A site-specific and dynamic modeling system for zoning and optimizing variable rate irrigation in cotton

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Amir Haghverdi

August 2015

Copyright © 2015 by Amir Haghverdi

All rights reserved.

DEDICATION

This dissertation is dedicated to my beloved wife, Somayeh Ghodsi, for her love and support throughout our marriage.

ACKNOWLEDGEMENTS

I would like to acknowledge my graduate advisory committee members who have provided guidance throughout this study. A special thanks to Dr. Brian Leib, my Major Professor, for supporting me in fulfilling my scientific dreams. I am especially appreciative of the financial support to this project that was provided by Cotton Incorporated and the USDA NRCS-CIG. I also appreciate Pugh Brothers Farms for allowing me to implement this project on their property.

ABSTRACT

Cotton irrigation has been rapidly expanding in west Tennessee during the past decade. Variable rate irrigation is expected to enhance water use efficiency and crop yield in this region due to the significant field-scale soil spatial heterogeneity. A detailed understanding of the soil available water content within the effective root zone is needed to optimally schedule irrigation. In addition, site-specific crop-yield mathematical relationships should be established to identify optimum irrigation management. This study aimed to design and evaluate a site-specific modeling system for zoning and optimizing variable rate irrigation in cotton. The specific objectives of this study were to investigate (i) the spatial variability of soil attributes at the field-scale, (ii) site-specific cotton lint yield-water relationships across all soil types, and (iii) multiple zoning strategies for variable rate irrigation scenarios. The field (73 ha) was sampled and apparent soil electrical conductivity (ECa) was measured. Landsat 8 satellite data was acquired, processed, and transformed to compare indicators of vegetation and soil response to cotton lint yields, variable irrigation rates, and the spatial variability of soil attributes. Multiple modeling scenarios were developed and examined. Although experiments were performed during two wet years, supplemental irrigation enhanced cotton yield across all soil types in comparison with rain-fed conditions. However, length of cropping season and rainfall distribution remarkably affected cotton response to supplemental irrigation. Geostatistical analysis showed spatial variability in soil textural components and water content was significant and correlated to yield patterns. There was as high as four-fold difference between available water content between coarse-textured and fine-textured soils on the study site. A good agreement was observed (RMSE = 0.052 cm^3 cm^{-3} [cubic centimeter per cubic centimeter] and r = 0.88) between predicted and observed water contents. ECa and space images were useful proximal data to investigate soil spatial variability. The site-specific water production functions performed well at predicting cotton lint yield with RMSE equal to 0.131 Mg ha^{-1} [megagram per hectare] and 0.194 Mg ha^{-1} in 2013 and 2014, respectively. The findings revealed that variable rate irrigation with pie shape zones could enhance cotton lint yield under supplemental irrigation in west Tennessee.

TABLE OF CONTENTS

Part 1: Introduction and Literature Review 1
1. Introduction 1
1.1 Irrigation in west Tennessee 1
1.2 Cotton response to supplemental irrigation 2
1.3 Site-specific cotton irrigation management 4
2. Root zone soil hydrology 6
2.1 Pedotransfer functions 6
2.2 Soil spatial prediction 7
2.3 On-the-go sensors 8
3- Management zones delineation 9
4. Water production functions 12
4.1 Yield prediction in conventional agriculture 13
4.2 Yield prediction in precision agriculture 15
References 18
Part 2: Studying effective root zone soil hydrology for site-specific irrigation 24
Abstract 25
Chapter 1: The spatial variation of soil physical properties within the effective root zone at a field scale in west Tennessee 26
Abstract 26
1. Introduction 26
2. Material and Methods 28
2.1 Study area 28
2.2 Soil data collection and lab analysis 29
2.3 Descriptive and spatial analysis of soil properties 29
2.4 Yield data collection and cleaning 30
3. Results and Discussion 31
3.1 Variation among soil layers 31
3.2 Spatial analysis of soil properties 31
3.3 Effectiveness of ECa as a proximal attribute 33
3.4 Soil heterogeneity and yield patterns 34
4. Conclusion 35
References 36
Appendix 1: Chapter 1 Figures and Tables 37
Chapter 2: High resolution prediction of soil available water content within crop root zone at a field scale 47
Abstract 47
1. Introduction 48

2. Material and Methods 50
2.1 Study area & data collection 50
2.2 Pseudo continuous pedotransfer function PC-PTF 50
2.3. Spatial modeling of soil available water content 52
2.3.1 Kriging, co-kriging and regression kriging 52
2.3.2 Geographically weighted regression 54
2.3.3 Artificial neural network 54
2.4 Yield data collection and cleaning 55
2.5 Performance evaluation 55
3. Results 56
3.1 Pseudo continuous pedotransfer function 56
3.2 Spatial prediction of soil available water content 56
3.3 Soil water retention, yield and ECa variation 58
4. Discussion 59
4.1 Model performance analysis 59
4.2 Practical challenges and findings 61
5. Conclusion 62
References 64
Appendix 2: Chapter 2 Figures and Tables 66
Part 3: Perspectives on cotton supplemental irrigation in west Tennessee 79
Abstract 80
Chapter 3: Studying crop yield response to supplemental irrigation and the spatial heterogeneity of soil physical attributes in a humid region 81
Abstract 81
1. Introduction 82
2. Material and Methods 84
2.1 Study area 84
2.2 Irrigation experiment 85
2.3 Instrumentation and measurements 86
2.4 Satellite data processing and transformation 87
2.5 Temporal stability of yield patterns 89
3. Result and Discussion 90
3.1 Cotton lint yield 90
3.2 Soil water status 91
3.3 Tasseled cap transformation and vegetation indices 92
3.4 Temporal stability of yield patterns 94
4. Discussion 95

4.1 Yield patterns, soil heterogeneity and supplemental irrigation....95
4.2 Application of space images and soil moisture sensors....96
5. Conclusion....98
6. Acknowledgment....98
References....99
Appendix 3: Chapter 3 Figures and Tables....101
Part 4: Toward site-specific irrigation management in west Tennessee....116
Abstract....117
Chapter 4: Perspectives on delineating management zones for variable rate irrigation....118
Abstract....118
1. Introduction....119
1.1 Precision farming and management zone delineation....119
1.2 Irrigation management zones delineation....120
2. Material and Methods....122
2.1 Field of study....122
2.2 Proximal data....122
2.3 Management zone delineation....123
2.3.1 K-means....124
2.3.2 ISODATA - Maximum Likelihood....124
2.3.3 Gaussian Mixture Model....125
2.3.4 Integer linear programing....125
2.4 Performance Evaluation....126
3. Results....127
3.1 Zoning using ECa and satellite images....127
3.2 Spatiotemporal changes in soil available water content....130
3.3 Management zone delineation, phase two....130
4. Discussion....132
4.1 Management zones delineation methods....132
4.2 Optimum number of zones....132
4.3 Application of proximal data....133
4.4 Temporal variability and role of crop....135
5. Conclusion....135
References....137
Appendix 4: Chapter 4 Figures and Tables....139
Chapter 5: Studying uniform and variable rate center pivot irrigation strategies with the aid of site-specific water production function....154
Abstract....154

1. Introduction....155
1.1 Precision farming and variable rate irrigation....155
1.2 Site-specific water production functions....156
2. Material and Methods....158
2.1 Field of study....158
2.2 Irrigation experiment....158
2.3 Input attributes and data preparation....159
2.4 Modeling and simulation....160
2.4.1 Phase 1: Establishing water production functions....160
2.4.2 Phase 2: Simulating irrigation and zoning strategies....162
2.5 Evaluation procedure....163
3. Results....163
3.1 Principal component analysis....163
3.2 Performance of WPFs....164
3.3 Irrigation and zoning scenarios....164
4. Discussion....167
4.1 Site-specific water production functions....167
4.2 Zoning and irrigation strategies....168
5. Conclusion....170
6. Acknowledgment....171
References....172
Appendix 5: Chapter 5 Figures and Tables....174
Part 5: Conclusion....184
5-1 Root zone soil hydrology....185
5-2 Cotton response to supplemental irrigation....185
5-3 Irrigation strategy and zone delineation....186
References....188
Vita....189

LIST OF TABLES

Table 1-1. Descriptive statistics for selected soil properties from different soil sampling layers. 43

Table 1-2. Semivariogram and Moran's I parameters of soil properties for different soil layers. 44

Table 1-3. Correlation coefficient between ECa (mS m^{-1}) data and soil basic information. 45

Table 1-4. Correlation coefficient between cotton / soybean yield data and soil information. 46

Table 2-1. Properties of the soils selected from the UNSODA data set and collected from the field of study. 75

Table 2-2. The RMSE (cm^3 cm^{-3}) for the models (Figure 4) predicting water content at FC and PWP at different layers. 76

Table 2-3. The MBE (cm^3 cm^{-3}) for the models (Figure 4) predicting water content at FC and PWP at different layers. 77

Table 2-4. Correlation coefficients between ECa and yield data with water retention data at different layers, predicted by PC-PTF for samples (n= 100 (points) × 4 (depths) = 400). 78

Table 3-1. Detailed information on the irrigation programs for two center pivots in the study area for one revolution. 110

Table 3-2. Growing season summary of weather and supplemental irrigation data in the study area in 2013 and 2014, in comparison to the 30-year mean for these variables. 111

Table 3-3. Summary of satellite images collected for this study from Landsat 8. 112

Table 3-4. Correlation coefficient between cotton lint yield data (2013 and 2014 cropping seasons), soil properties at 4 depths, and fertilizer application. 113

Table 3-5. The correlation coefficients between field-measured cotton lint yield maps and remotely sensed maps of crop indices. 114

Table 3-6. Descriptive statistics on yield data (Mg ha^{-1}) at the field of study. 115

Table 4-1. Descriptive statistics for soil water content from different soil layers (from chapters 1 and 2). 151

Table 4-2. Summary of zoning algorithms and attributes. 152

Table 4-3. All of the pie shape zones underneath a center pivot irrigation system. 153

Table 5-1. Input/output attributes for deriving site-specific WPFs. 182

Table 5-2. Performance of k-NN, ANN and MLR WPFs for 2013 and 2014 cropping seasons. 183

LIST OF FIGURES

Figure 1-1. Long term variation in rainfall (dash line) and temperature (column) data in west Tennessee close to the area of study. 37

Figure 1-2. Sampling scheme in the field of study located in west Tennessee. 38

Figure 1-3. Scattering of basic soil properties against volumetric water content where the darker colors corresponds to the deeper layers. 39

Figure 1-4. Soil textural distribution of samples at different layers where the darker colors corresponds to the deeper layers. 40

Figure 1-5. Interpolated maps by kriging method. 41

Figure 1-6. Cotton and soybean yield maps where green and red represent high and low yield, respectively. 42

Figure 2-1. Soil sampling scheme within the field of study located in Dyer County, Tennessee. 66

Figure 2-2. Illustration of topology of point, pseudo continuous and parametric neural network pedotransfer functions 67

Figure 2-3. Soil texture distribution; samples from the field of study (panel 1), selected samples from UNSODA (panel 2). 68

Figure 2-4. Spatial models to predict the soil water retention. 69

Figure 2-5. Performance of pseudo continuous PTFs in predicting the WRC of UNSODA samples. 70

Figure 2-6. The WRCs at different soil textures predicted by PC-PTF for the soil samples (n= 400) collected from field of study. 71

Figure 2-7. Scattering of water content at 10 kPa and 33 kPa predicted by PC-PTF against water content at the time of sampling. 72

Figure 2-8. The semivariograms for water retention data. 73

Figure 2-9. AWC maps for different layers 74

Figure 3-1. The 73-ha study farm and its 2 center pivots in the Dyer County humid region of west Tennessee. 101

Figure 3-2. The supplemental irrigation experimental design 102

Figure 3-3. Irrigation and soil type effect on cotton lint yield for selected regions across the field. 103

Figure 3-4. Temporal changes [days after planting (DAP)] in soil water status (SF) throughout the 2013 growing season. 104

Figure 3-5. Temporal changes [days after planting (DAP)] in soil water status (SF) throughout the 2014 growing season. 105

Figure 3-6. The growing season dynamics of Landsat 8 106

Figure 3-7. The growing season dynamics of Landsat 8 107

Figure 3-8. The October to November of 2013 and May to September of 2014 growing season's normalized difference vegetation index (NDVI). 108

Figure 3-9. The cotton lint yield map time series from 2012 to 2014 (a - c) and the mean (d, using Eq. 6) and standard deviation yield maps (e, using Eq. 7). 109
Figure 4-1. Field of study within the state of Tennessee. 139
Figure 4-2. Soil texture distribution for the samples (n= 400) collected from the field of study. 140
Figure 4-3. Performance of clustering methods based on percent reduction in AWC variance and within-zone average yield. 141
Figure 4-4. Gray-scale maps of panchromatic spectral band during 2013-2014 from the field of study. 142
Figure 4-5. Distribution of standardized brightness values for panchromatic spectral band during 2013-2014. 143
Figure 4-6. AWC variance against number of zones using different satellite images. 144
Figure 4-7. Spatial arrangement of MZs. 145
Figure 4-8. Percent of average yield within MZs clustered with k-means considering ECD as input (bar colors are related to the zones in Figure 4-7). 146
Figure 4-9. Within-season temporal variation in user defined MZs considering available water content as input... 147
Figure 4-10. Performance of irrigation MZ delineation methods based on AWC standard deviation reduction. 148
Figure 4-11. Irrigation MZs for center pivot systems located at east and west parts of the field. 149
Figure 4-12. Kappa coefficient for the zones delineated by the integer linear programming (ILP) method. 150
Figure 5-1. Field of study in Dyer County, State of Tennessee. 174
Figure 5-2. Three methods to establish site-specific water production functions. 175
Figure 5-3. PCA result. Panel a: cumulative Hotelling's T2 against PCs. 176
Figure 5-4. The k-NN WPFs performance over both cropping season for different combinations of k and p values. 177
Figure 5-5. Scatter plots of measured versus predicted cotton lint yield using k-NN, ANN and MLR WPFs. 178
Figure 5-6. Cotton lint yield prediction using site-specific k-NN WPF for multiple irrigation scenarios. 179
Figure 5-7. Cotton lint yield maps predicted by site-specific k-NN WPF under different irrigation scenarios. 180
Figure 5-8. Spatial arrangement of the optimum pie shape zones (n = 4). 181

Part 1: Introduction and Literature Review

1. Introduction

1.1 Irrigation in west Tennessee

The growing demand for food and fiber production along with the intrinsic uncertainty in rainfall patterns, due to climate variability, has focused great attention on irrigation. World agricultural production over the last fifty years has grown between two and four percent per year on average while irrigation has doubled in the same time period. Irrigated agriculture has the highest rate of consumption of the fresh water resources (about 70%) and produces more than 40% of the food supply while using 17% of the agricultural land area (FAO, 2013; Fereres and Connor, 2004). In arid and semiarid climates, irrigation is essential for crop production where a crop failure or a significant reduction in the amount of yield would most likely occur without irrigation. The growing demand for water and intense water scarcity creates a philosophy of "more crop per drop" making deficit irrigation a common practice that by definition means to deliberately apply water below the evapotranspiration (ET) requirement. Overall, irrigation optimization has been a key area of research for a long time in arid and semiarid regions.

Irrigation management in moderately-humid and humid regions is a different concept where optimized water allocation is not the prime concern and "more crop per drop" does not define the irrigation approach. Supplemental irrigation is the relevant practice in such areas (irrigate if needed), which is an irrigation strategy that attempts to stabilize maximum yield production by irrigating whenever the rainfall is not sufficient to fulfill the crop water requirement. Crops in humid regions produce some yield without supplemental irrigation, since precipitation is the prime source of moisture rather than irrigation. Irrigation management for optimizing yield in a humid region is more complicated than that in an arid region for several reasons. Excess water content in the root zone could occur due to overlapping irrigation events with rainfall. This may cause yield reduction either because of lack of aeration or by unnecessarily increasing crop biomass but not yield. On the other hand, severe in season drought conditions for a short period of time are likely to occur in a humid region when lack of irrigation can significantly reduce yield. Stabilizing yields and the recent high commodity prices are the reason that row crop irrigation is rapidly expanding in humid regions.

In west Tennessee, irrigation has been increasing exponentially during the last 2 decades such that there was approximately 2000 acres per year increase from 1992 to 1999 while this number was doubled by a rate approximately equal to 4000 acres per year from 2002 to 2007 (NASS, 2010). Informal statistics shows the growth rate increased to 60,000 acres per year in 2013. Although more producers recently utilized some irrigation systems in west Tennessee, there are still lots of potential users who have not yet established an irrigation system. There are some studies available focusing on irrigation management in west Tennessee, while more information is needed to help producers optimize their irrigation management.

1.2 Cotton response to supplemental irrigation

Cotton is one of the major crops in west Tennessee and it is also vital to the economy of the United States since it is an essential export-oriented product. Cotton supplemental irrigation is growing fast in west Tennessee. The temporal pattern of rainfall changes from year to year with unexpected drought periods likely to occur each growing season. Consequently, producers are more willing to invest in irrigation systems to prevent any yield loss. Gwathmey et al. (2011) investigated the cotton responses to supplemental irrigation in Jackson TN in a 4-year study. They observed irrigation treatment significantly improved lint yield (i.e. 38 % in average at the 2.54 cm wk^{-1} irrigation rate) in comparison with the rainfed treatment in 3 of the 4 years. Several factors such as heat unit accumulation, nutrient level, and determinacy of varieties affect cotton growth and yield and also influence the cotton response to irrigation.

Full irrigation, which is an irrigation strategy to fulfill ET demand, is expected to produce the highest amount of yield in most agricultural crops. Cotton irrigation in humid regions with a short season environment is one of a few exceptions to this rule. Unlike dry regions, the response of cotton to irrigation in humid regions is not consistent due to several reasons. Applying enough water to satisfy cotton ET demand causes an excessive vegetative growth but most likely reduces yield since there is not usually enough time for all of the bolls to open by harvesting time (Morrow and Krieg, 1990). Bajwa and Vories (2007) demonstrated that excessive irrigation in wet weather conditions decreased cotton lint yield in Arkansas. That is why, the quantity of optimum ET (to maximize cotton lint yield) is directly proportional to length of cropping seasons; in longer seasons higher irrigation is expected to produce the highest cotton lint yield (Orgez et al., 1992). Therefore, deficit irrigation is the suggested irrigation discipline for cotton in a humid short season environment to maximize lint yield, yet there might be some minor

issues due to irrigation as opposed to rainfed agriculture. For example, an increase in insect control problems may occur due to cotton irrigation and also irrigation may cause plants to lodge. Moreover, irrigated cotton needs higher levels of management than non-irrigated cotton (Parks et al., 1978).

Magnitude and timing of irrigation affect cotton yield, maturity and fiber quality. High water stress reduces cotton lint yield mainly by decreasing the number of bolls per unit growth area. Pettigrew (2004) observed more bolls produced in irrigated cotton at higher plant nodes than dryland cotton. In general, irrigation delays maturity while water stress improves earliness. Pettigrew (2004) observed 20% taller plants as well as a delayed cutout due to irrigation in relation to dryland plots. Pettigrew (2004) showed that irrigation can cause a longer vegetative growth period after initiation of reproductive growth compared to dryland. The effect of water stress and irrigation on fiber quality components is inconsistent in the literature. Pettigrew (2004) stated that irrigation had no substantial consistent effect on fiber quality relative to dryland plots but in most of the years produced slightly longer fiber. Snowden et al. (2013) also observed both increase and decrease in micronaire due to irrigation in different cropping seasons.

The optimum cotton ET to maximize lint yield varies by the different cultivars. Orgez et al. (1992) found cotton crop water production functions were curvilinear in shape for most of the cultivars. They found a linear function for a cultivar with early maturity that is most likely due to less immature bolls by harvesting. Long season cultivars may end up with a big portion of the bolls immature by harvest time. This difference was smaller for longer seasons. Pettigrew (2004) reported different lint yield across cultivars, yet the general response to moisture deficit was similar among them. They observed inconsistent early flower production in dryland plots. Snowden et al. (2013) found a strong irrigation by cultivar interaction for micronaire by studying six cotton cultivars under four irrigation levels within two years on the Texas High Plains.

In most of the crops, it is expected to see biomass reduction and relatively constant harvest index for small to medium water deficits while reduction in both (harvest index and biomass) is anticipated for greater water deficits (Fereres et al., 2007). Cotton is one of a few crops that does not follow this general relationship between biomass and harvest index. It was reported that applying small to medium water deficits reduces cotton biomass, while varying by variety either increases the harvest index or does not affect it (Orgaz et al., 1992, Carmona et al., 2007). This is because a small to medium water deficit predominantly reduces vegetative parts of the cotton biomass but does not affect or only slightly decreases lint yield.

When water is adequate, the available heat-units is the prime factor limiting cotton yield. Gwathmey et al. (2011) showed that the failure to accumulate adequate heat-unit, may significantly affect cotton lint yield and the applied water relationships in west Tennessee with short season humid environments. On the other hand, water stress can significantly decrease heat unit-lint yield correlation (Peng et al., 1989). Determinacy of varieties as well as distribution of heat units across growth stages affect the heat unit-lint yield relationship. The impact of collected heat units is not uniform at all cotton growth stages. Peng et al. (1989) found a linear relationship between heat unit accumulation and lint yield for indeterminate and moderate determine cultivars but a curvilinear relationship for determinate cultivar. As a result, under cool conditions more determinate cultivars may produce higher lint yield while the inverse is correct under dry conditions. The evidence in literature points toward flowering and yield formation as the most sensitive growth stages to heat unit accumulation. The early reproductive development period highly affects the number of fruiting sites hence lint yield. Peng et al. (1989) reported early fruiting as the most sensitive growth stage to accumulated heat-unit. In another study, Kerby (1986) showed that heat unit accumulation at the beginning of flowering is highly correlated with the amount of lint yield per boll. Morrow and Krieg (1990) found a positive correlation between lint yield and heat unit accumulation before flowering while the opposite was true after flowering. They mentioned that considering the data of a large number of dryland plots was the reason for getting the negative correlation after flowering.

1.3 Site-specific cotton irrigation management

Precision agriculture (PA) is moving forward in line with the significant achievements in instrumentation, measurements and data-processing. This revolution is changing the concept of an agricultural unit from farm to sub-field by providing precision management opportunities for farmers. The National Research Council definition of PA comprises three components: gathering data, analyzing data and subsequently managing the farm, based on the results at an appropriate scale and time (Oliver, 2010).

With PA technology, farmers have the means to collect the information needed to optimize site-specific variable-rate irrigation. In fact, they have the facilities to conduct customized research that has traditionally been reserved for research farms with small plots. The parameters that are needed for site-specific irrigation are related to components in the soil-water-crop-atmosphere continuum. Most farmers in west Tennessee have access to precision farming equipment. Precision farming provides a unique opportunity to continuously produce valuable sources of

information. An enormous amount of data is obtained with extremely valuable information about temporal and spatial changes in each field. Yield maps are the most readily available information, while other data such as soil apparent electrical conductivity (ECa) maps and elevation maps are also, to a great extent, available. If not, they can be collected without spending a considerable amount of time and money.

Conventional irrigation management tries to answer *when* and *how much* to irrigate. Given field soil variation, variable rate irrigation management should address *where* to irrigate as well. It has been proven in various studies that soil characteristics such as soil water holding capacity (WHC) and depth of soil significantly affect crop yield, thus irrigation strategies should be adjusted in regard to the soil type. Soil variation is common in west Tennessee with respect to soil hydraulic and physical properties. Duncan (2012) illustrated that a single irrigation decision cannot optimize the cotton lint yield for soils with significantly different WHC in this region. Given soil spatial variation under irrigation systems in this region, Duncan (2012) concluded variable rate irrigation may be the optimum irrigation scenario to enhance cotton lint production. Hedley and Yule (2009b) compared variable rate irrigation and conventional uniform irrigation and illustrated that 9–19 % of irrigation water was saved which in turn reduced nitrogen leaching. Precision irrigation center pivots that pulse banks of sprinklers creating variable flow have been commercially available for a while. In addition, most of the available center pivots in west Tennessee are able to vary irrigation across a field by changing their travel speed and creating limited pie shape zones.

Overall, optimizing irrigation has received less attention in practice than other PA applications such as precision fertilization. This fact is revealed by a rapid assessment of abstract titles from a recent PA conference (http://www.ecpa2013.udl.cat/) in Spain (July 7th-11th 2013). One possible reason is the inherent complexity of irrigation management compared with the more straightforward nature of the other precision applications. Another practical challenge is to provide all of the necessary information from the soil-water-crop-atmosphere continuum for managing variable rate irrigation. An example is the spatial and temporal changes in soils. Spatial variation is inherent to soil and will most likely be different in any region and at various scales. Less pronounced temporal soil variation also can be caused by human activities, such as agriculture and soil management, or natural phenomena like soil shrinkage-swelling and soil crusting (Wösten et al., 2001).

Soil moisture is the most widely used indicator for irrigation scheduling (Leib et al., 2012; Leib et al., 2003). Recent advances in wireless communication makes it more feasible to monitor soil water status in multiple locations within a field which is required for scheduling variable rate irrigation systems. Pan et al. (2013) established a

framework to manage irrigation in a field with soil spatial variation by means of information available in PA (i.e. field elevation and ECa), wireless sensing technology, and site-specific derived equations. In another study, Hedley and Yule (2009a) produced soil hydraulic maps using regression equations from high resolution ECa maps. They added a daily time step to the generated map to spatially estimate soil water status across the field by means of a network of wireless soil moisture sensors. Hedley and Yule (2009b) found that daily soil water content mapping could be utilized to manage a variable rate irrigation system.

2. Root zone soil hydrology

2.1 Pedotransfer functions

The most widely accepted method to fulfill the need for soil water retention information is called a pedotransfer function (PTF). PTF concept was introduced by Bouma and van Lanen (1987) as a mathematical relationship between easy collected soil data and hard to measure characteristics. In fact, PTF allows us to transfer the data we have to data we need. The majority of the established PTFs target either saturated hydraulic conductivity or the water retention curve (WRC) as desired output. Similar to most of the empirical models, PTFs were developed using regression techniques during the 1980s while the data mining (DM) procedures became more prevalent in the 90s. Artificial neural networks (ANNs) were the first DM method that was utilized by Schaap and Bouten (1996) for developing a PTF. ANNs became standard tools to model complex systems in the 1990s. ANNs are able to capture the behavior of a system by adjusting their parallel components which provide them with very flexible structures. ANNs are often called "universal function approximators" because they can estimate any continuous function to any desired degree of accuracy. This characteristic makes ANNs very appropriate for establishing PTFs (Pachepsky and Rawls, 2004; Haghverdi et al., 2010, 2011). The successful use of ANN-PTF encouraged researchers to examine other DM algorithms. Group Method of Data Handling (Pachepsky and Rawls, 1999), *k*-nearest neighbor techniques (Nemes et al., 2006a, 2006b) and support vector machines (Lamorski et al., 2008; Twarakavi et al., 2009) are among the most recent data mining methods which were applied to derive PTFs.

Traditionally there are two types of PTFs for estimating water retention data: point and parametric PTFs. A point PTF predicts the water retention data at some predefined water retention points on the curve (such as water retention at field capacity and permanent wilting point) while a parametric PTF is able to predict the whole WRC. A parametric PTF, in fact, estimates the parameters of a soil hydraulic equation, mostly the one by van Genuchten

(1980), and those parameters are then used to predict the WRC through the chosen equation. Tomasella et al. (2003) showed that point PTFs are more accurate than parametric PTFs based on the van Genuchten equation. The desire to be able to predict the entire WRC encouraged researchers to establish parametric PTFs, yet there are some drawbacks associated with this type of PTF. Parametric PTFs predetermine which equation the user ought to use, which is for most PTFs the van Genuchten model. There is no guarantee that the real shape of the WRC is similar to the shape of the chosen equation for all soils. Furthermore, Minasny and McBratney, (2002) mentioned overparametization as a potential problem resulting from van Genuchten-based-parametric PTFs. Recently, a new type of ANN- PTF has been introduced by Haghverdi et al. (2012, 2014b) which is able to predict the entire WRC without the need for a predetermined soil hydraulic model. Haghverdi et al (2012) showed that this model was able to perform as well as classical PTFs and sometimes outperform them. One of the highlighted features of this applied model is its unique structure which makes it able to establish a PTF for soils that have uneven available water retention data.

In summary, well-known PTFs (e.g. Rosetta by Schaap et al., (2001) and kNearest by Nemes et al., (2008)) are useful tools to convert basic soil information to soil hydraulic data with a reasonable degree of accuracy. However, there are two major concerns hampering their application for providing enough data for variable rate irrigation: (i) PTFs are not spatial tools and therefore cannot generate a map and (ii) the routine input predictors of PTFs, e.g. detailed soil texture information, bulk density and sometimes organic matter content, cannot be classified as easily collected data needed for management of variable rate irrigation.

2.2 Soil spatial prediction

Grid sampling and subsequently using geostatistical methods is the most accepted way to map the soil hydraulic properties. In this process, PTF would convert the basic soil properties to soil hydraulic properties in a grid format while an interpolation technique will be used to convert discrete points to a map. Kriging is one of the most acceptable interpolating techniques to determine the possible spatial correlation among the data. A remarkable advantage of kriging over other interpolation techniques is that the former provides the semivariogram of the target attribute(s) as well as associated uncertainty. An alternative option is to use DM procedures to model the spatial correlation among the data. Behrens et al. (2005) found ANNs a notable tool for digital soil mapping. Fu et al. (2011) clarified the precise spatial prediction of dissolved organic carbon using ANNs comparing with spatial

regression kriging. In another study, Motaghian and Mohammadi (2011) found kriging combined with ANNs performed best for spatial estimation of saturated hydraulic conductivity from terrain attributes.

The density and distribution of the grid samples would significantly affect the performance of all interpolation techniques. A relatively large number of samples are needed to establish accurate high-resolution soil maps (Zhao et al., 2009) which is costly and time consuming (Saey et al., 2009). In practice, however, it is not feasible to take lots of samples. Including extra input variables is a possible solution to obtain high accuracy and to minimize the field work. Recently, Levi (2012) tried to estimate high resolution physical soil properties using a raster approach via digital soil mapping. He utilized remotely sensed auxiliary data and a variety of geostatistical techniques. Afterwards, a PTF was used to predict water retention at the landscape scale. Sharma et al. (2006) utilized topography and vegetation attributes for developing PTFs. They found some improvement with certain input combinations of basic soil properties along with the topography and vegetation attributes. In another study, Leij et al. (2004) added the topographical attributes to the basic input predictors of PTF for predicting water retention along a hillslope transect. They observed up to 26% improvement in modeling results due to additional input variables.

2.3 On-the-go sensors

Readily available dense data can be collected via on-the-go sensors which are able to gather the information while traversing a field. Ground penetrating radar and soil Veris machines are the most widely used means to map the soil variability in PA applications. Yield maps also can be utilized to distinguish spatial soil differences. Available means to collect ground penetrating radar are more suitable to collect data in small scale projects. ECa maps through Veris machines and yield data are easily collected and dense enough to be used to generate high resolution maps, including large scale maps, and hence in theory can help to minimize or even to avoid the soil sampling work. The practical challenge is how to convert these data to soil hydraulic maps.

Yield maps are affected by lots of temporal and spatial attributes so it is not easy to specifically quantify the role of soil hydraulic properties on them. Florin et al. (2011) examined an inverse meta modeling approach to estimate water retention via yield maps. They found that having a number of yield maps together with a crop growth model can aid estimation of high resolution soil hydraulic properties maps. EC is an indicator showing salinity in arid regions. However, in humid region, for which salinity is not a major factor, EC may provide some useful information about clay percentage, cation exchange capacity (CEC) and water content (Sudduth et al. 2005).

Nevertheless, the reported relationships between ECa data and soil hydraulic properties are not consistent. Duncan (2012) showed that the depth to sand layer and soil available water in an irrigated cotton field in west TN was strongly related to EC data. Sudduth et al. (2005) tried to relate ECa to soil properties across the north-central USA. They showed a relatively high correlation between ECa with clay and CEC across all study fields but this correlation was only available in a limited number of fields between soil moisture and ECa. McCutcheon et al. (2006) reported a weak temporal uniformity in ECa data when mapped in a dryland field over time. They found volumetric soil water content was the dominant factor which affects the ECa variability in both spatial and temporal manners. Some studies showed that it is better to conduct ECa mapping under wet conditions because the relationship between ECa and texture is more pronounced near field capacity (Auerswald et al., 2001; Taylor et al., 2003).

There are few studies relating spatial ECa data to soil hydraulic properties. Saey et al. (2009) developed a simple regression-based-PTF to convert ECa data to clay percentage under a non saline condition. They found a good correlation (R^2=0.81) for their PTF. Abdu et al. (2008) predicted soil texture and WHC in the subsurface soil of a 38 ha watershed using ECa data by deriving a simple regression equation. The input predictors were normalized ECa and the latitude (northing). Their model performed well having an R^2 = 0.86 and a root mean square error (RMSE) = 4.4 % for predicting clay percentage and an R^2 = 0.75 and a RMSE = 0.01 m^3 m^{-3} for predicting WHC. The predicted values of soil WHC ranged from 0.07 cm^3 cm^{-3} to 0.22 cm^3 cm^{-3} across the watershed, while only a uniform value of 0.13 m^3 m^{-3} could be estimated by means of soil survey maps. Abdu et al. (2008) emphasized the need for additional studies to appropriately relate ECa data into hydrological attributes.

There are many approaches for developing PTFs yet the options are more limited when the goal is to produce a high resolution soil hydraulic property map. The evidence in the literature seems to point toward dense data sets (such as on-the-go sensing of ECa, yield maps and topography) with calibration and verification from limited soil core data. Therefore, a next generation of PTFs should be developed as site-specific models capable of using easily obtainable data to predict soil hydraulic properties.

3- Management zones delineation

In conventional agriculture, each field is considered as a uniform unit, by purposely ignoring the heterogeneity across the field. Thus, decision-making is based on an estimation of average conditions. The motivation for site-specific farming was first addressed by researchers during late the 80s and early 90s (Arslan and Colvin, 2002). As

such, PA methodology is a way to look at field management by taking the within field variation into account and involving that variability into management decisions. There are different definitions of the PA available in literature. The definition of PA used in this study is a farming system which uses information technology to do site-specific crop management in which decisions on resource application are modified in regard to within field variation of components such as soil, water and crop (Whelan and Taylor, 2013). Variable-rate application is required where the within-field heterogeneity is caused by both temporal and spatial variation of a variety of factors such as climatic, topographic and biological (Córdoba et al., 2013). Map-based and sensor-based approaches are two major methods to practice variable-rate application. In the map-based method, application maps are prepared using site-specific information such as yield data and soil data prior to implementation. In the sensor-based method, a real time decision on application rate is made using data collected via sensors and pre-developed application algorithms (Thöle et al., 2013).

One of the steps in PA is to delineate management areas within fields where it is expected that applying identical treatments will cause significant yield differences. Management zones in precision farming are field areas with similar soil-landscape attributes (Schepers et al., 2004). A corollary expectation is that varying the treatment of management zones will facilitate optimizing yield. Other goals such as protecting the environment, saving resources, and keeping agriculture sustainable may also be achievable through precision farming. Moral et al. (2010) used regression kriging, principle component analysis (PCA) and fuzzy cluster classification to delineate management zones using soil texture information and electrical conductivity as ancillary data. There are several methods to delineate management zones. Applying unsupervised clustering techniques and user-defined zoning are the main procedures. Clustering techniques group similar data points (cells) into distinct classes. Methods such as *k*-means and fuzzy *k*-means have been widely used to identify management zones (Córdoba et al., 2013).

A field can be zoned based on a single soil-crop variable or multiple attributes which are expected to affect yield (Khosla et al., 2010). Yield maps, topography, aerial photographs, canopy images and ECa are among suggested attributes to delineate management zones. Yield maps are useful sources of information reflecting within-field variation. However, some difficulties have been reported to delineate zones solely by yield maps (Khosla et al., 2010). Temporal inconsistency among yield maps from year to year is probably the main cause of this problem. Schepers et al. (2004) reported that temporal climate variability in an irrigated cornfield significantly affected yield spatial variability from year to year. Combining yield data with other ancillary information or averaging yield data

over years help better explain spatial variation and in turn can provide more trustable zones. Promising results have been reported by the studies that have utilized several years of yield data to create management zones. However, stability of such zones has to be tested in each individual field (Khosla et al., 2010). Calculating temporal variance can be helpful to verify zone stability (Basso et al., 2012). There are some methods to screen potential attributes regarding their importance. Hornung et al. (2006) suggested assigning different weights to individual layers based on their importance to the variation in crop yield. PCA can be utilized to linearly transfer original variables to new, independent variables. Researchers apply PCA to understand the characteristics of a data set along with the relative importance of each individual variable (Fraisse et al., 2001). Finally, the most important PCs will be used to delineate management zones.

Necessary information to analyze spatial soil variability and in turn to delineate management zones can be obtained via multiple methods: using soil survey maps, sampling and interpolation, on-the-go sensors and remote sensing techniques. Application of remote sensing is especially attractive because it is noninvasive and relatively inexpensive (Schepers et al., 2004). ECa data are relatively easy and inexpensive to collect, and have thus become one of the most widely studied attributes to zone fields. However, reported results in the literature on efficiency of ECa for zoning are not consistent because of several factors affecting ECa, though these factors do not necessarily influence crop productivity. ECa alone may not be suitable under all cropping systems but there is a potential for using it combined with other soil and crop attributes (Khosla et al., 2010).

The software MZA was developed by Fridgen et al. (2004) for delineating management zones. The MZA uses a fuzzy c-means unsupervised clustering algorithm for dividing a field into management zones and tests the result to evaluate how many zones to create in a given field. The optimum number of the zones and their shape is dependent on the target input and available equipment. For instance, there is more freedom to precisely apply fertilizer even in very small regions in a field but this is not the case with variable rate irrigation. Zhang et al. (2010) developed a web-based decision support system for zoning using satellite imagery and field data. More recently, Cid-Garcia et al. (2013) proposed an integer linear programming management zone delineation method to make rectangular shaped zones which facilitates the work and operation of machinery.

Delineating irrigation management zones is challenging because there are many parameters that simultaneously affect crop available water and crop water requirement. There are only a few studies on deriving management zones for variable rate irrigation. For instance, Jiang et al. (2011) used the physical properties of soil as

the data source to delineate irrigation zones. They utilized management zone analysis software which uses a fuzziness performance index and normalized classification entropy for identifying the least number of subzones. In another study, Bereuter (2011) studied zoning techniques on irrigated corn in Nebraska. Nine soil and landscape attributes were chosen as potential factors for zoning. Results showed that different combinations of selected attributes were suitable for zoning at different sites.

In a simplified conceptual framework, the right amount of water within the root zone would maximize the yield. However, applying deficit or excess irrigation may cause either drought stress or aeration problems and both can reduce the yield potential. As such, both the soil and the crop are crucial parameters to consider. Different types of soils hold varying amounts of water with only a portion of that water being readily available for the crop. Some crops may tolerate more severe drought stress and/or poorer soil aeration than the others and some may use excess water to grow vegetatively rather than producing more yields from reproductive parts of the plant. Therefore, the following knowledge is required to delineate variable rate irrigation zones: (i) the spatial variation of soil hydraulic properties as well as (ii) the response of a target crop to different irrigation regimes for each soil type.

4. Water production functions

Modeling crop production based on water availability is an important step toward optimizing irrigation and therefore is a critical issue for farmers, governmental agencies and consulting companies. Unfortunately, it is not easy to obtain a precise estimation of yield at the beginning of a cropping season because a variety of factors, such as water usage, weather and rainfall variation, seed quality, topographic attributes, soil properties and pest control, simultaneously affect crop yield in a very complex manner. That is why, in most studies researchers only focus on understanding the effect of one or a few attributes on yield while trying to keep other parameters unchanged in both spatial and temporal manners. However, this concept of yield modeling has been changing over time as practices move from conventional farming to PA. In conventional agriculture, one predicts average yield across the field and ignores yield variation within the field. In PA, however, the goal is to consider and to manage within-field yield variation. As a result, understanding and modeling the effect of different parameters on crop yield in a spatiotemporal scheme becomes a crucial research topic in PA.

4.1 Yield prediction in conventional agriculture

Yield prediction in conventional agriculture is a prevailing subject of interest. In a majority of studies yield variation is simulated either by crop growth models or empirical equations. These are useful methods to quantify the irrigation effect on crop yield while there are inherent advantages and disadvantages associated with each of them.

There are lots of crop models available, mathematically relating yield to some parameters such as soil, water, weather and fertilizer. The crop models are divided to three branches with regard to the resources used for estimating the crop growth rate. The resources are water, solar radiation and carbon dioxide. Aquacrop (Steduto et al., 2009; Raes et al., 2009) and CropSyst (Stöckle et al., 2003) are two well-known water driven models. There are some studies on cotton yield prediction using these models. Sommer et al. (2008) evaluated the performance of the CropSyst model to simulate cotton yield. García-Vila et al. (2009) utilized AquaCrop to optimize deficit irrigation of cotton. They concluded that AquaCrop is a useful model to help managers identifying the optimum irrigation decisions. Farahani et al. (2009) also studied full and deficit irrigation of cotton by means of AquaCrop model. These models were shown to be precise yet they need lots of input data. In practice, running these models is time consuming because of all the necessary input parameters; therefore, their application is mostly limited to research projects rather than practical situations.

The regression-based empirical equations are alternative options for identifying crop yield responses to environmental and management parameters. The term production function (PF) may be assigned to any mathematical relationship between crop yield and input components such as water, fertilizers and energy (De Juan, 1996). Almost all of the available PFs require crop water use as an independent variable. The PFs predict total dry matter (or marketable product of each crop as dependent variable while the independent variables are transpiration (T), evapotranspiration (ET) or amount of applied water during irrigation (IW). These functions are distinguished into two groups based on independent variable: (i) crop water production functions (CWPFs) which use ET and (ii) water production functions (WPFs) which use IW. IW may consist of different components such as crop water requirement, pre-plant irrigation, leaching requirement, and rainfall (Igbadun et al., 2007). From a temporal point of view, PFs are divided into seasonal and dated PFs. Seasonal PFs consider the total in season IW or ET as independent variable but dated PFs divide the growing seasons into different stages and consider an individual independent variable for each of the stages. In dated PFs (Rao et al., 1988) the role of water at different growth stages varies but are related to the others. The well-known dated PFs connect the independent variables either by

adding them up or multiplying them. A multiplicative PF would predict yield failure if lack of ET or IW occurred even at a single growth stage but an additive PF is formulated differently. Recently, some studies have aimed to revisit and rebuild the concept of PF. Some examples are: a WPF for water logging stress on corn (Kuang et al., 2012) and rice water – fertilizer PF (Ai-hua et al., 2012). Tong and Guo (2013) also tried to involve an estimation of uncertainty in CWPF along with the optimal allocation of water resources in an irrigation area. There are some studies on cotton yield prediction by means of PFs. Wang et al., (2007) derived cotton and wheat water-salinity PFs. Dinar et al. (1986) derived cotton CWPF under saline conditions in California. They stated that PFs were good models to improve irrigation management.

Classical PFs are useful tools for irrigation management and economic analysis of yield reduction due to deficit irrigation, but there are some shortcomings associated with them. PFs, like other regression-based equations, are relatively easy to build but are mostly linear and not powerful enough to model complex ecological systems (Dai et al. 2011). Recent studies (e.g. Fortin et al., 2010; Haghverdi et al., 2014a) are looking for more robust and non-linear techniques to predict yield. In most of agricultural related studies, machine learning methods became the favorable tools because they are powerful empirical tools for modeling complex systems. Data mining could be defined as the process of capturing important and useful information from large data sets (Mucherino et al., 2009). An extensive review on data mining methods and their application in agricultural related studies was gathered by Huang et al. (2010). Machine learning algorithms were employed to some extent to predict yield of different crops. Fortin et al (2010) used ANNs for predicting potato tuber growth as well as its in-field variations in Canada. They reported that with an ANN model using enough data they can precisely model site-specific tuber growth. Dai et al. (2011) adopted ANN and multi-linear regression models to simulate the response of sunflower yield to soil moisture and salinity. ANNs appeared to be the model with higher precision than regression. They concluded that ANN is a useful tool for modeling relationship between crop yield and soil moisture and salinity at different crop growth periods. Haghverdi et al. (2014a) derived some novel WPFs by comparing data mining-based methods with traditional regression procedures. They utilized ANN and decision tree as modeling algorithms for deriving water salinity PF for spring wheat. Haghverdi et al. (2014a) suggested data mining methods to derive WPFs since they are accurate and flexible in their structure.

In summary the balance between simplicity and accuracy of the models should be considered as a critical issue when applying them in broad practice (Farahani et al., 2009). Empirical PFs are only valid for a single crop at a

specific location. Using a previously derived PF is not recommended for a new condition because there are significant temporal and spatial differences among crops and climatic zones (Al-Jamal et al., 2000). The performance of PFs for different crops and different locations should be tested carefully in advance to use them for irrigation planning and establishment of water management plans (Igbadun et al., 2007). On the other hand, sophisticated crop growth models are reliable tools only if the massive input variables are available at a reasonable level of accuracy. However, in practice, poorly calibrated crop growth models are not a better option than empirical tools.

4.2 Yield prediction in precision agriculture

Yield mapping using yield monitors is one of the most widely implemented PA technologies. The first commercialized yield monitor was released in 1992 for a grain combine while after 6 more years the first cotton yield monitor became commercially available. Yield monitoring systems generate spatially dense data during harvesting as a combine harvester moves through a field. The data is simultaneously geo-referenced by means of an appropriate GPS device and can be displayed and post-processed via a suitable GIS system. Producers utilize yield monitoring systems mostly for grains, oilseeds and cotton. In the US, 28% of maize areas in 2005 and 22% of soybean areas in 2002 were harvested using yield monitoring systems (Griffin, 2010). However, there are lots of potential users who have not adopted yield monitor technology because the methods for making use of these extensive data sets are not yet adequate. Yield monitoring and mapping is not only output from the production system but a trustable source of information for understanding spatial and temporal variability (Guo et al., 2012). Yet interpreting it is challenging because yield variability is caused by a variety of factors. Both crop growth models and site-specific empirical equations are adoptable from conventional farming to predict yield in PA and their application will be reviewed in the following segment.

There are several studies on applying crop growth models in PA. For example, Florin et al. (2011) employed an inverse-modeling step to estimate the soil physical properties over a grain farm in Australia using yield map data. They then used soil data in a meta-modeling process to predict spatially- and temporally-dense yield. They found implicit information, during modeling, about the interaction of climate with soil, crop and landscape which they stated needs to be identified. In another study, Basso et al. (2007) used crop growth models to analyze spatiotemporal stability of maize, soybean and wheat yield. They utilized 5 years of yield monitor data to form some

management zones. Afterwards, 14 years of simulation data were generated to analyze the effects of climate on the spatial and temporal variation. Basso et al. (2007) stated that combination of crop growth models and GIS could be useful to identify the temporally stable zones. Link et al. (2006) also applied the CERES-Maize crop growth simulation model along with a decision support system to study groundwater nitrate concentration. They separated the field of study into 30 grid cells and ran the model at each point. They were able to explain approximately 60% of spatial and temporal yield variability by their method.

Some site-specific empirical models also were developed for yield prediction at a high spatial resolution. Wong and Asseng (2006) tried to predict wheat yield by means of available soil water which was estimated from ECa data. They derived different equations based on the quantity of the other related factors such as rainfall, initial soil moisture and nitrogen application. The promising result of their study revealing the dominant factors causing spatiotemporal yield variation were interactions of seasonal rainfall, plant available soil water storage capacity and N fertilizer applications. There are also some studies on spatial prediction of cotton yield. Guo et al. (2012) tried to mathematically explain spatial and temporal variation of cotton yield on the Southern High Plains of Texas by means of ECa, soil brightness, and topography data. They observed that a combination of the above attributes could explain up to 70.1 % of cotton yield variability. Guo et al. (2012) mentioned the soil texture as one of the greatest factors affecting cotton yield. They found a relatively stable spatial pattern of yield over time, although yield and soil properties had a stronger relationship in dry years than wet seasons. There are some studies proving the positive correlation of yield with topographic attributes (Jigang and Thelan, 2004; Kaspar et al., 2004; Andales et al., 2007; Green et al., 2007). The modeling attempts to relate those attributes to yields varied from regression analysis (Jigang and Thelan, 2004; Andales et al, 2007) to more complex machine learning models (Green et al. 2007; Ruß and Brenning, 2010, Jiang et al., 2009).

Despite all of the advancements, it is difficult to find a yield modeling study in PA that focuses on irrigation effect on yield variation. Crop growth models are robust to deal with temporal changes but they are point-based and unable to cover spatial variation in their calculations. Dividing the field into homogenous sub-units, using interpolation techniques and model parameterization at a high spatial resolution are possible methods to solve this issue (Florin, 2008). However, all of these methods are time-consuming and costly and not practical in a majority of situations. Empirical PFs are easier to establish which make it more feasible to derive them.

As previously discussed, the primary problem with PFs in conventional farming is that as empirical models they are difficult to reuse in new places or even at the same place over time. This is a significant issue in conventional farming because producing enough data to derive site-specific PFs is not feasible. In fact, producing data is limited to classical irrigation research studies in which different irrigation treatments are applied, most likely in a multiple year experiment within a research station, following design of experiment (DOE) principles. This lack of data is automatically solved by yield monitoring systems where each producer can produce enough data to derive an empirical PF each year for a specific field. Furthermore, since the derived PFs are site-specific, transferring them to a new location is not an issue. Under these circumstances, PFs can become a practical tool to investigate site-specific irrigation-yield relationships. However there are new challenges that need to be adequately addressed. First, there is the question of how to incorporate spatial soil variation into the PF. The second is how to deal with significant temporal changes in weather patterns that obviously affect irrigation scheduling from year to year and require a dynamic modeling framework to update and incorporate each year's data. Thirdly, if the PF is going to be derived for each cropping season separately, how will the PF be used after the fact to improve irrigation management in upcoming years? Obviously, restructuring of the classical PFs is required to incorporate new site-specific input variables covering not only variation in ET/IW but variations in all of the other major parameters which spatially affect the yield such as soil hydraulic properties.

This dissertation consists of three parts; part 1: "*studying effective root zone soil hydrology for site-specific irrigation*", part 2: "*perspectives on cotton supplemental irrigation in west Tennessee*", and part 3: "*toward site-specific irrigation management in west Tennessee*". The objective of part one was to study soil physical and hydraulic properties within the effective root zone and to generate a high resolution field-scale soil available water map. The objective of part two was to investigate cotton response to varying irrigation regimes across different soil types. The objective of part three was to delineate irrigation management zones and quantify and investigate different variable rate irrigation scenarios.

References

Abdu, H., Robinson, D. A., Seyfried, M., and Jones, S. B. (2008). Geophysical imaging of watershed subsurface patterns and prediction of soil texture and water holding capacity. *Water resources research* 44(4).

Ai-hua, S., Shi-jiang, Z., Ya-fen, G., and Zhong-xue, Z. (2012). Jensen model and modified Morgan model for rice water-fertilizer production function. *Procedia Engineering* 28(0), 264-269.

Al-Jamal, M.S., Sammis, T.W., Ball, S., and Smeal, D. (2000). Computing the crop water production function for onion. *Agricultural Water Management* 46, 29–41.

Andales, A. A., Green, T. R., Ahuja, L.R., Erskine, R.H., and Peterson, G.A. (2007). Temporally stable patterns in grain yield and soil water on a dryland catena. *Agricultural Systems* 94, 119-127.

Arslan, S., and Colvin, T.S. (2002). Grain yield mapping: Yield sensing, yield reconstruction, and errors. *Precision Agriculture* 3(2), 135-154.

Auerswald, K., Simon, S., and Stanjek, H. (2001). Influence of soil properties on electrical conductivity under humid water regimes. *Soil Science* 166(6), 382–390.

Bajwa, S. G., and Vories, E. D. (2007). Spatial analysis of cotton (Gossypium hirsutum L.) canopy responses to irrigation in a moderately humid area. *Irrigation Science* 25(4), 429-441.

Basso, B., Bertocco, M., Sartori, L., and Martin, E.C. (2007). Analyzing the effects of climate variability on spatial pattern of yield in a maize–wheat–soybean rotation. *European Journal of Agronomy* 26(2), 82-91.

Basso, B., Fiorentino, C., Cammarano, D., Cafiero, G., & Dardanelli, J. (2012). Analysis of rainfall distribution on spatial and temporal patterns of wheat yield in Mediterranean environment. *European Journal of Agronomy* 41, 52-65.

Behrens, T., Forster, H., Scholten, T., Steinrucken, U., Spies, E.D. and Goldschmitt, M. (2005). Digital soil mapping using artificial neural networks. Journal of Plant Nutrition and Soil Science-Zeitschrift Fur Pflanzenernahrung Und Bodenkunde 168(1), 21-33.

