Geosciences and Engineering, Vol. 10, No. 15 (2022), pp. 37–49.
https://doi.org/10.33030/geosciences.2022.15.037

INVESTIGATION OF THE POSSIBILITY OF EXTENDING SOIL MOISTURE SENSOR DATA

TAMÁS DEÁK, RÉKA DIÁNA BERTÓTI, ENDRE DOBOS

Institute of Geography and Geoinformatics, Faculty of Earth Science and Engineering, University of Miskolc

Abstract: Smart irrigation requires soil moisture (SM) data in order to determine the water needs of each point within the agricultural field. The aim of the study is to develop methodologies for testing different approaches and to present the provisional results. A complex methodology was applied to process and analyze the sensor data. First spatial variability was assessed using regression and geostatistical interpolation. A vertical distribution estimation was performed by extending the SM values of each sensor for all their depth levels. The latest results of this method show that the statistical correlations between the different depths can be used to characterize the soil types and their corresponding vertical SM distribution.

Keywords: *Soil science, soil moisture, statistics, Sentek, soil moisture sensor*

1. INTRODUCTION

In Hungary, drought periods are getting longer due to the deteriorating soil conditions caused by the agricultural sector since the first industrial revolution and as a result of the decreasing amount of rainfall [1]. More and more farmers are considering building their own irrigation systems in order to supply the water necessary for keeping their crops alive. Although Hungary has a great supply of water, the price over the years will increase and the question of "how much water do we need for the crop" will become more frequent and difficult to answer. Precision agriculture might have an answer. It uses spatially specific data to make decisions like where, when, why and how we should water the crops [2].

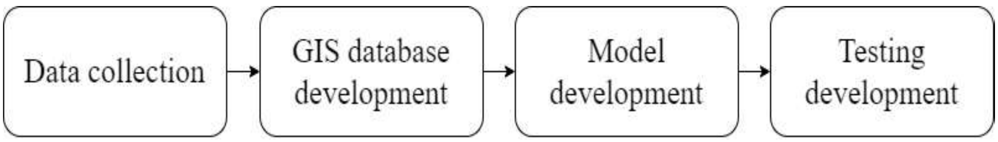

Figure 1
Soil moisture extension process

Our institute is involved in a new project for creating an automated irrigation system based on spatial data. To fulfill this, we need soil moisture data which can tell us the volume of the available water within the soil and also answer how much water the plants can take up with all their important nutrients. Many researchers have already used different methods regarding soil moisture mapping: Jackson used microwave radiometry [3], Paloscia used Sentinel-1 images [4], Wagner used an ERS scatterometer in western Africa [5] and also Kibirige and Dobos by integrating environmental data [6].

The soils of Hungary are very diverse [7]. Due to this, there can be a significant variability in soil moisture levels and the water storage capacity. This paper will discuss two types of methodologies to evaluate geostatistical interpolation.

