Range Mgmt. & Agroforestry 38 (1) :17-26, 2017
ISSN 0971-2070

Modeling distribution of vegetation types in arid and semiarid rangelands of Iran through binary logistic regression and canonical correspondence analysis techniques

Mohammad Mousaei Sanjerehei*
Department of Plant Protection, Yazd Branch, Islamic Azad University, Yazd, Iran
*Corresponding author e-mail: mmusaei@iauyazd.ac.ir
Received: 15th January, 2016 Accepted: 28th January, 2017

Abstract

The distribution of nine vegetation types in arid and semiarid Nodushan rangelands was predicted based on the indicator species by using binary logistic regression (BLR) and canonical correspondence analysis (CCA) techniques. Nine soil variables (salinity, pH, C/N, available water, gravel, texture, gypsum, organic matter and lime) in two depths (0-20 and 20-60 cm) and three topographic variables (elevation, slope and aspect) were used for modeling. The habitat suitability was determined using the maximum sensitivity plus specificity threshold. Results indicated that the indicator species in each vegetation type was effective and efficient for modeling the distribution of the vegetation type. BLR models provided more accurate predictions than CCA models in most of the vegetation types. The vegetation type whose distribution was well modeled by CCA was also well modeled by BLR but not vice-versa. None of the techniques provided accurate results for the vegetation type that was under grazing disturbance.

Keywords: Binary logistic regression, Canonical correspondence analysis, Indicator species, Vegetation types

Introduction

Modeling the distribution of species and communities is of high interest due to a number of reasons like for analyzing the relationships between environment and species attributes such as species distribution (Guisan *et al*., 2002) and composition (Ohmann and Gregory, 2002), identifying the potential of an area for supporting a particular species or a high level of biodiversity, successfully implementing restoration measures and to predict the impact of climate change and land use change on the distribution and diversity of species (Riordan and Rundel, 2014). A variety of statistical models have been used to relate the environmental factors to distribution of plant species (Guisan *et al*., 1999; Chahouki and Ahvazi, 2012; Sahragard *et al*., 2015; Zare chahouki and Esfanjani, 2015), plant communities (Zimmermann and Kienast, 1999; Chahouki *et al*., 2010), biodiversity (Rodriguez-Castaneda *et al*., 2012) and life zones (Mousaei Sanjerehei, 2014). Canonical correspondence analysis (CCA) and binary logistic regression (BLR) are the most efficient used methods for predicting the distribution, composition and structure of species and communities. CCA is unique among the ordination methods in that the ordination of the main matrix representing species data is constrained by a multiple regression on environmental variables included in the second matrix. CCA is appropriate for data set with many zeros (i.e. presence-absence data) (Guisan *et al.,* 1999; Aspinall, 2002; Calef *et al*., 2005; Chahouki *et al*., 2012). The distribution of species can be modeled using a regression equation made with the canonical coefficients of the environmental predictors. To predict the potential distribution of the species, Euclidean distance to the centroid (weighted average) of each species in canonical space is calculated. The centroid of a species indicates the position of the species distribution along an environmental variable (Ter Braak, 1987). For presence-absence data, the centroid is simply the mean of the ordination axis values over the sampling points selected for each plant species (Ter Braak and Looman, 1986). Binary logistic regression uses maximum likelihood to estimate the probability of a dichotomous outcome. The probability can take any value between 0 and 1 (Guisan *et al*., 2002). Most parts of arid and semiarid rangelands of Iran are covered by *Artemisia sieberi* due to the adaptation of this species to a wide range of environmental conditions (Mousaei Sanjerehei *et al*., 2011; Hosseini *et al*., 2013). Therefore, for modeling the distribution of the study vegetation types whose first dominant species was *A. sieberi* (Fig. 1), the second dominant species was selected as the indicator species representing the local edaphic, topographic and climatic conditions. The objectives of this study were to determine the most important environmental drivers of the distribution of 9 vegetation types in Nodushan arid and semiarid rangelands, to map the potential distribution of the vegetation types using CCA and BLR and to compare the efficiency of the two models for predicting the vegetation distribution.

