IOSR Journal of Environmental Science, Toxicology and Food Technology (IOSR-JESTFT)
e-ISSN: 2319-2402,p- ISSN: 2319-2399.Volume 9, Issue 12 Ver. III (Dec. 2015), PP 60-65
www.iosrjournals.org

Identify potential habitat of awooden species in temporal forestusingLogistic Regressionmodel

D.Rostam Salam Aziz
University of koya-faculty of Education .Dep. Geography

Abstract: *Growth of plant species in an area is affected by environmental factors and biological area.To determine the potential habitat of plant species, studies on the ecological of the species and the Climatic and edaphic characteristics of the study area are necessary.*

In this study, potential habitat of Crataegus pontica in Central Zagros (Lorestan province) was modeled using geographic information systems (GIS) and Logistic Regression model. In order to collect the different maps such as topographic, geological, climate and soil of various organizations, these maps were digitized to provide the required data layers.

Land use maps such as rock outcrops, forests, cropland, grassland, and shadow data were prepared from Landsat 8. Then the presence and absence of species were identified. To do this study, samples were taken from case study using 500 line transects.The coordinates of the 505 sites as places of presence and absence of species were recorded. In 250 transects, Crataegus ponticawere detected. Also in the 250transect the species was not observed. Then the presence and absence points of species assessed with maps of the landform, geology, soil and climate were combined by GIS procedures. Afterward the data was export to Logistic Regression model for habitat modeling of C.pontica. Thereafter using Arc GIS10.1, the potential distribution map of C. pontica species in the study area was determined.

The results showed thatthere is a kind of significance of correlation among the parameters and the availability of c. pontica, however their impact may vary for example Elevation comes first and Precipitation comes second in terms of the impact on thec. pontica where (Mean temperature, maximum temperature, minimum temperature) have a restricted impact on it, on the contrary (slope, aspect) are non -significance. And also to determine the total assessment reliability including 150presence/absence points, we have made are overlay process on CCA equation we could bring about the result the reliability assessment of the map result is %67.

Keyword: *Habitat modeling,Logistic Regression, GIS,Crataegus pontica,Species, Zagros*

I. Introduction

The common idea for many people is that forests are just a collection of trees. However, forests are much more than just that. They are complex, functional systems of interacting and often interdependent biological, physical, and chemical components, the biological part of which has evolved to perpetuate itself through the production of new organic matter created by consecutive generations of plants and animals (Kimmins 2004).

Forest cover 33 million square kilometers %26 of the Earths land surface More than 1.6 billion people depend to varying degrees on forests for their livelihoods (Sustaining forests : a development strategy 2004). Forests Multiple-benefit play very important role economic functions, they help to preserve agricultural land from the danger of erosion, protect water sources, and reduce the risks of natural disasters such as landslides and flooding.

Forests cover about 12 million hectares in Iran, An area about of 5-milion hectares in the mountainous Zagros region, From north west to southwest, has covered by Zagros forests that are one of the most important floristic zones in Iran(Heydari 2013). But in spite of that the forests in that region may face to many problems, Increasing population, low level of development and high dependence of local communities on forests for their primary livelihood needs, are the main reasons of this destruction (Fattahi 1994).The most species scattered in this region area which is are Quercus spp. (oaks), Pistaciamutica (wild pistachio), Crategus spp. and Pyrusspp(Jazirehi and Ebrahimi Rostaghi 2003). Depending density and Horizontal diversity and Vertical diversity of forest in the region the characteristics ofaltitude, and distance from sea, topography Abiotic factors, such as elevation, slope, aspect, soil texture, climate.The study of plant diversity showed that there are 165 woody species (tree and shrub) in Zagros forest (Jazirehi and Ebrahimi Rostaghi 2003). C. pontica is one these Zagrosian species. Studies concerned the classification diversity forest showed approximately 280 species types of genus Crataegus, widely distributed throughout the Northern temperate region of the world (Blumenthal et al., 2000)mostly between latitudes 30° and 50° .there are17 species of that genus in Iran (Dehkordi and Ghannadi 2012). C. pontica is a small tree or shrub, 4-6 m. Leaves coriaceous, light green, softly pubescent

from the both sides, entire, obovate or rhomb-shaped, 3-lobed with cuneate base, deeply 3-6-partite. Fruits yellow, rounded ovate 12-18 mm. (Mozaffarian 2005)

The genus C. pontica belongs to the family Rosacea comprises of a complex group of tree contain of carbohydrates, flavonoids, cyanogen tics, phenolic compounds, saponins, phytosterols,tannins, fats and fixed oils. The important diagnostic features of the powder include anomocytic Stomata .(Yazdinezhad et al. 2014) .

