Mapping Soil Salinity Using Soil Salinity Samples and Variograms: Case Study in the Lower Arkansas Basin

Ahmed Eldeiry[1] and Luis A. García
Decision Support Group, Department of Civil and Environmental Engineering, Colorado State University
Fort Collins, CO 80523-1372

Abstract. The objective of this study was to develop a methodology to generate high accuracy soil salinity maps with the minimum number of soil salinity samples. Variograms are used in this study to estimate the number of soil salinity samples that need to be collected. A modified residual kriging model was used to evaluate the relationship between soil salinity and a number of satellite images. Two datasets, one representing corn fields where Aster, Landsat 7, and Ikonos images were used, and the other representing alfalfa fields where the Landsat 5 and Ikonos images were used. The satellite images were acquired from different sources to check the correlation between measured soil salinity and remote sensing data. Two strategies were applied to the datasets to produce subset samples. For the corn fields dataset, nine subsets of the data ranging from 10% to 90% of the data in 10% increments were produced. For the alfalfa fields dataset, three subsets of the data 75 %, 50%, and 25% of the data were produced. A modified residual kriging model was applied to the reduced datasets for each image. For each combination of satellite image and subset of the data, a variogram was generated and the correlation between soil salinity and the remote sensing data was evaluated. The results show that the variograms can be used to significantly reduce the number of soil salinity samples that need to be collected.

1. Introduction

Soil salinity is a severe environmental hazard that increasingly impacts crop yields and agricultural production. Soil salinity refers to the presence in soil and water of various electrolytic mineral solutes in concentrations that are harmful to many agricultural crops (Hillel, 2000). Natural salinization or primary salinization results from the long-term influence of natural processes. In contrast, human-induced salinization or secondary salinization is the result of salt stored in the soil profile being mobilized by extra water provided by human activities such as irrigation (Szabolcs, 1989). In 1999, Postel (1999) stated that worldwide, one in five hectares of irrigated land suffers from a build-up of salts in the soil, and vast areas in China, India, Pakistan, Central Asia, and the United States are losing productivity. Postel (1999) estimates that soil salinization costs the world's farmers $11 billion a year in reduced income and warns that the figure is growing. The spread of salinization, at a rate of up to 2 million hectares a year, is offsetting a good portion of the increased productivity achieved by expanding irrigation. It has been estimated (Ghassemi et al., 1995) that close to 1 billion hectares (about 7% of the earth's landscape) are affected by primary salinity, while about 77 million hectares have been salinized as a consequence of human activities, with 58% of these concentrated in irrigated areas. On average, 20% of

[1] Civil and Environmental Engineering Department
Colorado State University
Fort Collins, CO 80523-1372
Tel: (970) 491-7620
e-mail: aeldeiry@engr.colostate.edu

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. The development of saline soils is a dynamic phenomenon that needs to be monitored regularly in order to secure up-to-date knowledge of their extent, spatial distribution, nature and magnitude (Ghassemi et al., 1995).

Remote sensing of surface features using aerial photography, videography, infrared thermometry and multispectral scanners has been used intensively to identify and map salt-affected areas (Robbins and Wiegand, 1990). Multispectral data acquired from platforms such as Landsat, SPOT, and the Indian Remote Sensing (IRS) series of satellites have been found to be useful in detecting, mapping and monitoring salt affected soils (Dwivedi and Rao, 1992). Procedures for using soil salinity, plant information, and digitized color infrared aerial photography and videography have been developed to help with determining soil salinity (Wiegand et al., 1994). Other related approaches such as using spectral brightness coefficients and image photodensities for areas known to have specific characteristics have also been developed (Golovina et al., 1992). For mapping surface land salinity, color and thermal infrared aerial photography, spectral image interpretation techniques, such as satellite imagery (Landsat TM, or SPOT), and other airborne remote sensing techniques are used (Spies and Woodgate, 2004). Other techniques, such as gamma radiometrics (Wilford et al., 2001), are useful for mapping soils and shallow sub-soil materials that can assist with interpretation of likely recharge and discharge areas.

Geostatistical methods provide a means to study the heterogeneous nature of the spatial distribution of soil salinity. The results of a study by Pozdnyakova and Zhang (1999) suggest that sampling costs can be dramatically reduced and estimation can be significantly improved by using cokriging. Sample variograms of soil electrical conductivity can be a useful tool in selecting the distance between soil sampling points for laboratory electrical conductivity determination (Utset et al., 1998). In geostatistical theory, the range of the variogram is the maximum distance between correlated measurements (Journel and Huijbregts, 1978; Webster, 1985; Warrick et al., 1986). This means that samples separated at smaller distances are generally not needed (Nielsen et al., 1983). Therefore, the range of soil salinity variograms can be an effective criterion for the selection of a sampling design in mapping soil salinity. Sampling incorporates concepts of survey intensity, spatial variability, and mapping scale, and is usually the most costly aspect of a survey (Webster and Oliver 1990). In a conventional soil survey, sampling sites are selected subjectively by surveyors to support their mental predictive model of soil occurrence, a so-called free survey (White 1997). Such design are purposive and non-random, and do not provide statistical estimates. By contrast, a pedometric soil survey (McBratney et al., 2000) aims at statistical modeling of soil cover, including uncertainty about the predictions using objective techniques.