Materials and Methods

Study area: Three topographic (elevation, slope and aspect) and nine edaphic factors (soil texture, organic carbon, pH, EC, lime, gypsum, nitrogen, available water and gravels) were used to model the distribution of 9 vegetation types (*A. sieberi-S. tomentosa*, *A. sieberi-E. strobilacea*, *A. aphylla-A. sieberi, A. sieberi-S. arbusculiformis, A. sieberi-P. harmala, A. sieberi-Z. eurypterum, A. sieberi-E. ceratoides, A. sieberi-I. songarica* and *A. aucheri*) in Nodushan rangelands of Yazd, Iran, using BLR and CCA. Arid and semiarid rangelands of Nodushan are located in the northwest of Yazd province in central Iran (31°462 N, 52°242 E to 32°152 N, 53°472 E) at elevations between 1540 m and 3250 m. The mean annual precipitation ranges from 75 to 170 mm. The mean annual temperature is 12.6 °C (7.3-16.9 °C).

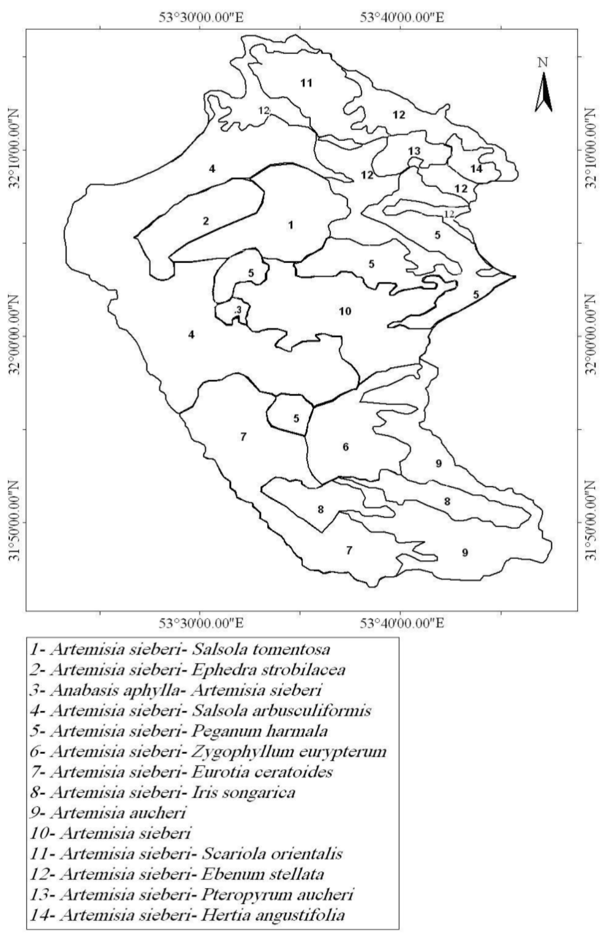

Fig 1. The actual vegetation map of Nodushan rangelands

Sampling methods: Sampling was performed in homogenous units identified by overlapping the elevation, geologic and the ETM+ cluster map (Fig. 2) of the study area. To produce the cluster map, a scene of landsat 7 ETM+ image (Path: 162, Row: 038) covering Nodushan rangelands was acquired under clear atmospheric conditions on August 06, 2006. Clustering (unsupervised classification) was performed using ETM+ bands 3 (red) and 4 (visible near infrared) images. The images of these bands effectively represent the reflectance attributes of the soil and vegetation in arid environments (Mousaei Sanjerehei, 2014). In each unit, six to ten 100-m transects were located approximately equidistantly. Soil samples were taken in the middle of each transect from the depth of 0-20 and 20-60 cm (Brunetto *et al*., 2011; Ghorbani *et al*., 2015). The percentage cover of each species was estimated in 30 equidistantly located 2×2 m quadrates along each transect. The coordinate of sampling points was registered for determining the elevation, slope and aspect using the topographic maps (Fig. 2). The maps of soil properties were produced interpolating the point data for each variable based on the kriging algorithm (Fig. 3). *Artemisia sieberi* was the most dominant species in the area due to its adaptation to a wide range of environmental conditions, and it was excluded from the analysis (Table 1). The vegetation types were modeled based on their second dominant species as indicator species where *A. sieberi* was the first dominant species.

