International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-4, November 2019

Semi-Arid Region Soil Moisture Prediction using Multivariate Regression

Roohul Abad Khan, Javed Mallick, Rachida El Morabet

Abstract*: The Water Scarcity is a prominent feature in Arid and Semi-Arid region. Soil moisture content is significant factor in deciding vegetation growth and also affects the performance of any water harvesting system in place. This paper evaluates the interrelationship of Soil properties with Soil Moisture content. The study covers about 13 soil Samples from Single Watershed. The soil properties covered in the study are Conductivity, pH, Bulk Density, Dry Density, Specific gravity, organic content, void ratio, and Moisture Content. Multiple linear regression analysis was done to determine significance of each soil properties for soil moisture content as individual and as whole. Modelling was done based on soil characteristics to predict Soil Moisture. Principal Component Analysis was performed to identify most significant soil properties responsible for variation of prediction of Soil Moisture content. The Correlation between location topography and Moisture Content was obtained through Cluster Analysis.

***Keywords* : Moisture Content, Regression Analysis, Principal Component Analysis, Cluster Analysis**

I. INTRODUCTION

Soil is a natural element formed over long period of time owing to erosion of rocks and mountains. The major three elements of soil are clay, silt and sand. The percentage constitution of these elements determines soil group. Based on soil group, appropriateness of soil for relevant use or work can be inferred. While little work has been on soil properties of semi-arid region especially seasonal streams. Semi-arid regions have unique land surface characteristics where the soil surface layer is dominated by soil crusts, which could significantly reduce infiltration capacities and increase surface runoff. Becker et. al. (2018) analysed impact of surface and subsurface properties of soil in terms of hydraulic conductivity. Multiple machine learning technique was employed to generate digital map based on soil properties. (Zeraatpisheh et al. 2018). Bracken et. al. (2019) employed simulated EnMAP data to identify soil erosion in Semiarid Mediterranean environment. Aguilera et. al. (2016) predicted soil moisture in order to manage Semiarid wetlands in dry periods. While Schoener et. al. (2019) identified soil moisture impact on semiarid catchment runoff. Eisele et al. (2015) used thermal infrared to quantify and detect soil properties in semiarid area. Alidoust et. al. (2018) estimated soil carbon sequestration potential in semi-arid region under different land use. However, literature on taking into account Initial moisture content of soil prior to any water harvesting activity is still lacking. The study intends to predict soil moisture content using multiple regression analysis for Semi-arid region Bisha Saudi Arabia based on various other parameters of soil. The prediction of initial soil moisture content will determine initial water loss for any future harvesting system undertaken owing to loss from absorption by soil.

Manuscript published on November 30, 2019.
* Correspondence Author
Roohul Abad Khan*, Faculty of Science and Engineering, Himalayan University, Arunachal Pradesh, India Email: environmental629@gmail.com
Javed Mallick, Department of Civil Engineering. King Khalid University, Abha, Saudi Arabia Email: jmallick@kku.edu.sa
Rachida El Morabet, Department of Geography, University Hassan II Casablanca Morocco.

II. DATA AND METHOD

A. Study Area

The study area Bisha is located at 20°0'0" N 42° 36'0" E, in Asir Province of Saudi Arabia. The 240 villages and 58 bigger settlements are located on both sides of Bisha valley which is also the largest valley/stream of Saudi Arabia. It is located 610 meters above Mean Sea Level. The average maximum temperature is 32°C while minimum average temperature is 17.1°C. the annual precipitation is 130 mm.

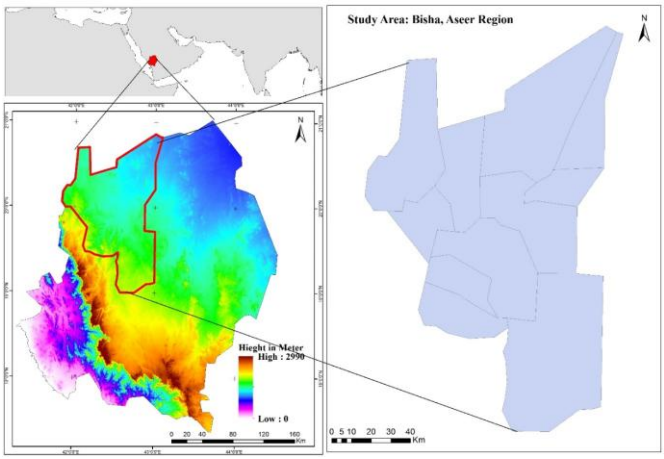

Figure 1 Study Area Watershed

B. Soil Sampling and Analysis

The soil samples were collected from stream beds in proximity to groundwater wells. Thirteen soil samples were obtained from a depth of 0-30 cm. the locations of the soil samples are presented in Table 1. Soil samples were collected in polythene bags, labelled and transported to Soil Mechanics and Foundation Engineering lab for sample preparation for testing and Analysis.