Geostatistical methods such as kriging are becoming commonly used estimation techniques to generate soil maps. Kriging has been applied to quantify spatial variability of a number of parameters in soil science. Tabor et al. (1984, and 1985) used variograms and kriging to determine the spatial variability of nitrates in cotton petioles and analyzed spatial variability of soil nitrate and correlated variables. Istok and Cooper (1988) and Cooper and Istok (1988a, b) applied kriging to study groundwater contamination. Yates et al. (1993) used geostatistics in the description of salt-affected soils. 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. Yates et al. (1986 a, b) used disjunctive kriging to present spatial distributions and corresponding conditional probability maps of soil electrical conductivity.

The approach presented in this paper involve integrating remotely sensed data, GIS, and field observations of soil salinity to evaluate the most appropriate spatial interpolation techniques to use to develop high quality soil salinity maps. The approach was tested on soil salinity data observed in the Lower Arkansas River Valley near the Kansas border in Colorado. The correlation between soil salinity data and the satellite images was based on the crop cover reflection as an indicator of soil salinity. For the corn fields dataset the methodology was applied to nine subsets of the data ranging from 10% to 90 % of the data in 10% increments that were randomly sampled to evaluate the influence of sample size on the ability to spatially interpolate soil salinity data. For the alfalfa fields dataset, three subsets of the data 75 %, 50%, and 25% of the data were produced. The range of each variogram was used to evaluate the distance between the collected soils salinity samples.

2. Site Description

This research is part of Colorado State University's Arkansas River Basin Salinity Mapping Project, a project which began in 1999. The study area for this project is shown in figure 1 and is located in southeastern Colorado. Fields in this area are cultivated with alfalfa, corn, wheat, onions, cantaloupe and other vegetables and are irrigated by a variety of systems including a mixture of border and basin, center pivot, and furrow. Salinity levels in the canal systems along the river, which provide much of the region's irrigation water, increase from 300 ppm total dissolved solids (TDS) near Pueblo to over 4,000 ppm at the Colorado-Kansas border (Gates et al. 2002).

For the particular research dealing with remote sensing of salinity in corn and alfalfa fields described in this paper, the study area consists of a number of fields located in an area of about 20 miles in length and 10 miles in width. Soil salinity was measured in these fields at the beginning and at the end of the irrigation season. The location of the corn and alfalfa fields is shown in Figure 1.

3. Methodology

3.1. Using an EM-38 for Soil Salinity Readings

Soil salinity was measured in the fields using an EM-38 electromagnetic probe. The EM-38 takes vertical and horizontal readings that can be converted to soil salinity estimates. When collecting geo-referenced soil salinity data, the EM-38 probe was used in conjunction with handheld GPS units (with 95% accuracy of approximately 2 meters) to obtain geographic coordinates of the observed soil salinity data points. The EM-38 can cover large areas in a fairly short amount of time without the need for ground electrodes and provides depths of exploration of 1.5 meters and 0.75 meters in the vertical and horizontal dipole modes respectively.

EM-38 readings are affected by soil moisture and soil temperature. Therefore, the EM-38 readings must be calibrated. For the calibration, soil moisture, soil temperature, and EM-38 readings were taken in a number of fields in the study area. The following equation was developed for use in the study area for the calibration of the EM-38 taking into consideration both soil moisture content and soil temperature (Wittler et al., 2006). The vertical reading of the EM-38, EM-38v, was corrected as follows:

$$EM\text{-}38v_c = EM\text{-}38v * f_{tc} \quad (1)$$

where f_{tc} is the temperature correction factor and obtained by using the following equation:

$$f_{tc} = 1.8509 - 0.0516951(T) + 0.000858442(T^2) - 0.00000613535(T^3) \quad (2)$$

where T is the corrected soil temperature and equals to ($T_{measured}$ -25)/10. $T_{measured}$ is the soil temperature (^oC) measured in the field in ^oC, (Richards, 1954).