2. GENERAL DESCRIPTION OF THE METHODOLOGY

2.1. Description of the physical geographical aspects of the study area

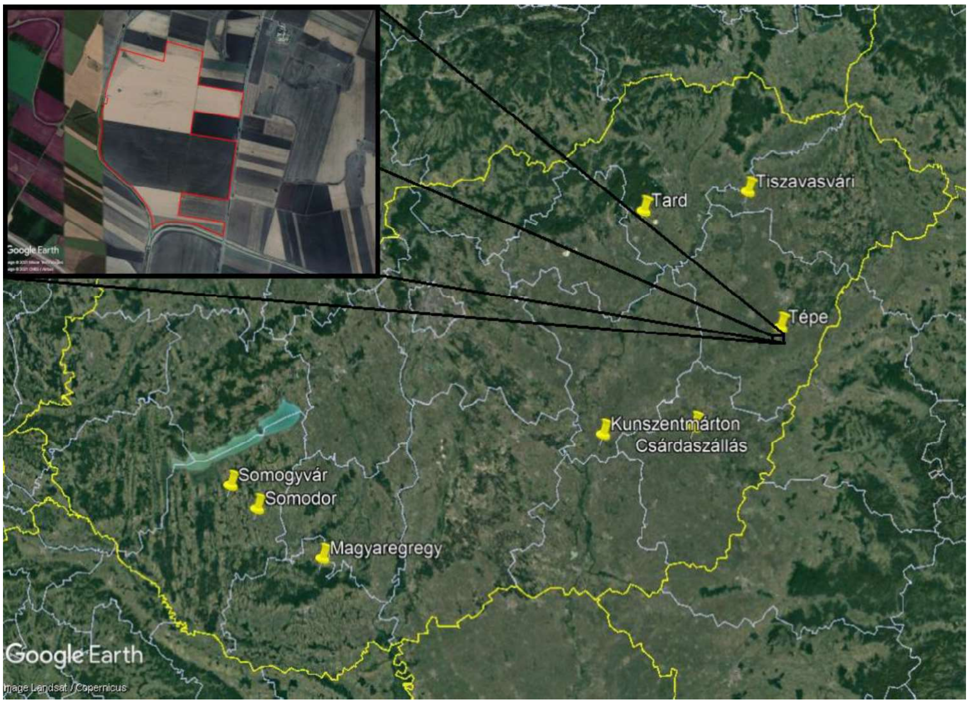

Figure 2

Locations of the Sentek EnviroSCAN sensors placed in Hungary;
inset: the pilot area of the study

For this study, an agricultural field at Tépe has been selected as a pilot area, which is one of several locations where the SM data collecting sensors are placed. Tépe is a settlement in Hajdú-Bihar County, 26 km south of Debrecen *(Figure 2)*. It is located in the Berettyó–Kálló micro-region, which is part of the Berettyó–Körös region. The elevation difference within the field is only a few meters [8] so at first glance, there should not be any significant differences as for geomorphological factors. However, due to the heavy agricultural use and the relative size of the pilot area, surface morphology is a significant factor, and finding the right parameters that will have the most correlation with the observed data is crucial in order to predict new SM values of the area.

Surface forms are mainly of a riverine origin, while the primary surface forming force is fluvial erosion, which brings all riverbed forms, reinforced by anthropomorphic impacts of agricultural use and surrounding channel networks. The geology of the micro-region is based on a 2.5 km deep, metamorphic bedrock formation. Three-quarters of the area is a Holocene floodplain with marsh mud and clay [8] where an increasing refinement of sand material can be observed in the East-West direction.

As for climate, the area has moderately hot and dry summers. During the vegetation season the average temperature is 17 °C, while the annual average temperature is around 10 °C. Annual average rainfall is 550 mm, with 320 mm of this in the vegetation season. The main wind directions are eastern and southern winds (just like throughout the whole Eastern Hungarian region) [8]. The flat surface allows wind to go unobstructed, thus the area is vulnerable to wind erosion. When checking the national averages for the amount of rainfall, the area does not belong to the driest nor to the wettest. Due to this, and as mentioned previously, canal and river networks are very common, e.g. the Eastern main canal, Berettyó and other smaller channels are surrounding the area.

Groundwater levels move between 2–4 meters. Stratified water is rare so there are a number of artesian wells in the area that can go down to 200 meters. The surface is formed mainly by river water, which results in various soil conditions at a micro-relief level. There are three typical main soil types to be mentioned here: chernozems on loess, gleysols and salt effected soils on lower lying, more clayey surfaces with groundwater impact [8].

2.2. Sentek data collection

The institute has been collecting SM data over three years now, nationwide. For this, we are using Sentek EnviroSCAN sensors [9]. These units are stand-alone, real time, soil moisture monitoring devices. Originally, they were used for general research purposes, but over the time they became critical units of irrigation systems for monitoring SM [10]. The equipment is not a single module but can be customized with many setups, at different depth levels. All of our units are calibrated to collect SM data at 10, 20, 30, 40, 60 and at 100 cm depths. The sensors use the dielectric constant values of the soil, water, and air to calculate soil moisture *(Figure 3)* [11]. Each sensor is placed in a different soil type that has its own electric capacity values and physical attributes; therefore, it is mandatory to calibrate each of the sensors properly before starting the data collection.

