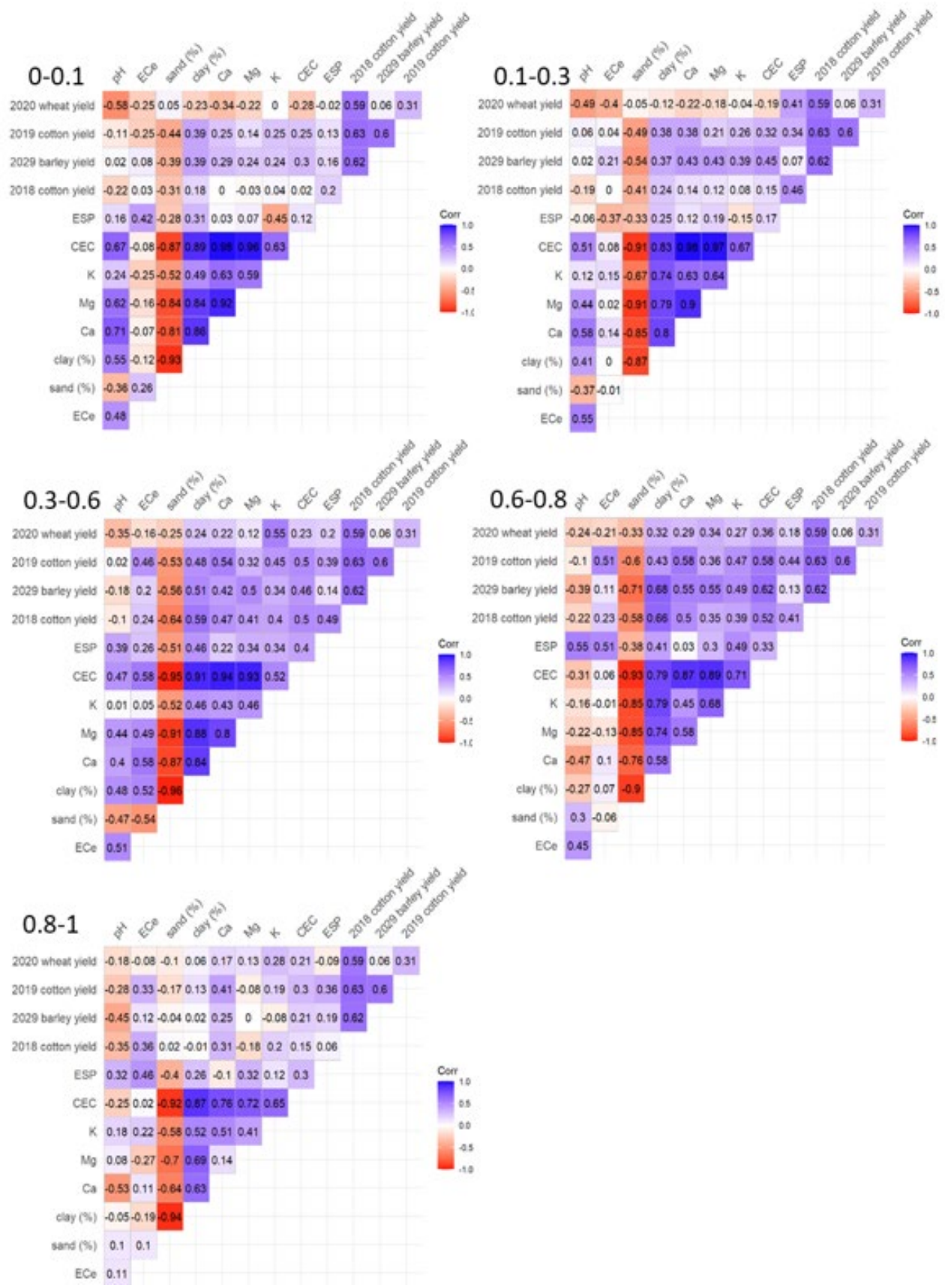


Figure 5.18. Point data correlation matrices between measured soil properties and crop yields at the measured depths (0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1) for sample points at Farm C. Values approaching 1, dark blue, show a strong positive correlation while values approaching -1, dark red, indicate a strong negative correlation. A value of 0 indicates there is no correlation between data.

5.3.4. Soil spatial variability

5.3.4.1. Final model covariates and validation statistics

To allow for concurrent predictions at multiple depths, layer mid-depth (0.05, 0.2, 0.45, 0.7, 0.9 m) was included as a predictor variable in all models. At Farm A, the model for ECe contained only mid-depth and ECa 300 as predictor variables (Table 5.2). In all other models at Farm A mid-depth, ECa 300 and Gamma U were the predictors that remained following the VIF function and stepwise deletion of variables (Table 5.2). Using a LOSOCV the highest quality linear models generated at Farm A were for CEC, ESP and sand (%), with a moderate LCCC of 0.65 (Table 5.2). The lowest quality model, as indicated by the LCCC obtained through the LOSOCV, was for ECe, with a LCCC of 0.45 (Table 5.2). All models exhibit high RMSE with signs of over and under prediction for sand and clay content, respectively (Table 5.2). The CEC dataset was transformed prior to modelling. This occurred as the initial prediction of CEC were poor with a LOSOCV LCCC of 0.51 which improved to 0.65 after undergoing a logarithmic transformation.

Table 5.2. Model quality statistics for the linear models developed at Farm A. * Indicates that data was transformed

Attribute	Model covariates	LCCC	RMSE	Bias
pH	Mid-depth ECa 300 Gamma U	0.55	0.53	-0.01
ECe	Mid-depth ECa300	0.45	0.76	0.01
CEC*	Mid-depth ECa 300 Gamma U	0.65	0.96	-0.12
ESP	Mid-depth ECa 300 Gamma U	0.65	7.51	0.56
Sand (%)	Mid-depth ECa 300 Gamma U	0.65	22.68	2.06
Clay (%)	Mid-depth ECa 300 Gamma U	0.57	17.97	-1.32

A wider variety of predictor variables remained in the final models at Farm C (Table 5.3). All soil property models included depth, ECa 50 and Gamma K (Table 5.3). Unlike Farm A, remotely sensed covariates remained in models following the VIF function and stepwise deletion. Silica was a predictor variable in all models, excluding ESP; TWI was a predictor for ECe, CEC, ESP and clay (%); and NIR was used in models for sand (%), ESP and ECe (Table

5.3). Linear models produced better model statistics than RF models for all soil attributes at Farm C and were, therefore, chosen to make predictions for all soil properties (Table 5.3). Using a LOSOCV the highest quality model was for CEC, with a LCCC of 0.73, while the lowest quality linear model was for ESP, with a LCCC of 0.62 (Table 5.3). There is negligible bias for each soil attribute model. With the exclusion of pH, the RMSE is lower than for the respective models at Farm A (Table 5.2; Table 5.3). The RMSE cannot be compared between farms for the CEC models as the data at Farm A had been transformed.

Table 5.3. Model quality statistics for both linear and random forest models developed to predict soil attributes at Farm C. ^Indicates that an outlier was removed prior to modelling

Attribute	Model covariates	Model type	LCCC	RMSE	Bias
pH	Mid-depth Elevation ECa 50 Gamma K Gamma Th Silica	Linear	0.68	0.57	-0.01
		Random Forest	0.62	0.56	-0.03
ECe^	Mid-depth ECa 50 Gamma K Gamma Th NIR TWI Silica	Linear	0.72	0.61	0.02
		Random Forest	0.59	0.68	0.03
CEC	Mid-depth Elevation ECa 50 Gamma K TWI Silica	Linear	0.73	4.80	0.19
		Random Forest	0.50	5.75	0.11
ESP	Mid-depth ECa 50 Gamma K Gamma Th NIR TWI	Linear	0.62	2.95	0.09
		Random Forest	0.39	3.32	0.12
Sand (%)	Mid-depth Elevation ECa 50 Gamma K NIR Silica	Linear	0.70	7.49	-0.10
		Random Forest	0.47	8.76	-0.42
Clay (%)	Mid-depth Elevation ECa 50 Gamma K TWI Silica	Linear	0.69	6.97	0.41
		Random Forest	0.50	7.94	0.50

5.3.4.2. Digital soil maps

5.3.4.3. Farm A

There is a clear spatial relationship between the maps of CEC, sand content and clay content, with sand content lower where CEC and clay content are higher (Figure 5.19). This trend matches yield variability and is consistent with cluster variability and point correlations. For these three soil properties, while there is significant spatial variability, there is minimal change predicted vertically between the five mapped depths (Figure 5.19). The maps of ESP, pH and ECe showcase a similar trend spatially but with greater differentiation vertically. As depth increases, the predicted values of ECe, ESP and pH also increase (Figure 5.20).

5.3.4.4. Farm C

There is a strong spatial relationship between the CEC, sand and clay content maps, however, unlike Farm A there is clearer change in predicted values with depth (Figure 5.21). The CEC, clay and sand maps show slight banding relating to the elevation covariate data, however, the models still appropriately predict trends consistent with point data measurements (Figure 5.21). There is predicted variability both spatial and vertically for ESP, ECe and pH with each attribute predicted to increase down the soil profile (Figure 5.22).

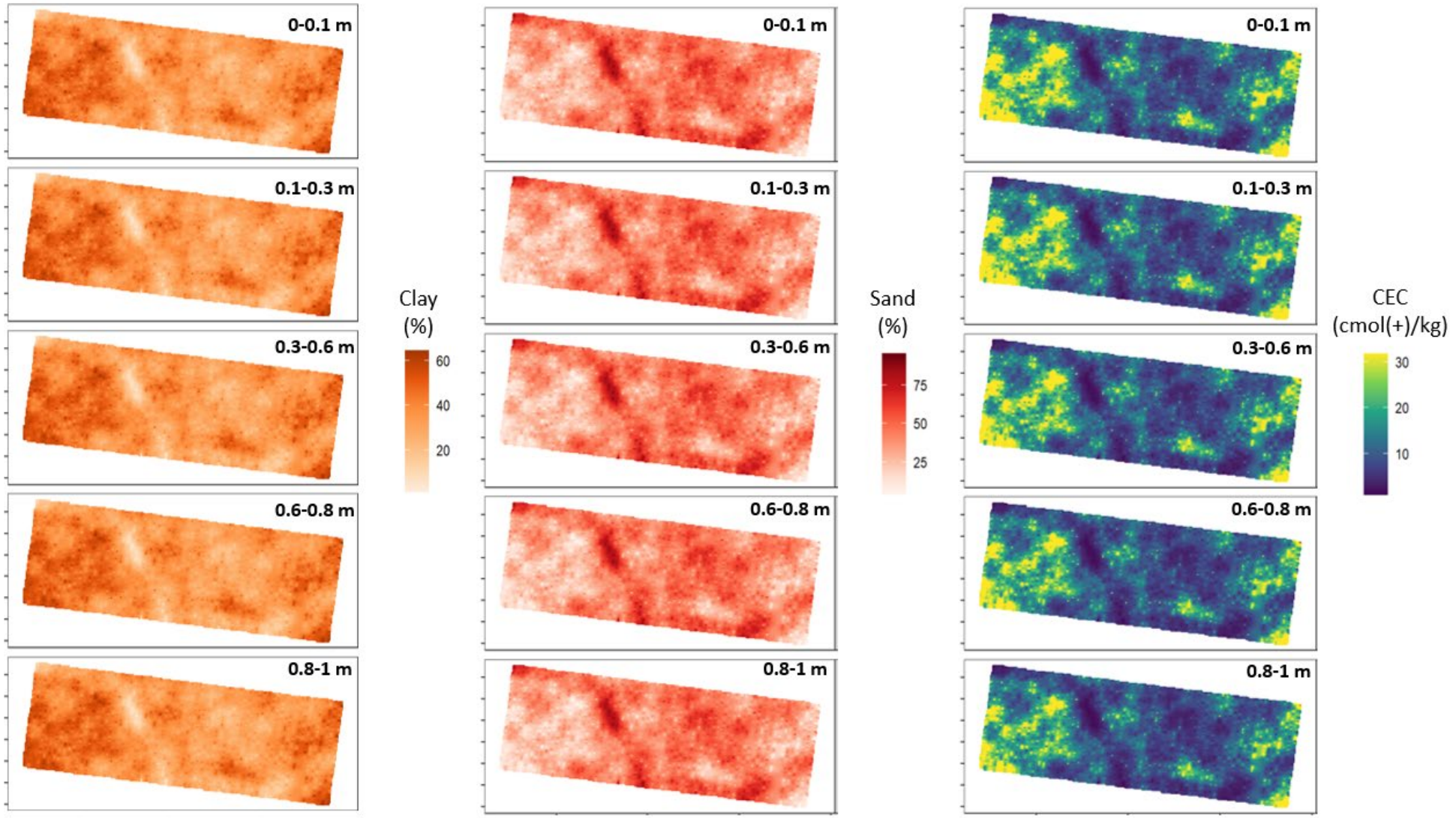


Figure 5.19. Digital soil maps at a 5 m resolution for clay (%) (left), sand (%) (centre) and cation exchange capacity (CEC) (cmol(+)/kg) (right) at Farm A

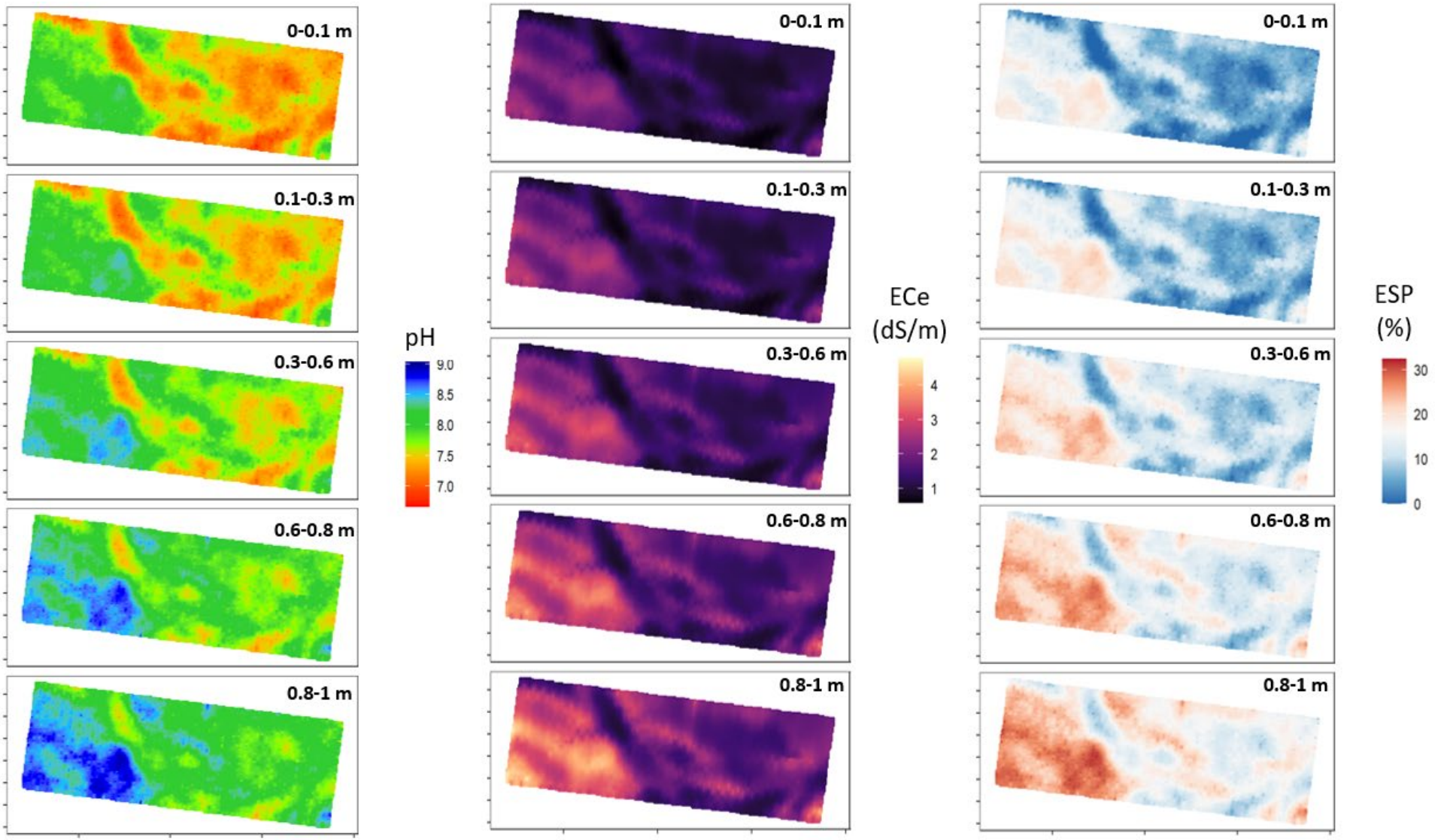


Figure 5.20. Digital soil maps at a 5 m resolution for pH (left), electrical conductivity of the extract (ECe) (dS/m) (centre) and exchangeable sodium percentage (ESP) (right) at Farm A.

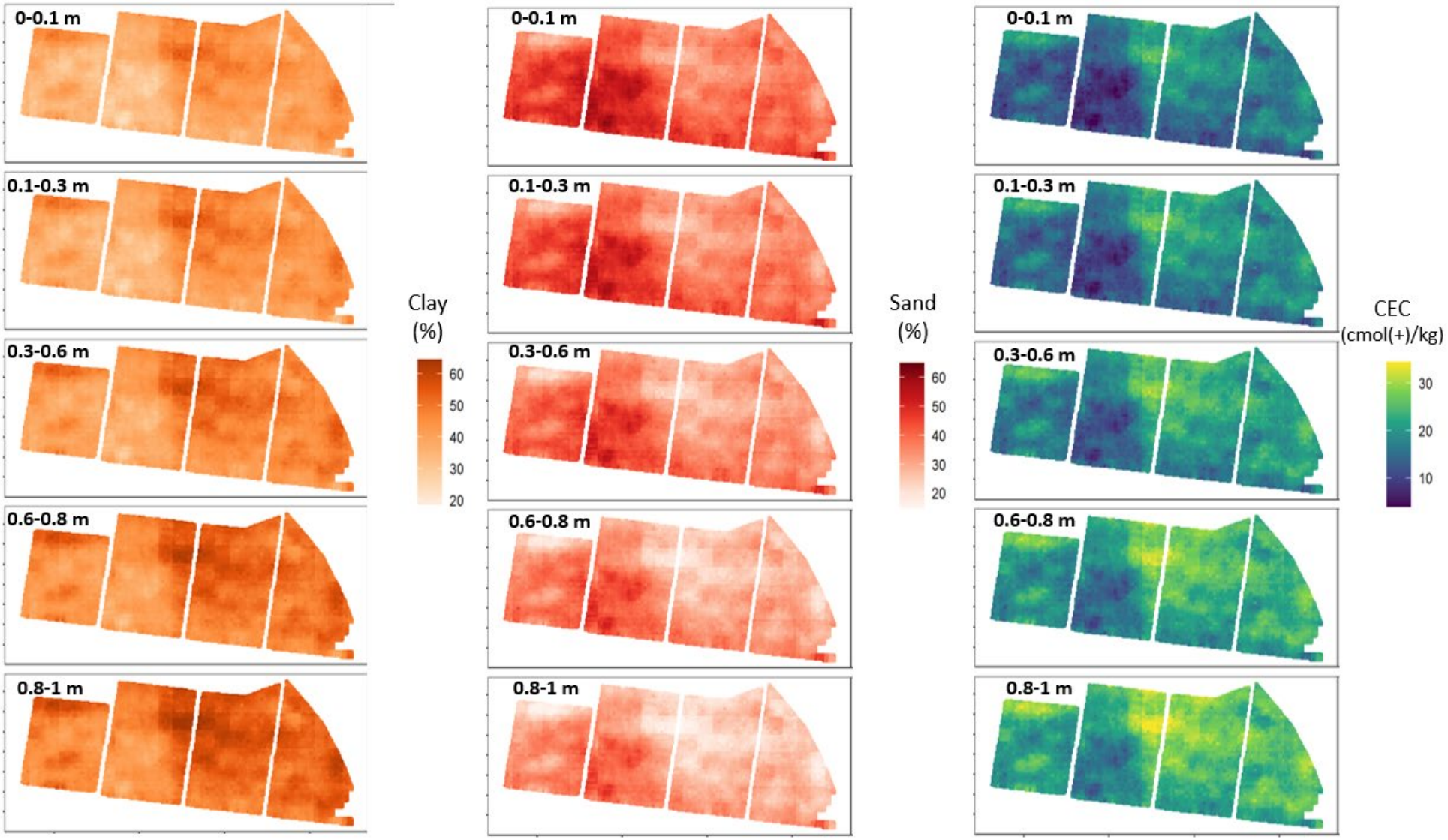


Figure 5.21. Digital soil maps at a 5 m resolution for clay (%) (left), sand (%) (centre) and cation exchange capacity (CEC) (cmol(+)/kg) (right) at Farm C.

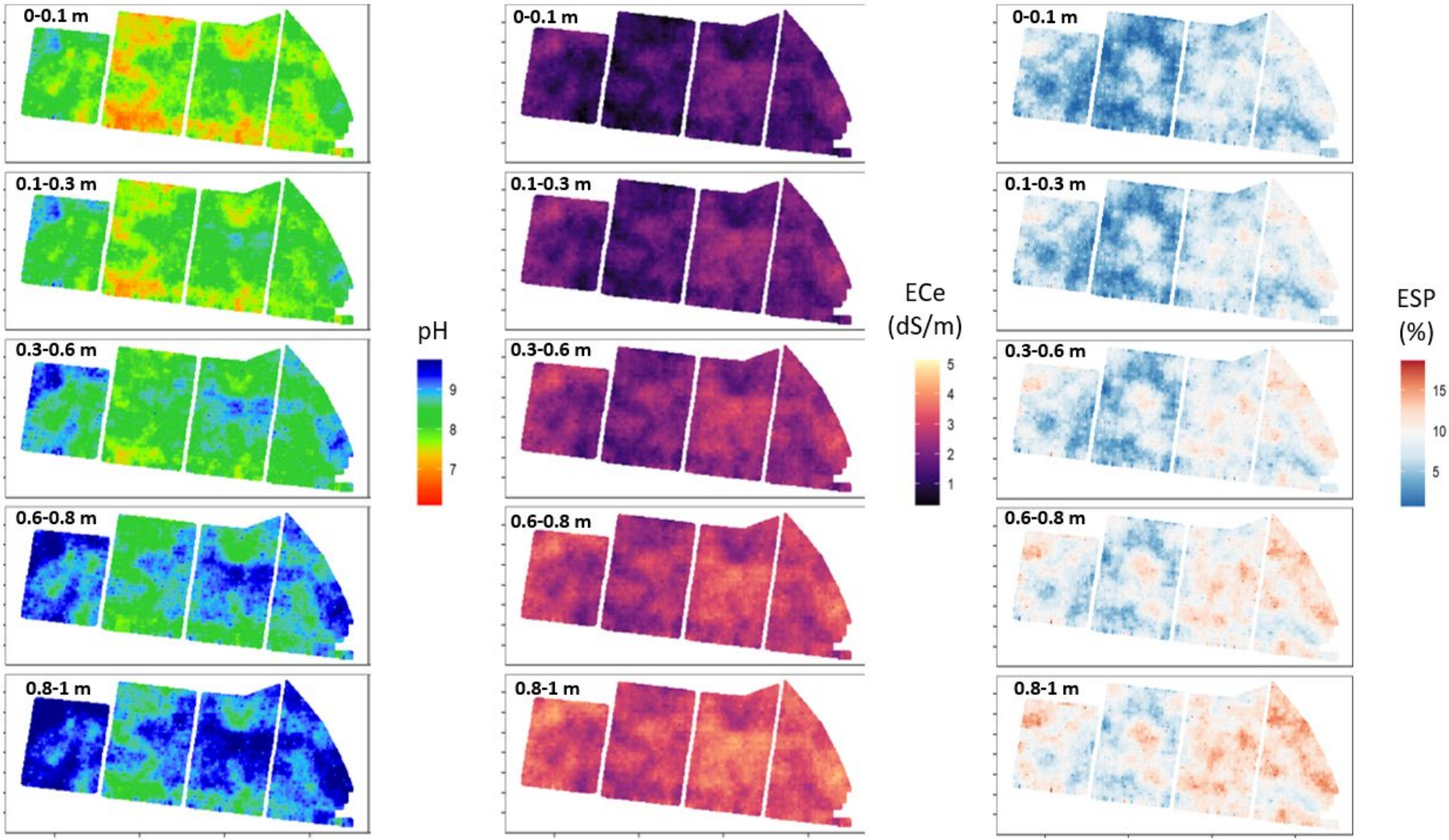


Figure 5.22. Digital soil maps at a 5 m resolution for pH (left), electrical conductivity of the extract (ECe) (dS/m) (centre) and exchangeable sodium percentage (ESP) (right) at Farm C.

5.4. Discussion

5.4.1. Influence of palaeochannel systems on soil variability

The influence of palaeochannel systems on soil morphological variability has been discussed in the preceding chapter. This study identifies that this variability influences cotton lint yields. The characteristics of the palaeochannel system determines the extent of soil variability, resulting in a clearer differentiation of soil attributes between clusters at Farm A. This aligns with the morphological descriptions of the deep sandy and heavy clay soils that were dominant at this site where a terminal palaeochannel anabranch associated with the Gum Creek system (Page & Nanson, 1996) has incised across the paddock. Due to the significant contrast between the heavy clay and deep sandy soils, there is greater spatial variability in measured soil properties, excluding the Ca:Mg, at Farm A than Farm C. At Farm C there is greater variability vertically within soil profiles, suggesting areas with more pedologically developed profiles, attributed primarily to the texture contrast or transitional earth soils located centrally within the paddock.

5.4.2. Production zones

The production zones developed using yield data created two clusters that broadly align with noted morphological variability. While there is an indication of irrigation tail drain and head ditch effects, the poorer yielding clusters generally align with the identified sandier areas of each paddock. These production zones were created only to assess yield variability across each field. When developing management zones, however, there are other approaches that warrant investigation that incorporate proximally and/or remotely sensed data. Within literature there is conjecture around which approach is optimal. Hornung et al. (2006) found that a soil-colour-based management zone approach incorporating only bare-soil imagery, topography and farmer knowledge was better at classifying grain yield patterns than an approach which also incorporated CEC, soil texture and the previous year's yield data. Conceição et al. (2024) present a methodology targeted to farmers which delineates management zones utilising remotely sensed data, proximally sensed electrical conductivity and sampled soil sand content and exchangeable calcium. Leo (2022) found that management zones derived from remotely sensed vegetation indices outperformed those developed using soil data or soil data and vegetation indices. Such methodological considerations are important as not all farmers maintain or have access to quality yield or proximally sensed data.

5.4.3. Soil texture and drivers of cotton yield variability

This study has identified the significant impact that inherent soil properties have on irrigated cotton yields. Despite the moderate-to-strong correlations between 'manageable' soil properties, including pH and ESP, the point correlations and variability between production zones show yield is primarily driven by soil texture. This relationship is stronger at Farm A than Farm C, however, across both paddocks higher yields are positively correlated to clay content with the inverse true for sand content. Within-field variability in soil texture has multi-level impacts on the soil-yield relationship.

The primary impact on crop yields is through water holding capacity and hydraulic conductivity, with more clayey soils able to hold water more tightly between soil particles (Hake and Grimes, 2010). Water also moves through the profile more rapidly in the sandier soils (Hazleton & Murphy, 2007). Both paddocks in this study are flood irrigated either using siphons over banks and furrows or a pontoon system. Flood irrigation is well suited to soils with low-medium infiltration rates and higher water holding capacities (North & Schultz, 2017). Both fields, however, do not have textural homogeneity, with this study showing that under these systems sandy areas exhibit lower yields. Soil texture also has secondary impacts on crop yield correlations. As CEC is strongly positively correlated to clay content, for example, it also exhibits a strong positive correlation to cotton lint yields. A lower CEC also reduces the buffering capacity of soil, resulting in potential increases to pH following fertiliser application and base cation leaching (Lesturgez et al., 2006). This may explain the more acidic soils in the lower yielding, Cluster 2 at Farm A, where pH is positively correlated to cotton yield. When comparing this site to Farm C, the impact of extreme alkalinity is witnessed. At Farm C, the subsoil pH in both clusters approaches a limiting threshold of 9 (Hazleton & Murphy, 2007). Where this occurs the positive correlation between pH and yield at the surface becomes negative at depth (>0.6 m).

In considering these results more broadly, the impact of soil properties is not consistent between these two farms and will not be consistent across the entire region. Instead, the relationship will depend on the management system in association with the full suite of soil attributes. Voires et al. (2021) identified that more sandy soils produced lower cotton seed yields under pivot irrigation. In contrast, Ping et al. (2007) identified, on loamy soils in the Texas high plains, that a low pH and high sand content were associated with higher cotton yielding areas grown under centre-pivot irrigation. In a rainfed system with silty clay loam soils in Mississippi, Cox

et al. (2005), however, found that neither sand nor clay content were significantly correlated to lint yield in two paddocks over two growing seasons.

This study highlights the importance of not examining individual soil attributes, and their relationship to crop yield, in isolation. Along with alkalinity two constraints commonly researched in association with crop production are salinity and sodicity (Cattle & Field, 2013; Dagar et al., 2022; Dang et al., 2022; Orton et al., 2018). At all depths and in both paddocks ECe and ESP, measures of salinity and sodicity, respectively, are positively correlated to cotton yield (excluding ECe at 0-0.1 m). At Farm A, specifically, all samples in the higher yielding cluster at depths >0.6 m are considered extremely sodic while multiple samples also exceed the cotton salinity threshold (Hazleton & Murphy, 2007; Tilse et al., 2022). If these levels were examined in isolation they would be considered highly limiting and having surpassed levels where amendments for sodicity are likely to improve yield (Dodd et al., 2013b). In these instances, however, a high ESP is positively correlated to cotton yield. The observations here are different to those of Nachimuthu et al. (2024) who found a higher ESP significantly reduced cotton yields in a paired field comparison near Griffith. In their study both sites exhibited an ESP ≤10 and were uniformly clayey (>45%), classified as Dermosols (Nachimuthu et al., 2024). In more homogeneously textured fields it may be appropriate to focus on sodicity, however, where texture is highly variable an emphasis on this is required. This is because ESP itself is strongly related to clay content, with this study demonstrating that the latter has a stronger influence on cotton lint yields when texture is highly variable. Future work could subset the correlations by production zones or individual sample points to examine this relationship more deeply.

Also of note is that the strongest point correlations between soil attributes and yield are seen in the subsoil. This is of interest as commercial sampling operations often focus on the 0-0.3 m soil layer due to the cost and labour-intensive nature of this exercise. The significance of this depth is likely due to changing water requirements and plant rooting depths throughout the growing season (CRDC & CottonInfo, 2023). Early in the growing season the plant is only drawing water from the upper layers of the soil. Throughout the season, however, the increased rooting depth and plant water requirements, which are greatest 110-130 days after sowing, require more water to be extracted from stores at depth (CRDC & CottonInfo, 2023; Datta et al., 2019). In more clayey soils this is possible, however, where there is a lower clay content there is less moisture stored between irrigation cycles limiting plant growth and eventual lint yield. Previous work by Ulfa et al. (2023) established that constraints in the subsoil impact late

season crop growth while Tilse et al. (2022) examined the production losses resulting from reductions in available water capacity, albeit resulting from chemical constraints and not texture, on crop yields. Thorp et al. (2020) also suggested that the first irrigation, when plant roots are shallowest, is not as strongly linked to fibre quality or yield as subsequent irrigations.

5.4.4. Digital soil models and maps

This study highlights the ability of simple linear models to predict soil properties with moderate-to-good quality when assessed using the LCCC from a LOSOCV. Linear models are not widely considered in recent DSM literature, with most focussing on machine learning methods (Khaledian & Miller, 2020). They may, however, present an opportunity for bespoke DSMs with small datasets. The high error (RMSE) seen in linear models for CEC, sand and clay content at Farm A is expected given the small sample size and significant measured variability. At Farm C, linear models outperform RF models when assessed using the LOSOCV and produced superior model statistics than their equivalents at Farm A. This highlights the potential impact of sample size on model quality, suggesting the benefits of increased data points. In a practical sense, however, this may not always be financially feasible so the development of products from smaller datasets is important and linear models are well suited to this. This increase in sample size may have also allowed for better differentiation of variability in remotely sensed covariates across the paddock, resulting in a more diverse range of spatial predictors being included in the final models at Farm C. This is a potential source of model improvement, with Filippi et al. (2024) highlighting that bespoke models incorporating remote and proximally sensed covariates outperform those that use only one or the other. When visually assessed, the resulting maps align with anecdotal understandings of variability. There have not been any similar studies within the region to directly compare these results to, with Kidd et al. (2020) noting that the development of DSMs at the farm or paddock scale is often undertaken by private industry. This is, however, an area that could be of significant value to farmers, especially where high soil variability is present.

5.4.5. Agronomic and management considerations

While previously mapped plant limiting thresholds for sodicity, alkalinity and salinity (Tilse et al., 2022) have been identified in one or both paddocks, at the paddock scale these are not the most significant drivers of yield. As such, it would not be in the economic interest of farmers to fixate on these attributes. This supports ideas conveyed by Bennett et al. (2014) who made

clear the need to assess the economic feasibility of ameliorating soil constraints. So, while chemical data suggests that amelioration of chemical constraints may be required, the primary focus should first be on managing the most limiting areas of each paddock, these being the sandy soils.

The first consideration should be options to implement overhead centre pivot or lateral move systems, especially at Farm A, which are capable of variable rate irrigation that is not possible under the presently used surface systems. On the Texas plains, sand content was identified as a positive contributor to cotton lint yields irrigated using pivots on loamy soils (Ping et al., 2007). The ability to apply water more precisely under these systems is important as the texture-caused limitations cannot be overcome by simply irrigating more regularly, with Neupane et al. (2021) identifying that soils with >50% sand content do not respond to higher irrigation rates. Further efficiencies may result from the implementation of variable rate irrigation, considerations for which are described by Neupane and Guo (2019) and McCarthy et al. (2023). Irrigation decisions can then be informed through moisture sensors in morphologically distinct zones of each paddock to account for soil variability (Sui, 2017). Despite their potential applicability to the region, there has not been specific research examining these systems in the Murrumbidgee. Alongside improving water use efficiency due to their suitability for lighter textured soils there is also the capacity to improve nitrogen use efficiency through fertigation (Antille, 2018; North, 2017; Scheer et al., 2023). Benefits of this and zonal application of inputs is discussed in more detail below.

There is the capacity to increase the value of the developed bespoke DSM products. One approach to improve the accuracy of paddock-scale models is to collate data from multiple paddocks, farms or available data sources as undertaken by Filippi et al. (2019a). While this study has mapped five depths concurrently in each field, this may not be of maximum value to stakeholders. One option to reduce the dimensionality of this data, while retaining necessary detail, is through the development of decision-making tools such as depth to constraint maps (Filippi et al., 2019b). At the farm scale, Tilse et al. (2022) highlighted the potential for depth to constraint maps to be used to quantify yield losses and constrained plant available water capacity by identifying the depth where constraints are reached. There is the potential for such maps of be of particular use at Farm C, where pH is demonstrated to exceed the plant limiting alkalinity threshold (pH>9) and have negative correlations to yield below 0.6 m.

The identification of strong relationships between inherent soil properties and cotton yields in the subsoil is important. While needed to inform management decisions, the cost and intensive

nature of soil sampling can be inhibiting factors for landowners (Condon, 2019; Viscarra Rossel & McBratney, 1998). Bennett et al. (2021) provided a detailed discussion on this in relation to the cost-benefits of DSM, also noting that there has been too strong a focus on nutrient status in the topsoil, with the subsoil and inherent properties often overlooked. The implementation of management zones through the utilisation of yield, soil, proximal or remotely sensed data can assist with this (Mallarino & Wittry, 2004; Nawar et al., 2017). By clustering representative areas of the paddock, as undertaken in this study, there is the capacity for samples to be taken at fewer sites and lower costs while still representing the entire area. Consequently, this may allow for a greater number of subsoil samples to be extracted without increasing the overall number samples analysed. It is important to communicate this with stakeholders so that they understand the potential benefits associated with zonal, as distinct from grid, sampling.

The increased homogeneity within developed production or management zones can also be of financial benefits resulting from targeted agronomic decisions (Whelan & McBratney, 2000). An application of zonal, or site specific, management in these specific paddocks may begin by determining the maximum yield potential based on each zone's inherent soil properties (Whelan & McBratney, 2003). Management decisions can then be made to achieve target yields based on these benchmarks, rather than attempting to achieve homogeneous yields, which is an impossible task given the soil variability and demonstrated influence on lint yield. In these paddocks lower, but achievable, yield targets can be set in the sandier areas, with higher targets in the more clayey zones. Variable rate input applications can then be applied zonally based on this production capacity, with lower rates in sandy zones resulting in environmental and economic benefits. This is a particularly important consideration as over-applications of nitrogen are common in the Australian cotton industry (Macdonald et al., 2018). This results in negative financial, through reduced nitrogen use efficiency, and environmental, including through surface runoff, impacts (Brackin et al., 2018; Macdonald et a., 2016). Increasing fertiliser use efficiency by avoiding over-application can also reduce greenhouse gas emissions at the farm scale (Badgery et al, 2024; Grace et al., 2016; Snyder et al., 2009). This is likely to become an important consideration for farmers as they may be required to account for their emissions as food and fibre supply chains continue toward net-zero emissions goals (Badgery et al, 2024).

5.5. Conclusions

This study has shown that on the highly variable soils of the lower Murrumbidgee valley the relationship between soil properties and cotton yield is complex, with soil texture identified as the primary driver of cotton lint yields. It is deduced that the main impact of texture variability is through water holding capacity. This reduces the efficacy of flood, including furrow and pontoon, irrigation systems. A point correlation (r) analysis was able to identify this trend. At Farm A, clay content exhibited a >0.89 correlation with cotton lint yield at all examined depths. This analysis also showed that the strongest correlations between soil properties and crop yields occurred at a depth of 0.6-0.8 m, suggesting a need to sample and analyse the subsoil. Interestingly, soil properties including ESP and ECe approached, and on occasion, exceeded previously identified plant thresholds of 15% and 7.7 (dS/m), respectively. Despite this, generally positive correlations to crop yield were observed. For ESP across both farms and at all depths the positive correlation (*r*) to cotton lint yield varied from 0.06-0.75. This highlights the need to understand the soil in its entirety, not just soil attributes individually. The development of production zones utilising yield data was able to partition the paddock into clusters that were representative of the soil-yield relationship, whereby, lower yielding clusters were sandier. The utilisation of these zones can allow for improved overall productivity by applying inputs based on the production potential of each zone. Lastly, it was shown that simple linear models can predict soil properties at the within-field scale to a moderate-to-good quality. At Farm A, LCCC values obtained from a LOSOCV ranged from 0.45-0.65. At Farm C linear models outperformed random forest models, with LCCCs in the former varying from 0.62-0.73. At Farm C, an increased sample size corresponded to increases in both model quality and the suite of covariates retained during the modelling process. Future work should examine the potential of variable rate irrigation through overhead pivot or lateral move systems in southern NSW, where the soil and resulting yield variability identified in this study is widespread.

Chapter 6

6. Using digital soil mapping as a tool to understand the spatial distribution of soil properties in the lower Murrumbidgee valley

6.1. Abstract

The evolution of digital soil mapping (DSM) has seen it progress from a research topic to a tool that can be implemented to understand the spatial variability of soil properties at different scales. Previous studies have identified that state, national, or global DSM products may be ineffective at identifying fine scale soil variability, an understanding of which is essential for developing sustainable management practices. This study uses DSMs as a tool to map the variability of soil pH, exchangeable sodium percentage (ESP), soil electrical conductivity of the extract (ECe), cation exchange capacity (CEC), sand content and clay content at the regional level to a 90 m resolution in the lower Murrumbidgee of southern NSW. Extreme Gradient Boosting (XGBoost) models were developed using selected covariates and soil data from 153 sites at five depths (0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8, 0.8-1 m) to make predictions of the soil properties. Layer mid-depth was incorporated as a predictor, allowing for '3D' models to be produced that predicted each of the depth increments concurrently. Models were assessed using a leave-one-site-out-cross-validation (LOSOCV) with the Lin's Concordance Correlation Coefficient (LCCC) used to measure model quality. The quality of the different models varied, with LCCCs of 0.45 (ECe) to 0.58 (clay). An external point support validation was performed on presently available state, national and global DSMs for clay content at depths of 0.3-0.6 and 0.6-1 m using the data collected for this study. The state, national and global models performed more poorly than their reported values, with LCCCs of the three products ranging from 0.11-0.5 and 0.07-0.36 at depths of 0.3-0.6 and 0.6-1 m, respectively, in the external point validation. At both depths the global product performed the worst, and the state product the best. This study showcases the potential of more specific, regional DSMs to accurately identify landscape variability at finer scale. It also highlights the importance of assessing pedological validity of DSM outputs and utilising landscape knowledge when collating spatial covariates.

6.2. Introduction

An understanding of individual soil properties at a given location is essential in guiding sustainable management, implementing precision agricultural practices and allowing for informed policy decisions. Broadscale digital soil map (DSM) products, however, have been demonstrated as ineffective at identifying variability at this support (Han et al., 2022). A regionally focused DSM approach was identified by Searle et al. (2021) as a means to improve accuracy at the finer farm support, while Brungard et al. (2021) found that physiographic sub-region specific models outperformed those for a wider area. Lowland areas, a term encompassing coastal areas as well as deltas and floodplains such as the lower Murrumbidgee valley, are rarely the target of digital soil mapping activities as noted by Adeniyi et al. (2024). Holland and Eastwood (2014) identified that an improved, accurate understanding of the spatial variability of soil properties was required within the lower Murrumbidgee valley.

