**Studying uniform and variable rate center pivot irrigation strategies with the aid of
site-specific water production functions**

Amir Haghverdi^{1,2*}, Brian G. Leib^{2*}, Robert A. Washington-Allen³, Michael J. Buschermohle²; Paul D. Ayers²

¹ Department of Biological Systems Engineering, University of Nebraska-Lincoln, Panhandle Research and Extension
Center, 4502 Avenue I, Scottsbluff, NE 69361-4939, United States.

²Department of Biosystems Engineering & Soil Science, University of Tennessee, 2506 E.J. Chapman Drive, Knoxville,
TN 37996-4531

³Department of Geography, University of Tennessee, Burchfiel Geography Building, Knoxville, TN 37996-0925

* Correspondence author, ahaghverdi2@unl.edu, (308) 632-1246

Abstract

Irrigation management has evolved into a top priority issue since available fresh water resources are limited. 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 (k -NN), multiple linear regression, and neural networks), (ii) simulating yield maps for uniform, sector control VRI, and zone control VRI center pivot systems using the site-specific WPFs, (iii) using the best WPF to investigate different cotton irrigation and zoning strategies using integer linear programming, and (iv) comparing soil-based and WPF-based zones for sector control VRI systems. 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 -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, respectively. The result indicated that variable rate irrigation with limited sector control capability could enhance cotton lint yield under supplemental irrigation when field-level spatial soil heterogeneity 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. We believe site-specific WPFs are useful empirical tools for on-farm irrigation research.

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

1. Introduction

1.1 *Variable rate irrigation*

The demand for food and fiber is increasing as the world's population grows. Irrigation management has evolved into a top priority issue since available fresh water resources are limited. Row crop supplemental irrigation is rapidly expanding in the moderately humid region of west Tennessee because of a trend of stabilizing yields by irrigation and high commodity prices. However, the soils of west Tennessee exhibit high spatial and temporal variability of soil characteristics that relate to moisture availability for crop production at the field-level. 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). Consequently, variable rate irrigation (VRI) is recommended as the appropriate irrigation scenario to address field-level spatiotemporal heterogeneity and thus how, when, and where to irrigate (Duncan, 2012). VRI is expected to improve water-use efficiency, increase productivity, save fuel and decrease nutrient leaching (Pan et al., 2013). Several authors have stressed the potential water conservation that has been observed using VRI systems (Evans et al., 1996; Duke et al., 1997; Sadler et al., 2005, Hedley et al., 2009). For instance, Hedley et al. (2009) compared VRI and conventional uniform irrigation methods and found that 9–19 % of irrigation water was saved, which in turn reduced nitrogen leaching. The state-of-the-art zone control VRI systems, allow farmers to adjust irrigation applications across fields, by pulsing individual sprinklers or banks of sprinklers on and off. The sector control VRI systems cut pie shape zones simply by adjusting the revolution speed of the center pivot between given pairs of angles. 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. Recently, Haghverdi et al (2015b) used integer linear programming (ILP) to find the optimum number and spatial arrangement of irrigation management zones for center pivot systems with limited sector control capability. They used soil properties to delineate management zones, yet suggested to investigate crop response to water amount across different soil types as a means to provide crop-specific irrigation management zones.

1 In practice, the level of irrigation management being followed by the majority of producers is relatively
2 low, while experience with successful zone control VRI management is rare among producers and service
3 providers. Initially, the larger price tag, accompanied with the fear of complexity that is associated with using
4 zone control VRI, are the two main stumbling blocks experienced by farmers. The sector control VRI
5 systems require less capital investment and the operation/maintenance is also easier than that of zone control
6 VRI. Most of the center pivots that farmers use today are equipped to allow a limited number of pie shape
7 sectors to be made, without any additional control panel upgrades. The knowledge and full use of this
8 capability is the perfect and most appropriate step in bridging the gap between current practice and cutting-
9 edge zone control technology (Haghverdi et al., 2015b). Currently, it is not easy to quantify the potential
10 production benefits of zone control and sector control VRI systems over conventional uniform irrigation
11 systems in a given field, further contributing to producers' present hesitancy and reluctance, with regards to
12 investing in VRI systems.

13 However, before the benefits of VRI can be achieved, it is necessary to identify the spatial heterogeneity
14 of soil within a given field and delineate homogenous management zones wherein identical inputs of water
15 can be applied. In practice, the number of zones depends on the target input, available equipment and the
16 crop(s) planted. However, a corollary expectation is that varying the inputs within the management zones will
17 facilitate optimal yields. This desire leads to the development of water production functions (WPF) where
18 crop yield is mathematically related to the amount of irrigation water.

19 20 1.2 *Water production functions*

21 The term production function (PF) may be assigned to any mathematical relationship between crop yield
22 and input components such as water, fertilizers and energy (De Juan et al., 1996). In practice, almost all the
23 derived PFs require crop water use as an independent variable. PFs predict total dry matter (or marketable
24 product of each crop) as a dependent variable, while the independent variables are transpiration,
25 evapotranspiration (ET) or the amount of applied water during irrigation (IW). The PFs are categorized into
26 crop water production functions (CWPFs) and water production functions (WPF) that use ET and IW as

independent variables, respectively. IW may consist of different components such as the crop water requirement, pre-planting irrigation, leaching requirement, and rainfall (Igbadun et al., 2007).

