EVALUATING THE SOIL MOISTURE CONTENT THROUGH DIFFERENT INTERPOLATION METHODS

D. Yıldırım[1], B. Cemek[2], A. Ünlükara[3]

[1]Black Sea Agricultural Research Institute, Irrigation and Land Reclamation Unit, Samsun, Turkey
[2]Department of Agricultural Structures and Irrigation, Faculty of Agriculturem, Ondokuz Mayıs University, Samsun, Turkey
[3]Department of Biosystems Engineering, Seyrani Faculty of Agriculture, Erciyes University, Kayseri, Turkey

Abstract
Water is vital for the plant growth. An adequate amount of soil moisture content is required in order to increase plant growth and yield. The spatial distribution can be determined using different methods for different depths of soil moisture content. In this study the spatial distribution is created at four different soil depths (30, 60, 90 and 120 cm) using deterministic and stochastic methods. In order to determine the most appropriate methods, Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values were compared between the methods. The lowest RMSE (11.296) and MAE (7.821) values were obtained for 0-30 cm depth of soil moisture content by Ordinary Kriging method. As for the depth of 30-60 cm, the lowest RMSE (13.682) and MAE (8.444) values were obtained through Inverse Distance Weight (IDW). As for the depth of 60-90 cm, the lowest RMSE (17.767) and MAE (11.473) values were obtained through the Radial Basis Function (RBF). As for the depth of 90-120 cm, the lowest RMSE (20.24) and MAE (14.18) values were obtained through the IDW method. The soil moisture content maps have been prepared for 0-30 cm, 30-60 cm, 60-90 cm and 90-120 cm soil depths based on these methods.

Key words: interpolation methods, Ordinary Kriging, radial basis function, inverse distance weight, soil moisture.

INTRODUCTION

The soils can vary as a result of natural soil formation and the effects of the human activities on different spatial and temporal scales. A source of information including the basic information such as soil, physiography, climate, vegetation, soil moisture content and land use is needed to develop the planning strategies to ensure sustainable use of agricultural lands as well as to carry out modelling of the environmental topics. The change to occur should be monitored in temporal and spatial terms in order to determine, map and plan the changes of soil characteristics of agricultural lands and to make these characteristics efficient and sustainable. Soil variability is frequently seen as a result of the data obtained from soil survey, soil and plant tests and product yield. However, assuming that their lands are homogeneous, most of the farmers implement homogeneous soil and plant management throughout all the land. As a result of these practices, excessive or insufficient practices are implemented in certain parts of the lands. This situation leads to an increase in the cost, decrease in net economic gain and unnecessary energy use as well as pollution in surface and ground waters (CASTRIGNANO ET AL., 2000).

Today, land use is more intensive and there is a need for a more precise evaluation of the distribution and behaviors of the soil as a result of that. Users want to know the amount and kinds of the variability of the soil mapping units in order to make their management decisions accurately within confidence limits. For example, it is essential to know the mean values of the mapping units within certain confidence limits. In addition, the information about the structure of the spatial variability can be forecasted through mapping units (NORDT ET AL., 1991). Depending on the soil variability, the use of geo-statistical methods has been more widespread rather than the traditional methods including different approaches in order to make these forecasts.

In parallel with the developments in computer technology, Geo-statistics is a technique that allows generalizing land properties using the relationships between the parameters by spending much less time, laboring and money (WARRICK ET AL., 1986; ZHANG ET AL., 1995). The concept of geo-statistics generally refers to the stochastic methods used for determining and forecasting the distinctive characteristics of the spatially referenced data.

Traditional statistics assumes that the variability in soil is random and it does not have a spatial correlation; the samples to be taken from each point of the land are independent. It uses the parameters such arithmetic mean, standard deviation and variation coefficient by calculating from the samples. The samples taken from the points that are close to each other are similar compared to the samples taken from distant points. In modelling studies (water flow modelling and non-point source (uncertain) pollution studies, etc.), intensive terrain data are needed for the verification of the models. In this type of use, the soil data, which are spatially variable in the soil in an area, layer or volume, should be grouped, processed, averaged or transformed to a scale that can reveal the similarities and differences in soil properties. The information about the spatial variable components of the soil is required to perform such groupings within the confidence limits to an extent (YOUNG ET AL., 1998; ROGOWSKI AND WOLF, 1994; BROWN AND HUDDLESTON, 1991).