Improvements in computational capacity and the accessibility of quality spatial data has seen studies implementing DSM increase significantly since the year 2000 (Minasny and McBratney, 2016). While maps produced through DSM serve the same purpose as traditional soil surveys, that being the conveyance of information at a given location, these products do not rely on conceptual understandings of landscapes and are not as cost and labour intensive to produce. In early research, McBratney et al. (2003) formalised a quantitative approach built on empirical relationships between measurements relating to soil properties or classes and spatial data. This ‘scorpan’ model allows for the prediction of a soil property, or class, as a function of soil (s), climate (C), organisms (o), relief (r), parent material (p), age (a) and spatial position (m) (McBratney et al., 2003). There are numerous resources available detailing the methodologies for producing digital soil maps and reviewing the field more broadly (Arrouays et al., 2020; Kidd et al., 2020; Malone et al., 2017a ; McBratney et al., 2003). Diverse modelling methods can be applied to DSM exercises, with the most appropriate contingent on the study’s aims. Over the last decade there has been an exponential increase in machine learning (ML) approaches being used for DSM, with random forest (RF) algorithms the most reported in literature (Khaledian and Miller, 2020). In studies specifically focused on lowland areas, akin to that of this study, Adeniyi et al. (2024) note that RF algorithms are those most used. Extreme gradient boosting (XGBoost) models have also been demonstrated as effective in modelling a range of soil properties (Chinchmalatpure et al., 2023; Hengl et al., 2017b; Ramcharan et al., 2018; Shirazi et al., 2024; Zhang et al., 2022), with Meier et al. (2018) noting that a particular benefit of XGBoost models is their capacity to work with sparse data.

DSM allows for spatial covariates to be used to represent the features of the 'scorpan' equation. Richer-de-Forges et al. (2023) review the role of remotely sensed data in DSM; Maino et al. (2022) discuss the specific role of radiometric covariates in predicting soil texture; while Adeniyi et al. (2024) note that covariates representing organisms (O) and relief (R) are the most common in lowland DSM studies. Given the breadth of publicly available data, specific considerations should be taken when selecting covariates and drawing conclusions on their model impacts. Wadoux et al. (2019) emphasise the need for caution, where, just because a relationship exists does not mean it is related to soil processes, while Jones et al. (2022) showcase the potential of Shapley Additive exPlanations (SHAP), an approach based on game theory, to assist in identifying the specific contributions of each predictor variable to the model.

Rather than developing models for each individual depth increment (2D) the incorporation of depth as a predictor variable allows for multiple soil depths to be predicted concurrently, referred to as '2.5' or '3D' modelling (Brus et al., 2016; Rentschler et al., 2019; Roudier et al., 2020; Zhang et al., 2020). This can improve models for certain soil properties, where depth is linked to pedological processes, allowing them to be more dynamic with multiple depths predicted concurrently, potentially increasing the interpretability of outputs to meet stakeholder needs (Roudier et al., 2020). The impact of this approach on model quality varies depending on the attribute being modelled, with care required as models may exhibit larger prediction uncertainties (Roudier et al., 2020).

Various methods are available to validate model performances, with Piikki et al. (2020) reviewing the prolificity of methods in literature and noting cross-validation to be most common. They also emphasised issues in interpreting the reported statistics where essential information is missing (Piikki et al., 2020). The implementation of different validation methods is known to impact the model quality statistics, as shown by Pozza at el. (2022). Khaledian and Miller (2020) emphasised the importance of undertaking a cross validation to test model performance and allow comparisons between modelling projects. An expert consideration of the spatial patterns resulting from models is also important, with high quality model statistics not necessarily meaning the map produced will be pedologically valid (Padarian Campusano, 2014; Bui et al., 2020; Kidd et al., 2020). The optimal resolution, or support, for predictions to be made varies depending on numerous factors including the input covariates, soil attribute, soil heterogeneity and study aims (Bishop et al., 2015; Piedallu et al., 2022).

Across the lower Murrumbidgee valley there are three main DSM products available, which can be considered as state (Gray 2023), national (Grundy et al., 2015; Viscarra Rossel et al.,

2014) and global (Hengl et al., 2017a; Poggio et al., 2021). While these DSMs report moderate to good validation statistics, Han et al., (2022) identify that these models performed poorly when validated against external datasets form 14 different farms across eastern Australia.

This study will use DSM as an operational tool at the regional level to improve understandings of the spatial variability of specific soil properties at depths of 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m across the lower Murrumbidgee valley. XGBoost models will be developed to predict pH, exchangeable sodium percentage (ESP), soil electrical conductivity of the extract (ECe), cation exchange capacity (CEC), sand content and clay content at the five depths concurrently. Alongside statistical analysis, expert landscape knowledge obtained through reviewing geomorphological resources will be used to guide covariate selection and assist in determining the most optimal model. Lastly, a brief analysis will occur to assess the accuracy of the presently available state, national and global DSMs at the point support within this study area.

6.3. Materials and methods

6.3.1. Soil sampling and soil property datasets

This study utilised the 'full-dataset' as described in Chapter 3, containing 153 soil cores from across the study area. All data processing and analysis for this study occurred using the program 'R' (R Core Team, 2023). Once returned to the laboratory cores were morphologically described before being subsampled at five depth increments (0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m). Each subsample was then air dried (40ºC) before being ground to pass through a 2 mm sieve.

Soil pH and electrical conductivity (EC) were measured in a 1:5 soil:water solution using a SevenCompactTM s220 pH/Ion meter and S230 Conductivity meter, respectively. Soil EC was converted to electrical conductivity of the extract (ECe) using a conversion factor based on soil texture. For each subsample, soil was sent for external analysis of base cations (Ca^{2+}, Mg^{2+}, K^{+} and Na^{+}) using the 15C1 method of Rayment and Lyons (2011) from which cation exchange capacity (CEC) and exchangeable sodium percentage (ESP) were determined. After being finely hand ground, all samples were scanned using a contact probe on an ASF Agrispec spectrophotometer (Malvern Panalytical, Boulder, Co, USA). This spectral dataset was used to identify a representative subset of data to undergo particle size analysis testing using the pipette method (Glendon & Or, 2002). Utilising this spectral and measured data, cubist models were

developed to predict clay (%) and sand (%) for all samples. A detailed discussion of the methodologies used in this study is provided in Chapter 3 of this thesis.

Once curated into a data frame, summary statistics for each soil property at each measured depth were generated.

6.3.2. Covariate selection

A set of predictor variables (Table 6.1) was obtained based on the *SCORPAN* approach to digital soil mapping (McBratney et al., 2003). More detail surrounding the obtaining of covariates and this approach is provided in Chapter 3. The covariates in this study can be considered as satellite attributes (Figure 6.1), radiometric attributes and geological data (Figure 6.2) and terrain attributes (Figure 6.3). Once downloaded, each predictor variable was resampled on a 90 m grid using the 'bilinear' method. While DSMs are now commonly produced at finer resolutions, 90 m was chosen as Han et al. (2022) demonstrate that just because predictions are at a fine scale does not mean models are improved. Once at a 90 m resolution, a data cube with measured soil properties and predictor variables was created through the extraction of each covariate at each sample location. A variance inflation factor (VIF) function was then implemented with a threshold of 10 (Fox and Monette, 1992; Fox and Weisberg, 2019). This function identifies and removes variables that exceed the set threshold with the purpose of reducing multicollinearity amongst the covariates in the predictor dataset.

Table 6.1. Obtained covariate data prior to the removal of variables based on exceedance of the VIF threshold. Each covariate can be categorised as a terrain, radiometric, geological or satellite attribute with the original resolution of each layer varying. Note: 1 arc second represents approximately 30 m.

Data category	Data layer	Data source	Original resolution
Terrain	DEM (m)	Gallant et al. (2011)	1 arc second
	Slope (%)	Gallant and Austin (2012a)	3 arc seconds
	MrRTF	Gallant et al. (2013)	3 arc seconds
	MrVBF	Gallant et al. (2012)	3 arc seconds
	TWI	Gallant and Austin (2012c)	3 arc seconds
Radiometric	Dose rate	Minty et al. (2009)	100 m
	Uranium (%)	Minty et al. (2009)	100 m
	Potassium (%)	Minty et al. (2009)	100 m
	Thorium (%)	Minty et al. (2009)	100 m
Geological	Silica (%)	Gray et al. (2016)	1:250,000
Satellite	Barest earth blue	Roberts et al., 2019	25 m
	Barest earth red	Roberts et al., 2019	25 m
	Barest earth green	Roberts et al., 2019	25 m
	Barest earth NIR	Roberts et al., 2019	25 m
	Barest earth SWIR	Roberts et al., 2019	25 m
	Barest earth SWIR 2	Roberts et al., 2019	25 m

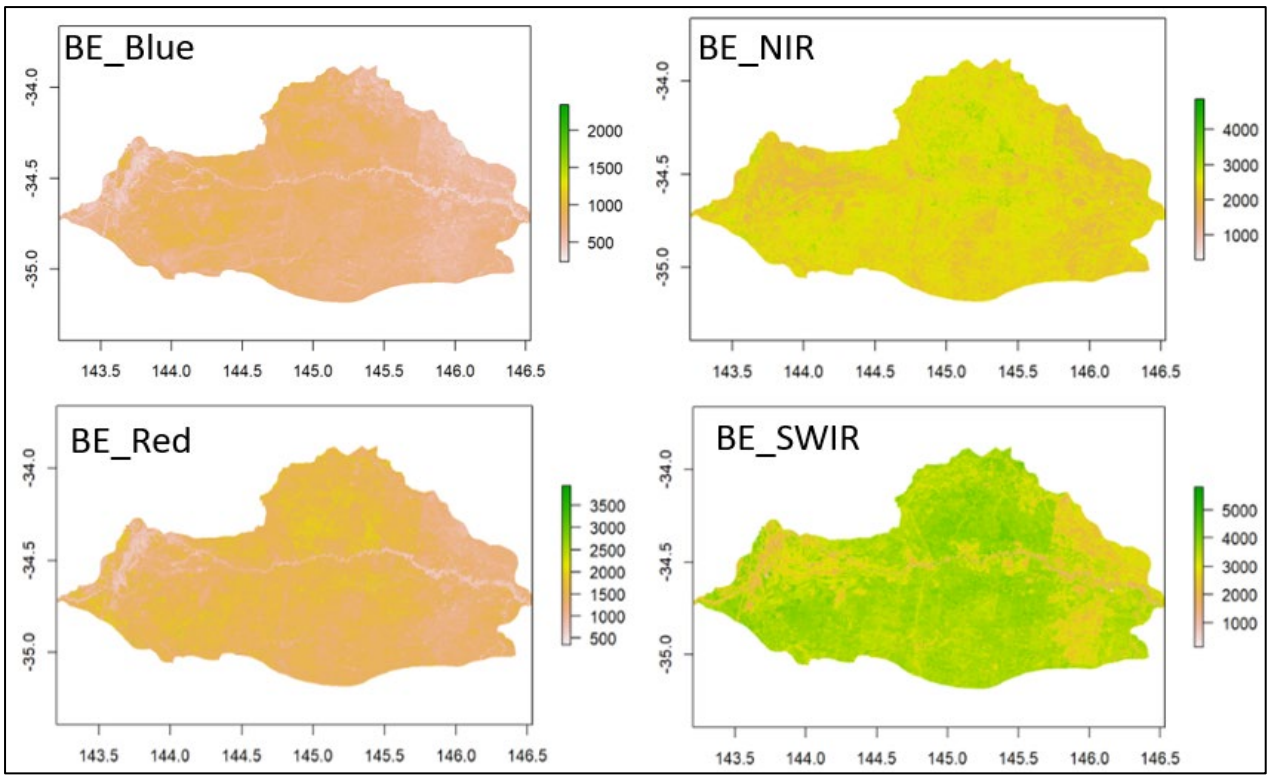


Figure 6.1. Maps of the satellite covariates remaining following deletion due to collinearity including barest earth (BE): blue, red, NIR and SWIR at a 90 m resolution across the study area.

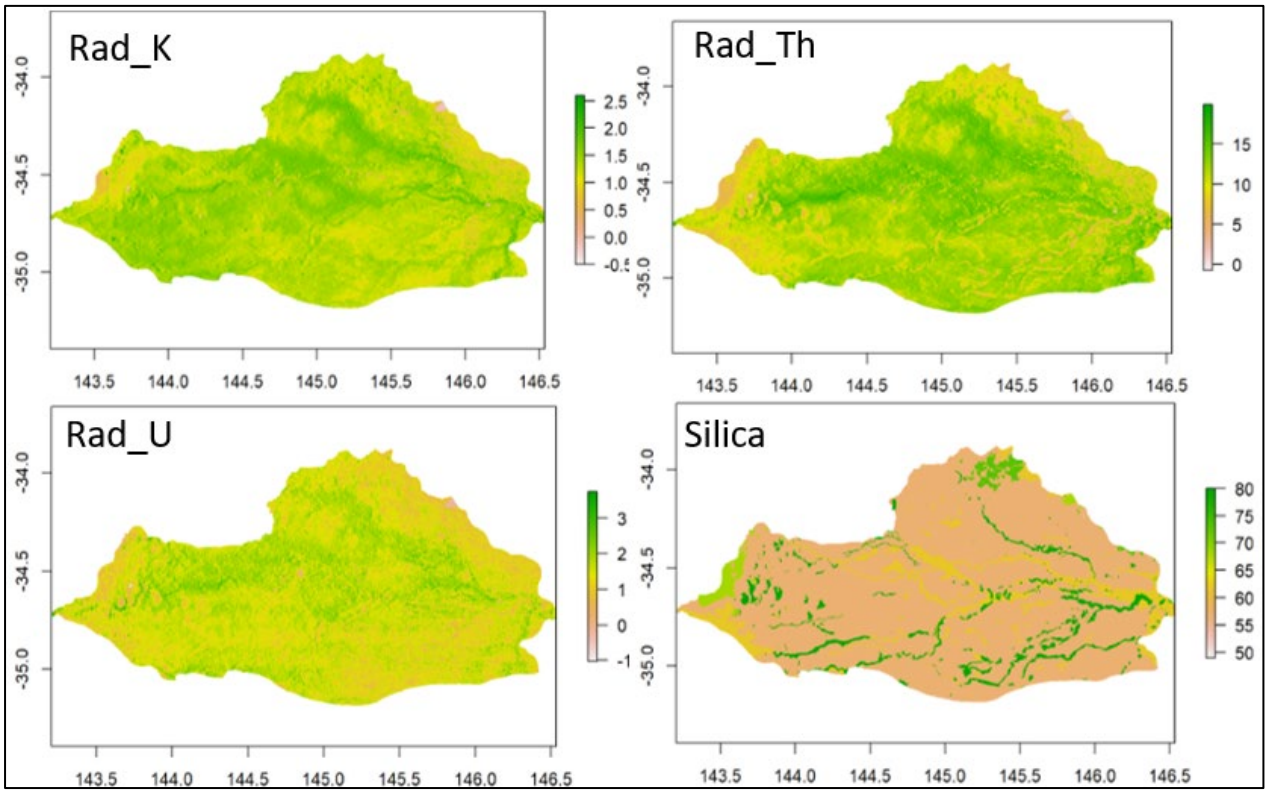


Figure 6.2. Maps of the remaining radiometric Potassium (K) (%), Thorium (Th) (ppm) and Uranium (ppm) attributes as well as the geological attribute silica at a 90 m resolution across the region following the removal of radiometric dose based on collinearity.

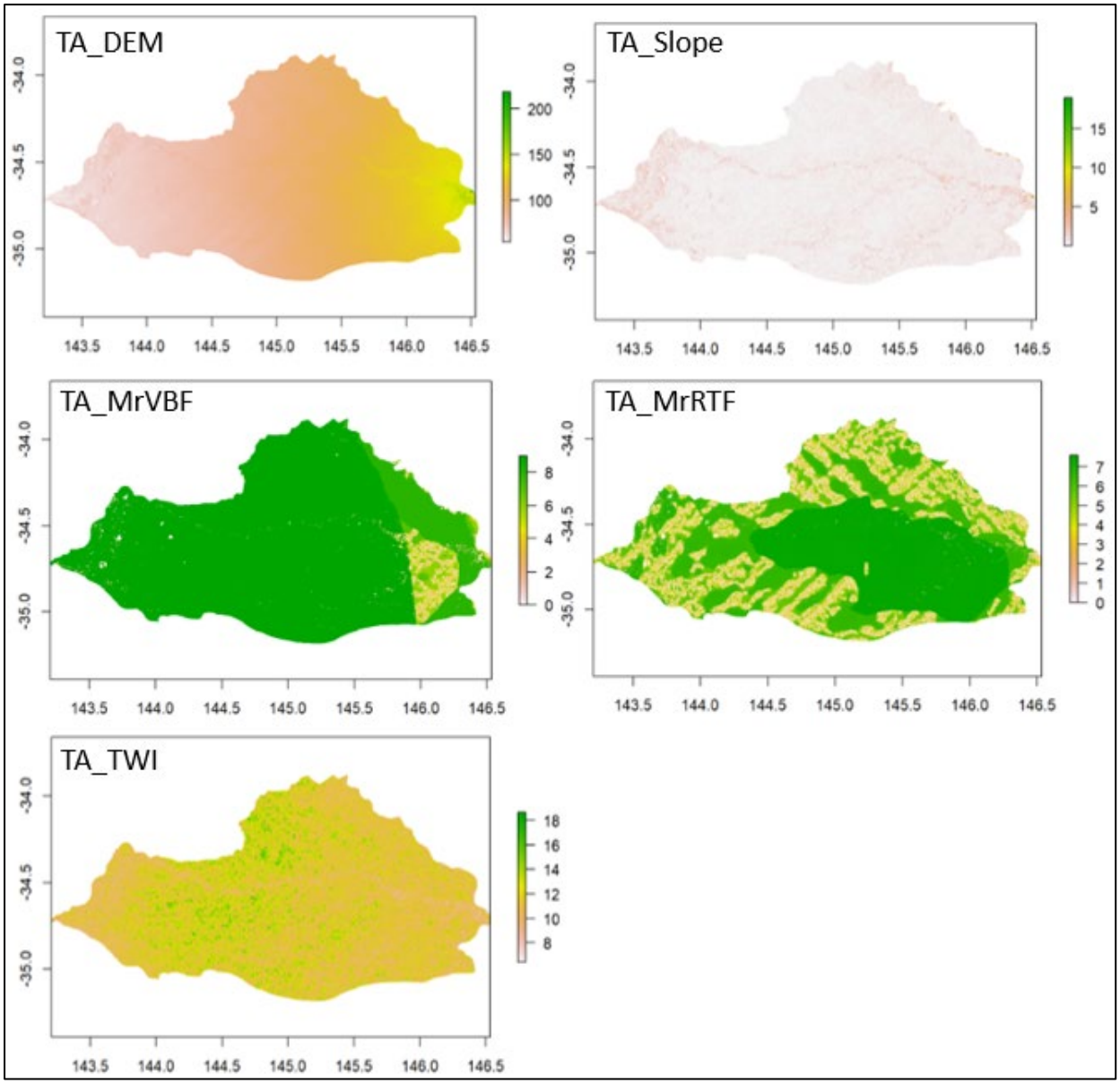


Figure 6.3. Maps of DEM (m), Slope (%), MrVBF, MrRTF and Topographical Wetness Index (TWI) at a 90 m resolution.

6.3.3. Soil property model development and map production

Models were developed to predict pH, ECe, ESP, CEC, sand content and clay content using the obtained soil datasets and covariate stack. For each soil property, a '3D' modelling approach was utilised to predict the five sampled depths concurrently. As such, the mid-depth of the layers being modelled (0.05, 0.2, 0.45, 0.7 and 0.9 m) was utilised as a predictor variable. As the outputs of this research are directly targeted to increase stakeholder understanding of the region's soils, the interpretability and cohesivity gained through this process is important. Through this method a total of 764 samples were used to develop models of pH, CEC and ESP, while 762 samples were used to develop models for clay (%), sand (%) and ECe.

6.3.3.1. Development of extreme gradient boost (XGB) models

Following the removal of spatial covariates on the basis of collinearity and visual assessments, Extreme Gradient Boosting (XGBoost) models were built utilising the remaining spatial covariates and layer mid-depth. XGBoost models utilise a gradient boosting framework to produce final predictions that are the result of numerous decision trees (Chen and Guestrin, 2016; Zhang et al., 2022). It is an ensemble model that works to train itself in an additive way through the redefining of parameters through learned patterns (Chen and Guestrin, 2016; Meier et al., 2018). In refining itself, the construction of new trees is based on prediction errors in previous trees allowing for prediction error in the final model to be minimised (Chen and Guestrin, 2016; Zhang et al., 2022). In this instance the 'xgboost' function, the simple wrapper function contained within the 'xgboost' R package, was used (Chen et al., 2024). Settings of 6 and 0.3 for the maximum depth of a tree (max_depth) and controlled learning rate (eta), respectively, were used. The maximum number of boosting iterations, 'nrounds', was set at 500 and 'regression with squared loss' set as the specified objective.

6.3.3.2. Variable importance

Model interpretability is a vital aspect in understanding, auditing and explaining machine learning models. In this instance, SHapley Additive exPlanation (SHAP) values were used to understand the impact of each variable on model predictions (Molnar, 2023). These values were computed using the 'SHAPforxgboost' package within R (Liu and Just, 2023). Shapley values were first proposed in game theory by Shapley (1953). Lundberg et al. (2020) proposed a method for SHAP aimed specifically towards tree-based and boosted models, while Padarian et al. (2020) demonstrated and discussed their role in interpreting DSMs from a game theory perspective which allows for a 'global' interpretation of models, this being the understanding of how well the model makes decisions across the entire dataset. 'Beeswarm' plots were used to visualise this relationship by combining 'local' point interpretations graphically to understand the variable-to-predictor relationship wholly, or as the 'global' interpretation.

In the beeswarm summary plot the y-axis shows the model variables, with variables ranked based on their importance in the model determined as the average of absolute SHAP values (Molnar, 2023). The x-axis represents the SHAP values with the colour, or feature value, detailing the nature of the relationship at each sample point (Molnar, 2023). Each dot within the plot represents the SHAP value for an individual data point. Dependence plots were used to interrogate specific predictor variables. These plots also show SHAP values on the y-axis against the predictor attribute value on the x-axis, however, a colour ramp can be generated to

represent the attribute value of a second predictor, allowing for a greater understanding of the relationship between predictors.

6.3.3.3. Model quality assessment and selection

For each soil attribute, the model was initially fitted using all covariates. A manual stepwise backwards deletion was undertaken, removing the lowest ranking predictor at each iteration until a model was produced with three spatial covariates and layer mid-depth, the latter being required to predict multiple depths concurrently. Three covariates were chosen as the endpoint as this was the point where model quality significantly worsened. In each iteration of this process the spatial covariate having the smallest model influence was removed.

The quality of each model (n=8), for each individual soil property (n=6), was calculated using the out-of-bag predictions from a leave-one-site-out-cross-validation (LOSOCV). The LOSOCV method trains the model on n-1 sites (n=153) and then makes predictions for all depths at the excluded site, in this instance a soil core. This method can be considered beneficial when working with smaller datasets as it does not rely on the splitting of data into calibration and validation subsets. Model quality was assessed through the Lin's Concordance Correlation Coefficient (LCCC). The LCCC statistic ranges from 0 to 1, taking into account precision and accuracy, where a value closer to 1 indicates a higher quality model (Lin, 1989; Akoglu, 2018). The model bias and Root Mean Square Error (RMSE) are also presented.

The final model was selected based on parsimony, that being the simplest model with the greatest explanatory power. Consequently, the most parsimonious models contained the lowest number of predictor variables while not reducing the model quality.

6.3.3.4. Map production

The selected model for each soil property was used, with the relevant covariates, to predict onto a 90 m resolution grid of the study area. For each soil property the five depths were mapped concurrently with the same legend to assist with interpretations of change between depths.

6.3.3.5. Visual examination

The model development procedure thus far has focused on selecting models to develop maps on a purely statistical basis. A visual assessment using expert knowledge is recommended to confirm that the maps produced by the model are pedologically realistic and are not negatively impacted by artefacts, this being the error observed in a digital signal (Padarian Campusano, 2014). Consequently, each map was interrogated for anomalies and unrealistic representations.

If present, model covariates were eliminated, and different combinations tested until the most parsimonious model resulting in an appropriating spatial representation was found.

6.3.3.6. Area of Applicability

The Area of Applicability (AoA), as published in Meyer and Pebesma (2021), is determined by whether a calculated Dissimilarity Index (DI) exceeds a threshold. It is used to identify where reliable predictions can be made based on existing soil observations and their relationship to environmental covariates (Pozza et al., (2022). The AoA was calculated using the 'aoa' function contained within 'CAST' package in 'R' (Meyer, 2021; R Core Team, 2023). The DI is "the normalized and weighted minimum distance to the nearest training data point divided by the average distance within the training data" (Pozza et al., 2022). The DI can range from 0 to ∞, with a value of 0 indicating that the new datapoint has identical predictor properties to a point in the training data (Meyer & Pebesma, 2021). A lower DI is desirable, indicating that new datapoints are more similar to training datapoints (Meyer & Pebesma, 2021). Utilising these values, the AoA is derived by considering if the DI at each point has surpassed a set threshold. Predictions are considered unreliable if the DI value at a point exceeds the set threshold, in which instance the AoA is classed as unapplicable. This study takes an applied approach to these methods. The DI summary statistics will be shown for each developed soil model. The DI will then be mapped for ECe to showcase how this varies spatially. While it is important to identify the AoA for each soil property due to variable combinations of covariates, each combination in this study was similar, so the DI for the lowest quality model will be shown as an example. For ECe, the AoA will be presented for two thresholds, the 95% and 50% quantile of the DI for ECe. This will showcase, spatially, where the most unreliable 5% and 50% of predictions are being made.

6.3.4. Point validation of external digital soil map products

Presently there are three major, publicly available and accessible DSMs that cover the entirety of New South Wales (NSW). For each of these products the predicted clay content was chosen to be assessed at the point scale with the clay content data layer extracted for the bounds of this study area at two depths: 0.3-0.6 m and 0.6-1 m. At these two depths the predicted clay content was extracted at each point location where samples were taken in this study. To assess each of the models at the point support, measured soil data from this study was used to validate extracted DSM data for each of the three products separately at both depths. Goodness of fit

plots were generated and LCCC, RMSE and bias statistics were used to assess the performance of each model.

6.3.4.1. State (NSW OEH)

The NSW DSM, developed by the Office of Environment and Heritage (OEH), covers the entirety of the state at a 100 m resolution (Gray, 2023). Modelled soil properties contained within this resource were produced using linear regression models or, in the case of clay content, cubist linear piecewise decision tree models (Gray, 2023). Data was subset to validate these models, with 80% partitioned as a training dataset and 20% as a validation dataset (Gray, 2023). At a soil depth of 0.3-0.6 m, validation statistics, determined on 1159 validation samples, report a LCCC of 0.53 and RMSE of 14.5% while at a depth of 0.6-1 m the model was validated on 875 samples reporting a LCCC of 0.43 and RMSE of 17.2% (Gray, 2023).

6.3.4.2. National (Soil and Landscape Grid of Australia)

The SLGA is a nationwide DSM of Australia available at a 3 arc second (approximately 90 m) resolution (Grundy et al., 2015). At each point on this grid there are predictions of 20 soil attributes at depths of 0-0.05, 0.05-0.15, 0.15-0.3, 0.3-0.6, 0.6-1 and 1-2 m (Grundy et al., 2015; Viscarra Rossel et al., 2014). The modelling for each of these attributes occurred using a cubist algorithm with validation undertaken on bootstrap out-of-bag samples with a 10-fold cross validation with data partitioned into training and testing subsets (Viscarra Rossel et al., 2014). For the clay model, 5117 sites from NSW were used to train the model while 2160 were used to test it. In total, 16258 sites were used to train and test the model nationally with the most common Australian Soil Classification (ASC) order of these sites being sodosols (n=3573) (Viscarra Rossel et al., 2014). For the 0.3-0.6 m depth increment a test validation LCCC of 0.56 and RMSE of 14.27% are presented while at 0.6-1 m soil depth these validation statistics are 0.53 and 14.23%, respectively (Viscarra Rossel et al., 2014).

6.3.4.3. Global (SoilsGrids250m)

The SoilGrids250m product is available at a 250 m resolution with organic carbon, bulk density, CEC, pH, soil texture fractions and coarse fragments predicted at depths of 0-0.05, 0.05-0.15, 0.15-0.3, 0.3-0.6, 0.6-1 and 1-2 m (Hengl et al., 2017a; Poggio et al. 2021). This product consists of a ‘top-down’ approach, where a single model is used to produce a global map as opposed to a ‘bottom-up’ approach where local maps are joined to produce global maps (Han et al., 2022). The updated version of this product increased the covariates used from 158 to 500 (Hengl et al., 2017a; Poggio et al. 2021). These were used to fit machine learning

methods including random forest, gradient boosting and/or multinomial logistic regression. Validation statistics are presented at each depth, for each soil property, calculated through a 10-fold cross validation across different spatial strata (Poggio et al., 2021). There is difficulty with the interpretability of the model statistics. Concordance is reported as the model efficiency coefficient (MEC), with values of 0.41 and 0.40 for 0.3-0.6 and 0.6-1 m depths, respectively (Poggio et al., 2021). RMSE is reported as a combined statistic for all depths. This value is stated as 0.13, however, the unit of this statistic is unclear (Poggio et al., 2021).

6.4. Results

6.4.1. Exploratory analysis of soil datasets

Significant soil variability is seen for all measured soil properties (Table 6.2). Interpretations of these measurements throughout the results and discussion are in accordance with descriptions presented by Hazleton and Murphy (2007).

The texture of soils in the lower Murrumbidgee valley varies significantly, with some soils considered heavy clays (>50%) throughout the first metre of the soil profile while others exhibit a low clay content (<5%) throughout (Table 6.2). It is observed that clay content is strongly, positively correlated with CEC; as one value increases so to does the other, while the inverse is true for sand content (Table 6.2).

Across the dataset there are samples that range from being considered highly saline (ECe > 8 dS/m) to non-saline (ECe < 2 dS/m) at each depth (Table 6.2). Variability in the dataset increases with depth, as indicated by higher standard deviation statistics at 0.6-0.8 and 0.8-1 m (Table 6.2). Similarly, some sites in this study are strongly sodic throughout the first metre of the soil profile (ESP>14) while there were also samples, at each depth, with ESP values <1. (Table 6.2).

Soils within the region are generally neutral to alkaline at the surface with alkalinity increasing with depth (Table 6.2). At each depth, however, there are samples that are slightly acidic to acidic (Table 6.2).

Table 6.2. Summary statistics for pH, ECe, ESP, CEC, sand and clay at the five examined depths 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m). Represented for each soil property, at each depth (m), is the number of samples (n) and the mean, standard deviation, minimum and maximum measured values.

Soil attribute	Depth	n	Minimum	Mean	Maximum	Standard deviation
pH	0-0.1	153	5.27	7.4	9.1	0.8
	0.1-0.3	153	4.95	7.9	9.8	0.8
	0.3-0.6	153	6.34	8.4	10.2	0.7
	0.6-0.8	153	6.11	8.5	10.2	0.8
	0.8-1	152	6.13	8.5	10.2	0.8
ECe (dS/m)	0-0.1	153	0.14	1.6	13.1	1.5
	0.1-0.3	152	0.03	2.9	25.1	4.5
	0.3-0.6	153	0.04	5.1	37.2	7.5
	0.6-0.8	153	0.04	7.5	49.1	10.1
	0.8-1	150	0.02	9.2	42.1	10.7
ESP (%)	0-0.1	153	1.80	16.0	1.8	6.3
	0.1-0.3	153	1.50	18.4	1.5	6.7
	0.3-0.6	153	0.94	20.5	0.9	6.9
	0.6-0.8	153	0.77	22.0	0.8	8.6
	0.8-1	152	0.88	22.7	0.9	10.6
CEC (cmol(+)/kg)	0-0.1	153	0.20	5.4	26.3	4.8
	0.1-0.3	153	0.90	9.6	38.1	8.3
	0.3-0.6	153	0.80	13.9	45.2	10.8
	0.6-0.8	153	0.60	14.6	45.5	10.3
	0.8-1	152	0.20	14.1	48.3	9.5
Clay (%)	0-0.1	153	4.30	41.0	64.5	12.9
	0.1-0.3	152	1.20	45.3	70.0	12.4
	0.3-0.6	153	1.30	46.5	63.0	12.4
	0.6-0.8	153	3.36	47.2	63.9	12.2
	0.8-1	150	1.50	45.8	70.3	12.7
Sand (%)	0-0.1	153	19.10	39.5	89.8	14.4
	0.1-0.3	152	16.71	35.8	93.9	15.0
	0.3-0.6	153	11.60	33.7	90.1	15.2
	0.6-0.8	153	0.91	32.4	90.1	14.8
	0.8-1	150	1.01	33.1	90.5	15.4

6.4.2. Soil property model development and map production

6.4.2.1. Removal of covariates prior to modelling

Three variables; barest earth SWIR 2, barest earth green and radiometric dose were removed from the initial covariate stack due to collinearity. In visually examining the remaining rasters, the decision was made to remove the terrain attribute covariates MrRTF and MrVBF (Figure 6.3). Consequently, 11 predictor variables were used in the initial modelling process prior to the stepwise deletion.

6.4.2.2. Optimal soil property model selection

When using the most parsimonious model, following a stepwise deletion, predictions of pH, ECe and ESP resulted in an unrealistic spatial pattern. By further analysing covariate combinations, this was attributed to DEM which was the most influential predictor in each model. While removing this predictor from the model reduced the LCCC (Appendix 1), it resulted in a more accurate representation of the landscape. This effect was not observed in the models for clay, sand and CEC. As a result of this, no manual covariate deletion occurred, and the most parsimonious model was used for these attributes.

There are different combinations of predictor variables for each of the six soil property models (Table 6.3). Barest earth blue and radiometric potassium (%) were retained in all models (Table 6.3). The model for soil clay content contained the greatest number of predictor variables (n=11). TWI and slope were the variables most commonly removed during the deletion process.

Table 6.3. The covariates utilised in the development of models to predict pH, ECe, ESP, CEC, sand and clay following a stepwise deletion of predictor variables. A 'Y' indicates that the variable was utilised in the final model while a '-' indicates that it was not included.

Covariate	**Soil property model**					
	pH	**ECe (dS/m)**	**ESP (%)**	**CEC (cmol(+)/kg)**	**Sand (%)**	**Clay (%)**
Layer mid depth	Y	Y	Y	Y	Y	Y
DEM (m)	-	-	-	Y	Y	Y
Slope (%)	Y	-	Y	-	-	-
TWI	-	Y	-	-	-	Y
Uranium (%)	Y	Y	-	Y	Y	Y
Potassium (%)	Y	Y	Y	Y	Y	Y
Thorium (%)	Y	-	-	Y	Y	Y
Silica (%)	-	-	-	Y	Y	Y
Barest earth blue	Y	Y	Y	Y	Y	Y
Barest earth red	Y	Y	Y	-	-	Y
Barest earth NIR	-	-	Y	Y	Y	Y
Barest earth SWIR	-	Y	Y	-	Y	Y

6.4.2.3. Soil property model statistics

The validation statistics generated using a LOSOCV show each model to be of moderate quality with LCCCs ranging from 0.45 (ECe) to 0.58 (clay) (Figure 6.4). Due to differences in the units of measurement the RMSE statistics cannot be compared in the same way. With the exclusion of sand content, each model shows a low, negative bias (Figure 6.4). The vertical scatter of observed samples with a high sand content (Figure 6.4) suggests that these values are not well represented in the model.

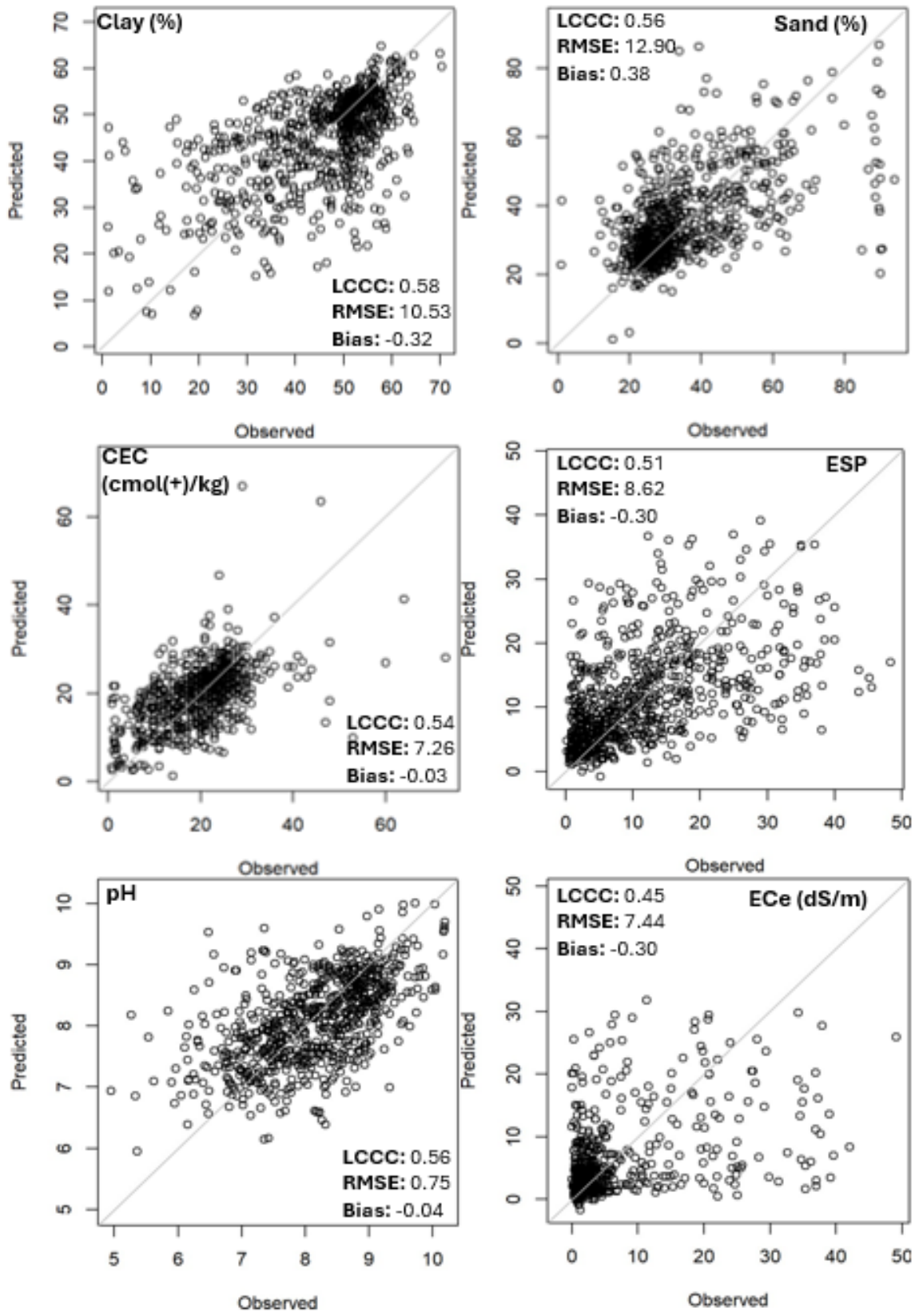


Figure 6.4. Leave one-site-out-cross-validation (LOSOCV) statistics for modelled soil clay, sand, CEC, ESP, pH and ECe. The reported statistics for each soil property model are the Lin's Concordance Correlation Coefficient (LCCC), Root Mean Square Error (RMSE) and bias.

6.4.2.4. Variable importance and model influence

The variables included in the final model for, and SHAP rankings relating to, CEC as well as clay and sand content are similar. For both clay and sand content, radiometric thorium is the highest ranked predictor (Figure 6.5). When assessing these plots, the relationship is inversed so that where higher thorium positively contributes to clay content, it negatively contributes to sand content (Figure 6.5). When examined using the dependence plot the local interpretation is clearer, with a relatively gradual trend, from low thorium values positively contributing to sand content, to high thorium values negatively contributing (Figure 6.7). For both sand and clay content, variables ranked above DEM showcase a high data spread indicating a broad range of influence (Figure 6.5).

A similar spread is shown when examining the CEC beeswarm plot (Figure 6.6). The highest ranked variable in this instance, layer mid depth, exhibits a clear monotonic relationship whereby a greater layer mid depth, or lower position in the soil profile, positively contributes to CEC. There are multiple points where a high mid-depth has a strong, positive model impact (Figure 6.6). These specific relationships likely represent data from lower in the soil profile, reflecting a greater mid-depth, from sites sampled in the west of the region that had a very high CEC (Table 6.2). The dependence plots can also allow for examinations of relationships between predictor variables at each datapoint (Figure 6.8). In this instance, the x and y axes show the relationship between different elevations and the resulting SHAP value for predicted datapoints, respectively (Figure 6.8). The colouration of each point represents the barest earth blue value at said point, showing that generally as elevation increases barest earth blue values decrease (Figure 6.8). As this occurs, higher DEM values initially positively contribute to clay content, before generally negatively contributing beyond a value of 120 m. The vertical distribution of some sample points is due to multiple depths being modelled concurrently and instances where cores were taken in close proximity to each other or from similar longitudes, as the region follows a gentle east-west gradient (Figure 6.8).