Recently, some studies have aimed to revisit and rebuild the concept of PF including the development of a WPF for water logging stress on corn (Kuang et al., 2012) and a rice water – fertilizer PF (Ai-hua et al., 2012). Tong and Guo (2013) included 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 (Ahuja et al., 2000) and historical weather data to develop average corn CWPFs across years and locations. Sadler et al. (2002) applied WPFs at field level and observed significant differences in corn response to irrigation within and across soil map units. Holan et al. (2008) formulated spatially dependent WPFs that later were used by Stone and Sadler (2015) who reported high accuracies for the spatial modeling of corn yield. Wang et al., (2007) derived cotton and wheat water-salinity PFs. Dinar et al. (1986) derived cotton CWPF under saline conditions in California 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. The behavior of complex ecological systems is non-linear, but most PFs are based on easily built linear regression-based equations and are thus not particularly powerful representations of these systems (Dai et al., 2011; Haghverdi et al., 2014b).

However, recent studies have investigated the use of more robust and non-linear techniques to predict yield and model complex systems, particularly powerful empirical algorithms derived from machine learning data mining tools (e.g. Fortin et al., 2010; Haghverdi et al., 2014b).

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 if enough data is input to an ANN model, 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. This resulted in ANN models having higher precision than regression for modeling relationships between crop yield and soil moisture and salinity at different crop growth periods (Dai et al. 2011). Haghverdi et al. (2014b) derived some novel WPFs and compared data mining-based methods with traditional regression procedures. They utilized ANN and decision tree modeling algorithms to derive water salinity PFs for spring wheat and found the results promising.

1.3 *Site-specific water production functions & variable rate irrigation*

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 environments following the statistical principles behind design of experiments. However, there are some shortcomings. Given the time consuming and labor intensive nature of field experiments, this approach becomes unrealistic for the near future (Drummond et al., 2003). In addition, this method relies on scaling and transforming small plot research findings to operational management of farmers' fields. In reality, each field has its own spatial and temporal heterogeneity of main attributes and this complexity requires site-specific irrigation management solutions. An alternative option is to simulate the behavior of the fields using mechanistic crop growth models. However, the drawback of these models is that they are time-consuming and expensive to develop because of the numerous inputs that need to be collected to run them (Drummond et al., 2003).

Precision agriculture (PA) enables farmers to collect numerous site-specific data. Empirical analysis of these data may be the key to aid farmers in 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, and is 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 in providing a comprehensive understanding of irrigation-yield relationship for different crops across different soil types.

Robust algorithms are needed for identifying site-specific crop yield responses to environmental and management parameters and helping to find either optimum practices or alternative scenarios. The increased cost and complexity of variable rate irrigation (VRI) systems compared to conventional systems are the main obstacles barring widespread adaptation by farmers. VRI related research projects have mainly focused on (i) designing sprinklers to spatially vary irrigation rates and (ii) developing software/hardware for guiding the VRI system (Pan et al., 2013). There thus remains by necessity, novel models and techniques that simplify irrigation zone delineation, and provide an estimation of attainable production benefits of VRI systems, the main goal of this study.

The objectives of this study are to:

- (i) Use multiple data mining techniques to design customized site-specific WPFs for cotton.
- (ii) Determine the best site-specific WPF for cotton.
- (iii) Simulate yield maps for uniform, sector control VRI, and zone control VRI center pivot systems using the site-specific WPFs.
- (iv) Formulate an optimization problem for sector control VRI systems using integer linear programming and water production functions to maximize yield.
- (v) Compare the spatial arrangement of the optimum soil-based and WPF-based irrigation management zones for sector control VRI center pivot systems.

2. Material and Methods

2.1 *Field of study*

A 73-ha cotton field that contained two center pivots for irrigation within which the resident farmer practiced uniform nitrogen (N) application and variable rate potassium and phosphorus application was selected as the study area (Figure 1). The field is located in southwestern Dyer County in west Tennessee within a moderately humid region along the Mississippi river. The fields are Mississippi river terrace alluvial deposits that have produced Robinsonville loam and fine sandy loam, Commerce silty clay loam, and Crevasse sandy loam soils (Figure 1). The mean monthly growing season precipitation and temperature is 97

mm month⁻¹ and 21 °C from May to November, respectively. Soil sampling and modeling of the fields different soil constituents, particularly soil available water content, showed substantial spatial heterogeneity and as much as a four-fold increase in the variability of available soil water content within the effective root zone across the field (Haghverdi et al., 2015c). Given the spatial variability of these soil properties, this field was an ideal site for a study of the spatial optimization of a VRI system.

2.2 Irrigation experiment

A two-year experiment (2013-2014) was implemented to study the irrigation-cotton lint yield relationship across the different soil types. The field was no-tillage planted to cotton ('PHY375' variety) on the 30th of May, 2013, and to cotton ('Stoneville 4946' variety) on the 5th of May, 2014. Crop pest management was implemented following state extension recommendations. In the first year, the field was harvested on December 2 -3, 2013 and in the 2nd year it was harvested on October 18-20, 2014.