DING ET AL. (2011) also used IDW, Ordinary Kriging and Spline methods in order to determine the distribution of soil moisture content. The distribution of soil moisture was determined by each method and Ordinary Kriging method was determined to be more appropriate for the spatial distribution upon comparing the lowest RMSE values of each method. ARSLAN ET AL. (2011) evaluated the spatial distribution of the exchangeable sodium percentage using different interpolation methods and they determined Simple Kriging as the best interpolation method.

The study has been conducted in a 6-decare area. The moisture distributions have been mapped by means of geostatistical techniques for different soil depths (30 cm, 60 cm, 90 cm and 120 cm) and soil moisture distributions have been determined for different layers. Geostatistical techniques can be used to obtain and map the soil properties.

MATERIALS AND METHODS

Vezirköprü District of Samsun Province is geographically situated between 41° 00' – 41° 19' North latitudes and 35° 48' – 35° 01' East longitudes. The long annual average highest temperatures have been respectively recorded as 29.8 °C (July) and 30.7 °C (August). The average precipitation has been recorded as 520 mm based on the average of long years; April and May are the rainy months (Tab. 1). The study area is about 6.5 decares. The soil moisture content in the field has been measured by neutron gauge device through the 40 accesstubes placed on certain points. The access tubes were placed at 120-cm depth and soil moisture content has been measured at every 30 cm (0-30, 30-60, 60-90 and 90-120 cm). The tubes have been placed in the study area as shown in Fig. 1. A system has been established based on the precipitation in this experimental area. In this system, the rainwater is stored under the soil. The area has not been irrigated during the experiment and melon and watermelon have been cultivated. An experimental area has been formed to examine the effects of precipitation on the changes of soil moisture content and crop yield. The experimental subjects have been placed in parallel to the levelling curves and an accesstube has been placed in each subject in order to determine soil moisture content. Thanks to this system, the effects of precipitation on the changes of soil moisture content have been examined at different depths (0-30 cm, 30-60 cm, 60-90 cm and 90-120 cm). As a result of the study, soil moisture maps have been drawn using the interpolation methods at the depths of 0-30 cm, 30-60 cm, 60-90 cm and 90-120 cm. The interpolation methods used in this study are stochastic (Ordinary Kriging and CoKriging) and deterministic (IDW and RBF) methods.

