THESIS

USING NONLINEAR GEOSTATISTICAL MODELS FOR SOIL SALINITY AND YEILD MANAGEMENT

Submitted by

Ahmed Eldeiry

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2013

Master's Committee:

Advisor: Luis A. Garcia
Co-Advisor: Robin M. Reich

Neil S. Grigg

ABSTRACT

USING NONLINEAR GEOSTATISTICAL MODELS FOR SOIL SALINITY AND YEILD MANAGEMENT

Crop production losses associated with soil salinity on irrigated lands are significant. The genetic complexity of crops with regards to salt tolerance has limited the success of improving salt tolerance through conventional breeding programs. In the meantime, land reclamation and leaching can be expensive and sometimes impractical when fresh water sources are scarce or not readily available. This research introduces a geostatistical approach for the management of crop yield under current soil salinity conditions. It uses three nonlinear geostatistical models – disjunctive kriging (DK), indicator kriging (IK), and probability kriging (PK) – to manage soil salinity and crop yield. The nonlinear models were applied to selected irrigated fields in a study area located in the south eastern part of the Arkansas River Basin in Colorado where soil salinity is a problem in some areas. The overall objectives of this research are: 1) estimate soil salinity in irrigated fields using nonlinear gestatistical models; 2) develop conditional probability (CP) maps that divide each field into zones with different soil salinity levels; 3) estimate the expected yield potential (YP) for several crops at different zones in fields under multiple soil salinity thresholds; 4) evaluate the performance of the nonlinear geostatistical models in developing the interpolated and CP maps provide guidance to farmers and researchers by considering the output of this research as input for precision management of agriculture; and 5) provide guidance to farmers and decision makers in precision management of agriculture.

The three nonlinear geostatistical models DK, IK, and PK were used to develop CP maps based on soil salinity thresholds for different crops. These CP maps were compared with actual yield data taken while conducting a soil salinity survey for two fields cultivated with

alfalfa and corn. The CP maps divide each field of interest into zones with different probabilities to reach a specific YP for a given crop at a specific soil salinity threshold. Different crops were selected to represent the dominant crops grown in the study area: alfalfa, corn, sorghum, and wheat. Six fields were selected to represent the range of soil salinity levels in the area. Soil salinity data were collected in the fields using an EM-38 and the location of each soil salinity sample point was determined using a GPS unit. Datasets of soil salinity collected in irrigated fields were used to generate the CP maps and to evaluate different scenarios of the expected YP% of several crops at multiple soil salinity thresholds. These datasets were selected to represent a wide range of soil salinity conditions in order to be able to evaluate a wide variety of crops (larger set of crops than those grown in the study area) according to their soil salinity tolerances. Yield data were collected at the same fields to compare the actual data with that estimated by the models. The crops were used for evaluation were selected based on two criteria: dominant in the study area, and represent high, moderate, and low soil salinity tolerances. Different scenarios of crops and salinity levels were evaluated. Semivariograms were constructed for each scenario to represent the different classes of percent yield potential based on soil salinity thresholds of each crop.

The results of this research show the nonlinear geostatistical models are efficient in assessing the impact of soil salinity on the spatial variability yield productivity. The comparison of the actual yield data with the estimated yield from the three models shows good agreement where most of the yield samples were located at the appropriate zones estimated with the models. The IK and PK models generated very similar estimates for each of the zones. However, the zones generated by both of these models are slightly different to the zones generated using the DK model. Wheat and sorghum show the highest expected yield potential based on the different soil salinity conditions that were evaluated. Expected net revenue for alfalfa and corn are the highest under the different soil salinity conditions that

were evaluated. The CP maps generated using the DK technique give an accurate characterization and quantification of the different zones of the fields. Upon the knowledge of the YP% of different areas, a management decision action can be taken to manage the productivity of a field by selecting another crop or adjusting the inputs such as fertilizer, seeding rates and herbicides in low productivity areas. The information provided by the models about the variability and hotspots can be used for the precision management of agricultural resources. The IK model can be used to generate guidance maps that divide each field into areas of expected percent yield potential based on soil salinity thresholds for different crops. Zones of uncertainty can be quantified by IK and used for risk assessment of the percent yield potential.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 GENERAL INTRODUCTION

Approximately 25-30% of the irrigated lands in the United States have crop yields that are negatively affected by high soil salinity levels (Tanji, 1990; Postel, 1989; Ghassemi et al., 1995; Wichelns, 1999). Worldwide crop production losses associated with salinity on irrigated lands are estimated to be around US $11 billion annually and increasing (Ghassemi et al., 1995). The Arkansas River is one of the most saline rivers in the United States (Tanji, 1990; Miles, 1977). The Arkansas River drains approximately 25% of the state and is the state's largest river basin. Soil salinity problems exist when the buildup of salts in a crop's root zone is significant enough that it results in a loss in crop yield. Soil salinity negatively affects crop growth by increasing the osmotic potential of the soil solution (Jones and Marshall, 1992), which decreases a plant ability to extract water and results in suppressed plant growth and decreased yield. The development of saline soils is a dynamic phenomenon that needs to be monitored regularly in order to secure up-to-date knowledge of its extent, spatial distribution, nature and magnitude (Ghassemi et al., 1995).

Geostatistical methods have been widely used for sampling and mapping soil salinity. They provide means to study the heterogeneity of the spatial distribution of soil salinity (Pozdnyakova and Zhang, 1999). Kriging is a collection of linear regression techniques that takes into account the stochastic dependence among data (Olea, 1991). Kriging remains the best choice as a spatial estimation tool since it provides a single numerical value that is best in some local sense (Deutsch and Journel, 1998). The results of spatial prediction generate reasonable estimates of soil salinity regardless of what interpolation method was used (Triantafilis et al., 2006). Kriging models estimate the values at unsampled locations by a weighted averaging of nearby samples where the correlations among neighbouring values are modelled using variograms (Miller et al., 2007). Studies have shown that semivariograms of electrical conductivity can be a useful tool in determining the spacing between soil samples

for laboratory electrical conductivity determination (Utset et al., 1998). Samra and Gill (1993) used kriging results to assess the variation of pH and sodium adsorption ratios associated with tree growth on a sodium-contaminated soil.

The variogram is the key function in geostatistics as it is used to fit a model of the spatial correlation of the observed phenomenon and provides a unique spatial study. Given a collection of data, a variogram reveals the type of spatial structure inherent to a spatial phenomenon. In addition, the variogram reveals the amount of noise present in the data, known commonly as the nugget (Carr et. al., 1985). Recent research shows that this noise can substantially mask prominent spatial autocorrelation and result in what appears to be a purely random spatial process. When a variogram is used to describe the correlation of different variables it is called a cross-variogram. Cross-variograms are used in co-kriging. If the variable being analysed is binary or represents classes of values, this is referred to as indicator variograms. Nonparametric geostatistical techniques such as IK offer immeasurable power for analysis of data quality (Journel, 1983).

Linear kriging methods such as simple, ordinary and universal kriging are well established for predicting soil variables at unsampled locations and have been used widely in soil and water science. Eldeiry and Garcia (2008a; 2008b; 2010) used different linear kriging techniques to estimate soil salinity using remote sensing data. Burgess and Webster (1980b) and Webster and Burgess (1980) demonstrated the use of block and universal kriging. Triantafilis et al. (2001) used ordinary kriging, regression kriging, three-dimensional kriging, and cokriging to predict soil salinity from electromagnetic induction data in irrigated cotton. However, correctly assessing prediction uncertainty (conditional probability) has the same importance as predicting a variable at unsampled locations. Nonlinear kriging methods provide estimates of the conditional distribution of a variable quantity. There are two groups of nonlinear kriging techniques where the conventional linear kriging estimators are applied

to the data after a nonlinear transformation. The first group is indicator methods (Journal, 1983), where the nonlinear transform to data is a discrete (binary) indicator variable. These techniques have been widely applied (e.g. Van Meirvenne and Governs, 2001; Halvorson et al., 1995; Eldeiry and Garcia 2011). The second group of techniques, which is discussed in this study, involves a nonlinear transformation of the data to a continuous (Gaussian) variable. This approach is exemplified by Disjunctive Kriging (DK) (Matheron, 1976) and has found widespread use in soil science (e.g. Wood et al., 1990; von Steiger et al., 1996).

The document is organized into three main chapters, each containing a self-contained presentation in which specific components of the objectives presented above are addressed. The organization of the thesis is as follows:

Chapter 2 presents a practical method to manage soil salinity and yield in order to obtain maximum economic benefits using Indicator Kriging (IK) technique. The IK was applied to six irrigated fields in the study area located in the south eastern part of the Arkansas River Basin in Colorado where soil salinity is a problem. Different scenarios of crops and salinity levels were evaluated, while the IK technique was applied to each scenario. The generated maps show the expected percent yield potential areas and the corresponding zones of uncertainty for each of the different classes. Throughout the results section, the expected crop net revenue for each scenario was calculated and all the results were compared to determine the best scenarios. The results show that IK can be used to generate guidance maps that divide each field into areas of expected percent yield potential based on soil salinity thresholds for different crops. Zones of uncertainty can be quantified by IK and used for risk assessment of the percent yield potential. The results section discussed and evaluate the expected net revenue of several crops under different soil salinity conditions.

In Chapter 3, the Disjunctive Kriging (DK) technique was applied to two datasets of soil salinity in the same study area to generate the conditional probability (CP) maps. Different

scenarios of the expected YP% of several crops at multiple soil salinity thresholds were discussed and evaluated. From the discussion and evaluation it was concluded that the CP maps generated using the DK technique give an accurate characterization and quantification of the different zones of the fields. The CP maps can be used to assess the expected YP% of whole fields for several crops under multiple soil salinity thresholds. The information provided about the YP% of different areas can be used to support the management decision action that can be taken to manage the productivity. This management could be through the selection of another crop or adjusting the inputs such as fertilizer, seeding rates and herbicides in low productivity areas.

In Chapter 4, instead of discussing and evaluating the third nonlinear geostatistical technique, Probability Kriging (PK), the three nonlinear models – DK, IK, and PK – were compared and evaluated. Also, instead of generating the CP maps of different crop scenarios, actual crop samples for alfalfa and corn were collected in order to compare the actual data with that estimated by the models. The three nonlinear models were used to develop the CP maps based on soil salinity thresholds for alfalfa and corn and the CP maps were compared with actual yield data taken while conducting a soil salinity survey for alfalfa and corn. The comparison of the actual yield data with the estimated CP maps from the three models shows good agreement where most of the yield samples were located at the appropriate zones estimated with the three geostatistical models. The IK and PK models generated very similar estimates for each of the zones. However, the zones generated by both of these models are slightly different to the zones generated using the DK model. The information provided by the models about the variability and hotspots can be used for the precision management of agricultural resources.

A summary with general remarks is presented in Chapter 5, highlighting the most relevant findings of this research.

1.1 REFERENCES

Burgess, T. M. and Webster, R. (1980b). "Optimal interpolation and isarithmic mapping of soil properties, 2. Block Kriging." *J. Soil Science* **32**, 315-331.

Carr, J. R., Bailey, R. E., and Deng, E. D. (1985). "Use of Indicator Variograms for an Enhanced Spatial Analysis." *Mathematical Geology*, **17(8)**, 797 – 811.

Deutsch, C.V., and Journel, A.G. (1998). GSLIB, Geostatistical software library and user's guide. 2nd ed. Oxford University Press, New York.

Eldeiry, A., and Garcia, L. A. (2008a). "Detecting Soil Salinity in Alfalfa Fields using Spatial Modeling and Remote Sensing." *Soil Sci. Soc. Am. J.*, **72(1)**, 201-211.

Eldeiry, A., and Garcia, L. A. (2008b). "Spatial Modeling of Soil Salinity Using Remote Sensing, GIS, and Field Data." *VDM Verlag, Saarbruken, Germany.*

Eldeiry, A., and Garcia, L. A. (2010). "Comparison of Ordinary Kriging, Regression Kriging, and Cokriging Techniques to Estimate Soil Salinity Using Landsat Images." *J. Irrig. Drain. Engin.*, **136(6)**, 355-364.

Eldeiry, A., and Garcia, L. A. (2011). "Using Indicator Kriging Technique for Soil Salinity and Yield Management." *J. Irrig. Drain. Engin.*, **137(2)**, 82-93.

Ghassemi, F., Jakeman, A.J., Nix, H.A. (1995). Salinisation of Land andWater Resources: Human Causes, Extent, Management, and Case Studies. *University of New SouthWales Press Ltd., Sydney, Australia.*

Halvorson, J.J., Smith, J.L., Bolton, H., Rossi, R.E. (1995). "Evaluating shrub-associated patterns of soil properties in a shrub-steppe ecosystem using multiple-variable geostatistics." *Soil Sci. Soc. Am. J.* **59**, 11476-11478.

Jones, R., Marshall, G. (1992). "Land salinisation, waterlogging and the agricultural benefits of a surface drainage scheme in Benerembah irrigation district. Rev. Market." *Agric. Econ.* **60 (2),** 173–189.

Journal, A. G. (1983). “Nonparametric Estimation of Spatial Distributions.” *Mathematical Geology,* **15(3)**, 445–468.

Journel, A. G. (1983). “Nonparametric Estimation of Spatial Distributions.” *Mathematical Geology*, **15(3)**, 445 – 468.

Matheron, G. (1976). A simple substitute for conditional expectation: The disjunctive Kriging, in Advanced Geostatistics in the Mining Industry, M. Guarascio, M. David, and C. Huijbregts, eds., pp. 221-236, *D. Reidel Publishing, Dordrecht, Holland.*

Miles, D.L. (1977). Salinity in the Arkansas Valley of Colorado. *Region VIII, Environmental Protection Agency, Denver.* 80 p.

Miller, J., Franklin, J., and Aspinall, R. (2007). “Incorporating spatial dependence in predictive vegetation models.” *Ecol. Modelling,* **202(3-4)**, 225-242.

Olea, R.A., (1991). “Geostatistical glossary and multilingual directionary”. Oxford Univ. Press, New Yourk.

Postel, S. (1989). Water for Agriculture: Facing the Limits. *Worldwatch Paper 93 Worldwatch Institute, Washington, DC.*

Pozdnyakova, L. and Zhang, R. (1999). “Geostatistical Analyses of Soil Salinity in a Large Field.” *Precision Agriculture,* **1**, 153-165.

Samra, S., and Gill, H.S. (1993) “Modeling of variation in a sodium-contaminated soil and associated tree growth.” *Soil Sci.*, **155(2)**, 148.

Tanji, K. K., ed. (1990). Agricultural Salinity Assessment and Management. ASCE Manuals and Reports on Engineering Practice No. 71, ASCE, New York, N.Y.

Triantafilis, J., Odeh, I. O. A., and McBratney, A.B. (2001). "Five Geostatistical Models to Predict Soil Salinity from Electromagnetic Induction Data Across Irrigated Cotton." *Soil Sci. Soc. Am. J.*, **65**, 869-878.

Triantafilis, J., Odeh, I.O.A, and McBratney, A.B. (2006). "Five geostatistical models to predict soil salinity from electromagnetic induction data across irrigated cotton." *Soil Science Society of America Journal*, **65**, 869-878.

Utset, A., Ruiz, M.E., Herrera, J., and Ponce de Leon, D. (1998). "A geostatistical method for soil salinity sample site spacing." *Geoderma*, **86**, 143-151.

van Meirvenne, M., Goovaerts, P. (2001). "Evaluating the probability of exceeding a site-specific soil cadmium contamination threshold." *Geoderma* **102**, 75-100.

von Steiger, B., Webster, R., Schulin, R., Lehmann, R. (1996). "Mapping heavy metals in polluted soil by disjunctive kriging." *Environ. Pollut.* **94**, 205-215.

Webster, R. and Burgess, T. M. (1980). "Optimal interpolation and isarithmic mapping of soil properties, 3. Changing drift and universal Kriging." *J. Soil Sci.* **32**, 505-524.

Webster, R., Oliver, M.A. (1989). "Optimal interpolation and isarithmic mapping of properties. VI. Disjunctive kriging and mapping conditional probability." *J. Soil Sci.* **40**, 497-517.

Wichelns, D. (1999). "An economic model of waterlogging and salinization in arid regions." *Ecol. Econ.* **30,** 475–491.

Wood, G., Oliver, M.A., and Webster, R. (1990). "Estimating soil salinity by disjunctive kriging." *Soil Use Manage.* **6**, 97-104.

2 USING INDICATOR KRIGING TECHNIQUE FOR SOIL SALINITY AND YIELD MANAGEMENT

2.1 Summary

This chapter presents a practical method to manage soil salinity and yield in order to obtain maximum economic benefits. The method was applied to a study area located in the south eastern part of the Arkansas River Basin in Colorado where soil salinity is a problem in some areas. The following were the objectives of this study: 1) generate classified maps and the corresponding zones of uncertainty of expected yield potential for the main crops grown in the study area; 2) compare the expected potential productivity of different crops based on the soil salinity conditions; 3) assess the expected net revenue of multiple crops under different soil salinity conditions.

Four crops were selected to represent the dominant crops grown in the study area: alfalfa, corn, sorghum, and wheat. Six fields were selected to represent the range of soil salinity levels in the area. Soil salinity data were collected in the fields using an EM-38 and the location of each soil salinity sample point was determined using a GPS unit. Different scenarios of crops and salinity levels were evaluated. Indicator-variograms were constructed for each scenario to represent the different classes of percent yield potential based on soil salinity thresholds of each crop. Indicator kriging (IK) was applied to each scenario to generate maps that show the expected percent yield potential areas and the corresponding zones of uncertainty for each of the different classes. Expected crop net revenue for each scenario was calculated and all the results were compared to determine the best scenarios. The results of this study show that IK can be used to generate guidance maps that divide each field into areas of expected percent yield potential based on soil salinity thresholds for different crops. Zones of uncertainty can be quantified by IK and used for risk assessment of

the percent yield potential. Wheat and sorghum show the highest expected yield potential based on the different soil salinity conditions that were evaluated. Expected net revenue for alfalfa and corn are the highest under the different soil salinity conditions that were evaluated.

2.2 Introduction

Soil salinity refers to the presence in the soil and water of various electrolytic mineral solutes in concentrations that can be harmful to many agricultural crops (Hillel, 2000). Salts decrease the availability of water to plants due to increase osmotic potential and have direct adverse effects on the plant metabolism (Douaik, 2003; Greenway and Munns, 1980). Increasing soil salinity is offsetting a good portion of the increased productivity achieved by expanding irrigation (Postel, 1999). On average, 20% of the world's irrigated lands are affected by salts, but this figure increases to more than 30% in countries such as Egypt, Iran and Argentina (Ghassemi et al., 1995). Crop yield reduction in fields in the Lower Arkansas Valley due to salinization is estimated to vary between 0 to 75% with a total revenue loss ranging from $0-$750/ha based on 1999 crop prices (Gates et al., 2002).

A careful selection of thresholds in assigning an indicator function can yield an indicator variogram which reveals underlying spatial autocorrelation. Problems arise when dealing with highly variant phenomena where the data present long-tailed distributions with a coefficient of variation in the range of 2-5. Raw variograms become extremely sensitive to tail data, and are basically useless (Journel, 1983). Indicator variograms are not affected by outliers, since they do not call for the data values themselves but rather for their rank order (indicator values) with regard to a given cutoff. Data are used through their rank order with regard to a given cutoff, allowing for a more comprehensive structural analysis, and are yet

more robust with regard to outliers. The influence of outliers is removed from the distribution and data sensitivity to different thresholds can be uniquely studied. The indicator approach, whereby the data are used through their rank order, allows a nonparametric approach to study the bivariate distribution of the data (Journel, 1983). This rich structural information allows a nonparametric risk-qualified analysis of the data as well as an estimation of local and global spatial distributions.

IK provides a non-parametric distribution estimated directly at fixed thresholds by considering indicator transforms of conditioning data in the form of cumulative distribution functions (Richmond, 2001). The power of multi-variable IK as a tool is that it is flexible and can be modified to fit specific management or research goals by modifying the critical threshold criteria (Smith et al., 1993). IK makes no assumptions on the underlying invariant distribution, and 0:1 indicator transformation of the data makes the predictor robust to outliers (Cressie, 1993). At an unsampled location, the values estimated by IK represent a probability that the value is less than a specified threshold. That is, the expected value at the location derived from indicator data is equivalent to the cumulative distribution function of the variable (Smith et al., 1993). Mapping of uncertainty zones for individual phases is one advantage of using a geostatistical approach to characterize the morphology of quantitative variables (Soares, 1992). Smoothing effects occurring around zero thickness investigation sites can be reduced significantly by the use of a combined ordinary-IK approach (Marinoni, 2003). Solow et al. (1986) used simple IK to estimate the conditional probability that a sample point belongs to one type or another. Their results show that simple IK performed well, and in some cases can be exact.

IK provides a way to use depth to water table data to quantify the probability of saturation and evaluate the predicted spatial distributions of runoff generation risk (Lyon et. al., 2006). Spatial principal component analysis and IK were used to estimate the geochemical

distributions by utilizing their statistical and spatial properties (Panahi et al., 2004). Indicator variables and multi-level thresholds were used to analyse the Arsenic concentration probability in the coastal aquifer in Yun-Lin, Taiwan (Liu et al., 2004). Using this technique allowed them to solve the problem of data scarcity and provided multi-level thresholds in the probability estimation of contamination. IK geostatistics was also used successfully to identify the areas where mercury concentration was higher than the median in southern Portugal, and to produce an index that combines mercury contamination across trophic levels (Figueira et al., 2009). Mapping of uncertainty zones for individual phases is one advantage of using a geostatistical approach to characterize the morphology of quantitative variables (Soares, 1992). Western et al. (1998) examined soil moisture patterns through indicator semivariograms and showed good spatial structure for high soil moisture conditions.

IK has also been frequently applied to the pollution of soil by heavy metals. For example, Smith et al. (1993) and Oyedele et al. (1996) used multivariate IK to analyze the quality of soil; Lin et al. (2002) applied IK to delineate the variation and pollution sources of heavy metals in agricultural land; Juang and Lee (1998), Castrignano' et al. (2000) and van Meirvenne and Goovaerts (2001) adopted multi-level-threshold IK to estimate the probability distribution of heavy metal pollution in a field. Geostatistical indicator methods have also been applied in the lithological classification of rocks (McCord et al., 1997; Fogg et al., 1999) and in the estimation of probability of contamination in groundwater aquifers (Istok and Pautman, 1996).

Several studies have been carried out using IK in soil science. Bierkens and Burrough (1993a) showed the application of IK to predict categorical soil data. Bierkens and Burrough (1993b) also applied IK to water-table mapping and land suitability assessment. Goovaerts (1994) compared the performance of cokriging, simple kriging and multiple indicator kriging (MIK) in predicting soil quality indicators. Triantafilis et al. (2003) used MIK to produce

conditional probability maps of deep drainage risk in an irrigated cotton field in the lower Gwydir valley in southeastern Australia. Triantafilis et. al, (2004) used IK, MIK and disjunctive kriging (DK), to assess the current status and potential threat of soil salinity using data from soil and water surveys in the lower Namoi valley of northern New South Wales, Australia.

The geostatistical approach presented in this paper uses IK to provide farmers with a tool to estimate the potential maximum economic benefit under the current conditions of their fields. Crops with different soil salinity tolerances have significantly different crop yield potential under the same soil salinity conditions. Therefore, depending on the soil salinity conditions of a field, some crops will have higher yields than others. A classified map of expected yield potential based on soil salinity thresholds of different crops can help in selecting the appropriate crop that maximizes the potential yield for a specific area. In this research a set of scenarios generated from combinations of different crops and fields was analyzed using the soil salinity data for each field. Each field was classified into different thresholds to produce the following crop yield potentials: 100%, 90%, 75%, 50%, < 50% & > 0 %, and 0%. Indicator-variograms were constructed for each of the scenarios and IK was applied to each scenario to generate maps that show the expected percent yield potential as well as zones of uncertainty for different parts of each field. Expected crop net economic revenue for each scenario was calculated. The expected yield potential maps can be used by farmers to determine which crop would maximize the yield and the economic benefits of their fields under the current soil salinity conditions.

2.3 Data and Methodology

2.3.1 Study Area and Data Collection

The study area is located in the south eastern part of the Arkansas River Basin in Colorado near the cities of Rocky Ford and La Junta Figure 2.1. Farmers in this area are facing decreasing crop yields due in part to high levels of salinity in their irrigation water. In some areas, land is being taken out of production due to unsustainable crop yields. This is due in part to the fact that the Arkansas River is one of the most saline rivers in the United States (Tanji, 1990; Miles, 1977). In a survey of the region, 68% of producers stated that high salinity levels were a significant concern (Frasier et. al., 1999). Farmland along the lower Arkansas River Basin has been continuously irrigated since the 1870's and began to develop shallow, saline water tables by the beginning part of the twentieth century (Miles, 1977). Average water table depths in this region have risen towards the surface approximately 0.3 – 1.22 m between 1969 and 1994 which has only exacerbated the salinity problems because of increasing amounts of upflux of saline groundwater.

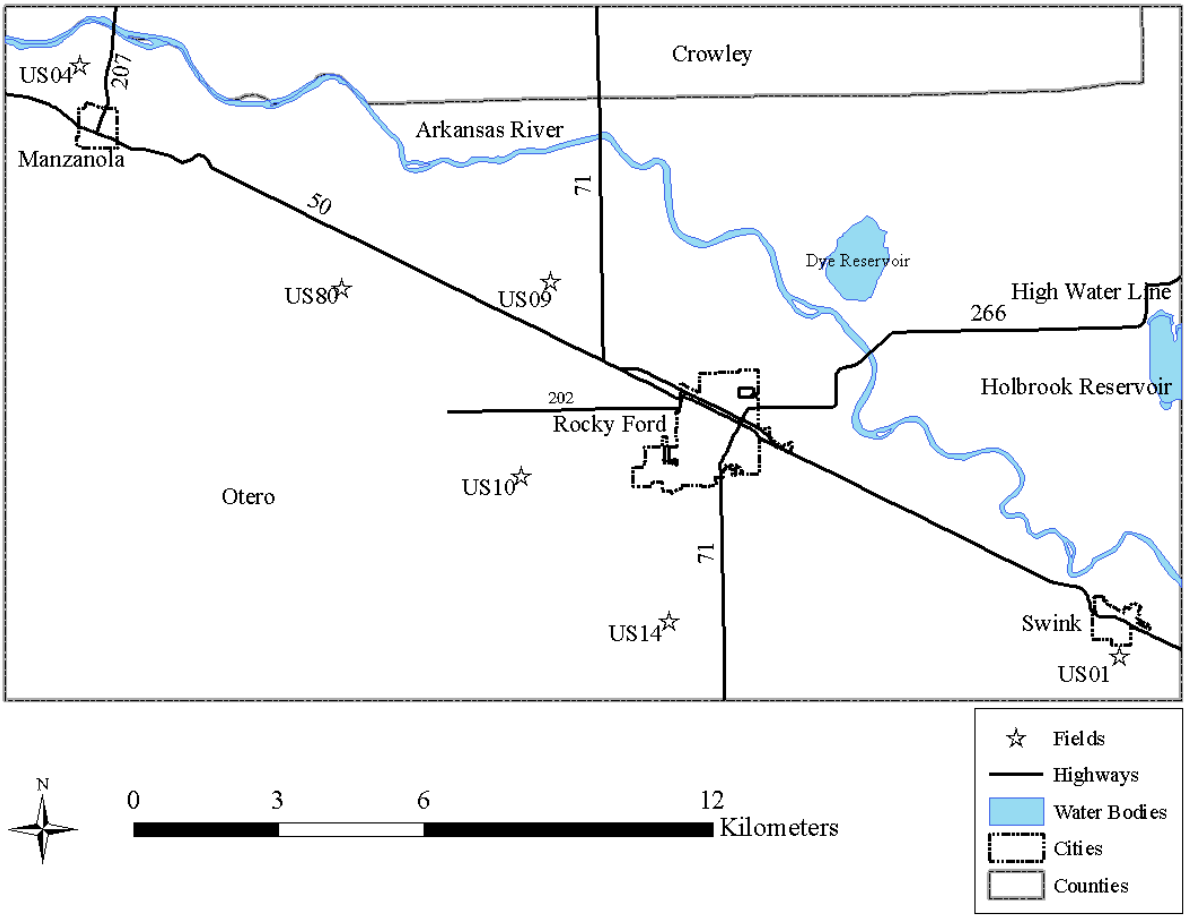


Figure 2.1: The study area in the south eastern part of the Arkansas River Basin in Colorado.