There is some overlap between the two center pivots that cover the majority of the field (Figure 2). Irrigation scheduling was carried out by the producer since the optimum practical goal was to enhance knowledge and expertise of the farmer over the course of the experiment. We provided sufficient information for the farmer to irrigate appropriately using a wireless network of soil sensors installed throughout the field of study, as well as an on-site weather station to estimate ET. Additionally, the Management of Irrigation Systems in Tennessee (MOIST) program (<http://bioengr.ag.utk.edu/weather/>) was employed to assist the farmer with irrigation scheduling. MOIST is an irrigation decision support tool that delivers irrigation recommendations that consider both ET and soil water status. We simulated spatial variability in water application across different soil types (Figure 2) by programming the pivots and partially swapping the sprinklers. The pivot's control panels were Valley Select2 module (Valmont Industries, Inc) that were programmable up to 9 sectors, thus allowing one to vary the location, frequency, and amount of irrigation across the field by changing the travel speed of the irrigation system. Consequently, the sprinklers were swapped to apply approximately 50 % more/less irrigation water (IW) than the farmer's decision. The IW on sector zones was between 25-30 % higher/lower than the IW that was set by the farmer.

2.3 Input attributes and data preparation

Table 1 summarizes the data used in this study to derive the WPFs. The field was sampled on March 21 and 22, 2014 where 400 undisturbed **core soil** samples from 100 sampling locations at 4 different depths were collected to measure soil texture, soil water content, and bulk density. Then, the soil water retention curves were predicted for samples using pseudo-continuous pedotransfer functions (Haghverdi et al., 2014a) and interpolated to generate high-resolution maps. 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 and modeling is provided by Haghverdi et al. (2015c). The cotton lint yield maps were obtained from the yield monitors. A quality assessment of the raw cotton lint yield data was conducted to remove outliers and bad data. The study field was gridded into 5-m $\times$ 5-m cells or pixels (25m² area per cell) and for each cell, all attributes in Table 1 (inputs and outputs) were obtained. Ordinary kriging was used for interpolation. Empirical semivariograms were calculated using the Geostatistical Analyst toolbox in ArcGIS 10.2.2 (ESRI Inc., Redlands, California) considering 12 lags and an omnidirectional stable model. A total of 16,000 cells were prepared after removing cells with missing/bad data. Since neither pivot was shut off in the overlap zone, the water applied in that zone was considered to be the summation of the applied water by the both pivots.

In many situations, there are only a few major driving forces that govern a system. As the number of input predictors increases, the likelihood of redundancy or high correlation between variables increases with more than one attribute capable of explaining the behavior of the system. Principal component analysis (PCA) is a statistical method that is used to reduce the dimensionality or the number of high correlations in a given data set. PCA finds hidden patterns within a data set and extracts the important information from the raw data and replaces it with 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, we used Matlab 2015a (MathWorks,

Inc., Natick, Mass.) to perform PCA. The Hotelling's T^2 , 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

Phase 1: Establishing site-specific water production functions

Multiple Linear Regression water production function, MLR WPF

Figure 3 illustrates the three methods that were developed to establish site-specific WPFs. Traditionally linear regression is used to derive WPFs that typically result in a curvilinear representation of water-yield relationship. We followed the recommendations of Haghverdi et al. (2015b) and used k -means cluster analysis to divide the data into 4 homogenous soil property sub-groups. Then, for each sub-group, a separate MLR-WPF was developed considering IW, IW² and IW×PCs as input predictors. The ten most important PCs were used as input for clustering.

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 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 agricultural studies (Mucherino et al., 2009). The k -NN considers a distance function, such as the Euclidean distance, to find so-called closer observations with higher similarity. Then, it combines the response of the k nearest surrounding observations to predict the response for new observations.

Figure 3 provides a visual step-by-step guide to deriving the site-specific k -NN WPFs, where first the raw inputs (Table 1) were converted to PCs. Then, the first 10 PCs and the IW were converted to *input predictors* with a zero mean and a standard deviation of one, and then the ranges of input predictors were

tuned following the method recommended by Nemes et al. (2006). 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):

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

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, a 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. Haghverdi et al., (2015a)

recommended that the number of ensemble values be 40. The modeling was done using Matlab R2014a (MathWorks, Inc., Natick, Mass.).

Artificial Neural Network water production function, ANN WPF

The past performance of artificial neural networks (ANNs) suggests that they have potential to establish reasonable WPFs (Haghverdi et al., 2014b). A three-layer perceptron model was selected for this study. SPSS Modeler 15 (SPSS Inc., Chicago, IL) automatically computed the best number of neurons to be used in the hidden layer to derive ANN WPFs. The first 10 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.

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 increments. The optimum cotton lint yield maps were delineated for the three irrigation strategies including (i) uniform irrigation, (ii) VRI with limited sector control capability and (iii) a hypothetical situation in which each cell (5×5 m²) to be irrigated individually at optimum level.

We adopted and modified the integer linear programming (ILP) optimization process developed by Haghverdi et al (2015b) to find the optimum arrangement of management zones to maximize cotton lint yield for VRI systems with limited sector control capability. Each center pivot irrigation system was considered to consist of 360 sectors each 1 degree wide. A total of 12,241 possible zones were obtained by all combinations of 1 degree sectors into sectors from 1 to 360 degrees long (Table 2). 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:

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

subject to

$$\sum c_{ij} x_i = 1 \quad (5)$$

$$\sum_{i=1}^n x_i \leq P_{\max} \quad (6)$$

$$lx \geq L_{\min} \quad (7)$$

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

where x_i is the decision variable for zone i , y is the optimum yield per unit area 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.