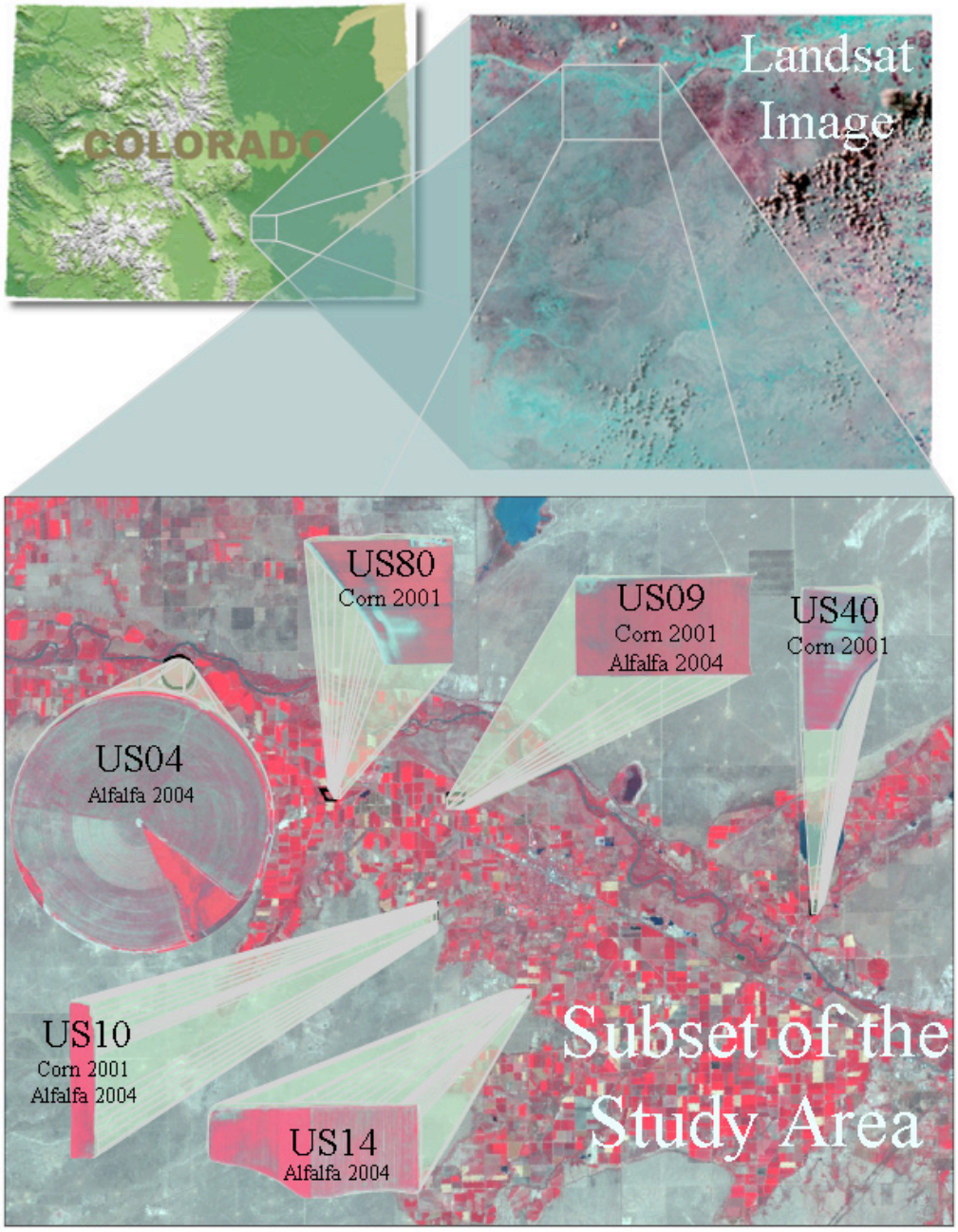

Figure 1. The location of the corn and alfalfa fields.

Finally, soil salinity (SS_a in dS/m) is obtained by adjusting the EM-38v_c as follows:

$$SS_a = 0.45 + 7.23 * \left(\frac{EM - 38v_c}{100}\right)^{1.78} + 19.54 * (GMC) - 34.06 * \left(\frac{EM - 38v_c}{100}\right) * (GMC) \quad (3)$$

where GMC, is the gravimetric moisture content of the soil sample.

3.2. Satellite Imagery

Four satellite image types were evaluated for their ability to estimate soil salinity. The ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) (Yamaguchi et al. 1998) sensor is an imaging instrument flown on the Terra satellite launched in December 1999. ASTER is a cooperative effort between NASA and Japan's Ministry of Economy and has been designed to acquire land surface temperature, emissivity, reflectance, and elevation data. An ASTER scene covers an area of approximately 60 km by 60 km and consists of 14 bands of data, three bands in the visible and near infrared (VNIR) with a 15m resolution, six bands in the short wave (SWIR) with a 30m resolution, and five thermal bands (TIR) with a 90m resolution. The Aster image was acquired on August 16th, 2001 and all bands were resampled to 30 meter resolution. Landsat 7 images have three visible bands (blue, green, and red), 1 near infrared band (NIR), and 2 shortwave infrared bands (MIR-1, MIR-2) at 30m resolution; a thermal infrared band (TIR) at 60m resolution; and a panchromatic (PAN) band with 15m resolution. The Landsat 7 image was acquired on July 8th, 2001 and was also resampled to 30 meter resolution. Ikonos images have three bands in the visible and one in the near infrared with a resolution of 4m, and a panchromatic band with 1m resolution. The Ikonos image was acquired on July 11th, 2001. The Landsat 5 images contain seven bands including three visible bands (band1, band2, and band 3) with 30 meter resolution, two NIR bands (band4 and band5) with 30 meter resolution, one thermal band (band6 with 120 meter resolution), and a Mid IR (band7 with 30 meter resolution). The Landsat 5 image was obtained on August 9th, 2004. An additional Ikonos image was acquired on August 1st, 2004.

The four types of satellite images are highly variable in spectral and spatial resolution, with a range of four to fourteen spectral bands and 1m to 90m spatial resolution. This variability provides the opportunity to explore the use of spatial and spectral resolution for predicting soil salinity.