Several fields were selected to carry out the soil salinity assessment in the study area. Soil salinity data were collected using EM-38 electromagnetic probes and the location of the samples was determined using global position systems (GPS) units. The EM-38 electromagnetic probes provide vertical and horizontal readings while the GPS units provide X and Y coordinates for each sample point. A calibrated equation which was developed for the study area by Wittler et al. (2006) was used to convert the EM-38 electromagnetic probe readings to EC (dS/m). Soil moisture content and soil temperature were used for the calibration equation. A detailed description of using the EM-38 electromagnetic probe in combination with GPS in collecting soil salinity can be found in Eldeiry and Garcia (2008)

and Eldeiry et al. (2008). Six fields were selected to represent the different soil salinity ranges: low, moderate and high.

Table 2.1: Description of the fields of the study area and the collected soil salinity samples.

Field	Area (ha)	Number of Soil Salinity Samples	Min.	Max.	Mean
US01	16.20	318	2.38	7.19	3.32
US04	93.19	316	2.38	41.23	8.41
US09	28.92	369	1.57	3.49	2.30
US10	4.19	132	3.04	31.26	6.82
US14	12.73	254	2.66	11.26	4.45
US80	11.26	178	2.86	12.33	4.21

Table 2.1 shows a description of the fields used in this study. The table contains the area, number of samples, minimum, maximum, and mean values of the soil salinity that were collected in each field. These fields were selected to represent different soil salinity ranges (low, medium and high) since soil salinity is an important factor which can significantly affect crop yield. Table 2.1 shows that the selected fields represent a wide range of soil salinity levels from 1.57 to 41.23 dS/m.

2.3.2 Soil Salinity Classification

Table 2.2 shows the percent yield potential and the corresponding soil salinity EC (dS/m) for alfalfa, corn, sorghum, and wheat (adapted from Ayers and Westcot, 1976). Ayers and Westcot carried out an experiment where soil salinity was measured based on the electrical conductivity of the saturated paste extract taken from a root zone soil sample (ECe) measured in (dS/m).

Table 2.2: Yield potential and the corresponding soil salinity (dS/m) for selected crops (adapted from Ayers and Westcot, 1976).

Crop	Yield Potential %, Soil Salinity (dS/m)				
	100%	90%	75%	50%	0%
Corn	1.7	2.5	3.8	5.9	10.0
Alfalfa	2.0	3.4	5.4	8.8	16.0
Sorghum	4.0	5.1	7.2	11.0	18.0
Wheat	6.0	7.4	9.5	13.0	20.0

For barley and wheat, during the germination and seedling stages, ECe should not exceed 4 to 5 dS/m except for certain semi-dwarf varieties. For beets, during germination ECe should not exceed 3 dS/m. Many crops have little tolerance for salinity during seed germination, but significant tolerance during later growth stages. The crops shown in Table 2.2 were sorted based on their tolerance to soil salinity from low to high: corn, alfalfa, sorghum, and wheat. Table 2.2 shows that wheat can reach up to 100% of yield potential at a soil salinity of 6 dS/m while corn can only reach 50% of yield potential at a soil salinity of 5.9 dS/m.

2.3.3 Preparing the Data

The soil salinity data for each field were sorted and classified into different thresholds to produce the following crop yield potential classes: 100%, 90%, 75%, 50%, < 50% & > 0 %, and 0%. For each of these six fields, the classification was done for each of the four selected crops. For high soil salinity tolerant crops such as wheat or sorghum, in fields with low soil salinity levels such as US09, there is no need for IK since the whole field has 100% expected yield potential. However, with the same crops in fields with moderate soil salinity levels, crops can reach a high yield potential from 75% to 100% while classes with yield potential ranging from 0% to 100% can be present in fields with high soil salinity levels. For crops with moderate and low soil salinity tolerance such as alfalfa and corn, a wide range of yield potentials is represented in the selected fields for this study.

2.3.4 Constructing the Indicator Variograms

From the data combinations of crops and fields, twenty four scenarios were created (combinations of four crops and six fields). For each scenario, data were analysed using the S+ statistical software package and the indicator variograms were decided based on the number of classes or thresholds of yield potential for each scenario. For example the scenario of planting alfalfa in field US04 has five classes: 90%, 75%, 50%, <50% & > 0%, and 0% of percent of yield potential. The best model variogram among the Exponential, Gaussian, and Spherical was chosen based on the smallest Akaike Information Corrected Criterion (AICC). AICC is a measure of the goodness of fit of an estimated statistical model. It is a tradeoff between bias and variance in model construction. It is not a test of the model in the sense of hypothesis testing; rather it is a test between models (a tool for model selection). AICC was defined by McQuarrie and Tsai (1998) as:

$$AICC = \ln \frac{RSS}{n} + \frac{n+k}{n-k-2} \tag{2.1}$$

where: *RSS* is the residual sum of squares, *k* is the number of parameters, and *n* is the number of samples.

Indicator variograms were constructed for each of the scenarios using the model with the smallest AICC value. Each phase of the variograms represents one class of percent yield potential. The indicator variograms contain six phases or less depending on the tolerance of the crop and the soil salinity in the field. The indicator variogram (Soares, 1992) is defined as the probability that x and $x+h$ belong to different classes K_i:

$$\gamma(h) = \tfrac{1}{2} E\left\{ \sum_i [K_i(x) - K_i(x+h)]^2 \right\} \tag{2.2}$$

where: x and $x+h$, represent a pair of sample locations separated by distance h and i is the number of k classes of soil salinity.

2.3.5 Applying IK

IK was applied to each scenario to generate classified maps that show the expected percent yield potential. The number of classes in each map depends on the number of phases of the indicator variograms of that scenario. One of the advantages of IK is that it has the power to quantify the zones of uncertainty for different parts of each field. Zones of uncertainty exist around the borders of classes and these areas have the probability of belonging to either of the classes. Assessing zones of uncertainty can be very beneficial for the accuracy of the generated maps since it can produce more information about the risk assessment. The essence of the indicator approach is the binomial coding of soil salinity data into either 1 or 0 depending upon its relationship to the thresholds of soil salinity for each crop. For a given value of $z(x)$:

$$i(x{:}z_k) = \begin{cases} 1 \; if \; z(x) \geq z_k \\ 0 \; if \; z(x) < z_k \end{cases} \tag{2.3}$$

where z_k is the soil salinity threshold for a specific crop (Lyon et al. 2006). More detailed description of IK can be found in Soares (1992).

2.3.6 Zones of Uncertainty

The indicator variable can be described as the probability of exceeding a given threshold. Therefore, the estimation of the indicator variable at unsampled locations produces probability maps (Reis et al. 2005b). Zones of uncertainty between soil salinity classes can be obtained by identifying locations with low probability, for a given threshold, of belonging to a specific soil salinity level. For example a zone of uncertainty can be defined as being the lowest 25% of the probabilities of belonging to a particular soil salinity level. To generate a map of uncertainty, the first thing to do is to obtain some information regarding the

distribution of probabilities associated with each soil salinity class, such as identifying the threshold representing the lowest 25% of the probabilities.

Consider an attribute Z that must be conditionally simulated and the information available consists of z values at n locations x_i, $z(x_i)$, i = 1, 2, ... , n. The uncertainty about the soil salinity value at an unsampled location x is modeled by the conditional cumulative distribution function (ccdf) of the random variable Z(x):

$$F(x,z) = \text{Prob}\{Z(x) \leq z(x)\} \tag{2.4}$$

The function *F(x,z)* represents the probability that the unknown soil salinity does not exceed a threshold z. The ccdfs are modeled using a non-parametric (IK) approach, which estimates the probability for a series of K threshold values z_k discretizing the range of variation of Z (Froideveux, 1993; Saito and Goovaerts, 2002; Reis et al. 2005a):

$$F(x, z_c) = \text{Prob}\{Z(x) z_k \mid (n)\} \qquad k = 1, ..., K \tag{2.5}$$

where k is the number of samples within a specific class K.

The calculated probabilities are recoded into 0 and 1 in order to obtain binary maps with two levels, the areas with uncertainty and the areas without uncertainly, while considering a confidence interval.

2.3.7 Net Revenue

Expected crop net economic revenue for each scenario was calculated based on the Colorado State University Extension (Agriculture and Business Management) 2007 crop budget estimates. The total revenue includes the revenue of the crop without taking into account the costs. The costs include the operations associated with pre-harvest, harvest, property ownership and cost, and for some crops a factor payment. Net revenue is the revenue after the costs are taken into account. The expected crop net economic revenue can be used as guidance for the growers to determine which crop would maximize the potential economic benefits from their fields under the current soil salinity conditions.

Each field is composed of a number of areas according to soil salinity classes. Each area produces a specific yield potential based on its soil salinity class for each crop. The total revenue, cost, and net revenue for alfalfa, corn, sorghum, and wheat are based on the Colorado State University Extension (Agriculture and Business Management) 2007 crop budget estimates. The crop budget takes the averages and does not take into account the distribution within each field. However, the actual net revenue of each field depends on how many soil salinity classes are present in each field and the yield potential percentage of each class for a particular crop being considered. Therefore, the following equation is used to adjust the net revenue of each field:

$$Adjusted\ Net\ \text{Re}venue = \frac{\sum_{i=1}^{n} Net\ revenue \times Area\ of\ class \times \%\ \ of\ class yield\ potential}{Area\ of\ the\ field} \tag{2.6}$$

where: n represents the number of different yield potential classes, i.e. n represents five classes when field US04 is planted with alfalfa.

2.4 Model Validation

This study focused on three levels of soil salinity (low, moderate and high) and each level was represented by two fields. Out of each set of two fields, one field was used to construct the indicator variogram while the other was used for validation of the indicator variogram. Four different scenarios of planting alfalfa, corn, sorghum, and wheat were evaluated for each field. Therefore, four different indicator variograms were constructed for each field, and then applied to the other field in the same soil salinity level (validation field). Fields US01, US14, and US04 were used to construct the indicator variograms for low, moderate, and high soil salinity levels respectively. Fields US09, US80, and US10 were used for validation of the same levels of soil salinity, respectively. The criteria used for selecting a field for constructing the variogram or validating it was based on the range of soil salinity in each of the two fields. The field with the larger soil salinity range was chosen for constructing the indicator variogram while the other one was used for validation. Therefore, if the validation field has fewer classes for indicator variograms, the extra classes are removed.

2.5 Model Performance

IK performance with the different crops and fields is measured using the following criteria:

Model precision: The *RMSE* is used to measure the prediction precision (Dobermann et. al., 2006; Triantafilis et. al., 2001) and is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Z_i - Z_i^*)^2} \tag{2.7}$$

where Z_i is the observed value of the i^{th} observation, Z_i^* is the predicted value of the i^{th} observation, and n is the number of points collected.

The *RMSE* tends to place more emphases on larger errors and, therefore, gives a more conservative measure than the mean absolute error MAE.

Smoothing effect: Interpolation usually leads to a smoothing of the observations and thus to a loss of variance. To assess the ability of the interpolation method to preserve the variance, the ratio of the variance of the estimated values to the variance of the observed values is used (Haberlandt, 2006):

$$RVar = \frac{Var[Z_i^*(u)]}{Var[Z_i(u)]} \quad \textbf{(}2.8\textbf{)}$$

The closer *RVar* is to 1, the better the ability of the interpolation method to preserve the observed variance.

Model effectiveness: The effectiveness of the model was evaluated using a goodness-of-prediction statistic, G (Agterberg 1984; Kravchenko and Bullock 1999; Guisan and Zimmermann 2000; Schloeder et al. 2001). The *G*-value measures how effective a prediction might be relative to that which could have been derived by using the sample mean (Agterberg 1984):

$$G = \left(1 - \left\{\sum_{i=1}^{n}\left[Z_i - Z_i^*\right]^2 \Big/ \sum_{i=1}^{n}\left[Z_i - \bar{Z}\right]^2\right\}\right) \quad (2.9)$$

$\bar{Z}$ is the sample mean. A *G*-value equal to 1 indicates perfect prediction, a positive value indicates a more reliable model than if the sample mean had been used, a negative value indicates a less reliable model than if the sample mean had been used, and a value of zero indicates that the sample mean should be used.

2.6 Results

This section presents the process of selecting the indicator variograms of IK based on the Akaike Information Corrected Criteria (AICC) statistical parameter. Examples of IK maps for

different scenarios of crops and fields are provided. Examples of zones of uncertainty are also presented to quantify the risk associated with each of these zones. Finally, an estimate of the net economic revenue for each of the scenarios is provided.

Table 2.3: Akaike Information Corrected Criteria (AICC) of the Exponential, Gaussian, and Spherical variogram models for indicator kriging when evaluating alfalfa, corn, sorghum, and wheat as possible crops.

Field	Alfalfa			Corn		
	Sph.	Exp.	Gau.	Sph.	Exp.	Gau.
US01	55.7	55.6	55.0	72.3	72.4	72.0
US04	65.6	46.8	44.4	68.0	48.5	46.9
US14	63.9	58.5	61.0	67.6	67.7	68.8
	Sorghum			Wheat		
US01	61.5	61.5	54.5	N/A	N/A	N/A
US04	64.0	60.3	60.3	84.0	84.2	84.0
US14	62.5	62.4	63.3	64.5	64.6	64.5

*N/A: the total area of the field has 100% yield potential.

Table 2.3 shows the AICC values of the Exponential, Gaussian, and Spherical variogram models for the different combinations of crops and fields. The variogram model with the smallest AICC is considered the best. In most of the scenarios the Gaussian model performance is the best since the AICC values are the smallest. The performance of the Spherical and Exponential models is very similar. The average AICC values of the Spherical, Exponential, and Gaussian models for all the scenarios are: 66.3, 62.0, and 61.3 respectively. Fields US01, US04, and US14 were used to construct variograms for the different crop scenarios while fields US09, U10, and US80 were used for validation.

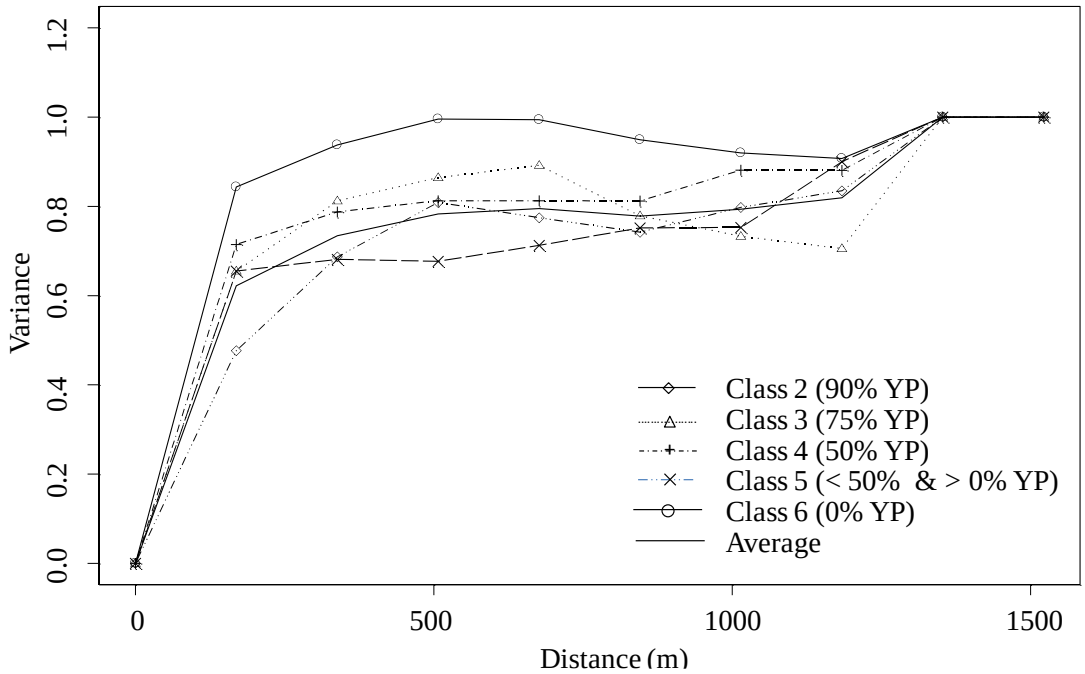


Figure 2.2: Example of indicator variograms for field US04 for alfalfa.

Figure 2.2 shows an example of the indicator variograms for field US04 for a scenario of planting alfalfa. From the data presented in Table 3, the AICC value of the Gaussian model is the smallest; and therefore it was used to construct the indicator variogram by sorting the collected soil salinity data for that field from low to high. Five classes were assigned to the sorted soil salinity data according to the percent yield potential of alfalfa to represent the following yield potentials: 90%, 75%, 50%, < 50% ~ > 0%, and 0%.

Table 2.4: Different classes and zones of uncertainty for the selected fields planted with different scenarios of growing alfalfa, corn, sorghum, and wheat were evaluated

	US01		US04		US09		US10		US14		US80	
	Alfalfa											
YP	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc
100%					3.83	0						
90%	11.68	0.24	18.10	0.21	25.53	0.07	1.92	0.24	3.01	0.22	7.08	0.16
75%	4.30	0.27	11.57	0.28			0.66	0.32	7.48	0.24	3.25	0.23
50%	0.22	0.17	32.56	0.17			0.57	0.38	2.04	0.28	0.53	0.20
<50%			20.54	0.17			0.72	0.23	0.20	0.32	0.40	0.25
0%							0.33	0				
	Corn											
YP	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc
100%					1.98	0.15						
90%					16.89	0.17						
75%	5.98	0.20	20.56	0.20	10.04	0.17	2.33	0.25	5.29	0.26	8.57	0.15
050%	8.04	0.19	18.54	0.15			0.43	0.27	6.11	0.18	2.08	0.21
<50%	2.18	0.18	29.74	0.19			0.43	0.27	1.34	0.13	0.61	0.22
0%			24.34	0			1.00	0				
	Sorghum											
YP	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc
100%	15.66	0.03	25.08	0.24	28.92	0	2.42	0.25	7.28	0.25	9.52	0.20
90%	0.35	0.25	4.40	0.35			0.15	0.32	3.27	0.25	0.71	0.23
75%	0.19	0.20	20.96	0.21			0.34	0.35	1.40	0.25	0.61	0.25
50%			24.80	0.22			0.44	0.21	0.79	0.32	0.32	0.28
<50%			10.85	0.18			0.72	0.24			0.09	0.29
0%			7.09	0.24			0.11	0.30				
	Wheat											
YP	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc	A(ha)	Unc
100%	16.20	0	40.66	0.29	28.92	0	2.87	0.23	12.28	0.18	10.98	0.21
90%			12.48	0.25			0.11	0.24	0.01	0	0.15	0.19
75%			13.32	0.26			0.11	0.16	0.34	0.26	0.13	0.18
50%			11.84	0.25			0.54	0.22	0.11	0.23		
<50%			9.06	0.28			0.56	0.24				
0%			5.83	0.22								

YP: Yield potential; Unc: Zone of Uncertainty percentage.

Table 2.4 shows the yield potential areas of each class and the corresponding zones of uncertainty for all the scenarios of the selected crops and fields.

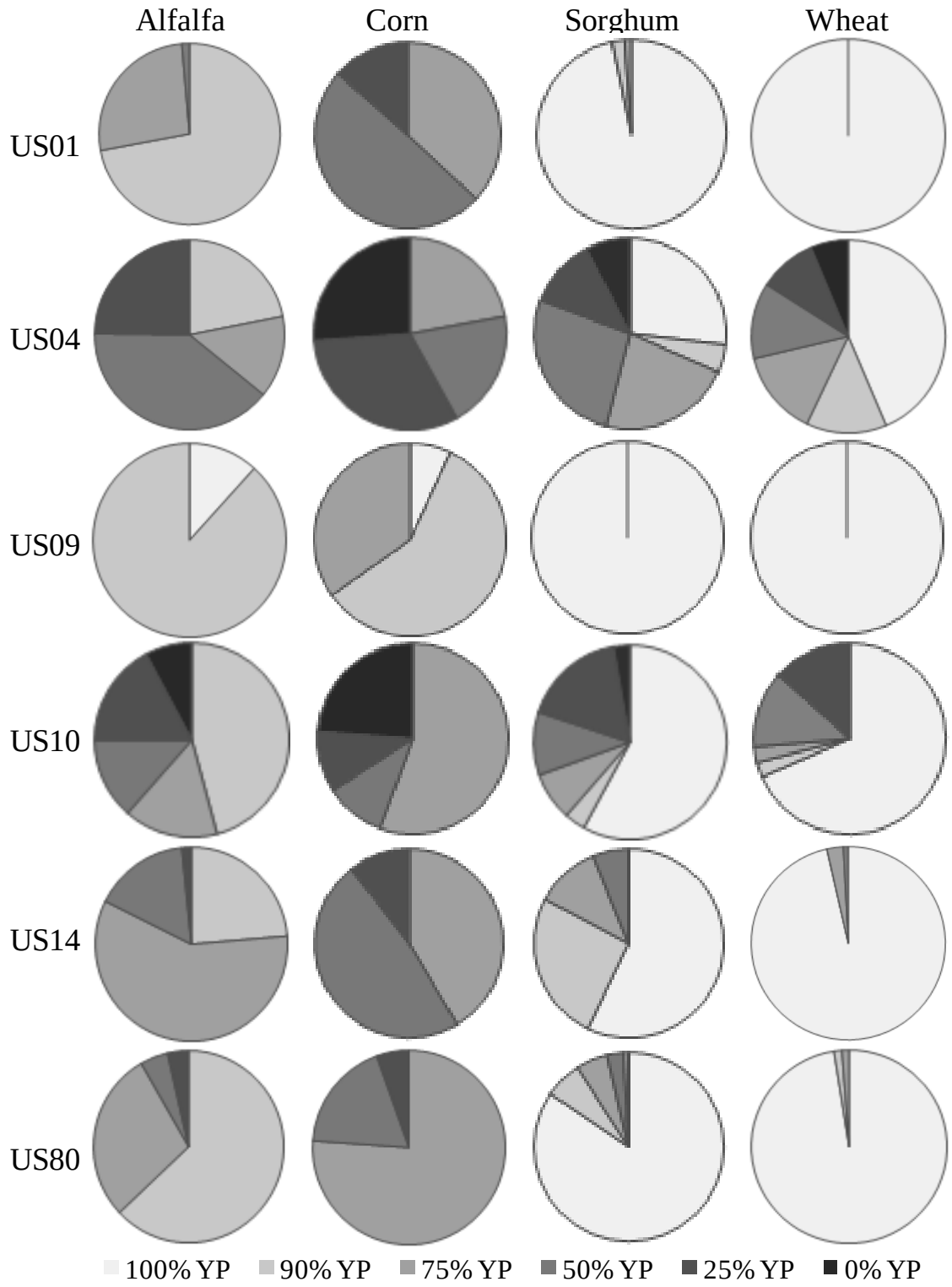


Figure 2.3: Pie charts showing different categorical kriging areas for the different fields when different crops are evaluated.

In addition, Figure 2.3 shows pie-charts that summarize all scenarios of the different combinations of crops and fields. The same color scheme used with the maps was also used to produce the pie-charts where colors go from light to dark to represent productivity from high to low. Both Table 4 and Figure 3 show that fields with low soil salinity ranges (US01 and US09) can reach the maximum production for all crops. However, with moderate and high salinity fields only sorghum and wheat start with 100% yield potential areas. Alfalfa has good production and in most scenarios, it starts with 90% yield potential areas. Corn has moderate production and in most cases, it starts with 75% yield potential areas.

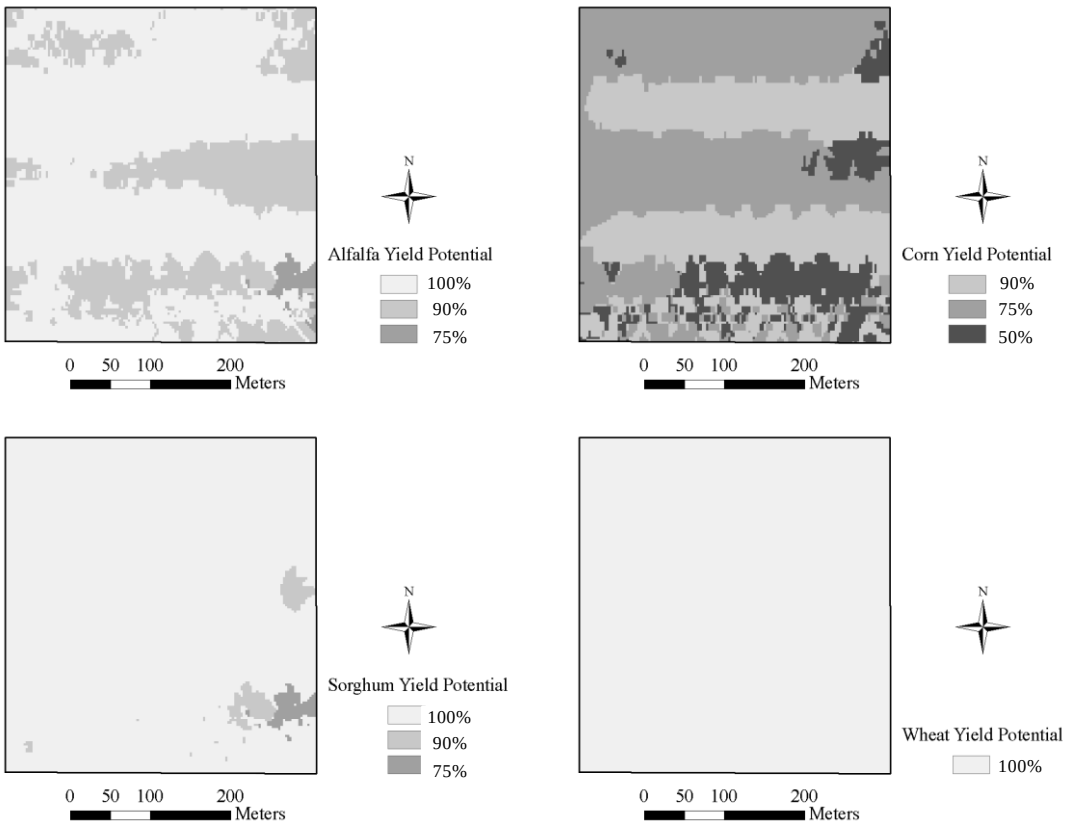


Figure 2.4: IK maps for field US01 (low soil salinity) when different crops are evaluated.

Figure 2.4, Figure 2.5, and Figure 2.6 show three examples of IK maps for fields US01, US14, and US04 which represent low, moderate, and high soil salinity ranges when alfalfa,

corn, sorghum, and wheat are assumed to be grown. Figure 2.4 shows the IK maps for field US01, which has low soil salinity. The whole area of field US01 can reach the maximum expected productivity (100% yield potential) for wheat, while the expected production of sorghum is quite high with the majority of the field having the potential to produce 100% of yield potential with small areas of 90% and 75% of yield potential. Alfalfa expected production is high with most of the field expected to produce between 90 and 100% of yield potential and very small areas expected to produce 75% of yield potential. Corns expected production is moderate where the production is between 90% and 50% of yield potential.

