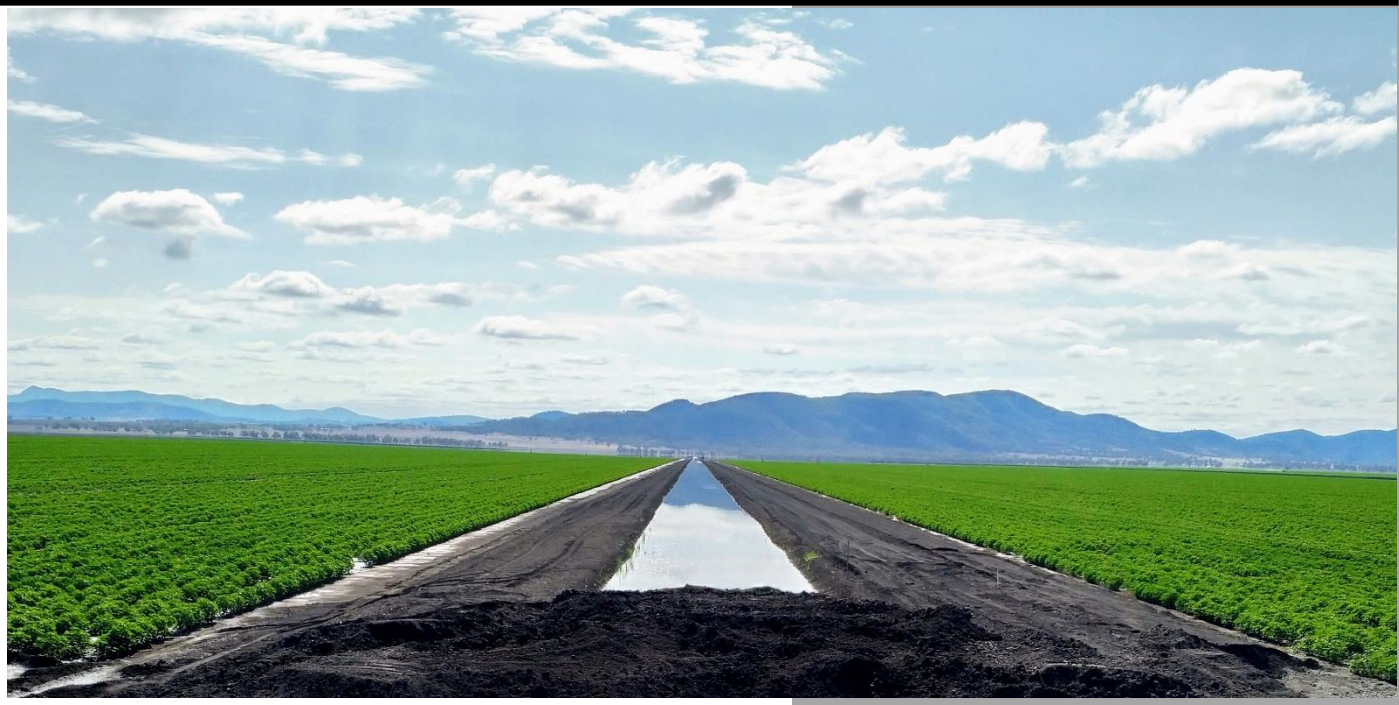
School of Biological, Earth and Environmental Sciences

Faculty of Science

Digital Mapping of Soil Water Properties Across the NSW Cotton Growing Areas



A thesis submitted in

fulfilment of the requirements

for the degree of

Doctor of Philosophy

Ehsan Zare

#UNSWSOILSCIENCECENTRAL

March 2017-December 2020

Abstract

Climate change forecasts and modelling for south-east Australia, where irrigated cotton growing occurs, suggest hotter conditions and reduced rainfall with less water for evapotranspiration. There are also ever increasing and competing demands for water from environment, mining and domestic applications. To improve soil-water management, knowledge on soil-water dynamics and soil properties which influence water holding capacity in cotton growing areas is required at the field, farm and district level. Given the heavy-clay nature of the soil (i.e., Vertosols), detailed characterization of soil moisture is a challenging task. Recently, apparent soil electrical conductivity (EC_a) data, measured by electromagnetic (EM) induction instruments, is increasingly being used with inversion modelling to understand the dynamics in volumetric soil moisture content (θ). This is because it is non-invasive, quick and inexpensive. In this thesis, electromagnetic conductivity images (EMCI), generated by inverting EC_a data (σ), has been used to map the spatial and temporal variations of θ along an irrigated cotton field located near Narrabri, New South Wales, Australia. Furthermore, EMCI was applied to create 2-dimensional maps of different soil properties in order to identify more permeable prior stream channels along a 4km section of a farm supply channel and to better understand how soil physical (e.g., clay, silt and sand) and chemical properties (e.g., EC_e and CEC) influence soil water holding capacity. In addition, the use of additional sources of digital data (i.e., gamma-ray spectrometry) was examined to understand spatial variation of available water content (AWC) at the field scale and landscape units across larger spatial scales, in concert with other digital data (i.e., EC_a and DEM). The results showed that the collection of EC_a and its inversion to produce estimates of σ is a useful tool to create digital soil maps (DSM) of θ and soil physical (e.g., clay, sand) and chemical (e.g., EC_e and CEC) properties using machine learning modelling (i.e., support vector machine) with good accuracy (Lin's Concordance > 0.8). Moreover, it was evident that DSM of AWC and soil profile classes can be created using EC_a in combination with gamma-ray spectrometry data. It was concluded that DSM techniques using easy and cheap to acquire digital data such as EC_a and gamma-ray spectrometry data can aid precision irrigation and agriculture by providing quantitative information and digital maps of soil properties for the farmers and decision makers.

Acknowledgment

My sincere thanks to my supervisor Associate Professor John Traintafilis for his generous patience, intuitive suggestions, support and the fabulous discussions we had throughout the duration of my PhD candidature. I have learnt a great deal not only about digital soil mapping but academic life over the past 4 years through his uninhibited help and motivation and during conferences, trips, fieldworks and meetings.

I would like to acknowledge Australian Federal Government for providing me with Research Training Program Scholarship and Cotton Research and Development Corporation (CRDC) for making available to me a top-up scholarship during my PhD candidature. I would also like to acknowledge CRDC for financially support all the projects presented in this thesis.

I would like to express my gratitude to a number of colleagues at #UNSWSOILSCIENCECENTRAL2017-2020 who provided invaluable support to this PhD thesis including Dr Jingyi Huang, Muddassar Muzzamal, Triven Koganti, Nan Li, Tibet Kongnawang, Dongxue Zhao, Maryam Arshad, Xueyu Zhao and Jie Wang. In addition, I would like to thank Mr Jonathan Russell the HDR officer at school of BEES for his countless assistance during my PhD candidature.

Finally, I would like to thank my beautiful parents, Azar and Mohammad and my lovely wife, Negar for their continuous encouragement and support during my PhD candidature.

Table of Contents

List of Tables

List of Figures

List of publications from this thesis

Chapter 3:

Zare, E., Li, N., Arshad, M., Nachimuthu, G., & Triantafilis, J. (2020). Time-lapse imaging of soil moisture in a flood irrigation field monitored using electromagnetic conductivity imaging: Wetting phsase. *Soil Science Society of America Journal*. doi.org/10.1002/saj2.20192

Chapter 4:

Zare, E., Arshad, M., Zhao, D., Nachimuthu, G., & Triantafilis, J. (2020). Two-dimensional time-lapse imaging of soil wetting and drying cycle using EM38 data across a flood irrigation cotton field. *Agricultural Water Management*, 241, 106383. doi.org/10.1016/j.agwat.2020.106383

Chapter 5:

Zare, E., Li, N., Khongnawang, T., Farzamian, M., & Triantafilis, J. (2020). Identifying Potential Leakage Zones in An Irrigation Supply Channel by Mapping Soil Properties Using Electromagnetic Induction, Inversion Modelling and a Support Vector Machine. *Soil Systems*, 4(2), 25. doi.org/10.3390/soilsystems4020025

Chapter 6:

Zare, E., Wang, J., Zhao, D Arshad, M., & Triantafilis, J. (2020). Scope to map available water content using proximal sensed electromagnetic induction and gamma-ray spectrometry data. *Agricultural Water Management*, 247, 106705. doi.org/10.1016/j.agwat.2020.106705

Chapter 7:

Zare, E., Ahmed, M. F., Malik, R. S., Subasinghe, R., Huang, J., & Triantafilis, J. (2018). Comparing traditional and digital soil mapping at a district scale using residual maximum likelihood analysis. *Soil Research*, 56(5), 535-547. doi.org/10.1071/SR17220

Chapter 1: Introduction

Soil is a complex system which supports the growth of plants with an adequate supply of nutrients as well as water and air. It does this through pore spaces, which allows roots to explore nutrients, water and air for plant respiration. One of the most agriculturally productive and versatile Soil Orders of the Australian Soil Classification are the Vertosols (Triantafilis et al., 2001). This is because they are inherently fertile (Gray et al., 2002) owing to the high clay content (i.e. > 35 %) and smectite mineralogy, with very high cation exchange capacity (i.e. 80 < CEC < 150 cmol(+) kg^{-1}). Moreover, because the interlayer of smectites (e.g., montmorilionite) are expandable, these are referred to as "swelling clays" because they can undergo as much as 30 % volume change (Bronswijk and Evers-Vermeer, 1990). In the rice and cotton growing areas of Australia, including southern and northern New South Wales (NSW), respectively, Vertosols can provide this pore space via these inherent shrink-swell properties (see Figure 1a).

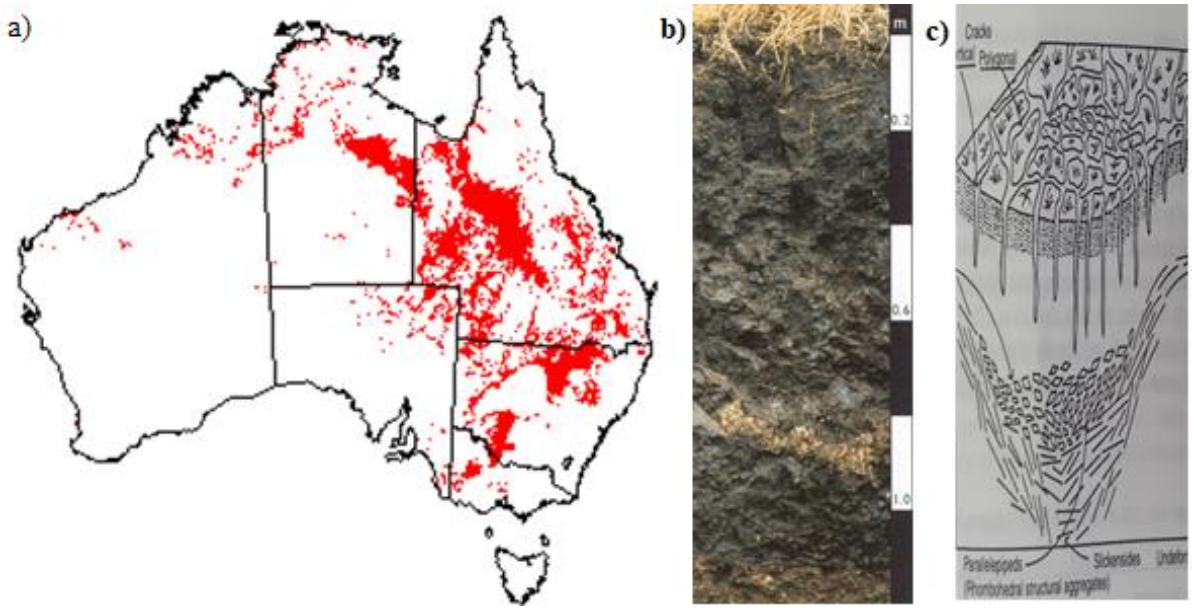


Figure 1.1 a) Location of Vertosols of the Australian Soil Classification, b) Vertosol image (after McKenzie et al., 2004) and c) vertical cracks and angular blocky structure in the topsoil (0-0.3 m) and subsurface (0.3-0.6 m) and wedge/lenticular structure and slickensides in the subsoil (> 0.6-0.9 m) (Isbell, 2002).

Specifically, and as shown in Figure 1.1b) and c), the large vertical cracks in the topsoil (0-0.3 m) and subsurface (0.3-0.6 m) allow passage of water and air readily, however, the underlying wedge/lenticular shaped structure in the subsoil (~ > 0.6 m) can slow the flow and allow the soil profile to fill with water. Owing to the strong swelling behaviour, the expanding montmorilionite clay seals in the water due to this self-mulching character. This property makes the Vertosols highly prized, because in furrow irrigated farming systems, the soil profile can be rapidly filled through these vertical cracks. In terms of cotton habit, the large tap root then allows the plant to grow into the stored soil moisture and reopen the cracks. However, where water is not used, water may be lost via deep drainage (Silburn et al., 2014); and this has been identified as a water use inefficiency in irrigated cotton systems (Willis et al., 1997).

Whilst there have been vast improvements in water use efficiency (Roth et al., 2014) via soil/water/crop management strategies, climate change forecasts and modelling (Dai, 2013) for south-east Australia, where irrigated cotton growing occurs, suggest reduced rainfall, hotter conditions and therefore less water for evapotranspiration. There are also ever increasing and competing demands for water from environment (Gawne et al., 2010), mining and domestic applications (Roth et al., 2014). To improve soil-water management, knowledge of the availability of water to cotton growing plants in these Vertosols is required at the field scale and district level. This includes information about field capacity (FC) which is the maximum water that can be stored and used by plants where there is enough air-filled pore space to allow aerobic microbial activity. Another information is permanent wilting point (PWP) which is the minimum moisture at which a plant wilts and cannot recover turgidity.

Given the heavy-clay nature of Vertosols, and as shown in Figure 1.2a), at FC the soil volumetric water content (θ, $cm^3\,cm^{-3}$) which is the measure of water in soil, is large (i.e., 0.32-0.40 $cm^3\ cm^{-3}$) (Gardiner and Miller, 2004). Conversely, and as shown in Figure 1.2b), θ at PWP is intermediate (0.20-0.24 cm^3cm^{-3}). However, the water held between FC and PWP.,

(i.e., Available Water Content - AWC) is small (0.12-0.16 cm^3 cm^{-3}) (Gardiner and Miller, 2004).

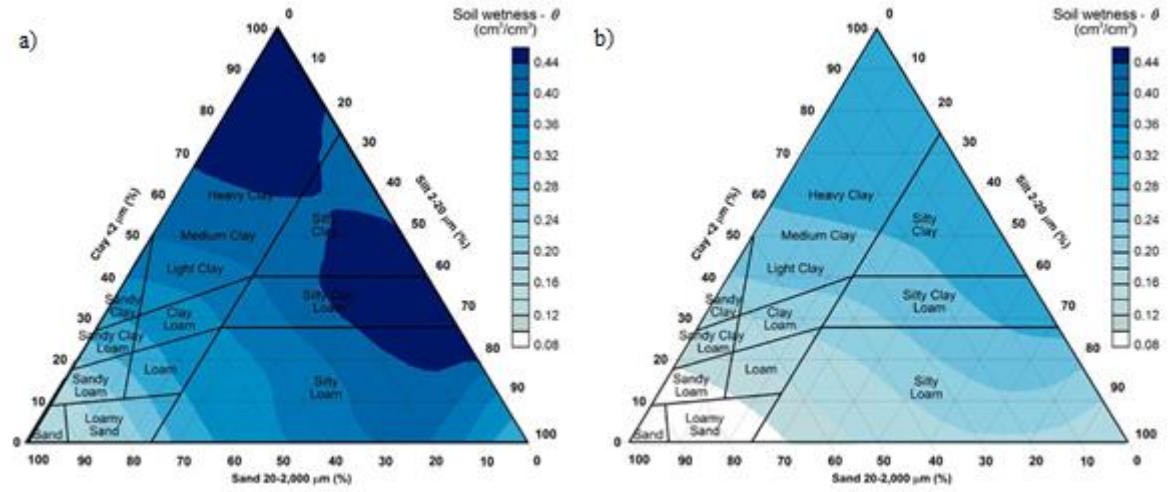


Figure 1.2 Volumetric soil moisture (θ, m^3 m^{-3}) at a) field capacity (FC) and b) permanent wilting point (PWP) as a function of the three particle size fractions (i.e. clay silts and sand) and soil textures (e.g. heavy, medium and light clay) as shown in terraGIS (see http://www.terragis.bees.unsw.edu.au/, last accessed 22 September, 2020).

Owing to the small window of AWC, furrow irrigated cotton production requires precise information about θ. In the 70's this was provided by neutron moisture meters (see Figure 1.2a). However, there were several disadvantages, including the sphere of influence of the neutron probe varies with the drying cycle of soil as well as the surface effect. Nevertheless, it was less time-consuming, non-destructive (after access tube installation) and was independent of temperature and pressure. However, there were serious associated health hazards from radiation and increasing regulatory burden and training for licensee and users meant these instruments were phased out and replaced by time-domain reflectometry (TDR) instruments (HV technologies, Inc., Virginia, USA). While more accurate and not requiring calibration, the major disadvantage is that a fair amount of expertise is needed to set-up and operate a TDR system (Wraith et al., 2005). Moreover, accuracy beyond 2.5 cm along the waveguides (see Figure 1.2b), requires excellent probe/soil contact with technical expertise required to set up the equipment in the field (Wraith et al., 2005).

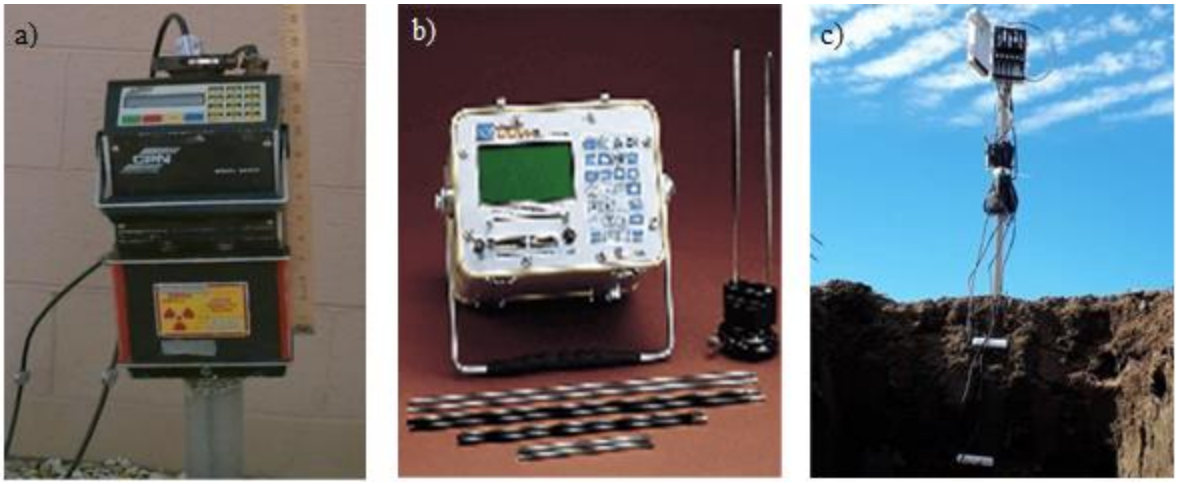


Figure 1.3 Commercially available soil moisture measuring equipment including a) Neutron probe (InstroTek, North Carolina, USA), b) time-domain reflectometry (TDR) (HV Technologies, INC., Virginia, USA) and c) Decagon GS3 moisture sensors (METER Group, Washington, USA).

More recently, and also taking advantage of the dielectric constant (K_a) of water (80), solids (4) and air (1) of the TDR, Figure 1.3c) shows that smaller stainless-steel probes are increasingly being used, including Decagon GS3 sensors (METER Group, Washington, USA). Despite these instrumental advances, these instruments still only provide point source information. To value add to these limited measurements of θ and to better understand the wetting and drying cycles of furrow irrigated cotton farming systems and potentially improve soil/water/crop management in Vertosols, information about θ is required in space and time.

A potentially useful approach is the use of direct current resistivity (DCR) meters and as shown by Kelly et al. (2011) who monitored wetting phase of irrigation along a 32-m furrow and over a 1-day period by inverting the DCR data to produce electrical resistivity tomography (ERT). While the data which can be generated using this approach is useful, the DCR technique requires physical insertion of steel electrodes that need to be connected using wires to connect the electrodes. This is time consuming and expensive. Moreover, the equipment once installed cannot be moved and obtaining information across an entire field or along a long transect is not possible. Recently, it has been shown the single frequency and multi-coil electromagnetic (EM) induction instruments could be used as an alternative. This is because EM instruments do not require direct physical contact. It is also because they measure the inverse of resistivity,

that is the soil apparent electrical conductivity (EC_a, mS m^{-1}), and if the soil clay content and mineralogy is the same and the salinity is constant, differences in EC_a can be related to changing θ across a field or with time (Robinson et al., 2012).

EM instruments are increasingly being used, with either 2- or 3-dimentional inversion modelling, to map soil properties such as clay content (Xueyu Zhao et al., 2019), cation exchange capacity (Khongnawang et al, 2020) and salinity (Koganti et al., 2018). Of interest herein, is the use of EC_a and inversion to mimic the ERT and to estimate the electrical conductivity (σ, mS m^{-1}) at specified depths which can be directly related to θ at the same depths. The equivalent to ERT, in this sense, are electromagnetic conductivity images (EMCI). This approach was used by Huang et al. (2017) to create 2-dimensional time-lapse images of soil-moisture dynamics beneath a centre-pivot irrigated lucerne field in San Jacinto, in southern California. Moreover, at a single time-step Huang et al. (2016) showed how 2-dimensional images from multiple transects could be combined to compile a single layer depth of θ status across a part of the same field.

Owing for the need to understand the temporal and spatial distribution of soil θ in space and time in irrigated cotton growing areas and also to understand where there are inefficiencies in water delivery and also the requirement for knowledge across larger spatial scales and across the landscape, this PhD thesis has the following brief aims.

Chapter 2 is a review of the literature into the theory of operation of various field-based sensors to measure θ and the use of geophysical data to value add to θ, using models and their application, with particular emphasis on inversion of EC_a data to predict θ and other soil physical (i.e., clay, silt and sand) and chemical (e.g., CEC) properties.

In Chapter 3, in a small furrow irrigation field a single-frequency multiple-coil array instrument (i.e., DUALEM-421) was used to measure EC_a, with inversion used to estimate σ (mS m^{-1}) and make a relationship between σ and θ using an artificial neural network (ANN).

Validation is used to test if the approach could be used to monitor θ during wetting-phase at the start of an irrigation cycle and with level of agreement (i.e., Lin's correlation coefficient).

Following experimenting the use of a heavy and complex EM instrument (i.e., DUALEM-421) during wetting-phase in Chapter 4, in the same field but at a later stage of the irrigation cycle, a single-frequency single-coil EM38 instrument was used to collect EC_a in various modes of operation and at multiple heights during wetting and drying-phase. The estimates of σ (mS m^{-1}) were used to compare various linear and machine learning (e.g., Cubist) models to make a relationship between σ and θ and to determine which has largest prediction agreement to monitor θ during wetting and drying phases throughout a 10-day period of a cotton growing season.

In Chapter 5, using a similar inversion approach used in the in Capters 3 and 4 the aim is to determine if the EC_a data from a DUALEM-421 can be inverted to estimate σ, which can be related to and used to predict soil physical (clay and sand) and chemical (EC_e and CEC) properties using a machine learning model (e.g., support vector machine-SVM) and whether the 2-dimensional maps can be used to identify more permeable prior stream channels along a 4km section of a farm supply channel.

In Chapter 6, the aim is to see if DSM of FC and PWP could be developed from γ-ray spectrometry (i.e. RS-700) data in addition to EC_a (i.e., DUALEM-421) data which was used in Chapters 3 to 5 and using a stepwise multiple linear regression (MLR) model and across a small (35 ha) dryland cotton field. The final DSM of AWC within the root-zone (0-0.90 m) was estimated subtracting PWP from FC and the results discussed in terms of soil-water management.

Using the formerly employed sources of digital data (i.e., EC_a and γ-ray data) and contrasting the previous research chapters which was performed in field scale, in Chapter 7, the aim is to create the DSM in district scale. This was conducted across two

geomorphologically different areas (i.e., Trangie and Warren) in the central-west of New South Wales, Australia. The effects of clustering either in combination or individual digital data including airborne γ-ray and proximal EC_a data to map soil profile classes is discussed. Also, the results are compared against a traditionally derived pedoderm map (McKenzie, 1992), in terms of predicted soil physical (e.g., clay, FC, PWP, AWC) and chemical (e.g., pH, $EC_{1:5}$ and CEC) properties.

Chapter 8 concludes the main results, research contributions and the significance of this thesis. It also discusses the limitations of this thesis and raises suggestions, improvements, and recommendations for further research.

1.1 References

Bronswijk, J. J. B., & Evers-Vermeer, J. J. (1990). Shrinkage of Dutch clay soil aggregates. *NJAS wageningen journal of life sciences*, *38*(2), 175-194.

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature climate change*, *3*(1), 52-58.

Gardiner, D. T., & Miller, R. W. (2004). *Soils in our environment* (pp. 126-165). NJ.

Gawne, B., Crase, L., & Watson, A. S. (2010). Can a collaborative focus on solutions improve our capacity to achieve sustainable water management?. *Marine and Freshwater Research*, *61*(7), 814-820.

Gray, J. M., & Murphy, B. W. (2002). Predicting soil distribution. NSW Department of Land and Water Conservation (DLWC) in collaboration with ASSI.

Huang, J., Scudiero, E., Choo, H., Corwin, D. L., & Triantafilis, J. (2016). Mapping soil moisture across an irrigated field using electromagnetic conductivity imaging. *Agricultural Water Management*, *163*, 285-294.

Huang, J., Scudiero, E., Clary, W., Corwin, D. L., & Triantafilis, J. (2017). Time-lapse monitoring of soil water content using electromagnetic conductivity imaging. *Soil Use and Management*, *33*(2), 191-204.

Isbell, R. F. (2002). The Australian Soil Classification., Rev. edn (CSIRO Publishing: Melbourne).

Kelly, B. F. J., Acworth, R. I., & Greve, A. K. (2011). Better placement of soil moisture point measurements guided by 2D resistivity tomography for improved irrigation scheduling. *Soil Research*, *49*(6), 504-512.

Khongnawang, T., Zare, E., Srihabun, P., & Triantafilis, J. (2020). Comparing electromagnetic induction instruments to map soil salinity in two-dimensional cross-sections along the Kham-rean Canal using EM inversion software. *Geoderma*, *377*, 114611.

Koganti, T., Narjary, B., Zare, E., Pathan, A. L., Huang, J., & Triantafilis, J. (2018). Quantitative mapping of soil salinity using the DUALEM-21S instrument and EM inversion software. *Land Degradation & Development*, *29*(6), 1768-1781.

McKenzie, N., Jacquier, D., Isbell, R., & Brown, K. (2004). Australian soils and landscapes: *an illustrated compendium*. CSIRO publishing.

Robinson, D. A., Abdu, H., Lebron, I., & Jones, S. B. (2012). Imaging of hill-slope soil moisture wetting patterns in a semi-arid oak savanna catchment using time-lapse electromagnetic induction. *Journal of Hydrology*, *416*, 39-49.

Roth, G., Harris, G., Gillies, M., Montgomery, J., & Wigginton, D. (2014). Water-use efficiency and productivity trends in Australian irrigated cotton: a review. *Crop and Pasture Science*, *64*(12), 1033-1048.

Silburn, D. M., Foley, J. L., Biggs, A. J. W., Montgomery, J., & Gunawardena, T. A. (2014). The Australian Cotton Industry and four decades of deep drainage research: a review. *Crop and Pasture Science*, *64*(12), 1049-1075.

Willis, T. M., Black, A. S., & Meyer, W. S. (1997). Estimates of deep percolation beneath cotton in the Macquarie Valley. *Irrigation Science*, *17*(4), 141-150.

Wraith, J. M., Robinson, D. A., Jones, S. B., & Long, D. S. (2005). Spatially characterizing apparent electrical conductivity and water content of surface soils with time domain reflectometry. *Computers and Electronics in Agriculture*, *46*(1-3), 239-261.

Zhao, X., Wang, J., Zhao, D., Li, N., Zare, E., & Triantafilis, J. (2019). Digital regolith mapping of clay across the Ashley irrigation area using electromagnetic induction data and inversion modelling. *Geoderma*, *346*, 18-29.

Chapter 2: Literature Review

2.1 Introduction

Knowledge about the spatial distribution of soil volumetric water content (θ, $m^3\ m^{-3}$) and soil-water related physical (i.e., clay, silt and sand), chemical (e.g., CEC, EC_e and pH) and hydrological (i.e., field capacity, permanent wilting point and available water content) properties is required for effective soil and water use and management. This is particularly the case in highly productive cotton growing areas of New South Wales (NSW), Australia where Vertosols are the dominant soil type. While information on spatiotemporal variation of θ at the field scale is a key variable for improving water use efficiencies and crop yield optimisation, knowledge about physical and chemical properties can assist identifying underlying coarser textures (i.e., prior stream channels) which are susceptible to deep drainage. Moreover, information on hydrological properties can aid understanding the hydraulic characteristics of the soil and the amount of stored moisture which is important for precision agriculture and in particular water use and management.

However, acquiring soil property information at a level beneficial for soil/water use and management is a challenging job. This requires extensive soil sampling and laboratory analysis which is time-consuming and labour intensive. One solution that is being increasingly utilised is the application of digital soil mapping (DSM) techniques to value-add to limited soil property data. This includes use of proximal or remotely sensed digital data collected using either active (e.g., electromagnetic induction) or passive (e.g., gamma-ray spectrometry) sensors. This is less-challenging specifically when the digital data can be correlated to the soil and/or water properties. However, when the soil properties are not directly correlated with digital data, more complex techniques come into practice. One successful way to do this is to numerically cluster (e.g., Fuzzy K-means) the digital data to generate classes and then test those classes against the measured soil properties to find out if they are statistically significant.

Another approach is to use machine learning algorithms which can account for non-linear relationships.

This chapter initially reviews the traditional laboratory and field methods for determination of various hydrological (i.e., volumetric moisture content, field capacity, permanent wilting point and available water content), physical (i.e., clay, silt and sand), and chemical (e.g., CEC, EC_e and pH) properties and then describes the different sources of digital data which can potentially be used in association with mathematical methods to create DSM of these soil properties. Moreover, various research gaps are identified and relative to the application of DSM methods at the field scale to measure and map θ and particularly soil water-related properties and identify soil management classes.

2.2 Laboratory methods for determination of soil moisture

The thermogravimetric method is traditionally used for measuring θ across a field and at various depths. The oldest established and the only truly direct method to measure θ is the thermogravimetric method (Smith, 2000). This method is labour intensive and time consuming. In addition, real-time θ measurement is not possible using this method. This is because this method requires drying the soil samples in the oven at 105 ^{o}C until there is no further weight loss. Basically, a soil sample is placed in a metal container of known weight and weighed. It is then placed in an oven for 24 hours and weighed again. For some soil types, and especially wet soil samples, further drying might be required and the samples should be placed in the oven for further 4 hours. When the weight loss does not exceed 0.1% of the mass at the start of the 4 hours, the dry mass is attained. The gravimetric water content (w, g g^{-1}), of the sample is calculated as follows:

$$w = \frac{\text{Mass of wet soil } - \text{Mass of dry soil}}{\text{Mass of dry soil}}$$

If the intact soil sample volume is unknown, the bulk density is required, whereby θ (m^3 m^{-3}) can be calculated as follows:

$$\theta = w \times \text{bulk density (g cm}^{-1}\text{)}$$

In terms of soil-water management, the classification of soil θ is based on the availability of water to plants in the soil environment. Terms used to describe this include; saturation (Gravitational water), field capacity (FC), permanent wilting point (PWP), and available water capacity (AWC). At saturation, soil porosity corresponds to θ. However, FC approximates the water that is held in soil after it has been fully wetted, and all gravitational water has been drained away and it is reached about one to two days after heavy rainfall or irrigation ceases. For example, for a clay soil texture, θ at FC is large (i.e., 0.32-0. 4 m^3 m^{-3}). In practical terms, the soil holds the maximum amount of water that can be stored and can be used by plants. Moreover, there is sufficient air-filled pore space which allows for aeration (i.e., aerobic microbial activity and plant growth). However, and from a tillage standpoint, soil turns to mud above FC is unworkable. In the laboratory condition the suction that defines this value varies from soil to soil but it is around –10 kPa.

Permanent wilting point (PWP) is defined as the minimum soil moisture at which a plant wilts and can no longer recover its turgidity. In the laboratory condition the suction that defines this value is –1500 kPa. In the field, PWP will appear dusty and dry. However, water is still present in very thin films around soil particles and in micropores. For example, for a clay soil texture, θ at PWP is small (i.e., 0.2-0.24 m^3 m^{-3}). Nevertheless, soil water is held tightly by adhesion forces (i.e., about 10 molecules thick) and is generally unavailable for use by microbes and/or plant roots. The amount of water held between FC and PWP is known the Available Water Content (AWC), and it is a measure of the amount of water in the soil that is "potentially" available to plants. In the example of a clay soil, the difference between FC large

(i.e., 0.32-0.4 m^3 m^{-3}) and PWP (i.e., 0.2-0.24 m^3 m^{-3}) is quite narrow (0.12-0.16 m^3 m^{-3}) and therefore to measure and monitor this in real time is quite crucial, particularly in irrigated farming systems. To measure and monitor this in real-time requires in-situ measurement, whereas to manage this practically in the field requires spatial information at the field, farm and/or district scale.

2.3 Field methods for θ measurement

Several geophysical instruments have been developed by scientists and manufacturers to aid the measurement of θ in-field and monitor it in real-time. These include, neutron probes, time domain reflectometry (TDR), frequency domain reflectometry (FDR) and capacitance probes. Nevertheless, probe installation is required for most of these instruments, therefore they are labour intensive, and cost associated with that. The main elements of these methods are compared in Table 2.1.

Table 2.1 Summary of various field instruments used for measuring soil water content (after Huang 2017).

Instrument	Measuring property	Accuracy	Limitation	Invasive	Efficiency	Mobile	Hazard	Price
Neutron probes	Neutrons	High	Not applicable for soil surface	Yes	Low	No	Yes	High
TDR	Permittivity and bulk electrical conductivity	High	Affected by soil texture	Yes	Intermediate	No	No	Intermediate
FDR	Dielectric constant	High	Affected by temperature	Yes	Low	No	No	Intermediate
Capacitance probes	Dielectric constant	High	Affected by soil texture, temperature, and salinity	Yes	Low	No	No	Intermediate

2.3.1 Neutron probe

A neutron probe is a device used to either quantitatively predict θ or estimate it qualitatively. The former requires calibration. The probe operates using fast neutrons, specifically, when these neutrons collide with hydrogen nuclei present in the soil, they lose

energy and consequently slowdown. The probe can estimate the amount of hydrogen present in the soil by detecting the slow neutrons returning to the device. Since water contains two atoms of hydrogen per molecule, this provides a measure of the amount of water present in the soil. Access tube installation is required to measure the subsurface and subsoil θ (Figure 2.1). Neutron probes can provide an accurate prediction of θ when they are calibrated with θ determined at any depth with high spatial resolution (Al-Ain et al., 2009). Nevertheless, it is essential to calibrate the device for different soil types and zones over a period with different soil moisture (Robock et al., 2000).

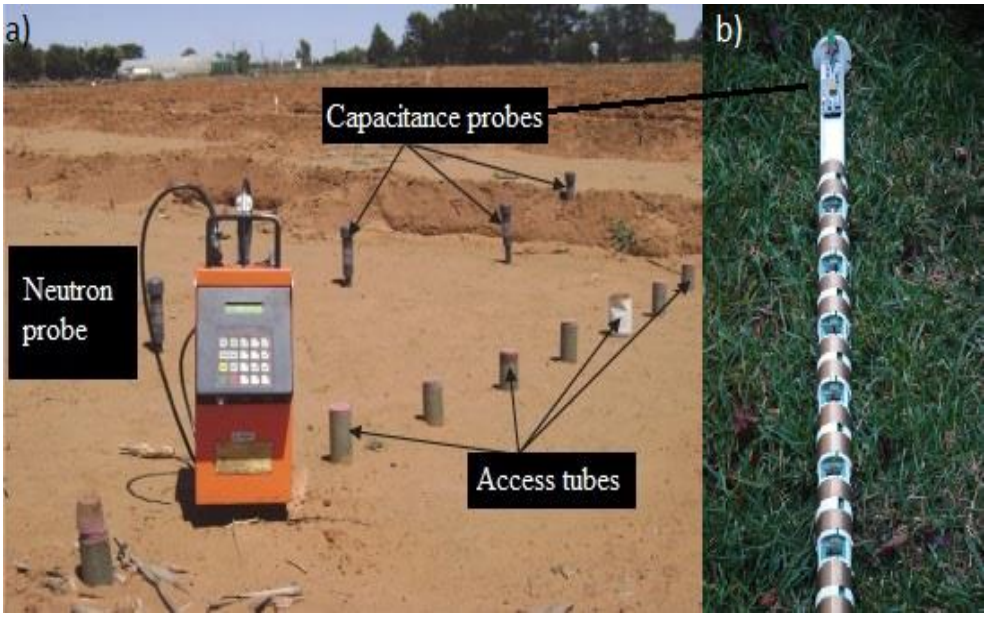


Figure 2.1 a) Neutron and capacitance probe and access tubes (after Zerizghy et al., 2013) and b) EnviroSCAN capacitance probes (Sentek Technologies, Stepney, Adelaide, South Australia) (accessed from http://www.americanlaboratory.com/914-Application- Notes/38090-Purchasing-Considerations-for-Field-Soil-Water-Content-Profile-Probes/).

However, neutron probes do not provide reliable predictions because of the low spatial resolution associated with the sphere of influence which varies as a function of either the drying cycle, textural differences (i.e., particularly in duplex soil) and surface effect due to lack of enough returning slow neutrons (Robinson and Dean, 1993). However, the major concern of neutron probes is the health hazard of exposure to the radioactive source that is needed for neutron scattering. Therefore, there is a large regulatory burden which requires training for

licensee. Moreover, it requires doubly locked storage and cannot be used or left unattended. In addition, calibration is essential and although measurements can be done repeatedly, continuous logging of θ is not been possible (Zerizghy et al., 2013).

2.3.2 TDR

To overcome the health hazard and regulatory requirements of neutron probes, time domain reflectometry (TDR) instruments were developed in the 1980's. A TDR instrument actually measures the dielectric constant (K_a) of the soil using active electromagnetic waves. The theory of operation is that the instrument converts the travel time of high frequency of an electromagnetic (EM) pulse into θ. This is achieved by generating a fast-rise pulse sent at speed of light down a transmission line consisting of two parallel waveguides. The wave is reflected from open ends of the waveguides and returns along original path. A microprocessor measures the travel time of the wave, which is then used to directly calculate K_a of soil. This is possible because θ (K_a = 80) dramatically slows speed of EM pulse as compared with solids (K_a = 2-4) and air (K_a = 1). A TDR wave form is displayed on screen for storage and interpretation, with time delay correlated to θ also displayed.

Figure 2.2 Acclima TDR-315 (Meridian, Idaho USA) moisture sensor (access from https://soilsensor.com/articles/time-domain-reflectometry-tdr/).

Since the 1980's, technological innovations in instrument design and waveform analysis, has enabled TDR instruments to become one of the most accurate tools for in-field θ measurement (Ferré et al., 2002; Robinson et al., 2003). This is because it can provide θ values which range from 0-0.5 cm^3 cm^{-3}, with accuracy better than 0.01-0.20 cm^3 cm^{-3} without calibration. TDR also almost eliminates site specific calibration and requirement for costly, specialized licensed personnel associated with neutron probes. However, they also need to be installed in the soil profile (Figure 2.2). The issue herein is that while multiple depths and more discrete measurements can be made with probes installed at the desired depth of measurement, extensive cabling is required, and a multiplexer is also needed to be located close by to enable waveform analysis and interpretation. For farming enterprises, the most important disadvantage is that a fair amount of expertise is needed to set up and operate a TDR system with commercially available equipment. This makes them not suitable for large spatial extent θ monitoring.

With all the above said, TDR has been used broadly for in-field θ monitoring and studies related to solute movement and infiltration. In one study, Bargar et al. (1999) used TDR measurements of θ at 0.5-h intervals following natural rainfall events to determine the pattern of infiltration and soil water recharge for uncropped row and furrow positions in long-term ridge-tillage fields. The idea was to modify the pattern of θ and solute movement to reduce NO_3 leaching from corn production. They measured θ under adjacent rows and three adjoining furrows for 103 days in 1992 on a Clarion loam (fine-loamy, mixedmesic, Typic Hapludoll) near Boone, Iowa, USA and 96 days on Monona silt loam (fine-silty, mixed mesic, Typic Hapludoll) near Treynor, Iowa, USA. The study sites were a 16 by 16 m area at Treynor and 9.1 m wide by 16 m long area at Boone.

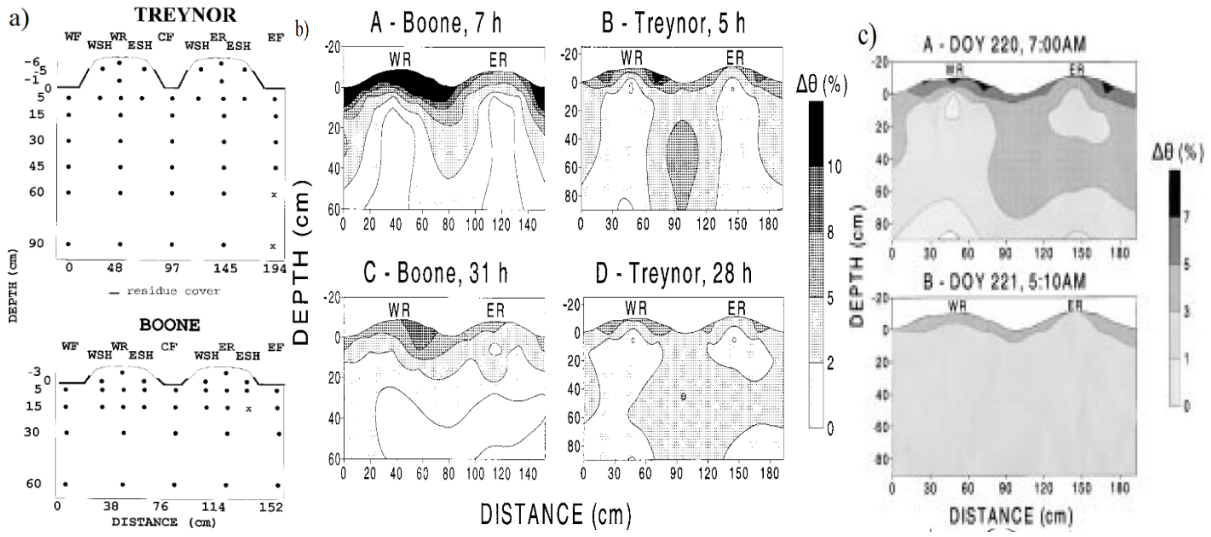


Figure 2.3 a) Sensor placement at Boone and Treynor sites b) Increase in volumetric soil water content (θ), compared with antecedent θ measured 1 h before start of rainfall, under ridges and furrows at indicated elapsed time in hours following start of rainfall events of 34 mm at Boone (0250-0950 h) and 21 mm at Treynor (0035-0845 h) DOY 184. c) Increase in θ, compared with antecedent θ measured 1 h before start of rainfall, under ridges and furrows at time indicated following 22 mm rainfall ending 0555 h DOY 220 at Treynor (after Barger et al., 1999).

Figure 2.3a shows θ was measured in ridge and shoulder positions on two adjacent rows and three adjoining furrows at 0.5-h intervals by TDR and weekly by neutron thermalization. Thirty-three TDR probes were installed horizontally to the 60 cm depth in a trench at Boone and 40 TDR probes were installed horizontally to the 90 cm depth at Treynor. Figure 2.3b shows the redistribution and water budget. Figure 2.3b, (part C and D) shows soil water redistribution was less uniform between row and furrow positions at Treynor than at Boone after 24 h. Vertical redistribution at Treynor is shown by the decrease in θ for the centre furrow and upper ridge, shoulder, and furrow depths.

Moreover, lateral soil water movement was indicated by the concurrent increase in θ under the right-hand side row by 24 h after the start of rainfall (Figure 2.3b, part C and D). Figure 2.2c shows on day 220 of the year (DOY 220) at Treynor θ averages 0.04 m^3 m^{-3} greater across 1-m profile than DOY 184 (Figure 2.2b, B) and the response time for row and lower furrow positions was more rapid than on DOY 184, possibly due to greater hydraulic conductivity and smaller possible maximum θ changes for DOY 220 than DOY 184. Greatest

increases in θ were measured 0.5 m below the surface of ridge (26-cm soil depth) and shoulder (25-cm soil depth) positions. The initial increase in θ was generally greater in the furrow than at the same elevation in the row for depths below 0.5 m, but following redistribution, the average increase was similar for row and both tracked and untracked furrow positions for these depths (Figure 2.3c, part B).

They concluded that the precision and frequency of 0.5-h field TDR measurements of θ allowed a qualitative assessment of the direction of water movement between naturally occurring rainfall events and quantitative assessment of θ under rows and furrows over time. They used this to estimate infiltration direction and redistribution amounts between rainfall events across the measurement depth under rows and furrows. However, patterns of infiltration and redistribution were less obvious from daily average θ measurements and were difficult to determine from weekly TDR or neutron probe measurements.

2.3.3 FDR

Frequency domain reflectometry (FDR) sensors continuously measure θ by transmitting an electromagnetic wave along probes and record the frequency of the reflected wave. They also provide the measurements of θ based on the K_a of the soil through a capacitor in which the water molecules are polarised and aligned in a dipolar electric field (Muñoz-Carpena and Duckes, 2015). The capacitor measures the K_a via two parallel electrodes which create a sinusoidal waveform with means of an electronic oscillator (Evett and Cepuder, 2008). The waveform interacts with the soil around the capacitor and will be affected by the soil bulk electrical permittivity and the dipoles respond to the frequency of the electric field that leads to know the inverse relationship between the frequency of oscillation and θ (Hernandez et al., 2018).

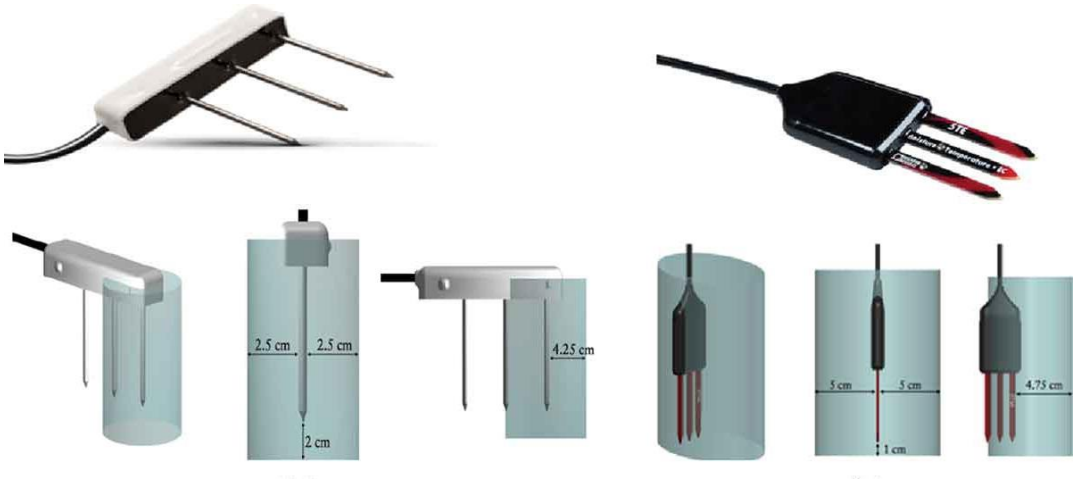


Figure 2.4 FDR sensing electrodes and idealized measurement volumes of Decagon GS3 (left) and Meter 5TE (right) moisture sensors (after Son et al., 2017).

Recently FDR sensors, and as shown in Figure 2.4, including the Decagon GS3 and Meter 5TE (METER Group, Inc. USA) sensors, have been used in several studies because they operate at 70 MHz. This minimises the effects of soil salinity and texture on θ measurements. In one study, Son et al. (2017) examined the performance of the GS3 and 5TE sensors under laboratory and field conditions for estimating θ of saline-alkali reclaimed land in South Korea. They also calibrated both sensors using different soils from different textural groups (i.e., Silt, Silt loam, Silt clay) and compared that with the universal factory calibration. They showed that the factory calibration for θ of the GS3 and 5TE sensor were accurate for all soil samples (RMSE = 3.02-3.59%) tested, except that the 5TE factory calibration was not accurate for silt soil (RMSE=7.35%).

These findings suggest accurate calibration for all the soil samples tested in the laboratory experiments. The factory calibrations of the 5TE sensors did not meet the statistical goals for accuracy. However, the factory θ calibration of the GS3 sensors closely matched the measured values, suggesting accurate calibration (Figure 2.5). They also concluded that testing of the GS3 and 5TE sensors showed very good agreement between actual θ and θ readings generated using the field-based calibration. Therefore, either sensor could be used to collect accurate θ in the field.

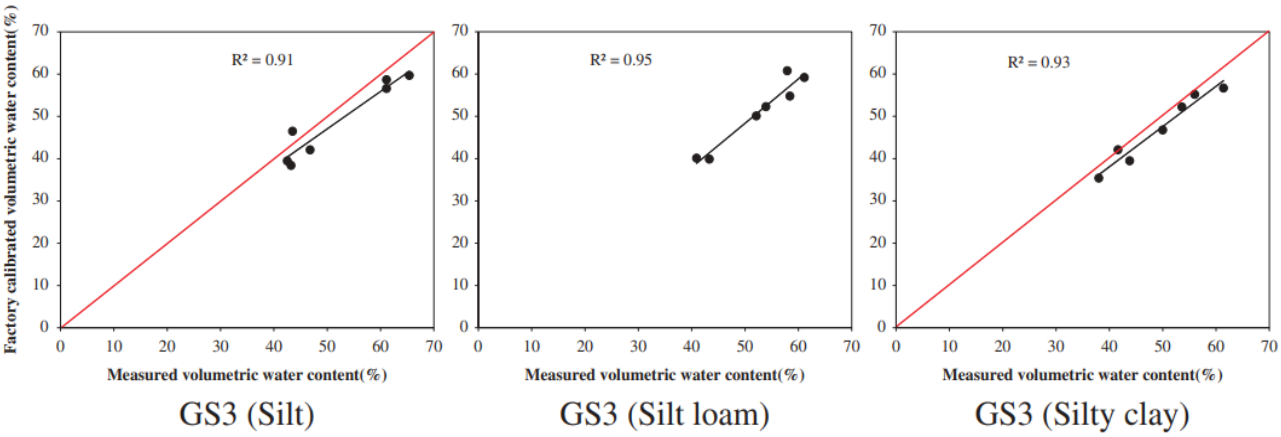


Figure 2.5 Field comparison of the factory-calibrated and measured θ results for the Decagon GS3 sensor (after Son et al., 2017).

2.3.4 Capacitance probes

Capacitance probes are another type of FDR sensor that work based on substantial difference in K_a between water with soil (i.e., 2-4) and air (i.e., 1). This instrument has been proposed by Dean et al. (1987) for in-field θ measurement. It has been used for continuous (Cheng et al., 2013) θ monitoring and at various depths (Paltineanu and Starr, 1997; Fares and Alva, 2000). A capacitance probe usually consists of several sensor rings mounted on a vertical probe (Figure 2.6) column, which is installed in the soil via a waterproof access tube (Figure 2.6a). Each sensor consists of an oscillator and a number of electrodes either two circular rings or an array of parallel metal spikes. The soil adjacent forms a dielectric of capacitance, which completes an oscillating circuit.

Capacitance probes have some advantages over TDR instruments. Most sensors operate at low frequencies (<100MHz) and can therefore detect "bound" water in fine particle soil. This bound water is strongly attracted to surface of soil particles and can constitute more than 10% soil θ. Much of it is available to plants but is not detected effectively by TDR systems operating frequencies (> 250MHz). Therefore, less critical measurements required than TDR, and no specific knowledge of analysing waveforms is necessary. While capacitance probes provide accurate measurements, many of them are required to be installed across larger spatial

extent to monitor θ dynamics since they provide a point measurement. Furthermore, soil texture and salinity can affect the measurements of capacitance probes more than TDR or neutron probe (Baumhardt et al., 2000). Therefore, empirical calibration is required. Also, installation is critical because it requires good contact with soil.

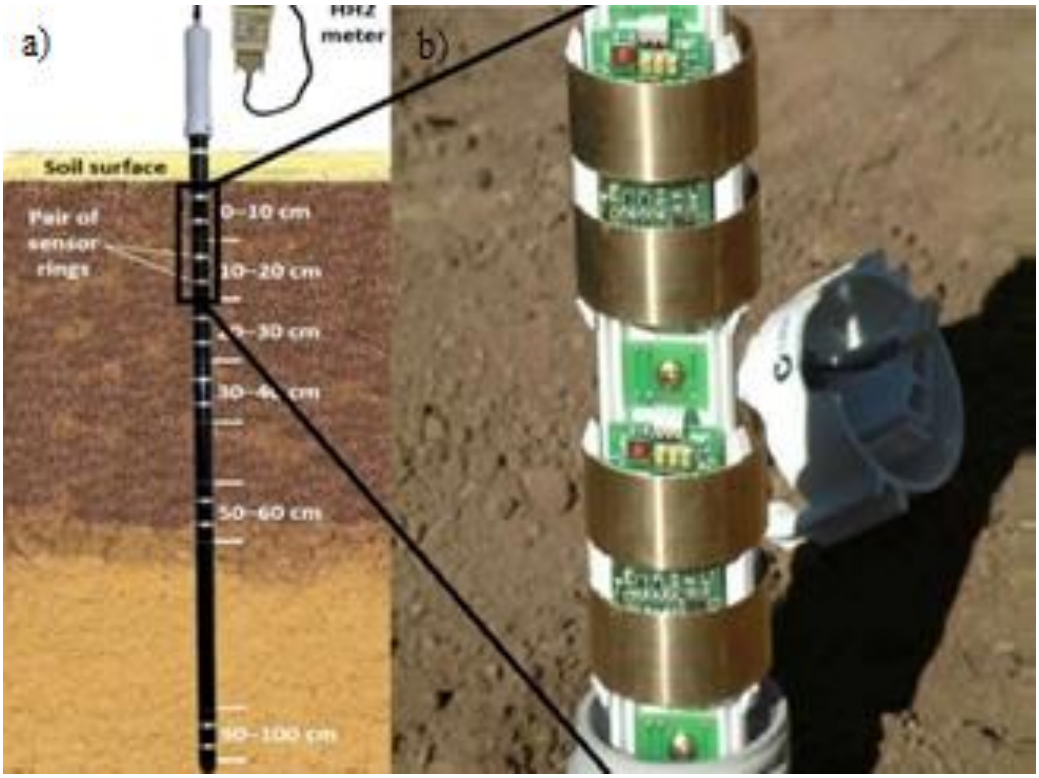


Figure 2.6 a) Capacitance probe installed in access tube and b) close-up of capacitance rings (after Dhakal et al., 2018)

Nevertheless, capacitance probes have been used widely for soil moisture monitoring and irrigation optimization studies. For example, Fares and Alva, (2000) used capacitance probes to determine irrigation set points under cropped field conditions. Moreover, they aimed to evaluate the performance of capacitance probes for citrus irrigation scheduling in deep sandy soils. They monitored θ using EnviroSCAN capacitance probes (Sentek Pty Ltd, Environmental Innovations, Kent Town, South Australia). They installed one PVC access tube for each probe with capacitance sensors at 10, 20, 40, 70, and 110 cm depths. In total they placed three probes at the drip line of three randomly selected trees and monitored θ during and after two long irrigation events (total of 19 h) in this experiment.

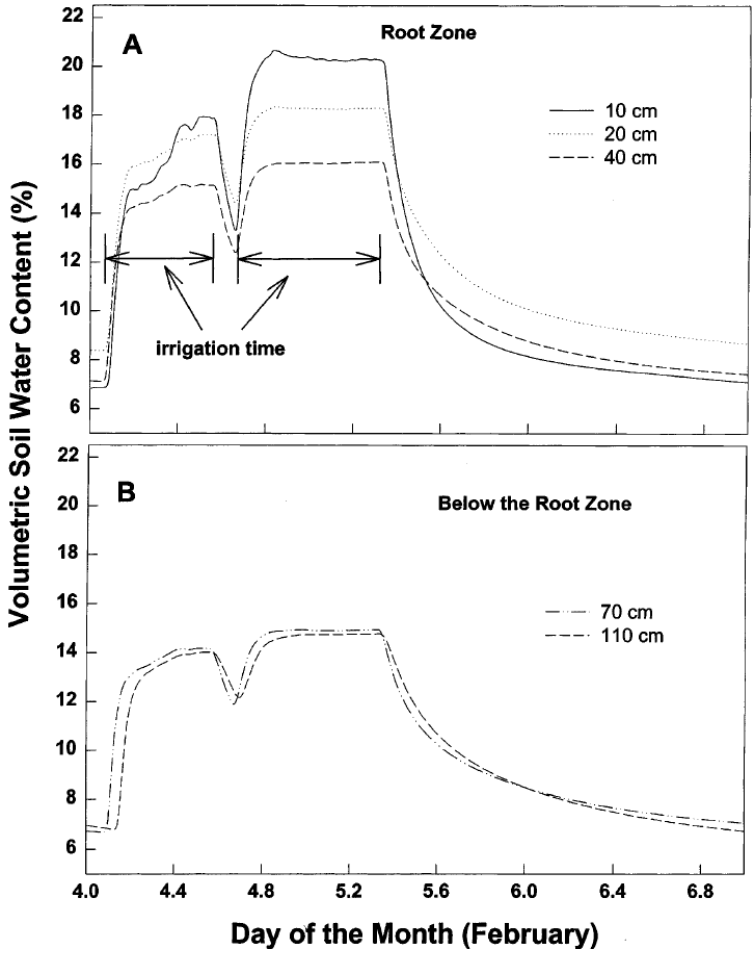


Figure 2.7 Volumetric soil water content (θ, m^3 m^{-3}) at various depths of Candler fine sand following 133 mm irrigation in two consecutive events (after Fares and Alva, 2000).

Their comparison between θ determined by EnviroSCAN and by the gravimetric method showed no significant differences. This was expected for them because the soil was relatively homogenous within the study soil profile. Their results indicated the greatest rate decrease in θ was observed in the top 10 cm of soil (Figure 2.7A). In addition, after irrigation ceased, water continued to move down the soil profile and the top part of the profile was continuously losing water to the soil below it. The lower soil horizons were leaching water into the horizon below (70-110 cm) but at the same time were receiving water that had drained from the horizon immediately above (Figure 2.7B), which resulted in lower rates of decrease in θ than that of the top horizon (10 cm). Their results also showed, the changes in water storage in the soil below the rootzone (40 cm) gave an indication of the effectiveness of irrigation, since any increase in water storage below the rootzone was due to excess water leached from the

rootzone. Considering lag time, variations in water storage below the rootzone followed patterns similar to that of the water storage in the rootzone.

They concluded that it was possible to monitor θ in the rootzone during and after two irrigation events. Although θ in the rootzone exceeded the full point during periods of high irrigation, it drained rapidly within 24-48 h after the end of such irrigation events. Using soil moisture depletion in the rootzone during periods of low water application to estimate citrus evapotranspiration (ET), the calculated daily average ET during 10-day period was 1.33 mm day^{-1}. While these methods have advantages and disadvantages and provide good measurement of θ, these point measurements can be value added using digital data and mathematical models.

2.4 Geophysical methods for θ prediction

To value add to point measurements of θ, geophysical methods such as direct current resistivity (DCR), electrical resistivity tomography (ERT), ground penetration radar (GPR) and electromagnetic induction (EMI) have showed great potential for mapping the transient nature of θ in space or time and specially across larger spatial extends. The reason for that is, field methods usually provide point information, and the measurements are not transferable to larger scales and similarly, laboratory measurement are too time consuming and labour intensive. As such, geophysical methods are being applied to value add to this limited data. The main elements of these methods are compared in Table 2.2.

Moreover, gamma-ray (γ-ray) data can also potentially be used to make DSM either alone or in combination with these geophysical data (e.g., EM data) to map FC, PWP and AWC (by difference). In the following section, the theory of operation of these geophysical methods are briefly introduced, along with their advantages and disadvantages in terms of invasiveness and efficiency. The theory of operation, as well as some of the commercially available instruments available, are briefly introduced.

Table 2.2 Summary of various geophysical methods used for measuring soil water content (after Huang 2017).

Instrument	Measuring property	Accuracy	Limitation	Invasive	Efficiency	Mobile	Hazard	Price
DCR and ERT	Resistivity	Intermediate	Affected by soil types	Yes	Low	No	No	Intermediate
GPR	Dielectric constant	Intermediate	Not applicable for conductive soils	No	High	Yes	No	High
EMI	Apparent electrical conductivity	Intermediate	Affected by soil texture, temperature, and salinity	No	High	Yes	No	Intermediate

2.4.1 DCR and ERT

Direct current resistivity (DCR) is one of the most experimented geophysical methods for geological and hydrological studies. The basis of DCR is based on Ohm's law, which states that when a voltage (V) source is applied the amount of current (I) in a conductor is proportional to the amount of resistance (R) and between two points:

$$V = I \times R$$

The constant of proportionality is known as resistance (R-Ω) and is a function of a conductor's length (L), cross-sectional area (A), temperature and the inherent material resistivity (ρ, Ωm):

$$R = \rho LA$$

The later attribute allows for determination of the ground material by DCR. Generally, a DCR survey is conducted by injection of direct current (DC) using two current electrodes, labelled as C1 (source) and C2 (sink) and two potential electrodes (P1 and P2) used to measure the potential difference (Figure 2.8).

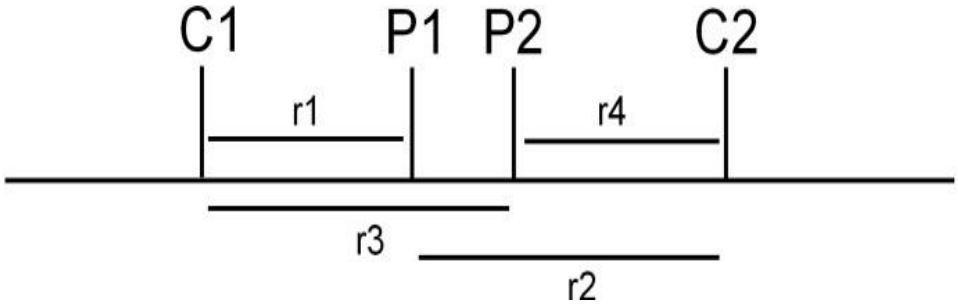


Figure 2.8 Schematic representation of electrode configuration used in resistivity surveys.

Based on that, the apparent resistivity (ρ_a) can be determined as follows:

$$\rho_a = \frac{2\pi\Delta V}{I}\left[\frac{1}{\frac{1}{r1}-\frac{1}{r2}-\frac{1}{r3}+\frac{1}{r4}}\right] = \frac{K\Delta V}{I}$$

In which K is the geometric factor and can be calculated as follow:

$$K = 2\pi\left[\frac{1}{\frac{1}{r1}-\frac{1}{r2}-\frac{1}{r3}+\frac{1}{r4}}\right]$$

However, the value of ρ_a only provides information about the average resistivity between C1 and C2. Therefore, this value should be inverted to get the estimates of true resistivity which provides information about the distribution of resistivity with depth and helps to determine the variability in the ground material. Several computer software packages have been developed to provide estimates of true resistivity. These software packages are very useful, especially when presence of moisture is being studied. Using these software packages, a weighted average of the resistance of earth materials to current flow can be translated to the spatial distribution of soil ρ_a (Figures 2.9B) and then converted to θ using Archie's law (Archie, 1942).

Electrical resistivity tomography (ERT) is a form of DCR, which is widely used for in-field continuous θ dynamics monitoring. A typical ERT instrument for carrying out θ dynamics monitoring consists of a current source, ammeter, voltmeter, electrodes and cables. Most common ERT survey employ multi-electrode systems (Figures 2.9A) for 2- or 3-dimensional electrical imaging. This is because ERT enables automatic collection of multiple resistance data by creating an electrical gradient between two source electrodes and measuring the resultant potential distribution at two or more receiving electrodes along lines of equidistantly spaced electrodes which provides spatially comprehensive data along transects (Wellbrock et al., 2018).

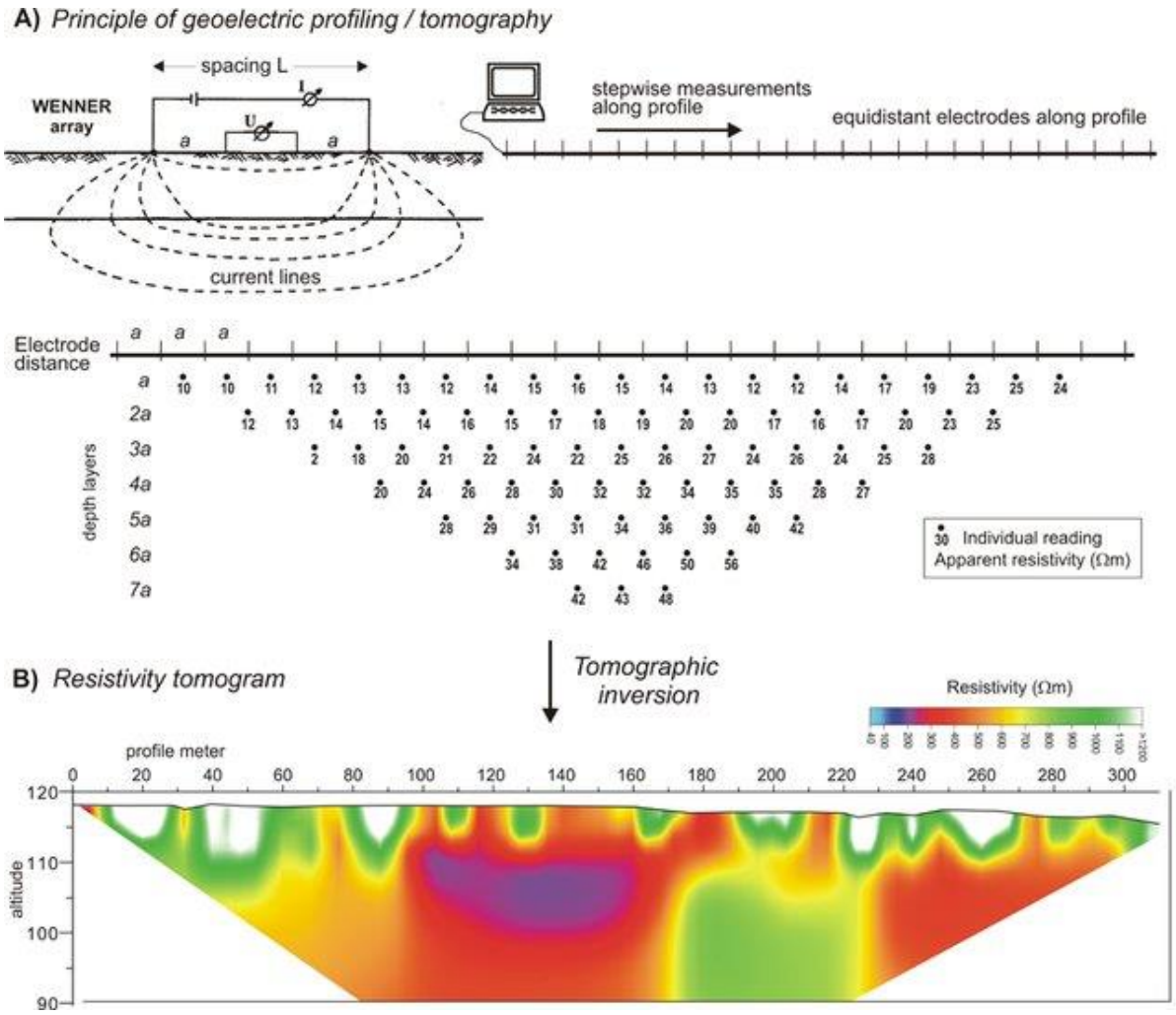


Figure 2.9 Geoelectric profiling/electrical resistivity tomography (ERT). A) Principle of measurement with 4-electrode array (left). I: current, U: voltage. Measurement is automated along a profile with fixed electrodes at equidistant spacing. Result of measurement: two-dimensional section of apparent resistivities. a: spacing between two adjacent electrodes. B) Tomographic inversion gives the final resistivity tomogram for interpretation. (after Wellbrock et al., 2018).

2.4.2 GPR

Ground penetration radar (GPR) works based on transmission and reflection of an electromagnetic wave in the soil (Chanzy et al.,1998, Figure 2.12). Therefore, it does not require direct contact between the instrument and the soil. Like TDR, the K_a of the soil properties influences the propagation velocity of the electromagnetic waves of GPR. Hence, θ plays a significant role here (Urliksen, 1982). Theoretically, GPR has the potential to provide rapid measurement of θ by mounting on a mobile system across large spatial extent.

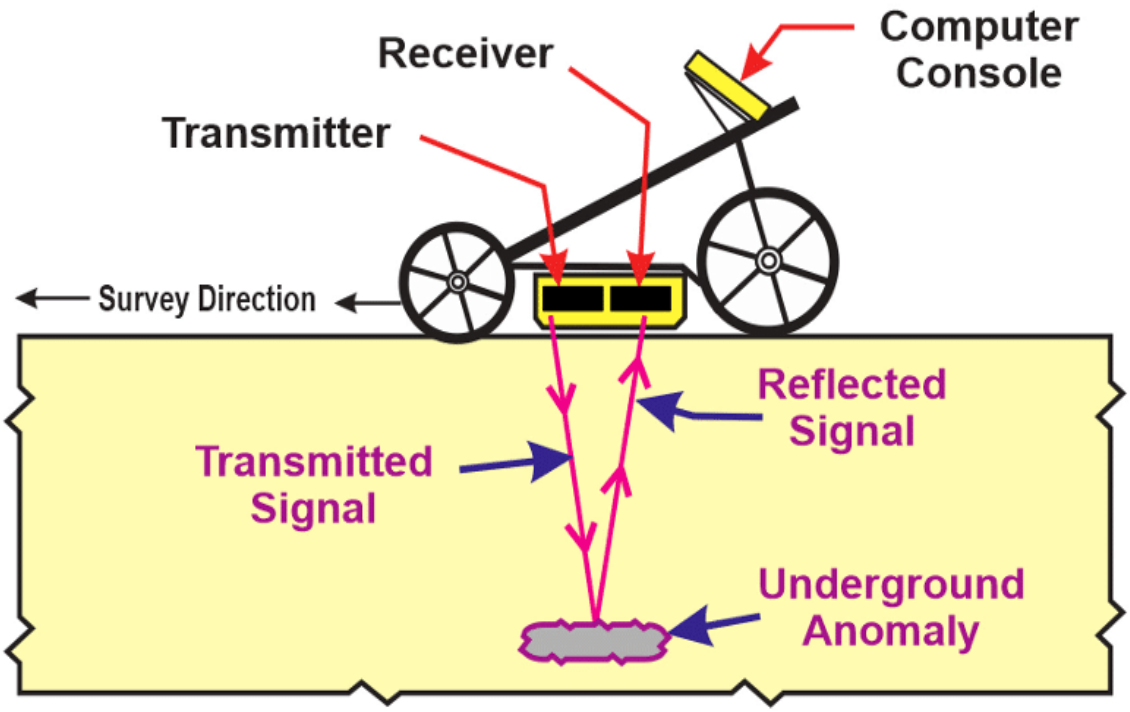


Figure 2.10 Schematic GPR system mounted on a cart (accessed from https://www.topographix.com/ground-penetrating-radar/).