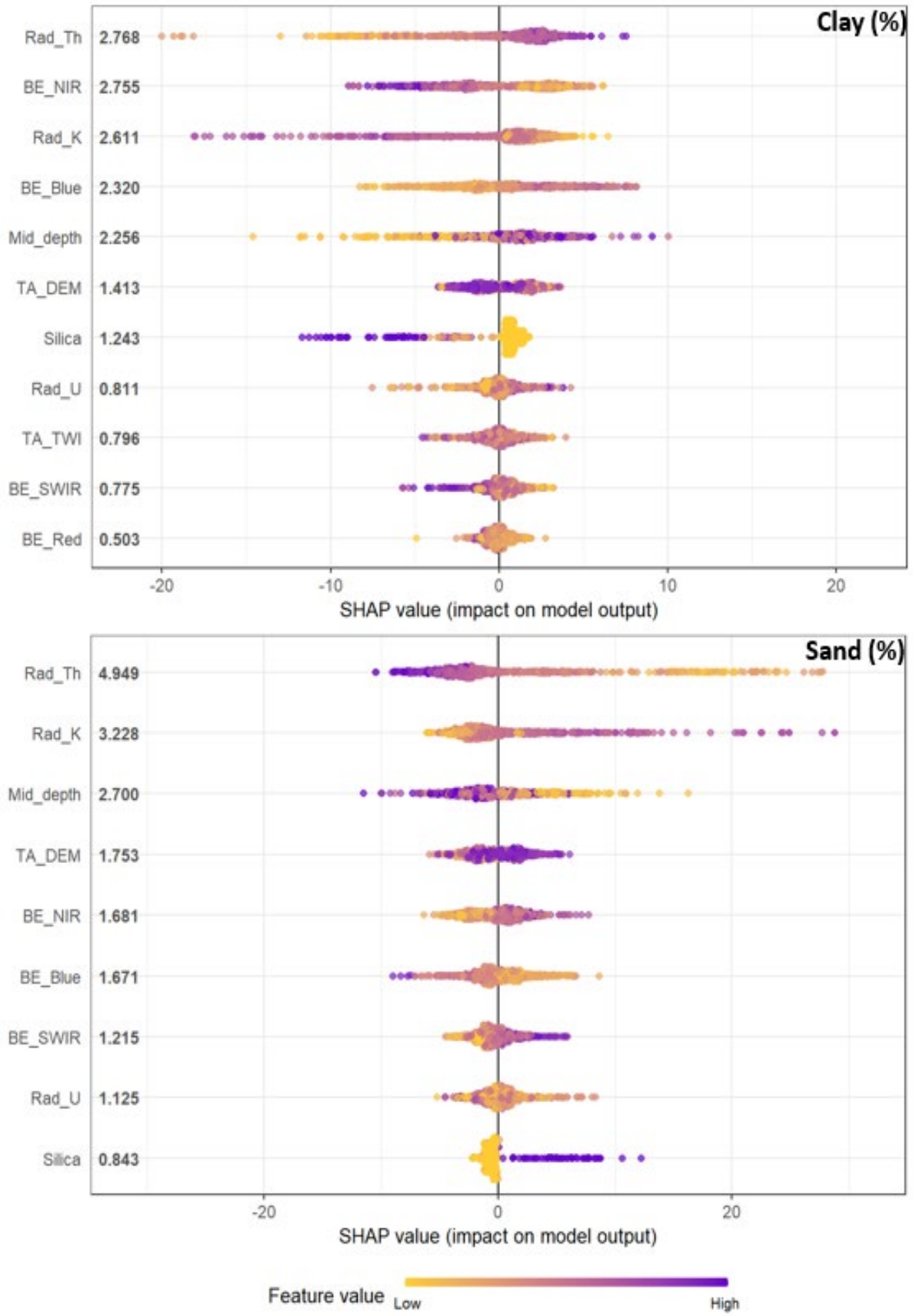


Figure 6.5. SHAP beeswarm plots for soil clay (%) (top) and sand (%) (bottom) content.

Figure 6.6 SHAP beeswarm plots for pH (top) and CEC (cmol(+)/kg) (bottom).

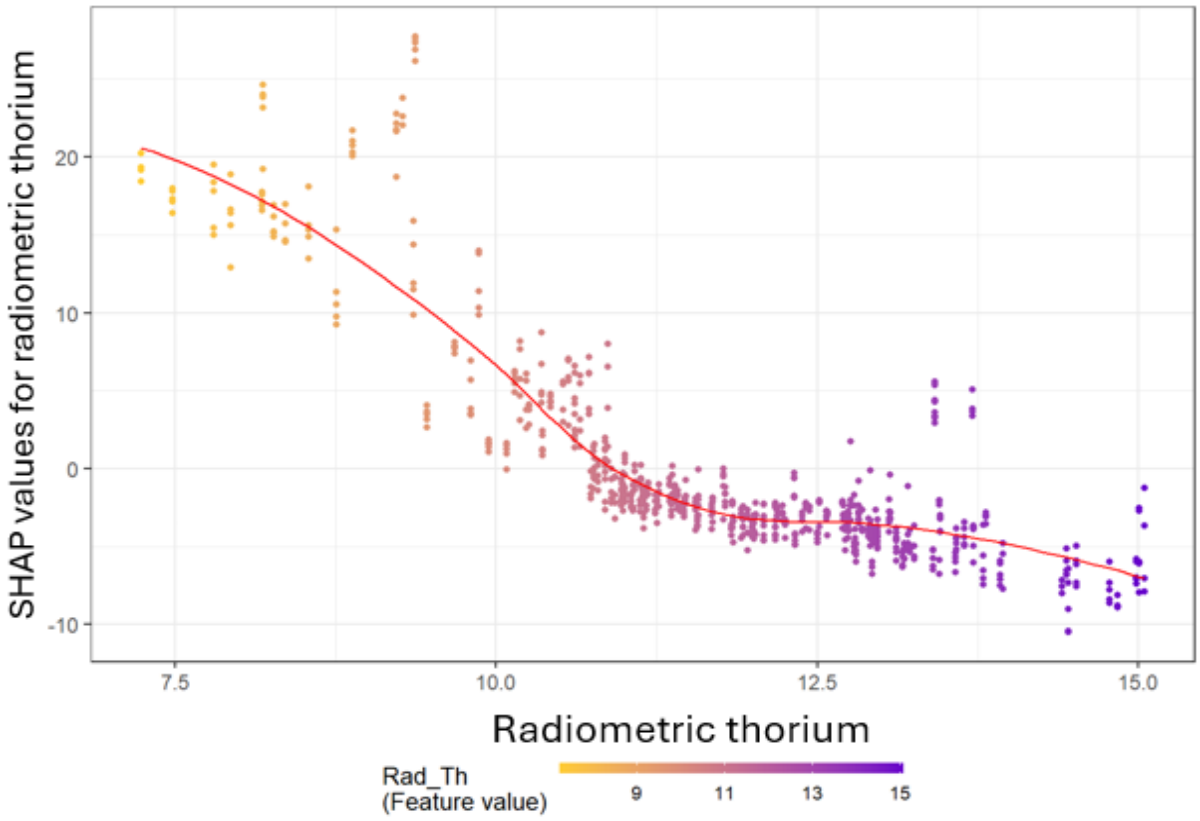


Figure 6.7. A SHAP dependence plot for radiometric thorium in the model predicting soil clay content. The x-axis and colour of datapoint show the radiometric thorium measurement and the y-axis shows the SHAP value. Each point on the plot is sample from this study.

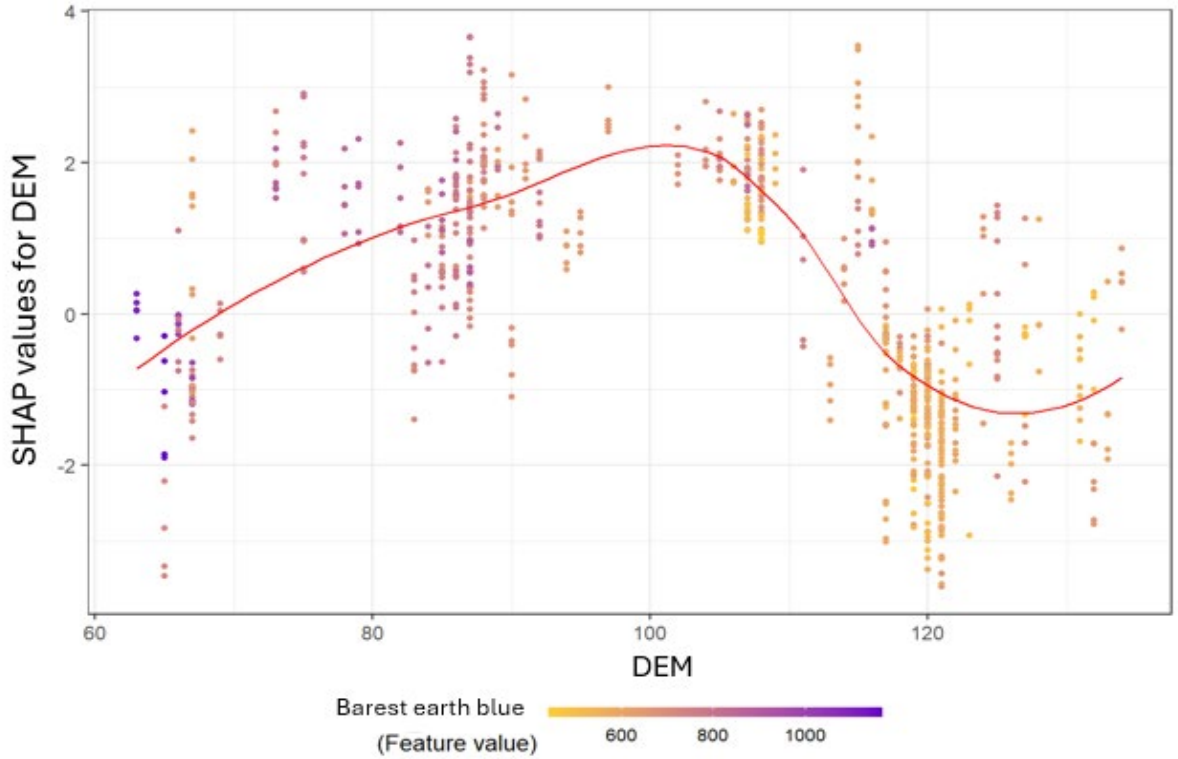


Figure 6.8. A SHAP dependence plot for DEM (elevation) in the model for clay. The x-axis shows the elevation (m), the y-axis shows the corresponding SHAP value at datapoints. The colour scale shows the barest earth value at each sample point.

Apart from the mid-depth variable, DEM was originally the highest ranked predictor variable for pH, ECe and ESP (Appendix 2). This, however, resulted in a contour-like banding effect when maps were generated (Appendix 3). In removing this variable, the ranking of the previously lower spatial covariates changes, with barest earth blue becoming the highest ranked variable, with the exception of layer mid-depth, for pH, ESP and ECe (Figure 6.6; Figure 6.9). The highest ranked predictor variable, mid-depth, is shown to positively contribute to pH (Figure 6.6), ECe (Figure 6.9) and ESP (Figure 6.9). The dependence plot for mid-depth in relation to pH provides more detail on this, with the positive contribution of layer mid-depth to predicted pH values initially increasing before plateauing, as indicated by the SHAP values (Figure 6.10). This is an expected occurrence as these soil properties generally increase at greater depths within the profile.

The beeswarm plot for pH (Figure 6.6) shows that high feature values for barest earth blue both positively and negatively contribute to predicted values. The remaining variables showcase clustering of points around a SHAP value of zero, however, commonly exhibit extended tails. Despite clustering the lowest values of radiometric potassium, for example, produce negative SHAP scores suggesting a negative contribution to pH predictions (Figure 6.6). The inverse of this is true for radiometric uranium and thorium, where high values of these features generally produce low SHAP scores (Figure 6.6). Despite significant clustering (Figure 6.6), slope shows two clear points where a high slope percentage results in high SHAP scores, positively contributing to pH largely irrespective of layer mid-depth (Figure 6.11).

Barest earth satellite covariates are generally more highly ranked in predicting ESP than ECe, where, in the latter, radiometric covariates assume more influence (Figure 6.9). In predicting ESP, high attribute values for barest earth blue and SWIR are associated with positive contributions despite clustering in the latter. A similar relationship is observed in predictions for ECe (Figure 6.9), however, in this instance barest earth SWIR is the lowest ranked predictor. Despite significant clustering for all predictors of ECe, there is a high data spread, especially for positive SHAP values, suggesting a broad range of influence. Monotonic, or linear, trends are less common for pH, ECe and ESP than they are for CEC, sand and clay content. This may be the result of greater variability in these properties in the subsoil impacting relationships between predictor variables.

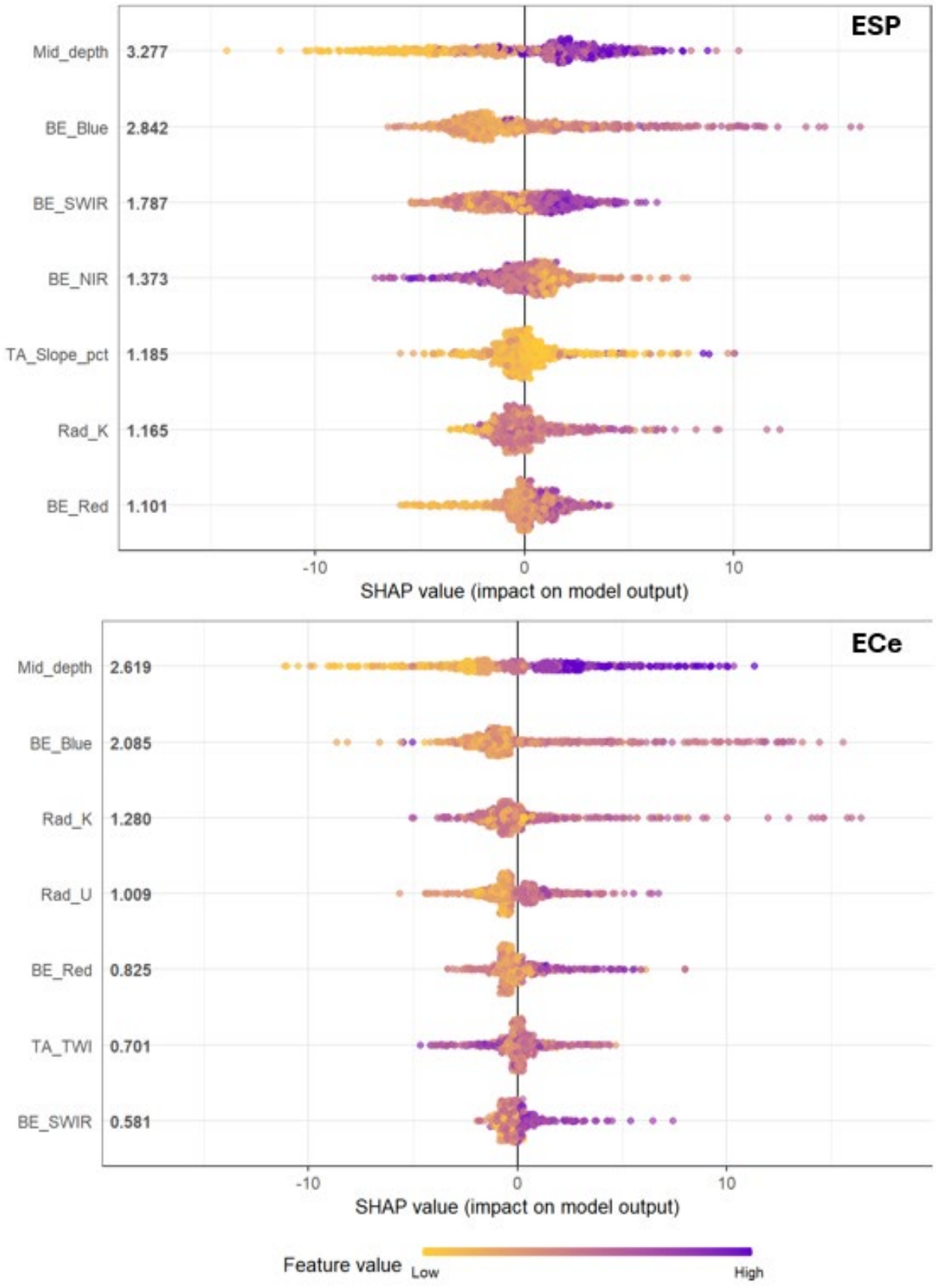


Figure 6.9. SHAP beeswarm plots for ESP (%) (top) and ECe (dS/m) (bottom).

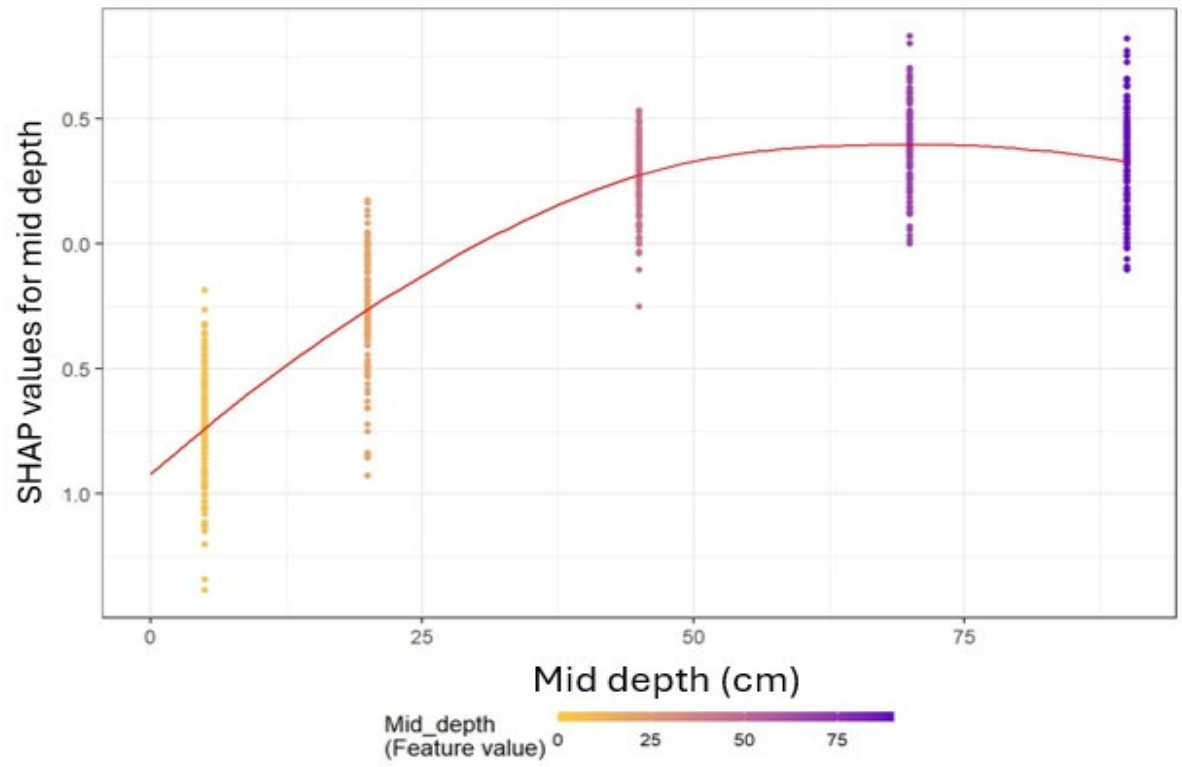


Figure 6.10. A SHAP dependence for layer mid-depth in the model predicting soil pH. The x-axis and colour bar represent the mid-depth attribute value while the y-axis shows the corresponding SHAP value.

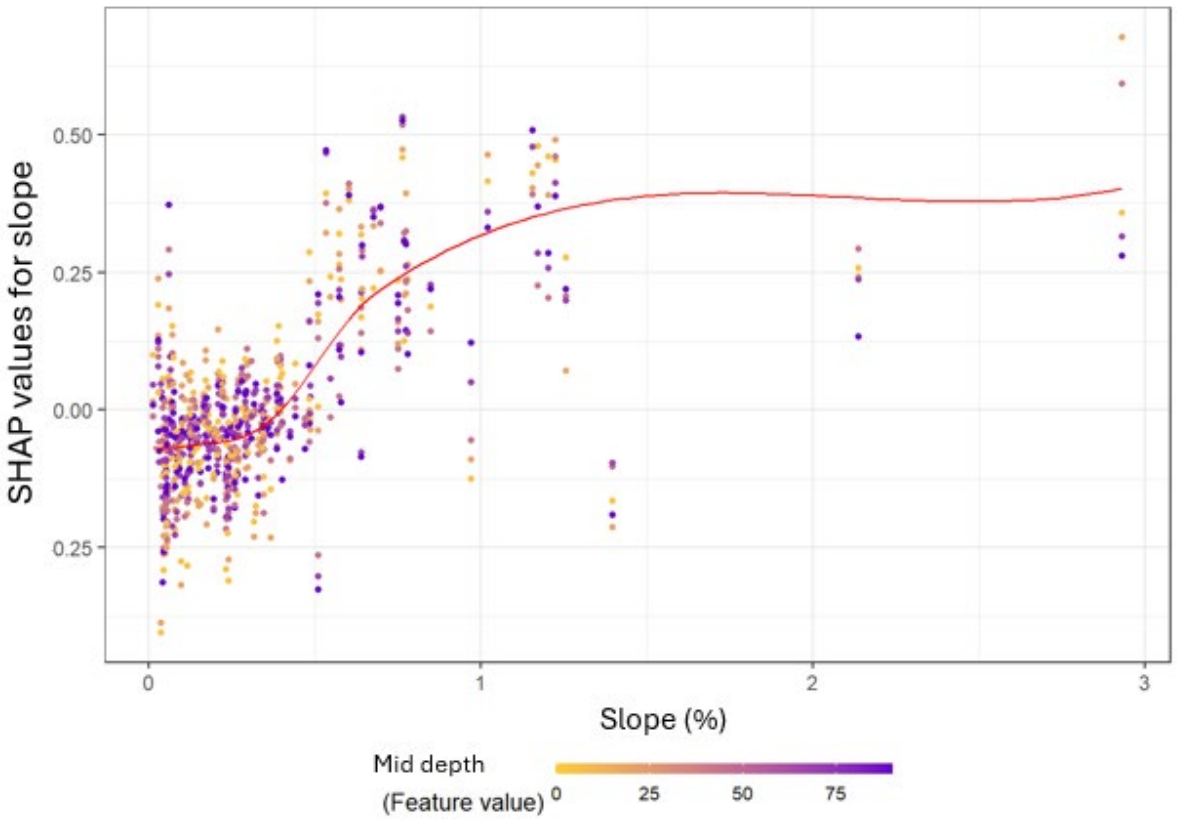


Figure 6.11. A SHAP dependence plot for slope in the model predicting pH. The x-axis shows the slope (%) value while the y-axis shows the corresponding SHAP values. The colour bar and colouring of each sample is the mid-depth (cm).

6.4.3. Digital soil maps

The initial maps that included DEM as a predictor for pH, ECe and ESP produced an unrealistic banding effect (Appendix 3). These abrupt changes are not a realistic reflection of the landscape. By removing DEM, a more gradual and pedologically appropriate trend for soil pH (Figure 6.12), ECe (Figure 6.13) and ESP (Figure 6.14) is produced. There is still a slight increase in each attribute along the east-west axis, however, the soil properties are more closely associated with paleochannel systems. In the east and throughout the centre of the region there are pockets of acidity (pH<7) predicted in the surface layer (Figure 6.12). These generally align with land that has been under irrigated agricultural production for an extended period. There is significant alkalinity in the west and north of the region, with alkaline tracts of soil also observed in the southeast (Figure 6.12). Across the entirety of the region a trend of pH increasing with depth is predicted (Figure 6.12).

A similar trend with depth is seen in the predicted ECe (Figure 6.13) and ESP (Figure 6.14), with the most sodic soils predicted to occur at 0.3-0.6 m depth between longitudes of 144-144.5º (Figure 6.14). At these locations ECe values (Figure 6.13) are also predicted to be highest, however, unlike ESP they continue to increase below 0.6 m (Figure 6.14). Based on these spatial trends, soils at some point in the first metre of the profile, will be generally considered strongly sodic (ESP>14) with the exclusion of areas in the regions far east (Figure 6.14). Similarly, most soils west of 145.5º would be considered extremely saline (ECe>16 dS/m) at some point within the first metre of the profile (Figure 6.13). Unlike sodicity, however, the surface soil is not predicted to cross the salinity threshold as frequently (Figure 6.13).

The produced maps for clay content (Figure 6.15) and sand content (Figure 6.16), as well as CEC (Figure 6.17) showcase similar spatial trends whereby a high clay content is mirrored by a high CEC and low sand content. Patterns of predicted low clay content, low CEC and high sand content soils follow the pathways of palaeochannel systems (Figure 6.15, Figure 6.16, Figure 6.17). Outside of these areas, soils are generally clayey with minimal differentiation between the five measured depths (Figure 6.15). In the far west of the region loamy soils with a clay content of 20-30% (Figure 6.15) and higher sand content (Figure 6.16) are predicted to be common. There is a trend of CEC increasing to become very high (>40 cmol(+)/kg) at depths of 0.6-0.8 and 0.8-1 m in the west of the region between longitudes of 144-144.5º (Figure 6.17). At these locations, ESP and ECe are also predicted to be high.

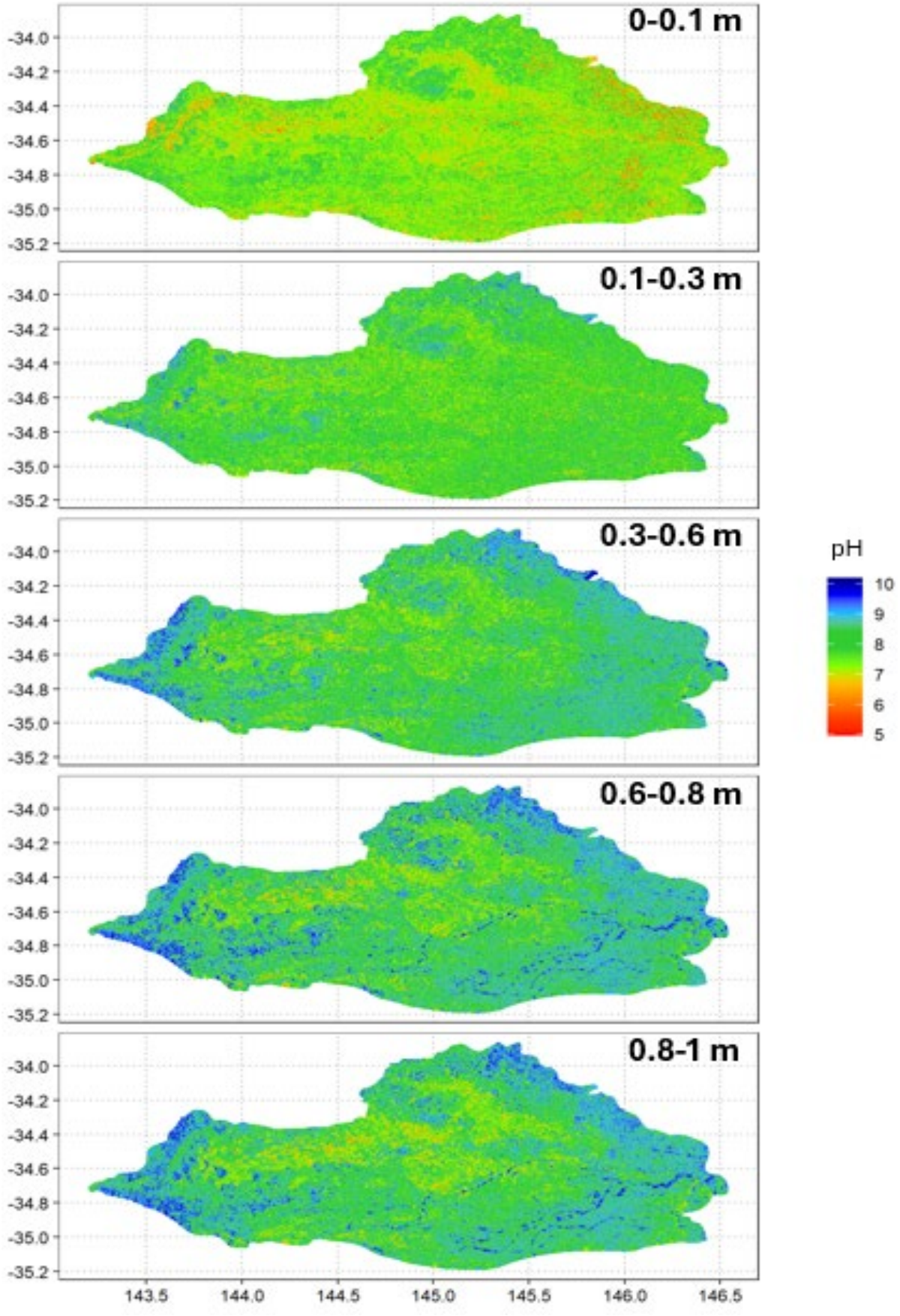


Figure 6.12. A digital soil map of pH across the study area at depths of 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m with a continuous scale.

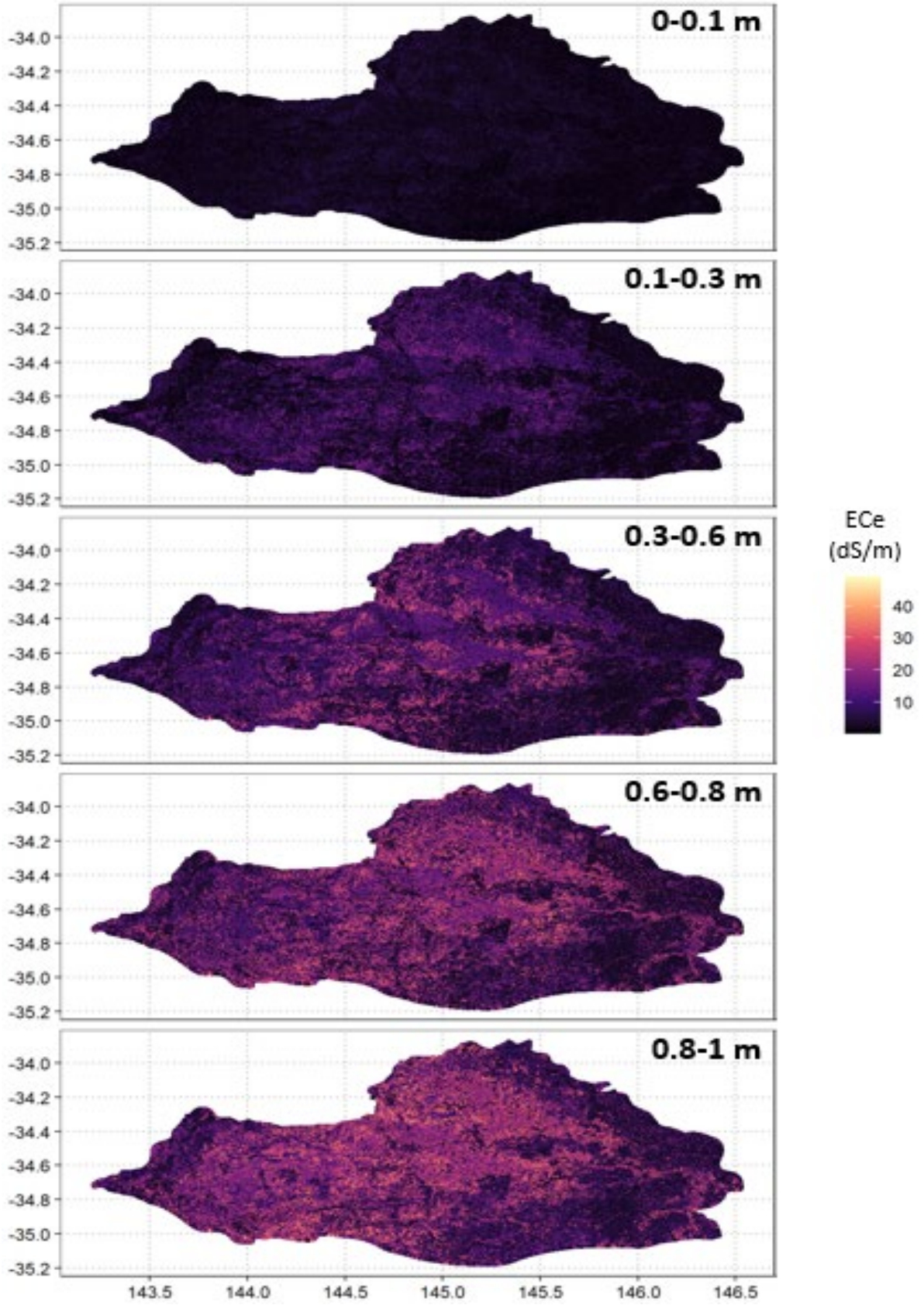


Figure 6.13. A digital soil map of soil electrical conductivity of the extract (ECe) across the study area at depths of 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m with a continuous scale.

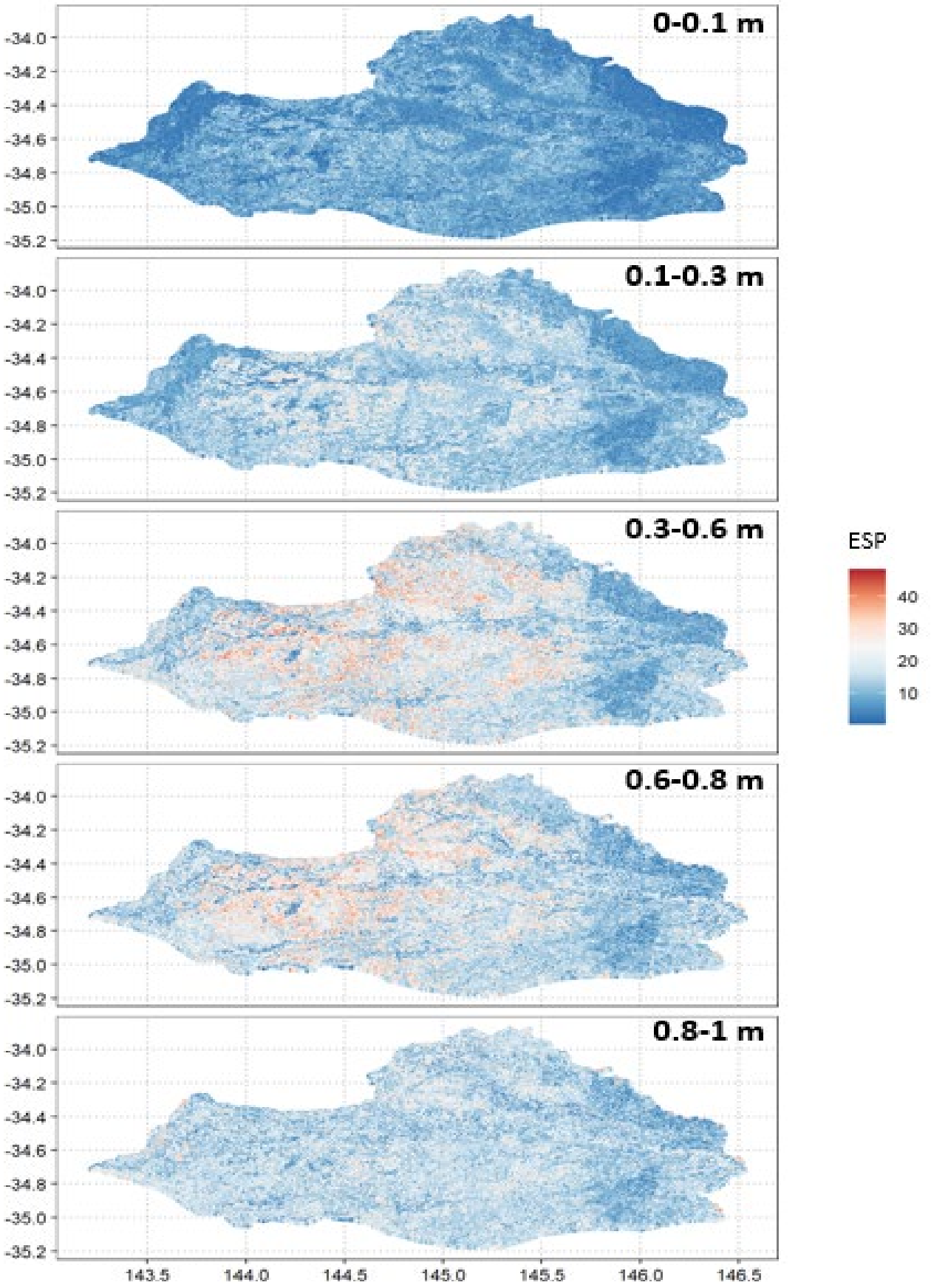


Figure 6.14. A digital soil map of exchangeable sodium percentage (ESP) across the study area at depths of 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m with a continuous scale.

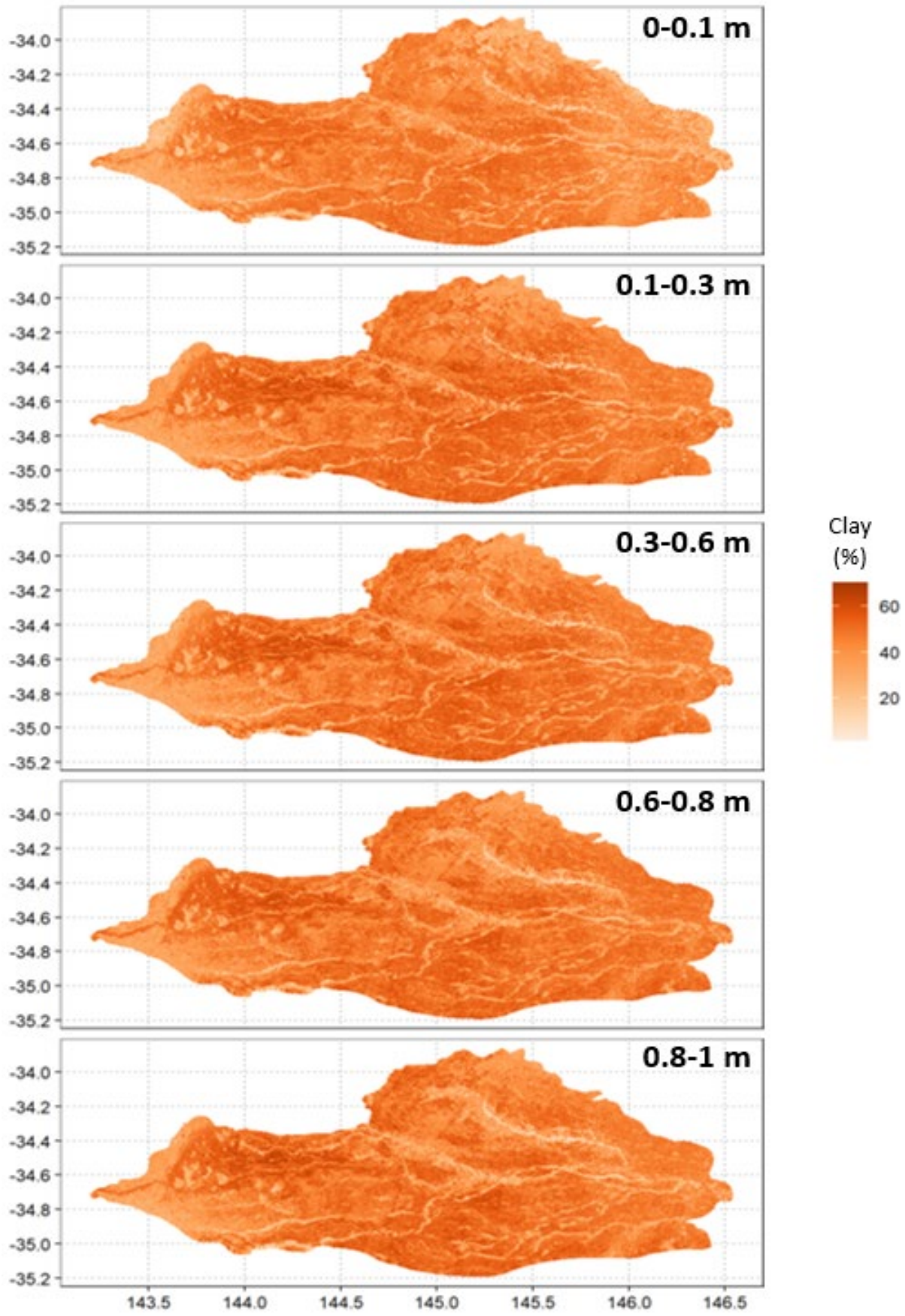


Figure 6.15. A digital soil map of clay content (%) with a continuous scale for concurrently modelled depths of 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m.