Figure 3
Overview of the Sentek EnviroSCAN sensor unit
(Source: sentekusa.com)

There are a total of 21 Sentek EnviroSCAN sensors scattered throughout the country, placed on agricultural areas of 8 settlements, each with different SM values, geomorphological features, soil types and cultivation methods currently used on the fields. For this present case, we have chosen Tépe as the study area of our analysis *(Figure 5)*. It is worth mentioning that each sensor position has its own corresponding soil, containing a number of new variables such as physical variability, chemical composition, etc. that can be useful for finding correlations between the sensor and soil data.

2.3. Development of the GIS database

Several environmental parameters were gathered to explain the soil moisture variability. This starts with the construction of the digital elevation model (DEM), which is created from a set of points, captured by RTK corrected GPS equipped machinery, containing many spatial variables like elevation, fuel consumption, etc. For this present case, elevation values were used to create an interpolated raster image with 5 m spatial resolution *(Figure 4)*. All of the geomorphological parameters – slope, aspect, flow accumulation and relative relief – are derived from this DEM using ArcGIS. Other dynamic raster datasets are also collected for the area, such as

Sentinel-1B backscattering images, preprocessed with VV polarization that has also been converted from linear to dB values using ESA SNAP [12].

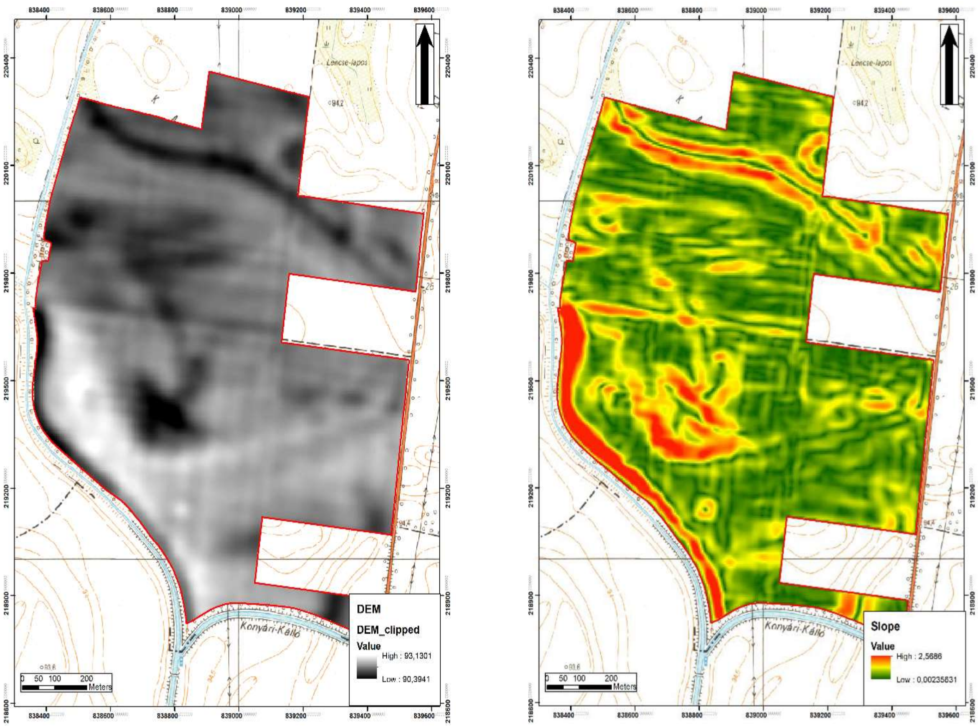

Figure 4

Digital elevation model (meters above sea level, left) and slope (degrees, right) geomorphologic parameters of the study area

Normalized difference vegetation indexes (NDVI) are also used from the Sentinel-2B bands. The satellite images were collected from the Copernicus Open Access Hub, made available by the European Space Agency.

The Hungarian vegetation season – between March and September – has been chosen as the timeline of the study, and the satellite images were selected to be as close as possible to the examined period of SM values. Based on all this, May 3rd, June 17th, July 4th and August 30th were picked as the sample dates.