Several methods have been developed for modeling spatial distribution models(SDM) presence/absence data, Logistic regression models is biogeographically important in estimating the niche and distribution of the species. This model is normally used in small areas (Guisan and Thuiller 2005). See for example(Finley et al. 2009)&(Ovaskainen et al. 2010)&(Yang et al. 2006)&(Özkan et al. 2015)(Chahouki and Zare Chahouki 2011).

In this research we try to apply this model in a large area (approximately 1,300,000 ha) for a forest tree species (C. pontica) in central Zagros.

The researchers can summarize the goals of this paper in three points as following:

1- To evaluate the potential of Logistic regression model in determining potential habitat of a forest species in large distribution area.
2- To define the potential habitat of Crataeguspontica in a Zagros region as a temporal forest.

II. Materials And Methods

Study area

Lorestan Province is a province of western Iran in the Zagros Mountains. In the wider sense it consists of that part of western Iran coinciding with the province of Ilam and extending for about 650 km on a northwest to southeast axis from Kermanshah to Fars, with a breadth of 150–180 km. The terrain consists chiefly of mountains, with numerous ranges, part of the Zagros chain, running northwest to southeast. The central range has many summits that almost reach the line of perpetual snow, rising to 4000 m and more.The climate is generally sub-humid continental with winter precipitation, a lot of which falls as snow. At Khorramabad, the average annual precipitation totals 530 millimeters (21 inches) of rainfall equivalent, while up to 1270 millimeters (50 inches) may fall on the highest mountains. The months June to September are usually absolutely dry, but Khorramabad can expect 4 inches of rainfall equivalent in December and January.Temperatures vary widely with the seasons and between day and night. At Khorramabad, summer temperatures typically range from a minimum of 12 °C (54 °F) to a hot maximum of 32 °C (90 °F). In winter, they range from a minimum of -2 °C (28 °F) to a chilly maximum of 8 °C (46 °F).Figure 1 shows the location of the study area.The area of the present study is about 1,300,000 ha.

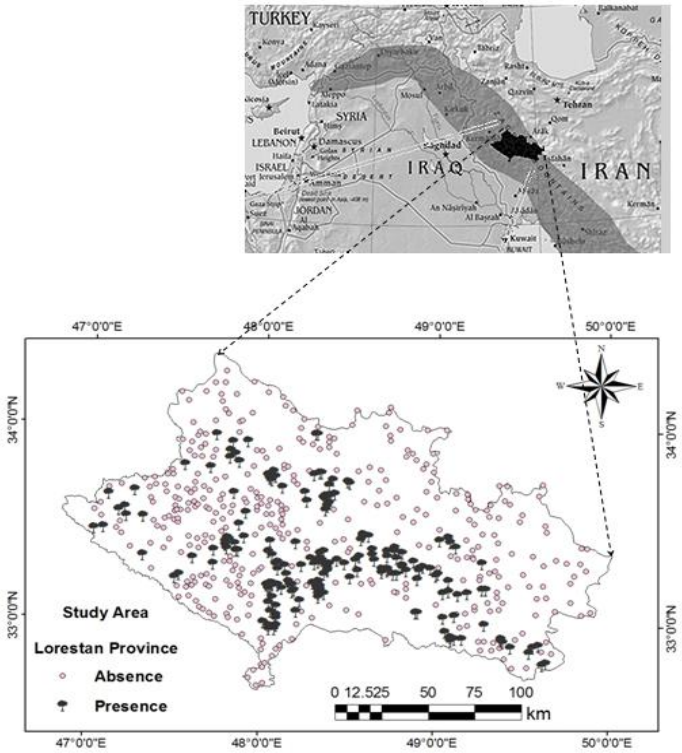

Figure 1.the location of the study area.