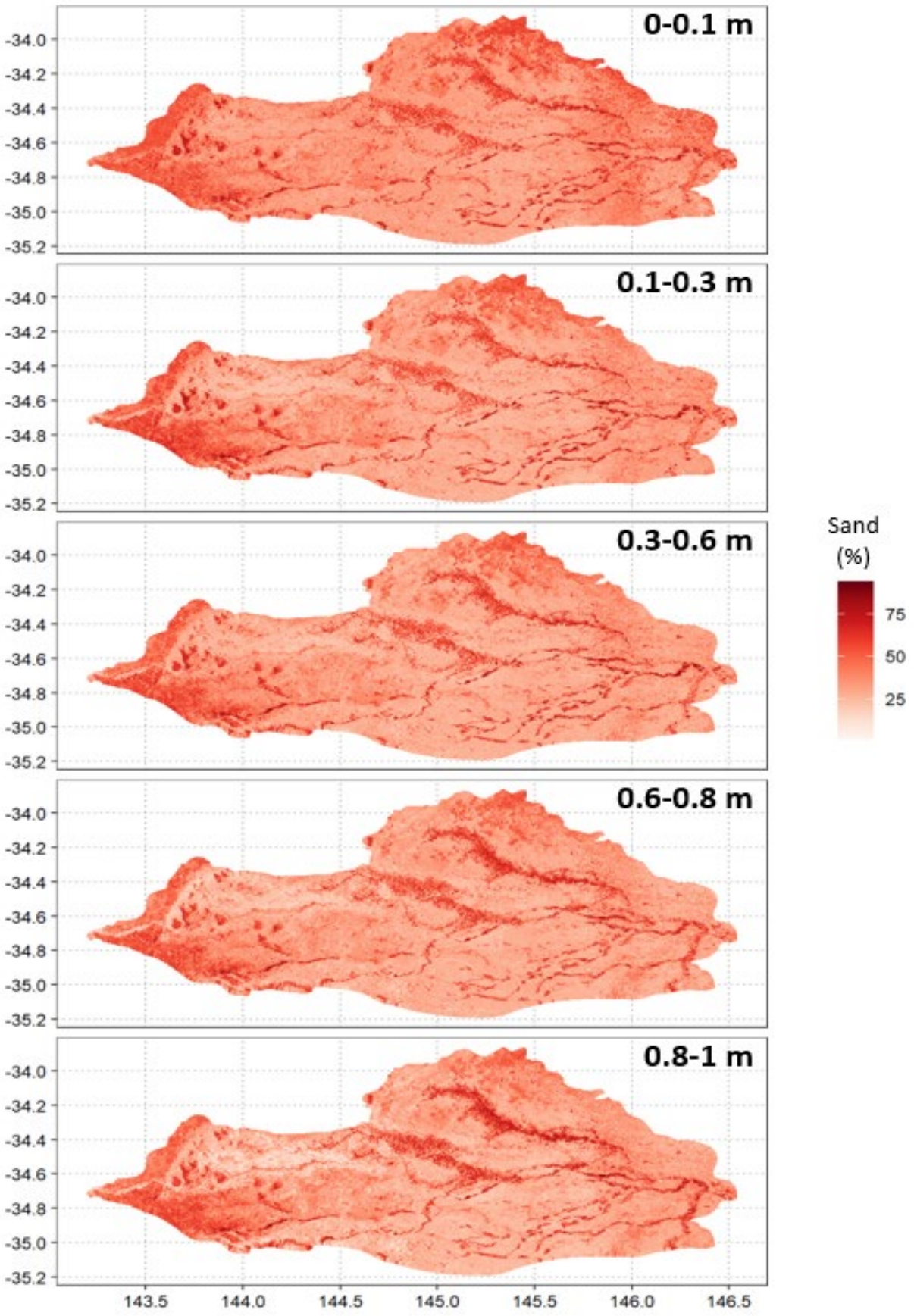


Figure 6.16. A digital soil map of sand content (%) with a continuous scale for concurrently modelled depths of 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m.

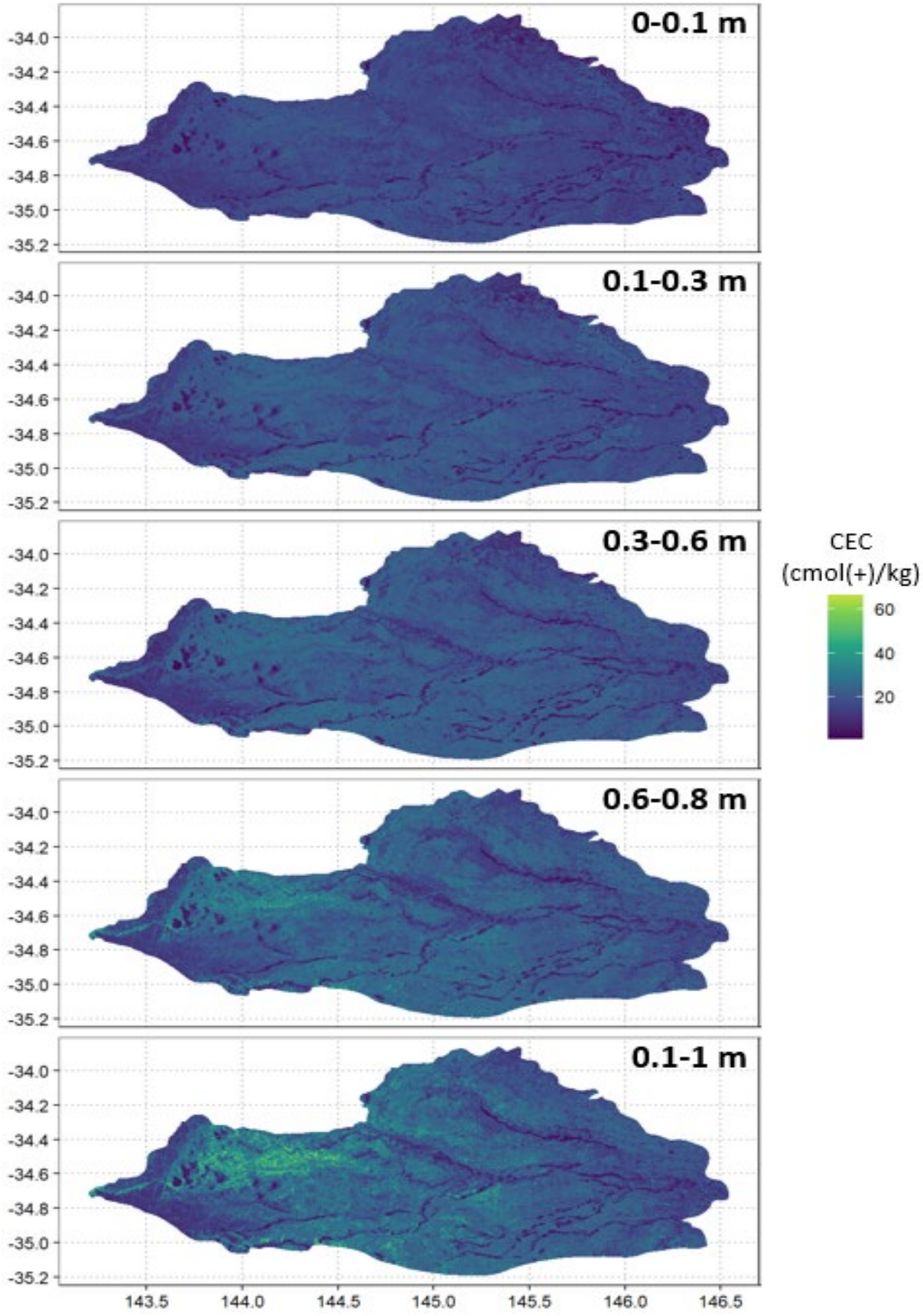


Figure 6.17. A digital soil map of cation exchange capacity (cmol(+)/kg) with a continuous scale for concurrently modelled depths of 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m.

6.4.4. Area of applicability

Each model was shown to exhibit a similar DI range (Table 6.4), with only the models for pH (11.4) and ESP (7.36) showcasing a maximum DI significantly greater than the other soil property models. A DI closer to zero is considered desirable, with the resulting map showing that areas with a greater dissimilarity are generally located near watercourses (Figure 6.18). When utilising the DI to map the AoA with a 95% and 50% threshold, each soil property showed similar spatial trends as these values are related to the distance between point predictions and sample sites as well as relationships with environmental covariates. In the former this was constant while in the latter instance the assembled covariate stacks were similar. Consequently, only the AoA for ECe, the model with the poorest quality statistics (Figure 6.4), is shown. The subsequent maps show that areas where the model is considered unreliable are most frequently associated with areas very closer to the Murrumbidgee River, along a paleochannel systems in the north of the region or in the far west of the region (Figure 6.19; Figure 6.20).

Table 6.4. The median, mean, 95% quantile and maximum Dissimilarity Index (DI) for each of the developed soil models.

Soil property model	Median DI	Mean DI	95% quantile DI	Maximum DI
pH	0.20	0.24	0.61	11.4
ECe	0.25	0.27	0.48	4.94
ESP	0.21	0.25	0.49	7.36
Clay content	0.30	0.34	0.61	3.45
Sand content	0.26	0.30	0.58	3.90
CEC	0.29	0.33	0.58	3.55

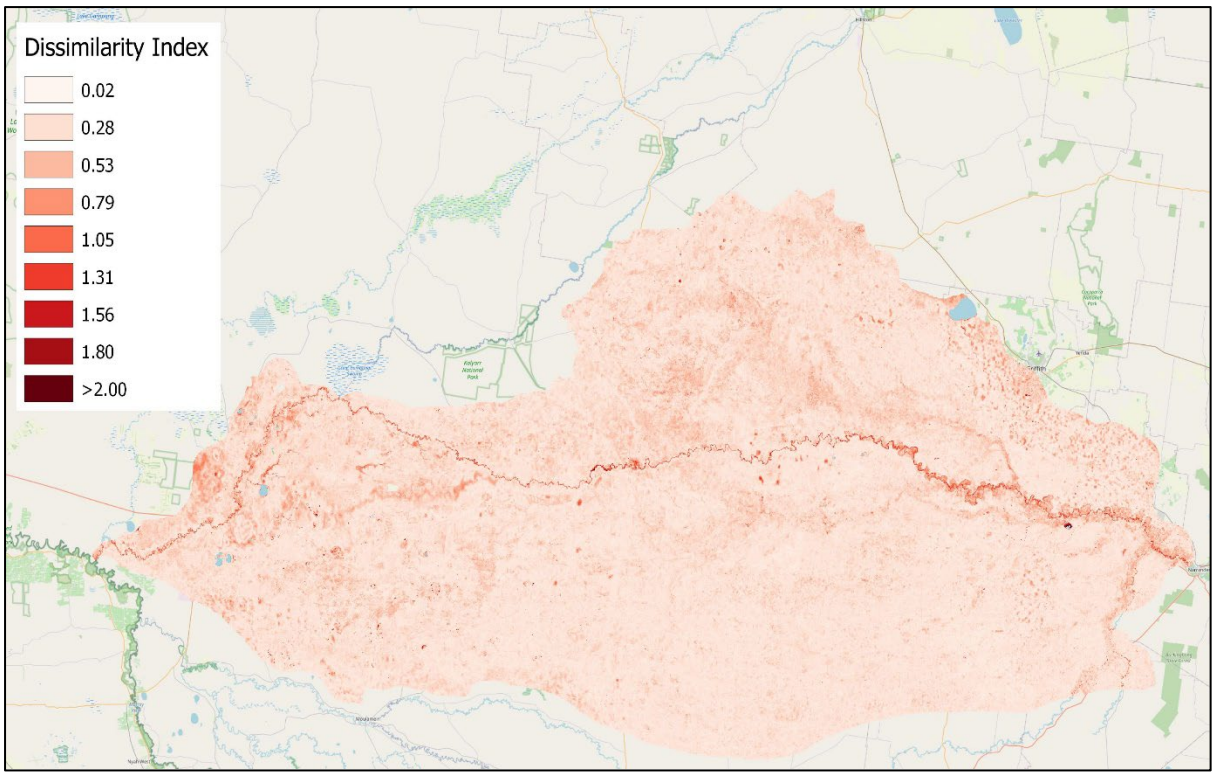


Figure 6.18. A map showing variability in the Dissimilarity Index (DI) values obtained for the ECe model.

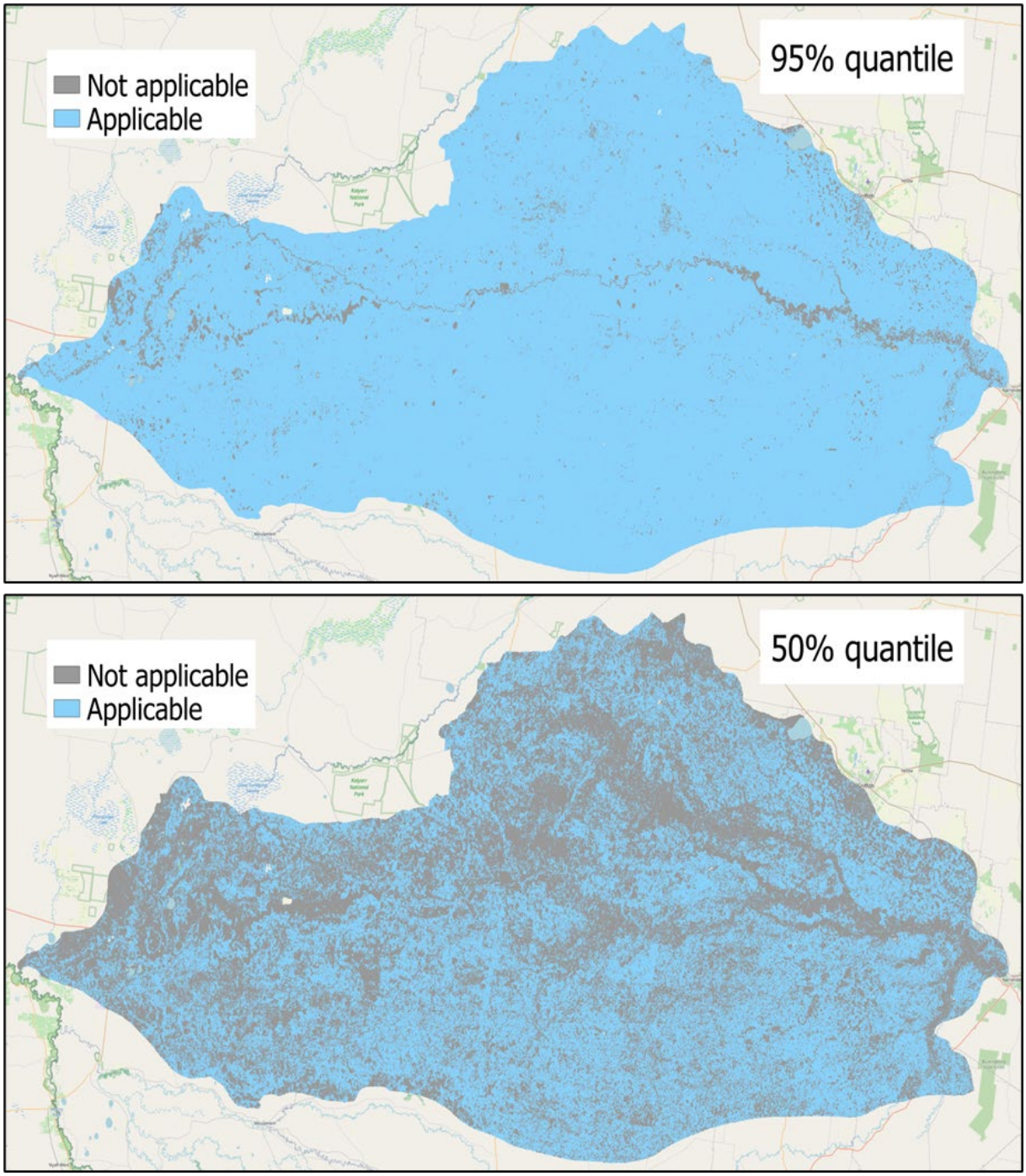


Figure 6.19. Area of applicability maps determined utilising the 95% quantile (top) and 50% quantile (bottom) of the Dissimilarity Index (DI) for model predictions of soil electrical conductivity of the extract (ECe).

6.4.5. Point validation of external DSMs

The accuracy of the state (NSW) model in predicting clay content at each sampled point was moderate at the 0.3-0.6 m depth, with an LCCC of 0.50 and low at the 0.6-1 m depth, with an LCCC of 0.36 (Figure 6.20). At both depths a positive bias of 2.51% and 3.24%, respectively, indicates a propensity for the model to overfit (Figure 6.20). While there is a clear horizontal pattern in the scatter plots where the model predicts between 40-60 % clay, there is some variability, with a capacity to predict a lower clay content (Figure 6.20).

The accuracy of the point predictions for the national (SLGA) DSM was low, with LCCCs of 0.16 and 0.15 at depths of 0.3-0.6 and 0.6-1 m, respectively (Figure 6.20). As shown by the horizontal scatter of points between 40-60%, there is little variation in predictions when the measured clay content is low (<40%). Consequently, a positive bias statistic highlights that overprediction is occurring with both depths exhibiting similar RMSEs (Figure 6.20).

The global (SoilsGrid250m) DSM predicts poorly at the 0.3-0.6 m depth, with an LCCC of 0.11 and very poorly at the 0.6-1 m depth, with an LCCC of 0.07 when assessed at the measured points (Figure 6.20). Like the other two assessed DSMs there is little variation in the predictions as indicated by the horizontal scatter of points at both depths (Figure 6.20).

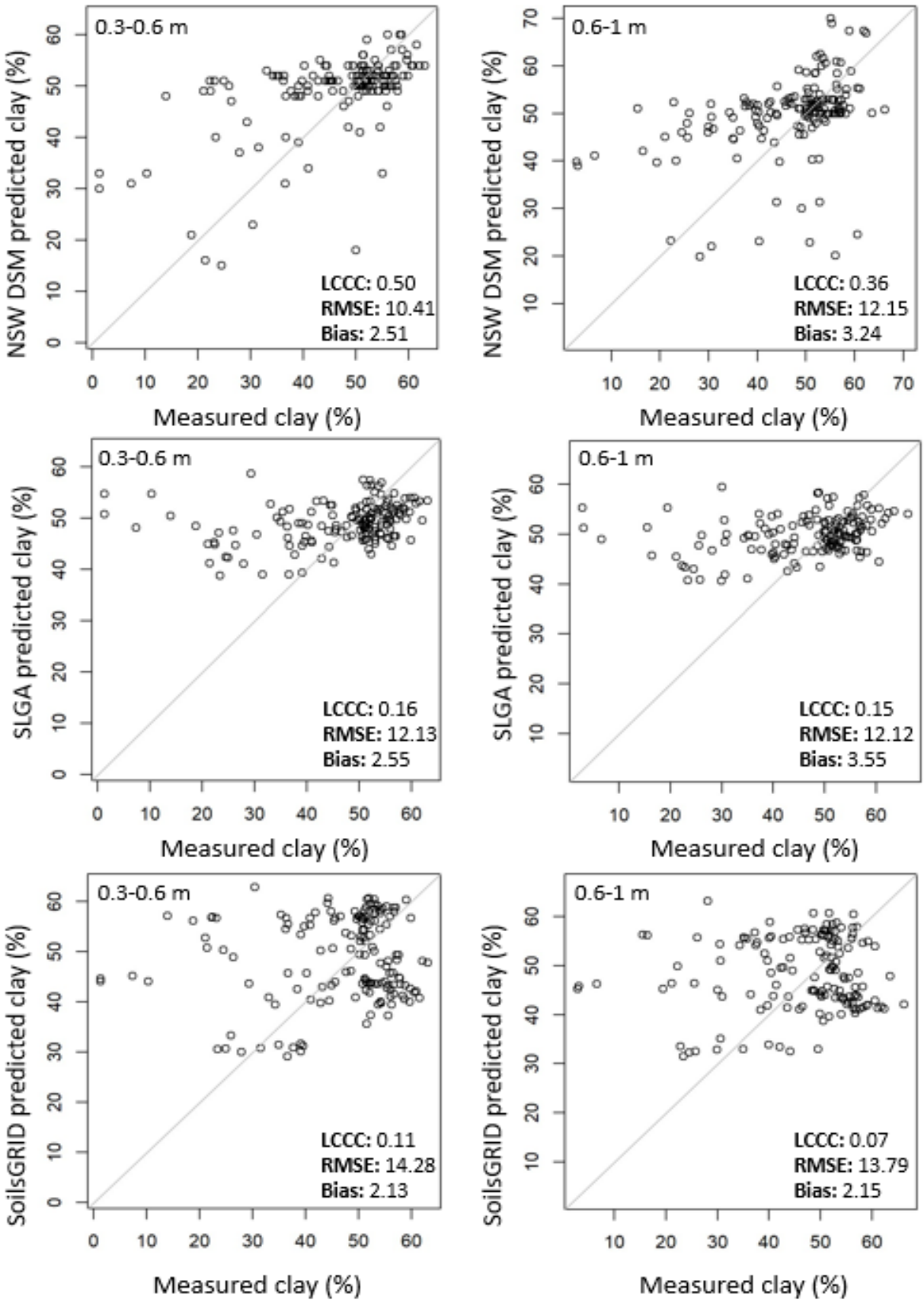


Figure 6.20. Goodness of fit plots for the state (top), national (centre) and global (bottom) DSM predictions of clay content (%) validated against measured clay content from sites in this study at 0.3-0.6 m (left) and 0.6-1 m (right).Reported goodness of fit statistics are the Lin's Concordance Correlation Coefficient (LCCC), Root Mean Square Error (RMSE) and bias.

6.5. Discussion

6.5.1. Variable selection and importance

Covariate reduction to produce the most parsimonious models was undertaken in this study, with the overall number of predictor variables, including mid-depth, reduced from 11 to between 10 (clay) and 7 (pH, ECe, ESP, CEC). These reductions decrease complexity and increase interpretability. In an extreme example of this, Kasraei et al. (2024) show that a reduction from 70 to 7 variables in a model predicting pH did not impact quality, with both models reporting a concordance of 0.74. Similarly, Roudier et al., (2020) removed 16 of 29 covariates prior to mapping soil pH across New Zealand using a random forest (RF) model.

The use of SHAP values allowed for a clear interpretation of relationships between the predictor and target variables. For individual covariates the dependence plots allowed for the identification of trends at specific attribute values and between predictor variables. The examination of pH in relation to slope SHAP values showed multiple points where a high slope was positively contributing to pH. These sites would be associated with the Calcarosols sampled in the west of the region on easterly facing slopes at the interface of the Mallee Dunefields and the lower Murrumbidgee valley. When considering the global SHAP explanations, the ranking of predictor variables showed that mid-depth was the highest ranked predictor for pH, ECe and ESP, attributes where the measured data also varied with depth. Mid-depth was, however, only moderately ranked for sand content, clay content and CEC. This is expected as these attributes varied more spatially than vertically.

The variables included in the final models and their relationships to the target variable are consistent with previous studies, also. In this study all soil attribute models contained at least two satellite (barest earth) covariates which is consistent with discussions from Richer-de-Forges et al. (2023) on their effectiveness in DSM. A positive relationship between clay content and radiometric thorium and uranium was observed in this study, with radiometric thorium the highest ranked SHAP predictor with higher measurements positively contributing to clay. The inverse was demonstrated for sand content, where higher radiometric thorium values negatively contributed to sand content. This relationship has been previously described in literature (Mahmood et al., 2013; Maino et al., 2022; Van Der Klooster et al., 2011). Specifically, Maino et al. (2022), studying palaeochannel-influenced landscapes in Italy, showed negative correlations to sand content for radiometric potassium (-0.62) and thorium (-0.56). In their study, R^2 values of 0.52 and 0.49 were obtained from non-linear machine learning algorithms

predicting sand and clay content, respectively, using potassium and thorium abundance as input variables (Maino et al., 2022).

Issues with the incorporation of DEM as a predictor for regional scale DSMs are not new, as discussed by Adeniyi et al. (2024) where small, gradual gradients reduce the efficacy of this predictor. While it is noted that recently released finer resolution DEMs have improved mapping outcomes (Adeniyi et al., 2024), as seen for CEC, sand content and clay content in this study, this was not the case for pH, ESP and CEC. Generally, care needs to be taken when interpreting predictor relationships (Wadoux et al., 2019). Local knowledge is important in understanding these interactions, with the removal of DEM in this instance aligning with notions that DSMs may produce good validation statistics but, from a pedological standpoint, be spatially implausible (Bui et al., 2020; Kidd et al., 2020). The importance in careful covariate selection on a case-by-case basis is seen when viewing the CEC, sand and clay models where DEM was a moderate ranking variable and appropriately aided in representing variability spatially.

While climate is recognised as an important factor in soil formation (McBratney et al., 2003) it follows a similar trend to terrain and satellite attributes across this specific region. Further there is minimal variability in climate averages across the study area (BOM, 2023). It was on this basis that climatic predictor variables were not included.

6.5.2. Model quality

The models in this study, validated using a LOSOCV at the point support, produced moderate quality statistics with LCCCs ranging from 0.45-0.58. For each soil model there was negligible bias indicating that models were not over or under-fitting. The goodness of fit plots for sand content and ECe visually demonstrate issues predicting across a dataset of this size that has extreme variability.

Ma et al. (2021) identified that, in mapping EC in northwest China, XGBoost models outperformed RF and classification and regression tree (CART) models with an R^2 and RMSE of 0.59 and 11.99, respectively. In their study, the XGBoost model was developed from a combination of topographic and vegetation indices validated through a five-fold cross validation (Ma et al., 2021). In employing a 3D approach, as undertaken in this study, Roudier et al. (2020) used RF models to predict soil pH nationally across NZ. When tested on a

validation dataset the model produced an R^2 of 0.65, broadly comparable to that obtained in this study.

In modelling ESP across the Murray Darling Basin, Pozza et al. (2022) present an independent test LCCC of 0.66 for their RF model, however, when using a spatial cross validation, a poorer LCCC of 0.45 was obtained. This suggests a decrease in quality where datapoints are sparse, a known issue in the western reaches of the lower Murrumbidgee valley where there are few publicly available data points. This spatial cross validation statistic of Pozza et al. (2022) is comparable to the LCCC value, obtained through a LOSOCV, in this study. Employing a similar approach, Filippi et al., (2020) in mapping ESP in the Namoi catchment of northern NSW, obtained a similar model accuracy with an LCCC and RMSE of 0.55 and 6.70%, respectively, also comparable to the LCCC of 0.51 obtained for the model in this study.

Across a similarly variable dataset obtained from a plains region in the Jiangsu Province, China, Qu et al. (2024) produced RF and cubist models to predict sand content, achieving R^2 values of 0.56 and 0.49, respectively. Zhao et al. (2022) showcased that XGBoost models performed slightly poorer when modelling clay content across an area of 390 km^2 using proximally sensed data in northern NSW. In the subsoil (0.9-1.2 m) the XGBoost model achieved an LCCC of 0.54 compared to the cubist (LCCC=0.6) and RF (LCCC=0.58) models.

The model statistics for sand and clay content achieved in this study are comparable to those achieved in the state (NSW) and national (SLGA) DSMs where different depths were modelled individually (Gray, 2023; Grundy et al., 2015; Viscarra Rossel et al., 2014). Model statistics, however, tell only part of the story. As explained by Kidd et al., (2020) these statistical approaches do not consider the resulting spatial pattern and if it is pedologically plausible. Further, the choice of validation method can result in statistics that are not representative of how the data will be predicted over a broader area (Piikki et al., 2020).

6.5.3. Spatial distribution of soil properties

Assessing if produced DSMs are spatially appropriate from a pedological standpoint is essential in considering their efficacy. Despite the models in this study being of moderate quality, the spatial patterns of the resulting maps accurately identify variability in relation to known landscape characteristics. Given their intercorrelations, the CEC, sand content and clay content maps can be described together. Patterns for each of these maps directly relate to, and appropriately represent, the pathways of palaeochannel systems which were active over four

distinct periods during the last 100,000 years (Page and Nanson, 1996; Page et al., 2009). The relic streambeds, and occasional adjacent sand dunes, of these systems result in deep sands (>1 m) that are incised within heavy clays (>35 % clay) of different systems. Therefore, there is minimal variability in texture down the profile with the exclusion of the far east and west of the region. The former represents an area with known 'hillslope' (Butler, 1979) or texture contrast soils, with slightly acidic topsoils overlaying an alkaline subsoil. The latter occurs at the interface of the Murrumbidgee and Mallee regions where calcareous soils are prevalent (Bowler & Magee, 1978; Mitchell, 2002). Visually, the maps of these soil properties showcase greater fine scale variability than the presently available, broader scale, state (Gray 2023), national (Grundy et al., 2015; Viscarra Rossel et al., 2014) and global (Hengl et al., 2017a) DSMs.

The seams of predicted high clay and low sand content soils in the south-southeast of the region align with an older palaeochannel system (Page & Nanson, 1996). Consequently, it is likely that these soils are more pedologically developed. Schematics of the region presented by Butler (1958) and Langford-Smith (1960) note that along such systems there will be wider tracts of loamy soil that, despite having a high sand content, are not deep sands (>90% sand). These loamy soils are theorised to contain lime at depth, resulting in high alkalinity seen in the pH map.

Unlike sand content, clay content and CEC; soil pH, ECe and ESP are more prone to change as a result of land clearing and agricultural management (Cattle & Field, 2013). The clearest example of this is seen for ESP, where there is an abrupt increase at approximately 145.75ºC. This increase is also seen in the map of Pozza et al. (2022) and occurs at the western boundary of the Murrumbidgee and Coleambally Irrigation Areas, which have been intensely farmed under irrigation for 90 and 60 years, respectively. To the west, irrigation occurs only in pockets with land primarily used for the grazing of native vegetation. This is interesting, as it has previously been reported that irrigated agriculture can increase sodicity as well as salinity (Rengasamy & Olsson, 1993; Cattle & Field, 2013). It is possible that, in seeking to optimise agricultural production, management decisions such as deep ripping combined with chemical amelioration have been used to reduce sodicity (Jayawardane and Chan, 1994). Irrespective of management, a gradual westerly trend of increasing salinity (ECe) and sodicity (ESP) was also identified by Talsma (1968). Prior to the completion of the Coleambally Irrigation Area, Talsma (1968) found ESP in this subregion was generally <7%, however, increased to 7-22% in the topsoil and 15-30% in the subsoil further west at approximately 144.75º (Talsma, 1968). Sodic

soils are predominantly alkaline (Dagar et al., 2022) and this is supported by similar spatial trends of the pH and ESP maps for the region.

The decision to remove DEM from the pH, ECe and ESP models has been discussed. It is believed that the models without DEM are able to more accurately represent spatial trends across the landscape, not just at the points that the model was built on, a key objective of this research and DSMs more broadly.

6.5.4. Area of applicability

The Area of Applicability assessment assists with identifying where model predictions may be unreliable and should be interpreted with caution. This results from a lack of data meaning the model needs to extrapolate further, reducing reliability. In this instance, these 'unreliable' areas are closely associated with the Murrumbidgee River, palaeochannel systems and the dunefields soils in the far west. Each of these areas is associated with sandier soil types. From the perspective of covariate relationships this is also not surprising as there is significantly different vegetation near the Murrumbidgee River compared to across the plain more broadly. Given an element of the sampling plan employed in this study that was purposively targeted towards cotton soils, which are generally more clayey, this is not surprising. This map shows that in the future sampling should occur in these areas.

6.5.5. Point validation of external DSMs

It has been presented that the models in this study appear to identify greater fine scale variability than presently available DSM products. Consequently, this element of the study sought to assess three DSMs at a point support using independent datasets containing 153 (0.3-0.6 m) and 150 (0.6-1 m) observations. Each of the assessed models performed more poorly than their reported model statistics for clay content. At 0.3-0.6 m depth the NSW DSM produced a point scale external validation LCCC of 0.50 and RMSE of 10.4%, marginally lower than the reported statistics of 0.53 and 14.5v%, respectively (Gray, 2023). At the 0.6-1 m an external validation LCCC of 0.36 was obtained, worse than the reported LCCC of 0.43 (Gray, 2023).

The SLGA model reports LCCC statistics of 0.56 and 0.53, significantly higher than the external validation statistics obtained in this study of 0.16 and 0.15 for depths of 0.3-0.6 and 0.6-1 m, respectively. At these depths the SLGA presents a validation RMSE of 14.3% and

14.2% for the two depth increments, both slightly higher than the corresponding values of 12.1% and 12.1%, obtained in this external validation.

There are issues in comparing the poor and very poor external validation LCCC values obtained for the SoilsGrid250m model here with the reported statistics. Poggio et al. (2021) reported model efficiency concordances of 0.41 and 0.40 for depths of 0.3-0.6 and 0.6-1 m, respectively. This cannot be directly compared to the LCCC obtained in this validation, or the R^2 of 0.73 reported in the previous iteration of Hengl et al. (2017a), as these metrics measure different aspects of model performance. Further, it is not clear which units are associated with the RMSE of 0.13, however, it is stated that the model statistics are slightly lower than those obtained by Hengl et al. (2017a) where an RMSE of 19.3% was presented (Poggio et al., 2021). It is possible that this statistic is related to the additive log ratio transformation that was performed on the data prior to modelling (Poggio et al., 2021). Irrespective, it is demonstrated that at a point support when validated against external datasets in the lower Murrumbidgee valley, the SoilsGrid250m DSM performs poorly.

The horizontal scatter observed in the goodness of fit plots show an inability for the models, especially the SLGA and SoilsGrids250m, to predict higher sand contents at depth. In this region there are numerous sand seams, tracts of land extending across the landscape with a sand content of >90% to depth. Given the fine scale at which these occur it is possible that they are not being identified due to the broader resolutions of the SoilsGrid250m product or that the NSW and SLGA models are not able to detect this fine variability.

The trend of external validation statistics worsening from the state to national to global DSMs suggest that, not surprisingly smaller, specific models may be better at understanding variability at finer spatial supports. A potential reason for this is the model sensitivity to change. Using DEM as an example, in this study models were developed for a region where elevation only ranged from 53-227 m. A national product for Australia, however, must be sensitive to elevation changes from -9 m, in the Lake Eyre basin, to 2228 m. Because of this, it may not be able to identify fine scale variability. In the lower Murrumbidgee valley this is important, as across the plain a change in elevation of 2-3 m can represent a change from a heavy clay to a gradational loam on a source bordering dune. Similarly, radiometric covariates may be able to differentiate between changes in sand and clay content in the streambed sands, allowing for their identification. Ultimately, these external validation statistics show the value in regionally specific DSMs, a conclusion also reached by Lemercier et al. (2022) who went a step further,

suggesting investment in local DSMs due to the underperformance of regional, national and global products.

Han et al. (2022) undertook a similar study with a focus on farms, examining both soil organic carbon and clay content at a depth of 0.3-0.6 m at the farm, strata (field) as well as point spatial supports. At 14 sites, 12 of which were in NSW, the NSW, SLGA and SoilsGrids250m DSMs exhibited significantly worse external validation statistics at the point support (Han et al., 2022). When aggregated to the larger strata (field management) and farm spatial supports the validation statistics (LCCC, R^2 and RMSE) generally improved (Han et al., 2022). This is similar to findings of Bishop et al. (2015) who identify that the prediction quality is best at a soil-land use complex, or 'block' support. As such, the statistics presented in this study are likely a 'worst case' scenario and considerations must be given to the optimal support for DSM products that balances prediction accuracy with the identification of landscape variability. Piedallu et al. (2022), however, in developing DSMs for pH and C/N in north-east France identify that model performance decreases when the prediction resolution increases beyond 1000 m^2.

6.6. Conclusion

This study has shown the potential of regionally developed DSMs to predict the spatial variability of relevant soil properties at five depth increments concurrently. The importance of understanding the pedological plausibility of the resulting maps was also demonstrated, with an understanding of landscape characteristics important in covariate selection and output interpretation. XGBoost models were built to predict pH, ECe, ESP, CEC, clay and sand content at depths of 0-0.1, 0.1-0.3, 0.3-0.3, 0.6-0.8 and 0.8-1 m. Model quality was moderate for all attributes, with the LCCC obtained from a LOSOCV ranging from 0.45 (ECe) to 0.58 (clay). Pleasingly, the resulting maps produced were able to identify landscape trends associated with the past geomorphological processes of the lower Murrumbidgee valley. The plains were appropriately predicted as having a high clay content throughout the first metre of the soil profile, with veins of sandier soils predicted along pathways of palaeochannels. Predicted alkaline soils in the west of the region relate to dunefield Calcarosols while in the east slight surface acidity is present. Both soil ECe and ESP increase along the east-west axis, with strong sodicity and extreme salinity predicted to occur frequently within the first metre of the soil profile. A visual assessment, incorporated alongside the model statistics, is an important step in evaluating the efficacy of maps in displaying realistic spatial trends with SHAP values shown to be effective in identifying the relationship between covariates and the modelled attribute at a local and global scale. It was determined that, when validated against data collected in this study, the presently available state, national and global DSMs for clay content underperformed compared to their reported model statistics. At depths of 0.3-0.6 and 0.6-0.1 m, the state model produced external point support validation LCCCs of 0.5 and 0.36, respectively, compared to the corresponding values of 0.11 and 0.07 for the global product and the LCCC of 0.58 obtained for the clay content model developed in this study. This shows the capacity of more specific DSMs to outperform broader products within regions at the point scale.

Chapter 7

7. Applying a digital soil assessment to classify the suitability of lower Murrumbidgee valley soils for irrigated cotton production

Abstract

Digital soil assessments (DSAs) have the capacity to leverage soil spatial data to answer specific agricultural or environmental questions. This study applied multiple DSA approaches to determine the suitability of lower Murrumbidgee valley soils for irrigated cotton production. Depth to constraint maps, a secondary application of digital soil mapping (DSM) products, were produced for soil alkalinity (pH) and sodicity (exchangeable sodium percentage - ESP). These maps were assessed to determine the most limiting soil chemical constraint and the depth where this is reached in the soil profile. Pedotransfer functions were used to make predictions of plant available water capacity (PAWC) using DSMs of cation exchange capacity (CEC), sand content and clay content. After incorporating a digital map of landscape slope, a regionally specific suitability ruleset was created with four classes: well suited, suited, marginally suited and unsuited. The most limiting factor (soil chemical constraints, slope or PAWC) was identified, and the overall suitability for irrigated cotton production was classed with 90 m resolution maps produced. Results showed that sodicity was the most limiting soil constraint, with soil chemical suitability the most limiting factor by area, negatively impacting suitability for irrigated cotton production across 52.9% of the region. The second most limiting factor by area was PAWC, reducing the suitability class across 24.9% of the lower Murrumbidgee valley. Overall, 16.7% of the area was classed as well suited to irrigated cotton production and 18.4% was classed as unsuited, with the remainder comprised of suited (31%) or marginally suited (34%) land. Further work should investigate the development of continuous, rather than categorical, classification metrics as these may better represent gradual landscape trends and agronomic management considerations rather than relying on hard cutoffs.

7.1. Introduction

There is the potential to enhance the value of digital soil map (DSM) products by using them to address specific agricultural and environmental challenges. Such applications can include inferring soil attributes, such as plant available water capacity (PAWC), identifying the depth to particular soil constraints (Pozza et al., 2022) and classifying land or enterprise suitability, all of which can be considered digital soil assessments (DSAs) (Carre et al., 2007).

DSAs were defined by McBratney et al. (2012) as processes which "translate DSM into decision-making aids that are framed by the particular, contextual, human-value system which addresses the question/s at hand." In the context of their framework, this study can be considered as a stakeholder-led assessment (McBratney et al., 2012). Carre et al. (2007) were the first to present a framework for DSAs where, following rapid advancements in the field, DSMs were not to be considered the end point of studies, instead, they are a precursor to DSAs (Carre et al., 2007). DSAs can be more applicable to specific end users, capable of accounting for social and economic factors, better guiding decisions from land management to policy (Carre et al., 2007; Kidd et al., 2015). Since this framework was formalised (Carre et al., 2007) there has been a significant increase in publications relating to the topic globally. Searle et al. (2021) provide a detailed discussion on this. They focus on how DSM and DSAs have progressed to a point where they can be implemented on a broad scale to meet global, national, and local challenges including ecosystem enhancement, climate change mitigation and the safeguarding of food production (Searle et al., 2021).

Within Australia, recent DSA applications to classify land for agriculture include the work of Kidd et al. (2015) in Tasmania and Thomas et al. (2015, 2018) with Harms et al. (2015) in northern Australia. These projects implemented DSAs to identify new land for development (Harms et al., 2015; Thomas et al., 2015, 2018) or to understand crop suitability in newly commissioned irrigation areas (Kidd et al., 2015). This study, however, aims to examine the suitability of irrigated cotton production in a region where it already had demonstrated success. This allows for the targeted creation of suitability rulesets as was undertaken by Kidd et al. (2015), who consulted industry experts to assist in ruleset development.

Both depth-to-constraint and PAWC maps can be considered as forms of DSAs as they use DSMs to produce decision making aids. Depth-to-constraint maps, as opposed to maps showing soil attributes at different depth intervals, have the potential to provide more interpretable resources. Recently, there have been multiple studies mapping the depth to soil

constraints, and their impact, across the Murray Darling Basin. In the Namoi catchment, Filippi et al. (2020) mapped the depth to sodicity constraints with ESP thresholds of 6%, 10% and 15% at a 30 m pixel size with a 1 cm vertical support. At the within field scale, within the Namoi catchment also, Filippi et al. (2019) mapped the depth to which an alkalinity constraint (pH>9) is reached. Across the entirety of the Murray Darling Basin, Pozza et al. (2022) mapped the depth to sodicity thresholds of 6%, 10% and 15% with the probability of threshold exceedance at specific depths and the area of applicability shown. By combining depth to constraint maps with PAWC, Tilse et al. (2022) quantified the constrained-PAWC, and potential yield loss, resulting from chemical constraints. Identification of the most limiting constraint, and the depth at which it occurs, allows for informed decisions on the economics of amelioration or management.

This study will showcase the ability for regionally developed DSMs, produced in Chapter 6, to be used for DSAs. Firstly, a regionally specific ruleset for soil and landscape factors impacting irrigated cotton suitability will be developed. Secondly, DSMs of soil pH and ESP will be used to generate spatial resources showing the depth to alkalinity and sodicity constraints across the region. Thirdly, DSMs of CEC, sand and clay content will be used within pedotransfer functions (PTFs) to predict PAWC to 1 m. Lastly, the soil or landscape parameter most limiting to irrigated cotton production will be identified at each pixel, and, finally, the suitability of the region for irrigated cotton production will be classified.

7.2. Materials and methods

7.2.1. Study area

The lower Murrumbidgee valley, in southern NSW, extends from Narrandera in the east to Balranald in the west. The region is known for its agricultural diversity and includes the Murrumbidgee and Coleambally Irrigation Areas. Soils of the region are primarily alluvial, having developed from Cenozoic era sediments deposited by palaeochannel systems (Brown and Stephenson, 1991; Page et al., 1996). There is considerable variability, ranging from heavy clay soils, considered Vertosols within the Australian Soil Classification (ASC), to deep sandy soils, considered Arenosols (Isbell & NCST, 2021). While the diversity of soils presents challenges, it allows for a diverse agricultural sector. Included within this is the irrigated production of perennial and annual crops, including cotton, which utilise water from the Murrumbidgee River or groundwater aquifers. Decisions surrounding the viability of

establishing perennial plantations or the sowing of annual crops, such as cotton, are strongly influenced by annual water allocations. Major population centres of the region include Griffith, Leeton, Hay, Coleambally, and Balranald.

7.3. Irrigated cotton suitability ruleset

The rulesets for irrigated cotton suitability were developed to be applied to the DSMs created in Chapter 6. This study has adopted the four-class system, as per the work of Kidd et al. (2015), whereby soils are classed as:

- **Well suited** – no soil-landscape based limitations to productivity,
- **Suited** – minor limitations to productivity,
- **Marginally suited** – moderate limitations to productivity, and,
- **Unsuited** – severe productivity limitations.