2.4. Methodology to describe the spatial soil moisture distribution

The aim of the collected observed SM, static and dynamic spatial data was to estimate SM values for the entire field. The research methodology of Daniel Kibirige and Endre Dobos dealt with processing and analyzing low-cost, soil moisture sensor data in order to reach the same goal [13]. In their study, they used linear regression, the Ordinary Least Square method and cokriging, based on the observed SM data.

The model itself needs at least one or more independent variables, along with a dependent – observed – one.

The residual values were created by using the Ordinary Least Square (OLS) tool inside ArcGIS's Geostatistical Analysis toolbox. The "Square Root" and "Root Mean Square Error" were also calculated to measure the difference between the predicted model and observed data.

Ordinary kriging is a geostatistical interpolation method based on spatially dependent variance, which gives unbiased estimates of variable values at target location in space, using the known sampling values at surrounding locations [14]. One of the biggest advantages of cokriging compared to ordinary kriging is that it can use more than one variable. Like in the OLS model, slope was the independent variable and the observed SM data was the dependent variable for the cokriging model.

2.5. Methodology to describe the vertical soil moisture distribution

After reviewing how well the spatial distribution worked, the conclusion was that the available observed sensor data – in terms of the number of spatial samples – were not enough to do any kind of advanced spatial analysis, which has been confirmed by both the linear regression and cokriging results. Due to the high cost of the equipment, they had to be placed in low density, mainly to represent the different soil units within the area. This makes spatial extension difficult, because the points do not represent the same continuity. To overcome this limitation, we decided to characterize the vertical distribution of soil moisture within the upper meter, in order to make the different locations correlatable and interpolatable. At this point, spatial expansion was not possible, therefore a new methodology was needed, and the main goal had to be renewed as well. Examining the sensor data again proved that data usability lies in the vertical analysis, where each SM sensor needs to be analyzed with all of its depth levels.

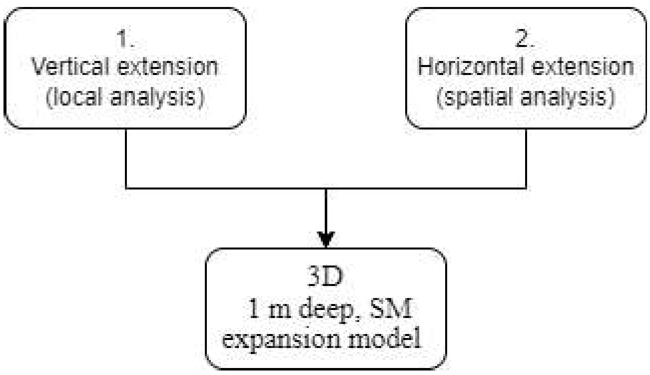

Figure 5
Flowchart of the vertical distribution process

A new goal has been set: creating a 3D, 1 m deep SM model *(Figure 5)*. Two steps need to be made in order to reach this: First, every sensor needs to be analyzed individually to find out the correlation between the different depths, and identify the soil type related relationships (1 in *Figure 4*). The analytical aspect needs to be changed into a vertical direction by doing descriptive statistics first. A Spline function was applied to to estimate the SM values for the entire profile [15] (2 in Figure).

3. THE RESULTS

3.1. Statistical correlation study to select the appropriate variable

The real problem of adapting the methodology of Kibirige and Dobos [13]for this case was the small sample size – 5 points, which is a huge disadvantage throughout the whole study and allows one independent variable for the model. A Pearson correlation analysis was done [16] to find out which static or dynamic variable would correlate best with the observed SM data. The results show that the slope variable had the best correlation *(Table 1)*.

