Abstract

Vegetation is the most important factor affecting the stability and balance of natural ecosystems and its importance in controlling erosion and runoff is fully accepted. Understanding the environmental factors affecting the establishment and expansion of plants، including soil، can allow the ecological management of species by helping to identify native species in each area. The purpose of this study is to estimate the spatial distribution of plant species in Khezri D.B. region of South Khorasan using machine learning methods. Regarding the conditions of the region and field visits، sampling of vegetation was done systematically at a level of about 14500 hectares. 110 soil samples were taken at the beginning and the end of each transect and randomly collected in the range of each type with respect to the root depth of the depth of 0 to 30 cm. Physical and chemical tests were carried out on soil samples using standard methods to determine the parameters of acidity، soil texture، organic carbon and electrical conductivity. 16 environment variable includes the features of the land، vegetation index and the index of salt as an estimator for the maps to be used; It should be noted that 80% of the data was used to train models and 20% of the data was used for testing. Finally، by using random forest methods، tree classification and regression، and generalized population، important variables were identified in the distribution of plant species. The R-software was used to run the models and the spatial distribution map of the plant species was prepared. The significance of variables in the random forest model was determined by the method of average reduction of accuracy in generalized population models and tree classification and regression with denominator coefficient.

In the randomized forest model، the slope، electrical conductivity، gravel، silt، channel network، topographic variables in the distribution of plant species of the region are affected. Tree classification and regression model showed longitudinal curvature variables، vertical distance to canal network، slight difference in slope، valley depth and arable index in all habitats effective in occurrence of plant species. The general population model of the vertical spacing variables to the channel network and the base level of the channel network were more important in the distribution of the species in the studied types. Kappa coefficient was used to evaluate the accuracy of the models.

The Kappa coefficient for all types is more than 80 percent، which indicates good conditions for implementing models. In generalized population model، Kappa coefficient in all types is more than 95%، but it is linear and non-linear in variables. In general population method، *Tamarix serotina* with a kappa coefficient of 98% had the best performance. In tree classification and regression model، the best performance was for *ephedra sinica* site with 82% kappa coefficient. Kappa coefficient has the best performance for a random forest model with a Kappa coefficient of 91% on the *Ephedra sinica* site. In general، the results of this research show good performance of machine learning methods in estimating spatial distribution of rangeland species.

Keywords: Rangeland species، Potential dispersal، Machine learning models، Khezri pastures.

University of Zabol
Graduate School
Faculty of Water and Soil
Department of Range and Watershed Management

Estimation of spatial distribution of rangeland species using machine learning methods (Case study: Khezri Dasht Bayaz ranges South Khorasan)

Supervisor:

Dr Hosein Pirisahragard

Advisors:

Dr Mohamad Reza Pahlavanrad

Dr Reza yari

By:

Malihe Keyghobadi

February 2019