Previous studies, including those of Kidd et al. (2015) and Harms et al. (2015), incorporate climate parameters in their enterprise suitability rulesets. It was decided that this study would focus entirely on soil-landscape attributes. This was on the basis that the climate is known to be suitable for cotton production. There is minimal variability in the summer mean maximum (32.2-32.5ºC) and minimum (16-17.9 ºC) temperatures across the region that fall within the ranges of 27-32ºC (daytime maximum) and 16-20ºC (overnight minimum), the temperatures considered ideal for cotton production (BOM, 2023; CRDC and CottonInfo, 2023). The prolific expansion of the industry also suggests that climactic uncertainty is not a key driver of decisions. As this study is only considering the suitability for irrigated cotton production, rainfall is irrelevant with the upper catchment, located beyond the bounds of the study area, being the source of this water. Similarly, other attributes such as the soil depth and rockiness were not considered as these are not relevant to this specific region. The developed suitability rulesets (Table 7.1) were applied to produce a rating for each attribute. A most limiting factor approach was employed where the lowest suitability class of any parameter at each pixel was labelled as the overall suitability (Kidd et al., 2015; Klingebiel and Montgomery, 1961). All analysis related to the creation and application of rulesets occurred using ‘R’ (R Core Team, 2023) and maps were formatted using ‘QGIS’ (QGIS.org, 2023).

Table 7.1. The suitability ruleset for irrigated cotton production in the lower Murrumbidgee valley.

Suitability class	Soil chemical suitability (Depth to soil constraint (m))	Slope (%)	PAWC in the first metre of the soil profile (mm)
Well suited	>0.8	<0.5	>125
Suited	0.6-0.8	0.5-1	100-125
Marginally suited	0.3-0.6	1-2	75-100
Unsuited	0-0.3	>2	<75

7.3.1. Depth to soil constraints and soil chemical suitability parameters

Rather than applying a ruleset for pH and ESP it was decided to identify the depth to any soil chemical constraint and utilise this as a parameter. While the impact of soil constraints on plant growth variers, irrespective of the constraint, the shallowest depth at which it occurs influences plant growth by limiting the effective rooting depth (Dang et al., 2006). DSMs for the two soil attributes were modelled in Chapter 6 of this thesis, with mapped depths of 0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m for each attribute. A function was developed to identify the upper most depth (0, 0.1, 0.3, 0.6 or 0.8 m) where a cotton plant growth threshold was exceeded. Thresholds for pH and ESP were chosen primarily on the soil chemical interpretations of Hazleton and Murphy (2007), alongside discussions with local agronomists.

In depth to soil constraint studies the upper threshold for sodicity (ESP) is often set to 15% (Filippi et al., 2020; Pozza et al., 2022). Chapter 5, however, has demonstrated that cotton can be grown successfully on soils exceeding this level. Further, discussions with local stakeholders revealed that irrigated agriculture on these sodic soils is common practice as is seen when consulting land use maps (Figure 7.1). As a result, the threshold for ESP was set to 20% as a balance between previous studies (Filippi et al., 2020; Pozza et al., 2022), published considerations of sodicity (Dodd et al., 2010, 2013b) and local knowledge. The threshold for pH was set to 9 in accordance with published depth to constraint studies (Filippi et al., 2019; Hazleton and Murphy, 2007; Tilse et al., 2022), as this is the point where nutrient deficiencies can be induced (Dang et al., 2006; Gupta & Abrol, 1990).

Salinity was not considered in the in the constraint analysis as it has been previously demonstrated that irrigated cotton production can reduce soil salinity (Filippi et al., 2018a). This trend was observed in the ECe maps, with lower predicted salinity in areas that have been developed for irrigation. Consequently, it was believed that salinity is likely to reduce under

irrigated cotton production and therefore it would not have been appropriate to classify areas as unsuitable on account of this. The inclusion of ECe would have also classified extensive areas in the west of the region, where cotton is known to be successfully grown (Figure 7.1), as unsuitable. While this was the case, to a lesser extent, for ESP the 15% threshold was able to be increased on account of findings in literature (Dodd et al., 2013b). Given maps for both ESP and ECe predict similar spatial trends, it was decided to incorporate ESP only.

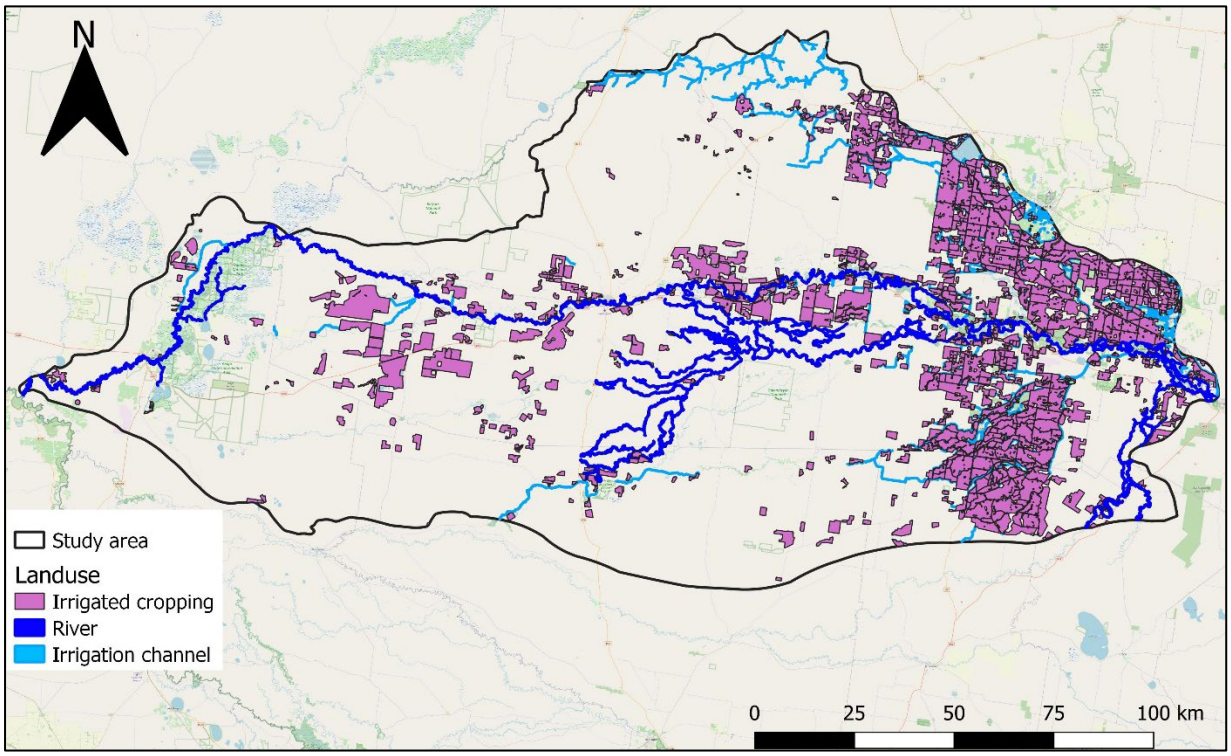


Figure 7.1. The area within the lower Murrumbidgee valley used for broadacre irrigated cropping. Also shown is the pathway of the Murrumbidgee River, distributary streams and irrigation channels (NSW DCCEEW, 2023). Much of the irrigated cropping land, especially in the central and west of the region, includes cotton as a central part of the crop rotation.

Maps of the depth to constraint, where each pixel was assigned a depth value, were then produced for each soil attribute and visually assessed. If these attribute thresholds were not predicted to be reached at any point in the first metre of the profile a value of 1 (metre) was assigned. The rasters for the depth to pH and ESP constraints were stacked to identify the most limiting constraint, that being the constraint predicted to occur closest to the surface in each pixel. A function was then employed on the raster stack to identify the depth to any constraint. Soil chemical suitability classes of well suited (>0.8 m), suited (0.6-0.8 m), marginally suited (0.3-0.6 m) and unsuited (<0.3 m) were assigned based on this depth (Table 7.1).

7.3.1.1. Slope suitability parameters

A raster of slope (%) (Gallant and Austin, 2012a) at an original resolution of 30 × 30 m was extracted on a 90 m grid for the region. A low slope is required to minimise the soil that must be moved through land-reformation to achieve a target gradient of approximately 0.3% for irrigated cotton production (CRDC and CottonInfo, 2023). The suitability classes for this parameter (Table 7.1) were decided upon by consulting previous studies of Thomas et al. (2018) in northern Australia and Walke et al. (2012) in India. After discussions with stakeholders on what slope would be uneconomical to reform, final classes of well suited (<0.5%), suited (0.5-1%), marginally suited (1-2%) and unsuited (>2%) were decided upon (Table 7.1).

7.3.1.2. Plant available water capacity (PAWC) suitability parameter

Soil PAWC to a depth of 1 m was calculated using pedotransfer functions (PTFs) developed by Padarian Campusano (2014). These PTFs were applied to each depth increment (0-0.1, 0.1-0.3, 0.3-0.6, 0.6-0.8 and 0.8-1 m) incorporating DSMs for cation exchange capacity (CEC), sand content and clay content developed in Chapter 6. The utilisation of PTFs to predict PAWC has been previously demonstrated at farm (Tilse et al., 2022), regional (Malone et al., 2009) and national (Hong et al., 2013) extents.

For each depth increment, the drained upper limit (DUL) ($cm^3/100cm^3$) was estimated using clay content, sand content and cation exchange capacity (CEC) (*Equation 1*). This PTF reported an R^2 and RMSE of 0.75 and 4.53%, respectively (Padarian Campusano, 2014). The crop lower limit (CLL) ($cm^3/100cm^3$) for each depth increment was then estimated using a PTF incorporating clay and sand content (*Equation 2*). This PTF reported an R^2 of 0.65 and RMSE of 4.37% (Padarian Campusano, 2014). The PAWC was determined for each depth increment by subtracting the CLL from the DUL. To calculate the PAWC for each horizon, this value was multiplied by the horizon depth (mm) before being multiplied by 10 (conversion to mm). The total PAWC (mm) to a depth of 1 m was determined by summing each of the five horizons.

Equation 1. Pedotransfer function for drained upper limit (DUL) (Padarian Campusano, 2014):

$$DUL(cm^3/100cm^3) = 0.2358 + 0.002572CEC + 0.001001clay - 1.70 \times 10^{-7}sand^3$$

Equation 2. Pedotransfer function for crop lower limit (CLL) (Padarian Campusano, 2014):

$$CLL(cm^3/100cm^3) = 0.1476 + 9.002 \times 10^{-5}clay^2 - 0.00115sand - 9.752 \times 10^{-7}clay^3$$

The suitability parameters for PAWC were based on the studies in northern Australia (Thomas et al., 2018), Queensland (DSITI & DNRM, 2015) and India (Walke et al., (2012). There is, however, no consistently defined boundary between PAWC suitability classes in literature

despite its importance (CRDC & CottonInfo, 2023). For example, SOILpak (McKenzie, 1998), the 'best practice soil management manual', states specific target thresholds for attributes such as pH, however, for water storage and PAWC, the target is only stated as 'maximise'. As a higher water holding capacity is crucial in the selection and efficacy of irrigation systems, and the variability in key determinants of water holding capacity in this region has been demonstrated, the parameters from previous studies were tightened. Rather than a minimum PAWC of 50 mm (DSITI & DNRM, 2015; Thomas et al., 2018) being considered the threshold for a soil to be unsuited, the final level was set to 75 mm (Table 7.1). The final ruleset thresholds were: well suited (PAWC >125 mm), suited (PAWC 100-125 mm), marginally suited (PAWC 75-100 mm) and unsuited (<75 mm) (Table 7.1).

7.3.2. Cotton suitability classification and the most limiting constraint

Following their individual suitability classification, the three rasters for soil chemical suitability (determined through depth to any constraint), slope and PAWC suitability were stacked. A function was developed to identify the most limiting factor at each pixel. In situations where the three attributes were equally limiting, PAWC was allocated the most limiting factor. This was on the basis that PAWC is the result of inherent soil properties not easily altered by management practices. Soil chemical constraints, however, can potentially be ameliorated or managed while slope can be altered through land reformation. To determine the overall suitability a function was employed to identify the least suitable class among the three parameters at each pixel, with this being set as the overall suitability. To improve interpretability, maps of the most limiting factor and overall suitability were smoothed using a majority filter over a 5 × 5 moving window.

7.4. Results and discussion

7.4.1. Depth to soil constraints and soil chemical suitability

There are only isolated pockets of land where soil pH is alkaline to the point of exceeding plant limiting thresholds (pH > 9) within the first metre of the soil profile (Figure 7.2). Consequently, most of the region's soils can be considered unconstrained by pH (Figure 7.2). Points where pH is likely to impact cotton production are limited to calcic soils in the west, palaeochannel-adjacent soils in the south and hillslope soils (Butler, 1979) in the east (Figure 7.2).

While Australian soils are formally considered 'sodic' when ESP exceeds 6% (Isbell & NCST, 2021), this study mapped a soil ESP threshold of 20% on the basis that cotton is known to be grown on these soils (Figure 7.1). The ESP map suggests that soils in the east are not constrained within the first metre of the profile (Figure 7.2). As ESP increases along the east-west axis, the depth to a constraint being reached decreases with increasing distance to the west. Consequently, the depth to any constraint is more closely tied to ESP than pH (Figure 7.2). At points in the centre and west of the region constraints are frequently encountered within the first 0.3 m of the soil profile (Figure 7.2). This is consistent with initial observations of the region's soils by Talsma (1968) who noted ESP to be lower in the eastern areas, increasing along a westerly trend. The spatial pattern is also consistent with recent work of Pozza et al. (2022).

In a study within the Lachlan Valley, Bennett et al. (2014) identified that to reduce ESP below 6% from 0-0.4 m at a specific site, where the pre-treatment ESP ranged from 11-38%, would be economically unviable. Consequently, when soils are extremely sodic and have exceeded a point where amelioration is economically viable a focus should be placed on management (Ghosh et al., 2010; Hulugalle et al., 2010). The Murrumbidgee valley serves as an example of this, where sodic soils are farmed effectively (Figure 7.1). One reason for this could be the mechanisms through which sodicity influences yield. Dodd et al. (2013) showed that, up to an ESP of 19, sodicity primarily impacts cotton yield through the soil's structural collapse. Beyond this point, however, it becomes chemically constraining through toxicity (Dodd et al. 2013b). Dodd (2010) also found that cotton plant growth is largely unaffected by ESP values of up to 15%. Targeted irrigations can also illicit a positive response from certain crops in extremely sodic soils (Sharma et al., 1990). Considering the above, a focus on management to prevent further sodification and targeted irrigations, as outlined by Dagar et al. (2022), should occur.

In regard to cotton production, the majority of the region's east is unconstrained by pH or ESP (Figure 7.3). Overall, 37.7% of the area is considered unconstrained in the first metre of the profile (Figure 7.3), with these areas considered as well suited to cotton production and often following pathways of palaeochannel deposits (Figure 7.4). Centrally within the region the palaeochannel influence is clear, with veins of well suited land with suitability decreasing as the perpendicular distance increases (Figure 7.4) For 17.6% of the region, primarily in the far west and east (Figure 7.3), pH is the most limiting soil chemical constraint. In these instances, soils are most likely to be classed as marginally suited to cotton production on account of pH (Figure 7.4). Overall, ESP is the most limiting soil chemical constraint for 44.7% of the region's

soils, particular in the area's centre (Figure 7.3). As the depth to constraint varies, the overall suitability of these soils, on account of soil chemical constraints varies also (Figure 7.4). In areas where ESP is most limiting, soils are commonly classed as marginally suited or unsuited to cotton production on account of this (Figure 7.3, Figure 7.4)

Overall, when applying the soil chemical suitability ruleset to the depth to any soil constraint data, 40.9% of the lower Murrumbidgee is considered well suited while 9.5% is considered suited (Figure 7.4). In total, 17.5% of the region is considered unsuited to cotton production while a further 32.2% are marginally suited based on soil chemical suitability (Figure 7.4).

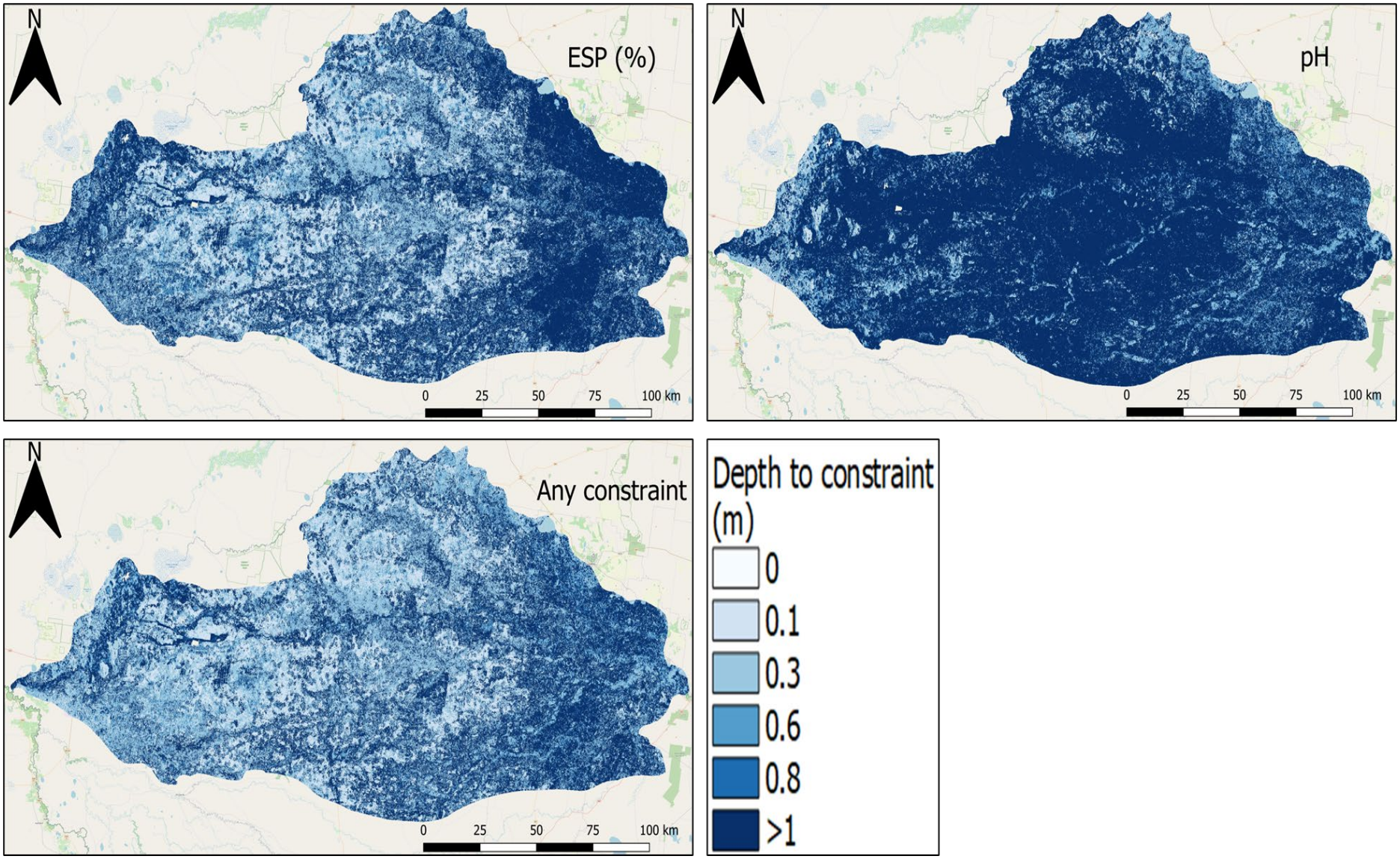


Figure 7.2. The depth where a plant limiting constraint is reached for ESP (20%) and pH (9). Also shown is the shallowest depth to any soil chemical constraint.

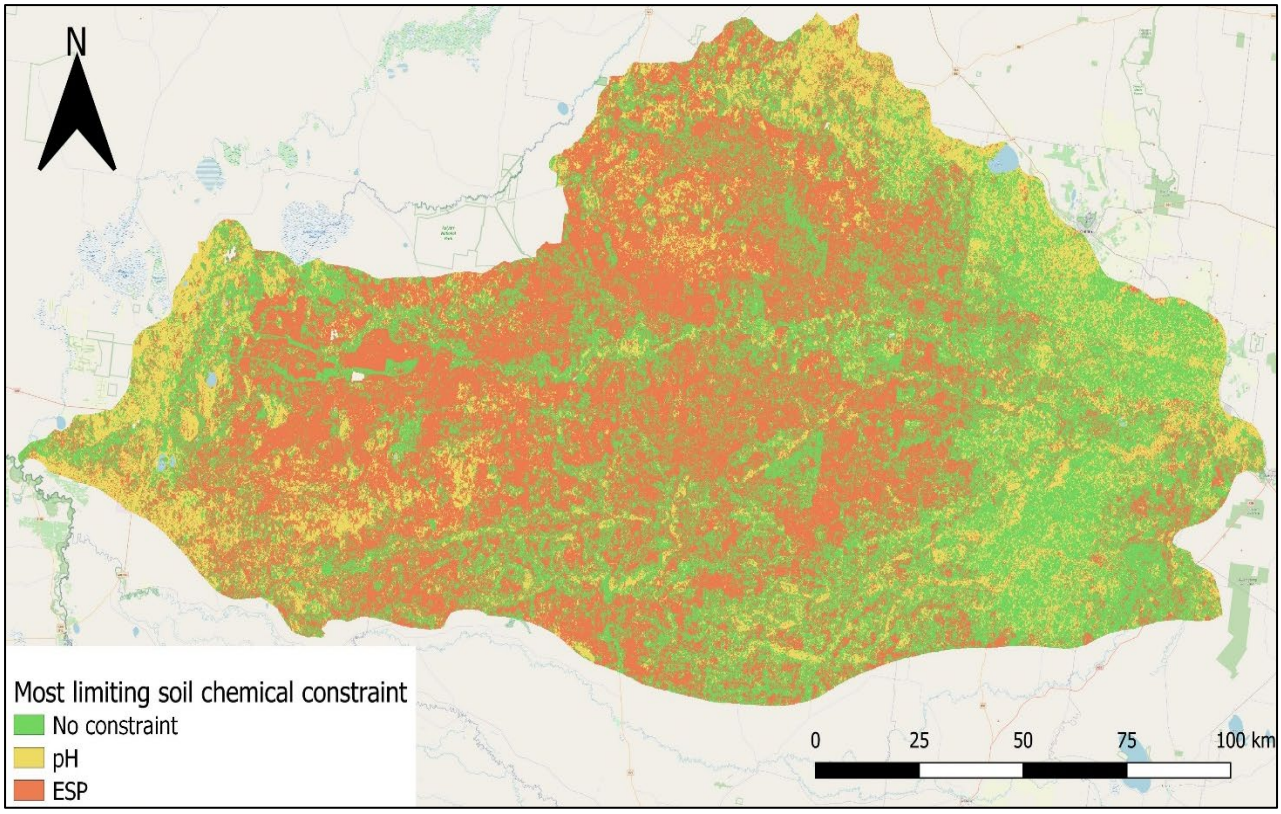


Figure 7.3. The most limiting soil chemical constraint determined in this research for the study area. Where no limits were exceeded for pH or ESP within the first metre of the soil 'no constraint' was assigned.

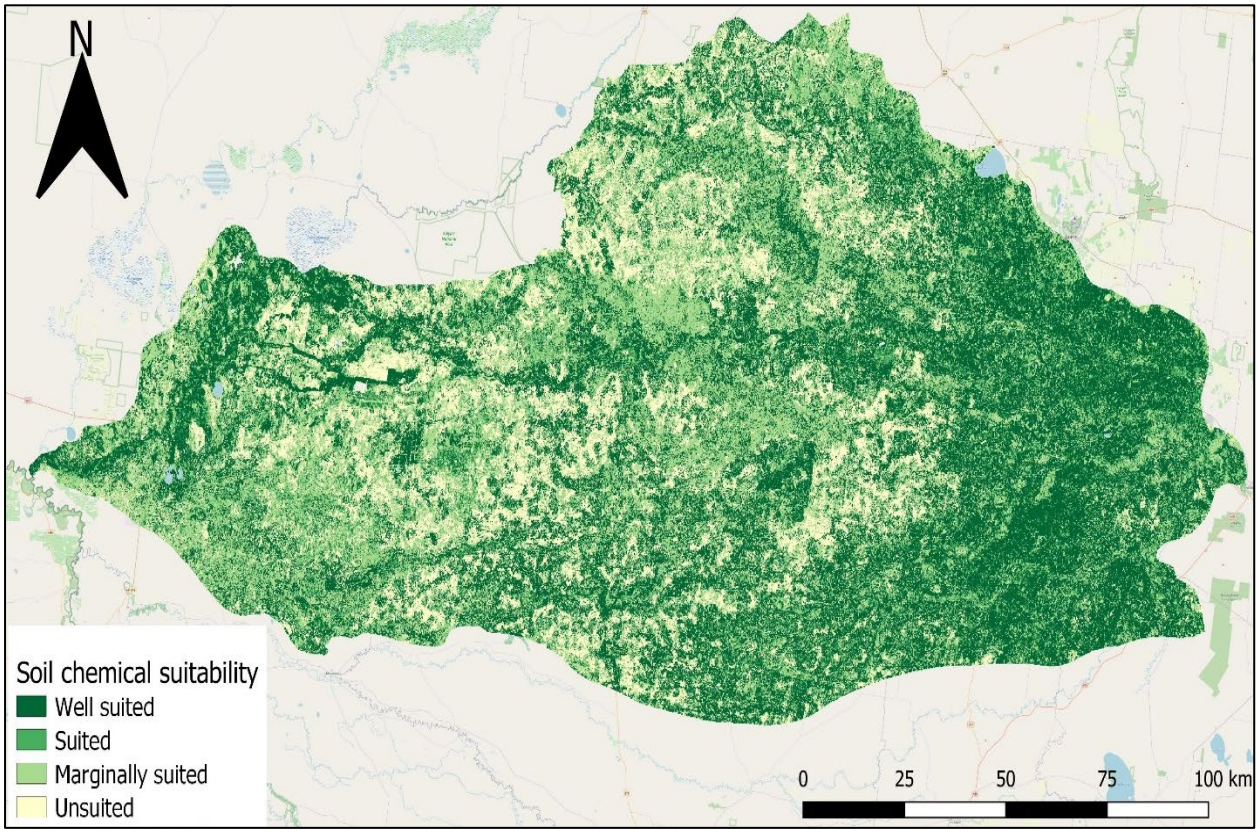


Figure 7.4. The chemical suitability of lower Murrumbidgee valley soils for irrigated cotton production determined by identifying the depth where pH or ESP will constrain plant growth.

7.4.2. Slope suitability

Across most of the region slope is not a factor limiting cotton production (Figure 7.5). In total, 94.1% of the region's slopes are at a level considered well suited (<0.5%) or suited (0.5-1%) to cotton production. This is expected given the region falls on the Riverine Plain, an alluvial landscape with only a gradual east-west slope. In the south, pockets of marginally suited (slope 1-2%) or unsuited (slope >2%) land are linked to palaeochannel systems and their associated dunes (Figure 7.5). In the west of the region greater slopes are linked to dune systems from the Mallee landscape, with this landscape described as the 'Mallee Cliffs Sandplains' (Mitchell, 2002). In the far east, marginally suited or unsuited land is linked to hillslope soils (Figure 7.5).

The consideration of slope as a limiting factor is based upon the amount of soil required to be moved to achieve a target gradient of approximately 0.33% (CRDC and CottonInfo, 2021). Abdullaev et al. (2007) identify, over a three-year study in Tajikistan, that appropriately levelled fields can significantly decrease water use, reducing runoff by 24% and increasing gross margins by 92%, on average, over the three-year period of their study. The initial slope, however, should require minimal earth movement to minimise costs and soil impacts (Miao et al., 2021). Cay and Cattle (2005) identify that land reformation on cotton soils in the neighbouring Lachlan Valley significantly impacts pedologically differentiated soil types, with the scalping, or mixing, of the topsoil potentially exposing constraining subsoils. This, in turn, will impact the suitability, decreasing the depth to which a soil chemical constraint is reached.

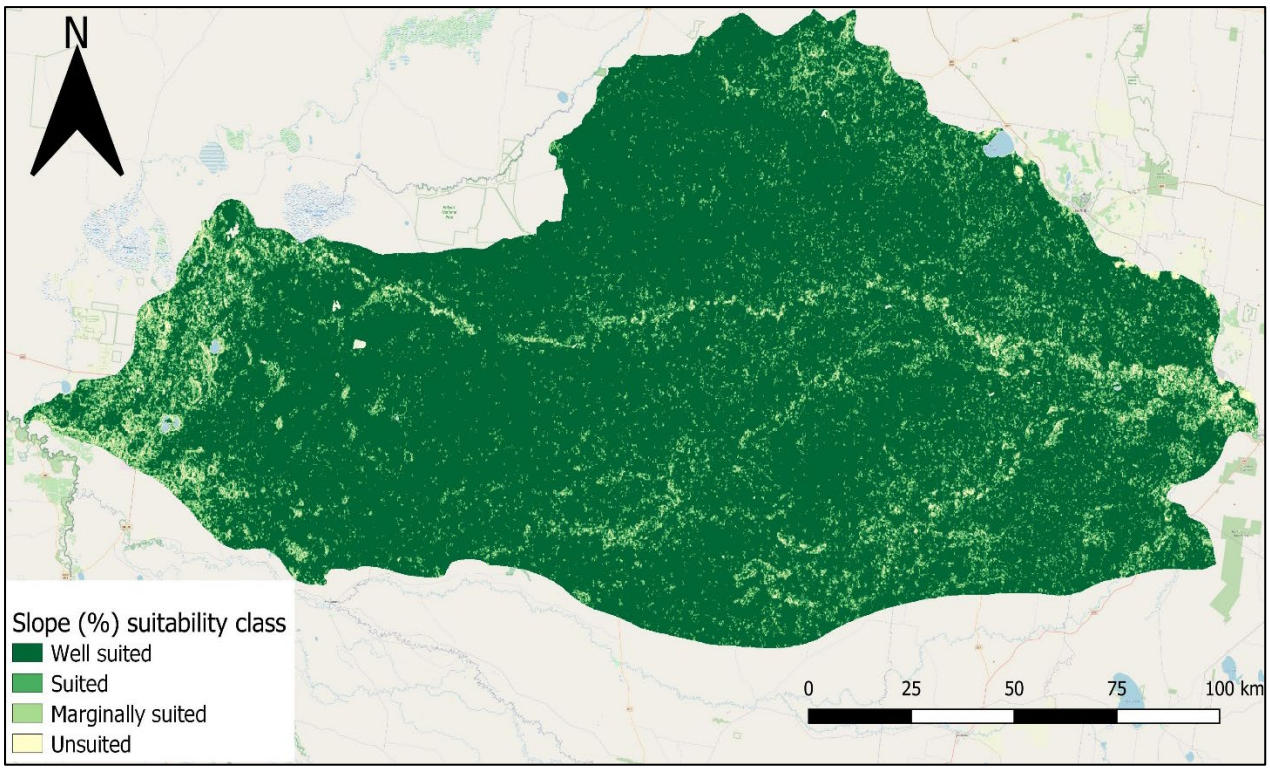


Figure 7.5 Classes of suitability for landscape slope (%), for irrigated cotton production, across the lower Murrumbidgee valley.

7.4.3. Plant available water capacity (PAWC) suitability

The predicted PAWC across the region was consistent with observations of soil classes. The prominence of clay soil types resulted in little variability in PAWC across the central and western regions (Figure 7.6). Lower PAWC predictions are associated with the pathways of paleochannel systems where streambed or dune sands are likely to be present (Figure 7.6). The heavier clay soils were generally classified as well suited to irrigated cotton production on account of PAWC (Figure 7.7). Overall, these soils account for 51.9% of the region while a further 47.2% of the area's soils are classed as having a PAWC suited to cotton production. This is expected given the deep, alluvial soils that account for much of the region. Where the veins of sandy soils are present, either in palaeochannel streambeds or on associated sand dunes, PAWC may be reduced to levels considered marginally suited or unsuited to irrigated cotton production (Figure 7.7). In total 0.8% and 0.1% of the area is considered marginally suited or unsuited due to PAWC, respectively. Low PAWC soils are agronomically concerning due to difficulties in managing irrigation systems when this variability appears at the within-field scale, an occurrence already demonstrated in this thesis (Chapter 5). Given the narrow form of streambed sands in the west of the region, it is possible that low PAWC soils may be

under-represented. Given this, at a finer spatial support PAWC may become a more limiting factor.

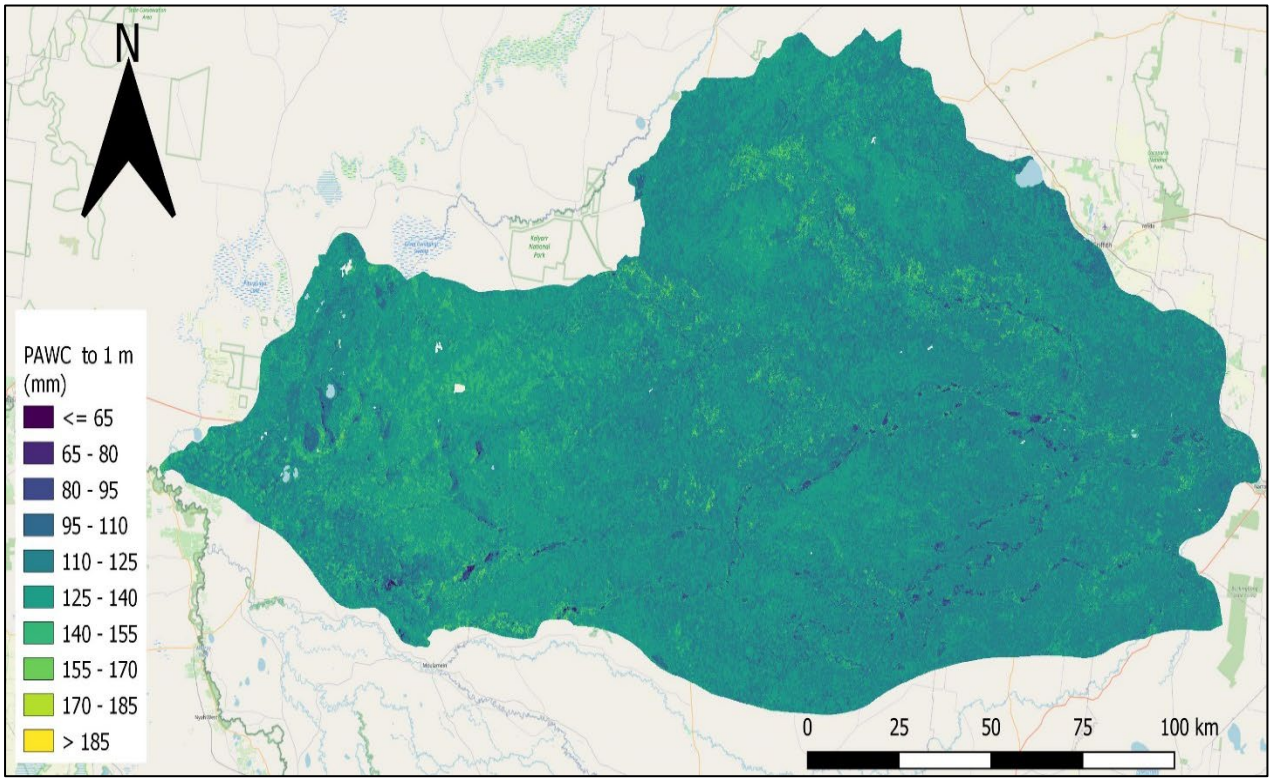


Figure 7.6. Predicted PAWC (mm) to a depth of 1 m across the lower Murrumbidgee valley.

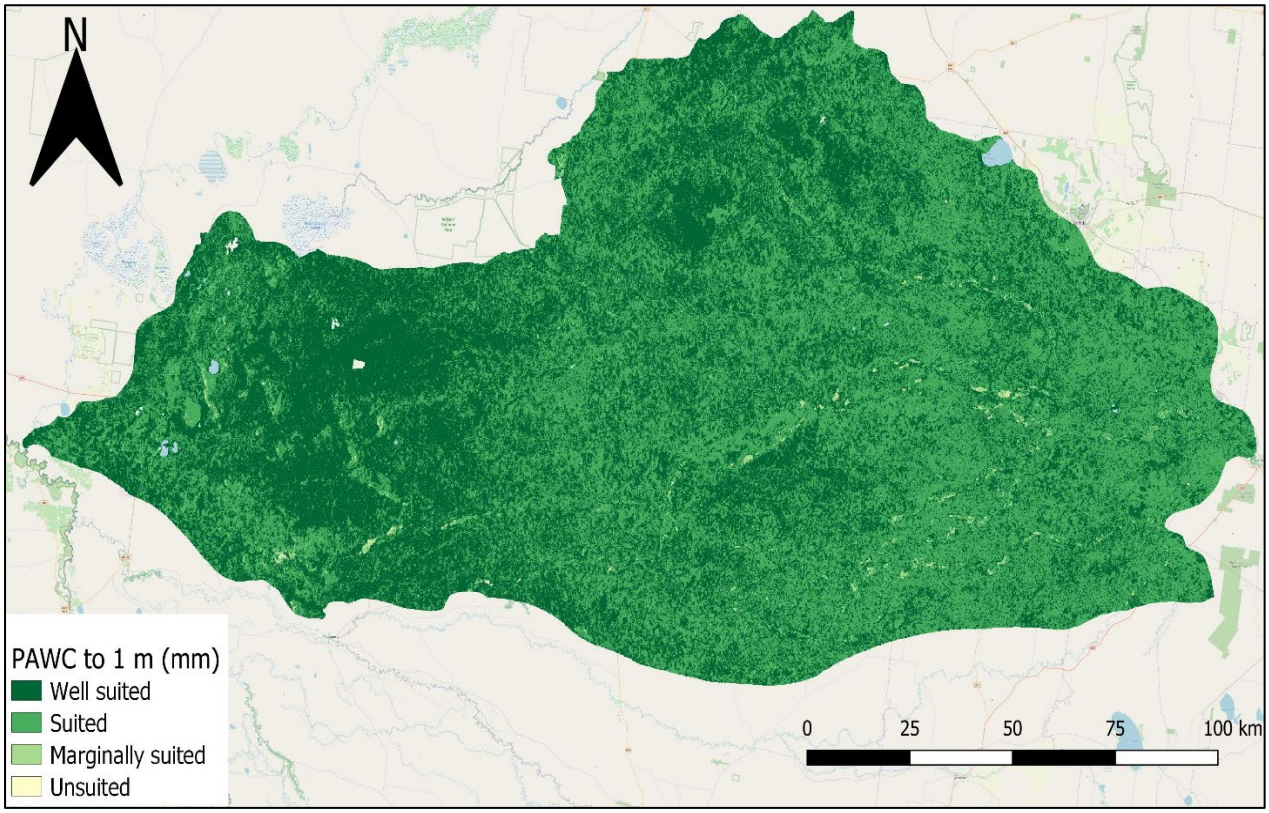


Figure 7.7. Classes of suitability for irrigated cotton production for PAWC (mm) to a depth of 1 m across the lower Murrumbidgee valley.

The PAWC is critical in assessing soils for cotton production, however, parameters for PAWC are not clearly defined. To maximise lint yield a cotton crop requires approximately 800 mm of water during the growing season (CRDC and CottonInfo, 2023) and a sufficiently large PAWC to store moisture between irrigation cycles and through fallow periods. The parameters set in this study seek to balance past literature with regionally specific observations. This study has calculated the total PAWC to 1 m, however, Figure 7.4 identifies that 62.3% of the study area is constrained by alkalinity or sodicity within the first metre of the soil profile. This can result in stored water being unavailable for plant uptake. Tilse et al. (2022) quantified interactions between the depth-to-shallowest constraint to determine the constrained available water capacity (AWC). Across 80,000 ha in northern NSW they determined that the water made unavailable due to constraints, calculated as PAWC minus the constrained PAWC, ranged from 0-206.4 mm in the upper 1.2 m of soil the profile (Tilse et al., 2022). While this study has applied these two parameters separately in the suitability analysis, there is the potential to synthesise these resources into a single constrained PAWC map at either the valley or farm scale.

7.4.4. Classification and mapping of soil suitability for irrigated cotton production

In total, 16.7% of land, is considered well suited to irrigated cotton production, with these areas spread throughout the region (Figure 7.8). A further 31% of the area is considered suited for cotton production (Figure 7.8). The location of these areas, particularly in the east of the region and near the Murrumbidgee River, align with understandings of where cotton can, and has, been successfully grown (Figure 7.1). While the industry has only expanded significantly in the last 15 years, soils of the region's east have been successfully farmed for up to a century. Suitability generally decreases along an east-west axis (Figure 7.8) which is consistent with locations where the depth to a soil chemical constraint becomes shallower (Figure 7.2) and soil chemical constraints become the most limiting factor (Figure 7.9). Overall, 34% of the area is considered as marginally suited while a further 18.4% is considered unsuited to irrigated cotton production (Figure 7.8). Many of these unsuited or marginally suited areas, particularly in the centre of the study area, have been used to successfully grow cotton, suggesting further interrogation of the most limiting factor and management options are required (Figure 7.1).

Suitability, especially low suitability, is spatially linked to sodicity (Figure 7.3; Figure 7.8). This is not surprising given soils in this region have previously been identified as extremely sodic, especially in the west. Interestingly, some patterns of suitability align with schematic

landscape diagrams presented over 60 years ago (Figure 2.16). In the central north of the region, channels of well suited and suited land are flanked by unsuited areas. The distance over which this occurs is similar to the initial landscape models of Butler (1958) and Langford-Smith (1960). This is not surprising given that landscape models of this kind informed the initial development of irrigated land within the region. When considering areas that are constrained by sodicity these resources should be considered alongside agronomic knowledge. This is because it has been demonstrated that cotton, under certain management, can be grown successfully in strongly sodic soils (Chapter 5). As overall plant production is influenced by a complex suite of factors, these products, and suitability maps in general, should be considered alongside local agronomic resources to assess the practical implications of the 'hard rulesets' employed.