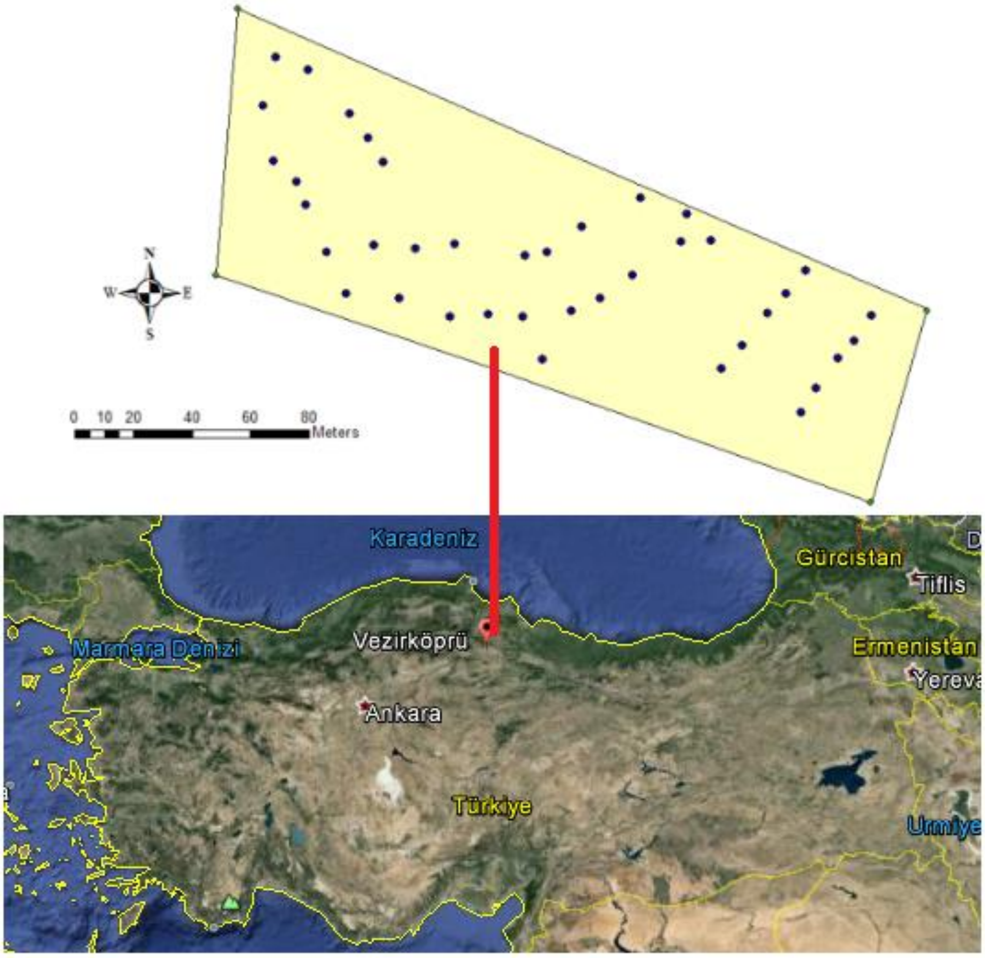

Fig. 1. – Study Area and Sample Points

Tab. 1. – Some of the long average annual climate parameters in Vezirköprü District

Meteorological Data	Months											
	Jan.	**Feb.**	**March**	**April**	**May**	**June**	**July**	**August**	**Sept.**	**Oct.**	**Nov.**	**Dec.**
Average Temp. (°C)	2.1	3.4	7.5	12.2	16	19.7	22.1	22.5	18.9	13.8	7.6	3.6
Highest Temp. Avr. (°C)	13.4	16.9	21	22.4	25.4	27	29.8	30.7	27.1	22.1	17.2	16
Lowest Temp. Avr. (°C)	-7.7	-8.8	-5.8	2.8	4.8	12.3	14.6	16	9.6	4.1	-0.8	-8.1
Relative Humidity Avr. (%)	75.5	70.5	64.1	63.5	63.8	61.7	56.6	55.5	57.7	65.1	74.1	76.3
Total Precipitation Avr. (mm)	36.5	31.7	37.3	56.7	72	65.8	30.5	22.1	23.2	52	46.2	46.2

Interpolation Methods

Today in Geographical Information Systems (GIS), the spatial interpolation methods are used to interpret the data collected from the points whose coordinates are known. There are two different interpolation methods in literature; these are stochastic and deterministic interpolation methods (ISAAKS AND SRIVASTAVA, 1989; ESRI, 2003). Stochastic methods process the data taking into consideration both mathematical and statistical functions to reveal the uncertainty and errors in forecasting. The stochastic methods also called as geo-statistical methods are basically known as Kriging. Kriging is an interpolation method that takes the spatial changes (variograms) into consideration. Kriging is used as an overall definition for the lowest generalized squared regression algorithms that form the basis of DenieKrige's studies (WEBSTER AND OLIVER, 2001). All Kriging predictors are derived from the following equation.

$$\hat{Z}(x_o) - \mu = \sum_{i=1}^{n} \lambda_i [Z(x_i) - \mu(x_o)] \quad (1)$$

μ given in the equation is a known constant average that is assumed as constant throughout the whole area and calculated as the average of the data (WACKERNAGEL, 2003). λ_i refers to Kriging weighting and it is dependent on the size of search window and the number of the samples determined to use in calculating n. $\mu(x_o)$ denotes the average of the samples in the search window. Kriging methods are subcategorized as Ordinary Kriging (OK), Block Kriging (BK), Disjunctive Kriging (DK), Universal Kriging (UK), Indicator Kriging (IK), and CoKriging (CoK). The traditional statistical methods are unable to provide information about the spatial relationship of the sampled points in a given area. CoKriging method is based on using detailed and incomplete secondary data and considering prominently the spatial cross-correlation between the primary and secondary variables (GOOVAERTS, 1999). It is used to estimate a variable by means of the other variable that are related to each other (eg: bulk density of the soil and soil moisture content). IDW, RBF and Kriging are the most commonly used interpolation methods (SUN ET AL., 2009). IDW is based on the distance from a point with an unknown value to the point with the known value by means of an unprejudiced weight matrix using the following equation

$$Z = \left[\sum_{i=1}^{n} \left(Z_i / d_i^m \right) / \sum_{i=1}^{n} \left(1 / d_i^m \right) \right] \quad (2)$$