Spatial distortion of the images was corrected using a geometric correction in ERDAS Imagine 8.7 (ERDAS, 2006), to guarantee that the points on the image match with the same points on the ground. A dark object correction was used to compensate for the effect of atmospheric scattering (Song 2001). The IKONOS images were mosaicked because each individual image covers a small portion of the study area, while the portion of the Landsat 7 image that covers the study area was subset. The Normalized Difference Vegetation Index (NDVI) was added to the bands of the three images. The NDVI uses the contrast between red and infrared reflectance as an indicator of vegetation cover and vigor. The Normalized Difference Vegetation Index (NDVI) was developed to provide an indicator of the amount of vegetation in each of the fields (Wiegand et al., 1994; Hill and Donald, 2003).

3.3. Using Variograms With the Different Datasets

For each dataset the variogram was generated to evaluate the distance between correlated soil salinity samples. Depending on how well the variogram fits, it can be used to indicate how important it is to keep a given sample or remove it. The observed datasets were tested first with the OLS model and the derived equation was applied to the combination of bands to generate the predicted surface using the OLS model. Then from the whole dataset nine subsets were generated and the OLS model was applied to these subsets. For each subset, an equation was derived for the OLS model using only the subset points and then this equation was applied to the combination of bands. The kriged residual surface was generated for each subset and then combined with the OLS surface.

For each set of data a variogram was produced. This variogram was used to evaluate the distance between the correlated soil salinity samples. For corn each time the dataset was reduced by 10%, the range of the variogram and how well it fit provided an indication if the sample points that were removed were needed given the distance between these points and the range of the variogram. The same thing was done to produce viriogram for alfalfa and at each time the dataset was reduced by 25%.

3.4. Modeling Approach

Multiple regression analysis was used to explore the coarse-scale variability in the soil salinity as a function of the Satellite images bands.

$$\hat{z}(s_o) = \hat{\beta}_0 + \hat{\beta}_1 x_1(s_o) + \ldots + \hat{\beta}_k x_k(s_o) \qquad (4)$$

where $\hat{z}(s_o)$ is the predicted soil salinity at spatial location, s_0, $\hat{\beta}_i$ are estimated regression coefficients and x_i are the independent variables at spatial location, s_0. A stepwise procedure was used to identify the best subset of satellite bands to include in the regression models that minimized the AICC (Akaike, 1997; Brockwell and Davis, 1991).

The spatial structure of the residuals from the ordinary least squares (OLS) multiple regression models were analyzed using a geostatistical method, the variogram, which has been widely used to analyze spatial structures in ecology (Phillips 1985, Robertson 1987). The sample variogram, $\hat{\gamma}(h)$ is estimated using the following equation:

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[\varepsilon(s_i) - \varepsilon(s_i + h)\right]^2 \qquad (5)$$

where $\hat{\varepsilon}(s_i)$ 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 distance *h*, *N(h)* is the total number of pairs of samples separated by distance *h*. The empirical variogram, which is a plot of the values of $\hat{\gamma}(h)$ as a function of *h*, gives information on the spatial dependency of the variable. Exponential, Gaussian and Spherical models were fitted to the sample variograms using a weighted least squares method (Robertson 1987) as shown in figure 3. The variogram model with the smallest AICC was selected to describe the spatial dependencies in the salinity data.

If the residuals were spatially correlated, ordinary kriging was used to model the spatial distribution of salinity in the fields. At every location where a soil salinity sample was not taken, estimates of the true unknown residuals, $\varepsilon(s_o)$ at spatial location, s_o, were obtained

using a weighted linear combination of the available soil salinity samples at spatial locations, s_i:

$$\hat{\varepsilon}(s_o) = \sum_{i=1}^{n} w_i \hat{\varepsilon}(s_i) \quad (6)$$

where the set of weights, w_i takes into consideration the distances between soil salinity sample locations and spatial continuity, or clustering between the soil salinity samples. The best fitting variogram model was used to describe the spatial continuity in estimating the kriging weights.

4. Results And Analysis

The stepwise procedure was used to identify the best combinations bands of each image that correlates with soil salinity. In most instances the residuals from the OLS multiple regression models were spatially correlated. Kriging the residuals generally improved the predictive performance of the models.

4.1. Variograms and Number of Collected Samples

Figure 2 is an example of the different variograms used in the kriged residuals of the 2001 corn fields with the Landsat 7 image for all datasets. The figure shows that variograms can fit well with all datasets but it cannot be fitted for the datasets of 30%, 20%, and 10% which prevents the use of kriging for these three datasets. This means that the whole dataset through 40% of the data produce reasonable variograms. These results show that there is a significant number of points collected that can be removed without having a large impact on the accuracy of the interpolation technique. Figure 2 also shows that there is no significant difference among the three variogram models.