A selection of environmental and climatic conditions limits the use of GPR, however, including; Landscape condition and vegetation cover can lead to poor performance of GPR; High conductive areas can decay the radar energy (Doolittle and Collin, 1995); Large trees reflect the electromagnetic waves (Schrott and Sass, 2008); Large-size antennas are required in steep and rocky slopes (Dobriyal et al., 2012). Thus, when it comes to choosing GPR for θ measurements selecting and verifying the origin and depth needs careful consideration.

2.4.3 Electromagnetic induction

Since the 1970s, electromagnetic (EM) induction instruments have been broadly utilised by researchers in soil and hydrological studies. This is because, EM instruments can measure the apparent electrical conductivity (EC_a, m Sm^{-1}) of the soil which is influenced by salinity, clay content, type of clay and water content. The EM instruments contain a transmitter coil along with either of one or multiple receiver coil(s). The orientation of the receiver coil can be horizontal or perpendicular to the transmitter. The transmitter coil which is located at

one end of an instrument requires an alternating current to be introduced. This creates a primary electromagnetic field which penetrates the soil and the regolith. According to Faraday's electromagnetic theory, in the presence of conductors in the soil a secondary electromagnetic field will be induced by circular eddy current loops (Figure 2.14). A receiver coil, located at the other end of an instrument, senses both primary and secondary magnetic fields. McNeill (1980), describes how a depth weighted EC_a is measured by the instrument as follows;

$$EC_a = \frac{4}{\omega \mu_0 s^2} \left| \frac{H_s}{H_p} \right|$$

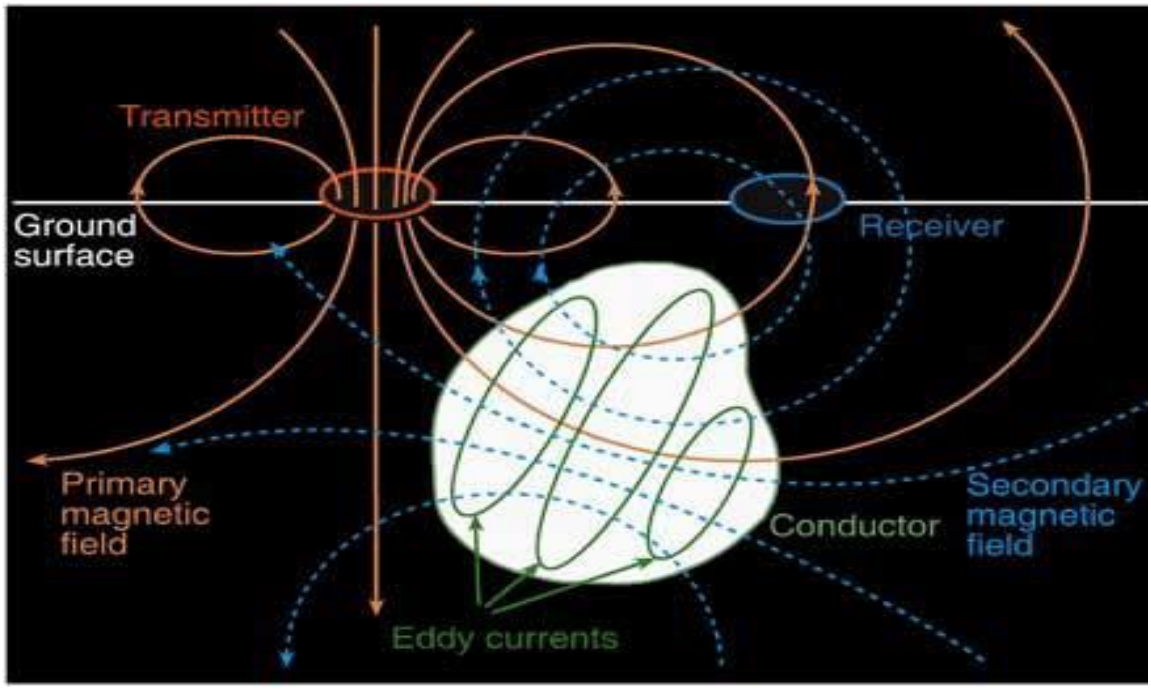


Figure 2.11 Principles of EM induction (accessed from http://www.terragis.bees.unsw.edu.au).

where ω is angular frequency of the instrument, μ_0 is the vacuum permeability, s is coil spacing between the transmitter and the receiver and H_p and H_s are the primary and secondary magnetic field, respectively. The height of the instrument, the distance and orientation of the transmitter and receiver coils affect the secondary field (Hendrickx and Kachanoski, 2002).

Additionally, as a function of soil properties, the amplitude and phase of the secondary field vary from those of the primary field. Thus, EM instruments have been used to map various

soil properties directly (e.g., salinity, clay content, cation exchange capacity and water content) and indirectly (saturation percentage, bulk density, organic matter) (Corwin and Lesch, 2005; Doolittle and Brevik, 2014).

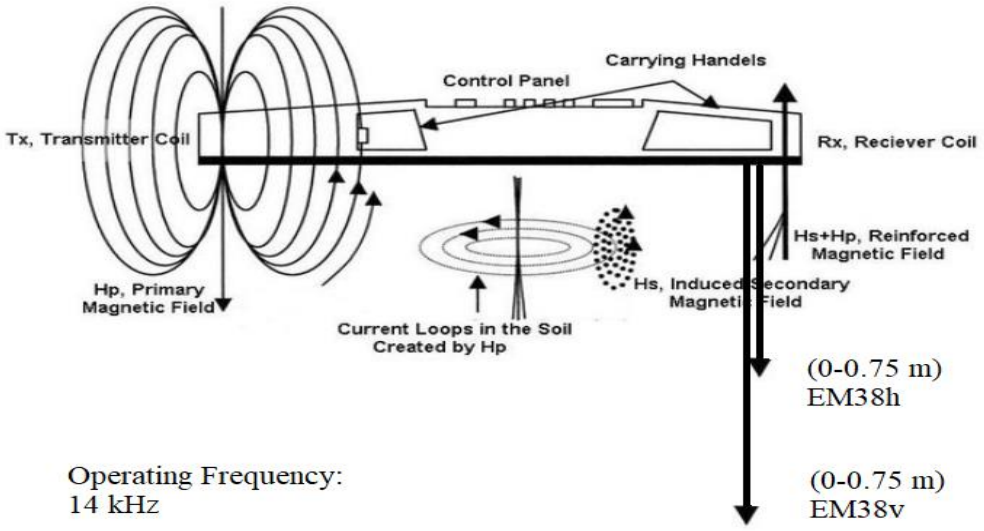


Figure 2.12 Schematic diagram of the EM38, which is 1 m in length. Tx is the transmitting coil and Rx is the receiving coil (after Robinson et al., 2004).

One of the first manufacturers that developed EM instruments for proximal sensed soil studies is Geonics Limited (Geonics Limited, Mississauga, Ontario, Canada); The most popular instruments include; EM31, EM34-3, EM38, and more recently EM38-MK2. Among them EM38 (Figure 2.12), has been extensively used for soil moisture monitoring. The transmitter of this instrument operates at 14 kHz and has one receiver 1 m apart from the transmitter and in analogous orientation. The instrument has two modes of operation. The depth of exploration (DOE) for horizontal mode (EM38h) is 0-0.75 m, while the DOE for vertical mode (EM38v) is 0-1.5 m (Table 2.2). Given the DOEs of this instrument, EM38's have been widely used for characterizing the variation in soil properties within root-zone.

Another Geonics Limited EM induction instrument which has been used for hydrogeological interaction studies is EM34-3 (Figure 2.13). It has two independent coils with 3 inter-coils spacings which are 10 m, 20 m and 40 m, respectively. The inter-coil spacing determines the frequency at which the instrument works. For 10 m spacing it is 6.4 kHz and it

is 1.6 kHz and 0.4 kHz for 20 m and 40 m spacing, respectively. This instrument also has two modes of operation (HMD and VMD). In VMD, measurement depth is greater compared to HMD mode of operation, and can explore to a maximum depth of 0-60 m. The DOEs varies with the inter-coil spacing and the mode of operation as shown in Table 2.3.

Table 2.3 Parameters of the Geonics EM conductivity meters

Instrument	Spasing (m)	Frequency (kHz)	DOE (m)	
			Horizontal	Vertical
EM38	1	14	0-0.75	0-1.5
EM 34-3	10	6.4	0-7.5	0-15
EM 34-3	20	1.6	0-15	0-30
EM 34-3	40	0.4	0-30	0-60

Figure 2.13 EM34-3 Geonics Limited, Mississauga, Ontario, Canada.

DUALEM Inc. (DUALEM Inc., Milton, ON, Canada) is another manufacturer of frequency domain EM instrument.One Unlike early Geonics models, DUALEM instruments are designed to enable collection of EC_a data using a multi-coil receiver array. For example, a DUALEM-421 instrument simultaneously measures EC_a to six different depths; enabled by three pairs of receiver coils (Rx) that are situated; 1, 2 and 4 m apart from the transmitter coil (Tx, operates at 9.0 kHz) (Figure 2.14). The orientation of the two receiver coils in each pair is

opposing. The transmitter coil and one of the receiver coils have horizontal windings which form a horizontal coplanar array (HCP). Whereas the other coil in each receiver pair, have vertical windings with the transmitter coil and receiver forming perpendicular arrays (PRP). The notations Hcon and Pcon represent EC_a of the receiver pairs with the 1mHcon and 1mHPcon corresponding to DOEs of 0-1.5 and 0-0.5m, respectively. The 2mHcon and Pcon correspond to DOEs of 3 and 1 m, respectively while 4mHcon and 4mPcon correspond to DOEs of 6 and 2m, respectively (Table2.4).

Table 2.4 Parameters of DUALEM sensors (DUALEM-421S Manual, 2008).

Instrument	Spasing (m)		DOE (m)	
	Pcon	Hcon	Pcon	Hcon
DUALEM-1	1.1	1	0-0.5	0-1.5
DUALEM-2	2.1	2	0-1.0	0-3.0
DUALEM-4	4.1	4	0-2.0	0-6.0

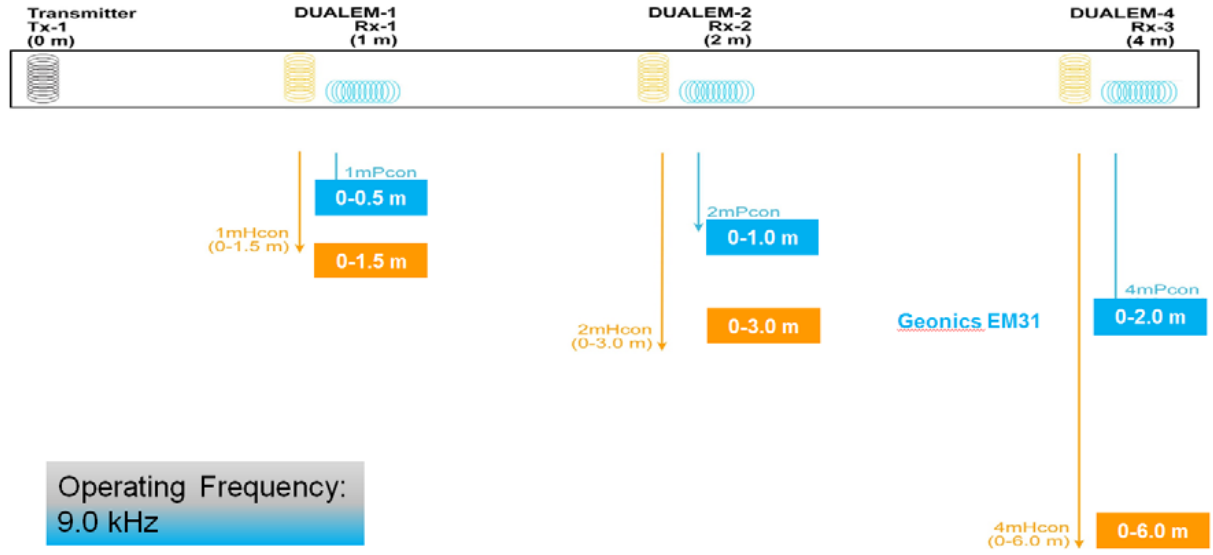


Figure 2.14 Schematic diagram of the DUALEM-421(DUALEM Inc., Milton, ON, Canada), which is 4m in length (after Triantafilis et al., 2011).

2.4.1 Gamma-ray (γ-ray) spectrometry

Gamma-ray (γ-ray) spectrometry, or radiometrics, measures the natural radiation in the earth's surface. Geologists/geophysicists routinely use (γ-ray) spectrometry surveys as mapping tools to identify where certain rock types change. Increasingly it is being used by geomorphologist and soil scientists to understand the distribution of unconsolidated sediments

and soil types, respectively. This is being facilitated by remote (airborne surveys) and proximal survey instruments. This is because g-ray spectrometry is a passive method which detects natural radiation. There are many isotopes of various elements which are unstable in nature and show radioactivity. The three most useful elements include, potassium (K), uranium (U) and thorium (Th). Typically, isotopes of these elements first undergo an alpha (α) or beta (β) decay.

An α decay or α ray is like a helium (He) nuclei being emitted from a nucleus. Owing to their atomic mass (i.e., two protons and two nuetrons) and charge (i.e., +2), these particles are absorbed within a relatively short distance in air (i.e., a couple of cm's). One element that undergoes this form of decay is uranium (i.e., $^{238}U_{92}$), whereby the α-particle ($^{4}He_{2}$) emitted leaves behind a thorium ($^{234}Th_{90}$) atom, with the decay summarised as follows:

$$^{238}U_{92} \rightarrow {}^{234}Th_{90} + {}^{4}He_{2}$$

A β decay results in a net negative charge (+1) increase in a nucleus with β-particles (of negative charge) not as readily absorbed and travelling further in air (i.e., up to 1 m). They usually result because a radioelement nucleus remains unstable after an α-decay. Specifically, a neutron decays into one of the following three products, including a proton, electron or neutrino. An example of β-decay is when thorium ($^{234}Th_{90}$) converts into protactinium ($^{234}Pa_{91}$) by increasing in atomic number by one (i.e., 91) with an electron emitted along with a neutrino as follows:

$$^{234}Th_{90} \rightarrow {}^{234}Pa_{91} + e + n_e$$

Moreover, the unstable $^{234}Pa_{91}$ may undergo another β-decay to become uranium ($^{234}U_{92}$), with an additional neutron converting to a proton with another electron and neutrino emitted and as follows;

$$^{234}Pa_{91} \rightarrow {}^{234}U_{92} + e + n_e$$

However, $^{234}U_{92}$ is still in an excited state. In this regard, a γ-ray is emitted (i.e., electromagnetic radiation) to release the excess energy and because a γ-ray has no charge or mass, it can penetrate through 0.3 m of rock and/or soil and hundreds of meters into the air.

This makes g-rays a very beneficial tool in the radiometric analysis of rocks, sediments and soil materials. A γ-ray spectrometer has the capability to sense and measure the γ-rays and consists of a detector, a photomultiplier and a spectrometer. The main part of the detector is a transparent crystal of sodium iodide (NaI) (Figure 2.15). The γ-rays which are emitted from the ground surface elements, absorbs in the NaI crystal, and stimulate a little pulse of light. Here a photomultiplier tube measures the emitted γ-rays by converting the light flash to a voltage of a proportional intensity.

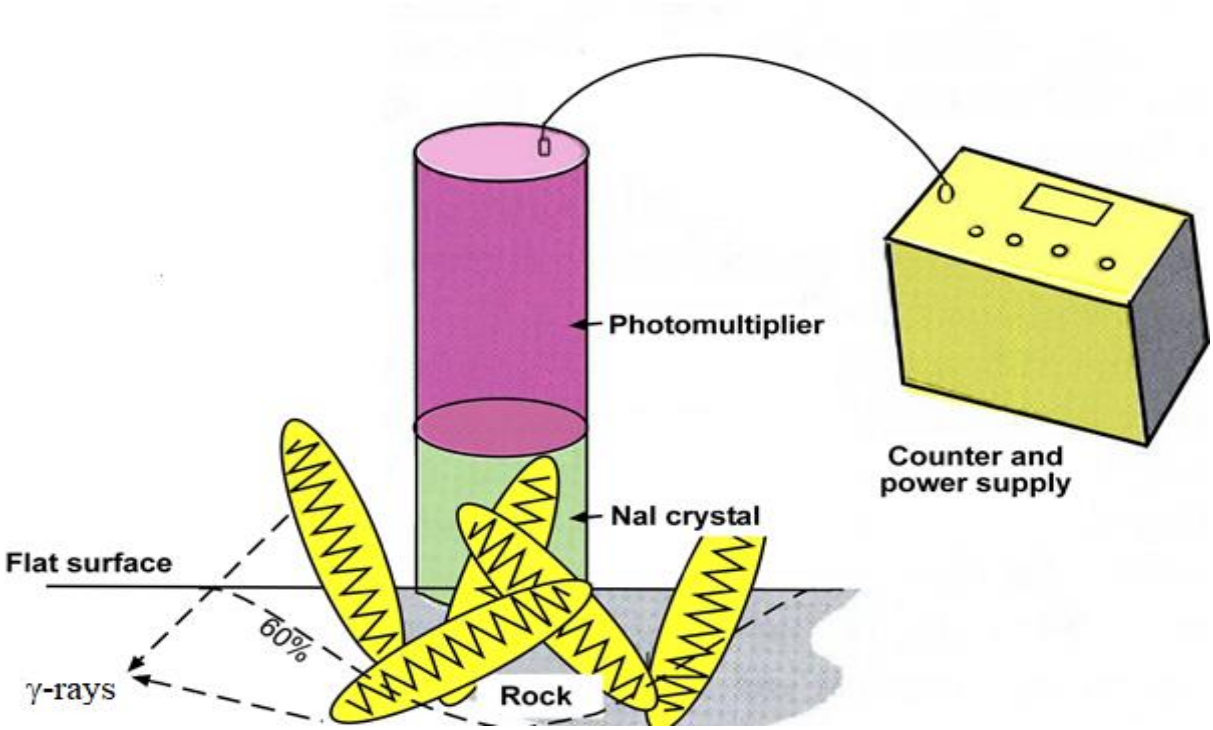


Figure 2.15 Components of a γ-ray spectrometer and including the detector and photomultiplier chamber and counting device (i.e., spectrometer).

The counting device separates this voltage into several magnitude-dependent classes that represent the energy spectrum of γ-rays. These classes are based on selectivity record windows recommended by the International Atomic Energy Agency (Grasty, 1991) an as; 1.37-

1.57 MeV for K, 1.66-1.86 MeV for U, 2.41-2.81 MeV for Th and 0.41-2.81 MeV for the total counts (TC) (Figure 2.16).

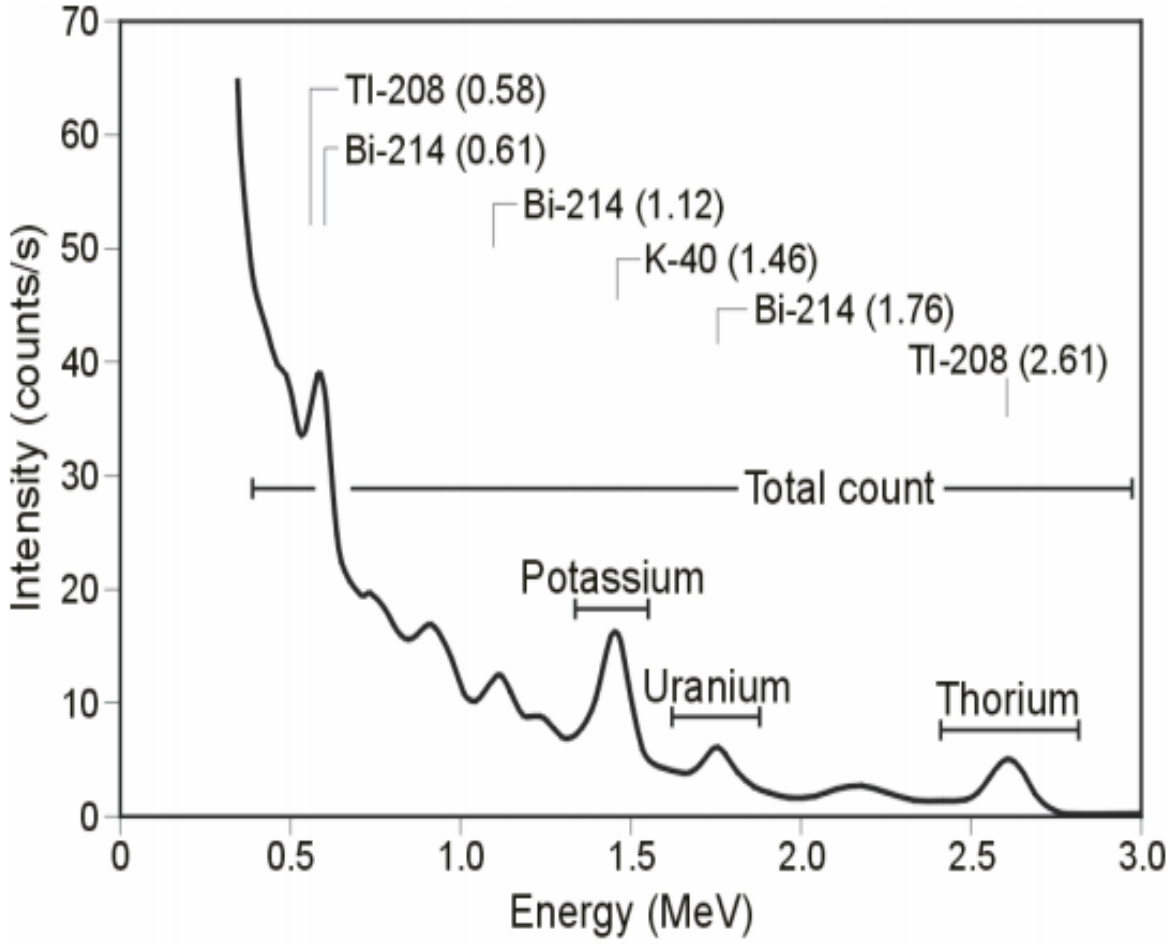


Figure 2.16 Recommended γ-ray spectrum windows (after Grasty et al., 1991).

The measuring unit for the total count (TC) is counts per second (cps) and parts per million (ppm) for K, U, and Th. However, percent (%) is also used for K as well. The depth of penetration of γ-ray is around 0.4 m (Jaques et al., 1997) from soils with a density of approximately 1.5 g cm^{-3} (Grasty, 1975). Various types of soils and rocks can be discovered based on the radiometric values, since within small areas γ-ray concentration can differ greatly. Table 2.5 shows the average radioelements of the soils and rocks in Australia.

Table 2.5 Average value radioelement content of Australian rocks and parent materials of soils (after Dickson and Scott, 1997).

Rock type	Rock			Soil		
	K (%)	U (ppm)	Th (ppm)	K (%)	U (ppm)	Th (ppm)
Intrusives						
granitoids	2.4	3.3	16	2.1	2.7	13
gneissic rock	2.4	2.5	15	1.3	1.3	12
pegmatite	3.7	0.7	2			
quartz-feldspar porphyry	2.9	1.7	13			
Intermediate intrusives	2.7	0.8	2.4	1.6	1.9	5.6
mafic intrusives	0.4	0.3	1.2			
Extrusives						
felsic volcanics	3.7	2.4	17	2.4	2.1	13
intermediate volcanics	2.7	2.3	9	1.9	2.1	10
low-K andesites	0.8	1.6	5	1.1	1.3	5
mafic volcanics	0.9	0.7	3.0	0.7	1.6	7.9
Ultramafic volcanics	0.4	0.6	1.2	0.6	2.0	6
Sedimentary rocks						
Archaean shales	0.9	0.9	2.7	0.8	1.2	3
other shales	2.6	2.6	19	1.5	2.3	13
carbonates	0.2	1.6	1.4			

The γ-ray concentrations of soils are generally correlated to the corresponding parent rocks. For example, radiometric values are comparatively large both in granitic rocks (i.e., K-2.4 %; U-3.3 ppm and Th-16 ppm) and soils (i.e., K-2.1 %; U-2.7 ppm and Th-13 ppm). Moreover, it is apparent from the table that K (%), U (ppm) and Th (ppm) are large in felsic volcanic rocks and soils. Contrarily, small radiometric values are associated with mafic extrusive rocks (i.e., K-0.4 %; U-4 ppm and Th-1.2 ppm) and soils of mafic parent material (i.e., K-0.6 %; U-2.0 ppm and Th-6 ppm).

2.5 Application of geophysical methods

2.5.1 Application of DCR and ERT

One of the most well-established techniques for geologic, hydrologic and engineering studies is DCR. It has been used widely for groundwater investigations as well as locating the geological structures which regulate the movement of groundwater (Gibson and George, 2004). Moreover, several environmental topics such as; detect the position of former waste disposal

sites (Chambers et al., 2006), locate pollutant pathways (Rucker et al., 2009), detect salt water intrusions (Ingham et al., 2006), determine the depth of bed rock (Zhou et al., 2000) and detection of cavities and cave systems (Metwaly and AlFouzan, 2013) have been investigated using this technique. Increasingly, and to investigate dynamics in θ, ERT has been used.

The accuracy of DCR and ERT methods is not as good as the above-mentioned methods for in-situ θ measurement. However, they have been broadly used for agricultural water management due to their capability of θ dynamics monitoring across large spatial extent at different depths without permanent installation and disturbance of the soil profile. User-defined flexibility (Furman et al., 2004) in terms of number of electrodes, direction and configuration is another advantage of these methods. For example, Kelly et al. (2011) inferred the spatial heterogeneity in θ along and beneath an irrigated cotton field using DCR before and after an irrigation event. They did this by installing 32 electrodes with 1 m spacing, to achieve an inverted model depth of approximately 5 m. They collected the DCR data on two days continuously using an automatic meter and used a capacitance probe to collect the calibration data at one point within the field at depths of 0.2, 0.3, 0.4, 0.6, and 0.8 m.

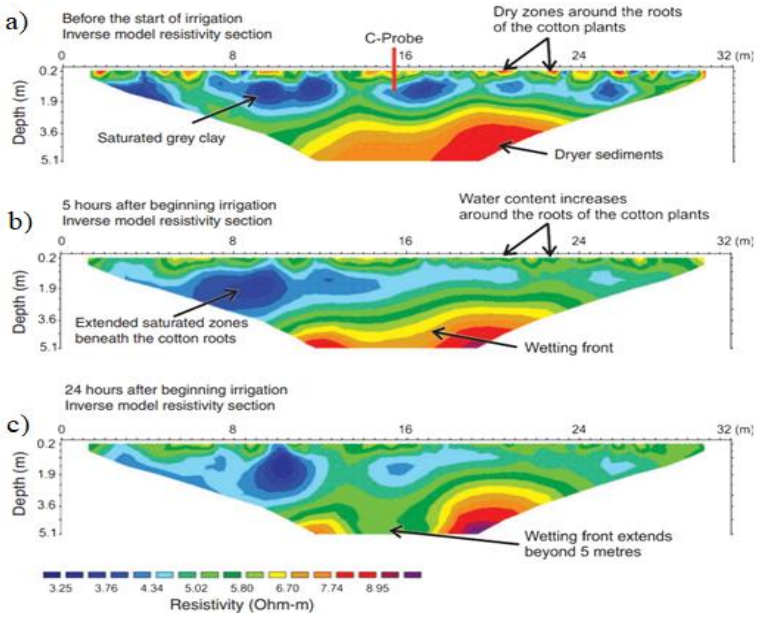


Figure 2.17 ERT time-lapse results of deep drainage beneath a cotton crop a) before b) 5 hours and c) 24 hours after flood irrigation in furrows (after Kelly et al, 2011).

Their results showed, dry soil zones, indicated by the orange/red semi-circles at shallow depths, highlight the positions of some of the cotton plants (Figure 2.17a). Also, from Figure 2.17a, it can be observed that the capacitance probe (C-probe) was positioned in a portion of the field that was relatively dry in the upper 0.6 m and moist at the base of the C-probe. Had the C-probe been placed at 13.5 m (wetter in the shallower profile and drier at the base) or 12 m (wetter throughout the profile), the irrigation scheduling would have been very different. Figure 2.17b shows the ERT results 5 hours after the start of irrigation. The wetting of the upper soil is indicated by the diminishing size of the semi-circles around the roots of the plants (more of the shallow portion of the image is green). The wetting front is indicated by the downward extension of the zone where the soil resistivity is less than 6 ohm-m (green). Figure 2.10b also shows that the wetting front, which indicates deep drainage, has reached a depth of approximately 4 m after 5 hours, and extends beyond 5 m after 24 hours (Figure 2.17c).

They concluded that ERT images provide information on the uniformity of the soil/sediment at depth along the furrow, detect major zones of infiltration associated with cracking clays or palaeochannels, and differentiate between clayey and sandy soils in areas of uniform pore water salinity. Their study also shows the importance for an electromagnetic or resistivity survey prior to the placement of the C-probe in order to determine a representative location for the moisture measurements that are used to schedule irrigation.

In another study, Zhou et al. (2001) installed 88 electrodes to generate 3-dimensional time-lapse images of θ across a 3.5 m by 3.5 m block after a rain event. They conducted an inversion calculation of the 3-dimensional soil resistivity based on these potential data. They also collected 46 soil samples at representative positions in the square and measured the parameters in the Archie's relationship in the laboratory. Then, they obtained 3-dimensional distributions of the parameters by a distance weight interpolation method. Finally, based on

Archie's relationship and the 3-dimensional distribution of the soil resistivity and the related parameters, they calculated the 3-dimensional distributions of θ.

Figure 2.18 shows an example of their 3-dimensional predicted θ. Each image represents a different depth range. For the depth range of 0.0-0.25 m, there is a relatively higher θ region in the northwest area. When depth increases, this relatively higher θ region gradually expands. Then, at the southwest area of 0.5-0.75 m depth range, there appears another relatively higher θ region. With further increasing depth the relatively higher θ region in the northwest area gradually moves toward the north, while that in the southwest area seems to remain constant. These results clearly show that θ not only changes in the vertical direction but also varies in a horizontal plane. They attribute this to the spatial heterogeneity nature of the soil of their study block.

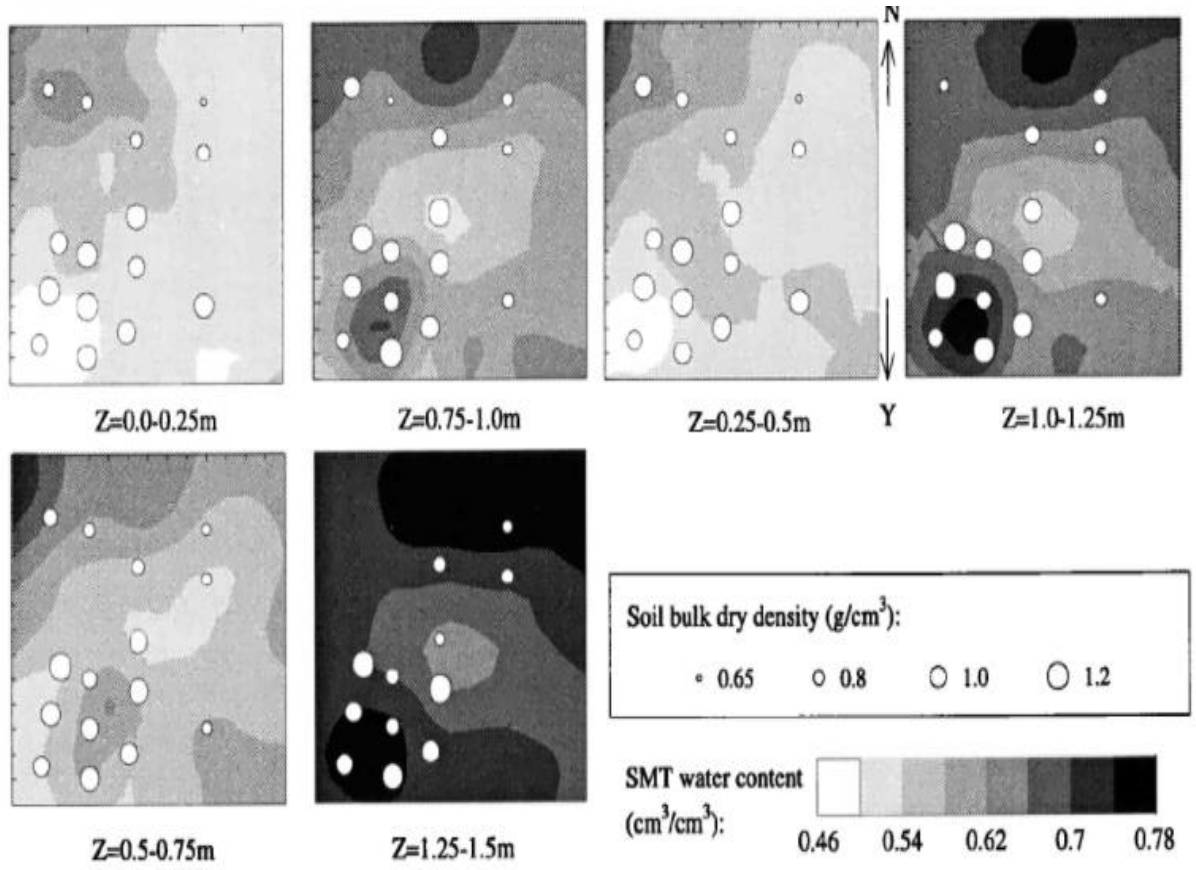


Figure 2.18 The 3-dimensional distribution of calculated θ, and the spatial distribution of soil bulk dry density. The soil bulk dry density data came from the linear interpolation of the measured density profiles (after Zhou et al, 2001).

They concluded that there are errors (0.10 m^3 m^{-3}) in the calculated θ, but between the calculated and the measured θ there exists a good linear relationship; It is possible to use the calculated θ to analyse the very general 3-dimensional distribution characteristics of the soil moisture and investigate the 3-dimensional rainfall infiltration process, the redistribution of soil water after rain, and other hydrological processes in the field. Nevertheless, the proposed method is preferred for porous media where the water resistivity is relatively stable.

While the DCR and ERT approach is useful for field θ monitoring, the technique requires physical insertion of steel electrodes that need to be connected using wires. This is time consuming and expensive. Moreover, the equipment once installed cannot be mobilised during the experiment. Therefore, obtaining information across larger spatial extends is challenging.

2.5.2 Application of GPR

The GPR method is rapid and continuous and provides in situ measurements of underground media; consequently, it has been applied in a variety of disciplines, such as hydrology (Algeo et al., 2016), hydrogeology (Annan, 2005), archaeology (Goodman and Piro, 2013), glacier and frozen soil research (Woodward and Burke 2007), soil science (Wang et al, 2016), and environmental science (Hubbard et al., 2005). The GPR has also been used to determine the shallow soil water content at the field scale (Galagedra et al., 2003). Although, GPR does not directly measure θ, it has shown potential to describe the relationship between θ and the soil dielectric permittivity (Husiman et al., 2003).

For example, Galagedra et al. (2003) compared GPR-measured θ data with TDR-measured data using 6 vertical TDR probes of different lengths. Using a regression analysis, they showed that the GPR-measured θ was not different from θ measured by TDR at 0-0.1 m.

Their results showed that the GPR direct ground wave did not penetrate beyond 0.1 m under their experimental conditions.

The study of Zhou et al. (2019) confirmed these results who conducted a laboratory approach to evaluate the use of a high-frequency GPR antenna to determine the θ of various soil textures. They collected soil samples (0-0.2 m) from six soil textures, including of loamy sand, clay, and silty loam and used a high-frequency (2-GHz) GPR antenna to determine the θ. They used GPR reflection two-way travel times (TWTT) and the known depths of reflectors to calculate the average soil dielectric permittivity above the reflectors and establish a spatial relationship between the soil dielectric permittivity and θ, which was used to estimate the depth-averaged θ.

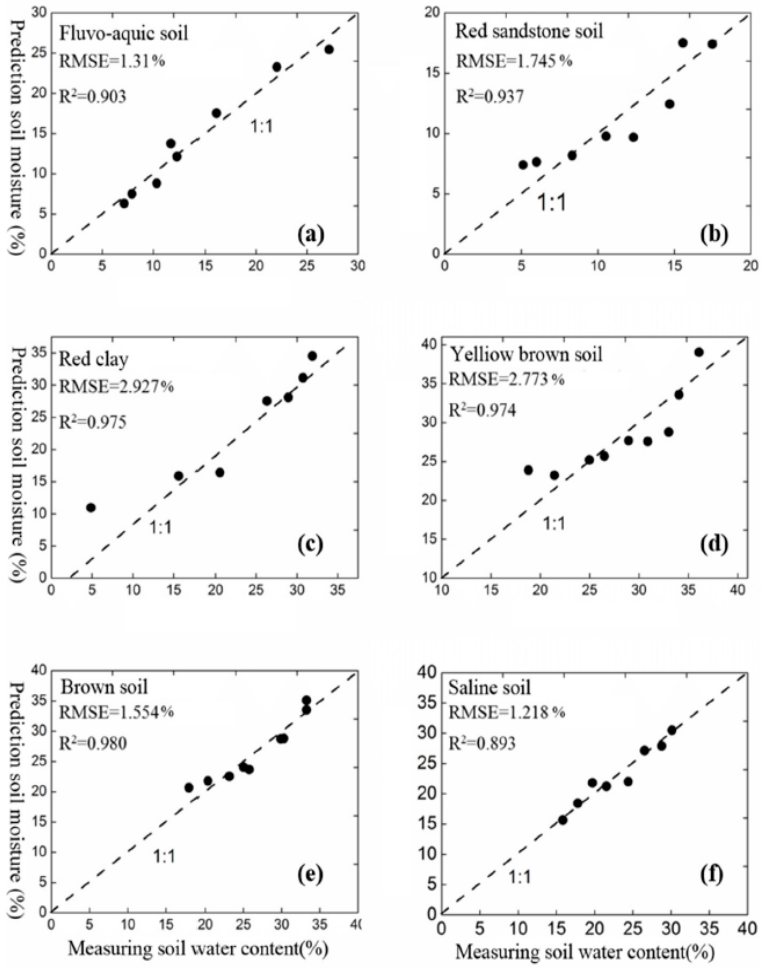


Figure 2.19 Regression analysis of the measured θ and model-retrieved θ of (a) fluvo-aquic soil (b) red sandstone soil (c) red clay (d) yellow brown soil (e) brown soil; (f) saline soil (after Zhou et al., 2019).

Their results show that the calculated and retrieved θ values of the soil types have relatively good correspondence (Figure 2.19). However, the RMSE of the red clay is the highest (2.92%, Figure 2.19c). They attribute this to the fact that the content of clay has a large influence on the soil dielectric permittivity, which increases with the soil clay content. As the clay content increased, both the real and imaginary parts of this, relative to soil permittivity increased. The reasons for the low soil dielectric permittivity at high soil clay contents are the high-volume expansion, high moisture absorption capacity, and high electrical conductivity, which result in high dielectric losses and consequently poorer prediction of θ.

They concluded that under laboratory conditions, GPR provided feasible and reliable estimates of shallow (0-0.2m) θ. However, the conductivity, magnetic soil, soil texture, soil organic matter, and soil temperature have substantial effects on the soil dielectric permittivity, which consequentially affects the prediction of θ. Moreover, using this approach in the field can be difficult because it requires the reflectors to be buried without disturbing the surrounding ground or excavating a natural object.

2.5.3 Application of electromagnetic induction in mapping θ

Volumetric soil moisture content (θ, m^3 m^{-3}) is an important factor in agriculture and environment because it controls the exchange of water and nutrients between the land and vegetation. Therefore, information on spatiotemporal variation of θ in field scale is a key variable not only for improving water use efficiencies and crop yield optimisation but for controlling the effect of fertilisers and pesticides on groundwater and environment. In practice, estimating the variation in θ from the water balance equation is required: Δθ = water inputs (Precipitation + Irrigation) – water output (Runoff + Evapotranspiration + Deep Drainage).

Moreover, θ is a dynamic factor in space and time and is also affected by soil texture (McCutcheon et al., 2006).

The first requirement for precision irrigation implementation is mapping and monitoring spatiotemporal variation of θ. Although gravimetric methods can be used for θ determination in the laboratory (Reynolds, 1970), several in situ invasive and non-invasive electrical and electromagnetic geophysical methods have drawn great attention to characterise θ in space and time (Zhou et al., 2001) (Table 2.1). They allow for measuring soil electrical conductivity (EC_a, mS m^{-1}) or resistivity (ρ, mS m^{-1}), which can be converted to θ by using different empirical (e.g., Archi, 1942; Rhoades, 1976) or mathematical methods (e.g., artificial neural networks, ANN).

Among all empirical models Rhoades et al., (1976) has been used frequently by soil scientists because of the simplicity of the model;

$$EC_a = (\alpha\,\theta^2 + \beta\,\theta)\,EC_\omega + EC_s$$

whereby α and β are characteristic soil parameters. EC_ω electric conductivity of soil water and EC_s surface conductivity. Later, this was modified by Rhoades et al. (1989) to enable a more flexible description between the relationship of EC_a and θ. Nevertheless, obtaining several numbers of empirical parameters, limit the applicability of the latter model (Friedman, 2005). For example, to obtain characteristic parameters based on soil type, requires the identification of subsoil structure. This is a labour intensive, time consuming and demanding job.

Herein, another issue might come into consideration; And this is the effect of shrink-swell characteristic of cracking clay soils (e.g., Vertosols) and/or hysteresis in water retention curve on EC_a values. Because they can be problematic in mapping various soil properties that directly (e.g., salinity, clay content, cation exchange capacity) and indirectly (saturation

percentage, bulk density, organic matter) correlated with EC_a values. However, this is not the case for mapping θ, because shrink-swell characteristic of cracking clay soils (Dinka et al., 2013) and/or hysteresis of water retention curve (Van Genuchten, 1980) are concurrent phenomena of θ changes and they are controlled by soil water content fluctuations. The electric properties are dominated by any water that may be present, and so the EC_a increases significantly with a small initial increase in water content (Nobes 1996).

In one of the primary studies, Kachanoski et al. (1988) employed a EM38 and EM31 and modelled a curve-linear relationship between EC_a and θ measured by TDR in rootzone. By using a second-order regression equation they observed a strong relationship (R^2=0.96) between θ and EC_a. Later, Kachanoski et al (1990), using a neutron probe to measure θ confirmed this relationship. They concluded that EC_a can be used to predict θ at the field scale for soils with low concentrations of dissolved electrolytes. However, these results were not quite comparable with previous research undertaken by Rhoades et al. (1976), who instead of using EM data used electrode induced EM induction for their EC_a measurements in a soil with high-level of clay content (~ 44.4%). Archie's Law (Archie, 1942) also suggests a different calibration between EC_a and θ. All these disagreements confirm that there is not a universal relationship between θ and EC_a. This is because θ is not uniformly distributed within the sample volume of the EM induction instrument (Callegary et al., 2012).

As a result, researchers have mainly used EM instruments to map θ at the field scale largely associated with a uniform soil texture or soil type. For instance, a linear regression relationship (Figure 2.20) was recognized by Sheets and Hendrickx (1995) between EM31 EC_a and average θ within the rootzone (0-1.5 m) along a transect of predominantly sandy soil. Similarly, Sherlock and McDonnell (2003) used EC_a from an EM31 and EM38 to map topsoil

(0-0.2 m) θ across a loamy texture hill slope. Likewise, Akbar et al. (2005) calibrated EM38 EC_a against soil θ in a non-saline heavy clay Vertosol.

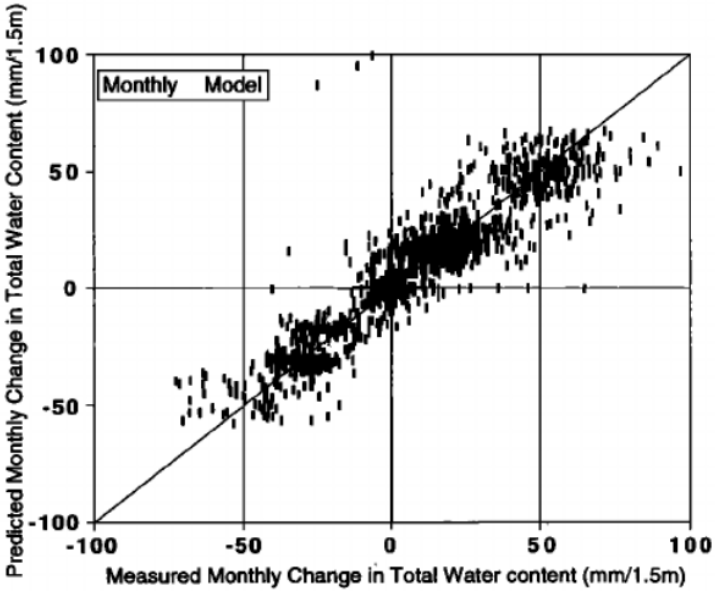


Figure 2.20 Plot of the predicted monthly change in total water content on the basis of EM31 measurements versus the one measured with the neutron probe for monthly models (after Sheets and Hendrickx, 1995).

Figure 2.21 Polymer ladder device for multi-height EM38 measurements (after Hossain et al., 2010)

Measuring EC_a at multiple heights is another technique that researchers have tried to incorporate in their studies. Hossain et al. (2010) found moderate correlations (R^2 = 0.37-0.57) (Figure 2.17) between average θ in two average depths (0-0.8 and 0-1.2 m) and EC_a collected by EM38h and v at various heights in a deep Vertosol soil (Figure 2.21).

Similarly, Stanley et al. (2014) conducted a study across a 96-ha field of Grey Vertosols 30km west of Moree, NSW, Australia. They monitored θ at thirty sites from three EC_a zones within the study field (Figure 2.22a). They reported good correlations (r > 0.85) between θ measured using a neutron probe at 0.2 to 1.4m depth and all combination of EM38h and v held 0 to 1.2 m above the soil surface. They also found the best correlations of EC_a readings to θ when the EM38 was held at 0.2 m above the soil surface and the θ was averaged to a depth of 0.4 or 0.6 m (Figure 2.22b).

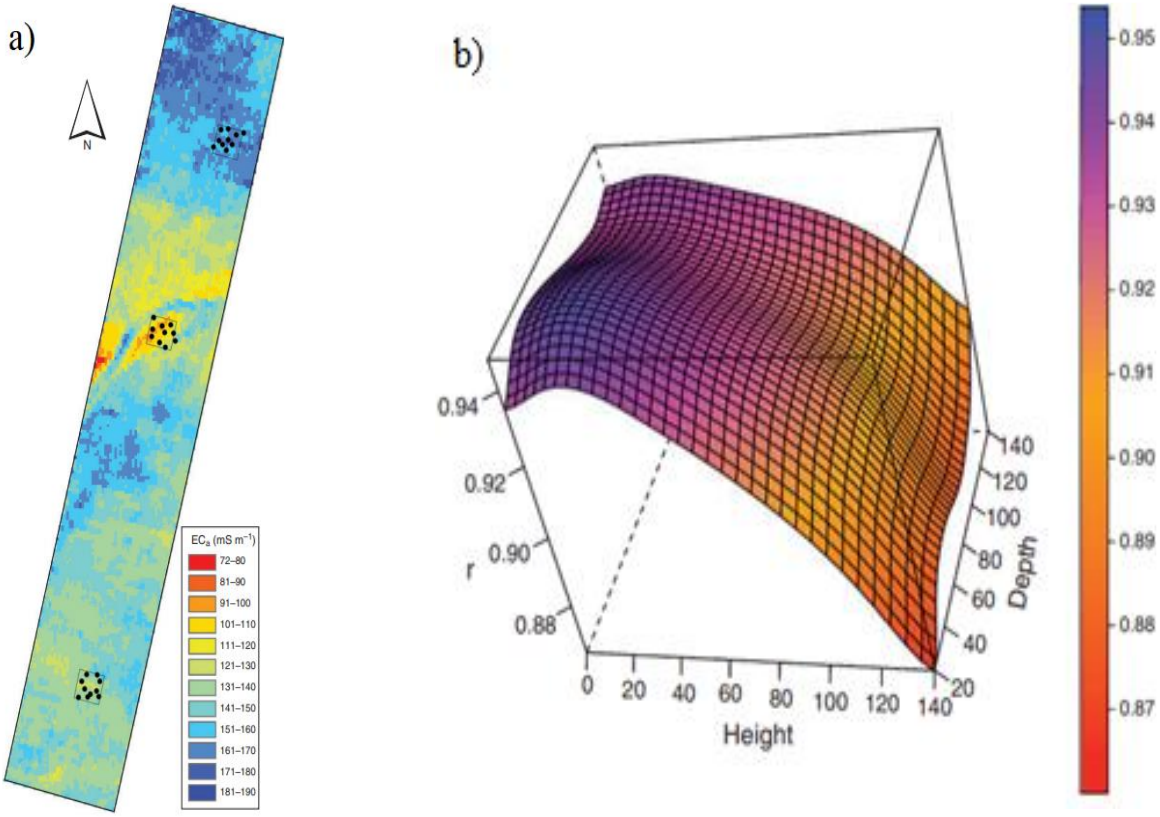


Figure 2.22 a) Initial EMI survey using an EM38v (2.4 by 0.4 km), also showing the three, 1-ha, square plots of high, medium and low EC_a levels that generated the range of background EC_a levels for the neutron probe access tube installations, indicated by (.) b) a three-dimensional surface (multi-level B spline) representing the Pearson

correlations of EC_a to θ for each combination of EM38 sensor height (and neutron probe depth (Depth) for all access tube-sites (after Stanley et al., 2014).

They suggested that the calibration of EC_a to θ may perhaps be influenced by a two-point measurement of the soil, namely at both upper (FC) and lower (PWP) EC_a, and site-specific calibration model can be produced by extending these point measures to whole-field surveys.

In order to represent spatial-temporal variation in θ, Robinson et al. (2012) collected DUALEM-1 data at various times across a 4 ha hillslope in a semi-arid oak savanna catchment in Stanford foothills reserve, CA, USA. They estimated difference in θ relative to changes in soil EC_a as a function of moisture. They did this by subtracting the EC_a data of the driest part of the season from the EC_a collected during subsequent moister periods and in doing so remove the effects of EC_s and other factors. They used georeferenced TDR data to collect enable calibration data to be collected.

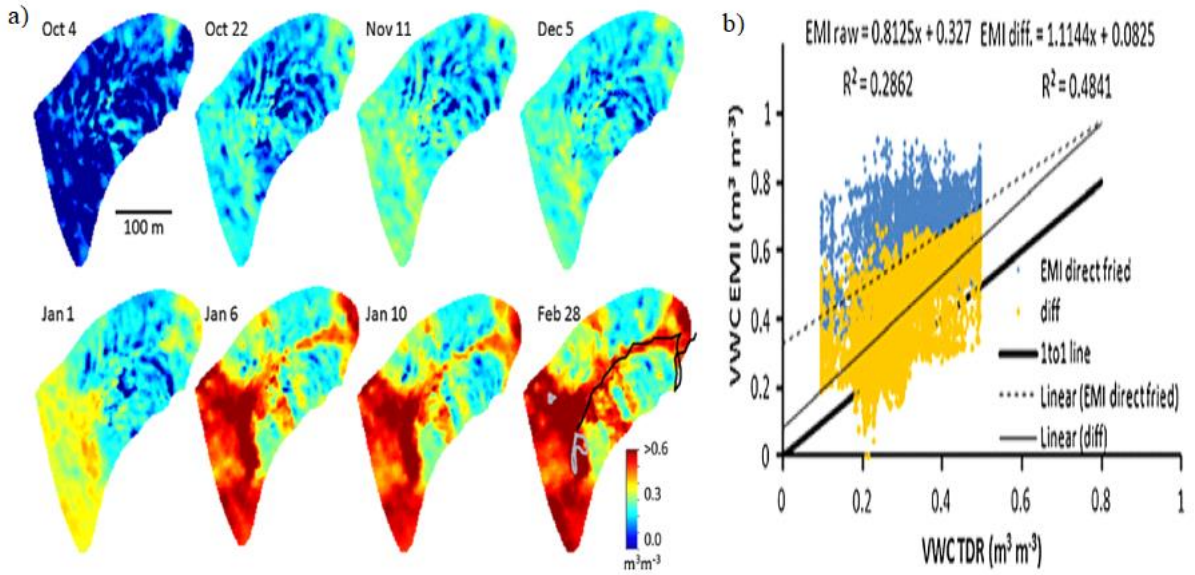


Figure 2.23 a) Evolution of θ pattern estimates using the EMI differencing approach on eight different days during 2007 and 2008 b) Comparison of θ determined from (i) EMI using direct estimation from raw data (EMI raw) and from (ii) the differencing approach (EMI diff) (after Robinson et al 2012).

Their results show that patterns of θ were essentially random when dry and dominated by soil texture, with increasing continuity and spatial correlation with topography when moist

(Figure 2.23a). The patterns suggest there is reasonable uniformity in θ across the catchment up until January, and that this changes to strong, distinct texture and topographically controlled patterns from February onwards. Their differencing approach improved prediction of average θ (0-0.15 m) with R^2 increasing significantly from 0.28 to 0.48 (Figure 2.23b).

Generally, EM instruments work well in mapping θ when there is a uniform soil texture. However, a few researchers have attempted to map θ when the soil texture varies. This has been enabled by applying data mining algorithms and developing more complex machine learning models. Moreover, the inclusion of other factors, including elevation, trend surface parameters (i.e., Easting and Northing), rainfall have also been considered. Recently, Hedley et al., (2013) used a machine learning approach (i.e., random forest) to monitor average θ in the topsoil (0-0.5 m) and as a function of water table depth across an irrigated sandy field. By incorporating elevation and rainfall, along with EM38 and EM31 EC_a, they achieved a good prediction (Figure 2.24) of water table depth (R^2 = 0.91) and θ (R^2 = 0.94). However, the overall correlation between EC_a and θ was moderate (R^2 = 0.77 and 0.71). They attributed this to the influence of other factors affecting EC_a including, clay content, salinity and temperature.

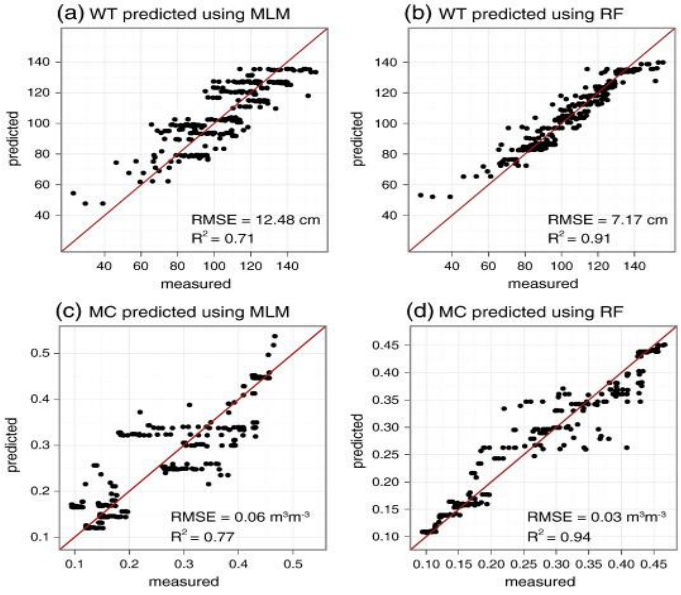


Figure 2.24 Predicted vs. measured water table depth (WT) (a, b) and soil moisture content at 50 cm (MC) (c, d). (MLM = multivariate linear model; RF = random forest model) (after Hedley et al., 2013).

Despite the fact that there is not a unique relationship between EC_a and θ, over the past decade, EM induction instruments are increasingly being used to monitor θ dynamics. This is because of the capability of these instruments in being non-invasive, providing repeatable data and collecting continuous measurements across large spatial extents. Thus, there is an exchange between the precision and the temporal and spatial resolution of θ measurement.

2.5.3.1 Application of the EM inversion algorithm in mapping θ dynamics

Recently, a site-specific calibration technique has been increasingly used for mapping θ dynamics. In this approach in order to obtain the model parameters a selection of reference points is chosen in the field to monitor EC_a and θ simultaneously using electromagnetic (EM) induction instruments and in-field probes (i.e., TDR, FDR), respectively. Afterwards, electromagnetic conductivity images (EMCI) can be generated by inverting EC_a which produce estimates of the depth specific true electrical conductivity (σ, mS m^{-1}). Subsequently, several mathematical and machine learning (i.e., ANN) models are fitted to the time-laps data to optimise model selection based on goodness of fit (Huang et al. 2016). In addition, different soil (i.e., texture) and environmental parameters (i.e., DEM) (Mahrooghy et al. 2016 and Xing et al. 2017) can be used as predictors to characterize the non-linear relationship between σ and θ.

In one study, Moghadas et al. (2017) mapped the 2-dimentional spatio-temporal variation of θ (Figure 2.21) based on estimates of σ values using petrophysical (i.e., Rhoades et al., 1976) relationship between σ and θ. They measured EC_a data using a CMD mini-Explorer sensor (GF Instruments, S.R.O., Brno, Czech Republic) along a 10m transect of a maize field over a 6-day period and used 5TE (METER Group, Inc. USA) sensors to measure θ at the end of the profile via an excavated pit with sandy topsoil (0-0.4 m). They applied a Differential Evolution

Adaptive Metropolis algorithm (DREAM) to invert the EC_a data and create the time-lapse estimates of σ.

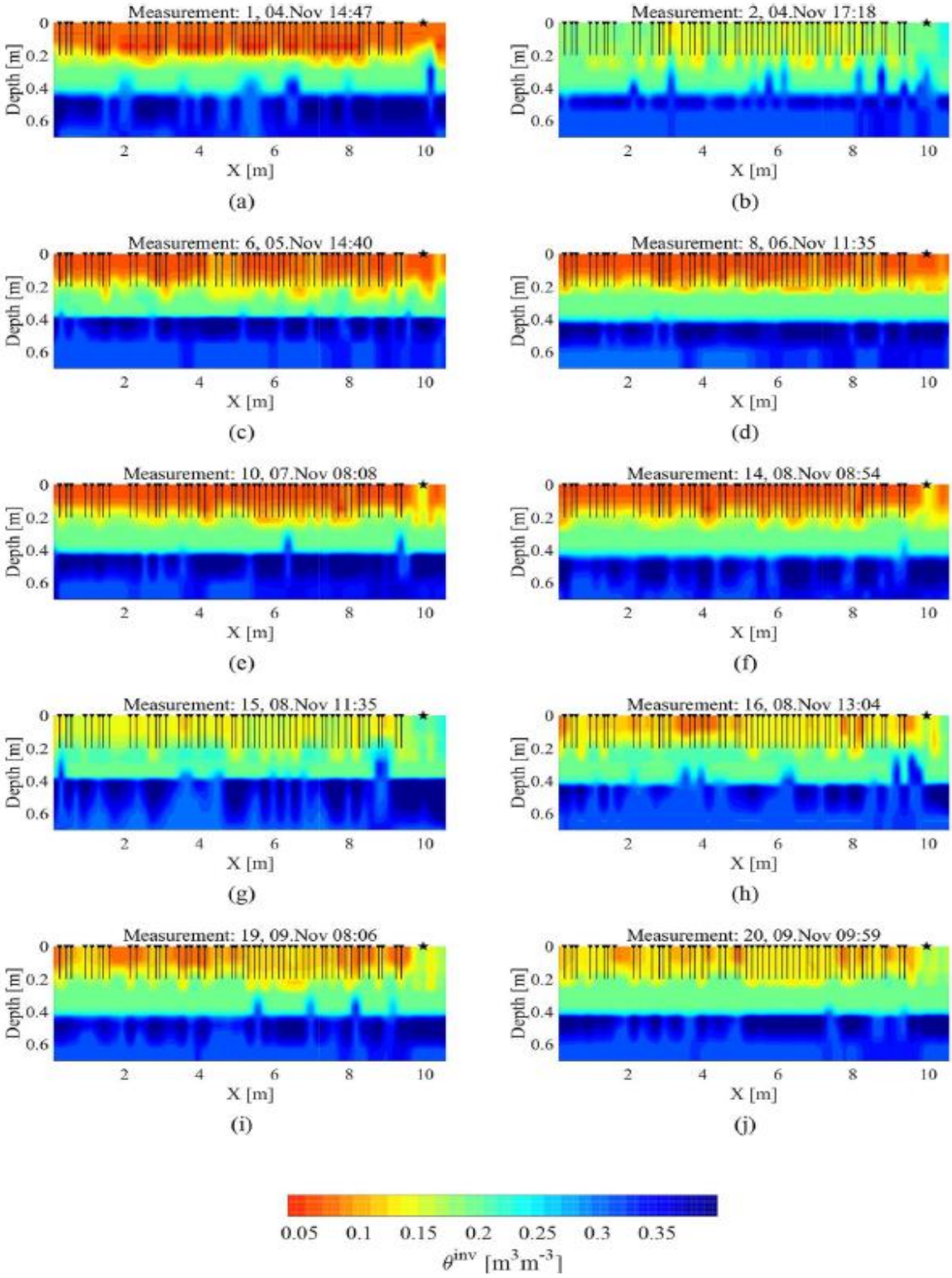


Figure 2.25. Spatiotemporal variations of θ derived from depth-specific σ for 10 different measurements. The location of the excavation pit and maize plantings are illustrated by black star and triangles, respectively (after Moghadas et al., 2017).

As shown in Figure 2.25, the 2D depth profile models illustrated lateral variations of θ along the transect. Three layers with relatively sharp θ contrast can be recognized. Measurements from first irrigation on November 4 (Figure 2.25a and b) and second irrigation on November 8 (Figure 2.25f, g and h) clearly show the effect of irrigation as indicated by small θ ranges after irrigation (Figure 2.25a and f) in comparison with the measurements before irrigation (Figure 2.25b and g). The results also show after the second irrigation, the drying pattern commenced (Figure 22.2i and j), although its intensity was lower than the first drying process (Figure 2.25 c to f). They concluded that despite some small discrepancies between inversely estimated and reference values, the proposed approach allows for the monitoring of the spatiotemporal variation of soil electrical conductivity and θ. In particular, the predicted θ profile rendered more relevant results than depth specific electrical conductivity to monitor subsurface flow processes. Moreover, the proposed EMI measurement and modelling technique allowed a delineation of the temporal wetting front and rootzone soil θ which can aid evaluation of uniformity of irrigation in the field as well as optimizing water use efficiency by providing an adequate control on site specific irrigation management.

In another study, Huang et al., (2016) established a calibration between σ and θ across 2-dimnetional depth slices of a centre-pivot irrigated Lucerne field with sandy soil in San Jacinto, CA, USA. They collected EC_a data using a single frequency multi-coil DUALEM-421 along 13 transects and determined an optimal set of inversion parameters (i.e., forward modelling, inversion algorithm and damping factor-λ) using EM4Soil (EMTOMO, V.21, Lisbon, Portugal) software package. The soil samples were collected at 0.3 m intervals to a depth of 1.5 m from 20 calibration sites. Thermogravimetric method was used in this study for θ measurement. Additionally, they also aimed to see if all the coil arrays of the DUALEM-421 are necessary.

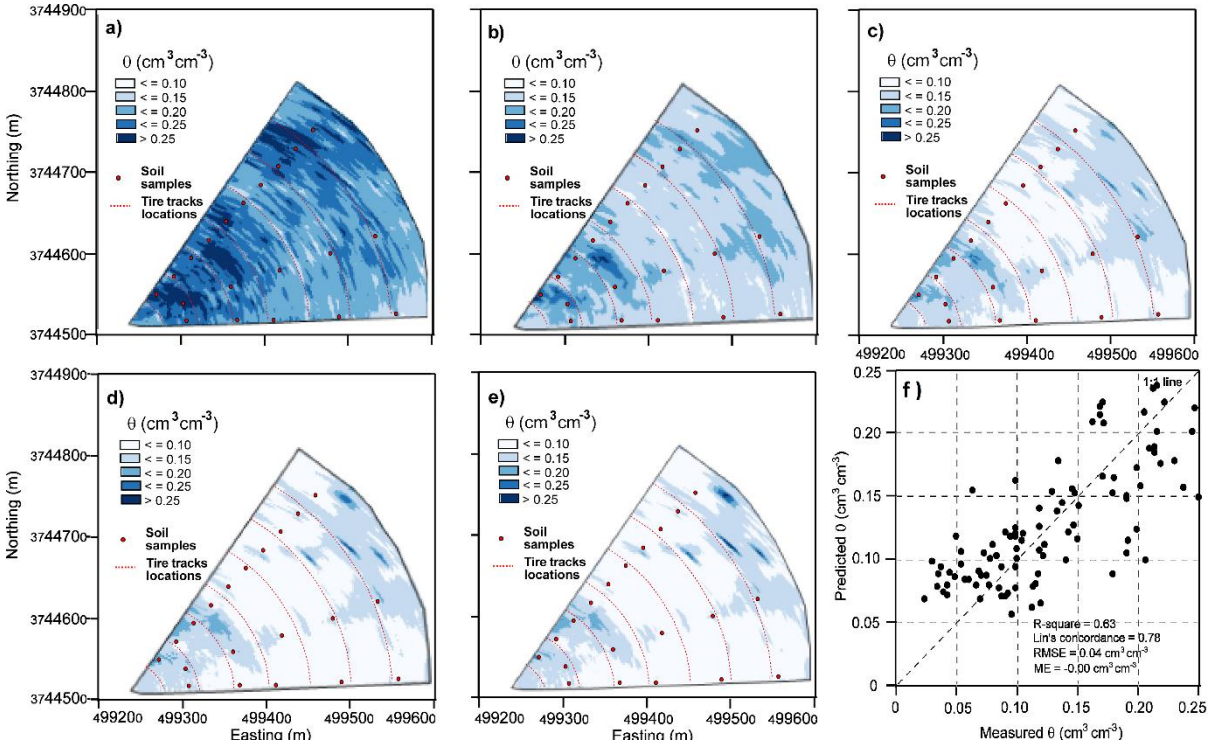


Figure 2.26 Predicted volumetric water content (θ, cm^3 cm^{-3}) from the inversion of bulk apparent electrical conductivity (EC_a,mS m^{-1}) collected using a DUALEM-421 and at various depth slices a) 0-0.3, b) 0.3-0.6, c) 0.6-0.9, d) 0.9-1.2, and e) 1.2-1.5 m; and f) plot of measured versus predicted soil θ generated by leave-one-out cross validation (after Huang et al., 2016).

Their results show variation in irrigation coverage. Figure 2.26a and b) show predicted topsoil and subsurface θ. In both cases the wettest areas (θ > 0.20 m^3 m^{-3}) are within 150 m of the location of the centre pivot. This is most likely because the sprinklers closer to the centre pivot cover much less ground compared to those located furthest away and near the edge of the field; due to the shorter and longer arc of travel, respectively. As suggested by intermediate (0.15-0.20 m^3 m^{-3}) subsoil θ (Figure 2.26e), the part of the field closest to the pivot may be being over-irrigated, which is leading to saturated soil conditions, increased hydraulic conductivity and deep drainage.

In terms of predicting θ they concluded the Lin's concordance (0.73) between measured and predicted θ was good and the resulting 2-dimesional depth slices (Figure 2.22) give insights into the spatial distribution of θ, thereby allowing the inference of depth of saturated soil and location of the wetting front which potentially enable the identification of areas where deep drainage may be problematic. Regarding the second aim of their study, they

concluded that while the DUALEM-1 produces a larger R^2 (0.59) in terms of calibration between σ and θ, the use of the DUALEM-421 data was better (R^2 = 0.56), because the total model misfit (4.70 mSm^{-1}) was smaller and better accounts for θ in the subsoil.

Recently, Moghadas et al. (2019) showed the advantages of using an Artificial neural network (ANN) model over the commonly used Rhoades et al. (1976) model to create time-lapse images of θ from σ. Similarly, Huang et al. (2017) used an ANN model to predict θ dynamics for different soil types from σ. They repeated EC_a data collection using a DUALEM-421 on 20 days over a 40-day period along a 480-m transect with varying soil texture profiles. The field was cropped with Lucerne and irrigation was supplied by sprinklers. They used Decagon GS3 moisture sensors to measure θ at different depths in 8 equally distanced locations along the transect. In this study they explored the potential to combine machine learning models (i.e., ANN) with soil physical models using data assimilation such as ensemble Kalman filter.

Figure 2.27b to e) shows the contour plots of predicted θ at different times. In general, θ was intermediate-small (0.2-0.3 m^3 m^{-3}) within the loam (sites 1 and 8), intermediate (0.3-0.4 m^3 m^{-3}) in duplex profiles (sites 2 and 7), and intermediate-large (0.3-0.4 m^3 m^{-3}) within uniform clays (sites 3-6), with θ increasing with depth. In addition, at sites 4 and 6, topsoil θ was larger. Their results showed that predicted θ was consistent with soil texture (Figure 2.27a). At sites 3, 4, 5 and 6 where clay (>40%) in the subsoil was large, θ was largest in comparison with sites 1, 2, 7 and 8 where clay was intermediate (35 - 30%), intermediate small (30 - 25%) and very small (>25%). They confirmed that predicted θ was also consistent with measured θ (Figure 2.27b to e). For example, at site 6, the measured large θ (> 0.35 m^3 m^{-3}) in the topsoil was predicted to be large (> 0.35 m^3 m^{-3}). Similarly, at site 1, the measured and predicted θ in the subsoil were both small (< 0.2 m^3 m^{-3}). However, it was not the case for topsoil at site 5 where θ was over-estimated.

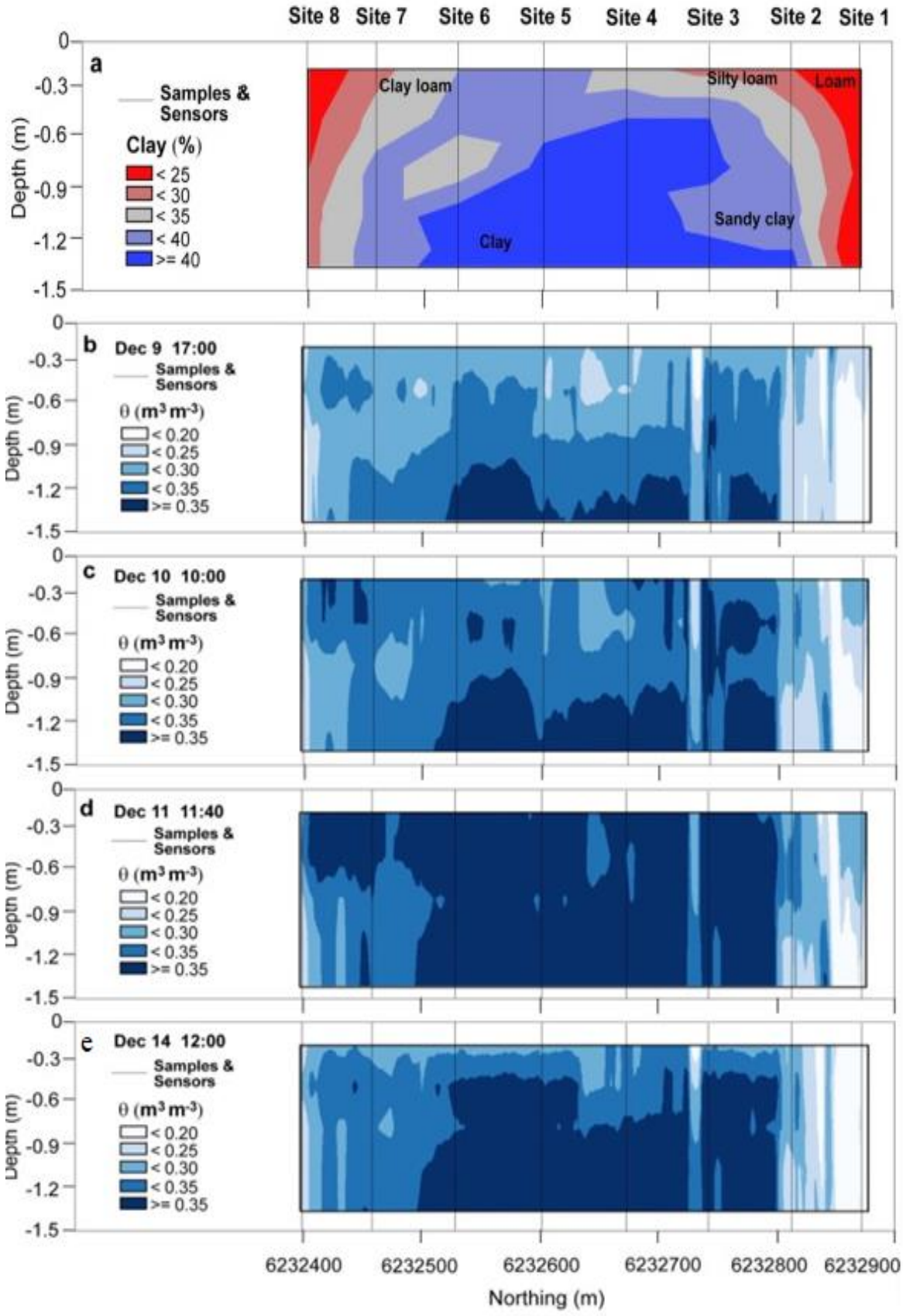


Figure 2.27 Contour plots of measured clay content (%) and predicted volumetric water content (θ, $m^3\ m^{-3}$) by the ANN model at various time steps including b) 17:00 h on 9 December 2015, c) 10:00 h on 10 December 2015, d) 11:40 h on 11 December 2015, e) 12:00 h on 14 December 2015 (after Huang et al., 2017).