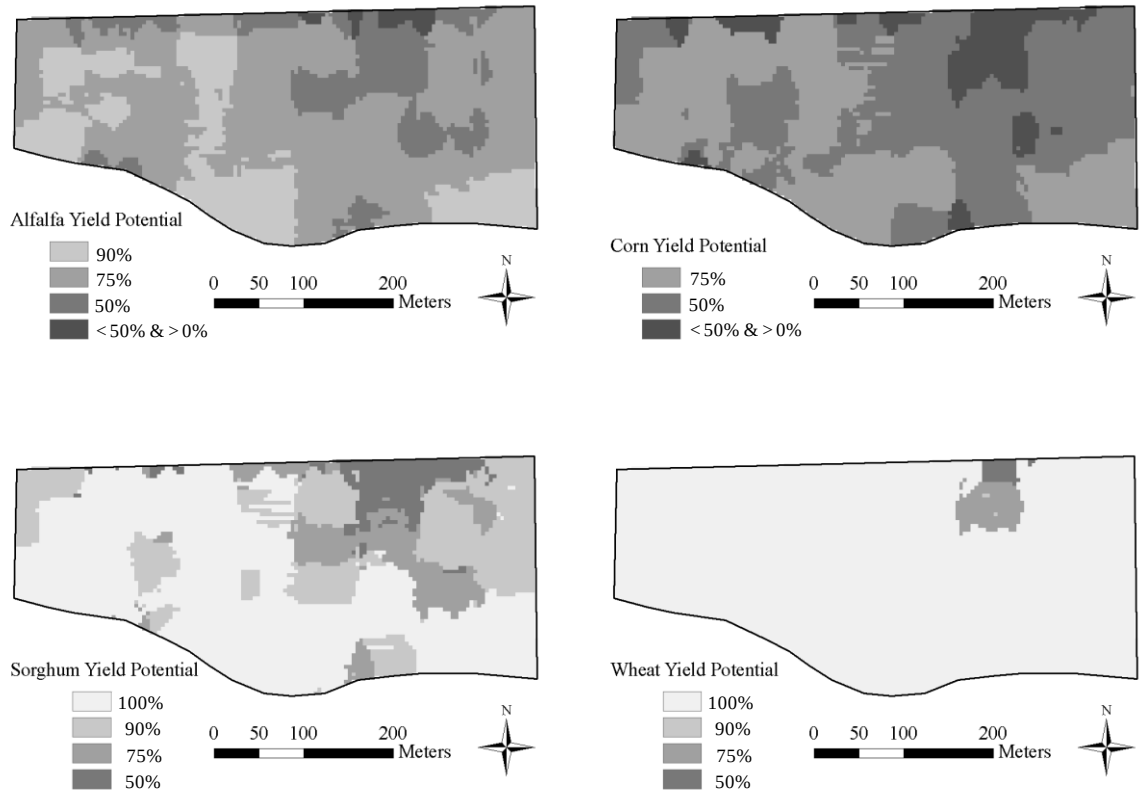


Figure 2.5: IK maps for field US14 (moderate soil salinity) when different crops are evaluated.

Figure 2.5 shows IK maps for field US14, with moderate soil salinity range when the scenarios of planting alfalfa, corn, sorghum, and wheat are applied. The wheat expected yield in field US14 is high with large areas represented by 100% of yield potential and very small

areas represented by 90% and 75% of yield potential. Alfalfas expected production is reasonably good where the expected yield production is between 90% and less than 50% of yield potential. Sorghum has moderate production where large areas in the field are represented by 100% and 90% of yield potential and some areas are represented by 75% and 50% of yield potential. Corn is moderate where the expected production is between 75% and less than 50% of yield potential.

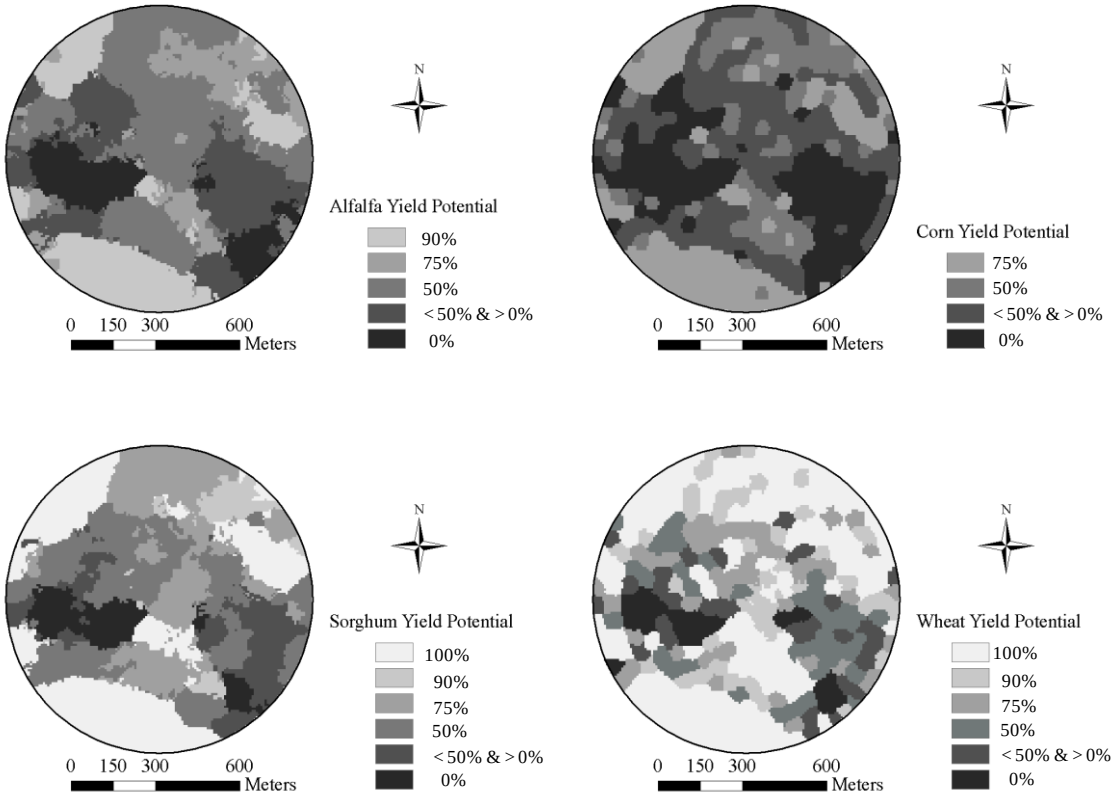


Figure 2.6: IK maps for field US04 (high soil salinity range) when different crops are evaluated.

Figure 2.6 shows IK maps for field US04, with high soil salinity range when the scenarios of planting alfalfa, corn, sorghum, and wheat are applied. Even though field US04 has relatively high soil salinity, the expected production of wheat is relatively high with a large percent of the area represented by 100% and 90% of yield potential. The expected production

for sorghum and alfalfa is moderate where the production of sorghum covers a range between 100% and 0% of yield potential while alfalfa covers a range between 90% and 0% of yield potential. Corn expected production is poor with a few areas represented by 75% of yield potential and the majority of the areas have 50% or less of yield potential.

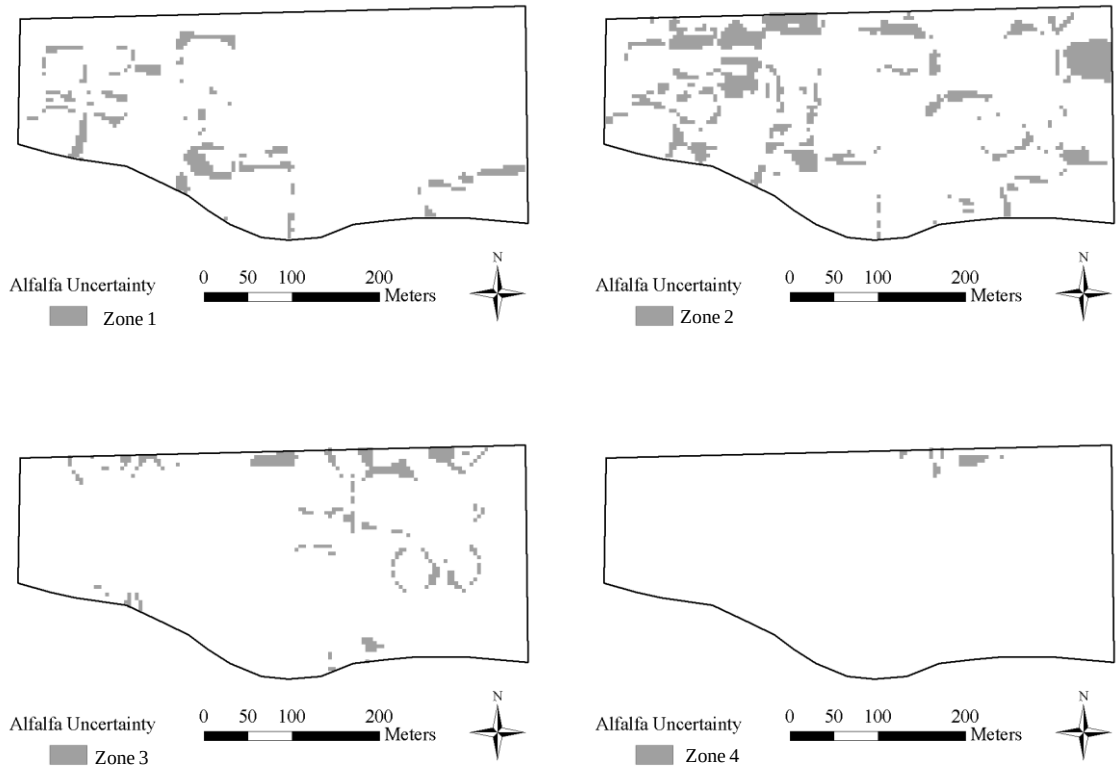


Figure 2.7: Zones of uncertainty for field US14 for alfalfa.

Figure 2.7 shows an example of zones of uncertainty for field US14 when the scenario of planting alfalfa is applied. One of the advantages of IK is that it can provide a risk-assessment tool for high-risk regions in a field. Figure 2.7 as well as Table 2.4 show how these areas can be quantified. As shown in Table 2.4, the areas of the different zones of uncertainty vary between 0% and 35% of the class area.

Table 2.5: Total revenue, cost, and net revenue per hectare ($/ha) of alfalfa, corn, sorghum, and wheat

Crop	Alfalfa	Corn	Sorghum	Wheat
Total revenue ($/ha)	1,780	1,780	381	863
Cost ($/ha)	751	724	161	403
Net Revenue ($/ha)	1,028	1,055	220	460

Table 2.6: Adjusted net revenue ($/ha) with and without risk of alfalfa, corn, sorghum, and wheat under the different conditions of soil salinity at the selected fields.

Field ID	Alfalfa		Corn		Sorghum		Wheat	
	Net Rev. ($/ha)		Net Rev. ($/ha)		Net Rev. ($/ha)		Net Rev. ($/ha)	
	Without Uncert.	With Uncert*	Without Uncert*	With Uncert.	Without Uncert*	With Uncert.	Without Uncert*	With Uncert.
US01	878	661	589	447	219	212	460	460
US04	511	407	364	297	141	108	346	250
US09	937	880	902	745	220	220	460	460
US10	658	478	522	266	169	74	380	168
US14	758	574	609	609	201	151	455	371
US80	837	686	714	599	211	168	458	363

*With Uncertainty - the net revenue was calculated using the percentage of the uncertainty **z**ones

Table 2.5 shows the total revenue, cost, and net revenue for alfalfa, corn, sorghum, and wheat based on the Colorado State University Extension (Agriculture and Business Management) 2007 crop budget estimates. The total revenue includes the final revenue of the crop without taking into account the costs. Net revenue is the revenue after the costs are taken into account. The net revenue and cost of alfalfa and corn are high while both are low for sorghum and wheat. For one hectare of alfalfa, in order to gain a net revenue of $1,028, a grower needs to spend $751 while they only need to spend $161 for sorghum but they only gain a net revenue of $220.

Table 2.6 shows the adjusted net revenue with and without risk according to the indicator kriging maps of yield potential of each crop based on the soil salinity thresholds of each field. The net revenue of the different crops has the following order: alfalfa, corn, wheat, and

sorghum. The net revenue of alfalfa and corn are highly affected by the soil salinity levels while sorghum and wheat are slightly affected.

Table 2.5 shows that there is a slight difference between the net revenue of alfalfa and corn while Table 2.6 shows that there is a significant difference in the adjusted net revenue for alfalfa and corn among different fields due to the sensitivity of these crops to salinity and the salinity levels in each field. The difference between the net revenue of alfalfa and corn is significant in all fields except for field US09. That is due to the fact that soil salinity in field US09 allows for 100% of yield potential and there is a big portion of the field with 90% of yield potential while the salinity levels in other fields allows only for 75% or less of yield potential. Uncertainty zones sometimes have significant impact and sometimes marginal impact. Therefore, taking uncertainty zones into consideration provides farmers with more support when making a selection on the crop that has the potential to generate higher net revenue.

Table 2.7: Performance parameters: *RMSE*, *RVar*, and *G* values of indicator kriging when evaluating alfalfa, corn, sorghum, and wheat as possible crops

	RMSE	RVar	G	RMSE	RVar	G
	US01			US04		
Alfalfa	0.59	0.95	-0.12	0.78	1.05	0.61
Corn	0.85	0.82	-0.49	0.06	1.00	1.00
Sorghum	0.46	0.54	0.02	1.00	1.04	0.57
Wheat	N/A	N/A	N/A	0.06	1.00	1.00
	US09 (Validation)			US10 (Validation)		
Alfalfa	0.60	0.40	-0.44	0.45	1.05	0.89
Corn	0.36	0.99	0.64	0.33	1.04	0.93
Sorghum	N/A	N/A	N/A	0.44	1.03	0.93
Wheat	N/A	N/A	N/A	0.61	1.06	0.84
	US14			US80 (Validation)		
Alfalfa	0.49	0.77	0.57	0.49	0.90	0.62
Corn	0.40	0.87	0.64	0.44	0.70	0.61
Sorghum	0.73	0.77	0.45	0.69	0.79	0.47
Wheat	0.63	0.42	0.02	0.46	0.25	0.38

Table 2.7 shows the performance parameter values of IK when evaluating alfalfa, corn, sorghum, and wheat as possible crops under different soil salinity conditions. N/A means that the whole field can produce 100% of yield potential which applies to the crops with high tolerance to soil salinity when planted in the fields with low soil salinity levels. G values are positive for the fields with high and moderate soil salinity (US04, US10, US14, and US80) while it does not perform as well in fields with low soil salinity (US01 and US09). In some cases the G value reaches 1 or close to 1 which means that the model is perfect such as corn and wheat in US04 and corn in US09. The RVar values are closest to 1 in fields with a high range of soil salinity (US04 and US10). In cases where the RVar values are small such as wheat in US14 and US80, this means that the model was not able to overcome the smoothing effects problem. The RMSE values are reasonable in all fields since all values are equal or less than 1.

As previously mentioned, several studies have been carried out using IK applications in soil science. However, none of them used IK as a tool to manage soil salinity with crop productivity to maximize the benefit. As presented in the results above, IK was used to determine which crops to grow in order to maximize the potential net benefits taking into account the variability of soil salinity in the fields. In this study, different thresholds were made in the soil salinity data based on the salt tolerance of different crops. Therefore, instead of representing different soil salinity thresholds, the resulting indicator variograms represent different yield potentials. To improve the usability of the resulting maps, the different areas of yield potential as well as the corresponding zones of uncertainty of produced maps were quantified and evaluated. The yield potential and the uncertainty zones can provide a management tool for selecting the crops that have the potential to generate the highest yields. However, high yield of specific crop is not guaranteed to provide the maximum net revenue

due to the market price, therefore, the net revenue values were evaluated and adjusted to incorporate the variability in soil salinity in fields.

2.7 Crop Selection Recommendations

The following are practical recommendations for farmers and technicians to be used as guidelines for crop selection based on the variability of soil salinity:

Case 1: Low level of soil salinity where no significant impact on most crops. In this case, no restrictions for crop selection and high profit crops should be considered. Alfalfa would be strongly recommended as the first choice while corn would be recommended as the second choice. The expected net revenue from examples presented in this study for low levels of soil salinity can provide a net revenue for alfalfa of approximately $900/ha while the expected net revenue for corn is approximately $750/ha. Wheat and sorghum are not recommended since the net revenue for both of them is low compared to those of alfalfa and corn.

Case 2: Moderate level of soil salinity where the impact of soil salinity is slight on moderate sensitive crops such as alfalfa and corn and no impact on moderate tolerant crops such as wheat and sorghum. Alfalfa is strongly recommended as the best choice. The examples in this study for moderate level of soil salinity fields shows that the net revenue of alfalfa would be $800/ha while the net revenue of corn would be $660/ha.

Case 3: High level of soil salinity where its impact on moderate sensitive crops is significant while the impact on moderate tolerant crops such as wheat and sorghum is slight. Even though the high level of soil salinity has significant impact on moderate sensitive crops, these crops still provide high net revenue. The examples presented for high soil salinity levels shows a net revenue of approximately $580/ha while the second choice would be corn with a net revenue of approximately $440/ha.

It is clear that the market price has greater impact on crop selection rather than soil salinity impact. As presented earlier, the expected yield of alfalfa and corn is less than the expected yield of wheat and sorghum. However, the market price of alfalfa and corn is higher than those of wheat and sorghum which makes alfalfa and corn better selections. There was a significant difference between the prices of alfalfa and corn versus wheat and sorghum in spite of the reduction in the productivity of alfalfa and corn due to the impact of soil salinity. To make the results of this study more general, the crops presented: alfalfa, corn, wheat, and sorghum can be replaced by crops with similar tolerance to soil salinity in order to accommodate other crop selections. Therefore, since alfalfa and corn are moderate sensitive to soil salinity, they can be replaced by some other crops that have the same tolerance to soil salinity such as broccoli, cabbage, celery, cucumbers, and tomatoes. Wheat and sorghum are moderate tolerant to soil salinity, therefore, they can be replaced by grapes, pineapples, squash, and sugar beets.

2.8 Conclusion

A geostatistical approach (indicator kriging), which makes no assumptions regarding the normality of the dataset and is essentially a non-parametric model, was used in this study. IK uses the behavior and correlation structure of the transformed data instead of the data itself. It uses a series of threshold values between the smallest and largest data values in the dataset. This advantage allows incorporating soil salinity with crop yield potential where soil salinity values were transformed into yield potential classes. Therefore, IK was successful in generating classified maps of expected yield potential of the main crops grown in the study area. In addition to generating the classified maps, the results show that IK has the power to generate the corresponding zones of uncertainty. Providing farmers with information

regarding the uncertainty associated with each zone, which can help them in their decision making process. The fields used in this study were selected to represent different levels of soil salinity from low, to moderate, to high. Soil salinity values for some fields were homogeneous with small ranges such as field US09 while others had high ranges such as US04. The outcomes of this study show how to obtain the maximum productivity for a particular field under its current soil salinity conditions. However, to reach a high potential productivity may be a target; but to maximize the expected net revenue under different soil salinity conditions should be the optimal target. The results presented in this study show that wheat and sorghum provide the highest expected yield potential while alfalfa and corn provide the highest expected net revenue under the same conditions of soil salinity. Therefore, this study can be used to develop management strategy guidelines for crop selections in order to maximize the economic benefit based on the soil salinity of fields.

2.9 REFERENCES

Agterberg, F.P. (1984). "Trend surface analysis in spatial statistics and models". *Ed. G.L. Gaile and C.J. Willmott* 147-171. (Reidel: Dordrecht, The Netherrlands).

Ayers, R. S., and Westcot, D. W. (1976). *Water Quality for Agriculture*. Irrigation and Drainage Paper 29. FAO, Rome.

Bierkens,M.F.P., Burrough, P.A., (1993a). "The indicator approach to categorical soil data. I. Theory." *J. Soil Sci.* **44**, 361–368.

Bierkens,M.F.P.,Burrough, P.A., (1993b). "The indicator approach to categorical soil data. II. Application tomapping and land use suitability analysis." *J. Soil Sci.* **44**, 369–381.

Castrignano A., Goovaerts P., Lulli L., and Bragato G. (2000). "A geostatistical approach to estimate probability of occurrence of Tuber melanosporum in relation to some soil properties." *Geoderma*, **98**, 95 –113.

Cain, D. (1985). Quality of the Arkansas River and irrigation-return flows in the lower Arkansas River Valley, Colorado. U.S. Dept. of the Interior. 85 p.

Cressie N. (1993). *Statistics for spatial data*. New York: Wiley.

Dobermann, A., and Ping, J. L. (2004). "Geostatistical Integration of Yield Monitor Data and Remote Sensing Improves Yield Maps." *Agron. J.*, **96**, 285-297.

Douaik, A., Van Meirvenne, M., and Toth, T. (2003). "Spatio-temporal kriging of soil salinity rescaled from bulk soil electrical conductivity." *In: Sanchez-Vila X, Carrera J, Gomez-Hernandez J (eds) Quantitative geology and geostatistics, GeoEnv IV: 4th European conference on geostatistics for enironmental applications*, Kluwer Academic Publishers: Dordrecht, The Netherlands, pp. 413–424.

Eldeiry, A., and Garcia, L. A. (2008). "Detecting Soil Salinity in Alfalfa Fields using Spatial Modeling and Remote Sensing." *Soil Sci. Soc. Am. J.*, **72(1)**, 201-211.

Eldeiry, A., Garcia, L. A., and Reich, R. M. (2008). "Soil Salinity Sampling Strategy Using Spatial Modeling Techniques, Remote Sensing, and Field Data." *J. Irrig. Drain. Engin.*, **134(6)**, 768-777.

Figueira, R. Tavares, P. C., Palma, L., Beja, P., and Sergio, C. (2009). "Application of IK to the complementary use of bioindicators at three trophic levels." *Environmental Pollution,* **157,** 2689-2696.

Fogg, G. E., LaBolle, E. M., and Weissman, G. S. (1999) "Groundwater vulnerability assessment: hydrologic perspective and example from Salinas Valley, California." *Assessment of non-point source pollution in the Vadose zone,* American Geophysical Union, pp. 45 –61 (Geophysical Monograph 108).

Frasier, W.M., Waskom, R.M., Hoag, D.L., Bauder, T.A., (1999). Irrigation Management in Colorado: Survey Data and Findings. Technical Report TR99-5 Agricultural Experiment Station, Colorado State University. Fort Collins, CO.

Gates, T. K., Burkhalter, J. P., Labadie, J. W., Valliant, J. C., and Broner, I. (2002). "Monitoring and modeling flow and salt transport in a salinity-threatened irrigated valley," *Journal of Water Resources Planning and Management,* **128(2)**, 87-99.

Ghassemi, F., Jackeman, A. J., and Nix, H. A. (1995). *Salinization of land and water resources: human causes, extent, management and case studies.* CAB International, Wallingford Oxon, UK.

Goovaerts, P., (1994). Comparative performance of indicator algorithms for modeling conditional probability distribution functions. *Mathematical Geology* **26**, 389–411.

Greenway, H., Munns, R. (1980). "Mechanisms of salt tolerance in nonhalophytes." *Annu. Rev. Plant Physiol.*, **31**, 149-190.

Guisan, A., and Zimmermann, N. E. (2000). "Predictive habitat distribution models in ecology." *Ecological Modeling,* **135**, 47-186.

Haberlandt, U. (2006). "Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event." *Journal of Hydrology,* **322**, 144-157.

Hillel, D. (2000). Salinity management for sustainable irrigation: integrating science, environment, and economics. The World Bank: Washington D.C.

Istok, J. D., and Pautman, C. A. (1996). "Probabilistic assessment of groundwater contamination 2.Results of case study." *Ground Water,* **34(6)**, 1051 –1064.

Journel, A. G. (1983). "Nonparametric Estimation of Spatial Distributions." *Mathematical Geology*, **15(3)**, 445 – 468.

Juang, K. W., and Lee, D. Y. (1998). "Simple indicator kriging for estimating the probability of incorrectly delineating hazardous areas in a contaminated site." *Environ. Sci. Technol.*, **32**, 2487 – 2493.

Kravchenko, A., and Bullock, D. G. (1999). "A comparative study of interpolation methods for mapping soil properties." *Agron. J.*, **91,** 393–400.

Lin, Y. P., Chang, T. K., Shih, C. W., and Tseng, C. H. (2002). "Factorial and indicator kriging methods using a geographic information system to delineate spatial variation and pollution sources of soil heavy metals." *Environ. Geol.*, **42**, 900 –909.

Liu, C., Jang, C., and Liao, C. (2004). "Evaluation of arsenic contamination potential using indicator kriging in the Yun-Lin aquifer (Taiwan)" *Science of the Total Environment,* **321**, 173-188.

Lyon, S.W., Lembo Jr., A.J., Walter, M.T., and Steenhuis, T.S. (2006). "Defining probability of saturation with indicator kriging on hard and soft data." *Advances in Water Resources*, **29**, pp. 181-193.

Marinoni, O. (2003). "Improving geological models using a combined ordinary-indicator kriging approach." *Engineering Geology*, **69,** 37-45.

McCord, J.T., Gotway, C.A, Conrad, S.H. (1997). "Impact of geologic heterogeneity on recharge estimation using environmental tracers: numerical modeling investigation." *Water Resour. Res.*, **33(6)**, 1229 –1240.

McQuarrie, A.D.R., and Tsai, C.-L., (1998). "Regression and Time Series Model Selection". Singapore, World Scientific.

Miles, D.L. (1977). "Salinity in the Arkansas Valley of Colorado. Region VIII." Environmental Protection Agency, Denver. 80 p.

Oyedele, D. J., Amusan, A.A., and Obi, A. (1996). "The use of multiple-variable indicator kriging technique for assessment of the suitability of an acid soil for maize." *Tropical agriculture*, **73(4),** 259 –263.

Panahi, A., Cheng Q., and Bonham-Carter, G.F. (2004). "Modelling lake sediment geochemical distribution using principal component, indicator kriging and multifractal power-spectrum analysis: a case study from Gowganda, Onatario." *Geochemistry: Exploration, Environment, Analysis*, **4**, 59-70.

Postel, S. (1999). *Pillar of Sand: Can the Irrigation Miracle Last?* W.W. Norton and Co.: New York, NY.

Reis, A.P., Ferreira Da Silva; E., Sousa, A.J., Matos, J., Patinha, C., Abenta, J. & Cardoso Fonseca, E. (2005) "Combining GIS and Stochastic Simulation to Estimate Spatial Patterns of Variation for Lead at the Lousal Mine, Portugal." *Land Degradation Development*, **16(2)**, 229-242.

Reis, A.P., Sousa, A.J., Ferreira Da Silva, E. and Cardoso Fonseca, E. (2005) "Application of geostatistical methods to arsenic data from soil samples of the Cova dos Mouros mine (Vila Verde-Portugal)." *Environmental Geochemistry and Health*, **27(3),** 259-270.

Richmond, A. (2001). "An alternative implementation of indicator kriging." *Computers & Geosciences,* **28**, 555-565.

Smith, J.L., Halvorson, J.J., and Papendick, R.I. (1993). "Using Multiple-Variable Indicator Kriging for Evaluating Soil Quality." *Soil Sci. Soc. Am. J.*, **57,** 743-749.

Soares, A., (1992). "Geostatistical Estimation of Multi-Phase Structures." *Math Geol.*, **24(2)**, 149-160.

Solow, A.R., (1986). "Mapping by Simple Indicator Kriging." *Mathematical Geology*, **18(3)**, 335-351.

Schloeder, C.A, Zimmermann, N.E, and Jacobs, M.J. (2001). "Comparison of methods for interpolating soil properties using limited data". *Soil Sci. Soc. Am. J.*, **65**, 470-479.

Tanji, K. K., ed. (1990). Agricultural Salinity Assessment and Management. ASCE Manuals and Reports on Engineering Practice No. 71, ASCE, New York, N.Y.

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

Triantafilis, J., Odeh, I. O. A., and McBratney, A.B. (2001). "Five Geostatistical Models to Predict Soil Salinity from Electromagnetic Induction Data Across Irrigated Cotton." *Soil Science Society of America Journal*, **65**, 869-878.

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

van Meirvenne, M., and Goovaerts, P. (2001). "Evaluating the probability of exceeding a site-specific soil cadmium contamination threshold." *Geoderma*, **102**, 75 –100.