Based on this analysis, the most suited land not presently used for irrigated cotton production is located in the region's south and central north (Figure 7.8). Based on the trends of suitability, these pockets of land spatially align with the pathways of younger palaeochannel systems (Page & Nanson, 1996; Page et al., 2009). This suggests that understanding the pathways of these systems and their characteristics can directly assist in understanding suitability. This linkage is not surprising as it was the foundation of initial land use considerations within soil surveys in the region's east and has underpinned the ongoing success of these irrigation areas (Butler, 1979; van Dijk, 1961; van Dijk & Talsma, 1964).

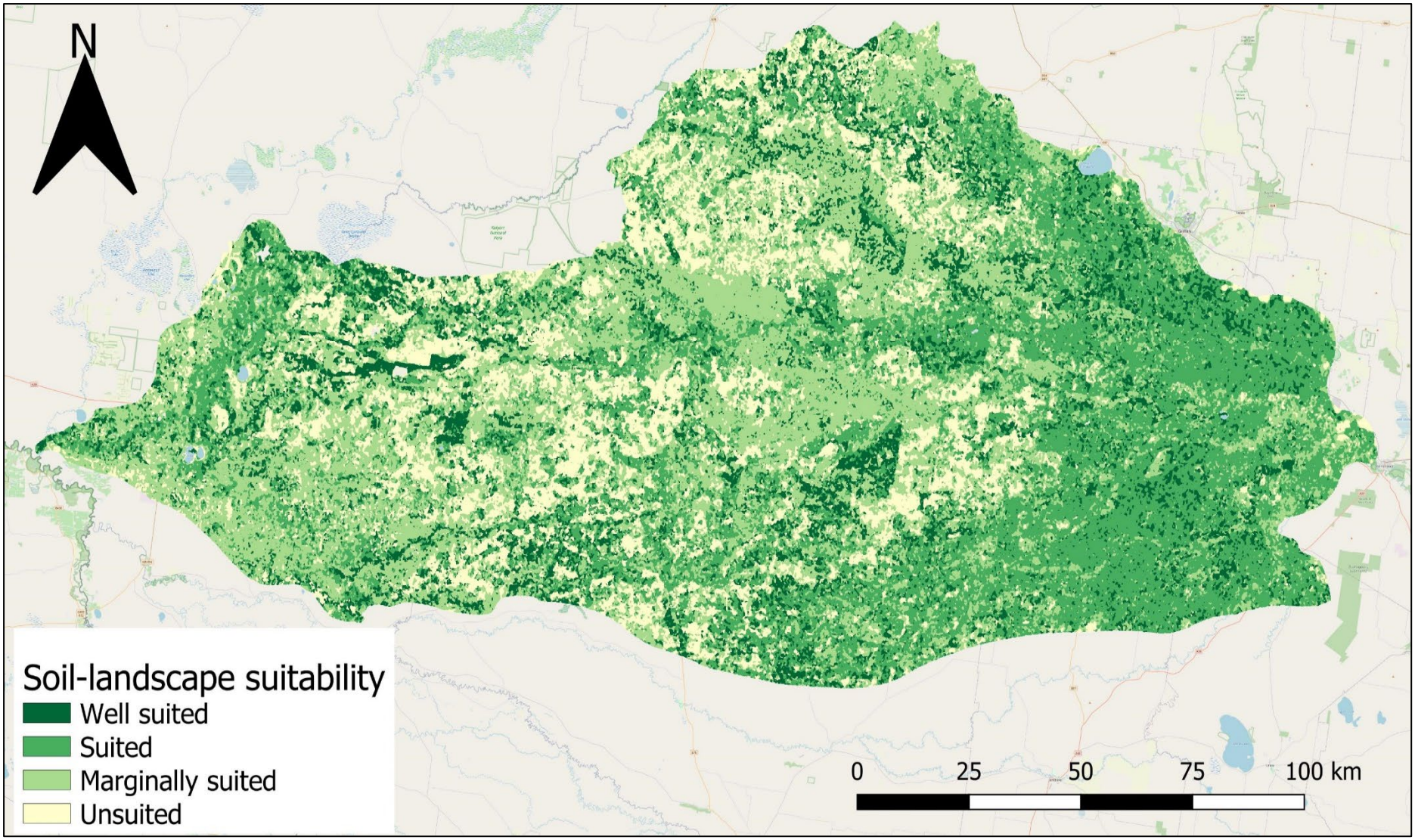


Figure 7.8. Soil-landscape suitability for irrigated cotton production across the lower Murrumbidgee valley in southern NSW. Soils are classified as: well suited, suited, marginally suited, or, unsuited.

7.4.5. Most limiting factor

The identification of the most limiting factor alongside overall suitability is important in assessing land management options. Across the lower Murrumbidgee valley, soil chemical constraints are the most limiting factor by area, limiting suitability for 52.9% of the region. Soil PAWC is the second most limiting factor, covering 24.9% of the area primarily in the east and along tracts of paleochannel systems (Figure 7.9). Slope is the least limiting factor by area, primarily contained to land surrounding the Murrumbidgee River and on slopes in the far east (Figure 7.9). Areas where PAWC is the most limiting factor may be equally limited by slope as many of the region's sandier soils are associated with dunes (Stannard, 1970). A total of 16.7% of the lower Murrumbidgee is unconstrained by the assessed parameters, with pockets of this land spread throughout the region (Figure 7.9).

It is important to consider the most limiting factor alongside the overall suitability (Figure 7.8) and the suitability of individual parameters. This is because not all limiting factors have the same impact. For example, while PAWC is the most limiting factor for roughly a quarter of the region (Figure 7.9) this is primarily in the east where PAWC is still classed as suited for cotton (Figure 7.8). Conversely, many areas where soil chemical constraints are the most limiting factor are classified as unsuited or marginally suited (Figure 7.4). A focus may be placed on managing the latter, however, PAWC and slope are largely inherent and generally unable to be sustainably or economically altered. Further, a low PAWC cannot simply be addressed by irrigating more frequently, with Neupane et al. (2021) identifying that cotton yields do not respond to increased water application in soils with >50% sand content. At the within-field scale PAWC may be a more significant local limitation in areas otherwise regarded as suitable. While not a significant factor due to the landscape of the lower Murrumbidgee valley, where greater topographical variability is present slope will more frequently be a limiting factor, as demonstrated by Thomas et al. (2018) in northern Australia.

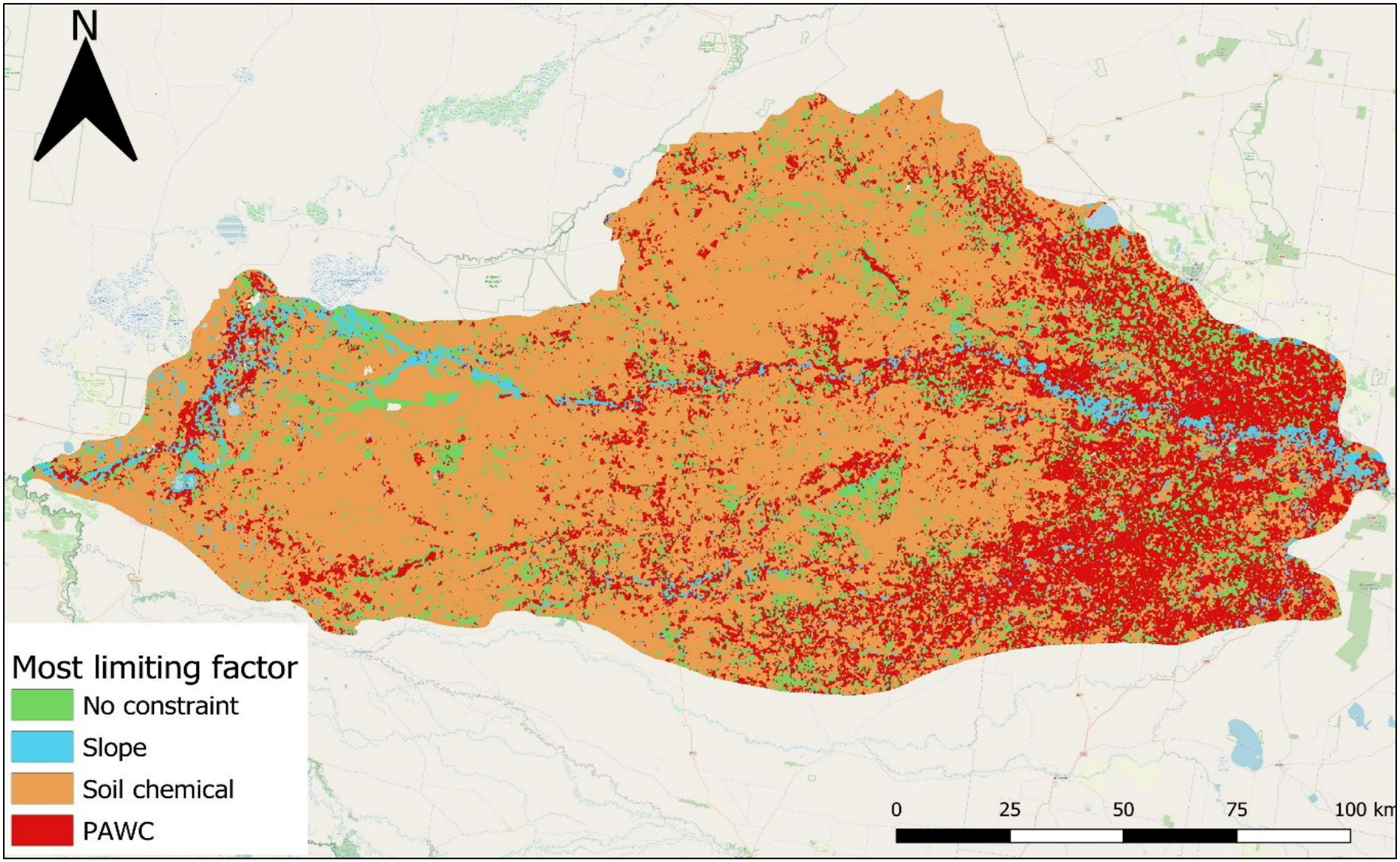


Figure 7.9. The factor identified in this study to be most limiting to irrigated cotton suitability across the lower Murrumbidgee valley.

7.5. Opportunities to enhance this study

Incorporating local knowledge in parameter selection and subset determination was important as if chemical suitability parameters from other studies were used, a higher proportion of land known as capable of producing cotton would have been considered marginally suited or unsuited. Despite this, these parameters remain somewhat subjective, and classes are hard to quantify in terms of actual cotton production. Consequently, the maps produced, while providing value to the industry by leveraging soil spatial information, are a guide only and there are opportunities for improvement. While categorical classifications as applied in this study are common, hard distinctions may not be representative of the continuity of land (Triantafilis et al., 2001). Further, it is difficult to realistically justify that a PAWC of 99 and 101 mm are classed separately despite their similarity while values of 101 and 124 mm, despite being more different, are placed in the same class. From a management perspective a difference of 2 mm is unlikely to govern decisions. Malone et al. (2015) presented a similar perspective, highlighting a lack of sensibility in classing that plants will grow in a subsoil with a pH of 5.5 but not 5.4. With respect to soil chemical constraints specifically, the depth to constraint approach employed in this study may be a more appropriate option than hard thresholds for each soil property. Another alternative to these 'hard rulesets' is to develop continuous classes, an example of which is outlined by Triantafilis et al. (2001) who employed a fuzzy approach as an alternative method to classify land on a scale of 0 (unsuitable) to 1 (suitable). Further, cotton is grown within the region in rotations with other crops. Consequently, decisions on landuse and practices are seldom based on cotton only with considerations for wheat, barley and corn, among others, required.

Quantifying the uncertainty of DSMs and incorporating this into DSAs can increase the value and reliability of the resource. There are various methods available for undertaking this and understanding prediction confidence spatially. Across Australia it is common for soil sample datapoints to be densely concentrated in some areas but spatially sparse in others, as visualised by Searle et al. (2021). This can lead to misleading model validation statistics resulting from spatial autocorrelation. To avoid this, spatial cross-validations can be used, as demonstrated by Pozza et al. (2022). Their study also implemented the area of applicability, a method developed by Meyer and Pebesma (2020), to spatially identify where reliable predictions can be made. Carre et al., (2007) outline approaches to assess the accuracy of soil information, environmental

covariates, soil geographic space inference systems and DSMs, allowing for the accounting of different types of error, from each of these sources, prior to the undertaking of DSAs.

The benefits of incorporating uncertainties into suitability assessments are demonstrated by Malone et al. (2015) who, in a Tasmanian study assessing land suitability for hazelnut production, showed an assumption of error-free inputs predicted 18% of land as unsuited (Malone et al., 2015). In contrast, when the suitability was determined using the limits of the 90% confidence prediction interval, 66% of land was predicted as unsuited (Malone et al., 2015). It was discussed that this was because the root mean square error (RMSE) for pH and ECe, two main contributors, were near the rigid level of their respective parameter thresholds (Malone et al., 2015). Carre et al. (2007), however, note even if a map has a high uncertainty, it is a good starting point in DSA exercises.

This study demonstrated the capacity of DSAs and applications of this kind should not be limited to a single industry. In Tasmania, Kidd et al. (2015) assessed 20 different crops while Harms et al. (2015) evaluated land for 76 different uses in the Flinders and Gilbert catchments of northern Australia. A similar undertaking could be of benefit in the lower Murrumbidgee given the highly variable soils and diverse agricultural industries already present. For larger land holders, this may allow greater diversification with confidence that the ‘most suitable’ crop is being grown on the ‘most suited’ land. In irrigated areas, such resources are valuable given the cost outlay often required to develop or change irrigation systems.

Assessing the suitability for all potential landuses may also provide the opportunity to follow the pathway outlined by Kidd et al. (2015) involving the creation of an enterprise versatility index or the median potential gross margin, allowing the most versatile or highest valued land, respectively, to be identified (Kidd et al., 2015). Regardless of future approaches, DSAs can provide valuable support resources to guide decisions. This is of critical importance considering the increased pressure placed upon farmers from increasing input and operational expenses, fluctuating commodity prices, variable climatic conditions and water scarcities.

7.6. Conclusion

This study conducted a digital soil assessment (DSA) to showcase how soil suitability for irrigated cotton production varies in the lower Murrumbidgee valley of southern NSW. Land was classified into four categories: well suited, suited, marginally suited, or unsuited. This classification was based on regionally specific rulesets for three parameters: the depth to soil chemical constraints, slope, and PAWC. Overall, 16.7% of the region is considered as well suited for irrigated cotton production, 31% is considered suited, 34% is considered marginally suited and 18.4% is considered unsuited. Much of the area considered well suited has already been used to grow cotton, particularly in the east of the lower Murrumbidgee valley. Across 52.9% of the region soil chemical constraints were the soil-landscape factor most limiting to irrigated cotton production. Due to the impact that soil constraints have on the effective rooting depth of plants, suitability on account of this factor was determined by identifying the depth to any soil constraint. These depth-to-constraint maps are a further application of DSAs, effective in streamlining the dissemination of information to stakeholders. Across 44.7% of the region an ESP of 20% was reached within the first meter of the soil profile, while 37.7% of the region was considered unconstrained by pH or ESP. Although PAWC was the most limiting factor for a quarter of the area, most of this land was still classed as suited for irrigated cotton production, with only 0.9% of the lower Murrumbidgee considered marginally suited or unsuited due to this parameter. This highlights the need for multiple levels of information including the overall suitability, most limiting factor, and specific influence of this factor, to fully interpret the information. These products should be considered alongside agronomic knowledge of the region as it is known that cotton is able to be grown in areas where soil chemical constraints are present. If further land was to be developed for irrigated cotton production, this analysis demonstrated the most suited land for this to occur is in the south and central north of the lower Murrumbidgee valley. This land has likely not been developed due to the distance from the Murrumbidgee River or irrigation supply channels. Access to groundwater, however, may see this land developed in the future.

Chapter 8

8. General discussion and concluding remarks

8.1. Introduction

This project was designed to meet the needs of, among other industry members, farmers, agronomists and land managers. This chapter is a two-part discussion synthesising takeaway messages relating to within-field variability and considerations on data at a valley-wide level. Within each of these discussions, future research directions will be touched on.

Broadly speaking, future research in this region does not need to rely on creating new methods. Farmers and agronomists are eager to learn more about their soils and how to optimise management decisions. Support tools, therefore, can be created by applying and appropriately using a combination of existing methodologies as has been done here. In undertaking any future work, it is imperative to work directly with the industry and its stakeholders. Despite a lack of published data and/or soils information within the region, most farms have multi-generational owners and the farmers have a strong understanding of their soils. The missing link is undertaking basic and applied research to place a scientific foundation behind this anecdotal knowledge. To conclude this thesis, each of the research questions and associated aims will be addressed before final concluding remarks are made.

8.2. General discussion

It has been known for almost a century, by researchers and farmers alike, that the soils of the lower Murrumbidgee valley are highly variable. Despite this, until now most research work has focused solely on the eastern irrigation areas, which was initially appropriate given the high infrastructure investment in these locations. The expansion of the cotton industry has, however, highlighted that more information relating to soils and how they should be managed is required for the lower Murrumbidgee valley in its entirety.

8.2.1. Moving forward: Soil variability in relation to cotton production

This research demonstrated that morphologically distinct deep sandy and heavy clay soils occur directly adjacent to each other. When cotton is grown under surface or flood irrigation systems, including using siphons and furrows or pontoon systems, and this variability is present, it was demonstrated that sandier areas produce lower cotton lint yields. This likely results from the impact of texture on soil hydrology and water holding capacity, with sandier soils exhibiting higher infiltration rates, saturated conductivity and a lower water holding capacity (Hazleton & Murphy, 2007). This contrast in textures also presented interesting findings regarding sodicity. Across the lower Murrumbidgee valley, sodicity was shown to be prevalent, especially

in the west. While this has appropriately been identified as an area of concern (Holland and Eastwood, 2014), at both Farm A and Farm C a higher ESP was positively correlated with cotton lint yield. An extreme example of this was seen at Farm A where, in the higher yielding production zone, ESP approached and surpassed 30% in the subsoil. This highlights the importance of understanding the full suite of soil properties. While some sodic areas of the paddock may benefit from specific management, such as gypsum application, at a paddock-level the strongest influence comes from soil texture. The variability observed here is anecdotally understood to occur across the region. An opportunity to address this, while also understanding the impact of sodicity in the higher yielding areas of fields, is through the development of production/management zones.

This study demonstrated the capacity for production zones to be derived from yield data obtained from a commercial cropping operation. The maintaining of quality yield data, however, is imperative in allowing this process to occur. From a statistical perspective, there are other methodologies that allow for the development of production or management zones, including using only remotely sensed data (Leo, 2022). The derivation of zones allows specific areas of paddocks to be managed more appropriately, as homogenous units (Whelan & McBratney, 2000). Each unit can then be optimally farmed, with inputs matched to the capacity of the soil (Whelan & McBratney, 2000). In this study, where texture was shown to be primarily driving yield variability, this would mean setting lower yield targets in sandier areas that are incapable of supporting higher yields under flood irrigation systems. A reduction in inputs will be of both economic and environmental benefit. For example, in areas where soil texture is most limiting through its influence on soil hydrology, increased applications of nitrogen will have no positive impact and may have undersides consequences downstream from eutrophication.

Deciding on which methods to employ to understand within-field variability spatially can be difficult. Digital soil mapping (DSM) can be complex, with Filippi et al. (2024) noting a lack of expertise as a reason it is not commonly undertaken at the field scale. An alternative method is the simple spatial interpolation of data between sampling points. As a consequence of this, grid sampling is a common technique used by some consultants, as a high density of datapoints is required for the interpolation process. Both the cost of sampling and the time it takes increases under such a strategy, reducing the capacity of additional subsoil samples to be taken. Spatially informed zonal sampling may be considered, however, the initial drawbacks regarding expertise or a lack of data may be realised, and grid sampling is chosen again.

As this study demonstrated, linear models may be an alternative approach to understanding spatial variability at the paddock scale. Where machine learning methods are complex to implement and interpret, while also requiring more training data, linear models are relatively intuitive, interpretable and can be developed with fewer samples. By removing the need to interpolate data spatially, random stratified or targeted sampling can occur with models then developed and predictions made using spatial covariates. By removing the need to interpolate between points, management or production zones can be used to develop representative sampling plans, eliminating the need for grid sampling (Mallarino & Wittry, 2001; Tilse et al., 2022). This has the potential to reduce the number of samples and cost outlay, required to accurately understand spatial variability (Lawrence et al., 2020; Tilse et al., 2022). The reduction in samples taken from using a zonal approach may also allow for more subsoil samples to be analysed within the same budget. The importance of this is demonstrated in this study, with the strongest soil-yield correlations present in the subsoil. It is at these depths where later season crop growth and cotton boll development will likely be impacted as plant roots explore deeper in the soil profile (CRDC & CottonInfo, 2023). The sampling of the subsoil can also allow for the concurrent modelling of soil to depth, as demonstrated here. This presents opportunities to develop within-field depth-to-constraint maps, as described by Tilse et al. (2022). As demonstrated at the valley-scale in this study, these can reduce the dimensionality of data, simply showing the target variable of interest.

To achieve desired model quality, strong consideration should be given to the number of sample sites required. This will vary depending on the heterogeneity of each paddock. This research showed that an increase in sample size, at Farm C, increased model quality and also resulted in the inclusion of more spatial covariates, both proximally and remotely sensed, in the final model. In contrast, at Farm A only proximally sensed covariates (electromagnetic induction and gamma radiometric data) and layer mid-depth were included in the final model. This study did not seek to examine whether proximal or remotely-sensed covariates produce superior models, so it is not appropriate to make statements on which combination of these covariates will best serve landholders. However, given the inclusion of proximally sensed covariates in all field-scale models, and their selection as sole predictors at Farm A, there may be value in obtaining this data, especially if less soil samples are to be taken.

Considerations should be given to the efficacy of surface irrigation systems where there is extreme variability at the within-field scale. Roth et al. (2013) state that gravity-surface irrigation systems have traditionally been the mainstay of the Australian cotton industry, but

there has been a shift to increased adoption of overhead pivot or lateral-move irrigation systems. These systems can allow for the integration of variable rate irrigation technology, providing the capacity to improve both water and nitrogen use efficiency, the latter through fertigation (Antille, 2018; McCarthy et al., 2023; Neupane & Guo, 2019). Given the significant costs associated with these inputs, and the extent to which textural variability occurs in lower Murrumbidgee valley soils, these systems warrant further investigation. Future research should seek to assess the viability and potential benefits to be gained from these overhead systems and/or variable rate irrigation in southern NSW specifically. Bankless channel irrigation systems have also become more common recently and may be a 'middle of the road' alternative between flooded furrow and overhead irrigation systems.

To conclude, the cotton industry in southern NSW has gained by learning from the industry in the north of the state. The soils in the lower Murrumbidgee valley, however, are highly variable and distinctly different from those in the Gwydir or Namoi valleys where much research has been focused (Cattle & Field, 2013). Given the extent of within-field variability in this region, and the impacts that this has on yield, a new regional focus is needed. This paradigm shift may include taking a zonal approach to management, with yield targets set based on the inherent capacity of the soils or considering new irrigation systems. Future research should target the latter; different irrigation systems may be required to maximise efficiency in the lower Murrumbidgee valley compared to northern growing regions. To make informed decisions on management there is also a need to understand the full suite of soil attributes to depth. Zonal sampling and bespoke DSM methods present an opportunity here, however, these benefits need to be clearly communicated to stakeholders. A key focus should be this communication, both through extending scientific knowledge but also by listening to what stakeholders require from researchers.

8.2.2. Improving digital soil mapping at the regional scale

The quality of a DSM output is dependent on the soil data, the spatial covariate data as well as practitioner expertise and the methods employed in its development. In the latter two instances, technological advancements have seen DSM advance rapidly and there are opportunities to improve practitioner skill. To see significant improvements at a local level a greater quantity of quality data is needed. This is not to say that there isn't data available, as testing and private surveys are common in the lower Murrumbidgee valley. If data and information were to be

shared, however, new knowledge can be derived for stakeholders in the cotton industry and land managers in the lower Murrumbidgee valley.

This study has shown the potential of regionally-developed DSMs, reinforcing other research that has suggested that regional, or even local, DSM approaches outperform state, national and global models (Lemercier et al., 2022). There is the capacity to improve the accuracy of the soil mapping products by increasing the data incorporated into the modelling process. A future direction should look towards data sharing at an industry or regional level to further improve resources targeted specifically to that industry or region. While some farmers may be apprehensive about sharing data due to privacy concerns, there are appropriate means with which this can be anonymised with codes now developed for sharing data under agreed terms and conditions (Australian Research Data Commons [ARDC], 2024). This can then be used to build or improve models, with raw data not publicly released. If led by known industry members, trust may increase also.

There are multiple potential benefits from this. The first of these is through improved general spatial resources, such as the models and maps produced here, the quality of which would likely increase with more data. Improved models then have the capacity to be of benefit at the farm, as well as valley scale. This could occur through spatial downscaling (Malone et al., 2017b) or through using the regional model and data in association with on-farm data to improve bespoke farm or paddock scale DSMs. The benefits of this have been demonstrated by Filippi et al. (2020) in the Namoi valley.

This thesis also showcased the potential of DSMs to be used as inputs for digital soil assessments (DSAs), such as depth-to-constraint maps and suitability classifications for numerous crops. These resources can assist with guiding investment in infrastructure or can be used to make decisions on diversifying production systems. There is also the capacity, as outlined by Kidd et al. (2015), to develop enterprise versatility indexes or enterprise specific gross margins. The latter can provide an understanding of land value spatially based on production capacity. This may further investments in off-farm infrastructure, such as processing or value adding facilities, by quantifying the productive capacity of land.

Lastly, a point for consideration. Initial soil surveys of the Murrumbidgee Irrigation Area, almost a century ago, classified suitability for horticultural crops, laying the foundation for one of Australia's most productive 'food-bowls'. It guided investments in appropriate infrastructure targeted to specific crops based on soil type. There is no reason, given the 'data age' we are

living in, that similar but more technological approaches cannot be employed again. This is of particular importance given increased pressure placed on farmers through rising input costs and fluctuating water allocations. Farmers are also seeking to maximise water use efficiency which may result in land re-development or changes to the production system entirely. Improved access to spatial information can assist in decision making, including selecting the most sustainable land to develop or reform.

Ultimately there are benefits to be gained from improving understandings of the spatial distribution of soil properties. Irrespective of the target soil property, an increase in data will be of benefit. Presently, many farmers have access to a wealth of valuable soil data that, if harnessed, could develop products to advance agricultural industries in the lower Murrumbidgee valley.

8.3. Addressing the aims and research questions of this thesis

1. To morphologically classify lower Murrumbidgee valley soils and discuss their occurrence in relation to palaeochannel systems at a regional and within-field scale
   a. What are the typical morphological characteristics of lower Murrumbidgee valley soils?

While the morphological characteristics of the region's soils vary, it was demonstrated that in the west of the region, clay soils are dominant. This provides a more targeted insight into the lower Murrumbidgee valley, with prior research focused on the eastern areas where earthy soils are more prolific. Sodicity and salinity were shown to be common, especially in the clay soils of the region's west. While soil profile descriptions to the Suborder level of the Australian Soil Classification (ASC) are provided, these were reliant on average data where significant variability was present. Pedogenic lime was common as nodules, veins or pedogenic segregations, especially in clay soil classes, as was the presence of subdominant colours unrelated to mottling. While present, soil coarse fragments are not at a size where they will impact agricultural tillage practices.

   b. Can soil cores extracted from across the entirety of the region be appropriately classified using previously-developed, region-specific, soil groups?

It was decided that classifying the 153 soil cores using a previously developed system was not appropriate. Firstly, soil groupings related to past surveys utilised localised geographic labels to create associations. Cores from this study were taken from beyond these regions and it is believed that general, not localised, classes are more appropriate. Secondly, prior surveys have occurred in smaller pockets in the region's east, accounting for only 5.6% of the lower Murrumbidgee valley. Consequently, the previous classifications are based on data from these subregions where 'Earth' soils are dominant. This study identified that across the previously unsurveyed central and west of the region, which accounts for most of the total area, clay soils are dominant. It was deemed that a greater focus be placed on examining and then differentiating between these clay soil types. A new classification scheme was therefore created, classing the 153 soil profiles as: Red-brown clays (n=16), Red-brown calcic clays (n=31), Red-brown gypsic clays (n=10), Grey clays (n=17), Grey calcic clays (n=26), Grey gypsic clays (n=7), Transitional earth soils (n=9), Texture contrast soils (n=6), Deep sands (n=3) and exceptions, which did not fall into the prior nine classes (n=28).

Adopting the developed classification scheme also allowed for Texture contrast and Transitional soils to be differentiated. If these soils were classified using the ASC this distinction, based on the depth where a clear, sharp or abrupt texture change occurs, would not have been possible. It is believed that this separation is important as, historically, the transitional soils have been considered to be more similar to heavy clay soils and successfully used for rice production. These systems require management considerations distinct from the previously classed hillslope soils, which are often subplastic, and fall within the Texture contrast class.

c. How is the distribution of soil classes related to palaeochannel systems?

The distribution of soils is determined by the pathways of palaeochannel systems. Regionally, this can be witnessed in the soil group map where pathways of transitional earth soils cross the landscape in an east-west direction. In the north of the area, grey coloured clays associated with a younger stream have been deposited atop red-brown clays of an older stream. This interaction can be witnessed within soil profiles where there are sharp or abrupt colour changes from overlying grey, to underlying red-brown, sediments. This is also witnessed in soil texture, where deposits of younger sediments have resulted in texture contrast soils that have not

developed through pedogenic processes. Deep sands occur either within the streambeds of, or on the dunes adjacent to, paleochannel streams.

d. What is the extent of soil morphological variability at the within-field scale?

Morphological variability was particularly evident at Farm A where a streambed sand, associated with a palaeochannel anabranch, has incised within the paddock. This causes Deep sands to occur directly adjacent to heavy clays, with no discernible difference in topography. This variability will significantly impact irrigated agricultural production, which this paddock has been developed for. Understanding the trends observed at Farm C is complicated by land reformation potentially cutting areas where sandier topsoils were present. When compared to Farm A, Farm C showcases the increased textural uniformity present on soils adjacent to more significant paleochannel systems.

2. To assess how the variability of inherent soil properties at the within-field scale impacts irrigated cotton production
    a. What is the extent of soil variability between yield-derived production zones within paddocks used to grow cotton?

At Farm A there was significant variability in measured soil properties between the higher and lower yielding zones. In the higher yielding cluster, soils were significantly more clayey and had a higher ESP which approached 30% at depth. The clusters also aligned spatially with morphological variability and surficial trends in soil colour. There was greater within-cluster variability at Farm C, however, both median clay content and ESP were greater in the higher yielding cluster. The exceedance of plant-limiting sodicity thresholds in the higher yielding zones indicates that this is not the most limiting factor at a paddock-wide scale. Future work could seek to derive a greater number of clusters or to examine point correlations within each cluster. Irrespective, this analysis still provides a clear picture of trends within the paddocks.

b. How are inherent soil properties correlated to cotton lint yields?

A point correlation analysis identified soil texture, represented by sand and clay content, as the soil property most correlated to lint yield. This was demonstrated by moderate-to-strong correlations between clay content and lint yield in the topsoil and subsoil. The inverse was

displayed for sand content, which was moderately-to-strongly negatively correlated to lint yields. Across both farms correlations were generally strongest in the subsoil (0.6-0.8 m). At this depth at Farm A, clay content exhibited a correlation (*r*) of 0.95 with lint yield. Exchangeable sodium percentage (ESP) was positively correlated to cotton lint yield across all years, at all depths, at both farms. This is likely due to the positive correlation exhibited between ESP and clay content. This highlights the complexity of soil-crop relationships and the need to avoid examining individual soil properties in isolation.

c. Can paddock scale digital soil maps accurately represent soil variability spatially and vertically?

Simple linear models were able to predict soil properties at multiple depths concurrently to a moderate-to-good quality when assessed using the Lin's Concordance Correlation Coefficient (LCCC) in a leave-one-site-out-cross-validation (LOSOCV). At Farm A, model LCCCs ranged from 0.45-0.65 with improved performance at Farm C, with LCCCs of 0.62-0.73. This improvement at Farm C corresponded with a greater number of sample sites and a wider range of proximal as well as remotely sensed covariates being retained through the model selection process.

3. To model and map multiple soil properties across the lower Murrumbidgee valley at a 90 m resolution
    a. Can digital soil mapping be used as a tool to accurately model and then map soils at multiple depths concurrently?

Extreme Gradient Boosting (XGBoost) models were demonstrated as capable of predicting soil properties at multiple depths concurrently, with moderate quality, across the region. In a LOSOCV, LCCCs ranged from 0.45 (ECe) to 0.58 (clay). Pleasingly, when assessed visually the produced maps identified known landscape variability, following trends of previously mapped paleochannel systems. Soil texture was generally predicted to be clayey throughout the first metre of the soil profile, excluding sandier tracts of land associated with the pathways of paleochannels. ECe and ESP were predicted to increase along a westerly trend, with extreme salinity or strong sodicity commonly encountered in the first metre of the profile. A key step in this exercise was the visual assessment to ensure maps were pedologically appropriate.

b. How do presently available digital soil mapping products perform when validated against external datasets from the lower Murrumbidgee valley at the point scale?

The state, national and global DSM products each performed worse than their published model statistics when validated at the point scale using the data collected in this study. This aligns with previous findings questioning the capacity of these 'broader' products to accurately identify variability at the point or farm support. This reinforces that these products, while identifying landscape trends, are not currently capable of providing accurate information to assist with within-field decisions in the lower Murrumbidgee valley. This is likely the result of sparse data within the region and the broad 'global' scale that these models are built on. Visually, too, these maps do not detect the fine scale variability known to be present.

4. To identify the most limiting factor to, and classify the overall suitability for, irrigated cotton production in the lower Murrumbidgee valley
    a. How does the depth to which soil chemical constraints are reached vary across the region?

There is a clear spatial trend in the depth to any soil constraint and the most limiting soil chemical constraint, with large areas in the east of the region unconstrained by ESP or pH in the first metre of the soil profile. Soil ECe was not included in this analysis as trends within the region and prior research suggest salinity will likely decrease under irrigated cotton production. Areas where an alkalinity threshold of 9 is reached in the first metre of the soil profile occur adjacent to major palaeochannel streams; on calcareous soils in the west, at the interface of the Mallee Dunefields; or in the northeast. Overall, 17.6% of the region reaches this alkalinity threshold within the first metre of the profile while 37.7% of land is considered unconstrained to the same depth. ESP is the most limiting soil constraint for the remaining 55.3% of land, with the depth to constraint decreasing along the east-west axis. In these central and western regions, where ESP is most limiting, there are tracts of land associated with younger palaeochannel systems that are unconstrained to 1 m.

b. What are the soil-landscape factors most limiting to cotton production?

Following the development of rulesets adjusted specifically for the lower Murrumbidgee valley, it was shown that soil chemical constraints were the most limiting factor by area, limiting suitability for irrigated cotton production across 52.9% of the region. The influence of

plant available water capacity (PAWC) highlights the importance of viewing all data concurrently. While suitability was limited by PAWC across 24.9% of the area, only 0.9% of the lower Murrumbidgee valley is considered marginally suited or unsuited on account of PAWC. This occurs primarily in the east of the region, where PAWC reduces land from being classed as well suited to suited. In contrast, much of the land limited by soil chemical constraints is considered marginally suited or unsuited. Thus, not only does the most limiting constraint need to be identified, but the extent to which it is limiting is crucial.

c. What proportion of the lower Murrumbidgee is considered suited to irrigated cotton production?

Using a most limiting factor approach the suitability of land in the lower Murrumbidgee valley for irrigated cotton production was classed as: well suited, suited, marginally suited or unsuited. Overall, 17% of land is considered well suited while an additional 31% of the area was classified as suited. Much of the land classed as well suited either is or has been used to grow cotton. Of the remaining 52% of land, 34% was considered marginally suited with 18% of the land area classed as unsuited for irrigated cotton production. Increasing the sodicity threshold was important to avoid areas known to be capable of producing profitable crops being classed as unsuited on account of this. The trends of suitability are related to pathways of paleochannel systems where veins of well suited or suited land are flanked by unsuited areas or vice versa. Like soil morphological observations, the suitability is spatially variable. Generally, in the east of the area, including the Murrumbidgee and Coleambally Irrigation Areas, land is classed as suited or well suited. This classification used a categorical ruleset, however, this may set unrealistic hard cutoffs. For example, an ESP of 19% is considered unconstrained while 21% is considered constrained. Therefore, adopting a continuous classification scheme may be of benefit.

8.4. Concluding remarks

It has been demonstrated that there is variability in soils at the sub-paddock scale while there are spatial trends and patterns in soils aligned to palaeochannels that infer variability across the lower Murrumbidgee valley. As a result of sediment deposition from temporally distinct palaeochannel systems, morphologically described heavy clay and deep sandy soils can occur directly adjacent to each other. When this occurs at a within-field scale, cotton crops yield substantially less in the sandy areas. Presently available state, national and global digital soil mapping (DSM) products for subsoil clay content did not account for this variability when assessed at the point support using data collected in this study. A regional DSM approach was able to map soil properties of interest across the region with moderate model quality statistics and pedologically appropriate spatial patterns. Utilising these DSM products, a digital soil assessment (DSA) was undertaken. This predicted the most limiting factor for irrigated cotton production to be soil chemical constraints, of which sodicity was the constraint predicted to be most frequently encountered at the shallowest depth. This, however, is a constraint at a broader scale, with it demonstrated that when there is variability between heavy clay and deep sandy soils within-field, high sand content is most limiting. Consequently, when considering these suitability products an understand of agronomic relationships is also required. Overall, the majority of the lower Murrumbidgee valley is considered either suited (31%) or marginally suited (34%) to irrigated cotton production. It is predicted that 17% of the region is well suited to irrigated cotton production while 18% is considered unsuited, with a large proportion of the well-suited areas already used for broadacre cropping.

References

Abdullaev, I., Ul Hassan, M. & Jumaboev, K. (2007). Water saving and economic impacts of land leveling: the case study of cotton production in Tajikistan. Irrigation and *Drainage Systems, 21*(1), 251–263. https://doi.org/10.1007/s10795-007-9034-2.

Ackerson, J.P., Demattê, J.A.M. & Morgan, C.L.S. (2015). Predicting clay content on field-moist intact tropical soils using a dried, ground VisNIR library with external parameter orthogonalization. Geoderma, 259-260(1), 196-204. https://doi.org/10.1016/j.geoderma.2015.06.002.

Adeniyi, O.D., Bature, H., Mearker, M. (2024). A Systematic Review on Digital Soil Mapping Approaches in Lowland Areas. *Land, 13(3)*, 379. https://doi.org/10.3390/land13030379.

Ali, A., Rondelli, V., Martelli, R., Falsone, G., Lupia, F. & Barbanti, L. (2022) Management Zones Delineation through Clustering Techniques Based on Soils Traits, NDVI Data, and Multiple Year Crop Yields. *Agriculture. 12*(2), Article 231. https://doi.org/10.3390/agriculture12020231.

Al-Shammari, D., Whelan, B.M., Wang, C., Bramley, R.G.V., Fajardo, M. & Bishop, T.F.A. (2021). Impact of spatial resolution on the quality of crop yield predictions for site-specific crop management. *Agricultural and Forest Meteorology, 310*, Article 108622. https://doi.org/10.1016/j.agrformet.2021.108622.

Anderson, G.C., Pathan, S., Hall, D.J.M., Sharma, R. & Easton, J. (2021). Short- and Long-Term Effects of Lime and Gypsum Applications on Acid Soils in a Water-Limited Environment: 3. Soil Solution Chemistry. *Agronomy, 11*(5), 826. https://doi.org/10.3390/agronomy11050826.

Antille, D.L. (2018). Evaluation of fertigation applied to furrow and overhead irrigated cotton grown in a Black Vertosol in Southern Queensland, Australia. *Applied Engineering in Agriculture, 34*(1), 197-211. https://doi.org/10.13031/aea.12519.

Andres, S.E., Powell, J.R., Emery, N.J., Rymer, P.D. & Gallagher, R.V. (2021). Does threatened species isting status predict climate change risk? A case study with Australian *Persoonia* (Proteaceae) species. *Global Ecology and Conservation, 31*, Article e01862. https://doi.org/10.1016/j.gecco.2021.e01862

Arrouays, D., Grundy, M.H., Hartemink, A.E., Hempel, J.W., Hervelink, G.B.M., Hong, S.Y., Lagacherie, P., Lelyk, G., McBratney, A.B., McKenzie, N.J., Mendonca-Santos, M.d.L., Minasny, B., Montanarella, L., Odeh, I.O.A., Sanchez, P.A., Thompson, J.A. & Zhangm G. (2014). Chapter Three – GlobalSoilMap: Toward a Fine-Resolution Global Grid of Soil Properties. *Advances in Agronomy, 125*(1), 93-134. https://doi.org/10.1016/B978-0-12-800137-0.00003-0.

Arrouays, D., Poggio, L., Guerrero, O.A.S. & Mulder, V.L. (2020). Digital soil mapping and GlobalSoilMap. Main advances and ways forward. Geoderma Regional, 21, Article e00265. https://doi.org/10.1016/j.geodrs.2020.e00265.

Australian Bureau of Statistics [ABS]. (2020). Agricultural Commodities 2018-19, Australia. ABS. https://www.abs.gov.au/statistics/industry/agriculture/agricultural-commodities-australia/2018-19.