The objective function (4) maximizes cotton lint yields 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 sectors is greater than the minimum user-defined (i.e. 5 degree in this study). If a sector 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):

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

$$\text{MBE} = \frac{1}{n} \sum_{i=1}^n (E_i - M_i) \quad (9)$$

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

where E_i and M_i are the estimated 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 estimated and the mean of actual cotton 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 in Table 1 (Figure 4). A cumulative Hotelling's T^2 was plotted against the number of PCs where about 96 % of the inputs were explained by considering 10 PCs, but adding more PCs did not improve nor explain most of the variability. Our goal was to explain at least 95 % of the variability in our data set using selected PCs. Consequently, we used only the first 10 PCs in our WPFs. However, to provide an overall insight regarding performance of the PCA approach in this study, we only discussed the coefficient of the first two PCs in detail.

The PCs coefficients for the first two PCs (Figure 4, 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 4b). 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 4, panel d). There was a clear distinction between (i) points with lower ECa (corresponding to 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 (corresponding to fine-textured soil with higher water holding capacity) 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 presents the performance of k -NN WPFs (i.e. the average RMSE as well as the standard deviation of RMSE values among 10 subsets of data used for model evaluation) 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 $k=5$, 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 among 10 subsets 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 3 summarizes the performance evaluation statistics for the WPFs. Figure 6 illustrates the estimated 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 agreement with well scattered cloud of points around 1:1 line (Figure 6). The ANN (MLR) exhibited 60 % (73 %) and 52 % (65 %) increase in RMSE compared to the corresponding k -NN WPFs in 2013 and 2014 cropping seasons, respectively. The ANN WPFs had a better performance (i.e. 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 equal to 14 % and 13 % and r equal to 0.65 and 0.78 for 2013 and 2014 lint yield data, respectively.

3.3 Irrigation and zoning scenarios

Figure 7 (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 at optimum level) 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. We attribute this 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) $1.09 \pm 0.06 \text{ Mg ha}^{-1}$ and $0.98 \pm 0.06 \text{ Mg ha}^{-1}$ for east and west 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 $1.19 \pm 0.05 \text{ Mg ha}^{-1}$ and $1.11 \pm 0.06 \text{ Mg ha}^{-1}$ 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 compared 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 $1.52 \pm 0.12 \text{ Mg ha}^{-1}$ and $1.62 \pm 0.12 \text{ Mg ha}^{-1}$, underneath the east and west pivots, respectively. When yield in each cell was optimized with a distinct level of IW, the predicted yield increased to $1.71 \pm 0.10 \text{ Mg ha}^{-1}$ and $1.80 \pm 0.09 \text{ Mg ha}^{-1}$ underneath the east and west pivots, respectively.

Figure 7 (panels c and d) depicts the average irrigation and yield for optimum arrangement of the sectors. Yield was always higher in 2014 than in 2013. In both years there was a constant increase in yield by adding to the number of sectors, yet the highest boost occurred when pivot area was divided into two sectors. The yield was higher under the east pivot than under west pivot in 2013, but the opposite was true in 2014. In 2013, as number of zones increased, there was a slight reduction in optimum level of IW underneath the west pivot, but it 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 west pivot and same was true under the east pivot but less pronounced. Compared 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. Sector control VRI added about 4 % to the expected yield in both years. For west pivot, compared 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 sector control VRI added 4 % in 2013 and 7 % in 2014 to the expected yield boost.

Figure 8 illustrates the thematic yield maps predicted by k -NN WPFs for the rainfed condition along with multiple irrigation strategies (i.e. best uniform irrigation, cutting 4 sectors, and irrigating each cell individually at optimum level) for both pivots and cropping seasons. Haghverdi et al. (2015b) suggested four sectors as optimum to minimize soil spatial variation underneath pivots. To compare the zoning results across the studies, we selected 4 sectors here too. By visual assessment of the yield maps, the biggest change occurred from rainfed to uniform irrigation followed by the change from sector control VRI 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 by Haghverdi et al. (2015c). They showed that coarse-textured regions with low available water content were located in the middle, eastern and southwestern parts of the eastern pivot as well as in northern and northwestern of the western 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 were 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 sector control VRI scenario exhibited lower minimum but higher maximum than that for the corresponding uniform irrigation maps. Moving from sector control VRI into the hypothetical scenario of irrigating each 25 m² 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 9 shows the spatial arrangement of the optimum sectors from this study and the previous study by Haghverdi et al (2015b). 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. Haghverdi et al. (2015b) found the best spatial arrangement of sector zones using ILP optimization and soil related attributes without considering the

irrigation-yield relationships. They used soil available water content, soil ECa, yield data and Landsat 8 panchromatic satellite data when the objective was to minimize soil spatial variation within each zone. The spatial arrangement of optimum pie shape sectors 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 from Haghverdi et al (2015b) study. In 2014 for the east pivot, three sectors were located on the coarse-textured region with low soil available water content, while a big sector covered rest of the pivot. There were some similarities between 2014 zones for the west pivot and zoning by ECa and AWC data from Haghverdi et al (2015b) study. In 2013, sectors in northern part of the west pivot were somewhat similar to sectors from AWC data from Haghverdi et al (2015b) study. The spatial arrangement of optimum zones using yield data (by Haghverdi et al., 2015b) was different than the configuration of zones in our 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 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 we used the entire data set to train the ANN models. This is in line with the results reported by Haghverdi et al. (2014b) and Drummond et al. (2003) who found ANN a better option than MLR models to predict yield. However, ANN in our 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 our 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., 2015a), 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 the spatial variability of crop yields (Sudduth et al., 1996), yet the k -NN WPF designed in our study performed satisfactory because it acts as a smart search

1 engine to select only a few similar data points for prediction. This approach helps k -NN WPF to take
2 advantage of numerous data points (cells) provided by the yield monitoring systems without necessarily
3 having a good approximation of the non-linear relationships between inputs and yield for all cells across the
4 field of study. The k -NN WPF makes use of similarities to distinguish patterns instead of fitting coefficients
5 and this approach could be favorable whenever the form of relationships between input and output of a
6 system is not known in advance (Nemes et al., 2006). One should note that, k -NN WPF introduced in this
7 study only provides a pseudo-continuous prediction of yield over a range of applied irrigation. In other
8 words, it gives trustable results only in the neighborhood of measured IWs. Hence, its application for
9 continuous prediction of yield-water relationship is only suggested if IW is applied at multiple levels such that
10 the available data sufficiently cover the desired prediction range.