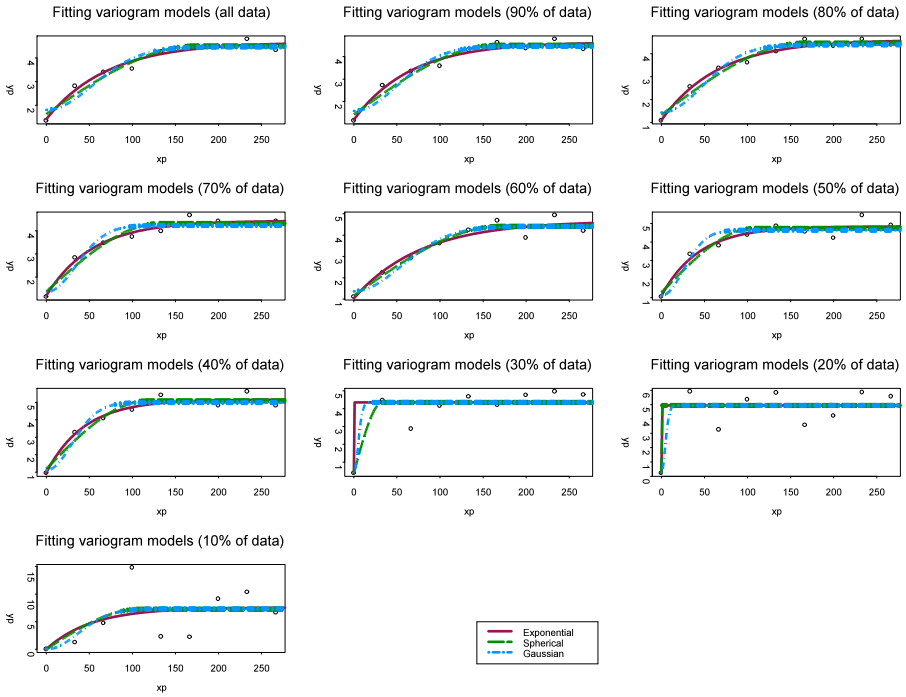

Figure 1: Variogram models for the Landsat7 image with different sets of data.

Table 1: The variogram models parameters using 2001 corn datasets with the Aster, Landsat7 and Ikonos images.

	Aster		Ikonos		Landsat 7	
All data						
	Range	AICC	Range	AICC	Range	AICC
Gaussian	235.3	33.5	73.1	26.2	132.9	27.8
Spherical	280.7	27.4	124.1	24.7	174.2	21.9
Exponential	181.6	26.2	40.8	24.0	71.1	18.2
90% of data						
Gaussian	236.6	35.7	62.3	27.5	123.7	28.1
Spherical	298.8	29.2	96.1	26.9	166.0	22.1
Exponential	167.0	26.4	35.2	26.3	66.2	18.9
80% of data						
Gaussian	184.1	36.4	61.7	22.9	116.7	26.4
Spherical	262.2	31.0	93.6	22.2	166.0	20.1
Exponential	119.6	26.1	61.7	45.3	66.5	14.7
70% of data						
Gaussian	143.1	36.8	61.8	27.7	84.2	27.5
Spherical	198.5	32.3	94.9	27.2	132.0	24.1
Exponential	91.3	25.0	36.3	26.3	53.9	17.4
60% of data						
Gaussian	194.9	40.7	106	25.7	135.9	28.4
Spherical	281.4	36.9	142.4	22.0	170.7	25.9
Exponential	114.5	31.7	49.2	21.5	78.4	27.9
50% of data						
Gaussian	278	32.1	88.3	29.3	63.0	33.1
Spherical	371.7	26.4	133	26.6	107.5	31.6
Exponential	388.5	26.1	42.2	26.2	43.7	28.3
40% of data						
Gaussian	231.7	39.3	91.2	27.8	73.5	31.4
Spherical	300.8	33.0	134.4	23.7	116.5	29.0
Exponential	181.1	30.5	47.1	21.2	44.2	25.9
30% of data						
Gaussian	456.1	46.5	137.6	33.2	12.3	39.3
Spherical	379	45.4	164.8	30.9	33.3	39.3
Exponential	310.2	45.3	70.7	32.0	0	39.3
20% of data						
Gaussian	150	56.5	33	41.8	11.0	48.2
Spherical	150	56.5	115.2	42.8	0	48.2
Exponential	150	56.5	24.4	41.6	0.4	48.2
10% of data						
Gaussian	18315	57.3	67.1	57.7	91.1	72.5
Spherical	17.9	55.1	78.0	57.8	108.0	72.6
Exponential	59279	57.9	33.7	58.3	52.8	73.1

Table 2: Variogram models parameters using the Ikonos and Landsat 5 images for 2004 alfalfa field datasets.

	Ikonos		Landsat 5	
	Range	AICC	Range	AICC
All data				
Gaussian	248.4	48.6	130.9	48.9
Spherical	315.6	46.0	151.9	48.1
Exponential	269.8	45.8	84.6	47.6
75% of data				
Gaussian	300.4	49.6	171.7	55.5
Spherical	428.4	48.2	164.1	55.3
Exponential	699.3	48.3	122.0	54.1
50% of data				
Gaussian	298.2	56.0	169.6	61.2
Spherical	391.8	55.5	160.3	61.1
Exponential	658.6	55.6	153.0	59.9
25% of data				
Gaussian	141.6	76.4	297.2	63.1
Spherical	168.3	75.7	347.4	63.2
Exponential	681.6	75.6	681.3	63.4

Table 1 shows the parameters (range, and AICC) of the Gaussian, Spherical, and Exponential models for the 2001 corn datasets (Aster, Landsat 7 and the Ikonos) for the different sample sizes. Table 2 show the same parameters for the 2004 alfalfa datasets (Landsat 5 and Ikonos). The most important parameter in selecting the model is the AICC (Akakie Information Corrected Criteria) value. The AICC of the exponential model has the smallest value among the three variograms, in most cases, which makes it the best choice. The AICC values of the three models increase as the datasets decreases. The values of the AICC of the 2001 corn datasets are generally less than those of the 2004 alfalfa datasets. The Ikonos images for both the 2001 corn datasets and the 2004 alfalfa datasets are less than those of the other images. In most of the datasets, the values of AICC and range increase as the sample size decreases.