where Z is the estimated value, Zi is the measured sample value at point i, di is the distance between Z and Zi, and m is the weighting power that defines the rate at which weights fall off with d_i, with a typical m value of 1–5 (KESHAVARZI AND SARMADIAN, 2012). RBF methods are radial-based methods and the features of the interpolated surfaces precisely pass through the data points and they have the lowest curves. The RBF method contains different functions for the forecasted surfaces. These functions are Completely Regularized Spline (CRS), Spline with Tension (ST), Thinplate Spline (TPS), Multiquadratik (M) and Inverse Multiquadratic (IM) (XIE ET AL., 2011).

The traditional statistical methods do not provide exactly accurate results since they make estimations based on the sampled points in a given area irrespective of the spatial dependence (MULLA AND MC BRATNEY, 2001). In the current study, four different interpolation methods have been used to determine the soil moisture distribution. The soil moisture data has been tested to see whether it has a normal distribution at different depths or not. In the current study, the spatial distribution of the soil moisture for 4 depths have been determined according to the most appropriate method using IDW and RBF of the deterministic methods and Ordinary (second order) Kriging and CoKriging of the geostatistic methods. In CoKriging method, the height has been used as the second variable to determine the effect of the second variable. The sequence followed in the study is shown in Fig. 2.

Evaluation of the Methods

The samples taken from a total of 40 points were measured using neutron gauge device. The soil moisture content values related to the depths of 0-30 cm, 30-60 cm, 60-90 cm and 90-120 cm have been used to evaluate the results of the most appropriate interpolation model. Different comparison techniques are used to determine the relationship between the measured and forecasted values. There are different comparison methods in literature to investigate the relationship between the measured and estimated values of soil moisture content and to choose the most appropriate model providing the results that are the closest to the measured values. Root mean square error (RMSE) and mean absolute error (MAE) are among the most widely used parameters. The model providing the lowest RMSE and MAE values has been chosen as the most appropriate model based on the equations 3 and 4.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{m} (yi - Oi)^2}{n}} \quad (3)$$

$$MAE = \left| \frac{\sum_{i=1}^{n} \left(y_i - O_i \right)}{n} \right| \quad (4)$$

where n is the number of the tested data, Oi is the forecasted precipitation data and yi is the measured precipitation data.