11 We had substantial spatial variation of soil properties in the study field. Long-term analysis of yield maps
12 for different crops revealed that the spatial variability of soil attributes can seriously impact crop yields
13 (Haghverdi, 2015). Guo et al. (2012) tried to mathematically explain the spatial and temporal variation of
14 cotton yield on the Southern High Plains of Texas by means of ECa, soil brightness, and topography data.
15 They observed that these attributes together could explain up to 70.1 % of the variability in cotton yield. Guo
16 et al. (2012) mentioned soil texture as one of the greatest factor affecting cotton yield. They found a relatively
17 stable spatial pattern of yield over time although yield and soil properties had stronger relationship in dry
18 years than wet seasons. The results of this study showed a considerable impact of the spatial and temporal
19 pattern of rainfall pattern on the spatial distribution of cotton yield. We realized that rainfall pattern can
20 mask/change the effect of supplemental irrigation practices on yield. Receiving high amounts of rainfall with
21 random temporal patterns over growing seasons is expected in western TN. There were 712 mm and 633 mm
22 of rainfall over the 2013 and 2014 growing seasons, respectively. The total seasonally applied irrigation water
23 by the farmer for the east and west pivots was 133-mm and 93 mm in 2013 and 65 mm and 50 mm in 2014,
24 respectively. This meant that irrigation was less than 20 % of precipitation in 2013 and less than 10 % of
25 precipitation in 2014. Wong and Asseng (2006) introduced interactions of seasonal rainfall, plant available soil
26 water storage capacity, and N fertilizer applications as dominant factors causing spatiotemporal wheat yield

variation. In 2013, wet soil and colder temperatures forced the farmer to postpone planting cotton. Furthermore, heavy rainfall throughout the growing season caused excessive vegetative growth and larger number of bolls on soils with higher available water content. We believe that planting delay caused an inability in cotton to accumulate adequate heat-units and delayed its maturity leading to a reduction in yield, especially over soils with higher water holding capacity, where 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 required 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 under different conditions, hence they should be tested for different crops, climatic conditions and the spatial variation of soils. 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 we assume that coarse resolution topography and slope have negligible influence on cotton lint yield. However, if topography and slope impact the spatial distribution of water within a field they should be considered in the development of site-specific WPFs.

4.2 Zoning and irrigation strategies

The majority of the farmers in west Tennessee practice uniform irrigation despite spatial variability in soil properties that exists in their fields. Site-specific WPFs can help them to move towards optimum uniform irrigation strategies for different crops and climatic condition. Rainfall was abundant over the course of this experiment, but some dry periods did occur and supplemental irrigation was crucial to fulfilling ET demand, especially for coarse soils with lower AWC. Thus, there was a positive cotton lint yield response to irrigation predicted by the use of k -NN WPFs under both center pivots when compared to rainfed yields. In 2013, delayed planting affected heat unit accumulation/distribution and maturity which in turn suppressed the yield

1 across the field, while yield reduction in soils with higher available water was more pronounced. This was
2 shown in WPFs prediction as a general slight increase in yield as IW increased. Adding more water reduced
3 the drought stress for soils with lower AWC but increased overall yield until a peak point. Analysis of yield
4 data showed that even for coarse-textured soils with low WHC, cotton yield reduction occurred due to over
5 irrigation. That is why, a yield reduction was predicted by ϵ -NN-WPF for IWs of more than 120 mm. In
6 2014, the overall yield-irrigation relationship across the field was somewhat different than that for 2013, yet
7 closer to what we expected to see in most of the years. As in 2013, there was a peak at a high irrigation level
8 (IW = 110 mm) which is attributed to the optimum level of IW for soil with lower AWC. Unlike 2013,
9 however, in 2014 there was also a second peak for lower IWs which turned out to be the best uniform cotton
10 irrigation strategy according to ϵ -NN WPF predictions. This was recognized as the point of optimum IW
11 level for soils with higher AWC suggesting that in a more typical year like 2014 (a year without delayed
12 planting), lower irrigation for fine-textured soils produced higher yields that compensated for the yield
13 reduction over water-stressed coarse-textured soils in different regions of the field. More years of data are
14 needed to confirm this as a long-term trend.