Western, A.W., Bloschl, G., and Grayson, R.B. (1998). "How well do indicator variograms capture the spatial connectivity of soil moisture?" *Hydrological processes*, **12(12),** 1851-86.

Wittler, J.M., Cardon, G.E., Gates, T.K., Cooper, C.A., and Sutherland, P.L. (2006). "Calibration of Electromagnetic Induction for Regional Assessment of Soil Water Salinity in an Irrigated Valley." *J. Irrig. Drain. Eng.*, **132(5)**, 436-444.

3 USING DISJUNCTIVE KRIGING AS A QUANTITATIVE APPROACH TO MANAGE SOIL SALINITY AND CROP YIELD

3.1 Summary

Disjunctive kriging (DK) is a nonlinear geostatistical model that provides unbiased estimates of the conditional probability (CP) that the true value of the property of interest doesn't exceed a defined threshold. It has important implications in aiding management decisions by providing growers with a quantitative input that can be used for evaluating the variability of the crop productivity at different zones in fields. The objectives of this study are: 1) identify the yield potential percentage (YP%) for several crops at different zones in fields under multiple soil salinity thresholds; 2) evaluate the excepted YP% of whole fields for several crops under multiple soil salinity thresholds; and 3) provide guidelines to growers to help them decide which crops to grow. To achieve these objects, the DK technique was applied to data from a project conducted in the south eastern part of the Arkansas River Basin in Colorado to generate CP maps. Two datasets of soil salinity (316 and 136 points) collected in two fields in 2004 and 2005 were used to generate the CP maps and to evaluate different scenarios of the expected YP% of several crops at multiple soil salinity thresholds. These datasets were selected to represent a wide range of soil salinity conditions in order to be able to evaluate a wide variety of crops (larger set of crops than those grown in the study area) according to their soil salinity tolerances. The following crops were used for evaluation: field crops (barley, sorghum, and corn); fruit crops (pomegranate, apples, and strawberries); vegetable crops (beets, tomatoes, and lettuce); and forage crops (barley (hay), crested wheat grass, and alfalfa). This selection was set to include three crops of each type to represent high, moderate, and low soil salinity tolerances. Scenarios were created for each of the above mentioned crops and the DK technique was applied to each scenario in order to generate CP

maps and evaluate the expected YP%. The results of this study show that the CP maps generated using the DK technique give an accurate characterization and quantification of the different zones of the fields. CP maps can be used to assess the expected YP% of whole fields for several crops under multiple soil salinity thresholds. Upon the knowledge of the YP% of different areas, a management decision action can be taken to manage the productivity of a field by selecting another crop or adjusting the inputs such as fertilizer, seeding rates and herbicides in low productivity areas.

3.2 Introduction

Disjunctive Kriging (DK), unlike other geostatistical methods such as ordinary kriging, can be used as a quantitative method for making management decisions if the conditional probability (CP) information is available. DK has several advantages over linear estimation methods. It provides a more accurate estimate of the property of interest and can generate an estimate of the CP for that property (Yates and Yates, 1988). This CP can be used as an input to a management decision making model to provide a quantitative means for determining whether management actions are necessary (Yates and Yates, 1988). Whenever the value of a property in a region is larger than the cutoff level at a probability equal to, or greater than, the critical probability level, this indicates that an action should be taken. Such management decisions may often be based on threshold values of a soil property. There may be threshold concentrations of contaminants specified by regulators that land managers are obliged to maintain. The management of soil nutrients may also be based on threshold values. For example, if the concentration of available (Bray-1) phosphorus in the soil is larger than 15 mg kg^{-1} then no phosphorus is needed according to the University of Nebraska recommendations (Ferguson et al., 2000). There are other examples where threshold values of other soil

properties are of importance for management. If the concentration of cobalt in the pasture soils of Scotland is smaller than 0.25 mg kg^{-1} then action should be taken to avoid cobalt deficiency in grazing livestock (Webster and Oliver, 1989). Land use planning may also refer to threshold values of soil properties. Wood et al. (1990) used a DK technique to estimate and map the soil salinity in the Bet Shean Valley of Israel from measurements of electrical conductivity. Zirschky (1985), Zirschky and Harris (1986) and Zirschky et al. (1985) investigated the use of geostatistics for determining reclamation strategies for the cleanup of hazardous waste sites. The kriged estimates of the concentrations of contaminants may be used to plan soil remediation, for example, estimates of the concentration of a nutrient may be used to plan spatially variable application of fertilizers (Schepers et al., 2000). Russo (1984a, b) described a method for using geostatistics to aid in managing the soil salinity of a heterogeneous field.

In addition to kriging techniques, other authors have used delineation of management zones for either soil salinity or yield management. Fridgen et al. (2000) used elevation, soil salinity, and slope to create management zones for wheat. Fleming et al. (1999) used bare soil color, farmer's perception of yield, and field topography to classify fields into three productivity zones (high, moderate, and low). Fraisse et al. (1999) used cluster analysis to identify areas that have similar landscape attributes, soil properties and plant parameters, to quantify patterns of variability and to reduce the empirical nature of defined management zones. Stafford et al. (1998) used fuzzy clustering of combined yield monitor data to divide a field into potential management zones. Boydell and McBratney (1999) divided a field into management zones using cotton yield estimates from satellite imagery.

Most of the previous studies that used geostatistical techniques were able to provide different approaches to assess soil salinity. However, most of these studies do not provide techniques that integrate soil salinity and crop yield to improve crop production.

Geostatistical techniques have been used for management of soil nutrients, land use and reclamation (Ferguson et al., 2000; Webster and Oliver, 1989; Wood et al., 1990; Zirschky, 1985; Zirschky and Harris, 1986; and Zirschky et al., 1985; Schepers et al., 2000). Only a few studies have utilized geostatistical techniques to manage soil salinity (Eldeiry and Garcia, 2011; Wood et al., 1990; Russo,1984a; and Russo,1984b). Eldeiry and Garcia (2011) used Indicator Kriging (IK), a non-linear technique, for soil salinity and yield management to maximize the economic benefits. They applied IK to different scenarios of crops and soil salinity thresholds to generate maps that show the expected percent yield potential areas and the corresponding zones of uncertainty. The DK, a non-linear technique, is used in this study, to provide unbiased estimates of the conditional probability (CP) that the true value of the property of interest doesn't exceed a defined threshold. Even though two different techniques (DK and IK) have been used in these two studies, both provide management tools to maximize the crop productivity under the current soil salinity conditions. The main contributions of this study are: 1) several crops are evaluated under different soil salinity thresholds which provide growers with a variety of crop selections; 2) CP maps were generate which can be used to quantify the variability of YP% in different soil salinity zones of fields; and 3) the CP maps can be used as management tools to increase crop productivity based on the current soil salinity of different fields.

3.3 Data and Methodology

3.3.1 Study Area

The study area is located in the south eastern part of the Arkansas River Basin in Colorado near the cities of Rocky Ford and La Junta (Figure 3.1). Farmers in this area are facing decreasing crop yields due in part to high levels of salinity in their irrigation water. In

some areas, land is being taken out of production due to unsustainable crop yields. This is due in part to the fact that the Arkansas River is one of the most saline rivers in the United States (Tanji, 1990; Miles, 1977). Farmland along the lower Arkansas River Basin has been continuously irrigated since the 1870's and began to develop shallow, saline water tables by the beginning part of the twentieth century (Miles, 1977). Average water table depths in this region have risen towards the surface approximately 0.3 – 1.3 m between 1969 and 1994 (Cain, 1997) which has only exacerbated the salinity problems because of increasing amounts of upflux of saline groundwater. In a survey of the region, 68% of producers stated that high salinity levels were a significant concern (Frasier et al., 1999). Crop yield reduction due to salinity in fields in the Lower Arkansas Valley has been estimated to be between 0 and 75% with a total revenue loss ranging from $0-$750/ha based on 1999 crop prices (Gates et al., 2002).

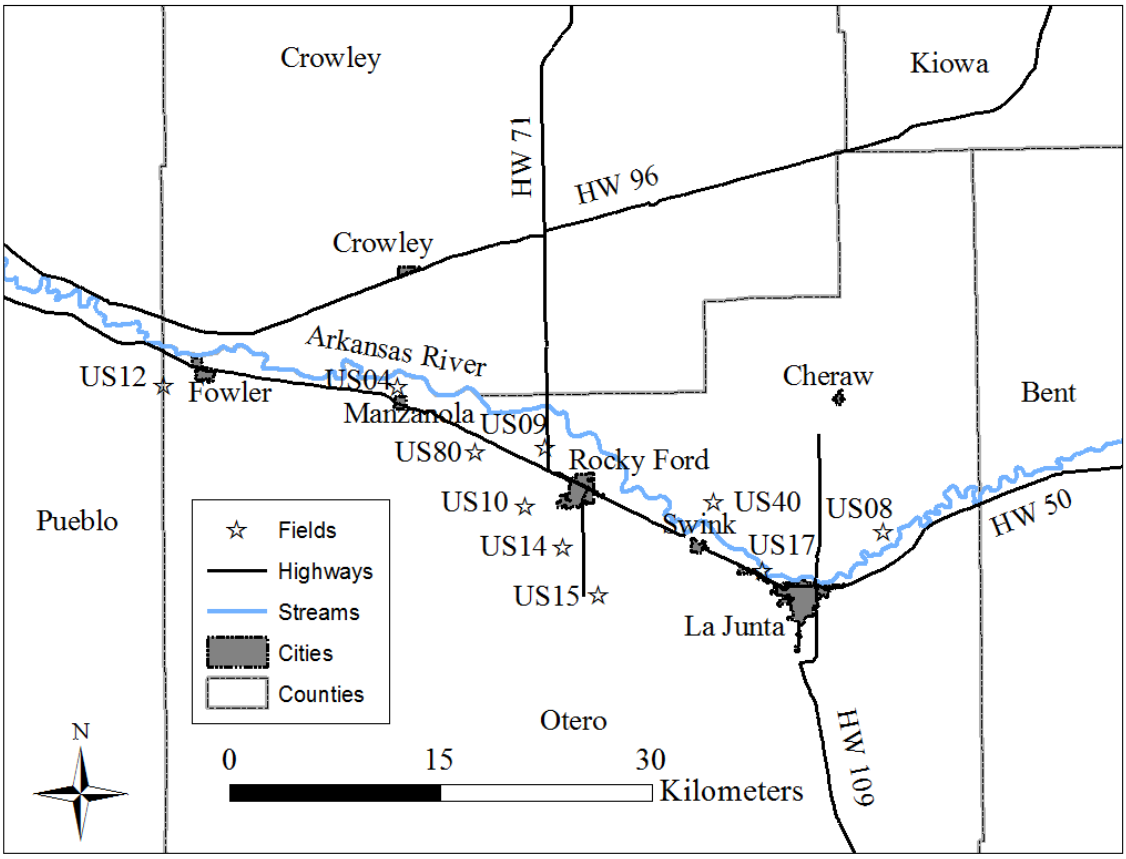


Figure 3.1: The study area in the south eastern part of the Arkansas River Basin in Colorado.

3.3.2 Selected fields and crops

Two datasets of soil salinity points (316 and 136) were collected in two fields during the 2004 and 2005 growing seasons. These two datasets were selected to represent a high range of soil salinity which allows evaluating a wide variety of crops with different soil salinity tolerances. The first dataset consists of 316 points with a minimum soil salinity value of 2.38 dS/m, a maximum value of 41.23 dS/m, and a variance of 42.21 dS/m. The second dataset consists of 132 points with a minimum soil salinity value of 3.04 dS/m, a maximum soil salinity value of 31.26 dS/m, and a variance of 31.38 dS/m. Soil salinity data were collected using EM-38 electromagnetic probes and the location of the samples was determined using global position system (GPS) units. The EM-38 electromagnetic probes provide vertical and horizontal readings while the GPS units provide X and Y coordinates for each sample point.

A calibrated equation to convert the EM-38 electromagnetic probe readings to EC (dS/m) was developed for the study area by Wittler et al. (2006) and was used in this study. Soil moisture content and soil temperature were used for the soil salinity calibration equation. A detailed description of using the EM-38 electromagnetic probe in combination with GPS in collecting soil salinity can be found in Eldeiry and Garcia (2008) and Eldeiry et al. (2008). The evaluated crops were selected to represent field, fruit, vegetable, and forage crops where three crops of each category were selected to represent high, moderate, and low soil salinity tolerances. The following crops were evaluated as part of this study: field crops (barley, sorghum, and corn); fruit crops (olive, apples, and strawberries); vegetable crops (beets, tomatoes, and lettuce); and forage crops (barley (hay), crested wheat grass, and alfalfa). Different scenarios using each of these crops were created based on the soil salinity thresholds for each of them. These scenarios provide growers with a wide selection of crops according to the level of soil salinity in their fields. In addition to the selected crops in this study, other crops can be evaluated based on their similarity in soil salinity tolerance to one of the crops evaluated in this study.

Table 3.1: Soil salinity threshold values (dS/m) of different YP% for the selected crops.

Crop		YP %				
Common name	Botanical name	100	100-90	90-75	75-50	50-0
Field Crops		Soil Salinity (dS/m)				
Barley	Hordeum vulgare	8.0	10.0	13.0	18.0	28.0
Sorghum	Sorghum bicolor	4.0	5.1	7.2	11.0	18.0
Corn	Zea mays	1.7	2.5	3.8	5.9	10.0
Fruit Crops						
Olive	Olea europaea	2.7	3.8	5.5	8.4	14
Apples	Pyrus malus	1.7	2.3	3.3	4.8	8.0
Strawberries	Fragaria spp.	1.0	1.3	1.8	2.5	4.0
Vegetable Crops						
Beets	Beta vulgaris	4.0	5.1	6.8	9.6	15
Tomatoes	Lycopersicon esculentum	2.5	3.5	5.0	7.6	12.5
Lettuce	Lactuca sativa	1.3	2.1	3.2	5.2	9.0
Forage Crops						
Barley (hay)	Hordeum vulgare	6.0	7.4	9.5	13.0	20.0

CWG	Agropyron desertorum	3.5	6.0	9.8	16.0	28.5
Alfalfa	Hedicago sativa	2.0	3.4	5.4	8.8	15.5

Table 3.1 shows the YP% and the corresponding soil salinity for the selected crops from field, fruit, vegetable, and forage crops. The YP% values according to soil salinity levels were adapted from Ayers and Westcot, (1976). They mentioned that for barley during the germination and seedling stages, soil salinity should not exceed 4 to 5 dS/m except for certain semi-dwarf varieties. However, Storey and Jones (1978) mentioned that barley is most sensitive to salinity at germination and young seedling stage and that it exhibits increased tolerance with age. Salinity tolerance at germination and seedling stages determines the stand density in the field under saline conditions. Therefore, the impact of salinity on barley during germination and seedling stages can be mitigated by increasing the seed density. Ayers and Westcot (1976) also mentioned that for beets, during germination electrical conductivity should not exceed 3 dS/m. Many crops have little tolerance for salinity during seed germination, but significant tolerance during later growth stages**.** Table 1 shows the significant impact of soil salinity on productivity and how some crops can reach high productivity while others barely grow under the same conditions. For example, at a specific area of a field where soil salinity is 8.0 dS/m, the expected yield of barley is 100% while the expected yield of apples would be between 50% and 0%.

3.3.3 DK equations:

A brief description is provided in this study to explain the basic equations of DK. A more comprehensive explanation can be found in Matheron (1976), Journel and Huijbregts (1978), Yates et al. (1986a, b), and Yates (1986).

To obtain the DK estimator, the original soil salinity data must be transformed into a new variable, $Y(x)$, with a standard normal distribution where pairs of sample values are bivariate normal. The function, $\phi[Y(x)]$, which describes this transformation is:

$$\phi[Y(x)] = Z(x) = \sum_{k=0}^{\infty} C_k H_k[Y(x)] \tag{3.1}$$

where the values for $Y(x)$ are obtained by taking the inverse of the data, $Y(x) = \phi^{-1}[Z(x)]$ and $H_k[Y(x)]$ is a Hermite polynomial of order *k*. The C_k'*s* are the Hermitian coefficients, which are determined using the properties of orthogonality, and are generally determined using numerical integration, as follows:

$$C_k = \frac{1}{k!\sqrt{(2\pi)}} \sum_{i=1}^{j} w_i \phi(v_i) H_k(v_i) \exp[-v_i^{\,2}/2] \tag{3.2}$$

where v_i, and w_i, are the abscissa and weight factors for Hermite integration (Hochstrasser, 1965).

The DK estimator is calculated from a sum of unknown functions of the transformed sample values, $Y(x_i)$. It is required that each unknown function, $f_i[Y(x_i)]$ depend on only one transformed value, $Y(x_i)$. The DK estimator is calculated using the following equation:

$$Z^*_{DK}(x_o) = \sum_{i=1}^{n} f_i[Y(x_i)] = \sum_{i=1}^{n} \sum_{k=1}^{\infty} f_{ik} H_k[Y(x_i)] \tag{3.3}$$

where f_i is the unknown function with respect to the transformed variable, and *n* is the number of samples.

An unbiased estimator with the minimum estimation variance can be obtained using the following equations:

$$Z^*_{DK}(x_o) = \sum_{k=0}^{K} C_k H^*_k[Y(x_o)] \tag{3.4}$$

where

$$H_k^*[Y(x_o)] = \sum_{i=1}^{n} b_{ik} H_k[Y(x_i)] \tag{3.5}$$

where the series in Eq. [4] has been truncated to K terms and b_{ik} are the DK weights. The $H_k^*[Y(x_o)]$ represents the estimated value of the Kth Hermite polynomial at the estimation location. The sum of these estimates multiplied by the coefficient, C_k [which transforms $Y(x)$ into $Z(x)$] makes up the DK estimate at x_o. To obtain an estimated value for the Hermite polynomial, the DK weights, b_{ik}, must be found by solving the linear kriging equation for each k as follows:

$$\sum_{i=1}^{n} b_{ik} (\rho_{ij})^k = (\rho_{oj})^{k.} \qquad j = 1,2,3,...,n. \tag{3.6}$$

when $k = 0$, Eq. [6] represents the unbiased condition, that is, the sum of the weights equals unity. The disjunctive kriging covariance can be calculated using the following equation:

$$\sigma_{DK}^2 = \sum_{k=1}^{K} k! C_k^2 \left[1 - \sum_{i=1}^{n} b_{ik} (\rho_{oi})^k \right] \tag{3.7}$$

One advantage of the DK method is that an estimate of the CP that the value at an estimation site is greater than an arbitrary critical value, y_c, can be calculated. This CP is a useful means for determining the risk of various management alternatives. The CP is obtained by defining an indicator variable that is equal to unity if $Y(x_i) \geq y_c$ and is zero otherwise (see Yates et al., 1986a, b). This allows the CP to be written in terms of the conditional expectation and gives the estimator of the CP as:

$$P_{DK}^*(x_o) = 1 - G(y_c) + g(y_c) \sum_{k=1}^{K} H_{k-1}(y_c) H_k^*[Y(x_o)] / k! \tag{3.8}$$

where $G(y_c)$ and $g(y_c)$ are the cumulative and probability density functions, respectively, for a standard normal variable, and $H_k^*[Y(x_o)]$ is found using Eq. [5]. The estimated CP density function, $pdf_{DK}^*(x_o)$, is found by taking the derivative of Eq. [8] with respect to y_c and is:

$$pdf_{DK}^*(x_o) = g(u)\left\{1 + \sum_{k=1}^{K} H_k(u) H_K^*[Y(x_o)]/k!\right\} \tag{3.9}$$

3.3.4 Applying DK technique on soil salinity datasets:

Data transformation:

Data transformations should be performed before using DK. Transformations are used to make the data normally distributed where pairs of sample values are bivariate normal. There are several transformations methods and the appropriate method should be chosen. For all transformations, the predictions are automatically back-transformed to the original values before a map is produced. There are many forms of transformations such as: square-root which is a special case of the Box-Cox and is usually used when data is counts; Log transformation which is used for data with a skewed distribution; arcsine which is used with data that is proportions or percentages; and the normal score transformation which is used with simple, disjunctive, and cokriging. DK ranks the dataset, from lowest to highest values, and matches these ranks to equivalent ranks from a normal distribution. The normal score transformation was used for the data of this study since it is the best for the DK technique.

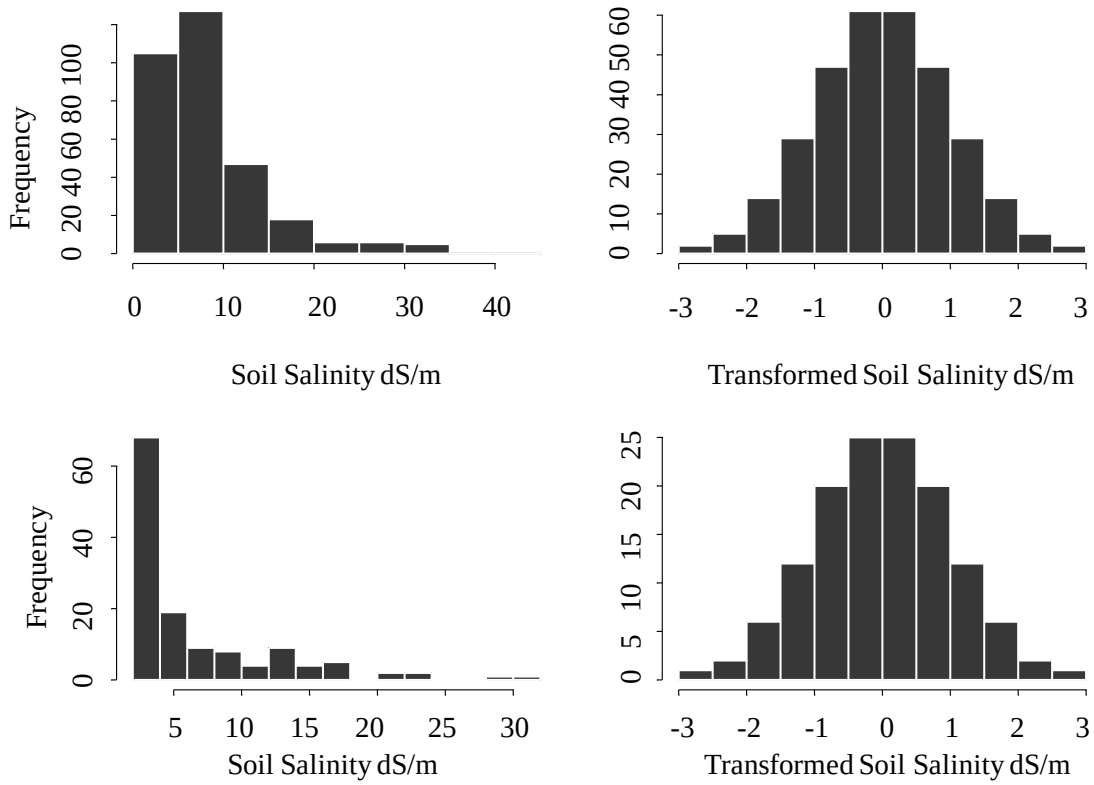


Figure 3.2: Histograms of the collected and transformed soil salinity data for the two datasets.

Figure 3.2 shows the histogram plots of the observed and transformed soil salinity data for the two datasets. Histograms can provide information about the mode and its frequency, an indication of the overall variation, and the shape of the distribution. Both datasets are transformed using normal score transformations. When the first dataset was transformed, the mode value was set to zero with a frequency of 60, the overall variation is between -3 and 3 dS/m, and the distribution is normal. When the second dataset was transformed, the mode value was set to zero with a frequency of 25, the overall variation is between -3 and 3 dS/m, and the distribution is normal.

Generating the CP maps:

An advantage of DK is the ability to generate CP maps that a value at an estimation site is greater than an arbitrary critical value. CP maps are generated by specifying a threshold as a

condition of probability that the values exceed (or do not exceed) the specified threshold. The level or quantity of the property that is being studied must be known in order to use DK effectively. This value is called the cutoff or critical level and values of the property that are larger than this level represent the event being considered. Also, the probability level that spurs a management action must be known. This is the critical probability level where the levels of the property being investigated will no longer be tolerated (Yates and Yates, 1988). Maas and Hoffman (1977) concluded that crops will generally be unaffected by salinity up to some threshold at which time yield will begin to decrease linearly as soil salinity levels increase. This correlation between soil salinity and crop productivity was used in this study to produce CP maps for YP% under different conditions of soil salinity thresholds. Each crop has different thresholds that can determine its YP% levels according to its tolerance to soil salinity. For example, sorghum can produce 100, 100-90, 90-75, 75-50, and 50-0 YP% when soil salinity values do not exceed 4, 5.1, 7.2, 11, and 18 dS/m respectively. Another example, corn can produce 100-90, 90-75, 75-50, and 50-0 YP% when soil salinity values do not exceed 2.5, 3.8, 5.9, and 10 dS/m respectively. These soil salinity threshold values were used as conditions to produce CP maps for different YP% of the selected crops. For each condition, a CP map was generated from 0% to 100% probability with 20% intervals. Therefore, in order to generate a CP map of sorghum that reaches 90-75 YP%, a condition must be set such that the soil salinity values should not exceed 7.2 dS/m while such a condition for corn should not exceed 3.8 dS/m to generate similar CP maps. For soil salinity sensitive crops such as strawberries, a higher CP cannot be produced since the condition requires that soil salinity must be very low. For example, to produce a CP map for strawberries that has a 100% YP, a condition must be set that soil salinity values should not exceed 1 dS/m.

Assessing the crop productivity from the CP maps:

The spatial analyst in ArcGIS was used to reclassify the resulting CP raster maps into six classes for each crop scenario with each dataset. These six classes of the CP maps represent data at 20% intervals from 0% to 100%. This was implemented in ArcGIS using the manual classification and setting the category values as: 0, 0.2, 0.4, 0.6, 0.8, and 1. Contour maps were used for visual illustration and they were generated using the ArcGIS surface analysis option of the spatial analyst. The "tabulated area" option in the ArcGIS toolbox was used to calculate the total area of each class, i.e., to quantify the CP maps. When a condition was set, i.e., soil salinity values do not exceed 4 dS/m, a condition for sorghum to have 100 YP%, the resulting CP map has contour lines of probability from 0% to 100% with 20% intervals that represent the 100 YP% of sorghum. The area contained within the 100% contour lines represents the area of the field that has 100% probability to produce 100% YP. There is one scenario for each threshold of each crop that produces a specific CP map. For example, there are five scenarios for sorghum while only four for corn based on the soil salinity tolerance of each crop (no 100 YP% class for corn). The following is an example of how the areas of different contour lines are calculated for the CP map that has a condition that the soil salinity values do not exceed 7.2 dS/m, which is the condition for sorghum to have 90-75 YP%. The area contained within the 100% contour line represents the area of the field that has 100% probability to produce 90-75 YP%. The area contained within the 100% and 80% contour lines represents the area of the field that has the 80-100% probability to produce 90-75 YP%. The area contained within the 80% and 60% contour lines represents the area of the field that has the 60-80% probability to produce 90-75 YP%, and so on. After calculating the areas of the different classes, each class area was divided by the total area of the field to obtain the percentage of that class from the total area of the field. To obtain the cumulative probability

for each scenario, the percentage of each class was multiplied with its probability and all of them were summed.

3.3.5 Model Evaluation:

Cross-validation was used to evaluate the DK geostatistical model for the different scenarios of the selected crops at different thresholds. Cross-validation removes each data location one at a time and predicts the associated data value and compares the measured and predicted values for all points. The statistics used in cross-validation serve as diagnostics to indicate whether the performance of the model is acceptable. The following statistical measures were set to guarantee that the prediction is unbiased, as close as possible to the measured value, and the variability of the prediction is correctly assessed:

The mean prediction error was used to check if the model is unbiased (centered on the measured values), these values should be near zero to guarantee that the model is unbiased.

The mean prediction error depends on the scale of the data; therefore the mean standardized prediction error was also used to check if the model is unbiased. These values should be close to zero to guarantee the model is unbiased.