From figure 2 and tables 1 and 2, the Exponential model is the closest to the points while the Gaussian and the Spherical models deviate more from the data. In most of the cases, the Exponential model performed the best.

4.2. Comparison Among the Predicted Data from the Five Images

Table 3 and 4 show the MAE for the 2001 cron data (Aster, Landsat 7, and Ikonos) and the 2004 alfalfa data (Landsat 5 and Ikonos) of the predicted soil salinity for all datasets. There is no data for field US10 with Ikonos image since it was not covered by that image. The tables show that both the MAE values are smaller when using larger datasets and increase when using smaller datasets. For fields with a low range of soil salinity the MAE values are small, regardless of the size of the datasets such as US09 while it getting higher for fields with high range of soil salinity such as US04.

Table 3: Mean absolute error (MAE) values (dS/m) of the predicted soil salinity from the 2001 corn fields using Aster, Landsat 7, and the Ikonos images for the different sets of data.

	All	90%	80%	70%	60%	50%	40%	30%	20%	10%
Aster										
US09	0.79	0.52	0.38	0.39	0.39	0.78	0.41	1.33	1.07	1.79
US10	1.45	1.41	1.39	1.34	1.38	1.85	1.52	2.16	2.50	3.10
US40	1.01	1.07	1.00	0.94	1.02	1.37	0.98	1.44	1.50	1.96
US80	0.74	0.78	0.67	0.68	0.76	1.09	0.70	1.29	1.69	1.96
All	0.98	0.90	0.81	0.79	0.83	1.22	0.85	1.53	1.61	2.15
Landsat 7										
US09	0.28	0.26	0.27	0.33	0.43	0.44	0.42	0.53	0.71	1.70
US10	1.21	1.22	1.21	1.21	1.28	1.40	1.57	1.69	1.92	2.14
US40	0.89	0.87	0.88	0.92	0.89	1.12	1.18	1.45	1.46	2.20
US80	0.79	0.77	0.77	0.93	0.90	1.05	1.18	1.55	2.17	2.77
All	0.73	0.73	0.73	0.79	0.82	0.94	1.01	1.22	1.46	2.14
Ikonos										
US09	0.52	0.53	0.53	0.57	0.68	0.68	1.13	0.74	0.66	1.14
US10										
US40	0.61	0.65	0.72	0.79	0.96	1.07	2.94	1.32	1.33	2.03
US80	0.37	0.43	0.44	0.47	0.62	0.73	3.28	0.83	0.95	1.48
All	0.51	0.54	0.57	0.62	0.75	0.81	2.27	0.95	0.94	1.50

Table 4: Mean absolute error (MAE) values (dS/m) of the predicted soil salinity from the 2004 alfalfa fields using Landsat 5, and Ikonos images for the different sets of data.

	All	75%	50%	25%
Landsat 5				
US04	1.94	2.70	3.45	4.73
US09	0.38	0.49	0.61	0.88
US10	1.65	1.75	1.85	2.91
US14	0.69	0.76	0.95	1.33
All	1.16	1.43	1.71	2.46
Ikonos				
US04	1.76	2.61	2.73	4.54
US09	0.30	0.34	0.46	0.67
US10	1.08	1.37	1.36	1.74
US14	0.49	0.63	0.77	1.86
All	0.91	1.24	1.33	2.20

4.3. Example of Predicted Maps

Figures 3 show examples of the observed and predicted surfaces of soil salinity (dS/m) for field US40 using the Aster image for the corn fields in 2001 for the observed data and all datasets predicted from Aster image. The number of points collected in this field is 79, the area is 8.2 hectares and the range of soil salinity is 3.0 – 12.2 dS/m. The generated of the predicted surface using all the observed points captured the range of variability. The predicted soil salinity surfaces captured less variability as the number of dataset points de-

creases. The predicted generated surfaces are acceptable until it reached the dataset of the 30% of data.