Table 1

Pearson correlation matrix with the observed SM values

	aspect	**slope**	**rela_ref**	**DEM**	**flow_acc**	**SigmaVVdB**	**NDVI**
10 cm	0.310	0.586	–0.605	–0.379	–0.049	–0.214	–0.451
20 cm	–0.174	0.599	–0.58	–0.469	–0.469	–0.121	–0.364
30 cm	–0.252	0.398	–0.377	–0.286	–0.503	–0.164	–0.366
40 cm	0.247	0.656	–0.651	–0.352	–0.117	–0.178	–0.397
60 cm	0.259	0.917	–0.924	–0.772	–0.073	–0.078	–0.292
100 cm	0.236	0.96	–0.947	–0.644	–0.188	0.052	–0.217
MEAN	0.104	0.686	–0.681	–0.484	–0.233	–0.117	–0.348

3.2. Spatial soil moisture distribution

The error statistics for the OLS model are listed in *Figure 6*. Each sensor showed the largest difference at depths of 10–40 cm and 60+ cm, while it became zero between 40–60 cm. This applies to the upper (10–40 cm) and lower (60+ cm) water supply layers, which are completely separated by an impermeable clay layer, originating from agricultural cultivation as a result of mixing the upper soil depths. This assumption was validated with soil excavation, where the characteristic soil layers were referring to the residual values.

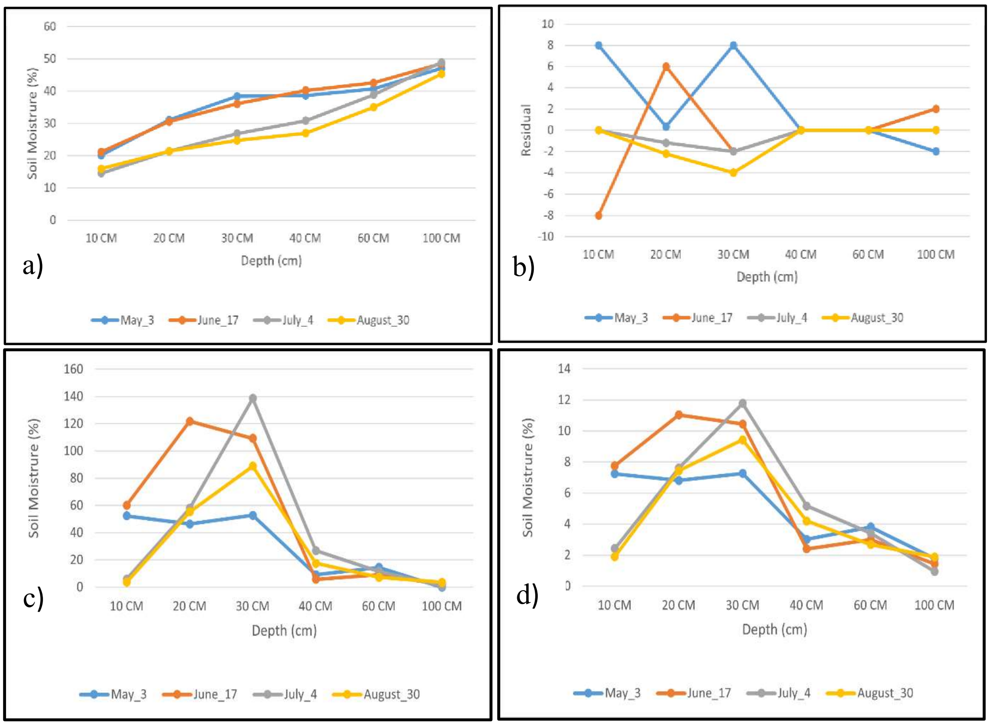

Figure 6

Ordinary Least Square model results: Observed SM values (a), residual for OLS predicted SM values (b), root square (c), and root mean square error (d)

Cokriging showed weak interpolations due to the small number of sample points. Each examined dataset with all depths showed mixed results, but in general, the 10 cm values had the best range of values, and it also shows the “least worse” results *(Figure 7)*. Nevertheless, there are still some correlations to be seen from these images. They show increasing SM values to the North-Northwestern direction, which indicates a mild inclining feature for the whole landscape. This has been validated by checking the topology maps showing the inclination, based on the contour line values *(Figure 7)*.

3.3. Vertical soil moisture distribution

Descriptive statistical analysis for all sensor locations at Tépe have been completed to show a general overview of the data *(Table 2)*. Correlation matrices were created to characterize the vertical moisture movements between the different depths *(Table 3)* and boxplots to identify the outlier values *(Figure 8)*.