The parameters analysed covered soil pH, Conductivity, Moisture content, Organic Content, Bulk Density, Dry Density, Porosity, Void Ratio and Specific gravity.Soil Moisture content was obtained in accordance with ASTM D4959-07 (2007) standard. While pH of soil was obtained using digital pH meter (ASTM G51-95 2012). Soil conductivity was determined using conductivity meter, the values were expressed in Ohm cm. Organic content was obtained based on Loss on Ignition method. Porosity was determined using the equation proposed by Hillel (2004) and was expressed in decimal. Void ratio was obtained using empirical formula based on porosity. Bulk Density was done using gravimetric core method.

C. Statistical Analysis

Multivariate regression Analysis was done to determine the relationship between Moisture Content and other soil parameters of interest. The obtained coefficient correlation results were analysed based on Pearson correlation and factor analysis to explain interrelationship of soil sample variables and parameters. PCA (Principal Component Analysis) based on data reduction multivariate technique processes original data set without altering inherent interdependencies of variables of parameters. Combination of Varimax method along with Kaiser normalization was employed to rotate Principal Component. CA (Cluster Analysis) is an approach to group or classify parameters with similar properties in the same group or class. Variable/samples in a specific cluster are similar to each other but distinct from other clusters. To measure similarity between soil parameters cluster analysis was performed employing combination of Ward's linkage and squared Euclidean distance.

Table 1. Details of Soil Samples

Table 1. Details of Soil Samples

Sample Name	Date	Latitude	Longitude	Elevation
Harjab	13-01-2018	18 12 43 N	42 31 101 E	1368
Kafarat	14-01-2018	19 47 597 N	42 24 130 E	1245
Alabla	14-01-2018	20 05 138 N	41 53 652	1438
Qutna	13-01-2018	19 20 457 N	42 49 193 E	1464
Aloqlain	13-01-2017	20 25 250 N	42 51 305 E	1029
Najd waras	14-01-2018	20 06 298 N	41 56 502 E	1447
Tabala-2	14-01-2018	20 17 197 N	42 00 521 E	1322
Almaboos	14-01-2018	19 51 588 N	42 16 269 E	1341
Tabala-1	14-01-2018	14 58 770 N	42 10 727 E	1322
Sarash	13-01-2018	14 28 881 N	42 31 591 E	1377
Quba	14-01-2018	14 33 158 N	42 17 172 E	1347
Sarash	14-01-2018	20 12 532 N	42 03 568 E	1359
Danan	13-01-2018	19 32 554 N	42 46 654 E	1370

III. RESULT AND DISCUSSION

A. *Laboratory Analysis of Physio-Chemical Properties of soil sample*

The result of soil investigated in Lab for various parameters are presented in Table 2. The pH of soil ranges from 7.72 - 8.52 giving alkaline characteristic to the soil. The conductivity of the soil ranges between 52.8 – 377.9. The soil sample at Tabala-2 depicted highest conductivity exceeding other soil samples more than other soil samples (<129 Mohm) this can be attributed to high agricultural activity in the catchment area but it is not reflected in terms of pH. The specific gravity ranges between 2.57 - 2.95 which provides a relatively similar range of data. The porosity n, ranged **from** 6.9% (BD = 1.77 g/cm3) to 25% (BD = 1.59 g/cm3). The soil organic content ranged from 0.063 % to 1.42%. The Moisture Content ranged between 0.067 % to 3.71 %. The soil sample at Najdwaras depicted highest Moisture content while other soil samples depict moisture content not more than 0.51%. This may be the result from local alteration.