Australian Bureau of Statistics [ABS]. (2023). Australian Agriculture: Broadacre Crops 2022-23. ABS. https://www.abs.gov.au/statistics/industry/agriculture/australian-agriculture-broadacre-crops/latest-release.

Australian Bureau of Statistics [ABS] (2022). 2021 Census Data. https://www.abs.gov.au/census/find-census-data/search-by-area, accessed 21 May 2024.

Australian Research Data Commons [ARDC] (2024). Data Sharing Initiative. https://doi.org/10.47486/DC102

Badgery, W., Simmons, A., Eckard, R. & Grace, P. (2024). Reducing GHG emissions in cropping systems – responding to drivers for change. *GRDC Update Papers,* March 2024.

Beecher, H.G. (1991). Effect of saline water on rice yields and soil properties in the Murrumbidgee Valley. *Australian Journal of Experimental Agriculture, 31*(6), 819-823. https://doi.org/10.1071/EA9910819.

Ben-Dor, E., Levin, N., Singer, A., Karnieli, A, Brain, O. & Kidron, G.J. (2006). Quantitative mapping of the soil rubification process on sand dunes using an airborne hyperspectral sensor. *Geoderma, 131*(1-2), 2-21. https://doi.org/10.1016/j.geoderma.2005.02.011.

Bennett, J.Mcl, Cattle, S. R., & Singh, B. (2014). The Efficacy of Lime, Gypsum and Their Combination to Ameliorate Sodicity in Irrigated Cropping Soils in the Lachlan Valley of New South Wales. *Arid Land Research and Management, 29*(1), 17–40. https://doi.org/10.1080/15324982.2014.940432.

Bennett, J.Mcl, Roberton, S.D., Ghahramani, A. & McKenzie, D.C. (2021). Operationalising soil security by making soil data useful: Digital soil mapping, assessment and return-on-investment. *Soil Security, 4*, Article 100010. https://doi.org/10.1016/j.soisec.2021.100010.

Blackmore, A.V. (1976). Subplasticity in Australian Soils. IV* Plasticity and structure Related to Clay Cementation. *Australian Journal of Soil Research, 14*(1), 261-272. .

Bishop, T.F.A., Horta, A. & Karunaratne, S.B. (2015). Validation of digital soil maps at different spatial supports. *Geoderma, 241-242*(1), 238-249. https://doi.org/10.1016/j.geoderma.2014.11.026.

Bouma, J. (1986). Using soil survey information to characterize the soil-water state. *Journal of soil science, 37*(1), 1-7. https://doi.org/10.1111/j.1365-2389.1986.tb00001.x.

Bowler, J.M. (1978). Quaternary Climate and Tectonics in the Evolution of the Riverine Plain, Southeastern Australia. In J.L. Davies & M.A.J. Williams (Eds.), *Landform Evolution in Australasia* (pp. 70-113). Australian National University Press, Canberra, Australia and Norwalk Connecticut.

Bowler, J.M., Hope, G.S., Jennings, J.N., Singh, G. & Walker, D. (1976). Late Quaternary climates of Australia and New Guinea. *Quaternary Research, 6*(3), 359-394. https://doi.org/10.1016/0033-5894(67)90003-8.

Bowler, J.M. & Magee, J.W. (1978). Geomorphology of the Mallee Region in semi arid northern Victoria and western New South Wales. *Proceedings of the Royal Society of Victoria, 90*(1), 27-42. .

Brackin, R., Buckley, S., Pirie, R. & Visser, F. (2018). Predicting nitrogen mineralisation in Australian irrigated cotton cropping systems. Soil Research, 57(3), 247-256. https://doi.org/10.1071/SR18207.

Brewer, R. & Blackmore, A.V. (1976). Subplasticity in Australian Soils. II* Relationships between Subplasticity Rating, Optically Oriented Clay, Cementation and Aggregate Stability. *Australian Journal of Soil Research, 14*(1), 237-248. .

Bricklemyer, R.S. & Brown, D.J. (2010). On-the-go VisNIR: Potential and limitations for mapping soil clay and organic carbon. Computers and Electronics in Agriculture, 70(1), 209-216. https://doi.org/10.1016/j.compag.2009.10.006.

Brown, C.M. (1989). Structural and stratigraphic framework of grounwater occurrence and surface discharge in the Murray Basin, southeastern Australia. *BMR Journal of Australian Geology and Geophysics, 11*, 127-146. .

Brown, C.M. & Stephenson A.E. (1991). *Geology of the Murray Basin, southeastern Australia*. Australian Bureau of Mineral Resources, Geoogy and Geophysics, Bulletin 235.

Brungard, C., Nauman, T., Duniway, M., Veblen, K., Nehring, K. White, D., Salley, S. & Anchang, J. (2021). Regional ensemble modeling reduces uncertainty for digital soil mapping. *Geoderma, 397*, Article 114998. https://doi.org/10.1016/j.geoderma.2021.114998.

Brus, D.J. (2019). Sampling for digital soil mapping: A tutorial supported by R scripts. *Geoderma, 338*(1), 464-480. https://doi.org/10.1016/j.geoderma.2018.07.036.

Brus, D.J., Yang, R.M & Zhang, G.L. (2016). Three-dimensional geostatistical modeling of soil organic carbon: A case study in the Qilian Mountains, China. *CATENA, 141*(1), 46-55. https://doi.org/10.1016/j.catena.2016.02.016.

Bui, E. N., Searle, R. D., Wilson, P. R., Philip, S. R., Thomas, M., Brough, D., Harms, B., Hill, J.V., Holmes, K., Smolinski, H.J. & Van Gool, D. (2020). Soil surveyor knowledge in digital soil mapping and assessment in Australia. *Geoderma Regional, 22*, Article e00299. https://doi.org/10.1016/j.geodrs.2020.e00299.

Bullard, J. E., & McTainsh, G. H. (2003). Aeolian-fluvial interactions in dryland environments: examples, concepts and Australia case study. Progress in Physical Geography: Earth and Environment, 27(4), 471-501. https://doi.org/10.1191/0309133303pp386ra.

Bureau of Meteorology [BOM]. (2023). Climate Data Online. Retrieved October 15, 2023, from: http://www.bom.gov.au/climate/data/index.shtml.

Butcher, K., Wick, A.F., DeSutter, T., Chatterjee, A. & Harmon, J. (2016). Soil Salinity: A Threat to Global Food Security. *Agronomy Journal, 108*(6), 2189-2200. https://doi.org/10.2134/agronj2016.06.0368.

Butler, B.E. (1950). A theory of prior streams as a casual factor of soil occurrence in the Riverine Plain of southeastern Australia. *Australian Journal of Agricultural Research, 1,* 231-252. https://doi.org/10.1071/AR9500231.

Butler, B.E. (1958). Depositional systems of the Riverine Plain in South-Eastern Australia in Relation to soils. *Soil Publication 10.* Commonwealth Scientific and Industrial Research Organisation, Melbourne, Australia.

Butler, B.E. (1976). Subplasticity in Australian Soils. Introduction. *Australian Journal of Soil Research, 14*(1), 225-226. .

Butler, B.E., Blackburn, G., Bowler, J.M., Lawrence, C.R., Newell, J.W. & Pels, S. (1973). A *Geomorphic map or the Riverine Plain of Southeastern Australia*. Australian National University Press, Canberra, Australia.

Butler, B.E. & Hutton, J.T. (1956). Parna in the Riverine Plain of South-Eastern Australia and the soils thereon. *Australian Journal of Agricultural Research, 7*, 536-553. https://doi.org/10.1071/AR9560536.

Butler, B.E. (1979). *A Soil Survey of Horticultural Soils in the Murrumbidgee Irrigation Areas, New South Wales*. *A Revised Edition.* CSIRO, Sydney, Bulletin No. 289.

Carré, F., McBratney, A.B., Mayr, T. & Montanarella, L. (2007). Digital soil assessments: Beyond DSM. *Geoderma, 142*(1-2), 69-79. https://doi.org/10.1016/j.geoderma.2007.08.015.

Cattle, S.R. & Field, D.J. (2013). A review of the soil science research legacy of the triumvirate of cotton CRC. *Crop and Pasture Science, 64*(1), 1076-1094. https://doi.org/10.1071/CP13223.

Cattle, S.R., McTainsh, G.H. & Elias, S. (2009). Æolian dust deposition rates, particle-sizes and contributions to soils along a transect in semi-arid New South Wales, Australia. *Sedimentology, 56*(1), 765-783. https://doi.org/10.1111/j.1365-3091.2008.00996.x.

Cattle, S.R. & Smith, C.M.S. (2018). Fabric of soil derived from parna and the riddle of transported pellets. *Soil Research, 56*(3), 219-234. https://doi.org/10.1071/SR16343.

Cay, E., Cattle, S. (2005). The effects of landforming on soil profile characteristics of an irrigated cotton-producing area of southeastern Australia. *Soil and Tillage Research, 84*(1), 76-86. https://doi.org/10.1016/j.still.2004.09.019.

Chen, S., Arrouays, D., Mulder, V.L., Poggio, L., Minasny, B., Roudier, P., Libohova, Z., Lagacherie, P., Shi, Z., Hannam, J., Meersmans, J., Richer-de-Forges, A.C. & Walter, C. (2022). Digital mapping of *GlobalSoilMap* soil properties at a broad scale: A review. *Geoderma, 409*, Article 115567. https://doi.org/10.1016/j.geoderma.2021.115567.

Chen, Y., Gao, S., Jones, E.J. & Singh, B. (2021). Prediction of Soil Clay Content and Cation Exchange Capacity Using Visible Near-Infrared Spectroscopy, Portable X-ray Fluorescence, and X-ray Diffraction Techniques. Environmental Science and Technology, 55(8), 4629-4637. https://doi.org/10.1021/acs.est.0c04130.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY, USA: ACM. https://doi.org/10.1145/2939672.2939785.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y., Li, Y. & Yuan, J. (2024). xgboost: Extreme Gradient Boosting. R package version 1.7.7.1. https://doi.org/10.32614/CRAN.package.xgboost.

Chinchmalapure, S., Deshpande, D. & Shinha, A. (2023). Enhancing accuracy of machine learning model in digital soil mapping. In H. Shankar, P. Thangaraj & K. Mohama Sundaram (Eds.), Third International Conference on Advances in Physical Sciences and Materials: ICAPSM 2022, Volume 2901(1) (050010). https://doi.org/10.1063/5.0178821.

Churchward, H.M. (1956). *The soils of the East Murrakool district, New South Wales, and their relation to land use under irrigation*. CSIRO, Division of Soils, Melbourne.

Churchward, H.M. (1958). *The soils and land use of the Denimein Irrigation District, New South Wales*. CSIRO, Division of Soils, Melbourne.

Churchward, H.M. & Flint, S.F. (1956). *Jenargo Extension of the Berriquin Irrigation District, NSW.* CSIRO, Australia Soils and Land Use Series No. 18.

Cockroft, B. & Olsson. K. A., (2000). Degradation of soil structure due to coalescence of aggregates in no-till, no-traffic beds in irrigated crops. *Australian Journal of Soil Research, 38*(1), 61–70. https://doi.org/10.1071/SR99079.

Conaty, W. C., Broughton, K. J., Egan, L. M., Li, X., Li, Z., Liu, S., Llewellyn, D. J., MacMillan, C. P., Moncuquet, P., Rolland, V., Ross, B., Sargent, D., Zhu, Q. H., Pettolino, F. A., & Stiller, W. N. (2022). Cotton Breeding in Australia: Meeting the Challenges of the 21st Century. Frontiers in plant science, 13, Article 904131. https://doi.org/10.3389/fpls.2022.904131.

Condon, J. (2019). Effective soil sampling – high and low cost options to gain soil fertility information for management. *GRDC Update papers*, February 2019.

Constable, G.A. & Bange, M.P. (2015). The yield potential of cotton (*Gossypium hirsutum* L.). *Field Crops Research, 182*(1), 98-106. https://doi.org/10.1016/j.fcr.2015.07.017.

Conyers, M.K., Hume, I., Summerell, G., Slinger, D., Mitchell, M. & Cawley, R. (2008). The ionic composition of the streams of the mid-Murrumbidgee River: Implications for the management of downstream salinity. *Agricultural Water Management, 95*(5)*,* 598-606. https://doi.org/10.1016/j.agwat.2008.01.007.

Conceição, L.A., Silva, L., Valero, C., Loures, L. & Maçãs, B. (2024) Delineation of Soil Management Zones and Validation through the Vigour of a Fodder Crop. *AgriEngineering. 6*(1), 205-227. https://doi.org/10.3390/agriengineering6010013.

Cook, F. J., Jayawardane, N. S. & Blackwell, J. (1989). Effects of amelioration on infiltration characteristics of a transitional red-brown earth. *New Zealand Journal of Crop and Horticultural Science, 17*(2), 183-188. https://doi.org/10.1080/01140671.1989.10428029.

Cotton Research and Development Corporation [CRDC] & CottonInfo (2023). Australian Cotton Production Manual. CRDC. Narrabri, NSW.

Crocker, R. L. (1946*). Post-Miocene climatic and geologic history and its significance in relation to the genesis of the major soil types of South Australia*. Council for Scientific and Industrial Research, Bulletin 193, 5–65.

CSIRO (2024). ANSIS – Australian National Soil Information System. https://www.csiro.au/en/research/natural-environment/land/soil/ansis. Accessed 30 May 2024.

CSIRO Publishing & The National Committee on Soil and Terrain [NCST] (Australia). (2009). Australian soil and land survey field handbook. Collingwood, Victoria. CSIRO Publishing.

Cox, M.S., Gerard, P.D. & Reynolds, D.B. (2005). Selected Soil Property Variability and Their Relationships With Cotton Yield. Soil Science, 170(11), 928-937. http://dx.doi.org/10.1097/01.ss.0000196766.67036.81.

Dagar, J.C., Rai, A.K., Basak, N. & Yadav, R.K. (2022). Soil Alkalinity/Sodicity: Degradation Processes, Constraints for Crop Production and their Management. In. Y. Dang, N. Menzies & R. Dalal (Eds), *Soil Constraints on Crop Production* (pp. 116-138). Cambridge Scholars Publishing.

Dang, Y.P., Dalal, R.C., Routley, R., Schwenke, G.D. & Daniells, I. (2006). Subsoil constraints to grain production in the cropping soils of the north-eastern region of Australia: an overview. *Australian Journal of Experimental Agriculture, 46*(1), 19-35. https://doi.org/10.1071/EA04079.

Dang, Y.P., Page, K.L., Dalal, R.C. & Menzies, N.W. (2022). Soil Constraints to Crop Production: An Overview. In. Y. Dang, N. Menzies & R. Dalal (Eds), Soil Constraints on Crop Production (pp. 116-138). Cambridge Scholars Publishing.

Daniells, I. G. (2012). Hardsetting soils: a review. *Soil Research, 50(5*), 349–359. https://doi.org/10.1071/SR11102.

Department of Agriculture, Fisheries and Forestry (2024, May, 22). National Soi Action Plan. https://www.agriculture.gov.au/agriculture-land/farm-food-drought/natural-resources/soils/national-soil-action-plan. Accessed 30 May 2024.

Department of Climate Change, Energy, the Environment and Water (2023). *NSW Landuse 2017 v1.5*. Accessed from The Sharing and Enabling Environmental Data Portal https://datasets.seed.nsw.gov.au/dataset/nsw-landuse-2017-v1p5-f0ed-clone-a95d. doi: 10.25948/6fcf-gc02.

Department of Climate Change, Energy, the Environment and Water [DCCEEW] (2024), *Interim Biogeographic Regionalisation for Australia (IBRA), Version 7 (Regions).* https://datasets.seed.nsw.gov.au/dataset/8e242336-7d10-4630-ae81-e1b6e7464f3c.

Datta, A., Ullah, H., Ferdous, Z., Santiago-Arenas, R. & Attia, A. (2019). Water Management in Cotton. In K. Jabran & B.S. Chauhan (Eds.), Cotton Production (pp. 47-59). https://doi.org/10.1002/9781119385523.ch3.

Devkota, K. P., Yadav, S., Humphreys, E., Kumar, A., Kumar, P., Kumar, V., Malik, R. K., & Srivastava, A. K. (2021). Land gradient and configuration effects on yield, irrigation amount and irrigation water productivity in rice-wheat and maize-wheat cropping systems in Eastern India. *Agricultural Water Management, 255*, Article 107036. https://doi.org/10.1016/j.agwat.2021.107036.

Dodd, K., Guppy, C.N., Lockwood, P.V. & Rochester, I.J. (2010). The effect of sodicity on cotton: plant response to solutions containing high sodium concentrations. *Crop and Pasture Science, 64*(8), 816-824. https://doi.org/10.1071/CP13093.

Dodd, K., Guppy, C.N., Lockwood, P.V. & Rochester, I.J. (2013a). Impact of waterlogging on the nutrition of cotton (Gossypium hirsutum L.) produced in sodic soils. *Crop and Pasture Science, 64*(8), 806-815. https://doi.org/10.1071/CP13078.

Dodd, K., Guppy, C.N., Lockwood, P.V. & Rochester, I.J. (2013b). The effect of sodicity on cotton: does soil chemistry or soil physical condition have the greater role? *Plant and Soil, 330*(1), 239-249. https://doi.org/10.1007/s11104-009-0196-6.

DSITI & DNRM (2015). *Guidelines for agricultural land evaluation in Queensland (Second edition*). Queensland Government Department of Science, Information Technology and Innovation and Department of Natural Resources and Mines, Brisbane, Queensland.

Ellili-Bargaoui, Y., Malone, B. P., Michot, D., Minasny, B., Vincent, S., Walter, C. & Lemercier, B. (2020). Comparing three approaches of spatial disaggregation of legacy soil maps based on the Disaggregation and Harmonisation of Soil Map Units Through Resampled Classification Trees (DSMART) algorithm, *SOIL, 6*(1), 371–388, https://doi.org/10.5194/soil-6-371-2020.

Elms, M. K., Green, C. J., & Johnson, P. N. (2001). Variability of Cotton Yield and Quality. *Communications in Soil Science and Plant Analysis, 32*(3–4), 351–368. https://doi.org/10.1081/CSS-100103012.

Esfandiari, M. & Maheshwari, B.L. (2001). SW-Soil and Water: Field Evaluation of Furrow Irrigation Models. *Journal of Agricultural Engineering Research, 79*(4), 459-479. https://doi.org/10.1006/jaer.2001.0717.

Filippi, P. (2014) Monitoring and modelling spatio-temporal soil change in a semi-arid irrigated cotton-growing region of south-west NSW – The impacts of land use and climatic fluctuations. [PhD Thesis, The University of Sydney]. Sydney, NSW, Australia. http://hdl.handle.net/2123/17900.

Filippi, P., Cattle, S., Bishop, T., Pringle, M. & Jones, E. (2018a). Monitoring changes in soil salinity and sodicity to depth, at a decadal scale, in a semiarid irrigated region of Australia. Soil Research, 56(7), 696-711. https://doi.org/10.1071/SR18083.

Filippi, P., Cattle, S., Bishop, T., Odeh, I. & Pringle, M. (2018b). Digital soil monitoring of top- and sub-soil pH with bivariate linear mixed models. *Geoderma, 322*(1), 149-162. https://doi.org/10.1016/j.geoderma.2018.02.033.

Filippi, P., Jones, E.J. & Bishop, T.F.A. (2020). Catchment-scale 3D mapping of depth to soil sodicity constraints through combining public and on-farm soil databases – A potential tool for on-farm management. *Geoderma, 371*, Article 114396. https://doi.org/10.1016/j.geoderma.2020.114396.

Filippi, P., Jones, E.J., Ginns, B.J., Whelan, B.M., Roth, G.W. & Bishop, T.F.A. (2019b) Mapping the Depth-to-Soil pH Constraint, and the Relationship with Cotton and Grain Yield at the Within-Field Scale. *Agronomy, 9*(5), 251. https://doi.org/10.3390/agronomy9050251.

Filippi, P., Jones, E.J., Wimalathunge, N.S., Somarathna, P.D.S.N., Pozza, L.E., Ugbaje, S.U., Jephcott, T.G., Paterson, S.E., Whelan, B.M. & Bishop, T.F.A. (2019a). An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. *Precision Agriculture, 20*(5), 1015-1029. https://doi.org/10.1007/s11119-018-09628-4.

Filippi, P., Minasny, B., Cattle, S. & Bishop, T. (2016). Monitoring and Modelling Soil Change: The Influence of Human Activity and Climatic Shifts on Aspects of Soil Spatiotemporally. In D.L. Sparks (Ed.), *Advances in Agronomy, Volume 139*, (pp. 153-214). London: Academic Press. https://doi.org/10.1016/bs.agron.2016.06.001

Filippi, P., Whelan, B. & Bishop, T. (2024). Proximal and remote sensing – what makes the best farm digital soil maps? *Soil Research, 62*(2), Article SR23112. https://doi.org/10.1071/SR23112.

Fox, J. & Monette, G. (1992). Generalized Collinearity Diagnostics. *Journal of the American Statistical Association, 87*(417), 178–183. https://doi.org/10.2307/2290467.

Fox, J. & Weisberg, S. (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA. https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Fried, A.W. (1993). Late Pleistocene river morphological change, southeastern Australia: the conundrum of sinuous channels during the Last Glacial Maximum. *Palaeogeography, Palaeoclimatology, Palaeoecology, 101*(3-4), 305-316. https://doi.org/10.1016/0031-0182(93)90021-A.

Fuentes, I., Padarian, J., Iwanaga, T. & Vervoort, R.W. (2020). 3D lithological mapping of borehole descriptions using word embeddings. *Computers and Geosciences, 141*(1), 104516. https://doi.org/10.1016/j.cageo.2020.104516.

Gallant, J., & Austin, J. (2012a). Slope derived from 1" SRTM DEM-S. v4. CSIRO. Data Collection. https://doi.org/10.4225/08/5689DA774564A.

Gallant, J., & Austin, J. (2012b). Slope Relief Classification derived from 1" SRTM DEM-S. v3. CSIRO. Data Collection. https://doi.org/10.4225/08/57512079C1A93.

Gallant, J., & Austin, J. (2012c). Topographic Wetness Index derived from 1" SRTM DEM-H. v2. CSIRO. Data Collection. https://doi.org/10.4225/08/57590B59A4A08.

Gallant, J., Dowling, T. & Austin, J. (2012d). Multi-resolution Valley Bottom Flatness (MrVBF). v3. CSIRO. Data Collection. https://doi.org/10.4225/08/5701C885AB4FE.

Gallant, J., & Austin, J. (2012d). Aspect derived from 1" SRTM DEM-S. v6. CSIRO. Data Collection. https://doi.org/10.4225/08/56D778315A62B.

Gallant, J., & Austin, J. (2012e). Topographic Position Index derived from 1" SRTM DEM-S. v6. CSIRO. Data Collection. https://doi.org/10.4225/08/5758CCC862AD5

Gallant, J., & Austin, J. (2013).Multi-resolution Ridge Top Flatness (MrRTF). v2. CSIRO. Data Collection. https://doi.org/10.4225/08/56EA312A5E63B.

Ghosh, S., Lockwood, P., Hulugalle, N., Daniel, H., Kristiansen, P. & Dodd, K. (2010). Changes in Properties of Sodic Australian Vertisols with Application of Organic Waste Products. *Soil Science Society of America Journal, 74*(1), 153-160. https://doi.org/10.2136/sssaj2008.0282.

Glendon, W.G. & Dani, O. (2002). Particle-Size Analysis. In J.H. Dane & G.C. Topp (Eds.), Methods of Soil Analysis: Physical Methods, Part 4, (pp. 272-278). Soil Science Society of America, Madison, Wisconsin, United States of America.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment, 202*(1), 18-27. https://doi.org/10.1016/j.rse.2017.06.031.

Gozukara, G., Akça, E., Dengiz, O & Adak, A. (2022). Soil particle size prediction using Vis-NIR and pXRF spectra in a semiarid agricultural ecosystem in Central Anatolia of Türkiye. *CATENA, 217*, Article 106514. https://doi.org/10.1016/j.catena.2022.106514.

Grace, P., Shcherbak, I., Macdonald, B., Scheer, C. & Rowlings, D. (2016). Emission factors for estimating fertiliser-induced nitrous oxide emissions from clay soils in Australia's irrigated cotton industry. *Soil Research, 54*(5), 598-603. https://doi.org/10.1071/SR16091.

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

Gray, J.M., Bishop, T.F.A. & Wilford, J.R. (2016). Lithology and soil relationships for soil modelling and mapping. CATENA, 147(1), 729-440. https://doi.org/10.1016/j.catena.2016.07.045.

Greene, R. S. B., Cattle, S. R., & McPherson, A. A. (2009). Role of eolian dust deposits in landscape development and soil degradation in southeastern Australia. Australian *Journal of Earth Sciences, 56*(sup1), S55–S65. https://doi.org/10.1080/08120090902871101.

Green, D.L., Petrovic, J., Burrell, M. & Moss, P. (2011). *Water resources and management overview: Murrumbidgee catchment.* NSW Office of Water, Sydney.

Grundy, M.J., Viscarra Rossel, R.A., Searle, R.D., Wilson, P.L., Chen, C. & Gregory, L.J. (2015). Soil and Landscape Grid of Australia. *Soil Research, 53*(8), 835-44. https://doi.org/10.1071/SR15191.

Gupta, R.K. & Abrol, I.P. (1990). Salt affected soils: their reclamation and management for crop production. In R. Lal & B.A. Stewart (Eds), *Advances in Soil Science Volume 11-Soil Degradation* (pp, 223-288). Springer, New York.

Gupta, M., & Hughes, N. (2018). Shift from rice to cotton production in NSW Murrumbidgee region. *Agricultural Commodities, 8*(1), 74–76. https://search.informit.org/doi/10.3316/informit.489033721678180.

Haan, S., Harianto, J., Butterworth, N. & Bishop, T.F.A (2023). Geodata-Harvester: A Python package to jumpstart geospatial data extraction and analysis. Journal of Open Source Software, 8(89), 5205. https://doi.org/10.21105/joss.05205.

Hake, K. & Grimes, D. (2010). Crop Water Management to Optimize Growth and Yield. In J.M. Stewart, D.M.Oosterhuis, J.J. Heitholt, & J.R. Mauney (Eds.), *Physiology of Cotton*. Springer, Dordrecht. https://doi.org/10.1007/978-90-481-3195-2_23.

Hamza, M.A. & Anderson, W.K. (2005). Soil compaction in cropping systems: A review of the nature, causes and possible solutions. *Soil and Tillage Research, 82*(2), 121-145. https://doi.org/10.1016/j.still.2004.08.009.

Han, S.Y., Filippi, P., Singh, K., Whelan, B.M. & Bishop, T.F.A. (2022). Assessment of global, national and regional-level digital soili mapping products at different spatial supports. European Journal of Soil Science, 73(5), e13300. https://doi.org/10.1111/ejss.13300.

Harms, B., Brough, D., Philip, S., Bartley, R., Clifford, D. Thomas, M., Willis, R. & Gregory L. (2015). Digital soil assessment for regional agricultural land evaluation. *Global Food Security, 5*(1), 25-36. https://doi.org/10.1016/j.gfs.2015.04.001.

Hartemink A.E. & Minasny, B. (2014). Towards digital soil morphometrics. *Geoderma, 230-231*(1), 305-317. https://doi.org/10.1016/j.geoderma.2014.03.008.

Harvey, N., Guillaume, J. H. A., Merritt, W., Ticehurst, J., & Thompson, K. (2023). How could managed aquifer recharge be feasible in the Coleambally Irrigation Area? *Australasian Journal of Water Resources, 28*(1), 86–100. https://doi.org/10.1080/13241583.2023.2180830.

Hazelton, P.A. & Murphy, B.W. (2007) Interpreting Soil Test Results: What Do All the Numbers Mean? CSIRO Publishing, Australia. https://doi.org/10.1071/9780643094680.

He, D., Wang, E. (2019). On the rlation between soil water holding capacity and dryland crop productivity. *Geoderma, 353*(1), 11-24. https://doi.org/10.1016/j.geoderma.2019.06.022.

Hengl, T., de Jesus, J.M., Heuvelink, G.B.M., Gonzalez, M.R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevera, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S. & Kempen. (2017a). SoilGrids50m: Global gridded soil indormation based on machine learning. *PLOS ONE, 12*(2), Article e0169748. https://doi.org/10.1371/journal.pone.0169748.

Hengl T., Leenaars J.G.B., Shepherd K.D., Walsh M.G., Heuvelink G.B.M., Mamo T., Tilahun H., Berkhout E., Cooper M., Fegraus E., Wheeler I. & Kwabena N.A. (2017). Soil nutrient maps of Sub-Saharan Africa: assessment of soil nutrient content at 250 m spatial resolution using machine learning. *Nutrient Cycling in Agroecosystems, 109*(1), 77-102. https://doi.org/10.1007/s10705-017-9870-x.

Hills, E. S. (1939). The physiography of north-western Victoria. *Proceedings of the Royal Society of Victoria*, *51*, 297–322. .

Holland J. & Eastwood, C. (2014). *An evaluation of the current understanding of cotton-growing soils and soil management practice issues in Southern NSW (DAN1408)*. NSW Department of Primary Industries. https://www.insidecotton.com/evaluation-current-understanding-cotton-growing-soils-and-soil-management-practice-issues-southern.

Hong, S. Y., Minasny, B., Han, K. H., Kim, Y., & Lee, K. (2013). Predicting and mapping soil available water capacity in Korea. *PeerJ, 1*, Article e71. https://doi.org/10.7717/peerj.71.

Hornbuckle, J. & Christen, E. (1999). *Physical Properties of Soils in the Murrumbidgee and Coleambally Irrigation Areas*. CSIRO Land and Water Griffith, Technical Report 17/99.

Hornbuckle, J., Christen, E., Thacker, J., Muirhead., W. & Stein, T.M. (2008). *Soils of the Murrumbidgee, Coleambally and Murray Irrigation Areas of Australia II: Physical properties*. CSIRO Land and Water Science Report 23/08.

Hornbuckle, J., Car, N., White, B. & Christen, E. (2008). *Soils of the Murrumbidgee, Coleambally and Murray Irrigation Areas of Australia I: A user guide to accessing and identifying soils using digital soil maps and Google Earth™*. CSIRO Land and Water Science Report 22/08.

Hornung, A., Kholsa, R., Reich, R., Inman, D. & Westfall, D.G. (2006). Comparison of Site-Specific Management Zones: Soil-Color-Based and Yield-Based. *Agronomy Journal, 98*(2), 407-415. https://doi.org/10.2134/agronj2005.0240.

Hulugalle, N. R., McCorkell, B. E., Weaver, T. B., & Finlay, L. A. (2010). Managing Sodicity and Exchangeable K in a Dryland Vertisol in Australia with Deep Tillage, Cattle Manure, and Gypsum. *Arid Land Research and Management, 24*(3), 181–195. https://doi.org/10.1080/15324981003741731.

Ibáñez-Asensio, S., Marqués-Mateu, A. & Moreno-Ramón, H. (2013). Statistical relationships between soil colour and soil attributes in semiarid areas. *Biosystems Engineering, 116*(2), 120-129. https://doi.org/10.1016/j.biosystemseng.2013.07.013.

Isbell, R.F. & National Committee on Soil and Terrain (2021). *The Australian Soil Classification*. 3rd. CSIRO Publishing, Melbourne.

Javadi, S. H., Guerrero, A., & Mouazen, A. M. (2022). Clustering and Smoothing Pipeline for Management Zone Delineation Using Proximal and Remote Sensing. Sensors (Basel, Switzerland), 22(2), 645. https://doi.org/10.3390/s22020645.

Jayawardane, N.S. & Chan, K.Y. (1994). The management of soil physical properties limiting crop production in Australian sodic soils - a review. *Australian Journal of Soil Research, 32*(1)*,* 13-44. https://doi.org/10.1071/SR9940013.

Jenny, H. (1994). *Factors of soil formation: a system of quantitative pedology*. Courier Corporation.

Jessop, R.S., Macleod, D.A., Hulme, P.J. & McKenzie, D.C. (1985). The effects of landforming on crop production on a red-brown earth. *Australian Journal of Soil Research, 23*(1), 85-93. https://doi.org/10.1071/SR9850085.

Johnston, E.J. (1952). *Soils of the Deniboota Irrigation District and their classification for Irrigation*. CSIRO Soils and Land Use Series No. 5.

Jones, E.J., Bishop, T.F.A., Malone, B.P., Hulme, P.J., Whelan, B.M. & Filippi, P. (2022). Identifying causes of crop yield variability with interpretive machine learning. *Computers and Electronics in Agriculture, 192*, Article 106632. https://doi.org/10.1016/j.compag.2021.106632.

Kandasamy, J., Sountharanajah, D., Sivabalan, P., Chanan, A., Vigneswaran, S. & Sivapalan, M. (2014). Socio-hydrologic drivers of the pendulum swing between agricultural development and environmental health a case study from Murrumbidgee River basin, Australia. *Hydology and Earth Systems Sciences, 18*(1), 1027-1041. https://doi.org/10.5194/hess-18-1027-2014.

Kasraei, B., Schmidt, M.G., Zhang, J., Bulmer, C.E., Filatow, D.S., Arbor, A., Pennell, T. & Heung B. (2024). A framework for optimizing environmental covariates to support model interpretability in digital soil mapping. *Geoderma, 445*, Article 116873. https://doi.org/10.1016/j.geoderma.2024.116873.

Kemp, R.A. (1985). The cause of redness in some buried and non-buried soils in eastern England. *Journal of Soil Science, 63*(3), 329-334. http://dx.doi.org/10.1111/j.1365-2389.1985.tb00339.x.

Khaledian, Y. & Miller, B.A. (2020). Selecting appropriate machine learning methods for digital soil mapping. *Applied Mathematical Modelling, 81I*(1), 401-418. https://doi.org/10.1016/j.apm.2019.12.016.

Khan, S., Best, L. & Wang, B. (2002). *Surface-Groundwater Interaction Model of the Murrumbidgee Irrigation Area (Development of the Hydrogeological Database).* CSIRO Land and Water, Griffith NSW, Technical Report 36/02.

Khan, M.A., Wahid, A., Ahmad, M., Tahir, M.T., Ahmed, M., Ahmad., S. & Hasanuzzaman, M. (2020). World Cotton Production and Consumption: An Overview. In S. Ahmad & M. Hasanuzzaman (Eds.) *Cotton Production and Uses* (pp. 1-7). Springer, Singapore. https://doi.org/10.1007/978-981-15-1472-2_1.

Kidd, D., Searle, R., Grundy, M., McBratney, A., Robinson, N., O'Brien, L., Zund, P., Arrouays, D., Thomas, M., Padarian, J., Jones, E., Bennett, J.M., Minasny, B., Holmes, K., Malone, B.P., Liddicoat, C., Meier, E.A., Stockman, U., Wilson, P., … Triantafilis, J. (2020). Operationalising digital soil mapping – Lessons from Australia. *Geoderma Regional, 23*0, Article e00335. https://doi.org/10.1016/j.geodrs.2020.e00335.

Kidd, D., Webb, M., Malone, B., Minasny, B. & McBratney, A. (2015). Digital soil assessment of agricultural suitability, versatility and capital in Tasmania, Australia. *Geoderma Regional, 6*(1), 7-21. https://doi.org/10.1016/j.geodrs.2015.08.005.

Klingebiel, A. A., & Montgomery, P. H. (1961). *Land-capability classification (No. 210)*. Soil Conservation Service, US Department of Agriculture.

Lagacherie, P. (2008). Digital Soil Mapping: A State of the Art. In A.E. Hartemink, A.B. McBratney, & M.D. Mendonça-Santos (Eds.) *Digital Soil Mapping with Limited Data* (pp. 3-14). Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-8592-5_1.

Lang, A.R.G. & Hicks, H.M.R. (1975). *Bibliography on soils of the N.S.W. Riverina.* CSIRO, Division of Irrigation Research.

Langford-Smith, T. (1960). Deposition on the Riverine Plain of southeastern Australia. *Geographical Review, 50(3)*, 368-389. https://doi.org/10.2307/212281.

Lawes R., Hochman Z., Jakku E., Butler R., Chai J., Chen Y., Waldner F., Mata G. & Donohue R. (2023). Graincast™: monitoring crop production across the Australian grainbelt. *Crop & Pasture Science 74*(1), 509-523. https://doi.org/10.1071/CP21386.\

Lawrence, P.G., Roper, W., Morris, T.F. & Guillard, K. (2020). Guiding soil sampling strategies using classical and spatial statistics: A review. Agronomy Journal, 112(1), 493-510. https://doi.org/10.1002/agj2.20048.

Leizica, E., Buss, M.E.F. & Noellemeyer, E. (2022). Geomorphology as a tool to digitize homogeneous management zones based on soil properties in the semiarid central Argentinian Pampas. *Geoderma Regional, 28*, Article e00458. https://doi.org/10.1016/j.geodrs.2021.e00458.

Lemercier, B., Lagacherie, P., Amelin, J., Sauter, J., Pichelin, P., Richer-de-Forges, AC. & Arrouays, D. (2022). Multiscale evaluations of global, national and regional digital soil mapping products in France. *Geoderma, 425,* Article 116052. https://doi.org/10.1016/j.geoderma.2022.116052.

Leo, S. (2022). Potential of Remote and Proximal Sensing, Publicly Available Datasets and Machine Learning for Site-Specific Management in Australian Irrigated Cotton Systems. [PhD Thesis, Queensland University of Technology]. Brisbane, QLD, Australia. https://eprints.qut.edu.au/235383/.

Leo, S., Migliorati, M.D.A., Nguyen, T.H. & Grace, P.R. (2023). Combining remote sensing-derived management zones and an auto-calibrated crop simulation model to determine optimal nitrogen fertilizer rates. *Agricultural Systems, 205*, Article 103559. https://doi.org/10.1016/j.agsy.2022.103559.

Lesturgez, G., Poss, R., Noble, A., Grünberger, O., Chintachao, W. & Tessier, D. (2006). Soil acidification without pH drop under intensive cropping systems in Northeast Thailand. *Agriculture, Ecosystems and Environment, 114*(2-4), 239-248. https://doi.org/10.1016/j.agee.2005.10.020.

Liaw, A. & Wiener, M. (2002). Classification and Regression by randomForest. *R News, 2*(3), 18-22. https://doi.org/10.32614/CRAN.package.randomForest.

Lin, L.I. (1989) A Concordance Correlation Coefficient to Evaluate Reproducibility. Biometrics, 45(1), 255-268. http://dx.doi.org/10.2307/2532051.

Liu, Y. & Just, A. (2023). SHAPforxgboost: SHAP Plots for 'XGBoost'. R package version 0.1.3. https://doi.org/10.32614/CRAN.package.SHAPforxgboost.

Ma, G., Ding, J., Han, L., Zhang, Z. & Ran, S. (2021). Digital mapping of soil salinization based on Sentinel-1 and Sentinel-2 data combined with machine learning algorithms. *Regional Sustainability, 2*(2), 177-188. https://doi.org/10.1016/j.regsus.2021.06.001.

Ma, Y., Minasny, B. (2023). Mapping and Modelling Soil Constraints. In N.S. Bolan & M. B. Kirkham (Eds.), *Soil Constraints and Productivity,* (pp. 409-418). Boca Raton: CRC Press.

Ma, Y., Minasny, B., Malone, B.P. & McBratney, A.B. (2019). Pedology and digital soil mapping (DSM). *European Journal of Soil Science, 70*(1), 216-235. https://doi.org/10.1111/ejss.12790.

Macdonald, B.C.T., Chang, Y.F., Nadelko, A., Tuomi, S. & Glover, M. (2016). Tracking fertiliser and soil nitrogen in irrigated cotton: uptake, losses and the soil N stock. *Soil Research, 55*(3), 264-272. https://doi.org/10.1071/SR16167.

Macdonald, B.C.T., Latimer, J.O., Schwenke, G.D., Nachimutu, G. & Baird, J.C. (2018). The current status of nitrogen fertiliser use efficiency and future research directions for the Australian cotton industry. *Journal of Cotton Research, 1*, Article 15. https://doi.org/10.1186/s42397-018-0015-9.

Mahmood, H.S., Hoogmoed, W.B. & Van Henten, E.J. (2013). Proximal Gamma-Ray Spectroscopy to Predict Soil Properties Using Windows and Full-Spectrum Analysis Methods. *Sensors 13I*(1), 16263-16280. https://doi.org/10.3390/s131216263.

Maino, A., Alberi, M., Anceschi, E., Chiarelli, E., Cicala, L., Colonna, T., De Cesare, M., Guastaldi, E., Lopane, N., Mantovani, F., Marcialis, M., Martini, N., Montuschi, M., Piccioli, S., Raptis, K.G.S., Russo, A., Semenza, F. & Strati, V. (2022). Airborne Radiometric Surveys and Machine Learning Algorithms for Revealing Soil Texture. *Remote Sensing, 14*(5), 3814. https://doi.org/10.3390/rs14153814.

Mallarino, A.P. & Wittry, D.J. (2004). Efficacy of Grid and Zone Soil Sampling Approaches for Site-Specific Assessment of Phosphorus, Potassium, pH, and Organic Matter. *Precision Agriculture 5*(1), 131–144 https://doi.org/10.1023/B:PRAG.0000022358.24102.1b.

Malone, B. P., Kidd, D. B., Minasny, B., & McBratney, A. B. (2015). Taking account of uncertainties in digital land suitability assessment. *PeerJ, 3*, Article e1366. https://doi.org/10.7717/peerj.1366.

Malone, B.P., McBratney, A.B., Minasny, B. & Laslett, G.M. (2009). Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma, 154*(1-2), 138-152. https://doi.org/10.1016/j.geoderma.2009.10.007.

Malone, B.P., McBratney, A.B., Minasny, B. & Wheeler, I. (2012). A general method for downscaling earth resource information. *Computers & Geosciences, 41*(1), 119-125. https://doi.org/10.1016/j.cageo.2011.08.021.

Malone, B.P., Minasny, B., Odgers, N.P. & McBratney, A.B. (2014). Using model averaging to combine soil property rasters from legacy soil maps and from point data. *Geoderma 232-234*, 34-44. https://doi.org/10.1016/j.geoderma.2014.04.033.