Bereuter, A. M. (2011). Management Zone Delineation Techniques on Irrigated Corn in Nebraska. Thesis, University of Nebraska - Lincoln. http://digitalcommons.unl.edu/agronhortdiss/40.

Bouma, J., and van Lanen, J.A.J. (1987). Transfer functions and threshold values: From soil characteristics to land qualities. p. 106–110. In K.J. Beek et al (ed.) Quantified land evaluation. International Institute Aerospace Surv. Earth Sci. ITC publ. 6. Enschede, the Netherlands.

Carmona, M.A., Mateos, L., and Fereres, E. (2007). Riego defi citario del algodón. p. 51–61. In L. Mateos (ed.) Metodologías e instrumentos para la planficación y la gestión sostenible del riego en condiciones de escasez de agua. Proyecto MIPAIS, Cordoba, Spain.

Cid-Garcia, N. M., Albornoz, V., Rios-Solis, Y. A., and Ortega, R. (2013). Rectangular shape management zone delineation using integer linear programming. *Computers and Electronics in Agriculture* 93, 1-9.

Córdoba, M., Bruno, C., Costa, J., and Balzarini, M. (2013). Subfield management class delineation using cluster analysis from spatial principal components of soil variables. *Computers and Electronics in Agriculture* 97, 6-14.

Dai, X., Huo, Z., and Wang, H. (2011). Simulation for response of crop yield to soil moisture and salinity with artificial neural network. *Field Crops Research* 121(3), 441-449.

De Juan, J.D., Tarjuelo, J.M., Valiente, M., and Garcia, P. (1996). Model for optimal cropping patterns within the farm based on crop water production functions and irrigation uniformity I: Development of a decision model. *Agricultural Water Management* 31(1), 115-143.

Dinar, A., Knapp, K. C., and Rhoades, J. D. (1986). Production function for cotton with dated irrigation quantities and qualities. *Water Resources Research* 22(11), 1519-1525.

Duncan, H. A. (2012). Locating the variability of soil water holding capacity and understanding its effects on deficit irrigation and cotton lint yield. Master's Thesis, University of Tennessee. http://trace.tennessee.edu/utk_gradthes/1286.

FAO, F. (2013). Agriculture Organization of the United Nations, 2012. FAO statistical yearbook.

Farahani, H.J., Izzi, G., and Oweis, T.Y. (2009). Parameterization and evaluation of the AquaCrop model for full and deficit irrigated cotton. Agronomy *Journal* 101(3), 469-476.

Fereres, E., and Connor, DJ. (2004). Sustainable water management in agriculture. In: Cabrera E, Cobacho R, eds. Challenges of the new water policies for the XXI century. Lisse, The Netherlands: A.A. Balkema, 157–170.

Fereres, E., and Soriano, M. A. (2007). Deficit irrigation for reducing agricultural water use. *Journal of experimental botany* 58(2), 147-159.

Florin, M. (2008). Towards precision agriculture for whole farms using a combination of simulation modelling and spatially dense soil and crop information. Ph.D. thesis. http://hdl.handle.net/2123/3169.

Florin, M., McBratney, A., Whelan, B., and Minasny, B. (2011). Inverse meta-modelling to estimate soil available water capacity at high spatial resolution across a farm. *Precision Agriculture* 12(3), 421-438.

Fortin, J.G., Anctil, F., Parent, L-É., and Bolinder, M.A. (2010). A neural network experiment on the site-specific simulation of potato tuber growth in Eastern Canada. *Computers and Electronics in Agriculture* 73(2), 126-132.

Fraisse, C. W., Sudduth, K. A., and Kitchen, N. R. (2001). Delineation of site-specific management zones by unsupervised classification of topographic attributes and soil electrical conductivity. *Transactions of the ASAE* 44(1), 155-166.

Fridgen, J. J., Kitchen, N. R., Sudduth, K. A., Drummond, S. T., Wiebold, W. J., and Fraisse, C. W. (2004). Management zone analyst (MZA). *Agronomy Journal* 96(1), 100-108.

Fu, Y., Zeng, X., and Lu, X. (2011). Spatial prediction of dissolved organic carbon using GIS and ANN modeling in river networks. Seventh International Conference on Computational Intelligence and Security. IEEE. DOI 10.1109/CIS.2011.96.

García-Vila, M., Fereres, E., Mateos, L., Orgaz, F., and Steduto, P. (2009). Deficit irrigation optimization of cotton with AquaCrop. *Agronomy Journal* 101(3), 477-487.

Green, T.R., Salas, J.D., Martinez, A., and Erskine, R.H. (2007). Relating crop yield to topographic attributes using Spatial Analysis Neural Networks and regression. *Geoderma* 139(1-2), 23-37.

Guo, W., Maas, S.J., and Bronson, K.F. (2012). Relationship between cotton yield and soil electrical conductivity, topography, and Landsat imagery. *Precision Agriculture* 1-15.

Gwathmey, C.O., Leib, B.G., and Main, C.L. (2011). Lint yield and crop maturity responses to irrigation in a short-season environment. *Journal of Cotton Science* 15(1), 1-10.

Griffin, T. W. (2010). The spatial analysis of yield data. In Geostatistical Applications for Precision Agriculture (pp. 89-116). Springer Netherlands.

Haghverdi, A., Ghahraman, B., Khoshnud Yazdi, A.A., and Arabi, Z., (2010). Estimating of water content in FC and PWP in north and north east of Iran's soil sampels using k-nearest neighbor and artificial neural networks. Journal of Water and Soil 24 (4), 804-814 (In Persian).

Haghverdi, A., Mohammadi, K., Mohseni Movahed, S.A., Ghahraman, B., and Afshar M. (2011). Estimation of soil salinity profile in Tabriz irrigation and drainage network using SaltMod and ANN models. Journal of Water and Soil. 25 (1), 174-186 (In Persian).

Haghverdi, A., Cornelis, W.M., and Ghahraman, B. (2012). A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. *Journal of Hydrology* 442, 46-54.

Haghverdi, A., Ghahraman, B., Leib, B. G., Pulido-Calvo, I., Kafi, M., Davary, K., and Ashorun, B. (2014a). Deriving data mining and regression based water-salinity production functions for spring wheat (Triticum aestivum). *Computers and Electronics in Agriculture* 101, 68-75.

Haghverdi, A., Öztürk, H. S., and Cornelis, W. M. (2014b). Revisiting the pseudo continuous pedotransfer function concept: Impact of data quality and data mining method. *Geoderma* 226, 31-38.

Hedley, C. B., and Yule, I.J. (2009a). A method for spatial prediction of daily soil water status for precise irrigation scheduling. *Agricultural Water Management* 96(12), 1737-1745.

Hedley, C.B., and Yule, I.J. (2009b). Soil water status mapping and two variable-rate irrigation scenarios. *Precision Agriculture* 10(4), 342-355.

Hornung, A., Khosla, R., Reich, R., Inman, D., and Westfall, D. G. (2006). Comparison of site specific management zones: soil color based and yield based. *Agronomy Journal* 98, 405–417.

Huang, Y., Lan, Y., Thomson, S.J., Fang, A., Hoffmann, W.C., and Lacey, R.E. (2010). Development of soft computing and applications in agricultural and biological engineering. *Computers and Electronics in Agriculture* 71(2), 107-127.

Igbadun, H.E., Tarimo, A.K.P.R., Salim, B.A., and Mahoo, H.F. (2007). Evaluation of selected crop water production functions for an irrigated maize crop. *Agricultural Water Management* 94(1-3), 1-10.

Jiang, P., and Thelen, K. D. (2004). Effect of soil and topographic properties on crop yield in a North-Central corn–soybean cropping system. *Agronomy Journal* 96, 252–258.

Jiang, P., He, Z., Kitchen, N. R., and Sudduth, K. A. (2009). Bayesian analysis of within-field variability of corn yield using a spatial hierarchical model. *Precision agriculture* 10(2), 111-127.

Jiang, Q., Fu, Q., and Wang, Z. (2011). Delineating site-specific irrigation management zones. *Irrigation and Drainage* 60(4), 464-472.

Kaspar, T. C., Pulido, D. J., Fenton, T. E., Colvin, T., Karlen, D. L., Jaynes, D. B., and Meek, D. (2004). Relationship of corn and soybean yield to soil and terrain properties. *Agronomy Journal* 96, 700-709.

Kerby, T.A. (1986). Boll growth and development. Cotton Comments, vol. 21. Cooperative Extension, University OF California, Berkeley, pp. 1-4.

Khosla, R., Westfall, D. G., Reich, R. M., Mahal, J. S., and Gangloff, W. J. (2010). Spatial variation and site-specific management zones. In Geostatistical applications for precision agriculture (pp. 195-219). Springer Netherlands.

Kuang, W., Xianjiang, Y., Xiuqing, C., and Yafeng, X. (2012). Experimental study on water production function for waterlogging stress on corn. Procedia Engineering 28, 598-603.

Lamorski, K., Pachepsky, Y., Slawihski, C., and Walczak, R.T. (2008). Using support vector machines to develop pedotransfer functions for water retention of soils in Poland. *Soil Science Society of America Journal* 72(5), 1243-1247.

Leib, B. G., Jabro, J. D., and Matthews, G. R. (2003). Field evaluation and performance comparison of soil moisture sensors. *Soil Science* 168(6):396-408.

Leib, B. G., Perry, C., and Fisher, K. (2012). Section 6: Sensor based scheduling. Cotton Irrigation Management for Humid Regions. Cotton Incorporated, Cary, NC, Pp 21 to 31.

Leij, F. J., Romano, N., Palladino, M., Schaap, M. G., and Coppola, A. (2004). Topographical attributes to predict soil hydraulic properties along a hillslope transect. *Water Resources Research* 40(2).

Levi, M. R. (2012). Predictive soil mapping in southern Arizona's basin and range. PhD Dissertation, Department of Soil, Water and Environmental Science, University of Arizona. Pages: 207.

Link, J., Graeff, S., Batchelor, W.D., and Claupein, W. (2006). Evaluating the economic and environmental impact of environmental compensation payment policy under uniform and variable-rate nitrogen management. *Agricultural Systems* 91(1), 135-153.

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

Minasny, B., and McBratney, A.B. (2002). The Neuro-m method for fitting Neural Network parametric pedotransfer functions. *Soil Science Society of America Journal* 66, 352–361.

Moral, F. J., Terrón, J. M., and Silva, J. R. (2010). Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. *Soil and Tillage Research* 106(2), 335-343.

Morrow, M. R., and Krieg, D. R. (1990). Cotton management strategies for a short growing season environment: Water-nitrogen considerations. *Agronomy Journal* 82(1), 52-56.

Motaghian, H. R., and Mohammadi, J. (2011). Spatial estimation of saturated hydraulic conductivity from terrain attributes using regression, kriging, and artificial neural networks. *Pedosphere* 21(2), 170-177.

Mucherino, A., Papajorgji, P., and Pardalos, P.M. (2009). A survey of data mining techniques applied to agriculture. *Operational Research* 9(2), 121-140.

NASS. (2010). 2007 Census of Agriculture: Farm and Ranch Irrigation Survey, National Agricultural Statistics Service, USDA. http://www.agcensus.usda.gov/index.php.

National Research Council. (1997). Precision agriculture in the 21st century. Washington DC, USA: National Academy Press.

Nemes, A., Rawls, W.J., and Pachepsky, Y.A. (2006a). Use of the nonparametric nearest neighbor approach to estimate soil hydraulic properties. *Soil Science Society of America* 70:327–336.

Nemes, A., Rawls, W.J., Pachepsky, Y.A., and van Genuchten, M.T. (2006b). Sensitivity analysis of the Nonparametric Nearest Neighbor Technique to estimate soil water retention. *Vadose Zone Journal* 5:1222–1235.

Nemes, A., Roberts, R.T., Rawls, W.J., Pachepsky, Y.A., and van Genuchten, M.T. (2008). Software to estimate− 33 and− 1500kPa soil water retention using the non-parametric k-Nearest Neighbor technique. *Environmental Modelling and Software* 23(2), 254-255.

Oliver, M. (2010) Geostatistical Applications for Precision Agriculture, pp. 1-34, Springer.

Orgaz, F., Mateos, L., and Fereres, E. (1992). Season length and cultivar determine the optimum evapotranspiration deficit in cotton. *Agronomy Journal* 84(4), 700-706.

Pachepsky, Y., and Rawls, W.J. (Eds.). (2004). Development of pedotransfer functions in soil hydrology (Vol. 30). Access Online via Elsevier.

Pachepsky, Y., and Rawls, W.J. (1999). Accuracy and reliability of pedotransfer as affected by grouping soils. *Soil Science Society of America Journal* 63:1748–1757.

Pan, L., Adamchuk, V.I., Martin, D.L., Schroeder, M.A., and Ferguson, R.B. (2013). Analysis of soil water availability by integrating spatial and temporal sensor-based data. *Precision Agriculture* 1-20.

Parks, W.L., Overton, J.R., and Graves, C.R. (1978). Cotton Irrigation. Tenn. Agric. Exp. Sta. Bull. 581. Knoxville, TN.

Peng, S., Krieg, D. R., and Hicks, S. K. (1989). Cotton lint yield response to accumulated heat units and soil water supply. *Field Crops Research* 19(4), 253-262.

Pettigrew, W. T. (2004). Moisture deficit effects on cotton lint yield, yield components, and boll distribution. *Agronomy Journal* 96(2), 377-383.

Raes, D., Steduto, P., Hsiao, T.C., and Fereres, E. (2009). AquaCrop—the FAO crop model to simulate yield response to water II. Main algorithms and soft ware description. *Agronomy Journal* 101, 438–447.

Rao, N.H., Sarma, P.B.S., and Chander, S. (1988). A simple dated water-production function for use in irrigated agriculture. *Agricultural Water Management* 13(1), 25-32.

Ruß, G., and Brenning, A. (2010). Data mining in precision agriculture: management of spatial information. In Computational Intelligence for Knowledge-Based Systems Design (pp. 350-359). Springer Berlin Heidelberg.

Saey, T., Van Meirvenne, M., Vermeersch, H., Ameloot, N., and Cockx, L. (2009). A pedotransfer function to evaluate the soil profile textural heterogeneity using proximally sensed apparent electrical conductivity. *Geoderma* 150(3), 389-395.

Snowden, C., Ritchie, G., Cave, J., Keeling, W., and Rajan, N. (2013). Multiple irrigation levels affect boll distribution, yield, and fiber micronaire in cotton. *Agronomy Journal* 105(6), 1536-1544.

Schaap, M.G., and Bouten, W. (1996). Modeling water retention curves of sandy soils using neural networks. *Water Resources Research* 32(10), 3033-3040.

Schaap, M.G., Leij, F.J., and van Genuchten, M.T. (2001). ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *Journal of hydrology* 251(3), 163-176.

Schepers, A. R., Shanahan, J. F., Liebig, M. A., Schepers, J. S., Johnson, S. H., and Luchiari, A. (2004). Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. *Agronomy Journal* 96(1), 195-203.

Sharma, S.K., Mohanty, B.P., and Zhu, J. (2006). Including topography and vegetation attributes for developing pedotransfer functions. *Soil Science Society of America Journal* 70(5), 1430-1440.

Sommer, R., Kienzler, K., Conrad, C., Ibragimov, N., Lamers, J., Martius, C., and Vlek, P. (2008). Evaluation of the CropSyst model for simulating the potential yield of cotton. *Agronomy for sustainable development* 28(2), 345-354.

Steduto, P., Hsiao, T.C., Raes, D., Fereres, E. (2009). AquaCrop—the FAO Crop Model to Simulate Yield Response to Water I. Concepts and Underlying Principles. *Agronomy Journal* 101, 426–437.

Stöckle, C.O., Donatelli, M., Nelson, R. (2003). CropSyst, a cropping systems simulation model. *European Journal of Agronomy* 18, 289–307.

Sudduth, K.A., Kitchen, N.R., Wiebold, W.J., Batchelor, W.D., Bollero, G.A., Bullock, D.G., Clay, D.E., Palm, H.L., Pierce, F.J., Schuler, R.T. and Thelen, K.D. (2005). Relating apparent electrical conductivity to soil properties across the north-central USA. *Computers and Electronics in Agriculture* 46(1–3), 263-283.

Taylor, J. C., Wood, G.A., Earl, R., and Godwin, R.J. (2003). Soil factors and their influence on within-field crop variability; part II: spatial analysis and determination of management zones. *Biosystems Engineering* (Special Issue on Precision Agriculture), 84(4), 441–453.

Thöle, H., Richter, C. and Ehlert, D. (2013). Strategy of statistical model selection for precision farming on-farm experiments. *Precision Agriculture* 1-16.

Tong, F., and Guo, P. (2013). Simulation and optimization for crop water allocation based on crop water production functions and climate factor under uncertainty. *Applied Mathematical Modelling* 37(14), 7708-7716.

Tomasella, J., Pachepsky, Y., Crestana, S., and Rawls, W.J. (2003). Comparison of two techniques to develop pedotransfer functions for water retention. *Soil Science Society of America Journal* 67(4), 1085-1092.

Twarakavi, N.K.C., Simunek, J., and Schaap, M.G. (2009). Development of pedotransfer functions for estimation of soil hydraulic parameters using support vector machines. *Soil Science Society of America Journal* 73(5), 1443-1452.

Van Genuchten, M.Th. (1980). A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal* 44:892–898.

Wang, Y. R., Kang, S. Z., Li, F. S., Zhang, L., and Zhang, J. H. (2007). Saline water irrigation scheduling through a crop-water-salinity production function and a soil-water-salinity dynamic model. *Pedosphere* 17(3), 303-317.

Whelan, B., and Taylor, J. (2013). Precision agriculture for grain production systems. CSIRO PUBLISHING. Pages 208.

Wong, M. T. F., and Asseng, S. (2006). Determining the causes of spatial and temporal variability of wheat yields at sub-field scale using a new method of upscaling a crop model. *Plant and Soil* 283(1-2), 203-215.

Wösten, J.H.M., Y.A, Pachepsky., and W.J, Rawls. (2001). Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics. *Journal of Hydrology* 251:123–150.

Zhao, Z., Chow, T.L., Rees, H.W., Yang, Q., Xing, Z. and Meng, F.-R. (2009). Predict soil texture distributions using an artificial neural network model. *Computers and Electronics in Agriculture* 65(1), 36-48.

Zhang, X., Shi, L., Jia, X., Seielstad, G., and Helgason, C. (2010). Zone mapping application for precision-farming: a decision support tool for variable rate application. *Precision agriculture* 11(2), 103-114.

Part 2: Studying effective root zone soil hydrology for site-specific irrigation

Abstract

Irrigation has been rapidly growing in west Tennessee during the recent decade. Farmers tend to invest in center pivot systems to avoid crop yield loss due to unpredictable dry periods. The spatiotemporal variation in soil and weather conditions needs to be studied for irrigation scheduling. If spatial soil variation is significant, variable rate irrigation may be the optimum option. A detailed spatial understanding of soil available water content within the effective root zone is needed to optimally schedule irrigation in this region. However, it is impractical to directly measure soil basic physical/hydraulic properties at field-scale because of the time consuming and expensive nature of lab/in situ techniques. This study was carried out to investigate the field-scale variation of soil available water content and its effect on yield patterns. The field (73 ha) was sampled (100 samples) and apparent soil electrical conductivity data was collected. Soil basic information including sand, silt and clay percentages and bulk density as well as soil water content were measured at four different depths (i.e. 0-25 cm, 25-50 cm, 50-75 cm and 75-100 cm) across the field. Multiple modeling scenarios were examined to indirectly predict high resolution soil available water content maps within the effective root zone with the aid of soil apparent electrical conductivity as proximal data. Geostatistical analysis showed spatial variability in soil textural components and water content was significant and moderately correlated to the yield patterns. There was as high as four-fold difference between available water content of coarse-textured and that of fine-textured soils on the study site. The result indicated that ECa was a useful proximal data to investigate soil spatial variability in the field of study. There was a good agreement (RMSE = 0.052 cm^3 cm^{-3} and r = 0.88) between predicted and observed water contents.

Chapter 1: The spatial variation of soil physical properties within the effective root zone at a field scale in west Tennessee

(Submitted to *Precision Agriculture* for publication)

Amir Haghverdi*, Brian G. Leib, Wesley C. Wright, Tim Grant, Muzi Zheng, Phue Vanchiasong

Department of Biosystems Engineering & Soil Science, University of Tennessee, 2506 E.J. Chapman Drive, Knoxville, TN 37996-4531

* Corresponding author, ahaghver@vols.utk.edu, Cellphone: (865) 235-9694

Abstract

Irrigation has been rapidly growing in west Tennessee during the recent decade. Farmers tend to invest in center pivot systems to avoid yield loss due to unpredictable dry periods. The spatiotemporal variation in soil and weather conditions needs to be studied for irrigation scheduling. If spatial soil variation is significant, variable rate irrigation may be the optimum option. This study aimed to investigate field-scale soil spatial variation for a typical agricultural field in West Tennessee. The field (73 ha) was sampled and apparent soil electrical conductivity data was collected. Soil basic information including sand, silt and clay percentages and bulk density as well as soil water content were measured at four different depths across the field. Geostatistical analysis showed spatial variability in soil textural components and water content was significant and correlated to yield patterns. The result showed ECa was a useful proximal data to investigate soil spatial variability in the field of study.

Keywords: apparent electrical conductivity, variable rate irrigation, water holding capacity

1. Introduction

It is vital to understand the distribution of soil physical and hydraulic properties at the field-scale in order to enhance water use efficiency and evaluate the effects of different irrigation strategies on environmental quality. The soil physical and hydraulic properties inherently differ within or among fields in nature. In addition, the management practices may induce the spatial heterogeneity. Soil usually exhibits structured variation in space, meaning it is

likely to find closer samples more similar to one another than those further apart. Classical statistical methods are not suitable to study spatial attributes because they require the data to be independent. Geostatistical techniques have become standard tools to study soil spatial dependency and distribution and to predict soil properties at unsampled locations (Khosla et al., 2010).

The spatial heterogeneity of soil attributes in West Tennessee is likely substantial with one responsible factor being parent material where for example, alluvial soils and windblown loess deposits can be found across much of the region. The soil specifically varies in its physical properties governing the soil water holding capacity and readily available water for crops, which in turn affects yield patterns. The vertical arrangement and thickness of deposited sediments also may vary; it is not unusual to see in a single field a sandy layer deposited over silt loam and clay sediments on top of a sandy layer. This vertical variation may be due to several factors including parent material, internal drainage, physiographic location and magnitude of erosion (Graveel et al., 1989). The long term historical use of conventional tillage together with steep slopes has further increased the soil spatial heterogeneity (Duncan, 2012). Iqbal et al. (2005) stressed the importance of studying differences among the distribution of surface, subsurface and deep soil horizons where in depth soil variation is significant.

The need to avoid yield loss from drought stress along with increased commodity prices have caused a significant conversion of rainfed agriculture to irrigated production in the last decade in moderately humid west Tennessee (Duncan, 2012). Irrigation management in this region is challenging due to significant season to season variability in rainfall patterns. It is likely to find numerous soil types and thus drastic spatial changes in soil physical attributes under a single irrigation system. The vertical soil variation even further changes the pattern of crop available water within the root zone at different growth stages. Duncan, (2012) showed that there was no single optimum irrigation decision for all of the soil types in this region and suggested variable rate irrigation as a desired strategy to enhance water use efficiency and optimize yield. The first step towards this goal is to study field scale soil spatial variation and find practical methods to derive a high resolution map of soil physical and hydraulic properties within the effective root zone.

Crop yield has been proven to be strongly related to soil physical properties. For example, Wong and Asseng (2006) considered available soil water as an input predictor of wheat yield. They reported plant available soil water storage capacity as one of the dominant factors governing spatiotemporal yield variation. In another study, Guo et al. (2012) mentioned the soil texture as one of the greatest factor affecting cotton yield. They found a relatively stable

spatial pattern of yield over time although yield and soil properties had stronger relationship in dry years than wet seasons. Graveel et al. (1989) studied the response of corn to soil variation in west Tennessee. They observed a large yield difference between soils with different erosion levels emphasizing a need to distinguish between soil profiles with sandy versus silty texture layers.

Soil survey maps, soil sampling, on-the-go sensors and remote sensing from field, airborne, and satellite sensors are the most widely used methods to obtain information on the spatial distribution of soil attributes (Khosla et al., 2010). Soil sampling at field-scale provides valuable information on the spatial variation of soil attributes, but to collect this data has become laborious and expensive. Soil apparent electrical conductivity (ECa) is a proximal attribute which has created substantial interest for soil mapping and management zone delineation in precision agriculture. The ECa is measured in a simple and in-expensive way by inducing an electrical current into the soil while traversing a field. ECa may provide useful information on soil physical and hydraulic properties (Gooley et al., 2013) including soil texture and soil water content where salinity is not an issue as the main factor influencing it. However, there are some inconsistency in the literature on factors affecting ECa variation for non-saline fields (McCutcheon et al., 2006). This suggests the need to investigate the practical utility of using ECa for site-specific management in different regions.

Consequently, this study was carried out to investigate the degree of spatial heterogeneity of soil properties including soil texture, bulk density, ECa, and AWC at the landscape spatial scale of an irrigated cotton field in western Tennessee. The objectives were (i) to determine the degree of spatial variability of soil physical properties at different layers using geo-statistical techniques, (ii) to evaluate the usefulness of ECa as a proximal attribute to map soil variability and (iii) to assess the relationship of the spatial distribution of soil attributes to yield patterns.

2. Material and Methods

2.1 Study area

The study area (about 73 ha) was an irrigated agriculture field located in west Tennessee along the Mississippi river. There were two center pivot systems available for irrigation. The field has been planted in cotton, soybeans and corn during the past years. Figure 1-1 illustrates the long term variation in regional weather data. Rainfall is relatively high even in dry years. Temperature changes are less pronounced and to some extent inversely proportional to

rainfall. Farmers practice supplemental irrigation in this region since rainfall events usually are not temporally well-scattered to fulfill crop water requirement over the entire growing season.

2.2 Soil data collection and lab analysis

Two rules of thumb were considered based on Kerry et al. (2010) (i.e. sampling at half the semivariogram range and collecting at least 100 samples) to calculate the number of samples and to design the sampling scheme. To obtain a rough estimation of spatial dependency of soil properties, semivariogram of available proximal attributes, including space image of the bare soil and elevation information, was analyzed. A total of 86 samples were gathered on a grid sampling scheme where samples were about 100 m apart (i.e. half the mean semivariogram range of proximal attributes). The rest of the samples (=14) were randomly collected from underneath the center pivot circles. Field sampling occurred after rainfall events when soil water status was assumed to be close to the field capacity. One hundred undisturbed samples (100 cm deep) were collected by a truck mounted soil sampler between March 21 and 22, 2014. Figure 1-2 shows the locations of samples within the field of study.

During sampling, each core was visually assessed and divided into four sub samples. The depth of subsamples (diameter = 67 mm) was set to 25 cm increments (i.e. 0-25 cm, 25-50 cm, 50-75 cm and 75-100 cm) yet was adjusted in respect to the available horizons if needed. The soil texture was measured by the hydrometer method (Blake et al., 1986). Water content of subsamples were calculated by subtracting after-sampling from oven-dried weights. The soil dry bulk density (BD) was estimated as the oven dry weight to volume of each subsample. The ECa data were collected using a Veris 3100 (Veris Technologies, Salina, KS) instrument on March 20, 2014 with about 10 m and 20 m distance between adjacent points in a same row and adjacent rows, respectively. This instrument had six rolling coulters for electrodes and collected two simultaneous ECa measurements from shallow and deep layers.

2.3 Descriptive and spatial analysis of soil properties

The Microsoft Excel 2013 was used to descriptively analyze soil data. The soil texture triangle was plotted for each layer. The relationship between volumetric water content at the time of sampling and soil basic properties, i.e. sand, silt and clay percentages and bulk density, was investigated. The average depth across samples after adjustment were fairly close to 25 cm for all the layers (i.e. 24.45 cm, 22.84 cm, 24.92 cm and 24.88 cm from first to the last

increment, respectively). All of the analysis was done individually for four different layers (each approximately 25 cm thick) to explore within root zone soil variation where the word *layer* was used to distinguish among subsamples rather than the real soil horizons. The relationship between ECa data and soil physical information, obtained from soil samples, was studied. To match ECa and soil basic data, the ECa data were interpolated to each sample using ordinary kriging approach.

The spatial analysis was done using the ARCGIS 10.2.2 (ESRI, Redlands, California). To examine the spatial autocorrelation of the attributes, the semivariogram (equation 1) and Global Moran's I statistic (equation 2) were obtained as follow:

$$\gamma(h)=\frac{1}{2N(h)}\left\{\sum_{i=1}^{N(h)}[Z(x_i+h)-Z(x_i)]^2\right\} \quad (1)$$

where $\gamma(h)$ is the semivariance, h is interval class, $N(h)$ is the number of pairs separated by lag distance, $Z(x_i)$ and $Z(x_i + h)$ are measured attributes at spatial location i and $i+h$, respectively.

$$I=\frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{i,j}}\times\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{i,j}z_i z_j}{\sum_{i=1}^{n}z_i^2} \quad (2)$$

where z is the deviation of an attribute from its mean, $w_{i,j}$ is the spatial weight between ith and jth point and n is equal to the number of points.

The ordinary kriging was applied to generate maps using sampling and ECa data. Finally, the maps were visually assessed against each other.

2.4 Yield data collection and cleaning

Soybean and cotton yield data were available from 2009 and 2012 cropping seasons, respectively. The raw yield data were cleaned prior to drawing yield maps. First, a thematic map was generated using raw data to investigate the GPS tracks and yield pattern across the field. Then, the swath width, distance between the points, speed of harvester and change in speed were carefully monitored. The data greater/smaller than 3 times the standard deviation from the mean were assumed statistically unexpected and removed unless there was a scientific evidence acting against this assumption. The yield data was converted to maps by the ordinary kriging approach. The location of yield data did

not exactly match to the ECa and sampling locations thus the interpolated maps were used to obtain the expected yield values at the given locations.

3. Results and Discussion

3.1 Variation among soil layers

Table 1-1 contains descriptive statistics for selected soil properties. The BD had its highest mean value at the deepest layer while the mean value was almost identical among other layers. The compaction of fine particles due to movement of heavy machinery along with available macroporosities by root channels may affect vertical distribution of BD. The mean volumetric water content (VWC) decreased with depth while its standard deviation (SD) slightly increased. Higher water content in surface layer could be due to textural differences among layers and also rainfall events prior to sampling which built the moisture level up within top layers but perhaps did not penetrate to deeper layers. The mean sand percentage increased with depth which was inversely proportional to decline in silt and clay. The mean and SD of the deep ECa readings was greater than those of shallow readings. Standard deviation among deep ECa reading was almost twice as that of shallow readings. The same result was reported by Sudduth et al. (2005) on differences between SD and distribution of shallow versus deep ECa readings. Deep ECa readings were positively skewed hence was log transformed.

The soil texture drastically varied across the field such that almost the entire triangle was covered by the collected samples except silt and clay textures (Figure 1-4). There was a shift from fine to coarse textures by depth with sand as the primary particle increased. The sand had the highest absolute correlation with VWC (Figure 1-3). There was a weak negative correlation between BD and VWC showing soil texture was the prime attribute governing water content. There was a clear pattern in clay and silt percentage plots versus VWC; the majority of the samples with lower clay and silt content belonged to deeper layers (a cluster of black dots) while samples from shallower layers were more likely to have higher clay and silt content. The opposite was seen in sand versus VWC plot.

3.2 Spatial analysis of soil properties

Table 1-2 presents the semivariogram and Global Moran's Index parameters for selected attributes for each soil layer. The nugget effect, sill and range are the basic parameters of a semivariogram to describe spatial structure. The

nugget effect mostly represents sampling/measurement error and variation at scales smaller than sampling interval. The total variance is called sill and the range is the maximum distance at which variables are spatially dependent. A higher positive Moran's Index for an attribute indicates stronger spatial structure. The z-score changes in line with the Moran's Index. A z- score from -1.65 to 1.65 shows spatial pattern does not significantly differ than random. A z- score, less than -1.65 is an indicator of a dispersed process while a Z score greater than 1.65 displays a spatially clustered attribute.

The highest range did not belong to the same layer across soil properties. The average range varied from 200 m to 300 m among attributes, 2 to 3 times greater than sampling intervals. The percent of nugget ranged from 18 to 50 % among soil properties in the study of Iqbal et al. (2005) who investigated spatial variability of soil physical properties of alluvial soils in a 162 ha cotton field at the Mississippi. This was somewhat similar to what was found for all of the attributes except BD which reached a nugget percent as high as 73 percent. The z-scores revealed all of the attributes except BD within different layers had clustered patterns. BD only showed a clustered pattern at the third layer and had a random pattern at other layers.

Figure 1-5 shows maps interpolated by kriging. The white strip expanding from northwest to southeast of the field is a surface drainage pathway. There were three major sandy regions within the field of study at surface layer located at: (i) center of the center pivot at eastern part of the field, (ii) south of the field mostly outside of the irrigated zones and (iii) northwest part of the field. The sequence of sand maps from first to fourth layers illustrated how these coarse soil regions expanded across the field by depth such that sand covered the majority of the field in deeper layers. The sandy regions could be either river flood-induced sand boils or earthquake-induced sand blows.

The vertical arrangement of soil textures was not consistent across the field. The clay had its highest influence from 0-50 cm yet sand was the dominant particle from 50-100 cm. This depth to sand across field ranged from 15-75 cm with average depth of 40 cm for almost 40 % of the sampling spots. For the rest of the samples (60%) either there was not a clear immediate change from fine to coarse texture or sand appeared at the surface soil. The silt contributed highly in subsurface layers (25-75 cm) where it reached its highest quantity and SD (Table 1-1). The majority of the samples from subsurface layers (50-75 cm) with high silt content were compacted to some extent. This compaction was also projected in relatively higher BD values from the same layers (Table 1-1). The BD map of the third layer corresponded well to the textural pattern where higher BD matching coarse samples. However, it was

difficult to pick up a trend from the rest of the BD maps as was expected from the result of the spatial analysis (Table 1-2).

The VWC map of the surface soil clearly showed the sandy spots as regions with lower water content. Moving to the deeper layers, VWC maps almost exactly matched the pattern of sand maps. It occurred because coarse soils tend to dry out faster and also hold less water than fine texture soils. The soil water content is a dynamic property of soil with time. However, it is expected that one time measurement of water content across a field provides a useful insight to relative spatial pattern of soil hydraulic properties (Corwin et al., 2003). This suggests measured VWC may be mathematically transformable to soil water holding capacity and in turn into crop available water. Overall, VWC map for the entire profile (0-100 cm) was similar to those of individual layers. From variable rate irrigation point of view, however, minor differences may matter. It might be required to practice a dynamic zoning strategy considering available water for crop within its effective root zone during the growing seasons.

3.3 Effectiveness of ECa as a proximal attribute

The ECa maps tended to follow the same general spatial patterns as for soil basic properties and VWC (Figure 1-5). Table 1-3 illustrates the correlation coefficient between ECa values and soil basic data. The ECa data were moderately correlated to soil texture and VWC information. The lowest correlation was obtained between BD and ECa data. The correlation between shallow ECa and other attributes declined from layer 1 to 4 as expected while the opposite was true for ECa deep readings. Sudduth et al. (2005) showed that 90 % of the shallow and deep readings responses in Veris machines were obtained from the soil above the 30 cm and 100 cm depth, respectively. The sand increased with depth hence regions with high conductivity became less pronounced in deep ECa map as opposed to the shallow ECa map.

These result show that ECa was a useful surrogate map of both soil texture and water content for the field of study hence its application is suggested for fields with similar conditions in west Tennessee. Sudduth et al. (2005) studied ECa readings on 12 fields in 6 states of the north-central United States. They found a good relationship between ECa data and soil cation exchange capacity as well as clay content at different times and locations hence introducing the possibility of establishing a general calibration to relate cation exchange capacity and clay content to ECa readings. They found the most variation in ECa values in the Iowa fields, which had the widest range in soil texture, from loam to clay loam. In contrast, McCutcheon et al. (2006) reported water content as the main factor

influencing ECa readings in a dry land field but they did not find textural parameters as significant predictors of ECa. They found a weak correlation between VWC and clay content and introduced other factors such as elevation and organic matter as parameters that may govern water content. In theory, multiple factors including the relative fractions occupied by soil, water and air, geometry and distribution of particles, and soil solution attributes affect ECa (Friedman, 2005). The introduced current to measure the ECa of a soil, in fact, travels through liquid, soil-liquid and solid pathways (Rhoades et al., 1999). In this study site, variation in soil texture was the main factor governing spatial distribution of volumetric water content. This might be the main reason ECa performed well in the study site.

The result of this study demonstrated that ECa can guide the sampling schemes and help to minimize field works. It may even be feasible to make use of ECa information in modeling soil hydraulic properties through pedotransfer functions and co-kriging techniques. However, an accurate understanding of soil texture distribution and water content is crucial to interpret ECa maps. Brevik et al. (2006) studied temporal stability of ECa data with respect to soil water content. They observed a strong influence of water content on ECa readings and found that ECa's power to differentiate soils was proportional to soil moisture. They mentioned that soil water content should be reported as an essential part of ECa studies. In-depth monitoring and modeling of soil spatiotemporal water status in west Tennessee is recommended for future studies.

3.4 Soil heterogeneity and yield patterns

Figure 1-6 illustrates the interpolated yield maps. Table 1-4 shows the correlation coefficient between soil properties at different layers with yield data. The spatial pattern of yield maps was fairly similar to that of soil properties; generally, fine textured soils tended to be more productive than coarse textured soil due to their higher water holding capacity. The ECa data had a moderate correlation with both cotton and soybean yield. Guo et al. (2012) investigated the relationship between cotton yield and ECa in the Southern High Plains of Texas. They reported a consistent positive correlation, up to 0.29 for shallow and 0.44 for deep ECa readings, between yield and ECa across years. In another study, Crowin et al. (2003) found ECa to be a useful indicator of cotton yield in the San Joaquin Valley in central California. They also observed a positive correlation between clay content (0.36), water content (0.42) and yield and a less pronounced negative correlation for soil BD (-0.29). The results of this study supports their findings.

The available drastic variation in soil physical and hydraulic properties was an influential factor in spatial yield variation across the field. However, it is known that other attributes related to management, water, crop and climate affect crop yield in a complex manner. The crop yield maps per se only provide limited information about the influence of each attribute, so are not sufficient for recommending variable rate applications (Corwin et al., 2003). The simple correlation analysis and visual assessment of maps provides a general understanding of the importance of different soil properties on its productivity. However, the establishment of a mathematical relationship between yield and soil related parameters without adjusting for spatial autocorrelation is not suggested. Yield modeling and prediction in respect to crop, weather, soil and management conditions can provide useful insight to evaluate different variable rate scenarios for the field of study.

4. Conclusion

The purpose of this study was to explore soil physical and hydraulic variation in one field with typical soil types for west Tennessee. Analyzing yield against soil maps revealed that soil physical and hydraulic properties to a great extent influenced yield patterns. This is expected because plant available water is a function of soil water holding capacity. This suggests that variable rate irrigation is the appropriate irrigation scenario for this mixture of soils within a typical west Tennessee field. There is evidence in the literature showing ECa may not be as effective for some conditions. The findings of this study showed ECa was a good proximal attribute to understand spatial variation of alluvial soils in that region. There is another type of agricultural field in west Tennessee with totally different conditions where soil textural differences are minor yet slope and elevation differences affect infiltration and redistribution of water within root zone hence available water for plants. The findings of this study may not be transferable to this type of fields.

References

Blake, G.R., and Hartge. K.H. (1986). Bulk density. In: Klute, A. (Ed.), Methods of Soil Analysis. Part 1, second ed. Agron. Monogr. 9. ASA and SSSA, Madison, WI, pp. 363-375.

Brevik, E. C., Fenton, T. E., and Lazari, A. (2006). Soil electrical conductivity as a function of soil water content and implications for soil mapping. *Precision Agriculture* 7(6), 393-404.

Corwin, D. L., Lesch, S. M., Shouse, P. J., Soppe, R., and Ayars, J. E. (2003). Identifying soil properties that influence cotton yield using soil sampling directed by apparent soil electrical conductivity. *Agronomy Journal* 95(2), 352-364.

Duncan, H. A. (2012). Locating the variability of soil water holding capacity and understanding its effects on deficit irrigation and cotton lint yield. Master's Thesis, University of Tennessee. http://trace.tennessee.edu/utk_gradthes/1286.

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

Gooley, L., Huang, J., Pagé, D., and Triantafilis, J. (2013). Digital soil mapping of available water content using proximal and remotely sensed data. *Soil Use and Management* 30, 139-151.

Graveel, J. G., Fribourg, H. A., Overton, J. R., Bell, F. F., and Sanders, W. L. (1989). Response of Corn to Soil Variation in West Tennessee, 1957–1980. *Journal of production agriculture* 2(4), 300-305.

Guo, W., Maas, S. J., and Bronson, K. F. (2012). Relationship between cotton yield and soil electrical conductivity, topography, and Landsat imagery. *Precision Agriculture* 13(6), 678-692.

Iqbal, J., Thomasson, J. A., Jenkins, J. N., Owens, P. R., and Whisler, F. D. (2005). Spatial variability analysis of soil physical properties of alluvial soils. *Soil Science Society of America Journal* 69(4), 1338-1350.

Kerry, R., Oliver, M. A., and Frogbrook, Z. L. (2010). Sampling in precision agriculture. In: Oliver, M. A. (Ed), Geostatistical applications for precision agriculture (pp. 35-63). Springer Netherlands.

Khosla, R., Westfall, D. G., Reich, R. M., Mahal, J. S., and Gangloff, W. J. (2010). Spatial variation and site-specific management zones. In: Oliver, M. A. (Ed), Geostatistical applications for precision agriculture (pp. 195-219). Springer Netherlands.

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.

Rhoades, J.D., Corwin, D.L., and Lesch, S.M. (1999). Geospatial measurements of soil electrical conductivity to assess soil salinity and diffuse salt loading from irrigation. In: Corwin, D.L., Loague, K., Ellsworth, T.R. (Eds.), Assessment of Non-point Source Pollution in the Vadose Zone. Geophysical Monograph 108. American Geophysical Union, Washington, DC, USA, pp. 197–215.

Sudduth, K. A., Kitchen, N. R., Wiebold, W. J., Batchelor, W. D., Bollero, G. A., Bullock, D. G., Clay, D. E., Palm, H. L., Pierce, F. J., Schuler, R. T., Thelen, K. D. (2005). Relating apparent electrical conductivity to soil properties across the north-central USA. *Computers and Electronics in Agriculture* 46(1), 263-283.

Wong, M. T. F., and Asseng, S. (2006). Determining the causes of spatial and temporal variability of wheat yields at sub-field scale using a new method of upscaling a crop model. *Plant and Soil* 283(1-2), 203-215.

Appendix 1: Chapter 1 Figures and Tables

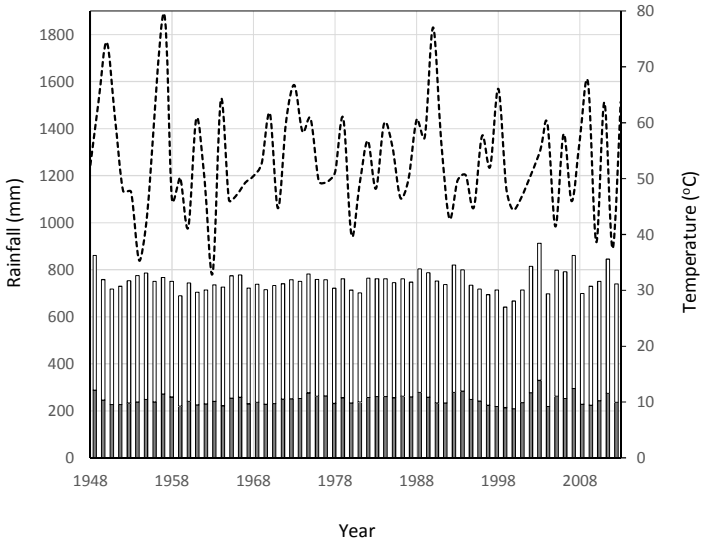

Figure 1-1. Long term variation in rainfall (dash line) and temperature (column) data in west Tennessee close to the area of study. Columns show average monthly mean minimum and average monthly mean maximum temperature data.

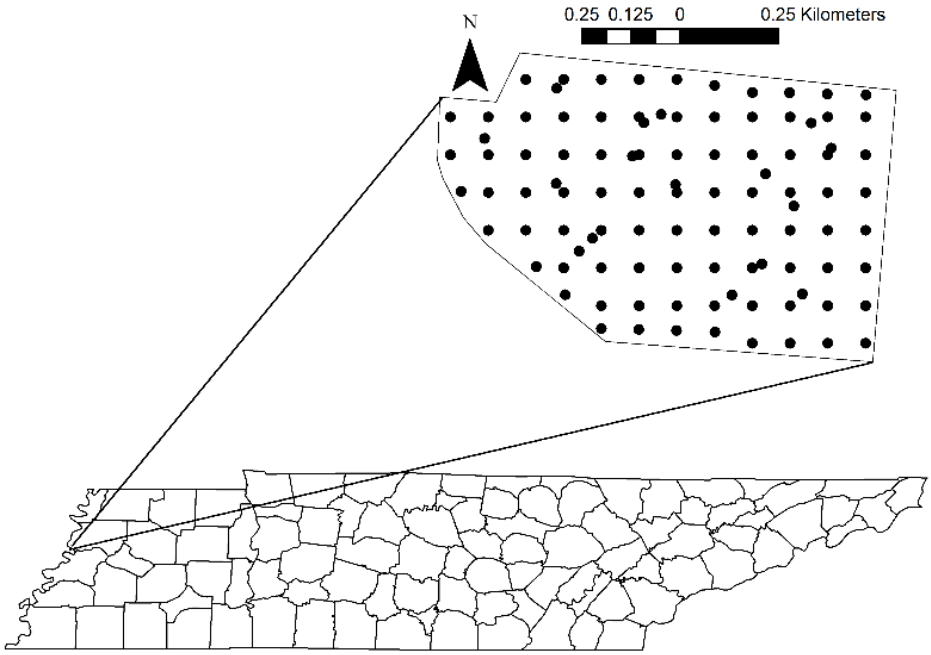

Figure 1-2. Sampling scheme in the field of study located in west Tennessee.

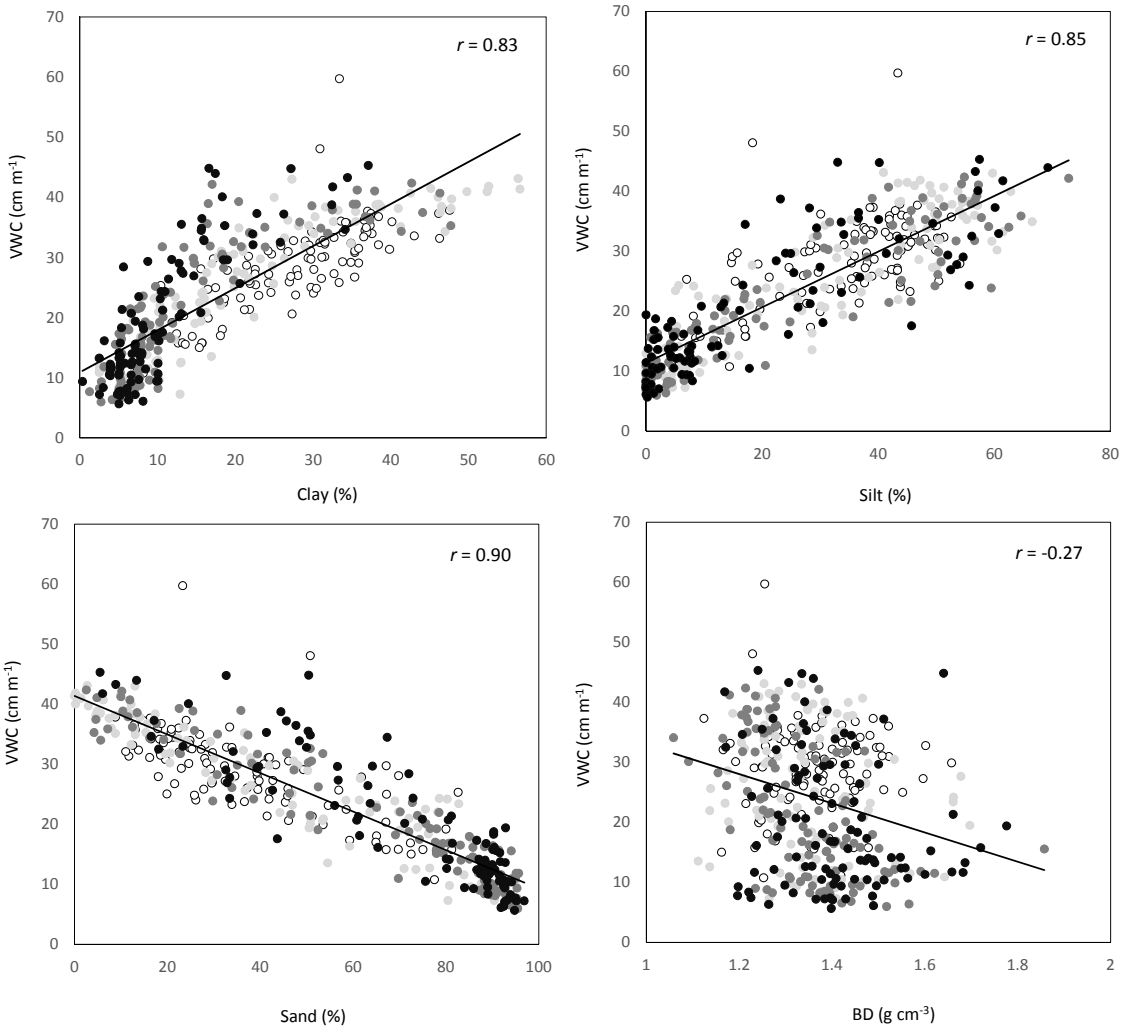

Figure 1-3. Scattering of basic soil properties against volumetric water content where the darker colors corresponds to the deeper layers.

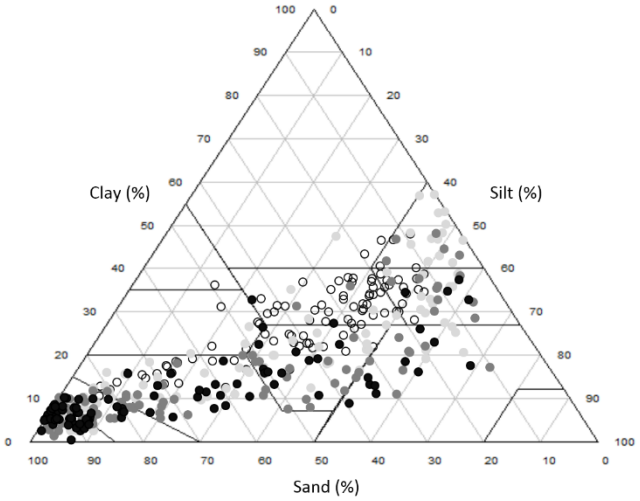

Figure 1-4. Soil textural distribution of samples at different layers where the darker colors corresponds to the deeper layers.