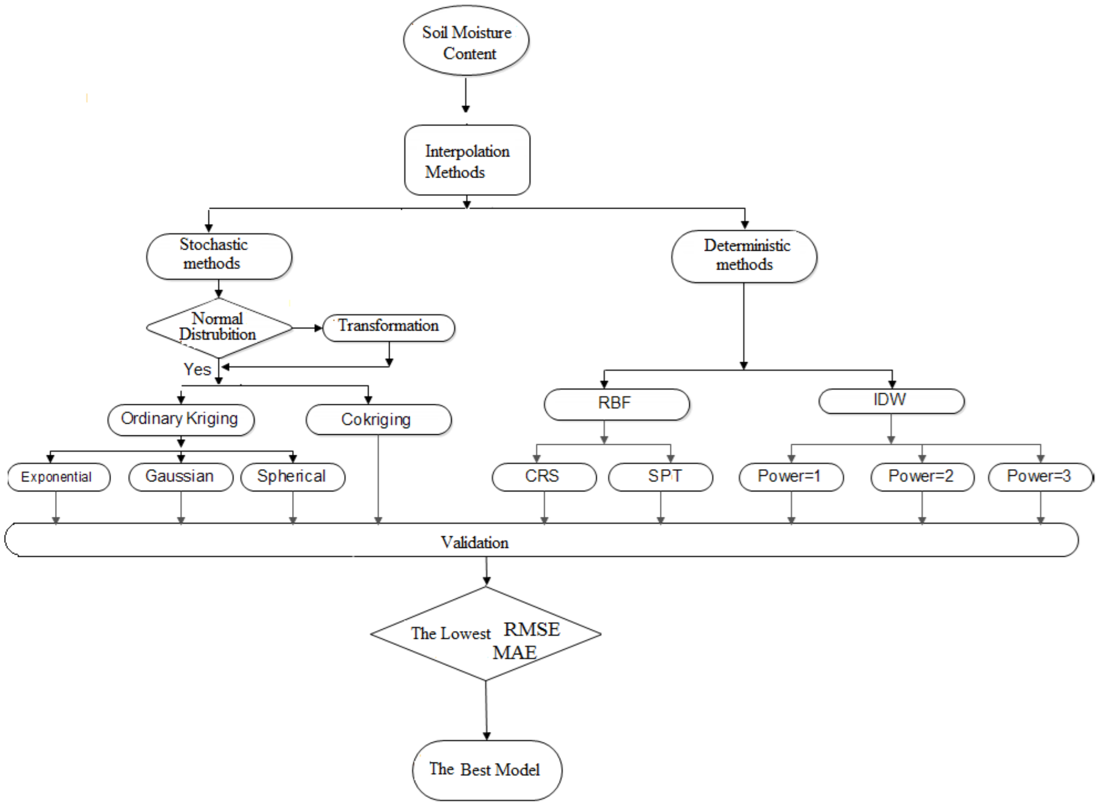

Fig. 2. – Flowchart of the interpolation methods used for the spatial distribution of soil moisture content

RESULTS AND DISCUSSION

In the current study, the soil moisture value has been determined at the depths of 0-30 cm, 30-60 cm, 60-90 cm and 90-120 cm. The soil depth maps have been drawn using deterministic and stochastic methods. The descriptive statistical values of the soil moisture content measured for each depth are shown in Fig. 2. The CV values of the soil moisture at the depths of 0-30 cm and 30-60 cm have been found lower than 15 (10.7 and 13.8) and they are categorized as "less variable". The soil moisture has a moderate variability (CV value between 15 and 35) at 60-90 cm and 90-120 cm depths. It has been found that the skewness values of the 4 depths do not have normal distribution; so, logarithmic distribution has been ensured performing transformation. Skewness value of the dataset is considered normal up to 0.5. When the skewness value is between 0.5 and 1, square root transformation is performed; however, when it is greater than 1, logarithmic transformation is performed. With these two transformation methods, the data can have a normal distribution (WEBSTER, 2001). Logarithmic transformation is performed to make data distribution closer to normal, to linearize the non-linear relationship and to decrease the heterogeneity and extreme/deviation values of the variances.

Comparison of the Interpolation Methods

In the current study, 8 different combinations of Ordinary Kriging (second-order), CoKriging, IDW, RBF interpolation methods have been used. Exponential, Gaussian and Spherical models have been tested in Ordinary Kriging method that is a stochastic method. As for the deterministic methods, three different weight powers have been tested for IDW and two different functions have been tested for RBF. 10 classifications have been used for soil moisture distribution. RMSE and MAE values have been evaluated for each method. The most accurate mapping method has been chosen between the deterministic and stochastic methods based on the lowest RMSE and MAE values. The lowest RMSE and MAE values were respectively obtained from Gaussian model (Ordinary Kriging), power 1 variogram (IDW), RBF method and power 1 variogram (IDW) for the depths of 0-30 cm, 30-60 cm, 60-90 cm and 90-120 cm. It has been determined that the most accurate mapping can be done based on these methods. The results of these methods are listed in Tab. 3. The spatial distribution maps of the soil moisture at different depths are shown in Fig. 3 based on the best methods.

Tab. 2. – Basic statistical results for all data

Depth	Highest	Lowest	Mean	Standard Dev.	CV	Kurtosis	Skewness
0-30	122.82	64.37	109.70	11.81	10.77	5.42	-2.05
30-60	123.01	52.18	104.61	14.48	13.84	5.36	-2.20
60-90	136.83	45.20	108.15	19.49	18.02	3.51	-1.69
90-120	141.81	28.82	113.11	21.07	18.63	5.67	-1.84

Tab. 3. – Results of the soil moisture at four depths based on different interpolation methods

	OK-G		IDW-1		RBF-ST		CoK-Spherical	
	RMSE	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**
0-30	11.296	7.821	11.506	7.923	11.938	8.055	11.35	7.919
30-60	13.813	8.709	13.682	8.444	13.98	8.742	13.823	8.631
60-90	21.695	13.97	17.846	11.653	17.767	11.473	18.557	12.728
90-120	21.06	14.6	20.24	14.18	20.46	13.92	21.223	14.468