They concluded that the ANN model was considered accurate given the large concordance coefficient between the measured and predicted θ (0.90). In terms of the physical model which was fitted using a simultaneous optimization and data assimilation, the performance was moderate given Lin's concordance (0.74). Based on these two models, they applied EnKF to 8 profiles separately and predicted θ dynamics with an overall concordance coefficient of 0.89. The EnFK modelling was superior to the physical model and equivalent to the ANN model within loam, clay and duplex soil profiles.

A few studies have also employed different machine learning models, such as Cubist (Zhao et al. 2018), support vector machines (SVM; Viscarra Rossel and Behrens, 2010) and random forests (RF; Grimm et al., 2008) for digital soil mapping (DSM) of other soil properties (i.e., clay, organic carbon). However, the potential of various machine learning models have not been comprehensively investigated to relate σ and θ. Therefore, further investigation is required. Moreover, there is a necessity to evaluate the different component of the site-specific calibration technique, which was described in this section, for the soils from heavier textural groups (i.e., Vertosols) since most of the previous studies have been conducted in the lighter textured soils (e.g., sandy, sandy loam).

2.5.3.2 Application of the EM inversion algorithm in mapping other soil properties

Recently, use of depth-specific true electrical conductivity (σ) generated by inversion of EC_a for mapping soil properties in 2 and 3 dimensions has been studied by researchers including but not limited to, salinity (Huang et al., 2015, Figure 2.28a), cation exchange capacity (Zhao et al., 2020, Figure 2.28b) and clay content (Khongnawang et al., 2019, Figure 2.28c). Herein, the 2 and 3-dimensional maps have been created based on a calibration model between σ and the limited measurements of the soil properties.

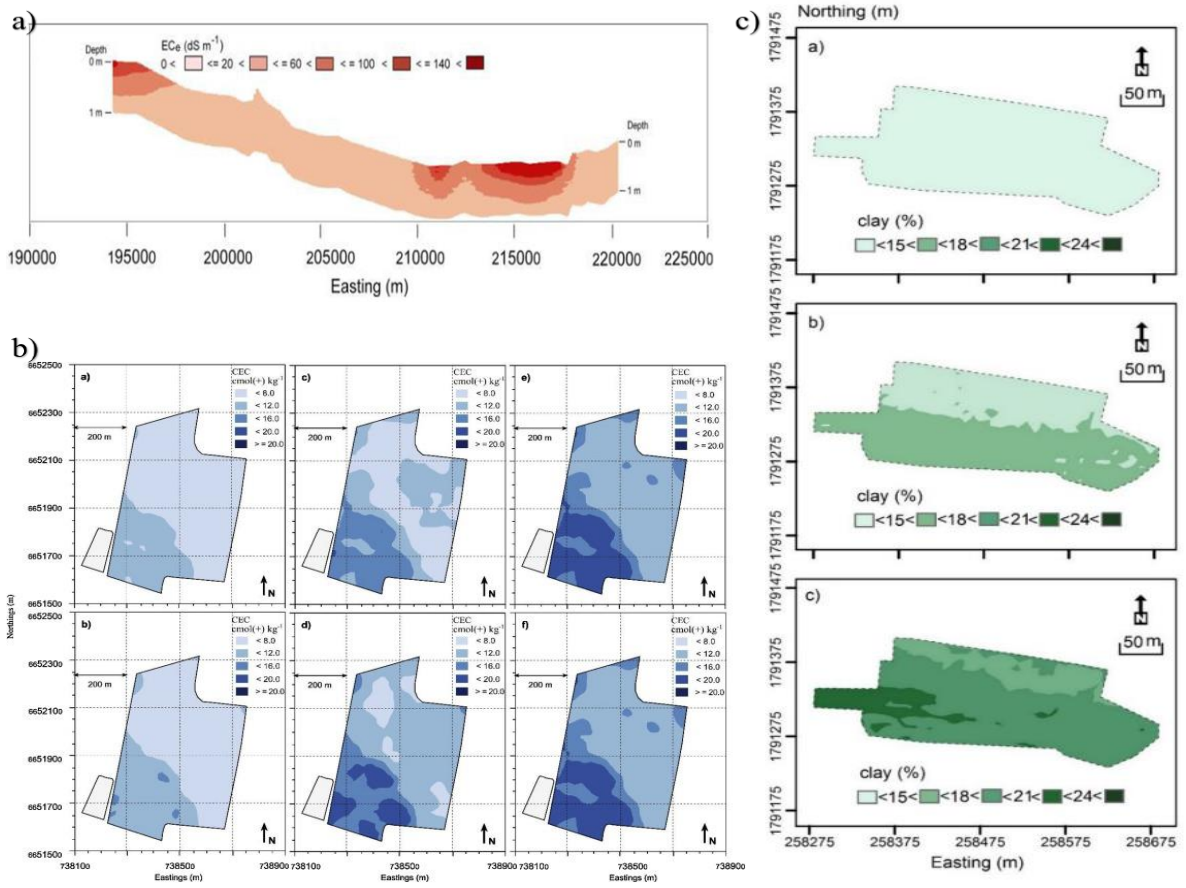


Figure 2.28 Contour plots of a) estimated electrical conductivity of a saturated soil-paste extract (EC_e) in 2d along a transect with elevation (after Huang et al., 2015) b) cation exchange capacity (CEC, cmol(+) kg^{-1}) (after Zhao et al., 2020) and c)clay content (after Khongnawang et al., 2019) with depth using inversion approach.

In one study, Zhao et al., (2019) using the same approach attempted to identify leakage zones associated with storage reservoirs by mapping clay. They believed that leakage has led to point source salinization in the cotton growing areas of Asley, NSW, Australia, therefore, information about the areal distribution of clay can assist farmers to identify where existing water storages may be better relocated. They inverted the EM data from an older EM survey using EM4Soil software package. Firstly, they developed a quasi-two-dimensional (Q-2D) model to estimate σ from the EM data along a single transect of EM38 and EM34 EC_a. Subsequently, they established a linear regression (LR) between σ and clay collected from 10 soil sample locations to a depth of 12 m along the same transect. Secondly, they used this LR to predict clay from σ estimated using a quasi-three-dimensional (Q-3D) model developed from EM38 and EM34 EC_a data collected on an approximate 500-1000 m grid across

50,000 ha. They showed that a validation set of clay, collected from 34 soil sample locations, indicated a substantial Lin's concordance (0.80).

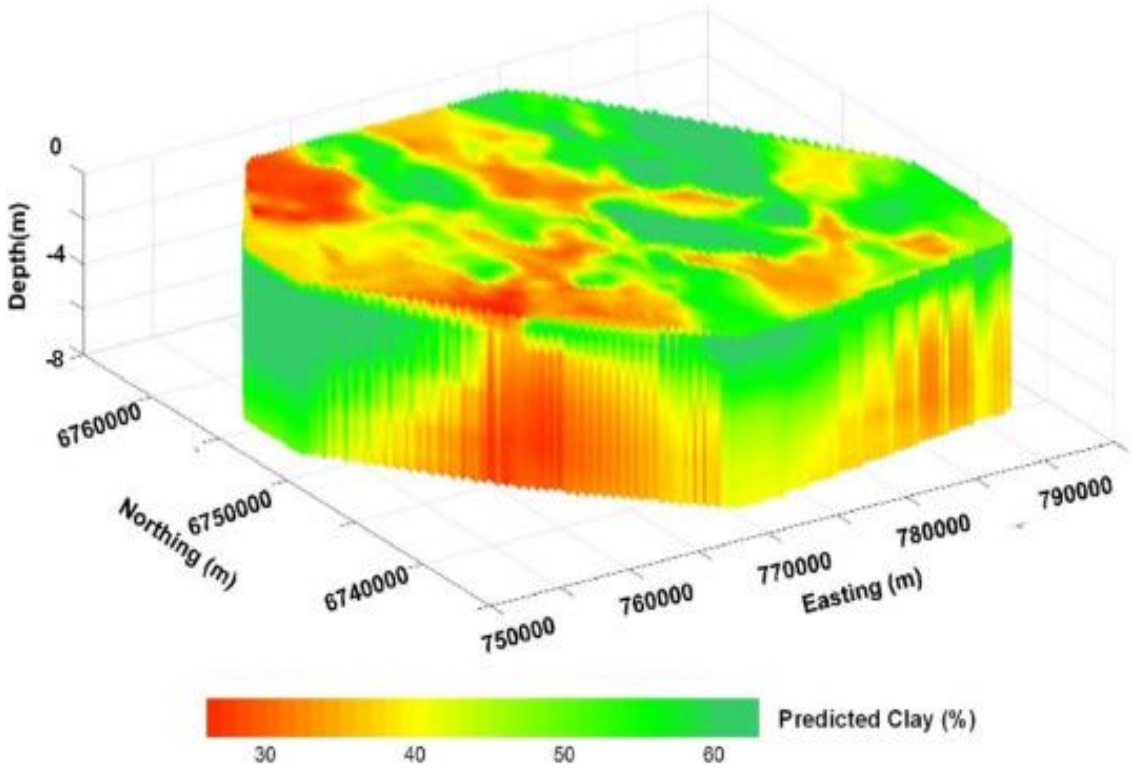


Figure 2.29 Areal distribution of predicted clay content in 3D view in Asley NSW, Australia (after Zhao et al., 2019).

Using the 3-dimensional map of clay (Figure 2.29) they were able to identify the location of the prior stream channels and their interconnectedness to underlying migrational deposits. These areas where associated with deep sands of Red-Orthic Tenosol. They concluded that their proposed approach has some merits and a better initial Q-2D calibration model might be achieved if they could collect EM31 or DUALEM-421 EC_a and combine this with the EM38 and EM34 EC_a. This could assist with improving their ability to resolve the depth of the clay plain and the location of the migrational channel sediments. Additionally, collection of more EC_a at smaller grid could lead to better prediction of clay, especially where the prior and present stream channels are contrasted against the clay alluvial plain.

In another study, Khongnawang et al. (2020) showed how they used a similar inversion approach to identify leaking areas along 1.6 km cross section of an irrigation canal

located in northeast Thailand. They showed that secondary soil salinization is associated with areas where leakage from the canal is happening and recharge water interacts with an underlying sequence of geological units rich in salts (Figure 2.30). They also evaluated the performance of three EM instruments including, Geonics EM38, EM34 and DUALEM-421 to generate 2-dimensional models of estimates of σ distribution with depth, against soil properties, including, EC_e.

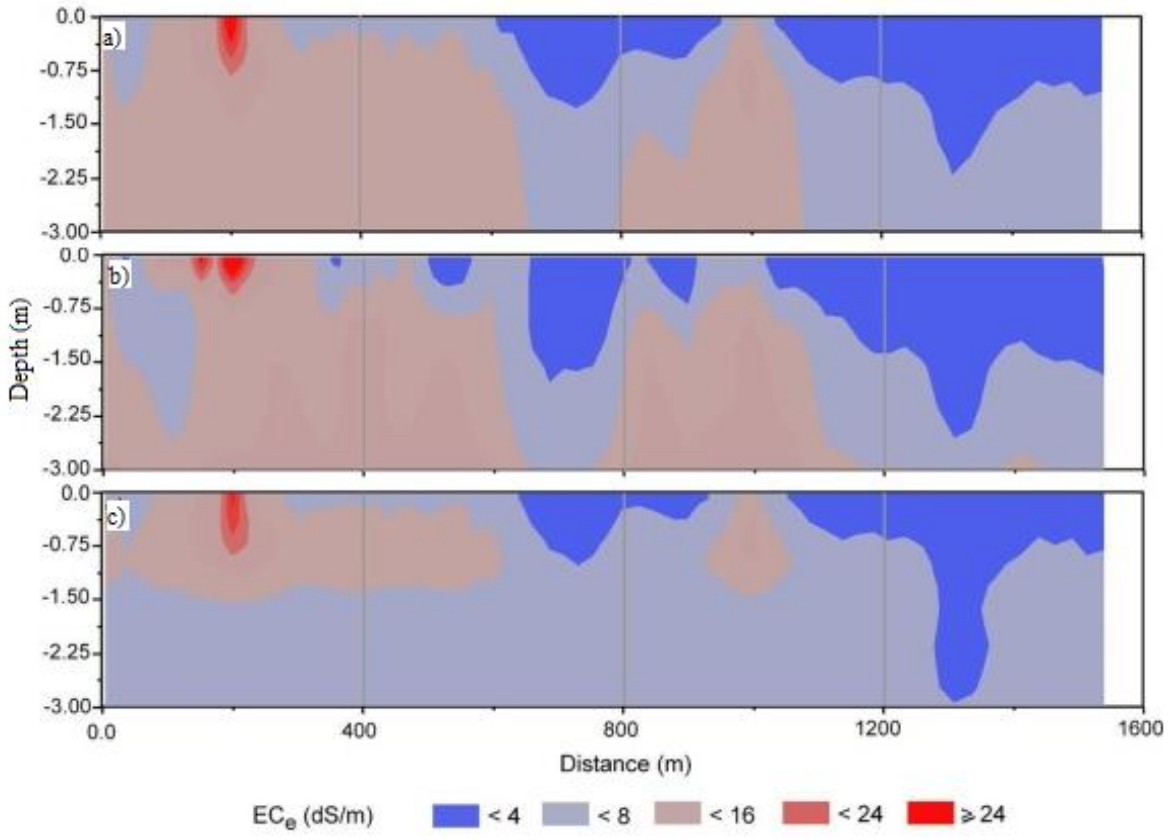


Figure 2.30 Spatial distribution of predicted electrical conductivity of a saturated soil paste extract (EC_e, dS m^{-1}) using estimates of electrical conductivity (σ, mS m^{-1}) inverted from measured apparent electrical conductivity (EC_a, mS m^{-1}) from a) EM38 b) DUALEM-421 and c) EM34 (after Khongnawang et al, 2020)

They found a strong coefficient of determination (R^2) between σ and EC_e by inverting EC_a using DUALEM-421 (0.72) moderate R^2 by using EM38 (0.66) and generally weak R^2 achieved using EM34 (0.51). They concluded that the recharge zones are associated with the areas with large silt (>70%), moderately sodic (6 < ESP < 12%) and where the soil was slightly saline (EC_e < 4 dS m^{-1}). They also suggest further soil physical and chemical characterisation can be performed to achieve better identification of recharge areas.

2.5.4 Application of gamma-ray (γ-ray) spectrometry

In recent times, use of γ-ray spectrometry data in soil science has been brought to attention in several mineral-rich countries such as Australia, United States, UK, and China. The studies are mainly about soil texture, identifying parent materials, soil types and more recently in mapping AWC and identifying soil management units. Previously, geologists and geophysicists routinely used radiometric surveys as mapping tools to identify where certain rock types change. Use of γ-ray spectrometry, is common for applications in study of mineral exploration (Foote and Humphrey, 1976), geological survey (Galbraith and Saunders, 1983) and radioelement contaminant evaluation (Roca et al., 1989). The γ-ray surveys can be performed proximal, using ground based γ-ray spectrometers (e.g., Radiation Solution-RS700, Figure 2.31) on finer grides or remotely using airborne γ-ray device mounted on a helicopter or an aeroplane flying at an altitude of 60-100 m above the surface. In Australia, generally airborne γ-ray surveys are conducted at 100-400 m line spacing. The γ-ray surveys across earth's surface tell us about the distribution of certain soil and rock types.

Figure 2.31 Proximal RS-700 γ-ray spectrometer (Radiation Solution INC, Ontario, Canada) being towed on a PVC sled.

2.5.4.1 DSM of available water content

Several studies have explored the use of DSM techniques to model soil AWC at the continental (Padarian et al., 2014), country (Hong et al., 2013) and district scale (Malone et al., 2009 and Gooley et al., 2014). In these studies, different combinations of environmental covariates have been used to produce maps of AWC including, EMI, DEM, NDVI, γ-ray spectrometry, slope, topographic wetness index (TWI), Multi-resolution Valley Bottom Flatness (MRVBF) and Landsat 7 bands (i.e., air temperature, rainfall, evapotranspiration, Prescott Index). From these studies a few has been conducted in cotton growing areas of NSW, Australia where Vertosol is the dominant soil type.

In one study, Malone et al., 2009 combined equal-area spline depth functions with digital soil mapping techniques to predict the vertical and lateral variations of AWC across the 1500 km^2 area of Edgeroi district, NSW, Australia. They constructed Neural network models for AWC to model its relationship with a set of environmental factors derived from DEM, γ-ray spectrometry data and Landsat 7 imagery. The soil legacy profile dataset they used consisted of 341 soil profiles. 210 were sampled on a systematic, equilateral triangular grid with a spacing of 2.8 km between sites at depth increments of 0-0.1, 0.1-0.2, 0.3-0.4, 0.7-0.8, 1.2-1.3 and 2.5-2.6 m. The other 131 soil profiles were distributed more irregularly or on transects. They fitted the spline to the values for the soil property (i.e., AWC). From this they derived the mean value of the soil property (i.e., AWC) for each legacy soil profile to the maximum depth of 1 m with 0.1 m increments. Then, they implemented their model framework to derive relationship between the training dataset and the environmental covariates (i.e, DEM, γ-ray and Landsat7).

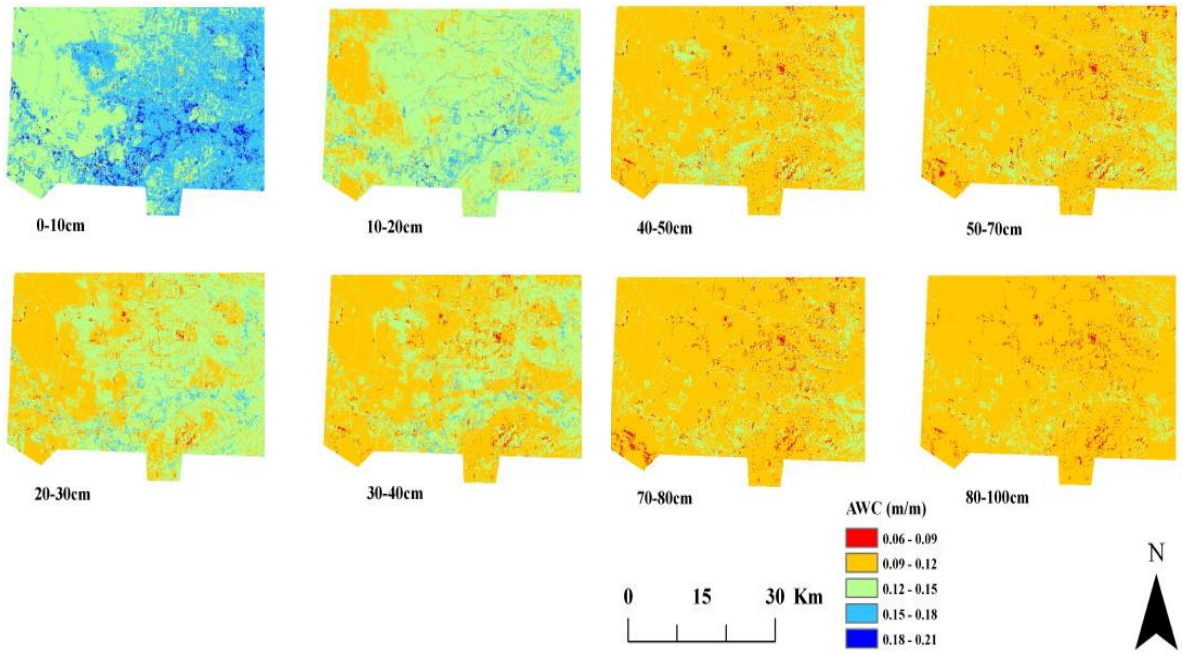


Figure 2.32 Predicted soil profile available water capacity (m^3 m^{-3}) to 1 m displayed in 8 profile layers (after Malone et al., 2009)

They concluded that, their prediction model of AWC (Figure 2.32) was able to account for 44% of the variation of this property across the study area. However, the validation results (R^2= 8-29%) indicated that the predictions were not as good as those generated by the prediction model. They attribute this to the fact that they used a group of surface environmental variables to predict soil depth functions. This shows the importance of the use of a digital data that can account for soil depth functions such as EMI.

In another study which was also conducted in cotton growing areas of NSW, Australia, Gooley et al., 2014 intended to create a DSM of AWC at the district scale. They did this by determining AWC with the difference between laboratory measured (i.e., pressure plate apparatus) permanent wilting point (PWP) and field capacity (FC). Subsequently, The PWP and FC data were coupled to remotely sensed γ-ray spectrometry and proximal sensed EM38 and EM34 data and two trend surface parameters (i.e., Easting and Northing). They predicted PWP and FC using a hierarchical spatial regression (HSR) across the areas of Warren and Trangie in the lower Macquarie valley, Australia. They also aimed to find the sufficient set of covariates to predict the properties (i.e., FC and PWP).

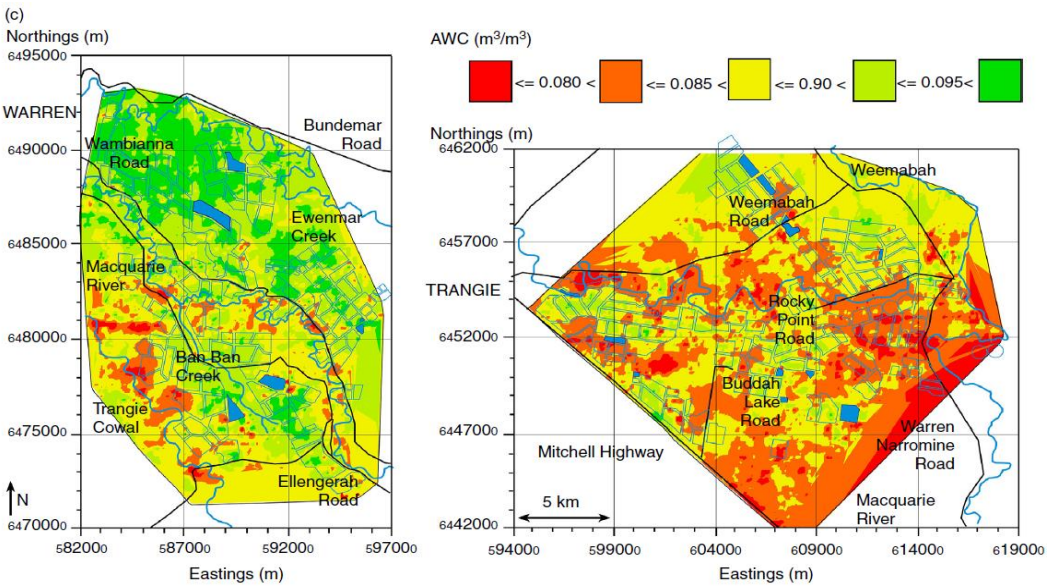


Figure 2.33 Spatial distribution of predicted average (0-0.90 m) available water content (AWC, $m^3\ m^{-3}$), across Warren and Trangie areas as calculated by difference from predicted FC and PWP (after Gooley et al., 2014).

They concluded that, Th, U, EM38v and EM34-20 and two trend-surface vectors (i.e., Easting and Northing) were sufficient covariates to predict FC and PWP. In addition, they showed that the resultant DSM of AWC was consistent with Pedoderms (Figure 2.33). The Pedoderms with sandy topsoil textures (e.g., Trangie Cowal) hold the smallest amount of water ($\leq$0.085 $m^3\ m^{-3}$). Conversely, the heavy textured pedoderms (e.g., Trangie Cowal) hold larger amounts (>0.095 $m^3\ m^{-3}$) of water. They also suggested that errors might be associated with their overall approach. This was because the ancillary data they used were collected at different scales and used to estimate PWP and FC on a fine grid. To improve precision of prediction a better approach might have been to develop separate maps of FC and PWP at 0.30 m depth increments.

Most of the previous studies in DSM of AWC were conducted in large scale, this might benefit the land holders, corporate stakeholders and governmental departments which need access to quantitative soil information in order to perform ecological and economical sustainable development. However, from a perspective of precision agriculture and irrigation management farmers need access to maps of AWC at field level. Therefore, further investigation is required.

2.5.4.2 DSM of soil management classes

Since precision farming has become important, application of ancillary data to identify soil management classes has earned broad attention. This is particularly because in many instances a direct correlation between digital data and soil data is often not linear. Several researchers have attempted to account for soil property variation by using the digital data to establish soil management classes by applying either EM data alone or in combination with other data such as γ-ray spectrometry data.

In one of the first studies, Cambouris et al., (2006) aimed to identify management classes in a 13.8 ha commercial potato production field where potato yield (5.9 t per ha) was in disagreement with the soil properties (i.e., AWC). They used EM38 data, obtained on a regular grid and considering various physical (e.g., AWC and texture) and chemical (e.g., pH) properties. Using a k-means clustering approach they partitioned the kriged EM data and found that two classes were adequate to account for within field variation of various soil properties. Similarly, Vitharana et al. (2006) considered top and subsoil clay content and fuzzy k-means clustering (FKM) to identify the soil management classes using EM38 data, in an 11.5-ha field located in northwest Flanders in Belgium.

Another method is to partition multiple EM instruments and/or types of ancillary data to discern management classes. Zhu et al. (2010) considered three different EM instruments (i.e., DUALEM-2, EM38 and EM31) in different modes of operation (i.e., Horizontal-h and Vertical-v) to distinguish classes based on the soil silt content and depth to bedrock. They found that EM31h could better account for silt content while it performed better in vertical mode to best resolve depth to bedrock. Anderson-Cook et al., (2002) used the crop yield data in addition to EM38 data in a 24-ha field in Virginia, USA. Using recursive binary classification trees, they generated coherent results to understand different soil types when they used both crop yield and EM38 data together in contrast to using either yield data or EM38 data alone. In

another study, Kitchen et al., 2005 combined elevation data with EC_a obtained by EM38 and Veris3100 (Veris Technologies, Inc., Kansas, USA) within two fields covering 41 ha in Missouri, USA. They demonstrated productivity classes obtained from FKM clustering of the ancillary data are in agreement with the classes attained by the clustering of the yield maps independently.

Recently, γ-ray spectrometry data, in combination with the EM data, has been used to identify management classes at the field (Dennerley et al., 2018) and multiple-field (Arshad et al., 2019) scale. Moreover, Huang et al. (2014) numerically (i.e., FKM) clustered EC_a data acquired by a DUALEM-1 and proximal sensed γ-ray spectrometry (i.e., GPX-21, GPX Services Pty Ltd, Perth, Australia) data (i.e., K, U, Th and Total counts) to identify contiguous classes across seven fields located in Nottingham, UK. They showed that the resultant maps of 7 (Figure 2.34a) and 8 classes were statistically distinct and resolved most of the soil variation (Figure 2.34b) in terms of clay and pH, respectively.

More studies have provided similar results including, Altdorff and Dietrich (2012), Van Miervenne et al. (2013), Islam et al (2011) and Castrignano et al (2012). Remote sensing platforms can also possibly be used in development of DSM of management classes at district level if the resolution is sufficient. For example, Jing et al. (2017) used a DSM approach, where remotely sensed data, obtained from an air-borne γ-ray spectrometer survey, including K, Th, U and TC were combined with proximal sensed data measured using EM38 to identify soil management classes across a farming district near Gunnedah, NSW, Australia. They used digital data as surrogate for morphological data, with soil types identified by fuzzy k-means and a linear mixed model with measured physical (i.e., clay, silt and sand) and chemical (i.e., EC_e, pH, CEC, ESP and exchangeable cations) properties from the topsoil (0-0.30 m) and subsoil (0.9-1.2 m).

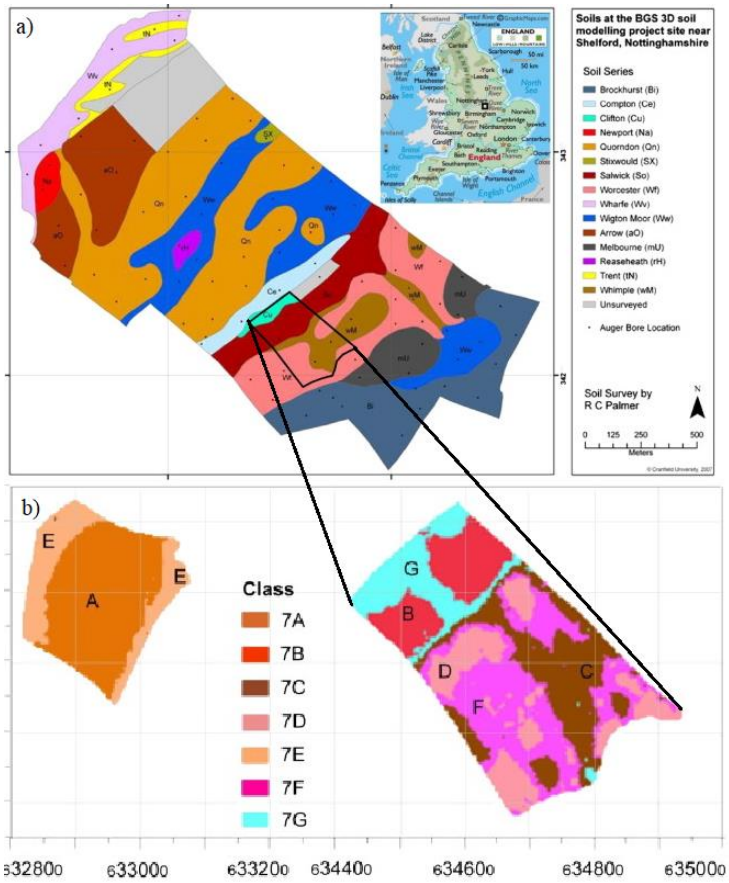


Figure 2.34 a) Location of study area east of Nottingham and River Trent and the soil series map and b) spatial distribution of FKM classes for K=7 (after Huang et al. 2014).

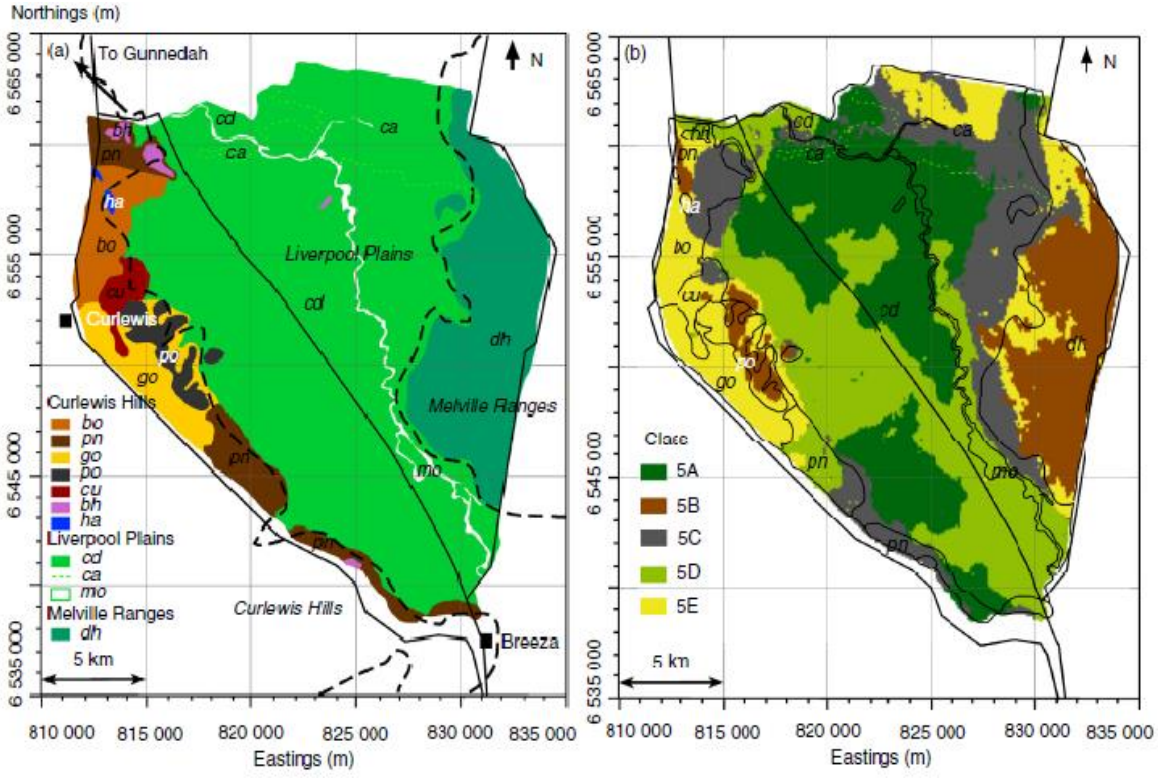


Figure 2.35 a) Physiographic and landscape units identified by Banks (1995), and Spatial distribution of b) fuzzy k-means class k = 5 near Gunnedah, Australia. (after Jing et al. 2017).

Figure 2.35a shows the physiographic map and spatial distribution of the landscape units across the farming district near Gunnedah, Australia. The largest unit is Liverpool Plains (e.g., cd), which is used for irrigated cotton and cereal, following by Melville Ranges (e.g., dh) and Curlewis Hills (e.g, po). Figure 2.35b shows the spatial distribution of fuzzy k-means classes when k = 5. Their results showed that k = 5 distinguished between three main physiographic areas and some of the soil landscape units (Banks, 1995). In Particular, po and go units on the Curlewis Hills were corresponding with 5B and 5E, respectively. While cd, ca and mo were appeared to be characterised by 5A, 5C and 5D on Liverpool Plain, respectively. On the Melville Ranges, dh was represented by 5B with 5E a transition class between 5C and 5B. However, it was apparent that their approach could not distinguish soil landscape units such as bh and cu. In the case of bh, they attributed this to the fact that the unit appears only as two small basalt outcrops in the northeast corner on the Curlewis Hills.

They concluded that their DSM approach was able to highlight subtle differences in chemical (i.e., CEC) properties in the agriculturally productive Liverpool Plains. In addition, the management units identified were representative of soil landscape units previously mapped using traditional methods (Banks, 1995) and based on geomorphic mapping of landforms.

The method can be used to update the traditional pedoderm maps in district scale. Therefore, it is worth teting the method by clustering proximal sensed EM data in combination with remotely sensed ancillary data such as γ-ray spectrometry or other sensor platforms which have the potential to productively map soil physical (e.g., clay), hydrological (e.g., AWC) and chemical (e.g., EC_e) properties in other agriculturally productive areas of NSW.

2.6 Conclusion

Tackling the challenges associated with agricultural water management is becoming more crucial due to increasing population and lack of sufficient amount of available water. This requires use of new techniques. DSM techniques is increasingly gaining attention to map soil water-related physical (i.e., clay, silt and sand), hydrological (i.e., AWC, FC and PWP) and chemical (e.g., CEC, EC_e and pH) properties in space and time. This is because it aids precision agriculture and potentially improves agricultural water management and crop yield optimization. Use of easy and cheap to acquire digital data along with spatial and non-spatial inference methods has enabled to map either soil water-related properties or management classes more rapidly and comprehensive. Among all types of digital data, EMI has been used largely at the field scale. However, use of this data in larger scales deserves more exploration. The benefit of EMI instruments (e.g., DUALEM-421, EM38 and 34) is they can provide information to account for depth functions. Besides, they have been employed widely in studies related to mapping soil types and properties of agricultural importance (e.g., clay, moisture, salinity). Another source of digital data is γ-ray spectrometry data. Whether it is being collected proximal or remote, it has shown potetial to be used as a surrogate to map soil properties related to soil texture (e.g., clay, AWC) or soil landscape units. Recently, this type of data is readily available in large spacial extents through airborne platforms. Thus, it is a useful tool to be used for identifyiung landscape unites and soil types at larger spatial scales (e.g., district scale).

2.7 Referencess

Akbar, M. A., Kenimer, A. L., Searcy, S. W., & Torbert, H. A. (2005). Soil water estimation using electromagnetic induction. *Transactions of the ASAE, 48*(1), 129-135.

Al-Ain, F., Attar, J., Hussein, F., & Heng, L. K. (2009). Comparison of nuclear and capacitance-based soil water measuring techniques in salt-affected soils. *Soil use and management*, *25*(4), 362-367. doi.org/10.1111/j.1475-2743.2009.00246.x

Algeo, J., Van Dam, R. L., & Slater, L. (2016). Early-time GPR: A method to monitor spatial variations in soil water content during irrigation in clay soils. *Vadose Zone Journal*, *15*(11), 1-9.

Altdorff, D., & Dietrich, P. (2012). Combination of electromagnetic induction and gamma spectrometry using K-means clustering: A study for evaluation of site partitioning. *Journal of Plant Nutrition and Soil Science*, *175*(3), 345-354.

Anderson-Cook, C. M., Alley, M. M., Roygard, J. K. F., Khosla, R., Noble, R. B., & Doolittle, J. A. (2002). Differentiating soil types using electromagnetic conductivity and crop yield maps. *Soil Science Society of America Journal*, *66*(5), 1562-1570.

Annan, A. P. (2005). GPR methods for hydrogeological studies. In *Hydrogeophysics* (pp. 185-213). Springer, Dordrecht.

Archie, G. E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the AIME, 146*(01), 54-62.

Arshad, M., Li, N., Zhao, D., Sefton, M., & Triantafilis, J. (2019). Comparing management zone maps to address infertility and sodicity in sugarcane fields. *Soil and Tillage Research*, *193*, 122-132.

Baumhardt, R. L., Lascano, R. J., & Evett, S. R. (2000). Soil material, temperature, and salinity effects on calibration of multisensor capacitance probes. *Soil Science Society of America Journal*, *64*(6), 1940-1946.

Callegary, J. B., Ferré, T. P., & Groom, R. W. (2012). Three-dimensional sensitivity distribution and sample volume of low-induction-number electromagnetic-induction instruments. *Soil Science Society of America Journal*, *76*(1), 85-91.

Cambouris, A. N., Nolin, M. C., Zebarth, B. J., & Laverdière, M. R. (2006). Soil management zones delineated by electrical conductivity to characterize spatial and temporal variations in potato yield and in soil properties. *American journal of potato research*, *83*(5), 381-395.

Castrignano, A., Wong, M. T. F., Stelluti, M., De Benedetto, D., & Sollitto, D. (2012). Use of EMI, gamma-ray emission and GPS height as multi-sensor data for soil characterisation. *Geoderma*, *175*, 78-89.

Chambers, J. E., Kuras, O., Meldrum, P. I., Ogilvy, R. D., & Hollands, J. (2006). Electrical resistivity tomography applied to geologic, hydrogeologic, and engineering investigations at a former waste-disposal site. *Geophysics*, *71*(6), B231-B239.

Chanzy, A., Tarussov, A., Bonn, F., & Judge, A. (1996). Soil water content determination using a digital ground-penetrating radar. *Soil science society of America journal*, *60*(5), 1318-1326.

Cheng, Q., Sun, Y., Qin, Y., Xue, X., Cai, X., Sheng, W., & Zhao, Y. (2013). In situ measuring soil ice content with a combined use of dielectric tube sensor and neutron moisture meter in a common access tube. *Agricultural and forest meteorology*, *171*, 249-255.

Corwin, D. L., & Lesch, S. M. (2005). Apparent soil electrical conductivity measurements in agriculture. *Computers and electronics in agriculture*, *46*(1-3), 11-43.

Dean, T. J., Bell, J. P., & Baty, A. J. B. (1987). Soil moisture measurement by an improved capacitance technique, Part I. Sensor design and performance. *Journal of Hydrology*, *93*(1-2), 67-78.

Dennerley, C., Huang, J., Nielson, R., Sefton, M., & Triantafilis, J. (2018). Identifying soil management zones in a sugarcane field using proximal sensed electromagnetic induction and gamma-ray spectrometry data. *Soil Use and Management*, *34*(2), 219-235.

Dickson, B. L., & Scott, K. M. (1997). Interpretation of aerial gamma-ray surveys-adding the geochemical factors. *AGSO Journal of Australian Geology and Geophysics*, *17*, 187-200.

Dinka, T. M., Morgan, C. L., McInnes, K. J., Kishné, A. S., & Harmel, R. D. (2013). Shrink–swell behavior of soil across a Vertisol catena. *Journal of Hydrology*, *476*, 352-359.

Dobriyal, P., Qureshi, A., Badola, R., & Hussain, S. A. (2012). A review of the methods available for estimating soil moisture and its implications for water resource management. *Journal of Hydrology*, *458*, 110-117.

Doolittle, J. A., & Collins, M. E. (1995). Use of soil information to determine application of ground penetrating radar. *Journal of applied geophysics*, *33*(1-3), 101-108.

Doolittle, J. A., & Brevik, E. C. (2014). The use of electromagnetic induction techniques in soils studies. *Geoderma*, *223*, 33-45.

EMTOMO, 2014. EM4Soil Version 2. http://www.emtomo.com/

Evett, S., & Cepuder, P. (2008). *Capacitance sensors for use in access tubes* (No. IAEA-TCS--30).

Fares, A., & Alva, A. K. (2000). Evaluation of capacitance probes for optimal irrigation of citrus through soil moisture monitoring in an entisol profile. *Irrigation Science*, *19*(2), 57-64.

Ferré, T. P., Nissen, H. H., & Šimůnek, J. (2002). The effect of the spatial sensitivity of TDR on inferring soil hydraulic properties from water content measurements made during the advance of a wetting front. *Vadose Zone Journal*, *1*(2), 281-288.

Foote, R. S., & Humphrey, N. B. (1976). Airborne radiometric techniques and applications to uranium exploration. In *Exploration for uranium ore deposits*.

Friedman, S. P. (2005). Soil properties influencing apparent electrical conductivity: a review. *Computers and electronics in agriculture*, *46*(1-3), 45-70.

Furman, A., Ferré, T. P., & Warrick, A. W. (2004). Optimization of ERT surveys for monitoring transient hydrological events using perturbation sensitivity and genetic algorithms. *Vadose Zone Journal*, *3*(4), 1230-1239.

Galbraith, J. H., & Saunders, D. F. (1983). Rock classification by characteristics of aerial gamma-ray measurements. *Journal of Geochemical Exploration*, *18*(1), 49-73.

Gibson, P. J., & George, D. M. (2004). *Environmental applications of geophysical surveying techniques*. Nova Biomedical Books.

Goodman, D., & Piro, S. (2013). *GPR remote sensing in archaeology* (Vol. 9, p. 233). New York: Springer.

Gooley, L., Huang, J., Page, D., & Triantafilis, J. (2014). Digital soil mapping of available water content using proximal and remotely sensed data. *Soil use and management*, *30*(1), 139-151.

Grasty, R. L. (1991). Airborne gamma-ray spectrometer surveying. *Tech. Rep.*, *97*.

Hedley, C. B., Roudier, P., Yule, I. J., Ekanayake, J., & Bradbury, S. (2013). Soil water status and water table depth modelling using electromagnetic surveys for precision irrigation scheduling. *Geoderma*, *199*, 22-29.

Hendrickx, J. M. H., Kachanoski, R. G., Dane, J. H., & Topp, G. C. (2002). Nonintrusive electromagnetic induction. *Methods of soil analysis. Part*, *4*, 1297-1306.

Hong, S. Y., Minasny, B., Han, K. H., Kim, Y., & Lee, K. (2013). Predicting and mapping soil available water capacity in Korea. *PeerJ*, *1*, e71.

Hossain, M. B., Lamb, D. W., Lockwood, P. V., & Frazier, P. (2010). EM38 for volumetric soil water content estimation in the root-zone of deep vertosol soils. *Computers and Electronics in Agriculture*, *74*(1), 100-109.

Huang, J., Lark, R. M., Robinson, D. A., Lebron, I., Keith, A. M., Rawlins, B., ... & Triantafilis, J. (2014). Scope to predict soil properties at within-field scale from small samples using proximally sensed γ-ray spectrometer and EM induction data. *Geoderma*, *232*, 69-80.

Huang, J., Taghizadeh-Mehrjardi, R., Minasny, B., & Triantafilis, J. (2015). Modeling soil salinity along a hillslope in Iran by inversion of EM38 data. *Soil Science Society of America Journal*, *79*(4), 1142-1153.

Huang, J., Scudiero, E., Choo, H., Corwin, D. L., & Triantafilis, J. (2016). Mapping soil moisture across an irrigated field using electromagnetic conductivity imaging. *Agricultural Water Management*, *163*, 285-294.

Huang, J., McBratney, A. B., Minasny, B., & Triantafilis, J. (2017). Monitoring and modelling soil water dynamics using electromagnetic conductivity imaging and the ensemble Kalman filter. *Geoderma*, *285*, 76-93.

Huang, J. (2017). Mapping and monitoring soil water dynamics using electromagnetic conductivity imaging. *UNSW Sydney. Dr. Thesis*

Huisman, J. A., Hubbard, S. S., Redman, J. D., & Annan, A. P. (2003). Measuring soil water content with ground penetrating radar: A review. *Vadose zone journal*, *2*(4), 476-491.

Ingham, M., McConchie, J. A., Wilson, S. R., & Cozens, N. (2006). Measuring and monitoring saltwater intrusion in shallow unconfined coastal aquifers using direct current resistivity traverses. *Journal of Hydrology (New Zealand)*, 69-82.

Islam, M. M., Cockx, L., Meerschman, E., De Smedt, P., Meeuws, F., & Van Meirvenne, M. (2011). A floating sensing system to evaluate soil and crop variability within flooded paddy rice fields. *Precision agriculture*, *12*(6), 850.

Jaques, A. L., Wellman, P., Whitaker, A., & Wyborn, D. (1997). High-resolution geophysics in modern geological mapping. *AGSO Journal of Australian Geology and Geophysics*, *17*(2), 159-173.

Jing, Y. Z., Huang, J., Banks, R., & Triantafilis, J. (2017). Scope to map soil management units at the district level from remotely sensed γ-ray spectrometry and proximal sensed EM induction data. *Soil Use and Management*, *33*(4), 538-552.

Kachanoski, R. G., Wesenbeeck, I. V., & Jong, E. D. (1990). Field scale patterns of soil water storage from non-contacting measurements of bulk electrical conductivity. *Canadian Journal of Soil Science*, *70*(3), 537-542.

Kachanoski, R. G., Wesenbeeck. V., & Gregorich, E. G. (1988). Estimating spatial variations of soil water content using noncontacting electromagnetic inductive methods. *Canadian Journal of Soil Science*, *68*(4), 715-722.

Kelly, B. F. J., Acworth, R. I., & Greve, A. K. (2011). Better placement of soil moisture point measurements guided by 2D resistivity tomography for improved irrigation scheduling. *Soil Research*, *49*(6), 504-512.

Khongnawang, T., Zare, E., Zhao, D., Srihabun, P., & Triantafilis, J. (2019). Three-Dimensional Mapping of Clay and Cation Exchange Capacity of Sandy and Infertile Soil Using EM38 and Inversion Software. *Sensors*, *19*(18), 3936.

Khongnawang, T., Zare, E., Srihabun, P., & Triantafilis, J. (2020). Comparing electromagnetic induction instruments to map soil salinity in two-dimensional cross-sections along the Kham-rean Canal using EM inversion software. *Geoderma*, *377*, 114611.

Nobes, D. C. (1996). Troubled waters: Environmental applications of electrical and electromagnetic methods. *Surveys in Geophysics*, *17*(4), 393-454.

McCutcheon, M. C., Farahani, H. J., Stednick, J. D., Buchleiter, G. W., & Green, T. R. (2006). Effect of soil water on apparent soil electrical conductivity and texture relationships in a dryland field. *Biosystems Engineering*, *94*(1), 19-32.

McNeill, J. D. (1980). Electromagnetic terrain conductivity measurement at low induction numbers.

Mahrooghy, M., Aanstoos, J., Nobrega, R. A., Hasan, K., & Younan, N. H. (2016). A neural network approach to soil electrical conductivity estimation on earthen Levees using spaceborne X-band SAR imagery. *Photogrammetric Engineering & Remote Sensing*, *82*(7), 509-520.

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.

Metwaly, M., & AlFouzan, F. (2013). Application of 2-D geoelectrical resistivity tomography for subsurface cavity detection in the eastern part of Saudi Arabia. *Geoscience Frontiers*, *4*(4), 469-476.

Moghadas, D., & Badorreck, A. (2019). Machine learning to estimate soil moisture from geophysical measurements of electrical conductivity. *Near Surface Geophysics*, *17*(2), 181-195.

Moghadas, D., Jadoon, K. Z., & McCabe, M. F. (2019). Spatiotemporal monitoring of soil moisture from EMI data using DCT-based Bayesian inference and neural network. *Journal of Applied Geophysics*, *169*, 226-238.

Muñoz-Carpena, R., & Dukes, M. D. (2015). Automatic irrigation based on soil moisture for vegetable crops.

Padarian, J., Minasny, B., McBratney, A. B., & Dalgliesh, N. (2014). Predicting and mapping the soil available water capacity of Australian wheatbelt. *Geoderma Regional*, *2*, 110-118.

Paltineanu, I. C., & Starr, J. L. (1997). Real-time soil water dynamics using multisensor capacitance probes: Laboratory calibration. *Soil Science Society of America Journal*, *61*(6), 1576-1585.

Poggio, L., Gimona, A., Brown, I., Castellazzi, M., 2010. Soil available water capacity interpolation and spatial uncertainty modelling at multiple geographical extents. *Geoderma*, 160 (2), 175–188.

Reynolds, S. G. (1970). The gravimetric method of soil moisture determination Part IA study of equipment, and methodological problems. *Journal of Hydrology*, *11*(3), 258-273.

Rhoades, J. D., Raats, P. A. C., & Prather, R. J. (1976). Effects of liquid-phase electrical conductivity, water content, and surface conductivity on bulk soil electrical conductivity. *Soil Science Society of America Journal*, *40*(5), 651-655.

Rhoades, J. D., Manteghi, N. A., Shouse, P. J., & Alves, W. J. (1989). Soil electrical conductivity and soil salinity: New formulations and calibrations. *Soil Science Society of America Journal*, *53*(2), 433-439.

Robinson, D. A., Jones, S. B., Wraith, J. M., Or, D., & Friedman, S. P. (2003). A review of advances in dielectric and electrical conductivity measurement in soils using time domain reflectometry. *Vadose Zone Journal*, *2*(4), 444-475.

Robinson, D. A., Lebron, I., Lesch, S. M., & Shouse, P. (2004). Minimizing drift in electrical conductivity measurements in high temperature environments using the EM-38. *Soil Science Society of America Journal*, *68*(2), 339-345.

Robinson, M., & Dean, T. J. (1993). Measurement of near surface soil water content using a capacitance probe. *Hydrological processes*, *7*(1), 77-86.

Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A., ... & Namkhai, A. (2000). The global soil moisture data bank. *Bulletin of the American Meteorological Society*, *81*(6), 1281-1300.

Rossel, R. V., & Behrens, T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, *158*(1-2), 46-54.

Rucker, D. F., Glaser, D. R., Osborne, T., & Maehl, W. C. (2009). Electrical resistivity characterization of a reclaimed gold mine to delineate acid rock drainage pathways. *Mine Water and the Environment*, *28*(2), 146-157.

Schrott, L., & Sass, O. (2008). Application of field geophysics in geomorphology: advances and limitations exemplified by case studies. *Geomorphology*, *93*(1-2), 55-73.

Sheets, K. R., & Hendrickx, J. M. (1995). Noninvasive soil water content measurement using electromagnetic induction. *Water resources research*, *31*(10), 2401-2409.

Sherlock, M. D., & McDonnell, J. J. (2003). A new tool for hillslope hydrologists: spatially distributed groundwater level and soilwater content measured using electromagnetic induction. *Hydrological Processes*, *17*(10), 1965-1977.

Smith, K. A. (2000). Soil and environmental analysis: physical methods, revised, and expanded. CRC Press.

Son, J. K., Shin, W. T., & Cho, J. Y. (2017). Laboratory and field assessment of the Decagon 5TE and GS3 sensors for estimating soil water content in saline-alkali reclaimed soils. *Communications in soil science and plant analysis*, *48*(19), 2268-2279.

Stanley, J. N., Lamb, D. W., Falzon, G., & Schneider, D. A. (2014). Apparent electrical conductivity (ECa) as a surrogate for neutron probe counts to measure soil moisture content in heavy clay soils (Vertosols). *Soil Research*, *52*(4), 373-378.

Swinbourne, M., Sparrow, E., Hatch, M., Bowden, T., & Taggart, D. (2014). Using near-surface geophysics to assist with the management of southern hairy-nosed wombats (Lasiorhinus latifrons) in South Australia. *The Leading Edge*, *33*(12), 1356-1362.

Triantafilis, J., Roe, J. A. E., & Monteiro Santos, F. A. (2011). Detecting a leachate plume in an aeolian sand landscape using a DUALEM-421 induction probe to measure electrical

conductivity followed by inversion modelling. *Soil use and management*, *27*(3), 357-366.

Ulriksen, C. P. F. (1982). Application of impulse radar to civil engineering. *Lund Univ. Dr. Thesis*.

Van Genuchten, M. T. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil science society of America journal*, *44*(5), 892-898.

Van Meirvenne, M., Islam, M. M., De Smedt, P., Meerschman, E., Van De Vijver, E., & Saey, T. (2013). Key variables for the identification of soil management classes in the aeolian landscapes of north–west Europe. *Geoderma*, *199*, 99-105.

Vitharana, U. W., Van Meirvenne, M., Cockx, L., & Bourgeois, J. (2006). Identifying potential management zones in a layered soil using several sources of ancillary information. *Soil Use and Management*, *22*(4), 405-413.

Wellbrock, K., Voß, P., Heemeier, B., Keilholz, P., Patzelt, A., & Grottker, M (2018). The water management of Tayama and other ancient oasis elements in the north-western Arabian Peninsula. *Preliminary Synthesis*.

Wang, P., Hu, Z., Zhao, Y., & Li, X. (2016). Experimental study of soil compaction effects on GPR signals. *Journal of Applied Geophysics*, *126*, 128-137.

Woodward, J., & Burke, M. J. (2007). Applications of ground-penetrating radar to glacial and frozen materials. *Journal of Environmental & Engineering Geophysics*, *12*(1), 69-85.

Xing, C., Chen, N., Zhang, X., & Gong, J. (2017). A machine learning based reconstruction method for satellite remote sensing of soil moisture images with in situ observations. *Remote Sensing*, *9*(5), 484.

Zerizghy, M. G., Van Rensburg, L. D., & Anderson, J. J. (2013). Comparison of neutron scattering and DFM capacitance instruments in measuring soil water evaporation. *Water SA*, *39*(2), 183-190.

Zhao, D., Zhao, X., Khongnawang, T., Arshad, M., & Triantafilis, J. (2018). A Vis-NIR Spectral Library to Predict Clay in Australian Cotton Growing Soil. Soil *Science Society of America Journal*, *82*(6), 1347-1357.

Zhao, D., Li, N., Zare, E., Wang, J., & Triantafilis, J. (2020). Mapping cation exchange capacity using a quasi-3d joint inversion of EM38 and EM31 data. *Soil and Tillage Research*, *200*, 104618.

Zhao, X., Wang, J., Zhao, D., Li, N., Zare, E., & Triantafilis, J. (2019). Digital regolith mapping of clay across the Ashley irrigation area using electromagnetic induction data and inversion modelling. *Geoderma*, *346*, 18-29.

Zhou, W., Beck, B. F., & Stephenson, J. B. (2000). Reliability of dipole-dipole electrical resistivity tomography for defining depth to bedrock in covered karst terranes. *Environmental geology*, *39*(7), 760-766.

Zhou, Q. Y., Shimada, J., & Sato, A. (2001). Three-dimensional spatial and temporal monitoring of soil water content using electrical resistivity tomography. *Water Resources Research*, 37(2), 273-285.

Zhou, L., Yu, D., Wang, Z., & Wang, X. (2019). Soil water content estimation using high-frequency ground penetrating radar. *Water*, *11*(5), 1036.

Zhu, Q., Lin, H., & Doolittle, J. (2010). Repeated electromagnetic induction surveys for determining subsurface hydrologic dynamics in an agricultural landscape. *Soil Science Society of America Journal*, *74*(5), 1750-1762.

Chapter 3: Time-lapse imaging of soil moisture in a flood irrigation field monitored using electromagnetic conductivity imaging: Wetting phase.

3.1 Introduction

The lower Namoi valley in northern New South Wales, is a highly productive agricultural area. In the 1950's it was synonymous for prime hard wheat production as well as sheep and cattle grazing. Its agricultural versatility (Triantafilis et al., 2001) is a function of the Vertosols, which have large clay content (>35 %), shrink-swell properties that exhibit strong cracking when dry and slickensides and/or lenticular structural aggregates at depth, that dominate the plains. These properties allow for easy passage of moisture deep into the subsoil before the self-mulching properties seal it in. In the 1960's, these Vertosol plains were recognised for their potential development for irrigation. Specifically, furrow irrigated cotton production. However, there have been issues with deep drainage (Woodforth et al., 2012) and threat of salinization (Huang et al., 2016); owing to water use inefficiencies (Triantafilis et al., 2003). In addition, there are competing demands for water from mining, environment and agricultural commodities. In addition, climate change forecasts for south-east Australia suggest reduced rainfall and hotter conditions (Dai, 2013).

To better manage the water resource and improve efficiencies in water-use, irrigation farmers need methods to measure the volumetric water content of the soil (θ, $m^3\ m^{-3}$), in real-time. This will provide information about when and how the water is being applied and allow for the monitoring of water-use to avoid water-logging and deep drainage. Instruments which can measure and monitor the θ have been used to assist irrigators. In the 1970's much research was carried out using neutron probes. However, issues related to ease of use, regulatory requirements with probes containing a radioactive source and poor resolution in the topsoil led to other instruments being adopted. In the early 1980's time-domain-reflectometry was *de riguer* with much work being carried out to understand water use efficiency (Wraith et al.,

2005). In addition, instrumentation which can be set-and-forget and left to log remotely would be advantageous.

In this regard, domain frequency sensors such as the Decagon GS3 (METER Group, Inc. USA) have shown to be useful in measuring and monitoring the soil θ continuously without site specific calibration. The Decagon GS3 sensor measures; θ, temperature (^{o}C) and bulk electrical conductivity (σ_b, mS m^{-1}). However, even these types of sensors are expensive, and installation require major disturbance of the topsoil (0-0.3 m), subsurface (0.3-0.6 m) and subsoil (>0.6 m). The information is also only available at point locations. Recently, electromagnetic (EM) induction instruments have been used to provide spatial data of the soil apparent electrical conductivity (EC_a, mS m^{-1}) and for applications in precision agriculture (Corwin et al., 2003). Examples include calibration of EC_a for soil moisture monitoring (Kachanoski et al., 1988). Although, EM instruments only provide a bulk measurement of EC_a, inversion of EC_a is providing useful estimates of the soil true electrical conductivity (σ, mS m^{-1}) which can be used to calibrate soil properties at the same depths at which they are measured. Examples include calibration for soil salinity (Zare et al., 2015), clay (Zhao et al., 2019) and cation exchange capacity (Koganti et al., 2018).

In terms of θ, Huang et al. (2017a) established a calibration between σ and θ along a single uniform loamy sand transect. The data collection was conducted over a 12-day period to monitor water-use beneath an overhead-irrigated Lucerne field in San Jacinto, southern California. A similar approach was used beneath a single olive tree to monitor θ in Cordoba Spain (Martinez et al., 2018). In another study near Camden in southern Sydney, NSW Australia Huang et al. (2017b) used an ensemble Kalman filter (EnKF) to monitor and nowcast the 3-dimensional θ dynamics. The study field had varying soil texture and sprinkler irrigation system. In a small furrow irrigated plot located at the Australian Cotton Research Institute, the aim was to use a similar approach to monitor θ across a uniformly heavy-clay soil by i)

installing θ sensors within the root zone, ii) collecting and correcting EC_a using a DUALEM-421 frequently on a one-day-irrigation event, iii) invert EC_a using EM4Soil to estimate σ in various depths; iv) understand relationship between σ and soil properties, v) establish a relationship between θ and σ; and, vi) map spatio-temporal variation of θ with depth.

3.2 Materials and methods

3.2.1 Study site

The study field is in the New South Wales Department of Primary Industries, Australian Cotton Research Institute (201 m above sea level; 30° 11′ 41″ S, 149° 36′ 20″ E) Australia (Figure 3.1a). The field is located at the southern end of field C1 and is located on an expansive alluvial floodplain of uniform topography (Stannard and Kelly, 1968). The uppermost sediments are a fine texture upon which grey-self mulching (vertic) clay (i.e., Grey Vertosols, Isbell 2002) soil has developed. It is used extensively for irrigated cotton (Gossypium sp.) production (Triantafilis et al., 2000). The plot sown with a cotton–wheat rotation on permanent beds where wheat stubble was retained as in situ mulch.

The climate is semi-arid, with mean annual rainfall of 643 mm and potential evaporation over 1,800 mm. The rainfall is slightly summer dominant with a mean-maximum in January (79.8 mm) and -minimum in August (37.5 mm). Temperature varies from a mean-maximum in January (35.3 ^{o}C) to a -minimum in July (17.0 ^{o}C). Given the soil type and availability of irrigation water, irrigation was introduced in the 1960's, with over 60,000 hectares now developed for furrow irrigated cotton production.

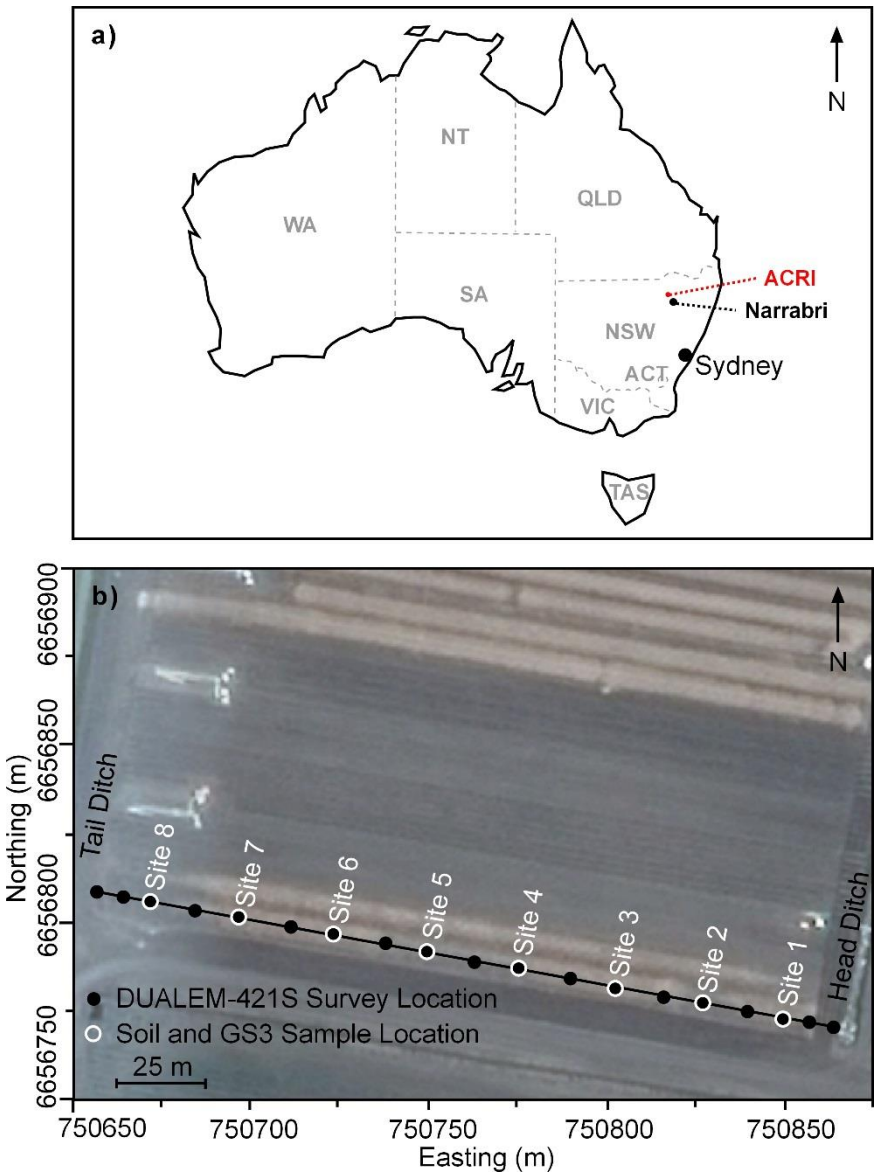


Figure 3.1 Location of a) study area at the Australian Cotton Research Institute (ACRI), and b) soil volumetric water content (θ, m^3 m^{-3}) sensor locations (8) and DUALEM-421 measurement locations.

3.2.2 Soil moisture sensors installation and soil samples collection

To measure and monitor θ Decagon GS3 soil moisture sensors were used. The GS3 also measures temperature (^oC) and bulk electrical conductivity (σ_b, mS m^{-1}). The reason it was selected was factory calibration for θ is accurate to ±0.03 m^3 m^{-3} in mineral soils, when the bulk soil EC is < 10 dS m^{-1}. This has been examined and confirmed by Son et al. (2017) in soil with different textures. The pits were spaced 20 m apart. They represent θ in the topsoil (0-0.3 m), subsurface (0.3-0.6 m), shallow- (0.6-0.9 m) and deep-subsoil (1.2-1.5 m). The GS3

sensors were connected to a Decagon Em50 data logger, which was configured to take measurements every 5 minutes. The field was irrigated (furrow) on 13 December 2017.

At the 8 selected sites, and from the head and tail-ditch, soil samples were also collected at depths of 0.15, 0.45, 0.75, and 1.35 m. The soil samples were sealed in plastic bags. The samples were oven-dried (60 °C) and ground to pass a 2-mm sieve. The hydrometer method was used (with 5 % sodium hexametaphosphate as the dispersing agent) to determine the particle size fractions (i.e., clay, silt and sand). Soil salinity (EC_e, dS m^{-1}) was estimated from $EC_{1:5}$ using conversion factor of 6.7 for heavy clay soils (DPI, 2014) and pH were also determined using a 1:5 soil:water extract.

3.2.3 EM instrumentation and EC_a data collection

To measure soil EC_a (mS m^{-1}) a DUALEM-421S was used. The DUALEM-421 (DUALM Inc., Milton, ON, Canada) simultaneously measures soil apparent electrical conductivity (EC_a, mS m^{-1}). It consists of a transmitting coil (Tx) that operates at 9.0 kHz and three pairs of receiver coils (Rx). The Tx and one Rx pair have horizontal windings which form a horizontal coplanar array (HCP). The distance between the Tx to the coplanar Rx are 1, 2, and 4 m. The notation 1mHcon, 2mHcon, and 4mHcon represent EC_a and correspond to measurements of approximately 0-1.5, 0-3.0, and 0-6.0 m, respectively. The other coils in each Rx pair have vertical windings and with the Tx forms perpendicular arrays (PRP). The distances between the Tx to the Rx are 1.1, 2.1, and 4.1 m, respectively. The respective EC_a measurements are represented by 1mPcon, 2mPcon, and 4mPcon with theoretical depth of corresponding to approximately 0-0.5, 0-1.0, and 0-2.0 m, respectively.

The DUALEM-421S sensor has internal thermometers and a thermal compensation system. Prior to shipping the instrument, multiple multi-hour tests are carried out to adjust the settings of the thermal compensation system until the effect of temperature on EC_a is

minimized. According to the technical specifications of the instrument the compensation system typically keeps the effect of a 5 C° temperature change on EC_a to a fraction (< 2 mS m^{-1}).

The first EC_a measurements were collected at the head-ditch at 10 m spacing, with 18 measurements made across the field (Figure 3.1b). A final measurement was made near the tail-ditch. The instrument was positioned directly on top and aligned in the direction of the bed (east-west). The first EC_a were collected before irrigation (0 hours). Later, five subsequent EC_a passes were made after irrigation commenced, specifically; 2, 4, 6, 9 and 24 hours. The irrigation ceased after 9 hours.

While the DUALEM-421S can be mounted on a sled and to speed up EC_a collection (Triantafilis et al., 2013), this approach was not used herein, since irrigated Vertosols are too adhesive and too challenging to drag a sled across a recently irrigated furrow. Moreover, because of the presence of small plants, it was not possible to drag the DUALEM-421S directly above the bed. As such, a station-by-station approach was adopted. To compare inversion with all six possible EC_a available from a DUALEM-421, the amount of EC_a data was also degraded by considering the EC_a that would be available from the following types of instruments, which can also be purchased from the manufacturer and including; DUALEM-21, DUALEM-1 and DUALEM-2.

3.2.4 Correction of ECa data based on soil temperature and ambient temperature

Soil temperature effects on EC_a need to be corrected. Corwin and Lesch (2005) found this could be done by multiplying the raw EC_a data using the following equation:

$$f_{\mathrm{T}} = 0.4470 + 1.4034e^{-T/26.815} \quad (1)$$

where *T* is the soil temperature. However, many of these studies (e.g., Ma et al., 2011) focused on the conventional EM38 meter (Geonics, Mississauga, Ontario, Canada), which has a depth

of exploration (DOE) of 1.5 m. Given the DOE of DUALEM-421S is up to ~6 m, Equation (1) might not satisfactorily account for soil temperature with location and with depth.

An alternative approach was proposed by Saey et al. 2009. In their research and to account for the heterogenous nature of soil temperature variation, they first inverted the EC_a data and then calibrated the depth-specific electrical conductivity (σ) using Equation (1). Herein, they assumed that the EC_a data collected at different times did not drift due to ambient (air) temperature.

In research carried out recently by Huang et al. (2017c), the diurnal effect on the drift of a DUALEM-421S was tested with regard to ambient and soil temperature. It was found that ambient temperature had a larger effect on stability of EC_a compared with soil temperature. The drift in EC_a was also found to be a function of temperature-dependent hysteresis, with the overall stability of EC_a increasing with coil array length.

Herein, the EC_a associated with the following coil arrays of 1mHcon, 2mPcon, 2mHcon and 4mPcon, were adjusted to a standard ambient temperature for the first survey (i.e., 24 °C). This was done using a step-wise multiple linear regression (MLR) modelling approach embedded in JMP 12.2 Software (SAS, 2015). The input predictors included ambient temperature, internal sensor temperature, average soil temperature, θ at each depth and EC_a measurements of 1mPcon, 4mHcon. The backward elimination technique was used to select useful predictors (with a probability of less than 0.001). Once the MLR models was determined, with 1mHcon, 2mPcon, 2mHcon and 4mPcon EC_a from all 5 surveys corrected using the same temperature condition.