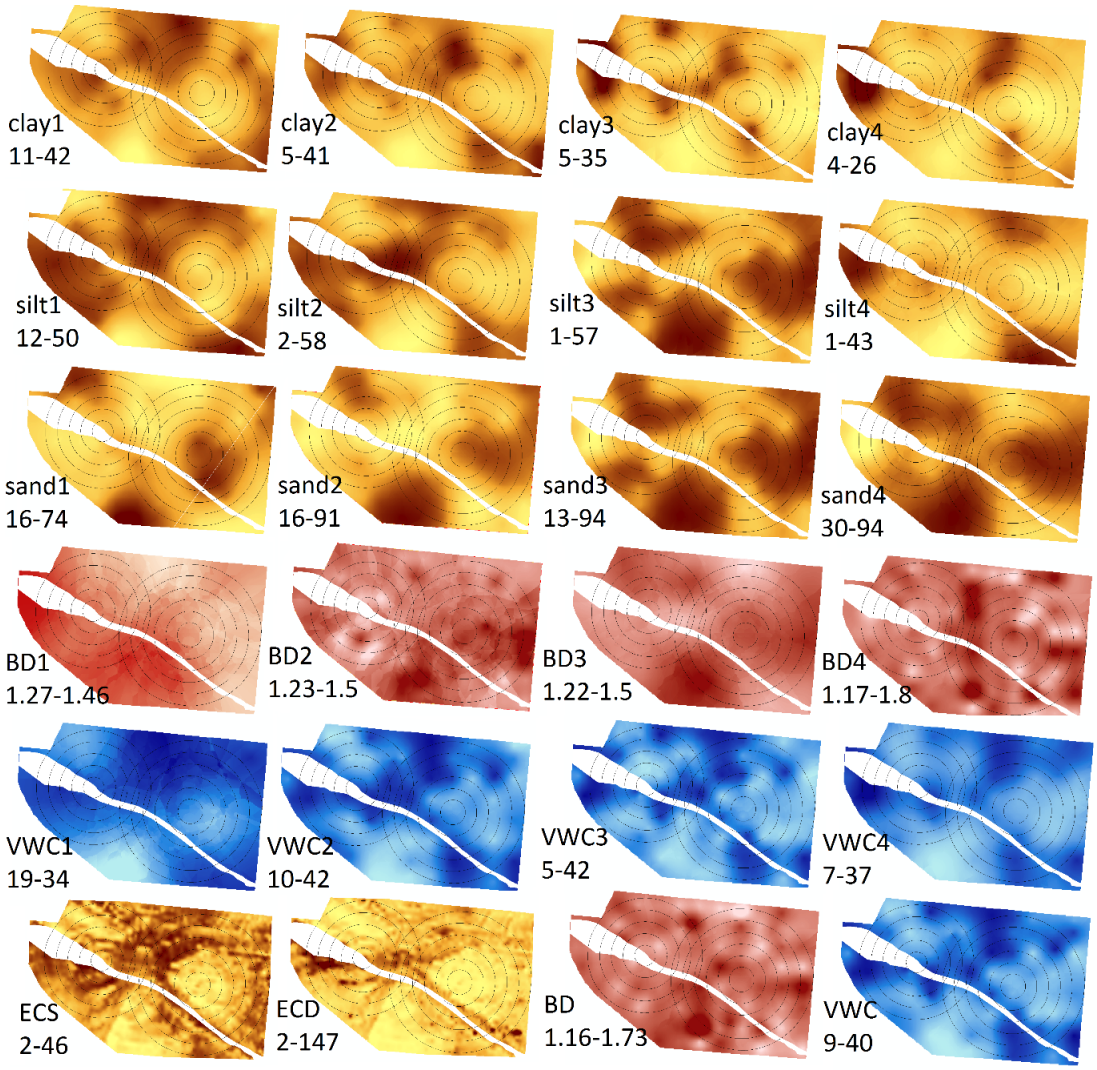

Figure 1-5. Interpolated maps by kriging method. Values underneath maps show the range of variation for each property. In all maps darker colors represent greater values. Silt (%), sand (%), clay (%); BD: soil bulk density (g cm^{-3}), VWC: volumetric water content (cm^3 cm^{-3}), ECS: apparent electrical conductivity shallow (mS m^{-1}), ECD: apparent electrical conductivity deep (mS m^{-1}).

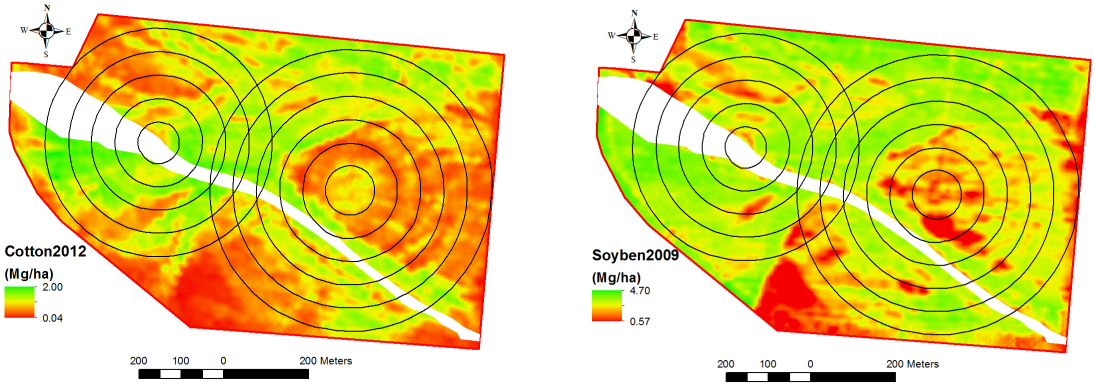

Figure 1-6. Cotton and soybean yield maps where green and red represent high and low yield, respectively.

Table 1-1. Descriptive statistics for selected soil properties from different soil sampling layers.

Variable	Layer	Min.	Max.	Mean	SD
BD, g cm^{-3}	1^{th}	1.12	1.66	1.36	0.10
	2^{nd}	1.11	1.70	1.35	0.12
	3^{rd}	1.06	1.86	1.34	0.12
	4^{th}	1.17	1.78	1.40	0.13
	total	1.06	1.86	1.36	0.12
VWC, cm^3 cm^{-3}	1^{th}	10.75	59.74	28.35	7.43
	2^{nd}	7.27	43.12	26.02	10.78
	3^{rd}	5.98	42.38	21.64	11.08
	4^{th}	5.67	45.32	20.18	11.15
	total	3.94	47.61	17.94	8.49
Sand, %	1^{th}	8.77	88.25	38.07	20.11
	2^{nd}	0.00	94.98	46.39	31.57
	3^{rd}	2.50	95.70	61.38	31.10
	4^{th}	5.46	96.86	69.90	26.09
Clay, %	1^{th}	7.37	47.56	27.55	9.04
	2^{nd}	2.50	56.60	22.18	14.17
	3^{rd}	1.26	47.72	14.27	11.44
	4^{th}	0.34	37.10	11.00	7.80
Silt, %	1^{th}	4.38	54.06	34.38	12.75
	2^{nd}	0.00	66.51	31.43	19.85
	3^{rd}	0.00	72.81	24.35	21.76
	4^{th}	0.00	69.23	19.10	19.83
ECa, mS m^{-1}	shallow	1.60	48.70	24.64	10.66
ECa, mS m^{-1}	deep	1.70	162.20	27.52	18.73

Table 1-2. Semivariogram and Moran's I parameters of soil properties for different soil layers.

Variable	Layer	Nugget	Sill	Range (m)	Moran's I	z-score
BD, g cm^{-3}	1^{th}	0.008	0.011	526	0.087	1.181
	2^{nd}	0.01	0.015	95	-0.086	-0.929
	3^{rd}	0.011	0.016	280	0.137	1.802
	4^{th}	0	0.017	100	0.091	1.221
	total	0	0.007	95	-0.007	0.038
VWC, cm^3 cm^{-3}	1^{th}	0	44	100	0.175	2.266
	2^{nd}	12	129	332	0.327	4.063
	3^{rd}	0	131	206	0.284	3.545
	4^{th}	56	125	212	0.284	3.556
	total	0	88	316	0.326	4.049
Sand, %	1^{th}	115	446	360	0.421	5.213
	2^{nd}	440	1119	300	0.365	4.510
	3^{rd}	401	1037	219	0.320	3.978
	4^{th}	413	717	260	0.300	3.747
Clay, %	1^{th}	19	92	389	0.392	4.861
	2^{nd}	123	215	428	0.239	3.016
	3^{rd}	68	138	177	0.321	4.034
	4^{th}	35	63	216	0.335	4.227
Silt, %	1^{th}	39	174	334	0.382	4.740
	2^{nd}	165	453	279	0.396	4.887
	3^{rd}	211	484	200	0.270	3.366
	4^{th}	6	10	341	0.266	3.332
ECa, mS m^{-1}	shallow	38	133	253	0.816	65.436
ECa, mS m^{-1}	deep	126	388	223	0.846	67.899

Table 1-3. Correlation coefficient between ECa (mS m^{-1}) data and soil basic information.

	Clay (%)				Sand (%)				Silt (%)			
layer	1	2	3	4	1	2	3	4	1	2	3	4
ECa, shallow	0.75	0.55	0.35	0.40	-0.75	-0.63	-0.45	-0.39	0.65	0.60	0.46	0.36
ECa, deep	0.59	0.61	0.52	0.57	-0.62	-0.73	-0.63	-0.63	0.56	0.72	0.63	0.60

	BD (g cm^{-3})				VWC (cm^3 cm^{-3})			
layer	1	2	3	4	1	2	3	4
ECa, shallow	-0.01	-0.15	-0.31	-0.02	0.66	0.61	0.47	0.47
ECa, deep	0.06	-0.20	-0.45	-0.13	0.63	0.71	0.64	0.65

Table 1-4. Correlation coefficient between cotton / soybean yield data and soil information.

	Cotton					Soybean				
Layer	1	2	3	4	total	1	2	3	4	total
BD, g cm^{-3}	-0.01	-0.25	-0.45	-0.12	-0.30	-0.12	-0.27	-0.33	-0.12	-0.29
VWC, cm^3 cm^{-3}	0.42	0.54	0.51	0.47	0.56	0.48	0.39	0.36	0.42	0.47
Sand, %	-0.43	-0.55	-0.50	-0.48		-0.52	-0.39	-0.35	-0.42	
Clay, %	0.35	0.39	0.39	0.37		0.46	0.25	0.23	0.31	
Silt, %	0.43	0.60	0.51	0.49		0.50	0.44	0.37	0.43	
ECS, mS m^{-1}					0.53					0.57
ECD, mS m^{-1}					0.66					0.50

* ECS: ECa shallow readings; ECD: ECa deep readings.

Chapter 2: High resolution prediction of soil available water content within crop root zone at a field scale

(Submitted to *Journal of Hydrology* for publication)

Amir Haghverdi[1*], Brian G. Leib[1], Paul D. Ayers[1], Michael J. Buschermohle[1]

[1]Dept. of Biosystems Engineering & Soil Science, UTK, 2506 E.J. Chapman Drive, Knoxville, TN 37996-4531

* Correspondence author, ahaghver@vols.utk.edu, Cellphone: (865)-235-9694

Abstract

A detailed spatial understanding of soil available water content within the effective root zone is needed to optimally schedule irrigation where field-scale spatial soil variation is significant. However, it is impractical to directly measure soil basic physical/hydraulic properties at field-scale because of the time consuming and expensive nature of lab/in situ techniques. In this study, multiple modeling scenarios were developed and evaluated to indirectly predict high resolution soil available water content maps within the effective root zone. The modeling techniques included kriging, co-kriging, regressing kriging, artificial neural networks (NN) and geographically weighted regression (GWR). The efficiency of soil apparent electrical conductivity (ECa) as proximal data in the modeling process was assessed. There was a good agreement (root mean square error (RMSE) = 0.052 cm^3 cm^{-3} and r = 0.88) between observed and point prediction of water contents using pseudo continuous pedotransfer function. GWR (mean RMSE = 0.062 cm^3 cm^{-3}) and regression kriging / co-kriging (mean RMSE = 0.063 cm^3 cm^{-3}) showed promising results and had higher accuracy than other methods for spatial prediction of water content maps. Up to 16 % improvement in the prediction accuracy was achieved by considering ECa as an ancillary attribute in the interpolation process. There was as high as a four-fold difference in available water content between coarse-textured and fine-textured soils on the study site. Further investigation is needed to evaluate the efficiency of variable rate irrigation scenarios to manage the spatial soil heterogeneity.

Keywords: apparent electrical conductivity, available water content, pseudo continuous pedotransfer function, water retention curve.

1. Introduction

The growing demand for the food and fiber production along with the uncertainty in rainfall patterns has focused a great attention on irrigation. If field-scale spatial soil variation is significant, variable rate irrigation becomes a desirable method to apply an optimum amount of water to each soil type in order to maximize yield. Duncan (2012) conducted a two-year cotton irrigation study and showed the optimum cotton supplemental irrigation strategy was different among plots with high, moderate and low water holding capacity (WHC). He emphasized that in the long term it is not possible to maximize yield by a uniform supplemental irrigation decision in a field with significant degrees of soil variability.

Soil hydraulic information is required for irrigation scheduling but is hard-to-obtain. Soil water retention and soil hydraulic conductivity are the prime factors governing water flow in the vadose zone. Irrigation management is highly dependent on soil available water content (AWC) within the root zone. Furthermore, soil hydraulic properties are essential inputs for most of the models linked with irrigation, drainage and hydrology. Obtaining soil hydraulic information is challenging due to the time-consuming and labor-intensive nature of the laboratory and in situ methods. When it comes to variable rate irrigation management, an additional difficulty is to provide a high resolution AWC map which precisely represents soil spatial distribution under each irrigation system.

Pedotransfer functions (PTFs) are a widely-used method to indirectly obtain soil hydraulic properties. The initial PTFs were derived using multiple regression technique while machine learning algorithms have become dominant in recent years (Vereecken et al., 2010). Traditionally soil basic data such as soil texture, soil bulk density and organic matter content are used as input predictors since they are often well correlated to soil hydraulic properties yet are easier to collect. However, obtaining high resolution soil basic data at field-scale is also time consuming and expensive thus impractical.

A combination of PTFs and interpolation techniques is usually required to generate a map of soil hydraulic properties. Ferrer Julià et al. (2004) generated a saturated hydraulic conductivity map of Spain, 1 km^2 resolution, using PTFs and subsequently kriging as an interpolation technique. They reported soil textural component as the most important input predictors while organic matter content showed a low influence on saturated soils. In recent years, alternative techniques have been introduced and evaluated to map spatial variability such as regression kriging, geographically weighted regression and machine learning-based spatial models (Eldeiry and Garcia, 2010; Li et al., 2011; Sharma et al., 2011). Herbst et al. (2006) examined different interpolation techniques to predict soil

hydraulic properties in a micro-scale catchment using terrain attributes. The result of their work indicated regression kriging as the most appropriate method with smallest average prediction error. They also reported up to 15% improvement in spatial predictions due to using terrain attributes as co-variables in comparison with ordinary kriging without co-variables.

Traditionally two modeling approaches were implemented to predict soil physical and hydraulic properties in a spatial manner. One can first run PTF at individual spots throughout the area of interest and then interpolate the PTF outputs to generate a map (i.e. 'calculate first, interpolate later', CI). An alternative way is to first interpolate the soil basic information (i.e. inputs for PTF such as bulk density, texture and organic matter content) across the study area and then convert soil basic maps to soil hydraulic properties map (s) by PTF (i.e. 'interpolate first, calculate later', IC). Many researchers (e.g. Sinowski et al. (1997), Heuvelink, and Pebesma (1999), Bechini et al. (2003)) compared different IC procedures against CI techniques, yet the reported results are different and do not indicate the supremacy of either of the methods.

The evidence in the literature points toward on-the-go sensing (e.g. Hedley and Yule, 2009a, 2009b; Hedley et al., 2013) and remote sensing (e.g. Jana and Mohanty, 2011) as suitable techniques to enhance the spatial prediction of soil hydraulic properties at field scale and beyond. Readily available dense data can be collected via on-the-go sensors which are able to gather the information while traversing a field. Apparent electrical conductivity (ECa) is a good example of such sensors which has been heavily used in precision agriculture. ECa may provide some useful information about physical and hydraulic properties of soil (Sudduth et al., 2005) when soil salinity is not a major factor. Duncan (2012) found a strong correlation between ECa data and depth to sand layer in an irrigated cotton field in a moderately humid region. Saey et al. (2009) tried to convert ECa data to clay percentage under non-saline conditions and found a good correlation for their model. In another study, Abdu et al. (2008) predicted soil texture and WHC in the subsurface soil of a small watershed using ECa data. Their model performed well with coefficient of determination equal to 0.86 and 0.75 for predicting clay percentage and WHC, respectively. Abdu et al. (2008) emphasized the need for additional studies to appropriately relate ECa data into hydrological attributes. The objectives of this study were (i) to predict AWC at field scale within the effective crop root zone using multiple modeling scenarios and (ii) to investigate the advantage of ECa data to improve spatial prediction of water retention.

2. Material and Methods

2.1 Study area & data collection

The field of study, approximately 73 ha, is located in west Tennessee close to the Mississippi River where two center pivot irrigation systems were available for supplemental irrigation (Figure 2-1). The field was planted to no-till cotton during recent cropping seasons. The field-work was scheduled after some rainfall events when soil was assumed to be close to its field capacity.

The shallow (approximately 0-30 cm) and deep (approximately 0-90 cm) ECa data (Sudduth et al., 2005) were collected (about 4,700 data points covering the entire field) by a Veris 3100 machine (Veris Technologies, Salina, KS) from the field of study on March 20, 2014. The Veris machine measures ECa using the principles of electrical resistivity. A small electrical current is introduced into the soil and the drop in voltage in two depths is measured. ECa is a function of the electrical conductivity of porous media solution, the porosity of soil and the cementation exponent. The current flowing through three different conductance pathways (i.e. liquid, soil-liquid and solid) affects ECa (Corwin and Lesch, 2005; Sudduth et al., 2005). The shallow ECa data exhibited a normal distribution but deep ECa data were skewed thus log transformed.

Then, soil sampling was done on March 21 and 22, 2014. Figure 2-1 shows the sampling scheme; one hundred undisturbed samples (0-100 cm) were collected using a truck mounted hydraulic probe. Each sample was divided into four segments. Hereafter, the word "layer" is used to distinguish among subsamples rather than the real soil horizons. The default length of subsamples was 25 cm though adjustment was made considering soil horizons. For instance, if there was a clear change between horizons at 20 cm we had the first layer from 0-20 cm but second layer started at 25 cm to keep the depth of samples as consistent as possible within each layer throughout the field. In the lab, soil texture, soil bulk density (BD) and gravimetric water content at the time of sampling were measured.

2.2 Pseudo continuous pedotransfer function PC-PTF

To predict water retention curves (WRCs) for the collected soil samples, NN network pseudo continuous PTF (PC-PTF) was developed (Figure 2-2). The concept of PC-PTF was first introduced by Haghverdi et al. (2012) as an alternative approach to point and parametric PTFs (respectively Type 2 and Type 3 PTFs in Wösten et al., 2001). Later, Haghverdi et al. (2014) investigated the impact of data mining procedure and data quality on the continuous performance of PC-PTF. Point PTFs only predict water retention at limited water retention points. Parametric PTFs

provide a prediction of parameters of a soil hydraulic equation (usually the one developed by van Genuchten, 1980) which then can be used to predict the whole WRC. PC-PTF relies on the power of machine learning algorithms to predict the WRC from a limited measured water retention points without using any soil hydraulic equation. It requires the logarithm of matric potential (e.g. for PWP the input equals to Log (15296 cm) = 4.18) as an input parameter, on top of soil basic information, where the output of the model is the corresponding volumetric water content. By using a wide range of soil matric potentials as input, a corresponding range of water contents will be predicted, and a (pseudo) continuous curve is generated (Haghverdi et al., 2014).

A subset (n=554) of UNSODA database (Nemes et al., 2001) was selected to establish PC-PTF. The UNSODA is a database of unsaturated soil hydraulic properties which contains a variety of information such as water retention and hydraulic conductivity as well as basic soil properties including particle-size distribution, bulk density and organic matter content. To derive a reliable PTF it is essential to have soils similar in their basic properties in training and prediction sets. Table 2-1 and Figure 2-3 illustrate the descriptive statistics and the textural distribution of the selected UNSODA samples along with samples collected from the field of study, respectively. The samples from UNSODA adequately covered most of the coarse textures of the textural triangle, meaning they were similar to the target soils from study area. The soil textural information, i.e. sand, silt and clay percentages, soil BD and logarithm of matric potential were considered as input predictors (Figure 2-2).

The NN was proven to be an appropriate model to derive PC-PTF (Haghverdi et al., 2012; 2014). A three layer feed forward back propagation NN was selected with sigmoid tangent hyperbolic and linear as the activation functions in hidden and output layers, respectively. Training was done by the Levenberg-Marquardt algorithm (Demuth and Beale, 2000). The UNSODA samples were divided into 10 almost equal subsets. Each subset was once assigned to the testing phase while the other 9 subsets were used to PTF *development*. The number of neurons in the hidden layer was changed from 1 to 20 when a cross-validation procedure was performed to identify the best structure. A sampling with replacement technique (Efron and Tibshirani, 1993) was applied on *development* data to create 50 statistically similar subsets of the same size for training. It was expected that each subset had about 63% of the parent data set (Schaap et al., 2001). The rest of the samples (i.e. 37%) were assigned to cross-validation. The training process was terminated by an increase in prediction error on cross-validation set. An average of the 50 bootstraps was considered as the output of the PTF. The Matlab R2014a environment (MathWorks, Inc., Natick,

Mass.) was used to build models. Once the PC-PTF was derived, a continuous prediction of water retentions over a wide range of matric potentials were obtained for the field of study soils.

2.3 Spatial modeling of soil available water content

The soil AWC is obtained as the difference between water content at the field capacity (FC) and permanent wilting point (PWP). The FC and PWP are the widely used thresholds for irrigation management. The FC is rather a qualitative parameter yet is a practical and understandable indicator of soil WHC (Romano and Santini, 2002). Water retentions at -10 and -33 kPa could be considered as FC for coarse and fine texture soils, respectively (Rivers and Shipp, 1972) while water content at -1500 kPa is usually chosen as the PWP for all the textures. For the sake of simplicity and convenience, in this study we considered identical matric potential (-10 ka) for the entire field whenever we made a spatial prediction of the water content at FC. However, three high resolution maps were predicted for water content at -10, -33 and -1500 kPa across the study site providing the opportunity to consider two matric potentials for water content at FC in further studies if it is needed.

Multiple spatial modeling scenarios, using kriging, co-kriging, regression kriging, GWR and NNs were examined (Figure 2-4). The objective was to convert point water retention data to continuous maps and to evaluate the efficiency of the ECa as a proximal attribute in this process. To evaluate the performance of the models, a cross-validation procedure was designed. First, soil samples were randomly divided into 5 groups (the groups were identical among soil layers). Soil water retention maps were then derived using 4 groups out of five and the left-over group was used to validate the performance of models. This task was repeated 5 times in order to involve all the samples in cross-validation process.

2.3.1 Kriging, co-kriging and regression kriging

Kriging is an advanced interpolation procedure which relies on the spatial autocorrelation information to model natural attributes. It assumes that spatial correlation is a function of the distance/direction between sample points. This technique provides predictions at unsampled locations, considering distance between points and overall spatial arrangement of data, and also evaluates the prediction uncertainty (Oliver, 2010). Cokriging is an extension of kriging technique which uses information on several attributes. It requires the semivariogram of the main attribute

plus its cross-semivariogram with proximal attribute(s) to improve predictions (Goovaerts and Kerry, 2010). The semivariogram and cross-semivariogram is estimated as follows:

$$\gamma_{uv}(h)=\frac{1}{2N(h)}\sum_{i=1}^{N(h)}\left\{\left[Z_u(\boldsymbol{x_i}+h)-Z_u(\boldsymbol{x_i})\right]\left[Z_v(\boldsymbol{x_i}+h)-Z_v(\boldsymbol{x_i})\right]\right\} \quad (1)$$

where γ_{uv} is the cross-semivariance when $u \neq v$ and is the semivariance when $u = v$; u is the main attribute; v is the proximal attribute; h is the interval class; $N(h)$ is the number of pairs separated by lag distance; and $Z(\boldsymbol{x_i})$ and $Z(\boldsymbol{x_i} + h)$ are measured attributes at spatial locations i and $i+h$, respectively. The general form of kriging (equation 2) and cokriging (equation 3) estimators are as follows:

$$z^*(\boldsymbol{x_p})=\sum_{i=1}^{I}\lambda_i z(\boldsymbol{x_i}) \quad (2)$$

where $z^*(\boldsymbol{x_p})$ is the kriging prediction at the target location $\boldsymbol{x_p}$; λ_i and $z(\boldsymbol{x_i})$ are weight of the attribute and the attribute value at ith location, respectively; and I is the number of measured values.

$$z^*(\boldsymbol{x_p})=\sum_{i=1}^{I}\lambda_i z(\boldsymbol{x_i})+\sum_{j=1}^{J}\lambda_j z(\boldsymbol{x_j}) \qquad \textbf{s.t.}\sum_{i=1}^{I}\lambda_i=1;\ \sum_{i=1}^{J}\lambda_j=0 \quad (3)$$

where i and j indicate the main and proximal attributes, respectively, and J is the number of measured values for proximal attribute; and other parameters as previously defined.

Regression kriging is a hybrid interpolation technique. It uses regression on proximal data and then kriging to interpolate residuals from the regression as follows (Hengl et al., 2007):

$$y(\boldsymbol{x_0})=\sum_{k=0}^{n}\hat{\beta}_k.q_k(\boldsymbol{x_0})+\sum_{i=1}^{m}\lambda_i.e(\boldsymbol{x_i}) \quad (4)$$

where y is the response predicted at target location $\boldsymbol{x_0}$, β_k is coefficient for the kth variable of regression model (drift model) and $q_k(\boldsymbol{x_0})$ is the value of proximal attribute at location $\boldsymbol{x_0}$, n is the number of proximal attributes, λ_i and $e(\boldsymbol{x_i})$ are weight of the residual and the residual value at location $\boldsymbol{x_i}$, respectively and m is the number of measured values.

Ordinary kriging, co-kriging and regression kriging were used in the models 1 to 6 (Figure 2-4). The ArcGIS 10.2.2 (ESRI Inc., Redlands, California) and Rstudio 0.98.1103 (RStudio, Inc., Boston, Ma, USA) were used to derive models. The Model 1 and Model 2 required PC-PTF for calculation and kriging to interpolate. Kriging was altered with co-kriging to establish the models 3 and 4 where shallow and deep ECa were considered as the proximal attributes for layers one and two/three/four, respectively. Regression kriging was used in models 5 and 6. The CI

sequence (i.e. calculating using PC-PTF followed by interpolating) was performed in the models 1, 3 and 5 as opposed to IC sequence (i.e. interpolating followed by calculating using PC-PTF) in the models 2, 4 and 6. If an attribute showed distribution very far from normal, the data were transformed to follow an approximate normal distribution. Then empirical semivariogram was calculated by Geostatistical Analyst toolbox in ArcGIS 10.2.2 (ESRI Inc., Redlands, California) considering 12 lags and omnidirectional stable model.

2.3.2 Geographically weighted regression

The ordinary least squares regression provides a global model to explore the overall data relationships in a desired area of study. However, it is probable to see different forms of dependency among variables and responses in different parts of the study area. In this case, the global model would not be reliable because it only represents an average of the non-stationary relationships. The geographically weighted regression (GWR) is a spatial regression technique which models local relationships among variables and responses. This technique defines a bandwidth for each data point, then establishes a regression equation using those data falling within the bandwidth (Fotheringham et al., 2002). The coefficients in GWR form a continuous surface representing the spatial variation:

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \quad (5)$$

Where y_i is the response predicted at point i; β_0 is the intercept and β_k is the coefficient for the kth variable; ε_i is the error associated with the point i; and (u_i,v_i) are the coordinates of the ith point in space.

In model 7, the GWR was used to predict soil texture components which subsequently were considered as input predictors in PC-PTF to predict water retention at FC and PWP. In model 8, the GWR was required to directly predict water retention at FC and PWP. Both shallow and deep (log transformed) ECa readings were considered as input predictors for all soil layers. To perform each local regression analysis, the optimum distance (fixed kernel) was automatically obtained using the Akaike information Criterion (AICc). The ArcGIS 10.2.2 (ESRI Inc., Redlands, California), was used to derive models.

2.3.3 Artificial neural network

The models 9 and 10 were similar to models 7 and 8, respectively, except that GWR was replaced with NN. Both shallow and deep (log transformed) ECa readings were considered as input predictors for all soil layers. The coordinates of samples were considered as extra input predictors, enabling the NN to learn the dependency of the

sampling locations hence working as a spatial model. The SPSS Modeler 15 (SPSS Inc., Chicago, IL, USA) was used to derive models. The best number of neurons in the hidden layer was automatically computed. The Bootstrap aggregating technique was implemented to enhance model stability and the number of ensembles was equal to 50. Thirty percent of the training data were assigned to cross-validation to avoid over-fitting.

2.4 Yield data collection and cleaning

The correlation between soil AWC throughout the effective root zone, ECa data and crop yield was obtained. Two available yield data sets (collected and stored by the farmer) were selected: soybean from 2009 and cotton from 2012. The majority of the field is covered by the two center pivots. The rainfall is usually high during cropping season yet farmer practices uniform supplemental irrigation whenever unpredicted dry periods occur. The raw yield data were pre-processed to remove outliers and bad data. The data greater/smaller than 3 times the standard deviation from the mean were assumed statistically unexpected and removed unless there was scientific evidence acting against this assumption. To calculate the correlation, the cleaned yield data first were converted to maps by the ordinary kriging method. Then, the expected yield values at the sampling locations were obtained. Same procedure was implemented to calculated ECa data at sampling locations.

2.5 Performance evaluation

The performance of the models was evaluated by three statistics: the root mean square error (RMSE), the mean bias error (MBE) and the correlation coefficient (r):

$$\mathrm{RMSE}=\left(\frac{1}{n}\sum_{i=1}^{n}\left(E_i-M_i\right)^2\right)^{0.5} \tag{6}$$

$$\mathrm{MBE}=\frac{1}{n}\sum_{i=1}^{n}\left(E_i-M_i\right) \tag{7}$$

$$r=\frac{\sum_{i=1}^{n}(M_i-M_m)(E_i-E_m)}{\left(\sum_{i=1}^{n}(M_i-M_m)^2\sum_{i=1}^{n}(E_i-E_m)^2\right)^{0.5}} \tag{8}$$

where E_i, and M_i are the predicted (i.e. the output of the spatial modeling process) and the actual water content (i.e. the outputs of the PC-PTF prior to performing the spatial modeling process) for the ith observation ($cm^3\ cm^{-3}$),

respectively; n is the number of actual water content points in test/validation set; E_m and M_m are the mean of predicted and the mean of actual water content ($cm^3\ cm^{-3}$), respectively.

3. Results

3.1 Pseudo continuous pedotransfer function

Figure 2-5 shows how PC-PTF performed in predicting water retention of UNSODA data set. At first the error sharply decreased as more neurons were added to the hidden layer, but leveled off after about 7-10 neurons. The model with 17 processing elements had the minimum RMSE (0.056 $cm^3\ cm^{-3}$), so was selected for the prediction. Figure 2-6 illustrates the WRCs predicted for the collected samples (n= 400) from the field of study. The shape of the curves followed the expected pattern: a steep slope at low and intermediate absolute matric potentials for more coarse-textured soils, as opposed to a constant decrease over a wider range of matric potentials for more fine-textured soils.

The field seemed to be close to its field capacity during sampling. We took advantage of this status to evaluate the performance of PC-PTF against measured soil water content data from the field of study. The water retention at matric potentials 10 kPa and 33 kPa was predicted and plotted against volumetric water content at the time of sampling (Figure 2-7). The points were color coded to distinguish among soil layers where darker circles correspond to samples from deeper layers. There was a very good agreement among predicted and observed water contents with a correlation coefficient (RMSE) equal to 0.87 (0.051 $cm^3\ cm^{-3}$) and 0.88 (0.052 $cm^3\ cm^{-3}$) at 10 kPa and 33 kPa, respectively. There were soils with low and high WHC at each layer. A cloud of dark circles at low water content, i.e. from 10 to 20 %, corresponds well to the presence of some sandy spots across the field of study. There was no evidence of over/under estimation since points were well scattered around 1:1 line. There was only one point with volumetric water content more than 60 % which was located very close to drainage path. The surface of the drainage path was still wet at the time of sampling, so at a higher water content than the FC.

3.2 Spatial prediction of soil available water content

Up to now, the WRCs were predicted using PC-PTF for the collected soil samples. Figure 2-8 depicts the spatial structure of the water content at FC and PWP along with AWC for different layers. The resulting semivariograms indicated mostly moderate to strong spatial structures. This strong soil spatial variability has likely been created by

depositional events, river flood-induced sand boils and/or earthquake-induced sand blows. The average *range* varied from 250 m to 279 m among layers, 2 to 3 times greater than sampling intervals. This shows the presence of spatial structure beyond sampling distance. With regard to FC, in general, the value for range decreased with depth. This trend was reversed, however, with regard to PWP while subsurface layer (25-50 cm) had the highest range. The *nugget* to *sill* ratio varied from 23 % to 60 % for water retention data at FC and PWP, except for the water content at PWP at the first layer which was 0. This could be an artifact of the variogram fitting and calculation since PWP at the first layer represented a weak spatial structure. On average, the subsurface layer (i.e. 25-50 cm) had the highest nugget to sill ratio. The AWC data exhibited some randomness and inconsistency in their spatial structure for the first and second layers. Therefore, AWC maps were produced for individual layers by subtracting interpolated FC maps (i.e. 10 kPa) from PWP maps.

Table 2-2 and 2-3 summarize performance of the models for each individual soil layer in terms of RMSE and MBE, respectively. The model 1, 3 and 5 almost always had a better performance and lower bias than model 2, 4 and 6, respectively. Therefore, applying PC-PTF to predict water retention on discrete sampling locations and subsequently interpolating those predictions to generate a map is suggested (i.e. CI). Co-kriging using ECa as ancillary attribute enhanced the accuracy of interpolation in comparison with ordinary kriging for both CI (i.e. 16 % lower mean RMSE for model 3 comparing to model 1) and IC (i.e. 16 % lower mean RMSE for model 4 comparing to model 2) approaches. The regression kriging showed higher accuracy (16 % lower mean RMSE) than ordinary-kriging for CI approach (model 5 versus model 1) for all layers and matric potentials except for predicting water content at 1500 kPa to deep layer, but this improvement was not consistent for IC approach across layers and matric potentials. In 11 cases out of 12, GWR based models worked more accurately than NN based models but the differences were minor such that average RMSE for models 7, 8, 9 and 10 were 0.063 $cm^3\ cm^{-3}$, 0.062 $cm^3\ cm^{-3}$, 0.064 $cm^3\ cm^{-3}$ and 0.066 $cm^3\ cm^{-3}$, respectively. There is no best model for all the matric potentials and layers. Overall, models 8 (RMSE = 0.062 $cm^3\ cm^{-3}$) following by models 7, 5 and 3 (RMSE = 0.063 $cm^3\ cm^{-3}$) showed the highest potential for mapping AWC. On average the MBE values were close to zero indicating systematic over/under estimation was not an issue.

Figure 2-9 shows the maps of AWC predicted for different layers and for the entire effective root zone (i.e. 1 m). The AWC maps of individual layers were produced using the model 8, considering water content at 10 kPa as FC for the entire field. To generate the AWC map for the entire effective root zone, AWC for each sampling

location was calculated then interpolated. Finally, the Natural Breaks (Jenks) method was employed to classify AWC map. In general, there were three regions within the field with low AWC: a region in the southern portion of the field somewhat outside of the both pivots circles, (ii) a triangular shaped region in the eastern of the field covering the pivot point of the east irrigation system and expanding to the eastern boarder of the field, and (iii) a region in the northwestern part of the field almost covering top one-third of the west irrigation system. The mean AWC was almost the same at first and second layers but dropped up to 20% at third and fourth layers. The standard deviation of the AWC was 0.54, 2.32, 3.62 and 3.28 % for the layers 1, 2, 3 and 4, respectively. The range of variation of AWC for the entire effective root zone was 5.34 to 20.95 % (mean = 11.45 % and SD = 3.28 %) showing a difference as great as almost fourfold between soil with low and high AWC across the field.

3.3 Soil water retention, yield and ECa variation

Table 2-4 summarizes the correlation of ECa and yield data with water retention at FC and PWP among different layers. Note, that the correlation information was only obtained for cores, hence water retention information was predicted by PC-PTF for cores (n= 100 (points) × 4 (depths) = 400) without interpolation. The ECa showed a moderate correlation, from 0.39 to 0.75, with water retention and AWC. As it was expected, the ECa-shallow had a greater correlation with water retention at first layer while ECa-deep indicated a higher correlation with water retention at second, third and fourth layers. On average, there was no difference between the correlations of ECa with water content at different matric potentials but yield showed higher correlation with water content at FC than with water content at PWP. Both cotton and Soybean had low to moderate correlation with water retention and AWC ranging from 0.35 to 0.65 and from 0.34 to 0.56, respectively. The highest correlation (r = 0.75) was observed between AWC in the entire effective root zone (considering water content at 10 kPa as the FC) and ECa-deep readings. The same was true for yield data such that the highest correlation (0.62 and 0.56 for cotton and soybean, respectively) occurred for AWC within the entire effective root zone. This is because ECa deep readings integrates the soil at more than one layer and yield is also affected more by the available water within the entire effective root zone in comparison with the water status at individual layers throughout growing season.

4. Discussion

4.1 Model performance analysis

The performance of PC-PTF was satisfactory; it had RMSE as low as 0.056 cm^3 cm^{-3} over a relatively large number of samples from UNSODA data set and even more importantly exhibited a high correlation (r= 0.89 and RMSE = 0.052 cm^3 cm^{-3}) with observed data from the field of study. Nebel et al. (2010) evaluated the performance of some published PTFs in respect to their ability to explain the spatial variability of water retention at FC and PWP. The reported correlation coefficient by them varied from 0.35 to 0.69 across PTFs. Our findings corroborated the result acquired by previous studies by Haghverdi et al. (2012, 2014) on the accuracy and reliability of the PC-PTFs. Haghverdi et al. (2014) reported 0.040 and 0.047 cm^3 cm^{-3} as RMSE of their PC-PTF that was applied to two data sets from Belgium and Turkey, respectively. The PC-PTF even showed a higher performance, RMSE = 0.028 cm^3 cm^{-3}, on a small data set from Iran (Haghverdi et al., 2012). The variation in PC-PTF performance among studies is related to the characteristics of the data sets. In previous studies, training and test soil samples belonged to small local data sets but we used UNSODA data set which included samples from different regions around the world. Deriving PTFs via large data sets comprising soils from different origins creates higher error than establishing a PTF by a local/regional dataset. For instance, Rosetta PTF and a support vector machine based PTF, when derived and tested on a big dataset, had RMSE equal to 0.053 cm^3 cm^{-3} and 0.067 cm^3 cm^{-3}, respectively (Twarakavi et al., 2009). Haghverdi et al. (2015) used the same subset of samples from UNSODA as we did to establish and test the accuracy of a novel non-parametric PTF which predicts the WRC using *k* nearest neighbor technique and van Genuchten soil hydraulic equation (Van Genuchten, 1980). Their PTF performed very similar (i.e. RMSE = 0.057 cm^3 cm^{-3}) to PC-PTF in our study. In another study, Schaap et al. (2001) reported RMSE equal to 0.068 cm^3 cm^{-3} for Rosetta PTF established and tested over a large number of soil hydraulic data.

The spatial arrangement of zones with low and high AWC matched our observation during sampling on soil texture and volumetric water content spatial heterogeneity. It suggests a combination of PC-PTF and interpolation as a useful method to produce high resolution maps of soil AWC. The error of the regionalization process comes from both calculation and interpolation steps, yet according to Sinowski et al., (1997) different error components, i.e. PTF and interpolation, may compensate each other. Herbst et al. (2006) used different interpolation techniques to predict some hydraulic properties of soil in a micro-scale catchment and found the error of different methods to be roughly in the same range. No method in our study consistently performed best for all layers and matric potentials. Zhang

and Srinivasan (2009) reported an increase in the precipitation spatial prediction accuracy when elevation and spatial coordinate were considered as ancillary data through interpolation process while no interpolation method consistently outperformed other approaches. On the other hand, Eldeiry and Garcia, (2010) used remotely sensed space images using Landsat satellite to estimate soil salinity using multiple techniques. They found ordinary kriging to perform better than regression kriging and co-kriging. They stated better auto correlation among soil salinity data comparing to cross-correlation between salinity data and remotely sensed data as a potential reason for the better results of the ordinary kriging. There was a considerable spatial agreement between ECa data and soil water content in our study. We collected the ECa data only one day before sampling when field was believed to be close to its field capacity. That is why ECa data provided useful information on spatial heterogeneity of soil water retention throughout the field of study. This confirms the efficiency of ECa as a proximal data in our study and in turn explains better performance of methods that made use of ECa in predication process such as GWR, co-kriging and regression kriging over ordinary kriging procedure. We observed up to 16 % improvement due to incorporating ECa data in our models which was in agreement with the reported result by Herbst et al. (2006) who observed up to 15% improvement by considering terrain attributes as ancillary data in the spatial prediction process.

In general, CI (i.e. calculation – interpolation) was a more accurate procedure than IC (i.e. interpolation – calculation) in this study. This is because for the former, water retention was the only variable interpolated, while for the latter each input variable had to be interpolated which increased the error in model application (Bechini et al., 2003), especially when one or some of the variables do not show a spatial structure. Moreover, some inconsistency of the spatial prediction of single fractions of the particle size distribution may occur (Herbst et al., 2006). In the study site, there was a weak spatial dependency among BD of the samples. Since BD was an input of the PC-PTF, associated error with interpolated BD likely contributed in higher error of IC procedure. However, CI may not be as efficient if a continuous realization of WRC over a wide range of matric potentials is needed. Saito et al. (2009) compared some fitting-interpolation scenarios in order to get a spatial interpolation of WRC. They reported lower prediction error for interpolation-fitting method. According to Heuvelink and. Pebesma (1999) in CI, the interpolation step does not fully utilize available information on spatial structure of inputs. Therefore, if many inputs are available with different spatial structures, IC might be a more accurate option.

4.2 Practical challenges and findings

The sampling density and scheme would affect the error associated with the spatial prediction process (Herbst et al., 2006). Sampling is time-consuming and expensive. Therefore, in practice, it is desired to minimize sampling density. Given the promising results achieved by incorporating ECa data in spatial modeling process along with the high degree of spatial agreement observed between ECa and soil physical data, it may be feasible to reduce the need for more difficult data collection (such as soil texture and BD) through sampling while maintaining the target level of accuracy. Debaene et al. (2014) studied the relationship between density of samples and prediction of several soil properties at the farm-scale using visible and near infrared spectroscopy method. They showed 1.5 samples per ha is adequate to predict soil organic matter and texture. Our sampling density (=1.42 samples per ha) was close to their recommendation. Taking a sample per ha is usually considered as the greatest sampling density that farmer can afford (Kerry et al., 2010). However, the extra difficulty for variable rate irrigation management is to collect deep samples to adequately cover effective crop root zone. Analyzing the effect of the density of collected soil samples on the performance of the spatial modeling techniques was out of the scope of the current study. Iqbal et al. (2005) recommended sampling interval at less than 100 m in order to detect boundaries of soil hydraulic properties in alluvial floodplain soil of Mississippi Delta. Further investigation is needed to determine the optimum number of samples to compromise between practical limitations and knowledge of the soil variability.

The soil texture distribution was the prime factor influencing AWC among different layers. We found a reduction in AWC with depth which was in line with an increase in sand percentage with depth. Moreover, a strong positional similarity was observed among AWC maps with our understanding of soil textural heterogeneity for the field of study which was somewhat expected since texture was the main driver of the PC-PTF. Iqbal et al. (2005) also found a great match among spatial distribution of water retention properties with that of clay and sand. ECa showed a good correlation with soil physical properties and soil hydraulic properties for the field of study. It also turned out to be a useful proximal attribute to map AWC; the accuracy of soil water retention maps enhanced when ECa was incorporated in the models as a proximal data. Abdu et al. (2008) showed that ECa can be useful to map soil textural patterns of the watershed. Using proximal data such as ECa also can be helpful to minimize the number of samples and in turn the associated cost with the mapping process.

Spatial variation in soil AWC was strong, i.e. up to fourfold among fine-textured and coarse-textured soils, across the field of study and also had a moderate influence on yield. Using a multi-year study with surface drip

irrigation system, Duncan (2012) showed that optimized cotton supplemental irrigation management is not identical across soils with low and high AWCs. Currently, most farmers practice uniform irrigation application rate regardless of the soil spatial variation across their fields. This management at least causes partial over / under irrigation, depending which soil they follow, and consequently yield reduction. Over irrigation in soils with high AWC causes cotton to increase its vegetative growth but was observed to decrease the yield. On the other hand, severe deficit irrigation in soils with lower AWC decreases the yield by influencing the boll number and retention (Gwathmey et al., 2011). The high spatial resolution maps of AWC could be used to delineate irrigation management zones for variable rate irrigation. Given the observed in depth variation of AWC, the spatial arrangement of management zones may even vary throughout growing season as roots penetrate into deeper layers. Therefore, a temporally dynamic zoning system may be needed to fulfill crop water requirement. The available center pivots at the field, like most of the typical center pivot irrigation systems, are capable of varying irrigation across field to some extent by changing their travel speed. Developing a variable rate irrigation strategy for this field is expected to increase the yield as well as water use efficiency. This, in turn, would reduce other potential problems including leaching, runoff and erosion. Further studies are needed to precisely investigate different irrigation managements and to economically compare the profitability of variable rate systems as opposed to an optimized management using available pivots with limited speed control ability.

5. Conclusion

Information on field-scale spatial heterogeneity of soil available water content within effective root zone could be used for irrigation zone delineation and variable rate irrigation scheduling. There are many methods in literature providing non-spatial point estimation of soil water retention, yet the options are more limited when the goal is to produce a high resolution soil hydraulic property map at the field-scale. The PC-PTF performed well for predicting WRC of soils with different textures. We found incorporating ECa data to be beneficial for interpolation of point predictions of WRC using methods such as regression kriging and co-kriging. ECa also showed promising performance as an input predictor in GWR models indicating the highest mean accuracy among models tested in our study. However, a firm understanding of soil characteristics affecting soil ECa is needed for each site prior to applying ECa in a modeling process. We leave this for future investigation to figure out how ECa could help to minimize number of samples yet maintaining the desired level of accuracy. We showed that drastic soil spatial

variable could affect yield distribution. Site-specific spatial models are required to investigate the optimum application of agricultural inputs such as irrigation and fertilizer in presence of field-scale soil spatial heterogeneity. The predicted information about soil spatial heterogeneity in this study could then be used as an input for deriving such models.

References

Abdu, H., Robinson, D. A., Seyfried, M., Jones, S. B., 2008. Geophysical imaging of watershed subsurface patterns and prediction of soil texture and water holding capacity. Water Resour. Res. 44(4).

Bechini, L., Bocchi, S., Maggiore, T., 2003. Spatial interpolation of soil physical properties for irrigation planning. A simulation study in northern Italy. Eur. J. Agron. 19(1), 1-14.

Corwin, D. L., Lesch, S. M., 2005. Apparent soil electrical conductivity measurements in agriculture. Comput. Electron. Agric. 46(1), 11-43.

Debaene, G., Niedźwiecki, J., Pecio, A., Żurek, A., 2014. Effect of the number of calibration samples on the prediction of several soil properties at the farm-scale. Geoderma, 214, 114-125.

Demuth, H., Beale, M., 2000. Neural Network Toolbox. Mathworks, Inc.

Duncan, H. A., 2012. Locating the variability of soil water holding capacity and understanding its effects on deficit irrigation and cotton lint yield. Master's Thesis, University of Tennessee. http://trace.tennessee.edu/utk_gradthes/1286.

Efron, B., R.J. Tibshirani., 1993. An Introduction to the Bootstrap. Monographs on Statistics and Applied Probability. Chapman and Hall, New York.

Eldeiry, A. A., Garcia, L. A., 2010. Comparison of ordinary kriging, regression kriging, and cokriging techniques to estimate soil salinity using LANDSAT images. J. Irrig. Drain. Eng. 136(6), 355-364.

Fotheringham, A. S., Brunsdon, C., Charlton, M., 2002. Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons. 269 p.

Ferrer Julià, M., Estrela Monreal, T., Sánchez del Corral Jiménez, A., Garcı́a Meléndez, E., 2004. Constructing a saturated hydraulic conductivity map of Spain using pedotransfer functions and spatial prediction. Geoderma 123(3), 257-277.

Goovaerts, P., Kerry, R., 2010. Using ancillary data to improve prediction of soil and crop attributes in precision agriculture. In: Geostatistical Applications for Precision Agriculture (pp. 167-194). Springer Netherlands.

Gwathmey, C. O., Leib, B. G., Main, C. L., 2011. Lint yield and crop maturity responses to irrigation in a short-season environment. J. of Cotton Sci. 15, 1-10.

Haghverdi, A., Cornelis, W. M., Ghahraman, B., 2012. A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. J. Hydrol. 442, 46-54.

Haghverdi, A., Öztürk, H. S., Cornelis, W. M., 2014. Revisiting the pseudo continuous pedotransfer function concept: Impact of data quality and data mining method. Geoderma. 226, 31-38.

Haghverdi, A., Leib, B. G., Cornelis W. M., 2015. A simple nearest-neighbor technique to predict the soil water retention curve. Transaction of the ASABE. 58(3).

Hedley, C. B., Yule, I. J., 2009a. A method for spatial prediction of daily soil water status for precise irrigation scheduling. Agric. Water Manag. 96(12), 1737-1745.

Hedley, C. B., Yule, I. J., 2009b. Soil water status mapping and two variable-rate irrigation scenarios. Precis. Agric. 10(4), 342-355.

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.

Hengl, T., Heuvelink, G. B., Rossiter, D. G., 2007. About regression-kriging: from equations to case studies. Comput. Geosci. 33(10), 1301-1315.

Herbst, M., Diekkrüger, B., Vereecken, H., 2006. Geostatistical co-regionalization of soil hydraulic properties in a micro-scale catchment using terrain attributes. Geoderma. 132(1), 206-221.

Heuvelink, G. B., Pebesma, E. J., 1999. Spatial aggregation and soil process modelling. Geoderma. 89(1), 47-65.

Iqbal, J., Thomasson, J. A., Jenkins, J. N., Owens, P. R., Whisler, F. D., 2005. Spatial variability analysis of soil physical properties of alluvial soils. Soil Sci. Soc. Am. J. 69(4), 1338-1350.

Jana, R. B., Mohanty, B. P., 2011. Enhancing PTFs with remotely sensed data for multi-scale soil water retention estimation. J. hydrol. 399(3), 201-211.

Kerry, R., Oliver, M. A., Frogbrook, Z. L., 2010. Sampling in precision agriculture. In Geostatistical applications for precision agriculture (pp. 35-63). Springer Netherlands.

Li, J., Heap, A. D., Potter, A., Daniell, J. J., 2011. Application of machine learning methods to spatial interpolation of environmental variables. Environ. Model. Softw. 26(12), 1647-1659.

Nebel, Á. L. C., Timm, L. C., Cornelis, W., Gabriels, D., Reichardt, K., Aquino, L. S., Pauletto, E. A., Reinert, D. J., 2010. Pedotransfer functions related to spatial variability of water retention attributes for lowland soils. Revista Brasileira de Ciência do Solo, 34(3), 669-680.

Nemes, A., Schaap, M.G., Leij, F.J., Wösten, J.H.M., 2001. Description of the unsaturated soil hydraulic database UNSODA version 2.0. J. Hydrol. 251(3), 151-162.

Oliver, M. A., 2010. An overview of geostatistics and precision agriculture. In: Geostatistical Applications for Precision Agriculture (pp. 1-34). Springer Netherlands.

Rivers, E.D., and R.F. Shipp. 1972. Available water capacity of sandy and gravelly North Dakota soil. Soil Sci. 113:74–80.

Romano, N., Santini, A., 2002. Field, In: Methods of Soil Analysis. Part 4, Physical Methods, Soil Sci. Soc. Am. Book Ser., vol. 5, edited by J. H. Dane and G. C. Topp, pp. 721– 738, Soil Sci. Soc. of Am., Madison, Wis.

Saey, T., Van Meirvenne, M., Vermeersch, H., Ameloot, N., Cockx, L., 2009. A pedotransfer function to evaluate the soil profile textural heterogeneity using proximally sensed apparent electrical conductivity. Geoderma. 150(3), 389-395.

Saito, H., Seki, K., Šimůnek, J., 2009. An alternative deterministic method for the spatial interpolation of water retention parameters. Hydrol. Earth Syst. Sci. 13(4), 453-465.

Schaap, M.G., Leij, F.J., van Genuchten, M. Th., 2001. ROSETTA: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. J. Hydrol. 251, 163–176.

Sharma, V., Kilic, A., Kabenge, I., Irmak, S., 2011. Application of GIS and Geographically Weighted Regression to Evaluate the Spatial Non-Stationarity Relationships between Precipitation vs. Irrigated and Rainfed Maize and Soybean Yields. Trans. ASABE. 54(3), 953-972.

Sinowski, W., Scheinost, A. C., Auerswald, K., 1997. Regionalization of soil water retention curves in a highly variable soilscape, II. Comparison of regionalization procedures using a pedotransfer function. Geoderma. 78(3), 145-159.

Sudduth, K.A., Kitchen, N.R., Wiebold, W.J., Batchelor, W.D., Bollero, G.A., Bullock, D.G., Clay, D.E., Palm, H.L., Pierce, F.J., Schuler, R.T., Thelen, K.D., 2005. Relating apparent electrical conductivity to soil properties across the north-central USA. Comput. Electron. Agric. 46(1), 263-283.

Twarakavi, N. K., Šimůnek, J., Schaap, M. G., 2009. Development of pedotransfer functions for estimation of soil hydraulic parameters using support vector machines. Soil Sci. Soc. Am. J. 73(5), 1443-1452.

Van Genuchten, M.Th., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated. Soil Sci. Soc. Am. J. 43, 892–898.