The root-mean-squared prediction error was used to check whether the prediction is close to the measured values, the smaller the root-mean-squared prediction error the closer the prediction is to the measured value.

The variability was assessed in two different ways:

Comparing the average standard error with the root-mean-squared prediction error. If the values are similar, then the variability in the prediction is correctly assessed. If the average standard error is greater than the root-mean-squared prediction error, then the variability of the predictions is overestimated. If the average standard error is less than the root-mean-squared prediction error, then the variability of the predictions is underestimated.

Evaluating the root-mean-squared standardized error value. If it is close to one, then the variability of the prediction is correctly assessed. If it is greater than one, then the variability of the prediction is underestimated. If it is less than one, then the variability of the prediction is overestimated.

3.4 Results

In this section, the use of the DK technique as a tool for the management of soil salinity and yield to achieve maximum productivity under existing soil salinity conditions is discussed. First, three examples of CP maps of YP% at different soil salinity thresholds are presented which represent: a sensitive crop (strawberries), a moderate sensitive crop (corn), and a moderate tolerant crop (sorghum). These examples are discussed and evaluated below to visualize the variation in the probability of YP% within a field. Second, the areas contained within the CP contours for all the selected crops were tabulated to evaluate the quantity of variation in the probability of YP% at different zones. Third, the cumulative probability of the whole field for each scenario was calculated to compare among the probabilities to reach different YP% for all the selected crops. Finally, some recommendations and guidelines for growers are presented based on the outcomes of this study to help them select specific crop(s) or to use agro-chemicals more efficiently in the different zones in their fields.

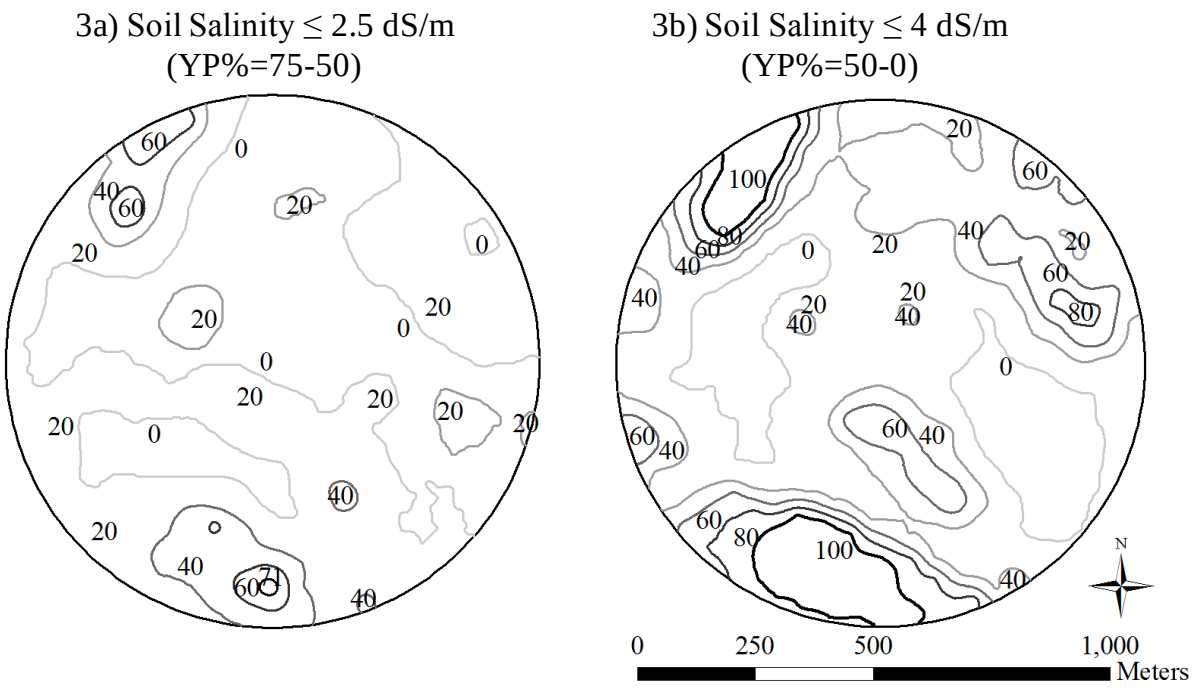


Figure 3.3: CP maps of YP% at different soil salinity thresholds of strawberries using the first dataset.

Figure 3.3 and Figure 3.4 show two examples of CP maps of YP% at different soil salinity thresholds using the first dataset to represent: a sensitive crop (strawberries) and a moderate sensitive crop (corn). The purpose of these two examples is to visualize the variation in the probability of YP% for different zones in fields when planting different crops with different soil salinity tolerances. Contour maps were used to display the CP where each line is labeled with its YP% value. The area contained within two contour lines represents the area of the field that has the range of probability of these two contours to reach a specific YP%. The first example (Figure 3.3) shows the scenario of planting a sensitive crop (strawberries) with the highest probability of productivity of less than 75%. The following thresholds of soil salinity ≤ 2.5, ≤ 4 dS/m were used as conditions to produce probability maps of 75-50, 50-0 of YP% of strawberries respectively. Figure 3.3a shows that the probabilities that strawberries can reach 75-50 YP% are very limited with the condition of soil salinity ≤ 2.5 dS/m, the contour lines with low probabilities cover the majority of the

field. The maximum probability that strawberries can reach 75-50 YP% is 71% which is represented by a very small area at the bottom of the field. Figure 3b shows that the probabilities that strawberries can reach 50-0 YP% is higher (than those of Figure 3.3a) with the condition of soil salinity ≤ 4 dS/m, the contour lines with high probabilities cover significant areas of the field.

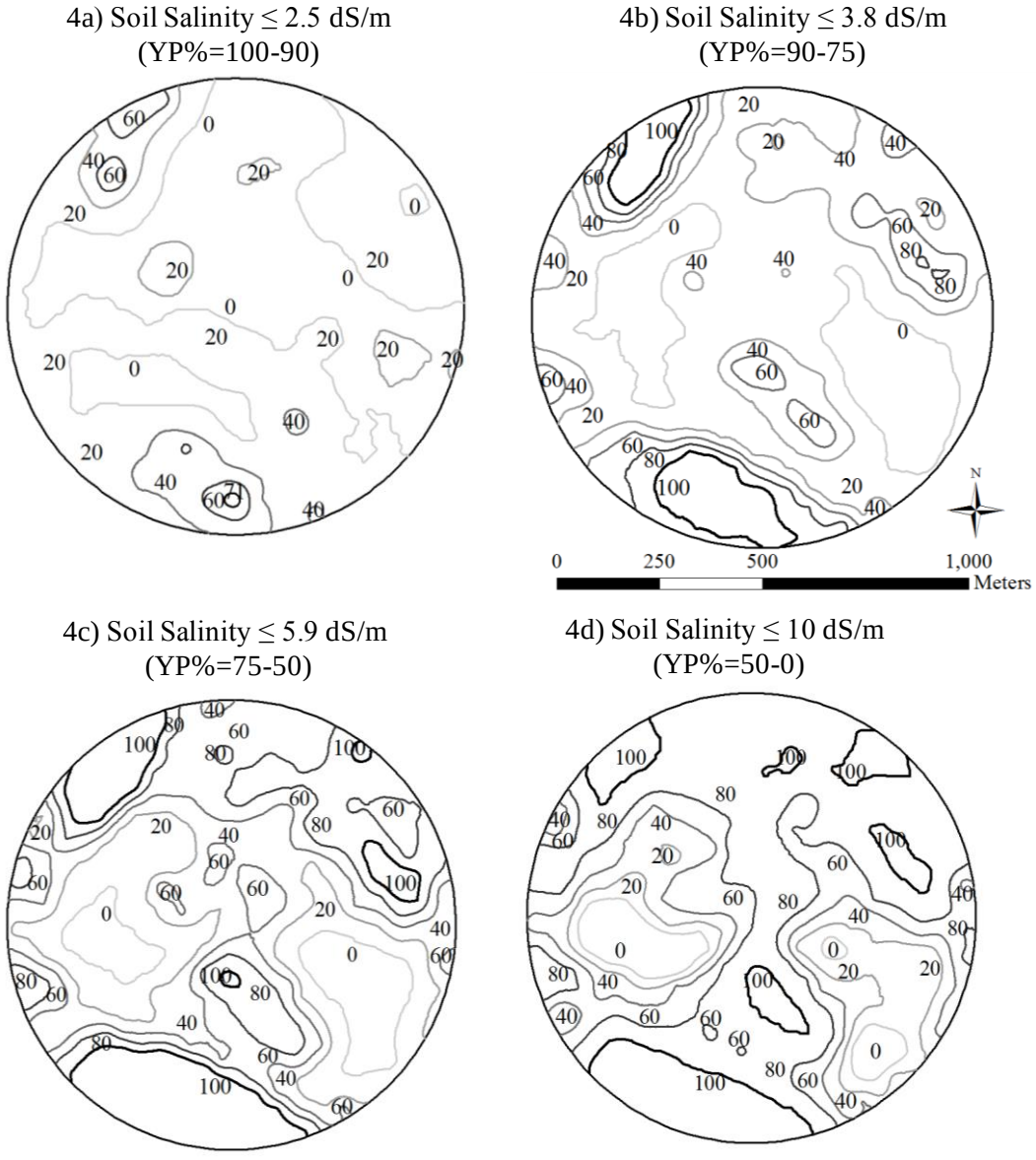


Figure 3.4: CP maps of YP% at different soil salinity thresholds of corn using the first dataset.

The second example (Figure 3.4) shows the scenario of planting a moderate sensitive crop (corn) with the highest probability of productivity of less than 100%. The following

thresholds of soil salinity ≤ 2.5, ≤ 3.8, ≤ 5.9, ≤ 10 dS/m were used as conditions to produce probability maps of 100-90, 90-75, 75-50, 50-0 of YP% of corn respectively. Figure 3.4a is similar to Figure 3.3a which means that corn can reach 100-90 YP% under the same condition as strawberries can reach 75-50 YP.

Table 3.2: Areas of different zones with different Conditional Probabilities (CP) for all scenarios of the selected crops under different soil salinity thresholds for the first dataset.

	CP						CP					
YP (%)	100%	80%	60%	40%	20%	0%	100%	80%	60%	40%	20%	0%
	Barley						Sorghum					
100	38.7	22.5	14.5	11.8	9.6	2.8	7.8	5.3	10.0	22.6	38.2	16.2
100-90	54.5	18.0	12.7	9.2	3.8	1.8	10.8	8.1	17.6	26.0	24.6	12.9
90-75	70.0	16.2	7.8	3.0	2.7	0.2	25.5	24.0	18.7	13.8	13.5	4.5
75-50	87.8	7.1	3.0	2.0	0.1	0.0	62.2	16.5	11.4	5.7	3.2	1.0
50-0	94.5	4.4	1.1	0.0	0.0	0.0	87.8	7.1	3.0	2.0	0.1	0.0
	Corn						Pomegranate					
100							0.1	2.6	5.8	7.7	51.5	32.3
100-90	0.0	0.1	2.1	7.5	47.0	43.3	7.1	5.1	8.8	21.4	40.9	16.8
90-75	7.1	5.1	8.8	21.4	40.9	16.8	12.0	9.7	19.2	25.3	22.1	11.8
75-50	14.4	15.8	22.2	20.4	17.7	9.5	42.4	21.6	14.1	11.5	7.7	2.6
50-0	54.5	18.0	12.7	9.2	3.8	1.8	74.3	15.1	5.7	2.8	2.2	0.0
	Apples						Strawberries					
100												
100-90												
90-75	4.2	5.5	6.1	16.7	48.5	19.1						
75-50	10.0	6.7	15.5	26.2	27.7	13.8	0.0	0.1	2.1	7.5	47.0	43.3
50-0	38.7	22.5	14.5	11.8	9.6	2.8	7.8	5.3	10.0	22.6	38.2	16.2
	Beets						Tomatoes					
100	7.8	5.3	10.0	22.6	38.2	16.2	0.0	0.1	2.1	7.5	47.0	43.3
100-90	10.8	8.1	17.6	26.0	24.6	12.9	5.2	5.2	6.8	18.2	46.3	18.3
90-75	21.2	23.1	20.4	14.9	14.7	5.7	10.5	7.5	16.9	26.1	25.8	13.2

75-50	51.9	18.9	12.7	10.1	4.4	2.0	30.8	23.9	16.9	12.7	12.3	3.3
50-0	78.0	13.7	4.0	2.8	1.4	0.0	68.5	16.4	8.7	3.2	2.9	0.3
	Lettuce						Barley (Hay)					
100							15.0	17.2	22.3	19.4	16.9	9.2
100-90							29.4	23.8	17.5	13.0	12.7	3.6
90-75	3.4	5.6	5.7	15.2	49.9	20.1	51.4	19.1	12.8	10.2	4.6	2.1
75-50	11.3	8.6	18.2	25.9	23.5	12.4	70.0	16.2	7.8	3.0	2.7	0.2
50-0	49.3	19.8	13.0	10.6	5.2	2.2	90.4	5.4	3.2	1.0	0.0	0.0
	Crested Wheat Grass						Alfalfa					
100	5.2	5.2	6.8	18.2	46.3	18.3						
100-90	15.0	17.2	22.3	19.4	16.9	9.2	4.8	5.3	6.4	17.6	47.2	18.6
90-75	54.5	18.0	12.7	9.2	3.8	1.8	12.0	9.7	19.2	25.3	22.1	11.8
75-50	81.0	11.9	3.4	2.8	0.9	0.0	45.2	20.9	13.8	11.0	6.6	2.4
50-0	95.0	4.1	0.8	0.0	0.0	0.0	80.9	11.8	3.4	2.8	0.9	0.0

Table 3.3: Areas of different zones with different Conditional Probabilities (CP) for all scenarios of the selected crops under different soil salinity thresholds for the second dataset.

	CP						CP					
YP (%)	100%	80%	60%	40%	20%	0%	100%	80%	60%	40%	20%	0%
	Barley						Sorghum					
100	62.0	11.2	7.8	7.5	6.4	5.2	42.7	10.6	5.7	8.7	18.9	13.5
100-90	71.5	8.5	7.2	3.9	6.1	2.8	51.8	6.8	7.0	10.0	14.4	10.0
90-75	82.7	5.8	3.4	4.1	4.0	0.0	60.2	10.3	8.4	7.2	8.0	5.9
75-50	91.6	4.6	2.7	1.1	0.0	0.0	75.1	8.9	5.0	3.6	6.3	1.0
50-0	97.3	2.6	0.1	0.0	0.0	0.0	91.6	4.6	2.7	1.1	0.0	0.0
	Corn						Pomegranate					
100												
100-90							40.9	11.0	6.2	8.0	19.3	14.5
90-75	40.9	11.0	6.2	8.0	19.3	14.5	53.9	5.6	8.0	9.9	13.3	9.3
75-50	57.3	5.3	10.2	9.0	10.7	7.5	65.3	10.6	7.1	6.6	6.0	4.4
50-0	71.5	8.5	7.2	3.9	6.1	2.8	84.6	4.5	3.5	4.6	2.7	0.0
	Apples						Strawberries					
100												
100-90												
90-75	20.5	15.1	10.1	7.5	21.9	25.0						
75-50	49.7	8.2	5.8	10.3	15.4	10.6						
50-0	62.0	11.2	7.8	7.5	6.4	5.2	42.7	10.6	5.7	8.7	18.9	13.5
	Sugar Beets						Tomatoes					
100	42.7	10.6	5.7	8.7	18.9	13.5						

100-90	51.8	6.8	7.0	10.0	14.4	10.0	33.2	12.1	7.8	6.6	20.2	20.1
90-75	59.8	9.7	8.7	7.2	8.4	6.2	50.6	7.7	6.4	10.2	14.9	10.3
75-50	69.5	8.9	7.2	4.9	6.0	3.5	60.7	10.7	8.2	7.3	7.5	5.6
50-0	85.9	3.8	3.6	4.8	2.0	0.0	77.1	8.8	3.8	3.7	6.5	0.1
	Lettuce						Barley (Hay)					
100							57.7	6.0	10.0	8.8	10.3	7.3
100-90							60.7	10.7	8.2	7.3	7.5	5.6
90-75							68.4	9.2	6.9	5.6	6.0	3.8
75-50	51.8	6.8	7.0	10.0	14.4	10.0	82.7	5.8	3.4	4.1	4.0	0.0
50-0	68.4	9.2	6.9	5.6	6.0	3.8	91.6	4.6	2.7	1.1	0.0	0.0
	Crested Wheat Grass						Alfalfa					
100	33.2	12.1	7.8	6.6	20.2	20.1						
100-90	57.7	6.0	10.0	8.8	10.3	7.3	26.1	13.6	9.9	6.1	20.8	23.6
90-75	70.5	8.7	7.3	4.3	6.0	3.1	51.8	6.8	7.0	10.0	14.4	10.0
75-50	87.7	3.9	3.6	4.1	0.7	0.0	66.9	9.9	6.9	6.1	6.0	4.1
50-0	99.6	0.4	0.0	0.0	0.0	0.0	87.6	3.9	3.6	4.1	0.7	0.0

Table 3.2 and Table 3.3 show the areas with different CP that reach different YP% for all crop scenarios for both datasets. The purpose of presenting these tables is to provide a quantitative means to present the variation in the probability of YP%. Both tables show how the areas contained within the 100% CP contour lines increase while the areas contained within the 0% CP contour lines decrease with the decrease of YP%. This sequence of increase or decrease is not followed for the areas contained within contour lines between 100% and 0% of CP (80%, 60%, 40%, 20%). As an example, for the scenario of planting sorghum in Table 3.3 for the first dataset, the areas contained within the contour lines of 100% CP increase (8, 11, 25, 62, 88) as the YP% decrease (100, 100–90, 90–75, 75-50, 50–0) respectively. For the same scenario, the areas contained within the contour lines of 0% CP decrease (16, 13, 4, 1, 0). However, there is a tendency for the areas close to the 0% CP to decrease and the areas close to the 100% CP to increase. There is a transition zone among the areas that have a tendency to decrease and those that have a tendency to increase. This increase or decrease of the areas contained within contour lines with magnitudes between 100% and 0% CP depends on the values of the collected soil salinity data and their locations as well as the soil salinity threshold of each scenario.

Table 3.4: The cumulative CP% of the whole field for the two datasets at different levels of soil salinity thresholds (different YP%) of all the scenarios of the selected crops.

YP (%)	CP% for the whole field		CP% for the whole field	
	1st dataset	2nd dataset	1st dataset	2nd dataset
	Barley		Sorghum	
100	72.1	79.9	34.7	61.8
100-90	80.9	85.4	43.2	68.4
90-75	89.4	91.8	64.1	78.0
75-50	96.1	97.4	85.1	88.0
50-0	98.7	99.4	96.1	97.4
	Corn		Pomegranate	
100	0	0	19.0	0
100-90	13.7	0	33.1	60.5
90-75	33.1	60.5	45.8	69.8
75-50	52.0	73.4	74.3	81.9
50-0	80.9	85.4	91.3	92.7
	Apples		Strawberries	
100	0	0	0	0
100-90	0	0	0	0
90-75	28.6	46.0	0	0
75-50	40.7	66.9	13.7	0
50-0	72.1	79.9	34.7	61.8
	Beets		Tomatoes	
100	34.7	61.8	13.7	0
100-90	43.2	68.4	30.0	54.2
90-75	60.8	77.3	42.2	67.6
75-50	79.5	84.1	67.6	78.6
50-0	92.8	93.3	88.7	89.2
	Lettuce		Barley (Hay)	
100	0	0	53.3	74.0
100-90	0	0	66.7	78.6
90-75	27.4	0	79.3	83.4
75-50	44.2	68.4	89.4	91.8
50-0	78.1	83.4	97.0	97.4
	Crested Wheat Grass		Alfalfa	
100	30.0	54.2	0	0
100-90	53.3	74.0	29.4	49.5
90-75	80.9	84.8	45.8	68.4
75-50	93.9	94.8	75.9	82.7
50-0	98.8	99.9	93.7	94.7

Table 3.4 shows the cumulative probability of YP% for the whole field which include all zones of variable productivity. The cumulative probability increases with the increase of the soil salinity threshold values i.e., the decrease of YP%. For sorghum, the first dataset shows

that under the following soil salinity thresholds ≤4, ≤5.1, ≤7.2, ≤11, ≤18 dS/m, the cumulative probability for the whole field can reach 34.7, 43.2, 64.1, 85.1, 96.1 to achieve 100, 100-90, 90-75, 75-50, 50-0% YP respectively. However, for the second dataset the cumulative probability for the whole field can reach 61.8, 68.4, 78, 88, 97.4 to achieve the same YP%'s. For corn, the first dataset shows that under the following soil salinity thresholds values ≤2.5, ≤3.8, ≤5.9, ≤10 dS/m, the cumulative probability for the whole field can reach 13.7, 33.1, 52, 80.9 to achieve 100-90, 90-75, 75-50, 50-0% YP respectively. However, in the second dataset the cumulative probability for the whole field can reach 0, 60.5, 73.4, 85.4 to achieve the same YP%'s (there is no 100-90 YP% in the second dataset).

3.5 Model Evaluation:

Table 3.5 and Table 3.6 show the cross-validation parameters used to evaluate the DK geostatistical model. The prediction errors of the mean, root-mean-square (RMS), average standard error (ASE), mean standardized (MS), and root-mean-square standardized (RMSS) were used as cross-validation parameters. These parameters were obtained for each scenario of the selected crops at multiple soil salinity thresholds. The mean and mean standardized prediction errors were used to evaluate whether the model is unbiased or not.

Table 3.5: Cross validation parameters for the selected crops at different salinity thresholds for the first dataset.

	Mean	RMS	ASE	MS	RMSS	Mean	RMS	ASE	MS	RMSS
	Barley					Sorghum				
100	0.01	0.36	0.38	0.03	0.94	0	0.33	0.35	-0.01	0.94
100-90	0.01	0.34	0.36	0.02	0.95	0	0.37	0.38	0	0.98
90-75	0.01	0.30	0.30	0.03	0.98	0.01	0.38	0.39	0.02	0.97
75-50	0.01	0.21	0.22	0.03	0.96	0.01	0.32	0.33	0.02	0.96
50-0	0	0.15	0.16	0.03	0.98	0.01	0.21	0.22	0.03	0.96
	Corn					Pomegranate				

100						0	0.23	0.24	-0.01	0.91
100-90	0	0.19	0.19	0.02	1.01	0	0.32	0.34	0.01	0.92
90-75	0	0.32	0.34	0.01	0.92	0	0.38	0.38	0	0.98
75-50	0	0.39	0.39	0.01	1.00	0.01	0.35	0.38	0.02	0.93
50-0	0.01	0.33	0.35	0.02	0.95	0.01	0.28	0.29	0.03	0.99
	Apple					Strawberry				
100										
100-90										
90-75	0	0.29	0.32	0.01	0.91					
75-50	0	0.35	0.37	0	0.95	0	0.19	0.19	0.02	1.01
50-0	0.01	0.36	0.38	0.03	0.94	0	0.33	0.35	0.01	0.94
	Beet					Tomato				
100	0	0.33	0.35	-0.01	0.94	0	0.19	0.19	0.02	1.01
100-90	0	0.37	0.38	0	0.98	0	0.29	0.33	-0.01	0.87
90-75	0	0.39	0.39	0.02	0.99	0	0.36	0.38	0	0.96
75-50	0.01	0.34	0.36	0.02	0.97	0.01	0.37	0.39	0.02	0.96
50-0	0.01	0.25	0.27	0.03	0.94	0.01	0.30	0.31	0.02	0.97
	Lettuce					Barley (Hay)				
100						0.01	0.39	0.39	0.01	1.00
100-90						0.01	0.38	0.39	0.02	0.96
90-75	0	0.28	0.31	0.01	0.91	0.01	0.35	0.36	0.02	0.96
75-50	0	0.37	0.38	0	0.98	0.01	0.30	0.30	0.03	0.98
50-0	0.01	0.35	0.37	0.02	0.95	0	0.19	0.21	0.03	0.93
	Crested Wheat Grass					Alfalfa				
100	0	0.29	0.33	-0.01	0.89					
100-90	0.01	0.39	0.39	0.01	1.00	0	0.30	0.32	0.01	0.91
90-75	0.01	0.33	0.36	0.02	0.95	0	0.38	0.38	0	0.98
75-50	0.01	0.24	0.26	0.03	0.92	0.01	0.35	0.37	0.02	0.94
50-0	0	0.14	0.15	0.03	0.97	0.01	0.24	0.26	0.03	0.92

Table 3.6: Cross validation parameters for the selected crops at different salinity thresholds for the second dataset.

	Mean	RMS	ASE	MS	RMSS	Mean	RMS	ASE	MS	RMSS
	Barley					Sorghum				
100	0.01	0.22	0.27	0.02	0.81	0.00	0.20	0.30	0.00	0.68
100-90	0.01	0.20	0.25	0.02	0.80	0.00	0.21	0.29	0.01	0.72
90-75	0.01	0.19	0.21	0.04	0.89	0.01	0.23	0.27	0.02	0.84
75-50	0.00	0.21	0.14	0.02	1.47	0.01	0.19	0.24	0.03	0.79
50-0	0.00	0.13	0.09	0.01	1.50	0.00	0.21	0.14	0.02	1.47
	Corn					Pomegranate				
100										
100-90						0.01	0.20	0.30	0.02	0.68
90-75	0.01	0.20	0.30	0.02	0.68	0.00	0.21	0.29	0.02	0.72
75-50	0.01	0.24	0.28	0.02	0.86	0.01	0.19	0.26	0.02	0.73
50-0	0.01	0.20	0.25	0.02	0.80	0.00	0.21	0.20	0.02	1.01

	Apple					Strawberry				
100										
100-90										
90-75	0.00	0.30	0.28	0.00	1.05					
75-50	0.00	0.19	0.29	0.01	0.65					
50-0	0.01	0.22	0.27	0.02	0.81	0.00	0.20	0.30	0.00	0.68
	Beet					Tomato				
100	0.00	0.20	0.30	0.00	0.68					
100-90	0.00	0.21	0.29	0.01	0.72	0.00	0.26	0.29	-0.02	0.86
90-75	0.01	0.24	0.27	0.02	0.86	0.00	0.20	0.29	0.01	0.69
75-50	0.01	0.19	0.25	0.02	0.77	0.01	0.23	0.27	0.02	0.83
50-0	0.00	0.20	0.20	0.01	1.03	0.00	0.20	0.23	0.03	0.88
	Lettuce					Barley (Hay)				
100						0.01	0.24	0.28	0.02	0.85
100-90						0.01	0.23	0.27	0.02	0.83
90-75	-0.01	0.39	0.27	-0.03	1.47	0.01	0.18	0.26	0.02	0.71
75-50	0.00	0.21	0.29	0.01	0.72	0.01	0.19	0.21	0.04	0.89
50-0	0.01	0.18	0.26	0.02	0.71	0.00	0.21	0.14	0.02	1.47
	Crested Wheat Grass					Alfalfa				
100	0.00	0.26	0.29	-0.02	0.86	-0.01	0.30	0.29	-0.02	1.03
100-90	0.01	0.24	0.28	0.02	0.85	0.00	0.21	0.29	0.01	0.72
90-75	0.01	0.20	0.25	0.02	0.81	0.01	0.18	0.26	0.02	0.69
75-50	0.00	0.22	0.18	0.01	1.21	0.00	0.22	0.18	0.01	1.21
50-0	0.00	0.09	0.06	-0.01	1.51	0.00	0.22	0.18	0.01	1.21

Table 3.4 and Table 3.5 show that the mean and mean standardized prediction errors are almost zero for all the scenarios in both datasets. This means that the DK model is unbiased i.e., the prediction values are centered on the measured values for all scenarios. The root-mean-square prediction errors were used to check how close the predicted values were to the measured values, the smaller the error the closer the predicted values to the measured ones. The values of the root-mean-square prediction errors shown in both tables are small and close to zero which means that the DK model was successful in making the predicted values as close as possible to the observed values. However, the values in Table 3.6 for the second dataset are slightly less than the corresponding values of the first dataset, which means that the DK model was more successful when using the second dataset rather than the first dataset. Two ways were used to assess whether the variability in the predictions is correct, overestimated, or underestimated. First, the closer the values of the average standard errors

are to the values of the root-mean-squared prediction errors the better the assessment of the variability in the predictions; which is clear for all the scenarios. This means that the variability was correctly assessed in the predictions. Second, the values of the root-mean-square-standardized prediction errors should be 1 for the correct assessment of the variability; which is clear for most of the scenarios. Only in a few scenarios the root-mean-square-standardized prediction errors exceed 1 such as: barley at YP% < 75%, sorghum at YP% < 50%, barley (hay) at YP% < 50%, crested wheat grass at YP% < 75%, and alfalfa at YP% is < 75%. This means that the DK model underestimates the variability for these few scenarios.