3.2.5 Quasi-2D inversion of DUALEM-421 data

To invert the DUALEM EC_a, the EM4Soil software package (EMTOMO, 2014) was used. The algorithm works for EC_a collected at low induction numbers (Monteiro Santos et al., 2010). Primarily, the software generates electromagnetic conductivity images (EMCI) and

specifically estimates of σ by inverting EC_a obtained from EM instruments. The estimates of σ are constrained by neighbouring sites, with an optimal EMCI arrived at by comparing results from a number of parameters, including the choice of a forward model, inversion algorithm and damping factor (λ).

In forward modelling, forward calculations and derivatives consider the cumulative function (CF) or a full-solution (FS) model. The latter is preferable when EC_a data is acquired over conductive soil (i.e., > 100 mS m^{-1}). The modelling is conducted using a 1-dimensional laterally constrained approach (Auken et al., 2002), where 2-dimensional smoothness constraints are imposed. *The inversion* algorithms are variations of Occam regularization and include algorithm S1 and S2; the latter constrains EMCI variation around a reference model and is smoother than S1. With a forward model and algorithm selected a damping factor (λ) can also be varied. This generally tends to smooth the data with increasing larger values. A similar effect occurs when choosing the number of iterations.

Herein, the aim was to determine the optimal combination of inversion parameters (i.e., λ, S1 or S2, CF or FS) for the quasi-2D inversion of DUALEM-421 EC_a. The λ was set from 0.07, 0.3, and at 0.3 increments thereafter to a maximum of 3.0. Inversion of EC_a data collected on each day were generated with a maximum of 20 iterations. Owing to the large EC_a (~100 mS m^{-1}) of vertosol, the initial model was set of the σ to a value of σ = 100 $mS\ m^{-1}$.

To determine the most suitable parameters, the largest coefficient of determination (R^2) between σ and θ was identified. This had been performed by obtaining σ from the EMCI on the first survey and on each of the 5 surveys and correlating these with estimated θ on the corresponding time. Additionally, various combination of DUALEM EC_a was tested including DUALEM-21, DUALEM-1 and DUALEM-2. The ability of these data was compared to predict Decagon GS3 measured σ_b using Lin's Concordance and against estimated σ.

3.2.6 Establishing an empirical calibration model between θ and σ

To develop an empirical model between σ *and* θ, an Artificial Neural Network (ANNs) spatial modelling technique was chosen. This is because it is a data-driven approach, which simulates features of neural networks that address a range of difficult information processing, analysis and modeling problems (De Smith et al., 2007). One advantage of ANN is that it enables pattern recognition through classification tools that have the ability to generalize from input data that is imprecise (Porwal et al., 2003). When there is a non-linear relationship, this feature is useful; whereby ANNs allow predictive modelling of complex natural situations which are not readily observable.

The relationship between σ and θ may be complicated owing to the potential number of soil properties which influence σ. This includes, not only θ, but also potentially clay, mineralogy and salinity. Herein the attempt was to minimize this by considering a soil transect where the clay, mineralogy and EC_e variation was small. In terms of θ, the response with σ may also be complicated herein. This is because of differences of θ with depth prior to the first irrigation and owing to a dry topsoil and wetter subsurface and subsoil which were close to field capacity (θ at –10 to –33 kPa). In addition, θ behind the wetting front of irrigation would be close to saturated in the topsoil whilst in front, it would still be near permanent wilting point (θ at –1500 kPa).

Given the complexity of the relationship between σ and θ, an ANN modelling approach embedded in JMP 12.2 (SAS, 2015) was used. The JMP platform applies a penalty to the parameters in the optimization of the model to prevent the overfitting problem that can occur with Neural models. In this process, once further increasing or decreasing of the penalty parameter is no longer improving the model's loglikelihood, the algorithm terminates. This may require hundreds to thousands of iterations. The validation proceeds using a "holdback" method. Specifically, part of the data (1/3rd) was with-held automatically and used as a

validation set. The hyperbolic tangent (TanH) activation function with 1 hidden layer was used to fit the ANNs. TanH function is a sigmoid function and transforms values to be between -1 and 1. The hyperbolic tangent function is:

$$\frac{e^{2x} - 1}{e^{2x} + 1}$$

where x is a linear combination of the X variables. Here, 3 nodes were chosen because fewer nodes failed to reflect the complex inter-relationship and too many nodes led to over-fitting of the ANN model.

The performance of the predictions for calibration and validation was assessed using the coefficient of determination (R^2), root mean square error (RMSE), sum of squared error (SSE), log-likelihood, and mean absolute deviation (MAD).

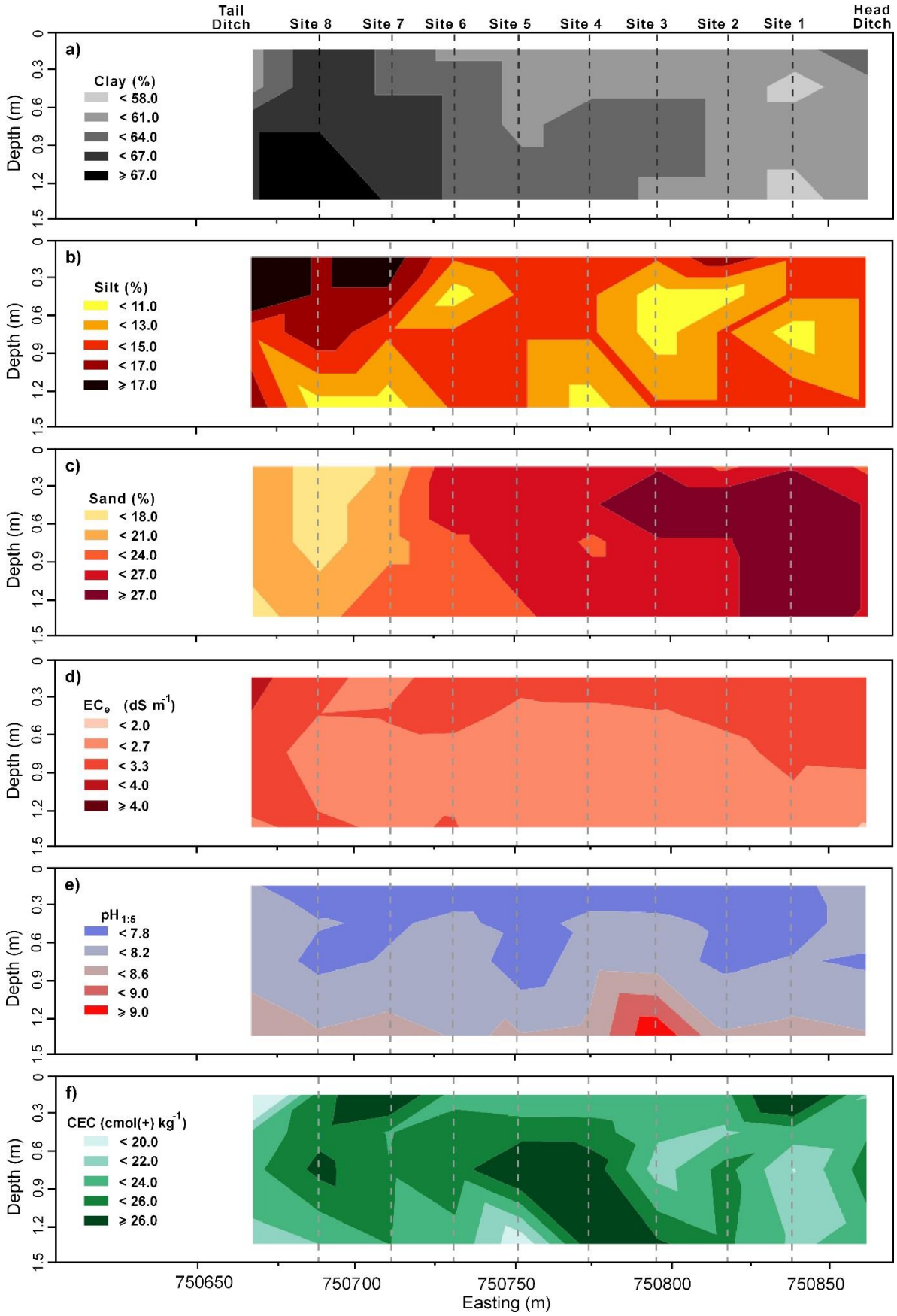


Figure 3.2 Contour plots of various soil properties measured at the 8 sampling sites and including, a) clay (%), b) silt (%), c) sand (%), d) estimated electrical conductivity of the saturation extract (EC_e, dS m^{-1}), and e) pH (1:5 soil:water extract).

3.3 Results and discussion

3.3.1 Distribution of soil properties along transect

Figure 3.2 shows contour plots of physical and chemical properties along the transect; *generated using Contour Plot Tool in JMP 12.2 (SAS, 2015).* Figure 3.2a, 2b and 2c, show the three particle size fractions (%); clay, silt and sand, respectively. Figure 3.2a, shows clay was small (<58 %) near the head-ditch, increasing towards the tail-ditch. Figure 3.2b shows silt was small and uniform (11-17%), with sand (18-27%) a little larger, but small compared to clay. As clay increased toward the tail-ditch, sand decreased (Figure 3.2c) from east (>=27 %) to west (< 18 %). The texture was heavy clay.

Figure 3.2d shows the distribution of EC_e. In the topsoil (0-0.3 m), EC_e (2.7-3.3 dS m^{-1}) was slightly saline (2-4 dS m^{-1}). This level of EC_e has minimal effect on tolerant crops like cotton, because salinity does not become problematic until much larger (i.e., > 7.7 dS m^{-1}) values (Mass et al., 1977). The slightly-saline EC_e was most likely a function of the water quality, given the slightly larger values in the topsoil. In the subsurface (0.3-0.6 m) and upper- (0.6 -0.9 m) and lower-subsoil (1.2-1.5 m), EC_e was non-saline (<2 dS m^{-1}). Figure 3.2e shows distribution of $pH_{1:5}$ was slightly alkaline (< 7.8) in the topsoil and subsurface. In the upper-subsoil, it increased to moderate alkalinity (8.2-8.6) at site 3. In the deeper-subsoil it was strongly alkaline (8.6-9). Figure 3.2f shows CEC was uniform and between 20 and 30 cmol(+) kg^{-1}; suggesting soil has moderate shrink-swell potential.

3.3.2 Summary statistics of raw and corrected EC_a and θ

Table 3.1 shows summary statistics of EC_a (mS m^{-1}) before irrigation and prior to data correction. The 1mPcon had the smallest mean (82.1 mS m^{-1}). This was followed closely by the 1mHcon (103.6) and 2mPcon (109.1). The largest mean EC_a was measured by 4mPcon (119.4), followed by 2mHcon (112.8) and then 4mHcon (97.6). It was noted the CV was larger

for EC_a at shallower theoretical depths, with 1mPcon largest (12.5) followed by 1mHcon (10.9). The data was not skewed, except 1mPcon (-1.2). These results can be attributed to the variable nature of topsoil soil properties; in particular interplay between transient θ, clay, and EC_e.

Table 3.1 Summary statistics of soil apparent electrical conductivity (EC_a, mS m^{-1}) measured by a DUALEM-421S prior to the commencement of irrigation (0 hours) in perpendicular (1mPcon, 2mPcon and 4mPcon) and horizontal coplanar (1mHcon, 2mHcon and 4mHcon).

Sensor	Num.	Min	Mean	Medium	Max	Std. Dev.	CV	Skewness
1mPcon	19	50.3	82.1	82.9	104.4	10.2	12.5	-1.2
2mPcon	19	85.6	109.1	108.6	132.5	11	10.1	0.4
1mHcon	19	83.3	103.6	102	127.2	11.3	10.9	0.6
4mPcon	19	106.2	119.4	116.6	137.9	8.9	7.5	0.5
2mHcon	19	98.6	112.8	110	133.6	8.3	7.9	0.8
4mHcon	19	88.2	97.6	97.9	104.9	5.2	5.3	-0.1

Table 3.2 Multiple linear regression models used to correct the drift of soil apparent electrical conductivity (EC_a mS m-1) measured by a DUALEM-421S prior to the commencement of irrigation (0 hours) in perpendicular (1mPcon, 2mPcon and 4mPcon) and horizontal coplanar (1mHcon, 2mHcon and 4mHcon).

Sensor	Model
1mPcon	N/A
2mPcon	(-36913.30) + 0.81 × 1mPcon + 1.04 × 4mHcon +0.05 * Easting + 2.80 × Soil Temp (0.3-0.6m) + -16.26 × σ_b (0.3-0.6m)
1mHcon	(-34095.89) + 0.7 × 1mPcon + 1.27 × 4mHcon + 0.05 × Easting + -1.24 × Soil Temp (0-0.3m) + 1248.93 × θ (1.2-1.5m)
4mPcon	(-6205.82) + 0.43 × 1mPcon + 1.49 × 4mHcon + 0.01 × Easting + 3.25 × Soil Temp (0.3-0.6m) + -25.67 × σ_b (0.3-0.6m)
2mHcon	(-36070.87) + 0.2 × 1mPcon + 1.70 × 4mHcon +0.05 × Easting + 35.59 × Soil Temp (1.2-1.5m) + -31.95 × σ_b (0.6-0.9m)
4mHcon	N/A

Table 3.2 shows the multiple linear regression (MLR) models used to correct EC_a (mS m^{-1}) before irrigation. No correction was applied to 1mPcon. It was also noted that in none of the MLR models developed, the ambient temperature, internal sensor temperature or average soil temperature were selected during the step-wise backward elimination.

With regard to 2mPcon, EC_a from the 1mPcon and 4mHcon as well as Easting, soil temperature (0.3-0.6 m) and σ_b (0.3-0.6 m) from the GS3 sensors were useful. To correct the deeper 1mHcon, 1mPcon, 4mHcon and Easting with different depths of soil temperature (0-

0.3 m) and σ_b (1.2-1.5 m) from the GS3 sensors were selected. The predictors for 4mPcon were identical with the 2mPcon. The predictors for the 2mHcon were 1mPcon and 4mHcon, Easting, soil temperature (1.2-1.5 m) and σ_b (0.6-0.9 m).

Table 3.3 shows the summary statistics of corrected EC_a (mS m^{-1}). The mean of the 2mPcon increased slightly from the measured (109.1 *mS* m^{-1}) and after correction (111.6). The same was the case for the mean of measured (103.6) and corrected (106) 1mHcon. The mean measured 4mPcon (119.4) decreased after correction (97.3), whereas the 2mHcon mean (112.8) and corrected (111.1) was mostly unchanged. Interestingly, the CV for all corrected EC_a decreased (except for 4mPcon) and the data was generally less skewed (except for 1mHcon).

Table 3.3 Summary statistics of corrected soil apparent electrical conductivity (EC_a, mS m^{-1}) for a DUALEM-421S prior to the commencement of irrigation (0 hours) in perpendicular (1mPcon, 2mPcon and 4mPcon) and horizontal coplanar (1mHcon, 2mHcon and 4mHcon).

Sensor	Num.	Min	Mean	Medium	Max	Std. Dev.	CV	Skewness
1mPcon	19	50.3	82.1	82.9	104.4	10.2	12.5	-1.2
2mPcon	19	95.5	111.6	110	132.5	8.9	8	0.3
1mHcon	19	90.8	106	103.4	127.7	9.1	8.6	0.7
4mPcon	19	81.1	97.3	95.7	118.1	9	9.2	0.3
2mHcon	19	94.7	111.1	110.2	125.3	8.3	7.4	-0.2
4mHcon	19	88.2	97.6	97.9	104.9	5.2	5.3	-0.1

Table 3.4 Summary statistics of corrected apparent electrical conductivity (EC_a, mS m^{-1}) measured by 1-m horizontal coplanar array (1mHcon) of DUALEM-421S at various times before (0 hours) and after (2, 4, 6, 9 and 24 hours) irrigation commenced.

hrs	Num.	Min	Mean	Medium	Max	Std. Dev.	CV	Skewness
0	19	90.8	106	103.4	127.7	9.1	8.6	0.7
2	19	97.4	106.72	105.5	124.3	6.8	6.4	1.1
4	19	97.4	109.6	114.1	124.8	7.6	6.9	0.1
6	19	97.7	110.2	110.2	125.1	7.7	7.2	0.5
9	19	100.8	113	111.5	132	8	8	0.8
24	19	101.9	114.6	114.1	131.4	8.4	7.3	0.4

Table 3.5 Summary statistics of a) average measured soil volumetric water content (θ, m^3 m^{-3}) and b) measured bulk electrical conductivity (σ_b) for all eight measurement locations at various depths, including topsoil (0.15 m), subsurface (0.45 m), upper subsoil (0.75 m) and deeper subsoil (1.35 m) and at various times before (0 hours) and after (2, 4, 6, 9 and 24 hours) irrigation commenced.

a) θ (m^3 m^{-3})	Depth (m)			
hrs	0.15	0.45	0.75	1.35
0	0.31	0.49	0.51	0.53
2	0.34	0.49	0.51	0.53
4	0.36	0.50	0.51	0.53
6	0.41	0.50	0.52	0.53
9	0.49	0.51	0.52	0.53
24	0.44	0.51	0.52	0.53

b) σ_b (mS m^{-1})	Depth (m)			
hrs	0.15	0.45	0.75	1.35
0	35	108	130	140
2	41	107	130	143
4	47	112	130	143
6	56	113	134	143
9	80	122	134	145
24	67	120	136	145

3.3.3 Summary statistics of time-lapse corrected EC_a, θ and σ_b

Table 3.4 shows the summary statistics of corrected 1mHcon at various survey times (i.e., 2, 4, 6, 9 and 24 hours) after irrigation commenced. As expected, the average 1mHcon was smallest before irrigation (106 mS m^{-1}) and reflects the drier topsoil. As the soil moisture increases with the advancement of the wetting front, 1mHcon gradually increases after each survey to an average maximum (114.6 mS m^{-1}) 24 hours after irrigation commenced.

Table 3.5a shows summary statistics of average GS3 θ for all sites at all depths and various times before (0 hours) and after (2, 4, 6, 9 and 24 hours) irrigation commenced. Prior

to irrigation, average topsoil θ (0.31 m^3 m^{-3}) was smallest. According to Gardiner and Miller (2004) this would be at the lower end of field capacity for θ for heavy-clay (~60 %). Conversely, subsurface (0.49 m^3 m^{-3}), upper- (0.51 m^3 m^{-3}) and deeper-subsoil (0.53 m^3 m^{-3}) θ was larger; akin to upper end of field capacity (0.42 m^3 m^{-3}) for heavy-clay (Gardiner and Miller, 2004).

The θ suggest there was enough subsurface and subsoil moisture available for the cotton crop. However, it also indicates the difficulty in resolving the difference in θ in a furrow irrigated field. This was because of the thin veneer of drier and therefore less conductive topsoil, which overlies a much larger volume of more conductive subsurface and subsoil θ. This can be seen in Table 3.5b, which shows equivalent σ_b data. Specifically, before irrigation commenced, upper- (130 mS m^{-1}) and deeper-subsoil σ_b was large (140 mS m^{-1}), whereas topsoil σ_b (35 mS m^{-1}) was much smaller.

In general, average topsoil θ along the transect increases slowly with time. This was expected, given at two hours (0.34 m^3 m^{-3}) the wetting front had passed the first two site locations of sensors and by nine hours had passed all eight site locations. This can be seen in Figure 3.4b and f, respectively, which show the contour plot of measured θ. At this point in time, θ was at field capacity (0.49 m^3 m^{-3}). The average subsoil θ did not change (0.51-0.53 m^3 m^{-3}), however. Table 3.5b, shows average σ_b in the subsoil was more or less consistent (130-145 mS m^{-1}).

3.3.4 Distribution of EC_a and measured θ along transect

Figure 3.3 shows corrected EC_a from various arrays, prior to irrigation and at various times. Figure 3.3a shows 1mPcon. Before irrigation, 1mPcon was intermediate-small (90 mS m^{-1}) near the head-ditch (site 1). It steadily increased toward the tail-ditch (site 8) to an

intermediate value (100 mS m^{-1}). After 2 hours of irrigation, 1mPcon increased, coalescing with 1mPcon near site 2 (750815); recorded before irrigation commenced. This indicates 1mPcon increasing as a function of increasing θ with irrigation advance along the transect.

The same effect can be seen with 1mPcon, 4 and 6 hours after irrigation commenced; whereby 1mPcon again coalesces with 0 hour 1mPcon near sites 4 and 6, respectively. These results suggest irrigation advance was 20 m per hour. The coalescing of EC_a also suggests the adopted approach to correct EC_a was robust. Figure 3.3b shows 2mPcon data also varies between 80-120 mS m^{-1}, although at the tail-ditch end, it was a little larger; exceeding 120 mS m^{-1}. This suggests subsoil conductivity was larger here and most likely a function of the slightly larger clay shown in Figure 3.2a.

Figure 3.3c shows 1mHcon. The change in EC_a from this array appears to have the largest variation prior to commencement, to 24 hours after irrigation. From here on, with equivalent or increasing depth of measurement, EC_a change was smaller, when comparing prior to commencement and 24 hours after irrigation. This was particular the case for the 4mHcon (Figure 3.3f).

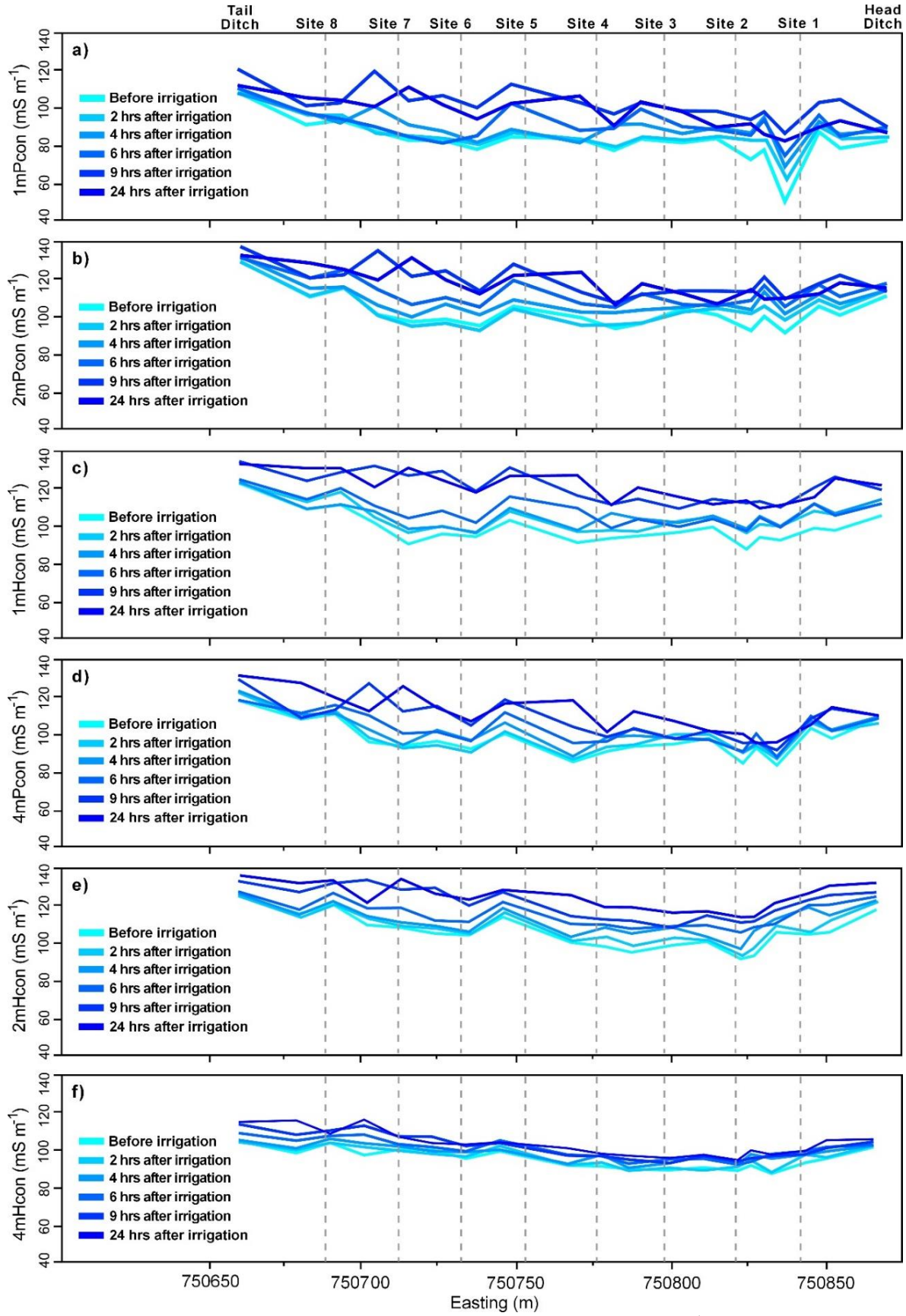


Figure 3.3 Distribution of measured apparent soil electrical conductivity (EC_a, mS m^{-1}) by DUALEM-421 along transect; including; a) 1mPcon, b) 2mPcon, c) 1mHcon, d) 4mPcon, e) 2mHcon and f) 4mHcon collected prior to commencement of irrigation (0 hours) and various times post-irrigation commencing (2, 4, 6, 9 and 24 hours). Note: Irrigation ceased after 9 hours.

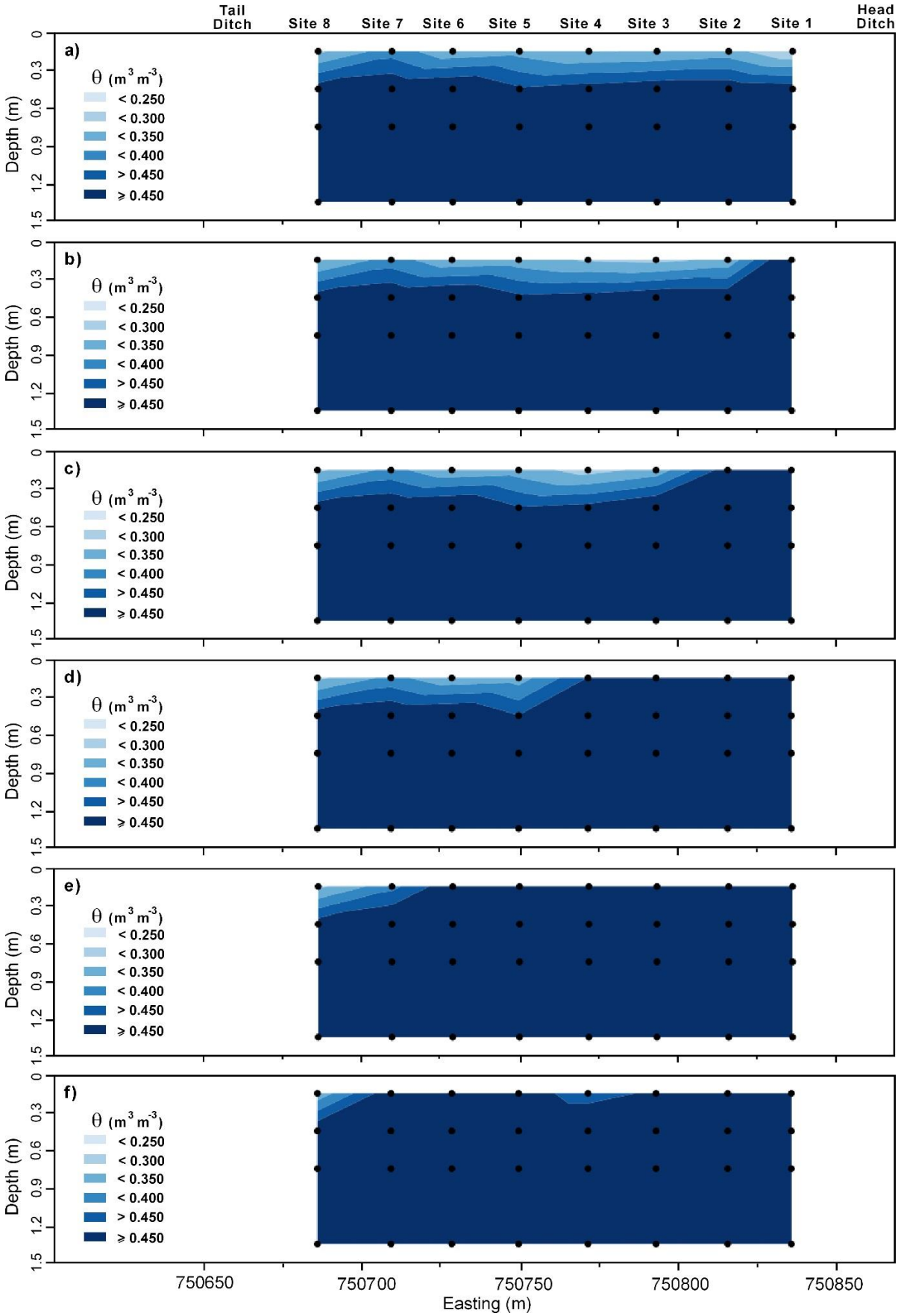


Figure 3.4 Distribution of measured soil volumetric water content (θ, m^3 m^{-3}) using Decagon GS3 sensors along transect; including; a) prior to commencement of irrigation (0 hours) and various times post-commencement and at b) 2, c) 4, d) 6, e) 9 and f) 24 hours. Note: Irrigation ceased after 9 hours.

Figure 3.4 shows θ along the transect and using the 8 sites where the GS3 sensors were installed. Figure 3.4a shows θ, before irrigation commenced (0 hour), was near saturation (>0.45 m^3 m^{-3}) below the topsoil (0-0.15 m). In the topsoil, θ was drier. Figure 3.4b shows θ two hours after irrigation commenced with Figure 3.4c, d, e and f showing θ at 2, 4, 6, 9 and 24 hours post-irrigation, respectively.

3.3.5 Coefficient of determination (R^2) between estimated σ and measured θ

Figure 3.5a shows coefficient of determination (R^2) between σ and measured θ at all depths (i.e., 0.15, 0.45, 0.75, 1.35 m) and all survey times (0, 2, 4, 6, 9 and 24 hours) when different sets of inversion parameters were considered; including, forward modelling (i.e., CF and FS), inversion (i.e., S1 and S2) algorithms and relative to different values of λ.

Regardless of which forward model or inversion algorithm was used, larger R^2 were achieved between σ and measured θ, when λ was small (0.3). This is because larger values of λ generally tending to smooth inversion models. The CF forward model generally produced consistently larger R^2 compared to the FS forward modelling algorithm and relative to λ.

The largest R^2 (0.46) between estimates of σ derived from DUALEM-421 EC_a and θ was achieved using CF forward model, S1 inversion algorithms and λ = 0.3. The usefulness of inversion of different combinations of DUALEM EC_a, including DUALEM-1, DUALEM-2 and DUALEM-21 were investigated. Table 3.6 shows inversion of DUALEM-2 estimated σ was unsatisfactory (R^2 = 0.07) to establish a correlation with θ. However, equivalent R^2 was achieved between σ and θ by inverting DUALEM-1 (0.47) and DUALEM-21 (0.46).

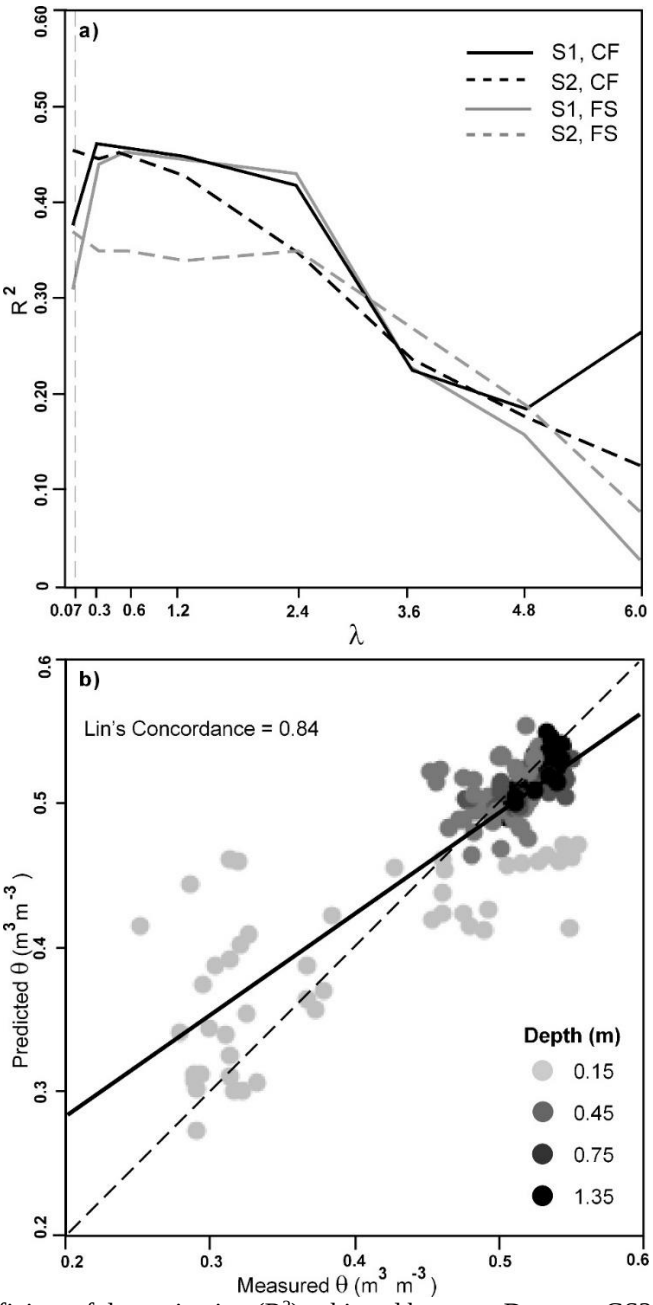


Figure 3.5 Plot of a) coefficient of determination (R^2) achieved between Decagon GS3 measured soil volumetric water content (θ, $m^3\ m^{-3}$) and estimated true electrical conductivity (σ, $mS\ m^{-1}$) generated by inverting DUALEM-421S EC_a using EM4Soil quasi 3-d model, cumulative function (CF) and full solution (FS), inversion algorithms S1 or S2 and varying damping factor (λ), b) measured soil θ ($m^3\ m^{-3}$) versus predicted soil θ ($m^3\ m^{-3}$) generated using the Artificial Neural Network (ANN) empirical model.

Given these equivocal results, the estimates of σ were compared by inversion of DUALEM-421, DUALEM-1 and DUALEM-21 EC_a, relative to measured σ_b from all the GS3 sensor and at all depths. It was found that the best agreement between measured σ_b and estimated σ was achieved when the inverted DUALEM-421 EC_a was used given it had the

largest Lin's concordance (0.84) with measured σ_b compared to DUALEM-21 (0.83) and DUALEM-1 (0.54). *Therefore, the DUALEM-421 EC_a data along with* using CF inversion algorithm, S1 forward model and $\lambda = 0.3$ *were selected to estimate σ and use it to predict θ from an artificial neural network (ANN) modelling approach.*

3.3.6 Understanding the relationship between soil properties and σ

To understand the relationship between soil properties and σ, Table 3.7 shows coefficient of determination (R^2) between measured soil properties and σ using EC_a data collected before irrigation commenced and using S1, CF and a $\lambda = 0.3$. The R^2 between σ *and* particle size fractions was small and not correlated with clay ($R^2 = 0.17$), silt ($R^2 = 0.15$) or sand ($R^2 = 0.01$). The R^2 with EC_e (0.52) and $pH_{1:5}$ ($R^2 = 0.36$) was larger but not significantly correlated. Additionally, σ was not correlated with CEC (R^2=0.01). The largest R^2 was between σ and θ ($R^2 = 0.80$). These results indicate that for the most part the change in EC_a and estimated σ was a function of changing θ.

Table 3.6 Coefficient of determination (R^2) between σ and θ for all eight measurement locations at various depths, including topsoil (0.15 m), subsurface (0.45 m), upper subsoil (0.75 m) and deeper subsoil (1.35 m) and at various times before (0 hours) and after (2, 4, 6, 9 and 24 hours) irrigation commenced.

σ (CF, S1,0.3)	DUALEM-1	DUALEM-2	DUALEM-21	DUALEM-421
R^2	0.47	0.07	0.46	0.46

Table 3.7 Coefficient of determination (R^2) between measured soil properties and estimated true electrical conductivity (σ, mS m^{-1}) using DUALEM-421S EC_a data collected before irrigation and using the cumulative function (CF) modelling algorithm, inversion model S1 and a damping factor (λ) = 0.3. Note: EC_e, soil salinity estimated from measured electrical conductivity of 1:5 soil:water extract; $pH_{1:5,}$ pH of 1:5 soil:water extract; CEC (cmol(+) kg^{-1}) Cation Exchange Capacity; θ (m^3 m^{-3}), volumetric water content.

Soil property	R^2
Clay%	0.17*
Silt%	0.15*
Sand%	0.01
EC_e	0.52***
$pH_{1:5}$	0.36***
CEC	0.01
θ	0.80***
*, P < 0.05; **, P < 0.01; ***, P < 0.0	

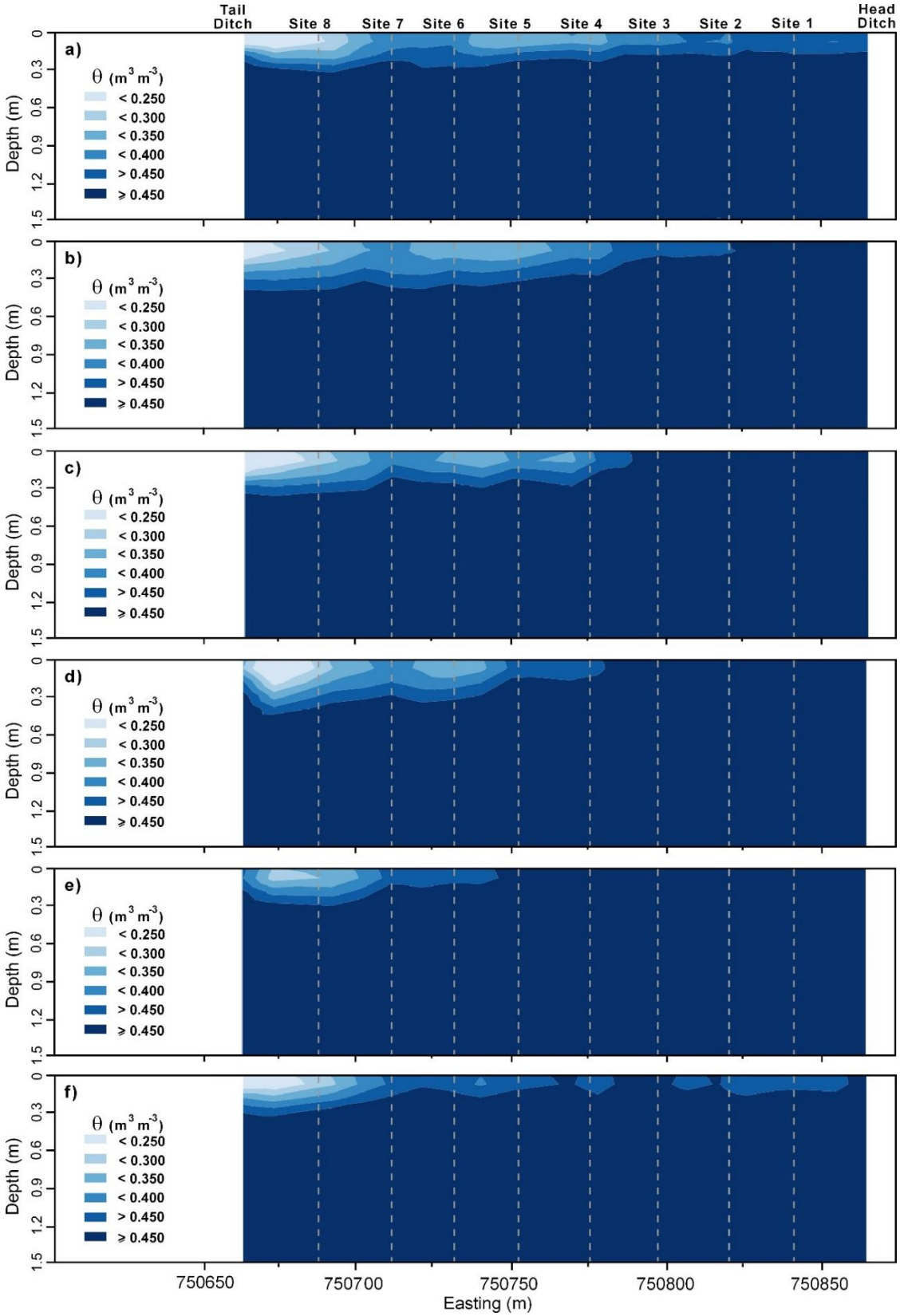


Figure 3.6 Distribution of predicted soil volumetric water content (θ, m^3 m^{-3}) using ANN model developed between estimated σ (mS m^{-1}) and measured θ (m^3 m^{-3}) using Decagon GS3 sensors along the transect; including; a) prior to commencement of irrigation (0 hours) and various times post-commencement and b) 2, c) 4, d) 6, e) 9 and f) 24 hours. Note: Irrigation ceased after 9 hours.

3.3.7 Artificial Neural Network modelling between σ and θ

Table 3.8 shows summary statistics of ANN model generated using σ, Easting and depth. In terms of calibration dataset, the model performance was good given the large coefficient of determination (R^2 = 0.74), accuracy (RMSE = 0.044) and mean absolute deviation (MA = 0.026). The model performance was also good for validation, whereby a similarly large coefficient of determination (R^2 = 0.71), accuracy (RMSE = 0.046) and mean absolute deviation (MAD = 0.025) was obtained.

Table 3.8 Summary statistics of the artificial neural network model (ANN) generated using σ (CF, S1,0.3), Easting and depth . Note: R^2, coefficient of determination; RMSE, root mean square error; MAD, mean absolute deviation; and SSE, sum of squared errors.

σ (CF, S1,0.3), Easting, depth	R^2	RMSE	MAD	Log likelihood	SSE	Sum Freq.
Calibration set	0.74	0.044	0.026	-273.2	0.308	160
Validation set	0.71	0.046	0.025	-132.9	0.169	80

3.3.8 ANN modelling of θ using σ, Easting and depth

Figure 3.5b shows measured versus predicted θ, generated using ANN model at various times before and after irrigation at various depths. The model was accurate, given large Lin's concordance correlation coefficient (0.84) and precise (RMSE = 0.04 $m^3\ m^{-3}$). For the most part, when measured θ was large (>0.5 $m^3\ m^{-3}$), predicted θ was also large (>0.5 $m^3\ m^{-3}$). This was the case in the deeper- (1.35 m) and upper-subsoil (0.75 m). The same was true for measured (>0.45 $m^3\ m^{-3}$) and predicted (>0.45 $m^3\ m^{-3}$) θ in the subsurface (i.e., 0.45 m). These results were attributable to θ being near saturation, prior to commencement of irrigation and because there was little change in subsoil θ.

Figure 3.5b shows there were a small number of points, where measured and predicted topsoil (0.15 m) θ were not near the 1:1 relationship. This was the case for four points where the measured topsoil θ was small (0.2-0.35 $m^3\ m^{-3}$) and predicted θ was intermediate (0.4-0.45 $m^3\ m^{-3}$). These θ were from the first three sites in the topsoil and near the head ditch. The over-

prediction was due to the DUALEM-421 measurement of EC_a, was influenced by water flowing past the sensors in the furrow and the saturated soil conditions therein.

When the inverted EC_a adopted, this leaded to large estimates in σ and therefore larger predicted θ. However, the topsoil sensors were placed in the seed-bed and above the furrow and at the time of measurement of θ, the water had not infiltrated to a close enough proximity to the sensor, for it to be measured. This was also the case for five points, with similar measured θ (0.275-0.325 $m^3 m^{-3}$) and slightly over-predicted θ (0.35-0.4 $m^3 m^{-3}$). These points were again associated with topsoil θ (0.15 m) but at sites 2, 3, 4, 5 and 6.

Another reason for over-predication was a function of the 10 m spacing between DUALEM-421 points. Given the modelling was conducted using a 1-dimensional laterally constrained approach (Auken et al., 2002), where 2-dimensional smoothness constraints were imposed, collecting EC_a on 10 m spacing leads to smoothing of σ. This would be most problematic when irrigation water has advanced past an EC_a measurement, but not the next moisture sensor down the furrow where the next EC_a measurement was made. As a result, EC_a would have been large at the first EC_a measurement site before the sensor and lead to over-prediction of σ, and therefore over-prediction of θ, at the moisture sensor site compared to measured θ.

In future research, a minimum 5 m EC_a sample spacing can be recommended, however, a major problem will still be the length and weight of carrying a DUALEM-421 and across a very wet Vertosol field. Alternatively, a much smaller and compact DUALEM-1HS or the use of a Geonics EM38Mk2 might be more appropriate.

3.3.9 Spatio-temporal distribution of predicted soil θ along the study transect

Figure 3.6 shows the spatial distribution of predicted θ using the ANN calibration and applied to the inverted EC_a data collected at the 19 measurement sites and at the various time-

steps. Figure 3.6a shows predicted θ before irrigation commenced (0 hour). Prior to irrigation, θ was near saturation (>0.45 m^3 m^{-3}) below the topsoil (0-0.15 m) and subsurface (0.15-0.30 m). In the topsoil, θ was drier at the tail-ditch end as compared to the head-ditch. Predicted θ was equivalent to measured θ (Figure 3.4a) Figure 3.6b shows predicted θ two hours after irrigation commenced. It was apparent that the wetting front can be discerned, where topsoil θ increases toward saturation (0.4 - 0.45 m^3 m^{-3}). It has approximately advanced 75 m down the transect (750800) and was predicted to be at site 2, given soil was saturated (>0.45 m^3 m^{-3}).

However, this does not align exactly with measured θ shown in Figure 3.4b. As mentioned above, this was because at the time of measurement of the DUALEM-421, the water had advanced to site 2 and was infiltrating and saturating the soil in the furrow but, as yet, had not infiltrated close enough to the GS3 sensor in the bed to enable measurement. Hence the over-prediction of θ. Figure 3.6c and Figure 3.4c show the same lag in wetting-front prediction and measurement was apparent where the former was predicted to be mid-way between site 3 and 4 and the latter between site 2 and 3.

Figure 3.6d shows θ six hours after irrigation commenced. It was apparent that the predicted location of the wetting front had approximately advanced 125 m (750750) and was near site 4. This location was equivalent to where the measured θ was at or near field capacity (>0.45 m^3 m^{-3}). From here the measured and predicted θ were in close alignment, although it was apparent that predicted θ was being slightly under-predicted with the wetting front lagging.

Figure 3.6e shows θ nine hours after irrigation commenced. It appears that predicted wetting front does not reach the end of the field but advanced to a maximum of 200 m (750600). Irrigation ceased at this point in time. Figure 3.6f shows predicted θ, twenty-four hours after irrigation commenced. The predicted θ in the topsoil and at the head-ditch was now draining and the soil was no longer saturated; perhaps approaching the drained upper limit or field

capacity. This was not reflected, however, in measured θ (Figure 3.4f). It was not immediately clear why there was such a disparity.

The accomplished results herein are comparable to those of others. For example, Huang et al. (2017a) were able to find a good relationship between σ and θ beneath an overhead pivot-irrigation system on sandy clay loam soil in San Jacinto, California. The reason for their success was a function of soil being much drier prior to commencement of overhead irrigation. In addition, irrigation was uniform beneath the pivot irrigation system because as soon as the sprinklers were turned on, irrigation commenced along the entire transect. When irrigation was completed and EC_a was collected, there was no difference in θ along the transect and as a function of the advance of irrigation water down a furrow and during the various time-steps of measurement. In addition, the EC_a data was collected over a 12-day period and there was markedly more variation in θ, because it was measured immediately after irrigation commenced and then during the drying-phase. The EC_a was also collected on an approximate 1 m spacing.

More recently, Huang et al. (2017b) were able to find a good relationship between σ and θ beneath a sprinkler-irrigation system on various soil textures across an alluvial plain near Camden, in western Sydney, Australia. The reason for their success was similarly a function of the soil being much drier, prior to the commencement of irrigation and EC_a being measured prior to irrigation commencing. As such, baseline data was available and near permanent wilting point. Again, irrigation was uniform beneath the pivot irrigation system, because as soon as the sprinklers were turned on, irrigation commenced along the entire transect. In addition, the EC_a was collected over a much longer period of time (i.e., 16 days over a 38 day period) and again there was more variation in θ, because it was measured during an intense wetting-up phase and then a prolonged drying-phase.

In both of these studies, the results were more convincing because the EC_a was collected on-the-go with spacing between measurements of approximately 1 m. This means that more location specific estimates of σ would be made with less smoothing in estimates of σ as a function of the 1-d laterally constrained inversion. As such, predicted θ would be more reliable relative to measured θ.

Given the difficulty of measuring soil EC_a using a DUALEM-421, owing to the length and weight of the instrument and trafficking across a saturated Vertosol field, to make it easier to collect EC_a data, smaller EM instruments such as the EM38-Mk2 or DUALEM-1HS could be considered. This would also be beneficial as EC_a data could be collected from 0.5 m coil spacing, which could potentially further assist in resolving topsoil and subsurface θ differences, leading to improved calibration and prediction.

3.4 Conclusions

The collection of DUALEM-421 EC_a and its inversion to produce estimates of σ provides useful information, which can be calibrated against θ across a furrow-irrigated cotton field. The calibration provided reasonable predictions of θ. The results could be understood in terms of monitoring θ during wetting-up phase of irrigation. In terms of irrigation management, this approach has potential benefits, which include ability to monitor water using time-lapse methods for; (i) improved timing of irrigation, (ii) evaluation of irrigation efficiency, and (iii) optimize water use efficiency by maintaining leaching of salts at a level matching leaching requirement.

However, given the difficulty of measuring soil EC_a using a DUALEM-421, owing to the length and weight of the instrument and trafficking across a saturated Vertosol field, EC_a could only be collected at 10 m spacing. It can be recommended the spacing needs to be reduced to at least 5 m. In addition, owing to the main variation in θ being confined to the topsoil (0.15

m) and subsurface (0.45 m), either a EM38-Mk2 or DUALEM-1HS instruments could be considered. This would be of benefit as more EC_a could be collected at 1 m spacing and from a 0.5 m coil spacing. In the case of the DUALEM-1HS, this would enable the collection of 0.5mPcon and 0.5mHcon data. This data would greatly assist in resolving topsoil and subsurface θ leading to improved calibration and prediction.

3.5 References

Auken, E., Foged, N., & Sørensen, K. I. (2002, September). Model recognition by 1-D laterally constrained inversion of resistivity data. In *8th EEGS-ES Meeting* (pp. cp-36). European Association of Geoscientists & Engineers.

Corwin, D. L., Kaffka, S. R., Hopmans, J. W., Mori, Y., Van Groenigen, J. W., Van Kessel, C., ... & Oster, J. D. (2003). Assessment and field-scale mapping of soil quality properties of a saline-sodic soil. *Geoderma*, *114*(3-4), 231-259.

Corwin, D. L., & Lesch, S. M. (2005). Characterizing soil spatial variability with apparent soil electrical conductivity: I. Survey protocols. *Computers and electronics in agriculture*, *46*(1-3), 103-133..

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature climate change*, *3*(1), 52-58.

Decagon Devices Inc, 2365 NE Hopkins Ct, Pullman, WA 99163, USA

De Smith MJ, Goodchild MF, Longley P (2007) Geospatial analysis: a comprehensive guide to principles, techniques and software tools. Troubador Publishing Ltd.

EMTOMO (2014) EM4Soil Version 2. EMTOMO, R. Alice Cruz 4, Odivelas, Lisboa, Portugal.

Gardiner DT, Miller RW (2004) Soils in Our Environment. 10th ed. Prentice-Hall, Inc., Upper Saddle River, New Jersey.

Huang, J., Prochazka, M. J., & Triantafilis, J. (2016). Irrigation salinity hazard assessment and risk mapping in the lower Macintyre Valley, Australia. *Science of the Total Environment*, *551*, 460-473.

Huang, J., Prochazka, M. J., & Triantafilis, J. (2016a). Irrigation salinity hazard assessment and risk mapping in the lower Macintyre Valley, Australia. *Science of the Total Environment*, *551*, 460-473.

Huang, J., McBratney, A. B., Minasny, B., & Triantafilis, J. (2017). 3D soil water nowcasting using electromagnetic conductivity imaging and the ensemble Kalman filter. *Journal of hydrology*, *549*, 62-78.

Huang, J., Minasny, B., Whelan, B. M., McBratney, A. B., & Triantafilis, J. (2017a). Temperature-dependent hysteresis effects on EM induction instruments: An example of single-frequency multi-coil array instruments. *Computers and Electronics in Agriculture*, *132*, 76-85.

Isbell RF (2002) The Australian soil classifi cation. Rev. ed. CSIRO Publ., Collingwood, VIC, Australia

Kachanoski, R. G., Wesenbeeck, I. V., & Gregorich, E. G. (1988). Estimating spatial variations of soil water content using noncontacting electromagnetic inductive methods. *Canadian Journal of Soil Science*, *68*(4), 715-722.

Koganti, T., Moral, F. J., Rebollo, F. J., Huang, J., & Triantafilis, J. (2017). Mapping cation exchange capacity using a Veris-3100 instrument and invVERIS modelling software. *Science of the Total Environment*, *599*, 2156-2165.

Martinez, G., Huang, J., Vanderlinden, K., & Triantafilis, J. (2018) Potential to monitor soil moisture beneath an olive tree using quasi-2d inversion modelling. *Soil Use and Managent 34*,236-248.

Ma, R., McBratney, A., Whelan, B., Minasny, B., & Short, M. (2011). Comparing temperature correction models for soil electrical conductivity measurement. *Precision Agriculture*, *12*(1), 55-66.

Maas, E. V., & Hoffman, G. J. (1977). Crop salt tolerance—current assessment. *Journal of the irrigation and drainage division*, *103*(2), 115-134.

Monteiro Santos, F. A., Triantafilis, J., Bruzgulis, K. E., & Roe, J. A. E. (2010). Inversion of multiconfiguration electromagnetic (DUALEM-421) profiling data using a one-dimensional laterally constrained algorithm. *Vadose zone journal VZJ*..

Porwal, A., Carranza, E. J. M., & Hale, M. (2003). Artificial neural networks for mineral-potential mapping: a case study from Aravalli Province, Western India. *Natural resources research*, *12*(3), 155-171.

SAS Institute (2015) JMP Version 12.2. SAS Institute Inc., Cary, North Carolina, USA.

Saey, T., Simpson, D., Vermeersch, H., Cockx, L., & Van Meirvenne, M. (2009). Comparing the EM38DD and DUALEM-21S sensors for depth-to-clay mapping. *Soil Science Society of America Journal*, *73*(1), 7-12.

Son, J. K., Shin, W. T., & Cho, J. Y. (2017). Laboratory and field assessment of the Decagon 5TE and GS3 sensors for estimating soil water content in saline-alkali reclaimed soils. *Communications in soil science and plant analysis*, *48*(19), 2268-2279.

Stannard, ME., Kelly, I.D. (1968) The irrigation potential of the lower Gwydir Valley. Water Conservation and Irrigation Commission.

Tenison, K., Wild, J., Madden, E., Draper, V., McMahon, L. (2014). Salinity Training Manual,320 pp. State of New South Wales through the Department of Trade and Investment,Regional Infrastructure and Services, Sydney.

Triantafilis, J., Laslett, G. M., & McBratney, A. B. (2000). Calibrating an electromagnetic induction instrument to measure salinity in soil under irrigated cotton. *Soil Science Society of America Journal*, *64*(3), 1009-1017.

Triantafilis, J., Ward, W. T., & McBratney, A. B. (2001). Land suitability assessment in the Namoi Valley of Australia, using a continuous model. *Soil Research*, *39*(2), 273-289.

Triantafilis, J., Huckel, A., & Odeh, I. (2003). Field-scale assessment of deep drainage risk. *Irrigation Science*, *21*(4), 183-192.

Triantafilis, J., & Santos, F. M. (2013). Electromagnetic conductivity imaging (EMCI) of soil using a DUALEM-421 and inversion modelling software (EM4Soil). *Geoderma*, *211*, 28-38.

Wraith, J. M., Robinson, D. A., Jones, S. B., & Long, D. S. (2005). Spatially characterizing apparent electrical conductivity and water content of surface soils with time domain reflectometry. *Computers and Electronics in Agriculture*, *46*(1-3), 239-261.

Woodforth, A., Triantafilis, J., Cupitt, J., Malik, R. S., Subasinghe, R., Ahmed, M. F., ... & Geering, H. (2012). Mapping estimated deep drainage in the lower Namoi Valley using a chloride mass balance model and EM34 data. *Geophysics*, *77*(4), WB245-WB256.

Zare, E., Huang, J., Santos, F. M., & Triantafilis, J. (2015). Mapping salinity in three dimensions using a DUALEM-421 and electromagnetic inversion software. *Soil Science Society of America Journal*, *79*(6), 1729-1740.

Zhao, X., Wang, J., Zhao, D., Li, N., Zare, E., & Triantafilis, J. (2019). Digital regolith mapping of clay across the Ashley irrigation area using electromagnetic induction data and inversion modelling. *Geoderma*, *346*, 18-29.

Chapter 4: Two-dimensional time-lapse imaging of soil wetting and drying cycle using EM38 data across a flood irrigation cotton field

4.1 Introduction

Vertosols, which have large clay content (>35 %), shrink-swell properties that exhibit slickensides (i.e., shiny ped surfaces) and/or lenticular structural aggregates at depth, dominate the clay alluvial plains of northern New South Wales. Owing to the cracking nature of the soil, Vertosols have been developed extensively for irrigated agricultural production. In the northern Murray-Darling Basin, this has mostly been for cotton production. However, owing to the issue of deep drainage (Silburn et al. 2014) and identification of water use inefficiencies (Roth et al. 2014) in flood irrigation, improvements in irrigation management are required. This is because climate change forecasts and modelling (Dai 2013) for south-east Australia, suggest reduced rainfall, hotter conditions and therefore less water for evapotranspiration. Moreover, there are increasing and competing demands for water from mining, environment and domestic applications.

To improve water use efficiency and irrigation management, farmers need to be able to measure and monitor soil volumetric water content (θ, $m^3\ m^{-3}$) during wetting and drying cycles. To generate the necessary information about the spatial variation in θ, data is required at multiple depths and from various locations. However, the standard method (i.e., laboratory measured gravimetric moisture), is time consuming and labour intensive. Moreover, it involves destructive sampling and requires determination of the soil bulk density (ρ, gcm^{-3}) (Brevik and Fenton 2004). In addition, θ measurements are required daily. Alternatively, field sensors which can be installed *a priori*, have been used to measure and monitor θ. For example, Kodikara et al. (2014) and Al-Ain et al. (2009) used neutron probes in urban and salt-affected areas, respectively. However, Chanasyk et al., (1996) indicated issues related to measuring θ in the near-surface and smoothing of θ across texture discontinuities. Time domain reflectometry

(TDR) or capacitance probes (CP), can provide depth specific measurements of θ (Baumhardt et al. 2000). Moreover, domain frequency sensors (e.g., Decagon GS3) have also been used. However, these sensors and data loggers are expensive given multiples are required to measure and monitor θ in topsoil (-0.15 m), subsurface (-0.45 m) and subsoil (< -0.6 m).

To value add to these sensors, time-lapse electrical resistivity tomography measurements can be used to extrapolate between their locations (Moghadas and Baddorreck, 2019). Electromagnetic (EM) induction instruments have also been proposed as a source of digital data for soil moisture monitoring. This data can be used to value add to the point θ data and extrapolate across at the field scale. This is because the soil apparent electrical conductivity (EC_a, mS m^{-1}), which is measured from EM instruments, is influenced by θ owing to the polarity of the water molecule. While EC_a is also related to other soil properties, such as clay content, mineralogy and salinity, if these three are similar across a given field, most of the response or changes in EC_a over time and during a wetting-drying cycle, can be modelled relative to θ. In terms of θ, Huang et al. (2017) established a calibration between estimates of true electrical conductivity (σ, mS m^{-1}) derived from inversion of DUALEM-421 EC_a and θ (at 5 sites) along a single uniform loamy sand transect and over a 12-day period to model and monitor water-use beneath an overhead-irrigated Lucerne field in San Jacinto, southern California. A similar 2-d approach was used at 8 sites but across a sprinkler irrigated field near Camden in southern Sydney, NSW Australia (Huang et al. 2017a). Therein an ensemble Kalman filter (EnKF) approach was developed in 2-d, and subsequently used to monitor and nowcast 3-dimensional θ dynamics on 20 separate days over a period of 40 days. On a plot scale, a similar approach (using 5 sites) was used beneath an olive tree to monitor θ in Cordoba Spain (Martinez et al. 2018) from estimates of σ (from a DUALEM-21).

In most research pertaining to EM inversion and monitoring of θ, single-frequency multi-coil sensors have been used to estimate σ and related to time-varying θ at multiple

calibration sites using either a DUALEM (e.g., Huang, et al. 2017) or CMD mini-Explorer (Moghadas et al., 2019). In addition, multiple sites have been used to develop a calibration using mostly simple-linear or curve-linear relationships. In this research, the aim was to determine if EC_a measured from a EM38 in horizontal (EM38h) or vertical (EV38v) modes at various heights (0, 0.2, 0.4, 0.6, 0.8 and 1.0 m) can be used to value add to limited θ. It was also determined if a relationship can be established between σ and θ using mathematical (i.e., multiple linear regression (MLR), random forest (RF), Cubist and support vector machine (SVM), Artificial Neural Networks (ANN)) models. Furthermore, the number of calibration sites required to make a satisfactory model was determined and evaluated.

4.2 Materials and methods

4.2.1 Study site

The study field is in the lower Namoi valley of northern New South Wales, Australia (Figure 4.1a). It is located at the New South Wales Department of Primary Industries, Australian Cotton Research Institute (ACRI) (201 m above sea level; 30° 11′ 41″ S, 149° 36′ 20″ E), which is 20 km to west of Narrabri. The experiment was sited at the southern end of experimental field C1, which is used for furrow irrigation to trial various cotton rotation farming systems; The plot is sown with a cotton–wheat rotation on permanent beds where wheat stubble was retained as in situ mulch. The field run is approximately 200 m long.

The climate is semi-arid, with mean annual rainfall of approximately 643 mm and potential evaporation over 1,800 mm. The rainfall is slightly summer dominant with a mean maximum in January (79.8 mm) and minimum in August (37.5 mm). Temperature varies from a mean maximum in January (35.3 ^{o}C) to a minimum in July (17.0 ^{o}C). Stannard and Kelly (1977) recognized three major physiographic units near the field, including clay plain, prior stream, and terminal prior stream deposits. The clay plain dominates, characterised by a grey-

self mulching (vertic) clay (i.e., Grey Vertosols) (Isbell 2002). These surround the ACRI and are used extensively for irrigated cotton production (Triantafilis et al. 2000).

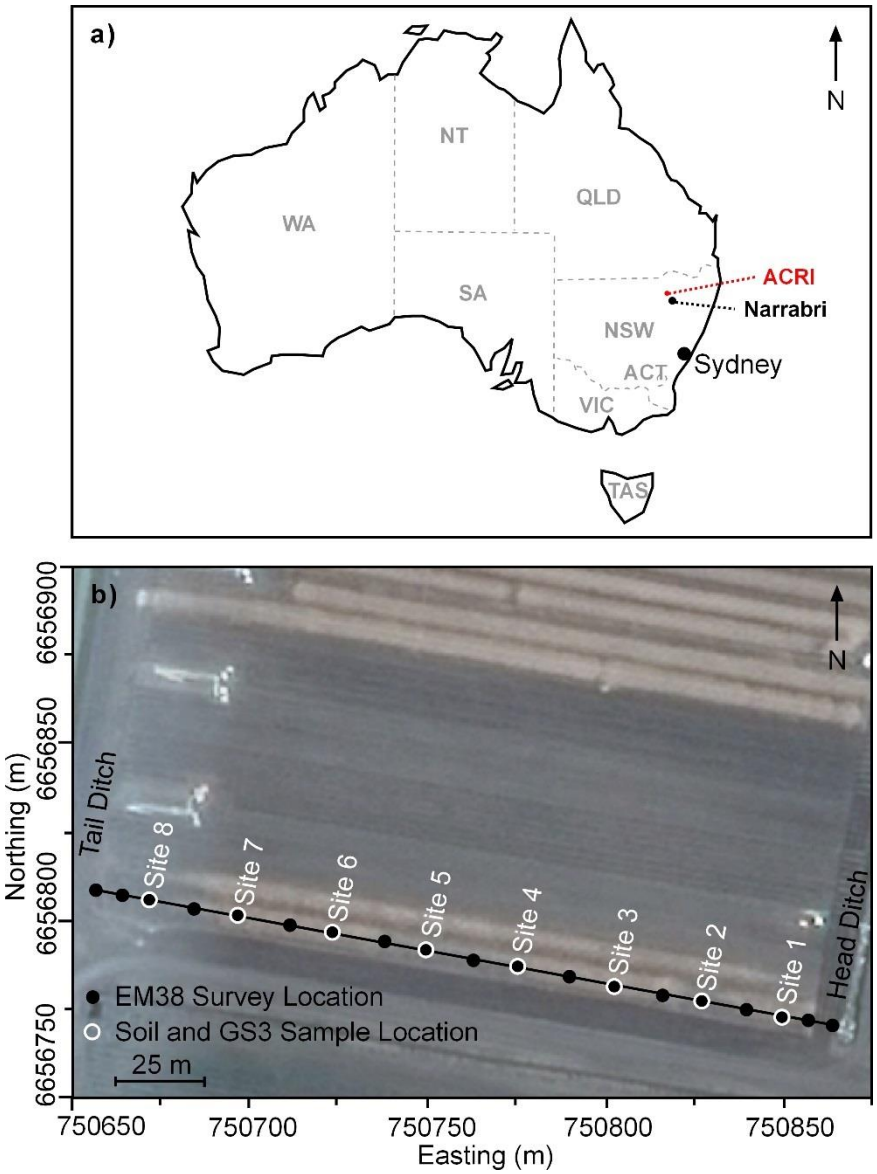


Figure 4.1 Location of a) study area at the Australian Cotton Research Institute (ACRI), and b) soil volumetric water content (θ, $m^3\ m^{-3}$) sensor locations (8) and EM38 measurement locations (19).

4.2.2 Soil moisture sensors installation and soil samples collection

To monitor θ ($m^3\ m^{-3}$), 8 sites located approximately 20 m apart were selected. Site 1 was located 20 m from the head-ditch at the eastern end of the field, while site 8 was positioned 20 m from the tail-ditch to the west. At each site, a large pit was excavated. Decagon GS3 (METER Group, Inc. USA) sensors were then installed at depths of 0.15, 0.45, 0.75, and 1.35

m. They approximately represent θ in the topsoil (0-0.3 m), subsurface (0.3-0.6 m), shallow- (0.6-0.9 m) and deep-subsoil (1.2-1.5 m), respectively. The reason that this type of sensor was selected is, factory calibration for θ is accurate to ±0.03 $m^3 m^{-3}$ in mineral soils, when the bulk soil EC is < 10 dS m^{-1}. This has been examined and confirmed by Son et al. (2017) in soils with different textures (e.g., silty clay). The sensors are capable of measuring not only θ (m^3 m^{-3}) but also temperature (^oC) and bulk electrical conductivity (σ_b , mS m^{-1}). These three measured data were collected and logged every 5 minutes using a Decagon Em50 during the entire duration of the experiment.

The field was irrigated (furrow) on 31 January 2018. Water was applied from the head-ditch in the east and flowed toward the tail-ditch under the influence of gravity. At the 8 selected sites and two extra points near the head and tail-ditch and from within the soil pits, soil samples were collected at depths of 0.15, 0.45, 0.75, and 1.35 m. The samples were oven-dried (60 °C) and ground to pass a 2-mm sieve. The hydrometer method was used (with 5 % sodium hexametaphosphate as the dispersing agent) to determine the particle size fractions (i.e., clay, silt and sand).

Soil salinity (EC_e, dS m^{-1}) was estimated from $EC_{1:5}$ using conversion factor (6.7) for heavy clay soil (Tenison et al., 2014). The pH was determined from a 1:5 soil:water extract. Tucker's (1974) method was used to determine the cation exchange capacity (CEC, cmol(+) kg^{-1}) using a mechanical leaching device. Briefly, samples were first washed with 60% ethanol to remove any soluble salts, followed by extraction with 1M NH_4Cl. The extracts were analyzed using an inductively coupled plasma optical emission spectrometry (ICP-OES).

4.2.3 EM instrumentation and EC_a data collection

Soil apparent electrical conductivity (EC_a) was collected using a Geonics EM38 (Mississauga Ontario, Canada). The EM38 EC_a was measured in horizontal (EM38h) and

vertical (EM38v) modes. Given operating frequency (i.e., 14.5 kHz) and coil spacing (1.0 m), the theoretical depth of exploration is, respectively, 0.75 ($EM38h_0$) and 1.5 m ($EM38v_0$) when the instrument is on the ground (McNeill 1990). In this study, EM38h and EM38v EC_a was measured at heights of 0.2, 0.4, 0.6, 0.8 and 1.0 m. These were delineated with subscripts, (e.g., 0.2 m height; $EM38h_{0.2}$ and $EM38v_{0.2}$).

The EC_a data were collected from the head-ditch to the tail-ditch and at 10 m spacing. In total, 19 positions were visited (Figure 4.1b). The EM38 was positioned directly above and in alignment with the bed (east-west). A graduated wooden rod enabled EM38 to be positioned at correct height. On the day of irrigation, the EC_a were collected before irrigation commenced (0 hours). Later, six subsequent EC_a measurement passes were made *during irrigation (*1, 2.5, 4.5, 6, 7.5 and 9 hours*)*. Irrigation ceased after 9 hours. From here and every day after irrigation commenced (2-10 days), EC_a data was collected at 9:00 am. Altogether 17 EC_a data measurement campaigns were conducted.

4.2.4 Correction of EC_a data based on soil temperature and ambient temperature

To account for soil temperature variation from one day to the next and to correct EC_a data to a standard soil temperature (i.e., 25 C°), Corwin and Lesch (2005) empirically showed that this could be done by multiplying the measured EC_a data with a factor f_T using the following equation:

$$f_T = 0.4470 + 1.4034e^{-T/26.815} \quad [1]$$

where T is the soil temperature. Herein, the EC_a data associated with the different EM38 channels and heights of measurement were adjusted to a standard ambient temperature (i.e., 25 °C) using Equation (1). The input soil temperature (T), for different depths and times, were provided from those measured by Decagon GS3 sensors simultaneously with θ measurements.

4.2.5 Quasi-2D inversion of EM38 data

The EM4Soil software package (2018) was used to invert EM38 EC_a and generate electromagnetic conductivity images (EMCI); specifically estimates of σ. The software operates for EC_a collected at low induction numbers (Monteiro Santos et al. 2010), with an optimal EMCI arrived at by varying parameters, including; forward model, inversion algorithm and damping factor (λ).

A EMCI is formulated from minimization of the difference between measured and calculated EC_a generated from a forward model. In forward modelling, calculations consider the cumulative function (CF) model proposed by McNeill (1980) or non-linear full solution (FS) model based on Maxwell's equations (e.g., Wait, 1982; Frischknecht, 1988). At low induction numbers, the CF model is preferred due to their simplicity and speed. When EC_a is acquired over conductive soil (i.e., > 100 mS m^{-1}), the full-solution (FS) is preferable.