BLR and CCA techniques: The presence/absence data for the plant species were used for CCA and BLR analysis. Binary logistic regression was calculated (Guisan *et al.,* 1999) as,

$$P = \frac{\exp(\ a + b_1 x_1 + b_2 x_2 + ... + b_n x_n)}{1 + \exp(\ a + b_1 x_1 + b_2 x_2 + ... + b_n x_n)}$$

This equation gives a probability value from 0 to 1, where *p* is the probability of the vegetation type occurrence, exp is the base of natural logarithms, a is the constant of the equation, *b* is the coefficient of the explanatory variables and $x_1 \ldots x_n$ are the variables. Backward stepwise logistic regression method was used to identify the statistically significant predictors (P<0.05). The amount of variance explained by the predictors in each model was determined by calculating Nagelkerke's R square.

For modeling the species distribution using CCA, two matrices were constructed. In the main matrix, presence/ absence data of the indicator species were used, and the second matrix was composed of the values of environmental variables (Ter Braak, 1986). The map of

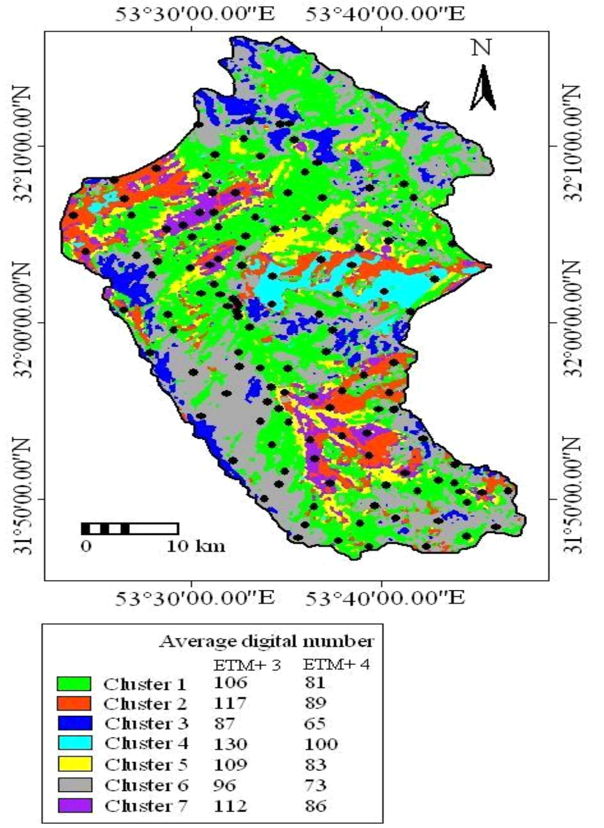

Fig 2. The classified map of Nodushan rangeland using unsupervised classification (clustering) algorithm based on the images of ETM+ bands 3 and 4. The sample points were selected based on overlaying the elevation, geologic and the cluster map. A 5×5 majority filter was applied to the clustered map.