3.6 Advantages and disadvantages of DK

DK which is a non-linear kriging technique has several advantages over linear estimation methods. First, it provides a more accurate estimate of the property of interest and can generate an estimate of the CP for that property (Yates and Yates, 1988). Second, the CP maps produced by DK can be used as an input to a management decision making model to provide a quantitative means for determining whether management actions are necessary (Yates and Yates, 1988). Third, the DK technique provides important implications in aiding management decisions by providing growers with a quantitative input that can be used for evaluating the variability of the crop productivity in different zones in fields. Fourth, DK performed better than IK because the continuous Hermite transform in DK retains all the information in the original data while the transform in other techniques, such as IK, use discrete transformations which inevitable looses information. The only disadvantage in using DK is the increased computational time (Yates et al. 1986a).

3.7 Discussion

It has become imperative to explore the potential of increasing the food production from saline lands due to the increasing pressure from growing populations. Thus combating land salinization problems is vital for food security through adoption of salinity and crop management strategies. Plants vary widely in their salinity tolerance. One method for addressing the soil salinity problem is to select and plant salt-tolerant crops in saline soil areas. This paper, introduces a technique on how to live with salinity in its current condition without leaching the soil salinity or doing other soil reclamation efforts. The critical or threshold value of the soil salinity is the value beyond which the crop productivity is negatively affected. In this study the threshold values were used as input values (conditional probability information) for the DK technique to generate CP maps for YP% under different conditions of soil salinity thresholds. The CP maps can be used as a quantitative method for making management decisions.

Non-linear kriging techniques can be used as a quantitative method for making management decisions for soil salinity and yield if the CP information is available. Non-linear kriging techniques include indicator kriging (IK) which involves a nonlinear transformation of the data to a discrete variable and DK which involves a nonlinear transformation of the data to a continuous variable. To the best of the authors knowledge, of the previous studies that used non-linear kriging techniques and targeted soil salinity and crop yield management, Eldeiry and Garcia (2011) is the closest to the current study. They applied the IK technique using indicator variograms to evaluate different scenarios of crops and salinity levels to generate maps that show the expected percent yield potential YP%. Their results show that IK can be used to generate guidance maps that divide fields into areas of expected percent yield potential based on soil salinity thresholds for different crops. In this paper, the DK technique was used to provide unbiased estimates of the conditional

probability (CP) that the true value of the property of interest doesn't exceed a defined threshold. The results of this study show that the CP maps generated using the DK technique provide an accurate characterization and quantification of the different areas of the fields. CP maps were used to assess the expected YP% of fields for several crops under multiple soil salinity thresholds. The methodologies used in both techniques (IK and DK) are different, however, both of them can be used as management decision tools to manage the productivity under current soil salinity conditions. Both techniques provide knowledge of the YP% of different areas. Based on the knowledge of YP% at different areas of a field, a decision can be taken to manage the productivity of these areas by selecting another crop or adjusting the inputs such as fertilizer, seeding rates and herbicides.

3.8 Conclusions

Decisions based on critical thresholds which use estimates of variability are subject to error. DK enables these errors to be converted to an estimated probability that the true value exceeds a given threshold, thereby giving decision makers a means to judge the risk associated with a particular estimate. DK provides minimum variance estimates of properties from nonlinear combinations of spatially correlated sample data. In addition it can be used to estimate the conditional probability that some critical threshold is exceeded. DK, a nonlinear kriging model, was used in this study to provide an unbiased estimator of the conditional probability that a given variable exceeds a threshold. DK assists in making management decisions by providing a quantitative input (CP maps), which can be used for evaluating the variability of different areas in fields. The data presented in the results show how the DK technique can generate valuable information for the growers for them to make decisions regarding which crop(s) to select or if they need to make more efficient use of agro-chemicals. Efficient use of agro-chemicals is beneficial for farmers as well as for the

environment. In this study, visual information of the variation in the probability of productivity in the different areas of fields, for different crop scenarios, were presented and discussed. This information enables growers to visualize the variability of the productivity in different areas of their fields. Tabulated information was also presented and discussed in order to provide growers with quantitative information about the probability of the productivity of different areas in their fields. This information enables growers to quantify the variability of the productivity in different areas of their fields. The DK technique presented in this study provides a tool to achieve spatial optimization of farm management which will increase productivity or reduce the amount of agro-chemicals applied.

3.9 REFERENCES

Ayers, R. S., and Westcot, D. W. (1976). Water Quality for Agriculture. *Irrigation and Drainage Paper 29. FAO, Rome.*

Boydell, B., McBratney, A.B. (1999). Identifying potential within-field management zones from cotton yield estimates. In: Stafford, J.V. (Ed.), Precision Agriculture '99. *Proceedings of the Second European Conference on Precision Agriculture. Denmark, 11–15 July. SCI, London, UK,* pp. 331–341.

Cain, D. (1997). U.S. Geological Survey Data Collection Center and Studies in the Arkansas River Basin. *Colorado Water: Newsletter of the Water Center at Colorado State University, Fort Collins, CO.*

Eldeiry, A., and Garcia, L. A. (2011). "Using Indicator Kriging Technique for Soil Salinity and Yield Management." *J. Irrig. Drain. Engin.,* **137(2)**, 82-93.

Eldeiry, A., Garcia, L. A., and Reich, R. M. (2008). "Soil Salinity Sampling Strategy Using Spatial Modeling Techniques, Remote Sensing, and Field Data." *J. Irrig. Drain. Engin.,* **134(6)**, 768-777.

Ferguson, R.B., Hergert, G.W., Penas, E.J. (2000). Corn. In: Ferguson, R.B. (Ed.), Nutrient Management for Agronomic Crops in Nebraska. *University of Nebraska, Lincoln, NE,* pp. 75– 83. EC 01-155.

Fleming, K. L., Westfall, D. G., Wiens, D. W., Rothe, L. E., Cipra, J. E., and Heerman, D. F. (1999). Evaluating farmer developed management zone maps for precision farming. *In: Proceedings of the Fourth International Conference on Precision Agriculture, 1998 (Robert P C; Rust R H; Larson W E, eds),* pp 335–344. *ASA, CSSA, SSSA, Madison, WI, USA.*

Fraisse, C., Sudduth, K., Kitchen, N., and Fridgen, J. (1999). Use of unsupervised clustering algorithms for delineating within-field management zones. *In ASAE, International Meeting, Toronto, Canada.* Paper 993043.

Frasier, W.M., Waskom, R.M., Hoag, D.L., Bauder, T.A. (1999). Irrigation Management in Colorado: Survey Data and Findings. *Technical Report TR99-5 Agricultural Experiment Station, Colorado State University. Fort Collins, CO.*

Fridgen, J., Fraisse, C., Kitchen, N., and Sudduth, K. (2000). Delineation and analysis of site-specific management zones. In Petoskey, B. J., editor, *Proceedings of the 2nd International Conference on Geospatial Information in Agriculture and Forestry,* pages 402–411, *Lake Buena Vista, Florida, USA. ERIM International, Inc.*

Gates, T. K., Burkhalter, J. P., Labadie, J. W., Valliant, J. C., and Broner, I. (2002). "Moitoring and modeling flow and salt transport in a salinity-threatened irrigated valley," *Journal of Water Resources Planning and Management,* **128(2)**, 87-99.

Hochstrasser, U.W. (1965). Handbook of mathematical functions. M. Abramowitz and I.A. Stegun (ed.) Dover Publ., New York. Journel, A.G., and Ch.J. Hujibregts. 1978. Mining geostatistics. *Academic Press, New York.*

Journel, A. G. and Huijbregts, Ch. J. (1978). Mining Characteristics, *Academic Press, New York,* 600 p.

Maas, E.V., Hoffman, G.J. (1977). "Crop salt tolerance—current assessment." *J. Irrig. Drain. Engin.*, **103 (IR2)**, 115–134.

Matheron, G. (1976). A simple substitute for conditional expectation: The disjunctive Kriging, in Advanced Geostatistics in the Mining Industry, M. Guarascio, M. David, and C. Huijbregts, eds., pp. 221-236, *D. Reidel Publishing, Dordrecht, Holland.*

Miles, D.L. (1977). Salinity in the Arkansas Valley of Colorado. *Region VIII, Environmental Protection Agency, Denver.* 80 p.

Russo, D. (1984b). "A geostatistical approach to the trickle irrigation design in a heterogenous soil: 2: A field test." *Water Resources Res.* **20(5)**, 543-552.

Russo, D. (1984a). "Design of an optimal sampling network for estimating the variogram." *Soil Sci. Soc. Am. J.*, **48**, 708-716.

Schepers, J.S., Schlemmer, M.R., Ferguson, R.B. (2000). "Site-specific considerations for managing phosphorus." *J. Environ. Qual.* **29**, 125-130.

Stafford, J.V., Lark, R.M., Bolam, H.C. (1998). Using yield maps to regionalize fields into potential management units. In: Robert, P.C. (Ed.), *Proceedings of the Fourth International Conference on Precision Agriculture. ASA, CSSA, SSSA, Madison,* **WI**, 225–237.

Storey, R., and Jones, R.G. (1978). Salt stress and comparative physiology in the *gramineae*. I. Ion relations of two salt-and water-stressed barley cultivars, California Mariout and Arimar. Aust. J. Plant Physiol., 5: 801-816.

Tanji, K. K., ed. (1990). Agricultural Salinity Assessment and Management. ASCE Manuals and Reports on Engineering Practice No. 71, ASCE, New York, N.Y.

von Steiger, B., Webster, R., Schulin, R., Lehmann, R. (1996). "Mapping heavy metals in polluted soil by disjunctive kriging." *Environ. Pollut.* **94**, 205-215.

Wittler, J. M., Cardon, G. E., Gates, T. K., Cooper, C. A., and Sutherland, P. L. (2006). "Calibration of Electromagnetic Induction for Regional Assessment of Soil Water Salinity in an Irrigated Valley." *J. Irrig. Drain. Eng.*, **132(5)**, 436-444.

Wood, G., Oliver, M.A., and Webster, R. (1990). "Estimating soil salinity by disjunctive kriging." *Soil Use Manage.* **6**, 97-104.

Yates, S. R. and Yates, M. V. (1988). “Disjunctive kriging as an approach to management decision making.” *Soil Sci. Soc. America J., 52, 1554-I 558.*

Yates, S.R., Warrick, A.W., Matthias, A.D., and Musil, S. (1988). “Spatial variability of remotely sensed surface temperatures at field scale.” *Soil Sci. Soc. Am, J.* **52**, 40-45.

Yates, S.R., Warwick, A.W., and Myers, D.E. (1986a). “Disjunctive kriging. 1. Overview of estimation and conditional probability.” *Water Resour. Res.* **22**, 615-622.

Yates, S.R., Warwick, A.W., and Myers, D.E. (1986b). “Disjunctive kriging. 2. Examples.” *Water Resour. Res.* **22**, 623-630.

Zirschky, J. (1985). “Geostatistics for environmental monitoring and survey design.” *Environ. Int.* **11**, 515-524.

Zirschky, J. H., Keary, G. P., Gilbert, R. O., and Middlebrooks, E. J. (1985). "Spatial Estimation of Hazardous Waste Site Data," *Journal of Environmental Engineering, ASCE* **Ill**, 777-789.

Zirschky, J., and Harris, D.J. (1986). “Geostatistical analysis of hazardous waste site data.” *J. Environ. Eng. (NY)* **112**, 770-784.

4 COMPARISON OF NONLINEAR GEOSTATISTICAL MODELS IN ESTIMATING THE IMPACT OF SOIL SALINITY ON THE SPATIAL VARIABILITY OF CROP YIELD

4.1 Summary

In order to manage soil salinity and crop yield, this research proposes that individual fields be divided into zones based on the conditional probability (CP) of each zone reaching a specific yield potential (YP). Three nonlinear geostatistical models – disjunctive kriging (DK), indicator kriging (IK), and probability kriging (PK) – were used to develop the CP maps based on soil salinity thresholds for two crops, alfalfa and corn. These CP maps were compared with actual yield data taken while conducting a soil salinity survey for two fields cultivated with alfalfa and corn. The CP maps divide each field of interest into zones with different probabilities to reach a specific YP for a given crop at a specific soil salinity threshold. The objectives of this study are as follows: (1) compare the performance of the DK, IK, and PK models in developing CP maps; (2) compare actual alfalfa and corn yield samples with the estimated YP by the three models; and (3) provide guidance by considering the output of the models used in this study as input for precision management of agriculture. The three nonlinear geostatistical models were applied to soil salinity datasets collected in two fields (alfalfa and corn) in the Lower Arkansas River Valley in Colorado. Yield data were collected at the same fields to compare the actual data with that estimated by the models. The results of this study show that the CP maps developed using the three geostatistical models are efficient in assessing the impact of soil salinity on the spatial variability of alfalfa and corn yield. The comparison of the actual yield data with the estimated CP maps from the three models shows good agreement where most of the yield samples were located at the appropriate zones estimated with the three geostatistical models.

The IK and PK models generated very similar estimates for each of the zones. However, the zones generated by both of these models are slightly different to the zones generated using the DK model. The information provided by the models about the variability and hotspots can be used for the precision management of agricultural resources.

4.2 Introduction

It is imperative that crop production increase to meet the increasing demand for food due to the growth of the world population. Soil salinity is a major limiting factor for crop yield in poorly drained soils (Rogers 2002; Patel et al., 2002). Based on crop biomass samples collected on corn and alfalfa fields in a study conducted from 2004-2008 in the Lower Arkansas River Valley of Colorado, there appears to be a clear trend of decreasing crop yield as soil salinity increases above threshold values of 3 to 5 dS/m (Gates et al., 2012). The worldwide salt-affected soils, including saline and sodic soils, were estimated to be 831 million hectares (Martinez-Beltran and Manzur, 2005), extending over all continents. Attempts to improve the salt tolerance of crops through conventional breeding programs have met with very limited success due to the complexity of the salt tolerance both genetically and physiologically (Flowers, 2004). Mapping and assessing soil salinity is the first step towards developing management strategies to optimize crop yield under current soil salinity conditions.

Linear kriging methods such as simple, ordinary, and universal kriging are well established for predicting soil variables at unsampled locations. Examples of using linear kriging in soil and water science are well documented (Burgess and Webster, 1980; Webster and Burgess, 1980; Triantafilis et al., 2001; Eldeiry and Garcia 2008a, 2008b, 2010). Assessing conditional probability (CP) of a specific variable is as important as predicting

value of the variable at unsampled locations, both of which can be achieved by using nonlinear kriging. Nonlinear kriging techniques depend on the nonlinear transformation of data, whether discrete or continuous.

The disjunctive kriging (DK) technique has found widespread use in soil science (Wood et al., 1990; von Steiger et al., 1996). The CP maps generated using the DK technique can be used as input to a management decision-making model in order to provide a quantitative means of determining whether management actions are necessary (Yates et al., 1988). Wood et al. (1990) used the DK technique to estimate soil salinity and mentioned that soil salinity thresholds are important in determining land suitability for different crops in Israel. Triantafilis et al. (2004) used indicator kriging (IK), multiple-indicator kriging and DK to assess the current status and potential threat of soil salinity. Eldeiry and Garcia (2012) used the DK technique to manage soil salinity and crop yield. The IK technique has been widely applied (Halvorson et al., 1995; Van Meirvenne and Goovaerts, 2001; Eldeiry and Garcia 2011, 2012). The IK technique is flexible and can be modified to fit specific management or research goals by modifying the critical threshold criteria (Smith et al., 1993). IK makes no assumptions on the underlying invariant distribution (Cressie, 1992). Solow (1986), used IK to estimate the conditional probability that a sample point belongs to one type of soil or another. The probability kriging (PK) technique is based on the cokriging estimator and is a shortcut of IK, which uses the order relation of observed values to recover information from all available attribute values (Carr, 1994; Carr and Mao, 1993). Juang and Lee (2000) mentioned that PK is based on the cokriging estimator and may be troubled by some intrinsic disadvantages of cokriging. The uniform value (the standardized rank), which denotes the order relation of observed values, is assigned as the only auxiliary variable in PK to improve the estimation of the probability of the attribute value being lower than the desired threshold (Juang and Lee, 2000).

Nonlinear kriging techniques have advantages over linear kriging techniques due to their ability to take data uncertainty into account, and therefore are often used to predict the CP for categorical data at unsampled locations (Goovaerts, 1994; Oyedele et al., 1996). Few studies have been published that use nonlinear geostatistical techniques to manage soil salinity and crop yield. The authors of this study previously investigated the CP maps generated using the DK nonlinear techniques for different crops. The results of their research motivated them to broaden their investigation to cover all nonlinear geostatistical techniques (DK, IK, and PK) in order to explore the differences between them. Therefore, the objectives of this study are: comparing the performance of the DK, IK, and PK models in developing CP maps; comparing actual alfalfa and corn yield samples with the estimated YP by the three models; and providing guidance by considering the output of the models used in this study as input for precision management of agriculture. The CP maps that were generated using the three nonlinear models divided the study fields into zones, with each zone having a probability that it would reach a specific YP according to the soil salinity thresholds for alfalfa and corn. Dividing fields into zones with different probabilities to reach a specific YP provides valuable information that can be used in the management of soil salinity for crop yield. Based on this information, farmers or managers can decide which crops to grow. Both visual and quantitative information presented in this study warns farmers about the potential for low yield in part or all of a specific field.

4.3 Data and Methodology

4.3.1 Study Area and Selected Datasets

The study area is located in the Lower Arkansas River Basin in Colorado, near the cities of Rocky Ford and La Junta (Figure 4.1). Farmers in this area are facing decreasing crop

yields due in part to high levels of salinity in the soil and irrigation water because the Arkansas River is one of the most saline rivers in the United Sates (Myers, 1982). Average water table depths in this region have risen toward the surface approximately 0.3-1.3 m between 1969 and 1994 (Cain, 1997). This has only exacerbated the salinity problems because of the increased upflux of saline groundwater. In a survey of the region, 68% of producers stated that high salinity levels were a significant concern (Frasier, 1999).

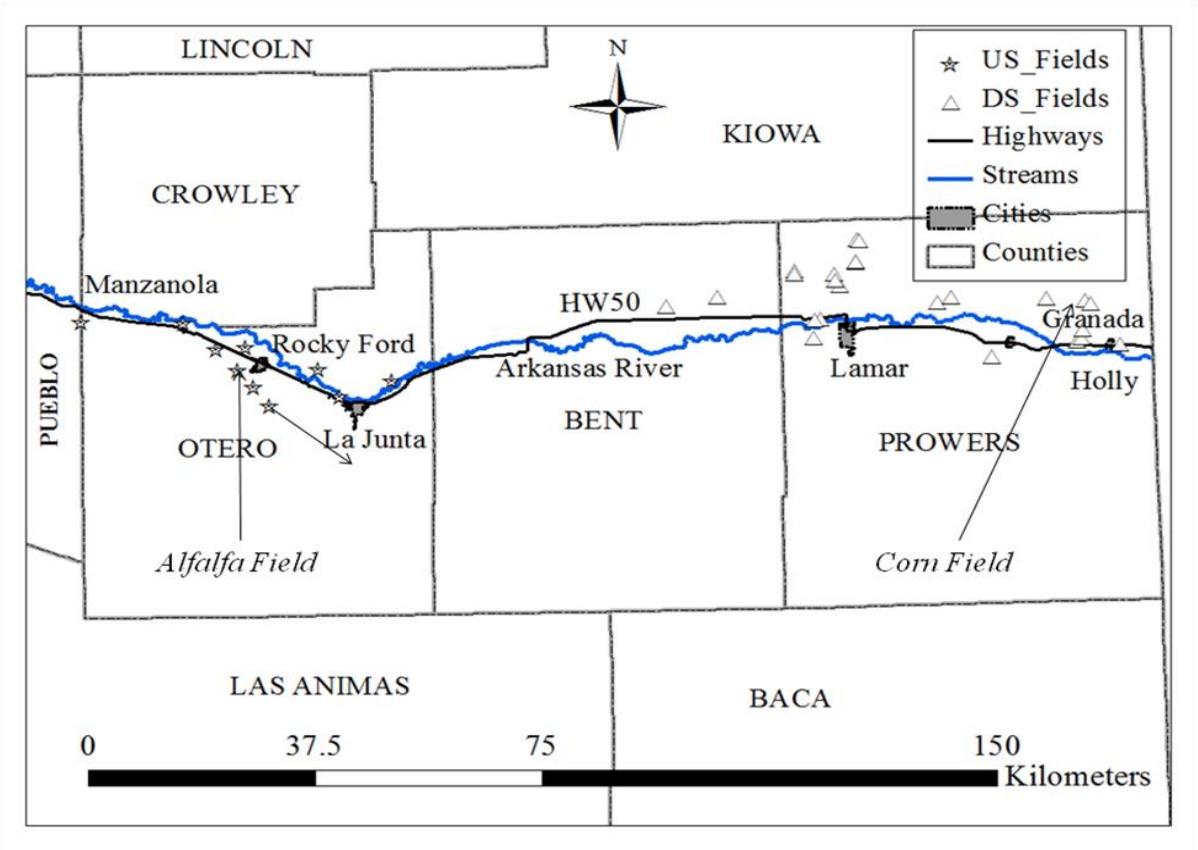


Figure 4.1: The study area in the Arkansas River Basin in Colorado, with the upstream (US) region on the left side and the downstream (DS) region on the right side.

Crop yield reduction due to salinity in fields in the Lower Arkansas River Basin in Colorado has been estimated to be between 0% and 75%, with a total revenue loss ranging from \$0-\$750/ha based on 1999 crop prices (Gates et al., 2012). Soil salinity datasets were

collected in two fields in the study area shown in Figure 4.1. Summary statistics of the soil salinity datasets collected in these fields are provided in Table 4.1.

Table 4.1: Summary statistics of the soil salinity datasets collected in the alfalfa and corn fields.

Dataset	N	Min.	Max.	Mean	Std_dev	Area (m^2)
Alfalfa	102	2.70	10.91	4.91	1.82	122,665
Corn	132	3.64	7.30	4.96	0.94	423,529

4.3.2 Data Collection

The study area soils, where soil salinity data were collected using an electromagnetic induction instrument (EM-38), can be termed as hyper-electrolytic, where the ratio of the EC_a/clay is larger than ~5 (EC_a measured in mS/m and clay in mass percentage) McBratney et al. (2005). Therefore, the EC_a measured in this soil can only be used for determining salinity levels, which was the main use of the EM-38 instrument developed by Rhoades et al. (1989). A calibrated equation was developed by Wittler et al. (2006) for the same area where the data for this study was collected. As part of their study a total of 414 sites were sampled, and EM-38 readings were taken at each site for calibration. At each calibration site, soil samples were collected at the surface, 0.3, 0.6, 0.9, and 1.2 m depths from the left, center, and right location of where the EM-38 readings were taken, for a total of 15 samples per calibration site. While collecting soil from the center hole, soil temperatures were taken at every depth and additional soil was collected at every depth for use in measuring the gravimetric soil-water content.

The major crops in the Lower Arkansas River Valley in Colorado in order of cropped area are alfalfa, corn, grass hay, wheat, sorghum, dry beans, cantaloupe, watermelon, and onions (USDA NASS Colorado Field Office, 2009). Alfalfa (*Hedicago sativa*) and corn (*Zea mays*) were selected as the crops to use in this study because they are the prevailing crops in

the study area. Also, alfalfa represents a moderately tolerant crop to soil salinity, while corn represents a moderately sensitive crop.

Table 4.2: Soil salinity threshold values (dS/m) of different yield potential (YP) for corn and alfalfa.

YP	Alfalfa Thresholds (dS/m)	Corn Thresholds (dS/m)
1.0	2.0	1.7
0.9	3.4	2.5
0.75	5.4	3.8
0.50	8.8	5.9
0.0	15.5	10.0

Table 4.2 shows the expected YP of alfalfa and corn that corresponds to the soil salinity thresholds of both crops. The soil salinity thresholds and the corresponding YP were adapted from Ayers and Westcot (1985). To estimate crop yield, six crop biomass samples were collected in each field. For alfalfa, biomass samples were collected at each sampling location by either of the following two methods: 1) If the crop was not cut, all vegetation was cut on a 1 meter square and placed in a mesh onion sack for greenhouse drying; and 2) If the crop was cut into windrows, a length of windrow was measured, collected and placed in a mesh onion sack for greenhouse drying. In addition, the distance from the centerline of the windrow to the centerline of an adjacent windrow was measured and recorded for the purpose of calculating the biomass/area. For corn, crop biomass samples were collected, by cutting a number of plants in each sampling location and placing them into a mesh onion sack for greenhouse drying. Crop biomass samples were allowed to air dry for a minimum of 3 weeks in a low humidity greenhouse environment. After drying, crop samples were weighed and normalized. Alfalfa biomass data were normalized by dividing by an estimated maximum yield per cutting of 7.4 ton/hectare while corn samples were divided by 81.5 ton/hectare, which was the maximum yield suggested by the Colorado Agricultural Experiment Station (2008).

4.3.3 Using Nonlinear Geostatistical Models to Generate CP Maps

In geostatistics, maps can be generated to estimate either a value or an indicator of the variable of interest (CP maps). Estimating a value of the variable of interest has been investigated thoroughly; however, estimating an indicator (CP maps) has not been given enough attention. This research investigates the CP maps of YP that were generated using three nonlinear geostatistical models: DK, IK, and PK. In order to generate CP maps, a variable of interest may be converted into a binary variable (0 or 1) by choosing a threshold. If values are above the threshold, they become 1, and if they are below the threshold, they become 0. Therefore, the interpolation is between 0 and 1, and the estimates can be interpreted as the probability of a variable being 1 (being in the class that is indicated by 1). If a threshold is used to create the indicator variable, then the resulting interpolated map shows the probabilities of exceeding or falling below the threshold. Crops are generally unaffected by soil salinity up to some threshold beyond which crop yield begins to decrease linearly as the soil salinity levels increase (Maas and Hoffman, 1977). This correlation between soil salinity and crop productivity was utilized to generate CP maps of YP. Soil salinity threshold levels were considered as conditions for a given crop to reach a specific YP. In each of the two selected datasets, the three models were applied in order to generate CP maps. These maps divide each field into zones with different probabilities to reach a specific YP according to the soil salinity thresholds of a crop (e.g., alfalfa and corn).