Inversion modelling is conducted using a one-dimensional laterally constrained approach (Auken et al. 2002), where two-dimensional smoothness constraints are imposed. *The inversion* algorithms are variations of Occam regularization. The S2 constrains EMCI variation around a reference model and consequently produces smoother inversions than S1. Also, *λ can be varied, which smooths data with increasing values.*

Herein, the aim was to determine the optimal EMCI by adjusting λ as a regularization parameter for the quasi-2D inversion of EM38 EC_a at the ground surface ($EM38h_0$ and $EM38v_0$) or at various heights (e.g., $EM38h_{0.2}$ and $EM38v_{0.2}$). The λ was set from 0.07, 0.3, and at 0.3 increments thereafter to a maximum of 6.0. Inversion of EC_a data collected on each day were generated with a maximum of 20 iterations. Owing to the large EC_a, the initial model was set of the σ to an average of σ = 100 mS m^{-1}.

To determine the most suitable parameters, the largest coefficient of determination (R^2) between σ and θ was identified. This had been done by obtaining σ from the EMCI on the first

survey and on each of the 16 subsequent surveys and correlating these with θ at the corresponding time.

4.2.6 Modelling methods

Five mathematical models were chosen to predict θ. The mathematical models were MLR, RF, Cubist, SVM and ANN. All the modelling in this study were performed using R software (Meyer et al. 2014).

4.2.6.1 MLR

An MLR model was used which is of the form:

$$y = a + b_1x_1 + b_2x_2 + \ldots + b_kx_k + \varepsilon \qquad [2]$$

where y is vector of the target property (i.e., θ) and where x_k represent the k_{th} predictor variable (i.e., EM38 EC_a data in the horizontal (EM38h) or vertical (EM38v) modes or estimates of σ from EM inversion of EC_a,) while a is the intercept, b the slope and ε is the model's residual.

To produce a model to predict θ at independent depths firstly the best MLR was determined which can be developed between EM38 EC_a data in the horizontal (EM38h) and vertical (EM38v) modes at various heights, including at the ground surface ($EM38h_0$ and $EM38v_0$). Also, MLR of EM38h and EM38v at various heights (0.2, 0.4, 0.6, 0.8 and 1.0 m) with the ground surface was compared. To find the best set of predictors, the fixed effects were determined prior to model building and by using backwards elimination until all predictors had a probability of being selected (<0.01).

4.2.6.2 RF

The RF is an ensemble of unpruned trees, whereby each tree is trained on a bootstrap sample of training data with prediction made by averaging their output (via regression). The algorithm for growing a RF of B trees is as follows:

(1) For b = 1 to B:

(a) Draw a bootstrap sample x_i of size N from the training data x (i.e., σ from inversion or EC_a).

(b) Grow a random-forest tree T_b to the bootstrapped data, by recursively repeating the following steps for each terminal node of the tree, until the minimum node size n_{min} is reached;

i. Randomly select m variables (i.e., σ and depth) from the p variables (i.e., σ or depth),

ii. Pick the best variable/split-point among the m, and

iii. Split the node into two daughter nodes.

(2) Make a prediction at a new location $\hat{y}$ (i.e., θ)

$$\hat{y}(x) = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \qquad [3]$$

The 'caret' package in R (Kuhn et al. 2017) was used to fit the RF model.

4.2.6.3 Cubist

The Cubist model is capable of first identifying which criteria are useful for partitioning the variables (i.e., σ and depth). Secondly, for each partition, it fits a MLR to predict θ. This approach is described as a decision tree (Adams and Sterling 1992), with the variables used for partitioning called nodes and the MLR under each partition called leaves. In addition, the Cubist software provides the attribute usage (relative importance) of each variable in each partition. The Cubist model was fitted using 'Cubist' package in R (Kuhn et al. 2014).

4.2.6.4 SVM

A SVM performs well on problems that are nonlinear, sparse and are high-dimensional. The SVM achieve this by using kernel functions to project the data onto a new hyperspace

where complex non-linear patterns can be simply represented (Gunn, 1998 and Williams, 2011). SVM predictions are calculated using the following function:

$$f(x) = \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) K(x_i, x_j) + d \qquad [4]$$

Where $f(x)$ is the target property (i.e., θ); $K(x_i, x_j)$ of a user-defined kernel function, where x is the feature vector (i.e., σ from inversion of EC_a) and d is a constant threshold. The kernel function used herein is an isotropic Gaussian radial basis function:

$$K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\gamma^2}} \qquad [5]$$

with the bandwidth parameter γ. The α_i^* and α_i are the weights (Lagrange multipliers), which determine the influence of each point in the model the weights α_i^* and α_i are obtained from the solution of a Quadratic Programming optimisation problem (Platt, 1999 and Smola and Schölkopf, 2004) of the form:

$$\max \alpha, \alpha^* - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(x_i, x_j)$$

$$-\varepsilon \sum_{i=1}^{N} (\alpha_i^* + \alpha_i) + \sum_{i=1}^{N} y_i (\alpha_i^* - \alpha_i) \qquad [6]$$

The constraints are given as follows:

$$\text{subject to} \begin{cases} \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) = 0 \\ 0 \le \alpha_i^*, \alpha_i \le C \; for \; i = 1, \ldots, N \end{cases} \qquad [7]$$

where ε is the specified non-negative contant; C is the regularization parameter that defines the trade-off between training error and model complexity (i.e., flatness of $f(x)$). The best parameters C and γ obtained using the training data were 5 and 0.1, respectively. In this work, the SVM was applied from the 'e1071' and 'caret' packages (Meyer et al. 2014) in R.

4.2.6.5 Artificial Neural Network (ANNs)

The artificial neural network (ANN) modelling technique was also chosen to fit the relationship between σ inverted from EC_a and θ. This is because it is a data-driven approach, which simulates features of neural networks that address a range of difficult information processing, analysis and modeling problems (De Smith et al. 2007). Neural networks achieve this by using hidden units between input and directs them through an activation function as the output. The units can be layers or nodes within the layers. This allows the hidden units to accomplish nonlinear transformations of the inputs entered in the network. Depending on their associated weights the hidden units may vary.

In this study, 3 hidden units were chosen (i.e., one layer with 3 nodes) because fewer units failed to reflect the complex inter-relationship and too many units led to over-fitting of the ANN model. The ANN was fitted using 'neuralnet' package in R (Günther and Fritsch 2010). In terms of validation, k-fold cross validation was adopted here. This method partitions the training set into *k* (i.e., 8) sets. The learner model trains k times for every model complexity, on every occasion it uses one of the sets as the validation set and the training data is the remaining sets. After calculating the average error over the *k* runs the model complexity with the smallest average error will be selected. It can return the model with that complexity, trained on all the data.

4.2.7 Validation and evaluating prediction accuracy, bias, and concordance

To compare various models, a leave-one-out cross validation was used. Specifically, a whole site was removed, including all depth and times. Then fitted the different models using the calibration dataset and applied to the point left out. Given there were 8 sites, this was done eight times. A number of indices, including coefficient of determination (R^2), root mean square error (RMSE) and mean error (ME), were considered, with the latter two a measure of

prediction accuracy and bias, respectively. In both cases, a smaller number, close to zero, is an indication of good performance. The Lin's concordance correlation coefficient is also used, herein as a measure of the agreement between measured and predicted θ. That is, how close predictions falls along a 45-degree line from the origin to the measured data (Lin, 1989).

4.3 Results and discussion

4.3.1 Distribution of soil properties along transect

Figure 4.2 shows contour plots of soil properties. *Figure 4.*2a shows clay was large, increasing from the head (< 58 %) toward the tail-ditch (> 67 %). *Figure 4.*2b shows silt was uniformly small (11-17%), with Figure 4.2c showing sand (18-27%) was slightly larger. The texture was uniformly heavy clay. Figure 4.2d shows EC_e was slightly saline (2.7-3.3 dS m^{-1}) in the topsoil (0-0.3 m), while in subsurface (0.3-0.6 m) and upper- (0.6-0.9 m) and lower-subsoil (1.2-1.5 m) it was non-saline (< 2 dS m^{-1}). Figure 4.2e shows topsoil and subsurface $pH_{1:5}$ was slightly alkaline (< 7.8) increasing to strongly alkaline (8.6-9). Figure 4.2f shows CEC was uniform and between 20 and 26 (cmol(+) kg^{-1}); suggesting the soil has moderate shrink-swell potential.

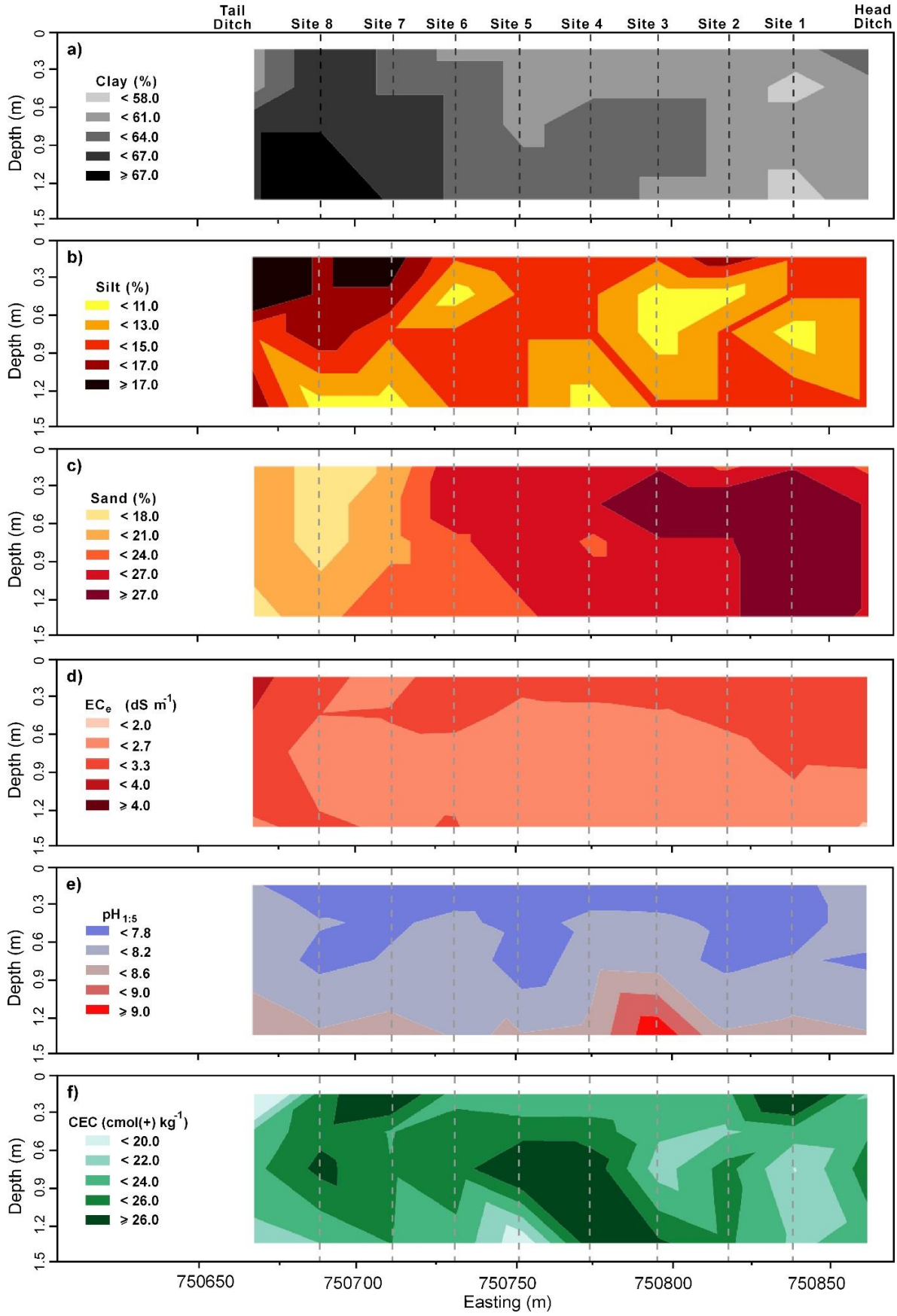


Figure 4.2 Contour plots of spatial distribution of measured soil properties at soil sampling sites including; a) clay (%), b) silt (%), c) sand (%), d) electrical conductivity of a saturated soil paste extract (EC_e, dS m^{-1}) estimated from 1 part soil 5 part water extract, e) $pH_{1:5}$ (pH of 1:5 soil:water extract) f) cation exchange capacity (CEC, cmol(+) kg^{-1}).

4.3.2 Summary statistics of EC_a

Table 4.1 shows summary statistics of EC_a (mS m^{-1}) from EM38 at various heights (i.e., 0, 0.2, 0.4, 0.6, 0.8 and 1.0 m). The $EM38v_0$ had the largest mean (103.0 mS m^{-1}), followed by the $EM38h_0$ (94.7). The mean EC_a decreased with increasing height. The EC_a was most skewed (-0.6) for $EM38h_0$, which also had largest (29 %) coefficient of variation (CV). With increasing height of EC_a measurement, CV and skewness decreased. These results can be attributed to the variable moisture content of topsoil, specifically, prior to irrigation, during irrigation and as the soil dried after irrigation ceased.

Table 4.1 Summary statistics of apparent electrical conductivity (EC_a, mS m^{-1}) measured using EM38 in a) horizontal (EM38h), and b) vertical (EM38v) modes and at various heights (0, 0.2, 0.4, 0.6, 0.8 and 1.0 m).

	EC_a (mS m^{-1})	Num.	Min	Mean	Medium	Max	Std. Dev.	CV	Skewness
a)	$EM38h_0$	323	32	94.7	99	153	27.5	29.0	-0.6
	$EM38h_{0.2}$	323	27	66.0	68	112	18.6	28.2	-0.2
	$EM38h_{0.4}$	323	20	47.8	49	78	12.0	25.0	0.0
	$EM38h_{0.6}$	323	17	36.0	37	58	8.6	24.0	-0.1
	$EM38h_{0.8}$	323	13	28.1	28	44	6.2	22.0	0.0
	$EM38h_{1.0}$	323	11	22.8	23	35	4.6	20.0	-0.1
b)	$EM38v_0$	323	45	103.0	104	162	23.7	23.0	0.0
	$EM38v_{0.2}$	323	40	89.1	90	145	21.5	24.1	0.1
	$EM38v_{0.4}$	323	37	73.2	74	117	16.5	22.5	0.3
	$EM38v_{0.6}$	323	31	59.5	60	94	12.5	21.0	0.3
	$EM38v_{0.8}$	323	27	48.9	49	75	9.5	19.4	0.3
	$EM38v_{1.0}$	323	23	40.8	41	62	7.4	18.2	0.3

4.3.3 Summary statistics of time-lapse θ

Table 4.2 shows summary statistics of mean Decagon GS3 measured θ (m^3 m^{-3}) for all sites at various depths (e.g., topsoil-0.15 m) and various times. The minimum (0.20 m^3 m^{-3}) topsoil θ, according to Gardiner and Miller (2004), would be close to permanent wilting point (0.25 m^3 m^{-3}) for heavy-clay (~60 %) textured soil. The suction that defines this value is –1500 kPa. This was expected given field had not been irrigated for two weeks. Nevertheless, the maximum (0.56 m^3 m^{-3}) subsoil θ was large and reflects θ beyond field capacity (0.42 m^3 m^{-}

^{3}) for heavy-clays (Gardiner and Miller 2004). The suction that defines this value varies from soil to soil but is generally around –10 kPa.

The large coefficient of variation (CV) of topsoil (27.2 %) and subsurface (14.41 %) θ reflects the variable nature of θ at various times before, during irrigation and days after irrigation commenced. The larger standard deviation (SD) of topsoil (0.10) and subsurface (0.07) θ, compared to the subsurface (0.06) and subsoil (0.06) shows the upper two layers were affected by the wetting and drying cycles. Despite the variation, soil θ was at or near field capacity for most of the experiment.

Table 4.2 Summary statistics of soil volumetric moisture (θ, m^3 m^{-3}) measured by the Decagon GS3 sensors at various depths: topsoil (0.15 m), subsurface (0.45 m), upper- (0.75 m) and deeper-subsoil (1.35 m).

Depth (m)	N	Min	Mean	Median	Max	Std. Dev.	CV%	Skewness
0.15	136	0.20	0.38	0.36	0.56	0.10	27.2	0.26
0.45	136	0.30	0.48	0.48	0.55	0.07	14.41	-1.01
0.75	136	0.32	0.49	0.50	0.56	0.06	12.47	-1.38
1.35	136	0.31	0.50	0.52	0.55	0.06	11.73	-2.04
All depths	544	0.20	0.46	0.50	0.56	0.09	19.51	-0.99

4.3.4 Distribution of measured EC_a along transect

Figure 4.3 shows the EM38 EC_a from ground surface (0 m) and at various times. Figure 4.3a shows $EM38h_0$ and $EM38v_0$. Before irrigation commenced and excluding the head-ditch and tail-ditch, $EM38h_0$ (~40 mS m^{-1}) was smaller than $EM38v_0$ (~ 60 mS m^{-1}). Figure 4.3b shows that the $EM38h_0$ (~75 mS m^{-1}) and $EM38v_0$ (~80 mS m^{-1}) increased near the head-ditch and as a function of water progressing down the field past site 1 and 1 hour after irrigation commenced. Fig 3c-g show a steady increased in EC_a between 2.5 and 9 hours when irrigation ceased.

Figure 4.3h-j show $EM38_{h0}$ (~75 mS m^{-1}) decreased and was equivalent to $EM38_{v0}$ (~80 mS m^{-1}) between days 2-4. Fig 3k-p shows that the $EM38h_0$ decreased much more rapidly than the $EM38v_0$ between days 5-9. This would be due to the soil drying as a function of the

evapotranspiration occurring in the topsoil, subsurface and upper subsoil; driven by the growing cotton crop. Fig 3q shows that the $EM38h_0$ was smaller than the $EM38v_0$ on day 10, but both measurements were not as small as day 0 and prior to irrigation commencing.

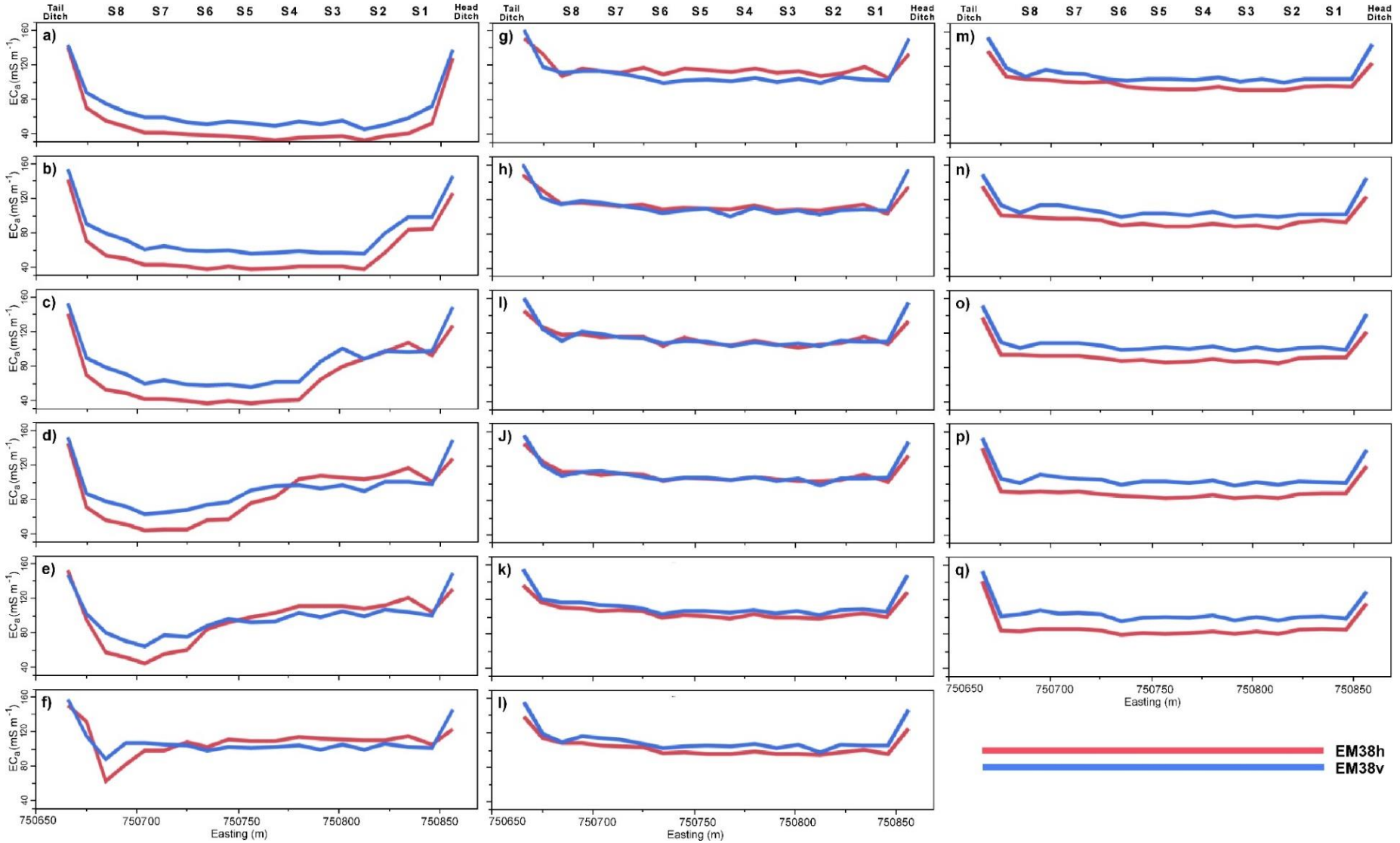


Figure 4.3 Distribution of measured apparent soil electrical conductivity (EC_a, mS m^{-1}) from a EM38 on the ground surface in horizontal ($EM38_{h0}$) and vertical ($EM38_{v0}$) mode a) 0 hours (before irrigation commenced), during irrigation at b) 1, c) 2.5, d) 4.5, e) 6, f) 7.5, g) 9 hours and after irrigation ceased on days h) 2, i) 3, j) 4, k) 5, l) 6, m) 7, n) 8, o) 9, q) 10.

4.3.5 Distribution of measured θ along transect

Figure 4.4 shows measured θ at various times at the 8 sites where the Decagon GS3 sensors were installed. Figure 4.4a shows that before irrigation commenced (0 hour), θ was at or near saturation (> 0.5 m^3 m^{-3}) only near the tail ditch (site 8) and in the upper- (0.75 m) and lower-subsoil (1.35 m). In the topsoil (0.15 m), θ was small (< 0.25 m^3 m^{-3}) at the head and tail ditch ends and close to permanent wilting point (0.25 m^3 m^{-3}) for a heavy-clay soil (Gardiner and Miller 2004). With increasing depth, θ increased (0.3-0.35 m^3 m^{-3}) and was wetter in the subsoil (0.3-0.35 m^3 m^{-3}).

Figure 4.4b shows a subsurface wetting front at site 1 and 1 hour after irrigation commenced. Figure 4.4c (2.5 hours) shows subsurface wetting front advanced to site 2. Figure 4.4d (4.5 hours) and Figure 4.4e (6 hours) show topsoil and subsurface wetting to site 4, with deeper subsoil wetting near site 5 and 6, respectively. Figure 4.4f (7.5 hours) and Figure 4.4g (9 hours) show wetting fronts at site 7 and 8, respectively. Figure 4.4h through to Figure 4.4q show measured θ after irrigation ceased. Figure 4.4h (day 2) shows topsoil was beginning to dry (0.45-0.5 m^3 m^{-3}), with Figure 4.4i (day 3) drying subsurface and subsoil (Figure 4.4j-q).

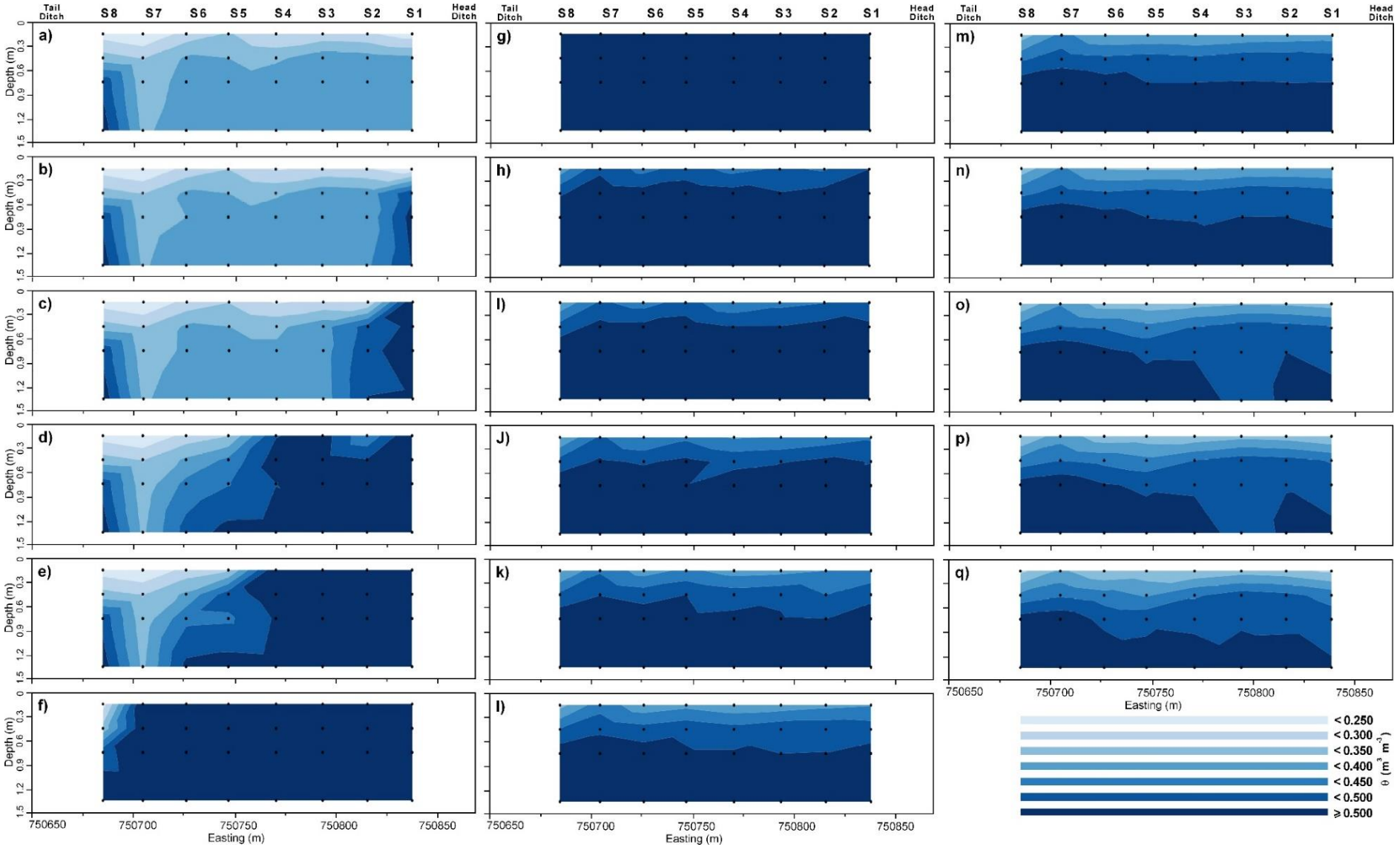


Figure 4.4 Distribution of measured soil volumetric water content (θ, m^3 m^{-3}) using Decagon GS3 sensors along transect, including; a) 0 hours (before irrigation commenced), during irrigation at b) 1, c) 2.5, d) 4.5, e) 6, f) 7.5, g) 9 hours and after irrigation ceased on days h) 2, i) 3, j) 4, k) 5, l) 6, m) 7, n) 8, o) 9, q) 10.

4.3.6 Coefficient of determination (R^2) between average EC_a and measured soil properties

Table 4.3 shows coefficients of determination (R^2) between average (0-1.5 m) soil properties and EM38 EC_a collected at ground level before irrigation. The R^2 between $EM38h_0$ was significant (*, $P < 0.05$*)* statistically for sand ($R^2 = 0.61$) and clay (0.55) but not silt (0.45). The R^2 between $EM38h_0$ and EC_e (0.07), $pH_{1:5}$ ($R^2 = 0.14$) and CEC (0.33) were small but larger for θ (*0.57*). The R^2 between $EM38v_0$ was largest for sand ($R^2 = 0.61$) and clay (0.57). The R^2 between $EM38h_0$ and θ was statistically significant and slightly larger (0.63) than sand.

Table 4.3 Coefficient of determination (R^2) between average (0-1.5 m) measured soil properties and bulk electrical conductivity (EC_a, mS m^{-1}) using EM38 at ground level before irrigation including in the horizontal ($EM38h_0$) and vertical ($EM38v_0$) modes of operation. Note: Soil salinity of saturated soil paste extract (EC_e) estimated from measured electrical conductivity from a 1 part soil to 5 parts water extract ($EC_{1:5}$); pH of 1:5 soil:water extract ($pH_{1:5}$); cation exchange capacity (CEC, cmol(+) kg^{-1}) and, volumetric water content (θ, m^3 m^{-3}).

Soil property	EM38 h_0	$EM38v_0$
Clay%	0.55*	0.57*
Silt%	0.45	0.49
Sand%	0.61*	0.61
EC_e	0.05	0.02
$pH_{1:5}$	0.05	0.06
CEC	0.33	0.32
θ	0.57*	0.63*
*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.0$		

4.3.7 Coefficient of determination (R^2) between EC_a and θ

Table 4.4a shows R^2 of four independent MLR relationships between measured θ and EM38h and EM38v. With respect to topsoil, the best R^2 was between θ and *$EM38h_0$ and $EM38v_0$* (0.76). This was also the case for subsurface (0.85) and deep subsoil (0.75) θ. With regards to shallow-subsoil θ, the R^2 was equivalent (0.85) when MLR was developed using both EM38h and EM38v at multiple heights of 0 and 0.2 m. There was no improvement in R^2 when the instrument was raised. Table 4.4b shows the MLR at one height and at ground surface. The results show R^2 were on par with when instrument was on the ground. Considering largest

R^2 were achieved with $EM38h_0$, and to a lesser extent $EM38v_0$, these two EC_a were the only two considered during inversion.

Table 4.4 Coefficient of determination (R^2) of multiple linear regression (MLR) between soil volumetric moisture (θ, $m^3\ m^{-3}$) measured by Decagon GS3 sensors at various depths and the measured apparent electrical conductivity (EC_a) in the horizontal (EM38h) and vertical (EM38v) modes, considering a) only one height and b) only one height and at the ground surface.

a)	Depth (m)	0 m	0.2 m	0.4 m	0.6 m	0.8 m	1.0 m
	0.15	0.76	0.56	0.56	0.45	0.35	0.39
	0.45	0.85	0.75	0.77	0.68	0.63	0.61
	0.75	0.85	0.85	0.84	0.82	0.80	0.77
	1.35	0.75	0.75	0.74	0.71	0.70	0.69
b)			0 & 0.2 m	0 & 0.4 m	0 & 0.6 m	0 & 0.8 m	0 & 1.0 m
	0.15		0.76	0.76	0.78	0.79	0.78
	0.45		0.84	0.84	0.84	0.85	0.85
	0.75		0.85	0.85	0.85	0.85	0.85
	1.35		0.76	0.75	0.75	0.75	0.75

Coefficient of determination (R^2) between estimated σ and measured θ

Figure 4.5 shows coefficient of determination (R^2) between estimates of true electrical conductivity (σ) and measured θ at all depths and survey times. *In estimating* σ, only $EM38h_0$ and $EM38v_0$ were considered along with using different sets of inversion parameters, including, forward model (i.e., CF and FS), inversion (i.e., S1 and S2) algorithm and different λ values.

Regardless of forward model or inversion algorithm, larger R^2 between σ and θ, were when λ was small (<1.2). The CF produced consistently larger R^2 compared to FS with S1 and relative to λ. This was attributed to EC_a approximating low induction numbers (~100 mS m^{-1}). *The largest R^2 (0.56) was between σ derived from* $EM38h_0$ and $EM38v_0$ *and θ using CF forward model, S1 inversion algorithm and λ = 0.6.* Given these results, calibration model development between σ and θ, was compared and contrasted *using estimates of σ derived from CF forward model, S1 inversion algorithms and λ = 0.6, using* RF, Cubist, SVM, ANN. Similarly, a MLR model between θ with σ and depth was developed.

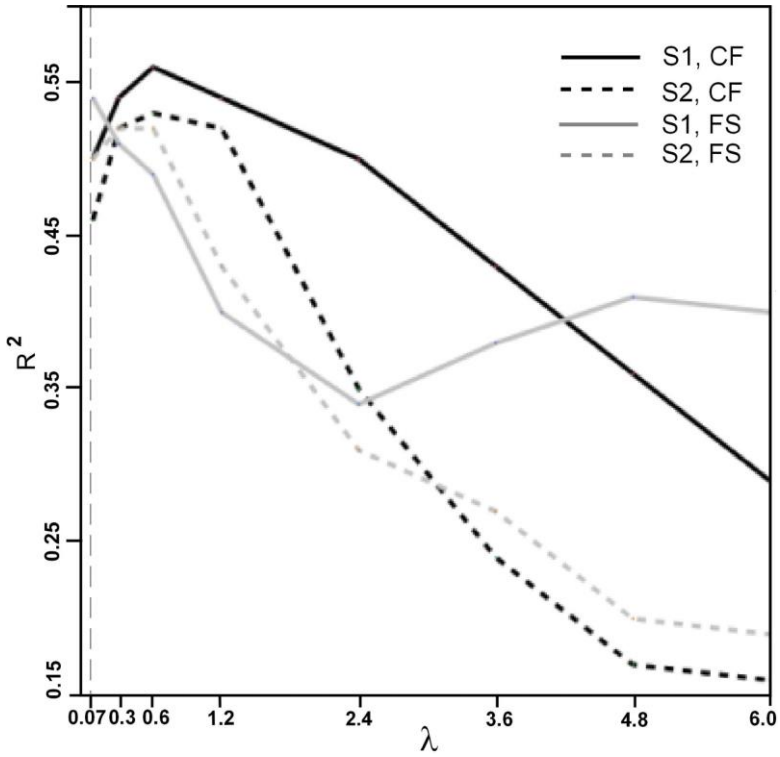


Figure 4.5 plot of coefficient of determinations (R^2) achieved between measured soil volumetric water content (θ, $m^3\ m^{-3}$) and estimates of true electrical conductivity (σ, mS m^{-1}) generated by inverting EM38 EC_a measured at ground level using EM4Soil quasi 2-d model, S1 and S2 inversion algorithm, cumulative function (CF) and full solution (FS) forward model, and varying damping factor (λ).

4.3.8 Coefficient of determination (R^2) between σ and measured soil properties

Table 4.5 shows the coefficient of determination (R^2) between measured soil properties at all depths and estimates of true electrical conductivity (σ, mS m^{-1}) estimated from EM38 EC_a collected at ground level in horizontal $EM38h_0$ and $EM38v_0$ and all time-steps. The R^2 between σ was small and not statistically significant for sand (0.03), silt (0.05) and clay (0.08) and σ with EC_e (0.07), $pH_{1:5}$ (R^2 = 0.14) and CEC (0.01). The R^2 for σ and θ (0.62) was statistically significant ($P < 0.001$). The results indicate while EC_a was influenced by average (0-1.5 m) clay, sand and θ, this was not the case for estimated σ, which was influenced by changes in θ.

Table 4.5 Coefficient of determination (R^2) between measured soil properties and estimates of true electrical conductivity (σ, mS m^{-1}) using EM38 EC_a at ground level including in the horizontal ($EM38_{h0}$) and vertical ($EM38_{v0}$) modes of operation before irrigation commenced and using the cumulative function (CF) modelling algorithm, inversion model S1 and a damping factor (λ) = 0.6. Note: Soil salinity of saturated soil paste extract (EC_e) estimated from measured electrical conductivity from a 1 part soil to 5 parts water extract ($EC_{1:5}$); pH of 1:5 soil:water extract ($pH_{1:5}$); cation exchange capacity (CEC, cmol(+) kg^{-1}) and, volumetric water content (θ, m^3 m^{-3}).

Soil property	R^2
Clay%	0.03
Silt%	0.05
Sand%	0.08
EC_e	0.07
$pH_{1:5}$	0.14**
CEC	0.01
θ	0.62***
*, P < 0.05; **, P < 0.01; ***, P < 0.0	

4.3.9 MLR between EC_a and θ

Table 4.6a shows the four different MLR equations used to predict θ from EC_a for different depths and considering only the $EM38h_0$ and $EM38v_0$. The intercepts (a) were uniform, albeit slightly larger in topsoil (0.273) and subsurface (0.278), compared to upper- (0.228) and lower-subsoil (0.233). The slope (b) was larger for $EM38h_0$ in topsoil (0.007) and the subsurface (0.003). The opposite was the case for $EM38v_0$. These results were consistent with theoretical depth of exploration. Specifically, $EM38h_0$ would be expected to relate to topsoil (0.15 m) and subsurface (0.45 m) θ, while $EM38v_0$ would be expected to be related to upper- (0.75 m) and lower-subsoil (1.35 m).

Table 4.6b shows the leave-one-out-cross validation of the use of the MLR equations when used to predict θ from EC_a for the four different depths and considering only the $EM38h_0$ and $EM38v_0$. The results show that for all depths, the prediction accuracy was small as was the bias. For example, topsoil prediction of θ was good, given RMSE (0.04) was less than half of the standard deviation (0.10) of measured θ. The ME (0) also indicated unbiased prediction. The leave-one-out cross validation show that the R^2 values were large (e.g., upper subsoil-

0.80). The 1:1 relationship was also large as exemplified by upper-subsoil Lin's correlation coefficient (0.88).

Table 4.6 a) Multiple linear regression (MLR) equations used to predict θ from the EM38 bulk soil apparent electrical conductivity (EC_a, mS m^{-1}) for different depths, and b) summary statistics of leave-one-site-out cross validation results of 4 MLR models generated using $EM38h_0$ and $EM38v_0$ EC_a for different depths. Note: Lin's Con., Lin's concordance correlation; R^2, coefficient of determination; RMSE, root mean square error; ME, mean error.

	Depth (m)	Equation			
a)	0.15 Topsoil	$0.273 + (0.007 \times EM38h_0) + (-0.005 \times EM38v_0)$			
	0.45 Subsurface	$0.278 + (0.003 \times EM38h_0) + (-0.001 \times EM38v_0)$			
	0.75 Upper subsoil	$0.228 + (0.001 \times EM38h_0) + (0.002 \times EM38v_0)$			
	1.35 Lower subsoil	$0.233 + (-0.0003 \times EM38h_0) + (0.003 \times EM38v_0)$			
		Lin's Con.	R^2	RMSE	ME
b)	0.15 Topsoil	0.82	0.71	0.04	0.00
	0.45 Subsurface	0.88	0.79	0.03	0.00
	0.75 Upper subsoil	0.88	0.80	0.02	0.00
	1.35 Lower subsoil	0.81	0.68	0.03	0.00

4.3.10 Modelling relationship between EC_a or σ and θ

Figure 4.6 shows measured versus predicted θ summary statistics of various models and considering a leave-one-out cross validation. Figure 4.6a shows four different MLR in Table 4.6a and using $EM38h_0$ and $EM38v_0$ and from the data shown in Table 4.6b. Overall, measured and predicted θ were in good accord (Lin's concordance = 0.90), with a strong correlation coefficient ($R^2 = 0.82$) with good accuracy (RMSE = 0.03) and unbiased predictions (ME = 0). Given RMSE was smaller than half the standard deviation (0.045) of measured θ, these predictions were satisfactory (Singh et al. 2005).

Figure 4.6b shows leave-one-out cross validation of RF using σ and depth. Overall, there was a good 1:1 fit, but Lin's concordance (0.81) and R^2 were smaller (0.67). Nevertheless, predictions of θ were unbiased (0) with RMSE (0.04) again less than half the standard deviation (0.045) of measured θ. Figure 4.6c shows Cubist using σ and depth was on parity with combined MLR using EC_a as indicated by Lin's concordance (0.89) and R^2 (0.79). Predictions

were also unbiased (0), however, RMSE (0.05) was a little larger than half the standard deviation (0.045) of measured θ.

What is evident in Figure 4.6b and 6c is that measured θ at large (0.45 m^3 m^{-3}) and intermediate (0.3-0.4 m^3 m^{-3}) values were under- and over-predicted. With respect to under-prediction, this was the case in the subsurface (0.45 m) and lower-subsoil (1.35 m). Over-prediction was evident in the lower-subsoil. This can be attributed to the fact that RF not performing well in time-series analysis, particularly when significant trends in increasing and decreasing values (Nau, 2015) exist. This was problematic herein, because all data had been removed at one point during leave-one-out cross validation. Since RF predicts an average value during calibration to the points it has not seen before (Nau, 2015), this causes under- and over-prediction. Additionally, the wetting front differs in the topsoil and subsoil and this causes the difficulty. With respect to Cubist, the problem arises because predictions generated using nearby points. Therefore, it performs poorly when all data at one point is removed during leave-one-out cross validation (Kuhn and Johnson 2013).

Figure 4.6d shows the results when the relationship between θ from σ and depth at all times was developed using SVM. The Lin's concordance (0.91) and the R^2 (0.83) were larger than any other model. Figure 4.6e shows the results when the ANN model was employed using σ and depth. The 1:1 fit (Lin's concordance = 0.86) for θ was smaller than MLR of EC_a. The R^2 was also smaller (0.76). The RMSE for the SVM (0.04) and ANN (0.04) were smaller than half the standard deviation (0.045) of measured θ at all depths.

In terms of overall ranking, the best approach was to invert the $EM38h_0$ and $EM38v_0$, with σ and depth modelled to θ using SVM. This was followed by MLR of $EM38h_0$ and $EM38v_0$ EC_a and θ. However, the latter approach does not allow for predictions to be made at depths where measurements in θ were not made; given depth specific nature of the calibrations. In this respect, and in-spite of not being as strong in terms of Lin's concordance and R^2, the

use of RF and ANN would not be preferred. This would also be the case for Cubist, which was superior to RF based on Lin's, but prediction accuracy (RMSE = 0.05) was larger than half the standard deviation (0.045) of measured θ.

Figure 4.6f shows if a simpler MLR model can be developed between θ using σ and depth. The summary statistics indicate that measured and predicted θ were in accord (Lin's concordance = 0.84) with a strong correlation coefficient (R^2 = 0.73). This approach was a slight improvement on RF but not as good as ANN. Nevertheless, accuracy (RMSE = 0.04) was good and smaller than half the standard deviation (0.045) of measured θ. In all modelling approaches the predictions were unbiased (ME = 0.00).

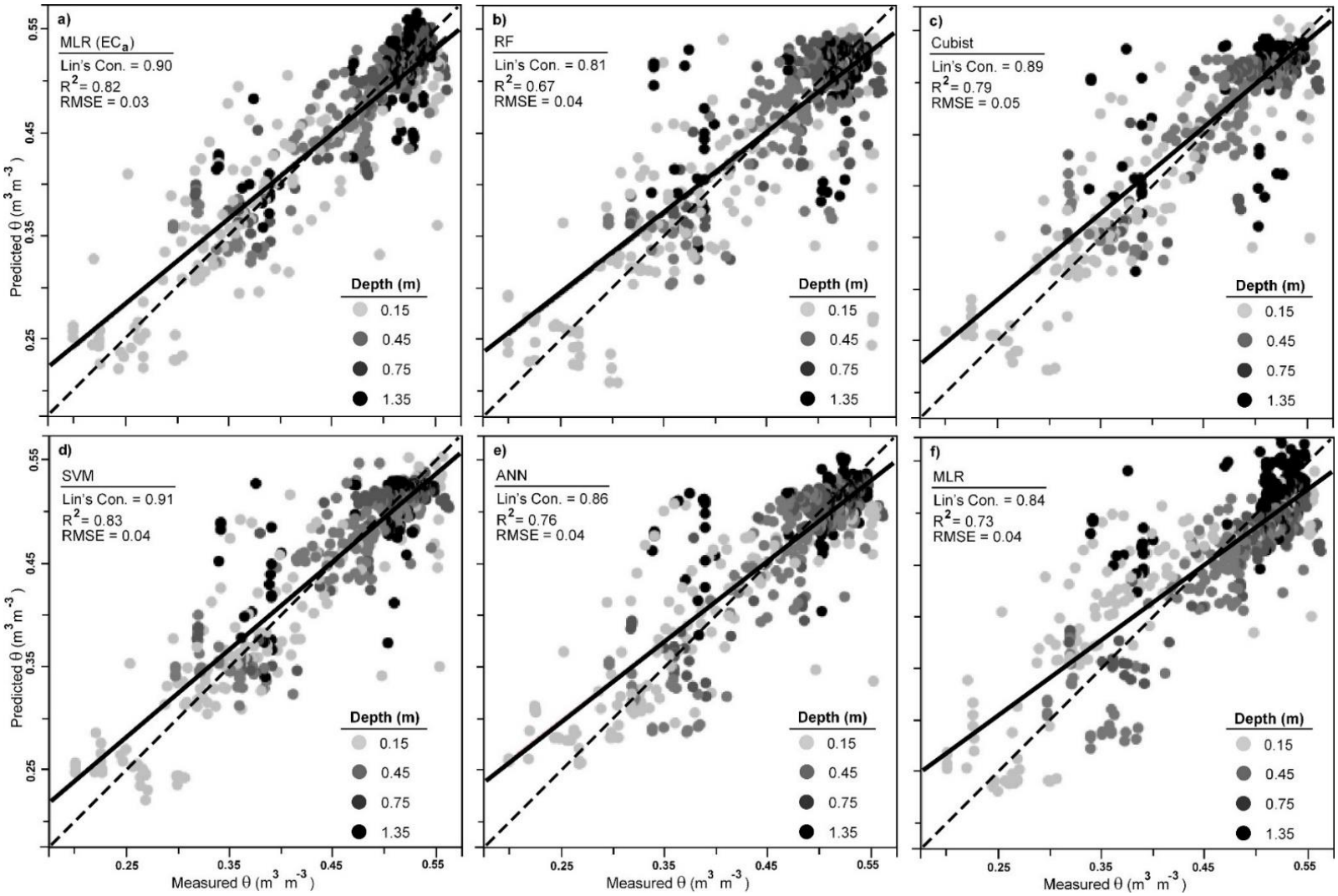


Figure 4.6 Plot of leave-one-out cross validation between measured soil volumetric water content (θ, m^3 m^{-3}) using a Decagon GS3 and predicted θ from a) multiple linear regression (MLR) of EM38 apparent electrical conductivity (EC_a , mS m^{-1}); and estimates of true electrical conductivity (σ, mS m^{-1}) and depth (m) generated by inverting EM38 EC_a using EM4Soil quasi 2-d model, cumulative function (CF), inversion algorithm S1 and damping factor (λ = 0.6), from model using b) random forest (RF), c) Cubist, d) support vector machine (SVM) e) Artificial Neural Network (ANN), and f) MLR. Note: leave-one-out cross validation of each site (all time-steps).

4.3.11 Modelling of θ using various modelling approaches

Figure 4.7a shows predicted θ during irrigation (6 hours) and using MLR of $EM38h_0$ and $EM38v_0$ EC_a. Fig 7b (day 5) and 7c (day 10) show predicted θ several days after irrigation commenced. The difference between plots of measured θ shown at 6 hours (Figure 4.4e), 5 (Figure 4.4l) and 10 (Figure 4.4q) days and predicted θ, was there were twice as many data points (i.e., 19) available, given EC_a was collected at 10 m intervals. There was an issue with short scale subsoil predicted θ (day 10) that was not reflected in measured θ.

Predicted θ using RF can be seen at 6 hours *(*Figure 4.7d), 5 (Figure 4.7e) and 10 (in Figure 4.7f) days. While the patterns were consistent with measured θ, under-prediction was evident in the subsurface (0.45 m) and lower-subsoil (1.35 m) behind wetting front. Over-prediction in advance of wetting-front was evident in lower-subsoil near the tail-ditch. Similar patterns were evident using Cubist (Figure 4.7g, h and i). The predicted θ for ANN (Figure 4.7m, n and o) and MLR (Figure 4.7p and q) of σ and depth were equivalent. However, the issue with these methods appears to be poor resolution of intermediate values of θ. This is particularly the case, in the subsurface and subsoil at 6 days and then resolving topsoil and subsurface drying; whereby the depth of drying seems to be over-predicted.

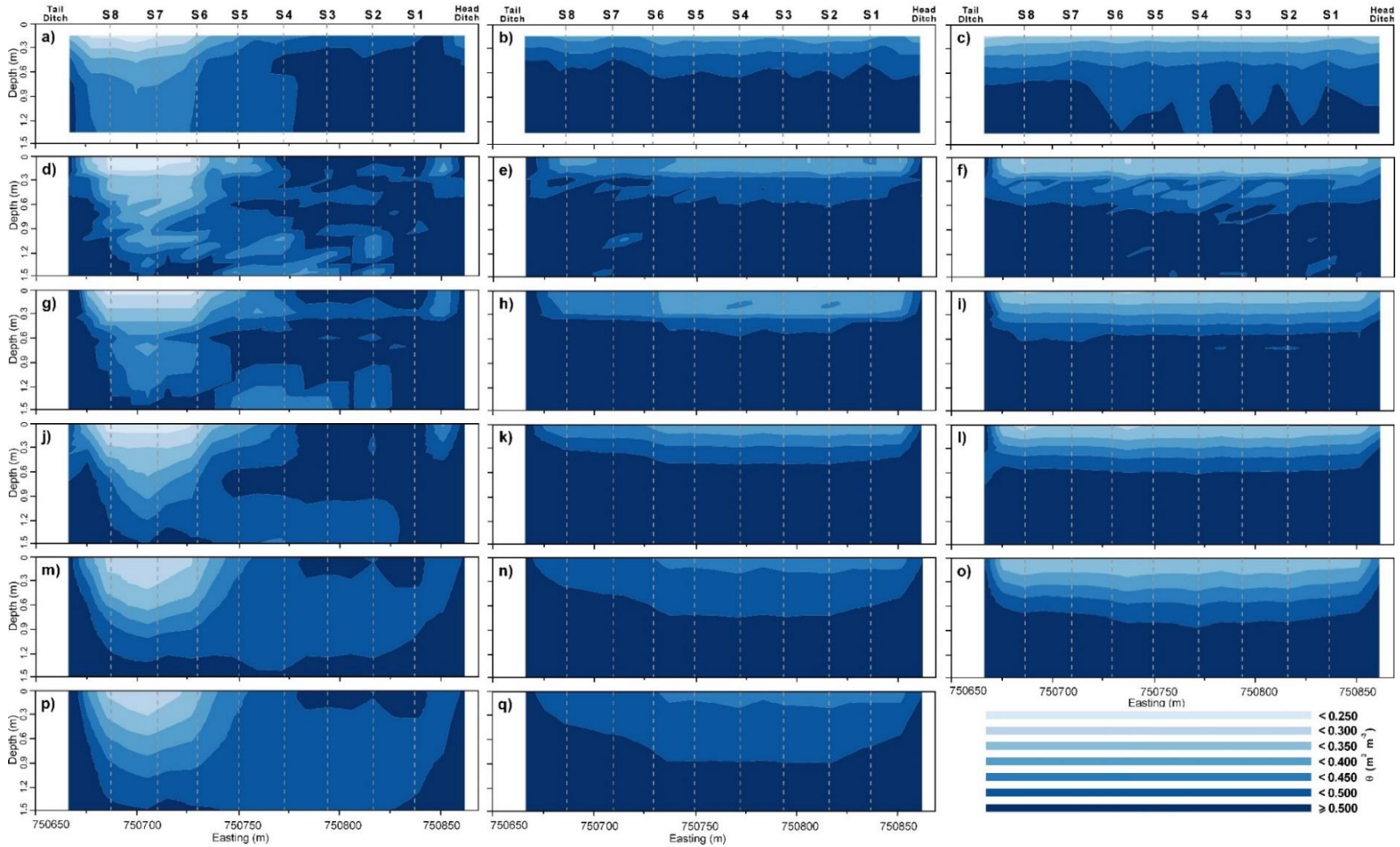


Figure 4.7 Contour plot of predicted soil volumetric water content (θ, $m^3\ m^{-3}$) along the transect at three time steps (6 hours and days 5 and 11) using multiple linear regression (MLR) of EM38 apparent electrical conductivity (EC_a, $mS\ m^{-1}$) and time steps of a) 6 hours, and days b) 5, and c) 11; and estimates of true electrical conductivity (σ, $mS\ m^{-1}$) and depth from model using: random forest (RF) and time steps of d) 6 hours, and days e) 5, and f) 11; Cubist and time steps of g) 6 hours, and days h) 5, and i) 11; support vector machine (SVM) and time steps of j) 6 hours, and days k) 5, and l) 11; Artificial Neural Network (ANN) and time steps of m) 6 hours, and days n) 5, and o) 11; and MLR and time steps of p) 6 hours, and day q) 5.

4.3.12 Modelling of θ using SVM

Figure 4.8 shows predicted θ prior to (Figure 4.8a), during (8b-g) and various days post-irrigation commencing (Figure 4.8h-q) using SVM. The advancing wetting front was well resolved and predicted. The better resolution was a function of the additional EC_a and also because EM4Soil considers neighboring sites. There was a slight issue in the subsoil prediction of θ and at 2.5 (Figure 4.8c) and 4.5 (Figure 4.8d) hours, behind the wetting front, however this appeared to be a function of the EM38 EC_a rather than modelling. Nevertheless, the modelling reflects the wetting and drying phases of the furrow irrigation system.

The results are also an advance on previous research carried out on a preliminary investigation by Zare et al. (2020), on this field. In that research, the Lin's concordance (0.84) was slightly smaller than that achieved herein using the SVM (0.91), with the most obvious reason being that in the previous research, the range in measured θ was smaller (0.21 m^{-3} m^{-3}) compared to that herein (0.36 m^{-3} m^{-3}). This was because of two reasons. In the first instance it was because EC_a was only collected during the wetting phase. Secondly, the θ prior to irrigation was closer to saturation in the subsoil and this made it difficult to resolve the rapidly changing topsoil θ.

To better model the subtle changes in θ between the topsoil, subsurface and subsoil, particularly given the fact that the subsoil wetting front advances ahead of the topsoil, the use of a DUALEM-1HS instrument would be appropriate. This is because it has four theoretical depths of measurement from the two coil pairs available. Specifically, 0.5mHcon (0-0.75 m) and 1mHcon (0-1.5 m) have equivalent depths of measurement as $EM38h_0$ (0-0.75 m) and $EM38v_0$ (0-1.5 m). However, the advantage of the DUALEM-1HS is the 0.5mPcon (0-0.3 m) and 1mPcon (0-0. 5 m) EC_a, which allows for better resolution of topsoil and subsurface σ and therefore prediction of θ in the upper- and lower-subsoil.

Using a light weight and relatively short DUALEM-1HS would enable the collection of EC_a data at more locations (i.e., 5 m spacing) than what that has undertaken here. The problem of the stickiness and poor trafficability of the Vertosols still needs to be resolved, however it was clear from this research that the collection of EC_a data at multiple heights was not advantageous or of any practical value to improving prediction, owing to the highly correlated nature of EC_a with measured θ, regardless of height. Moreover, the approach and the results achieved herein heralds its application to other irrigation systems, including travelling or lateral irrigation systems.

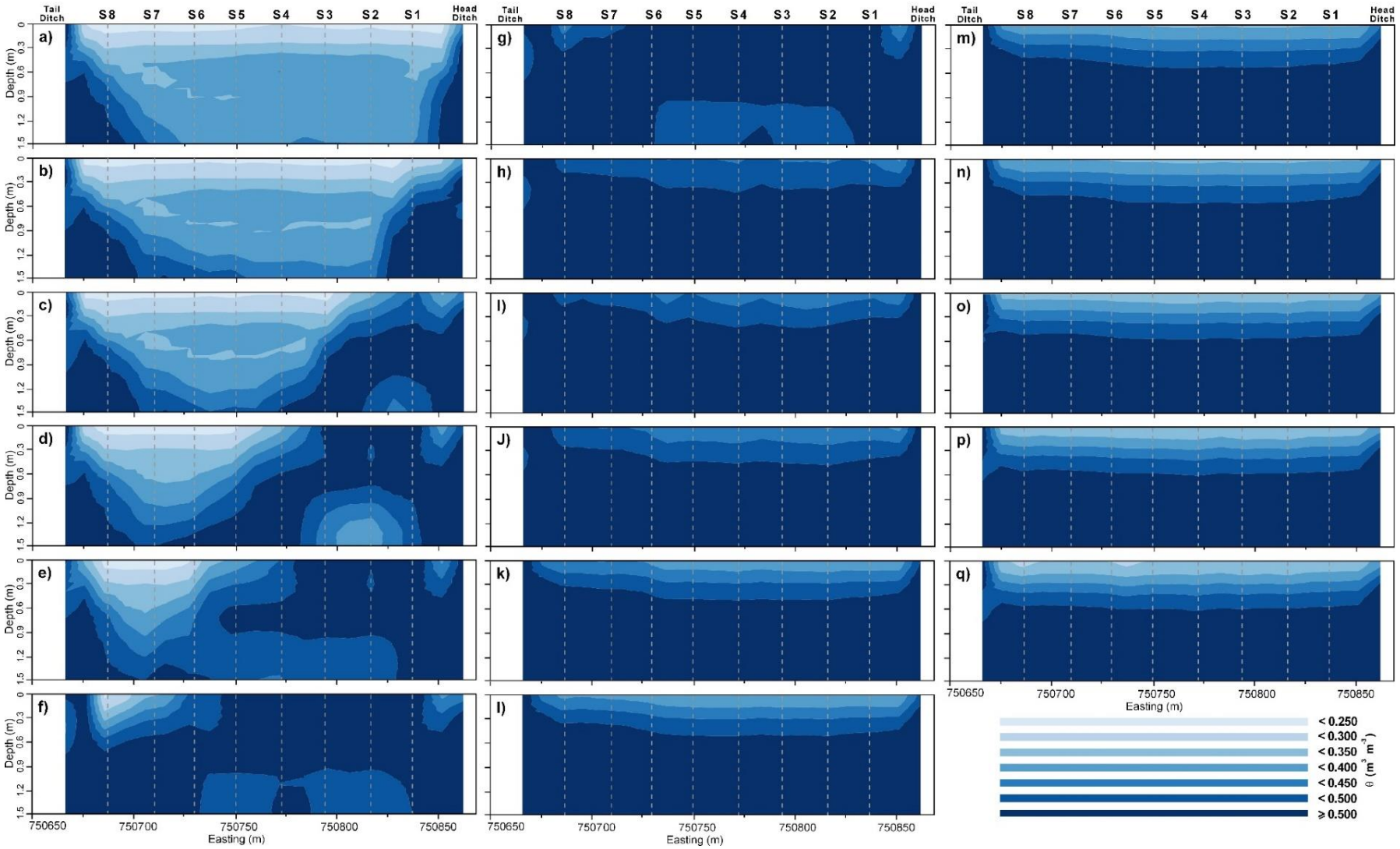


Figure 4.8 Contour plot of predicted soil volumetric water content (θ, $m^3\ m^{-3}$) along the transect generated support vector machine (SVM) and using estimates of true electrical conductivity (σ, $mS\ m^{-1}$) and depth at time steps of a) 0 hours (before irrigation commenced), during irrigation at b) 1, c) 2.5, d) 4.5, e) 6, f) 7.5 g) 9 hours and after irrigation ceased on days h) 2, i) 3, j) 4, k) 5, l) 6, m) 7, n) 8, o) 9, p) 10 and q) 11. Note: Irrigation ceased after 9 hours.

4.3.13 Modelling of θ using SVM and variable calibration sites

Table 4.7 shows the summary statistics of results for varying number of sites available to develop a calibration between θ with σ and depth using SVM. This had been based the comparison using varying numbers of diminishing calibration sites to predict θ at site 3. When 7 calibration sites were considered a good calibration (R^2 = 0.86) was achieved and a very well prediction of θ at the validation site (Lin's concordance = 0.88) attained. Equivalent results were achieved when the use of 6 and 5 sites were considered in terms of calibration (R^2 = 0.86), with the validation at site 3 producing strong 1:1 relationship between measured and predicted θ (Lin's concordance = 0.89 and 0.88, respectively). With decreasing calibration sites, the results actually improved slightly in terms of calibration. In fact, when only one site was considered for calibration (i.e., site 5) the largest correlation (R^2 = 0.93) was achieved. While the validation was the smallest, it was still very strong (Lin's concordance = 0.86).

Table 4.7 Summary statistics of drop-site-testing results of supported vector machine (SVM) models generated using σ and depth. Note: Lin's Con., Lin's concordance correlation; R^2, coefficient of determination; RMSE, root mean square error; ME, mean error.

ID of the calibration site(s)	Calibration set					Validation set (Site 3)				
	N	Lin's Con.	R^2	RMSE	ME	N	Lin's Con.	R^2	RMSE	ME
1, 2, 4, 5, 6, 7, 8	476	0.93	0.86	0.03	0.00	68	0.88	0.80	0.04	0.00
1, 2, 4, 5, 6, 7	408	0.92	0.86	0.03	0.00	68	0.89	0.81	0.04	0.00
2, 4, 5, 6, 7	340	0.92	0.86	0.03	0.00	68	0.88	0.79	0.04	0.00
2, 4, 5, 6	272	0.94	0.89	0.03	0.00	68	0.88	0.79	0.04	0.00
4, 5, 6	204	0.95	0.90	0.03	0.00	68	0.87	0.78	0.04	0.00
4, 5	136	0.93	0.89	0.03	0.00	68	0.87	0.79	0.04	0.00
5	68	0.95	0.93	0.02	0.00	68	0.86	0.77	0.04	0.00

4.4 Conclusions

The collection of EM38 EC_a data alone or its inversion to produce estimates of true electrical conductivity (σ, mS m^{-1}) provides useful information for soil moisture monitoring. This data can be calibrated against θ (m^3 m^{-3}) across a furrow-irrigated cotton field to predict and monitor θ during a wetting and drying cycle. In terms of irrigation management, the

approach has applications for understanding the spatio-temporal aspects of improving timing in irrigation and evaluation of efficiency. Another potential application is to use the approach to better understand the efficacy of lateral and bank-less channel irrigation systems, which are increasingly being deployed in Australian cotton growing areas.

Moreover, it can be recommended the spacing should be reduced to at least 5 m, to better account for the variation in θ and to capture the subsoil wetting front, which advances much more quickly down the furrow than the topsoil wetting front. This will enable less smoothing of predicted θ, because EM4Soil considers neighbouring sites during inversion. To better elucidate the topsoil (0.15 m) and subsurface (0.45 m) θ from subsoil θ (> 0.75 m) recommend a DUALEM-1HS instrument can be recommended. This is because more EC_a would be available from the 1m (*Pcon and Hcon*) coils, which are equivalent to those of the EM38 used herein, but also the 0.5 m (*Pcon and Hcon*) coils; because more EC_a would be available to better estimate σ in the topsoil and therefore predict θ.

The experiment also showed that satisfactory results could be achieved across the study transect from any number of calibration sites, however, owing to the uniform nature of the soil and the time-lapse nature of imaging this could be limited to a single calibration point. This has practical implications for irrigators and consultants in establishing a calibration between EM38 EC_a and θ. In addition and considering the results at day 10 when permanent wilting point was achieved in terms of θ, irrigation scheduling could be recommended based on the use of the $EM38h_0$ (80 mS m^{-1}) and $EM38v_0$ (100 mS m^{-1}) reached critical measurements. The actual inversion approach could also be used across the field to ascertain the uniform nature and efficacy of flood irrigation.

4.5 References

Adams, A., & Sterling, L. (Eds.). (1992). *Ai'92-Proceedings Of The 5th Australian Joint Conference On Artificial Intelligence*. World Scientific.

Al-Ain, F., Attar, J., Hussein, F., & Heng, L. K. (2009). Comparison of nuclear and capacitance-based soil water measuring techniques in salt-affected soils. *Soil use and management*, *25*(4), 362-367.

Auken, E., Foged, N., & Sørensen, K. I. (2002, September). Model recognition by 1-D laterally constrained inversion of resistivity data. In *8th EEGS-ES Meeting* (pp. cp-36). European Association of Geoscientists & Engineers.

Baumhardt, R. L., Lascano, R. J., & Evett, S. R. (2000). Soil material, temperature, and salinity effects on calibration of multisensor capacitance probes. *Soil Science Society of America Journal*, *64*(6), 1940-1946.

Brevik, E. C., & Fenton, T. E. (2004). The effect of changes in bulk density on soil electrical conductivity as measured with the Geonics EM-38. *Soil Horizons*, *45*(3), 96-102.

Chanasyk, D. S., & Naeth, M. A. (1996). Field measurement of soil moisture using neutron probes. *Canadian Journal of Soil Science*, *76*(3), 317-323.

Corwin, D. L., & Lesch, S. M. (2005). Apparent soil electrical conductivity measurements in agriculture. *Computers and electronics in agriculture*, *46*(1-3), 11-43.

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature climate change*, *3*(1), 52-58.

De Smith, M. J., Goodchild, M. F., & Longley, P. (2007). *Geospatial analysis: a comprehensive guide to principles, techniques and software tools*. Troubador publishing ltd.

Frischknecht, F. C. (1988). Electromagnetic physical scale modeling. Electromagnetic methods in applied geophysics-theory, 365-441.

Gardiner, D. T., & Miller, R. W. (2004). *Soils in our environment* (pp. 126-165). NJ.

Gunn, S. R. (1998). Support vector machines for classification and regression. *ISIS technical report*, *14*(1), 5-16.

Günther, F., & Fritsch, S. (2010). neuralnet: Training of neural networks. *The R journal*, *2*(1), 30-38.

Huang, J., McBratney, A. B., Minasny, B., & Triantafilis, J. (2017). 3D soil water nowcasting using electromagnetic conductivity imaging and the ensemble Kalman filter. *Journal of Hydrology*, *549*, 62-78.

Huang, J., Scudiero, E., Clary, W., Corwin, D. L., & Triantafilis, J. (2017). Time-lapse monitoring of soil water content using electromagnetic conductivity imaging. *Soil Use and Management*, *33*(2), 191-204.

Isbell, R. F. (2002). The Australian Soil Classification., Rev. edn (CSIRO Publishing: Melbourne).

JMP, S. (2012). Version 10.0. 2, SAS Institute Inc. *Cary, NC, USA*.

Kodikara, J., Rajeev, P., Chan, D., & Gallage, C. (2014). Soil moisture monitoring at the field scale using neutron probe. *Canadian Geotechnical Journal*, *51*(3), 332-345.

Kuhn, M., & Johnson, K. (2013). Regression trees and rule-based models. In *Applied predictive modeling* (pp. 173-220). Springer, New York, NY.

Kuhn, M., Weston, S., Keefer, C., Coulter, N., & Quinlan, R. (2014). Cubist: rule-and instance-based regression modeling. *R package version 0.0*, *13*.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., ... & Benesty, M. (2017). Caret: classification and regression training. 2016. *R package version*, *4*.

Lawrence, I., & Lin, K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 255-268.

Martinez, G., Huang, J., Vanderlinden, K., Giráldez, J. V., & Triantafilis, J. (2018). Potential to predict depth-specific soil–water content beneath an olive tree using electromagnetic conductivity imaging. *Soil Use and Management*, *34*(2), 236-248.

McNeill, J. D. (1990). Use of electromagnetic methods for groundwater studies. *Geotechnical and environmental geophysics*, *1*(5), 191-218.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C. C., & Lin, C. C. (2014). e1071: Misc functions of the Department of Statistics (e1071), TU Wien. *R package version*, *1*(3).

Moghadas, D. & Badorreck, A. (2019). Machine learning to estimate soil moisture from geophysical measurements of electrical conductivity. *Near Surface Geophysics* 17 (2), 181-195.

Moghadas, D., Jadoon, K. Z. & McCabe, M. F. (2019). Spatiotemporal monitoring of soil moisture from EMI data usingDCT-based Bayesian inference and neural network. *Journal of Applied Geophysics* 169, 226-238.

Monteiro Santos, F. A., Triantafilis, J., Bruzgulis, K. E., & Roe, J. A. E. (2010). Inversion of multiconfiguration electromagnetic (DUALEM-421) profiling data using a one-dimensional laterally constrained algorithm. *Vadose Zone Journal* 9, 117 - 125.

Nau, R. (2015). Statistical forecasting: notes on regression and time series analysis. *Notes and materials for an advanced elective course on statistical forecasting that is taught at the Fuqua School of Business, Duke University*.

Platt, J. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, *10*(3), 61-74.

Porwal, A., Carranza, E. J. M., & Hale, M. (2003). Artificial neural networks for mineral-potential mapping: a case study from Aravalli Province, Western India. *Natural resources research*, *12*(3), 155-171.

Roth, G., Harris, G., Gillies, M., Montgomery, J., & Wigginton, D. (2014). Water-use efficiency and productivity trends in Australian irrigated cotton: a review. *Crop and Pasture Science*, *64*(12), 1033-1048.

Silburn, D. M., Foley, J. L., Biggs, A. J. W., Montgomery, J., & Gunawardena, T. A. (2014). The Australian Cotton Industry and four decades of deep drainage research: a review. *Crop and Pasture Science*, *64*(12), 1049-1075.

Singh, J., Knapp, H. V., Arnold, J. G., & Demissie, M. (2005). Hydrological modeling of the Iroquois river watershed using HSPF and SWAT 1. *JAWRA Journal of the American Water Resources Association*, *41*(2), 343-360.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing*, *14*(3), 199-222.

Son, J. K., Shin, W. T., & Cho, J. Y. (2017). Laboratory and field assessment of the Decagon 5TE and GS3 sensors for estimating soil water content in saline-alkali reclaimed soils. *Communications in soil science and plant analysis*, *48*(19), 2268-2279.

Stannard, M. E., & Kelly, I. D. (1977). *The irrigation potential of the lower Namoi valley*. Water Resources Commission.

Tenison, K., Wild, J., Madden, E., Draper, V., McMahon, L. (2014). Salinity Training Manual,320 pp. State of New South Wales through the Department of Trade and Investment, Regional Infrastructure and Services, Sydney.

Triantafilis, J., Laslett, G. M., & McBratney, A. B. (2000). Calibrating an electromagnetic induction instrument to measure salinity in soil under irrigated cotton. *Soil Science Society of America Journal*, *64*(3).

Tucker, B. M. (1974). *Laboratory procedures for cation exchange measurements on soils* (No. 23 Tech. Pape).

Wait, J. (1982). Geo-Electromagnetism. Academic Press. New York.

Williams, G. (2011). Random forests. In *Data Mining with Rattle and R* (pp. 245-268). Springer, New York, NY.

Zare, E., Li, N., Arshad, M., Nachimuthu, G., & Triantafilis, J. (2020). Time-lapse imaging of soil moisture in a flood irrigation field monitored using electromagnetic conductivity imaging: Wetting phase. *Soil Science Society of America Journal*.

Chapter 5: Identifying potential leakage zones in an irrigation supply channel by mapping soil properties using electromagnetic induction, inversion modelling and a support vector machine.

5.1 Introduction

The Vertosols that characterise the extensive clay alluvial plains of the Namoi valley, New South Wales (NSW), make this a highly productive agricultural area. Up until the 1960's, their versatility enabled dryland wheat, sheep and beef cattle production (Triantafilis et al. 2001). The construction of Keepit Dam in the late 1960's provided water for irrigation and today, the Namoi valley has been extensively developed for irrigated cotton farming. In the lower Namoi this is mostly dependent on water supplied from Keepit Dam via the River. Moreover, a condition of an irrigation license is that the water needs to be stored on farm. This requires supply channels and water storages. Between Narrabri and Wee Waa, many supply channels were constructed and traverse the plain from south to north. However, the clay plains were developed upon alluvial sediments from prior stream channels, which traverse this area in a west to north-west direction (Stannard & Kelly, 1977). Because prior stream channels are characterised by more permeable Dermosols (Khongnawang et al. 2019; Triantafilis et al. 2013), water loss through leaking channels occurs where clay is < 40% (Triantafilis et al. 2004).