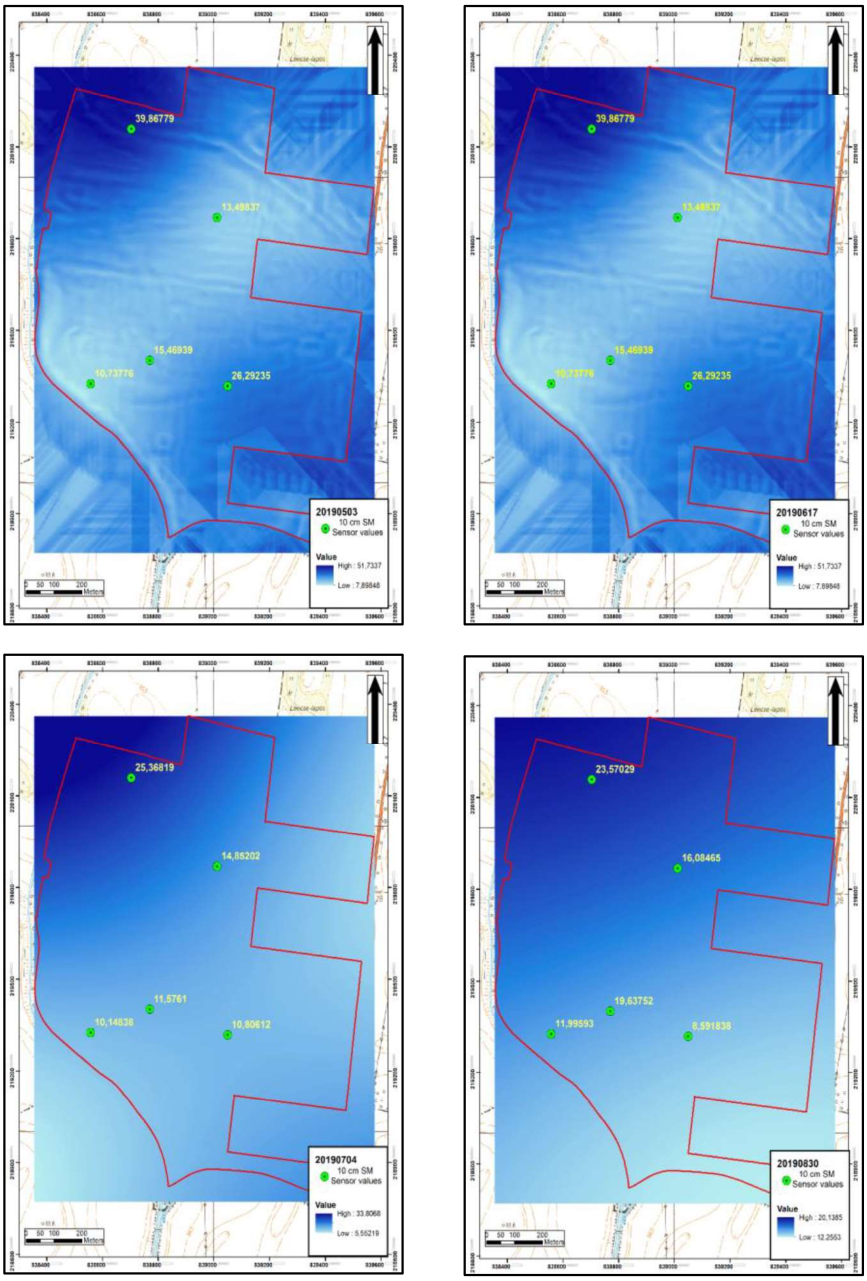

Figure 7
Predicted SM values with cokriging at 10 cm depth: 2019 May 3 (a), 2019 June 17 (b), July 4 (c), August 30 (d)

Table 2

Descriptive statistics of one of the SM sensors (ID 6, summer season)

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
10 cm	2208	32.064	11.873	43.937	20.586	6.370	40.582
20 cm	2207	28.586	5.753	34.339	12.764	8.001	64.022
30 cm	2208	39.309	4.997	44.306	15.777	12.537	157.167
40 cm	2206	26.845	17.340	44.185	31.112	10.346	107.032
60 cm	2208	13.015	27.082	40.097	33.453	4.577	20.950
100 cm	2208	6.101	38.382	44.483	41.956	2.169	4.706