Vereecken, H., Weynants, M., Javaux, M., Pachepsky, Y., Schaap, M. G., van Genuchten, M. Th., 2010. Using pedotransfer functions to estimate the van Genuchten–Mualem soil hydraulic properties: A review. Vadose Zone J. 9(4), 795-820.

Wösten, J.H.M., Pachepsky, Y.A., Rawls, W.J., 2001. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. J. Hydrol. 251, 123–150.

Zhang, X., & Srinivasan, R. 2009. GIS-Based spatial precipitation estimation: a comparison of geostatistical approaches. Am. Water Resour. Assoc. 45(4), 894-906.

Appendix 2: Chapter 2 Figures and Tables

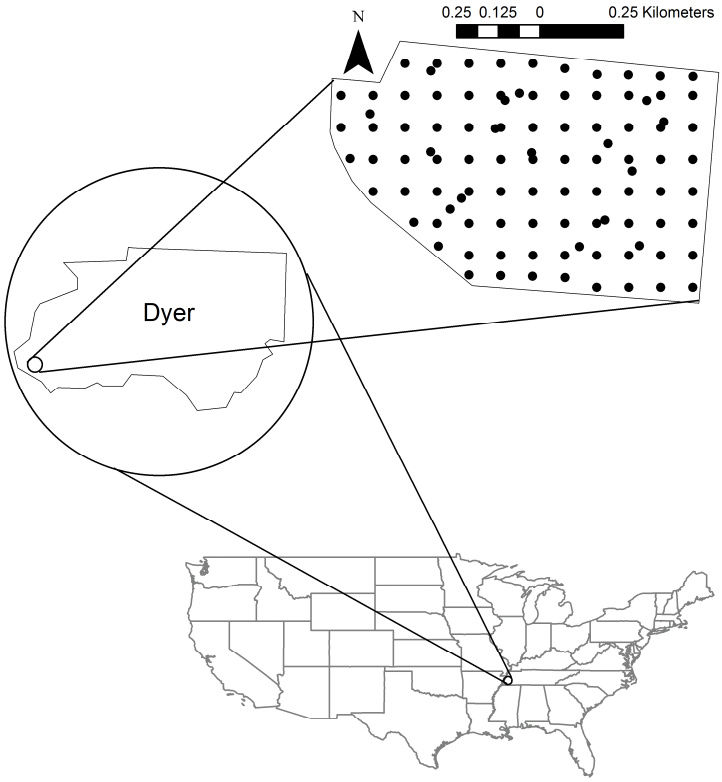

Figure 2-1. Soil sampling scheme within the field of study located in Dyer County, Tennessee.

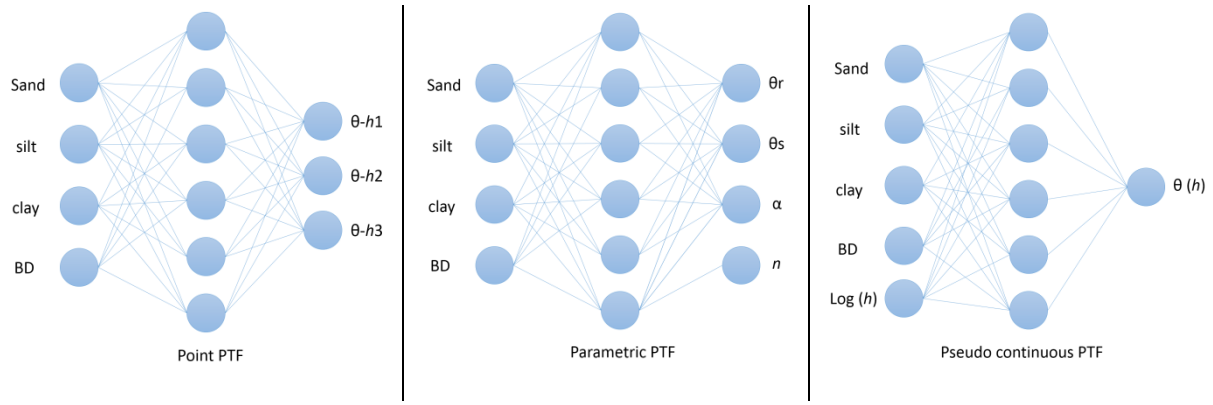

Figure 2-2. Illustration of topology of point, pseudo continuous and parametric neural network pedotransfer functions (adapted from Haghverdi et al., 2012). The point PTF predicts the water content at three different matric potentials (i.e. $h1$, $h2$ and $h3$). The parametric PTF predicts the inputs of van Genuchten equation (1980). The pseudo continuous PTF considers the soil matric potential as an extra input (i.e. Log (h)) and predicts the corresponding water content (i.e. $\theta(h)$).

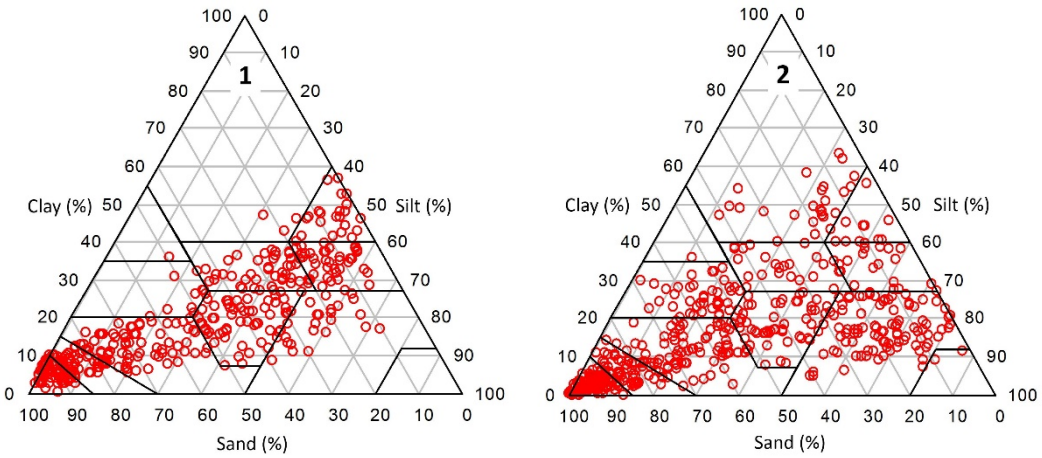

Figure 2-3. Soil texture distribution; samples from the field of study (panel 1), selected samples from UNSODA (panel 2).

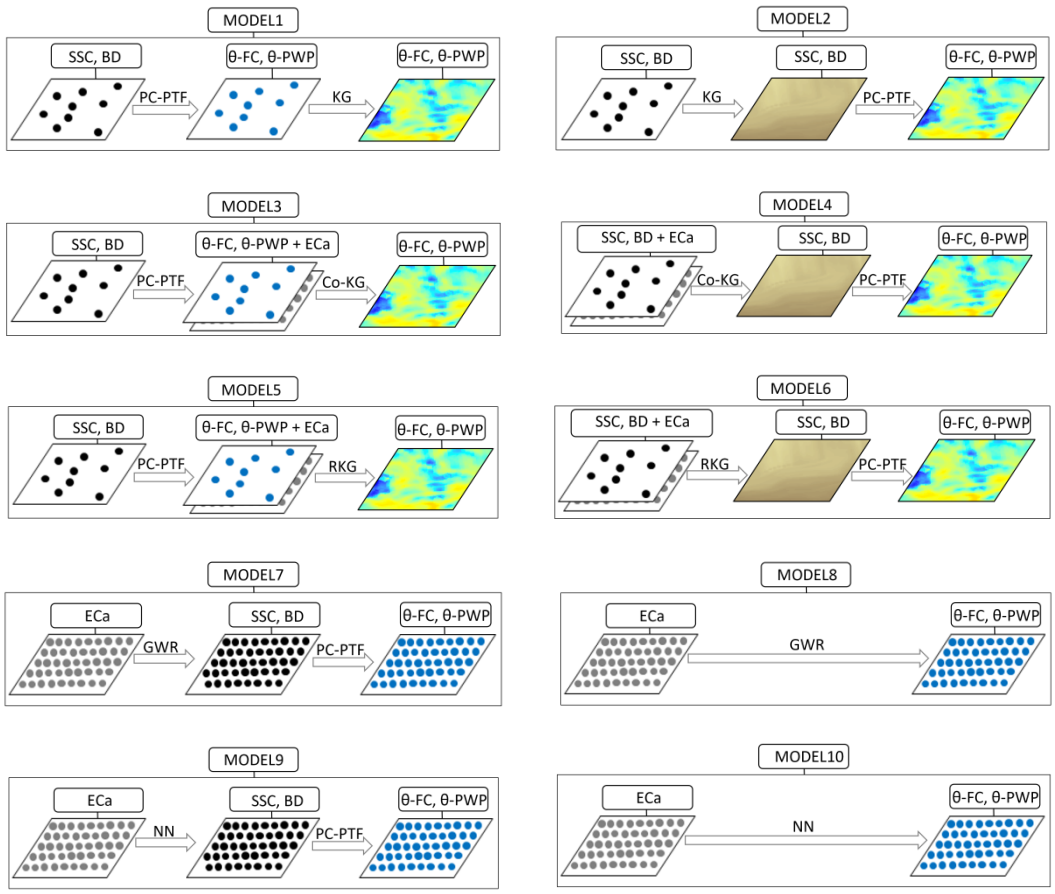

Figure 2-4. Spatial models to predict the soil water retention. SSC: sand, silt and clay (%), BD: bulk density (Mg m^{-3}), θ-FC: water content at field capacity (cm^3 cm^{-3}), θ-PWP: water content at permanent wilting point (cm^3 cm^{-3}), ECa: apparent electrical conductivity (mS m^{-1}), PC-PTF: pseudo continuous pedotransfer function, KG: kriging, Co-KG: co-kriging, RKG: regression kriging, GWR: geographically weighted regression, NN: neural network.

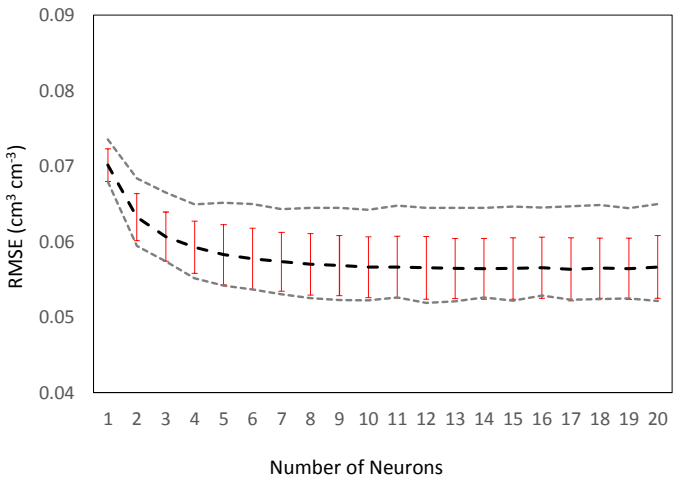

Figure 2-5. Performance of pseudo continuous PTFs in predicting the WRC of UNSODA samples. Thick dash line: mean RMSE; thin dash lines: min / max RMSE; error bars: standard deviation of RMSE values over the 10 subsets.

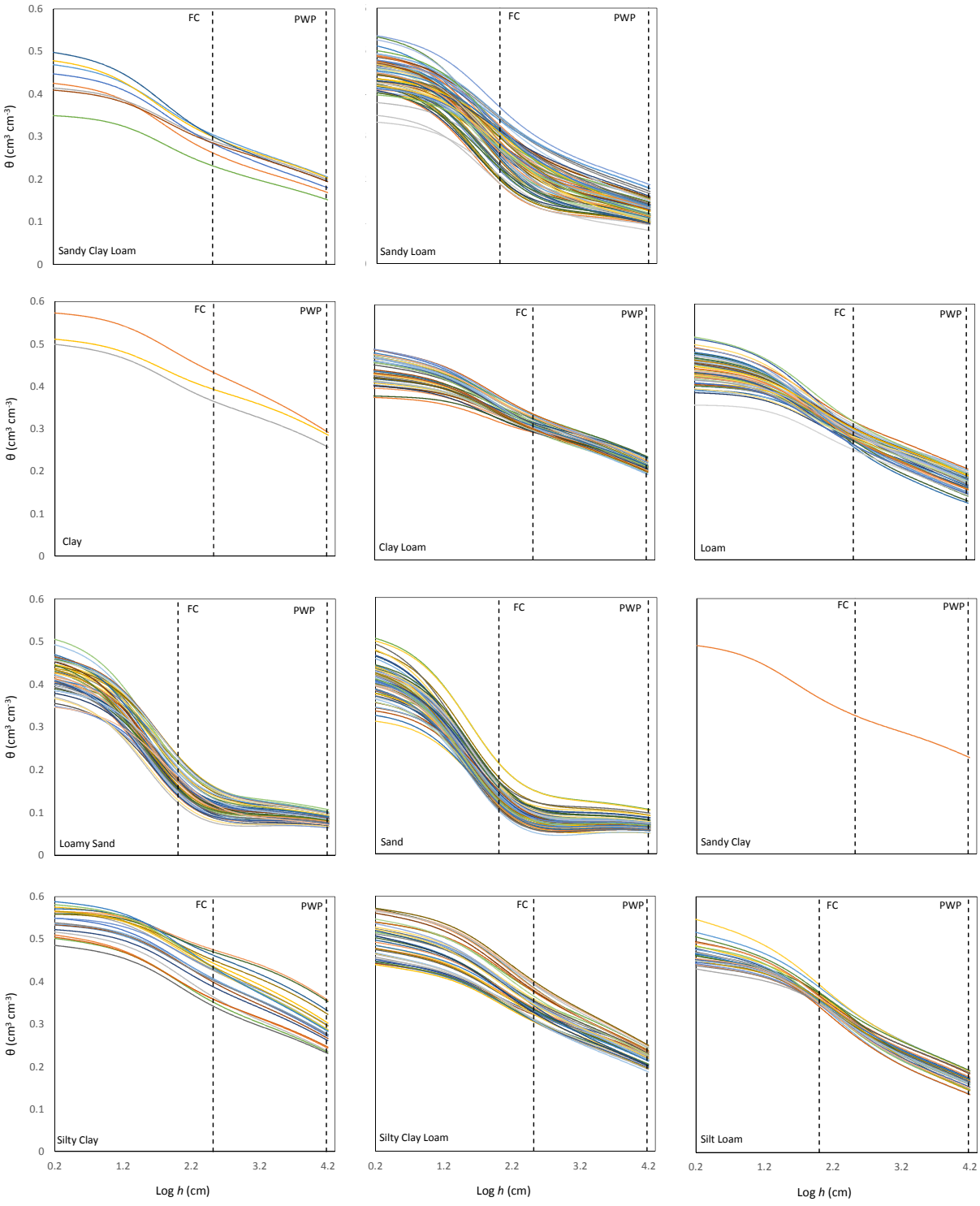

Figure 2-6. The WRCs at different soil textures predicted by PC-PTF for the soil samples (n= 400) collected from field of study.

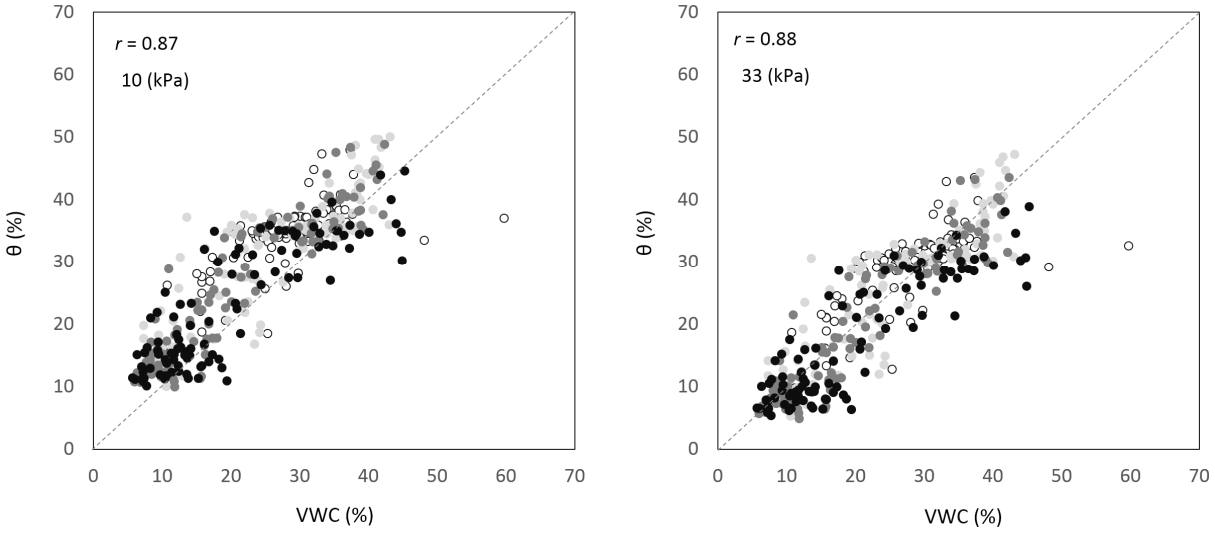

Figure 2-7. Scattering of water content at 10 kPa and 33 kPa predicted by PC-PTF against water content at the time of sampling. The darker dots represent deeper layers.

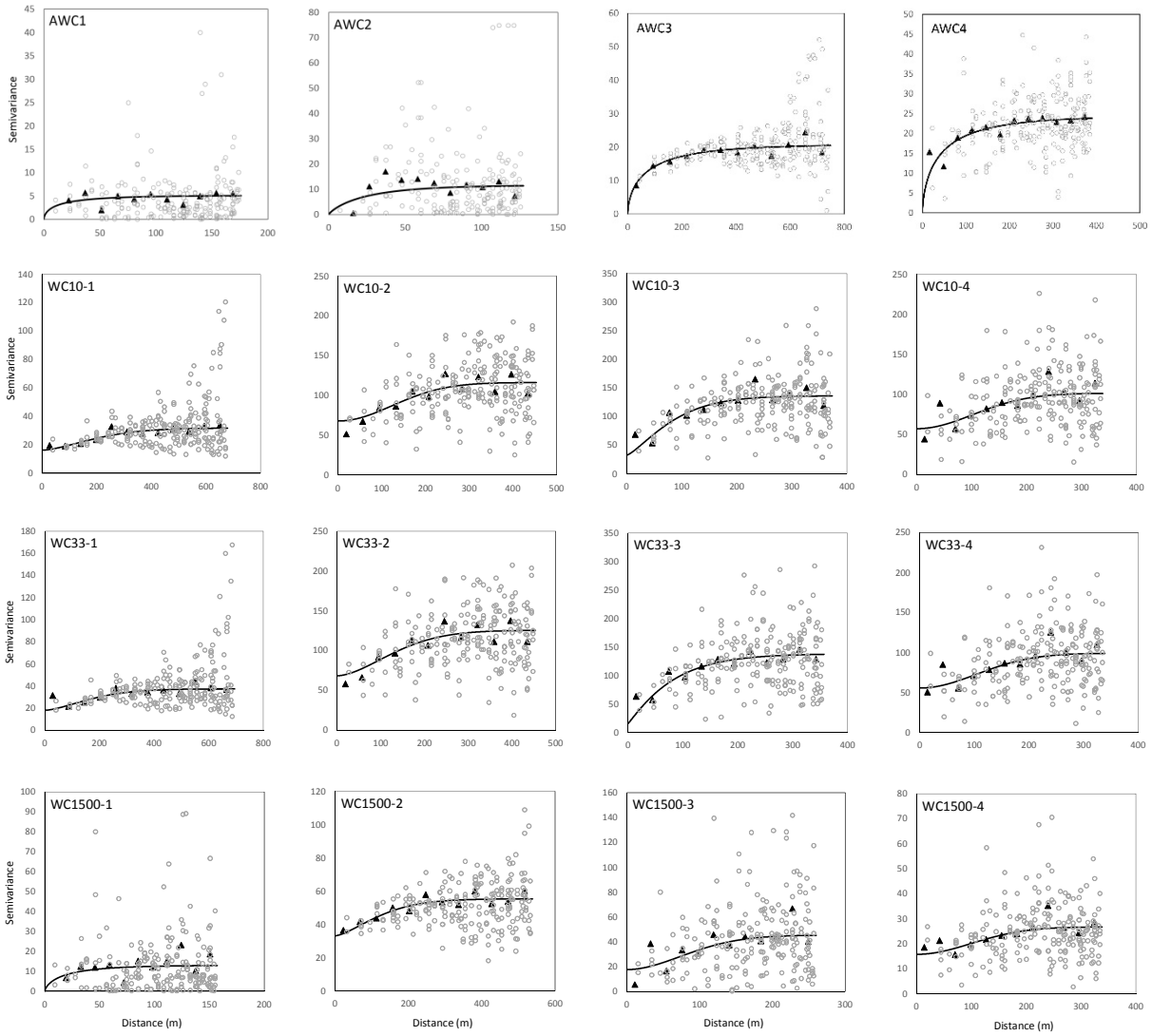

Figure 2-8. The semivariograms for water retention data. AWC*a*: available water content at layer *a*; WC*a-b*: water content at matric potential *a* (kPa) and layer *b*.

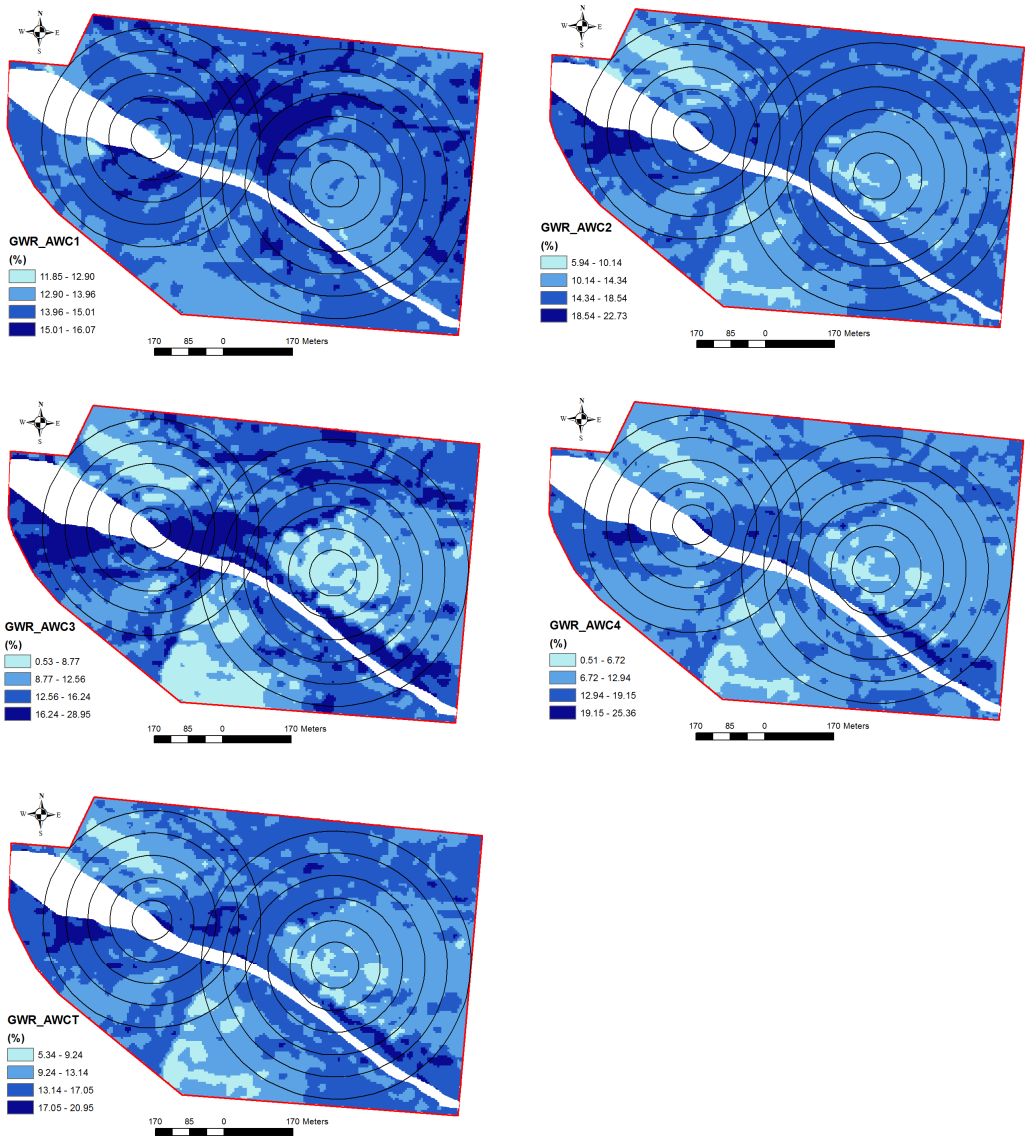

Figure 2-9. AWC maps for different layers (i.e. AWC1, AWC2, AWC3, and AWC4) and for the entire effective root zone (i.e. AWCT) generated using the model 8 (Figure 4).

Table 2-1. Properties of the soils selected from the UNSODA data set and collected from the field of study.

	UNSODA (*n*=554)				The field (*n*=400)			
	Min	Max	SD	Average	Min	Max	SD	Average
Sand (%)	0.1	99.6	31.2	53.8	0.0	96.9	30.2	53.9
Silt (%)	0.2	87.1	22.9	29.7	0.0	72.8	19.7	27.3
Clay (%)	0.0	63.0	13.6	16.6	0.3	56.6	12.7	18.8
BD (Mg m^{-3})	0.46	1.97	0.22	1.46	1.06	1.86	0.12	1.36

BD: bulk density, SD: standard deviation.

Table 2-2. The RMSE (cm^3 cm^{-3}) for the models (Figure 4) predicting water content at FC and PWP at different layers.

		Model									
(cm)	*h* (kPa)	1	2	3	4	5	6	7	8	9	10
0-25	10	0.049	0.050	0.042	0.046	0.041	0.076	0.042	0.042	0.044	0.043
	33	0.052	0.055	0.044	0.049	0.042	0.071	0.043	0.044	0.045	0.048
	1500	0.039	0.040	0.029	0.034	0.033	0.055	0.029	0.030	0.030	0.031
25-50	10	0.098	0.099	0.079	0.091	0.078	0.099	0.081	0.079	0.081	0.084
	33	0.101	0.103	0.082	0.094	0.084	0.097	0.083	0.082	0.083	0.088
	1500	0.069	0.070	0.058	0.063	0.058	0.070	0.058	0.058	0.059	0.061
50-75	10	0.103	0.112	0.085	0.100	0.084	0.092	0.090	0.084	0.095	0.088
	33	0.102	0.109	0.085	0.099	0.084	0.090	0.088	0.084	0.091	0.088
	1500	0.059	0.061	0.051	0.059	0.052	0.053	0.051	0.051	0.052	0.053
75-100	10	0.091	0.096	0.078	0.090	0.075	0.089	0.076	0.075	0.077	0.074
	33	0.090	0.093	0.077	0.089	0.074	0.087	0.073	0.073	0.074	0.077
	1500	0.047	0.048	0.040	0.047	0.050	0.059	0.041	0.040	0.041	0.052
RMSE-	**mean**	0.075	0.078	0.063	0.072	0.063	0.078	0.063	0.062	0.064	0.066

Table 2-3. The MBE (cm^3 cm^{-3}) for the models (Figure 4) predicting water content at FC and PWP at different layers.

		Model									
(cm)	*h* (kPa)	1	2	3	4	5	6	7	8	9	10
0-25	10	0.002	0.005	-0.004	0.000	0.000	-0.012	0.004	0.000	-0.003	-0.006
	33	0.001	0.009	-0.004	0.000	0.001	0.000	0.005	0.000	0.000	0.001
	1500	0.003	0.009	0.000	0.003	-0.005	0.023	0.005	0.000	0.005	-0.002
25-50	10	0.005	0.020	0.001	0.007	0.000	-0.013	0.005	0.001	0.005	0.000
	33	0.004	0.025	0.001	0.009	0.000	0.002	0.008	0.000	0.006	-0.004
	1500	0.003	0.016	0.000	0.007	-0.001	0.020	0.006	0.000	0.004	-0.002
50-75	10	0.005	0.039	-0.001	0.019	0.001	0.011	0.019	-0.002	0.024	0.001
	33	0.005	0.034	-0.001	0.016	0.001	0.012	0.019	-0.002	0.021	0.000
	1500	0.003	0.012	0.000	0.006	0.009	0.008	0.009	-0.001	0.009	-0.001
75-100	10	0.000	0.027	-0.004	0.013	0.000	0.016	0.019	0.001	0.012	-0.003
	33	-0.002	0.019	-0.005	0.008	-0.001	0.017	0.015	0.000	0.009	-0.008
	1500	-0.001	0.005	-0.002	0.003	0.011	0.012	0.005	0.000	0.000	0.014

Table 2-4. Correlation coefficients between ECa and yield data with water retention data at different layers, predicted by PC-PTF for samples (n= 100 (points) × 4 (depths) = 400).

layer	Matric potential, (kPa)	ECa -shallow (mS m^{-1})	ECa- deep (mS m^{-1})	Cotton 2012 (Mg ha^{-1})	Soybean 2009 (Mg ha^{-1})
1st	10	0.68	0.50	0.35	0.42
	33	0.72	0.54	0.36	0.44
	1500	0.75	0.58	0.36	0.45
2nd	10	0.59	0.67	0.50	0.40
	33	0.59	0.68	0.50	0.39
	1500	0.56	0.64	0.44	0.34
3rd	10	0.45	0.65	0.57	0.42
	33	0.45	0.65	0.55	0.40
	1500	0.42	0.61	0.51	0.34
4th	10	0.39	0.61	0.51	0.46
	33	0.40	0.62	0.50	0.44
	1500	0.39	0.59	0.44	0.38
AWCT*	10	0.58	0.75	0.64	0.55
AWCT	33	0.52	0.69	0.65	0.56

* AWCT: available water content throughout the effective root zone (=1 m).

Part 3: Perspectives on cotton supplemental irrigation in west Tennessee

Abstract

While irrigated acreage is shrinking in some arid regions due to increasing competition for water, supplemental irrigation is expanding in humid regions as a means to avoid sporadic periods of water stress and thus maintain high crop yields. The complexity of the interaction between the temporal variability of precipitation within a cropping season and the spatial heterogeneity of soil properties poses a problem for supplemental irrigation management. A 2-year (2013-2014) on-farm experiment on a 73-ha farm in humid West Tennessee was conducted with the goal of understanding cotton yield spatial patterns in relation to the heterogeneity of soil physical properties and supplemental irrigation regimes. We used a wireless network of soil moisture sensors and an annual time series of satellite remotely sensed indices including tasseled cap transformation (TCT) and normalized difference vegetation index (NDVI). In addition, we analyzed temporal stability of yield spatial patterns using 5 years of available historical yield data (corn, soybean and cotton) that had been collected by the farmer prior to our study. Supplemental irrigation increased cotton yields by 32% in 2013 and 41 % in the 2014 across all soil types when compared to rain-fed conditions. However, cotton lint yield had a 42% difference between 2013 and 2014, a result attributed to different planting dates in 2013 and 2014 that in turn affected crop growth and yield. Cotton lint yield maps correlated the highest with NDVI ($r = 0.82$) 93 days after planting. We showed, if appropriate space images are available, yield spatial patterns and pattern of soil spatial heterogeneity (in respect to soil available water content) is distinguishable by TCT. In addition, with TCT we successfully separated data points underneath an irrigation system using a space image taken shortly after an irrigation event. The analysis of yield data for several years for multiple crops indicated that the soil physical and hydraulic properties influenced crop growth and yield, hence variable rate supplemental irrigation is suggested for the field of study.

Chapter 3: Studying crop yield response to supplemental irrigation and the spatial heterogeneity of soil physical attributes in a humid region

(To be submitted to *Precision Agriculture* for publication)

Amir Haghverdi[1*], Brian G. Leib[1*], Robert A. Washington-Allen[2], Paul D. Ayers[1], Michael J. Buschermohle[1]

[1]Department of Biosystems Engineering & Soil Science, University of Tennessee, 2506 E.J. Chapman Drive, Knoxville, TN 37996-4531

[2] Department of Geography, University of Tennessee, Burchfiel Geography Building, 1000 Phillip Fulmer Way, Knoxville, TN 37996-0925

* Correspondence author, ahaghver@vols.utk.edu, (865)2359694

Abstract

While irrigated acreage is shrinking in some arid regions due to increasing competition for water, supplemental irrigation is expanding in humid regions as a means to avoid sporadic periods of water stress and thus maintain high crop yields. The complexity of the interaction between the temporal variability of precipitation within a cropping season and the spatial heterogeneity of soil properties poses a problem for supplemental irrigation management. A 2-year (2013-2014) on-farm experiment on a 73-ha farm in humid west Tennessee was conducted with the goal of understanding cotton yield spatial patterns in relation to the heterogeneity of soil physical properties and supplemental irrigation regimes. We used a wireless network of soil moisture sensors and an annual time series of satellite remotely sensed indices including tasseled cap transformation (TCT) and normalized difference vegetation index (NDVI). In addition, we analyzed temporal stability of yield spatial patterns using 5 years of available historical yield data (corn, soybean and cotton) that had been collected by the farmer prior to our study. Supplemental irrigation increased cotton yields by 32% in 2013 and 41 % in the 2014 across all soil types when compared to rain-fed conditions. However, cotton lint yield had a 42% difference between 2013 and 2014, a result attributed to different planting dates in 2013 and 2014 that in turn affected crop growth and yield. Cotton lint yield

maps correlated the highest with NDVI (r = 0.82) 93 days after planting. We showed, if appropriate space images are available, yield spatial patterns and pattern of soil spatial heterogeneity (in respect to soil available water content) is distinguishable by TCT. In addition, with TCT we successfully separated data points underneath an irrigation system using a space image taken shortly after an irrigation event. The analysis of yield data for several years for multiple crops indicated that the soil physical and hydraulic properties influenced crop growth and yield, hence variable rate supplemental irrigation is suggested for the field of study.

Keywords: apparent electrical conductivity, remote sensing, wireless soil moisture monitoring, yield map.

1. Introduction

Irrigated agriculture has been playing a globally significant role in providing nearly 30 percent of the total food and fiber supply [Food & Agriculture Organization (FAO), 2013]. While irrigated acreage is shrinking in some arid regions due to increasing competition for water, supplemental irrigation is expanding in humid regions as a means to avoid unpredicted periods of water stress and maintain high yields (National Agricultural Statistics Service (NASS), 2010). For example, in west Tennessee, row crop irrigation has expanded rapidly from twenty-five center pivot irrigation systems installed in 2007 to 270 systems installed in 2012. This represents an expansion of 16,000 ha of cropland per year under supplemental irrigation (Tennessee Farm Bureau Federation, 2013 which necessitates an essential demand to study supplemental irrigation management of different crops in this region.

Precipitation is the prime source of moisture in west Tennessee. However, severe in-season drought conditions for short periods are likely to occur which could substantially reduce yields under rainfed agricultural practices. Supplemental irrigation is an irrigation strategy that attempts to maintain maximum yield production by irrigating during periods of insufficient rainfall to fulfill the crop water requirements. The application of supplemental irrigation management is a spatially and temporally complex problem in west Tennessee where precipitation patterns are temporally variable within and across cropping seasons and interacts with the spatial mosaic of the physical and hydraulic attributes of alluvial and windblown loess deposited soils. Soil properties, such as texture and bulk density, greatly affect soil water retention and movement and govern readily available soil water for crop and irrigation management. Excess water content within the root zone could occur if irrigation events overlap with unpredicted rainfall. This may cause lack of sufficient aeration and consequently yield reduction. Moreover, soil

erosion and nutrient leaching are also prone to happen, which in turn increase the risk of contaminating nearby surface or ground waters.

Cotton is a major crop in west Tennessee that is grown in more than 15 states and is vital to the US economy because it is a critical export-oriented product. Currently, some 40% of US cotton is under irrigation with the area expanding throughout the mid to southern US. Given the limited water resources in many cotton-growing areas, a considerable amount of research has recently been performed on cotton irrigation to improve water use efficiency (Vellidis et al., 2014). However, inconsistent cotton yield have been observed in response to irrigation in the humid portion of the US (Pettigrew, 2004). Suleiman et al. (2007) studied the use of cotton deficit irrigation in a humid climate using FAO's 56-crop coefficient method in Georgia and suggested establishing a 90 % irrigation threshold for full irrigation of cotton in humid climates. Bajwa and Vories (2007) evaluated cotton canopy response to irrigation in a moderately humid area in Arkansas and found that under wet conditions excessive irrigation decreased the yield of cotton lint. A similar result was reported by Bronson et al. (2006), who also found excessive rainfall limited the yields from irrigation. Gwathmey et al. (2011) conducted a 4-year supplemental irrigation study in Jackson, Tennessee and found significantly improved lint yield by a mean of 38 % at the 2.54 cm wk-1 irrigation rate in comparison with 3 of 4 years rainfall. Duncan (2012) used surface drip irrigation system to investigate cotton yield response to irrigation across soil types with different water holding capacities (WHC). This study illustrated that uniform irrigation is not the optimum management decision for the cotton wherever field-scale soil heterogeneity affects spatial distribution of soil available water content (AWC).

Traditionally, irrigation studies were limited to small plots at research stations, mostly due to economical and computational limitations. Additionally, contemporary constraints to irrigation studies include the personnel time and expense for data collection as well as the limitations of conventional computational infrastructure and statistical methods to analyze the increasingly larger spatiotemporal datasets with inherent noise and uncertainty. In west Tennessee, the inherent heterogeneity and the spatiotemporal changes in soil and weather-related attributes of the region make it hard to extrapolate the results of design-based experiments on small plots to real field conditions. Supplemental irrigation scheduling is a site-specific irrigation management system where each field has its own irrigation management challenge that requires unique solution(s). On-farm experimentation is an alternative for design-based experiments since collecting site-specific information is becoming more and more common and affordable in US agriculture.

In contemporary agriculture, precision farming enables farmers to locally collect various site-specific information such as yield and soil apparent electrical conductivity data (ECa). Crop yield maps provide valuable quantitative information on crop production, change in production, and the response of crop production to different agricultural inputs, including irrigation and fertilizer. In addition, topographic maps of edaphic features including elevation, slope, and aspect, are usually freely available. If not, these site-specific attributes can be measured and mapped without spending a considerable amount of time and money. Recently new wireless technology has enabled progressive farmers to remotely and continuously monitor soil properties over time including soil temperature, AWC, and soil matric potential. Additionally, 40 years of historical archives of medium-resolution (30-m X 30-m pixel resolutions) satellite remote sensing data, specifically Landsat satellite imagery, are freely available to conduct annual and interannual time series analyses of the status of crop yields over a number of growing seasons. Consequently, the goal of this study is to understand the interactions of supplementary irrigation and the spatial variability of soil attributes such as AWC to crop (particularly cotton) growth and yield in west Tennessee. This study's objectives are to conduct an on-farm experiment to: (i) compare cotton lint yield under different supplemental irrigation regimes across different soil types, (ii) use a wireless network of soil moisture sensors and an annual time series of satellite remotely sensed indices to interpret cotton lint yield spatial patterns in relation to the heterogeneity of soil physical properties and supplemental irrigation and (iii) determine the temporal stability of spatial yield patterns using 7 years of yield data for different crops.

2. Material and Methods

2.1 Study area

We conducted the on-farm experiment on a 73-ha property in Dyer County, west Tennessee that is located on the flood plains of the Mississippi river (Figure 3-1). This field contains Robinsonville loam and fine sandy loam, Commerce silty clay loam, and Crevasse sandy loam soils that were formed in the loamy and sandy alluvium of the Mississippi river terraces (Soil Survey Staff, 2015). The study area's elevation ranges from 77 m to 80 m and is in the short season semi-humid region of west Tennessee where long term weather records indicate 97 mm monthly mean precipitation and 21 oC temperature throughout the May to November growing season (National Climate Data Center, 2015).

2.2 Irrigation experiment

The on-farm experiment was conducted for 2-years and designed to study the supplemental irrigation-cotton lint yield relationship across different soil types. The farmer no-tillage planted the field to 'PHY375' cotton variety on May 30, 2013 and to 'Stoneville 4946' on May 5, 2014. The farmer practiced variable rate potassium and phosphorus application, in accordance with soil test recommendations. However, nitrogen was applied uniformly. Crop pest management was implemented following state extension recommendations and the field was harvested between December, 2 and 3 2013 and in the 2nd year between October 18-20 2014.

One of the goals of this study was to engage the farmer in the irrigation decision making process, thus enhancing the producer's knowledge and expertise over the course of the study. Indeed, the producer was the prime decision maker on how much and when to irrigate using two center pivots for crop irrigation (Figure 3-1). At the same time, we wanted to make sure that the farmer was provided sufficient information to irrigate appropriately while maintaining statistical variability of the supplemental irrigation water applied (IW) across the field to fulfill our research purpose.

In 2013, the farmer adopted The Management of Irrigation Systems in Tennessee (MOIST) program (http://bioengr.ag.utk.edu/weather/) and installed a soil sensor network to schedule irrigation throughout the growing season. MOIST is an irrigation decision support tool that delivers irrigation recommendations by simultaneously measuring and monitoring soil water status and calculating water balance through a deployed wireless soil sensor network. We used MOIST to discuss the efficiency of irrigation management with the farmer. Additionally, we installed and provided online access to the farmer to another network of sensors from AgSpy (http://www.aquaspy.com/AgSpyAnalytics) and Decagon (http://www.decagon.com/). This network allowed us to monitor weather conditions, soil water status, and the performance of the irrigation systems throughout the irrigation season. In 2014, we started sending out a weekly MOIST report to the farmer.

Two different methods were used to vary supplemental IW across the field: 1) programming the two pivots and 2) partially swapping the sprinkler nozzles. Figure 3-2 shows the irrigation experimental design over the course of the study. Table 3-1 summarizes the information on irrigation programs at each pivot. The control panels of each pivot were Valley Select2 (Valmont Industries, Inc) that were programmable up to 9 different pie shape zones (e.g., Figure 3-2). The program changes the irrigation rate by adjusting the pivot's travel speed, where speeding up the pivot causes less irrigation and slowing it down applies additional irrigation. Two adjacent spans of each pivot were

renozzled to apply approximately 50 percent more and less water than the farmer's decision. The center pivots can be operated both clockwise and counter clockwise, but were programmed only in the clockwise direction (Table 3-1). Routine irrigations by the farmer were equal to 15.50 mm and 9.91 mm per revolution for the east and west pivots, respectively. The east (west) pivot panel was programmed to apply ±5.08 (±2.54) mm variation in irrigation per revolution on some pie shape zones (Figure 3-2a). The spans for renozzling were selected based on the pivots characteristics and soil spatial variation. For the east pivot, the high-irrigation (23.25 mm per revolution) ring was located between the second and third towers while the low-irrigation (7.75 mm per revolution) ring was located between the third and fourth towers. For the west pivot the low-irrigation ring (14.87 mm per revolution) was selected to be between the first and second towers while the high-irrigation (4.96 mm per revolution) ring was located between the second and third towers.

Table 3-2 summaries irrigation and weather data for the 2013 and 2014 cropping seasons. The growing degree days (GDD) were assumed to be the mean of the daily minimum and maximum temperature minus the base temperature of 15.6o C (Suleiman et al., 2007). The sensors were installed a couple of weeks after planting and were removed prior to the harvest period. Consequently, in situ data was not available for the whole cropping seasons. However, temperature and precipitation data from the closest weather station was obtained from National Climate Data Center to fill these gaps.

2.3 Instrumentation and measurements

We measured the following physical variables. Figure 3-2 illustrates the spatial arrangement of monitoring stations installed within the field of study. AgSpy soil moisture capacitance probes (3 posts in 2013, 4 posts in 2014) were 1-m in length and obtained measurements at 10 depths (from 10 cm to 100 cm with 10 cm increment). The sensor output is a dimensionless number in the range 0 to 100 called scaled frequency (*SF*) which is defined as:

$$SF = \frac{F_a - F_s}{F_a - F_w} \times 100 \quad (1)$$

where *Fa* is the frequency of oscillation in air (air count), *Fs* is the frequency of oscillation in soil (soil count) and *Fw* is the frequency of oscillation in water (water count). The *Fa* and *Fw* are calculated during the manufacturing of each sensor. The frequency of oscillation is related to capacitance between sensor plates that is in turn influenced by

the relative permittivity of the soil media. The relative permittivity of water is significantly larger than air and soil, thereby changes in soil water content will be accurately detected by the AgSpy (Agrilink, 2006).

An on-farm weather and soil monitoring station that contained an EM50G remote data logger, a vp-3 temperature and relative humidity sensor, an ECRN-100 high-resolution rain gauge and a pyranometer solar radiation sensor (Decagon Devices Inc., Pullman, WA) was installed in 2013 and run through 2014 using the MOIST program. The data was recorded once per hour, stored in the logger, and then transmitted to an on-line interface. The station also had two Mps-2 soil matric potential and temperature sensors installed approximately at 10 and 46 cm depths to monitor soil water status (Decagon Devices, Inc., 2014). A weekly report on soil water status and irrigation scheduling then was sent to the farmer based on soil sensors and water balance calculations using MOIST. MOIST calculated the daily evapotranspiration (ET) using Turc's 1961 equation (developed for regions with relative humidity > 50 %) as follows (Lu et al., 2005):

$$ET_p = 0.013\left(\frac{T}{T+15}\right)\left(R_s + 50\right) \quad (2)$$

where ET_P is the daily potential evapotranspiration (mm d^{-1}); Rs is the daily solar radiation (cal cm^{-2} d^{-1}) and T is the daily mean air temperature (°C).

Three additional Decagon posts were added in the second year, to monitor soil water status and center pivots performace, each had an EM50G remote data logger, an ECRN-100 high resolution rain gauge and two Mps-2 soil matric potential and temperature sensors (installed approximately at 10 and 46 cm).

2.4 Satellite data processing and transformation

The reflectance of a crop canopy and the bare soil background will change throughout a cropping season as plants go through different growth stages and soil water status is changed. Proxy indicators of crop phenological and yield dynamics can be derived from moderate spatial resolution (900-m^2 pixels) Landsat 8 satellite data. In this study, 18 cloud-free Landsat 8 Operational Land Imager (OLI) scenes for 2013 (7 scenes from April to December) and 2014 (11 scenes from February to September) of the study area were acquired (Table 3-3), subset to the study site, standardized to reflectance, corrected for atmospheric effects, and transformed to three different vegetation indices to track the growing season relationship of these proxies to contemporaneous field-measured yields of cotton lint. These remotely sensed vegetation or crop indices have been found to be proxies for vegetation biomass, canopy

cover, and productivity and include the simple ratio (SR, Birth and McVey 1968), the normalized difference vegetation index (NDVI, Rouse et al., 1974), and the green NDVI (GNDVI, Gitelson et al., 1996). These indices are calculated as:

$$SR = \frac{NIR}{R} \tag{3}$$

$$NDVI = \frac{NIR - R}{NIR + R} \tag{4}$$

$$GNDVI = \frac{NIR - G}{NIR + G} \tag{5}$$

where NIR, R, and G represent reflectance values of the near-infrared, red and green bands or channels, respectively.

This annual time series was also transformed to the newly developed Tasseled Cap Transformation (TCT) for Landsat 8 (Kauth and Thomas, 1976, Baig et al., 2015) in order to characterize and evaluate the spatiotemporal pattern in soil moisture and crop phenology. The TCT helps understand crop phenology and soil moisture dynamics over the cropping season by reducing the dimensionality or degree of correlation between the spectral bands of the original Landsat 8 data by consolidating the highly correlated bands with each other and increasing the degree of decorrelation between the new bands (Baig et al., 2015). Three primary new bands have been historically identified to provide site information on soils: the Soil Brightness (SBI) or Brightness index (BI), soil and vegetation moisture content: the Wetness index (WI), and vegetation cover or biomass: the Greenness index (GI). Plotting these indices in three-dimensional (3D) Cartesian space versus each other shows the "tassel cap space" or the plane of soils (i.e. Brightness versus Wetness), transition zone (i.e. Greenness versus Wetness) and plane of vegetation (i.e. Brightness versus Greenness). Plotting a time series of the Tassel Cap in 3D Cartesian space tracks the phenological dynamics of a crop (Kauth and Thomas, 1976, Baig et al., 2015). These remote sensing-based proxies can then be used to track crop dynamics within a growing season (Kauth and Thomas 1976). The yield data points for the 2013 and 2014 cropping seasons were averaged within each cell (i.e. 900-m^2 pixels) using ArcGIS 10.2.2 (ESRI Inc., Redlands, California) and then were correlated to crop indices values.

The field of study was subset from the 18 full scenes and its border cells and cells within the drainage pathways were removed. The image's digital numbers were converted to top-of-atmosphere reflectance using the

calculations described in the online USGS Landsat 8 User's Handbook (http://landsat.usgs.gov/Landsat8_Using_Product.php). The standardized reflectance time series was then atmospherically corrected using an automated procedure developed by Washington-Allen et al. (2006, http://www.gis.usu.edu/~doug/serdp/tutorial/PIF/index.html) that was based on concepts developed by Hall et al. (1991).

A secondary analysis was performed to examine the power of TCT planes for recognizing the spatial yield patterns and spatial heterogeneity of soil surface WHCs throughout the study site. Three zones of low, medium and high yields were delineated based on the statistical distribution of yield data using natural break (Jenks) classification method. Then, TCT planes were color coded for visual pattern recognition. The same zoning and color coding approach were implemented to divide the surface soil depth (i.e. 0-25 cm) into three zones with low, medium and high WHCs. We were also interested in finding any patterns in TCT planes illustrating usage of the irrigation system. We hypothesized that the TCT plane of soils, i.e., the SBI versus Wetness, should be useful to back-track irrigation practices if the satellite image had been taken immediately after an irrigation event by either of the pivots. To accomplish this goal, cells underneath the eastern and western pivots and the rainfed areas were color coded and analyzed within the TCT plane of soils. Finally, we performed a visual comparison between TCT planes of two images (also color coded to show soil and yield zones) in 2013 and 2014 cropping seasons to see if the TCT planes could help us explain differences in expected yield across years.

2.5 Temporal stability of yield patterns

Our main focus for this research was on cotton yield in two years, however, to better understand spatiotemporal dynamics of changes in yield, several years with different crops should be considered (Joernsgaard and Halmore, 2003). Except for 2011, yield data from 2007 to 2012 (i.e. corn 2007, corn 2008, soybean 2009, cotton 2010, cotton 2012) had been collected by the farmer using appropriate yield-monitor-equipped harvesters. We combined these data with the 2013 and 2014 yield data to analyze yield relative difference and temporal variance.

A multistep filtering process was designed and implemented in Microsoft Excel and ArcGIS 10.2.2 to process the data and produce yield maps. First, yield maps were visually assessed. A thematic map was generated to investigate the GPS tracks .The spatial yield pattern across the field was cross validated against the farmer knowledge and our experience on field situation. Second, multiple filters were designed (e.g. using swath width,

distance, speed of harvester, change in speed) to remove outliers and erroneous data points. Yield data that was within ± 3 standard deviations of the mean were assumed to be outliers and removed from the analysis. Then the field was divided into 100 m^2 cells, and relative yield difference (equation 6) and yield temporal variance (equation 7) across years were calculated as follows (Basso et al., 2007):

$$\bar{y}_i = \frac{1}{n}\sum_{k=1}^{n}\left[\frac{y_{i,k} - \bar{y}_k}{\bar{y}_k} \times 100\right] \quad (6)$$

where n is the number of years with yield data available, $\bar{y}_i$ is the average percentage yield difference at cell i, $\bar{y}_k$ is the average yield (Mg ha^{-1}) across cells at year k, and $\bar{y}_{i,k}$ is the yield value (Mg ha^{-1}) at cell i at year k.

$$\bar{\sigma}_i^2 = \frac{1}{n}\sum_{k=1}^{n}\left(y_{i,k} - \bar{y}_{i,n}\right)^2 \quad (7)$$

where $\bar{\sigma}_i^2$ is the temporal variance at cell i, $\bar{y}_{i,n}$ is the average yield across the n years and rest of the variables are as previously defined.

3. Result and Discussion

3.1 Cotton lint yield

Figure 3-3 shows cotton lint yield across different soil types and irrigation regimes in 2013 and 2014. The data for Figure 3-3 belongs to selected pie sections underneath both pivots and corners of the field that were rainfed. ECa information had a high correlation with soil hydraulic properties across the field of study, i.e. higher ECa corresponded to higher WHC and AWC (chapters 1 and 2), hence the selected regions were categorized into different groups. Table 3-4 summarizes the correlation coefficients between yield data and some soil properties across the field of study. To evaluate the effect of variable rate fertilizer application by the farmer on yield spatial variation, we also obtained correlation information for p and k. In general, there was an increase in yield from soils with lower AWC to soils with higher AWC, and one exception was the yield decline for regions with high AWC underneath west pivot at 2013. This is in line with a moderate correlation between water content and soil texture with the yield data in 2014. For coarse-textured soils, there was an overall positive response to supplemental irrigation which was more consistent at the east pivot. There was no consistent pattern for intermediate soil, but a

yield reduction for soils with higher AWC under higher irrigation regimes. In general, correlation coefficients were higher for 2014 data than for that of 2013. The correlations between *p*, *k* and yield data were negligible.