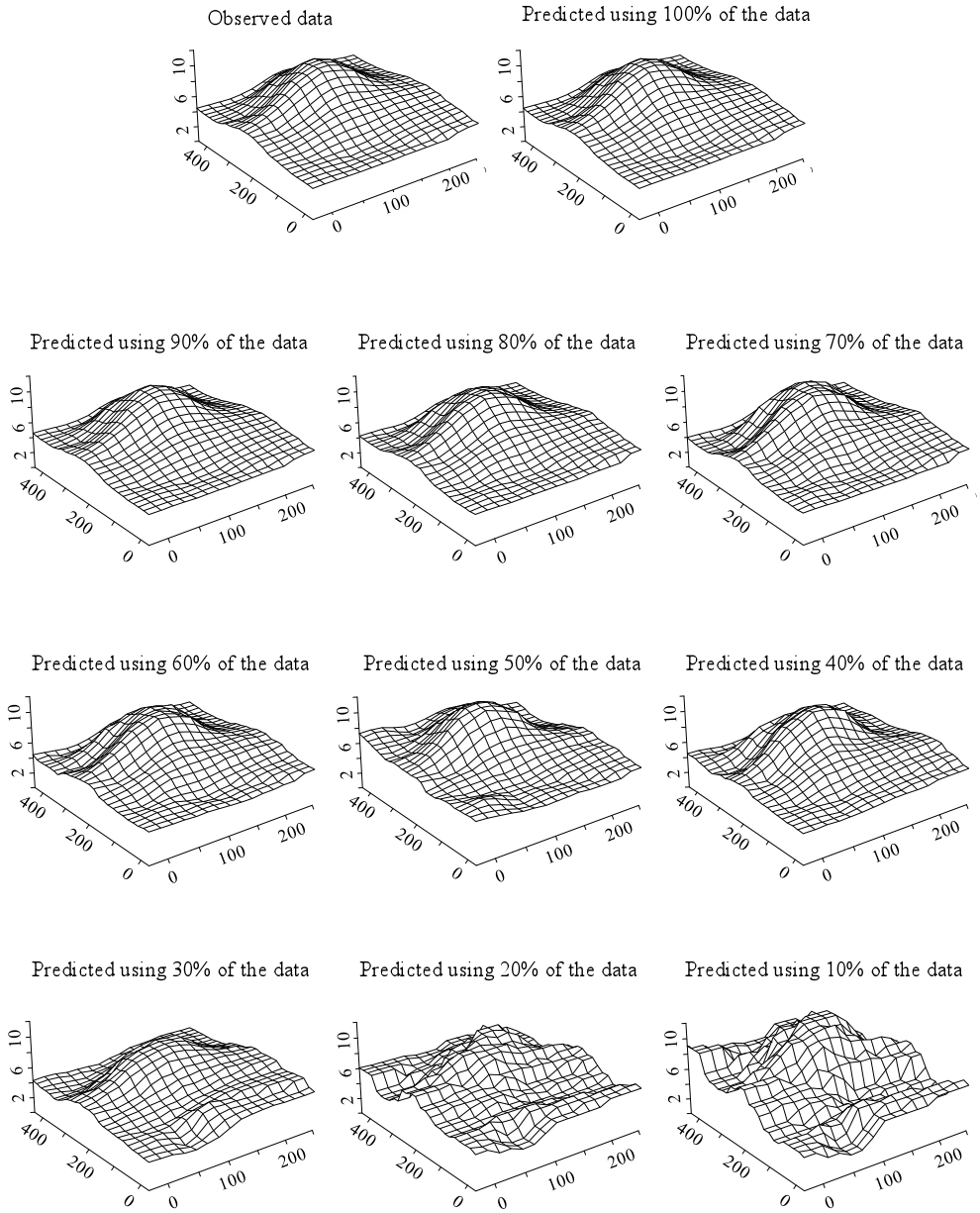

Figure 3: Observed and predicted surfaces of soil salinity (dS/m) of field US40 from the Aster 2001image using all the observed data and subsets.

4.4. Cross Validation

Table 5 shows the statistics of the observed dataset of the five images compared with the estimated data for corn and alfalfa fields. For the dataset of the corn fields for the year 2001, the datasets for the Aster and Landsat 7 have 326 points while the dataset has 257 points for the Ikonos image since field US10 was not covered by that image. Therefore, the statistics for the observed data are different and that is why they are separated. The stan-

dard deviation values of the observed and predicted data are very close for all datasets except the value of the Landsat 5 (2004) which is slightly higher than that of the observed data. The values of the coefficient of variation are 1.00 or less for all datasets which means that the distributions of the above datasets are considered to have low variance. The values of the Mean Square Errors (SMSE) for all datasets are 1.00 or less. The values of first quartile and third quartile (1^{st} Q and 2^{nd} Q) of the observed and predicted values are close to each other except the predicted values using the Landsat 7 (2001) and Ikonos (2001) images which are smaller than the observed value. The observed values of the mean compared to the predicted values of all datasets are very close to each other except the value of the Ikonos 2004 which is slightly higher than the observed data.

Table 5: cross validation parameters of datasets for all images.

Dataset	N	Stdev	CV%	SMSE	1^{st} Q	Mean	3^{rd} Q
Observed for Aster (2001) & Landsat 7 (2001)	326	3.1	0.60	NA	3.1	5.13	6.03
Predicted Aster (2001)	326	3.12	0.62	0.95	2.86	5.01	6.33
Predicted Landsat 7 (2001)	326	3.31	0.64	1	1.33	5.17	7.69
Observed for Ikonos (2001)	257	2.24	0.49	NA	3.00	4.55	5.4
Predicted Ikonos (2001)	257	2.64	0.56	1.00	0.41	4.69	6.34
Observed for Landsat 5 (2004) & Ikonos (2004)	256	4.46	0.78	NA	2.7	5.71	6.67
Predicted Landsat 5 (2004)	256	6.46	1.16	0.99	2.42	5.56	7.5
Predicted Ikonos (2004)	256	4.02	0.56	1	2.61	7.21	7.69

5. Conclusion

This research has shown that mapping and assessing soil salinity can be done by integrating field data, GIS, remote sensing, and spatial modeling techniques. However, any integration of field data, GIS, and remote sensing is considered weak unless strong statistical measures are introduced. The model that satisfies the assumptions, selection criteria and has no autocorrelation in the residuals is not considered the best unless the predicted values of soil salinity match up relatively well with the observed values. The results presented in this paper show the importance of considering the variability in the samples collected in the study field rather than the area of the study field itself. The area of the study field is not important when compared to the variability of the soil salinity in considering the number of collected points. This study introduces a methodology to collect and represent soil salinity accurately with less data.