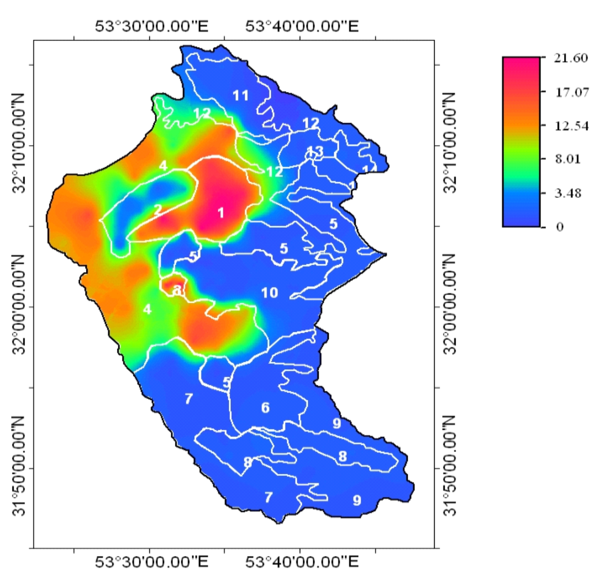

Fig 3. The map of EC (depth 0-20 cm) for Nodushan rangelands, an example of the produced map using kriging interpolation of point data

the environmental gradient for each canonical axis was produced for four first axes using the regression equation made with canonical coefficients. CCA was performed using Canoco 4.5 (ter Braak and Smilauer, 2002). The predicted probability of the species occurrence was classified based on the tolerance of the species. The tolerance based on presence-absence data is the standard deviation of the values of the ordination axis over the sampling points selected for each species (ter Braak, 1987). To identify the suitable = 1 and unsuitable = 0 habitat for each species based on CCA and BLR, the maximum sensitivity plus specificity threshold was calculated for each species model using the occurrence probability maps. This is one of the efficient methods for determining the threshold in species distribution modeling (Liu *et al.,* 2005). Sensitivity is calculated as a/(a+c) and specificity as d/(b+d), where a: true positives (or presences), b: false positives (or presences), c: false negatives (absences) and d: true negatives (or absences). Kappa coefficient was used to compare the predicted potential distribution maps with the actual vegetation map (Fig. 1) for obtaining the accuracy of the prediction. Kappa statistic strength of agreement is classified as: <0 poor, 0-0.20 slight, 0.21-0.40 fair, 0.41-0.60 moderate, 0.61-0.80 substantial and 0.81-1 almost perfect (Landis and Koch, 1977). The analysis and calculations were done using IlWIS 3.2 (ILWIS 3.2 Academic Version, 2004) and SPSS 15 (SPSS Inc., Chicago, IL).

Results and Discussion

BLR spatial modeling of the vegetation types: The significant predictors (P<0.05), odds ratio, coefficient of variables and the value of Nagelkerke's R square were obtained from the BLR modeling of the vegetation types (Table 2). The strongest relationship between the prediction and the predictors was found for *I. songarica* (Nagelkerke R^2 = 0.94), followed by *E. strobilacea* (0.91), *A. aphylla* (0.90), *S. arbusculiformis* (0.89), *A. aucheri* (0.88), *S. tomentosa* (0.87), *E. ceratoides* (0.87) and *Z. eurypterum* (0.71). The lowest R-squared (0.59) was related to *P. harmala*. C/N_1, $sand_1$, OM_2 and $silt_1$ were the significant predictors in predicting the occurrence of *I. songarica*. $Gypsum_{1,2}$ made a significant contribution to the prediction of the presence probability of *E. strobilacea.* $EC_{1,2}$ was the significant factor for *A. aphylla* and *S. tomentosa*, EC_1, C/N_1 and $gypsum_2$ for *S. strobilacea*, elevation for *A. aucheri*, C/N_1 for *E. ceratoides*, $lime_{1,2}$ and $sand_2$ for *Z. eurypterum* and OM_1, AW_2 and $clay_2$ for *P. harmala* model. Odds ratio (OR) is defined as the probability of occurrence over the probability of non-

Table 1. Mean values of environmental variables and cover of the plant species (%) in nine studied vegetation types of Nodushan rangelands