Table 3

Correlation matrix of one of the SM sensors (ID 6, summer season)

	10 cm	20 cm	30 cm	40 cm	60 cm	100 cm
10 cm	1	0.413	0.215	0.464	0.314	0.031
20 cm	0.413	1	0.948	0.765	0.764	0.476
30 cm	0.215	0.948	1	0.811	0.823	0.525
40 cm	0.464	0.765	0.811	1	0.940	0.606
60 cm	0.314	0.764	0.823	0.940	1	0.786
100 cm	0.031	0.476	0.525	0.606	0.786	1

The correlation table indicates that the uppermost 10 cm do not correlate well with the deeper horizons. this is the layer most exposed to the surface weather, and it shows large fluctuations which do not necessarily appear in the deeper horizons. The same is true for the 1 m depth, whch is quite different from all overlying layers. It also shows the expected correlation trend, namely the decreasing correlation toward the surface. The strongest correlations were found between the 20–30 cm and the 40–60 cm depths. It is also logical, because the 20–30 cm depths belong to the middle and deep part of the plough layer, insulated from the atmosphere by the upper 10 cm. 40 cm is the plough pan, which is compacted and has a low infiltration rate, causing a stagnic water body to form above it. This layer also separates the plough layer and the deeper horizons, namely the 40–60 cm depths, which represents a slightly separated water body. These results are common in all soil profiles. *Figure 8* explains these features from another aspect. The upper 30 cm always has a much higher range, with often hectic data distribution, meaning that this plough layer is more exposed to the surface and top the vegetative use of SM. The depths of 40, 60 and 100 cm are more stable with far fewer hectic features and outliers.

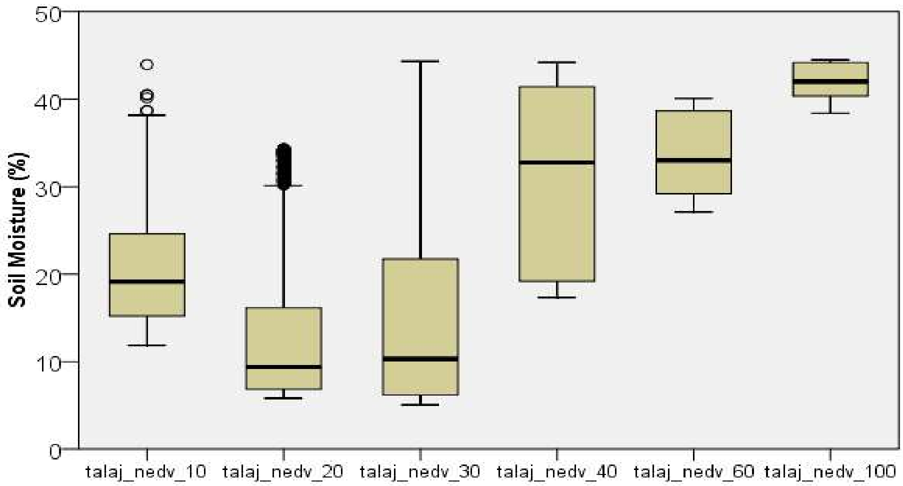

Figure 8

Boxplot of one of the SM sensors (ID 6, summer season)

4. Summary and Conclusion

Five soil sensors with six different depths of SM data were available for processing when examining the sample field at Tépe. The regression analysis had higher than acceptable error values, shown by the root square and root mean square error, which made this methodology irrelevant *(Figure 6)*. The geostatistical interpolation also gave weak results due to the limited amount of points inside the field (only five), resulting in maps with lower spatial resolution *(Figure 7)*.These are only preliminary results; a complete characterization of the methodology requires more sampling points and environmental covariances to better describe the spatial variability of the soil moisture. However, the data have great potential as separate, individual units, where the distribution and vertical extension could be done between the depths. We believe the latest methodology will bring more results in the future and will also answer the question of finding specific correlation levels between every soil type regarding variables and the observed SM depth values. This should allow associated soil types to be placed into water management type groups for the upcoming irrigation system.