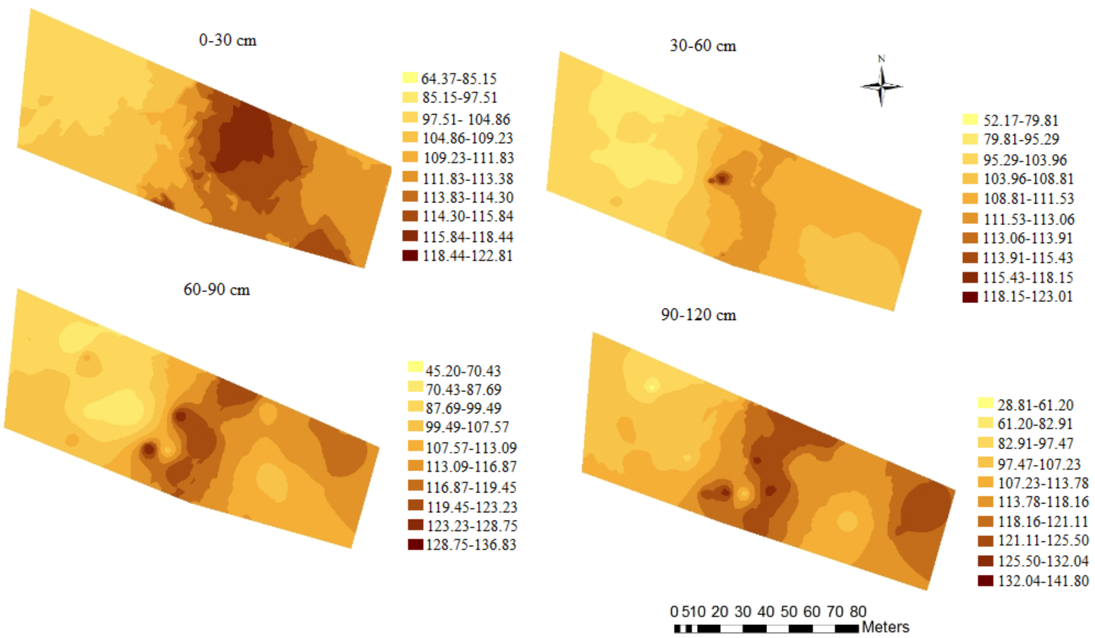

Fig. 3. – Spatial distribution maps of the soil moisture content at different depths

CONCLUSIONS

When compared, the lowest RMSE and MAE values have been respectively obtained from Gaussian model (Ordinary Kriging), power 1 variogram (IDW), RBF method and power 1 variogram (IDW) for the depths of 0-30 cm, 30-60 cm, 60-90 cm and 90-120 cm. The most accurate mapping has been done based on these methods. Cross-validation method has been performed to determine the best method. It has been seen that other methods provided similar results. In the current study, different results have been reached for each depth. Therefore, in the studies on soil properties and climate data, determining the most appropriate interpolation method and screening the distribution according to this method are considered to provide more accurate results in order to determine the spatial distribution properties of the soil moisture content from which the point data are obtained. Planning the agricultural production and hence the irrigation time is very difficult in the areas where the soil moisture content are not determined by depth. Therefore, it is important to produce and use the spatial distribution maps properly by means of these techniques. Especially, different interpolation methods are used to estimate the spatial distribution of the soil moisture content in the complex terrains in the catchment areas. XUELING ET AL. (2013) examined the appropriateness of different interpolation methods (Ordinary Kriging, IDW, linear regression, regression Kriging) in forecasting the soil moisture throughout the soil profile (1 m) based on 153 samples taken from the Loess Plateau of China. As a result, they found that there are significant differences between each method in terms of homogenous lands. It has been found that Environmental impact factors have discontinuous spatial dependence in Ordinary Kriging and IDW models and the soil moisture in the homogeneous terrains of the small basins has poor spatial predictive value. The spatial distribution of soil moisture cannot be sufficiently characterized by means of the direct measurements in the practical implementations; therefore, it is necessary to create new observation points and determine the best interpolation method for them. Most of the hydrological practices can affect the soil moisture and existence of different variations for soil moisture samples can lead to complicated results (PERRY AND NIEMANN, 2008). Before determining the soil parameters from which the point data are obtained and mapping the spatial distribution of the values related to soil properties, it is considered to have better results by firstly determining the most appropriate interpolation method and then interpreting the results based on the spatial distributions using this method. With this study, the spatial distributions of the soil moisture content have been determined at different soil depths and it has been found that the changes in soil moisture content can be spatially determined. As a result, this study is expected to be a guide for the further studies.