	A.sieberi-S. tomentosa	***A.sieberi-E. strobilacea***	***A. aphylla-A. sieberi***	***A.sieberi-S. arbusculiformis***	***A.sieberi-P. harmala***	***A.sieberi-Z. eurypterum***	***A.sieberi-E. ceratoides***	***A.sieberi-I. songarica***	***A.aucheri***
Environmental variable									
$Gravel_1$	17(4.58)	14(3.82)	12(3.97)	12(3.89)	11(4.69)	11(3.99)	16(3.86)	8(3.19)	8(3.9)
$Gravel_2$	19(3.79)	21(3.73)	18(3.94)	21(1.99)	19(4.54)	14(4.78)	9(4.5)	14(4.53)	10(4.87)
PH_1	7.3(0.24)	7.64(0.19)	7.43(0.23)	7.38(0.18)	7.59(0.19)	7.9(0.06)	7.06(0.07)	7.19(0.14)	7.67(0.13)
PH_2	7.83(0.08)	7.31(0.12)	7.31(0.12)	7.43(0.13)	7.78(0.09)	7.75(0.1)	7.72(0.17)	7.16(0.09)	7.69(0.09)
$Gyps_1$	0(0)	61.74(7.55)	2.06(0.73)	0(0)	0(0)	0(0)	1.01(0.21)	0(0)	0(0)
$Gyps_2$	0(0)	63.43(6.43)	0(0)	29.4(2.64)	0(0)	0(0)	0.11(0.03)	0(0)	0.21(0.06)
C/N_1	5.3(1.6)	3.94(1.71)	11.87(3.22)	1.38(0.55)	9.5(3.52)	11.75(4.56)	57.1(7.04)	6(1.89)	16.52(3.73)
C/N_2	4.76(1.93)	75.71(4.81)	3.51(1.44)	2.76(0.97)	4.52(2.11)	4.06(2.09)	21.07(3.62)	3.45(1.05)	12.45(4.17)
$Sand_1$	69.4(4.63)	66.4(3.94)	63.4(3.74)	57.4(4.03)	63.4(3.57)	53.4(2.89)	71.4(4.23)	43.4(3.83)	41.4(1.8)
$Sand_2$	73.4(4.29)	81.4(3.01)	68.4(3.86)	83.4(4.92)	77.4(4.01)	70.4(4.27)	77.4(3.58)	63.4(3.2)	41.4(2.79)
$Silt_1$	21.3(4.38)	23.3(3.78)	26.3(5.24)	33.3(4.18)	27.3(4.45)	37.3(4.72)	21.3(5.03)	47.3(5.3)	46.9(5.25)
$Silt_2$	17.3(3.74)	11.3(3.57)	21.3(3.75)	11.3(2.56)	17.3(3.26)	22.3(3.88)	16.3(3.46)	27.3(3.39)	45.3(3.7)
$Clay_1$	9.3(3.07)	10.3(2.72)	10.3(3.03)	9.3(2.62)	9.3(2.63)	9.3(2.91)	7.3(3.16)	9.3(3.56)	11.7(3.22)
$Clay_2$	9.3(1.81)	7.3(2.14)	10.3(1.98)	5.3(2.09)	5.3(2.47)	7.3(1.96)	6.3(2.51)	9.3(2.47)	13.3(2.61)
AW_1	0.1(0.02)	0.1(0.03)	0.11(0.02)	0.12(0.03)	0.11(0.02)	0.13(0.03)	0.09(0.03)	0.17(0.02)	0.17(0.02)
AW_2	0.09(0.03)	0.05(0.01)	0.1(0.03)	0.03(0.01)	0.06(0.02)	0.09(0.01)	0.07(0.02)	0.11(0.03)	0.16(0.02)
EC_1	16(3.14)	3.24(1.41)	16.5(3.77)	10.7(3.17)	0.673(0.24)	0.78(0.2)	0.72(0.24)	1.01(0.24)	0.624(0.24)
EC_2	1.98(0.95)	6.01(1.86)	30.37(4.42)	5.3(1.71)	1.17(0.5)	1.047(0.35)	0.87(0.27)	0.52(0.25)	1.6(0.45)
$Lime_1$	17.57(2.49)	9.07(3.27)	23.3(2.92)	22.8(4.06)	22.8(3.25)	34.8(3.52)	12.82(2.97)	34.8(2.5)	20.32(2.79)
$Lime_2$	13.57(4.51)	9.07(3.83)	18.57(2.74)	13.5(4)	23.57(4.05)	43.3(4.84)	15.3(4.18)	50.32(3.58)	20.07(4.6)
OM_1	0.075(0.01)	0.275(0.05)	0.229(0.05)	0.107(0.04)	0.458(0.12)	0.427(0.11)	0.275(0.05)	0.275(0.05)	0.917(0.16)
OM_2	0.275(0.06)	0.091(0.03)	0.229(0.06)	0.153(0.05)	0.229(0.06)	0.382(0.04)	0.305(0.06)	0.382(0.05)	0.962(0.13)
Elevation	1850(18)	1940(22)	2010(11)	2074(31)	2186(54)	2262(31)	2318(29)	2420(42)	2618(46)