While prior stream formations occur mostly in elevated areas, where the uppermost materials are much coarser than the finer clay plains, in some places the prior stream channels are in low-lying positions relative to the plain. Here, they act as preferential paths for floodwaters (i.e., floodways). Moreover, the upper sediments are of a fine texture and mask the underlying coarser textures, which are susceptible to deep drainage (Woodforth et al. 2012). Locating these channels based on topography alone is problematic, as the alluvial plain is for the most part flat. Moreover, characterising the thickness and extent of the low-lying channels is important to ascertain where and how water loss from supply channels can be minimised and to improve water-use efficiencies. To identify areas of concern, soil samples need to be taken to determine the depth of the fine clay and lateral extent of the coarser textured sediments.

However, it is not all together clear, where and how many soil samples need to be taken. Therefore, techniques need to be developed for delineating the spatial variability of soil properties that directly influence the hydraulic characteristics of the soil beneath supply channels. This will allow appropriate management strategies to be recommended.

Physical properties, such as clay and sand content, directly influence hydraulic characteristics. In addition, chemical properties, such as cation exchange capacity (CEC cmol(+) kg^{-1}), can indicate ability of soil to shrink and swell. For example, CEC less than 20 cmol(+)kg^{-1} indicates poor shrink-swell, whereas CEC between 20 and 40 cmol(+)kg^{-1} suggests moderate shrink-swell potential. Moreover, the electrical conductivity of saturated soil paste extract (EC_e, dS m^{-1}) can indicate recharge areas as a function of non-saline (EC_e < 2 dS m^{-1}) and discharge areas in highly-saline (8-16 dS m^{-1}) locations. Unfortunately, extensive soil characterisation is challenging, owing to time and cost-prohibitive nature of sampling and analysis. Increasingly, electromagnetic (EM) induction instruments have proven to be a good source of digital data to create digital soil maps (DSM), because the measured apparent electrical conductivity (EC_a, mS m^{-1}) is related to clay (Muzzamal et al. 2018), mineralogy (Nagra et al. 2017) as well as CEC (Li et al. 2018) and salinity (Zare et al. 2015).

However, given depth specific information is required, newer approaches are essential to predict soil properties with depth. Most recently, Zhao et al. (2019) (Zhao et al. 2019) showed how inversion of Geonics (Geonics Ltd., Ontario, Canada) EM38 and EM34 EC_a using EM4Soil (EMTOMO, Lisbon, Portugal) can characterise clay at 1 m depth increments up to 12 m and across 40,000 ha in the nearby lower Gwydir valley. In another study Khognawang et al. (2019) used depth specific estimates of true electrical conductivity (σ, mS m^{-1}) inverted from EC_a collected from an EM38 in order to develop a LR model to predict CEC at the field level and at 0.3 m increments to 0.9 m. In this study, the aim was to characterise soil variation by depth along a 4km section of a farm supply channel by; i) collecting DUALEM-421 EC_a data, ii) using EC_a to inform soil sample site selection; iii) using EM4soil software package to

make EMCI from inversion of EC_a, iv) determine best set of inversion parameters (e.g., forward modelling, inversion algorithm and damping factor) to account for soil physical (clay and sand content) and chemical (EC_e and CEC) properties, and v) use a support vector machine (SVM) to predict the soil properties from σ and understand how 2-dimensional images can be used to identify more permeable prior stream channels and the implication for water management.

5.2 Materials and methods

5.2.1 Study site

The study area is located within a corporate farm (204 m above sea level; 30° 19′ 60″S, 149° 66′ 76″ E), which is in the lower Namoi valley of New South Wales, Australia (Figure 5.1a). The climate is semi-arid, with mean annual rainfall of 643 mm and potential evaporation over 1,800 mm. The rainfall is summer dominant with a mean maximum in January (79.8 mm) and minimum in August (37.5 mm). The farm was developed for irrigated cotton production in the 1960's. To supply water for irrigation, a large earthen channel was constructed. It extends from the Namoi River in the south, toward a large water reservoir some 10 km away to the north. Herein, the first 4 km of the channel was investigated. Figure 5.1b shows the location of the channel relative to the Kamilaroi Highway and a floodway, which cut across the southern and northern thirds of the channel, respectively.

In a reconnaissance soil survey of this part of the lower Namoi valley, Stannard and Kelly (1977) identified four physiographic units. Figure 5.1c shows the location of these physiographic units, including the clay plain, prior stream formation, coarse-textured dissected flood plain and Pilliga scrub complex. The clay plain is the largest and is characterised by uniform topography, where uppermost sediments are fine-textured, upon which a self-mulching clay soil has developed. In the Australian Soil Classification (Isbell, 2016) these fall under Sub-Order of Grey or Black Vertosols.

The prior stream formations occur mostly in continuous belts of slightly elevated and undulating land. Figure 5.1d shows that the uppermost materials, in prior stream channels in the nearby Gwydir valley, are coarser and can be as wide as 300 m (Pietsch, 2006). The prior stream channels may be distinguished from the levees, with the former also underlain by coarse channel sediments and gravels. In some areas, the channels are in low-lying areas relative to the clay plain, where they act as preferential paths for floodwaters (i.e., floodways); where the upper sediments are of a fine texture and identifiable from normal effluent by wide meander belts and broad and shallow channels. In places, migrational channel deposits (palaeochannels) underlie these sediments and are characterized by coarse sands and gravels (Stannard & Kelly, 1968).

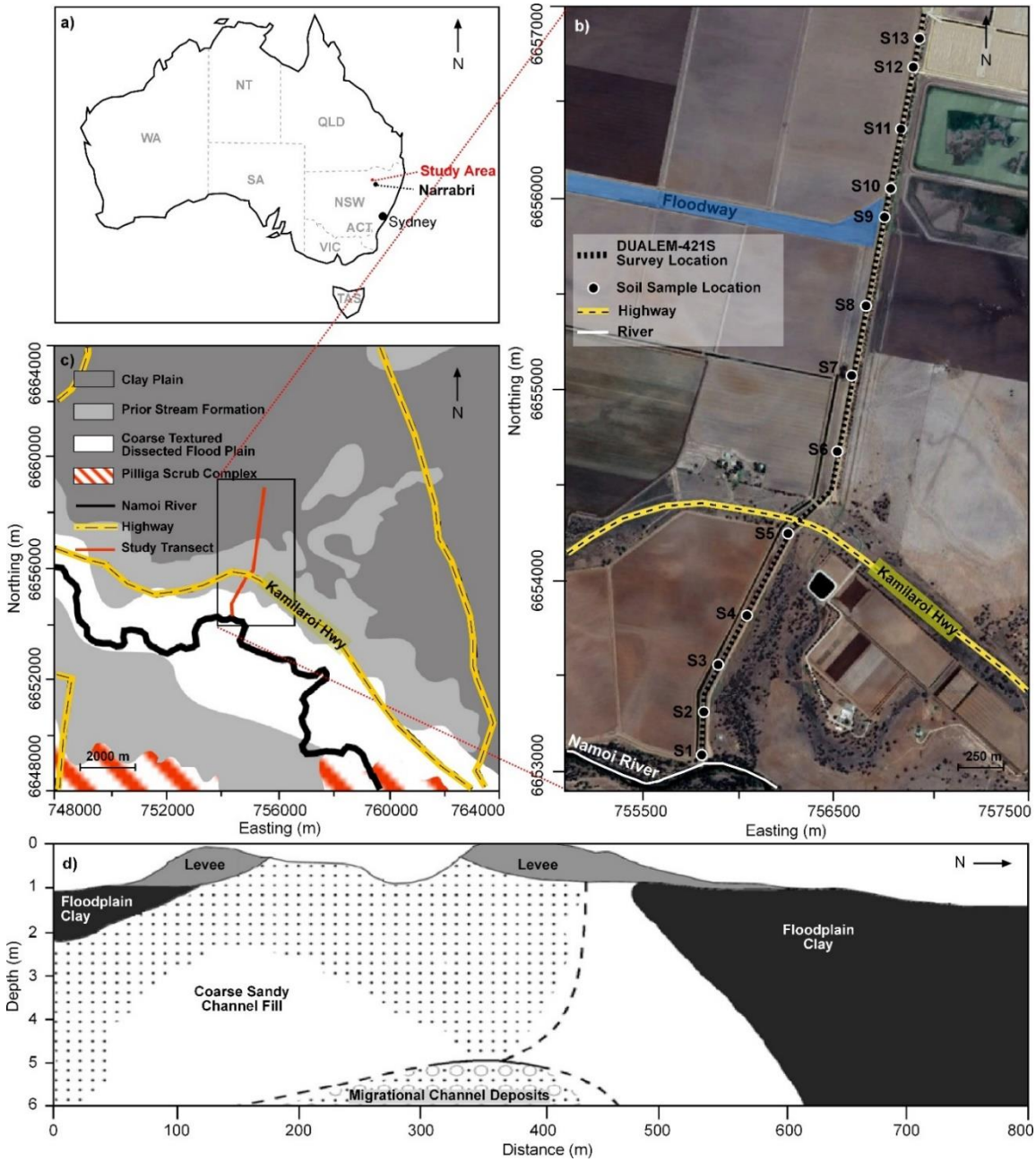


Figure 5.1 a) Location of the study area, b) air-photo (google earth) of the study supply channel either side of the Kamilaroi Highway and a local floodway with location of soil apparent electrical conductivity (EC_a, mS m^{-1}) survey transect and soil sample sites(13), c) physiography of the lower Namoi valley (after Stannard and Kelly, 1968), d) general shallow stratigraphy near prior stream channels, nearby Gwydir valley (after Pietsch, 2006).

5.2.2 DUALEM data collection

To measure soil EC_a (mS m^{-1}) a DUALEM-421 (DUALEM Inc., Milton, ON, Canada) was used. The DUALEM-421 operates at a low frequency (9 kHz) and includes a single transmitter and various horizontal co-planar (HCP) and perpendicular (PRP) receiver arrays. The transmitter is located at one end with HCP receivers 1, 2 and 4 m away. The instrument can measure EC_a (mS m^{-1}) at six depths of exploration (DOE). In HCP mode this includes

respectively, 0-1.5 m (1mHcon), 0-3.0 m (2mHcon) and 0-6.0 m (4mHcon). The distance from transmitter to PRP receivers is 1.1, 2.1 and 4.1 m, respectively and enable EC_a DOE of; 0-0.5 (1mPcon), 0-1.0 (2mPcon) and 0-2.0 m (4mPcon), respectively (Inc, 2008).

The EC_a from a DUALEM-421 was collected inside the channel (Figure 5.1b) starting from the northern end, with the last measurements logged adjacent to the Namoi River in the south where the pumps, pump the water into the channel. The survey was conducted on 15 October 2019. The instrument was mounted on plastic wheels, approximately 0.2 m above the ground and aligned in the orientation of the channel (north-south). The EC_a data were geo-referenced in latitude and longitude using a NovaTel SMARTV1 GPS (Novatel Inc., Hexagon Calgary, Alberta, Canada) antenna. The coordinates were later converted into Easting and Northing (UTM 84). Both EC_a and GPS data were collected using a Panasonic Toughbook CF-19 and logged using HGIS (v. 10.5) Software (StarPal Inc., Fort Collins, Colorado, USA).

5.2.3 Soil sampling and laboratory analysis

To evaluate the inversion modelling results soil samples were collected. Figure 5.1b shows the soil sample locations selected and spaced on average approximately 0.3 km apart along the transect. In total 13 sampling sites were selected and chosen to cover a wide range of EC_a. At each sampling site, the first sample was taken from the surface (0 m). From here soil samples were taken at 0.5 m increments to a maximum depth of 6 m. In total, 169 samples were collected. The samples were collected on 22 October 2019.

The samples were air-dried, ground and passed through a 2 mm sieve. The samples were oven dried (60 ^{o}C) for 48 hours. Laboratory analysis involved determination of soil physical properties, including soil particle size fractions (e.g., clay, silt and sand – %) and based on the hydrometer method (Beretta et al. 2014). Various chemical properties were also determined including the pH from a 1-part soil to 5-part water extract ($pH_{1:5}$). The soil electrical

conductivity was similarly determined from this extract ($EC_{1:5}$, mS m^{-1}). Subsequently EC_e (dS m^{-1}) was estimated from $EC_{1:5}$ using conversion factor for various soil textures, determined from the Australian Soil Texture Triangle (Moody & Phan, 2008), and as adopted from Tenison et al., (2014) as follows; clay ($EC_{1:5} \times 6.7$), clay loam (8.6) and sandy clay loam (9.5). Tucker's method (Tucker, 1974) was used to determine the CEC (cmol(+) kg^{-1}) using a mechanical leaching device. Briefly, samples were washed with 60% ethanol to remove any soluble salts, followed by extraction with 1M NH4Cl. The extracts were analyzed using an inductively coupled plasma optical emission spectrometry (ICP-OES). The exchangeable cations determine were summed to estimate the CEC.

5.2.4 Inversion modelling using EM4Soil

To invert the DUALEM EC_a, the EM4Soil software package (EM4Soil 2018) was used. The algorithm works for EC_a at low induction numbers (F. Monteiro Santos et al. 2010). Primarily, the software generates electromagnetic conductivity images (EMCI) and specifically estimates of σ by inverting EC_a obtained from EM instruments (F. A. Monteiro Santos et al. 2010). The estimates of σ are constrained by neighbouring sites, with an optimal EMCI arrived at by comparing results from several input parameters. To ascertain the best model, the following parameters was considered:

i) Forward modelling - forward calculations and derivatives consider either the cumulative function (CF) or a full-solution (FS). The latter is preferred when EC_a is acquired over conductive terrain (i.e., > 100 mS m^{-1}) and accounts for non-linearity. A 1-dimensional laterally constrained approach is used (Auken et al. 2002), where 2-dimensional smoothness constraints are imposed;

ii) Inversion algorithms: The inversion algorithms include S1 and S2. The latter constrains EMCI variation around a reference model and is generally smoother than S1. Both are variations of Occam regularization; and,

iii) Damping factor (λ) and iterations; as λ increases it tends to smooth the EMCI. Herein λ = 0.07, 0.3, and at 0.3 increments thereafter to a maximum of 3.0 was considered. A similar effect occurs when choosing the number of iterations. 10 iterations were selected.

The aim was to determine the optimal combination of inversion parameters (i.e., λ, S1 or S2, CF or FS) for the quasi-2D inversion of DUALEM-421 EC_a. To determine the most suitable parameters, the largest coefficient of determination (R^2) between σ and the soil physical properties (i.e., clay or sand) and chemical properties (i.e., EC_e and CEC) were identified. This had been done by obtaining σ from the EMCI and correlating these with the measured soil properties at the same depths σ was estimated.

5.2.5 Establishing a calibration model between σ and various soil properties

To develop a mathematical model between estimates of σ and various soil physical and chemical properties a support vector machine (SVM) model was chosen. For the reason that it was anticipated the EM instruments response in the soil would be a function, potentially, of four separate soil properties (i.e., clay content and mineralogy, salinity and moisture) and the relationship of each of these and σ would also be non-linear. In addition, a limited number of soil cores was available.

In these circumstances, a SVM can perform well on problems that are nonlinear, sparse and are highly-dimensional. This is because SVM uses kernel functions to project the data onto a new hyperspace where complex non-linear patterns can be simply represented (Gunn, 1998 and Williams, 2011). SVM predictions are calculated using the following function:

$$f(x) = \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) K(x_i, x_j) + c$$

where $f(x)$ is the target soil physical (e.g., clay) or chemical (e.g., EC_e) property; α_i^* and α_i are the weights (Lagrange multipliers), $K(x_i, x_j)$ is a user-defined kernel function (i.e., linear, polynomial, radial and sigmoid), and an error parameter c which is calculated from the best model, where x is the feature vector (i.e., σ from inversion).

The kernel function used herein is an isotropic Gaussian radial basis function:

$$K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\gamma^2}}$$

with the bandwidth parameter γ. The α_i^* and α_i are weights (Lagrange multipliers), which determine the influence of each point in the model and are obtained from the solution of a Quadratic Programming optimisation problem (Platt, 1999) , (Smola & Schölkopf, 2004) of the form:

$$\max \alpha, \alpha^* - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(x_i, x_j)$$

$$-\varepsilon \sum_{i=1}^{N} (\alpha_i^* + \alpha_i) + \sum_{i=1}^{N} y_i (\alpha_i^* - \alpha_i)$$

The constraints are given as follows:

$$\text{subject to } \begin{cases} \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) = 0 \\ 0 \leq \alpha_i^*, \alpha_i \leq C \text{ for } i = 1, \ldots, N \end{cases}$$

where ε is the specified non-negative contant; C is the regularization parameter that defines the trade-off between training error and model complexity (i.e., flatness of f(x)). The best parameters C and γ obtained using the training data were 5 and 0.1, respectively. In this work, the SVM from the 'e1071' and 'caret' packages (Meyer et al. 2014) in R was used.

In order to improve the performance of the SVM, the best parameters for the model with different couples of epsilons (ε) and cost parameter using the tune function in R was selected. In this study, the models were trained with ε = 0, 0.1, 0.2, ...,1 and cost = 2^1, 2^2,2^3, ..., 2^{100} which means it will train 1,100 models. Finally, the SVM with the best model values (ε = 0.3 and cost = 4) was successfully tuned.

5.2.6 Validation and evaluating prediction accuracy, bias, and concordance

Herein, a leave-one-site-out cross validation is used to validate the prediction of various soil physical and chemical properties. This was done owing to a limited number of sample sites. Specifically, a whole site was removed, including all depths. A model was fitted using the remaining calibration dataset of 12 sites and applied to the site removed. This was done 13 times.

Several indices, including root mean square error (RMSE) and mean error (ME), were considered, with the former a measure of prediction accuracy and the later bias, respectively. The RMSE is a frequently used measure of prediction accuracy and is calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_i - \hat{y}_i)^2}$$

and describes the difference between measured and predicted data. The smaller a RMSE value, the greater prediction accuracy. Moreover, if RMSE is less than half the standard deviation (SD) of the measured data, the prediction model is considered to produce satisfactory results (Singh et al. 2005).

The mean error (ME) is calculated as follows:

$$\mathrm{ME} = \frac{1}{n}\sum_{t=1}^{n}(y_i - \hat{y}_i)$$

For an unbiased prediction, their values should be near zero.

The Lin's concordance correlation coefficient (LCCC) is also used, herein, as a measure of the agreement between measured and predicted soil properties. That is, how close predictions fall along a 45-degree line from the origin to the measured data (Lin, 1989). It is a determinant of both accuracy and precision of the predictions and is calculated using following equation:

$$\mathrm{LCCC} = \frac{2\rho\sigma_y\sigma_{\hat{y}}}{\sigma_y{}^2 + \sigma_{\hat{y}}{}^2 + (\mu_y - \mu_{\hat{y}})^2}$$

where y and $\hat{y}$ are the measured and predicted soil physical or chemical property; ρ is the correlation coefficient (Pearson's r) between measured and predicted data; σ_y and $\sigma_{\hat{y}}$ are corresponding variances and μ_y and $\mu_{\hat{y}}$ are means for measured and predicted data. The agreement can then be considered in terms described by McBride (2005) (McBride et al. 2005) of; almost perfect (> 0.9), substantial (0.8-0.9), moderate (0.65-0.8) and poor (< 0.65).

Table 5.1 Summary statistics of apparent electrical conductivity (EC_a, mS m^{-1}) measured by a DUALEM-421 for the entire survey transect and at the 13 calibration sites.

DUALEM-421 Survey data	N	Min	Mean	Max	Median	CV (%)	SD	Skewness
1mHcon	2959	21	62	107	57	28	18	0.7
1mPcon	2959	11	32	57	31	31	10	0.4
2mHcon	2959	44	91	161	85	28	26	0.6
2mPcon	2959	24	61	111	56	28	17	0.7
4mHcon	2959	48	95	172	90	30	29	0.5
4mPcon	2959	40	89	161	83	30	26	0.6
Calibration data								
1mHcon	13	23	63	106	49	47	29	0.3
1mPcon	13	16	32	56	26	50	16	0.4
2mHcon	13	48	97	161	92	41	40	0.3
2mPcon	13	30	63	110	47	47	29	0.4
4mHcon	13	53	103	171	115	38	40	0.1
4mPcon	13	43	95	161	89	44	42	0.3

5.3 Results and discussion

5.3.1 Spatial distribution of EC_a

Table 5.1 shows summary statistics of EC_a collected along the channel. The mean of the 1mHcon (62 mS m^{-1}) was approximately twice as large as 1mPcon (32 mS m^{-1}). It also shows that while the difference between the mean of 2mHcon (91 mS m^{-1}) and 2mPcon (61 mS m^{-1}) was large, the mean of the latter was like that of 1mHcon. The mean of 4mHcon (95 mS m^{-1}) and 4mPcon (89 mS m^{-1}) were equivalent to 2mHcon. Table 5.2 shows summary statistics of EC_a at 13 calibration sites were equivalent and therefore were a good representation and valid to use for calibration/validation.

Figure 5.2a shows measured DUALEM-421 EC_a of 1mHcon (100 mS m^{-1}) and 1mPcon (50 mS m^{-1}) were largest within 600 m of the southern (sites 2 and 3) and northern (site 11) end. Conversely, 1mHcon (20 mS m^{-1}) and 1mPcon (10 mS m^{-1}) was smallest closest to the Namoi River (site 1) at the southern end and near the floodway (site 10). Figure 5.2b shows equivalent patterns for 2mHcon and 2mPcon. Whereas 2mHcon was generally larger along the entire transect than either 1mHcon or 1mPcon, 2mPcon was equivalent to 1mHcon. As indicated in Table 5.1, Figure 5.2c shows 4mHcon and 4mPcon were essentially equivalent to each other and 2mHcon.

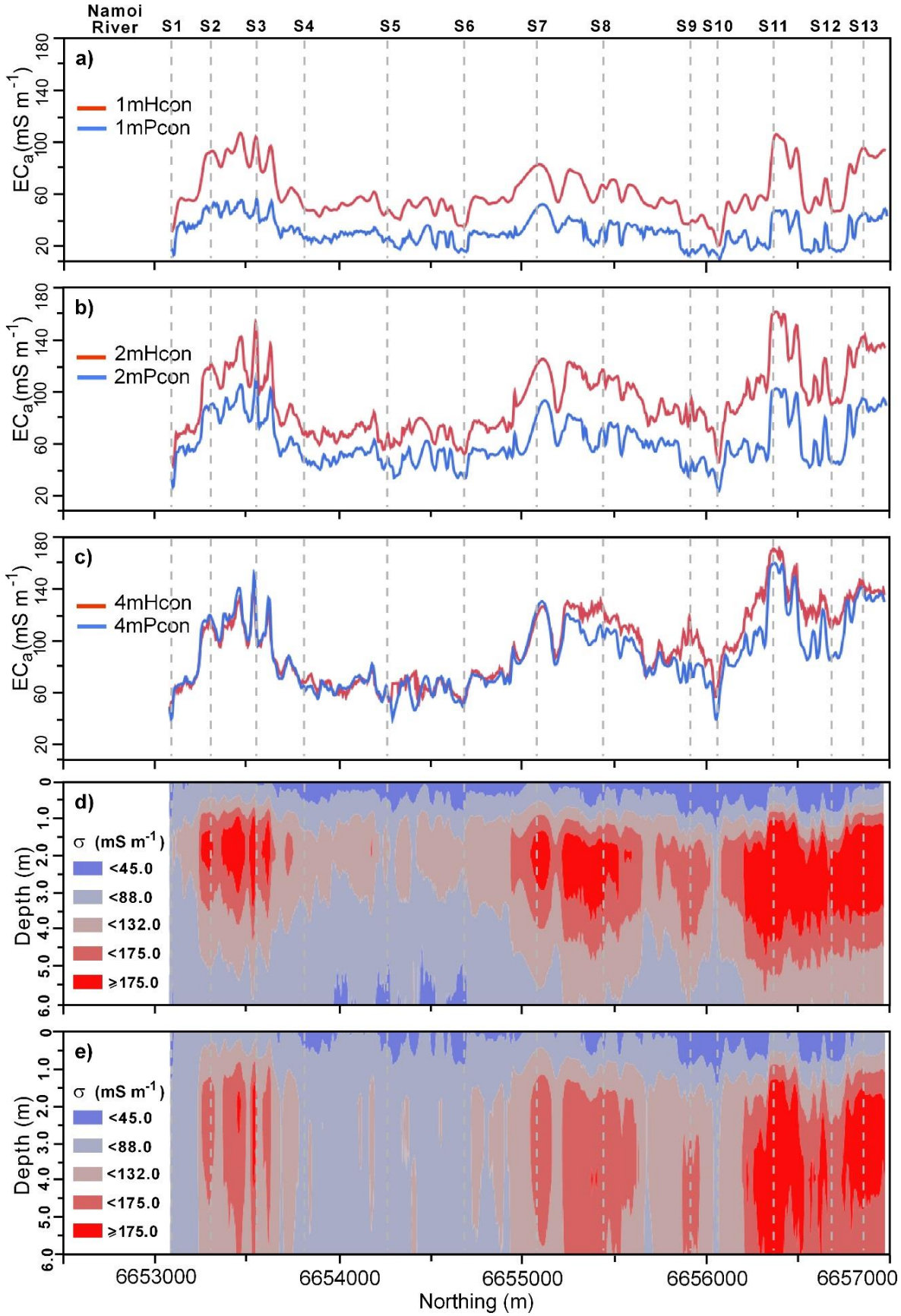


Figure 5.2 Measured soil apparent electrical conductivity (EC_a, mS m^{-1}) from DUALEM-421 in horizontal (Hcon) and perpendicular (Pcon) coplanar arrays including; a) 1mHcon and 1mPcon, b) 2mHcon and 2mPcon, c) 4mHcon and 4mPcon with electromagnetic conductivity images (EMCI) of estimates of true electrical conductivity (σ, mS m^{-1}) considering inversion of all DUALEM-421 EC_a and using cumulative function forward model (CF), S2 inversion algorithm and λ = d) 0.3 and e) 1.8.

Preliminary data analysis of soil physical and chemical properties

Table 5.2 shows the summary statistics of measured soil physical properties at the 13 soil sampling sites. The clay ranged from small (4 %) to large (77.3 %) with a mean clay (44.2 %) akin to a soil texture approximating a medium clay. The sand was similarly variable with minimum (5.9 %) and maximum (92.4 %) showing a larger range with a smaller mean (35.54 %) than clay. The standard deviation (SD) for clay (14.1 %) and sand (18.3 %) were comparable.

Table 5.2 also shows summary statistics of chemical properties. The EC_e minimum (0.1 dS m^{-1}) and mean (1.7 dS m^{-1}) were non-saline (< 2 dS m^{-1}) with the maximum (3.99 dS m^{-1}) slightly saline (2-4 dS m^{-1}). The CEC ranged between small (1.31 cmol(+) kg^{-1}) and large (47.7 cmol(+) kg^{-1}) that represent good shrink-swell behaviour, and on average the soil has a mean CEC (24.4 cmol(+) kg^{-1}), which indicates channel sediments have shrink swell behaviour. The SD of EC_e (0.9 dS m^{-1}) and CEC (8.43 cmol(+) kg^{-1}) were small. The $pH_{1:5}$ ranged between strongly acidic (4.5) and very strongly alkaline (9.65) with a mean of moderate alkalinity (8.5). Except for $pH_{1:5}$ (-1.16) the data were not skewed.

Table 5.2 Summary statistics of measured soil properties. Note: Soil salinity of saturated soil paste extract (EC_e, dS m^{-1}) estimated from measured electrical conductivity from a 1 part soil to 5 parts water extract ($EC_{1:5}$); cation exchange capacity (CEC, cmol(+) kg^{-1}); pH of a 1 part soil to 5 parts water extract ($pH_{1:5}$).

Soil Propeties	N	Min	Mean	Max	Median	CV (%)	SD	Skewness
Clay	169	4.00	44.18	77.33	42.33	31.99	14.13	0.02
Silt	169	0.73	20.29	41.23	19.77	38.77	7.88	0.19
Sand	169	5.90	35.54	92.40	31.87	51.51	18.31	0.63
EC_e	169	0.08	1.73	3.99	1.64	52.00	0.90	0.36
CEC	169	1.31	24.41	47.67	23.32	34.53	8.43	0.33
$pH_{1:5}$	169	4.47	8.30	9.65	8.4	9.15	0.76	-1.16

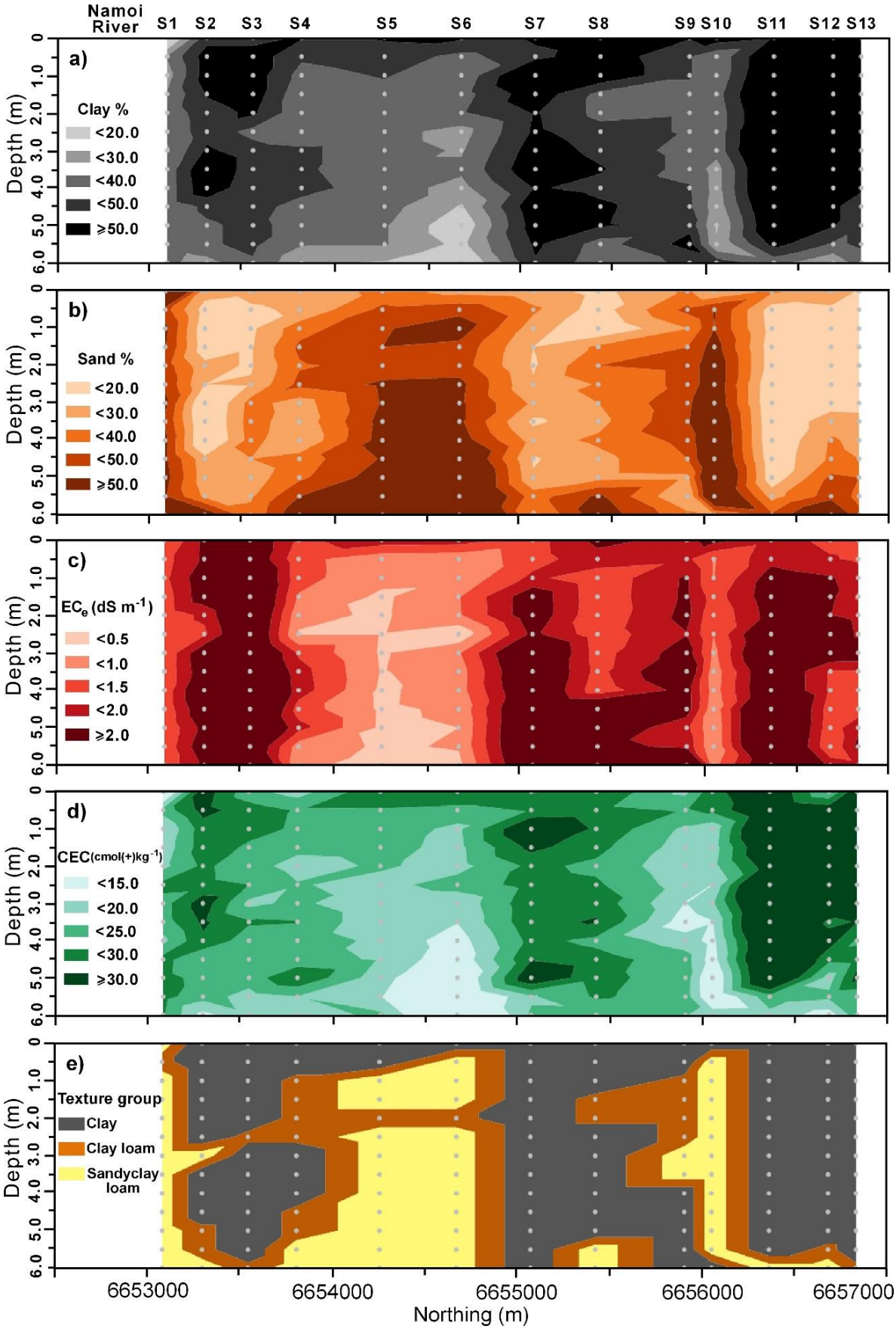


Figure 5.3 Contour plots of spatial distribution of soil physical and chemical properties at soil sampling sites (13) along the channel, including; a) clay (%), b) sand (%), c) electrical conductivity of a saturated soil paste extract (EC_e, dS m^{-1}) estimated from 1 part soil 5 part water extract, d) cation exchange capacity (CEC, cmol(+) kg^{-1}) and e) contour plot of soil texture at the soil sampling sites (×13).

Figure 5.3 shows contour plots, generated using JMP Software 13.2 (SAS Institute, Cary, NC, USA), of measured soil physical and chemical properties with depth and at the 13 calibration sites. Figure 5.3a shows clay was largest (≥ 50%) at the northern (11 to 13) end and in the middle of the transect (site 7) in the root (0-2 m) and vadose-zone (2-6 m). This was similarly the case at the southern end (sites 2 and 3). In this regard, 1mHcon was generally above 100 mS m^{-1}.

Beyond 2 m, clay was variable and generally decreased, however, to intermediate-small (20-30 %) clay. At the southern-most site 1, adjacent to the Namoi River, clay was uniformly intermediate-small beyond the rootzone, where it was small (< 20 %). This was also the case in the centre of the transect (sites 5 and 6) in the deeper vadose zone (4-6 m) and in the floodway (9 and 10). At these locations, 1mHcon was smaller and generally below 40 mS m^{-1}.

Figure 5.3b shows the spatial distribution of sand. For the most part it was the inverse of clay. Specifically, at the southern end (sites 2 and 3) where clay was large, sand was small (< 20 %). Conversely, where clay was small, sand was large (≥ 50%); at the southernmost end (site 1) next to the River, just to the south of the centre (sites 5 and 6) and where the channel intersects the northern part of the floodway (site 10). At these sites, large sand indicates the location of the current Namoi River, a prior stream channel and an active floodway, respectively. At these sites, it was noted that poorly sorted gravels were also present at various depths, particularly between 4-6 m. This indicates the presence of previous migrational channels or paleochannels.

Worthy to note, that at the last three sites (sites 5, 6 and 10), the location of coarser sediments was overlain in the topsoil (0-0.5 m) by intermediate-clay (40-50 %) which mask their presence. This was consistent with Stannard and Kelly (1977) who indicated prior stream channels in low-lying areas relative to the plain are characterised by coarse channel sediments, where upper sediments were of a fine texture. Silt (data not shown) did not vary as much as clay and sand with a narrow range (10-30%). The largest silt (20-30 %) was found at most

depths at the southern end (sites 2-4) but also just to the north of centre (sites 5-7) but at deeper depths (> 4 m).

Figure 5.3c, shows distribution of EC_e, which was non-saline (i.e., < 2 dS m^{-1}). Relatively speaking it was large (i.e., > 2.0 dS m^{-1}) in the subsoil (> 2 m) and in various parts of the channel where clay was large (e.g., sites 3, 7 and 11). In the areas where sand was large, the EC_e was small (< 0.5 dS m^{-1}). This suggests these areas were recharge areas and where water drains freely. Figure 5.3e shows CEC was similarly distributed as a function of clay. Of note was that at the northern end CEC was large (> 30 cmol(+) kg^{-1}) and most likely to have shrink swell properties and indicate the location of Vertosols. $pH_{1:5}$ (data not shown) was slightly alkaline in the sandy areas (< 7.5), but moderately (8-8.5) to strongly alkaline (≥ 8.5) in the clayier areas of the channel.

Figure 5.3e shows the spatial distribution of soil texture groups. This was calculated using the Australian Textural Soil Classification based on the laboratory measured particle size fractions of the soil samples. Essentially the transect was characterised by three soil texture groups, but for the most part was predominantly clay. However, next to the River (site 1) and south of the centre (i.e., sites 4-6) and in the floodway (sites 9 and 10), the dominant texture was a sandy clay loam, with clay loams sub-dominant. What is again worth pointing out is that in low-lying areas to the south of the centre and in the narrow floodway in the north, the veneer of finer texture soil (clay) which overlies and masks the coarser (sandy clay loam) soil associated with the prior stream formations.

5.3.2 Relationship between σ and soil variables

The values of σ were estimated from inversion of all EC_a measured using the DUALEM-421 instrument. By comparing various parameters (described in Section 2.4) the

best correlations coefficient between σ and soil physical and chemical properties were achieved by estimating σ using either CF forward model, S2 inversion algorithm and λ = 0.3 or λ = 1.8.

Figure 5.2 shows the EMCI from passing the DUALEM-421 EC_a through EM4Soil and using the CF forward model and S2 inversion algorithm. Figure 5.2d shows the EMCI using a λ = 0.3 and Figure 5.2e shows the EMCI using λ = 1.8. For all intents and purposes, and for both EMCI, the top 1 m was uniformly small in estimates of σ (45 mS m^{-1}). In the areas associated with the clayier parts of the channel, σ increased steadily to large values (>175 mS m^{-1}) at a depth of 1.5 m, which persisted until a depth of around 4 m when λ = 0.3 and to 6 m when λ = 1.8. Conversely, in the sandier part of the channel, σ increased to intermediate (132-175 mS m^{-1}) values before decreasing again to small σ.

Table 5.3 Coefficient of determination (R^2) between selected estimates of true electrical conductivity (σ, mS m^{-1}) and various soil properties.

Soil properties	R^2	
	0.3 S2 CF	1.8 S2 CF
Clay	**0.31*****	0.25***
Slit	0.05**	0.06**
Sand	**0.28*****	0.23***
EC_e	0.22***	**0.39*****
CEC	**0.26*****	0.17***
pH1:5	0.18***	0.32***
*, P < 0.05; **, P < 0.01; ***, P < 0.001		

Table 5.3 shows the best correlation coefficient (R^2) between σ and soil physical properties of clay and sand and chemical property of CEC was best with CF, S2 and λ = 0.3. All relationships, while small, were statistically significant (***, P < 0.001). The CF, S2 and λ = 1.8 produced the largest R^2 (0.39) between σ and EC_e. There was also a statistically significant (***, P < 0.001) correlation between σ and $pH_{1:5}$ (0.32) with estimates of σ from CF, S2 and λ = 1.8. Unlike in the study by Khongnawang et al. (2019) (Khongnawang et al. 2019), where a good R^2 was achieved between σ with clay (0.65) and CEC (0.68) when $EM38h_0$ and

EM38v$_0$ EC$_a$ was inverted, this was not the case here with either of the estimates of σ inverted from the DUALEM-421 EC$_a$ data available.

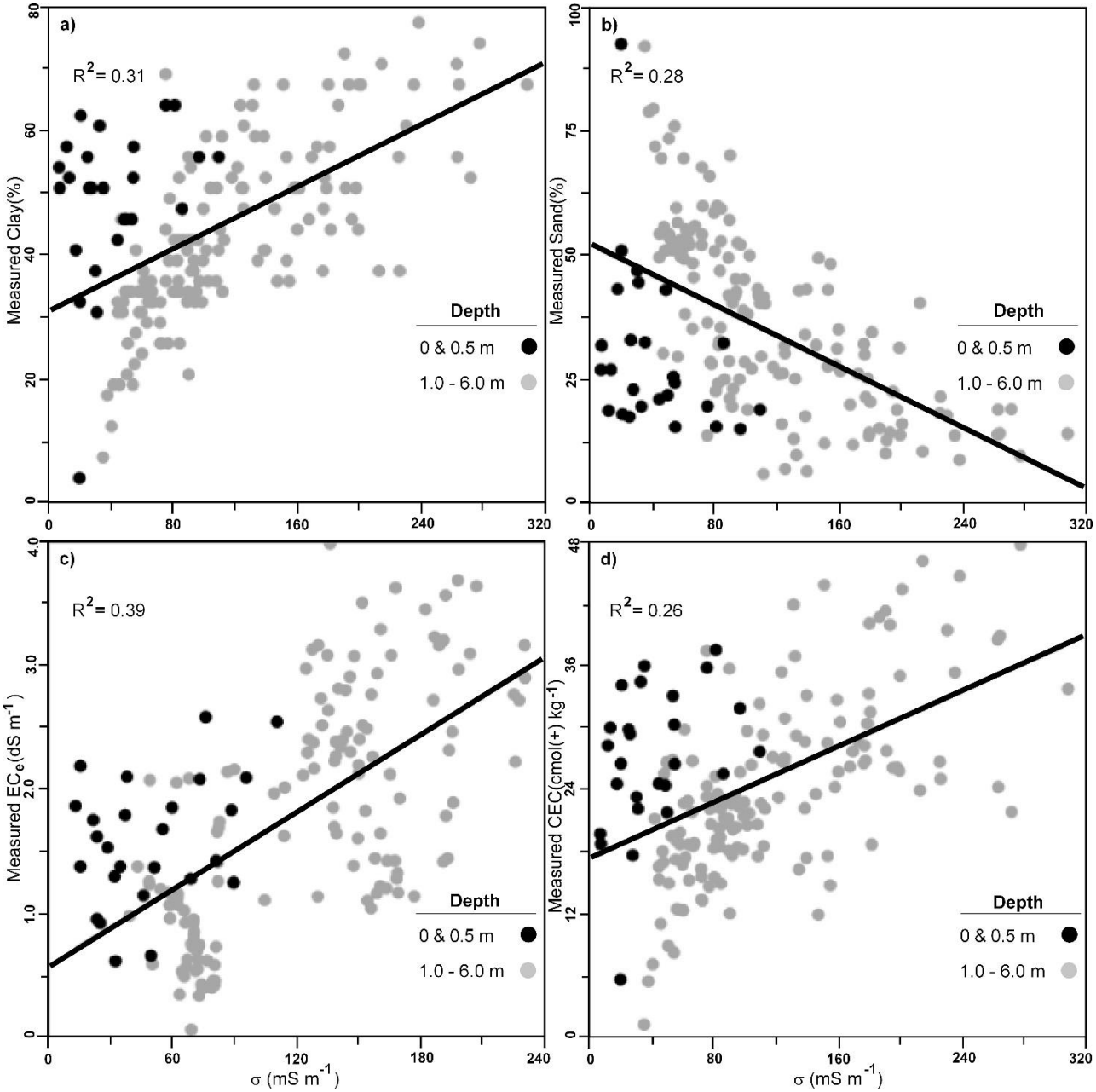


Figure 5.4 plots of estimates of true electrical conductivity (σ, mS m^{-1}) versus predicted a) clay (%), b) sand (%), c) electrical conductivity of a saturated soil paste extract (EC$_e$, dS m^{-1}) and d) cation exchange capacity (CEC, cmol(+) kg^{-1}).

To understand the lack of correlation further, Figure 5.4 shows the plots of σ, derived from the EMCI, against their respective physical and chemical properties. Figure 5.4a shows the plot of σ and clay. It shows that there was linearity between σ and clay at depths of 1-6 m. This was similarly the case with sand, EC$_e$ and CEC (Figure 5.4b, c and d, respectively). This

can be attributed to σ being jointly a function of the interrelationship between clay and CEC with EC_e, whereby the salts in the subsoil accumulate as a function of aeolian and cyclical deposition over time in this semi-arid landscape.

However, this was not the case in the near surface. Specifically, samples from the topsoil (0 m) and subsoil (0.5 m). This can be assigned to the fact that the large σ appears to be a function of the more reactive clay veneer (CEC > 25 cmol(+) kg^{-1}) above the prior stream and floodways and along the channel, because the EC_e herein was non-saline. Hence the smaller values of σ, relative to the subsoil clay which had relatively larger amounts of salts, but the same clay content and CEC.

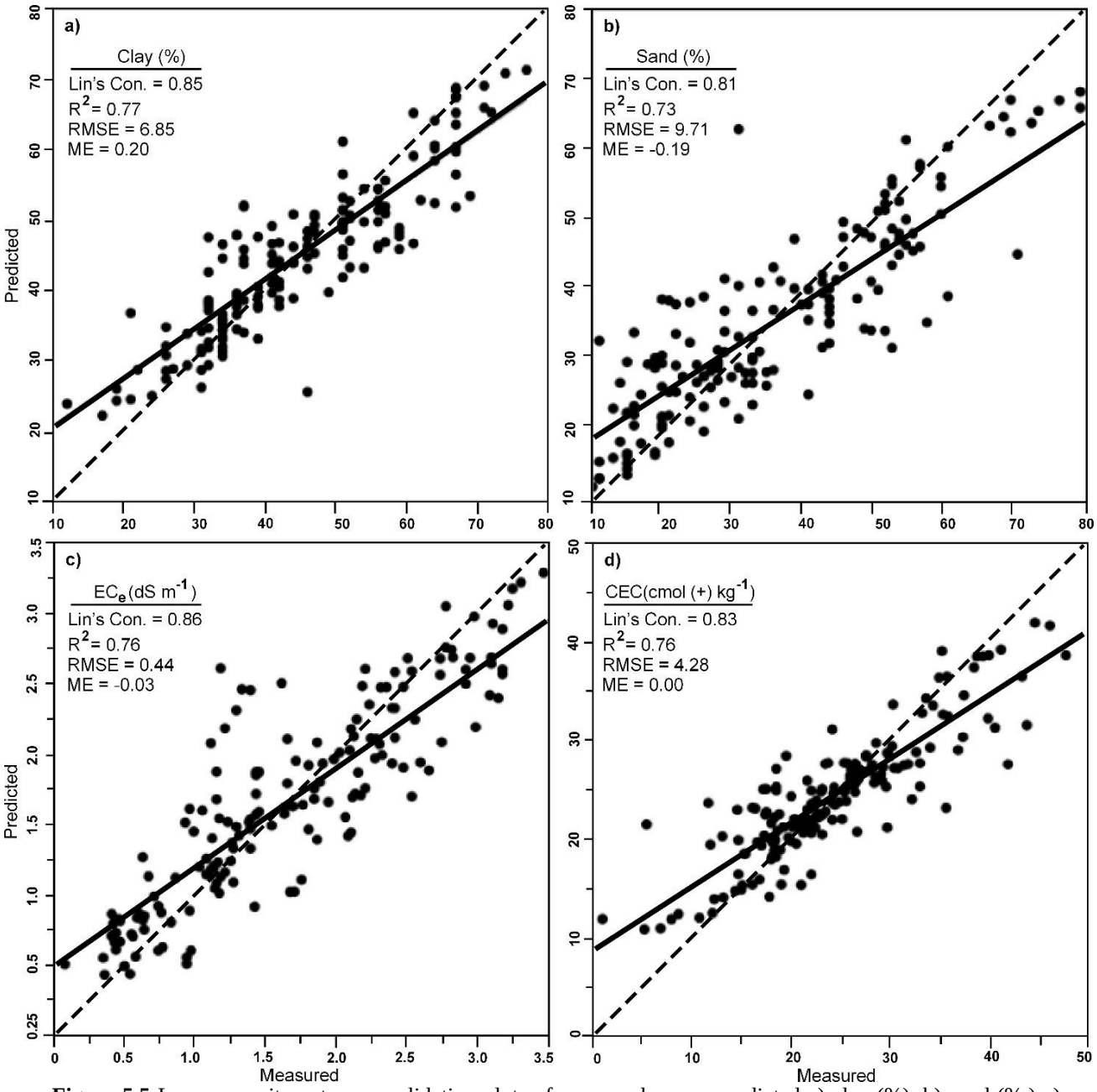


Figure 5.5 Leave-one-site-out cross validation plots of measured versus predicted a) clay (%), b) sand (%,) c) electrical conductivity of a saturated soil paste extract (EC_e, dS m^{-1}) and d) cation exchange capacity (CEC, cmol(+) kg^{-1}).

5.3.3 Measured versus predicted soil properties

Owing to the poor linear relationships and small R^2, and the inability to develop a linear or multiple linear regression model, the potential to develop various SVM models between the various soil physical (i.e., clay and sand) and chemical (EC_e and CEC) properties and estimates of σ was explored. Also, depth was considered in each SVM model. Figure 5.5 shows the leave-

one-site-out cross validation between measured and predicted soil properties using SVM models.

Figure 5.5a shows the predicted clay was in good agreement (Lin's = 0.85) with measured clay, with prediction precision (RMSE = 6.85) and bias (ME = 0.20) both being small and close to zero, respectively. Given the RMSE was less than half the SD for measured clay (7.07 %) these predictions of clay overall can be considered as satisfactory (Singh et al. 2005). They were also unbiased (ME = 0.20). Figure 5.5b shows a comparable result achieved for sand, and while the concordance (Lin's = 0.81) was smaller than clay, predicted sand was still in good agreement. Prediction precision (RMSE = 9.71 %) was also larger than clay, nevertheless, the prediction is still be considered satisfactory because the predicted sand was close to half the SD for measured sand (9.15 %).

With respect to the chemical properties of interest, similar conclusions can be drawn regarding their prediction using the leave-one-site-out cross validation results. In terms of salinity, Figure 5.5c shows for EC_e there was good concordance (Lin's = 0.86). Again, not only were the predictions precise (RMSE = 0.44 dS m^{-1}) and unbiased (ME = -0.03 dS m^{-1}), the precision was satisfactory given they were less than half the SD of measured EC_e (0.45 dS m^{-1}). Figure 5.5d shows that for CEC equivalent results were achieved with predictions showing good concordance (Lin's = 0.83), with precise (4.28 cmol(+) kg^{-1}) and unbiased (0) predictions.

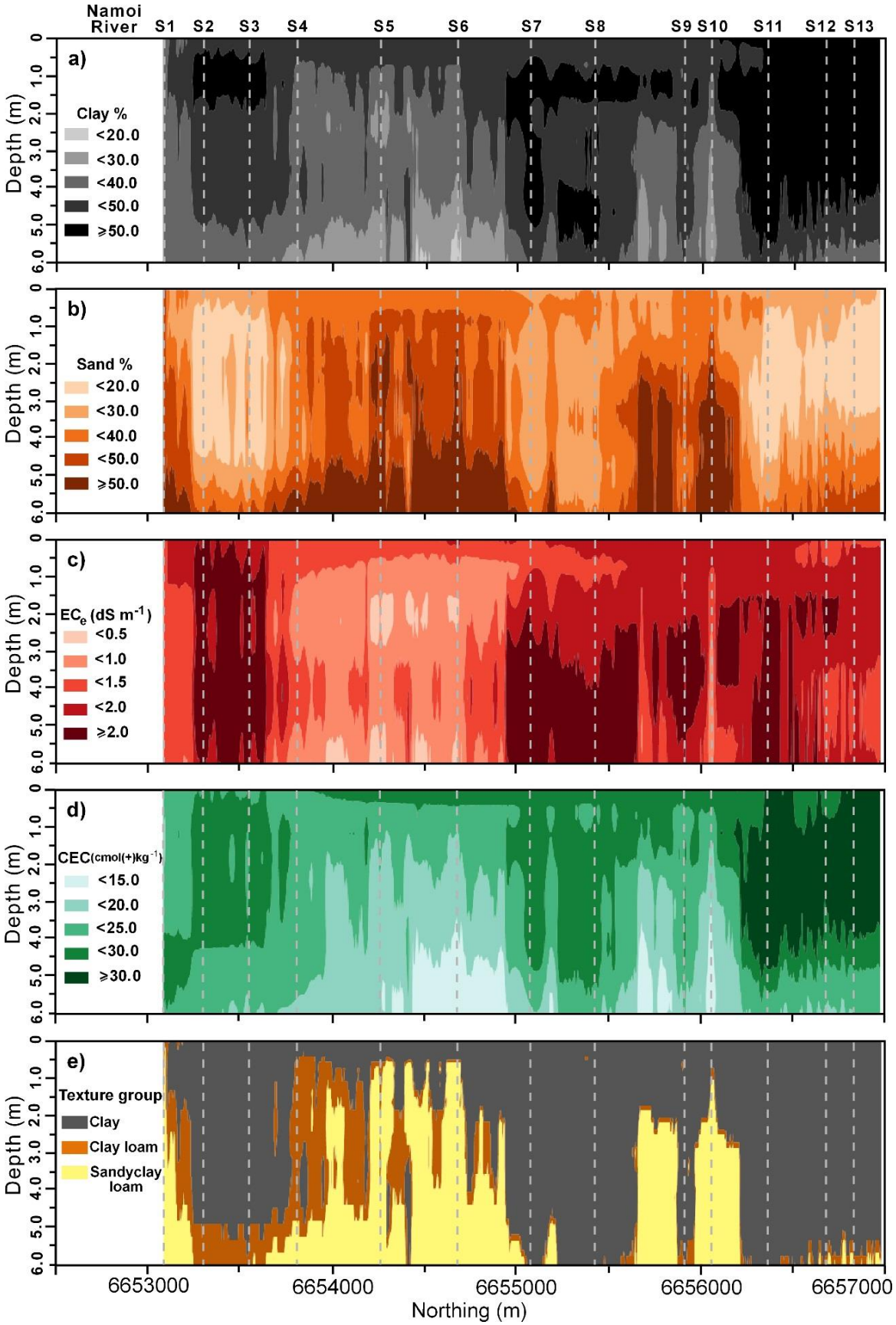


Figure 5.6 The spatial distribution of predicted a) clay (%), b) sand (%) c) electrical conductivity of a saturated soil paste extract (EC_e, dS m^{-1}), d) cation exchange capacity (CEC, cmol(+) kg^{-1}) generated using a support vector machine (SVM) model and from estimates of true electrical conductivity (σ, mS m^{-1}) of DUALEM-421 along with depth and e) The spatial distribution with depth of soil texture groups calculated using Australian Textural Soil Classification based predicted clay and sand.

5.3.4 Mapping soil properties with depth and along the channel

To understand the results from a pedological, geomorphological, soil and water use and management standpoint, predicted physical and chemical properties beneath the channel was plotted. This had been done by applying the SVM using the EMCI's generated from the various EC_a data, using CF, S2, with maximum iteration of 10 and $\lambda = 0.3$ (Figure 5.2d) for clay, sand and CEC and $\lambda = 1.8$ for EC_e (Figure 5.2e).

Figure 5.6a shows that for the most part, the pattern of predicted clay was consistent with measured clay (Figure 5.3a). What was evident was the uniform nature of predicted clay (40-50 %) in the topsoil (0 m) and subsoil (0.5 m) along most of the supply channel. Worthy to note, that clay was slightly under-predicted in the central-north part of the channel, however, where the heavy clay plain was evident (sites 7-10). Nevertheless, the heavy clay (> 50 %) of the plain in the northern part (sites 11-13) was well predicted. So too was the veneer of intermediate-large (40-50 %) clay, which overlies the prior stream channel in the centre and to the south (sites 4-6).

The location of a palaeo-channel, which commonly underlies the clay plain; herein at a depth of around 5-6 m where the clay (< 30 %) was intermediate-small, was also well resolved. These indicate places where previous migrational channel deposits (palaeochannels), characterized by coarse sands and gravels (Stannard and Kelly, 1968) shown in Figure 5.1c can be found. Figure 5.6b shows this more succinctly due to the large sand (> 50 %) evident (4.5-6 m) in the area of the prior stream channel just south of the centre (sites 4-6) and also the area of the floodway in the northern third of the channel (sites 9 and 10).

Figure 5.6c shows predicted EC_e was consistent with the slightly saline (2-4 dS m^{-1}) nature of the Vertosols, which characterise the clay plain. The non-saline (< 2 dS m^{-1}) and highly leached nature of the prior stream formations and the coarser sandy clay loam textures of the Dermosols were also well discerned. Figure 5.6d helps further explain these results as it

shows where the clay was most reactive and reinforces the notion that the clays with larger CEC (30 cmol(+) kg^{-1}) will be more reactive in terms of their shrink and swell and most likely help form a seal in the channel and hold the water sufficiently to enable its conveyance most efficiently along the northern section of the channel.

Conversely, the less reactive parts of the channel (< 20 cmol(+) kg^{-1}) indicates where the shrink-swell behaviour would be poor and where losses can be expected. This aligns well with the sandy clay loam soil textures which characterise the prior stream channels and floodway (see Figure 5.6e). This idea is generally supported by the generally non-saline nature of the soil and vadose zone, thereby indicating where recharge occurs along the channel. It was noted that the prior stream channel clay content was less than 40 % and this was consistent with where research in the nearby lower Gwydir valley showed these clay contents were susceptible to excess deep drainage at the field (Triantafilis et al. 2003) and district scale (Triantafilis et al. 2004).

The results described herein have implications for the remaining lower Namoi valley, where many other supply channels remain unsurveyed. This is because a similar approach can be used by an increasingly large number of consultants who commonly conduct EM surveys using the more popular Geonics EM38 or smaller versions of the DUALEM range of instruments, including DUALEM-1. In this regard, further research could be conducted to see if equivalent calibrations can be developed to predict the physical and chemical properties described herein but only using a smaller number of EC_a values. The transect studied herein could act as a calibration location. A similar approach can be used to develop equivalent calibrations to enable prediction in other cotton growing areas of Australia.

Moreover, the results described herein indicate that in a section of a supply channel, which extends some 300 m across a prior stream channel, future work needs to be undertaken to determine engineering solutions to seal the areas where the clay was predicted to be small (< 40 %) and indicative of where more permeable Dermosols were likely to exist.

5.4 Conclusions

Using a DUALEM-421 it was shown where soil samples sites can be located to characterise the soil types beneath a supply channel based on the measured EC_a. Using EM4Soil inversion software, it was also demonstrated how estimates of σ can be derived from the inversion of the EC_a from a DUALEM-421 and how these were correlated with soil physical properties of clay and sand and chemical properties of CEC when they were estimated using CF, S2 and λ = 0.3 and EC_e using CF, S2 and λ = 1.8. While these relationships were statistically significant (***, $P < 0.001$), the correlations were not strong ($R^2 < 0.4$). These poor correlations can be attributed to the fact that DUALEM-421 EC_a was a function of four separate soil properties (i.e., clay content and mineralogy, salinity and moisture) and the relationship of each of these and σ was non-linear. This was particularly the case for the topsoil (0 and 0.5 m) samples.

Given a limited number of soil cores was available, it was found that, by using a SVM model it is possible to establish a good model between σ and clay, sand, EC_e and CEC. Using the same EMCI, leave-one-site-out cross validation results produced good 1:1 agreement with clay (0.85), sand (0.81), EC_e (0.86) and CEC (0.83). The interpretation of predicted EC_e suggests the approach can identify recharge (non-saline) areas. These aligned with where the clay (< 40 %) was relatively small and in areas associated with prior stream channels that were typically underlain by migrational channels. This success can be attributed to the ability of SVM to perform well on data sets that are nonlinear, sparse and are highly dimensional. The results have practical implications for the farmer and where future work should be focussed along a 300 m section of the supply channel where a prior stream channel was identified. Herein, future work needs to be undertaken to determine a suitable engineering solution to seal the area where the clay was predicted to be too small (< 40 %).

5.5 References

Auken, E., Foged, N., & Sørensen, K. I. (2002). Model recognition by 1-D laterally constrained inversion of resistivity data. 8th EEGS-ES Meeting,

Beretta, A. N., Silbermann, A. V., Paladino, L., Torres, D., Bassahun, D., Musselli, R., & García-Lamohte, A. (2014). Soil texture analyses using a hydrometer: modification of the Bouyoucos method. *Ciencia e investigación agraria*, *41*(2), 263-271.

Gunn, S. R. (1998). Support vector machines for classification and regression. ISIS technical report, *14*(1), 5-16.

Isbell, R. (2016). The Australian soil classification. CSIRO publishing.

Khongnawang, T., Zare, E., Zhao, D., Srihabun, P., & Triantafilis, J. (2019). Three-Dimensional Mapping of Clay and Cation Exchange Capacity of Sandy and Infertile Soil Using EM38 and Inversion Software. *Sensors*, 19(18), 3936.

Li, N., Zare, E., Huang, J., & Triantafilis, J. (2018). Mapping soil cation-exchange capacity using Bayesian modeling and proximal sensors at the field scale. *Soil Science Society of America Journal*, *82*(5), 1203-1216.

McBride, R. B., McBride, G., McBride, G., MCBRIDE, G., McBride, J. B., & McBride, B. (2005). A proposal for strength-of-agreement criteria for Lins Concordance Correlation Coefficient.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C., & Lin, C. (2014). e1071: Misc functions of the Department of Statistics (e1071), TU Wien. R package version, 1(3).

Monteiro Santos, F., Triantafilis, J., Bruzgulis, K., & Roe, J. (2010). Inversion of multiconfiguration electromagnetic (DUALEM-421) profiling data using a one-dimensional laterally constrained algorithm. *Vadose zone journal* VZJ.

Monteiro Santos, F. A., Triantafilis, J., Taylor, R. S., Holladay, S., & Bruzgulis, K. E. (2010). Inversion of conductivity profiles from EM using full solution and a 1-D laterally constrained algorithm. *Journal of Environmental & Engineering Geophysics*, *15*(3), 163-174.

Moody, P. W., & Phan, T. C. (2008). Soil Constraints and Management Package (SCAMP): guidelines for sustainable management of tropical upland soils. Australian Centre for International Agricultural Research (ACIAR).

Muzzamal, M., Huang, J., Nielson, R., Sefton, M., & Triantafilis, J. (2018). Mapping soil particle-size fractions using additive log-ratio (ALR) and isometric log-ratio (ILR) transformations and proximally sensed ancillary data. *Clays and Clay Minerals*, *66*(1), 9-27.

Nagra, G., Burkett, D., Huang, J., Ward, C., & Triantafilis, J. (2017). Field level digital mapping of soil mineralogy using proximal and remote-sensed data. *Soil Use and Managemen*t, *33*(3), 425-436.

Pietsch, T. J. (2006). Fluvial geomorphology and late quaternary geochronology of the Gwydir fan-plain.

Platt, J. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. *Advances in large margin classifiers*, *10*(3), 61-74.

Singh, J., Knapp, H. V., Arnold, J., & Demissie, M. (2005). Hydrological modeling of the Iroquois river watershed using HSPF and SWAT 1. JAWRA *Journal of the American Water Resources Association*, *41*(2), 343-360.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. Statistics and computing, *14*(3), 199-222.

Stannard, M., & Kelly, I. D. (1968). The irrigation potential of the lower Gwydir Valley. Water Conservation and Irrigation Commission.

Stannard, M., & Kelly, I. D. (1977). The irrigation potential of the lower Namoi valley. Water Resources Commission.

Tenison, K., Wild, J., Madden, E., Draper, V., McMahon, L. (2014). Salinity Training Manual,320 pp. State of New South Wales through the Department of Trade and Investment,Regional Infrastructure and Services, Sydney.

Triantafilis, J., Huckel, A., & Odeh, I. (2003). Field-scale assessment of deep drainage risk. *Irrigation Science*, *21*(4), 183-192.

Triantafilis, J., Odeh, I., Jarman, A., Short, M., & Kokkoris, E. (2004). Estimating and mapping deep drainage risk at the district level in the lower Gwydir and Macquarie valleys, Australia. *Australian Journal of Experimental Agriculture*, *44*(9), 893-912.

Triantafilis, J., Ribeiro, J., Page, D., & Santos, F. M. (2013). Inferring the location of preferential flow paths of a leachate plume by using a DUALEM-421 and a Quasi-Three-Dimensional inversion model. *Vadose Zone Journal*, *12(*2).

Triantafilis, J., Ward, W., & McBratney, A. (2001). Land suitability assessment in the Namoi Valley of Australia, using a continuous model. *Soil Research*, *39*(2), 273-289.

Tucker, B. (1974). Laboratory procedures for cation exchange measurements on soils.

Williams, G. (2011). Data mining with Rattle and R: The art of excavating data for knowledge discovery. Springer Science & Business Media.

Woodforth, A., Triantafilis, J., Cupitt, J., Malik, R., Subasinghe, R., Ahmed, M., Huckel, A., & Geering, H. (2012). Mapping estimated deep drainage in the lower Namoi Valley using a chloride mass balance model and EM34 data. *Geophysics*, *77*(4), WB245-WB256.

Zare, E., Huang, J., Santos, F., & Triantafilis, J. (2015). Mapping salinity in three dimensions using a DUALEM-421 and electromagnetic inversion software. *Soil Science Society of America Journal*, *79*(6), 1729-1740.

Zhao, X., Wang, J., Zhao, D., Li, N., Zare, E., & Triantafilis, J. (2019). Digital regolith mapping of clay across the Ashley irrigation area using electromagnetic induction data and inversion modelling. *Geoderma*, *346*, 18-29.

Chapter 6: Scope to map available water content using proximal sensed electromagnetic induction and gamma-ray spectrometry data

6.1 Introduction

The Australian cotton industry is predominantly irrigated, however, approximately 10% is grown under dryland conditions. A fundamental management decision for dryland farmers is determining planting time, with a critical factor influencing whether to plant or not, being amount of stored moisture; because it can help dryland farmers optimize row spacing. Moreover, information about field capacity (i.e., FC-maximum water stored) and permanent wilting point (i.e., PWP-minimum moisture at which a plant wilts) is required. This is particularly the case for Vertosols, which characteristically have > 35 % clay content. While these heavy clays have large FC (i.e., $0.32 < \theta < 0.40$ m^3 m^{-3}), the available water-holding capacity (AWC) is constrained by the PWP also being large (i.e., ($0.20 < \theta < 0.24$ m^3 m^{-3}) (Gardiner and Miller, 2004). Therefore, knowing the AWC can help farmers make management decisions about planting time and row spacing.

The determination of the AWC is usually determined in the laboratory using either pressure plate (Cresswell et al., 2008), hanging column, filter paper (Leong et al., 2002), evaporation or vapour equilibration (Nam et al., 2010) methods. In-field measurements include the use of a tensiometer or resistance blocks. However, these methods are time consuming and only provide point source information. Because the AWC has been shown to be a function of soil texture (McCutcheon et al., 2006), and because the various soil particle size fractions vary spatially (Muzzammal et al., 2018), there is the potential to use a digital soil mapping (DSM) approach to provide this information. Specifically, digital data has been correlated with particle size fractions and mapped at the field (Weller et al., 2007), farm (Khongnawang et al., 2019) and district (Mora-Vallejo et al., 2008) scale.

To map particle size fractions, electromagnetic (EM) induction instruments have commonly been used, because they measure apparent soil electrical conductivity (EC_a), which has been shown to be influenced by clay (e.g., Saey et al., 2011). Moreover, other properties that influence AWC, including cation exchange capacity (CEC; Zhao et al., 2020), mineralogy (Nagra et al., 2017) and organic matter (Huang et al., 2017) have also been mapped. More recently, gamma-ray (γ-ray) spectrometry data has been used to map clay (Arshad et al., 2020), CEC (CEC; Li et al., 2018) and infer mineralogy differences (Triantafilis et al., 2013). With respect to DSM of AWC, Gooley et al. (2014) used a stepwise-multiple linear regression (stepwise-MLR) to predict average rootzone (0-0.9 m) FC and PWP, using ground based EM and airborne γ-ray data at the district scale (40,000 ha).

Recently, research has also shown that environmental indices from remote sensed platforms (e.g., Landsat 7, airborne γ-ray) can be used to map AWC at district (Malone et al., 2009), country (Hong et al., 2013), and continental (Padarian et al., 2014) scales. However, these DSM cannot be used by farmers to make decisions, owing to issues of spectral and spatial resolution. The aim of this study was to demonstrate how a DSM approach could be used to predict FC and PWP and by difference estimate AWC at the field scale in the topsoil (0-0.3 m), subsurface (0.3-0.6 m) and subsoil (0.6-0.9m). In the first instance, FC and PWP were determined using pressure plate apparatus. A multiple linear regression (MLR) and stepwise-MLR were developed between soil and digital data (i.e., EC_a and γ-ray spectrometry) with position (e.g., Easting) also considered. The best model calibration was determined considering adjusted coefficient of determination (adj-R^2) and accuracy (RMSE), with prediction agreement (Lin's concordance) of the DSM of FC and PWP compared.

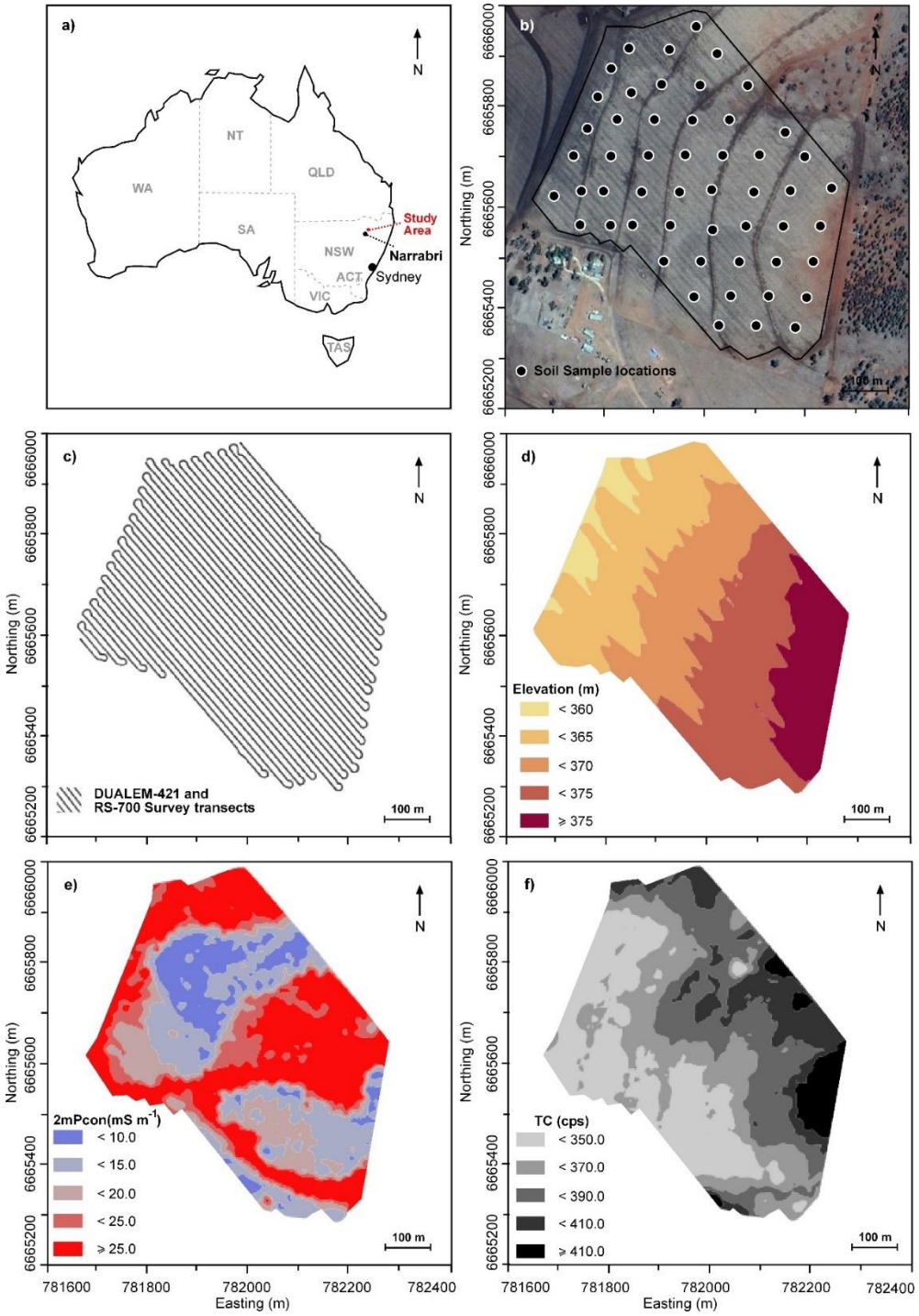


Figure 6.1 a) Location of the study area; b) soil sample locations; c) location of DUALEM-421 and Radiation Solution RS-700 measurement transects; and contour plot of d) elevation (m); e) apparent electrical conductivity (EC_a, mS m^{-1}) of 2m perpendicular coplanar array (2mPcon) collected using a DUALEM-421 and; f) Total Counts (TC, cps) collected using a Radiation Solution RS-700.