REFERENCES

1. ARSLAN, H., CEMEK, B., GÜLER, M., YILDIRIM, D.: Değişebilir Sodyum Yüzdesinin (ESP) Konumsal Dağılımının Farklı Enterpolasyon Yöntemleriile Değerlendirilmesi. II. Ulusal Sulamave Tarımsal Yapılar Sempozyumu. 24-25. Mayıs, 505-512., 2012: İzmir.
2. BROWN, R.B., HUDDLESTON, J.H.: Presentation of Statistics Data on Map Units to The User. In Spatial Variabilities of Soil And Landform. SSSA Special Pub.No:28, 1991: pp 127-147.
3. DING, Y., WANG, Y., MIAO, Q.: Research on the spatial interpolation methods of soil moisture based on GIS, International Conferences on Information Science and Technology, Jiansu, China, 2011: March 26-28.
4. ESRI: Using Arcview GIS. Environmental System Research Institute.Inc.:Redlans, Calif., 2003.
5. GOOVAERTS, P.: Using elevation to aid the geostatistical mapping of rainfall erosivity. Catena, 34, 1999: 227-242.
6. ISAAKS, E.H., SRIVASTAVA, R.M.: An Introduction toApplied Geostatistics. Oxford Uni. Press, Inc. New York, 1989, 561 pp.
7. KESHAVARZI, A., SARMADIANI, A.: Mapping of spatial distribution of soil salinity and alkalinity in a semi-aridregion. Annals of Warsaw University of Life Sciences Land Reclamation, 44(1), 2012: 3–14.
8. MULLA, D.J., MCBRATNEY, A.B.: Soil Spatial Variability, Handbook of soil science, CRS Pres., 2001.
9. NORDT, L.C., JACOB, J.S., WILDING, L.P.: Quantifying Map Unit Composition for Quality Control in Soil Survery.p., 1991, 183-197 In: Spatial Variabilities of Soil and Landforms. Editors: Mausbach, M.J. And Wilding, L.P. Soil Sci. Soc. of America, Inc, Madison Wisconsin, USA.
10. PERRY, M.A., NIEMANN, J.D.: Generation of soil moisture patterns at the catchment scale by EOF interpolation. Hydrol. Earth Syst. Sci., 12, 2008: 39-53.
11. ROGOWSKI, A. S., WOLF, K.J.: Incoporation Variability into Soil Map Unit Delineation. Soil Sci. Soc. Am. J., 58, 1994:163-174.
12. SUN, Y., KANG, S., LI, F: Comparison of interpolation methods for depth to ground water and its temporal and spatial variations in the Minqinoas is of North West China. Environmental Modelling& Software, 24, 2009: 1163–1170. Tsanis,
13. YOUNG, F.J., HAMMER, R.D., WILLIAMS, F.: Evaluating Central Tendency and Variance of Soil Properties within Map Units. Soil Sci. Soc. Am.J., 62, 1998: 1640-1646.
14. WARRICK, A.W., MYERS, D.E., NIELSEN, D.R.: Geostatistical Methods Applied to Soil Science, In: Methods of Soil Analysis, Part I. Physical and Mineralojical Methods, ASA and SSSA Agronomy Monografh no:9(2nded)., 1986: pp. 53-82, Madison.
15. WACKERNAGEL, H.,: Multivariate Geostatistics: An Introduction with Applications. Springer, Berlin, 2003: 387 pp.
16. WEBSTER, R., OLIVER, M.: Geostatistics for environmental Science. John Wiley and Sons, Ltd, Ch,chester, 2001: 271 pp.
17. WEBSTER, R.: Statistics To Support Soil Research And Their Presentation. European journal of Soil Science, 52, June 2001: 331-340.
18. XIE,Y., CHEN, T., LEI, M., YANG, J.: Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: accuracy and uncertainty analysis. Chemosphere, 82, 2011: 468–476.
19. ZHANG, R., RAHMAN, S., VANCE, G.F., MUNN, L.C.: Geostatistical Analysis of Trace Elements In Soil and Plants, Soil Science, 159(6), 1995: pp. 383-390.

Corresponding author:

Demet Yıldırım, Black Sea Agricultural Research Institute, Irrigation and Land Reclamation Unit, Samsun, Turkey