15 Given the spatial variability of soil attributes in the field, west Tennessee farmers are likely to depart
16 from uniform irrigation scenarios to some type of VRI. Sector control VRI is an easier and less expensive
17 irrigation strategy for those who possess center pivot irrigation systems. As previously stated, zone
18 delineation for irrigation management is complex. Site-specific WPFs could be beneficial to support such
19 conversion from uniform to sector control VRI systems, because they are useful for delineating crop-specific
20 irrigation zones and for identifying the optimum irrigation strategy for each sector. A field is likely to have
21 sectors with variation in soil attributes where real time irrigation scheduling becomes challenging. The site-
22 specific WPFs would provide insights into the right soil to follow within each section to maximize yield.
23 Haghverdi et al. (2015b) focused on soil properties to delineate irrigation management zones. We added crop
24 yields to the optimization process by using site-specific WPFs. While there were some similarities between
25 zones in this study and zones developed by Haghverdi et al. (2015b), the differences were more substantial
26 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 the interactions with soil types. We showed that the year-to-year temporal variation in rainfall patterns influences cotton yield-irrigation-soil relationships in west Tennessee. Farmer irrigation management techniques also influence 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 decisions. The more a farmer improves their irrigation management, the better k -NN WPF predictions reflect a field's yield potential. Consequently, longer experimentation is needed to see how our crop-based zoning strategy works across years and how different it would be from the soil-based zoning strategies introduced by Haghverdi et al. (2015b).

Based on the k -NN WPF predictions, irrigating each cell individually at optimum levels would have caused a moderate increase in yield in 2014 and a slight increase in 2013 compared to the best uniform irrigation strategy. In reality, this provided a cap for the potential yield boost expected from a site-specific VRI. Even in 2014, the expected boost in yield may not be high enough to compensate for the infrastructure and management expenses needed to convert to a VRI system. Obviously, the benefits of sector control and zone control VRIs is both site and crop specific and depends on a variety of factors including the magnitude of soil spatial heterogeneity throughout the field, weather conditions, and the efficiency of real time irrigation scheduling (Daccache et al., 2015).

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. However, the available data often have different spatial and temporal resolutions. Moreover, the observations may be spatially and temporally autocorrelated while also carrying inherent error and uncertainties. Therefore,

1 in order to convert these data to useful information, robust algorithms and models are needed. We designed
2 and evaluated site-specific WPFs. The results confirmed that site-specific WPFs may be powerful after-the-
3 fact tools that are used by farmers to enhance **future** site-specific irrigation management. We used site-specific
4 WPFs to evaluate different irrigation and zoning strategies. Currently, implementation and management of
5 VRI systems demands high capital costs (Evans et al., 2013), hence it is crucial to assess the potential yield
6 gain of zone control VRI over uniform and sector control VRI scenarios for each crop and a specific site.
7 Our results showed that crop-based zoning using WPFs produces different zones based on the annual
8 climatic conditions and farmer management practices. We believe a **long-term** data record is required that will
9 capture different climatic and irrigation management scenarios that lead to stable zones. A practical option is
10 to start with soil-based zoning for sector control VRI systems as proposed by Haghverdi et al. (2015b) and
11 then use site-specific k -NN-WPF to improve both zoning and within zone irrigation decisions in an iterative
12 process. For the study site and over the course of the experiment, the expected yield improvement by a zone
13 control VRI system was not substantially higher than that of sector control VRI. Note that results will
14 probably differ under drier conditions when greater yields are likely to occur under a zone control VRI
15 compared to a sector control VRI.

16 Consider that there are fundamental differences in characteristics and applications of the traditional
17 CWPF/WPF and proposed site-specific WPFs in this study. Classical PFs are mostly derived from multiyear
18 experiments that use research plots at different locations. The main goal of classical PFs is to eventually apply
19 the established functions for similar places and/or for the same place in the future. Contrary to **these studies**,
20 the site-specific WPFs that we propose are not designed to be used in other fields or even on the same field
21 in different years. These WPFs may provide insight, but are not ready for **real-time** VRI scheduling and
22 addressing related concerns such as runoff and deep percolation issues. The site-specific WPFs developed
23 here are after-the-fact tools to simulate different irrigation scenarios for a given-field over a specific cropping
24 season. They are tools to help farmers to monitor their irrigation management and provide a unique
25 opportunity that has only become possible due to recent advances and the spread of precision farming
26 equipment such as yield monitors and VRI systems.

1

2 **6. Acknowledgment**

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

4

References

- Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4), 433-459.
- Ahuja, L.R., K.W. Rojas, J.D. Hanson, M.J. Shaffer, and L. Ma (ed.). 2000. *The Root Zone Water Quality Model*. Water Resour. Publ., Highlands Ranch, CO.
- Ai-hua, S., Shi-jiang, Z., Ya-fen, G., & 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., & 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., & Vories, E. D. (2007). Spatial analysis of cotton (*Gossypium hirsutum* L.) canopy responses to irrigation in a moderately humid area. *Irrig. Sci.* 25(4), 429-441.
- Bannayan, M., & 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., & 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., & 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. *Mach. Learn.* 24:123–140.
- 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.
- Dai, X., Huo, Z., & 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., & 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., & 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., & 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.
- Duke, H.R., G.W. Buchleiter, D.F. Heermann, and J.A. Chapman. (1997). Site specific management of water and chemicals using self-propelled sprinkler irrigation systems. Pp. 273-280. In: *Precision Agriculture 1997*. J.V. Stafford (ed.) BIOS Scientific Publishers Ltd. Oxford, UK.

Evans, R.G., S. Han, S.M. Schneider, and M.W. Kroeger. (1996). Precision center pivot irrigation for efficient use of water and nitrogen. Pp.75-84. In: Proceedings of the 3rd International Conference on Precision Agriculture. P.C. Robert, R.H. Rust, and W.E. Larson (eds.) Agronomy Society of America, Crop Science Society of America, Soil Science Society of America, Madison, Wisconsin.

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.