References

[1] Alsafadi, K., Mohammed, S. A., Ayugi, B. et al. (2020). Spatial–Temporal Evolution of Drought Characteristics Over Hungary Between 1961 and 2010. *Pure and Applied Geophysics*, Vol. 177, pp. 3961–3978. https://doi.org/10.1007/s00024-020-02449-5

[2] Shannon, D. K., Clay, D. E., Kitchen, N. R. (2018). *Precision Agriculture Basics.* American Society of Agronomy Crop Science, Society of America Soil Science, Society of America, https://doi.org/10.2134/precisionagbasics.

[3] Jackson, T. J., Le Vine, D. M., Hsu, A. Y., Oldak, A., Starks, P. J., Swift, C. T., Isham, J. D., Haken, M. (1999). *Soil moisture mapping at regional scales using microwave radiometry: the Southern Great Plains Hydrology Experiment*. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37 (5), pp. 2136–2151, https://doi.org/10.1109/36.789610.

[4] Paloscia, S., Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L., Reppucci, A. (2013). Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation. *Remote Sensing of Environment*, Vol. 134, pp. 234–248. https://doi.org/10.1016/j.rse.2013.02.027

[5] Wagner, W., Scipal, K. (2000). *Large-scale soil moisture mapping in western Africa using the ERS scatterometer*. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 38 (4), pp. 1777–1782, https://doi.org/10.1109/36.851761.

[6] Kibirge, D., Dobos, E. (2021). *Estimation of surface soil moisture by integrating environmental data and remote-sensing satellites*. *Multidiszciplináris Tudományok*, Vol. 11, pp. 22–37, https://doi.org/10.35925/j.multi.2021.1.3.

[7] Stefanovits, P., Filep, Gy., Füleky, Gy. (1996). *Talajtan*. 4th edition. Budapest, Mezőgazda Kiadó.

[8] Dövényi, Z., Marosi, S., Somogyi, S. (2010). *Magyarország kistájainak katasztere*. 2nd edition. Budapest, MTA Földrajztudományi Kutatóintézet.

[9] *Sentek EnviroSCAN Soil Data Probe* [Online] Available at: https://sentek technologies.com/product-range/soil-data-probes/enviroscan [Accessed 7 Dec 2021].

[10] Al-Ghobari, H., Marazky, M., Aboukarima, A., Minyawi, M. (2016). Calibration of Soil Water Content Data from EnviroSCAN System Using Artificial Neural Network. *American Journal of Experimental Agriculture*, Vol. 12 (5), pp. 1–11, https://doi.org/10.9734/ajea/2016/26237.

[11] Al-Ghobari, H. M., Said, M., El, A. (2013): Field evaluation of EnviroSCAN performance for monitoring soil water content compared with other soil moisture sensors under arid conditions. *Wulfenia Journal*, Vol. 20 (4), pp. 54–70.

[12] Filipponi F. (2019): *Sentinel-1 GRD Preprocessing Workflow*. *Proceedings*. 18 (1), Art. 11, https://doi.org/10.3390/ECRS-3-06201.

[13] Kibirge, D., Dobos, E. (2021). Soil Moisture Estimation Using Citizen Observatory Data, Microwave Satellite Imagery, and Environmental Covariates. *Water*, Vol. 13 (6), p. 837, https://doi.org/10.3390/w13060837.

[14] Adhikary, S. K., Muttil, N., Yilmaz, A. G. (2017). Cokriging for Enhanced Spatial Interpolation of Rainfall in Two Australian Catchments. *Hydrological Processes*, Vol. 31 (12), pp. 2143–2161, https://doi.org/10.1002/hyp.11163.

[15] Voltz, M., Webster, R. (1990). A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. *Journal of Soil Science*, Vol. 41 pp. 473–490. https://doi.org/10.1111/j.1365-2389.1990.tb00080.x

[16] Benesty J., Chen J., Huang Y., Cohen I. (2009). Pearson Correlation Coefficient. In: *Noise Reduction in Speech Processing*. Springer Topics in Signal Processing, Vol 2, Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-00296-0_5