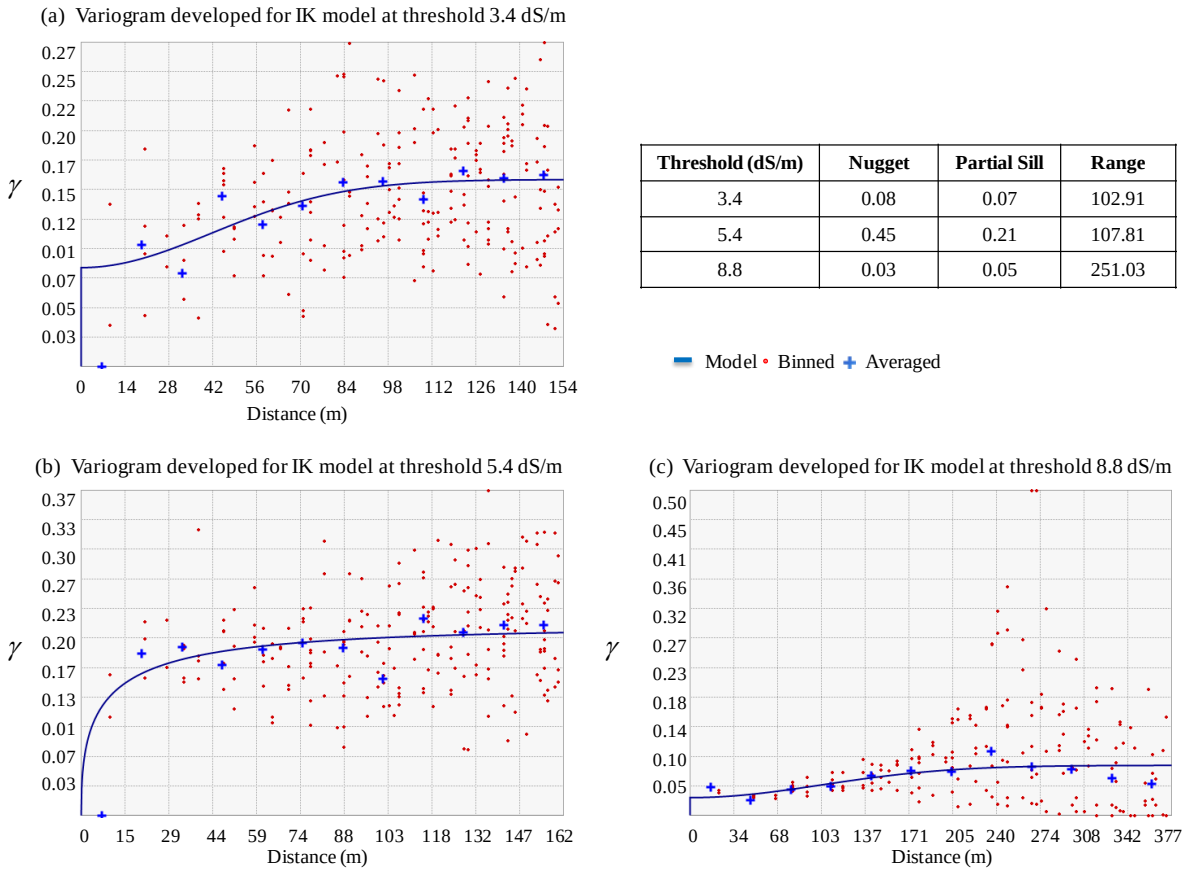
4.3.4 Using Variogram Models

The correlation among neighboring values are modeled as a function of the distance between the samples used in this study, defined as a variogram (Miller et al., 2007). The spatial distribution of the soil salinity data was analyzed using variograms, which have been widely used to analyze the spatial structures in ecology (Phillips, 1985; Robertson, 1987).

There are several terms that are often associated with the variogram. The height that the variogram reaches when it levels off is called the sill. It is often composed of two parts: the nugget, and the partial sill; added together, these give the sill. The distance at which the variogram levels off to the sill is called the range. Sample locations separated by distances closer than the range are spatially autocorrelated, whereas locations farther apart than the range are not. When the variogram is developed, the binned values are generated by grouping the variogram points together using square cells that are one lag wide, the average points are generated by binning the variogram points that fall within angular sectors, while the model shows the fitted curve to the points of the variogram. The binned points show local variation in the variogram values, whereas the average values show the smooth variogram variation. Geostatistical analysis can be used to optimize the variogram models, which are based on minimizing the mean square error. The variogram model parameters such as nugget and partial sill are optimized using cross-validation to estimate the range parameter. A variogram is estimated using the following equation:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [\hat{\varepsilon}(s_i) - \hat{\varepsilon}(s_i + h)]^2 \tag{4.1}$$

where $\gamma(h)$ is the variogram values, $\hat{\varepsilon}(s_i)$ and $\hat{\varepsilon}(s_i + h)$ are the estimated residuals from the multiple regression models at locations s_i and $s_i + h$, a location separated by a distance *h*; *N(h)* is the total number of pairs of samples separated by a distance *h*.



Threshold (dS/m)	Nugget	Partial Sill	Range
3.4	0.08	0.07	102.91
5.4	0.45	0.21	107.81
8.8	0.03	0.05	251.03

Figure 4.2: Examples of variograms developed for the indicator kriging (IK) model at soil salinity thresholds of 3.4, 5.4, and 8.8 dS/m for the alfalfa field.

Figure 4.2 shows three examples of the variograms developed for the IK model at soil salinity thresholds of 3.4, 5.4, and 8.8 dS/m for the soil salinity dataset collect in the alfalfa field. These three cases provided three different subsets of data, which require developing a specific variogram for each case. However, for the same data subset, only one variogram is developed for the three models. The values of the nugget, sill, and range were provided in Figure 4.2. The figure shows that case (b), where the soil salinity threshold was set to 5.4 dS/m, has a better structure that the ones in cases (a) and (c). When the soil salinity threshold was set to $\leq$ 5.4 dS/m, it has a better structure than the ones in cases (a) and (c). The reason for that is when setting a soil salinity threshold value of 5.4 dS/m, it is close to the mean of the collected soil salinity data (4.91 dS/m) or in the middle of the collected samples at the

alfalfa field (table 2). Therefore, there are several points that satisfy this condition and these points take the value of 1 as well as several points that don't satisfy the condition and take the value of zero to generate the CP maps. In this case there are enough points that have values of 1 as well as enough points that have values of zero, which make a good mixture of points that improve the structure of the variogram. However, setting soil salinity thresholds of 3.4 dS/m cases (a), and 8.8 dS/m case (c), these thresholds are close to the minimum and maximum of the collected soil salinity values collected at the alfalfa field (table 2). In both cases (a and b), only a few points have the value of 1 while the majority of the points have the value of zero (case a), or vise versa (case c). This makes the structure of the variogram not as good as in case (b). The table of the upper left of the Figure 2 shows the nugget, sill, and the range values for the three different variograms. The range values of cases (a and b) indicate that the autocorrelation among the data extend for almost the same distance while the autocorrelation extend more for case (c). The higher sill value, the nugget added to the partial sill, in case (b) than the ones in cases (a) and (c) indicate that the variogram in case (b) was able to capture more variations rather than that in cases (a) and (c), which improve the structure of the variogram in case (b). This should have an impact on the performance of the models and will be explained later in this manuscript. The better structure the variogram has, the better the model performs and the better the cross-validation statistics.

4.3.5 DK Model Equations

The DK model assumes an unknown constant mean and the general form can be written as follows:

$$f(Z(s)) = \mu_1 + \varepsilon(s) \tag{4.2}$$

where μ_1 is an unknown constant and $f(Z(s))$ is an arbitrary function of $Z(s)$. The DK model requires that the data have a bivariate normal distribution. In order to satisfy this

assumption, the original soil salinity data, $Z(x)$, must be transformed into a new variable, $Y(x)$, with a standard normal distribution where pairs of sample values are bivariate normal. The function $\phi[Y(x)]$, which describes this transformation, is as follows:

$$\phi[Y(x)] = Z(x) = \sum_{k=0}^{\infty} C_k H_k[Y(x)] \tag{4.3}$$

where the values for $Y(x)$ are obtained by taking the inverse of the data, $Y(x) = \phi^{-1}[Z(x)]$ and $H_k[Y(x)]$ are a Hermite polynomial of order k. The $C_k's$ are the Hermitian coefficients, which are determined using the properties of orthogonality, and are generally determined using numerical integration, as follows:

$$C_k = \frac{1}{k!\sqrt{(2\pi)}} \sum_{i=1}^{j} w_i \phi(v_i) H_k(v_i) \exp[-v_i^2/2] \tag{4.4}$$

where v_i, and w_i, are the abscissa and weight factors for the Hermite integration (Hochstrasser, 1965). The DK estimator is calculated from a sum of unknown functions of the transformed sample values, $Y(x_i)$. It is required that each unknown function, $f_i[Y(x_i)]$, depend on only one transformed value, $Y(x_i)$. The DK estimator is calculated using the following equation:

$$Z^*_{DK}(x_o) = \sum_{i=1}^{n} f_i[Y(x_i)] = \sum_{i=1}^{n} \sum_{k=1}^{\infty} f_{ik} H_k[Y(x_i)] \tag{4.5}$$

where f_i is the unknown function with respect to the transformed variable, and *n* is the number of samples.

An unbiased estimator with the minimum estimation variance can be obtained using the following equations:

$$Z^*_{DK}(x_o) = \sum_{k=0}^{K} C_k H^*_k[Y(x_o)] \tag{4.6}$$

where

$$H_{\mathrm{k}}^{*}[Y(x_{\mathrm{o}})]=\sum_{\mathrm{i}=1}^{n} b_{\mathrm{ik}} H_{\mathrm{k}}[Y(x_{\mathrm{i}})] \tag{4.7}$$

where the series in Eq. (4) has been truncated to K terms, and b_{ik} are the DK weights. The $H_{\mathrm{k}}^{*}[Y(x_{\mathrm{o}})]$ represents the estimated value of the K^{th} Hermite polynomial at the estimation location. The sum of these estimates multiplied by the coefficient, C_{k} [which transforms $Y(x)$ into $Z(x)$] makes up the DK estimate at x_{o}.

Using nonlinear geostatistical models, the conditional probability that the value at an estimation point is greater than a selected critical value (y_{c}) can be calculated. The CP is obtained by defining an indicator variable that is equal to unity if $Y(x_{\mathrm{i}}) \geq y_{\mathrm{c}}$ and is zero otherwise (Yates et al., 1986a, b). This allows the CP to be written in terms of the conditional expectation and gives the estimator of the CP as:

$$P_{\mathrm{DK}}^{*}(x_{\mathrm{o}})=1-G(y_{\mathrm{c}})+g(y_{\mathrm{c}})\sum_{\mathrm{k}=1}^{K} H_{\mathrm{k}-1}(y_{\mathrm{c}}) H_{\mathrm{k}}^{*}[Y(x_{\mathrm{o}})]/k! \tag{4.8}$$

where $G(y_{\mathrm{c}})$ and $g(y_{\mathrm{c}})$ are the cumulative and probability density functions, respectively, for a standard normal variable.

IK Model Equations

The IK model assumes an unknown constant mean and the general form can be written as follows:

$$I(s)=\mu+\varepsilon(s) \tag{4.9}$$

where μ is an unknown constant and $I(s)$ is a binary variable. The indicator function under a desired cutoff value z_{k} can be written as follows:

$$I(x, z_{\mathrm{k}})=\begin{cases} 1, \text{if } z(x) \leq z_{\mathrm{k}} \\ 0, \text{otherwise} \end{cases} \tag{4.10}$$

The IK model estimator $I(x_i; z_k)$ at the location can be calculated using:

$$I^*(x_o; z_k) = \sum_{i=1}^{n} \lambda_i I(x_i; z_k) \tag{4.11}$$

And the IK system given $\sum \lambda_i = 1$ is as follows:

$$\sum_{j=1}^{n} \lambda_j \gamma_I(x_j - x_i) = \gamma_I(x_o - x_i) - \mu \tag{4.12}$$

where λ_j is the weight coefficient, γ_I is the semivariance of the IK codes at the respective lag distance, and μ is the Lagrange multiplier.

PK Model Equations

The PK model assumes two unknown constants and the general form can be written as follows:

$$I(s) = I(Z(s) > c_t) = \mu_1 + \varepsilon_1(s) \tag{4.13}$$

where

$$Z(s) = \mu_2 + \varepsilon_2(s) \tag{4.14}$$

where μ_1 and μ_2 are unknown constants and $I(s)$ is a binary variable created by using a threshold indicator $I(Z(s) > c_t)$. There are two types of random errors, $\varepsilon_1(s)$ and $\varepsilon_2(s)$, which means that there is autocorrelation for each of them and cross-correlation between them. In other words, the PK model uses two variables, the main and auxiliary variables. $I(x; z_k)$ is assigned as the main variable while $U(x)$ is assigned as the auxiliary variable (uniform) in the cokriging estimator. The uniform value, also called the standardized rank, was reported in detail by Deutsch and Journel (1997) and is defined as follows:

$$U(x) = \frac{r}{n} \tag{4.15}$$

where r denotes the rank of the r^{th} order statistic $z_{(k)}$ located at x, and n is the total number of observations (Goovaerts, 1997).

The PK model estimator is defined by:

$$I^*(x_o; z_k) = \sum_{i=1}^{n} \lambda_{Ii} I(x_i; z_k) + \sum_{l=1}^{m} \lambda_{Ul} U(x_l) \tag{4.16}$$

where λ_{Ii} and λ_{Ul} are the weights associated with $I(x_i; z_k)$ and $U(x_l)$.

A more comprehensive explanation about the three models can be found in Matheron (1976), Yates et al. (1986a, b), and Juang and Lee (2000).

4.3.6 Model Evaluation

The models used in this study were evaluated based on two criteria: the accuracy, and the successfulness of the model in estimating the variability. The accuracy was evaluated by comparing the yield data with the estimated values from the models. The value of each yield sample for alfalfa or corn was compared with the corresponding zone estimated by the different models. If the value of the yield sample was located in a zone that complies with the estimates of Ayers and Westcot (1985) then the model was considered to be accurate. The accuracy of the models was also evaluated when selecting the variogram by using the root mean square error (RMSE) cross-validation statistic. The smaller the RMSE the closer the prediction was to the measured values. The model successfulness in assessing the variability was evaluated by using the root mean squared standardized error (RMSSE) cross-validation

statistic. If the RMSSE is close to 1, the variability of the prediction is correctly assessed. The RMSE and RMSSE are calculated using the following equations (Ramos et al., 2008):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[\hat{Z}(x_i) - Z(x_i)]^2} \tag{4.17}$$

$$RMSSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\frac{\hat{Z}(x_i) - Z(x_i)}{\sigma^2(x_i)}\right]^2} \tag{4.18}$$

where $\hat{Z}(x_i)$ is the predicted value at the cross-validation point, $Z(x_i)$ is the measured value at point (x_i), N is the number of measurements of the dataset, and $\sigma^2(x_i)$ is the kriging variance at cross-validation point (x_i).

4.4 Results

Table 4.3 shows the cross-validation parameter errors for the alfalfa and corn datasets when applying the three models at different soil salinity thresholds. The soil salinity dataset for the alfalfa field ranges allowed us to develop CP maps for three thresholds: 3.4, 5.4, and 8.8 dS/m. However, the soil salinity dataset for the corn field only allowed us to develop CP maps for two thresholds: 3.8 and 5.9 dS/m. The smaller the RMSE the closer the prediction is to the measured values (Robinson and Metternich, 2006). The RMSE values for the alfalfa dataset for the three models are smaller at the threshold of 5.4 dS/m than at the thresholds of 3.4 and 8.8 dS/m.

Table 4.3: Cross-validation parameter errors for the two datasets of alfalfa and corn when applying the DK, IK and PK models at different soil salinity thresholds.

Threshold (dS/m)	RMSE	RMSSE	Threshold (dS/m)	RMSE	RMSSE
Alfalfa; DK model			Corn; DK model		
3.4	0.80	2.42	3.8	0.77	2.28

5.4	0.73	1.92	5.9	0.81	2.42
8.8	0.94	4.52	10.0	NA	NA
Alfalfa; IK model			Corn; IK model		
3.4	0.82	2.41	3.8	0.70	1.66
5.4	0.69	1.60	5.9	0.85	2.67
8.8	0.94	5.27	10.0	NA	NA
Alfalfa; PK model			Corn; PK model		
3.4	0.81	2.46	3.8	0.70	1.66
5.4	0.72	1.64	5.9	0.85	2.67
8.8	0.93	4.32	10.0	NA	NA

This indicates that the middle threshold has a better performance than the lower and higher thresholds. However, the RMSE values for the corn dataset at the 3.8 dS/m threshold are better than at the 5.9 dS/m threshold. The closer the RMSSE values are to 1, the better the model performance (Robinson and Metternich, 2006). The RMSSE values for the alfalfa field dataset for the three models are smaller at the 5.4 dS/m threshold than at the 3.4 and 8.8 dS/m thresholds. Also, the RMSSE values for the corn dataset at the 3.8 dS/m threshold are better than the values at the 5.9 dS/m threshold.

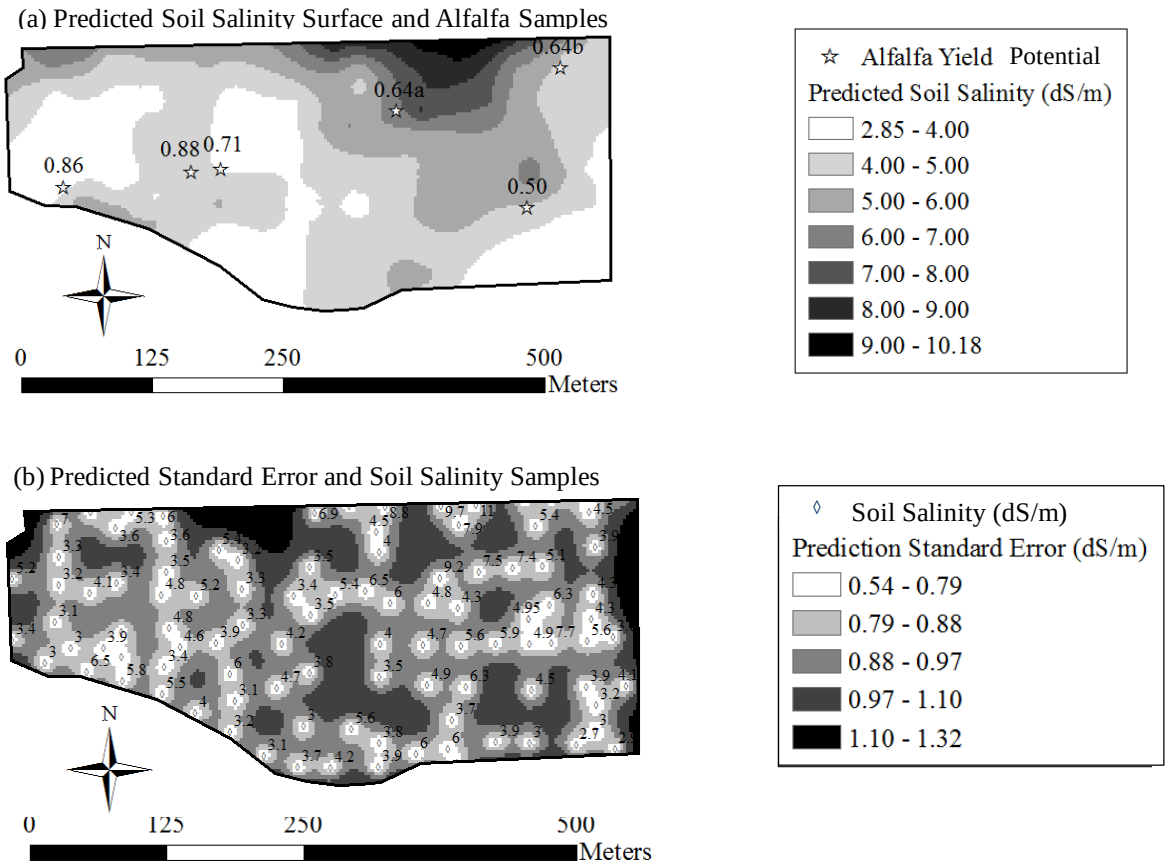


Figure 4.3: (a) Generated surface of predicted soil salinity using ordinary kriging (OK) and the alfalfa samples collected; (b) predicted standard error surface for the alfalfa field and the collected soil salinity samples.

Figure 4.3(a) shows the predicted soil salinity surface using the ordinary kriging (OK) and the collected alfalfa yield samples. At the three levels of soil salinity (3.4, 5.4, and 8.8 dS/m) that this field can reach, the expected alfalfa YP are 0.9, 0.75, and 0.5 respectively (Ayers and Westcot, 1985). Table 4.3 and Figure 4.3(a) show that some of the normalized alfalfa yield samples (0.88, 0.86, 0.64a) are located in soil salinity zones that match with the expected YP of these zones. However, other normalized alfalfa yield samples (0.71, 0.64b, and 050) are slightly less than the expected YP of these zones. There are several reasons that the yield of a particular sample might be less or more than expected. Soil salinity has the

potential to impact crop yield, however irrigation management, fertilizer application, pest control, and seed germination also have the potential to impact crop yield.

Figure 4.3(b) shows the prediction standard error surface of the generated soil salinity surface using the OK model and the collected soil salinity samples for the alfalfa field. The figure shows that the prediction standard error surface ranges from 0.54 to 1.32 dS/m. The prediction standard error values are small in the places where the soil salinity data were collected, while these values are larger in the places where no data were collected. Table 1 shows that the soil salinity dataset for the alfalfa field has a high range, where the minimum and maximum values are 2.70 and 10.91 dS/m respectively, and a high standard deviation of 1.82 dS/m. The high values of the range and the standard deviation can impact the values of the prediction standard error.

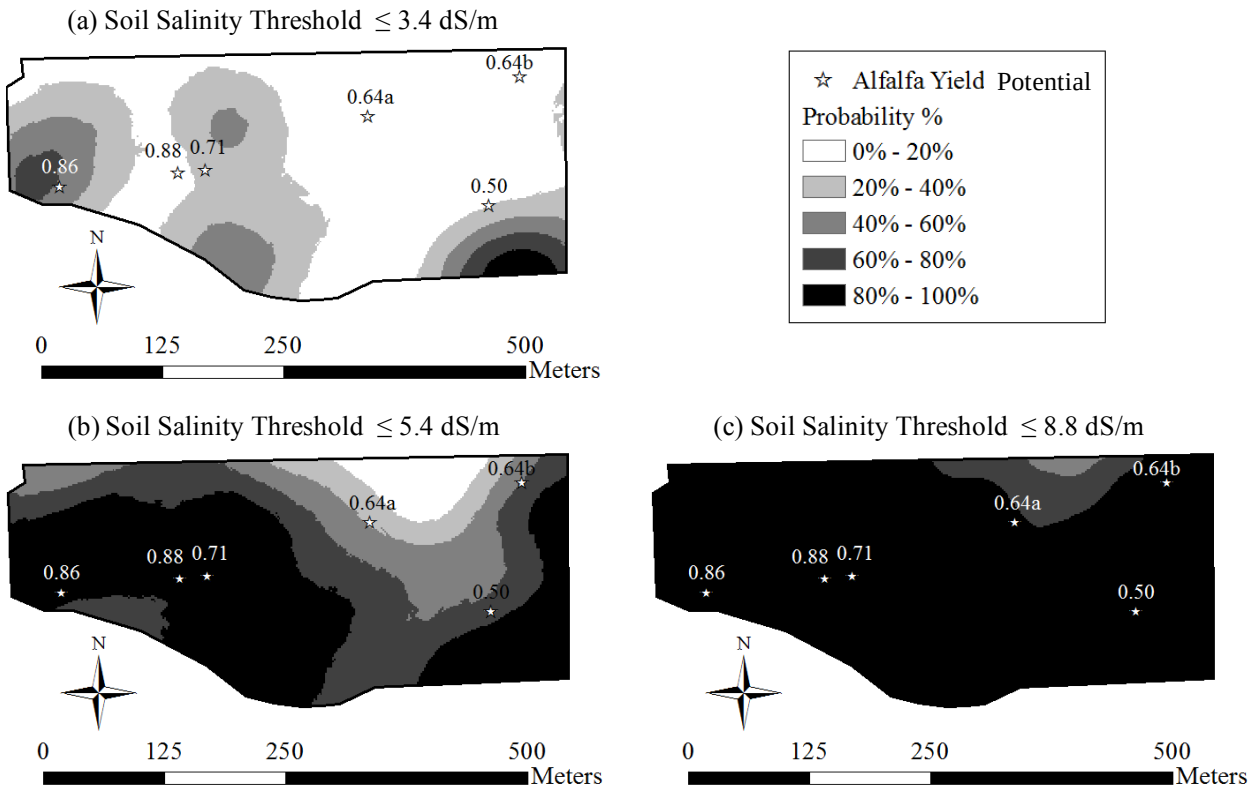


Figure 4.4: CP maps developed using the DK model at the soil salinity thresholds: (a) ≤ 3.4, (b) ≤ 5.4, and (c) ≤ 8.8 (dS/m), which are the conditions for alfalfa to reach: 0.9, 0.75, and 0.50 of YP respectively.

Figure 4.4 shows the CP maps developed using the DK model at the soil salinity thresholds of: (a) ≤ 3.4, (b) ≤ 5.4, and (c) ≤ 8.8 (dS/m), which are the conditions for alfalfa to reach: 0.9, 0.75, and 0.50 of YP respectively. The values of the YP of the alfalfa samples, the predicted soil salinity, and the probability of the expected YP using the DK model for the different soil salinity thresholds are summarized in Table 4.4. The three CP maps displayed in Figure 4.4(a, b, c) show that with a low soil salinity threshold, large areas of the field have low probability that alfalfa can reach a specified YP. However, as the soil salinity thresholds increase, the low probability areas decrease while the high probability areas increase. The reason is that for alfalfa to reach a high YP a low soil salinity threshold is needed and vise versa. Therefore, the probabilities are low for alfalfa to reach a high YP under the soil salinity

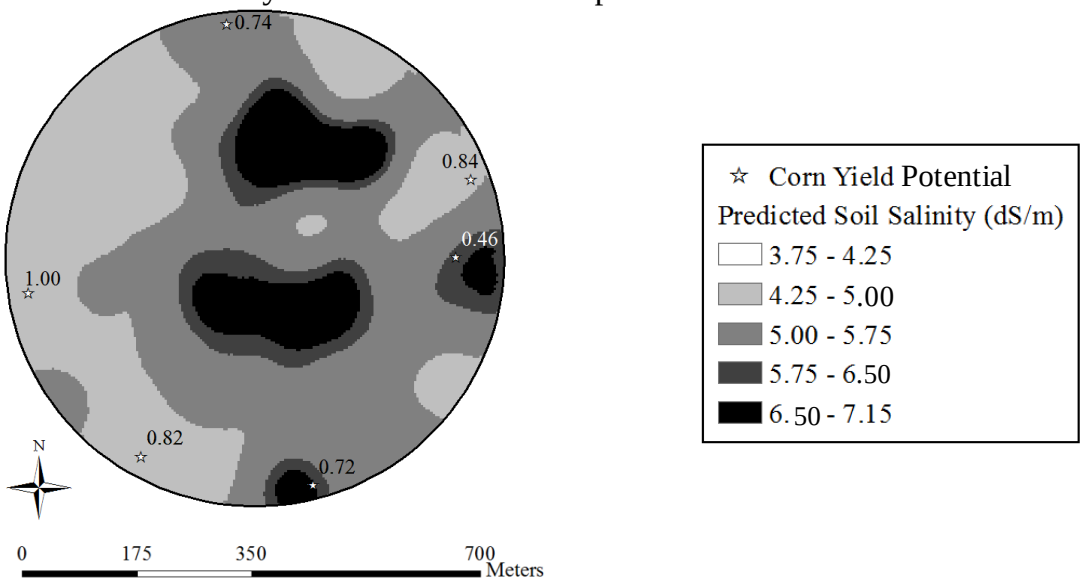
conditions present in that field. Table 4.4 shows that at a soil salinity threshold of ≤ 3.4 (dS/m) with an expected alfalfa YP of 0.90, the normalized alfalfa yield samples are less than expected.

Table 4.4: Summary of the normalized yield of alfalfa samples, the predicted soil salinity, and the probability of the expected YP at the locations of the alfalfa samples.