3.2 Soil water status

Figures 3-4 and 3-5 depict the dynamic of soil moisture at different depths and locations over the 2013 and 2014 growing seasons, respectively. There were some missing data and bad readings mostly in 2014. In 2013, posts *a*, *b* and *c* (Figure3-2a) received 163 mm, 36 mm and 133 mm seasonal IW, respectively. The average soil WHC throughout the effective root zone (i.e. 1 m) was 28 %, 22 % and 30 % for posts *a*, *b* and *c*, respectively. The pattern of soil moisture was dynamic and varied among sensors at different locations and depths. At post *a*, soil moisture depletion and replenishment occurred for sensors up to 40 cm deep during the monitoring period (i.e. DAP: 42-133). Soil water status stayed almost unchanged for deeper sensor for about 100 DAP, then gradually exhibited a reduction indicating roots started to pull out water from deeper layers as ET demand increased. At post *b*, rainfall plus irrigation kept the soil moisture at a fairly constant level up to about 80 DAP for all sensors, while fluctuation decreased by depth as expected. After that, there was a great depletion in soil moisture for sensors up to 50 cm, which even expanded to deeper sensors about 95 DAP. At post *c*, the overall trend was similar to that of post *a*. Toward end of the season, a big rainfall at DAP equal to 112 mm occurred that refilled the shallow layers for all posts and also penetrated to deeper layers such that there was an increase in soil moisture sensors at 30 and 40 cm and no decrease for deeper sensors up to the end of the monitoring period.

Posts *a*, *b*, *c* and *d* (Figure 3-2a) received 42 mm, 74 mm, 142, and 50 mm of IW, respectively during 2014 cropping season (Figure 3-5). Within the effective root zone (i.e. 1m) the average WHC for posts *a*, *b*, *c* and *d* were 33 %, 28 %, 19 %, and 23 %, respectively. Unlike 2013, most of deep sensors in 2014 showed some fluctuation starting about 70 DAP meaning crop started to use water from deeper layers as crop water requirement increased. We attribute this to bigger crop hence higher ET demand, lower irrigation in 2014 comparing to 2013 and lower irrigation under the west pivot (where we had sensors installed in 2014) comparing to east pivot (where we had sensors installed in 2013). Trends in posts *a*, *b* and *d* were similar except that some sensors (at 50 cm and 100 cm) in post b showed lower readings than expected. There was more fluctuation in shallow sensors in post *c*, since this sensor was irrigated by both pivots and was located on coarse-textured soil with lower AWC. In both years, heavy rainfall events were responsible for big changes in soil water status within soil profile and considering sensors

fluctuation they usually penetrated down to 50 cm. Irrigation events, however, mostly refilled shallow layer up to 20 cm and barley influenced sensors deeper than 30 cm.

3.3 Tasseled cap transformation and vegetation indices

Figures 3-6 and 3-7 depict the TCT (at-satellite reflectance) for 2013 and 2014, respectively, derived from Landsat 8 reflectance; the scatter plot of red against NIR band together with scatter plots of brightness, greenness and wetness are illustrated. Of the 7 scenes analyzed in 2013 (Table 3-3); 3 were taken before planting (i.e. April 13, April 29, May 15), 3 throughout the growing season (i.e. October 6, October 22 and November 7) and one after harvesting (i.e. December 25). In 2013, among images from bare soil the May 15 image had minimum average greenness and a relatively wide range of brightness which was somewhat close to what is called line of zero vegetation for bare soil. Other images from bare soil taken on April 13 and April 29, however, exhibited a different pattern with a higher range of variation for greenness. In fact, their pattern was somewhat similar to that of the first two images acquired throughout the growing season (on Oct 6 and Oct 22) with some slight differences; the images from cotton (on Oct 6 and Oct 22) had narrower red band but greater values for IR band which corresponded to a higher magnitudes of greenness in the plane of vegetation. Analysis of rainfall data revealed that bare soil images (on April 13 and April 29) were taken 2-3 days after moderate rainfall events when soil (at least partially across the field) was expected to be close to its field capacity in respect to its moisture status. Greenness was substantially reduced for last image over the growing season (November 7, DAP = 161) that is believed to be taken after defoliation, the application of chemicals to force cotton leaves to drop. After harvesting image (On December 25) had greenness close to zero and a narrow range of brightness.

There were 11 images for 2014 (Table 3-3) from which four images were taken from bare soil (i.e. February 27, March 31, April 15, May 2) and the rest of them were captured throughout the growing season (i.e. May 18, June 19, July 5, July 21, August 6, August 22 and September 23). In red against NIR scatter plot an inclined line (i.e. the line of zero vegetation for bare soil) was formed by data points belonged to images taken prior to planting (February 27, March 31, April 15, and May 2) and a few days after planting (May 18). As images progressed throughout the cropping season, data clusters exhibited a clear departure from the bare soil line in an orthogonal direction such that at the same time there was a gradual reduction and increase in red and NIR bands reflectance values, respectively.

There was a substantial reduction in NIR reflectance values from image taken on August 22 (DAP 109) to image taken on September 23 (DAP 141) that is believed to be taken after defoliation.

In 2014, as the growing season advanced, the plane of vegetation clearly indicated the growth of crop and canopy size when greenness gradually increased along with a reduction in range and magnitude of brightness. The greenness reached its highest value in images taken on Aug 6 (DAP = 93) and Aug 22 (DAP = 109) when cotton was close to its full canopy. The data cluster for full canopy (almost in parallel with y axis (greenness axis) with a narrow brightness range) was perpendicular to the soil line (which was in parallel with brightness axis with a narrow range on greenness). The data points belonged to the last three images (taken on August 6, August 22 and September 23) formed a line on the transition zone, while last image (taken after defoliation) exhibited a substantial reduction in both greenness and wetness level. On the plane of vegetation, it was even easier to distinguish the after defoliation image from the other images when data points were located approximately with a 45 degree angle difference in counter clockwise order from full canopy image (taken on August 22).

Figure 3-7 illustrates the results of our secondary analysis on application of TCT to recognize patterns and extract extra information on irrigation systems usage, spatial distributions of soil WHC and yield. No distinct pattern was found using TCT planes by available space images for 2013 yield data. We were able to visually distinguish among yield classes for 2014 yield data using the 3D plot of TCT planes and the space image taken on August 6 (Figure 3-7a). There was a continuous increase in both GI and WI from low yield to high yield classes while the range of brightness stayed the same among yield classes. Among available bare soil images the one taken on April 13, 2013 showed the highest separation among soil WHC classes (Figure 3-7b). Moving from soils with low WHC to soil with high WHC a continuous increase in WI and a less pronounced but continuous decline in BI were observed.

Considering DAP as criterion, we chose images taken on October 22, 2013 (DAP=145) and on September 23, 2014 (DAP=141) for point by pint comparison (Figure 3-7c and d). The data points of the images scattered significantly different on the plane of vegetation comparing to one another. As explained earlier, the pattern of data-points for 2013 image on October 22 (i.e. relatively lower brightness range and a broad greenness range) was associated with full canopy while the pattern of data-points for 2014 September 23 image (i.e. relatively wider range of BI and a decline in GI) was attributed to crop after defoliation. Despite the differences in patterns of data-points across these two images, it was possible to visually distinguish among soil classes in both of the images such that

higher greenness was associated with higher WHC (Figure 3-7c). Ta a great extent it was possible to visually distinguish between yield classes in 2014 image on September 23 but three yield classes were mixed and not separable for 2013 image on October 22.

Our records showed that the space image on July 21, 2014 had been taken after an irrigation event at eastern pivot. For this image, in the plane of soils (i.e. brightness against wetness scatter plot) data points formed a fork with two different clusters such that the cells underneath eastern pivot had lower brightness values than those for cells underneath western pivot and rainfed portions of the field (Figure 3-7e).

Correlation coefficient between yield data and crop indices are summarized in Table 3-5. Figure 3-8 depicts maps of NDVI from available space images throughout cropping seasons, 2013 and 2014. Overall, the differences among crop indices were negligible in terms of their correlation with yield data. In 2013, photos taken on October 6 and October 22 had higher max NDVI value equal to 0.68 and 0.67, respectively. In 2014, highest NDVI values belonged to photos taken on August 6 and 22 (=0.73). The lowest NDVI ranged from 0.07 to 0.49 and the minimum value belonged to a photo taken on June 19, 2014. The NDVI maps for 2013 showed spatial patterns similar to that of soil hydraulic properties, yet there was a weak correlation between crop indices and yield for 2013. We attribute this to delayed planting in 2013 (due to cold and wet condition) which in turn affected heat unit accumulation and distribution throughout growing season, hence effecting crop growth and yield. The farmer delayed harvesting, yet there were still some bolls unopened mostly located on region with soil with higher AWC. In 2014, there was a moderate to high correlation between crop indices and yield data for a handful of images. In 2014, as the cropping season progressed, correlation coefficients increased but diminished for the last image (on September 23) which was taken after defoliation. This is in line with the spatial arrangements in NDVI maps that better discriminated areas within the field as cropping season progressed. The highest correlation coefficient (r = 0.82) was obtained between yield data and crop indices derived from space image acquired on August 6 2014 (DAP = 93) which almost coincided with full canopy and cutout. In 2014, higher NDVI values belonged to regions with higher yield mostly located over soils with higher AWC.

3.4 Temporal stability of yield patterns

Figure3-9 illustrates the cotton yield maps for 2012, 2013 and 2014, long term mean yield map (from Eq. 6) and standard deviation yield maps (from Eq. 7). Table 3-6 summarizes the descriptive statistics about yield data

collected for different crops. We included 2012 yield map, a year before this experiment when uniform irrigation had been applied by the famer, to better represent the effect of soil spatial variation on cotton yield. Thematic cotton yield maps for 2012 and 2014 cropping seasons followed the same spatial distribution as soil hydraulic properties map (Figure 3-1b), but the 2013 yield map showed a different pattern. This finding agrees with low correlation between cotton yield data in 2013 and soil properties (Table 3-4). The spatial analysis of the long term mean yield map (ranged from -79 to -98) showed substantial similarity to soil AWC map (Figure 3-1b) developed for the field of study (chapter 2). There were three regions, in southern, eastern and north western parts of the field, with lower yield and all on coarse-textured soils with lower AWC. Highest temporal stability also belonged to part of those low yield regions with low WHC, (i) in the southern part of the field outside pivot coverage and (ii) in an area surrounding east pivot point. The temporal stability was lower for other parts of the field, but it was hard to pick up any cluster of cells with a similar temporal variance. We attribute this to different rainfall patterns and irrigation regimes across years, and their effect on yield across soil types. That is why mean yield varied substantially across years (Table 3-6). For instance, for cotton and corn there was as much as 43 % and 110 % temporal difference between mean yield across years, respectively. In 2012 and 2014, mean yield and standard deviation was higher than that of 2013 indicating a decline in yield on soil with higher AWC in 2013.

4. Discussion

4.1 Yield patterns, soil heterogeneity and supplemental irrigation

Cotton lint response to supplemental irrigation differed across soil types. For soil with low WHC there was a positive response to irrigation in comparison to rainfed where soil moisture deficit is expected to reduce the boll number and yield (Pettigrew 2004). The cotton response to irrigation was not consistent for soil with high WHC except that a yield reduction was occurred underneath both pivots for high irrigation rates in both cropping season. This is in line with reported results in literature indicating under wet conditions excessive irrigation decreased the yield of cotton lint (Bajwa and Vories,(2007; Bronson et al., 2006). In fact, availability of water within root zone did not guarantee cotton yield improvement for soils with medium to high WHCs from irrigation. For example, in 2013 cotton lint yield was only 12 % higher where we placed post *a* (IW= 163 mm and AWC = 28 %) in comparison to yield at post *b* (IW = 36 mm and AWC = 22 %), even though there was a remarkable difference in soil water status throughout growing season between the two spots (Figure 3-4). On the other hand, in 2014 the yield difference

between the exact same points with similar relative difference in irrigation raised by 44 %.We attribute this to wet season and delayed planting in 2013 which significantly affected cotton response to irrigation. Delayed planting influences heat unit accumulation and distribution which was underscored as an important factor for short season cotton response to supplemental irrigation by Gwathmey et al. (2011). Moreover, irrigation is expected to increase the number of bolls but delays cutout (cessation of flowering) since irrigation continues the vegetative growth for a longer period. We believe rapid canopy expansion occurred on soil with higher AWC in 2013 due to excessive water within effective root zone.

4.2 Application of space images and soil moisture sensors

Recent advances in modern instrumentation and measurement techniques such as on-the-go sensing, remote sensing and wireless network of sensors made site-specific on-farm experimentation possible for farmers. This is essential for cotton irrigation management in humid areas as a complex problem due to spatiotemporal variation in soil and weather. In this study we used a variety of information, mostly available for farmers as easy-collected data, to investigate cotton response to supplemental irrigation. We focused mainly on TCT to analyze the available space images.

In red-NIR space, especially for 2014 that we had more images available, it was possible to recognize the shape of a 'tassselled cap', a useful concept that was first introduced by Kauth and Thomas. We showed that TCT was not only useful to investigate spatial differences within a space image, but to temporally compare among images acquired at different times before, throughout and after growing seasons. Every point within red-NIR space was confined by the bare soil line (flat side of the cap) and the point of high vegetation (directly opposite to the soil line). Within this cloud of points each image had a distinct pattern mostly influenced by crop growth stage and soil water status. TCT showed a great potential for field-scale pattern recognition of soil moisture changes and cotton yield classes during irrigation events. It also revealed useful information on yield analysis across years by point by point comparison of two images that had been taken at a similar time after planting in 2013 and 2014. In 2014 both soil and yield classes were differentiable through plane of vegetation indicating that higher soil AWC improved yield on fine-textured soils with higher WHCs. On the other hand, in 2013 only soil classes were differentiable through plane of vegetation. This indicated that higher soil AWC on fine-textured soils in 2013 did not boost the yield and only resulted excessive vegetative growth and/or unopened bolls by the time of harvesting. For those images taken from

bare soil, soil water status clearly influenced reflectance values. When soil water status was close to FC, space images provided useful insight to relative spatial pattern of soil physical and hydraulic properties. Further investigation is needed to clarify if such space images could be used to predict maps of soil hydraulic properties.

Yield and NDVI maps from space images provided good information on field-scale spatiotemporal changes affecting cotton growth and yield. The highest agreement observed between yield map and the space image taken on 93 DAP. This was, from cotton growth stages point of view, somewhat similar to the result by Leon et al. (2003) who reported first bloom to first open boll as the optimum time window to explain cotton spatial variability using aerial photos. Vellidis et al. (2004) obtained 8-14 weeks after planting (9.9 weeks as the average time) as the optimum range when there was the best agreement between yield and aerial photos. They used visible bands for clustering. As season progressed closed canopy masked the spatial patterns, hence the yield pattern got less evident according to them.

Both 2013 and 2014 were wet years. In fact, the rainfall was always above the long term average except for July 2014. There were some heavy rainfall events during the growing seasons which caused significant increase in soil moisture content. This is a typical situation in west Tennessee, with unexpected rainfall events, where temporal changes in rainfall patterns significantly affect yield response to supplemental irrigation across years. Bajwa et al. (2007) reported the same complexity on cotton irrigation scheduling in a moderately humid area in Arkansas when rainfall was plentiful and caused yield reduction for high irrigated crops. Monitoring soil water status revealed that rainfall events refilled the top soil and penetrated into deeper layers while supplemental irrigation mostly influenced the shallow layer up to 20 cm. Therefore, any sustainable irrigation management in this region should take rainfall into account for irrigation scheduling. Sensors indicated fast depletion for soils with lower WHCs. This caused the crop to start using water from deeper layer as cropping season advanced and ET demand increased. The sensor located on overlap region of two pivots showed more frequent irrigation could prevent shallow soil layer from substantial depletion and possible yield reduction due to water stress and thus should be considered as a potential irrigation strategy for coarse-textured soils with low WHCs throughout the study site. This could be beneficial for soil with high WHCs since cotton usually responds favorably to periods of water stress adequate to reduce vegetation expansion.

5. Conclusion

Irrigation has been expanding across the humid areas of the Cotton Belt in the US for the last 20 years (Perry and Barnes, 2012). Stabilizing yields and the recent high commodity prices are the prime reasons for this growth in irrigation investment. In west Tennessee, irrigation management is significantly affected by spatial soil distribution and temporal weather variation. Overall, irrigation improved yield in comparison to rainfed throughout this study. However, we showed that the effect of cropping season length, rainfall pattern and heat unit accumulation /distribution on both cotton growth and development may change or even reverse the expected lint yield from an irrigation treatment for a specific soil type. While soil variation is inherent and not controlled by farmers, irrigation, if well scheduled, could be the key factor to orchestrate the whole cropping system toward optimum condition. Within this study we demonstrated how site-specific information collected by on-the-go sensing, remote sensing and wireless network of sensors could help farmers in irrigation management. In addition, analysis of such information throughout growing season provides insight to potential yield pattern, thus helps the farmer to modify other inputs allocation. Cotton responded differently to irrigation across soil types suggesting that variable rate irrigation (i.e. precision irrigation systems or speed control panels) would probably be beneficial for the field of study. More research, however, is required to precisely quantify the response of yield to different irrigation scenarios across soil types and to find the optimum zoning strategy for the study site.

6. Acknowledgment

This project was funded in part by Cotton Inc. and the United States Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS) Conservation Innovation Grants (CIG). The name of commercial products were only mentioned for information and does not imply any endorsement. We appreciate Pugh Brothers Farms for allowing us to implement this project on their property.

References

Agrilink, (2006). AquaSense™ Soil Moisture Probe. RS-485 Communications, Physical Layer and Simplified Protocol Description. Agrilink Holdings PTY Ltd.

Baig, M. H. A., Zhang, L., Shuai, T., & Tong, Q. (2014). Derivation of a tasselled cap transformation based on Landsat 8 at-satellite reflectance. Remote Sensing Letters, 5(5), 423-431.

Bajwa, S. G., & Vories, E. D. (2007). Spatial analysis of cotton (Gossypium hirsutum L.) canopy responses to irrigation in a moderately humid area. Irrigation Science, 25(4), 429-441.

Basso, B., Bertocco, M., Sartori, L., & Martin, E. C. (2007). Analyzing the effects of climate variability on spatial pattern of yield in a maize–wheat–soybean rotation. European Journal of Agronomy, 26(2), 82-91.

Birth, G.S., and G. McVey, 1968. Measuring the color of growing turf with a reflectance spectroradiometer, Agronomy Journal, 60:640-643.

Bronson, K. F., Booker, J. D., Bordovsky, J. P., Keeling, J. W., Wheeler, T. A., Boman, R. K., Parajulee, M. N., Segarra, E., & Nichols, R. L. (2006). Site-specific irrigation and nitrogen management for cotton production in the Southern High Plains. Agronomy Journal, 98(1), 212-219.

Duncan, H. A. (2012). Locating the variability of soil water holding capacity and understanding its effects on deficit irrigation and cotton lint yield. Master's Thesis, University of Tennessee. http://trace.tennessee.edu/utk_gradthes/1286.

Decagon Devices, Inc. (2014). MPS-2 & MPS-6 Dielectric Water Potential Sensors: Operator's Manual. Available online at: http://manuals.decagon.com/Manuals/13755_MPS-2and6_Web.pdf.

FAO (Food and Agriculture Organization of the United Nations). (2013). Statistical Yearbook 2013: World Food and Agriculture. FAO (Food and Agriculture Organization of the United Nations), Rome, Italy.

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment, 58(3), 289-298.

Gwathmey, C. O., Leib, B. G., & Main, C. L. (2011). Lint yield and crop maturity responses to irrigation in a short-season environment. Journal of Cotton Science 15: 1-10.

Haghverdi, A., Cornelis, W. M., & Ghahraman, B. (2012). A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. Journal of Hydrology, 442, 46-54.

Haghverdi, A., Öztürk, H. S., & Cornelis, W. M. (2014). Revisiting the pseudo continuous pedotransfer function concept: Impact of data quality and data mining method. Geoderma, 226, 31-38.

Hall, F.G., Strebel, D.E., Nickeson, J.E., & Goetz. J. (1991). Radiometric rectification: Towards a common radiometric response among multidate, multisensor images. Remote Sensing of Environment 35 (1), 11-27.

Joernsgaard, B., & Halmoe, S. (2003). Intra-field yield variation over crops and years. European Journal of Agronomy, 19(1), 23-33.

Kauth, R. J., and G. S. Thomas. (1976). The Tasselled Cap—A Graphic Description of the Spectral Temporal Development of Agricultural Crops as Seen by Landsat. Paper presented at the LARS Symposia, Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, IN, June 29–July 1, 4B41–4B51.

Leon, C. T., Shaw, D. R., Cox, M. S., Abshire, M. J., Ward, B., Wardlaw III, M. C., & Watson, C. (2003). Utility of remote sensing in predicting crop and soil characteristics. Precision Agriculture, 4(4), 359-384.

Lu, J., Sun, G., McNulty, S. G., & Amatya, D. M. (2005). A comparison of six potential evapotranspiration methods for regional use in the southeastern United States. Journal of the American Water Resources Association, 41, 621-633.

NASS (National Agricultural Statistics Service). (2010). 2007 Census of Agriculture: Farm and Ranch Irrigation Survey, National Agricultural Statistics Service, USDA. http://www.agcensus.usda.gov/index.php.

National Climate Data Center. (2015). National Climate Data Center home page. Available online at http://www.ncdc.noaa.gov/data-access/land-based-station-data. Accessed [6/13/2015].

Perry C. and Barns E., (eds.) 2012, Cotton irrigation management for humid regions. Cotton, Inc. Cary, N.C.

Pettigrew, W. T. (2004). Moisture deficit effects on cotton lint yield, yield components, and boll distribution. Agronomy Journal, 96(2), 377-383.

Rouse, J.W. Jr., R.H. Haas, D.W. Deering, J.A. Schell, and J.C. Harlan, 1974. Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation, NASA/GSFC Type III Final Report, Greenbelt, MD., 371p.

Soil Survey Staff. (2015). Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at http://websoilsurvey.nrcs.usda.gov/. Accessed [6/13/2015].

Suleiman, A. A., Tojo Soler, C. M., & Hoogenboom, G. (2007). Evaluation of FAO-56 crop coefficient procedures for deficit irrigation management of cotton in a humid climate. Agricultural water management, 91(1), 33-42.

Tennessee Farm Bureau Federation, (2013). Irrigation: Solving Potential Challenges: Policy Development 2013. Available online at: http://www.tnfarmbureau.org/sites/default/files/images/Irrigation%20PD%20paper%202013.pdf.

Vellidis, G., Liakos, V., Perry, C., Tucker, M., Collins, G., Snider, J., Andreis, J., Migliaccio, K., Fraisse, C., Morgan, K., Rowland, D., & Barnes, E. (2014, January). A Smartphone App for Scheduling Irrigation on Cotton. In 2014 Beltwide Cotton Conferences (January 6-8, 2014). Ncc.

Washington-Allen, R.A., Ramsey, R. D., West, N. E., & Efroymson, R. A. (2006). A remote sensing-based protocol for assessing rangeland condition and trend. Rangeland Ecology and Management 59(1), 19-29.

Appendix 3: Chapter 3 Figures and Tables

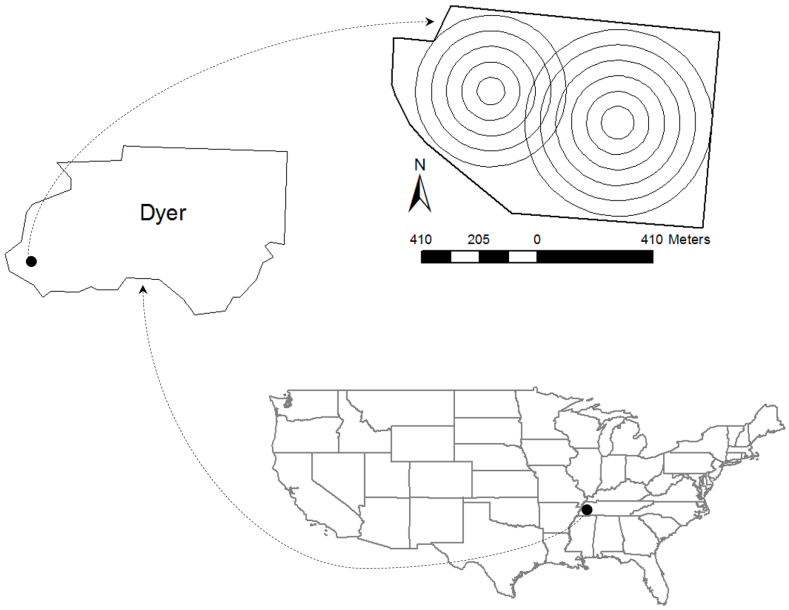

Figure 3-1. The 73-ha study farm and its 2 center pivots in the Dyer County humid region of west Tennessee.

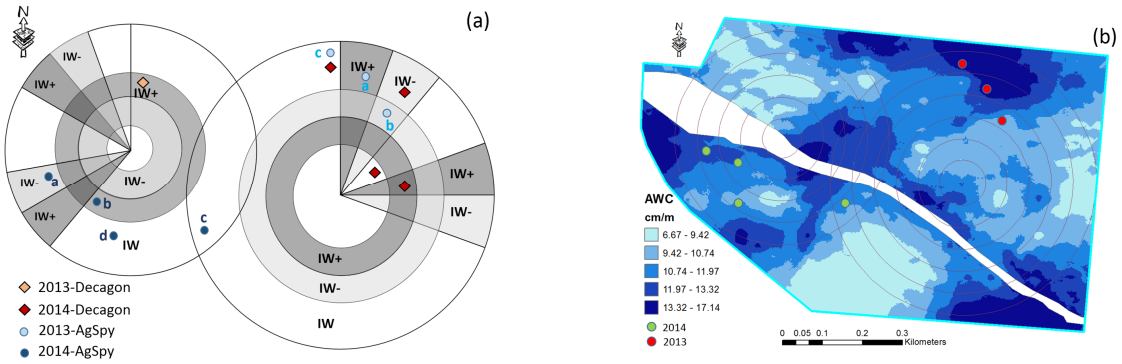

Figure 3-2. The supplemental irrigation experimental design (a) and its orientation in relation to the spatial distribution of soil available water content (AWC) (b). Within Panel a: IW+ and IW- are the regions with more and less irrigated water, respectively than the irrigation applied by the farmer (IW). The light blue (2013) and dark blue (2014) dots and light red (2013) and dark red (2014) diamonds show the Agspy and Decagon sensor installations. The letters assigned to the Agspy sensors locate that sensor's response over the two years. Within Panel b: red and green dots show the Agspy sensors installed during the 2013 and 2014 growing seasons, respectively.

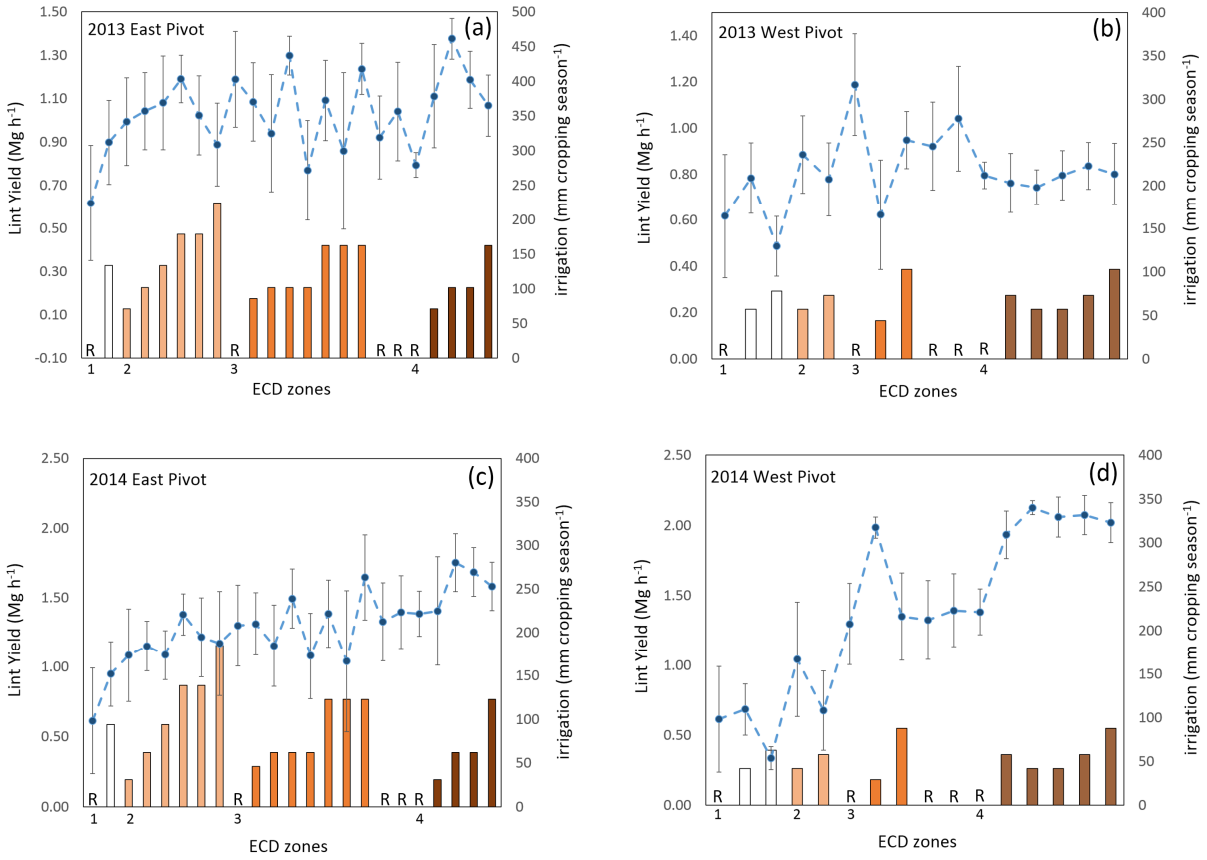

Figure 3-3. Irrigation and soil type effect on cotton lint yield for selected regions across the field. Panel a: East pivot in 2013, panel b: west pivot in 2013, panel c: East pivot in 2014; panel d: west pivot in 2014. Dash line: yield, bars: irrigation, R: rainfed, ECD: ECa deep readings (0-90 cm). ECD increased from zone 1 to zone 4.

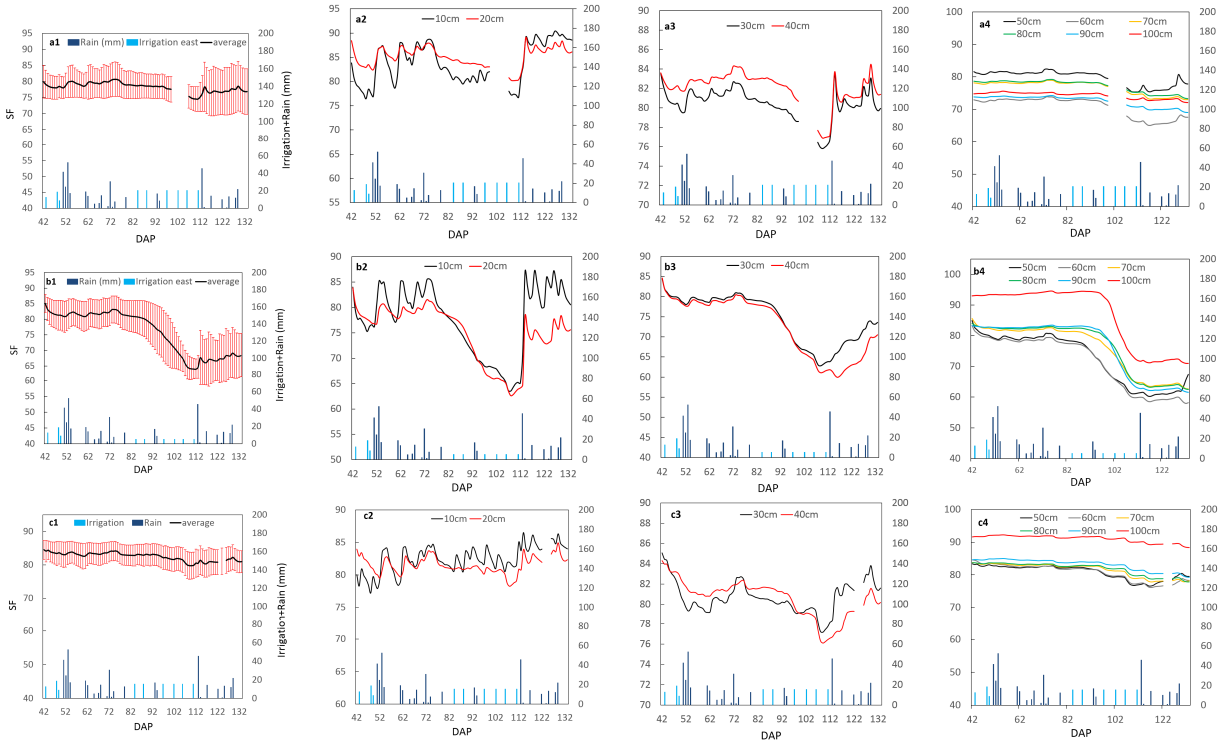

Figure 3-4. Temporal changes [days after planting (DAP)] in soil water status (SF) throughout the 2013 growing season. SF was measured using Agspy soil sensor probes at different locations (designated by letter, e.g., a) and depths (designated by number, e.g., a2: -10 cm and 20-cm).

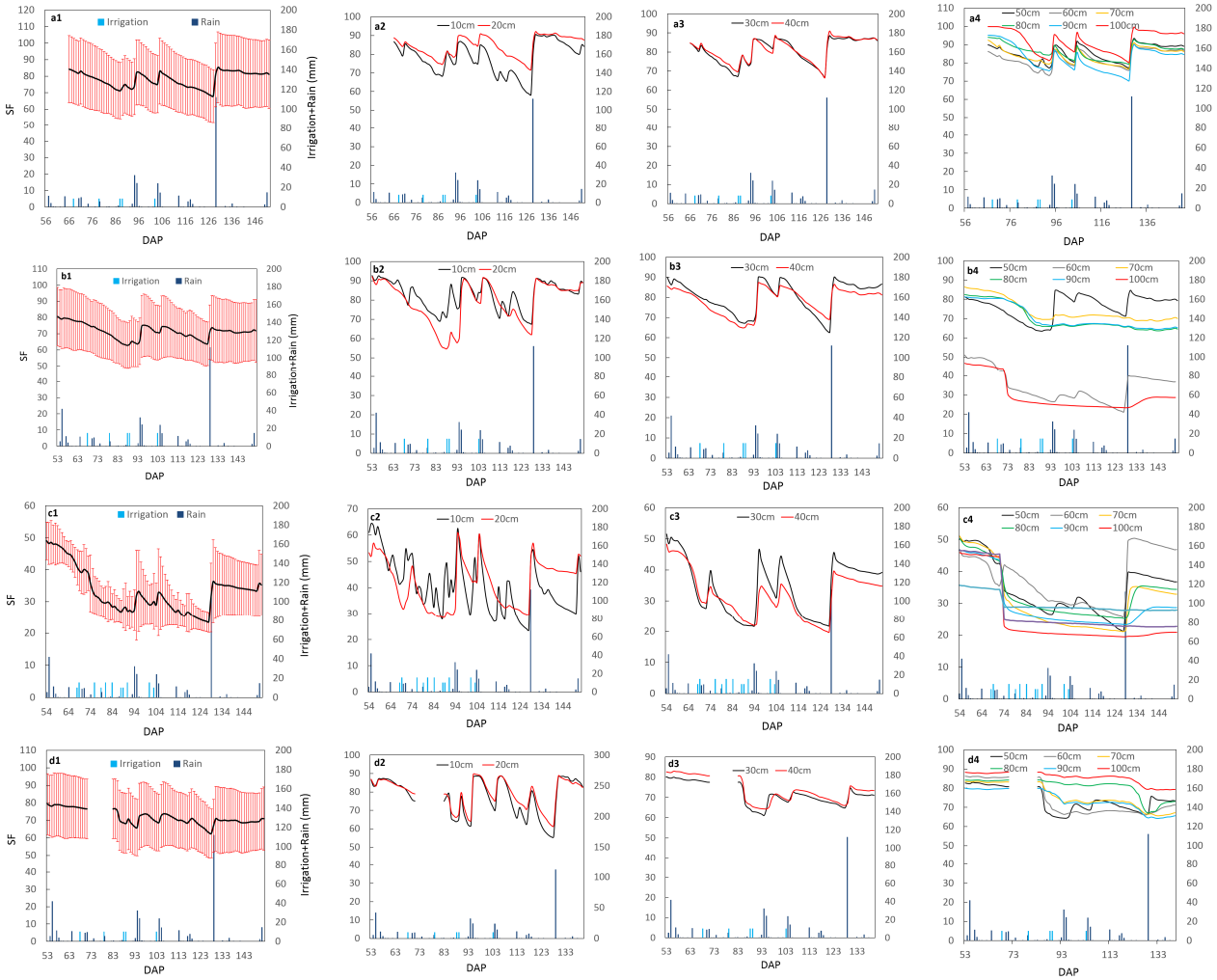

Figure 3-5. Temporal changes [days after planting (DAP)] in soil water status (SF) throughout the 2014 growing season. SF was measured using Agspy soil sensor probes at different locations (designated by letter, e.g., a) and depths (designated by number, e.g., a2: -10 cm and 20-cm).

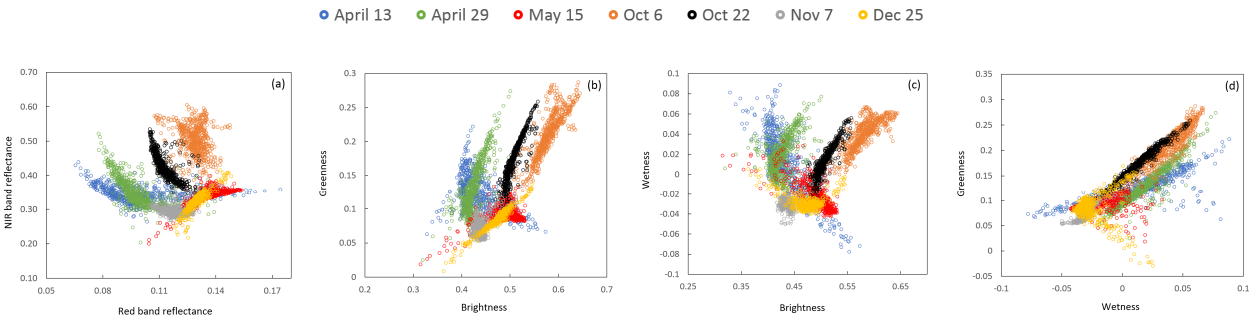

Figure 3-6. The growing season dynamics of Landsat 8 near infrared (NIR) versus red reflectance scenes (a) and combinations of Tasseled Cap Transformation (TCT) planes for Greenness, Soil Brightness, & Wetness (b - d) for seven periods between April 13th to December 25, 2013

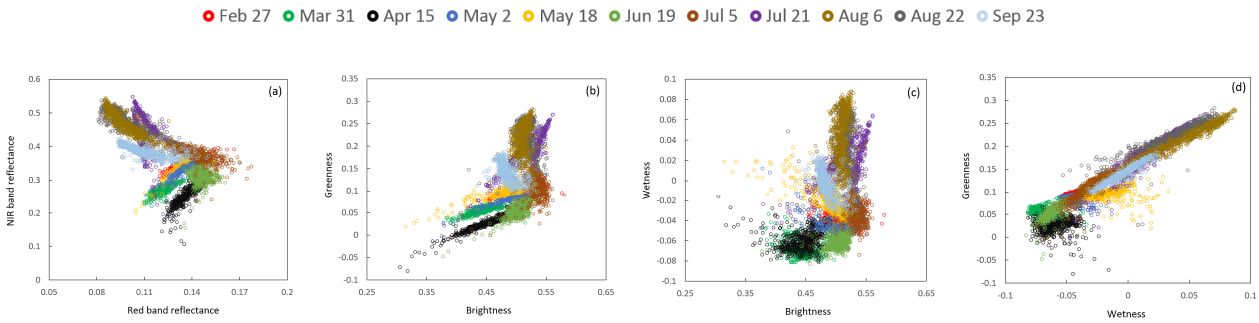

Figure 3-7. The growing season dynamics of Landsat 8 near infrared (NIR) versus red reflectance scenes (a) and combinations of Tasseled Cap Transformation (TCT) planes for Greenness, Soil Brightness, & Wetness (b - d) for eleven periods between February 27th to September 23, 2013.

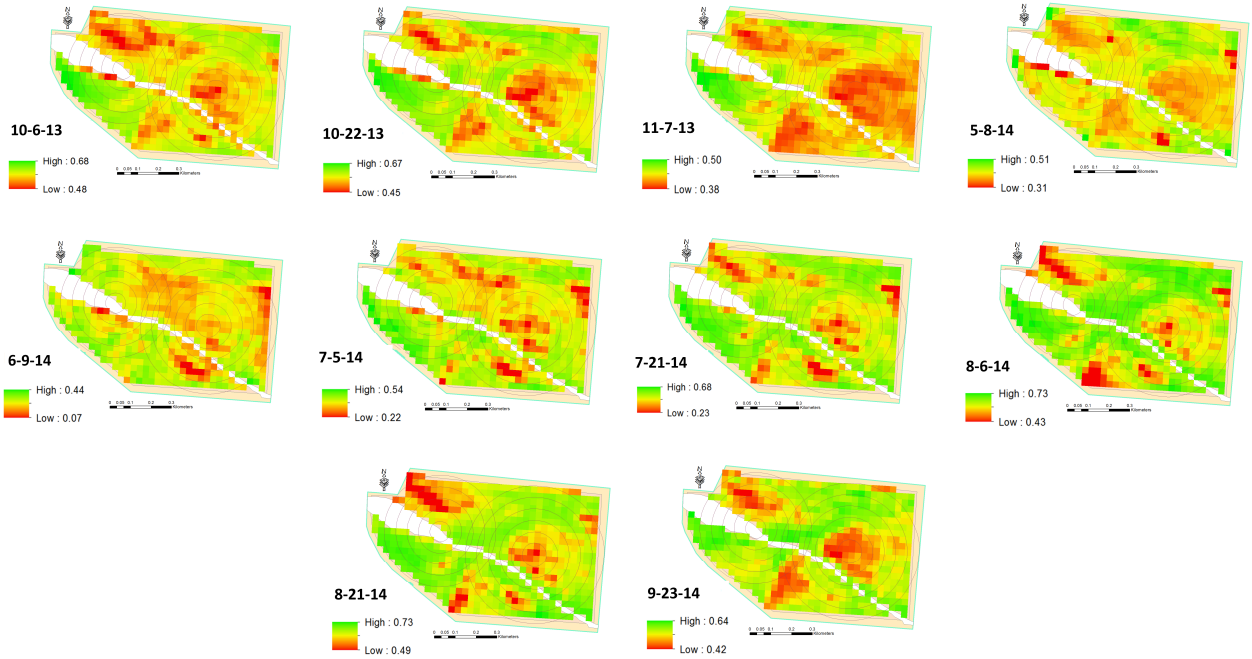

Figure 3-8. The October to November of 2013 and May to September of 2014 growing season's normalized difference vegetation index (NDVI).

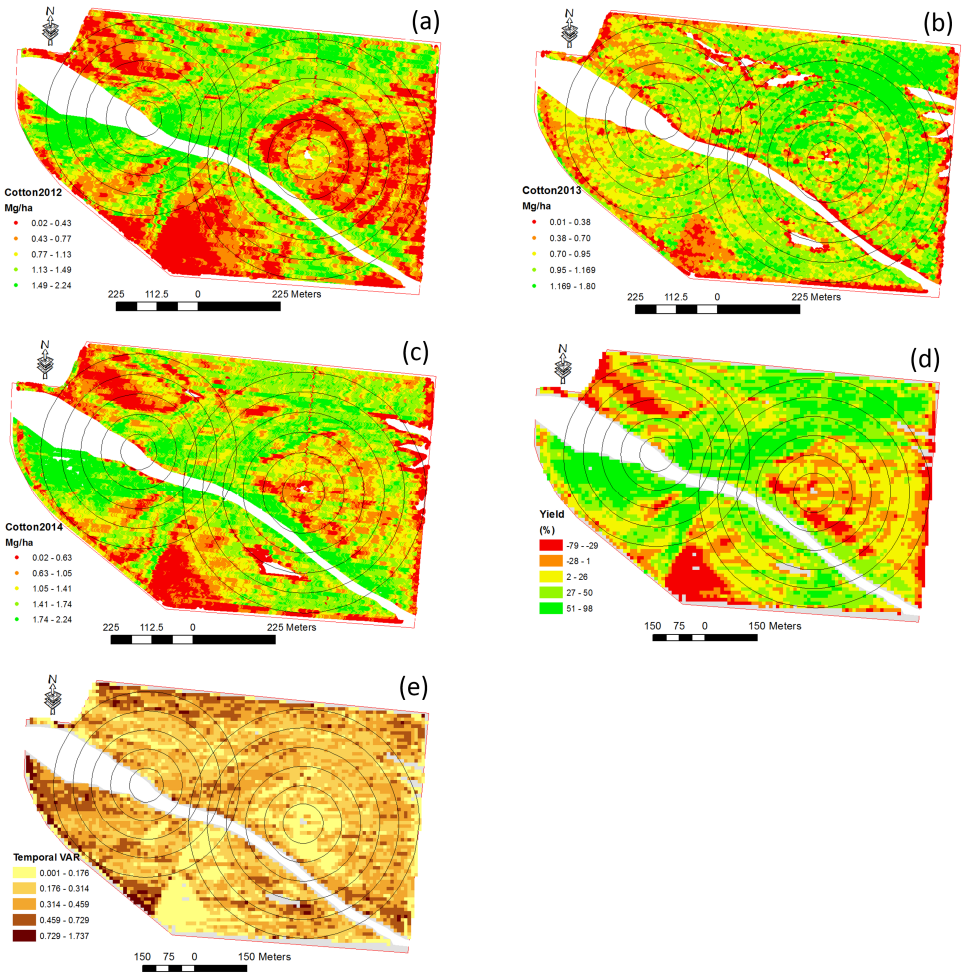

Figure 3-9. The cotton lint yield map time series from 2012 to 2014 (a - c) and the mean (d, using Eq. 6) and standard deviation yield maps (e, using Eq. 7).

Table 3-1. Detailed information on the irrigation programs for two center pivots in the study area for one revolution.

		East Pivot			West Pivot	
Program sector #	Start angle* (degree)	Stop angle (degree)	Depth of water (mm)	Start angle (degree)	Stop angle (degree)	Depth of water (mm)
1	90	110	10.41	275	315	9.91
2	110	0	15.49	315	335	11.68
3	0	20	20.57	335	355	8.38
4	20	40	10.41	355	235	9.91
5	40	70	15.49	235	255	11.68
6	70	90	20.57	255	275	8.38

* The zero degree was at north and pivot traveled clockwise.

Table 3-2. Growing season summary of weather and supplemental irrigation data in the study area in 2013 and 2014, in comparison to the 30-year mean for these variables.

Year	Variable	Month						
		May	June	July	August	September	October	November
2013	Rain, mm	23	150	190	95	79	112	63
	IW-East, mm			40	31	62		
	IW-West, mm			15	20	30		
	GDD, ^{o}C	10	300	308	360	292	123	12
	ET_P, mm day^{-1}			4.33	4.43	3.92	2.49	1.28
2014	Rain, mm	143	172	56	124	120	18	
	IW-East, mm			62	31			
	IW-West, mm			20	30			
	GDD, ^{o}C	170	289	254	308	175	58	
	ET_P, mm day^{-1}	4.15	4.42	4.86	4.51	3.47	2.94	
30 year	Rain, mm	120	101	102	74	82	82	117
	Tmean, ^{o}C	21	25	27	26	22	16	10

* GDD: growing degree days with 15.6 ^{o}C base, ET_P: potential evapotranspiration data that were calculated using Turc equation (equation 2) from 19 July 2013 (7 May 2014) to 30 November 2013 (5 October 2014), IW: irrigation water applied by the farmer. The 30-year mean data collected from the closest weather station (National Climate Data Center, 2015).

Table 3-3. Summary of satellite images collected for this study from Landsat 8.

**date of acquisition	DAP*	Sun Azimuth	Sun elevation
4/13/13		139.84	57.49
4/29/13		135.09	62.39
5/15/13		129.23	65.93
10/6/13	129	154.71	45.64
10/22/13	145	158.47	40.34
11/7/13	161	160.57	35.42
12/25/13		158.58	27.63
2/27/14		148.12	40.52
3/31/14		142.28	52.54
4/16/14		138.33	58.15
5/2/14		133.18	62.77
5/18/14	13	127.15	65.99
6/19/14	45	118.16	67.77
7/5/14	61	118.34	66.77
7/21/14	77	122.04	64.93
8/6/14	93	128.27	62.30
8/22/14	109	135.73	58.90
9/23/14	141	149.77	49.84

* DAP: days after planting, SD: standard deviation.

** For all images the path and row were 23 and 35, respectively.

Table 3-4. Correlation coefficient between cotton lint yield data (2013 and 2014 cropping seasons), soil properties at 4 depths, and fertilizer application.

	2013					2014				
Layer	1	2	3	4	total	1	2	3	4	total
*BD, g cm^{-3}	-0.16	-0.04	-0.09	-0.07		0.00	-0.23	-0.49	-0.18	
VWC, cm^3 cm^{-3}	0.22	0.08	0.03	0.12		0.47	0.51	0.46	0.51	
Sand, %	-0.12	-0.03	-0.03	-0.08		-0.44	-0.52	-0.50	-0.53	
Clay, %	0.16	0.03	-0.03	0.01		0.40	0.44	0.42	0.47	
Silt, %	0.07	0.03	0.06	0.10		0.42	0.53	0.51	0.53	
θ33	0.18	0.06	0.05	0.12		0.40	0.50	0.52	0.51	
θ1500	0.19	0.05	0.01	0.11		0.40	0.48	0.48	0.50	
ECa (shallow), mS m^{-1}					0.12					0.49
ECa (deep), mS m^{-1}					0.08					0.58
P, Mg ha^{-1}					-0.02					0.07
k, Mg ha^{-1}					-0.11					-0.23

* θ33 and θ1500: volumetric water content at matric potential -33 and -1500 kPa, respectively (predicted as explained in chapter 2); VWC: volumetric water content at the time of sampling. Layer 1, 2, 3 and 4 were from 0-25 cm, 25-50cm, 50-75 cm and 75-100cm, respectively. ECa shallow and deep readings represented approximately 0-30 and 0-90 cm of soil profile, respectively.

Table 3-5. The correlation coefficients between field-measured cotton lint yield maps and remotely sensed maps of crop indices.

Year	Date	*DAP	GNDVI	NDVI	SR
2013	6-Oct	129	0.30	0.30	0.26
	22-Oct	145	0.21	0.21	0.19
	7-Nov	161	0.03	0.02	0.01
2014	18-May	13	0.28	0.29	0.30
	19-Jun	45	0.05	0.11	0.09
	5-Jul	61	0.40	0.45	0.44
	21-Jul	77	0.63	0.67	0.68
	6-Aug	93	0.82	0.82	0.81
	22-Aug	109	0.74	0.74	0.73

* DAP: days after planting; NDVI, GNDVI and SR from Equations 3, 4 and 5, respectively.

Table 3-6. Descriptive statistics on yield data (Mg ha^{-1}) at the field of study.