Table 2. Physio Chemical Properties pH, CD, OC, PR of Soil samples

Location	pH	CD	OC	PR
HARJAB	8.31	105.6	0.14	0.10
KAFARAT	8.52	68.2	0.063	0.08
AlABlA	7.72	81.1	0.25	0.14
QUTNA	8.53	58.3	0.10	0.069
ALOQLAIN	8.16	70	0.15	0.25
NAJDWARAS	7.85	27.1	1.42	0.22
TABALA-2	8.27	377.9	0.482	0.16
ALMABOOS	8.7	128.2	0.07	0.11
TABALA-1	8.16	52.8	0.12	0.082
SARASH	8.44	83.1	0.18	0.187
QUBA	8.31	72.1	0.089	0.094
SARASH	8.44	78.6	0.19	0.109
DANAN	8.03	65.5	3.21	0.15

Table 3. Soil Properties of samples for VR< SG, MC, BD & DD

Location	VR	SG	MC	BD	DD
HARJAB	0.11	2.83	0.21	1.72	1.72
KAFARAT	0.093	2.81	0.067	1.86	1.86
AlABlA	0.16	2.95	0.51	1.92	1.93
QUTNA	0.074	2.77	0.14	1.77	1.77
ALOQLAIN	0.34	2.82	0.24	1.58	1.59
NAJDWARAS	0.29	2.57	3.71	1.47	1.52

TABALA-2	0.19	2.76	0.41	1.59	1.59
ALMABOOS	0.12	2.80	0.12	1.92	1.92
TABALA-1	0.089	2.88	0.15	2.09	2.09
SARASH	0.23	2.87	0.36	1.64	1.64
QUBA	0.10	2.87	0.16	1.43	1.43
SARASH	0.12	2.74	0.36	1.87	1.88
DANAN	0.17	2.96	0.29	1.89	1.90

B. *Regression Result Analysis*

The result of regression analysis are presented in Table 3. The regression analysis aids in modelling interrelationship of dependent variable MC with other soil parameters investigated: Specific Gravity (SG), Organic Content (OC), Porosity (PR), Void Ratio (VR), Bulk Density (BD), Dry Density (DD, pH and conductivity. The result reveal a positive relationship with MC for DD, pH, OC and SG. However, for PR and BD the relationship is negative. The relation with conductivity is 0% to cause any variation in Moisture content. Nevertheless, PR and VR indicate they cause 10.776 and 7.45% variation in soil moisture content. The developed modelling equation for existing relationship between MC and other parameters as obtained from regression analysis is

$$MC = 0.072DD + 0.182pH + 0.009OC + 7.451VR + 0.485SG - 10.776PR - 0.073BD - 2.513$$

Table 3 Analyzed Parameters Regression table

Variables	Coefficients	t-value
(Constant)	-2.513	
DD	.072	14.406
pH	.182	.053
CD	.000	.026
OC	.009	.008
PR	-10.776	-.642
VR	7.451	.634
BD	-.073	-14.828
SG	.485	.050

Based on the regression analysis the Moisture content of soil will vary at lower values of pH, OC, BD, DD SG and higher values of PR and VR. The result analysis provides a modelling equation which provides 5 soil parameters in positive correlation with moisture content while other 3 parameters are in negative correlation. However, particle size density, resistivity, tortuosity and pore structure may also affect the moisture content of the soil.

C. *Result of correlation coefficient, PCA and CA*

Table 4. presents degree of linear relationship between two soil quality parameters measured using Pearson's coefficient. A strong positive correlation between Moisture content with PR (0.521) and VR (0.521) was obtained. While Organic Content also corelates positively but is does not exhibits strong correlation 0.333. On the contrary negative correlation was obtained with pH (-0.517), CD (-0.193), BD (-0.431) and DD (-0.369). the expected correlation were obtained at 0.01 level.

Table 4. Soil Samples Coefficient Correlation in the study area

	MC	CD	pH	OC	PR	VR	BD	DD	SG
MC	1								
CD	-.193	1							
pH	-.517	.148	1						
OC	.333	-.098	-.454	1					
PR	.521	.076	-.465	.294	1				
VR	.521	.048	-.454	.271	.998**	1			
BD	-.431	-.172	.099	.012	-.522	-.528	1		
DD	-.369	-.191	.062	.038	-.499	-.505	.998**	1	
SG	-.706**	-.079	-.109	.101	-.269	-.286	.447	.409	1