Plant species	*A.sieberi-S. tomentosa*	*A.sieberi-E. strobilacea*	*A. aphylla-A. sieberi*	*A.sieberi-S. arbusculiformis*	*A.sieberi-P. harmala*	*A.sieberi-Z. eurypterum*	*A.sieberi-E. ceratoides*	*A.sieberi-I. songarica*	*A.aucheri*
Artemisia sieberi	6.6	5.1	2.9	9.0	5.2	12.9	10.4	16.6	-
Salsola tomentosa	1.5	0.1	-	0.15	-	-	-	-	-
Ephedra strobilacea	-	4.8	-	-	-	-	-	-	-
Anabasis aphylla	-	-	3.3	-	-	-	-	-	-
Salsola arbusculiformis	-	-	-	1.4	-	-	-	-	-
Peganum harmala	-	-	-	-	2.62	-	-	-	-
Zygophyllum eurypterum	-	-	-	-	-	2.6	-	-	-
Acantholimon sp.	-	-	-	-	-	1.9	-	2.1	-
Eurotia ceratoides	-	-	-	-	-	-	2.1	-	-
Astragalus glaucacanthus	-	-	-	-	-	-	0.52	-	-
Iris songarica	-	-	-	-	-	-	2.0	2.7	-
Stipa barbata	-	-	-	-	-	-	-	0.6	0.7
Artemisia aucheri	-	-	-	-	-	-	-	-	23
Astragalus micriphysa	-	-	-	-	-	-	-	-	1.5
Scariola orientalis	-	-	-	0.1	-	-	-	-	-

Indices 1 and 2 indicate the first (0-20 cm) and the second (20-60 cm) soil layer, respectively; Organic matter was calculated by multiplying % organic carbon by 1.724; values in parenthesis indicate standard deviations.

occurrence. The odds ratio of EC_1 (1.17) and EC_2 (1.48) for *A. aphylla* model indicated that for every one unit increase in EC_1 and EC_2 the odds of occurring this species increased by 17% and 48%, respectively. In addition the influence of EC_2 was greater than that of EC_1 on the occurrence probability of *A. aphylla*. The occurrence probability of *E. ceratoides* was positively associated with C/N_1 (OR: 1.22), and one unit increase in C/N_1 was associated with an increase of 1.22 times in the occurrence to non-occurrence probability of *E. ceratoides.* The odds ratio of elevation (1.02) in the *A. aucheri* model indicated that the probability of the presence to the absence of *A. aucheri* increased by 2% with an increase of one unit (meter) in the elevation. The calculated correlation coefficient (*r*) was 0.97 between the elevation and annual precipitation and -0.99 between the elevation and mean annual temperature in the study area. This indicated that temperature and precipitation are the important drivers of *A. aucheri* distribution and that the elevation can be considered as an efficient indicator of changes in temperature and precipitation for modeling vegetation distribution in the study area. The odds ratio of $gypsum_1$ (2.46) and $gypsum_2$ (1.56) in the *E. strobilacea* model indicated that the influence of $gypsum_1$ was more than that of $gypsum_2$ on the probability of the occurrence of *E. strobilacea* and that the occurrence to non-occurrence probability of this species increased by a factor of 2.46 and 1.56 with one unit increase in $gypsum_1$ and $gypsum_2$, respectively. The odds of *Z. eurypterum* occurrence were positively associated with $lime_1$ (OR: 1.61), $sand_2$ (OR: 1.3) and $lime_2$ (OR: 1.1) and increased by 61, 30 and 10% as $lime_1$, $sand_2$ and $lime_2$ increased by one unit, respectively. The probability of occurrence to non-occurrence of *S.*