Year	Crop	Mean	SD
2007	Corn	7.137	4.158
2008	Corn	3.420	0.903
2009	Soybean	3.221	0.860
2010	Cotton	0.947	0.306
2012	Cotton	0.913	0.494
2013	Cotton	0.871	0.329
2014	Cotton	1.244	0.493

Part 4: Toward site-specific irrigation management in west Tennessee

Abstract

Irrigation management has evolved into a top priority issue since demand for water is growing while available fresh water resources are limited. Collecting site-specific data has become feasible through precision farming equipment. Robust data mining approaches are needed in order to convert raw data to useful information which then could be used by farmers to enhance irrigation management and improve water use efficiency. This study consists of two parts aiming at delineating irrigation management zones and evaluating different uniform and variable rate cotton irrigation strategies. In the first part different clustering methods, including *k*-means, ISODATA and Gaussian Mixture, were selected to delineate irrigation management zones. In addition, a new zoning method, based on integer linear programming, was designed and evaluated for center pivot irrigation systems with limited speed control capability (i.e. pie shape zoning). The soil available water content (from second chapter) was used as the main attribute for zoning while soil apparent electrical conductivity (ECa), space-borne satellite images and yield data were required as ancillary data. A good agreement was observed among delineated zones by different clustering methods. The new zoning method based on integer linear programming explained up to 40 % of available water content variance underneath center pivot irrigation systems. The ECa achieved the highest Kappa coefficient (=0.79) among ancillary attributes, hence exhibited a considerable potential for irrigation zoning. Within the second part of this study, some novel water production functions (WPFs) were established and evaluated. WPFs (i.e. mathematical relationships between applied water and crop yield) are useful tools for irrigation management and economic analysis of yield reduction due to deficit irrigation. A two-year cotton irrigation experiment (2013-2014) was implemented and *k* nearest neighbors (*k*-NN), multiple linear regression, and neural networks were selected to derive site-specific WPFs. The site-specific *k*-NN WPFs showed the highest performance with root mean square error equal to 0.131 Mg ha^{-1} and 0.194 Mg ha^{-1} in 2013 and 2014 cropping seasons, respectively. The result showed that variable rate irrigation with pie shape zones could enhance cotton lint yield under supplemental irrigation when field-scale spatial soil variation is significant. The temporal changes in climate and rainfall patterns, however, had a great impact on cotton response to irrigation in west Tennessee, a moderately humid region with short season environment.

Chapter 4: Perspectives on delineating management zones for variable rate irrigation

(Submitted to *Computers and Electronics in Agriculture* for publication)

Amir Haghverdi[1*], Brian G. Leib[1], Robert A. Washington-Allen[2], Paul D. Ayers[1], Michael J. Buschermohle[1]

[1] Department of Biosystems Engineering & Soil Science, University of Tennessee, 2506 E.J. Chapman Drive, Knoxville, TN 37996-4531

[2] Department of Geography, University of Tennessee, Burchfiel Geography Building, Knoxville, TN 37996-0925

* Correspondence author, ahaghver@vols.utk.edu

Abstract

This study aimed at investigating the performance of multiple irrigation zoning scenarios on a 73 ha irrigated field located in west Tennessee along the Mississippi river. Different clustering methods, including *k*-means, ISODATA and Gaussian Mixture, were selected. In addition, a new zoning method, based on integer linear programming, was designed and evaluated for center pivot irrigation systems with limited speed control capability. The soil available water content was used as the main attribute for zoning while soil apparent electrical conductivity (ECa), space-borne satellite images and yield data were required as ancillary data. A good agreement was observed among delineated zones by different clustering methods. The new zoning method explained up to 40 % of available water content variance underneath center pivot irrigation systems. The ECa achieved the highest Kappa coefficient (=0.79) among ancillary attributes, hence exhibited a considerable potential for irrigation zoning.

Keywords: apparent electrical conductivity, integer linear programming, remote sensing, soil water retention, unsupervised clustering.

1. Introduction

1.1 Precision farming and management zone delineation

In conventional agriculture each field is considered as a uniform unit, by purposely ignoring the heterogeneity across the field, thereby decision-making is based on an estimation of average conditions. The motivation for site-specific farming was first addressed by researchers during the late 80s and early 90s (Arslan and Colvin, 2002). As such, precision agriculture (PA) methodology is a way to look at field management by taking the within field variation into account and incorporating that variability into management decisions. Within-field heterogeneity is caused by both temporal and spatial variation of a variety of factors such as climate, topography and biologic activity (Córdoba et al., 2013).

A management zone (MZ) is a sub-region of a field that is relatively homogeneous with respect to soil-landscape attributes. It is expected that variable rate application across MZs will help by saving the resources and optimizing yield (Schepers et al., 2004). Protecting the environment and keeping agriculture sustainable may also be achievable through precision farming. Sensor-based and map-based approaches are two major methods to practice variable-rate application. In the sensor-based method, a real time decision on application rate is made using data collected via sensors and a pre-developed application algorithms. In the map-based method, application maps are prepared using site-specific information such as yield data and soil data prior to implementation. It is critical to select appropriate attribute(s) and method to delineate robust zones (Thöle et al., 2013).

A field can be zoned based on a single soil-crop variable or multiple attributes which are expected to affect yield (Khosla et al., 2010). Yield maps, topography, satellite photographs, canopy images and soil apparent electrical conductivity (ECa) are among suggested attributes to delineate MZs. Application of remote sensing is especially attractive because it is noninvasive and relatively inexpensive (Schepers et al., 2004). Yield maps are useful sources of information reflecting within-field variation. However, some difficulties have been reported to delineate zones solely by yield maps (Khosla et al., 2010). Temporal inconsistency among yield maps from year to year is probably the main reason causing this problem. Schepers et al. (2004) reported that temporal climate variability in an irrigated cornfield significantly affects yield spatial variability from year to year. Combining yield data with other ancillary information or averaging yield data over years can help explain spatial variation better and in turn can provide more trustable zones. Promising results have been reported by the studies that have utilized

several years of yield data to create MZs. However, stability of such zones has to be tested for each individual field (Khosla et al., 2010). Calculating temporal variance can be helpful to verify zone stability (Basso et al., 2012).

There are some methods to screen potential attributes regarding their importance. Hornung et al. (2006) suggested assigning different weights to individual layers based on their importance to the variation in crop yield. Principal component analysis (PCA) can be utilized to linearly transfer original variables to new independent ones. Researchers apply this technique to understand the characteristics of a data set along with relative importance of each individual variable (Fraisse et al., 2001). Finally, the most important principal components will be used to delineate MZs.

There are several methods to delineate MZs. Applying unsupervised clustering techniques and zoning via user-defined thresholds are the main procedures. Clustering techniques group similar data points (cells), based on their inherent structure, into distinct classes. Methods such as *k*-means and fuzzy *k*-means have been widely used to identify MZs (Córdoba et al., 2013). These methods usually produce separated oval shape zones across a field. The software MZA was developed by Fridgen et al. (2004) for delineating MZs. The MZA uses fuzzy *k*-means for zoning and tests the result to evaluate how many zones to create in a given field. More recently, Zhang et al. (2010) developed a web-based decision support system for zoning using satellite imagery and field data. Moral et al. (2010) used regression kriging, PCA and fuzzy cluster classification to delineate MZs using soil texture information and electrical conductivity as ancillary data. Recently Cid-Garcia et al. (2013) proposed an integer linear programming MZ delineation method to make rectangular shaped zones which facilitates the work and operation of machinery.

1.2 Irrigation management zones delineation

Water has become the most valuable input for agriculture across the world. There has been a significant conversion from rainfed to irrigated production in humid regions as a safe guard against unpredictable severe drought periods within a cropping season which can cause yield reduction in such regions. At the same time, agriculture is under a great pressure to enhance its water use efficiency as other sectors such as industry and urban areas are demanding more water (Evans et al., 2013; Daccache et al., 2015). When within field soil spatial variation is significant, variable rate irrigation (VRI) comes into play as an engineering solution to manage spatial allocation of applied water through irrigation. Evidence exists in literature showing VRI could enhance water use efficiency and/or yield.

Hedley et al. (2009) mentioned up to 26 % annual water saving by VRI comparing to conventional uniform irrigation. King et al. (2006) reported 4% higher potato tuber yield (statistically not significant) under VRI in comparison to uniform irrigation.

Precision irrigation center pivots have been commercially available for a while yet their adoption by farmers has been slow to develop. The prime drawbacks of VRI systems are high capital costs of implementation and management which is hard to rationalize considering energy and water saving at current price (Evans et al., 2013). Most of the available center pivots, however, have control panels with speed control module, meaning farmers can vary the irrigation across their fields to some extent at no extra cost, if it is needed. Changing the pivot travel speed enables putting some pie shape zones where speeding the system up provides less time to irrigation and slowing the system down increases the quantity of applied irrigation at the selected pies.

There is a critical need to dynamically develop irrigation MZs in an accurate and inexpensive manner (Evans et al., 2013). Delineating zones for irrigation management is challenging. Soil physical and hydraulic properties govern plant available water and thereby directly affect irrigation scheduling. Unfortunately hydraulic properties of soil are not readily available at a field-scale to be used for delineating irrigation MZs. There is evidence in the literature that easily collected ancillary data may be spatially correlated to soil physical and hydraulic properties, thus be useful for irrigation zone delineation. In irrigation scheduling, plant available water is calculated within the effective root zone which expands to deeper soil layers as the cropping season progresses. If in depth soil variation is significant at the field-scale, theoretically, the spatial arrangement of irrigation MZs may change during the growing season. This is not, however, in agreement with the usual practice of static MZs sometimes for years. Given the limitation in number, size and location of speed control pie shape zones, an optimization procedure is needed to find the best combination of pies across a field.

The objectives of this study were: (i) to evaluate the performance of some zoning algorithms on an irrigated field with significant spatial soil variation located in west Tennessee (ii), to investigate the usefulness of proximal data, including ECa and space-born satellite images, to delineate irrigation MZs, (iii) to analyze stability of irrigation MZs within growing season in respect to soil available water content, and (iv) to develop and evaluate an irrigation zoning method for center pivot systems with speed control capability using integer linear programing technique.

2. Material and Methods

2.1 Field of study

An irrigated agriculture field, approximately 73 ha located in west Tennessee along the Mississippi river, was selected (Figure 4-1). The field was located in a short season, semi-humid region with high rainfall potential during a cropping season but some dry periods when supplemental irrigation is used to fulfill the crop water requirement. There were two center pivot systems for irrigation covering majority of the field with some overlap in their coverage. Cotton was planted in 2012, 2013 and 2014 cropping seasons in that field. The spatial variation in soil physical and hydraulic properties was significant (Table4-1 and Figure 4-2) which made this field a perfect case for the purpose of our study.

A total of 400 undisturbed samples were collected from 100 sampling locations at 4 different depths, 0-25 cm, 25-50 cm, 50-75 cm and 75-100 cm. The soil texture components as well as soil bulk density and volumetric water content were measured in lab. More details on sampling procedure and lab analysis is provided in chapter 1. The soil water retention curve for samples and high resolution soil available water content (AWC) maps were predicted by using pseudo continuous pedotransfer function (Haghverdi et al, 2012, 2014) and co-kriging (chapter 2). There was a decline in soil water holding capacity (i.e. water content at matric potential -33 kPa, θ_{33}) by depth (Table 4-1) which is in agreement with the expansion of the sandy regions across the field moving from surface to deeper layers.

2.2 Proximal data

It is appealing to delineate MZs by proximal data because they are low in cost, easy to collect and usually noninvasive. In this study soil ECa and space-borne satellite photos were considered as proximal information for zone delineation. The soil ECa measurements from shallow (i.e. ECS, 0-30 cm) and deep (i.e. ECD, 0-90 cm) layers were collected using a Veris 3100 (Veris Technologies, Salina, KS) instrument on March 20, 2014 (chapter 1). The ECS data showed a normal distribution but ECD information were log transferred to have a roughly Gaussian distribution. Then, the ECS and ECD data were interpolated using kriging to cover the entire field (with 2 m spatial resolution) with the exception of a drainage pathway starting from west of the field and ending to the southeastern corner of the field.

The satellite images were obtained from Landsat 8 Operational Land Imager (OLI) sensor during 2013 and 2014 (through http://earthexplorer.usgs.gov/). A total of 19 cloud-free images were selected for statistical analysis.

The selected images contain the reflectance at multiple bands (i.e. quantized and calibrated scaled Digital Numbers (DN)). The reflectance of a canopy and bare soil will change during a cropping season as a plant goes through different growth stages and soil water status is changed. The band 8, panchromatic band, with wavelengths ranged from 0.50- 0.68 μm for Landsat 8 has the highest spatial resolution (i.e. 15 m ×15 m spatial resolution) within available bands, thus was selected for this study. The field of study contained a total of 2320 cells.

The cotton yield data were available from 2010 to 2012 cropping seasons. First, the raw yield data were cleaned. More information on cleaning process could be find in chapter 1. Then, kriging was required to interpolate yield data (with 2 m spatial resolution) across the field with the exception of the drainage path. Finally, the yield data were standardized for each year and averaged across years to make a single yield map.

2.3 Management zone delineation

Multiple zoning strategies were required and evaluated through two phases (Table 4-2). In phase one, the entire field was zoned. The ECa and reflectance data (i.e. DN) were tuned to have a range between 0 and 1 by dividing each cell value by the overall maximum values across the field from all data points/cells. The ECa data were chosen as the prime attribute for zoning because they showed a good correlation with soil physical and hydraulic properties of the soil within the field of study (chapters 1 and 2). Three unsupervised clustering techniques, i.e. *k*-means, ISODATA-maximum likelihood and Gaussian mixtures, with different combination of ECa data were examined. Unsupervised clustering techniques group the data based on their inherent structure as opposed to supervised clustering methods which require a priori knowledge on data for training. More information on each clustering technique is provided later in this paper. Given the performance of the zoning strategies and distribution of the digital reflectance data, only *k*-means was required for zoning using satellite images.

The within-season temporal stability of MZs was also investigated in phase one through a user-defined zoning strategy considering AWC as input. The effective root zone (0-100 cm) was defined as a depth in which crop absorbs most of its water requirement. To mimic the dynamic of effective root growth, four crop available water maps were created for 0-25, 0-50, 0-75 and 0-100 cm. Each map, then, was divided into user-defined MZs and the spatial arrangement of zones were compared against each other.

In phase two, we focused on the area underneath the center pivot irrigation systems for which we developed and evaluated a new zoning approach using integer linear programing (ILP). The yield data were also included to

test how efficient this variable was in the zone delineation process. Most center pivot irrigators prefer to have less than 10 MZs in a field (Evans et al., 2013). Moreover, a majority of available center pivots at west Tennessee cannot put more than 10 pies underneath a center pivot system. Therefore, clustering procedures were repeated to delineate from 2 to up to 10 MZs.

Two software products were used for clustering: Matlab R2015a (MathWorks, Inc., Natick, Mass.) and ArcGIS 10.2.2 (ESRI Inc., Redlands, California).

2.3.1 K-means

The *k*-means is one of the most widely used methods for clustering due to its simplicity and efficiency. The *k*-means, in an iterative process, tries to partition data into *k* groups so that the differences among the features in a group, over all groups, is minimized. It moves the observations between clusters and monitors the sum of distances from each observation to the center of its cluster until group membership stabilizes. The *k*-means++ algorithm was used to assign initial centers to the clusters. This algorithm chooses the cluster centers in a random manner yet favors spreading them out by a distance-based weighting process (Jain, 2010; Jain et al., 1999; ESRI, 2014). The Euclidean metric was used to obtain distances. No spatial constraint was considered.

2.3.2 ISODATA - Maximum Likelihood

A well-known extension of *k*-means is called Iterative Self-Organizing Data Analysis Technique (ISODATA). The ISODATA starts with randomly assigning data into different classes. Then, over the course of an iterative process, the algorithm changes the membership of data points and tries to find the optimum clusters where the Euclidian distance between data points to cluster center is minimized. ISODATA applies some heuristics to adjust the number of clusters; clusters could be removed, divided or merged at the end of each iteration based on similarities between neighboring clusters and specified minimum class size (Jain, 2010; Jain et al., 1999; ESRI, 2014). In this study, functionality of ISODATA was combined with Maximum Likelihood which assigns cells to classes using Bayes' theorem. This method assumes that the distribution of a class sample is normal, hence uses mean vector and the covariance matrix to assign each cell to a class based on statistical probability.

2.3.3 Gaussian Mixture Model

A Gaussian Mixture (GM) model is a parametric probability density function which is often used for soft data clustering. If clusters have different sizes and correlation within them, GM may be more appropriate option for clustering than *k*-means. This method assumes data belong to a mixture of normal density components with unknown parameters. The GM parameters are estimated using an expectation maximization algorithm by assigning posterior probabilities to each component density. Each cluster corresponds to one of the Gaussian components, hence data are assigned to clusters such that posterior probability is maximized (Reynolds, 2009; Mathworks, 2014).

2.3.4 Integer linear programing

A linear programming (LP) problem is an optimization (maximization or minimization) problem of a linear objective function, subject to linear equality and/or inequality constraints. If some or all of the variables are limited to be integers, it is called an integer linear programing (ILP) problem. The LP has been a very popular technique to irrigation management and water resources system analysis and planning (Singh, 2012). Recently, the ILP has been successfully used for rectangular shape MZ delineation and crop planning optimization problems by Cid-Garcia et al. (2013; 2014).

The area underneath each center pivot irrigation system was divided into 360 pies each 1 degree wide. Then, a spatial join process was implemented to identify data points (cells) within each pie for each attribute. A total of 129,241 zones was created by combining 1 degree pies (Table 4-3). The standard deviation was calculated for data points within each zone with respect to the target attribute.

An integer linear programming model was designed to find the optimum spatial arrangement of pie-shape zones underneath each pivot. The general mathematical model was as follows:

$$\min \sum_{i=1}^{n} \sum_{k=1}^{m} \left(l \times a_k \times \sigma_{ik} \times x_i \right) \quad (1)$$

subject to

$$\sum c_{ij} x_i = 1 \quad (2)$$

$$\sum_{i=1}^{n} x_i \le P_{\max} \quad (3)$$

$$lx \geq L_{\min} \tag{4}$$

$$x_i \in \{0,1\}$$

where x_i is the decision variable for zone i, σ is the standard deviation for attribute k within zone i, a_k is a user-defined weighting factor based on the relative importance of an attribute in zoning process, m is the total number of attributes, n is the total number of the zones, l is the length of each zone in degrees, c_{ij} is a coefficient equal to 1 if angle j is covered by zone i otherwise equal to 0, $P_{\max}$ is the maximum number of desired zones.

Objective function (1) minimizes a weighted average of standard deviations across the zones. In this study we did not use more than one variable at each optimization attempt, thus k and a_k were set to one. Restriction (2) ensures that each degree is covered by only one zone and selected zones cover the entire area underneath the irrigation system. Constraint (3) ensures that total number of the pies is less than the maximum desired number which was set to 10 in this study. Constraint (4) ensures that the length of the pies is greater than the minimum desired value which was chosen to be 5 degree. For a pie, the decision variable of the model, x_i, was 1 if the pie was selected and 0 otherwise.

2.4 Performance Evaluation

The overall within-zone variance of AWC was calculated to assess zoning strategies and find the optimum number of zones. The AWC variances for each zone were weighted considering the zone-area (Fraisse et al., 2001):

$$S_Z^2 = \frac{1}{n_Z} \sum_{i=1}^{n_Z} \left(AWC_i - m\right)^2 \times \frac{n_Z}{n_T} \tag{5}$$

where S^2_z is weighted variance for zone z, AWC is the soil available water content for cell i, m is mean of soil available water values in zone z, n_Z and n_T are number of cells in zone z and total number of cells across field, respectively.

Total within-zone AWC variance across field (S_T^2) was obtained as the sum of weighted within-zone AWC variances:

$$S_T^2 = S_1^2 + S_2^2 + \ldots + S_Z^2 \tag{6}$$

The better zoning scenario was selected as the one with minimum total AWC standard deviation. For the ILP method, the kappa coefficient (Cohen, 1960) was also calculated to measure inter-classification agreement of

different zoning strategies by ancillary data against optimum zones using AWC data. The kappa coefficient equals 0 and 1 when the agreement is due to chance alone or when there is perfect agreement, respectively. A positive coefficient indicates agreement exceeds chance and the magnitude of coefficient reflects the strength of agreement.

3. Results

3.1 Zoning using ECa and satellite images

Figure 4-3 depicts the variance of AWC as well as average yield within zones for a different number of MZs and multiple clustering methods. The AWC variance across the field, not divided into zones, was chosen as a reference (i.e. one MZ with 100 % variance). For *k*-means, up to 11 %, 19 % and 16 % of variance was explained considering ECS, ECD and ECS plus ECD as input, respectively. Considering more than 4 zones only reduced the variance by an additional 1% (ECS input), 1% (ECD input) and 2% (ECS+ECD inputs). For ISODATA-ML, variance was reduced by 11%, 19% and 17 % considering ECS, ECD and ECS plus ECD as input, respectively. Given more than 4 zones, the variance decreased by an additional 1% (ECS input), 1% (ECD input) and 4% (ECS+ECD inputs). For GM method, 10% (ECS input), 17% (ECD input) and 16% (ECS+ECD inputs) of the variance reduction occurred by dividing the field into 4 zones while further division to up to 10 zones only decreased the total variance by an additional 1% (ECS input), 2% (ECD input) and 5% (ECS+ECD inputs).

Adding more than 4-5 zones only slightly improved the clustering results. In fact, most of the AWC variance was explained by dividing the field into two MZs. This is in line with changes observed in within-zone yield average; when the number of zones increased to more than 4-5, one starts to see more zones with similar productivity. The clustering methods performed similar. The highest reduction in variance was observed for GM with 10 zones and ECS plus ECD as inputs. However, considering 4 as the optimum number of zones, *k*-mean and ISODAT-ML performed slightly better than GM. Noting this similarity among results of clustering methods and given the simplicity and ease of use of *k*-means method, it was decided to only use it for the rest of the unsupervised clustering tasks during this study. Aggregating ECS and ECD did not decrease the within-zone variance of AWC for *k*-means and ISODATA-ML procedures. For an optimum number of zones (n= 4), the lowest variance was observed when ECD was used as an input. Therefore, the ECD was considered an attribute of high importance for explaining the variability found in AWC for the field of study.

Figure 4-4 illustrates the gray-scale panchromatic images taken by Landsat 8 during 2013 and 2014. Initial assessment of satellite images revealed the reflectance value of cells adjacent to the boundaries of the field were affected by the roads (located at north and west parts of the field) and trees (located at south and east parts of the field) and in turn influenced the spatial distribution of the entire field. Therefore, the border cells were filtered out from satellite images prior to running clustering analysis. The cells within drainage pathways were also removed from satellite images since they were missing from other attributes too. Moreover, they caused the same problem as border cells. The field was planted (harvested) in cotton on May 30, 2013 (December 2-3, 2013) and on May 5, 2014 (October 18-20, 2014), meaning some of the satellite images had been taken from bare soil but some from the cotton canopy in different growth stages. There were moderate temporal changes in spatial patterns of satellite photos. However, it was possible to visually associate brightness level to AWC spatial variation in a good number of images (i.e. photos taken on April 13, 2013; April 29, 2013; October 22, 2013; August 6, 2014; August 22, 2014; and September 23, 2014).

The coarse-textured soils with low AWC were mostly located over three spots across the field: close to the southern border of the field mostly outside of the outer pivot circles, (ii) underneath the eastern pivot approximately from the pivot point up to the right edge of the field, and (iii) a region in the northwestern part of the field above the drainage path adjacent to the northern border (chapters 1 and 2). In general, very bright cells (i.e. cells with high reflectance value/ digital number) were mostly associated with coarse-textured soils with low AWC values while darker-colored cells (i.e. cells with low reflectance value/digital number) were associated with higher AWC soils. The photo acquired in August 6, 2014 revealed considerable spatial agreement with our understanding of AWC patterns. Factors such as ponding water due to rainfall (i.e. photos taken on May 15, 2013; December 25, 2013; February 7, 2014; March 31, 2014; April 16, 2014, May 2, 2014; May 18, 2014; October 25, 2014 and December 12, 2014) and wet soil due to an irrigation event (i.e. photo taken on July 21, 2014) affected the spatial arrangement of satellite photos.

To understand temporal changes, the distribution of standardized brightness values for the satellite photos are illustrated in Figure 4-5 where similar distributions were roughly grouped by visual assessment. Coarse textured soils appeared as a long right tail in satellite photos grouped in panel A and D, while spots with ponding water formed a long left tail for images in panel C and caused negative skewness. The satellite images with less clear spatial pattern formed panel B. Images with no vivid spatial pattern showed high positive Kurtosis (panel E). The

image taken April 16, 2014 did not follow the distributions observed with the other images; with negative kurtosis it depicted both ponding water and coarse soils but did not well represent the entire spatial arrangement of AWC across the field. The better satellite images seemed to follow a log normal distribution (panel D) which was the distribution of raw ECD data as well.

Figure 4-6 shows the clustering results in terms of average within-zone AWC variance using satellite images as input. In 2014, images taken August 6 and September 23 showed the highest performance with 21 % and 14 % reduction in AWC variance, respectively. In 2013, images taken November 7 and April 29 exhibited the better result by explaining 12% and 10% of AWC variance, respectively. Identical to what we observed for ECa data, adding more than 4-5 zones only slightly improved explaining AWC variance. On average, the abovementioned satellite images with acceptable performance, considering more than 4 zones only helped to explain an additional 2% of AWC variance. The zones delineated from rest of the images explained less than 10 % of the AWC variance. The lowest performance belonged to images taken 19 July and 12 December 2014 which explained only 1 % of the AWC variance. In 2013, images taken on 22 October and 25 December exhibited the lowest efficiency by explaining 4 % and 3 % of AWC variance, respectively. With a few exceptions, our visual assessment of reflectance distribution (Figure 4-4) turned out to match the efficiency of images in zone delineation (Figure 4-5). Overall, the highest performance was observed for images in groups D and A with an average 11% and 10% reduction in variance, respectively. In contrast, images in groups E and C showed the lowest potential for clustering by only explaining 3 and 4 percent of AWC variance, respectively. The higher number of cloud free images available contributed in better zoning results in 2014 compare with 2013. We also attribute this to difference in field conditions; in 2013 planting was delayed due to cold weather and wet soil which affected cotton growth and maturity and also diminished the yield contrast across the field which is usually expected to occur between soils with high and low water holding capacities.

Figure 4-7 shows the optimum MZs (n= 4) using different clustering methods. Figure 4-8 depicts the average standardized yield for optimum number of zones (n= 4) for *k*-means approach. There was a good similarity among MZs delineated by different algorithms which also matched the spatial arrangement of soil physical and hydraulic properties (chapters 1 and 2). The delineated zones by *k*-means and ISODATA-ML were similar which was expected due to similarity between these clustering methods. The dark blue (zone 4) and yellow (zone 1) colors represent soils with low and high AWC, respectively and the other two zones illustrates the transient between these

extremes. The trend in average yield data across MZs matched the spatial arrangement of soil physical and hydraulic properties. For instance, higher yield belonged to soils with higher AWC and somewhat to the strip surrounded the drainage path which had good drainage condition hence did not suffer from excessive water content after heavy rainfall events which is likely to happen in west Tennessee. The yield on coarse-textured soil was 20 % below average while soils with higher AWC produced 10 % above average yield. In practice, small spots within clusters could easily be removed (for example by using a moving average window) to have more homogenous zones for variable rate applications.

3.2 Spatiotemporal changes in soil available water content

Temporal changes in spatial arrangement of user-defined irrigation MZs is shown in Figure 4-9. The selection of breaking levels between zones was arbitrary. We chose 2 cm water within effective root zone as the target range. The highest spatial change took place from first thematic map (0-25 cm) to second one (0-50 cm). After that, there was a good agreement among thematic maps up to 100 cm. This trend suggests the possibility of having different irrigation zones early in the cropping season when the crop is only using the available water from surface soil.

3.3 Management zone delineation, phase two

Figure 4-10 depicts the performance of ILP and *k*-means methods to delineate MZs underneath center pivots using different input attributes. The AWC variance for each pivot was considered as reference (100% level) which was equal to one MZ or uniform irrigation. Given AWC as input and *k*-means as clustering method, for both pivots within zone uniformity increased by adding more zones yet changes were minor as the number of zones became more than 4-5. The same trend was observed when ancillary data (i.e. ECa, reflectance value, and yield) were used for clustering but was less pronounced. This is in line with the observed trend when the entire field was clustered to MZs (i.e. phase one). Most of the reduction in AWC variance was obtained by dividing the pivot into two MZs. Given ECD and reflectance values as input, up to 30 % and 15 % of AWC variance was explained underneath the east and west pivots, respectively. Considering average yield as an input for clustering only helped to explain less than 10 % of AWC variance. Given AWC as input and ILP as the zoning method, up to 30 and 40 % of AWC variance was explained at the east and west pivots, respectively. In reality, this is the improvement expected to occur from uniform irrigation to VRI with limited speed control. Considering ILP as zoning method, both ECD and

reflectance values were able to efficiently zone both pivots. The average yield, however, performed weaker than other inputs for clustering at the east pivot but was as effective for the west pivot. The difference between ILP and *k*-means for optimum number of zones (n=4) was about 35 % for both pivots in respect to reduction in AWC variance. This represents difference expected between VRI with limited speed control and a hypothetical VRI with ability to irrigate each 4 m^2 cell individually. The ability of a real precision irrigation system in reducing AWC variance is expected to be somewhere in between.

Figure 4-11 represents the optimum irrigation MZs (n =4) for both pivots considering different inputs and clustering methods. The user-defined MZs using AWC data were considered as the reference to evaluate the performance of other zoning strategies using *k*-means. The same was true for zoning by ILP model when zones delineated using AWC were considered to show the optimum arrangement of pies and other zoning configurations were compared against it. For ILP with AWC as input, (i) zones 1-4 and zones 3-4 covered mostly coarse soils with low AWC underneath east and west pivots, respectively; (ii) zones 3 and 2 underneath east and west pivots, respectively, mostly consisted of fine-textured soils with high AWC; (iii) other zones (zone 2 and 1 underneath east and west pivots, respectively) contained a mixture of soils with high and low AWC.

The spatial arrangement of optimum zones provided by ILP techniques was not identical among models with different ancillary data as input (Figure 4-10). By visual assessment, there is a good and moderate agreement between pie zones using ECD and Band 8, respectively and that of AWC but using yield for zoning resulted somewhat different zones. The same trend occurred for the west pivot. The Kappa coefficient (Figure 4-12) confirmed our visual assessment. In the case of the east pivot ECD and yield data showed the highest and lowest kappa coefficient, respectively. The same trend was observed in the case of west pivot, but there was a consistent reduction in Kappa coefficient for all attributes in comparison with that of east pivot. It is clear that ECD and satellite images were good candidates for irrigation zone delineation as opposed to yield which exhibited less potential. This difference could be related to the spatial resolution of different soil attributes. The ECa data were the most exhaustive data, so carried information on soil spatial variation at a very fine scale. The satellite images were not point measurements but an integration of reflectance over 15 by 15 m cells. The AWC data were interpolated from soil cores.

4. Discussion

Recent research in precision agriculture has greatly focused on variable rate application and MZ as a method to enhance the efficiency of crop inputs usage. However, the phrase "management zones" remains uncertain unless additional information is included to clarify the goal in sub-dividing the field (Kitchen et al., 2005) which affects our choice of clustering technique and information used for zoning. The MZ delineation for variable rate application of fertilizer and seeding has extensively been studied in recent years. In contrast, there are only a handful of studies on zoning for variable rate irrigation. Given the ever increasing demand for fresh water, moving toward precise application of water for irrigation is unavoidable. The resulting zones should be simple, stable, accurate and inexpensive to identify, and enable within-field spatial variation to be managed (Khosla et al., 2010).

4.1 Management zones delineation methods

Guastaferro et al. (2010) compared different algorithms for the delineation of management zones and mentioned different pros and cons for each method. We observed a good agreement among delineated zones using *k*-means, Gaussian mixture and ISO-ML methods. It seems for a field with high spatial structure, all well-known clustering methods are able to recognize an accurate pattern for irrigation MZs in a substantial way. The ILP zoning strategy, introduced in this study, also showed promising results. This method was designed to delineate pie shape zones for center pivots with limited speed control ability. There are several benefits associated with this zoning method. First, most of the available center pivots located in west Tennessee and other irrigated areas across US have suitable control panels to create pie shape zones, so there is no need to upgrade available pivots in order to make use of this method. Second, applying ILP showed a significant decline in within-zone AWC variance, meaning it will help make better site-specific irrigation decisions, and improve the choice of timing and amount of irrigation, which in turn enhances WUE. In addition, it could promote VRI and act as a bridge to adopt more advanced precision irrigation systems if needed.

4.2 Optimum number of zones

It is extremely important to find the optimum number of irrigation MZs. We found 4-5 MZs to be sufficient to explain within field variability of AWC. By changing the number of zones from 1 to 10, the total within-zone variance drastically decreased but approached an asymptotic value slowly as the number of zones continued to

increase. A similar trend for variance reduction was found by Zhang et al. (2010). Fraisse et al. (2001) delineated 2 to 6 MZs for two fields located near Centralia, Missouri, and monitored the within zone variance of yield to identify the optimum number of zones. They observed a minimum variance for 5 zones when up to 32% of the yield variance was explained.

Delineating irrigation zones is the first step toward site-specific irrigation management. Then, an irrigation prescription map needs to be developed before each irrigation event throughout the cropping season. One may only use zones as a guide to adjust the timing of supplemental irrigation in west Tennessee, i.e. to delay starting irrigation or terminate irrigation earlier in a specific zone(s). A more sophisticated scenario is to calculate the irrigation for each zone independently by sensing soil/crop water usage/status. Therefore, over estimating the optimum number of MZs would make irrigation scheduling more complicated, time-consuming and costly. On the other hand, under estimation of management zones diminishes the irrigating efficiency and may cause yield reduction. There is less freedom to explain variability with pie shape zones. It is most likely to have some zones consisting a mixture of soils with low and high AWCs. This is an unavoidable problem which is due to the inherent limitation of center pivots with limited speed control ability. The critical question is how to manage this variation or which soil to follow during irrigation. Answering this question needs quantitative information on crop-soil-water interaction which was outside of the scope of this study but is suggested for further investigation.

4.3 Application of proximal data

The AWC is a logical basis to delineate irrigation MZs since it directly affects both plant growth and yield, and irrigation scheduling. According to Fraisse et al. (2001) crop production potential is strongly related to plant available water so its productivity can be approximately determined on the basis of soil physical properties and topographic characteristics, when topography changes significantly in the field. In the field of study more than 50 % of cotton and soybean yield variability was explained by soil physical and hydraulic properties (chapters 1 and 2). In practice, however, more easily obtained data should act as a surrogate for AWC information that is more difficult, time-consuming or expensive to obtain.

The ECa data exhibited a considerable potential for irrigation zoning. Moral et al. (2010) also found ECa to be a good proximal attribute for high resolution zone delineation. They reported ECD and percent of clay to be dominant factors explaining soil spatial variability based on PCA. Kitchen et al. (2005) found ECa a reliable

attribute for allocating productivity zones on claypan soil fields. Though less pronounced than ECa, the satellite images exhibited a good potential to be used for characterizing spatial variation in field-scale AWC. Our findings support the result reported by Guo et al. (2012) who observed a strong relationship between bare soil brightness and ECa with soil texture where low ECa and high brightness were associated with lower clay content. In another study, Song et al. (2009) used Quickbird commercial satellite data with 2.4 m resolution for MZ delineation and found it to be a reliable procedure. The main drawback of proximal sensor measurements is that they are complex and affected by multiple soil-crop properties (Corwin and Lesch, 2010). Therefore, they should show similar spatial distributions to those of soil physical and hydraulic characteristics, governing soil available water for crop, in order to be considered as an effective input to delineate irrigation zones. In addition, timing of capturing photos by satellite turned out to affect the practical application of satellite images for field-scale zone delineation. In fact, our results demonstrated that temporal variability alters the spatial pattern expressed by Landsat 8 panchromatic band and in turn its usefulness as an input to delineate MZs.

There were some similarities between zones delineated by ECa, satellite images and yield data, but yield data indicated less potential for zoning and exhibited some inconsistency. The reason is that different factors affect yield in a complex manner. In fact, it is not possible to distinguish between the variety of factors affecting crop growth and yield, such as physical and chemical properties of soil, irrigation regimes, pests and diseases and climate, using yield maps alone (Corwin and Lesch, 2010). We realized that even when visual assessment of yield maps indicates similar spatial arrangement as soil related maps the clustering result may be different. Moreover, for fields with no available yield map, using ECa and/or space-borne satellite images represents a great time saving. Averaging yield maps across years, as we did in our study, could be a safe guard against temporal changes and an attempt to find more stable scheme of yield distribution. On the other hand, if high temporal variance in yield spatial distribution exists, averaging yield data across years may mask information and cause misleading interpretation. We averaged cotton yield data from 2012 to 2014. In 2012, the field was uniformly irrigated by the farmer but an irrigation study was performed in 2013 and 2014. Visual assessment of maps against each other suggested that spatial soil differences strongly affected yield patterns during those years, yet this impact was less pronounced for 2013 cropping season where delayed planting and temperature distribution over the growing season depressed yield and influenced its spatial distribution to some extent.

4.4 Temporal variability and role of crop

We examined the necessity of having multiple zoning schemes within a cropping season in our study site. The result revealed if in depth soil variation is significant, the spatial arrangement of optimum irrigation MZs may vary within a cropping season. There is evidence in literature showing the effect of season to season climate variability in spatial arrangement of MZs, hence supporting the idea of dynamic zoning. For instance, Schepers et al. (2004) showed that temporal climate variability even under irrigation may change the yield spatial variation. Temporal changes may affect the spatial patterns of yield such that no consistent high/low yield zones were observed (Guastaferro et al., 2010). Rainfall patterns and available heat unit are the prime factors affecting cotton growth and yield in west Tennessee with a short season humid environment. For instance, if dry periods at sensitive growth stages occur, soil with higher water holding capacity provides better conditions for crop to avoid undergoing water stress, hence higher yield is expected for such soil. In contrast, a wet year with unexpected early heavy rainfall may delay planting hence reducing the chance for the crop to accumulate enough heat units. In such scenario, soils with high AWC may suffer from excessive vegetative growth and unopened bolls by the time of harvesting. Fraisse et al. (2001) found that number of MZs for precision farming declined when either enough water was provided throughout the growing season or a drought tolerant variety was planned. The optimum number of MZs is also influenced by the crop planted (Fraisse et al., 2001). Further study is required to investigate the crop response to water level across different soil types as a means to provide crop-specific irrigation MZs.

5. Conclusion

Precision agriculture is a farming system which uses information technology to do site-specific crop management in which decisions on resource application are modified with regard to within field variation of components such as soil, water and crop (Whelan and Taylor, 2013). Field-scale spatiotemporal variation is significant in west Tennessee, thus variable rate irrigation may enhance water use efficiency and crop yield. Conventional irrigation management tries to answer when and how much to irrigate. Given within field soil variation, variable rate irrigation management ought to address where to irrigate as well. Delineating management areas within fields is an important part of a precision farming system where it is expected that applying identical treatment will cause significant yield differences. Methods for delineating MZs vary widely in the information used as well as the techniques for creating

the zone boundaries. Use of on-the-go sensors and remote sensing technologies is appealing because it is easy to collect these data on a field-scale.

We evaluated the performance of several clustering methods for zone delineation. We also designed a new zoning strategy based on integer linear programming for center pivots with limited speed control ability. We used high resolution soil available water maps as standard input to zone the field. We also assessed the effectiveness of ECa, space-borne satellite images and yield data for zone delineation. The clustering methods performed similar in efficiently dividing the field into relatively homogenies zones in respect to soil hydraulic properties. The introduced integer linear programming method offers the optimum zoning strategy for center pivots with limited speed control ability to match water input to soil spatial patterns. However, further research is essential to investigate the optimum irrigation strategy over pies with a mixture of soils. The results of this study suggested that ECa and satellite images may be used to determine site-specific irrigation MZs in west Tennessee if a high spatial similarity is observed between these ancillary data with soil hydraulic properties. The spatial and temporal resolution and precision varies among different types of ancillary data and affect their efficiency for zone delineation. Temporal variability in soil moisture altered expression of spatial variability in space-born satellite images. Yield data did not show a high potential for zone delineation. We attribute this to various factors that affect yield in a complex manner and also temporal differences in yield data. In practice, farmer knowledge on field conditions could be extremely useful to choose the most appropriate data set for zone delineation.

References

Arslan, S., & Colvin, T.S. (2002). Grain yield mapping: Yield sensing, yield reconstruction, and errors. Precision Agriculture 3(2), 135-154.

Basso, B., Fiorentino, C., Cammarano, D., Cafiero, G., & Dardanelli, J. (2012). Analysis of rainfall distribution on spatial and temporal patterns of wheat yield in Mediterranean environment. European Journal of Agronomy, 41, 52-65.

Cid-Garcia, N. M., Bravo-Lozano, A. G., & Rios-Solis, Y. A. (2014). A crop planning and real-time irrigation method based on site-specific management zones and linear programming. Computers and Electronics in Agriculture, 107, 20-28.

Cid-Garcia, N. M., Albornoz, V., Rios-Solis, Y. A., & Ortega, R. (2013). Rectangular shape management zone delineation using integer linear programming. Computers and Electronics in Agriculture, 93, 1-9.

Cohen, Jacob (1960). "A coefficient of agreement for nominal scales". Educational and Psychological Measurement 20 (1): 37–46.

Córdoba, M., Bruno, C., Costa, J., & Balzarini, M. (2013). Subfield management class delineation using cluster analysis from spatial principal components of soil variables. Computers and Electronics in Agriculture, 97, 6-14.

Corwin, D. L., & Lesch, S. M. (2010). Delineating site-specific management units with proximal sensors. In Geostatistical Applications for Precision Agriculture (pp. 139-165). Springer Netherlands.

Daccache, A., Knox, J. W., Weatherhead, E. K., Daneshkhah, A., & Hess, T. M. (2015). Implementing precision irrigation in a humid climate–Recent experiences and on-going challenges. Agricultural Water Management, 147, 135-143.

Evans, R. G., LaRue, J., Stone, K. C., & King, B. A. (2013). Adoption of site-specific variable rate sprinkler irrigation systems. Irrigation Science, 31(4), 871-887.

ESRI. (2014). ArcGIS 10.2.2 Help. Esri, Redlands, CA.

Fraisse, C. W., Sudduth, K. A., & Kitchen, N. R. (2001). Delineation of site-specific management zones by unsupervised classification of topographic attributes and soil electrical conductivity. Transactions of the ASAE, 44(1), 155-166.

Fridgen, J. J., Kitchen, N. R., Sudduth, K. A., Drummond, S. T., Wiebold, W. J., & Fraisse, C. W. (2004). Management zone analyst (MZA). Agronomy Journal, 96(1), 100-108.

Guastaferro, F., Castrignanò, A., De Benedetto, D., Sollitto, D., Troccoli, A., & Cafarelli, B. (2010). A comparison of different algorithms for the delineation of management zones. Precision agriculture, 11(6), 600-620.

Guo, W., Maas, S. J., & Bronson, K. F. (2012). Relationship between cotton yield and soil electrical conductivity, topography, and Landsat imagery. Precision Agriculture, 13(6), 678-692.

Haghverdi, A., Cornelis, W. M., & Ghahraman, B., 2012. A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. Journal of Hydrology 442, 46-54.

Haghverdi, A., Öztürk, H. S., & Cornelis, W. M., 2014. Revisiting the pseudo continuous pedotransfer function concept: Impact of data quality and data mining method. Geoderma. 226, 31-38.

Hedley, C.B., & Yule, I.J. (2009). Soil water status mapping and two variable-rate irrigation scenarios. Precision Agriculture, 10(4), 342-355.

Hornung, A., Khosla, R., Reich, R., Inman, D., & Westfall, D. G. (2006). Comparison of site specific management zones: soil color based and yield based. Agronomy Journal, 98, 405–417.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31(8), 651-666.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. ACM computing surveys (CSUR), 31(3), 264-323.

Khosla, R., Westfall, D. G., Reich, R. M., Mahal, J. S., & Gangloff, W. J. (2010). Spatial variation and site-specific management zones. In Geostatistical applications for precision agriculture (pp. 195-219). Springer Netherlands.

King, B. A., Stark, J. C., & Wall, R. W. (2006). Comparison of site-specific and conventional uniform irrigation management for potatoes. Applied Engineering in Agriculture, 22(5), 677-688.

Kitchen, N. R., Sudduth, K. A., Myers, D. B., Drummond, S. T., & Hong, S. Y. (2005). Delineating productivity zones on claypan soil fields using apparent soil electrical conductivity. Computers and Electronics in Agriculture, 46(1), 285-308.

Mathworks. (2014). Statistics Toolbox™ User's Guide (R2014b). Retrieved November 10, 2011 from http://www.mathworks.com/help/pdf_doc/stats/stats.pdf

Moral, F. J., Terrón, J. M., & Silva, J. R. (2010). Delineation of management zones using mobile measurements of soil apparent electrical conductivity and multivariate geostatistical techniques. Soil and Tillage Research, 106(2), 335-343.

Reynolds, D. (2009). Gaussian mixture models. In Encyclopedia of Biometrics (pp. 659-663). Springer US.

Schepers, A. R., Shanahan, J. F., Liebig, M. A., Schepers, J. S., Johnson, S. H., & Luchiari, A. (2004). Appropriateness of management zones for characterizing spatial variability of soil properties and irrigated corn yields across years. Agronomy Journal, 96(1), 195-203.

Singh, A. (2012). An overview of the optimization modelling applications. Journal of Hydrology, 466, 167-182.

Song, X., Wang, J., Huang, W., Liu, L., Yan, G., & Pu, R. (2009). The delineation of agricultural management zones with high resolution remotely sensed data. Precision agriculture, 10(6), 471-487.

Thöle, H., Richter, C. & Ehlert, D. (2013) Strategy of statistical model selection for precision farming on-farm experiments. Precision Agriculture, 1-16.

Whelan, B., & Taylor, J. (2013). Precision Agriculture for Grain Production Systems. CSIRO PUBLISHING. Pages 208.

Zhang, X., Shi, L., Jia, X., Seielstad, G., & Helgason, C. (2010). Zone mapping application for precision-farming: a decision support tool for variable rate application. Precision agriculture, 11(2), 103-114.

Appendix 4: Chapter 4 Figures and Tables

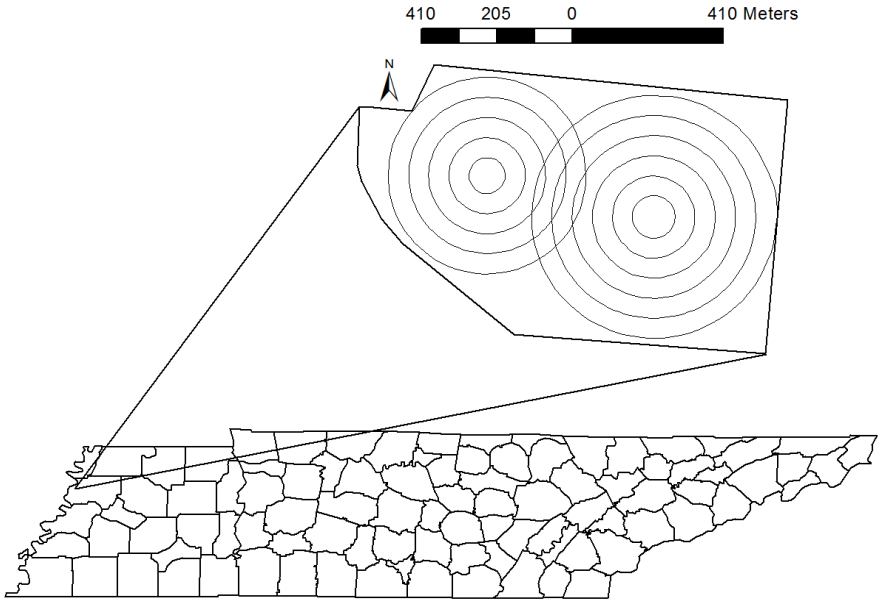

Figure 4-1. Field of study within the state of Tennessee.

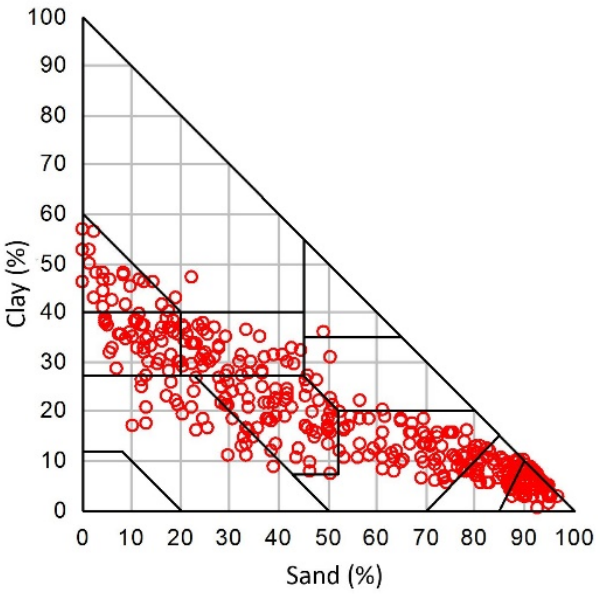

Figure 4-2. Soil texture distribution for the samples (n= 400) collected from the field of study.

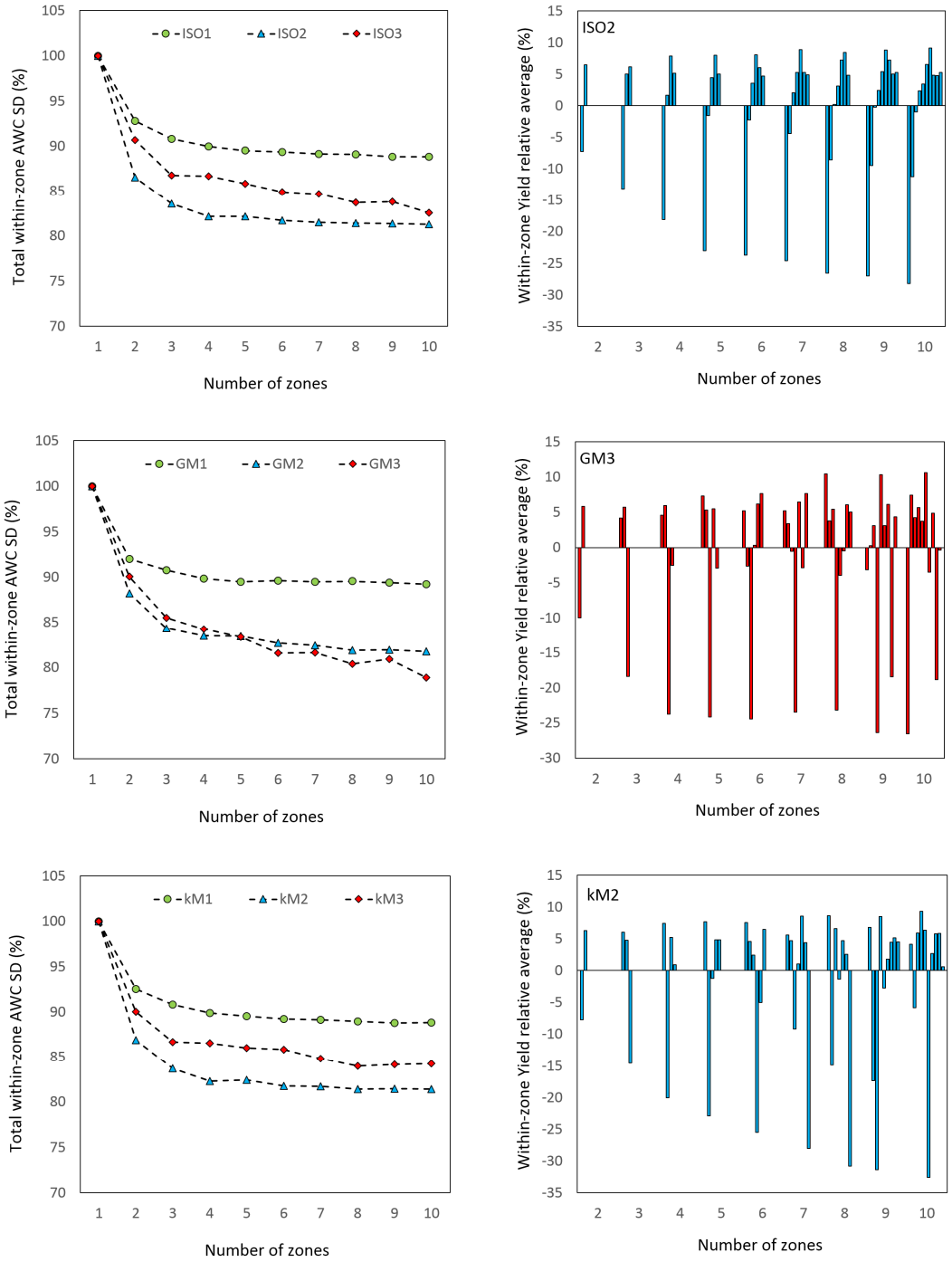

Figure 4-3. Performance of clustering methods based on percent reduction in AWC variance and within-zone average yield. kM: *k*-means; GM: Gaussian mixture; ISO: ISODATA-maximum likelihood; 1, 2, and 3: ECS, ECD and ECS plus ECD were used for clustering.