Fortin, J.G., Anctil, F., Parent, L-É., & 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., & 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., & Bronson, K.F. (2012). Relationship between cotton yield and soil electrical conductivity, topography, and Landsat imagery. *Precision Agriculture*, 1-15.

Haghverdi, A., (2015). A site-specific and dynamic modeling system for zoning and optimizing variable rate irrigation in cotton. PhD dissertation. University of Tennessee

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

Haghverdi, A., Ghahraman, B., Leib, B. G., Pulido-Calvo, I., Kafi, M., Davary, K., & Ashorun, B. 2014b. 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., Cornelis W. M., (2015a). A simple nearest-neighbor technique to predict the soil water retention curve. *Trans. ASABE*. 58(3).

Haghverdi, A., Leib, B. G., Washington-Allen, R. A., Ayers, P. D., Buschermohle, M. J., (2015b). Perspectives on delineating management zones for variable rate irrigation. *Comput. Electron. Agric.* 117, 154-167.

Haghverdi, A., Leib, B. G., Washington-Allen, R. A., Ayers, P. D., & Buschermohle, M. J. (2015c). High-resolution prediction of soil available water content within the crop root zone. *Journal of Hydrology*.

Holan, S., Wang, S., Arab, A., Sadler, E. J., & Stone, K. (2008). Semiparametric geographically weighted response curves with application to site-specific agriculture. *Journal of agricultural, biological, and environmental statistics*, 13(4), 424-439.

Hedley, C.B., & 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., & 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., & 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., & Yafeng, X. (2012). Experimental study on water production function for waterlogging stress on corn. *Procedia Engineering*, 28, 598-603.

- 1 Mathworks. (2014). Statistics Toolbox™ User's Guide (R2014b). Retrieved November 10, 2011 from
- 2 http://www.mathworks.com/help/pdf_doc/stats/stats.pdf
- 3 Mucherino, A., Papajorgji, P., & Pardalos, P. (2009). Data mining in agriculture (Vol. 34). Springer Science & Business
- 4 Media.
- 5 Nemes, A., Rawls, W. J., & Pachepsky, Y. A. (2006). Use of the nonparametric nearest neighbor approach to estimate
- 6 soil hydraulic properties. *Soil Science Society of America Journal*, 70(2), 327-336.
- 7 Orgaz, F., Mateos, L., & Fereres, E. (1992). Season length and cultivar determine the optimum evapotranspiration deficit
- 8 in cotton. *Agronomy Journal*, 84(4), 700-706.
- 9 Pan, L., Adamchuk, V. I., Martin, D. L., Schroeder, M. A., & Ferguson, R. B. (2013). Analysis of soil water availability by
- 10 integrating spatial and temporal sensor-based data. *Precision Agriculture*, 14(4), 414-433.
- 11 Pringle, M. J., Bishop, T. F. A., Lark, R. M., Whelan, B. M., & McBratney, A. B. (2010). The analysis of spatial
- 12 experiments. In *Geostatistical Applications for Precision Agriculture* (pp. 243-267). Springer Netherlands.
- 13 Sadler, E. J., Camp, C. R., Evans, D. E., & Millen, J. A. (2002). Spatial variation of corn response to irrigation.
- 14 *Transactions of the ASAE*, 45(6), 1869-1881.
- 15 Sadler, E. J., Evans, R. G., Stone, K.C., and Camp, C.R. (2005). Opportunities for conservation with precision irrigation.
- 16 *J. of Soil & Water Cons.* 60(6):371-379.
- 17 Saseendran, S. A., Ahuja, L. R., Ma, L., Trout, T. J., McMaster, G. S., Nielsen, D. C., Ham, J. M., Andales, A. A.,
- 18 Halvorson, A. D., Chávez, J. L., & Fang, Q. X. (2014). Developing and normalizing average corn crop water
- 19 production functions across years and locations using a system model. *Agricultural Water Management*, (The Jim
- 20 Oster Special Issue) 157: 65-77.
- 21 Sudduth, K. A., S. T. Drummond, S. J. Birrell, & N. R. Kitchen. (1996). Analysis of spatial factors influencing crop yield.
- 22 In *Proc. 3rd Int. Conf. on Precision Agriculture*, 129–140. P. C. Robert, R. H. Rust, and W. E. Larson, eds.
- 23 Madison, Wisc.: ASA–CSSA–SSSA.
- 24 Stone, K., & Sadler, E. 2015. Assessing spatial variation of corn response to irrigation using a bayesian semiparametric
- 25 model. *Transactions.2015 ASABE / IA Irrigation Symposium: Emerging Technologies for Sustainable Irrigation -*
- 26 *A Tribute to the Career of Terry Howell, Sr. Conference Proceedings* .(doi:10.13031/irrig.20152142961)
- 27 Tong, F., & Guo, P. (2013). Simulation and optimization for crop water allocation based on crop water production
- 28 functions and climate factor under uncertainty. *Applied Mathematical Modelling*, 37(14), 7708-7716.
- 29 Wang, Y. R., Kang, S. Z., Li, F. S., Zhang, L., & Zhang, J. H. (2007). Saline water irrigation scheduling through a crop-
- 30 water-salinity production function and a soil-water-salinity dynamic model. *Pedosphere*, 17(3), 303-317.
- 31 Wong, M. T. F., & Asseng, S. (2006). Determining the causes of spatial and temporal variability of wheat yields at sub-
- 32 field scale using a new method of upscaling a crop model. *Plant and Soil*, 283(1-2), 203-215.
- 33 Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou,
- 34 Z., Steinbach, M., Hand, D. J., & Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and*
- 35 *Information Systems*, 14(1), 1-37.
- 36

Figure 1. The two center pivot 73 ha cotton field that was used for this 2-years study is located in southwestern Dyer County, Tennessee, USA.