Providing data layers: Data layers required for this study was provided from two sources:thematic maps of the region that prepared by different organizations and the maps were generated using GIS for this study.At this stage, after preparation of topographic map with a scale of 1: 50000, the maps of elevation, slope and aspect were produced.Geological map with a scale of 1: 1,000,000 was obtained from the Geological Survey of Iran.Soli map of study area at a scale of 1: 1,000,000 was obtained from the Institute of Soil and Water.Layers of annual precipitation, annual mean temperature, annual mean maximum and minimum temperature, absolute maximum and minimum temperature, average temperature and average winter temperaturewere obtained from the National Weather Service of Iran.

In order to prepare a map of potential habitats of species of hawthorn tree (C. pontica), digital elevation model was developed. Then in the Arc GIS 10.1, degrees inclination and azimuth slope of maps was produced and were used as physiographic variables.The coordinates of the 1600 sites as places of presence and absence of species were recorded.In 800 transects, C. ponticawere detected. Also in the 800 transect the species was not observed.Then the presence and absence points of species assessed with maps of the landform, geology, soil and climate were combined by GIS procedures.Afterward the data was export to Logistic Regression model for habitat modeling ofC.pontica. Thereafter using Arc GIS10.1, the potential distribution map of C. pontica species in the study area was generated.

In order to have access to the prediction of potential habitat of C. pontica in the case study the following step will be considered:

1. The adaptation of transect with the distance of 1*1 km in the entire Lorestan province forests
2. In each transect the number of the C. pontica has been counted.
3. Each transect that has more than one tree of C. pontica has been taken as a present point which as 240 present points.
4. 480 absent point has been taken into consideration .240 absent point, have been considered within the forest, and 240 absent point were outside the forest area.
5. In order to have an accurate result, we have taken 50present points, and 50 absent points.
6. By the environment variable such as (perception, elevation, slope, aspect, mean temperature, maxima temperature,and minimum temperature) we have brought about a well-designed map that can represent them.
7. To do the training point the proportion amount of the environment data has been taken according to present/ absent
8. For determining threshold, Receiver operating characteristic (ROC) has been implemented which is about the probability of presence/absence, and then it has been reclassified.
9. We have made use of Canonical correspondence analysis(CCA) to assessment the reliability of the outcome of the probability presence/absence map.

III. Data Analysis:

In this study, to determine the factors affecting the wooden species distribution area and offers forecasting model, we were used the logistic regression model.In the next step using geostatistics and kriging interpolation, map of environment variables, including climatic factors were produced.After mapping the factors considered, using models derived from logistic regression, coefficients for each variable were produced. Then, in the GIS, map of potential habitats of wooden species were produced.

Due to the above, this study aimed to evaluate the ability of logistic regression to identify important variables affecting toC.ponticaspecies was done.Logistic regression(LR) is a kind ofgeneralized linear model (GLM) describes the relation between the response data are binary, usually the linear sum leading to predicts variables probabilities (Hosmer Jr and Lemeshow 2004).

The GLM model can be used to generate predictions of alliance presence for ensample locations based on their integration with the digital maps of the predictor variables. The implementation of the logistic models for predictions is straightforward-each predictor variable is multiplied by its model coefficient then summed to provide the linear predictor (LP) for the alliance. In order to obtain probability values between 0 and 1, a logistic transformation of LP is used.Occurrence probability of each wooden species is calculated with respect to the combined effect of site conditions with the following equation:

(1). Probe (Alliance) = $e^{LP}/(1+e^{LP})$. LP= $b_0+b_{1x1}+b_2x_2+\ldots+b_nx_n$

IV. Results:

This study illustrates the results of 570 present samples as well as the 7parameters environmentdatathey have significance correlation only with 5 parameters (Elevation, precipitation. Mean temperature, maximum temperature, minimum temperature)(table1).

Table1. Deviance Table

Source	DF	AdjDevAdj	Mean	Chi-Square	P-Value
Regression	5	182.641	36.528	182.64	0.000
Elevation	1	107.719	107.719	107.72	0.000
Precipitation	1	4.905	4.905	4.90	0.027
Mean temperature	1	67.479	67.479	67.48	0.000
Maximum temperature	1	24.089	24.089	24.09	0.000
Minimum temperature	1	112.216	112.216	112.22	0.000

According to the equation of liner Regression, the connection between the present /absence of this sort of tree will be achieved based on the following equation:

Regression Equation:

Equation2: Y' = 28.62 - 0.02981 Elevation + 0.00341 Precipitation + 1.100 Mean temperature - 0.894 Maximum temperature - 1.769 Minimum temperature.