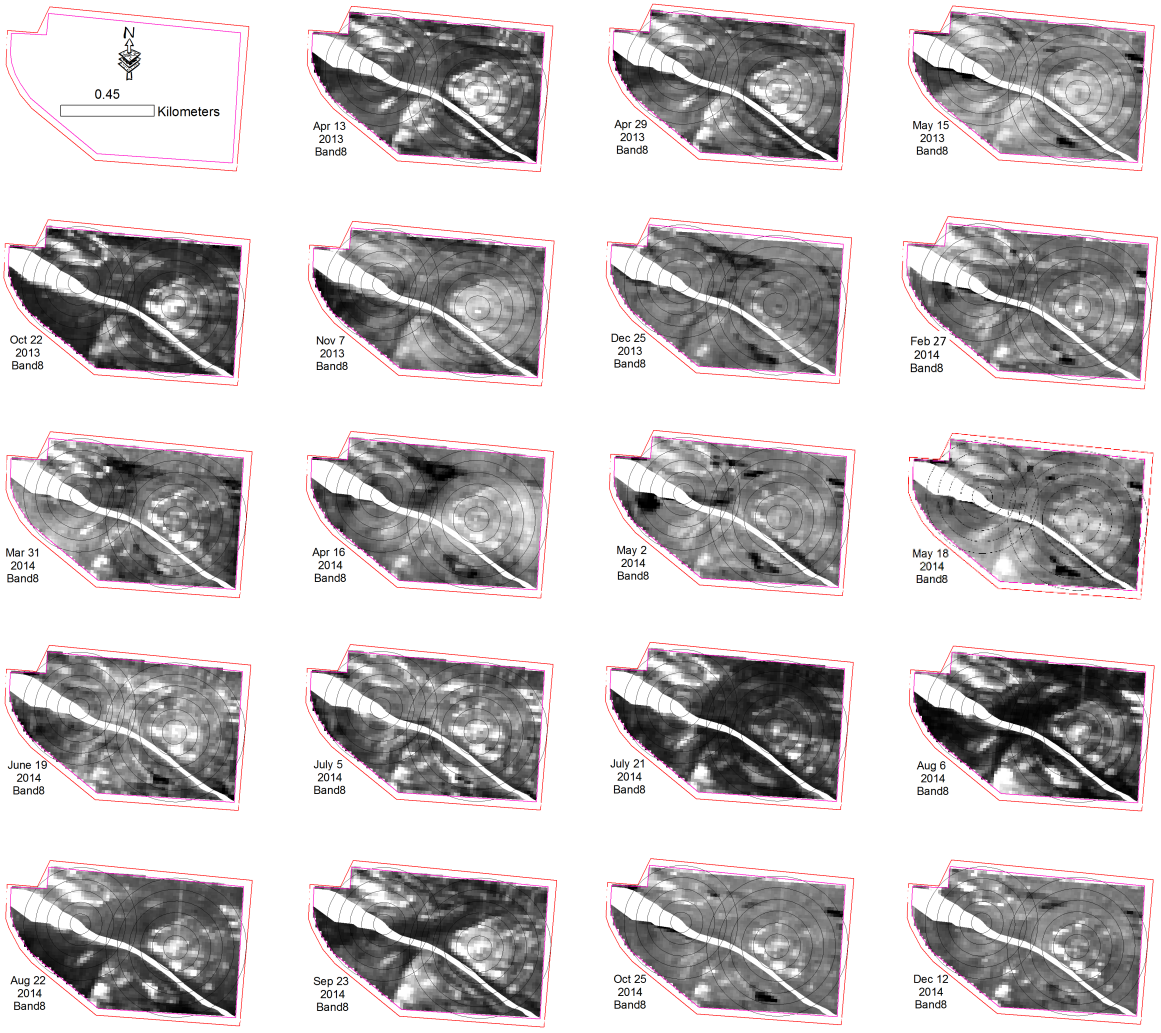

Figure 4-4. Gray-scale maps of panchromatic spectral band during 2013-2014 from the field of study.

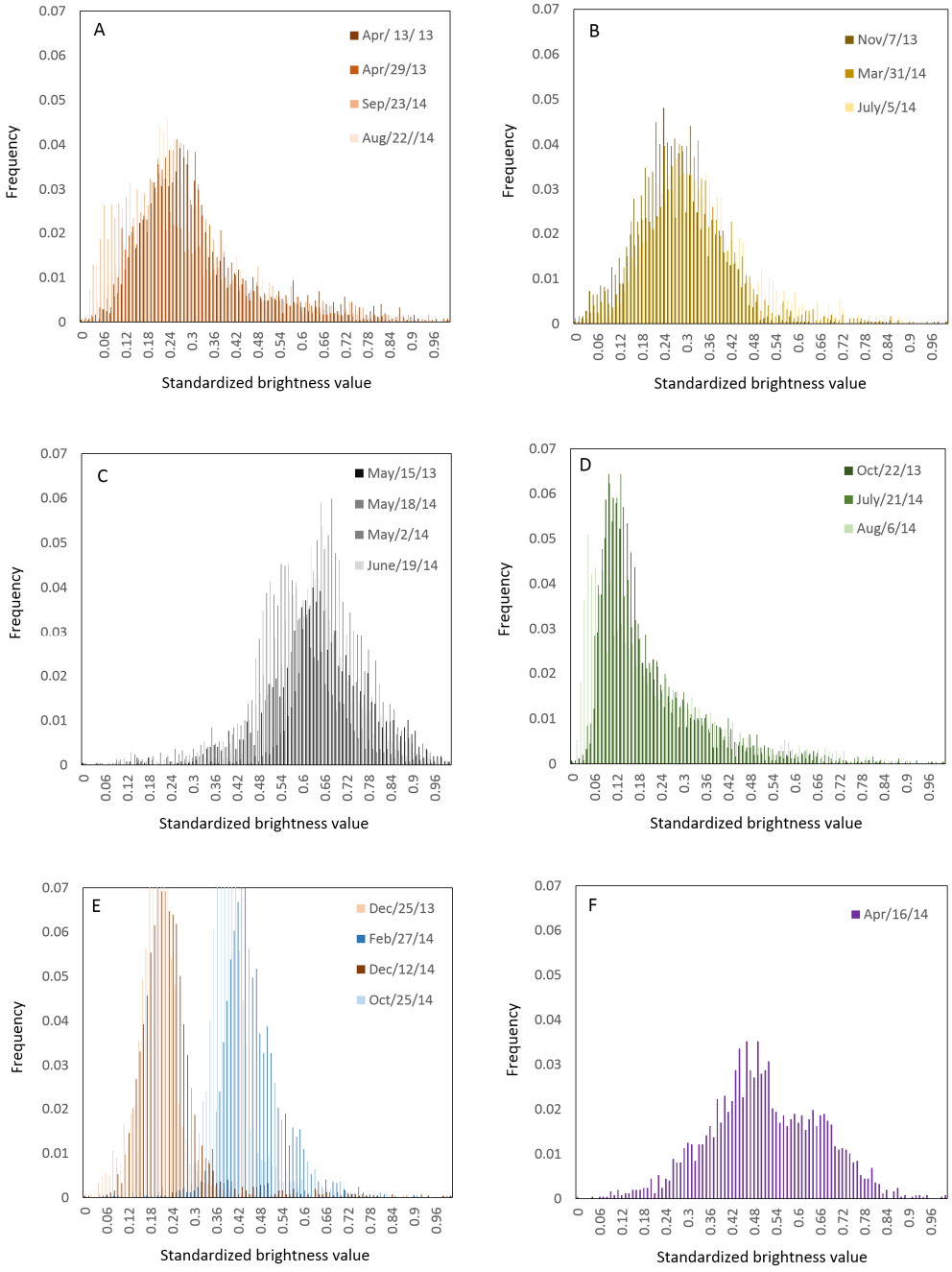

Figure 4-5. Distribution of standardized brightness values for panchromatic spectral band during 2013-2014.

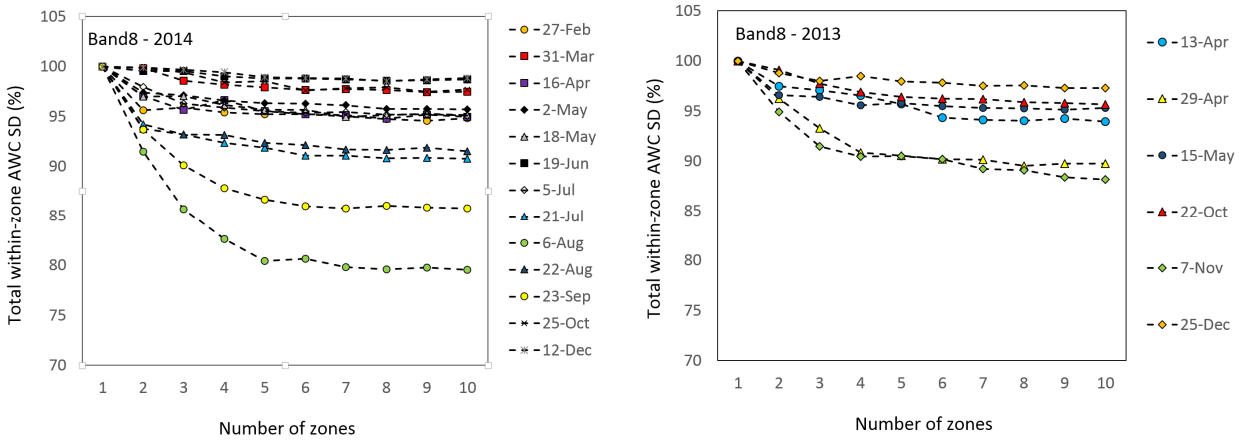

Figure 4-6. AWC variance against number of zones using different satellite images. Band 8: reflectance value for panchromatic band from images taken by Landsat 8.

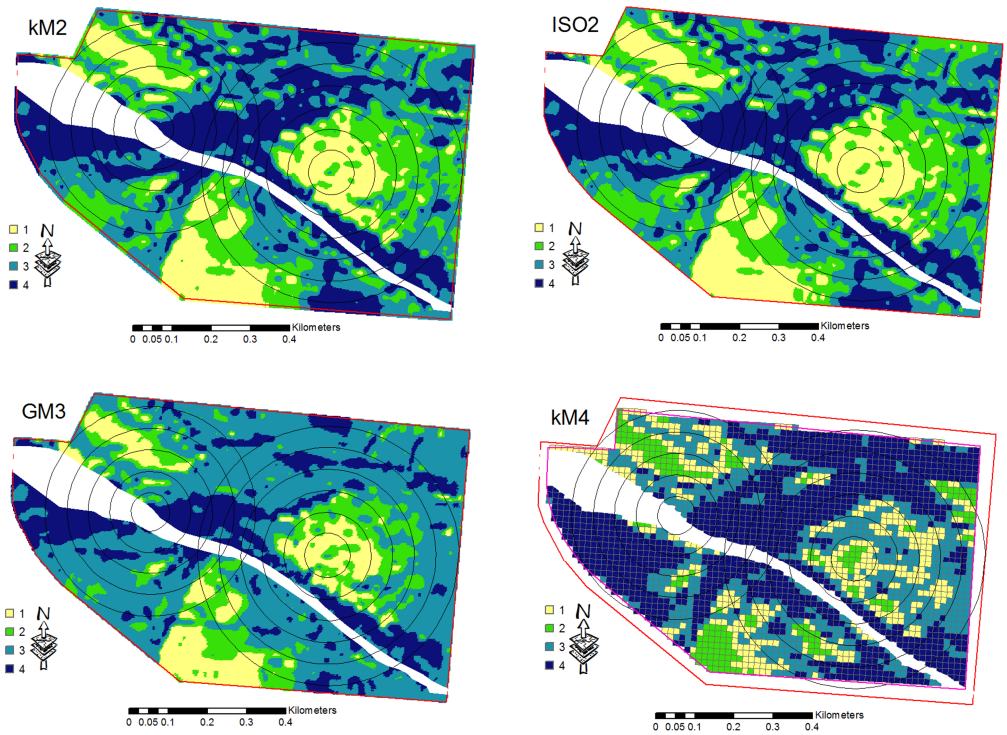

Figure 4-7. Spatial arrangement of MZs. kM2: *k*-means with ECD as input; ISO2: ISO-ML with ECD as input; GM3: Gaussian mixture with ECS and ECD as inputs, kM4: *k*-means with satellite image (Landsat 8 panchromatic band) as input.

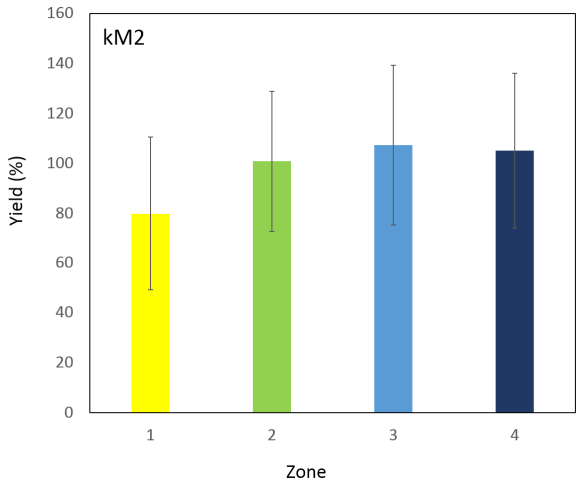

Figure 4-8. Percent of average yield within MZs clustered with *k*-means considering ECD as input (bar colors are related to the zones in Figure 4-7).

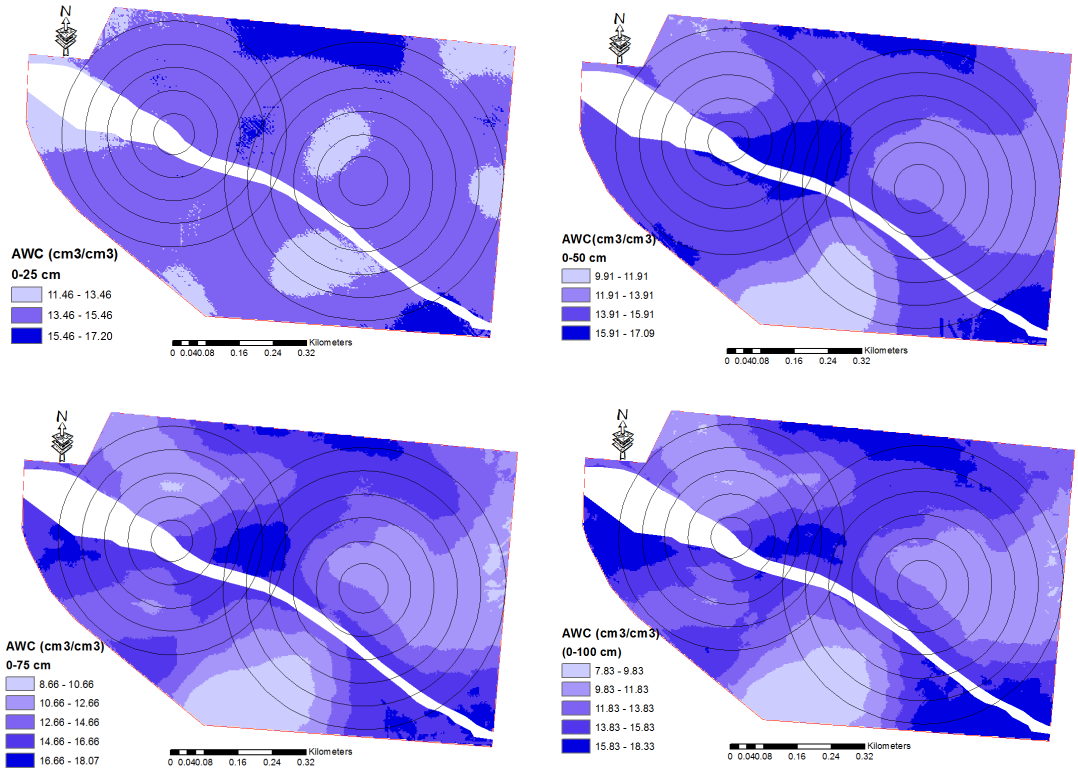

Figure 4-9. Within-season temporal variation in user defined MZs considering available water content as input.

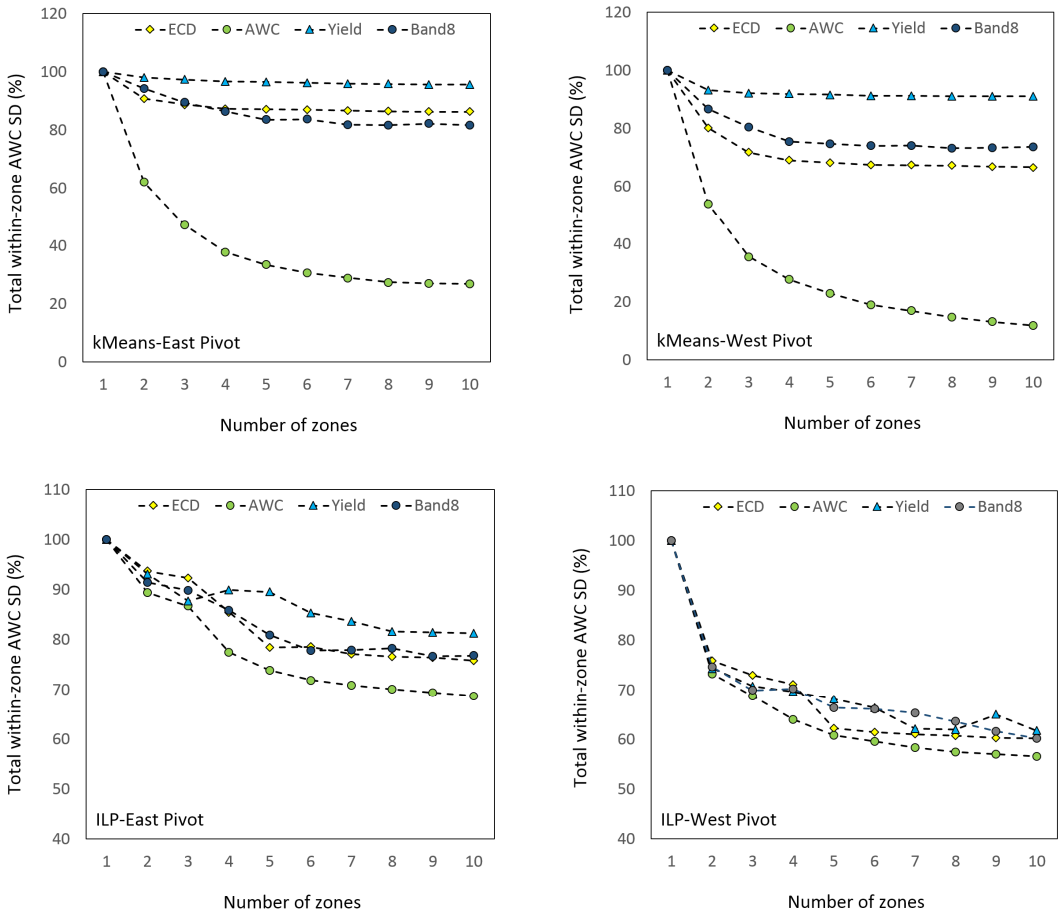

Figure 4-10. Performance of irrigation MZ delineation methods based on AWC standard deviation reduction. ECD: ECa deep; AWC: available water content; Band 8: reflectance value from satellite images taken by Landsat 8.

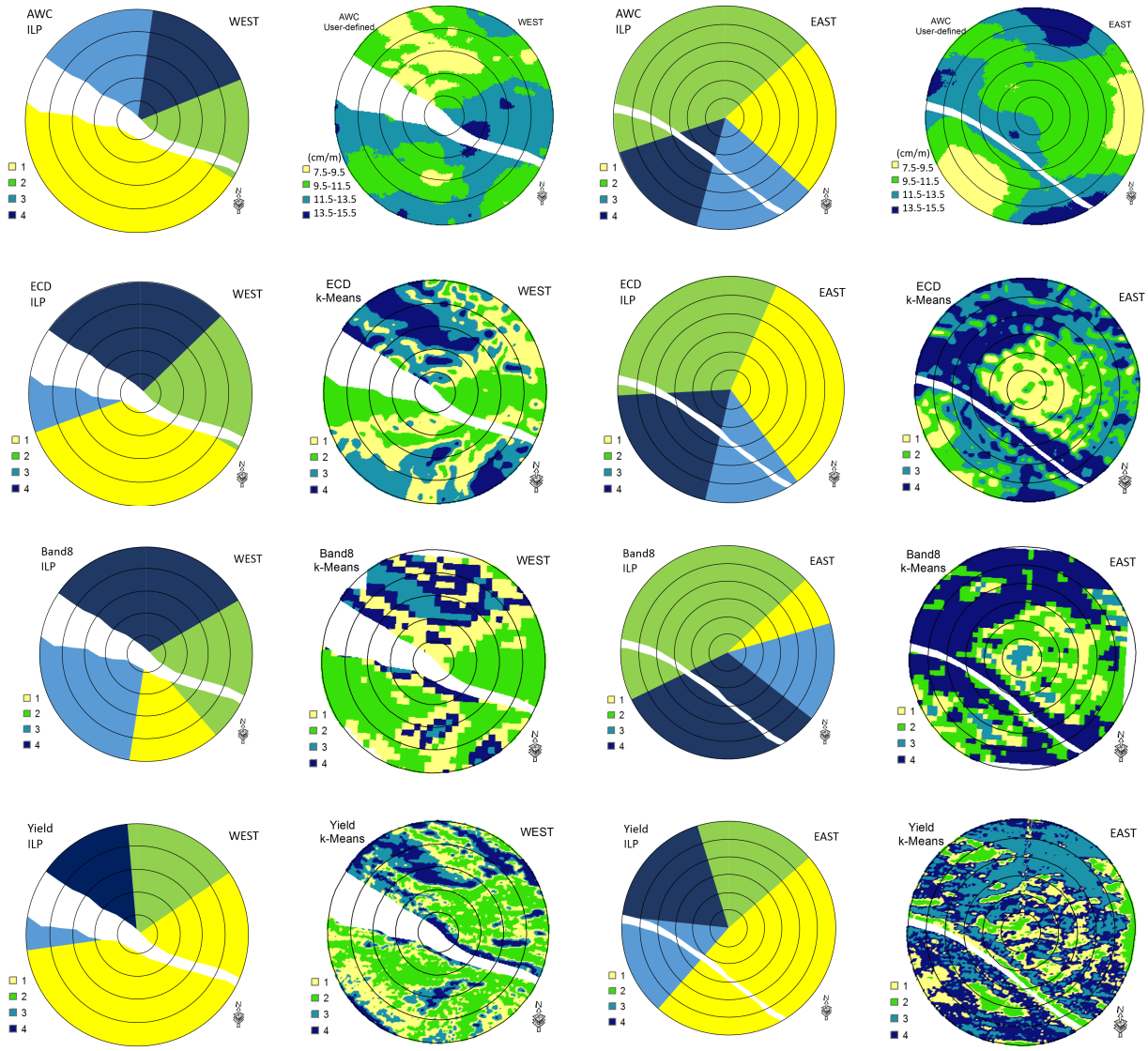

Figure 4-11. Irrigation MZs for center pivot systems located at east and west parts of the field. ILP: integer linear programming, Band 8: reflectance values from panchromatic band taken by Landsat 8 satellite; ECD: deep ECa.

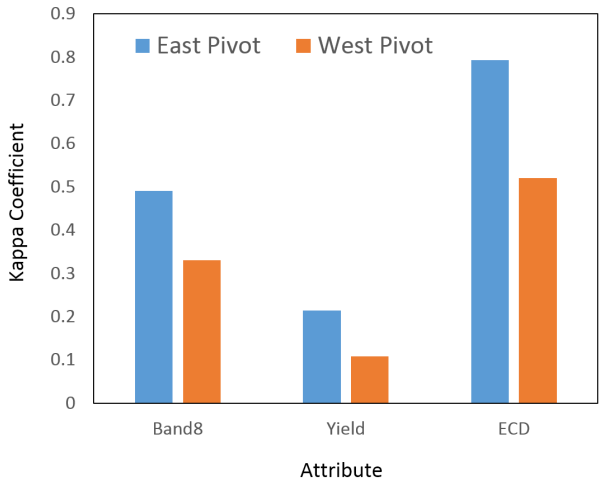

Figure 4-12. Kappa coefficient for the zones delineated by the integer linear programming (ILP) method. Band 8: reflectance values from panchromatic band taken by Landsat 8 satellite; ECD: deep ECa.

Table 4-1. Descriptive statistics for soil water content from different soil layers (from chapters 1 and 2).

Variable*	Depth	Min.	Max.	Mean	SD**
VWC, cm m^{-1}	0-25 cm	10.75	59.74	28.35	7.43
	25-50 cm	7.27	43.12	26.02	10.78
	50-75 cm	5.98	42.38	21.64	11.08
	75-100 cm	5.67	45.32	20.18	11.15
	0-100 cm	3.94	47.61	17.94	8.49
θ_{33}, cm m^{-1}	0-25 cm	9.86	43.57	29.43	5.83
	25-50 cm	5.26	47.20	26.03	10.75
	50-75 cm	4.93	43.54	20.55	11.26
	75-100 cm	5.38	38.95	17.31	9.71
θ_{1500}, cm m^{-1}	0-25 cm	7.72	29.61	19.69	4.28
	25-50 cm	5.23	35.49	16.98	7.22
	50-75 cm	4.57	29.32	13.14	6.47
	75-100 cm	4.75	24.52	11.13	5.09

* VWC: volumetric water content at the time of sampling, θ_{33}: water content at field capacity (-33 kPa); θ_{1500}: water content at permanent wilting point (-1500 kPa).

** SD: standard deviation.

Table 4-2. Summary of zoning algorithms and attributes.

	Clustering procedure	Attributes
Phase 1	*k*-means, ISODATA-ML, GM	ECS
	k-means, ISODATA-ML, GM	ECD
	k-means, ISODATA-ML, GM	ECS, ECD
	k-means	Satellite image
	User defined zones	AWC
Phase 2	ILP, *k*-means	ECD
	ILP, *k*-means	Satellite image
	ILP, *k*-means	AWC
	ILP, *k*-means	Yield

* GM: Gaussian Mixtures, ISO-ML: Iterative Self-Organizing Data Analysis Technique - Maximum Likelihood, ILP: Integer linear programing, AWC: available water content, ECS: shallow ECa, ECD: deep ECa.

Table 4-3. All of the pie shape zones underneath a center pivot irrigation system.

			Pie length	
Start angle	**1**	**2**	**…**	**360**
0	0-1*	0-2	**…**	0-0
1	1-2	1-3	**…**	1-1
2	2-3	2-4	**…**	2-2
.	**.**	**.**	**…**	**.**
.	**.**	**.**	**…**	**.**
.	**.**	**.**	**…**	**.**
359	359-0	359-1	**…**	359-359

* X-Y: start angle-end angle.

Chapter 5: Studying uniform and variable rate center pivot irrigation strategies with the aid of site-specific water production function

(Submitted to *Computers and Electronics in Agriculture* for publication)

Amir Haghverdi[1*], Brian G. Leib[1*], Paul D. Ayers1, Robert A. Washington-Allen[2], Michael J. Buschermohle[1]

[1]Department of Biosystems Engineering & Soil Science, University of Tennessee, 2506 E.J. Chapman Drive, Knoxville, TN 37996-4531

[2] Department of Geography, University of Tennessee, Burchfiel Geography Building, Knoxville, TN 37996-0925

* Correspondence author, ahaghver@vols.utk.edu, (865)2359694

Abstract

Water production functions (WPFs), mathematical relationships between applied water and crop yield, are useful tools for irrigation management and economic analysis of yield reduction due to deficit irrigation. This study aimed at (i) designing and evaluating site-specific WPFs (using *k* nearest neighbors, multiple linear regression, and neural networks) and (ii) using the best WPF to investigate different cotton irrigation and zoning strategies using integer linear programming. A two-year cotton irrigation experiment (2013-2014) was implemented to study irrigation-cotton lint yield relationship across different soil types. The site-specific *k* nearest neighbors WPFs showed the highest performance with root mean square error equal to 0.131 Mg ha^{-1} and 0.194 Mg ha^{-1} in 2013 and 2014, respectively. The result indicated that variable rate irrigation with pie shape zones could enhance cotton lint yield under supplemental irrigation when field-scale spatial soil variation is significant. The temporal changes in climate and rainfall patterns, however, had a great impact on cotton response to irrigation in west Tennessee, a moderately humid region with short season environment. The result indicated that site-specific WPFs were useful empirical tools for on-farm irrigation research.

Keywords: cotton, precision agriculture, principle component analysis, variable rate irrigation.

1. Introduction

1.1 Precision farming and variable rate irrigation

The demand for food and fiber is increasing as the world population grows. Irrigation management has evolved into a top priority issue since available fresh water resources are limited. Stabilizing yields and the recent high commodity prices are the prime reasons that row crop irrigation is rapidly expanding in west Tennessee, a moderately humid region. Given the field-scale spatial soil variation in this region, variable rate irrigation (VRI) was suggested as the appropriate irrigation scenario (Duncan, 2012). Conventional irrigation management tries to answer when and how much to irrigate. Given within field soil variation, VRI management should address where to irrigate as well. VRI is expected to improve water-use efficiency, increase productivity, save fuel and decrease nutrient leaching (Pan et al., 2013). It has been shown in a variety of studies that soil characteristics such as soil water holding capacity and depth of soil significantly affects crop yield, hence irrigation strategies should be adjusted in regard to the soil type (Duncan, 2012). Hedley et al. (2009) compared VRI and conventional uniform irrigation and illustrated that 9–19 % of irrigation water was saved which in turn reduced nitrogen leaching. VRI related research projects have mainly focused on (i) designing sprinklers to spatially vary irrigation rates and (ii) software/hardware for guiding the system (Pan et al., 2013). Precision irrigation center pivots that pulse banks of sprinklers to create variable flow have been commercially available for a while. Another option is to program the control panels to change travel speed of center pivot systems, thus creating some pie shape irrigation zones. This is a more affordable form of VRI that producers in west Tennessee and other places can easily practice with available center pivot irrigation systems.

One of the prime steps towards precision farming is to delineate management areas within fields where it is expected that applying identical treatment will cause significant yield differences. A corollary expectation is that varying the treatment of these areas will facilitate optimizing yield. In practice, the number of zones depends on the target input, available equipment and crop planted. For instance, there is more freedom to precisely apply fertilizer even in very small regions in a field, but this is not the case with VRI. Pan et al. (2013) analyzed soil water status through spatial and temporal data and suggested this method as a means to assess the potential of VRI and to schedule irrigation. In chapter 4 integer linear programing (ILP) was used to find the optimum number and spatial arrangement of pie shape zones for center pivots with limited speed control capability. Some soil properties were

required to delineate management zones, yet it is important to investigate crop response to water level amount across different soil types as a means to provide crop-specific irrigation management zones.

Traditional irrigation/agronomic work tries to address the irrigation related questions by means of small plot trials over multiple sites and across several years. This is probably the best way to collect the necessary data in controlled environment following statistical principles behind design of experiments. However, there are some shortcomings. Given the time consuming and labor intensive nature of a field experiment, this approach becomes unrealistic for the near future (Drummond et al., 2003). In addition, this method relies on transformation of the research findings from small plots to farmers' fields. In reality, each field has its own complexity and needs a site-specific irrigation management plan. An alternative option is to make use of mechanistic crop growth models. The drawback of these models is that they are time-consuming and expensive to develop because of numerous inputs that need to be collected to run these models (Drummond et al., 2003).

Precision agriculture (PA) enables farmers to collect numerous site-specific data. Empirical analysis of such data may be the key to help farmers learn from their occupational practice. PA is successful if accurate and detailed information about crop response to specific conditions is provided (Drummond et al., 2003). Quantifying crop response to water is the first step toward optimizing irrigation, therefore a critical issue for farmers, governmental agencies and consulting companies. Given the time consuming and expensive nature of irrigation studies, developing computer tools and models are very helpful to get a comprehensive understanding of irrigation-yield relationship for different crops across soil types. Robust algorithms are needed for identifying site-specific crop yield responses to environmental and management parameters and helping to find optimum practices.

1.2 Site-specific water production functions

Cropping systems are complex in their nature. A variety of factors, such as crop available water, weather and rainfall variation, seed quality, topographic attributes, soil properties and pest control, simultaneously affect crop growth and yield in a spatiotemporal manner. Traditional farming is based on average field condition which makes decision making and field operation easier yet usually demands more input and produces less income. Moving from conventional farming to PA, the goal has become to deal with within field variations. As a result, understanding and modeling the effect of different parameters on crop yield in a spatiotemporal scheme becomes a crucial research topic in PA.

The term production function (PF) may be assigned to any mathematical relationship between crop yield and input components such as water, fertilizers and energy (De Juan, 1996). In practice, almost all the derived PFs require crop water use as an independent variable. PFs predict total dry matter (or marketable product of each crop) as dependent variable while the independent variables are transpiration, evapotranspiration (ET) or amount of applied water during irrigation (IW). These functions are divided into two groups based on the independent variable: (i) crop water production functions (CWPFs) which use ET and (ii) water production functions (WPFs) which use IW. The IW may consist of different components such as crop water requirement, pre-plant irrigation, leaching requirement, and rainfall (Igbadun et al., 2007).

Recently, some studies have aimed to revisit and rebuild the concept of PF. Some examples are: a WPF for water logging stress on corn (Kuang et al., 2012) and rice water – fertilizer PF (Ai-hua et al., 2012). Tong and Guo (2013) also tried to involve an estimation of uncertainty in CWPF along with the optimal allocation of water resources in an irrigation area. Saseendran et al. (2014) established location-specific CWPFs using long term averaged data for corn in Colorado. They used the RZWQM2 model and historical weather data to develop average corn CWPFs across years and locations. There are some studies on cotton yield prediction by means of PFs. Wang et al., (2007) derived cotton and wheat water-salinity PFs. Dinar et al. (1986) derived cotton CWPF under saline condition in California. They stated that PFs were good models to improve irrigation management.

Classical PFs are useful tools for irrigation management and economic analysis of yield reduction due to deficit irrigation, but there are some shortcomings associated with them. PFs, like other regression-based equations, are relatively easy to build but are mostly linear and not powerful enough to model complex ecological systems (Dai et al., 2011; Haghverdi et al., 2014). That is why recent studies (e.g. Fortin et al., 2010; Haghverdi et al., 2014) are looking for more robust and non-linear techniques to predict yield. In fact, in most of agricultural related studies, machine learning methods became the favorable data mining tools because they are powerful empirical algorithms for modeling complex systems.

Data mining could be defined as the process of capturing important and useful information from large data sets (Mucherino et al., 2009). An extensive review on data mining methods and their application in agricultural related studies was gathered by Huang et al. (2010). Machine learning algorithms were employed to some extent for predicting yield of different crops. Fortin et al. (2010) used artificial neural networks (ANNs) for predicting potato tuber growth as well as its in-field variations in Canada. They reported that with an ANN model using enough data

one can precisely model site-specific tuber growth. Dai et al. (2011) adopted ANN and multi-linear regression models to simulate the response of sunflower yield to soil moisture and salinity; ANNs appeared to be the model with higher precision than regression. They concluded that ANN is a useful tool for modeling relationship between crop yield and soil moisture and salinity at different crop growth periods. Haghverdi et al. (2014) derived some novel WPFs and compared data mining-based methods with traditional regression procedure. They utilized ANN and decision tree as modeling algorithms for deriving water salinity PFs for spring wheat and reported promising result for ANN. This study aimed to (i) design and evaluate site-specific cotton WPFs and (ii) use the best WPF to investigate different cotton irrigation and zoning strategies.

2. Material and Methods

2.1 Field of study

The study site (approximately 73 ha) was located in west Tennessee, a short season semi-humid region, along the Mississippi river (Figure 5-1). The rainfall is usually high during cropping season, yet supplemental irrigation is practiced to fulfill crop water requirement when dry periods occur. There were two center pivot systems for irrigation covering majority of the field with some overlap in their coverage. Given the significant spatial soil variation, this field was an ideal site for VRI study.

2.2 Irrigation experiment

A two-year experiment (2013-2014) was implemented to study the irrigation-cotton lint yield relationship across different soil types. The goal was to conduct applied research by engaging farmer in irrigation decision making process, therefore both using and enhancing his knowledge and expertise over the course of the experiment. Two different methods were required to vary the irrigation throughout the field including programing the pivots and partially swapping the sprinkler nozzles. The pivot's control panels were Valley Select2 module (Valmont Industries, Inc) programmable to up to 9 sectors to vary the irrigation across the field. The program changed the irrigation rate by adjusting the pivot's travel speed where speeding the pivot up caused less irrigation and vice versa. The irrigation management and experiment was explained in details in chapter 3.

2.3 Input attributes and data preparation

Table 5-1 summarizes the data used in this study to derive WPFs. The field was sampled on March 21 and 22, 2014 (400 undisturbed samples from 100 sampling locations at 4 different depths) to measure soil texture, soil water content and bulk density. Then, the soil water retention curves were predicted for samples and interpolated to generate high resolution maps. More details on soil data analysis and modeling was provided in chapters 1 and 2. In addition, soil apparent electrical conductivity (ECa) was collected using a Veris 3100 (Veris Technologies, Salina, KS) instrument. More information on proximal soil data collection/sensing was provided by in chapter 1. The cotton yield maps were obtained from yield monitors. The raw yield data were cleaned as explained in chapter 3. The field was divided into $25m^2$ cells and for each cell all attributes (input and output) were obtained. The kriging was used for interpolation. A total of 16,000 cells were prepared after removing cells with missing/bad data. Considering overlap between coverage areas of pivots was outside of the scope of this research, hence pivots were considered as separate individual irrigation units throughout this study.

In many situation, there are only a few major driving forces governing a system. As number of input predictor increases, it is more likely to have redundancy/inter-correlation in data, meaning more than one attribute can explain the same driving principle controlling the behavior of the system. Prinicipal component analysis (PCA) is a statistical method aiming at reducing the dimension of a given data set, finding hidden patterns amongst the data and extracting the important information from the raw data through replacing them by a group of new input predictors called principle components. Principal components are orthogonal to each other and each of them is a linear combination of the original input attributes. (Mathworks, 2014; Abdi and Williams, 2010). In this study, Matlab 2015a (MathWorks, Inc., Natick, Mass.) was used to perform PCA. The Hotelling's T2, a statistical measure of the multivariate distance of each observation from the center of the data set, was calculated to find the most important PCs.

Prior to establishing WPFs, the data set was randomly divided into 10 almost equal-size subsets. The modeling was repeated 10 times where each time one subset was assigned to the test phase and the rest of them were used to develop the WPFs.

2.4 Modeling and simulation

2.4.1 Phase 1: Establishing water production functions

Multiple Linear Regression water production function, MLR WPF

Figure 5-2 illustrates the three methods designed to establish site-specific WPFs. Traditionally linear regression is used to derive WPFs. Evidence in literature shows that cotton WPFs differs across soils with different water holding capacities (Duncan, 2012). Therefore, *k*-means was used to divide the data into 4 sub-groups with more homogenous soil properties following the recommendations in chapter 4. The ten most important PCs were used as input for clustering. Then, for each sub-group a separate MLR WPF was developed considering IW, IW^2 and IW×PCs as input predictors.

K Nearest Neighbors water production function, k-NN WPF

The *k*-NN is a non-parametric lazy learning algorithm that processes all of the data in real-time when prediction of the response(s) for a new observation is required. The *k*-NN was considered to be among the top ten techniques used for data mining (Wu et al., 2008), but in general has been applied less than other data mining techniques in agriculture related studies (Mucherino et al., 2009). The *k*-NN considers a distance function, usually Euclidean distance, to find so-called closer observations with higher similarity. Then, it combines the response of *k* nearest surrounding observations to predict the response for a new observation. Figure 5-2 provides a visual step-by-step guide to derive site-specific *k*-NN WPFs.

First, raw inputs were converted to PCs. Then, first 10 PCs plus IW were converted to *input predictors* with a zero mean and a standard deviation of one and then the ranges were tuned. In order to predict yield for a target cell of the test data set at a specific irrigation level (I^*), the following steps were taken. First, distance from the target cell to every cell in the development data set was computed as the summation of the differences among their *input predictors*:

$$d_i = \sqrt{\sum_{j=1}^{N} \Delta^2 a_{ij}} \quad (1)$$

where d_i is the distance between *i*th cell in the development data set and the target cell, Δa_{ij} is the difference between *j*th attribute of the two cells and N=11 is the number of the input attributes.

Second, a *reference data set* was created consisting of those cells in the development data set that had irrigation levels within a 20 mm range from the target irrigation level (i.e. I^*). Third, the *k* nearest cells to the target cell were selected from the reference data set. Fourth, a distance-dependent weight was assigned to each *k* nearest cell (Nemes et al., 2006):

$$w_i = \frac{\left(\sum_{i=1}^{k} d_i \Big/ d_i\right)^p}{\sum_{i=1}^{k}\left(\sum_{i=1}^{k} d_i \Big/ d_i\right)^p} \quad (2)$$

where *k* is the number of selected nearest cells, *p* is the design parameter and w_i is the weight of *i*th cell. Finally, the cotton yield for the target cell (Y_p) was predicted as the weighted average of the yield values for the *k* nearest cells:

$$Y_p = \sum_{i=1}^{k} w_i \times Y_i \quad (3)$$

A cross-validation procedure was designed to identify the optimum *k* and *p* values. The parameter *k* was varied from 1 to 10 and *p* was changed from 0.5 to 3 with 0.5 increments. For each combination of *k* and *p* parameters, the bagging technique was implemented to compute the output of the *k*-NN WPF (Breiman, 1996). In other words, the output was the aggregation of the output of multiple WPFs which were established on different realizations of the reference dataset using sampling with replacement. The number of ensembles was chosen to be 40 considering the recommended values by Haghverdi et al., (2015). The modeling was done using Matlab R2014a (MathWorks, Inc., Natick, Mass.).

Artificial Neural Network water production function, ANN WPF

The artificial neural networks (ANNs) exhibited a good potential to establish WPFs (Haghverdi et al., 2014). A three layers perceptron model was selected in this study. The best number of neurons in the hidden layer was automatically computed by SPSS Modeler 15 (SPSS Inc., Chicago, IL) which was used to derive ANN WPFs. The PCs and IW were chosen as input predictors. The bagging technique was implemented to enhance model stability. The number of ensembles was chosen to be 40. Thirty percent of the training data were assigned to cross-validation to avoid over-fitting

2.4.2 Phase 2: Simulating irrigation and zoning strategies

The WPF with highest performance was selected to simulate irrigation and zoning strategies. Note, that both modeling and simulation phases were done separately for 2013 and 2014 data. The seasonal irrigation level was changed from 0 to 150 mm with 10 mm increment. Then, the optimum yield maps were delineated for three irrigation strategies including (i) uniform irrigation, (ii) speed control VRI and (iii) a hypothetical situation in which each cell (5×5 m^2) to be irrigated individually at optimum level.

The integer linear programing (ILP) optimization process developed in chapter 4 was adopted and modified to find the optimum arrangement of pie shape zones to maximize cotton lint yield. Each center pivot irrigation system was considered to consist of 360 pies each 1 degree wide. A total of 12,241 zones were obtained by combining 1 degree pies. Then, the optimum irrigation level to maximize cotton lint yield was predicted for each zone by means of *k*-NN WPF separately for each cropping season. The optimization process was formulated as follows:

$$\max \sum_{i=1}^{n} \left(l \times y_i \times x_i \right) \tag{4}$$

subject to

$$\sum c_{ij} x_i = 1 \tag{5}$$

$$\sum_{i=1}^{n} x_i \le P_{\max} \tag{6}$$

$$lx \ge L_{\min} \tag{7}$$

$$x_i \in \{0,1\}$$

where x_i is the decision variable for zone i, y is the optimum yield predicted by *k*-NN WPF within zone i, n is the total number of the zones, l is the length of each zone in degrees, c_{ij} is a coefficient equal to 1 if angle j is covered by zone i otherwise equal to 0, $P_{\max}$ is the maximum number of desired zones.

Objective function (4) maximizes cotton lint yield across the zones. Restriction (5) guarantees that each degree is covered only by one zone and the optimum zoning scheme covers the entire area underneath the center pivots. Constraint (6) keeps the total number of the zones less than the maximum user-defined number. Constraint (7) ensures that the length of the pies is greater than the minimum user-defined (i.e. 5 degree in this study). If a pie was selected, the decision variable of the model, x_i, was equal to 1 otherwise 0.

2.5 Evaluation procedure

The performance of the WPFs was evaluated using the root mean square error (RMSE), the mean bias error (MBE) and the correlation coefficient (*r*):

$$\mathrm{RMSE}=\left(\frac{1}{n}\sum_{i=1}^{n}\left(E_i-M_i\right)^2\right)^{0.5} \tag{8}$$

$$\mathrm{MBE}=\frac{1}{n}\sum_{i=1}^{n}\left(E_i-M_i\right) \tag{9}$$

$$r=\frac{\sum_{i=1}^{n}(M_i-M_m)(E_i-E_m)}{\left(\sum_{i=1}^{n}(M_i-M_m)^2\sum_{i=1}^{n}(E_i-E_m)^2\right)^{0.5}} \tag{10}$$

where E_i, and M_i are the predicted and the measured cotton lint yield (Mg ha^{-1}), respectively; n is the number of actual cells in test/validation set; E_m and M_m are the mean of predicted and the mean of actual lint yield (Mg ha^{-1}), respectively.

3. Results

3.1 Principal component analysis

PCA was performed to combine and summarize the variability in the input attributes (Figure 5-3). Cumulative Hotelling's T2 was plotted against the number of PCs; about 96 % was explained by considering 10 PCs, but adding more PCs barely helped explaining more variability. Therefore, it was decided to use the first 10 PCs in the WPFs. The PCs coefficients for the first two PCs (Figure 5-3, panel c) indicated soil texture and water content within the effective root zone followed by soil structure had the highest influence on PC1. On PC2, percent of clay and water content at FC and PWP of the shallow soil played the major role. ECa, irrigation and fertilizer produced medium coefficients for both PCs. The PC1 map to a great extent showed the expected pattern of soil physical and hydraulic properties across field (Figure 5-3b). The PC1 negative values corresponds to soils with lower water holding capacity and available water content which were formed by three regions at the mid-east, mid-south and north west of the field. The PC1 is plotted against PC2 when ECa data were used to color code the points (Figure 5-3, panel d). There was a clear distinction between (i) points with lower ECa (corresponded coarse soil with lower available

water content) which were skewed and clustered towards upper left corner of the figure and (ii) points with higher ECa which were skewed towards upper right corner of the plot. The transition section at the middle of the plot represented soil with medium ECa, corresponding to moderate available water content.

3.2 Performance of WPFs

Figure 5-4 presents the performance of *k*-NN WPFs for a range of design parameters, i.e. *p* and *k*. The trend was identical for both cropping seasons; the error started to decline as number of nearest neighbors increased, reached a minimum at *k*= 4-5 and started to increase again. There was almost no difference in performance of WPFs with different *p* values and *k* less than 4, but the WPFs with lower *p* values exhibited a higher increase in RMSE for *k* greater than 4. In general, the standard deviation of RMSE values consistently decreased with increasing *k* from 1 to 10 regardless of *p* values. The optimum parameters for *k* and *p* were equal to 4 and 1.5, respectively.

Table 5-2 summarizes the performance evaluation statistics for the WPFs. Figure 5-5 illustrates the predicted versus measured lint yield for both cropping seasons using all three WPFs. The *k*-NN exhibited the highest performance and accuracy for both 2013 and 2014 with RMSE equal to 0.131 Mg ha^{-1} and 0.194 Mg ha^{-1}, respectively which was in line with well scattered cloud of points around 1:1 line (Figure 5-5). The ANN (MLR) exhibited 60 % (73 %) and 52 % (65 %) increase in RMSE comparing to the corresponding *k*-NN WPFs in 2013 and 2014 cropping seasons, respectively. The ANN WPFs had a better performance, lower RMSE values and higher *r*, than MLR WPFs for both cropping seasons. The low MBE errors showed there was no systematic under/over estimation for any of the WPFs. However, both ANN and MLR exhibited partial over estimation tendency for plots with yield smaller than 0.40 Mg ha^{-1} and 0.60 Mg ha^{-1} in 2013 and 2014 cropping seasons, respectively. All WPFs showed a slightly better performance in 2014 in comparison with 2013 considering the correlation coefficient and normal RMSE (i.e. RMSE divided by range of yield variation) as the criteria. For instance, ANN WPF had normal RMSE (*r*) equal to 14 % (0.65) and 13 % (0.78) for 2013 and 2014 lint yield data, respectively.

3.3 Irrigation and zoning scenarios

Figure 5-6 (panels a and b) shows the cotton lint yield predicted by *k*-NN WPF for different irrigation scenarios (i.e. uniform irrigation and irrigating each cell individually) over the 2013 and 2014 cropping seasons. Yield was lower in 2013 than in 2014 for all irrigation levels. The predicted yield was slightly different between pivots, yet followed

the same general trend across irrigation levels. This is attributed to differences in soil spatial distribution underneath the pivots. Given uniform irrigation scenarios, in 2013 the highest predicted yield was (for 110 mm seasonal IW) 1.09 ± 0.06 Mg ha^{-1} and 0.98 ± 0.06 Mg ha^{-1} for east and wet pivots, respectively. In 2013 as IW increased, there was a small overall increase in yield up to the maximum level (i.e. IW = 110 mm) but it declined after that for higher IW amounts. If each cell was irrigated at optimum level, yield was improved to 1.19 ± 0.05 Mg ha^{-1} and 1.11± 0.06 Mg ha^{-1} underneath east and west pivots, respectively. In 2014, there were two peaks at 20 and 120 mm IWs. There was a sharp increase in predicted yield for small irrigation amount comparing to rainfed situation in 2014, then it decreased and leveled off for most of the intermediate irrigation levels and increased again at IW=120 mm yet continuously decreased after that. The highest predicted yields (IW = 20 mm) were 1.52 ± 0.12 Mg ha^{-1} and 1.62 ± 0.12 Mg ha^{-1}, underneath the east and west pivots, respectively. When yield in each cell was optimized with different IW the predicted yield increased to 1.71 ± 0.10 Mg ha^{-1} and 1.80 ± 0.09 Mg ha^{-1} underneath the east and west pivots, respectively.

Figure 5-6 (panels c and d) depicts the average irrigation and yield for optimum arrangement of pie shape zones. Yield was always higher in 2014 than in 2013. The yield predicted for 1 pie was equal to uniform irrigation. In both years there was overall constant increase in yield by adding to the number of pies, yet the highest boost occurred when pivot area was divided into two pies. The yield was higher under the east pivot than that under west pivot in 2013, but the opposite was true in 2014. In 2013, as number of zones increased, there was a slight decline in IW corresponding to the optimum yield underneath the west pivot, but IW remained almost constant across a different number of zones for the east pivot. In 2014, as the number of zones increased, higher IW was predicted to get the optimum yield underneath the west pivot and same was true under the east pivot but less pronounced. Compare to rainfed agriculture for the east pivot, the best uniform supplemental irrigation strategy was predicted to boost the yield up to 41 % and 47 % in 2013 and 2014 cropping seasons, respectively. Pie shape zoning added about 4 % to the expected yield in both years. For west pivot, comparing to rainfed, the best uniform supplemental irrigation scenario improved the yield as much as 25 % and 40 % in 2013 and 2014 cropping seasons, respectively. The pie shape zoning added 4 % in 2013 and 7 % in 2014 to the expected yield boost.

Figure 5-7 illustrates the thematic yield maps predicted by *k*-NN WPFs for rainfed scenario as opposed to multiple irrigation strategies (i.e. best uniform irrigation, cutting 4 pies and irrigating each cell individually at optimum level) for both pivots and cropping seasons. Four pies was suggested as optimum to minimize soil spatial

variation underneath pivots in chapter 4. To compare simulated yield zoning with zoning by soil, four zones was selected here and the results from chapter 4 are included in this figure. By visual assessment of the yield maps, the biggest change occurred from rainfed to uniform irrigation followed by the change from pie shape zoning to irrigating each cell individually. In 2014, the spatial yield pattern for the rainfed scenario is similar to soil texture maps and high resolution soil available water maps for the field of study illustrated and modeled in chapters 1 and 2: (i) underneath east pivot coarse-textured regions with low available water content located in the middle and eastern parts also in southwestern portion of the pivot (ii) underneath west pivot a sandy area with low water holding capacity in northern and northwestern of the pivot. In 2013, however, it is hard to match the yield pattern for the rainfed scenario against soil spatial distribution, yet yield patterns under uniform irrigation was somewhat similar to spatial pattern of soil physical and hydraulic properties within the effective root zone. For both pivots and cropping seasons, yield predicted under pie-shape zones exhibited lower minimum but higher maximum than that for corresponding uniform irrigation maps. Moving from pie shape zoning into the hypothetical scenario of irrigating each 25 m^2 cell individually yield improvement mostly occurred for cells with low expected yield. In other words, the maximum predicted yield remained almost unchanged but minimum predicted yield was substantially improved.