Normalized Yield of Alfalfa Samples	Predicted Soil Salinity (dS/m)	Probability of the Expected YP at Different Soil Salinity Thresholds		
		≤ 3.4 (dS/m) YP = 0.9	≤ 5.4 (dS/m) YP = 0.75	≤ 8.8 (dS/m) YP = 0.50
0.86	2.85 - 4.00	40% - 60%	80% - 100%	80% - 100%
0.88	4.00 - 5.00	0% - 20%	80% - 100%	80% - 100%
0.71	2.85 - 4.00	20% - 40%	80% - 100%	80% - 100%
0.64 (a)	6.00 - 7.00	0% - 20%	20% - 40%	80% - 100%
0.64 (b)	4.00 - 5.00	0% - 20%	40% - 60%	80% - 100%
0.50	4.00 - 5.00	0% - 20%	60% - 80%	80% - 100%

At a soil salinity threshold of ≤ 5.4 with an expected alfalfa YP of 0.75, the first two samples exceed the YP expectations, the third one is slightly less than expected, and the last three samples are less than the expected. At a soil salinity threshold of ≤ 8.8 with an expected alfalfa YP of 0.50, all samples exceed the YP expectations. A reason that some normalize crop yield samples exceed the YP expectation is that for any specific threshold, the actual soil salinity might be less than that the specified threshold, i.e., for the soil salinity threshold of ≤ 3.4 (dS/m), there is a chance that the soil salinity might be 2.0 dS/m or less, which gives a higher probability that the normalized yield sample will exceed the YP expectation at this location.

(a) Predicted Soil Salinity Surface and Corn Samples



(b) Predicted Standard Error and Soil Salinity Samples

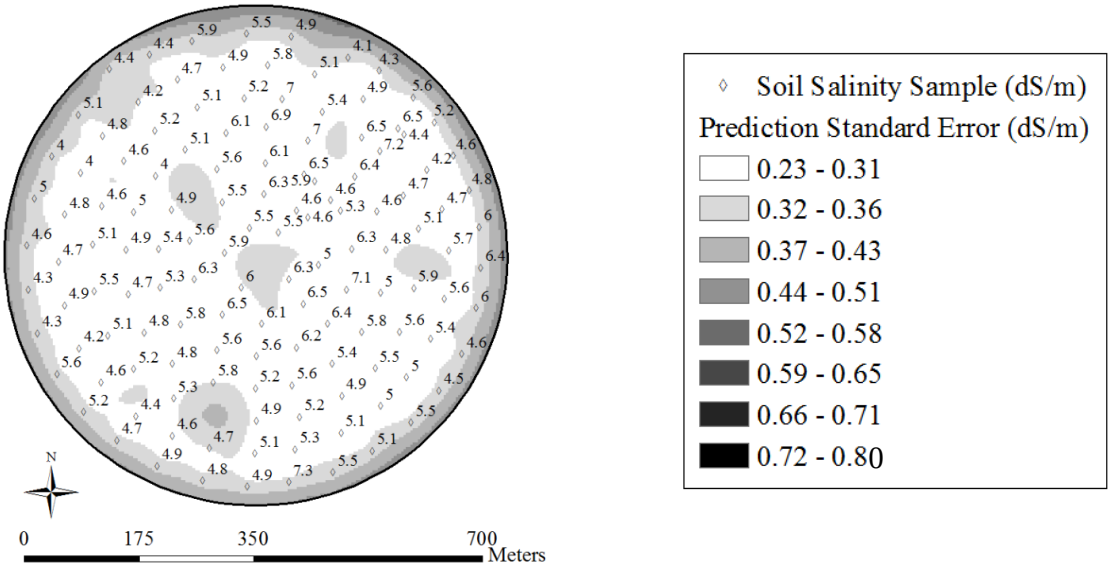


Figure 4.5: (a) Generated surface of predicted soil salinity using the OK model and the alfalfa samples collected; (b) predicted standard error surface for the corn field and the collected soil salinity samples.

Figure 4.5(a) shows the predicted soil salinity surface using the OK and the normalized corn yield samples. At the two levels of soil salinity (3.8 and 5.9 dS/m) that this field can reach, the expected corn YP are 0.75 and 0.5 respectively (Ayers and Westcot, 1985). Table 4.4Table 3.4 and Figure 4.5(a) show that the YP of all the corn samples exceed the expected

values. As discussed earlier with Figure 4.3(a), there are several factors that have the potential to impact YP.

Figure 4.5(b) shows the prediction standard error surface of the generated soil salinity surface using the OK model and the collected soil salinity samples for the corn field. The prediction standard error surface shows a range of values from 0.54-1.32 dS/m. The figure shows that the majority of the field has small values of prediction standard error which ranges from 0.23 to 0.34 dS/m. The outer part of the field have prediction standard errors that range from 0.35 to 0.56 dS/m, and some small areas in the outer part of the field range from 0.57 to 0.80. This can be explained by the fact that there are a lot of data points that were collected inside the field whereas there were no data points collected in the outer part of the field. Therefore the prediction standard error in the outer part of the field is based on extrapolation, which makes the prediction standard error higher. Table 4.1 shows that the soil salinity dataset for the corn field has a fairly low range of soil salinity (3.64 to 7.30 dS/m), while the standard deviation is 0.94 (dS/m). These values are relatively low which have a smaller impact on the prediction standard error compared to the values for the alfalfa field.

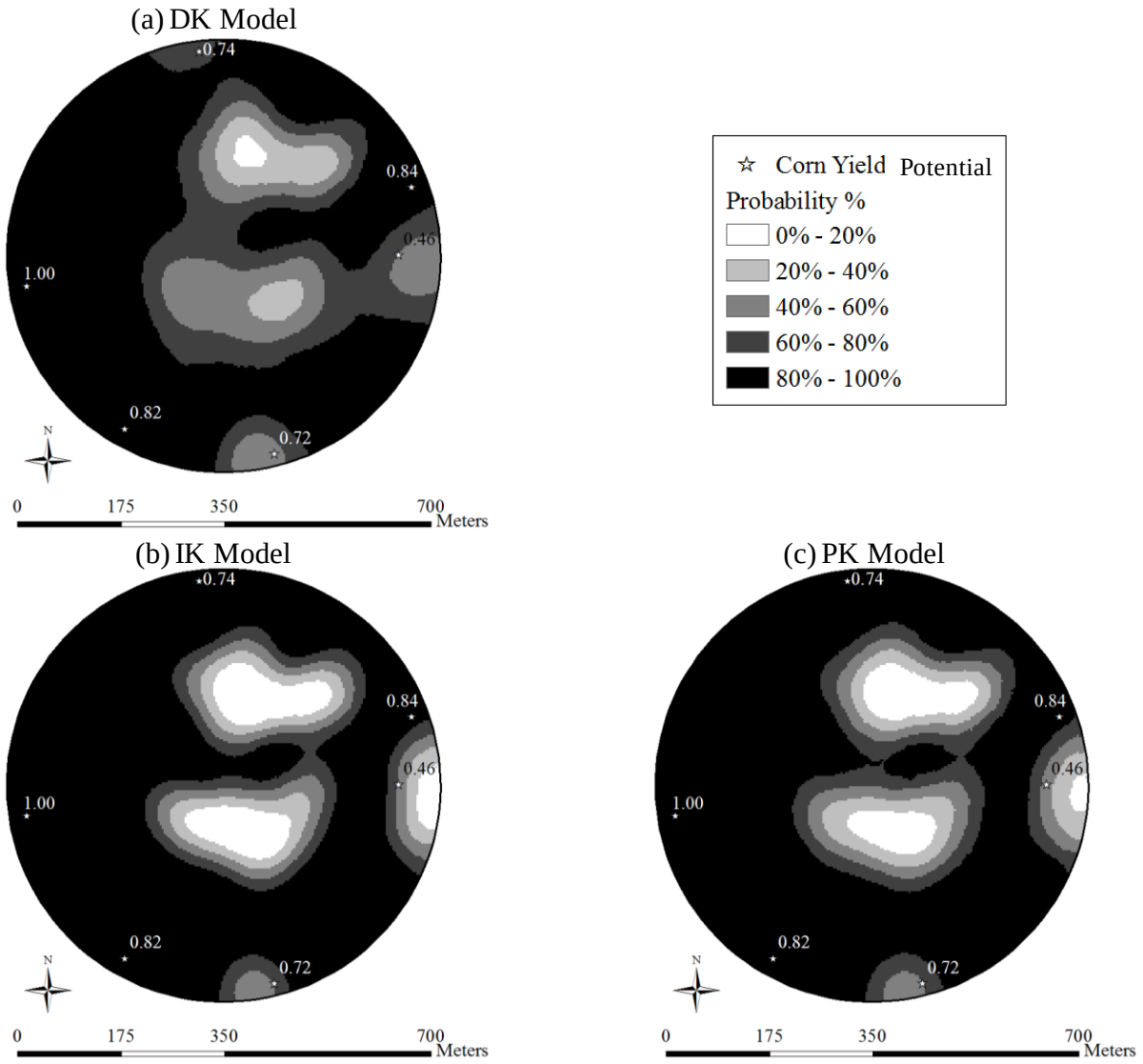


Figure 4.6: CP maps developed using the DK, IK, and PK models for the scenario of planting corn (a moderately salt sensitive crop) in field DS09 for a soil salinity threshold of ≤ 5.9 dS/m for it to reach a productivity of 0.50 of YP.

Figure 4.6 shows the CP maps developed using the DK, IK, and PK models at the soil salinity threshold of ≤ 5.9 dS/m, a condition for corn to reach a productivity of 0.50 of YP. The values of the normalized yield of the corn samples, the predicted soil salinity, and the probability of the expected YP using the three models at soil salinity threshold ≤ 5.9 dS/m were summarized in Table 4.5.

Table 4.5: Summary of the normalized yield of corn samples, the predicted soil salinity, and the probability of the expected YP using the three models (DK, IK, and PK) for a soil salinity threshold (≤ 5.9 dS/m) at the locations of the corn samples.

Normalized Yield of Corn Samples	Predicted Soil Salinity (dS/m)	Probability of the Expected YP Using the Three Models (DK, IK, and PK) for a Soil Salinity Threshold (≤ 5.9 dS/m)		
		DK	IK	PK
1.00	4.25 - 5.00	80% - 100%	80% - 100%	80% - 100%
0.74	5.00 - 5.75	60% - 80%	80% - 100%	80% - 100%
0.82	4.25 - 5.00	80% - 100%	80% - 100%	80% - 100%
0.72	6.50 - 7.15	40% - 60%	40% - 60%	40% - 60%
0.84	4.25 - 5.00	80% - 100%	80% - 100%	80% - 100%
0.64	5.75 - 6.50	40% - 60%	40% - 60%	40% - 60%

The three CP maps show that there are a lot of similarities among the three models. However, the IK and DK models are almost identical. The three maps show that there are a few areas of the field that have a low probability that corn can reach 0.50 of YP under a soil salinity threshold of ≤ 5.9 dS/m. Therefore, most of the field has a high probability that corn can reach at least a 0.5 of YP. Both Figure 6 and Table 5 show that all the normalized yield samples have higher yields than expected.

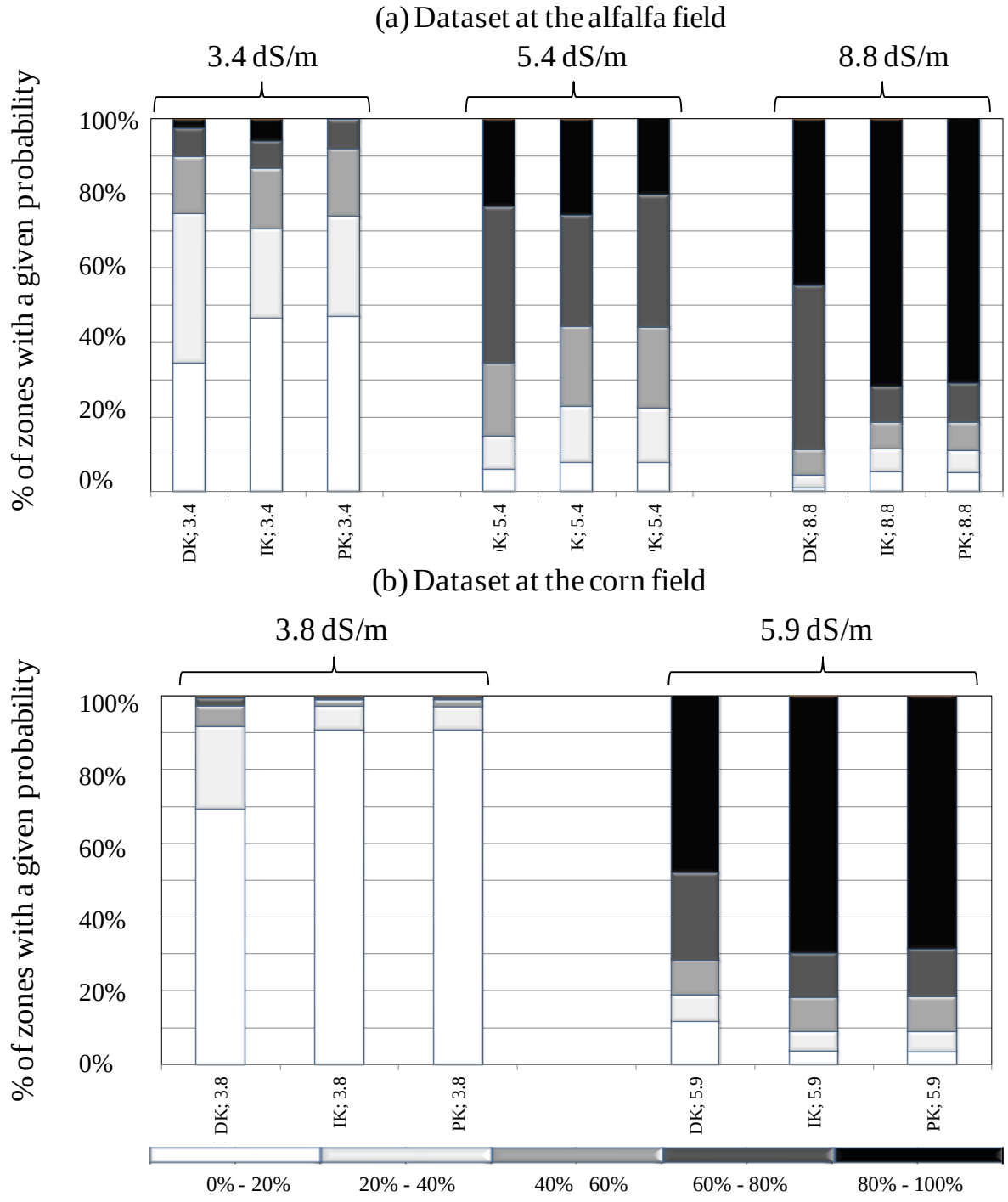


Figure 4.7: Percentages of zones of different probabilities generated from the CP maps developed for the two datasets of soil salinity collected for the alfalfa and corn fields when applying the three models at different soil salinity thresholds.

Figure 4.7 shows the bar charts of the estimated zones generated by the three models for both alfalfa and corn samples. Figure 4.7(a) shows the bar charts for the zones created using the three models for three soil salinity thresholds of 3.4, 5.4, and 8.8 dS/m for conditions of alfalfa reaching YPs of 0.9, 0.75, and 0.5. Figure 7b shows the bar charts for the zones created using the three models for two soil salinity thresholds of 3.8, and 5.9 dS/m as conditions for corn to reach YPs of 0.75, and 0.5. There is a clear trend when moving from low to high soil salinity thresholds for both alfalfa and corn when using the three models. High probability zones are small with low soil salinity thresholds and the zones get larger as the soil salinity thresholds increase. At the same time, low probability zones are large with small soil salinity thresholds and the zones get smaller as the soil salinity thresholds increase. With small soil salinity thresholds, the expectations are for high YP, which are not easy to accomplish, while with high soil salinity thresholds, the expectations are for low YP, which are easy to accomplish. The differences in the performance of the three models can be attributed to the fact that the DK model uses all the information in a dataset, while the IK and PK models do not use all of the data.

4.5 Conclusions

The CP maps generated using the nonlinear kriging techniques provide quantitative information about the expected yield of fields as well as the zones of risk for poor yield as a result of soil salinity. The CP maps were developed for soil salinity datasets at different thresholds for two dominant crops in the study area - alfalfa, a moderately tolerant crop to soil salinity, and corn, a moderately sensitive crop to soil salinity. Crop yield and soil salinity samples were collected for both alfalfa and corn in order to match the CP maps with the actual yield data. The models were evaluated using cross-validation, which helped to select a

well-structured variogram that led to improved model performance. The models were also evaluated by comparing the actual alfalfa and corn yield samples with the CP maps developed by the models. The results presented in this manuscript show that in most cases there is good agreement between the alfalfa and corn yield samples and the estimated CP maps of the three models. Most of the alfalfa and corn yield samples were located in zones that matched what was expected from the models according to Ayers and Westcot (1985). The results show a lot of similarities between the IK and PK models in estimating the zones of different probabilities of the CP maps. However, both the IK and PK models have a slight difference from the DK model in the areas of the estimated zones. The differences in the performance of the three models, can be attributed to the fact that the DK model makes use of all information in a dataset, IK uses most of it, while the PK recovered some of the information lost in the conversion of a continuous variable to a discontinuous one. The information provided in this study can be implemented by using the output of the models as an input for decision making in precision management of agricultural resources. The CP maps developed by the three models can provide information about the estimated yield for a field with a given crop and the zones that are at risk for producing a poor yield with certain soil salinity conditions. The visual and quantitative information presented in the CP maps provides information on the variability of the probability of YP in the different zones of a specific field. With this information, several actions can be considered to enhance the productivity. One possible action might be to switch to another crop. This action depends on several factors, such as: crop price, the initial and farming costs of the crop, water consumption of the crop and water availability, and marketability. Once a crop is selected, growers can improve the productivity by improving the management of low productivity zones to alleviate the impact of soil salinity. Fertilizer application and number of seeds planted can be applied based on the expected spatial yield variability.

4.6 REFERENCES

Ayers, R.S., and D.W. Westcot. 1985. Water quality for agriculture. Food and Agriculture Organization of the United Nations. Rome: Italy.

Burgess, T.M., and R. Webster. 1980. Optimal interpolation and isarithmic mapping of soil properties II: block kriging. Journal of Soil Science 31:333-341.

Cain, D. 1997. U.S. geological survey data collection center and studies in the Arkansas River Basin.

Carr, J.R. 1994. Order relation correction experiments for probability kriging. Mathematical Geosciences 26:605-621.

Carr, J.R. and N. Mao. 1993. A general form of probability kriging for estimation of the indicator and uniform transforms. Mathematical Geosciences 25:425-438.

Colorado Agricultural Experiment Station. 2008. Arkansas Valley Research Center 2006 report. Tech. Rpt. TR08-16 (http://www.colostate.edu/depts/avrc/pubs/tr08-16.pdf).

Cressie, N. 1992. Statistics for spatial data. Terra Nova 4:613-617.

Deutsch, C. 1989. DECLUS: a FORTRAN 77 program for determining optimum spatial declustering weights. Computers and Geosciences 15:325-332.

Deutsch, C.V. and A.G. Journel. 1997. GSLIB: Geostatistical Software Library and User's Guide. 2nd ed. New York City: Oxford U. Press.

Eldeiry, A.A. and L.A. Garcia. 2008a. Spatial modeling of soil salinity using remote sensing, GIS, and field data. VDM Verlag.

Eldeiry, A.A. and L.A. Garcia. 2008b. Detecting soil salinity in alfalfa fields using spatial modeling and remote sensing. Soil Science Society of America Journal 72:201.

Eldeiry, A.A. and L.A. Garcia. 2010. Comparison of ordinary kriging, regression kriging, and cokriging techniques to estimate soil salinity using LANDSAT images. ASCE Journal of Irrigation and Drainage Engineering 136:355.

Eldeiry, A.A. and L.A. Garcia. 2011. Using indicator kriging technique for soil salinity and yield management. ASCE Journal of Irrigation and Drainage Engineering 137:82-93.

Eldeiry, A.A. and L.A. Garcia. 2012. Using disjunctive kriging as a quantitative approach to manage soil salinity and crop yield. ASCE Journal of Irrigation and Drainage Engineering 138(3):211-224.

Flowers, T.J. 2004. Improving crop salt tolerance. Journal of Experimental Botany 55:307-319.

Frasier, W.M. 1999. Irrigation management in Colorado: survey data and findings.

Gates, T.K., L.A. Garcia., R.A. Hemphill, E.D. Morway, and A. Elhaddad. 2012. Irrigation practices, water consumption, and return flows in Colorado's Lower Arkansas River Valley. Field and model investigation. Completion Report No. 221, CAES Report No. TR12-10.

Goovaerts, P. 1994. Comparative performance of indicator algorithms for modeling conditional probability distribution functions. Mathematical Geology 26:389-411.

Goovaerts, P. 1997. Geostatistics for natural resources evaluation. Oxford University Press.

Halvorson, J.J., J.L. Smith, H.J. Bolton, and R.E. Rossi. 1995. Evaluating shrub-associated spatial patterns of soil properties in a shrub-steppe ecosystem using multiple-variable geostatistics. Soil Science Society of America Journal 59:1476-1487.

Hochstrasser, U.W. 1965. Handbook of mathematical functions, M. Abramowitz and I.A. Stegun, ed., Dover, Mineola, NY.

Juang, K.W., and D.Y. Lee. 2000. Comparison of three non-parametric kriging methods for delineating heavy-metal contaminated soils. Journal of Environmental Quality 29:197–205.

Maas, E., and G. Hoffman.1977. Crop salt tolerance: current assessment. ASCE Journal of the Irrigation and Drainage Engineering 103:115–134.

Martinez-Beltran, J., and C.L. Manzur. 2005. Overview of salinity problem. Proceedings of the international salinity forum. Riverside, California. April 2005 311-313

Matheron, G. 1976. A simple substitute for conditional expectation: the disjunctive kriging. In: M. Guarascio, M. David, and C. Huijbregts, editors, Advanced geostatistics in the mining industry. D. Reidel Publishing, Dordrecht, Holland. pp. 221-236.

McBratney A., B. Whelan, T. Ancev, and J. Bouma, 2005. Future directions of precision agriculture. Precis Agric 6, 7-23.

Miller, J., J. Franklin, and R. Aspinall. 2007. Incorporating spatial dependence in predictive vegetation models. Ecological Modelling 202(3-4):225-242.

Myers, D.E. 1982. Matrix formulation of cokriging. Mathematical Geology 14:249-257.

Oyedele, D., A. Amusan, and A. Olu Obi. 1996. The use of multiple-variable indicator kriging technique for the assessment of the suitability of an acid soil for maize. Tropical Agriculture 73:259-263.

Patel, R.S., R. Prasher, R. Bonnell, and R. Boughton. 2002. Development of comprehensive soil salinity index. ASCE Journal of Irrigation and Drainage Engineering 128:185-188

Phillips, J.D. 1985. Measuring complexity of environmental gradients. Vegetation 64:95-102.

Ramos, P., M. Monego, and S. Carvalho. 2008. Spatial distribution of a sewage outfall plume observed with an AUV. Proceedings of the MTS-IEEE Conference Oceans. Quebec, Canada, September 2008.

Rhoades, J.D., N.A. Manteghi, P.J. Shouse, and W.J. Alves. 1989. Soil electrical conductivity and soil salinity: new formulation and calibration. Soil Science Society of America Journal 53:433-439.

Robertson, G.P. 1987. Geostatistics in ecology: interpolating with known variance. Ecology, 68(3):744-748.

Robinson, T. P., and G. Metternich. 2006. Testing the performance of spatial interpolation techniques for mapping soil properties. Comput. Electron. Agric., 50(2): 97-108.

Rogers, M.E. 2002. Irrigating perennial pasture with saline water: effects on soil chemistry. Pasture production and composition. Australian Journal of Experimental Agriculture 42 (3): 265-272.

Smith, J.L., J.J. Halvorson, R.I. Papendick, and others. 1993. Using multiple-variable indicator kriging for evaluating soil quality. Soil Science Society of America Journal 57:743-749.

Solow, A.R. 1986. Mapping by simple indicator kriging. Mathematical Geology 18:335-352.

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

Triantafilis, J., I.O.A. Odeh, and A.B. McBratney. 2001. Five geostatistical models to predict soil salinity from electromagnetic induction data across irrigated cotton. Soil Science Society of America Journal 65:869-878.

United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) Colorado Field Office. 2009. Colorado agricultural statistics 2009. USDA, NASS, Colorado Field Office. Lakewood, CO.

van Meirvenne, M., and P. Goovaerts. 2001. Evaluating the probability of exceeding a site-specific soil cadmium contamination threshold. Geoderma 102:75-100.

von Steiger, B., R. Webster, R. Schulin, and R. Lehmann. 1996. Mapping heavy metals in polluted soil by disjunctive kriging. Environmental Pollution 94:205-215.

Webster, R., and T.M. Burgess.1980. Optimal interpolation and isarithmic mapping of soil properties III: changing drift and universal kriging. Journal of Soil Science 31:505-524.

Wittler, J.M., G.E. Cardon, T.K. Gates, C.A. Cooper, and P.L. Sutherland. 2006. Calibration of electromagnetic induction for regional assessment of soil water salinity in an irrigated valley. ASCE Journal of Irrigation and Drainage Engineering 132:436-444.

Wood, G., M.A. Oliver, and R. Webster. 1990. Estimating soil salinity by disjunctive kriging. Soil Use and Management 6:97-104.

Yates, S.R., A.W. Warrick, and D.E. Myers.1986a. Disjunctive kriging II: examples. Water Resources Research 22:623-630.

Yates, S.R., A.W. Warrick, and D.E. Myers.1986b. Disjunctive kriging I: Overview of estimation and conditional probability. Water Resources Research 22:615-621.

Yates, S., A. Warrick, A. Matthias, and S. Musil. 1988. Spatial variability of remotely sensed surface temperatures at field scale. Soil Science Society of America Journal 52:40-45.

5 SUMMARY AND FINAL REMARKS

The water of the Arkansas River is used to irrigate crops grown in its alluvial valley, and farmers are facing decreasing crop yields due in part to the high levels of salinity. In some areas, land is being taken out of production due to unsustainable crop yields. Attempts to improve the salt tolerance of crops through conventional breeding programs have met with very limited success due to the complexity of the salt tolerance both genetically and physiologically (Flowers, 2004). In the mean time, leaching of soil salinity requires a source of fresh or low salinity water, while land reclamation is expensive. The presented approach in this research is aimed to reach high productivity and live with soil salinity in its current condition without leaching or land reclamation efforts. Geostatistical methods are used in this research to map soil salinity and investigate the heterogeneity of its spatial distribution. The crop productivity is then investigated based on the spatial distribution of soil salinity and the crop tolerance. The threshold values of the investigated crops were used as input values for the geostatistical models to generate CP maps for YP% under different conditions of soil salinity thresholds. The CP maps were used as a quantitative method for making management decisions.

The findings of this research suggest that the CP maps developed using nonlinear geostatistical models can be used for effective assessment of the impact of the soil salinity on the spatial variability of YP. Throughout this research, the accuracy of the three nonlinear geostatistical models was evaluated using the cross-validation parameters. The results of this research indicated that most of the yield samples were located in the same zones as that estimated by the models. The results imply that there were a lot of similarities between the estimated zones of the IK and PK models, while these zones were slightly different from the zones generated using the DK model. According to these results, the similarities or the existing of slight differences among the three nonlinear geostatistical models suggest that any

of these models can be used to develop the CP maps. This is mainly due to the fact that all the three models have the ability to take data uncertainty into account (Goovaerts, 1994; Oyedele et al., 1996). However, the crops investigated in this research is limited to alfalfa and corn, which is not enough to cover all the crops in the study area. Moreover, soil salinity was considered to be the only variable that impacting the YP, which may be not in some instances. In addition, the possible impacting of some other parameters on YP such as fertilizer application, irrigation and drainage schemes, and insect infestations was neglected. Along with more field variables, there is still a great deal of future work to be done in this domain. This research can be improved by taking into consideration some other parameters in addition to soil salinity and by selecting a wide varieties of crops. Finally, the results of this research indicate that the CP maps of the nonlinear geostatistical models are potentially useful tool for farmers and researchers in order to control and manage YP with soil salinity.