Malone, B.P., Minasny, B. & McBratney, A.B. (2017a). Using R for Digital Soil Mapping. Progress in Soil Science. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-319-44327-0.

Malone, B.P., Styc, Q., Minasny, B. & McBratney, A.B. (2017b). Digital soil mapping of soil carbon at the farm scale: A spatial downscaling approach in consideration of measured and uncertain data. *Geoderma, 290(*1), 91-99. https://doi.org/10.1016/j.geoderma.2016.12.008.

McBratney, A., Minasny, B., Wheeler, I., Malone, B. & van der Linden, D. (2012). Frameworks for digital soil assessment. In B. Minasny, B.P. Malone & A.B. McBratney (Eds.) *Digital Soil Assessments and Beyond* (pp. 9-14). CRC Press.

McBratney, A. B., Santos, M. M., & Minasny, B. (2003). On digital soil mapping. *Geoderma, 117*(1-2), 3-52. https://doi.org/10.1016/S0016-7061(03)00223-4.

McCarthy, A., Foley, J., Raedts, P. & Hills, J. (2023). Field evaluation of automated site-specific irrigation for cotton and perennial ryegrass using soil-water sensors and Model Predictive Control. *Agricultural Water Management, 277*, Article 108098. https://doi.org/10.1016/j.agwat.2022.108098.

McIntyre, D.S. (1976). Subplasticity in Australian Soils. I* Description, Occurrence, and some Properties. *Australian Journal of Soil Research, 14*(1), 227-236. .

McKenzie, D.C. (1998). *SOILpak For cotton growers* (3rd ed.). NSW Agriculture. Orange.

Meier, M., Souza, E. D., Francelino, M. R., Fernandes Filho, E. I. & Schaefer, C. E. G. R. (2018). Digital soil mapping using machine learning algorithms in a tropical mountainous area. *Revista Brasileira de Ciência do Solo, 42*, Article e0170421. https://doi.org/10.1590/18069657rbcs20170421.

Meyer, H. & Pebesma, E. (2021). Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods in Ecology and Evolution, 12*(9), 1620-1633. https://doi.org/10.1111/2041-210X.13650.

Miao, Q., Gonçalves, J.M., Li, R., Gonçalves, D., Levita, T. & Shi, H. (2021). Assessment of Precise Land Levelling on Surface Irrigation Development. Impacts on Maize Water

Productivity and Economics. *Sustainability 2021, 13*(3), 1191. https://doi.org/10.3390/su13031191.

Minasny, B. & McBratney, A.B. (2016). Digital soil mapping: A brief history and some lessons. *Geoderma, 264(*B), 301-311. https://doi.org/10.1016/j.geoderma.2015.07.017.

Minty, B., Franklin, R., Milligan, P., Richardson, M. & Wilford, J. (2009). The radiometric map of Australia. Exploration Geophysics, 40(4), 325-333. https://doi.org/10.1071/EG09025.

Mitchell, P. (2002). *Descriptions for NSW (Mitchell) Landscapes Version 2 (2002)*. Department of Environment and Climate Change NSW.

Molnar, C. (2023). *Interpreting Machine Learning Models With SHAP: A Guide With Python Examples And Theory On SHAP Values*. Leanpub. .

Morgan, G. & Terrey, J. (1992). *Nature conservation in western NSW.* National Parks Association, Sydney. .

Mueller, D., Jacobs, Z., Cohen, T.J., Price, D.M., Reinfelds, I.V. & Shulmeister, J. (2018). Revisiting an arid LGM using fluvial archives: a luminescence chronology for palaeochannels of the Murrumbidgee River, south-eastern Australila. *Journal of Quaternary Science, 33*(7), 777-793. https://doi.org/10.1002/jqs.3059.

Mullins, C.E., MacLeod, D.A., Northcote, K.H., Tisdall, J.M. & Young, I.M. (1990). Hardsetting Soils: Behaviour, Occurrence, and Management. In R. Lal & B.A. Stewart (Eds.) *Advances in Soil Science, Volume 11*. Springer, New York. https://doi.org/10.1007/978-1-4612-3322-0_2.

Munns, R. (2002). Comparative physiology of salt and water stress. *Plant, Cell and Environment, 25*(2), 239-250. https://doi.org/10.1046/j.0016-8025.2001.00808.x.

Munsell Color (Firm). (2010). *Munsell soil color charts: with genuine Munsell color chips*. Grand Rapids, MI: Munsell Color,

Murphy, B.W., Eldridge, D.J., McKane, D.J. & Gray, J.M. (2000). Soil Landscapes of New South Wales. In P.E.V. Charman & P.W. Murphy, (Eds*.), Soils: Their properties and management* (2nd ed.) (pp. 150-165). Oxford University Press.

Murray Darling Basin Authority [MDBA] (2024). https://www.mdba.gov.au/, accessed 16 May 2024.

Murrumbidgee Irrigation [MI] (2010). Submission to the IPART review of prices for NSW water administration ministerial corporation (WAMC). NSW Office of Water.

Nachimuthu, G., Palmer, B., Hundt, A., Schwenke, G., Jamali, H., Knox, O. & Guppy, C. (2024). Soil property differences and irrigated-cotton lint yield—Cause and effect? An on-farm case study across three cotton-growing regions in Australia. *Soil Use and Management, 40*(2), Article e13065. https://doi.org/10.1111/sum.13065.

Nawar, S., Corstanje, R., Halcro, G., Mulla, D. & Mouazen, A.M. (2017). Chapter Four - Delineation of Soil Management Zones for Variable-Rate Fertilization: A Review. In D.L. Sparks (Ed.), *Advances in Agronomy* (pp.175-245). https://doi.org/10.1016/bs.agron.2017.01.003.

Neupane, J. & Guo, W. (2019). Agronomic Basis and Strategies for Precision Water Management: A Review. *Agronomy, 9*(2), 87. https://doi.org/10.3390/agronomy9020087.

Neupane, J., Guo, W., West, C.P., Chang, F. & Lin, Z. (2021). Effects of irrigation rates on cotton yield as affected by soil physical properties and topography in the southern high plains. *PLoS One, 16*(10), Article e0258496. https://doi.org/10.1371/journal.pone.0258496.

Norrish, K. & Tiller, K.G. (1976). Subplasticity in Australian Soils. V* Factors Involved and Techniques of Dispersion. *Australian Journal of Soil Research, 14*(1), 273-289. .

North, S. & Schultz, A. (2017). *Match Irrigation Systems to Soil Type Riverine Plains South-East Australia*. NSW Department of Primary Industries, Deniliquin.

NSW Department of Climate Change, Energy, the Environment and Water [DCCEEW] (2023). NSW Landuse 2017 v1.5, accessed from The Sharing and Enabling Environmental Data Portal https://datasets.seed.nsw.gov.au/dataset/nsw-landuse-2017-v1p5-f0ed-clone-a95d, Accessed 24-06-12. https://doi.org/10.25948/6fcf-gc02.

Office of Environment & Heritage [OEH] (2017) Digital soil mapping of key soil properties over NSW. Technical Report, NSW OEH, Sydney, NSW. http://www.environment.nsw.gov.au/Research-and-publications/Publications-search/Digital-soil-mapping-of-key-soil-properties-over-NSW, accessed 29 June 2020.

Olley, J. & Scott, A. (2002). *Sediment supply and transport in the Murrumbidgee and Namoi Rivers since European settlement*. CSIRO Land and Water, Canberra, Technical Report 9/02.

Onus, A., Cattle, S. & Odeh, I. (2003). How do Lachlan Valley cotton soils compare to northern NSW? *Australian Cottongrower 24*(1), 28–30. .

Orton, T.G., Mallawaarachchi, T., Pringle, M.J., Menzies, N.W., Dalal, R.C., Kopittke, P.M., Searle, R., Hochman, Z. & Dang, Y.P. (2018). Quantifying the economic impact of soil constraints on Australian agriculture: A case-study of wheat. *Land Degradation and Development, 29*(11), 3866-3875. https://doi.org/10.1002/ldr.3130.

Padarian Campusano, J.S. (2014) *Provision of soil information for biophysical modelling*. [Masters Thesis, The University of Sydney]. Sydney, NSW, Australia. http://hdl.handle.net/2123/12278.

Padarian, J.S., McBratney, A.B. & Minasny, B. (2020). Game theory interpretation of digital soil mapping convolutional neural networks. SOIL, 6(1), 389-397. https://doi.org/10.5194/soil-6-389-2020.

Pain, C. Gregory, L., Wilson, P., McKenzie, N (2011). T*he physiographic regions of Australia. Explanatory notes*. Australia: Australian Collaborative Land Evaluation Program. http://hdl.handle.net/102.100.100/104103?index=1.

Page, K.J., Dare-Edwards, A.J., Owens, J.W., Frazier, P.S., Kellett, J. & Price, D.M. (2001). TL chronology and stratigraphy of riverine source bordering sand dunes near Wagga Wagga, New South Wales, Australia. *Quaternary International, 83-85, 187-193. https://doi.org/10.1016/S1040-6182(01)00039-8.*

Page, K.J., Kemp, J. & Nanson, G.C. (2009). Late Quaternary evolution of the Riverine Plain palaeochannels, southeastern Australia. *Australian Journal of Earth Sciences, S19-S33. https://doi.org/10.1080/08120090902870772.*

Page, K.J. & Nanson, G.C. (1996). Stratigraphic architecture resulting from Late Quaternary evolution of the Riverine Plain, south-eastern Australia. *Sedimentology, 43*. 927-945. https://doi.org/10.1111/j.1365-3091.1996.tb01512.x.

Page, K.J., Nanson, G.C. & Price, D. (1996). Chronology of Murrumbidgee River palaeochannels on the Riverine Plain, southeastern Australia. *Journal of Quaternary Science, 11*(4), 311-326. https://doi.org/10.1002/(SICI)1099-1417(199607/08)11:4%3C311::AID-JQS256%3E3.0.CO;2-1.

Page, D., Vanderzalm, J., Gonzalez, D., Bennett, J & Castellazzi, P. (2023). Managed aquifer recharge for agriculture in Australia – History, success factors and future implementation. *Agricultural Water Management, 285*, Article 108382. https://doi.org/10.1016/j.agwat.2023.108382.

Pell, S.D., Chivas, A.R. & Williams, I.S. (2001). The Mallee Dunefield: development and sand provenance. *Journal of Arid Environments, 48*(2), 149-170. https://doi.org/10.1006/jare.2000.0751.

Pels, S. (1964). The present and ancestral Murray River system. *Australian Geographic Studies, 2*(1), 111-119. https://doi.org/10.1111/j.1467-8470.1964.tb00029.x.

Pels, S. (1966). Late quaternary chronology of the riverine plain of southeastern Australia. *Journal of the Geological Society of Australia, 13*(1), 111-119. https://doi.org/10.1080/00167616608728604.

Pels, S. (1968). Geology and Ground Water Hydrology. In *Environmental Studies of the Coleambally Irrigation Area and Surrounding Districts, Bulletin No. 2, Land Use Series* (pp. 3-26). Water Conservation and Irrigation Commission, New South Wales.

Pels, S. (1971). River systems and climatic changes in southeastern Australia. In D.J. Mulvaney & J. Golson (Eds.), *Aboriginal Man and Environment in Australia* (pp. 38-46). Australian National University Press, Canberra.

Piedallu, C., Pedersoli, E., Chate, E., Morneau, F., Seynave, I. & Gégout, J.-C. (2022). Optimal resolution of soil properties maps varies according to their geographical extent and location. *Geoderma, 412*, Article 115723. https://doi.org/10.1016/j.geoderma.2022.115723.

Piikki, K., Wetterlind, J., Söderström, M. & Stenberg, B. (2020). Perspectives on validation in digital soil mapping of continuous attributes—A review. *Soil Use and Management, 37*(1), 7-21. https://doi.org/10.1111/sum.12694.

Ping, J.L., Green, C.J., Bronson, K.F., Zartman, R.E. & Dobermann, A. (2004). Identification of Relationships between Cotton Yield, Quality, and Soil Properties. *Agronomy Journal, 96*(6), 1588-1597. https://doi.org/10.2134/agronj2004.1588.

Ping, J. L., Green, C. J., Zartman, R. E., Bronson, K. F., & Morris, T. F. (2007). Spatial Variability of Soil Properties, Cotton Yield, and Quality in a Production Field. *Communications in Soil Science and Plant Analysis, 39*(1–2), 1–16. https://doi.org/10.1080/00103620701758840.

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E. & Rossiter, D. (2021). SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty, *SOIL, 7*(1), 217–240. https://doi.org/10.5194/soil-7-217-2021.

Pozza, L.E., Filippi, P., Whelan, B., Wimalathunge, N.S., Jones, E.J. & Bishop, T.F.A. (2022). Depth to sodicity constraint mapping of the Murray-Darling Basin, Australia. *Geoderma, 428*(1), 116181. https://doi.org/10.1016/j.geoderma.2022.116181.

Prathapar, S.A., Lawson, S. & Enever, D.J. (1997). *Hydrogeology of the Coleambally Irrigation Area: A brief description for use with a groundwater simulation model*. CSIRO, Land and Water, Technical Report 3/97.

Pucillo, K. (2005). *Quaternary paleochannel evolution and groundwater movement in the Coleambally Irrigation District of New South Wales.* [Doctoral thesis, University of Wollongong]. Australia.

QGIS.org (2023). QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.org.

Qu, l., Lu, H., Tian, Z., Schoorl, J.M., Huang, B., Liang, Y., Qiu, D. & Liang, Y. (2024). Spatial prediction of soil sand content at various sampling density based on geostatistical and machine learning algorithms in plain areas. *CATENA, 234*, Article 107572. https://doi.org/10.1016/j.catena.2023.107572.

R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ramcharan A., Hengl T., Nauman T., Brungard C., Waltman S., Wills S. & Thompson J. (2018). Soil property and class maps of the conterminous United States at 100-meter spatial resolution. ISoiil Science Society of America Journal, 82*9*(1), 186-201. https://doi.org/10.2136/sssaj2017.04.0122.

Rengasamy, P. & Olsson, K.A. (1993). Irrigation and Sodicity. *Australian Journal of Soil Research, 31*(6), 821-837. https://doi.org/10.1071/SR9930821.

Rayment, G.E. & Lyons, D.J. (2011). Soil Chemical methods: Australasia. CSIRO Publishing, Melbourne, Australia.

Rentschler, T., Gries, P., Behrens, T., Bruelheide, H., Kühn, P., Seitz, S., Shi, X., Trogisch, S., Scholten, T. & Schmidt, K. (2019). Comparison of catchment scale 3D and 2.5D modelling of soil organic carbon stocks in Jiangxi Province, *PR China. PloS one, 14*(8), Article e0220881. https://doi.org/10.1371/journal.pone.0220881.

Richer-de-Forges, A.C., Chen, Q., Baghdadi, N., Chen, S.; Gomez, C., Jacquemoud, S., Martelet, G., Mulder, V.L., Urbina-Salazar, D., Vaudour, E., Weiss, M., Wigneron, J.P. & Arrouays, D. (2023). Remote Sensing Data for Digital Soil Mapping in French Research—A Review. *Remote Sensing, 15*(12), 3070. https://doi.org/10.3390/rs15123070.

Rifai, S. W., De Kauwe, M. G., Gallagher, R. V., Cernusak, L. A., Meir, P., & Pitman, A. J. (2024). Burn severity and post-fire weather are key to predicting time-to-recover from Australian forest fires. *Earth's Future, 12*, Article .e2023EF003780. https://doi.org/10.1029/2023EF003780.

Roberts, D., Wilford, J. & Ghattas, O. (2019). Exposed soil and mineral map of the Australian continent revealing the land at its barest. Natural Communications, 10(1), 5297. https://doi.org/10.1038/s41467-019-13276-1.

Roth, G., Harris, G., Gillies, M., Montgomery, J. & Wigginton, D. (2013). Water-use efficiency and productivity trends in Australian irrigated cotton: a review. *Crop and Pasture Science, 64*(12), 1033-1048. https://doi.org/10.1071/CP13315.

Roudier, P., Burge, O.R., Richardson, S.J., McCarthy, J.K., Grealish, G. & Ausseil, A.-G. (2020). National Scale 3D Mapping of Soil pH Using a Data Augmentation Approach. *Remote Sensing, 12*(18), 2872. https://doi.org/10.3390/rs12182872.

Sánchez-Marañón, M., García, P.A., Huertas, R., Hernández-Andrés, J. & Melgosa, M. (2011). Influence of Natural Daylight on Soil Color Description: Assessment Using a Color-Appearance Model. *Soil Science Society of America Journal, 75*(3), 984-993. https://doi.org/10.2136/sssaj2010.0336.

Scheer, C., Rowlings, D.W., Antille, D.L, Migliorati, M.D.A, Fuchs, K. & Grace, P. (2023). Improving nitrogen use efficiency in irrigated cotton production. *Nutrient Cycling in Agroecosystems, 125*(1), 95–106 https://doi.org/10.1007/s10705-022-10204-6.

Schumm, S.A. (1968). River adjustment to altered hydrologic regimen- Murrumbidgee River and palaeochannels, Australia, Geological Survey Professional Paper 598.

Searle, R., McBratney, A.B., Grundy, M., Kidd, D., Malone, B., Arrouays, D., Stockman, U., Zund, P., Wilson, P., Wilford, J., Van Gool, D., Triantafilils, J., Thomas, M., Stower, L., Slater, B., Robinson, N., Ringrose-Voase, A., Padarian, J., Payne, J., … Andrews, K. (2021). Digital soil mapping and assessment for Australia and beyond, A propitious future. *Geoderma Regional, 24*, Article e00359. https://doi.org/10.1016/j.geodrs.2021.e00359.

Shapley, L. (1953). A Value for n-Person Games. In H. Kuhn & A. Tucker (Eds.), *Contributions to the Theory of Games II* (pp.307-317). Princeton University Press, Princeton, 307-317. https://doi.org/10.1515/9781400881970-018.

Sharif, I., Aleem, S., Farooq, J., Rizwan, M., Younas, A., Sarwar, G., & Chohan, S. M. (2019). Salinity stress in cotton: effects, mechanism of tolerance and its management strategies. *Physiology and molecular biology of plants, 25*(4), 807–820. https://doi.org/10.1007/s12298-019-00676-2.

Sharififar, A., Sarmadian, F., Malone, B.P. & Minasny, B. (2019). Addressing the issue of digital mapping of soil classes with imbalanced class observations. *Geoderma, 150*(1), 84-92. https://doi.org/10.1016/j.geoderma.2019.05.016.

Sharma, D.K., Kumar, A. & Singh, K.N. (1990). Effect of irrigation scheduling on growth, yield and evapotranspiration of wheat in sodic soils. *Agricultural Water Management, 18*(3), 267-276. https://doi.org/10.1016/0378-3774(90)90048-4.

Shirazi, F.R.A., Shahbazi, F., Rezaei, H. & Biswas, A. (2024). Multi-property digital soil mapping at 30-m spatial resolution down to 1 m using extreme gradient boosting tree model and environmental covariates. *Remote Sensing Applications: Society and Environment, 33*, Article 101123. https://doi.org/10.1016/j.rsase.2023.101123.

Smith, R. (1945). *Soils of the Berriquin Irrigation District, N.S.W.* Council for Scientific and Industrial Research Melbourne.

Smith R., Herriot, R.I., Johnston, E.J. & Crocker, R.L. (1941). *Soil map, Parishes of Moulamein South, Burbagadah and Boyd, County of Wakool, New South Wales*. Council for Scientific and Industrial Research Australia, Division of Soils Adelaide.

Snyder, C.S., Bruulsema, T.W., Jensen, T.L. & Fixen, P.E. (2009). Review of greenhouse gas emissions from crop production systems and fertilizer management effects. *Agriculture, Ecosystems & Environment, 133*(3-4), 247-266. https://doi.org/10.1016/j.agee.2009.04.021.

Speer, M.S. (2018). Digitization of weekly Murrumbidgee River heights at Hay South Eastern Australia 1873-2017. *Geoscience Data Journal, 5*(1), 9-13. https://doi.org/10.5281/zenodo.1184869.

Stace, H.C.T., Hubble, G.D., Brewer, R., Northcote, K.H., Sleeman, J.R., Mulcahy, M.J. & Hallsworth, E.G. (1969). *A Handbook of Australian Soils*. Rellim Technical Publications, Glenside, S.A. https://doi.org/10.2136/sssaj1969.03615995003300020006x.

Stannard, M.E. (1968). Physiography. In *Environmental Studies of the Coleambally Irrigation Area and Surrounding Districts, Bulletin No. 2, Land Use Series* (pp. 27-33). Water Conservation and Irrigation Commission, New South Wales.

Stannard, M.E. (1970). *Morphological description of soils occurring in the Coleambally Irrigation Area.* Water Conservation and Irrigation Commission, Sydney.

Stephens, C.G. (1953). *A manual of Australian soils.* CSIRO, Melbourne, Australia.

Stern, H., de Hoedt, G. & Ernst, J. (2000). Objective classification of Australian climates. *Australian Meteorological Magazine,* 49(2), 87-96. .

Sui, R. (2018). Irrigation Scheduling Using Soil Moisture Sensors. *Journal of Agricultural Science, 10*(1), 1-11. *https://doi.org/10.5539/jas.v10n1p1*.

Sulaeman, Y., Minasny, B., McBratney, A.B., Sarwani, M. & Sutandi, A. (2013). Harmonizing legacy soil data for digital soil mapping in Indonesia. *Geoderma, 192*(1), 77-85. http://dx.doi.org/10.1016/j.geoderma.2012.08.005.

Talsma, T. (1968). Soil Salinity. In *Environmental Studies of the Coleambally Irrigation Area and Surrounding Districts, Bulletin No. 2, Land Use Series* (pp. 35-48). Water Conservation and Irrigation Commission, New South Wales.

Taylor, J.K. & Hooper, P.D. (1938). *A Soil Survey of Horticultural Soils in the Murrumbidgee Irrigation Areas, New South Wales*. Council for Scientific and Industrial Research. Melbourne, Australia, Bulletin No. 118. https://doi.org/10.25919/gjws-yr06.

Thacker, J., Hornbuckle, J., Christen, E., Muirhead, W. & Stein, T.M. (2008). *Soils of the Murrumbidgee, Coleambally and Murray Irrigation Areas of Australia III: Chemical properties*. CSIRO Land and Water Science Report 24/08.

Thomas, M., Clifford, D., Bartley, R., Philip, S., Brough, D., Gregory, L., Willis, R. & Glover, M. (2015). Putting regional digital soil mapping into practice in Tropical Northern Australia. *Geoderma, 241-242*(1), 145-157. https://doi.org/10.1016/j.geoderma.2014.11.016.

Thomas, M., Gregory, L., Harms, B., Hill, J.V, Morrison, D., Philip, S., Searle, R., Smolinski, H., Van Gool, D., Watson, I., Wilson, P.L. and Wilson, P.R. (2018). *Land suitability of the Fitzroy, Darwin and Mitchell catchments.* A technical report from the CSIRO Northern Australia Water Resource Assessment, part of the National Water Infrastructure Development Fund: Water Resource Assessments, CSIRO, Australia.

Tilse, M.J., Bishop, T.F.A., Triantafilis, J. & Filippi, P. (2022). Mapping the impact of subsoil constraints on soil available water capacity and potential crop yield. *Crop & Pasture Science, 73*(6), 636-651. https://doi.org/10.1071/CP21627.

Triantafilis, J., Ward, W.T. & McBratney, A.B. (2001). Land suitability assessment in the Namoi Valley of Australia, using a continuous model. *Australilan Journal of Soil Research, 39*(2), 273-289. https://doi.org/10.1071/SR99087.

Ulfa, F., Orton, T.G., Dang, Y.P. & Menzies, N.W. (2023). A Study of the Relationships between Depths of Soil Constraints and Remote Sensing Data from Different Stages of the Growing Season. *Remote Sensing, 15*(14), Article 3527. https://doi.org/10.3390/rs15143527.

USDA. (2018). *Australia Cotton and Products Annual Report*. United States Department of Agriculture Global Agricultural Information Network. .

USDA. (2024). *Australia Cotton and Products Annual Report.* United States Department of Agriculture Foreign Agricultural service. .

Van Der Klooster, W., Van Egmond, F.M. & Sonneveld, M.P.W. (2011). Mapping soil clay contents in Dutch marine districts using gamma-ray spectrometry. *European Journal of Soil Science, 62*(5), 743-753. https://doi.org/10.1111/j.1365-2389.2011.01381.x.

Van Dijk, D.C. (1958). *Principles of soil distribution in the Griffith-Yenda District, New South Wales*. CSIRO, Melbourne, Soil and Land Use Series No. 11.

Van Dijk, D.C. (1961). *Soils of the southern portion of the Murrumbidgee Irrigation Areas*. CSIRO, Melbourne, , Soil and Land Use Series No. 40.

Van Dijk, D.C. & Talsma, T. (1964). *Soils of portion of the Coleambally Irrigation Area, N.S.W.* CSIRO, Melbourne, Soil and Land Use Series No. 47.

Viscarra Rossel, R.A., Chen, C., Grundy, M.J., Searle, R., Clifford, D. & Campbell, P.H. (2015). The Australian three-dimensional soil grid: Australia's contribution to the *GlobalSoilMap* project. *Soil Research, 52*(8), 845-864. https://doi.org/10.1071/SR14366.

Viscarra Rossel, R.A. & McBratney, A.B. (1998). Soil chemical analytical accuracy and costs: implications from precision agriculture. *Australian Journal of Experimental Agriculture, 38*(7), 765-775. https://doi.org/10.1071/EA97158.

Vories, E., O'Shaughnessy, S., Sudduth, K. Evett, S., Andrade, M. & Drummond, S. (2021). Comparison of precision and conventional irrigation management of cotton and impact of soil texture. Precision Agriculture 22(1), 414–431 https://doi.org/10.1007/s11119-020-09741-3.

Wadoux A.M.J.C., Samuel-Rosa, A., Poggio, L. & Mulder, V.A. (2019). A note on knowledge discovery and machine learning in digital soil mapping. *European Journal of Soil Science, 71*(2), 133-136. https://doi.org/10.1111/ejss.12909.

Wadoux A.M.J.C., Malone B., Minasny B., Fajardo M. & McBratney A.B. (2021). Soil Spectral Inference with R: Analysing Digital Soil Spectra using the R Programming Environment. Progress in Soil Science. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-030-64896-1.

Walke, N., Obi Reddy, G.P., Maji, A.K. & Thayalan, S. (2012). GIS-based multicriteria overlay analysis in soil-suitability evaluation for cotton (Gossypium spp.): A case study in the black soil region of Central India. *Computers and Geosciences, 41*(1), 108-118. https://doi.org/10.1016/j.cageo.2011.08.020.

Walker, P.H. & Hutka, J. (1976). Subplasticity in Australian Soils. III* Particle-size Properties of Soil Materials of Varying Plasticity. *Australian Journal of Soil Research, 14*(1), 249-260. .

Wang, D., Wang, Z., Zhang, J., Zhou, B., Lv, T. & Li, W. (2021). Effects of Soil Texture on Soil Leaching and Cotton (Gossypium hirsutum L.) Growth under Combined Irrigation and Drainage. *Water, 13*(24) Article 3614. https://doi.org/10.3390/w13243614.

Wasson, R.J. (1986). Geomorphology and Quaternary environmental changes of the Australian continental dunefields. *Geographical review of Japan, 59*(1), 55-67. https://doi.org/10.4157/grj1984b.59.55.

Whelan, B.M. & McBratney, A.B. (2000). The "Null Hypothesis" of Precision Agriculture Management. *Precision Agriculture, 2*(1), 265-279. https://doi.org/10.1023/A:1011838806489.

Whelan, B.M. & Mcbratney, A. (2003). Definition and interpretation of potential management zones in Australia. In *Australian Society of Agronomy. "Solutions for a better environment". Proceedings of the 11th Australian Agronomy Conference, Geelong, Victoria.*

Zeraatpisheh, M., Bottega, E.L., Bakhshandeh, E., Owliaie, H.R., Taghizadeh-Mehrjardia, R., Kerry, R., Scholten, T. & Xu, M. (2022). Spatial variability of soil quality within management zones: Homogeneity and purity of delineated zones. *CATENA, 209*(Part 1), Article 105835. https://doi.org/10.1016/j.catena.2021.105835.

Zinck, J.A., Metternicht, G., Bocco, G. & Del Valle, H.F. (2016). *Geopedology: An integration of Geomorphology and Pedology for Soil and Landscape Studies.* Springer Cham. https://doi.org/10.1007/978-3-319-19159-1.

Zhang, G., Liu, F. & Song, X. (2017). Recent progress and future prospect of digital soil mapping: A review. Journal of Integrative Agriculture, 16(12), 2871-2855. https://doi.org/10.1016/S2095-3119(17)61762-3.

Zhang, Y., Ji, W., Saurette, D.D., Eacher, T.H., Li, H., Shi, Z., Adamchuk, V.I. & Biswas, A. (2020). Three-dimensional digital soil mapping of multiple soil properties at a field-scale using regression kriging. *Geoderma, 366*, Article 114253. https://doi.org/10.1016/j.geoderma.2020.114253.

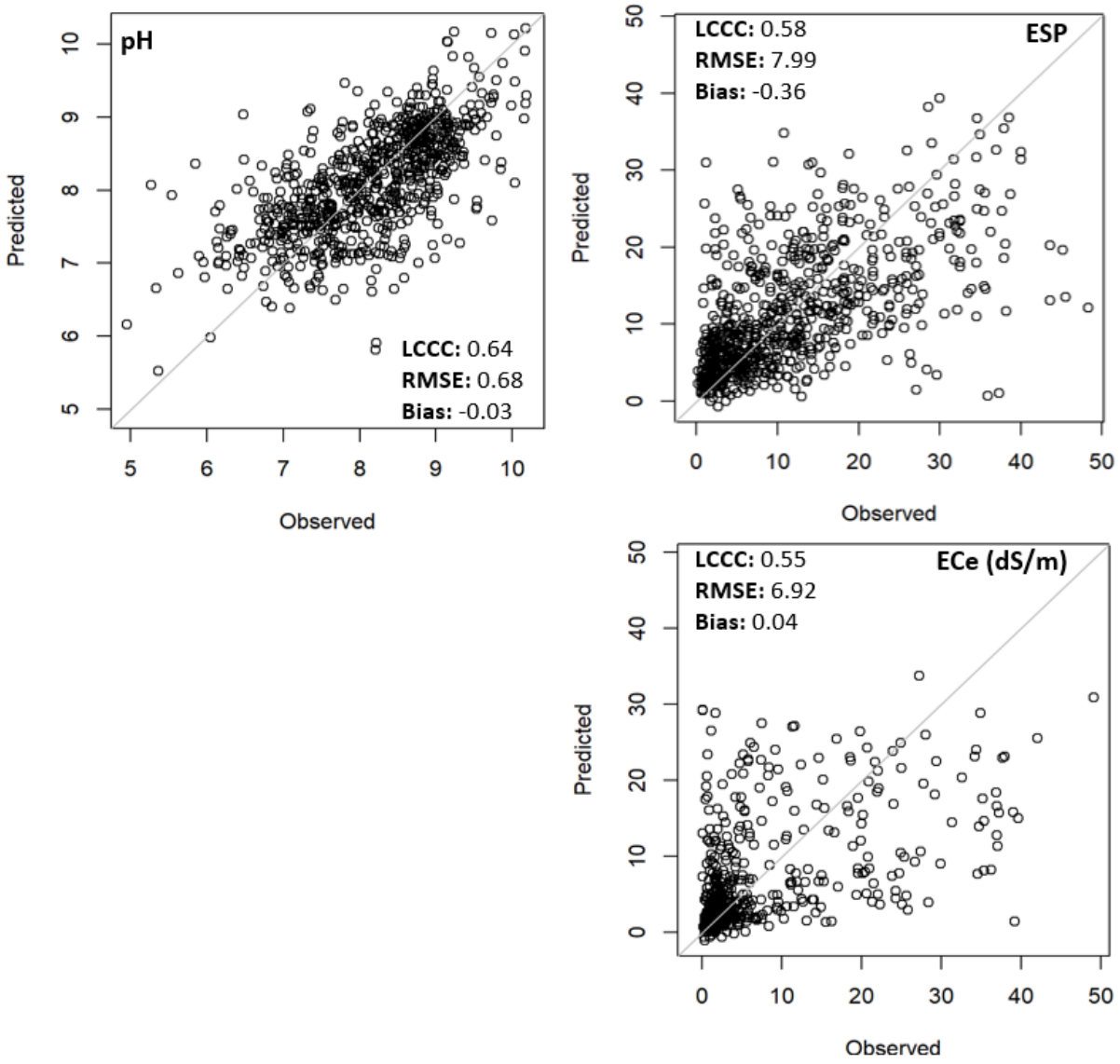
Zhang, J., Schmidt, M.G., Heung, B., Bulmer, C.E. & Knudby, A. (2022). Using an ensemble learning approach in digital soil mapping of soil pH for the Thompson-Okanagan region of British Columbia. *Canadian Journal of Soil Science, 102*(3), 579-596. https://doi.org/10.1139/cjss-2021-0091.

Zhang, N., Wang, M. & Wang, N. (2002). Precision agriculture—a worldwide overview. *Computers and Electronics in Agriculture, 36*(2-3), 113-132. https://doi.org/10.1016/S0168-1699(02)00096-0.

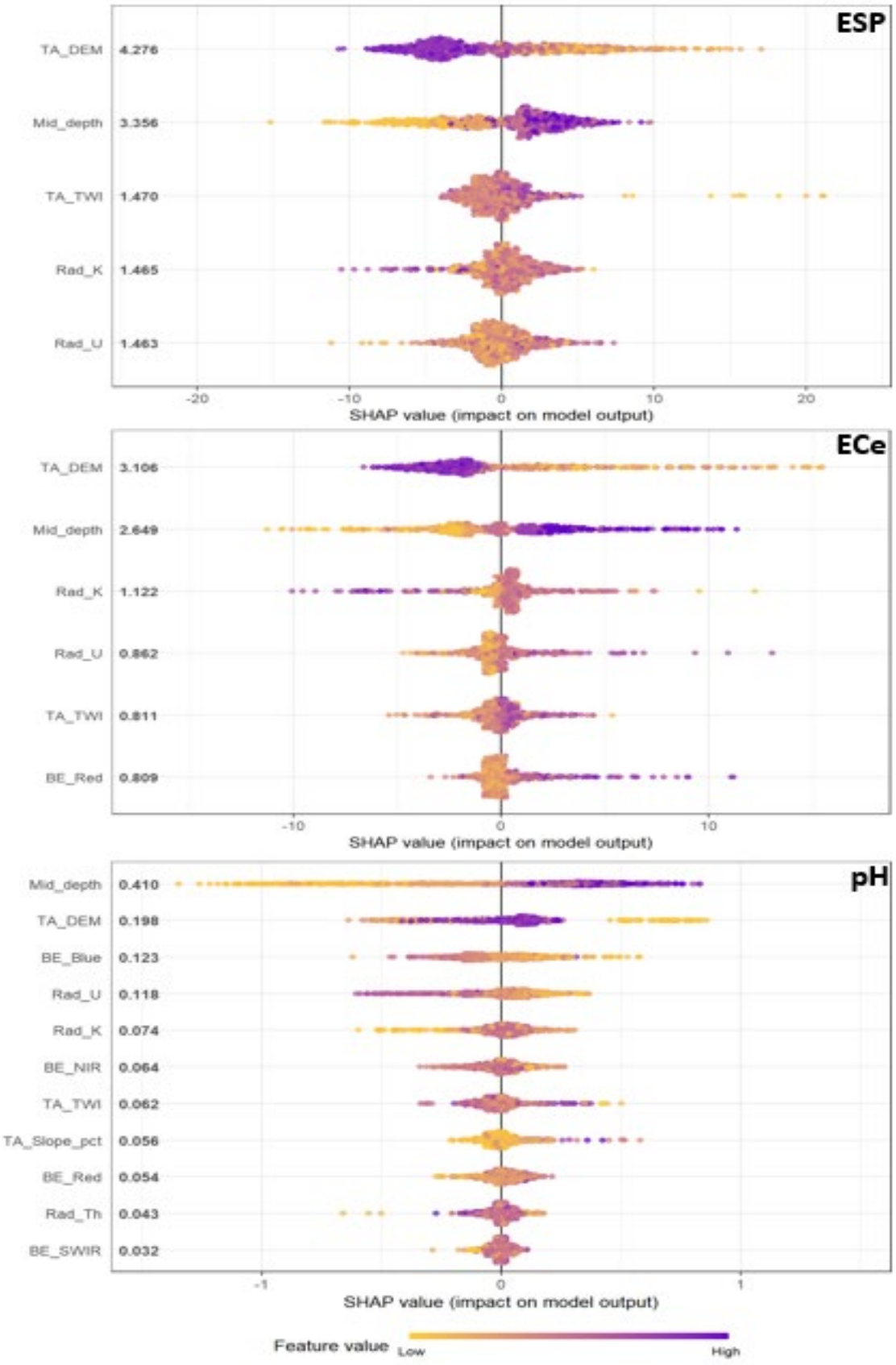
Zhao, T., Lio, S., Xu, J., He, H., Wang, D., Horton, R. & Liu, G. (2022). Comparative analysis of seven machine learning algorithms and five empirical models to estimate soil thermal conductivity. *Agricultural and Forest Meteorology, 323*, Article 109080. https://doi.org/10.1016/j.agrformet.2022.109080.

Zhou, D., Khan, S., Abbas, A., Rana, T., Zhang, H. & Chen, Y. (2009). Climatic regionalization mapping of the Murrumbidgee Irrigation Area, Australia. *Progress in Natural Science, 19*(12), 1773-1779. http://dx.doi.org/10.1016/j.pnsc.2009.07.007.

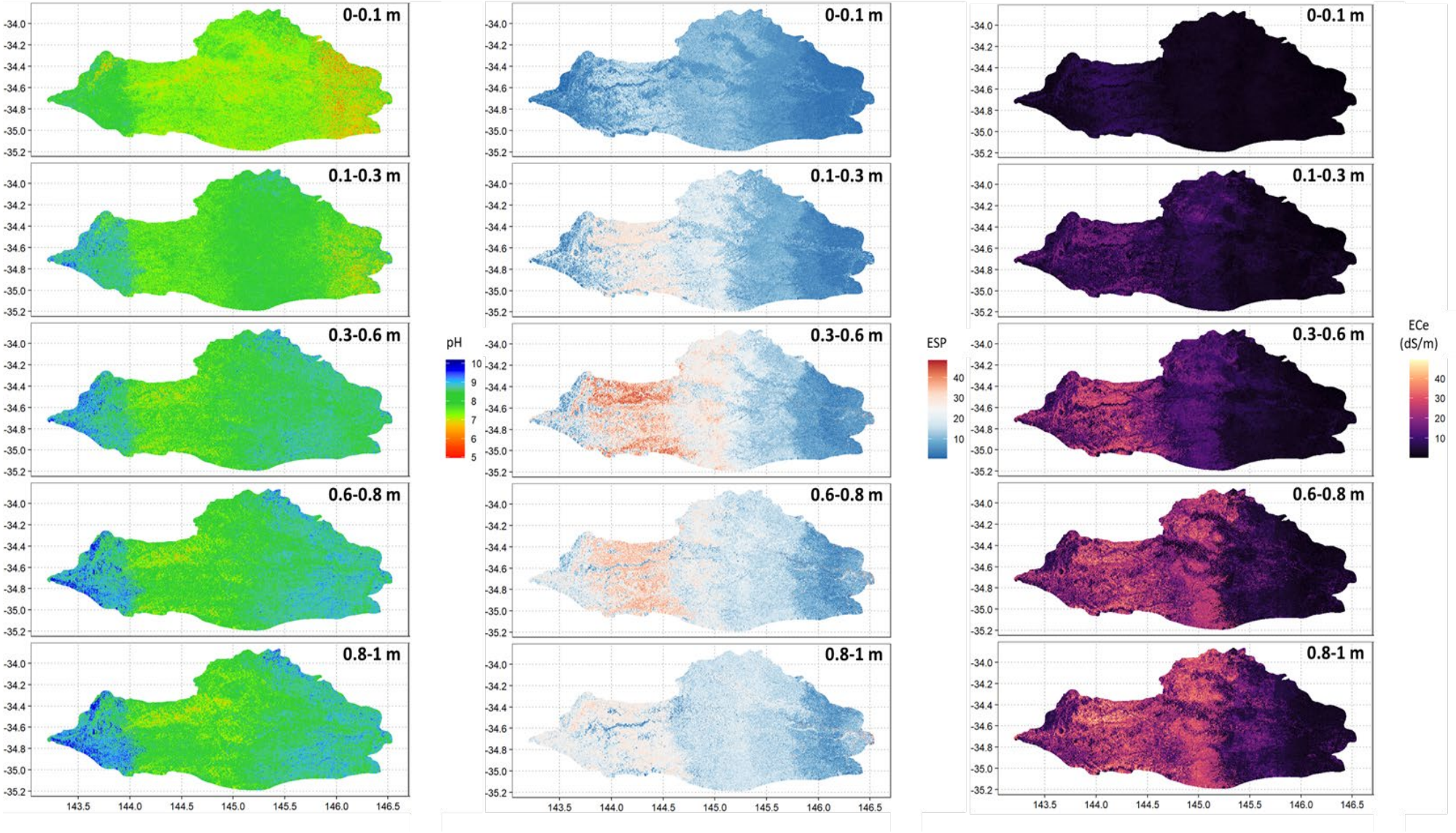
Appendix



Appendix 1. Model quality statistics for pH, ESP and ECe prior to the removal of DEM as a predictor variable.



Appendix 2.SHAP beeswarm plots for predictor variables in the models predicting ESP, ECe and pH prior to the removal of DEM from the model.



Appendix 3. Maps produced predicting pH, ESP and ECe prior to the removal of DEM as a predictor variable. A clear banding effect is seen.