Figure 2. The irrigation experimental design applied to the two center pivots in the 73 ha field. Irrigation water (IW) is the supplemental irrigation applied by the farmer while IW+ and IW- are the regions with more and less irrigated water, respectively (a). The center pivots is shown in relation to the spatial variability of soil available water content (AWC) within effective root zone of the study area (b). This figure is adapted from Haghverdi et al., (2015c).

Figure 3. The three workflow processes used in this study to develop customized site-specific water production functions (WPF) to predict cotton yields for 2 center pivots under in a 73 ha field in west Tennessee; k -NN: k nearest neighbors, PCA: principle component analysis, ANN: artificial neural networks.

Figure 4. The results of the principle components analysis (PCA) of input variables (Table 1) for 2 center pivots in a 73 ha field in west Tennessee. Cumulative Hotelling's T^2 for different number of PCs (a). The thematic map of PC1 (b). Principle component coefficients for PC1 and PC2 (c). Scatter plot of PC1 against PC2, color coded using soil ECa data (d).

Figure 5. The performance of the k nearest neighbors (k -NN) water production functions (WPFs) over the 2013 (a) and 2014(b) cropping season for 2 center pivots in a 73 ha field in west Tennessee for different combinations of k (i.e. 1 to 10) and p values (i.e. 0.5 to 3). **Left side panels:** average RMSE values among 10 test subsets, **right side panels:** standard deviation of RMSE values among 10 test subsets.

Figure 6. Scatter plots of field measured versus estimated cotton lint yields for 2 center pivots in a 73 ha field in west Tennessee are used to illustrate the performance of the three processes, including k -NN (k -nearest neighbor), ANN (artificial neural networks) and MLR (multiple linear regression), that were used in this study to produce water production functions (WPF) and predict cotton lint yields.

Figure 7. The cotton lint yield predictions in 2013 and 2014 cropping seasons was produced using site-specific k -NN (k -nearest neighbor) water production functions (WPF) in relation to multiple uniform and variable rate irrigation (VRI) scenarios for 2 center pivots in a 73 ha field in west Tennessee. Panels *a* and *b*: the dash lines with circles and triangles represent predicted yield under uniform irrigation for the east and west pivots, respectively. Individual circles and triangles above the line represent estimated yield for cell-based VRI scenario. Panels *c* and *d*: The dash lines with circles and triangles represent predicted yield for sector control VRI for east and west pivots, respectively. Bars are associated with right-side y axis and represent average irrigation applied by each pivot.

1 **Figure 8.** Cotton lint yield maps for 2 east (E) and west (W) center pivots in a 73 ha field in west Tennessee predicted
2 by site-specific k -NN (k -nearest neighbor) water production functions (WPF) under 16 different irrigation scenarios. On
3 the lower right part of each map the designation 13-W-R for example is 13 for the 2013 cropping season, W for the west
4 center pivot, and R for rainfed fields. The columns are ordered by irrigation status for each pivot as uniform irrigation
5 (U), sector control pie shape VRI (P), and individual cell (M).

6

1 **Figure 9.** Spatial arrangement of the predicted optimum pie shape sectors ($n = 4$) for variable rate irrigation for 2 east
2 and west center pivots in a 73 ha field in west Tennessee. Sectors delineated using ILP (integer linear programming) plus
3 site-specific k -NN (k -nearest neighbor) water production functions (WPFs) were from this study –i.e. the four circles up
4 in the first row- , but sectors delineated using ILP were from Haghverdi et al. (2015b). Band 8: panchromatic band from
5 space images taken by OLI sensor onboard Landsat 8; AWC: soil available water content; ECD: deep soil ECa readings.
6

Table 1. Input/output attributes for deriving site-specific water production functions (WPF) for 2 center pivots in a 73 ha field in west Tennessee.

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, ** θ_{10-1} , θ_{10-2} , θ_{10-3} , θ_{10-4} , θ_{33-1} , θ_{33-2} , θ_{33-3} , θ_{33-4} , θ_{1500-1} , θ_{1500-2} , θ_{1500-3} , θ_{1500-4}
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 ($\text{cm}^3 \text{ cm}^{-3}$), θ : water content predicted by PTFs and interpolated by co-kriging at soil matric potentials equal to -10, -33 and -1500 kPa ($\text{cm}^3 \text{ cm}^{-3}$), 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.

1 **Table 2.** All of the sector zones that could be delineated underneath a center pivot systems. This Table is adapted from
2 Haghverdi et al., (2015b).

		sector length			
		1	2	...	360
Start angle	0	0-1 ^a	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

3 ^a X-Y: start angle-end angle.

4

Table 3. The performance results (r = correlation coefficient, root mean square error = RMSE, and mean bias error = MBE) of the three workflow processes: k -NN (k -nearest neighbor), ANN (artificial neural networks) and MLR (multiple linear regression) that were used to produce site-specific water production functions (WPFs) for predicting cotton lint yields for the 2013 and 2014 cropping seasons in a 73 ha study field in west Tennessee.

	2013			2014		
	r	RMSE (Mg ha ⁻¹)	MBE (Mg ha ⁻¹)	r	RMSE (Mg ha ⁻¹)	MBE (Mg ha ⁻¹)
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