6. References

Akaike, H. (1997). On entropy maximization principal. In: *Application of Statistics.* P.R. Krishnaiah (ed.), pp. 27-41. Amsterdam, North-Holland.

Brockwell, P.J. & Davis, R.A. (1991). Time series: Theory and Methods. New York, 577 p.

Dwivedi, R.S., and Rao, B.R.M., (1992). "The selection of the best possible Landsat TM band combination for delineating salt-affected soils." *International Journal of Remote Sensing,* 13 (11), 2051–2058.

Gates, T.K., Burkhalter, J.P., Labadie, J.W., Valliant, J.C., & 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.

Golovina, N.N., Minskiy, D.Ye., Pankova, I., and Solov'yev, D.A. (1992). "Automated air photo interpretation in the mapping of soil salinization in cotton-growing zones." *Mapping Sciences and Remote Sensing,* 29, 262-268.

Hill, M.J., and Donald, G.E., 2003. "Estimating spatio-temporal patterns of agricultural productivity in fragmented landscapes using AVHRR NDVI time series." *Reomte Sensing of Environment,* 84 (3), 367-384.

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

Istok, J.D. and Cooper, R.M., geostatistics applied to groundwater pollution. III. Global estimates. Journal of environmental Engineering 114, 915 (1988)

Journel, A.G. & Huijbregts, Ch.J. (1978). Mining geostatistics. London: Academic Press.

McBratney, A.B. Odeh, I.O.A., Bishop, T.F.A., Dunbar, M.S., Shatar, T.M., 2000. An overview of pedometric techniques for use in soil survey. Geoderma 97, 293-327, doi:10.1016/S0016-7061(00)00043-4.

Nielsen, D.R., Tillotson, P.M., Viera, S.R., 1983. Analyzing field measured soil-water properties. Agriculture Water Management 6, 93-109.

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

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

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

Richards, L. A. (1954). Diagnosis and improvement of saline and alkali soils. *U. S. salinity laboratory staff agriculture handbook No. 60,* USDA, Washington, D.C.

Robbins, C.W., and Wiegand, C.L. (1990). "Field and laboratory measurements." *Agricultural Salinity Assessment and Management,* American Society of Civil Engineers, New York.

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

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

Song, C., Woodcock, C. E., Seto, K. C., Lenney, M. P., & Macomber, S. A. (2001). Classification and Change Detection Using Landsat TM Data: When and How to Correct Atmospheric Effects? *Remote Sensing of Environment*, *75,* 230-244.

Spies, B. and Woodgate, P.,2004. Salinity Mapping in the Australian context. Technical Report. Land and Water Australia. 153p.

Szabolcs, I. (1989). *Salt-Affected Soils.* CRC Press, Boca Raton, FL.

Tabor, J.A., Warrick, A.W., Myers, D.E., and Pennington, D.A., Spatial Variability of nitrate in irrigated cotton: I. Soil nitrate and correlated variables. Soil Science Society of America Journal. 49, 390 (1985).

Tabor, J.A., Warrick, A.W., Pennington, D.A., and Myers, D.E. Spatial Variability of nitrate in irrigated cotton: I. Petioles. Soil Science Society of America Journal. 48, 602 (1984).

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

Warrick, A.W., Myers, D.E., Nielsen, D.R., 1986. Geostatistical Methods Applied to Soil Science. SSSA, Agronomy Monograph No. 9.

Webster, R., Oliver M.A., "1990. Statistical methods in soil and land survey, Spatial information systems." (Oxford University Press: Oxford,UK)

White, R.E., 1997 'Principles and practice of soil science. The soil as a natural resource.' (Blackwell Science Ltd: Oxford)

Wiegand, C.L., Rhoades, J.D., Escobar, D.E., and Everitt, J.H. (1994). "Photographic and Videographic Observations for Determining and Mapping the Response of Cotton to Soil Salinity." *Remote Sens. Environ.*, 49, 212-223.

Wilford, J., R., Dent, D. L., Dowling, T. and Braaten, R., 2001. Rapid mapping of soils and salt stores using airborne radiometrics and digital elevation models. AGSO Res. News., May 2001, p. 33-40.

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

Yamaguchi, Y., Kahle, A. B., Tsu, H., Kawakami, T., & Pniel, M. (1998). Overview of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). *IEEE Transactions on Geoscience and Remote Sensing, 36,* 1062−1071.

Yates, S.R., Warrick, A.W., and Myers, D.E., Disjunctive kriging. I. Overview of estimation and conditional probability. *Water Resour. Res.* **22**, 615(1986a).

Yates, S.R., Warrick, A.W., and Myers, D.E., Disjunctive kriging. I. Examples. *Water Resour. Res.* **22**, 615(1986b).

Yates, S.R., Zhang, R., Shouse, P.J., and van Genuchten, M.Th. Use of geostatistics in the description of salt-affected soils, in Water Flow and Solute 1qTransport in soils, edited by Russo D. and Dogan G. (Springer-Verlag, New York, 1993), pp.283-304.