D. *Principal Component Analysis Results*

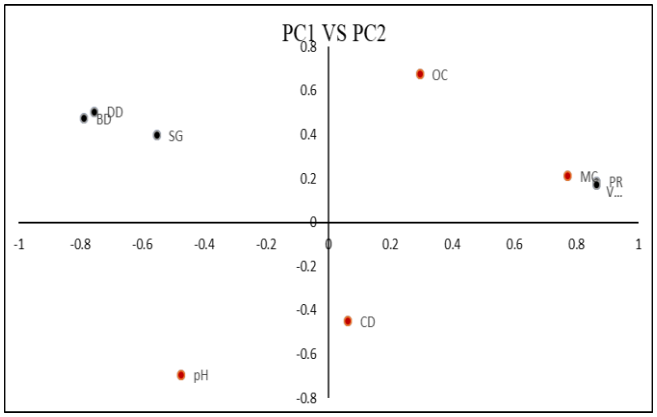

Figure 2 Principal Component Analysis of Soil properties for the study Area

Two principal components were identified using Scree plot. The two extracted principal components accounted for 63% of total variance. The PC1 accounted for 43.45% of total variance while PC2 accounted for 20.91% of total variance. The factor 1 (PC1) exhibited PR and VR to positively affecting soil moisture content, moderate variance was depicted by OC, while pH, BD, DD, SG depicted strong negative loadings. PCA analysis results validate positive correlation between MC with PR and VR. PC2 analysis exhibited strong negative loading of pH and CD, while strong positive loading was seen for OC, BD and DD. The results are in positive relation with the correlation analysis.

Table 5. Factor Loading and Eigen Values of Extracted Components

	Component	
	1	2
pH	-.476	-.693
CD	.063	-.449
MC	.771	.211
OC	.296	.674
PR	.865	.185
VR	.866	.173
BD	-.789	.474
DD	-.755	.504
SG	-.553	.396

IV. CONCLUSION

This study examined soil moisture variability of seasonal streams and the related soil properties of the obtained soil sample from the stream bed in proximity of wells. There is positive correlation between MC, OC, PR and VR. While pH, CD, BD, DD and SG exhibited negative correlations. the developed modelling equation exhibited PR and BD as having negative relationship with MC.

The PCA analysis identified two components accounting for 63% of total variance. The regression Analysis not only identify factors affecting target soil properties but can also help in modelling and predicting the parameters of soil. Thus, Multivariate analysis can be successfully employed in modelling soil properties and identifying their interrelationship.

REFERENCES

1. Aguilera, Héctor, Luis Moreno, Jan G. Wesseling, María E. Jiménez-Hernández, and Silvino Castaño. 2016. "Soil Moisture Prediction to Support Management in Semiarid Wetlands during Drying Episodes." Catena 147: 709–24. https://doi.org/10.1016/j.catena.2016.08.007.
2. Alidoust, Elham, Majid Afyuni, Mohammad Ali Hajabbasi, and Mohammad Reza Mosaddeghi. 2018. "Soil Carbon Sequestration Potential as Affected by Soil Physical and Climatic Factors under Different Land Uses in a Semiarid Region." Catena 171 (February): 62–71. https://doi.org/10.1016/j.catena.2018.07.005.
3. Becker, R., M. Gebremichael, and M. Märker. 2018. "Impact of Soil Surface and Subsurface Properties on Soil Saturated Hydraulic Conductivity in the Semi-Arid Walnut Gulch Experimental Watershed, Arizona, USA." Geoderma 322 (October 2017): 112–20. https://doi.org/10.1016/j.geoderma.2018.02.023.
4. Bracken, Ashley H, Alberta Canada, and Ashley H Bracken. 2014. "Detecting Soil Erosion in Semi-Arid Mediterranean Environments Using Siulated Enmap Data." Geoderma 340 (January): 164–74. https://doi.org/10.1016/j.geoderma.2019.01.026.
5. Eisele, Andreas, Sabine Chabrillat, Christoph Hecker, Robert Hewson, Ian C. Lau, Christian Rogass, Karl Segl, et al. 2015. "Advantages Using the Thermal Infrared (TIR) to Detect and Quantify Semi-Arid Soil Properties." Remote Sensing of Environment 163: 296–311. https://doi.org/10.1016/j.rse.2015.04.001.
6. Schoener, Gerhard, and Mark C. Stone. 2019. "Impact of Antecedent Soil Moisture on Runoff from a Semiarid Catchment." Journal of Hydrology 569 (December 2018): 627–36. https://doi.org/10.1016/j.jhydrol.2018.12.025.
7. Zeraatpisheh, Mojtaba, Shamsollah Ayoubi, Azam Jafari, Samaneh Tajik, and Peter Finke. 2018. "Digital Mapping of Soil Properties Using Multiple Machine Learning in a Semi-Arid Region, Central Iran." Geoderma 338 (January 2018): 445–52. https://doi.org/10.1016/j.geoderma.2018.09.006.