Then we will depend on the table of Goodness –of-fit tests, according to that table chi-square and P-value are significance (table 2)to have the presence probability of c.pontica in the work field have drawn on the following equation in the GIS program, after that in the same program GIS we made use of raster calculator to do the sample designing (fig 2).

Table2. Goodness-of-Fit Tests

Test	DF	Chi-Square	P-Value
Deviance	564	542.99	0.730
Pearson	564	541.58	0.744
Hosmer-Lemeshow	8	7.17	0.518

Equation2:P(1) = exp(Y')/(1 + exp(Y'))

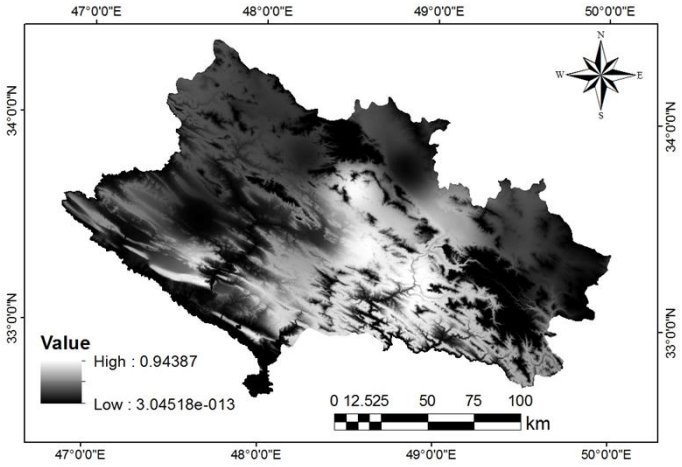

Fig2. Probability ofc.pontica, using the logistic regression model.

The result is the probability map of C. pontica presence in order to assess and indicate the rete of minimum potential percentage, we used (ROC) which is available in the Med CALC software as a result we have achieved optimal criterion, then we reclassified the potential maps c. pontica presence.(fig 3)

To assess the definite reliability ofc. pontica presence potential map, we have made use of (CCA)in which the potential reliabilityof the study is %67

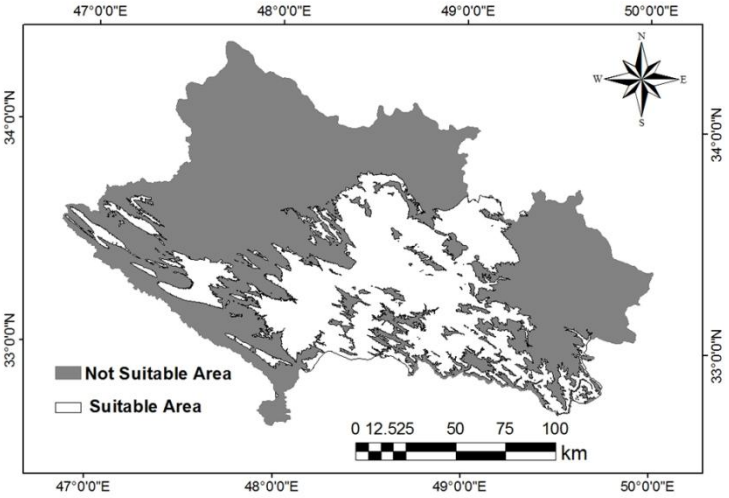

Fig3.Map of the habitats suitable for c.pontica.

V. Discussion:

According (fig) there is a kind of significance of correlation among the parameters and the availability of c. pontica, however their impact may vary for example Elevation comes first and Precipitation comes second in terms of the impact on thec. ponticawhere (Mean temperature, maximum temperature, minimum temperature) have a restricted impact on it, on the contrary (slope, aspect) are non -significance, therefore they are not given in the related table.

To assess the reliability of the relationship, according to P.valu, which is stated in the table (Goodness-of-fit tests) on the one hand and also according to the correlation which are recommended by Person, Hosmer, and Diviance in which P.value>0.5. it brings about that it has a significance correlation.