Table 2. Results of logistic regression using soil and topographic variables

Plant species	Nagelkerke R^2	Predictor	B	Sig.	Odds ratio
Iris songarica	0.94	C/N_1	0.887	0.05	2.43
		$Sand_1$	-1.11	0.04	0.33
		OM_2	-0.40	0.044	0.67
		$Silt_1$	0.41	0.038	1.50
		Constant	42.44	0.062	
Ephedra strobilacea	0.91	$Gypsum_1$	0.9	0.05	2.46
		$Gypsum_2$	0.468	0.044	1.56
		Constant	-5.06	0.001	
Anabasis aphylla	0.90	EC_1	0.157	0.04	1.17
		EC_2	0.389	0.03	1.48
		Constant	-7.135	0.004	
Salsola arbusculiformis	0.89	EC_1	0.322	0.05	1.38
		C/N_1	-0.55	0.034	0.58
		$Gypsum_2$	0.07	0.04	1.07
		Constant	9.98	0.43	
Artemisia aucheri	0.88	E	0.018	0.002	1.02
		Constant	-46.37	0.002	
Salsola tomentosa	0.87	EC_1	0.606	0.002	1.83
		EC_2	0.011	0.04	1.01
		Constant	-5.65	0.037	
Eurotia ceratoides	0.87	C/N_1	0.198	0.001	1.22
		Constant	-6.43	0.000	
Zygophyllum eurypterum	0.71	$Lime_1$	0.476	0.013	1.61
		$Lime_2$	0.093	0.045	1.10
		$Sand_2$	0.267	0.014	1.30
		Constant	-38.96	0.004	
Peganum harmala	0.59	OM_1	15.55	0.002	1.17
		AW_2	-79.33	0.003	0.45
		$Clay_2$	-0.45	0.050	0.64
		Constant	0.199	0.855	

Indices 1 and 2 indicate the first (0-20 cm) and the second (20-60 cm) soil layer, respectively *tomentosa* was increased by 83% and 1% with one unit increase in EC_1 and EC_2, respectively. The probability of occurrence of *S. arbusculiformis* was positively associated with EC_1 (OR: 1.38) and $gypsum_2$ (OR: 1.07) and negatively associated with C/N_1 (0.58). In this case, the odds of *S. arbusculiformis* presence increased by 1.38 and 1.07 times with one unit increase in EC_1 and $gypsum_2$, respectively and decreased by 0.58 times with one unit increase in C/N_1.