6.2 Materials and Methods

6.2.1 Study area

The study field is 23 ha in size and located within a commercial farm (370 m above sea level; 30°06′30″ S, 149°55′35″ E), which is near Narrabri in New South Wales, Australia

(Figure 6.1a). The climate is semi-arid, with a mean annual rainfall of 643 mm and potential evaporation over 1800 mm. The rainfall is summer-dominant, with a mean maximum in January (79.8 mm) and minimum in August (37.5 mm). Dryland cotton is the main cash crop and is grown in summer. Wheat is grown in the winter with chick-peas and other crops grown in rotation.

The study field is located to the west of and lies on the foot-slopes of the Nandewar Range. Specifically, it lies between the Couradda National Park to the north and the Bobbiwaa State Conservation area to the south. The footslopes are underlain by Pilliga Sandstone with the study field soil derived from second terraced fan of local alluvium (Triantafilis et al., 2001). The soil of the study field is predominantly Vertosol. Variations in subsoil colour occur from red brown subsoils on upper slopes (south east) and associated with Pilliga Sandstone to dark grey subsoils on mid-slopes to yellow subsoil on lower slopes (north west). As shown in Figure 6.1d, the elevation is highest in the south-east but falls away gradually to the north-west.

6.2.2 Digital data collection and co-location

The digital data considered includes acquisition from a proximally sensed EM instrument, γ-ray spectrometer, and elevation from a GPS. A DUALEM-421 (Mississauga, Ontario, Canada) was used to collect EC_a. Briefly, the instrument has a transmitter coil located at one end. It also includes a series of horizontal co-planar (HCP) and perpendicular (PRP) receiver array pairs. The distance between the transmitter to the HCP receivers are 1, 2 and 4 m and because the instrument operates at a low frequency (9 kHz) these spacings give a theoretical depth of exploration (DOE) of 0-1.5 m (1mHcon), 0-3.0 m (2mHcon) and 0-6.0 m (4mHcon), respectively. The distance between the transmitter and PRP coils are 1.1, 2.1 and 4.1 m, which gives theoretical DOE of 0-0.5 m (1mPcon), 0-1.0 m (2mPcon) and 0-2.0 m (4mPcon), respectively.

A Radiation Solutions RS-700 (Radiation Solutions, Mississauga, Ontario, Canada) measurement device was used to collect γ-ray spectrometry data from a Radiation Solutions crystal pack (RSX-1). The crystal pack, which is coated with NaI, measures the γ-rays being emitted from radioelements (and energy windows) including; potassium (K: 1.37-1.57 MeV), uranium (U: 1.66-1.86 MeV), thorium (Th: 2.41-2.81 MeV) as well as from the whole spectrum known at the Total Count (TC: 0.41-2.81 MeV). The various elements are presented as follows; K (%), U and Th (ppm) and TC (in counts per second-cps).

Figure 6.1c shows the digital data was collected along 42 transects. The transects were 12 m apart. The DUALEM-421S was strapped onto a PVC sled and towed behind a tractor, with a NovaTel SMARTV1 GPS (Novatel Inc., Hexagon Calgary, Alberta, Canada) antenna used to provide positional information. The RSX-1 crystal pack was situated in front of the same tractor and held approximately 1.5 m off the ground within a pre-fabricated aluminum frame. A Toughbook CF-19 (Panasonic, Kadoma, Osaka, Japan) was used to collect γ-ray data using RS software (Radiation Solutions, Mississauga, Ontario, Canada). A Radiation Solutions GPS antenna was used to provide positional information.

The coordinates were later converted into Easting and Northing (UTM 84). Both EC_a and GPS data were collected using a Panasonic Toughbook CF-19 and logged using HGIS (v. 10.5) Software (StarPal Inc., Fort Collins, Colorado, USA). The total number of γ-ray (11,331) and EC_a (11,415) was approximately equal. To co-locate the digital data, all three data (i.e., γ-ray, EC_a and elevation) were interpolated onto the same 5 × 5 m grid using local ordinary kriging using Vesper (v2.0, Australian Centre for Precision Agriculture) software (Minasny et al., 2005) and considering 20-30 neighbors.

6.2.3 Soil sampling and laboratory analysis

A soil sampling campaign was undertaken and included the collection of soil samples from 52 locations on an approximately regular grid, with samples taken at spacing of 70 m. Figure 6.1b shows the location of the soil samples. Soil samples were collected at 0.3 m depth increments and from topsoil (0-0.3 m), subsurface (0.3-0.6 m) and subsoil (0.6-0.9 m). The samples were air-dried, crushed and sieved (< 2 mm). Laboratory analysis involved determination of soil physical properties, including soil particle size fractions (e.g., clay, silt, and sand-%) and based on the hydrometer method (Beretta et al., 2014). Various chemical properties were also determined including the pH from a 1-part soil to 5-part water extract ($pH_{1:5}$). The soil electrical conductivity was similarly determined from this extract ($EC_{1:5}$-mS m^{-1}). Subsequently EC_e (dS m^{-1}) was estimated from $EC_{1:5}$ using conversion factor for various soil textures, determined from the Australian Soil Texture Triangle, and as adopted from Tenison et al., (2014) as follows; clay ($EC_{1:5} \times 6.7$), clay loam (8.6) and sandy clay loam (9.5).

Tucker's (1974) method was used to determine the CEC (cmol(+) kg^{-1}) using a mechanical leaching device. Briefly, samples were washed with 60% ethanol to remove any soluble salts, followed by extraction with 1M NH_4Cl. The extracts were analyzed using an inductively coupled plasma optical emission spectrometry (ICP-OES). The exchangeable cations determine were summed to estimate the CEC.

A pressure plate was used to determine FC and PWP (Cresswell et al., 2008). The samples were left to equilibrate several days before determination of gravimetric moisture content (w). The FC was determined at -10 kPa and PWP at -1500 kPa. Conversion into θ was achieved using an estimate of bulk density (ρ, g cm^{-3}) and considering the following pedo-transfer function (PTF) to calculate ρ from particle-size analysis data and depth;

$$\rho\ (g\ cm^{-3}) = (1813.6 + 111.27 \log (\text{depth (m)}) - 4.949\ (\text{clay \%} + \text{silt \%})) \div 1000 \quad [6.1]$$

This PTF was suggested and used by Minasny et al. (2006) to predict soil organic carbon from soil organic matter in the Edgeroi district and considering the shrink-swell characteristic of Vertosols. The data used for developing this PTF was derived from various studies in Australia (Minasny and McBratney 2002) where profile bulk density at several depths is complemented with particle-size analysis.

6.2.4 Multiple Linear Regression (MLR) modelling and interpolation

A multiple linear regression (MLR) model of the form:

$$(y) = \beta_0 + \beta_1 w_0 + \beta_2 w_1 + \beta_3 w_2 + \varepsilon \qquad [6.2]$$

where y is vector of the target property (i.e., FC and PWP) and w_k represent the kth predictor digital data (e.g., EC_a) while β_0 is the intercept, β is the slope and ε is the model's residual.

To develop a MLR model between the various sources of digital data and for example FC (0-0.3m), the following MLR model was fitted:

$$FC_{(0-0.3m)} = \beta_0 + \beta_1 w_{1mHcon} + \beta_2 w_{1mpcon} + \beta_3 w_{2mHcon} + \beta_4 w_{2mPcon} + \beta_5 w_{4mHcon} + \beta_6 w_{4mPcon} + \beta_7 w_K + \beta_8 w_{eU} + \beta_9 w_{eTh} + \beta_{10} w_{TC} + + \beta_{11} X_s + \beta_{12} Y_s + \beta_{13} ele + \varepsilon \qquad [6.3]$$

where w_{1mHcon} = 1mHcon, w_{1mPcon} = 1mPcon, w_{2mHcon} = 2mHcon, w_{2mPcon} = 2mPcon, w_{4mHcon} = 4mHcon, w_{4mPcon} = 4mPcon, w_K = radioelement K, w_U = eU, w_{Th} = eTh, and X_s = Easting, Y_s = Northing and ele = Elevation, with β_0 through β_{13} representing empirical regression model parameters. To find the best set of digital data as predictors, the fixed effects were determined prior to model building and by using backwards elimination until all predictors had a probability of being selected (< 0.05).

Validation and evaluating prediction accuracy, bias, and concordance

Herein, a leave-one-site-out cross validation is used to validate both FC and PWP prediction. This was done owing to a limited number of sample sites. Specifically, a whole site had been removed, including all depths. A model was fitted using the remaining calibration

dataset of 51 sites and applied to the site removed. This was done 52 times. Several indices, which measure prediction accuracy (i.e., root mean square error-RMSE), were considered, with respect to measured and predicted FC and PWP. The RMSE was calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (y_i - \hat{y}_i)^2} \qquad [6.4]$$

and considering the measured (y_i) and predicted ($\hat{y}_i$) number (n) of data. If a RMSE value is small (i.e., close to 0), the greater is the prediction accuracy. Similarly, unbiased predictions have ME near zero.

The Lin's concordance correlation coefficient (LCCC) is also used, as a measure of the agreement between measured and predicted PWP and FC. Specifically, how close do the predictions fall along a 45-degree line (Lin, 1989) and is calculated using the following equation:

$$\mathrm{LCCC} = \frac{2\rho\sigma_y\sigma_{\hat{y}}}{\sigma_y{}^2 + \sigma_{\hat{y}}{}^2 + (\mu_y - \mu_{\hat{y}})^2} \qquad [6.6]$$

where σ_y and $\sigma_{\hat{y}}$ are corresponding variances, ρ is the correlation coefficient (Pearson's r) between measured and predicted data; and μ_y and $\mu_{\hat{y}}$ are means for measured and predicted data. McBride (2005) described the agreement classes as; almost perfect (> 0.9), substantial (0.8-0.9), moderate (0.65-0.8) and poor (< 0.65) agreement.

With the intention of providing a better indication of trade-off between bias and variance in the models, the Akaike Information Criterion (AIC) was calculated; because it delivers a balance between goodness of fit and the number of parameters used in prediction of FC or PWP and calculated from the residual sum of squares (R) as follows,

$$\mathrm{AIC} = n\ln R + 2p \qquad [6.7]$$

where n is the number of prediction points and p the number of parameters in the model.

Table 6.1 Summary statistics of kriged digital data across the field and at the 52 soil sample locations. Note: values shown for soil apparent electrical conductivity (EC_a, mS m^{-1}) measured by a DUALEM-421 and Total Count (TC, cps), Potassium (K, %), Uranium (U, ppm) and Thorium (Th, ppm) measured by a RS-700 proximal gamma-ray (γ-ray) spectrometer.

Survey data	N	Min	Mean	Max	Median	CV (%)	SD	Skewness
1mHcon	10838	10	34	90	34	37	13	0.3
1mPcon	10838	1	10	40	10	42	4	0.8
2mHcon	10838	6	42	120	42	47	20	0.3
2mPcon	10838	3	21	81	21	46	10	0.7
4mHcon	10838	9	51	127	51	48	24	0.2
4mPcon	10838	5	38	122	38	49	19	0.4
TC	10838	303	368	451	364	7	25	0.7
K	10838	29	42	64	42	15	6	0.1
U	10838	7	9	13	9	7	1	+0
Th	10838	9	13	19	13	7	1	0.5
Calibration data								
1mHcon	52	12	37	66	38	35	13	0.2
1mPcon	52	3	11	24	11	37	4	1.0
2mHcon	52	10	46	84	48	42	19	0.1
2mPcon	52	7	23	54	24	42	10	0.7
4mHcon	52	12	55	97	58	40	22	-0.1
4mPcon	52	9	41	78	43	42	17	0.1
TC	52	330	367	425	364	7	25	0.6
K	52	30	42	59	41	15	6	0.3
U	52	8	9	11	9	6	1	1.0
Th	52	12	13	16	13	6	1	0.9

6.3 Results and discussion

6.3.1 Preliminary data analysis

Table 6.1a shows the summary statistics of the digital data collected. In terms of EC_a, the 1mPcon had the smallest mean (10 mS m^{-1}), followed by 2mPcon (21mS m^{-1}) and 4mPcon (38 mS m^{-1}) with the deeper measuring Hcon generally larger than equivalent Pcon. The smallest CV was for 1mHcon (37 %) and this indicates all EC_a had high CV (Wilding, 1985). The 1mHcon (0.3), 2mHcon (0.3) and 4mHcon (0.2) were all normally distributed (< 0.5) with 1mPcon (0.8), 2mPcon (0.7) and 4mPcon (0.4) slightly skewed (0.5-1.0). In comparison, the γ-ray data had small CV (< 15%), in particular TC, U and Th (7) with the skewness normally distributed for K (0.1), U (0) and Th (0.5). Table 6.1b shows the summary statistics of the digital data at the 52-soil sample locations, where the digital data had equivalent summary statistics to that collected in the field.

Table 6.2 Summary statistics of measured soil properties by depth, including topsoil (0-0.3 m), subsurface (0.3-0.6 m) and subsoil (0.6-0.9 m). Note: soil salinity of saturated soil paste extract (EC_e, dS m^{-1}) estimated from measured electrical conductivity from a 1 part soil to 5 parts water extract ($EC_{1:5}$); pH of a 1 part soil to 5 parts water extract ($pH_{1:5}$); cation exchange capacity (CEC, cmol(+) kg^{-1}); bulk density (BD, g cm^{-3}); field capacity (FC, m^3 m^{-3}) and permanent wilting point (PWP, m^3 m^{-3}).

Topsoil (0-0.3 m)								
Soil Properties	N	Min	Mean	Max	Median	CV (%)	SD	Skewness
Clay (%)	52	43	61	77	60	12	7.36	-0.04
Silt (%)	52	3.33	10.64	26.67	10.00	36.77	3.91	1.28
Sand (%)	52	13.33	28.39	46.67	26.67	26.67	6.87	0.23
EC_e	52	1.07	2.39	3.71	2.41	24.62	0.59	-0.23
$pH_{1:5}$	52	5.90	6.90	7.73	7.01	6.14	0.42	-0.39
CEC	52	2.69	33.29	59.17	35.14	33.07	11.01	-0.92
BD	52	1.17	1.25	1.34	1.24	2.73	0.03	0.24
FC	52	0.25	0.41	0.53	0.41	14.72	0.06	-0.21
PWP	52	0.11	0.22	0.36	0.20	31.49	0.07	0.52
Subsurface (0.3-0.6 m)								
Soil Properties	N	Min	Mean	Max	Median	CV (%)	SD	Skewness
Clay (%)	52	33.33	59.33	90.00	60.00	19.64	11.65	-0.22
Silt (%)	52	3.33	9.64	23.33	10.00	51.12	4.93	0.27
Sand (%)	52	6.67	31.14	60.00	30.00	34.60	10.77	0.45
EC_e	52	0.95	2.28	3.22	2.36	17.70	0.41	-0.85
$pH_{1:5}$	52	6.17	7.39	7.91	7.47	4.84	0.35	-1.30
CEC	52	2.73	36.12	51.31	36.93	23.62	8.53	-1.12
BD	52	1.26	1.38	1.52	1.38	3.92	0.05	0.40
FC	52	0.21	0.43	0.55	0.44	24.22	0.11	-0.05
PWP	52	0.08	0.23	0.38	0.22	36.59	0.08	0.01
Subsoil (0.6-0.9 m)								
Soil Properties	N	Min	Mean	Max	Median	CV (%)	SD	Skewness
Clay (%)	52	16.67	54.86	83.33	60.00	30.57	16.77	-0.43
Silt (%)	52	0.00	9.77	23.33	10.00	48.27	4.71	0.23
Sand (%)	52	13.33	35.50	80.00	31.67	43.18	15.33	0.72
EC_e	52	0.77	2.14	3.43	2.19	26.97	0.58	0.14
$pH_{1:5}$	52	6.70	7.81	8.30	7.89	3.98	0.31	-1.12
CEC	52	5.97	28.73	43.56	29.01	29.71	8.54	-0.63
BD	52	1.35	1.46	1.68	1.4	5.21	0.08	0.74
FC	52	0.21	0.39	0.51	0.39	27.15	0.11	0.17
PWP	52	0.10	0.23	0.35	0.23	38.08	0.09	0.14

Table 6.2 shows the summary statistics of the various soil properties. In terms of average clay, the topsoil (61 %) was largest and decreased with increasing depth with subsurface (59) and subsoil (55) also large clay. The sand increased with depth (e.g., topsoil mean = 28 %) with the silt content generally small (e.g., topsoil mean = 11 %) at all depths. On average the soil texture was heavy clay in the topsoil grading into a medium clay in the subsoil. With increasing depth, average pH increased from slightly acidic $pH_{1:5}$ in the topsoil (6.90) to slightly alkaline in the subsoil (7.81). At all depths, the average salinity was slightly saline (e.g., topsoil = 2.39 dS m^{-1}). Moreover, average CEC decreased with increasing depth

and was indicative of high clay (i.e., ~30 cmol(+) kg^{-1}) with moderate shrink-swell capacity and structural resilience.

6.3.2 Correlation coefficient between soil and digital data

Table 6.3 shows the coefficient of determination (R^2) between measured soil properties at all depths. According to Moore et al., (2007), the FC and PWP were moderately ($0.5 < R^2 < 0.7$) correlated with each other. With respect to the remaining soil properties, FC (0.53) and PWP (0.41) were moderately and weakly ($0.3 < R^2 < 0.5$) correlated with clay, respectively. While the R^2 between FC and PWP were statistically significant with other soil properties they were all very weak ($R^2 < 0.3$).

Table 6.3 Coefficient of determination (R^2) between different measured soil properties at 52 sampling points. Note: Soil salinity of a saturated soil paste extract (EC_e, dS m^{-1}) estimated from measured electrical conductivity from a 1 part soil to 5 parts water extract ($EC_{1:5}$); pH of a 1 part soil to 5 parts water extract ($pH_{1:5}$); cation exchange capacity (CEC, cmol(+) kg^{-1}); field capacity (FC, m^3 m^{-3}) and permanent wilting point (PWP, m^3 m^{-3}).

Properties	Clay	Silt	Sand	EC_e	$pH_{1:5}$	CEC	FC
Silt	0.09***						
Sand	0.86***	0.00					
EC_e	0.19***	0.02***	0.08***				
$pH_{1:5}$	0.02	0.00	0.01	0.00			
CEC	0.04*	0.01	0.07	0.02	0.00		
FC	0.53***	0.03*	0.52***	0.19***	0.01	0.05*	
PWP	0.41***	0.08**	0.36***	0.11***	0.01	0.02	0.60***
*, P < 0.05; **, P < 0.01; ***, P < 0.001							

Table 6.4 shows the R^2 between measured soil properties and digital data at the 52 soil sampling locations. While many of the soil properties were statistically correlated with FC the R^2 were generally weak. In the topsoil Table 6.4a shows FC was weakly correlated with 2mPcon (0.35) and TC (0.36). Table 6.4b shows subsurface FC was strongly ($R^2 > 0.70$) correlated (Moore et al., 2007) with 2mPcon (0.71) and moderately correlated with 4mPcon (0.68) and weakly correlated with TC (0.13). Table 6.4c shows subsoil FC was moderately correlated with 2mPcon (0.62) and again weakly correlated with TC (0.13). Table 6.4 also shows equivalent correlations between PWP and equivalent digital data described above, however, the correlations were generally weaker.

Table 6.4 Coefficient of determination (R^2) between measured soil properties and digital data at 52 soil sampling locations and for a) topsoil (0-0.3 m) b) subsurface (0.3-0.6 m) and c) subsoil (0.6-0.9 m). Note: Total Count (TC, cps), Potassium (K, %), Uranium (U, ppm), Thorium (Th, ppm); soil salinity of a saturated soil paste extract (EC_e, dS m^{-1}) estimated from measured electrical conductivity from a 1 part soil to 5 parts water extract ($EC_{1:5}$); pH of a 1 part soil to 5 parts water extract ($pH_{1:5}$) ; cation exchange capacity (CEC, cmol(+) kg^{-1}); field capacity (FC, m^3 m^{-3}) and permanent wilting point (PWP, m^3 m^{-3}).

a)	Topsoil (0-0.3 m)									
Properties	1mHcon	1mPcon	2mHcon	2mPcon	4mHcon	4mPcon	TC	K	U	Th
Clay	0.26***	0.28***	0.23***	0.21***	0.27***	0.21***	0.21***	0.22***	0.00	0.03
Silt	0.01	0.00	0.02	0.02	0.01	0.02	0.03	0.05	0.01	0.00
Sand	0.25***	0.21***	0.21***	0.17***	0.25***	0.19***	0.18***	0.17***	0.01	0.03
EC_e	0.07	0.06	0.09*	0.08*	0.08*	0.12*	0.00	0.00	0.00	0.02
$pH_{1:5}$	0.09***	0.08	0.14***	0.09*	0.16**	0.14	0.01	0.02	0.00	0.00
CEC	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.05
FC	0.25***	0.33***	0.33***	0.35***	0.30***	0.32***	0.36***	0.29***	0.06	0.11*
PWP	0.23***	0.35***	0.30***	0.33***	0.31***	0.32***	0.38***	0.32***	0.05	0.04
b)	Subsurface (0.3-0.6 m)									
Properties	1mHcon	1mPcon	2mHcon	2mPcon	4mHcon	4mPcon	TC	K	U	Th
Clay	0.48***	0.33***	0.47***	0.48***	0.43***	0.49***	0.09***	0.08***	0.02	0.00
Silt	0.03	0.01	0.07	0.03	0.07	0.06	0.01	0.04	0.00	0.02
Sand	0.46***	0.35***	0.40***	0.45***	0.36***	0.42***	0.10***	0.09***	0.01	0.00
EC_e	0.13**	0.04	0.12**	0.14**	0.08*	0.13**	0.07	0.04	0.04	0.06
$pH_{1:5}$	0.14**	0.03	0.17	0.13	0.13	0.16	0.23***	0.15***	0.08*	0.05
CEC	0.09*	0.01	0.08*	0.07	0.04	0.05	0.18**	0.20**	0.03	0.02
FC	0.63***	0.60***	0.63***	0.71***	0.59***	0.68***	0.13***	0.11***	0.00	0.00
PWP	0.60***	0.51***	0.56***	0.62***	0.54***	0.60***	0.14***	0.10***	0.04	0.00
c)	Subsoil (0.6-0.9 m)									
Properties	1mHcon	1mPcon	2mHcon	2mPcon	4mHcon	4mPcon	TC	K	U	Th
Clay	0.44***	0.36***	0.37***	0.45***	0.33***	0.38***	0.12**	0.10**	0.01	0.01
Silt	0.02	0.03	0.03	0.02	0.03	0.02	0.00	0.00	0.02	0.05
Sand	0.47***	0.38***	0.39***	0.49***	0.34***	0.41***	0.10**	0.09**	0.00	0.02
EC_e	0.14**	0.08*	0.12*	0.17**	0.09**	0.14	0.02	0.02	0.00	0.03
$pH_{1:5}$	0.01	0.00	0.02	0.01	0.03	0.03	0.00	0.01	0.05	0.07**
CEC	0.09*	0.02	0.07	0.10*	0.03	0.08*	0.20**	0.12*	0.02	0.14
FC	0.54***	0.55***	0.53***	0.62***	0.47***	0.55***	0.13***	0.09**	0.00	0.01
PWP	0.42***	0.38***	0.42***	0.49***	0.37***	0.43***	0.10**	0.07**	0.01	0.01

*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$

6.3.3 Spatial distribution of digital data

Figure 6.1e, shows the spatial distribution of 2mPcon, which was large (> 25 mS m^{-1}) in the northern part and in the lower landscape positions (<360 m). The 2mPcon was also large in a continuous belt which extends along the southern boundary to the southeast corner and through a large section in the center and represent the location of the clay rich (>35 %) Vertosols associated with the second terraced fan of local alluvium (Triantafilis et al., 2001). In between these large values, the 2mPcon decreases rapidly to intermediate-small (10-15 mS

m^{-1}) and small (<10 mS m^{-1}). These areas represent the loamy soil of the Pilliga Sandstone (Triantafilis et al., 2001).

Figure 6.1f shows the spatial distribution of TC, which was largest (>410 cps) along the eastern margin and associated with the highest elevation (>375 m). This was close to the contact with the Pilliga Sandstone foothills. From here and down the slope toward the northeast, the TC generally decreased gradually, where at the lower slope positions (i.e., < 360 m) the TC was small (350 cps).

6.3.4 Multiple Linear Regression (MLR vs stepwise-MLR)

Table 6.5a shows the MLR summary statistics considering all digital data used to predict topsoil FC. Specifically, it shows that the coefficient of determination (R^2 = 0.59) was moderate with the adj-R^2 smaller (0.44). The analysis of variance showed the RMSE was 0.045 m^3 m^{-3} with the AIC -147. Moreover, only the 1mHcon (0.046) and 2mPcon (0.0491) were significant ($P < 0.05$) statistically (not shown).

Table 6.5 Summary statistics for prediction of field capacity (FC, $m^3\ m^{-3}$) using a Multiple Linear Regression (MLR) model and stepwise-MLR modelling procedure of digital data including TC (Total Count, cps) for a) topsoil (0-0.3 m) b) subsurface (0.3-0.6 m) and c) subsoil (0.6-0.9 m).

Model summary statistics					
a)	Topsoil (0-0.3 m)				
	R^2	Adj. R^2	RMSE	AIC	
MLR	0.59	0.44	0.045	-147	
Stepwise -MLR	0.57	0.52	0.042	-171	
Variable	Estimate		Std. error	t-test	P > |t|
Intercept	1029.4344		265.0555	3.88	0.0003
2mPcon	0.01235		0.002897	4.26	<0.0001
Northing	-0.000154		$3.977e^{-5}$	-3.88	0.0003
TC	0.0012752		0.000374	3.41	0.0014
1mPcon	-0.009711		0.004281	-2.27	0.0280
1mHcon	-0.003171		0.001311	-2.42	0.0196
b)	Subsurface (0.3-0.6 m)				
	R^2	Adj. R^2	RMSE	AIC	
MLR	0.75	0.67	0.060	-118	
Stepwise -MLR	0.74	0.73	0.054	-150	
Variable	Estimate		Std.error	t-test	P > |t|
Intercept	0.0995172		0.0599	1.66	0.1030
2mPcon	0.0096709		0.000818	11.82	<0.0001
K	0.002654		0.001241	2.14	0.0375
c)	Subsoil (0.6-0.9 m)				
	R^2	Adj. R^2	RMSE	AIC	
MLR	0.75	0.65	0.062	-114	
Stepwise -MLR	0.73	0.70	0.058	-116	
Variable	Estimate		Std. error	t-test	P > |t|
Intercept	-1848.846		474.2617	-3.90	0.0003
Northing	0.0002775		$7.117e^{-5}$	3.90	0.0003
1mPcon	0.0328198		0.010108	3.25	0.0022
2mHcon	0.0053913		0.002	2.70	0.0098
TC	-0.001723		0.000709	-2.43	0.0190
2mPcon	-0.014803		0.007303	-2.03	0.0485

Table 6.5a also shows the stepwise-MLR summary statistics considering only the statistically significant digital data, including 2mPcon (<0.0001), Northing (0.0003), T.C. (0.0014), 1mPcon (0.0280) and 1mHcon (0.0196). This was determined using a forward sequential stepwise-MLR variable selection procedure, with significance level of 0.25 set as probability to enter and 0.05 to remove (SAS, 2012) the digital data. The R^2 (0.57), while moderate, was smaller than the MLR, but the adj-R^2 was larger (0.52). The analysis of variance

showed the RMSE (0.042) and AIC (-171) were also smaller. Given the stepwise-MLR outperformed MLR, the former was used to predict FC.

Table 6.5b shows the model summary statistics for the forward sequential stepwise-MLR to predict subsurface FC considering the statistically significant 2mPcon (<0.0001) and K (0.0375). The R^2 (0.74) was strong, and much larger than for the topsoil, with the adj-R^2 also larger (0.73). However, the analysis of variance showed the RMSE (0.054) and AIC (-150) were both larger than the topsoil FC stepwise-MLR. Table 6.5c shows the model statistics for the forward sequential stepwise-MLR to predict subsoil FC considering the statistically significant Northing (0.0003), 1mPcon (0.0022), 2mHcon (0.0098), TC (0.0190) and 2mPcon (0.0485). The R^2 (0.73) was again strong, with adj-R^2 larger (0.70) than the topsoil and equivalent to subsurface. The analysis of variance showed RMSE (0.058) and AIC (-116) were also larger than the topsoil and subsurface. These statistics indicate that with increasing depth the stepwise-MLR models have better correlations but larger RMSE.

Table 6.6a shows the model summary statistics for the MLR and forward sequential stepwise-MLR to predict topsoil PWP considering the statistically significant TC (<0.0001), 1mPcon (<0.0001) and Northing (0.0109). Interestingly, the TC was the most statistically significant digital data. Nevertheless, and as for the topsoil FC, the MLR model to predict PWP had a moderate R^2 (0.69), which was larger than to predict topsoil FC. Moreover, it was larger than the R^2 (0.57) for the stepwise-MLR model to predict topsoil FC. Likewise, the adj-R^2 was larger (0.62) for the stepwise-MLR. The analysis of variance also showed that the stepwise-MLR had RMSE (0.042) and AIC (-168) comparable with the topsoil FC (0.42 and -171) stepwise-MLR.

Table 6.6 Summary statistics for prediction of permanent wilting point (PWP, m^3 m^{-3}) using a Multiple Linear Regression (MLR) model and stepwise-MLR modelling procedure of digital data including TC (Total Count, cps) for a) topsoil (0-0.3 m) b) subsurface (0.3-0.6 m), and c) subsoil (0.6-0.9 m).

Model summary statistics (Stepwise)					
a)	Topsoil (0-0.3 m)				
	R^2	Adj. R^2	RMSE	AIC	
MLR	0.69	0.58	0.045	-148	
Stepwise -MLR	0.65	0.62	0.042	-168	
Variable	Estimate		Std. error	t-test	P > |t|
Intercept	-1137.263		472.1113	-2.41	0.0200
TC	0.0016858		0.000251	6.71	<0.0001
1mPcon	0.0060069		0.001408	4.27	<0.0001
Northing	$9.4912e^{-5}$		$3.583e^{-5}$	2.65	0.0109
b)	Subsurface (0.3-0.6 m)				
	R^2	Adj. R^2	RMSE	AIC	
MLR	0.73	0.63	0.050	-137	
Stepwise -MLR	0.71	0.68	0.046	-163	
Variable	Estimate		Std.error	t-test	P > |t|
Intercept	-0.336477		0.12623	-2.67	0.0105
2mPcon	0.0112208		0.002019	5.56	<0.0001
K	0.0048457		0.001478	3.28	0.0020
1mPcon	-0.010396		0.004833	-2.15	0.0367
c)	Subsoil (0.6-0.9 m)				
	R^2	Adj. R^2	RMSE	AIC	
MLR	0.76	0.67	0.050	-135	
Stepwise -MLR	0.73	0.69	0.049	-153	
Variable	Estimate		Std. error	t-test	P > |t|
Intercept	-849.1207		328.1165	-2.59	0.0129
Easting	-0.000339		$8.482e^{-5}$	-4.00	0.0002
1mPcon	0.0238242		0.006275	3.80	0.0004
Northing	0.0001672		0.000046	3.63	0.0007
4mHcon	0.0025985		0.000812	3.20	0.0025
2mPcon	-0.009673		0.003803	-2.54	0.0144

Table 6.6b shows the model summary statistics for the forward sequential stepwise-MLR to predict subsurface PWP considering the statistically significant 2mPcon (<0.0001), K (0.002) and 1mPcon (0.367). The R^2 (0.71) was strong and larger than for the topsoil, with the adj-R^2 also larger (0.68). The analysis of variance showed the RMSE (0.046) and AIC (-163) were also both larger than the topsoil PWP stepwise-MLR. Table 6.6c shows the model statistics for the forward sequential stepwise-MLR to predict subsoil PWP considering only Easting (0.0129), 1mPcon (0.0004), Northing (0.0007), 4mHcon (0.0025) and 2mPcon (0.0144). The R^2 (0.73) was strong and adj-R^2 was larger (0.69) than topsoil and equivalent to

subsurface. The analysis of variance again showed RMSE (0.049) and AIC (-153) were larger than the topsoil. These statistics indicate with increasing depth, the stepwise-MLR have better correlations but larger RMSE.

The results achieved herein are comparable to the results achieved by Gooley et al., (2014) who aimed to develop a calibration between proximal sensed EC_a (e.g., EM38) and remote sensed γ-ray data across two large irrigated cotton growing areas of the lower Macquarie valley. They also found that using stepwise-MLR was superior to model a relationship between average (0-0.9 m) FC (adj-R^2 = 0.55) and PWP (adj-R^2= 0.47) than MLR. Their calibration relationships were at the lower end of moderate and weak and were not as strong as herein. This was attributable to the PWP and FC data being averaged over three depths (i.e., 0-0.9 m) and that the γ-ray data was not point based but encompassed a large footprint owing to being measured from an aircraft.

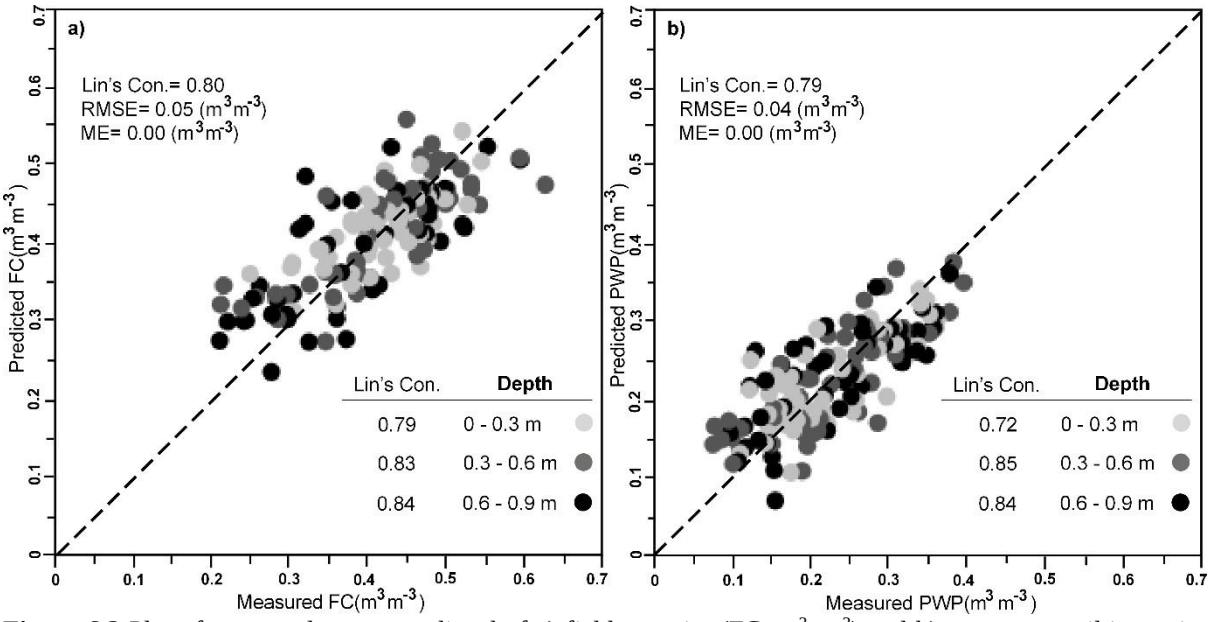


Figure 6.2 Plot of measured verses predicted of a) field capacity (FC, m^3 m^{-3}) and b) permanent wilting point (PWP, m^3 m^{-3}) using stepwise multiple linear regression (stepwise-MLR) models shown in Tables 5 and 6.

6.3.5 Model agreement: FC and PWP

Figure 6.2 shows the measured versus predicted agreement using the stepwise-MLR models. Figure 6.2a shows the results for FC. Overall, and considering all three depths combined the agreement (Lin's = 0.80) was substantial, with agreement also moderate when considering

only the topsoil (0.79) but it was substantial when considering only the subsurface (0.83) and subsoil (0.84) FC alone. It should be noted that the range in predicted topsoil FC was smallest (0.23 m^3 m^{-3}) with the range in subsoil FC (0.47 m^3 m^{-3}) largest. Figure 6.2b shows the results for PWP which were also moderate (0.79) considering all three depths combined, with agreement moderate in the topsoil (0.72) and again substantial in the subsurface (0.85) and subsoil (0.84) PWP. Again, the range in predicted topsoil PWP was smallest (0.24 m^3 m^{-3}) with the range in subsoil PWP (0.37 m^3 m^{-3}) largest. These results were attributable to the generally better correlation coefficients in the subsurface and subsoil.

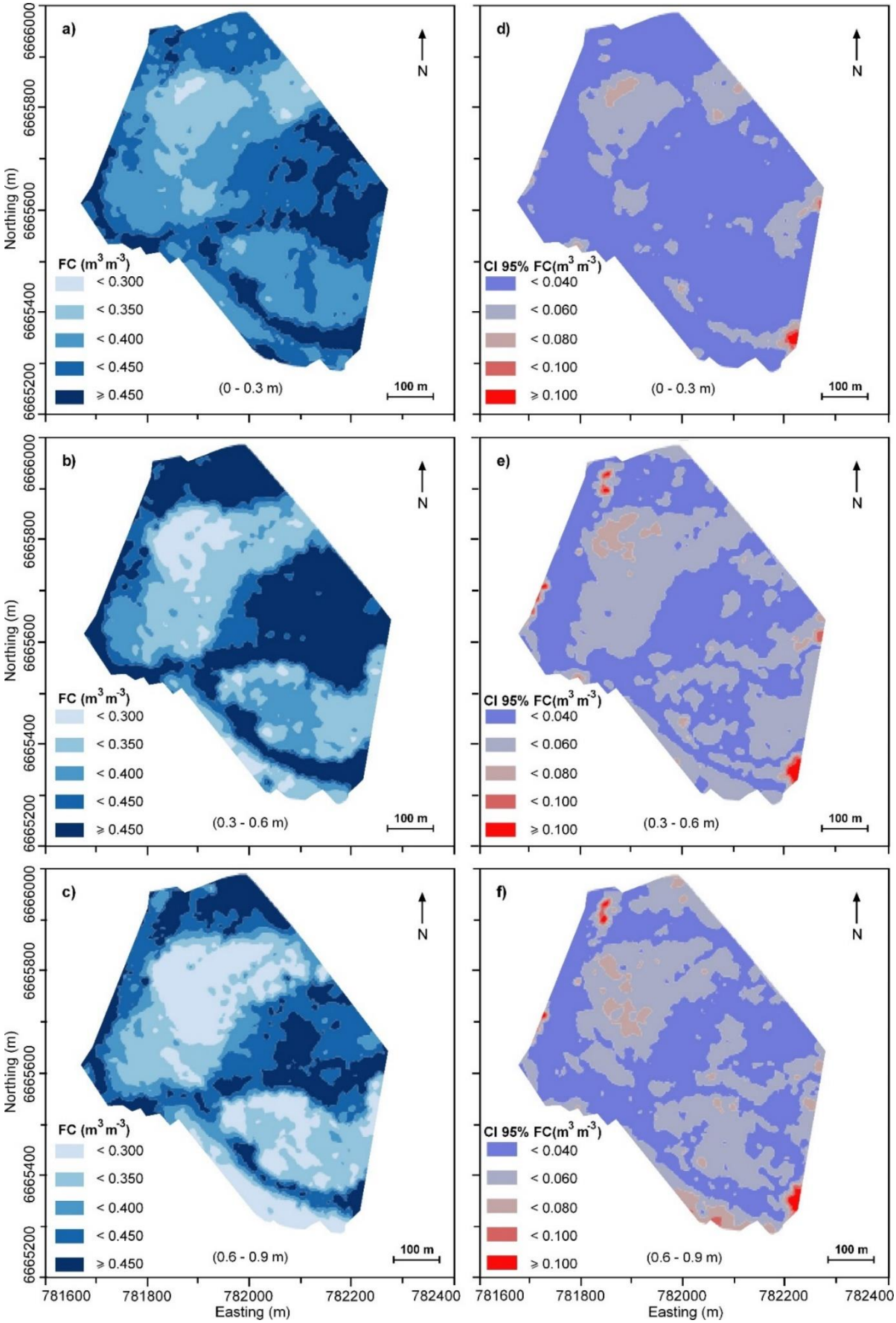


Figure 6.3 Digital soil map (DSM) of predicted field capacity (FC, m^3 m^{-3}) in the a) topsoil (0-0.3 m), b) subsurface (0.3-0.6 m) and c) subsoil (0.6-0.9 m); and, spatial distribution of 95% confidence intervals in the d) topsoil (0-0.3 m) e) subsurface (0.3-0.6 m) and f) subsoil (0.6-0.9 m).

6.3.6 Digital soil maps: FC and PWP

Figure 6.3a shows the DSM of predicted topsoil FC using the stepwise-MLR. In the northern part and in the lower landscape positions (<360 m), topsoil FC was intermediate-large (0.4-0.45 m^3 m^{-3}) or large (>0.45 m^3 m^{-3}). The FC was also large and/or intermediate-large in a continuous belt that extended along the southern boundary to the southeast corner and through a large section in the center. These values of FC are what would be expected for clay soil (Gardiner and Miller, 2004) and represent the location of the Vertosols associated with the second terraced fan of local alluvium. In between these large values, predicted FC decreases to intermediate-small (0.3-0.35 m^3 m^{-3}) and small (<0.30 m^3 m^{-3}), with the latter representing the loamy soil of the Pilliga Sandstone. Figure 6.3d shows the associated 95% confidence interval (CI) was generally small (< 0.04 m^3 m^{-3}) across the whole field and for a wide range of FC. There were a few areas where the CI was intermediate-small (0.04-0.06 m^3 m^{-3}) and larger. This was most evident in the lower lying area in the northern central part of the field where the FC was also intermediate-small or small. This was attributed to the fact that as shown in the spatial distribution of 2mPcon (Figure 6.1e) and TC (Figure 6.1f), both digital data exhibited different short scale variation. With respect to the former, its variation was in the effect of a halo radiating out from small values whereas the latter showed a trend of increasing values from north to south. Other areas with equivalent CI were apparent in small contiguous areas and in the southern corner in particular where the largest CI (> 0.1 m^3 m^{-3}) was associated with the drainage line where the largest values of FC (> 0.45 m^3 m^{-3}) were predicted. These values of CI can be attributed to the edge effect and the short scale variation associated with the narrow band of 2mPcon in this location, but also because the TC varied in a different way to the EC_a.

Figure 6.3b shows the DSM of predicted subsurface FC was equivalent, with the main difference being the greater extent of predicted large and small FC. Figure 6.3c, which shows the DSM of predicted subsoil FC, shows that the intermediate-large or large FC in the center

was generally equivalent to topsoil predicted FC, with the extent of small FC much larger than in the topsoil and subsurface. Owing to the greater RMSE of the stepwise-MLR with increasing depth, the CI, while generally exhibiting the same spatial pattern as for the topsoil, showed progressively large CI values with the subsoil having the greatest extent of large CI (> 0.1 m^3 m^{-3}). This was evident in the southern corner and in a small area in the north and western corners. In all instances, this was associated with areas where FC was large (> 0.45 m^3 m^{-3}) and because the EC_a (e.g., 2mPcon) and γ-ray (e.g., TC) data varied in different ways. Moreover, it was also a function of the short scale variation in both digital data and the proximity of these areas to the edges of the field. Nevertheless, and for either depth, the DSM of FC showed similar patterns to the measured EC_a, because the statistically significant digital data used in the stepwise-MLR was the EC_a data (e.g., 2mPcon) in the topsoil and subsurface.

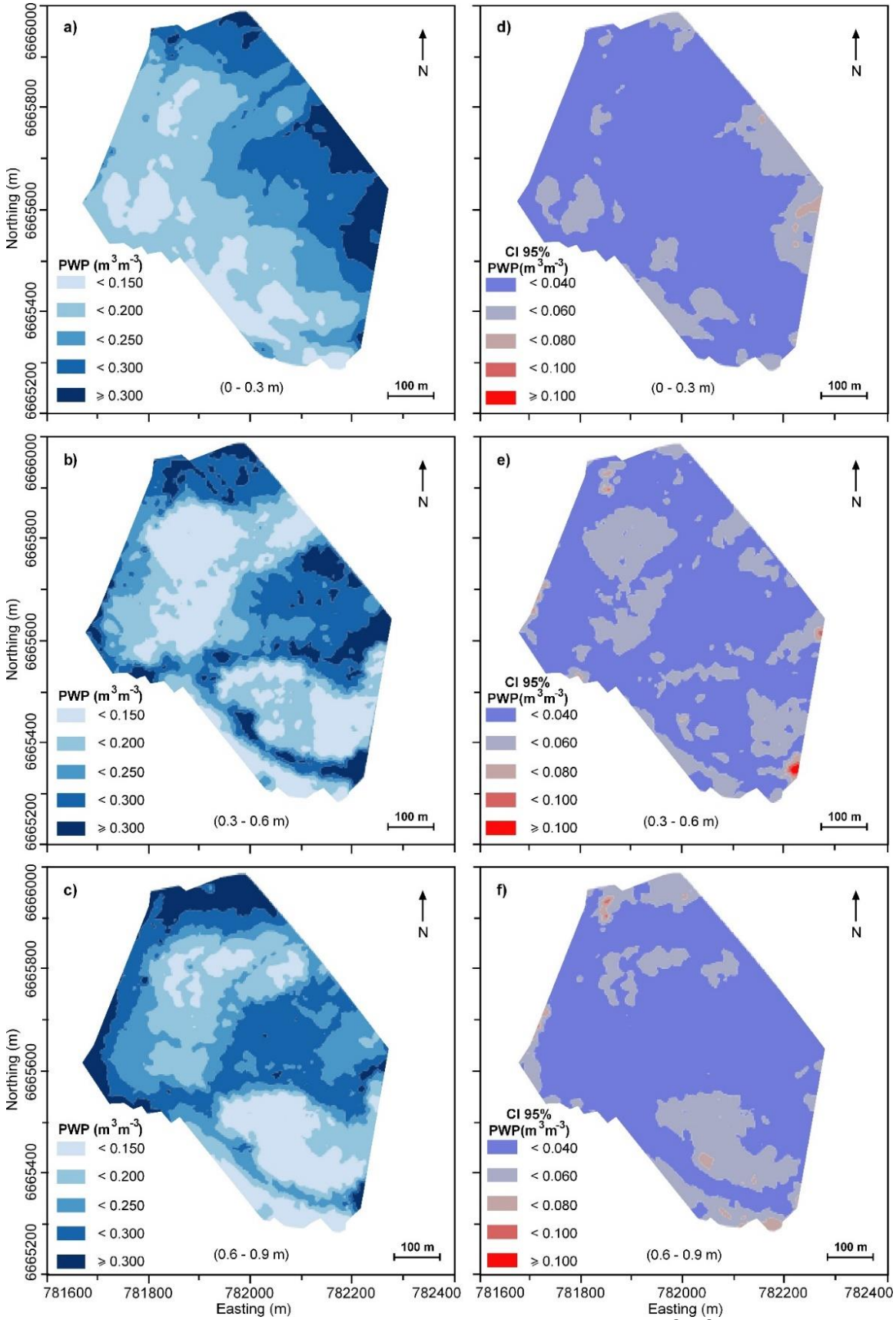


Figure 6.4 Digital soil map (DSM) of predicted permanent wilting point (PWP, $m^3\ m^{-3}$) in the a) topsoil (0-0.3 m), b) subsurface (0.3-0.6 m) and c) subsoil (0.6-0.9 m); and, spatial distribution of 95% confidence intervals in the d) topsoil (0-0.3 m) e) subsurface (0.3-0.6 m) and f) subsoil (0.6-0.9 m).

Figure 6.4a shows the DSM of predicted topsoil PWP using the stepwise-MLR. The spatial distribution of predicted topsoil PWP did not generally conform to that shown in Figures 3a-c for predicted FC. This was because the most statistically significant digital data was TC. As such, the predicted PWP reflected the spatial distribution of TC. The CI was generally small across the whole field, however, with increasing proximity to the field edges, the CI increased to intermediate-small. Figure 6.4b and c shows, respectively, the DSM of predicted subsurface and subsoil PWP, using the stepwise-MLR. The spatial patterns were equivalent to the patterns shown for the FC, however, and as expected the predicted PWP were smaller. The CI was also generally smaller across the whole field compared to the subsurface and subsoil FC CI. These results were attributed to the fact that the RMSE of the stepwise-MLR models developed to predicted PWP had a smaller RMSE as compared with those developed to predict FC.

6.3.7 Digital soil map of AWC

Figure 6.5 shows the DSM of AWC (m^3 m^{-3}) at three depths and predicted by subtracting the predicted PWP shown in Figure 6.4 from predicted FC shown in Figure 6.3. Figure 6.5a shows the DSM of topsoil AWC (> 0.3 m^3 m^{-3}) was largest adjacent to the southern margin of the field. These values of AWC are indicative of large clay content (Gardiner and Miller, 2004) soil, like the Vertosols which characterize the second terraced fan of local alluvium in the study field. The Vertosols herein were deposited in what appears to be a former drainage line located between the Pilliga Sandstone foothill to the south and another foothill in the eastern part of this field. A similar line or area of intermediate-large (0.25-0.3 m^3 m^{-3}) subsurface and subsoil AWC were evident in Figure 6.5b and 5c, respectively, and which show a second drainage line in an easterly direction,

Figure 6.5a also shows the smallest (< 0.15 m^3 m^{-3}) topsoil AWC was interspersed between the Vertosols and indicative of loamy soil textures of the Pilliga Sandstone soil. Figure

6.5b and 6.5c shows the DSM of subsurface and subsoil AWC, respectively, which more clearly align with the EC_a data and which better discern the clay rich Vertosols and the loamy soil of the Pilliga Sandstone. In terms of the latter, and as suggested by (Gardiner and Miller, 2004), these predicted AWC are in line with loamy sand (0.08-0.09 m^3 m^{-3}), sandy loam (0.12-0.14 m^3 m^{-3}), loam (0.13 m^3 m^{-3}).

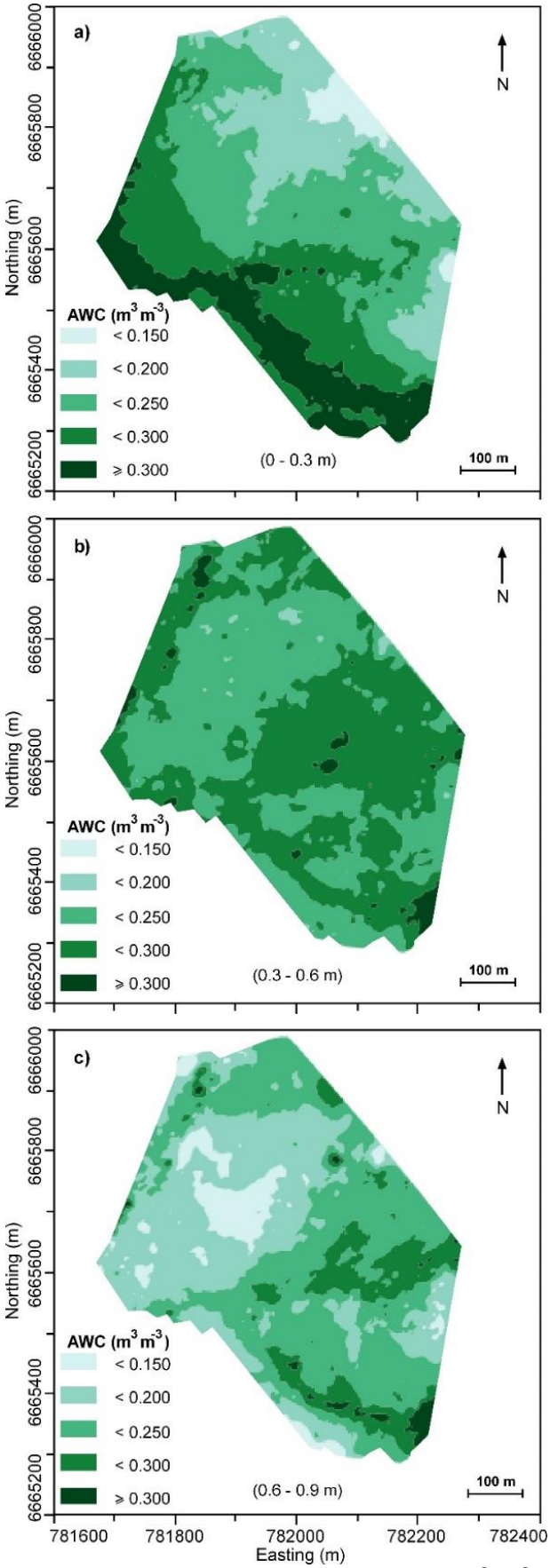


Figure 6.5 Digital soil map of predicted available water content (AWC, m^3 m^{-3}) in a) topsoil (0-0.3 m), b) subsurface (0.3-0.6 m) and c) subsoil (0.6-0.9 m), calculated by difference of the data shown in Figures 3 and 4.

6.4 Conclusions

The aim of this study was to develop DSM of the FC and PWP in the topsoil (0-0.3 m), subsurface (0.3-0.6 m) and subsoil (0.6-0.9 m) across a dryland cotton farm using digital data (i.e., EC_a, γ-ray spectrometry, elevation and positional data). A comparison of two modelling approaches (MLR and stepwise-MLR) showed that for both FC and PWP, the stepwise-MLR was superior at most depths compared to the MLR approach because the adj-R^2 was larger. For example, the adj-R^2 of MLR for prediction of topsoil FC was smaller (0.44) than the stepwise-MLR (0.52) which included only the statistically significant digital data (i.e., 2mPcon, Northing, T.C., 1mPcon and 1mHcon).

The final DSM of FC and PWP and by difference the DSM of AWC for each depth increment were indicative of the two main geomorphological/geological units which characterized the study field. Specifically, the larger AWC (> 0.3 m^3 m^{-3}) were associated with the southern margin of the field and other areas associated with various drainage lines, where the Vertosols are characterized by large clay content (Gardiner and Miller, 2004) and define areas of the field of the second terraced fan of local alluvium. Conversely, the smallest (< 0.15 m^3 m^{-3}) AWC was interspersed between the Vertosols and indicative of the loamy soil textures of the Pilliga Sandstone. Owing to the short scale variation and the fact that the local alluvium and Pilliga Sandsone were close together and the fact that the digital data varied in different ways, large CI values were evident where these areas were juxtaposed.

Nevertheless, the results here are an advance and more accurate than those achieve by Gooley et al. (2014) who created DSM of predicted average rootzone (0-0.9 m) FC and PWP with moderate to weak precision (R^2= 0.56 and 0.46) across two predominantly areas southeast of Warren and Trangie (Macquarie valley). Herein these results were superior because both digital data (i.e., EC_a and γ-ray) were collected proximal on detailed transects spaced only 12 m apart. In Gooley et al (2014), while they used proximal sensed EC_a (i.e., EM38 and EM34-

3), γ-ray and DEM data were acquired from a remote sensed platform (i.e., airborne). In the case of the former, the data was collected on an approximate 0.5-1 km grid, whereas the latter was collected on transects spaced 400 m apart. Moreover, the digital data were collected at different times.

6.5 References

Arshad, M., Li, N., Bella, L. D., & Triantafilis, J. (2020). Field-scale digital soil mapping of clay: Combining different proximal sensed data and comparing various statistical models. *Soil Science Society of America Journal*, *84*(2), 314-330.

Beretta, A. N., Silbermann, A. V., Paladino, L., Torres, D., Bassahun, D., Musselli, R., & García-Lamohte, A. (2014). Soil texture analyses using a hydrometer: modification of the Bouyoucos method. *International Journal of Agriculture and Natural Resources*, *41*(2), 263-271.

Cresswell, H. P., Green, T. W., & McKenzie, N. J. (2008). The adequacy of pressure plate apparatus for determining soil water retention. *Soil Science Society of America Journal*, *72*(1), 41-49.

Gardiner, D. T., & Miller, R. W. (2004). Soils in our environment (pp. 126-165). NJ.

Gooley, L., Huang, J., Page, D., & Triantafilis, J. (2014). Digital soil mapping of available water content using proximal and remotely sensed data. *Soil use and management*, *30*(1), 139-151.

Huang, J., Pedrera-Parrilla, A., Vanderlinden, K., Taguas, E. V., Gómez, J. A., & Triantafilis, J. (2017). Potential to map depth-specific soil organic matter content across an olive grove using quasi-2d and quasi-3d inversion of DUALEM-21 data. *Catena*, *152*, 207-217.

Hong, S. Y., Minasny, B., Han, K. H., Kim, Y., & Lee, K. (2013). Predicting and mapping soil available water capacity in Korea. PeerJ, *1*, e71.

Khongnawang, T., Zare, E., Zhao, D., Srihabun, P., & Triantafilis, J. (2019). Three-Dimensional Mapping of Clay and Cation Exchange Capacity of Sandy and Infertile Soil Using EM38 and Inversion Software. Sensors, *19*(18), 3936.

Leong, E. C., He, L., & Rahardjo, H. (2002). Factors affecting the filter paper method for total and matric suction measurements. *Geotechnical Testing Journal*, *25*(3), 322-333.

Li, N., Zare, E., Huang, J., & Triantafilis, J. (2018). Mapping Soil Cation-Exchange Capacity using Bayesian Modeling and Proximal Sensors at the Field Scale. *Soil Science Society of America Journal*, *82*(5), 1203-1216.

Lin, L.I-K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 255-268.

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.

McBride, G. B. (2005). A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient.. Hamilton, New Zealand: National Institute of Water and Atmospheric Research Ltd. McBride2005. pdf.

McCutcheon, M.C., Farahani, H., Stednick,, J.D., Buchleiter, G.W. & Green, T.R., 2006. Effect of soil water on apparent soil electrical conductivity and texture relationships in a dryland field. *Biosystems Engineering*, *94***,** 19-32.

Minasny, B., McBratney, A. B., & Whelan, B. M. (2005). A. 2 VESPER.

Minasny, B., & McBratney, A. B. (2002). The neuro-m method for fitting neural network parametric pedotransfer functions. *Soil Science Society of America Journal*, *66*(2), 352-361.

Minasny, B., McBratney, A. B., Mendonça-Santos, M., Odeh, I. O. A., & Guyon, B. (2006). Prediction and digital mapping of soil carbon storage in the Lower Namoi Valley. *Soil Research*, *44*(3), 233-244.

Minasny, B., & Hartemink, A. E. (2011). Predicting soil properties in the tropics. *Earth-Science Reviews*, *106*(1-2), 52-62.

Moore, D. S., Kirkland, S. (2007). The basic practice of statistics (Vol. 2). New York: WH Freeman.

Mora-Vallejo, A., Claessens, L., Stoorvogel, J., & Heuvelink, G. B. (2008). Small scale digital soil mapping in Southeastern Kenya. Catena, *76*(1), 44-53.

Muzzamal, M., Huang, J., Nielson, R., Sefton, M., & Triantafilis, J. (2018). Mapping soil particle-size fractions using additive log-ratio (ALR) and isometric log-ratio (ILR) transformations and proximally sensed ancillary data. *Clays and Clay Minerals*, *66*(1), 9-27.

Nagra, G., Burkett, D., Huang, J., Ward, C., & Triantafilis, J. (2017). Field level digital mapping of soil mineralogy using proximal and remote-sensed data. *Soil Use and Management*, *33*(3), 425-436.

Nam, S., Gutierrez, M., Diplas, P., Petrie, J., Wayllace, A., Lu, N., & Muñoz, J. J. (2010). Comparison of testing techniques and models for establishing the SWCC of riverbank soils. *Engineering Geology*, *110*(1-2), 1-10.

Padarian, J., Minasny, B., McBratney, A. B., & Dalgliesh, N. (2014). Predicting and mapping the soil available water capacity of Australian wheatbelt. *Geoderma Regional*, *2*, 110-118.

Rossel, R. V., Behrens, T., Ben-Dor, E., Brown, D. J., Demattê, J. A. M., Shepherd, K. D., ... & Aïchi, H. (2016). A global spectral library to characterize the world's soil. *Earth-Science Reviews*, *155*, 198-230.

Saey, T., Van Meirvenne, M., De Smedt, P., Cockx, L., Meerschman, E., Islam, M. M., & Meeuws, F. (2011). Mapping depth-to-clay using fitted multiple depth response curves of a proximal EMI sensor. *Geoderma*, *162*(1-2), 151-158.

Tenison, K., Wild, J., Madden, E., Draper, V., McMahon, L. (2014). Salinity Training Manual,320 pp. State of New South Wales through the Department of Trade and Investment,Regional Infrastructure and Services, Sydney.

Triantafilis, J., Ward, W. T., & McBratney, A. B. (2001). Land suitability assessment in the Namoi Valley of Australia, using a continuous model. Soil Research, *39*(2), 273-289.

Triantafilis, J., Gibbs, I., & Earl, N. (2013). Digital soil pattern recognition in the lower Namoi valley using numerical clustering of gamma-ray spectrometry data. *Geoderma*, *192*, 407-421.

Tucker, B. M. (1974). Laboratory procedures for cation exchange measurements on soils (No. 23 Tech. Pape).

Wilding, L. P. (1985). Spatial variability: its documentation, accomodation and implication to soil surveys. In Soil spatial variability, Las Vegas NV, 30 November-1 December 1984 (pp. 166-194).

Weller, U., Zipprich, M., Sommer, M., Castell, W. Z., & Wehrhan, M. (2007). Mapping clay content across boundaries at the landscape scale with electromagnetic induction. Soil *Science Society of America Journal*, *71*(6), 1740-1747.

Zhao, D., Li, N., Zare, E., Wang, J., & Triantafilis, J. (2020). Mapping cation exchange capacity using a quasi-3d joint inversion of EM38 and EM31 data. *Soil and Tillage Research*, *200*.

Chapter 7: Comparing traditional and digital soil mapping at a district scale using residual maximum likelihood analysis

7.1 Introduction

Conventional soil mapping uses field observations of morphological (e.g., structure and colour), physical (e.g., soil texture) and chemical (e.g., pH and salinity) properties to classify soil into predefined classification systems. The information is then coupled to aerial photographs for the purpose of extrapolation complemented with the expert judgments of the soil surveyor. Several criticisms have been made to this approach, including: i) the subjectivity of the surveyor in the classification of soil types (Viscarra-Rossel and McBratney 1998), and ii) the choice of morphological descriptors may not reflect soil properties of agricultural value (Triantafilis *et al.* 2000; 2001), or iii) the use of diagnostic horizons and diagnostic characteristics that are useful in agricultural management (Miltenyi *et al.* 2015). In addition, the process is time-consuming and very expensive.

Advances in computing and statistical analysis of data have produced an alternative means to map the soil. Known as digital soil mapping (DSM), it involves the creation of spatial soil information, but rather than using aerial photographs for extrapolation, soil information is coupled with cheaper and easier to acquire ancillary data using spatial and non-spatial models. Two types of ancillary data are being widely used. The first are gamma-ray (γ-ray) spectrometry data, which and it can reflect the soil's parent material (Wilford *et al.* 2001) and, therefore can differentiate soil- texture (Buchanan et al. 2012) and soil types (Triantafilis *et al.* 2013). Gamma-ray spectrometers measure topsoil (0-0.40 m) γ-rays generated during the natural decays of radioactive elements, including potassium (K), uranium (U) and thorium (Th) (Pickup and Marks 2000). Electromagnetic instruments, which measure the soil apparent electrical conductivity (EC_a), have proven useful in mapping soil properties (Doolittle and Brevik, 2014), including clay content (Saey *et al.* 2015), salinity (Huang *et al.* 2015), soil water content (Huang *et al.* 2017a), CEC (Triantafilis *et al.* 2009) and pH (Huang *et al.* 2014a).

While most of the literature describes the use of these data to map soil properties, there is an increasing amount of research which demonstrates how proximally sensed ancillary data can be used to identify soil types and, possibly associated management zones. For example, EC_a data have been used with yield maps (e.g., Anderson-Cook *et al.* 2002), topography (e.g., Fraisse *et al.* 2001), digital elevation models (e.g., Sommer *et al.* 2003) aerial photographs (e.g., Triantafilis *et al.* 2009) and Quickbird imagery (e.g., Guo *et al.* 2013) to identify types or zones at the field level. More recently, Castrignanò *et al.* (2012) combined γ-rays and EC_a data across an 80-ha cropping field in Western Australia and discriminated variations in soil properties (e.g., Colwell K). At a smaller scale, a 5-ha field in Germany, Altdorff and Dietrich (2012) used clustering of γ-ray and EC_a to discern management zones, an approach similar to that applied by Popp *et al.* (2013) and Huang *et al.* (2014b) to hill-slopes with heterogeneous soils. The former derived ecological zones, while the latter generated a map which resembled the local soil distribution of soil series discriminating variation in topsoil clay and pH.

More recently, larger-scale studies have demonstrated the potential of using remote-sensed γ-ray with reconnaissance EC_a data to identify soil landscape units across districts characterised predominantly by modern alluvium interspersed with either aeolian (Zare *et al.* 2016) or metamorphic and igneous landforms (Jing *et al.* 2017). The aim of this research was to further explore the use of this type of data in a DSM approach, in general, and, more specifically to identify a relevant number of soil classes with respect to the management of soil physical and chemical properties. It was also aimed to compare the results against a traditionally derived pedoderm map (McKenzie, 1992). The effects of clustering either in combination or individually airborne γ-ray and proximal EC_a data to map soil profile classes across two geomorphologically different areas in the central-west of New South Wales, Australia was also explored.

7.2 Materials and methods

7.2.1 Study area

The study areas are located in the Macquarie Valley southeast of Trangie and Warren (Figure 7.1a). The dominant land uses are dryland cropping for wheat (*Triticum aestivum L.*) and native pasture for wool and beef as well as irrigation for cotton (*Gossypium hirsutum L.*). The climate is semi-arid (526 mm mean annual rainfall) with a mean monthly temperature maximum of 31.5 ^{o}C in January and a minimum of 2.1 ^{o}C in July.

McKenzie (1992) identified various pedoderms using a traditional mapping approach (Figure 7.1b) where pedoderms are mappable units of soil, either at the land's surface or (partially) buried, which have specific physical characteristics. A pedoderm component distinguished between geomorphic forms (e.g., floodplain and hillslope). As shown in Figure 7.1b, the Trangie Cowal Pedoderm covers three pedoderm components (labelled as 7, 8, and 9) and is the largest in area, characterised by Wilga red-brown profiles (Red Chromosol) and Byron red-brown earth (Red Chromosol). Both have distinct clay content maxima between 0.30 m and 0.80 m. The second largest is Old Alluvium, which is consisted of three pedoderm components, including Meander Plain (labelled as 1), Backplain (labelled as 2) and an unknown pedoderm (labelled as 3). Meander Plain is characterised by the Mitchell class which is well-drained (Red Chromosol) owing to high coarse sand content. The Backplain is more variable, including Mullah-dark grey to black cracking clays (Black Vertosol), Snake-sodic grey cracking clays (Grey Vertosol), and Buddah (Red Vertosol) characterized by high clay. Of the remaining pedoderms, the Gin Gin Hills have two pedoderm components (labelled as 4 and 5) and consist of Gin Gin profiles (Red Dermosol) which are strongly weathered with a uniform to gradational texture profile. The Macquarie Alluvium Pedoderm has one pedoderm component (labelled as 6) and is made of the Macquarie class (Stratic Rudosol), which has minimal development and is characterised by fine sand and silt with a moderately high CEC and a permeable nature. The Contemporary Macquarie Pedoderm has only one pedoderm

component (labelled as 10) and is located along the Macquarie River. Key topsoil (0-0.2 m) and subsoil (0.6-0.8 m) properties are shown in Table 7.1 (McKenzie 1992).

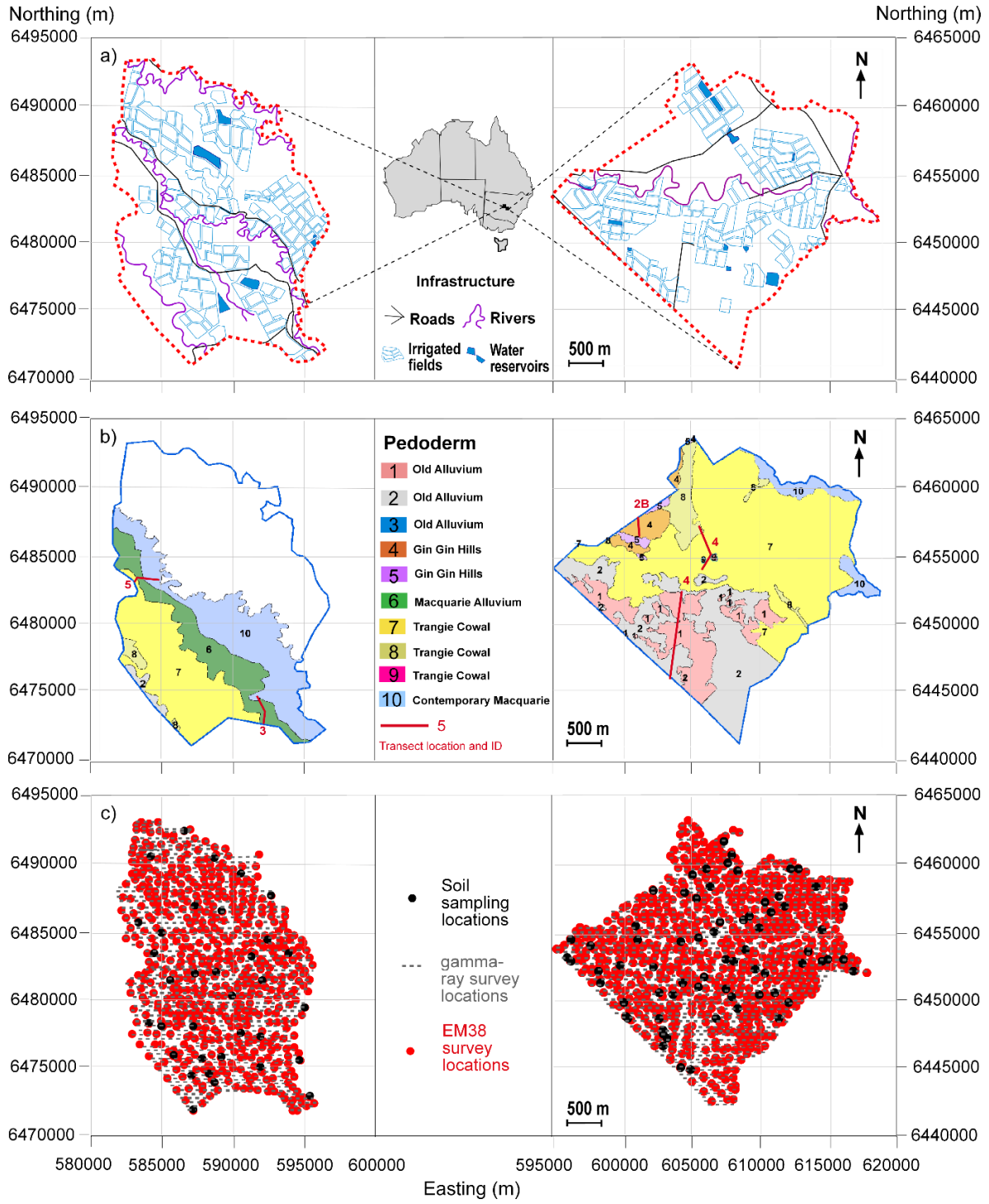


Figure 7.1 a) Infrastructure map of the Trangie and Warren and b) Soil pedoderm maps (after Mckenzie 1992) and c) locations of the EM38 measurements and soil sampling locations. The pedoderm components identified in Figure 7.1b include the following: 1. Meander Plain, 2. Backplain, 3. Infilled Channels, 4. Crests and Slopes, 5. Depressions, 6. Backplain Complex, 7. Alluvial Plain, 8. Depressions, 9. Source Bordering Dunes, 10. Contemporary Macquarie. The white area in the northern half of Warren area corresponds to an area for which no pedodem information was available.