To design the environment data map which are consisted of (7) parameters which only we could are made use of (5) of them which are (Elevation, precipitation. Mean temperature, maximum temperature, minimum temperature).

According to the ROC equation which has been stated in (fig) threshold is more than >0.36 and this result means that the potential presence rate which is more than %36 is regarded assignificance. For this purpose by GIS program in the (fig) we have reclassified to classes presence =1, absence =0.

To determine the total assessment reliability including 150presence/absence points, we have made are overlay process on CCA equation we could bring about the result the reliability assessment of the map result is %67.

VI. Conclusion

Several methods have been developed for modeling spatial distribution models(SDM) presence/absence data, Logistic regression models is biogeographically important in estimating the niche and distribution of the species. In this research we try to apply this model in a large area (approximately 1,300,000 ha) for a forest tree species (C. pontica) in central Zagros.

The results of Logistic regression models showed that there is a kind of significance of correlation among the parameters and the availability of c. pontica, however their impact may vary for example Elevation comes first and Precipitation comes second in terms of the impact on thec. pontica where (Mean temperature, maximum temperature, minimum temperature) have a restricted impact on it, on the contrary (slope, aspect) are non -significance. And also to determine the total assessment reliability including 150 presence/absence points, we have made are overlay process on CCA equation we could bring about the result the reliability assessment of the map result is %67.

The results of this study showed that Logistic regression, a method for mapping the distribution of woody species. And also the results show that this method can be used in large area. Therefore, it is suggested used the Logistic regression method for mapping the potential of distribution species of plants and wooden in large area.

These habitat models provide a platform on which future Eco physiologicalresearch can be based that will enable the constructionof more refined models, and more importantly, better conservationmanagement.

References

[1]. Chahouki Z, Zare Chahouki A. 2011. Predicting the distribution of plant species using logistic regression (Case study: Garizat rangelands of Yazd province). Desert.15:151-158.

[2]. Dehkordi NG, Ghannadi A. 2012. Crataegus pontica C. Koch; microscopical, macroscopical and chemical investigations and their uses in chemotaxonomy. Research in Pharmaceutical Sciences.7:S711.

[3]. Fattahi M. 1994. Study on Zagros oak forests and the most important their destruction causes. Institute of Forests and Rangelands Research press, Sanandaj, Iran.

[4]. Finley AO, Banerjee S, McRoberts RE. 2009. Hierarchical spatial models for predicting tree species assemblages across large domains. The annals of applied statistics.3:1052.

[5]. Guisan A, Thuiller W. 2005. Predicting species distribution: offering more than simple habitat models. Ecology letters.8:993-1009.

[6]. Heydari M. 2013. Floristic study of Dalab woodlands, north-east of Ilam province, west Iran. Iranian Journal of Science and Technology (Sciences).37:301-308.

[7]. Hosmer Jr DW, Lemeshow S. 2004. Applied logistic regression: John Wiley & Sons.

[8]. Jazirehi M, Ebrahimi Rostaghi M. 2003. Silviculture in zagros. University of Tehran.560.

[9]. Kimmins J. 2004. Forest Ecology; A foundation for Sustainable Forest Management and Environmental Ethics in Forestry. Thired Edition. In: Prentice Hall, New Jersey.

[10]. Ovaskainen O, Hottola J, Siitonen J. 2010. Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. Ecology.91:2514-2521.

[11]. Özkan K, ŞENTÜRK Ö, Mert A, Negiz MG. 2015. Modeling and mapping potential distribution of Crimean juniper (Juniperus excelsa Bieb.) using correlative approaches. Journal of Environmental Biology.36:9.

[12]. Sustaining forests : a development strategy. 2004.

[13]. Yang X, Skidmore AK, Melick DR, Zhou Z, Xu J. 2006. Mapping non-wood forest product (matsutake mushrooms) using logistic regression and a GIS expert system. ecological modelling.198:208-218.

[14]. Yazdinezhad A, Najafi F, Mousavi A. 2014. Pharmacognostic and phytochemical studies of leaves of Crataegus pontica C. Koch. Journal of Chemical & Pharmaceutical Research.6.