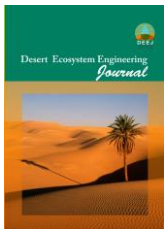
DEEJ (2019) 8 (24) 65-78

University of Kashan

Desert Ecosystem Engineering Journal

Journal homepage: http://deej.kashanu.ac.ir



Assessment of the Efficiency of Climatic factors and geomorphometry in predicting vegetation percentages based on machine learning processes

Zaynab Mirashkari[1], Majid Sadeghinia[2], Mostafa Shirmandi[3]*, Maryam Asadi[4]

Received: 24/12/2018 Accepted: 13/07/2019

Extended Abstract

Introduction: Rangelands are natural ecosystems having large genetic resources. Since plant vegetation is the bed of life on earth and changes under the influence of surrounding environmental elements, using environmental element can highly contribute to estimate vegetation percent more accurately. Two effective elements which can contribute to estimate the vegetation distribution are climatic elements and geomorphometric. Nowadays, one of new techniques which have attracted much attention to estimate vegetation percent is machine learning process which is able to establish a relationship between various variables of environmental conditions with the least costs and workforce. Therefore, in this study, geomorphometric and climatic elements and data mining techniques have been applied to address the vegetation percent.

Materials and Methods: The studied region is a part of Yazd-Ardakan plain and Nadoshan region. Sampling and measuring vegetation percent have been carried out in using transects and plots. In order to extract the geomorphometric elements, digital elevation models and SAGA software were utilized and also, seven meteorological stations were regarded to achieve the climatic elements. In the current research, to investigate the impact of different climatic and geomorphometric elements on vegetation percent estimate, data mining models such as artificial neural network, the nearest neighbor, support vector machine, decision tree, Gaussian process and linear regression were used. Artificial neural network: is one of computational models which can determine the relationships between inputs and outputs of one physical system and a network of connected nodes even if they are complicated and nonlinear. The nearest neighbor: involves selecting a certain number of data vectors and random sampling from the set-in order to simulate the time interval followed by a certain period. Support vector machine: is an efficient learning system based on theory of optimization applying the inductive principle of structural error minimization which leads to a total optimum response. Decision tree: is a method to estimate

1. M.Sc. Graduate of Range Management, Department of Nature Engineering, Faculty of Agriculture & Natural Resources, Ardakan University
2. Assistant Professor, Department of Nature Engineering, Faculty of Agriculture & Natural Resources, Ardakan University
3. Assistant Professor, Department of Horticultural Science, Faculty of Agriculture & Natural Resources, Ardakan University; shirmardi@ardakan.ac.ir
4. Ph.D. Student of Watershed Management, Department of Reclamation of Arid and Mountainous Regions Engineering, Faculty of Natural Resources, Tehran University
DOI: 10.22052/deej.2018.7.24.45

the discrete functions which are strong against the confused data and are capable to learn the terminology with two different fields. Gaussian process: is a random one consisted of random values in each point in a time or location domain so that each random variable has a normal distribution. Linear regression: is applied to model the value of a dependent quantitative variable based on a linear relationship with one or more independent variables. To assess the models and compare the results, such assessment criteria as RMSE, correlation coefficient and coefficient of determination have been used. Here, to weigh the input parameters of support vector machine algorithm, normal vector coefficients related to a linear support vector machine were specified as the weights.

Results: The study indicated that data mining models are able to estimate vegetation percent more accurately. Using geomorphometric elements, data mining models have shown that Gaussian process model had the most accuracy in the set of training and test data. As well, in applying the models on climatic data, it has been reported that decision tree had the most accuracy in the set of training and test data to estimate the vegetation percent.

Discussion and Conclusion: Vegetation is controlled by such environmental variables as geomorphometry and climate. The results have indicated that geomorphometric elements are of more impact on vegetation percent prediction as compared to climatic ones. Weighing results showed that such geomorphometric elements as distance to waterway, waterway baseline and elevation and such climatic ones as humidity affecting the vegetation growth rate were of the highest weigh and impact in vegetation percent prediction in the desired region.

Keywords: Coverage percentage, *Artemisia sieberi*, Climatic factors, Geomorphometric factors, machine learning.