Table 7.1 Medians of some soil properties of the soil classes identified in a conventional soil map (After McKenzie, 1992).

Soil	Mitchell	Mullah	Buddah	Snake	Gin Gin	Ellengerah	Wilga	Byron	Sandhill	Macquarie
Pedoderm Component	Meander Plain	Back plain	Back plain	Back plain	Crests and Slopes	Backplain Complex	Alluvium Plain	Depressions	Source Bordering Dunes	—
Pedoderm	Old Alluvium				Gin Gin Hills	Macq. Alluvium	Trangie Cowal			Contemporary Macquarie
Probable Age (years BP)	130,000 to 25,000				130,000 to 25,000	130,000 to 15,000	25,000 to 15,000			15,000 to 0
0.1 m										
Clay(%)	20	51	46	49	30	43	23	28	5	30
FC(m^3 m^{-3})	0.12	0.27	0.26	—	0.12	0.21	0.2	0.18	—	—
PWP(m^3 m^{-3})	0.07	0.19	0.16	—	0.07	0.16	0.07	0.09	—	—
AWC(m^3 m^{-3})	0.08	0.11	0.15	0.11	0.08	0.08	0.18	0.14	0.16	0.09
pH	6.8	7.7	8.3	9.0	6.2	6.8	7.1	6.6	7.3	6.7
$EC_{1:5}$(dS m^{-1})	0.04	0.10	0.13	0.18	0.04	0.06	0.05	0.06	0.04	0.08
CEC(cmol(+) kg^{-1})	4.0	14.1	19.7	29.8	4.4	13.7	5.7	5.3	2.7	13.0
0.7 m										
Clay (%)	48	51	50	51	41	47	33	42	4	38
FC(m^3 m^{-3})	0.19	0.27	0.28	—	0.14	0.20	0.21	0.23	—	—
PWP(m^3 m^{-3})	0.14	0.21	0.18	—	0.09	0.17	0.10	0.15	—	—
AWC(m^3 m^{-3})	0.08	0.09	0.15	0.16	0.05	0.05	0.16	0.12	0.18	0.11
pH	7.9	8.6	9.1	9.1	6.5	8.5	8.4	8.3	7.8	7.3
$EC_{1:5}$(dS m^{-1})	0.05	0.46	0.53	1.11	0.02	0.22	0.10	0.11	0.02	0.04
CEC(cmol(+) kg^{-1})	12.3	29.0	22.5	32.2	4.8	22.2	18.6	20.3	0.9	25.1

7.2.2 Ancillary data collection and data preparation

Two sources of ancillary data were acquired. Airborne γ-ray spectrometry data wasobtained from the New South Wales Department of Mineral Resources. The flight lines were predominantly east–west, with a spacing of 0.4 km and survey height of 80 m. Within a survey line, measurements were recorded every 60 m. The detector consisted of sodium iodide (NaI) treated with thallium in the form of single crystals of 4 L; three detector packages used, each containing four crystal detectors, giving a total volume of 48 L. The final data set includes 576,520 measurements locations, for each of which data of four windows (i.e., K, U, Th, and total count, TC) are available.

Proximal sensed EM induction data were acquired with an EM38 sensor used in horizontal (EM38h) and vertical (EM38v) modes, providing EC_a measurements to theoretical depths of exploration of 0.75 m and 1.5 m, respectively (McNeill 1990). In all, 755 and 564 measurements were made across Trangie and Warren (Figure 7.1c), respectively, on an approximate 500-m grid in irrigated and 1-km grid in dryland areas. The EM38 meter was

nulled and calibrated according to the user manual (Geonics 2010) prior to each day of the survey. Soil samples were taken at equivalent times. Coordinates were recorded in the Australian Map Grid (AMG94) using a Magellan NavPro5000 GPS.

The γ-ray spectrometry (K, U, Th and TC) and EC_a data (EM38h and EM38v) were interpolated onto a common 100-m grid, using ordinary kriging with a neighbourhood of 90-100 points for interpolation and local exponential variograms using VESPER software (Minasny *et al.* 1999). The ancillary data were standardised using

$$X_s = \frac{X_i - \overline{X}}{std(X)} \qquad (1)$$

where X_i is the i^{th} observation of the ancillary covariate (e.g., K); $\overline{X}$ is the mean of the covariate; $std(X)$ is the standard deviation of the covariate and X_s is the standardised covariate.

7.2.3 Soil sampling and laboratory analysis

Site selection for sampling in Trangie and Warren was based on Response Surface Sampling Design (RSSD) and the module of the EC_e Sampling, Assessment and Prediction (ESAP) software package was used to select locations based on the EC_a data (Lesch *et al.* 2000). In brief, the sampling design involves decorrelate EC_a data using a principal component transformation procedure and then centre and rescale the resulting *m* principal component vectors. This principal component data is then directly compared to a suitable response surface design composed of *n* design level combinations balanced across the *m* principal component vectors. Afterwards, the residual errors are assumed to be spatially correlated and an iterative adjustment in the sample site locations is attempted. The aims of the RSSD are to select a small set of *n* sampling sites to optimise the estimation of the regression parameters for predicting several soil properties with ordinary least squares estimation approach and to minimise the effects of the spatially dependent error structure on the estimation process (Triantafilis and Lesch, 2005).

As a result, 70 sample sites were selected in Trangie and 34 in Warren (Figure 7.1c). Samples were collected at 0.30 m depth increments to a depth of 1.5 m. Here only the samples for 0-0.3 m and 0.6-0.9 m were considered because McKenzie (1992) only reported the soil properties measured at equivalent depths.

All soil samples were air-dried, crushed, and passed through a 2-mm sieve. The particle size fractions were determined using the hydrometer method. For brevity, only the clay content (%) was considered. A pressure plate was used because of its ability to determine FC and PWP (Cresswell et al., 2008). Therefore, it became possible to compare the results directly with those reported by McKenzie (1992), FC was determined at -10 kPa and PWP at -1500 kPa. The samples were left to equilibrate (2 days) before determination of gravimetric moisture (w). Conversion into volumetric moisture (θ) was achieved using bulk density whereby each sample was located with reference to the pedoderm map, and the average value determined by McKenzie (1992) for each was used. AWC was determined as the difference between FC and PWP.

Soil pH was measured form a 1:5 soil: water extract. $EC_{1:5}$(dS m^{-1}) was also determined from this dilution to be consistent with the analysis undertaken by McKenzie (1992). The CEC was similarly determined using Tucker's (1974) method using a mechanical leaching device (Holmgren et al., 1977). The method was used because the Vertosols are alkaline and contain solid-phase carbonates (Loveday *et al.* 1972).

7.2.4 Fuzzy k-means (FKM) analysis

The FKM analysis has been previously described by McBratney *et al.* (1992). In brief, the similarity between an individual i and cluster c is measured to determine how much they are alike in multi-variate space (Bezdek 1981). The best outcome minimizes the objective function $J(\mathbf{M}, \mathbf{C})$:

$$J(\mathbf{M}, \mathbf{C}) = a \sum_{i=1}^{n} \sum_{c=1}^{k} m_{ic}^{\phi} d_{ic}^{2}(X_i, C_c) \qquad (2)$$

where $\mathbf{M} = m_{ic}$ is an $n \times k$ matrix of membership values (n denoting the number of objects), $\mathbf{C} = (C_{cv})$ is a $k \times p$ matrix of class centres (p denotes the number of variables), C_{cv} is the value of the center of class c for variable v, $X_i = (X_{i1},\ldots,X_{ip})^{\mathrm{T}}$ is the vector representing individual i, $C_c = (C_{c1},\ldots,C_{cp})^{\mathrm{T}}$ is the vector representing the centre of class c, and $d_{ic}^{2}(X_i, C_c)$ is the square distance between X_i and C_c according to a distance measure (d_{ic}^{2}). The fuzziness exponent (ϕ) determines the degree of fuzziness. When $\phi = 1$ this is equivalent to the hard partition.

The derivative of J($\mathbf{M}$, $\mathbf{C}$) versus ϕ can also be used to decide on the optimal values of ϕ and k (Bezdek 1981):

$$\frac{\mathrm{dJ}(\mathbf{M}, \mathbf{C})}{\mathrm{d}\emptyset} = \sum_{\mathrm{i=1}}^{\mathrm{n}} \sum_{\mathrm{c=1}}^{\mathrm{k}} \mathrm{m}_{\mathrm{ic}}^{\emptyset} \log(\mathrm{m}_{\mathrm{ic}}) d_{ic}^{2} \qquad (3)$$

Herein, the outcome of $J(\mathbf{M}, \mathbf{C})$ partitioning of the ancillary data into k = 2 to 10 was determined using increments in ϕ of 0.1 and between 1.2 and 2.4. FKM clustering of the kriged and standardised γ-ray spectrometry and EC_a data was conducted using FuzME 3.0 (Minasny and McBratney 2002). Euclidean distance was used given the data have been standardised.

7.2.5 Linear mixed model (LMM)

When soil data has not been collected randomly, it would be inappropriate to fit a linear model by ordinary least squares, since the residuals cannot be treated as independent random variables. Rather, a linear mixed model (LMM) to test the maps and of the form:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{\varepsilon} \qquad (4),$$

can be used, where $\mathbf{y}$ is a $n\times1$ vector of values of the target soil properties (i.e. clay, pH, $EC_{1:5}$ and CEC), $\mathbf{X}$ is a $n\times p$ design matrix, $\boldsymbol{\beta}$ is a $p\times1$ vector of fixed effect coefficients, $\boldsymbol{\eta}$ is a $n\times1$ vector, the elements of which are a realization of a spatially correlated random variable and $\boldsymbol{\varepsilon}$,

is a $n \times 1$ vector the elements of which are a realization of an independent and identically distributed random variable. The elements of the design matrix are predictor covariables and the fixed effects coefficients correspond to these.

The correlated random variable $\boldsymbol{\eta}$ is assumed to be normal and has a mean of zero and variance parameters, with an overall variance, σ^2_η, and a distance parameter for a selected variogram function (e.g., the range of a spherical variogram). The error variable $\boldsymbol{\varepsilon}$ also has zero mean and a variance σ^2_ε. Models of the form in Equation (4) were fitted for target soil variables and with the fixed effects from the class of maximum membership for the FKM clustering of the ancillary variables with k = 2-10. The fitting was done using the LME procedure from the NLME library for the R platform (Pinheiro et al., 2013; R Development Core Team, 2017).

Using this approach, variance parameters for the random effects are first estimated by residual maximum likelihood (REML) and the fixed effects coefficients are then estimated by weighted least squares. The null hypothesis that all class means are equal (where the fixed effects are classes) is tested by the Wald statistic. These methods are described by Lark *et al.* (2006).

7.2.6 Computation of the mean squared prediction error

The sum of the variances of the random effects in the model for k classes as the random effects, $\sigma^2_{\eta,k} + \sigma^2_{\varepsilon,k}$, was treated as the expected value of the variance of the random variable. This approach was used elsewhere to compute values for the variances of design-based sample estimates from the results of model-based analyses Lark (2011). The expected value of the mean squared prediction error for the classification into k classes for some sample size N was computed here using:

$$\sigma^2_{p,C}(N \mid k) = (\sigma^2_{\eta,k} + \sigma^2_{\varepsilon,k})(1+k/N). \qquad (5)$$

Where, $\sigma^2_{p,C}(N \mid k)$ was calculated for the maps for k = 2-10 using the topsoil and subsoil properties of the 104 soil samples across Trangie and Warren and thereby to assist in determining the optimal number of soil profile classes (k) (Huang *et al.* 2014b).

To further compare the FKM clustering with the conventional map of pedoderms and soil profile classes, the map from McKenzie (1992) was used. Because only 93 soil samples were present in the pedoderm map and the number of identified pedoderm components was 7, the mean square prediction error (i.e. $\sigma^2_{p,C}(N \mid k)$) for this map was calculated using only these 93 samples for the conventional soil map. Similarly, only these 93 samples were used for the selected k of the DSM (see Figure 7.4c). To evaluate the performance of different ancillary data, clustering using different combinations of ancillary data was also tested (i.e. EM only, and γ-ray only).

7.3 Results and discussion

7.3.1 Summary statistics of soil sampling data

Table 7.2 shows summary statistics for the soil samples collected herein and the Pearson correlation coefficients (r) between topsoil (0-0.30 m) and subsoil (0.60-0.90 m) properties and ancillary data. In terms of the topsoil properties, radioelement K was found to be correlated with $EC_{1:5}$ (r = 0.30) and CEC (0.27) while EC_a from EM38v was correlated with clay (0.58), FC (0.53), PWP (0.59), pH (0.46), CEC (0.67) and $EC_{1:5}$ (0.30). Similar correlations were found with the EM38h although they generally were slightly weaker, except for $EC_{1:5}$ which was slightly stronger. Equivalent correlations were achieved when considering the radioelements. Similar correlations were also identified between EC_a, K and TC and subsoil properties (0.6-0.9 m). It should be noted that coefficients of variation (CV) of soil $EC_{1:5}$ and CEC were relatively high compared to clay and pH.

Table 7.2 Summary statistics of soil samples from the topsoil (0-0.3 m) and subsoil (0.6-0.9 m).

	0-0.3 m							0.6-0.9 m						
	Clay (%)	FC (m^3 m^{-3})	PWP (m^3 m^{-3})	AWC (m^3 m^{-3})	pH	$EC_{1:5}$ (dS m^{-1})	CEC (cmol(+) kg^{-1})	Clay (%)	FC (m^3 m^{-3})	PWP (m^3 m^{-3})	AWC (m^3 m^{-3})	pH	$EC_{1:5}$ (dS m^{-1})	CEC (cmol(+) kg^{-1})
N	104	80	80	80	104	104	104	104	84	84	84	104	104	104
Min	15.4	0.13	0.07	0.05	5.6	0.02	5.2	9.2	0.10	0.16	0.04	5.2	0.03	4.8
Mean	42.0	0.24	0.15	0.09	8.0	0.28	18.6	46.3	0.27	0.19	0.08	8.6	0.38	19.7
Max	63.7	0.34	0.24	0.19	9.5	3.38	35.2	74.2	0.41	0.30	0.16	10.3	3.92	33.4
Medium	43.9	0.24	0.15	0.09	8.0	0.15	18.6	48.6	0.27	0.80	0.07	8.8	0.22	18.9
Std. Dev.	11.0	0.05	0.04	0.02	0.9	0.43	6.7	11.6	0.05	0.04	0.08	0.9	0.49	6.3
Skewness	-0.29	0.10	0.28	1.30	-0.27	4.92	0.04	-0.36	0.24	-0.14	0.98	-1.1	4.20	-0.01
Kurtosis	-0.51	-0.60	-0.30	3.44	-0.63	29.66	-0.75	-0.05	0.12	-0.27	3.63	2.3	25.69	-0.60
CV (%)	26.3	19.60	26.18	26.29	10.9	154.57	36.2	25.1	19.58	22.56	27.37	10.7	130.19	31.9
Pearson Correlation coefficient (*r*)														
K (%)	-0.21	-0.14	-0.33	0.26	-0.29	-0.30	-0.27	-0.38	-0.36	-0.29	-0.30	-0.02	-0.35	-0.33
U (ppm)	-0.07	0.01	-0.14	0.21	-0.31	-0.18	-0.11	-0.18	-0.27	-0.16	-0.20	-0.11	-0.18	-0.17
Th(ppm)	-0.02	-0.05	-0.07	0.21	-0.18	-0.10	-0.03	-0.14	-0.22	-0.17	-0.20	-0.07	-0.19	-0.15
TC (cps)	-0.11	-0.02	-0.21	0.29	-0.29	-0.22	-0.15	-0.24	-0.28	-0.19	-0.30	-0.05	-0.27	-0.21
EM38h (mS m^{-1})	0.53	0.53	0.58	0.12	0.41	0.35	0.64	0.50	0.58	0.62	0.19	0.28	0.48	0.54
EM38v (mS m^{-1})	0.58	0.53	0.59	0.08	0.46	0.30	0.67	0.53	0.61	0.66	0.18	0.37	0.46	0.59

7.3.2 Spatial distribution of proximally sensed data

Figure 7.2 shows the spatial distributions of the ancillary data. Figure 7.2a shows the variation for radioelement K. It is evident that intermediate-large K (1.0-1.4%) defines areas associated with the Trangie Cowal. Similar values characterise the Macquarie Alluvium and the Contemporary Macquarie. In drainage ways such as these, this suggests the presence of weathering alkali feldspars and micas which csn be responsible of a high K concentration (Wilford *et al.* 1997). In the Trangie Cowal Pedoderm and the Contemporary Macquarie, this is because the K has not been weathered from the soil and owing to the relatively young soil profile classes which have ages of approximately 25,000 to 15,000 and 15,000 to 0 years BP, respectively (McKenzie 1992).

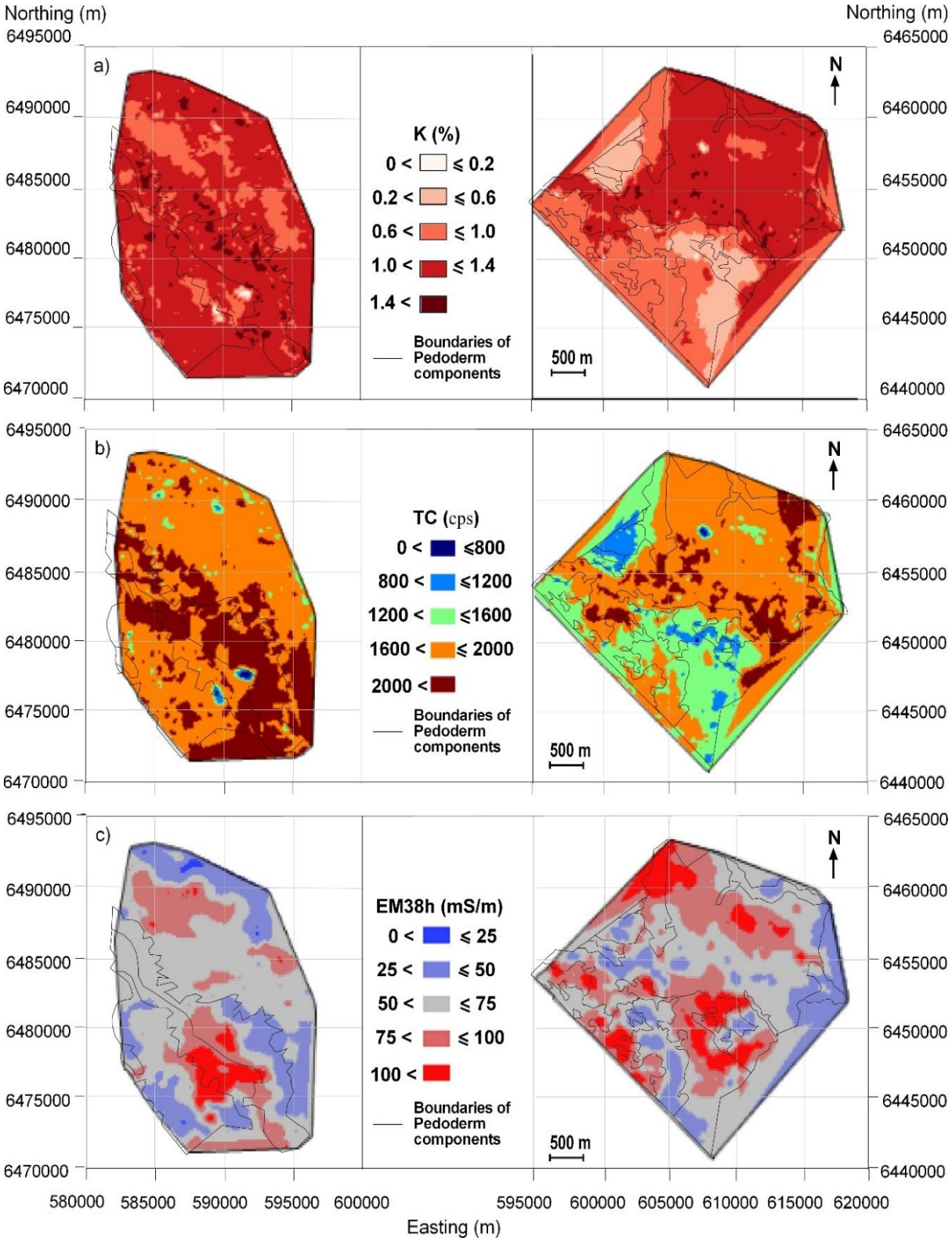


Figure 7.2 Spatial distribution of gamma-ray (γ-ray) spectrometry data including; a) potassium (K, %) b) total count (TC, counts per second), and electrical conductivity (EC_a, mS m^{-1}) of EM38h.

Conversely, intermediate-small K (0.2-0.6 %), characterised the Old Alluvium (Backplain) and Gin Gin Hills (Crests and Slopes), while intermediate K (0.6-1.0 %) defined the Old Alluvium (Meander Plain) and non-mapped areas in Warren. The smaller K can be

attributed to the leaching of soluble potash by chemical weathering, over a long period of time. This is consistent with the fact these pedoderms are older and have formed from deposits of approximately 130,000 to 25,000 years BP (McKenzie 1992). Figure 7.2b shows that the variation of TC (cps) is similar.

Figure 7.2c shows EM38h EC_a variation. Intermediate-large to large EC_a (> 75 mS m^{-1}) defined the northern part of the Trangie Cowal and Old Alluvium (Backplain) in Trangie. Similar values characterised the central south of Warren, including the Trangie Cowal, Macquarie Alluvium, and Contemporary Macquarie. These areas most likely represent the clayier Snake soil profile class (Old Alluvium, Back Plain), which are older and contain large accumulations of soluble salts, as a function of cyclical and aeolian deposition. Conversely, intermediate-small EC_a (25-50 mS m^{-1}) characterised the Old Alluvium (Meander Plain), the Gin Gin Hills (Crests and Slopes) in the Trangie area and associated with Trangie Cowal and Contemporary Macquarie in the Warren area.

7.3.3 FKM analysis

Figure 7.3 shows the plot of ϕ versus $-dJ(\mathbf{M}, \mathbf{C})/d\phi$. For any given value of k, McBratney and Moore (1985) suggest that ϕ is optimal when $-dJ(\mathbf{M}, \mathbf{C})/d\phi$ is a maximal. In most cases (i.e. k = 5-10) this occured when ϕ = 1.6. Given these results, ϕ = 1.6 was selected for calculating $\sigma^2_{p,C}$ of the various physical (i.e. clay) and chemical (i.e. pH, $EC_{1:5}$ and CEC) properties for k = 2 to 10.

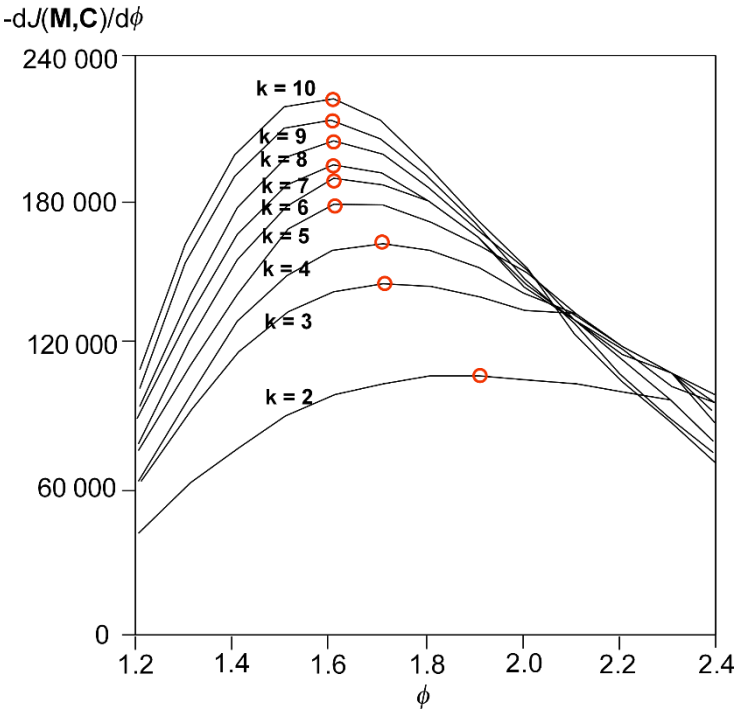


Figure 7.3 Fuzziness performance index (FPI), normalized classification entropy (NCE) and -dJ/dϕ of FKM clustering for ϕ = 1.2-2.4 when Euclidean distance was used. The red circles in Figure 7.3 illustrate the maximum of the -dJ/dϕ for each *k* value.

7.3.4 Digital soil maps

Figure 7.4a shows the DSM when *k* = 3. Class 3A covered the greatest area and coincides with the Trangie Cowal and Contemporary Macquarie Pedoderms. Table 7.3 shows that it was characterised by the largest radioelements concentration (e.g., K = 1.2 %; TC = 1,961 cps) and smallest EC_a (e.g., EM38h = 55 mS m^{-1}). Class 3C was second largest and was associated with Trangie Cowal and Macquarie Alluvium. Small radioelement (e.g., K = 0.7; TC = 1,382 cps) and intermediate EC_a (e.g., EM38h = 75.1 mS m^{-1}) characterise the class. The smallest in extent was 3B, defined by intermediate-large radioelement (e.g., K = 1.1 %; TC = 1,813 cps) and large EC_a (e.g., EM38h = 86 mS m^{-1}). It coincides with Old Alluvium and Gin Gin Hills.

Figure 7.4b shows the FKM clustering result when *k* = 6. The main difference was 3A divides into 6A and 6D, the latter having smaller radioelements (e.g., K = 1.2 %; TC = 1,905 cps) and EC_a (e.g., EM38h = 47.8 mS m^{-1}). Figure 7.4b also shows that 3C divides into 6C and 6E, with the latter characterising subtle variations in the Trangie Cowal; 6E having larger

radioelements (e.g., K = 1.0 %; TC = 1,733 cps) and EC_a (e.g., EM38h = 127 mS m^{-1}). Class 3B divided essentially into 6B and 6F; with 6B defining the Old Alluvium (Back Plain) with small radioelements concentrations (e.g., K = 0.6 %; TC = 1,266 cps) but a large EC_a (e.g., EM38h = 108 mS m^{-1}).

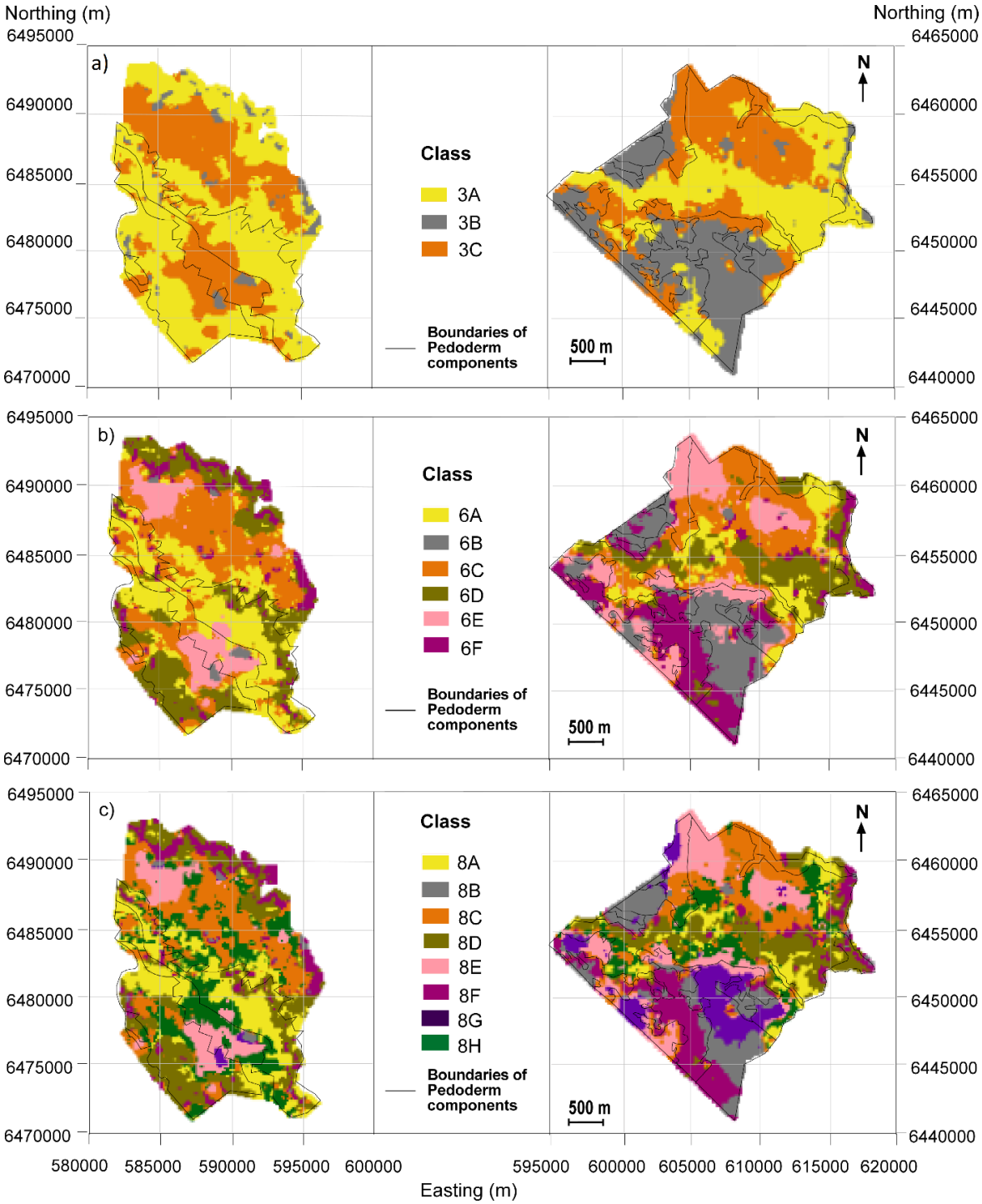


Figure 7.4 Spatial distribution of FKM classes when *k* = a) 3, b) 6 and c) 8 using γ-ray spectrometry and EC_a data, respectively. Note: Black lines show the borders of soil pedoderm components (after McKenzie 1992).

Table 7.3 Euclidean centroid values of proximally sensed ancillary data clustered using FKM when *k* = 3, 6 and 8. Note: centroids shown for K (%), U (ppm), Th (ppm), TC (cps-counts per second), EM38h and EM38v, respectively.

	Numbers	Centroid values
k = 3		K (%), U (ppm), Th (ppm), TC, EM38h, EM38v
3A	21,727	1.2, 2.1, 7.0, 1,961, 55, 66
3B	9,595	0.7, 1.6, 5.5, 1,382, 75, 91
3C	17,257	1.1, 1.9, 6.9, 1,813, 86, 106
k = 6		
6A	9,329	1.3, 2.3, 7.6, 2,106, 65, 82
6B	4,777	0.6, 1.5, 5.3, 1,266, 87, 108
6C	12,288	1.1, 1.9, 6.9, 1,832, 74, 91
6D	9,300	1.2, 2.0, 6.7, 1,905, 48, 54
6E	6,700	1.0, 1.9, 6.7, 1,733, 103, 127
6F	6,185	0.9, 1.7, 5.8, 1,556, 54, 61
k = 8		
8A	6,491	1.3, 2.4, 7.6, 2,133, 60, 74
8B	3,690	0.6, 1.5, 5.2, 1,259, 66, 79
8C	9,213	1.1, 1.9, 6.8, 1,781, 71, 88
8D	8,094	1.2, 2.1, 6.8, 1,920, 48, 54
8E	5,977	1.0, 1.9, 6.8, 1,748, 101, 125
8F	5,394	1.0, 1.7, 6.0, 1,628, 51, 57
8G	2,790	0.6, 1.5, 5.5, 1,348, 108, 135
8H	6,930	1.2, 2.1, 7.3, 1,980, 78, 97

The delineation of the Old Alluvium into its two components can be explained by considering the ancillary data and their response to various soil properties which characterised the different DSM classes. In terms of the γ-ray data, the Mitchell class of the Meander Plain has a small clay activity (i.e. 23-30 cmol(+) kg^{-1}) with kaolin and illite as dominant minerals (McKenzie 1992). This was consistent with 6F, which represented the Meander Plain, which had intermediate radioelement K (0.9%) as a function of the co-dominant illite. By comparison, 6B that was coincident with the Back Plain had much smaller K (0.6%). This appears consistent with the Mullah Grey profile class. Whilst it had larger activity (i.e. 69.9 cmol(+) kg^{-1}), the dominant mineral was smectite with kaolin and illite sub-dominant.

The EC_a data were also instructive because a larger EM38v EC_a (i.e. 108 mS m^{-1}) was consistent with a larger subsoil $EC_{1:5}$ (i.e. 3.12 dS m^{-1}) of the Back Plain Mullah soil profile class. Conversely, the Mitchell class, that characterised the Meander Plain, had a very small $EC_{1:5}$ (0.03 dS m^{-1}) which was reflected by much smaller EM38v (i.e. 61 mS m^{-1}) by comparison. The $EC_{1:5}$ and EC_a were also consistent with the occurrence of coarser bedload

deposits with low activity clays of the Mullah class and the finer over bank deposits which have developed into cracking clays of the Mitchell class, respectively (McKenzie, 1992).

Figure 7.4c shows the DSM for k = 8. Herein, there was a further division of class 6B into class 8B and 8G, with both associated with the Old Alluvium (Back Plain) in Trangie. Class 6A also divided into classes 8A and 8H. It was fairly apparent that whilst Figure 7.4a-c represented to some extent some of the larger pedoderms and components, it was difficult to understand on its own given the Trangie Cowal had now been divided into 5 different zones (i.e. 8A, C, D, E and H). However, this seemed to be consistent with the fact that of all the pedoderms identified by McKenzie (1992), the Trangie Cowal encompassed the most soil profile classes (i.e. Wilga Non-calcic, Wilga Calcic and Byron).

Because it was difficult to qualitatively compare the DSM approach with the traditional pedoderm approach mainly based on visual inspection, the optimal number of soil classes was determined for the DSM approach and used various topsoil and subsoil properties to quantitatively evaluate the mapping accuracies of the two methods. This will to be discussed in the following sections.

7.3.5 Determining the optimal classes

Table 7.4 shows the results of the calculated mean squared prediction error (i.e. $\sigma^2_{p,C}$) of the estimated class means for each of the k = 2 to 10 DSM maps using both γ-ray and EC_a data. Table 7.4a shows the results for topsoil properties (0-0.3 m). In terms of clay, $\sigma^2_{p,C}$ was minimal (75.50 $\%^2$) when k = 8. This was also the case for pH (0.66) and CEC (28.12 [cmol(+) $kg^{-1}]^2$). On the other hand, $EC_{1:5}$ was a minimal (2.68 [dS $m^{-1}]^2$) for k = 4, AWC (0.00084 ([m^3 $m^{-3}]^2$) for k = 7 and FC and PWP (0.00146 [m^3 $m^{-3}]^2$ and 0.00132 ([m^3 $m^{-3}]^2$, respectively) for k = 5. The next best DSM was k = 5 because it account for clay (77.82 $\%^2$), FC (0.00159 [m^3 $m^{-3}]^2$),

PWP (0.00143 [m^3 m^{-3}]^2) and CEC (28.97 [cmol(+) kg^{-1}]^2), while AWC (0.00087 [m^3 m^{-3}]^2) was best accounted for by k = 4, pH (0.67) by k = 6 and EC_e (2.74 [dS m^{-1}]^2) by k = 3.

Table 7.4b shows the equivalent results for $\sigma^2_{p,C}$ with respect to the measured subsoil (0.6-0.9 m) properties. In terms of physical properties, the minimum $\sigma^2_{p,C}$ for the clay content (98.81 %^{2}) was obtained for k = 7, while FC and PWP were both a minimum (0.00126 [m^3 m^{-3}]^2) when k = 6 and AWC (0.000420 [m^3 m^{-3}]^2) when k = 4. For the chemical soil properties, subsoil pH (0.80) was a minima when k = 3, with EC_e (3.86 [dS m^{-1}]^2) and CEC (28.06 [cmol(+) kg^{-1}]^2) a minima when k = 8. The next best $\sigma^2_{p,C}$ for clay (100.36 %^{2}), FC (0.00194 [m^3 m^{-3}]^2), PWP (0.00132 [m^3 m^{-3}]^2), AWC (0.000428 [m^3 m^{-3}]^2), was 7, 8, 9 and 6, respectively, while for pH (0.85), $EC_{1:5}$ (3.90 [dS m^{-1}]^2) and CEC (28.27 [cmol(+) kg^{-1}]^2) it was k = 6.

Table 7.4 Mean square prediction errors ($\sigma^2_{P,C}$) for FKM classification using γ-ray spectrometry and EM data calculated using 104 soil samples in the a) topsoil (0-0.3 m) and b) subsoil (0.6-0.9 m), respectively. Note: Values in bold show the minima $\sigma^2_{P,C}$ values that can be used as optimal soil classes; "–" indicated that the linear mixed model estimated contain extremely large error and was not used for estimate soil properties.

a) k	Clay (%^{2})	FC ([m^3 m^{-3}]^2)	PWP ([m^3 m^{-3}]^2)	AWC ([m^3 m^{-3}]^2)	pH	$EC_{1:5}$ ([dS m^{-1}]^2)	CEC ([cmol(+) kg^{-1}]^2)	No. of minima $\sigma^2_{P,C}$
2	–	0.00205	0.00158	0.00091	0.69	0.184	42.48	0
3	86.58	0.00181	0.00160	0.00088	–	0.184	32.91	0
4	84.29	0.00170	0.00153	**0.00087**	0.68	**0.176**	31.99	2
5	**77.82**	**0.00146**	**0.00132**	0.00089	0.68	0.185	**28.97**	4
6	78.79	0.00164	0.00142	0.00091	**0.67**	0.218	29.70	1
7	80.93	0.00171	0.00149	**0.00084**	0.77	0.207	30.34	1
8	**75.50**	**0.00159**	**0.00143**	0.00094	**0.66**	0.221	**28.12**	5
9	123.37	0.00230	0.00187	0.00087	0.78	0.188	46.24	0
10	119.67	0.00243	0.00195	0.00094	0.77	0.187	44.97	0

b) k	Clay (%^{2})	FC ([m^3 m^{-3}]^2)	PWP ([m^3 m^{-3}]^2)	AWC ([m^3 m^{-3}]^2)	pH	$EC_{1:5}$ ([dS m^{-1}]^2)	CEC ([cmol(+) kg^{-1}]^2)	No. of minima $\sigma^2_{P,C}$
2	120.52	–	0.00183	0.000423	0.92	0.225	34.60	0
3	**102.27**	0.00224	0.00143	0.000438	**0.80**	0.223	32.97	2
4	109.98	0.00219	0.00152	**0.000420**	0.86	0.211	30.33	1
5	113.47	–	–	0.000430	0.98	0.213	29.99	0
6	100.36	**0.00126**	**0.00126**	**0.000428**	**0.85**	**0.195**	**28.27**	6
7	**98.81**	0.00205	–	0.000435	1.02	0.201	29.40	1
8	99.67	**0.00194**	0.00132	0.000448	**0.88**	**0.192**	**28.06**	4
9	118.80	–	–	0.000480	0.91	0.236	38.51	0
10	125.25	0.00327	–	0.000467	0.93	0.233	40.40	0

There was no consistent optimal value for *k* to account for the variability of the different soil physical and chemical properties in the DSM approach. This contradicts the results of Huang *et al.* (2014b) where consistent optimal *k* was identified for clay content and pH, however, they only considered the topsoil at the field scale. Nevertheless, Zare *et al.* (2016) found that topsoil (0-0.3 m) and subsoil (0.9-1.2 m) physical (i.e., clay content) and chemical (i.e. pH, $EC_{e,}$ and CEC) properties were best accounted for by a DSM approach consisting of the same number of classes (i.e. k = 4) across an equivalent area but near Bourke, NSW Australia. In that case study, the landscape was not as complex, given it was predominantly alluvial, with less than 25% of the area characterised by aeolian sand deposits. Herein, 8 classes minimised the $\sigma^2_{p,C}$ of 9 topsoil and subsoil properties while 6 classes only minimised the $\sigma^2_{p,C}$ of 6 properties (see Table 7.4), 8 classes were selected for the following interpretation and comparison with the traditionally derived pedoderm map of Mckenzie (1992).

Table 7.5 Mean squared prediction error of soil properties in the a) topsoil (0-0.3 m) and b) subsoil (0.6 m-0.9 m) calculated for conventional Pedoderm map ($(\sigma^2_{P,C}(93|7)$) and for digital soil mapping (DSM) ($(\sigma^2_{P,C}(93|8)$) using FKM clustering with different combinations of ancillary data. Note: Values in bold show the smallest $\sigma^2_{P,C}$ value for each soil property.

a)	$\sigma^2_{P,C}$	Conventional Pedoderm mapping	γ–ray + EM	EM	γ–ray
	Clay (%^{2})	105.64	**68.83**	92.52	111.91
	FC (%^{2})	0.00215	0.00167	**0.00162**	0.00236
	PWP (%^{2})	0.0013	**0.00100**	0.00120	0.00160
	AWC (%^{2})	**0.00053**	0.00057	0.00061	0.00053
	pH	0.72	0.72	**0.69**	0.84
	$EC_{1:5}$ ([dS m^{-1}]^2)	0.193	0.241	**0.187**	0.189
	CEC ([cmol(+) kg^{-1}]^2)	39.25	**27.13**	27.42	43.89
	No. of minima $\sigma^2_{P,C}$	1	**3**	**3**	0
b)	$\sigma^2_{P,C}$	Conventional Pedoderm mapping	γ–ray + EM	EM	γ–ray
	Clay (%^{2})	132.36	**98.16**	108.92	113.16
	FC (%^{2})	0.0026	**0.00210**	0.00230	_
	PWP (%^{2})	0.0019	**0.00130**	0.00140	0.04100
	AWC (%^{2})	**0.00041**	0.00051	0.00055	0.00052
	pH	**0.88**	0.92	0.90	1.00
	$EC_{1:5}$ ([dS m^{-1}]^2)	0.234	**0.193**	0.203	0.227
	CEC ([cmol(+) kg^{-1}]^2)	35.41	29.26	**29.15**	39.59
	No. of minima $\sigma^2_{P,C}$	2	**4**	1	0

7.3.6 Comparison of mean squared prediction error between conventional soil mapping and DSM

Table 7.5 shows the result of the mean square prediction error ($\sigma^2_{p,C}$) for the conventional pedoderm map (Figure 7.1b) and the DSM (Figure 7.4c) developed using both γ-ray and EM data and considering k = 8. It can be assumed that the comparison between k = 8 generated by the DSM approach with 7 classes identified by the traditional pedoderm approach was reasonable because these classes were generated independently and objectively.

In general, the DSM approach, where FKM was used to cluster the γ-ray spectrometry and EC_a data, either in combination or individually, performed better than the conventional pedoderm map for most of the topsoil and subsoil properties. With respect to clay, the DSM approach considering the clustering of both the γ-ray and EC_a data was optimal given the $\sigma^2_{P,C}$ was smaller for the topsoil (68.83 $\%^2$) and subsoil (98.16 $\%^2$) as compared with the pedoderm map (respectively, 105.64 $\%^2$ and 132.36 $\%^2$). This was also the case for topsoil FC (0.00167 $[m^3\ m^{-3}]^2$) and PWP (0.00100 $[m^3\ m^{-3}]^2$) and subsoil FC (0.00210 $[m^3\ m^{-3}]^2$) and PWP (0.00130 $[m^3\ m^{-3}]^2$). In both cases, AWC was minimised considering the traditional pedoderm map.

However, this was not the case for pH. In terms of topsoil pH, the pedoderm map performed as well ($\sigma^2_{P,C}$ = 0.72) as the DSM developed using both γ-ray and EC_a but outperformed (0.88) the DSM (0.92) for subsoil pH. The results with respect to $EC_{1:5}$ were a little more equivocal as the pedoderm map outperformed the DSM for predicting topsoil $EC_{1:5}$ but underperformed the DSM for predicting subsoil $EC_{1:5}$. This can be attributed to the fact that soil salinity needed to take account of the texture, however, here only the $EC_{1:5}$ data was considered. The results achieved for CEC where similar to clay, given that the $\sigma^2_{P,C}$ for the DSM approach was smaller for the topsoil (27.13 $[cmol(+)\ kg^{-1}]^2$) and subsoil (29.26 $[cmol(+)\ kg^{-1}]^2$) as compared with the pedoderm map (respectively, 39.25 $[cmol(+)\ kg^{-1}]^2$ and 35.41 $[cmol(+)\ kg^{-1}]^2$).

7.3.7 Comparison of DSM results using different combination of ancillary data

Correspondingly, the performance of DSM results was compared using different ancillary data. As shown in Table 7.5a, when minimising topsoil clay the DSM generated by clustering the γ-ray and EM (68.83 $\%^2$) data was the best followed by using the EM (92.52 $\%^2$) data, with the pedoderm map (105.64 $\%^2$) better than the DSM produced by clustering γ-ray (111.91 $\%^2$) only. Equivalent results were achieved when considering the topsoil PWP and CEC. In two instances, topsoil pH (0.692) and $EC_{1:5}$ (0.187 $[dS\ m^{-1}]^2$) would be best predicted using the DSM developed from the EM data only. The only topsoil property for witch clustering γ-ray data outperformed the other ancillary data was for AWC (0.00053 $[m^3\ m^{-3}]^2$).

When considering the subsoil clay data (Table 7.5b), the γ-ray and EM (98.16 $\%^2$) was the best followed by EM (108.92 $\%^2$), with the γ-ray (111.91 $\%^2$) only outperforming the pedoderm map (132.36 $\%^2$). Similar results were evident for subsoil FC and PWP. The only time the EM only DSM outperformed the γ-ray and EM DSM was for pH (0.90^2) and CEC (29.15 $[cmol(+)\ kg^{-1}]^2$), with the γ-ray only derived DSM never out predicting the γ-ray and EM DSM. The results achieved here suggested that it was better to use both γ-ray and EM data for predicting most topsoil and subsoil properties. However, given that the $\sigma^2_{P,C}$ values were very similar, EM data alone could be sufficient for farmers and land managers to map and manage the soils across both study areas.

7.3.8 Soil management at the district scale using DSM

To better understand the DSM of k = 8 generated from clustering of the γ-ray and EC_a data, Table 7.6 shows the predicted mean and standard deviation of the topsoil and subsoil properties for the classes using DSM. When comparing these results, it can be done more-or-less with reference to the median of topsoil and subsoil properties of the soil profile classes

identified by McKenzie (1992) and shown in Table 7.1. The discussion was limited to a couple of pedoderms and the soil profile classes therein.

With respect to the Old Alluvium pedoderm, McKenzie (1992) identified four profile classes. The Mitchell class defined the Meander Plain component. It was characterised by a texture contrast profile (i.e. duplex) and indicated by the sharp difference in topsoil (20 %) and subsoil (48 %) clay. By way of comparison 8F, which characterised a similar area in Figure 7.4, was characterised by topsoil (33 %) and subsoil (44 %) clay. Whilst not definitive, this indicated some consistency between these soil properties. This was particularly the case for subsoil FC (0.19 m^3 m^{-3}), PWP (0.14 m^3 m^{-3}), AWC (0.08 m^3 m^{-3}), pH (7.9), $EC_{1:5}$ (0.05 dS m^{-1}) and CEC (12.3 cmol(+) kg^{-1}) for the Mitchell class and 8F, which were equivalent for subsoil FC (0.25 m^3 m^{-3}), PWP (0.18 m^3 m^{-3}), AWC (0.07 m^3 m^{-3}), pH (8.1), $EC_{1:5}$ (0.15 dS m^{-1}) and CEC (15.8 cmol(+) kg^{-1}). There were also some similarities between the topsoil and subsoil properties of the three Back Plain component classes identified by McKenzie (1992), including the Mullah, Buddah and Snake, and the DSM classes 8B, 8E and 8G which coincided with the area delineated as the Back Plain.

Table 7.6 Predicted mean and standard deviation values of soil properties for different classes when k = 8.

Class	8A	8B	8C	8D	8E	8F	8G	8H
0-0.3 m								
Clay (%)	48 (±4)	46 (±5)	47 (±4)	32 (±4)	47 (±4)	33 (±4)	49 (±4)	46 (±4)
FC (m^3 m^{-3})	0.25 (±0.02)	0.22 (±0.03)	0.26 (±0.02)	0.21 (±0.02)	0.29 (±0.02)	0.20 (±0.02)	0.27 (±0.02)	0.25(±0.02)
PWP(m^3 m^{-3})	0.16 (±0.02)	0.14 (±0.03)	0.16 (±0.02)	0.11 (±0.02)	0.18 (±0.02)	0.12 (±0.02)	0.18 (±0.02)	0.13 (±0.02)
AWC (m^3 m^{-3})	0.09 (±0.02)	0.08 (±0.02)	0.09 (±0.01)	0.10 (±0.01)	0.10 (±0.01)	0.08 (±0.01)	0.09 (±0.01)	0.11 a(±0.01)
pH	7.5 (±0.4)	8.0 (±0.5)	8.3 (±0.4)	7.5 (±0.3)	8.5 (±0.3)	7.6 (±0.3)	8.2 (±0.4)	7.6 (±0.3)
$EC_{1:5}$ (dS m^{-1})	0.14 (±0.20)	0.42 (±0.26)	0.40 (±0.21)	0.18 (±0.18)	0.38 (±0.18)	0.24 (±0.19)	0.72 (±0.21)	0.17 (±0.19)
CEC (cmol(+) kg^{-1})	20.3 (±2.3)	16.9 (±3.2)	22.8 (±2.6)	12.9 (±2.2)	21.6 (±2.2)	13.0 (±2.3)	24.9 (±2.3)	19.2 (±2.5)
0.6-0.9 m								
Clay (%)	45 (±4)	53 (±6)	50 (±5)	35 (±4)	50 (±4)	44 (±4)	53 (±5)	49 (±4)
FC (m^3 m^{-3})	0.25 (±0.02)	0.25 (±0.03)	0.28 (±0.03)	0.21 (±0.02)	0.30 (±0.02)	0.25 (±0.02)	0.31 (±0.02)	0.28 (±0.02)
PWP(m^3 m^{-3})	0.19 (±0.02)	0.18 (±0.03)	0.20 (±0.02)	0.14 (±0.02)	0.23 (±0.02)	0.18 (±0.02)	0.22 (±0.02)	0.21 (±0.02)
AWC (m^3 m^{-3})	0.07 (±0.01)	0.07 (±0.02)	0.07 (±0.01)	0.07 (±0.01)	0.08 (±0.01)	0.07 (±0.01)	0.09 (±0.01)	0.07 (±0.01)
pH	8.2 (±0.4)	8.1 (±0.5)	8.7 (±0.4)	7.9 (±0.3)	8.9 (±0.3)	8.1 (±0.4)	8.5 (±0.4)	8.8 (±0.4)
$EC_{1:5}$ (dS m^{-1})	0.11 (±0.19)	0.09 (±0.26)	0.26 (±0.22)	0.16 (±0.18)	0.56 (±0.18)	0.15 (±0.19)	0.92 (±0.19)	0.38 (±0.20)
CEC (cmol(+) kg^{-1})	21.8 (±2.3)	18.5 (±3.2)	23.0 (±2.6)	15.1 (±2.2)	21.2 (±2.2)	15.8 (±2.3)	25.9 (±2.4)	19.6 (±2.4)

Regarding the Trangie Cowal Pedoderm, this was characterised for the most part by two main profile classes; namely the Wilga and Byron classes. This was similar for Byron soils which were mostly located in the Trangie Cowal Pedoderm. The clay content of Byron soils was low (25.0 % and 41.2 % for topsoil and subsoil, respectively). These were like the predicted clay of 8D (31.8 % and 35.2 % for topsoil and subsoil, respectively). Similarly, the pH values of Byron soils differed greatly for topsoil (6.6) and subsoil (8.3), which was not very different from the pH values of 8D (7.5 and 7.9 for topsoil and subsoil, respectively). However, large differences of CEC existed for topsoil (5.3 cmol(+) kg^{-1}) and subsoil (20.3 cmol(+) kg^{-1}) of Byron soils while smaller differences of CEC occurred within Class 8D for topsoil (12.9 cmol(+) kg^{-1}) and subsoil (15.1 cmol(+) kg^{-1}). This may be because the locations of 8D and Byron soils were not exactly the same. Because the DSM approach delineated areas where soil CEC was more homogenous with depth than predicted from the traditional pedoderm map, similar management practices can be applied in these areas to avoid unnecessary cost (e.g., fertiliser).

Worthy to note that it should not be expected, nor is there a requirement, for statistical differences between all soil properties in each of the classes. This is because, while there is a similarity between soil properties such as clay content in the various soil profile classes, there are subtle differences in chemical properties in the topsoil and subsoil, which will guide appropriate soil use and management. It was also observed, that whereas previously there was no indication of the pedoderm distribution across the northern half of the Warren area, the DSM has provided some clarity in the types of soil which might be present and how they may be managed. The intricacies of the two main pedoderms and the distribution of soil profile classes identified by McKenzie (1992) were also resolved although this needs to be verified with a more extensive soil sampling scheme.

7.4 Conclusions

A digital soil mapping (DSM) approach across two districts was conducted using two types of ancillary data (i.e., γ-ray spectrometry and EC_a data) with fuzzy k-means clustering. Using a linear mixed model with a residual maximum likelihood estimation method and various topsoil and subsoil physical (i.e., clay) and chemical (i.e. pH, $EC_{1:5}$ and CEC) properties, it was shown that the DSM when k = 8 minimised the mean square prediction error ($\sigma^2_{P,C}$). By comparing the $\sigma^2_{P,C}$ of the k = 8 DSM, it was found that it was superior for predicting most of these soil properties relative to the traditional pedoderm map (McKenzie 1992). Equivalent results were achieved when a DSM was developed using only the EC_a data. In addition, the DSM was able to delineate potentially the soil profile classes which characterize the various soil pedoderms.

However, the results may be site-specific and future work is required to test the robustness of the method on a larger scale when a similar highly heterogeneous area is to be surveyed. In addition, FKM clustering using ancillary data from ground γ-ray data and EC_a might also be worth studying in order to evaluate the accuracy of different types of ancillary data (Huang *et al.* 2017b). The combination with other available ancillary data (i.e. digital elevation model and remote sensing data) might also be useful to further improve the mapping accuracy and robustness of the method at a large scale (Jing *et al.* 2017).

7.5 References

Altdorff D, Dietrich P (2012) Combination of electromagnetic induction and gamma spectrometry using Kmeans clustering: A study for evaluation of site partitioning. *Journal of Plant Nutrition and Soil Science 175*, 345–354.

Anderson-Cook CM, Alley MM, Roygard JKF, Khosla R, Noble RB, Doolittle JA (2002) Differentiating soil types using electromagnetic conductivity and crop yield maps. *Soil Science Society of America Journal 66*, 1562–1570.

Bezdek JC (1981) Pattern recognition with fuzzy objective function algorithms. Plenum Press, New York.

Buchanan SM, Triantafilis J, Odeh IOA (2012) Digital soil mapping of compositional particle-size fractions using proximal and remotely sensed ancillary data. *Geophysics 77*, WB201–211.

Castrignanò A, Wong MTF, Stelluti M, De Benedetto D, Sollitto D (2012) Use of EMI, gamma-ray emission and GPS height as multi-sensor data for soil characterisation. *Geoderma 175–176*, 78–89.

Doolittle JA, Brevik EC (2014) The use of electromagnetic induction techniques in soils studies. Geoderma *223*, 33–45.

Fraisse CW, Sudduth KA, Kitchen NR (2001) Delineation of site-specific management zones by unsupervised classification of topographic attributes and soil electrical conductivity. *Transactions of the American Society of Agricultural Engineers 44*, 155–166.

Gooley, L., Huang, J., Page, D., Triantafilis, J., 2014. Digital soil mapping of available water content using proximal and remotely sensed data. *Soil Use and Management, 30,* 139-151.

Guo Y, Shi Z, Li HY, Triantafilis J (2013) Application of digital soil mapping methods to identify salinity management classes in coastal lands of central China. *Soil Use and Management 29*, 445–456.

Holmgren GGS, Juve RL, Geschwender RC (1977) A mechanically controlled variable rate leaching device. *Soil Science Society of America Journal 41*, 1207–1208.

Huang J, Wong V, Triantafilis J (2014a) Mapping soil salinity and pH across an estuarine and alluvial plain using electromagnetic and digital elevation model data. *Soil Use and Management 30*, 394–402.

Huang J, Lark RM, Robinson DA, Lebron I, Keith AM, Rawlins B, Tye A, Kuras O, Raines M, Triantafilis J (2014b) Scope to predict soil properties at within-field scale from small samples using proximally sensed γ-ray spectrometer and EM induction data. *Geoderma 232*, 69–80.

Huang J, Mokhtari AR, Cohen DR, Monteiro Santos FA, Triantafilis J (2015) Modelling soil salinity across a gilgai landscape by inversion of EM38 and EM31 data. *European Journal of Soil Science 66*, 951–960.

Huang J, McBratney AB, Minasny B, Triantafilis J (2017a) Monitoring and modelling soil water dynamics using electromagnetic conductivity imaging and the ensemble Kalman filter. *Geoderma 285*, 76–93.

Huang J, Bishop T, Triantafilis J (2017b) An error budget for digital soil mapping using proximally sensed EM induction and remotely sensed γ-ray spectrometer data. *Soil Use and Management*.

Jing D, Huang J, Banks R, Triantafilis J, 2017. Scope to map soil management units at the district level from remotely sensed γ-ray spectrometery and proximal sensed EM induction data. *Soil Use and Management 33*, 538–552.

Lark RM, Cullis BR, Welham SJ (2006) On spatial prediction of soil properties in the presence of a spatial trend: the empirical best linear unbiased predictor (E-BLUP) with REML. *European Journal of Soil Science 57*, 787–799.

Lark RM (2011) Spatially nested sampling schemes for spatial variance components: Scope for their optimization. *Computers and Geosciences 37*, 1633–1641.

Lesch SM, Rhoades JD, Corwin DL (2000) The ESAP-95 version 2.01R User Manual and Tutorial Guide Research Report No 146 USDA-ARS. In: Brown GE Jr (ed) Salinity Laboratory, Riverside, CA.

Loveday J, Beatty HJ, Reeve R (1972) Comparison of current chemical methods for evaluating irrigation soils. CSIRO Division of Soils, Technical Paper No 14. (CSIRO, Australia).

McBratney AB, Moore AW (1985) Application of fuzzy-sets to climatic classification. *Agricultural Forest Meteorology 35*, 165–185.

McBratney AB, De Gruijter JJ, Brus DJ (1992) Spatial prediction and mapping of continuous soil classes. *Geoderma 54*, 39–64.

McKenzie NJ (1992) Soils of the Lower Macquarie Valley, New South Wales. CSIRO Division of Soils, Divisional Report No 117, Canberra, Australia.

McNeill JD (1990) Geonics EM38 Ground Conductivity Meter: EM38 Operating Manual. Geonics Limited, Ont., Canada.

Miltenyi GPL, Malte MC, Triantafilis J (2015) Assessing the Australian Soil Classification using cladistic analysis. *Soil Research 53*, 772-785.

Minasny B, McBratney AB, Whelan BM (1999) VESPER version 1.6, Precision Agriculture Laboratory, Sydney, Australia.

Minasny B, McBratney AB (2002) FuzME version 3.0, Precision Agriculture Laboratory, The University of Sydney, Australia.

Pickup G, Marks A (2000) Identifying large-scale erosion and deposition processes from airborne gamma radiometrics and digital elevation models in a weathered landscape. *Earth Surface Processes and Landforms 25*, 535–557.

Pinheiro J, Bates D, DebRoy S, Sarkar D (2013) nlme: Linear and Nonlinear Mixed Effects Models. R package version *3*.1-110.

Popp S, Altdorff D, Dietrich P (2013) Assessment of shallow subsurface characterisation with non-invasive geophysical methods at the intermediate hill-slope scale. *Hydrology and Earth System Sciences 17*, 1297–1307.

R Development Core Team (2017) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Geonics Limited (2010), Operating instructions, EM38-MK2, 1745 Meyerside Drive, Mississauga, Ontario, Canada.

Saey T, De Smedt P, Delefortrie S, Van De Vijver E, Van Meirvenne M (2015) Comparing one-and two-dimensional EMI conductivity inverse modeling procedures for characterizing a two-layered soil. Geoderma *241*, 12–23.

Triantafilis J, Gibbs ID, Earl NY (2013) Digital soil pattern recognition in the lower Namoi valley using numerical clustering of gamma-ray spectrometry data. *Geoderma 192*, 407–421.

Triantafilis J, Kerridge B, Buchanan SM (2009) Digital soil-class mapping from proximal and remotely sensed data at the field level. *Agronomy Journal 101*, 841–853.

Triantafilis J, Laslett GM, McBratney AB (2000) Calibrating an electromagnetic induction instrument to measure salinity in soil under irrigated cotton. *Soil Science Society of America Journal 64*, 1009–1017.

Triantafilis J, Ward WT, Odeh IOA, McBratney AB (2001) Creation and interpolation of continuous soil layer classes in the lower Namoi valley. *Soil Science Society of America Journal 65*, 403–413.

Triantafilis J, Lesch SM (2005) Mapping clay content variation using electromagnetic induction techniques. *Computers and Electronics in Agriculture 46*, 203–237.

Tucker BM (1974) Laboratory procedure for cation exchange measurements in soils. CSIRO Div. Of Soils, Tech. Paper No 23. (CSIRO, Australia).

Wilford JR, Bierwirth PN, Craig MA (1997) Application of airborne gamma-ray spectrometry in soil/regolith mapping and applied geomorphology. *Journal of Australian Geological Geophysics 17*, 201–216.

Wilford JR, Dent DL, Dowling T, Braaten R (2001) Rapid mapping of soils and salt stores: Using airbone radiometrics and digital elevation models. *AGSO Research Newsletter 34*, 33–40.

Zare E, Huang J, Triantafilis J, 2016. Digital mapping of soil landscape units at the district level using remote and proximal sensed data. *Computers and Electronics in Agriculture 127*, 510-520.

Chapter 8: Conclusions and future research

8.1 Conclusions

This thesis delves into the potential to utilize digital soil mapping (DSM) methods to characterize the temporal and spatial distribution of soil θ and soil water-related physical (i.e., clay, silt and sand), hydrological (i.e., permanent field capacity, permanent wilting point and available water content) and chemical (e.g., CEC, EC_e and pH) properties in space and time in irrigated cotton growing fields and districts where Vertosols are the dominate Australian Soil Order. Furthermore, it explores the possibility to recognize where there are inefficiencies in water delivery and also the requirement for knowledge across larger spatial scales and across the landscape.

In particular, this thesis studied the performance of different aspects of the DSM approach to couple the above soil data to digital data by comparison of mathematical models, including Artificial Neural Network (ANN), multiple linear regression (MLR), random forest (RF), Cubist, support vector machine (SVM), stepwise MLR, linier regression (LR), linear mixed model (LMM) and numerical fuzzy clustering (i.e., FKM). Moreover, the usefulness of digital data, including soil apparent electrical conductivity (EC_a) whether alone or inverted (σ), digital elevation model (i.e., elevation), proximal and remotely sensed gamma-ray (γ-ray) spectrometry data either alone or in combination was explored.

These model prediction performances were evaluated by leave-one-out cross validation and contrasted using statistical measures of agreement (Lin's concordance correlation coefficient-Lin's), coefficient of determination (R^2), root mean squared error (RMSE) and mean squared prediction error (MSPE). The agreement criteria were described by considering the agreement classification of McBride (2005), including poor (Lin's < 0.65), moderate (0.65 < Lin's < 0.8) and good (Lin's > 0.8). In addition, the prediction was considered acceptable if

the RMSE of predictions is less than half the standard deviation (½SD) of the measured data (Singh et al. 2005). In terms of the MSPE, the smaller the better.

In Chapter 3, by applying a quasi-2d non-spatial-temporal inversion algorithm to EC_a data collected along a 200m transect by a multicoil electromagnetic instrument (i.e., DUALEM-421) before and during an irrigation event, time-lapse images of θ were generated based on the correlation between θ and the estimates of true electrical conductivity (σ) collected over 7 consecutive time-steps. It was concluded that the distribution of θ at various depths and along the transect can be used for characterizing the wetting front using an ANN model. This facilitates the evaluation of the irrigation efficiency of the farrow irrigation system across the Vertosol field. One drawback of the approach however was that the data was only collected during the wetting-phase. Moreover, the length and weight of the DUALEM-421 made it impractical for predicting θ during the irrigation event. In addition, the research was undertaken during a singly irrigation event which occurred after a period of fallow where the soil θ was near field capacity in the subsoil and this made it difficult to resolve the change in θ as a function of the advance of the wetting front.

In chapter 4, to further test and understand the potential for spatial and temporal depth-characterisation of θ dynamics and as a function of wetting and drying front variations, EC_a data (i.e., EM38h and v) was used either alone or through inversion to estimate depth-specific σ, to make calibrations with depth-specific θ along the same transect used in Chapter 3. Herein, EC_a and θ data was collected over 10 continuous days during and after an irrigation event. Specifically, performance of different mathematical models was investigated, including MLR, RF, Cubist, SVM, ANN and LR. In terms of overall ranking, the best approach was to invert the EC_a, with σ and depth modelled to θ using SVM. This was followed by MLR of EC_a. The usefulness of collection of the EC_a at various heights (0, 0.2, 0.4, 0.6, 0.8 and 1.0 m) was also

investigated, but it was found that there was no improvement in predictions when the instrument was raised. Additionally, the number of calibration sites required to predict θ was determined. It was concluded that owing to the uniform nature of the soil and the time-lapse nature of imaging this could be limited to a single calibration site.

Chapter 5 focused on the identifying potential leakage zones beneath a 4 km cross-section of an irrigation supply channel by mapping soil properties to a depth of 6 m by employing a multicoil electromagnetic instrument (i.e., DUALEM-421). It was found that by using a SVM, a calibration model could be established between estimated σ (inverted from DUALEM-421 EC_a) and clay, sand, EC_e and CEC. The interpretation of predicted EC_e suggests the approach can identify recharge (non-saline) areas. These aligned with where the clay (< 40 %) was relatively small, sand was large (> 50%) and CEC was small (< 20 cmol(+) kg^{-1}) and in areas associated with prior stream channels that were typically underlain by migrational palaeochannels. It was concluded, with this calibration, the approach has the potential to be used to predict the soil physical and chemical properties within supply channels across the rest of the valley.

In Chapter 6, the aim was to develop a DSM of AWC across a dryland cotton field with Vertosol type of soil at 0-0.30 m depth increments within rootzone (0-0.9 m). To achieve this, a stepwise-MLR model was developed using ancillary data including proximally sensed EC_a (i.e., DUALEM-421) and γ-ray spectrometry (i.e., RS-700) data to predict field capacity (FC) and permanent wilting point (PWP). The final AWC maps were constructed by subtracting PWP from FC. While the γ-ray spectrometry data and the trend surface parameters added information relevant to soil mineralogy and perhaps soil age (i.e., degree of weathering) the EM relate FC and PWP to differences in soil texture. As, the AWC has been shown to be a function of soil texture and because the various soil particle size fractions vary spatially, there was the potential to use a DSM approach to provide this information. To better understand

where errors were large the maps of 95% confidence intervals were also provided. It was concluded that using the DSM approach it was possible to map AWC across a dryland cotton farm in topsoil (0-0.3 m), subsurface (0.3-0.6 m) and subsoil (0.6-0.9 m) with good accuracy (Lin's = 0.80). This can potentially help the dryland farmers to make management decisions about planting time and row spacing.

In Chapter 7, it was intended to study the effects of clustering either in combination or individually airborne γ-ray spectrometry and proximal EC_a (i.e., EM38h and v) data across two geomorphologically different areas in the central-west of New South Wales. Additionally, the DSM was compared with a Soil Order map (McKenzie 1992) derived using traditional methods. Using a linear mixed model with a residual maximum likelihood estimation method and various topsoil and subsoil physical (i.e. clay content), hydrological (i.e. FC, PWP and AWC) and chemical (i.e. pH, $EC_{1:5}$, CEC) soil properties, it was determined that the DSM when k = 8 minimised the mean square prediction error ($\sigma^2_{p,C}$). It was found that the $\sigma^2_{p,C}$ of the k = 8 DSM was superior for predicting most of these soil properties relative to the traditional pedoderm map (McKenzie 1992). Equivalent results were achieved when a DSM was developed using only the EC_a data. In addition, the DSM was able to delineate potentially the soil profile classes that characterise the various soil Pedoderms.

8.2 Implications for future research

Given the results presented in the thesis, it was concluded that the EM induction instruments can be applied to efficiently map and monitor θ dynamics and other soil properties (i.e., clay, EC_e, AWC) at the field scale associated with heavy clay, Vertosols. Moreover, it was shown that addition of γ-ray spectrometry data can provide information about the soil parent material and therefore can differentiate soil-texture and soil types. However, a number

of issues can be further investigated in the future to improve the robustness of the approaches used in this thesis.

Firstly, given the difficulty of measuring soil EC_a using a DUALEM-421, owing to the length and weight of the instrument and trafficking across a saturated Vertosol field, a smaller EM instruments such as DUALEM-1HS could be considered. This is because it has four theoretical depths of measurement from the two coil pairs available. Specifically, 0.5mHcon (0-0.75 m) and 1mHcon (0-1.5 m) have equivalent depths of measurement as EM38h (0-0.75 m) and EM38v (0-1.5 m). However, the advantage of the DUALEM-1HS is the 0.5mPcon (0-0.3 m) and 1mPcon (0-0. 5 m) EC_a, which allows for better resolution of topsoil and subsurface σ and therefore prediction of θ in the upper- and lower-subsoil. Moreover, this would also be beneficial as EC_a data could be collected from 0.5 m coil spacing, which could potentially further assist in resolving topsoil and subsurface θ differences, leading to improved calibration and prediction.

Secondly, it is worth carrying out EM surveys along with different irrigation practices. For example, EM surveys can be conducted at farms to evaluate and compare the irrigation efficiency of different irrigation systems (e.g., furrow, travelling, bankless and drip) and associated with different soil types. The results can be used to identify where over-irrigation or deep drainage occurs and provide guidance for maximizing the agricultural water use efficiency.

Thirdly, while the DSM approach used in chapter 5 has implication to identify potential leakage zones of the irrigation supply channels, for the remaining lower Namoi valley, it is worth investigating the use of cheaper and more popular EM instruments. This is because a similar approach can be used by an increasingly large number of consultants who commonly conduct EM surveys using the more popular Geonics EM38 or smaller versions of the

DUALEM range of instruments, including DUALEM-1. In this regard, further research could be conducted to see if equivalent calibrations can be developed to predict the physical and chemical properties described herein but only using a smaller number of EC_a values. The transect studied herein could act as a calibration location.

Lastly, in regard to FKM clustering of γ-ray spectrometry and EC_a data the results may be site-specific. Therefore, future work is required to test the robustness of the method on a larger scale when a similar highly heterogeneous area is to be surveyed. In addition, FKM clustering using ancillary data from ground γ-ray data and EC_a might also be worth studying in order to evaluate the accuracy of different types of ancillary data. The combination with other available ancillary data (i.e., digital elevation model and remote sensing data) might also be useful to further improve the mapping accuracy and robustness of the method at a large scale. Moreover, the inclusion of other data from nearby locations might also be useful to further improve the mapping accuracy and robustness of the method in multiple areas and thereby enable determination of how similar soil types are in various cotton growing areas.

8.3 References

McBride, G. B. (2005). A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. Hamilton, New Zealand: National Institute of Water and Atmospheric Research Ltd. McBride2005. pdf.

McKenzie NJ (1992) Soils of the Lower Macquarie Valley, New South Wales. CSIRO Division of Soils, Divisional Report No 117, Canberra, Australia.