Figure 5-8 shows the spatial arrangement of the optimum pie-shape zones, from this study and the previous study in chapter 4. In this study, *k*-NN WPF along with ILP optimization were used to find optimum zones when the objective was to maximize cotton lint yield. In chapter 4 the best spatial arrangement of pie shape zones was found using ILP optimization and soil related attributed without considering the irrigation-yield relationships. Soil available water content, soil ECa, yield data and panchromatic space images (i.e. Landsat 8 band 8) were required when the objective was to minimize soil spatial variation within each zone. The spatial arrangement of optimum pies shape zones to maximize lint yield changed from 2013 to 2014 for both pivots. The 2013 zones for east pivot were somewhat similar to zones by AWC and ECa data. In 2014 for the east pivot, three pies were located on the coarse-textured region with low soil available water content, while a big pie covered rest of the pivot. There were some similarities between 2014 zones for the west pivot and zoning by ECa and AWC data from chapter 4. In 2013, pies in northern part of the west pivot were somewhat similar to pies from AWC data (chapter 4). The spatial arrangement of optimum zones using yield data was different than the configuration of zones in this study.

4. Discussion

4.1 Site-specific water production functions

The *k*-NN performed better than MLR and ANN in predicting cotton yield. It seems even dividing the field into more homogenous regions did the not reduce the complexity enough to get the same level of accuracy by MLR WPFs as achieved by *k*-NN WPF. The ANN WPF showed a better performance than MLR, although the entire data set was used to train the ANN models. This is in line with the reported results by Haghverdi et al. (2014) and Drummond et al. (2003) who found ANN a better option than MLR models to predict yield. However, ANN in this study was not able to provide the same level of performance as *k*-NN WPF, meaning ANN could not entirely handle the huge amount of data and complex relationships between input and outputs. The efficiency of *k*-NN in this study is in line with the promising results reported in the literature: to model soil hydraulic properties (e.g. Nemes et al., 2006, Haghverdi et al., 2015), to estimate cultivar coefficient of crop models (Bannayan, and Hoogenboom, 2009) and for prediction of daily weather data (Bannayan, and Hoogenboom, 2008). In general, linear techniques are not suitable to model yield spatial variability (Sudduth et al., 1996), yet the designed *k*-NN WPF in this study acts as a smart search engine to selects only a few similar data points for prediction. This approach helps *k*-NN WPF to take advantage of numerous data points (cells) provided by yield monitoring systems without necessarily having a good approximation of complex relationships between inputs and yield for all cells across the field of study. The *k*-NN WPF makes use of similarities to distinguish patterns instead of fitting coefficients and this approach could be favorable whenever the form of relationships between input and output of a system is not known in advance (Nemes et al., 2006). One should note that, *k*-NN WPF introduced in this study only provides a pseudo-continuous prediction of yield over a range of applied irrigation. In other words it gives trustable results only in the neighborhood of measured IWs, hence its application for continuous prediction of yield-water relationship is only suggested if IW are applied at multiple levels such that the available data sufficiently cover the desired prediction range.

Soil spatial variation was substantial within the field. Long term analysis of yield maps with different crops revealed a great impact of soil variation on yield at the field of study (chapter 3). Guo et al. (2012) tried to mathematically explain spatial and temporal variation of cotton yield on the Southern High Plains of Texas by means of ECa, soil brightness, and topography data. They observed that a combination of the above attributes could explain up to 70.1 % of cotton yield variability. Guo et al. (2012) mentioned the soil texture as one of the greatest factor affecting cotton yield. They found a relatively stable spatial pattern of yield over time although yield and soil

properties had stronger relationship in dry years than wet seasons. The results of this study also showed a considerable impact of weather condition and rainfall pattern on cotton yield spatial distribution. It was observed that rainfall pattern can mask/change the effect of supplemental irrigation practices on yield. Both years were wet during the experiment when there were 712 mm and 633 mm of rainfall over growing seasons 2013 and 2014, respectively. The total of seasonal applied irrigation water by farmer by east (west) pivot was 133 (93) mm and 65 (50) mm in 2013 and 2014, respectively meaning that irrigation was less than 20 % of precipitation in 2013 and less than 10 % of precipitation in 2014. Wong and Asseng (2006) introduced interactions of seasonal rainfall, plant available soil water storage capacity and N fertilizer applications as dominant factors causing spatiotemporal wheat yield variation. In 2013, wet soil and cold weather forced the farmer to postpone planting cotton. Furthermore, heavy rainfall throughout the growing season caused excessive vegetative growth and larger number of bolls on soil with higher available water content. It seems that the planting delay caused cotton to not be able to accumulate adequate heat-unit and delayed maturity. This condition caused yield reduction, especially over soils with higher water holding capacity, since there was not enough time for all of the bolls to open by harvesting time. This trend affected the yield response to irrigation which in turn caused slightly higher error in corresponding WPFs. Bajwa and Vories (2007) demonstrated that excessive irrigation in wet weather conditions decreased cotton lint yield in Arkansas. That is why the quantity of optimum ET to maximize cotton lint yield is directly proportional to length of cropping seasons; in longer seasons, higher irrigation is expected to produce the highest cotton lint yield (Orgez et al., 1992).

The performance of site-specific WPFs may differ in different conditions, hence should be tested for different crops, climate and soil spatial variation. All attributes affecting crop available water need to be considered for deriving site-specific WPFs. For instance, there are some studies reporting the positive correlation of yield with topographic attributes (e.g. Andales et al., 2007; Green et al., 2007). The field of study was fairly flat, thus topography and slope had negligible influence on yield. However, if topography and slope impact the spatial distribution of water within a field, they should be considered in site-specific WPF derivation.

4.2 Zoning and irrigation strategies

The majority of the farmers in west Tennessee practice uniform irrigation despite the available field-scale soil spatial variation that exists in their fields. Site-specific WPFs can help them to move towards optimum uniform

irrigation strategies for different crops and climate condition. Rainfall was abundant over the course of this experiment but some dry periods occurred when supplemental irrigation was crucial to fulfill ET demand especially for coarse soils with lower AWC. Thus there was a positive cotton lint yield response to irrigation predicted by *k*-NN WPFs under both center pivots in comparison to rainfed. In 2013, delayed planting affected heat unit accumulation/distribution and maturity which in turn suppressed the yield across the field, while yield reduction in soils with higher available water was more pronounced. This was shown in WPFs prediction as a general slight increase in yield as IW increased. Adding more water reduced the drought stress for soils with lower AWC but increased overall yield until a peak point. Analysis of yield data showed that even for coarse-textured soils with low WHC, cotton yield reduction occurred due to over irrigation (chapter 3). That is why, a yield reduction was predicted by *k*-NN WPF for IW of more than 120 mm. In 2014 the overall yield-irrigation relationship across the field was somewhat different than that for 2013, yet closer to what was expected to occur in most of the years. Like 2013, there was a peak at a high irrigation level (IW = 110 mm) which is attributed to the optimum level of IW for soil with lower AWC. Unlike 2013, however, in 2014 there was also a second peak for lower IWs which turned out to be the best uniform cotton irrigation strategy according to *k*-NN WPF predictions. This was recognized as the point of optimum IW level for soils with higher AWC suggesting that in a more typical year like 2014 (a wet year without delayed planting) lower irrigation for the fine textured soils gained more yield compensating for the yield reduction over coarse textured regions due to water stress. More years of data are needed to confirm this as a long term trend.

Given the field-scale spatial soil variation in west Tennessee farmers are likely to depart from uniform irrigation scenarios to some type of VRI. Pie shape zoning is the easier and less expensive irrigation strategy for those who possess center pivot irrigation systems. Zone delineation for irrigation management is complex. Site-specific WPFs could be beneficial to support such conversion from uniform to pie shape zoning. First, they are useful to delineate crop-specific irrigation zones and second to identify the optimum irrigation strategy for each pie. It is likely to still have pies with soil variation and site-specific WPFs would be helpful to figure out which soil to follow in irrigation scheduling. In chapter 4 the focus was on soil properties to delineate irrigation management zones. Here the crop was added to the suggested optimization process with the aid of site-specific WPFs. While there were some similarities between zones in this study and zones developed in chapter 4, the differences were substantial both between years and methods. To a great extent, this difference reflected the inherent spatiotemporal

variation associated with crop growth and yield which in turn governed the spatial configuration of optimum crop-based zones. According to Basso et al. (2009) bias in the evaluation of homogenous management zones may occur due to the impact of weather patterns on both crop growth and interaction with soil type. It was shown that temporal rainfall pattern, which varies from year to year, influences cotton yield-irrigation-soil relationships in west Tennessee. Farmer irrigation management also influences the crop-based zoning result, because for each cropping season the *k*-NN WPF only combines the available observations to make a new predication in that year. Therefore, it is important to look at the zoning system in this study as an iterative process which helps farmers to improve their irrigation decision. The more farmer improves the irrigation management, the better *k*-NN WPF predications reflect yield potential within a field. As a result, longer experimentation is needed to see how the crop-based zoning strategy works across years and how different it would be from the soil based zoning strategy introduced in chapter 4.

Irrigating each cell individually at optimum levels caused a moderate increase in yield in 2014 and a slight increase in 2013 comparing to the best uniform irrigation strategy. In reality, this provides a cap for potential yield boost expected from a VRI for a specific site. Even in 2014, the expected yield boost may not be high enough to compensate for the infrastructure and management expenses to convert to a VRI system. This may differ for different places and even for the same site but on a drier year.

5. Conclusion

Design-based statistics were dominant for much of the last century when contemporary constraints on data collection and processing prevented on-farm experimentation (Pringle et al., 2010). Nowadays, however, most of the farmers in US are collecting numerous site-specific data useful for on-site experimentation. The available data usually have different spatial and temporal resolution. Moreover, the observations are spatially auto-correlated and carry some error and uncertainty. Therefore, in order to convert these data to useful information robust algorithms and models are needed. Some site-specific WPFs were designed and evaluated. The result confirmed that site-specific WPFs could be used as after the fact tools by farmers to enhance irrigation management and scheduling. The WPFs were used to evaluate different irrigation and zoning strategies. Currently, implementation and management of VRI systems demand high capital costs (Evans et al., 2013), hence it is crucial to assess the potential yield gain of a VRI over uniform and pie shape zoning scenarios for each crop and specific site. Pie shape zoning

seems to be a good option as opposed to uniform irrigation because it improves the yield, it is applicable with current center pivots and potentially can enhance water use efficiency and reduce leaching. The result showed that crop based zoning using WPFs produce different zones based on yearly climate condition and farmer management. Data from several years are needed under different climate and irrigation management conditions to derive stable zones. A practical option is to start with soil-based pie shape zoning proposed in chapter 4 and use site-specific *k*-NN WPF to improve both zoning and within zone irrigation decisions in an iterative process. For the study site and over the course of the experiment, the expected yield improvement by a VRI system was not substantially higher than that of pie shape zoning. Note, that this experiment was conducted on two wet years and this result may differ to some extent in a drier condition when more yield increase will likely occur by a VRI comparing to pie shaped zoning.

6. Acknowledgment

This project was funded in part by Cotton Incorporated and USDA NRCS CIG.

References

Abdi, H., and Williams, L. J. (2010). Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2(4), 433-459.

Ai-hua, S., Shi-jiang, Z., Ya-fen, G., and Zhong-xue, Z. (2012). Jensen Model and Modified Morgan Model for Rice Water-Fertilizer Production Function. *Procedia Engineering* 28(0), 264-269.

Andales, A. A., Green, T. R., Ahuja, L.R., Erskine, R.H., and Peterson, G.A. (2007). Temporally stable patterns in grain yield and soil water on a dryland catena. *Agricultural Systems* 94, 119-127.

Bajwa, S. G., and Vories, E. D. (2007). Spatial analysis of cotton (Gossypium hirsutum L.) canopy responses to irrigation in a moderately humid area. *Irrigation Science* 25(4), 429-441.

Bannayan, M., and Hoogenboom, G. (2008). Weather analogue: A tool for real-time prediction of daily weather data realizations based on a modified k nearest neighbor approach. *Environmental Modelling & Software* 23(6), 703-713.

Bannayan, M., and Hoogenboom, G. (2009). Using pattern recognition for estimating cultivar coefficients of a crop simulation model. *Field Crops Research*, 111(3), 290-302.

Basso, B., Cammarano, D., Chen, D., Cafiero, G., Amato, M., Bitella, G., Rossi, R., and Basso, F. (2009). Landscape position and precipitation effects on spatial variability of wheat yield and grain protein in Southern Italy. *Journal of Agronomy and Crop Science* 195(4), 301-312.

Breiman, L. (1996). Bagging predictors. *Machine Learning* 24:123–140.

Dai, X., Huo, Z., and Wang, H. (2011). Simulation for response of crop yield to soil moisture and salinity with artificial neural network. *Field Crops Research* 121(3), 441-449.

De Juan, J.D., Tarjuelo, J.M., Valiente, M., and Garcia, P. (1996). Model for optimal cropping patterns within the farm based on crop water production functions and irrigation uniformity I: Development of a decision model. *Agricultural Water Management* 31(1), 115-143.

Dinar, A., Knapp, K. C., and Rhoades, J. D. (1986). Production function for cotton with dated irrigation quantities and qualities. *Water Resources Research* 22(11), 1519-1525.

Drummond, S. T., Sudduth, K. A., Joshi, A., Birrell, S. J., and Kitchen, N. R. (2003). Statistical and neural methods for site-specific yield prediction. *Transactions of the ASAE* 46(1), 5.

Duncan, H. A. (2012). Locating the variability of soil water holding capacity and understanding its effects on deficit irrigation and cotton lint yield. Master's Thesis, University of Tennessee. http://trace.tennessee.edu/utk_gradthes/1286.

Evans, R. G., LaRue, J., Stone, K. C., and King, B. A. (2013). Adoption of site-specific variable rate sprinkler irrigation systems. *Irrigation Science* 31(4), 871-887.

Fortin, J.G., Anctil, F., Parent, L-É., and Bolinder, M.A. (2010). A neural network experiment on the site-specific simulation of potato tuber growth in Eastern Canada. *Computers and Electronics in Agriculture* 73(2), 126-132.

Green, T.R., Salas, J.D., Martinez, A., and Erskine, R.H. (2007). Relating crop yield to topographic attributes using Spatial Analysis Neural Networks and regression. *Geoderma* 139(1-2), 23-37.

Guo, W., Maas, S.J., and Bronson, K.F. (2012). Relationship between cotton yield and soil electrical conductivity, topography, and Landsat imagery. *Precision Agriculture* 1-15.

Haghverdi, A., Ghahraman, B., Leib, B. G., Pulido-Calvo, I., Kafi, M., Davary, K., and Ashorun, B. (2014). Deriving data mining and regression based water-salinity production functions for spring wheat (Triticum aestivum). *Computers and Electronics in Agriculture* 101, 68-75.

Haghverdi, A., Leib, B. G., and Cornelis, W.M. (2015) A. simple nearest-neighbor technique to predict the soil water retention curve. *Transaction of the ASABE* 58(3), 1-9.

Hedley, C.B., and Yule, I.J. (2009). Soil water status mapping and two variable-rate irrigation scenarios. *Precision Agriculture* 10(4), 342-355.

Huang, Y., Lan, Y., Thomson, S.J., Fang, A., Hoffmann, W.C., and Lacey, R.E. (2010). Development of soft computing and applications in agricultural and biological engineering. *Computers and Electronics in Agriculture* 71(2), 107-127.

Igbadun, H.E., Tarimo, A.K.P.R., Salim, B.A., and Mahoo, H.F. (2007). Evaluation of selected crop water production functions for an irrigated maize crop. *Agricultural Water Management* 94(1-3), 1-10.

Kuang, W., Xianjiang, Y., Xiuqing, C., and Yafeng, X. (2012). Experimental study on water production function for waterlogging stress on corn. *Procedia Engineering* 28, 598-603.

Mathworks. (2014). Statistics Toolbox™ User's Guide (R2014b). Retrieved November 10, 2011 from http://www.mathworks.com/help/pdf_doc/stats/stats.pdf

Mucherino, A., Papajorgji, P., and Pardalos, P. (2009). Data mining in agriculture (Vol. 34). Springer Science & Business Media.

Nemes, A., Rawls, W. J., and Pachepsky, Y. A. (2006). Use of the nonparametric nearest neighbor approach to estimate soil hydraulic properties. *Soil Science Society of America Journal* 70(2), 327-336.

Orgaz, F., Mateos, L., and Fereres, E. (1992). Season length and cultivar determine the optimum evapotranspiration deficit in cotton. *Agronomy Journal* 84(4), 700-706.

Pan, L., Adamchuk, V. I., Martin, D. L., Schroeder, M. A., and Ferguson, R. B. (2013). Analysis of soil water availability by integrating spatial and temporal sensor-based data. *Precision Agriculture* 14(4), 414-433.

Pringle, M. J., Bishop, T. F. A., Lark, R. M., Whelan, B. M., and McBratney, A. B. (2010). The analysis of spatial experiments. In Geostatistical Applications for Precision Agriculture (pp. 243-267). Springer Netherlands.

Saseendran, S. A., Ahuja, L. R., Ma, L., Trout, T. J., McMaster, G. S., Nielsen, D. C., Ham, J. M., Andales, A. A., Halvorson, A. D., Chávez, J. L., and Fang, Q. X. (2014). Developing and normalizing average corn crop water production functions across years and locations using a system model. *Agricultural Water Management* (The Jim Oster Special Issue) 157: 65-77.

Sudduth, K. A., S. T. Drummond, S. J. Birrell, and N. R. Kitchen. (1996). Analysis of spatial factors influencing crop yield. In Proc. 3rd Int. Conf. on Precision Agriculture, 129–140. P. C. Robert, R. H. Rust, and W. E. Larson, eds. Madison, Wisc.: ASA–CSSA–SSSA.

Tong, F., and Guo, P. (2013). Simulation and optimization for crop water allocation based on crop water production functions and climate factor under uncertainty. *Applied Mathematical Modelling* 37(14), 7708-7716.

Wang, Y. R., Kang, S. Z., Li, F. S., Zhang, L., and Zhang, J. H. (2007). Saline water irrigation scheduling through a crop-water-salinity production function and a soil-water-salinity dynamic model. *Pedosphere* 17(3), 303-317.

Wong, M. T. F., and Asseng, S. (2006). Determining the causes of spatial and temporal variability of wheat yields at sub-field scale using a new method of upscaling a crop model. *Plant and Soil* 283(1-2), 203-215.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z., Steinbach, M., Hand, D. J., and Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1), 1-37.

Appendix 5: Chapter 5 Figures and Tables

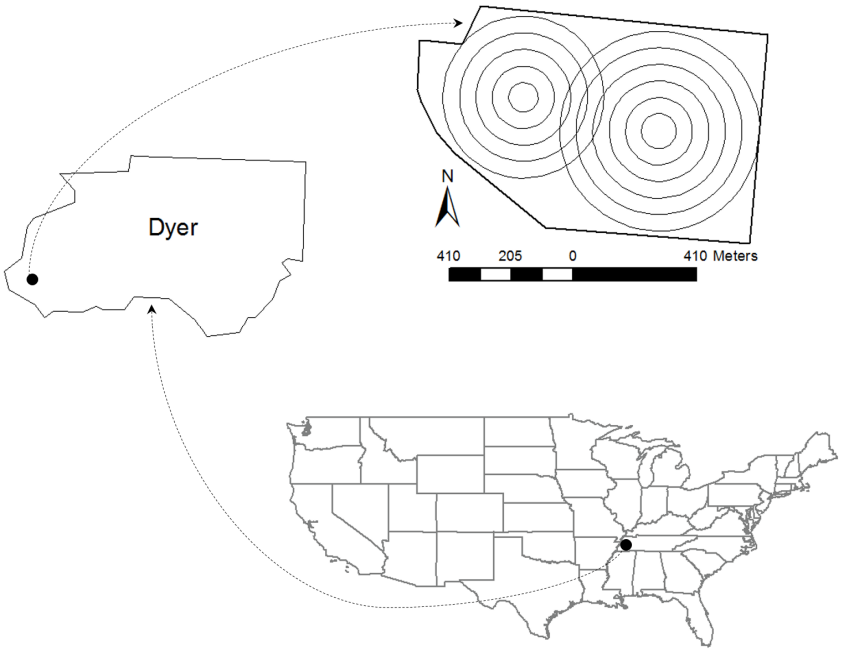

Figure 5-1. Field of study in Dyer County, State of Tennessee.

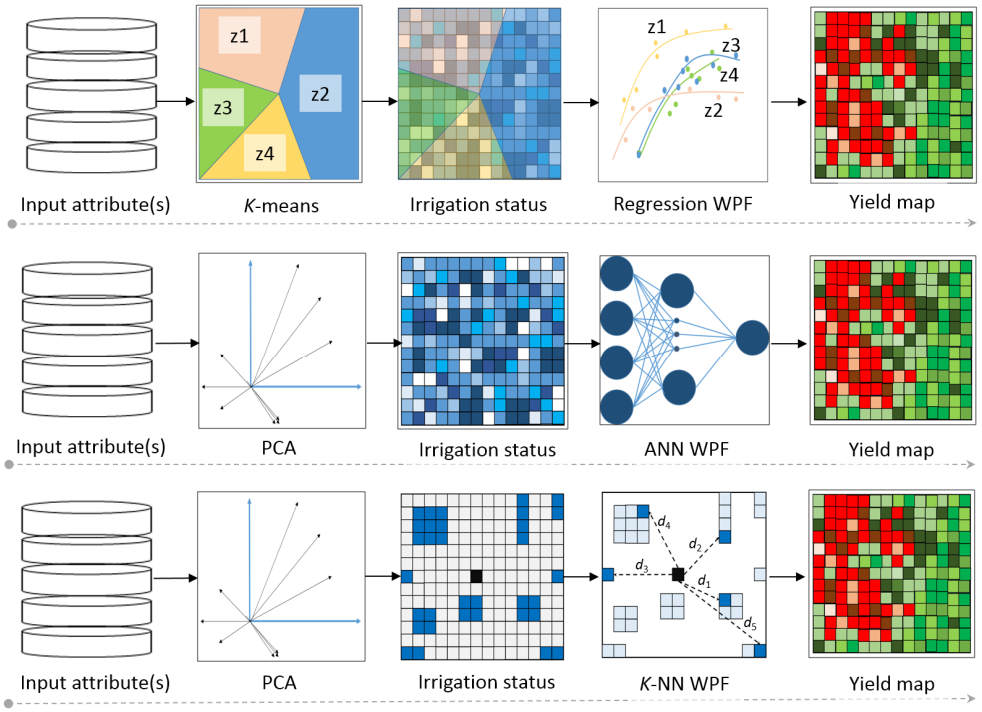

Figure 5-2. Three methods to establish site-specific water production functions. ANN: artificial neural network, *k*-NN: *k* nearest neighbors, PCA: principle component analysis.

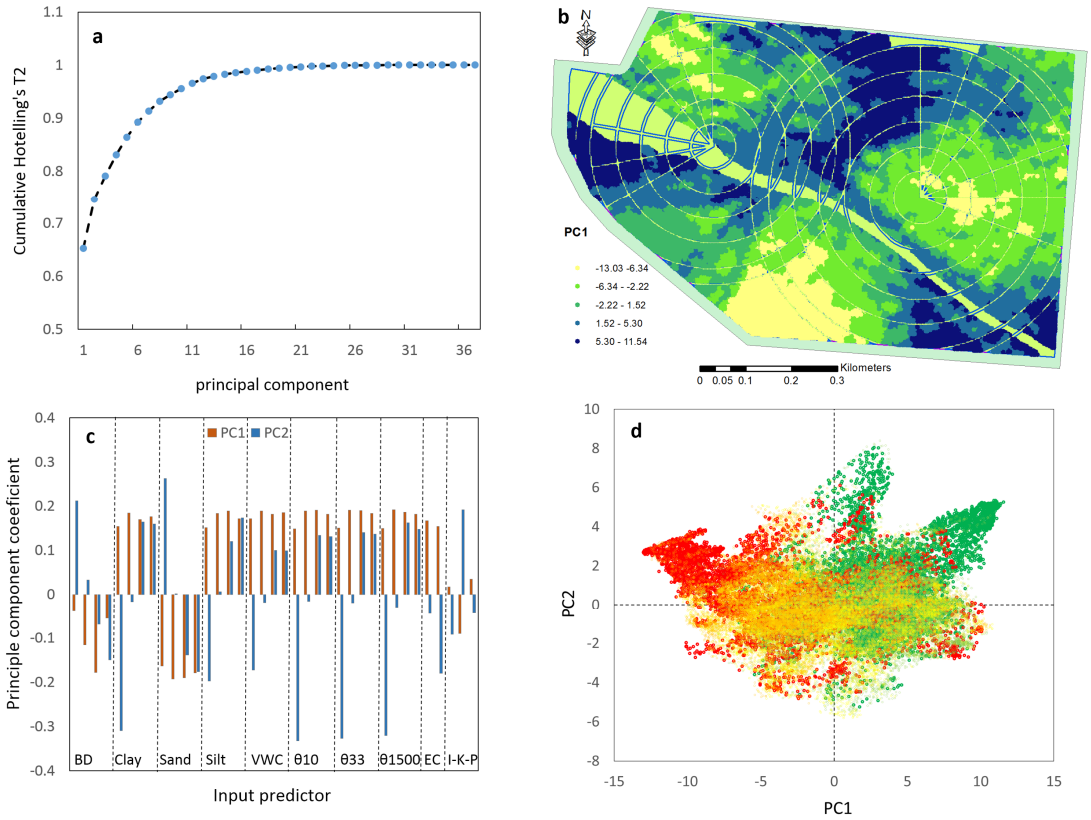

Figure 5-3. PCA result. Panel a: cumulative Hotelling's T2 against PCs. Panel b: thematic map of PC1. Panel c: principle component coefficients for PC1 and PC2. Panel d: Scatter plot of PC1 against PC2.

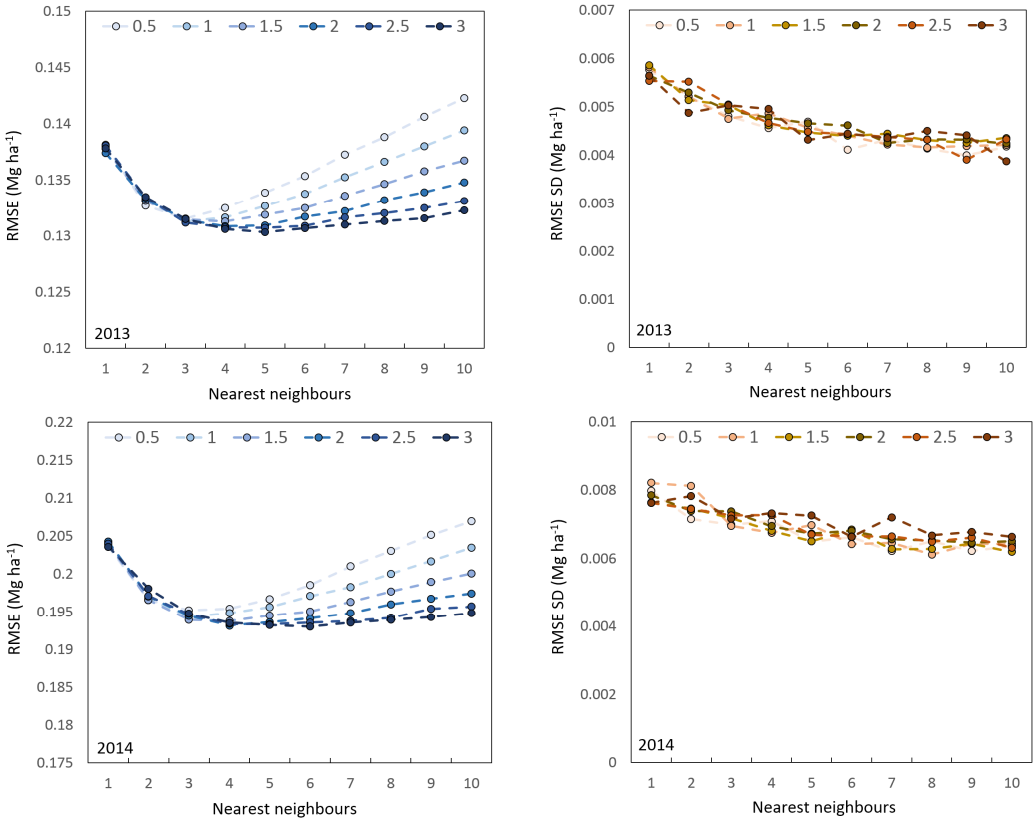

Figure 5-4. The *k*-NN WPFs performance over both cropping season for different combinations of *k* and *p* values.

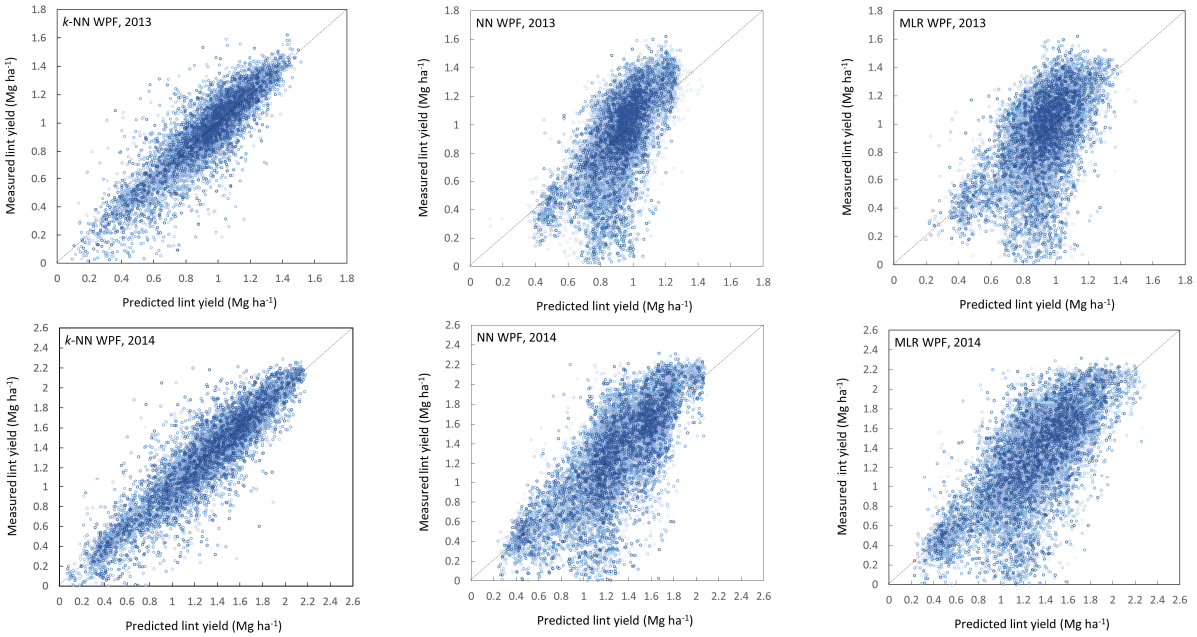

Figure 5-5. Scatter plots of measured versus predicted cotton lint yield using *k*-NN, ANN and MLR WPFs.

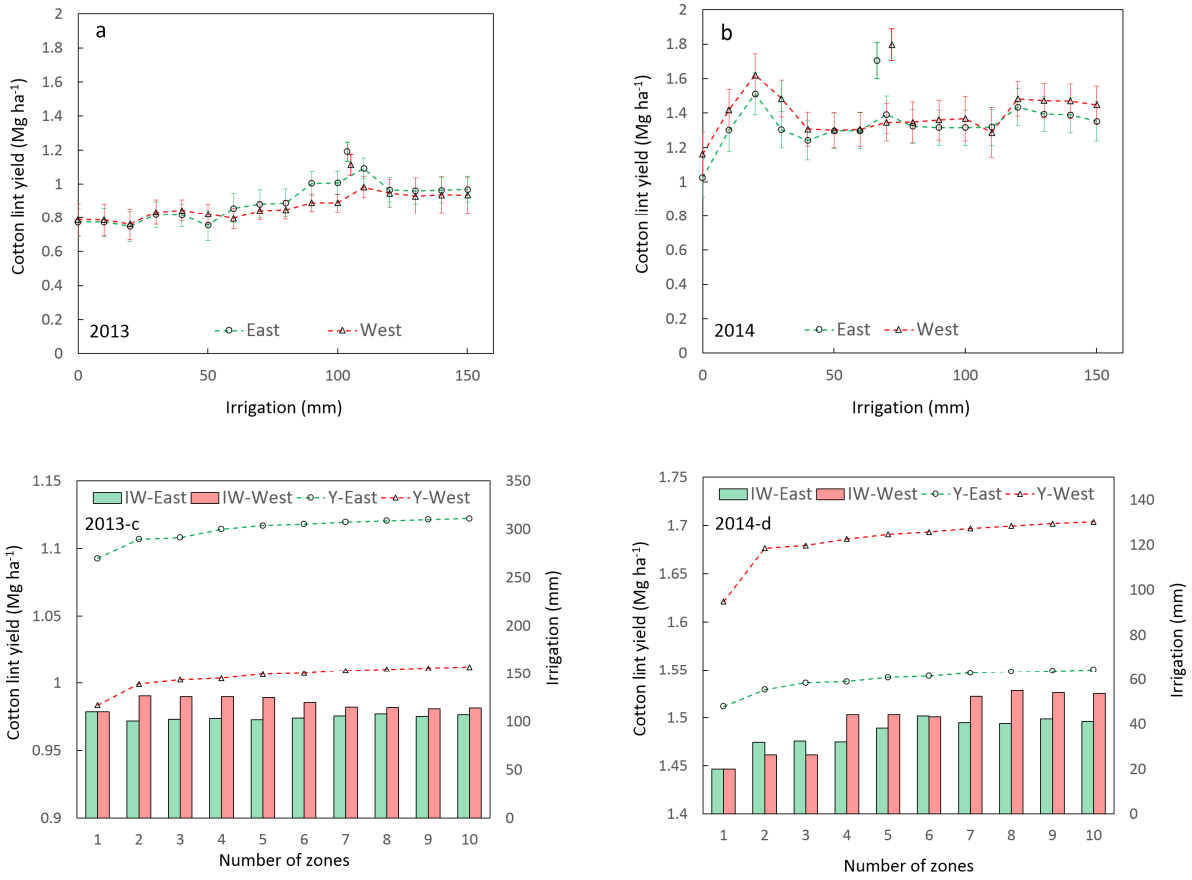

Figure 5-6. Cotton lint yield prediction using site-specific *k*-NN WPF for multiple irrigation scenarios. Uniform irrigation: panels *a* and *b*, dash lines with circles; Pie shape zoning: panels *c* and *d*, dash lines with circles; Irrigating each cell individually: panels *a* and *b*; individual circles. Bars in panels *c* and d are average irrigation applied by each pivot.

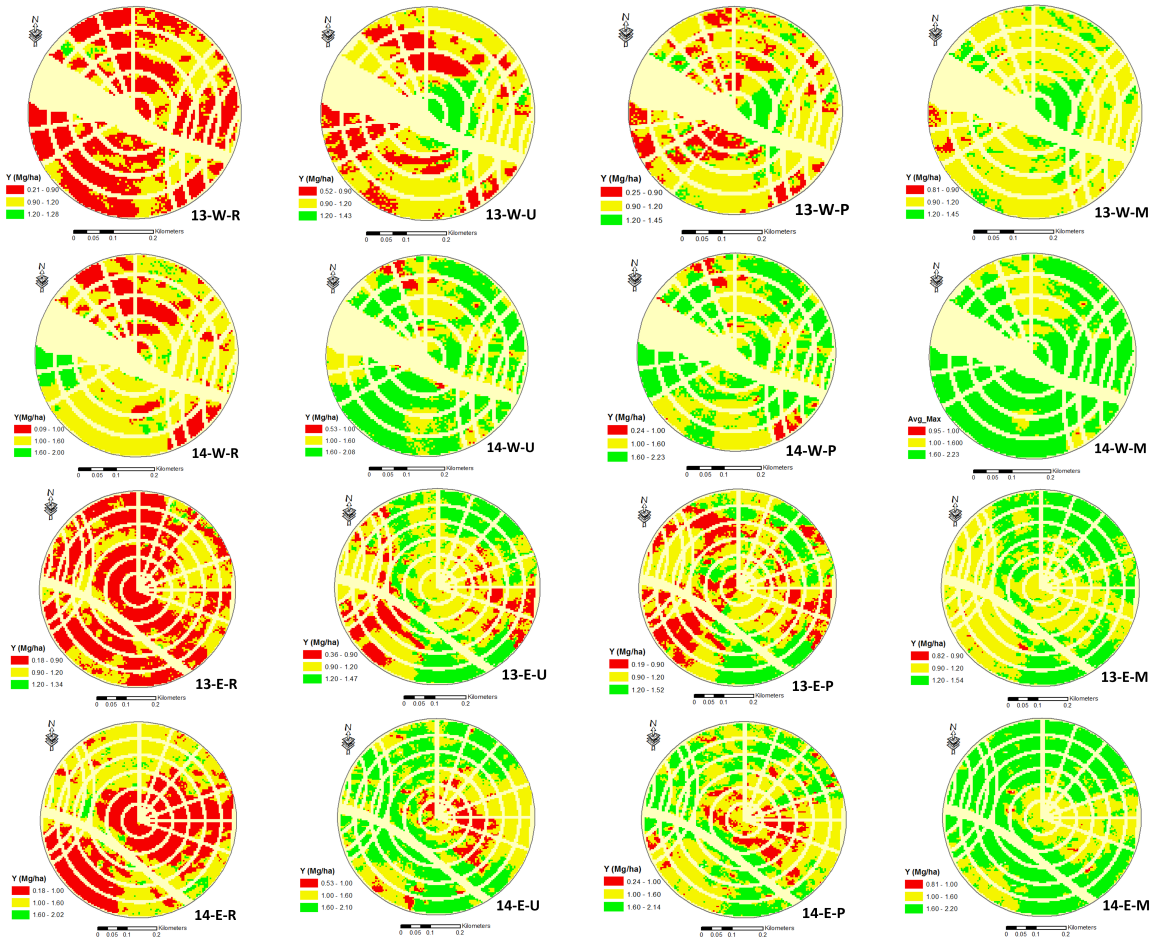

Figure 5-7. Cotton lint yield maps predicted by site-specific *k*-NN WPF under different irrigation scenarios. On the lower right part of each map: the numbers show cropping seasons (2013 versus 2014), the first letter distinguishes between pivots (i.e. E for east and W for west) and the last letter represents the irrigation status (R: rainfed, U: uniform irrigation, P: pie shape zoning (n = 4), M: irrigating each cell individually).

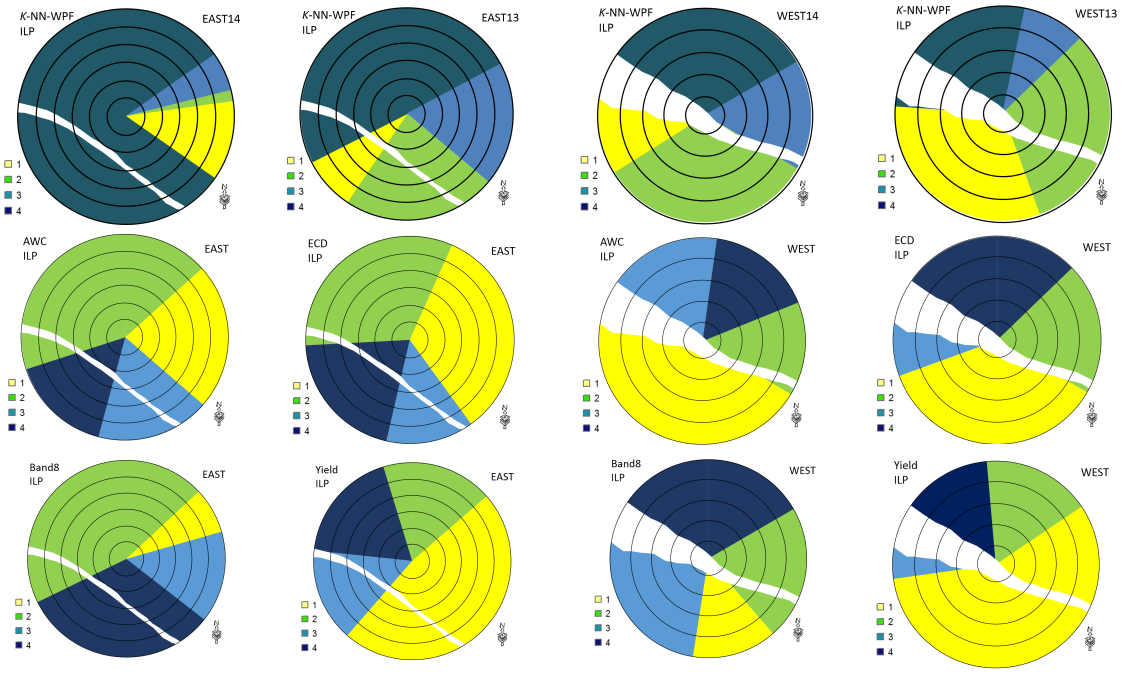

Figure 5-8. Spatial arrangement of the optimum pie shape zones (n = 4). Pies delineated using ILP plus k-NN WPFs were from this study, but Pies delineated using ILP were from chapter 4. Band 8: panchromatic band from space images taken by OLI sensor onboard Landsat 8; AWC: soil available water content; ECD: deep ECa readings.

Table 5-1. Input/output attributes for deriving site-specific WPFs.

Category	*attribute
Soil	ECS, ECD, SSC-1, SSC-2, SSC-3, SSC-4, BD-1, BD-2, BD-3, BD-4, VWC-1, VWC-2, VWC-3, VWC-4, **$\theta_{10\text{-}1}$, $\theta_{10\text{-}2}$, $\theta_{10\text{-}3}$, $\theta_{10\text{-}4}$, $\theta_{33\text{-}1}$, $\theta_{33\text{-}2}$, $\theta_{33\text{-}3}$, $\theta_{33\text{-}4}$, $\theta_{1500\text{-}1}$, $\theta_{1500\text{-}2}$, $\theta_{1500\text{-}3}$, $\theta_{1500\text{-}4}$
Crop	Yield 2013, Yield 2014
Water	IW 2013, IW 2014
Fertilizer	P, K***

* ECS: shallow ECa (mS m^{-1}), ECD: deep ECa (mS m^{-1}), SSC: sand, silt and clay (%), BD: bulk density (g cm^{-3}), VWC: volumetric water content at the time of sampling (cm^3 cm^{-3}), θ: water content predicted by PTFs and interpolated by co-kriging at soil matric potentials equal to -10, -33 and -1500 kPa (cm^3 cm^{-3}), IW: seasonal irrigation water (mm), P and K: phosphorus and potassium, respectively.

** The soil related inputs were measured and predicted for four layers (i.e. 0-25 cm, 25-50 cm, 50-75 cm, 75-100 cm) covering crop effective root zone.

*** The farmer practiced variable rate P and K, but nitrogen was applied uniformly.

Table 5-2. Performance of *k*-NN, ANN and MLR WPFs for 2013 and 2014 cropping seasons.

	2013			2014		
	r	RMSE*	MBE	*r*	RMSE	MBE
k-NN - WPF	0.88	0.131	0.002	0.91	0.194	0.001
ANN - WPF	0.65	0.209	0.000	0.78	0.295	0.000
MLR - WPF	0.55	0.227	0.000	0.73	0.321	0.000

* RMSE and MBE are in Mg ha^{-1}.

Part 5: Conclusion

Irrigation has been expanding across the humid areas of the Cotton Belt in the US for the last 20 years (Perry and Barnes, 2012). Stabilizing yields and the recent high commodity prices are the prime reasons for this growth in irrigation investment. There are two inherent difficulties for irrigation management in west Tennessee: (i) field-scale spatial soil variation and (ii) season-to-season variability in rainfall patterns. It was shown through a multi-year experiment in west Tennessee that optimum cotton irrigation was significantly different among the plots with different soil types (Duncan, 2012). Given temporal weather changes along with soil spatial variation, a variable rate irrigation scenario is needed to improve the efficiency of energy and water use as well as to increase sustainability of row crop agriculture in this region.

Design-based statistics were dominant for much of the last century when contemporary constraints on data collection and processing prevented on-farm experimentation (Pringle et al., 2010). Irrigation studies have been traditionally reserved for experimental farms with small plots when design-based statistics are used to analyses the data. This approach relies on transformation of the research findings from small plots to a farmers' fields. In reality, each field has its own complexity and needs a site-specific irrigation management plan. Nowadays most of the farmers in US are collecting numerous site-specific data useful for on-site experimentation. Statistical analysis of such data may be the key to move toward site-specific irrigation management. The available data usually have different spatial and temporal resolution. Moreover, the observations are spatially auto-correlated and carry some error and uncertainty. Therefore, robust algorithms and models are needed in order to convert these data to useful information.

The main objective of this study was to design a dynamic data analysis and modeling system to move toward optimum site-specific irrigation strategies in an iterative process when farmers learns from each year's results. Several site-specific models were developed and tested (i) to generate high resolution field-scale soil available water map within the effective root zone, (ii) to investigate cotton lint yield-irrigation relationships across different soil types, (iii) to delineate irrigation management zones and to (iv) to assess the performance of different irrigation scenarios from uniform to variable rate irrigation.

5-1 Root zone soil hydrology

The objective of first part of this study was to explore and model soil physical and hydraulic variation in one type of agricultural fields in west Tennessee where elevation and slope differences were minor and soil texture dictated the majority of spatiotemporal changes of water content. Analyzing yield against soil maps revealed that soil physical and hydraulic properties to a great extent influenced yield patterns. This is expected because plant available water is a function of soil water holding capacity. This suggests variable rate irrigation as the appropriate irrigation scenario for this typical form of fields in west Tennessee.

There are many methods in the literature providing non-spatial point estimation of soil water retention, yet the options are more limited when the goal is to produce a high resolution soil hydraulic property map at the field-scale. The result showed that incorporating pseudo-continuous pedotransfer functions (Haghverdi et al., 2012, 2014) and cokriging could be a reliable method to predict a soil water retention curve in the field-scale where drastic soil spatial variation is available. Given the result of this study, a combination of deep core sampling and ECa data is suggested to generate high resolution soil available water content map within the effective root zone. The sampling scheme was designed considering geostatistical principals. ECa maps may be used in advance to guide sampling scheme and perhaps reduce the number of samples. Further investigation is needed to find optimum sampling scheme in order to minimize the cost while maintaining the desired level of accuracy.

ECa turned out to be a useful proximal attribute to (i) understand spatial variation of alluvial soils and (ii) for converting point estimations of water retention to high resolution maps in west Tennessee. There is evidence in the literature showing ECa may not be as effective for some conditions. For instance, there is another typical form of agricultural fields in west Tennessee with totally different conditions where soil textural differences are minor yet slope and elevation differences affect infiltration and redistribution of water within root zone hence available water for plants. The findings of this study may not be transferable to this type of fields. A firm understanding of soil characteristics affecting soil ECa is needed for each site prior to applying ECa in a modeling process.

5-2 Cotton response to supplemental irrigation

A 2-year (2013-2014) on-farm experiment was conducted to understand the spatial and temporal dynamics of supplemental irrigation on cotton yields. Overall, irrigation improved yield in comparison to rainfed throughout this study. However, it was shown that cropping season length, rainfall pattern and heat unit accumulation /distribution

affected both cotton growth and development changing or even reversing the expected lint yield from an irrigation treatment for a specific soil type. While soil variation is inherent and not controlled by farmers, irrigation, if well scheduled, could be the key factor to orchestrate the whole cropping system toward an optimum condition. Within this study, it was demonstrated how site-specific information collected by on-the-go sensing, remote sensing and wireless network of sensors could help farmers manage irrigation. In addition, analysis of such information throughout the growing season provides insight to potential yield patterns, thus helps farmers to modify other inputs allocation. Cotton responded differently to irrigation across soil types suggesting that variable rate irrigation (i.e. precision irrigation systems or speed control panels) would probably be beneficial for the field of study.

5-3 Irrigation strategy and zone delineation

The performance of several clustering methods for zone delineation were evaluated. In addition, a new zoning strategy based on integer linear programming was designed for center pivots with limited speed control ability. High resolution soil available water maps were considered as standard input to zone the field. The effectiveness of ECa, space-borne satellite images and yield data for zone delineation were also assessed. The clustering methods performed similar in efficiently dividing the field into relatively homogenies zones in respect to soil hydraulic properties. The introduced integer linear programming method offers the optimum zoning strategy for center pivots with limited speed control ability to match water input to soil spatial patterns. However, further research is essential to investigate the optimum irrigation strategy over pies with a mixture of soils.

The results of this study suggested that ECa and satellite images may be used to determine site-specific irrigation management zones in west Tennessee if a high spatial similarity is observed between these ancillary data with soil hydraulic properties. The spatial and temporal resolution and precision varies among different types of ancillary data and affect their efficiency for zone delineation. Temporal variability in soil moisture altered expression of spatial variability in space-born satellite images. Yield data did not show a high potential for zone delineation. This is attributed to various factors that affect yield in a complex manner and also temporal differences in yield data. In practice, farmer knowledge of field conditions could be extremely useful to choose the most appropriate data set for zone delineation.

Currently, implementation and management of variable rate irrigation systems demand high capital costs (Evans et al., 2013), hence it is crucial to assess the potential yield gain of a VRI over uniform/pie shape zoning

irrigation scenarios for each crop and specific site. Pie shape zoning seems to be a good option as opposed to uniform irrigation because it improves the yield, is applicable with current center pivots and potentially can enhance water use efficiency and reduce leaching.

The efficiency of different irrigation strategies were evaluated using site-specific water production functions (WPFs). Different models were designed and evaluated and *k*-NN (i.e *k* nearest neighbors) technique exhibited the highest accuracy. The result confirmed that site-specific WPFs could be used as after the fact tools by farmers to enhance irrigation management and scheduling. The WPFs were also employed to test a crop-based zoning method. The integer linear programing zoning method was modified such that the objective was to maximize the overall yield throughout the field. The result indicated that crop-based zoning using WPFs produce different zones based on climate condition and farmer management. It seems data from several years are needed under different climate and irrigation management conditions to derive stable zones. A practical option is to start with soil-based pie shape zoning and later on use site-specific WPFs to improve both zoning and within zones irrigation decisions in an iterative process. For the study site and over the course of the experiment, the expected yield improvement by a variable rate irrigation system was not substantially higher than that of pie shape zoning. Note, that this experiment was conducted in two wet years and this result may differ to some extent in a drier condition when more yield increase will likely occur by a variable rate irrigation system comparing to pie shape zoning.

References

Duncan, H. A. (2012). Locating the variability of soil water holding capacity and understanding its effects on deficit irrigation and cotton lint yield. Master's Thesis, University of Tennessee. http://trace.tennessee.edu/utk_gradthes/1286.

Evans, R. G., LaRue, J., Stone, K. C., & King, B. A. (2013). Adoption of site-specific variable rate sprinkler irrigation systems. *Irrigation Science* 31(4), 871-887.

Haghverdi, A., Cornelis, W. M., & Ghahraman, B. (2012). A pseudo-continuous neural network approach for developing water retention pedotransfer functions with limited data. *Journal of Hydrology* 442, 46-54.

Haghverdi, A., Öztürk, H. S., & Cornelis, W. M. (2014). Revisiting the pseudo continuous pedotransfer function concept: Impact of data quality and data mining method. *Geoderma* 226, 31-38.

Perry C. and Barnes E., (eds.) 2012, Cotton irrigation management for humid regions. Cotton, Inc. Cary, N.C.

Pringle, M. J., Bishop, T. F. A., Lark, R. M., Whelan, B. M., and McBratney, A. B. (2010). The analysis of spatial experiments. In Geostatistical Applications for Precision Agriculture (pp. 243-267). Springer Netherlands.

Vita

Amir Haghverdi was born in Torbatejam, Iran in 1983. He earned his first PhD degree in Irrigation Engineering in Iran. He received his second Ph.D. in the Department of Biosystems Engineering and Soil Science at the University of Tennessee, Knoxville, where he expanded his research to precision agriculture. His primary subject is optimizing site-specific irrigation for cotton through remote sensing, GIS and GPS technologies, on-the-go sensors, and site-specific wireless sensing systems. Amir's research is in agricultural water management with irrigation engineering, soil hydrology, and spatiotemporal data mining as the main components. He has been collaborating with scientists in Belgium, Turkey, Spain, Germany, and Iran. He received the outstanding graduate student award from the International Society of Precision Agriculture in 2014. In addition, he was selected as the graduate student with professional promise by the faculty of Biosystems Engineering and Soil Science Department, UTK in 2015. Amir accepted a tenure-track assistant professor position in the Department of Biological Systems Engineering at University of Nebraska-Lincoln. Amir's goal is to conduct an integrated research-extension plan that incorporates computer modeling and laboratory/field experiments.