Table 3. Kappa coefficient for the predicted distribution of indicator plant species using BLR and CCA

Plant species	Kappa BLR model	Kappa CCA model
Iris songarica	0.65	0.40
Ephedra strobilacea	0.79	0.75
Anabasis aphylla	0.67	0.79
Salsola arbusculiformis	0.78	0.85
Artemisia aucheri	0.71	0.76
Salsola tomentosa	0.70	0.68
Eurotia ceratoides	0.90	0.54
Zygophyllum eurypterum	0.72	0.52
Peganum harmala	0.37	0.04

C/N_1 had the greatest influence on the occurrence probability of *I. songarica* followed by $sand_1$, $silt_1$ and OM_2. The odds of occurring *I. songarica* increased by a factor of 2.43 and 1.5 with one unit increase in C/N_1 and $silt_1$ and decreased by a factor of 0.33 and 0.67 with one unit increase in $sand_1$ and OM_2. The probability of presence to absence of *P. harmala* increased by 1.17-fold and decreased by 0.45 and 0.64-fold with one unit increase in OM_1, AW_2 and $clay_2$, respectively. The highest accuracy of prediction (Table 3) was found for *E. ceratoides* (kappa = 0.90) with an almost prefect agreement between the predicted distribution map and the actual vegetation map followed by *E. strobilacea* (0.79), *S. arbusculiformis* (0.78), *Z. eurypterum* (0.72) *A. aucheri* (0.71), *S. tomentosa* (0.7), *A. aphylla* (0.67) and *I. songarica* (0.65) with a substantial accordance between the predicted and the actual distribution. The lowest accuracy was related to the prediction of *P. harmala* occurrence (0.37) with a fair agreement.

CCA spatial modeling of the vegetation types: The ordination biplots for the nine indicator plant species showed the overall results of the CCA ordination. The CCA explained 49.2% of the variance in species presence-absence data. Axis 1, 2, 3 and 4 explain 12.5%, 12.3%, 12.2% and 12.2% of the variability in the species data, respectively. Of the 25 variables, 20 variables had the highest correlation with the first two axes. pH_2, $gypsum_{1,2}$, C/N_2 and $lime_{1,2}$ had the highest correlation with the first axes. High score of *E. strobilacea* on axis 1 indicated that this species is found on the soils with a high $gypsum_{1,2}$ and C/N_2 and a low $lime_{1,2}$ and pH_2 (Fig. 4). $Gypsum_{1,2}$ and C/N_2 appear to have more influence on the distribution of this species than $lime_{1,2}$ and pH_2 based on their correlation to axis 1. The relatively low scores of *Z. eurypterum* and *I. songarica* on axis 1 exhibited a preference for high-lime soils. $Gravel_{1,2}$, pH_1, $sand_{1,2}$, $silt_{1,2}$, EC_1, $OM_{1,2}$, $AW_{1,2}$, E and slope had the highest correlation with the second axes. The species were sorted out well along this axis. High score of *A.aucheri* on axis 2 indicated that this species is a high elevation and steep slope species and showed a preference for silty loam soils with high $OM_{1,2}$ and $AW_{1,2}$ and low $gravel_{1,2}$ and $salinity_1$. *A. aphylla, S. tomentosa* and *S. arbusculiformis* on axis 2 indicated that these species are low elevation species and had a preference for sandy loam soils with high $salinity_1$ and $gravel_{1,2}$. Although EC_2 had the highest correlation with axis 4 but its high correlation to axis 2 indicated that this factor is also an important driver in the occurrence of *A. aphylla, S. tomentosa* and *S. arbusculiformis. I. songarica* on axis 1 and 2 indicated that this species is a relatively high elevation species with a preference for silty loam soils. *P. harmala, E. ceratoides* and *Z. eurypterum were* dominant at middle elevation. Axis 3 was dominated by C/N_1. *S. arbusculiformis* was sorted out well along this axis indicating that this species exhibited a preference for a low C/N_1. Axis 4 was dominated by $clay_{1,2}$ and EC_2. In addition, EC_1 had also a high correlation with this axis. High score of *A. aphylla* on axis 4 indicated that this species occurred on saline soils and preferred a heaver textured soil (Fig. 5). However, EC was appeared to be the most important driver of *A. aphylla* distribution. High score of *E. ceratoides* on axis 3 and its low score on axis 4 showed that the preference of this species for growing was the loamy sand soils with a high C/N_1 and low salinity. Low score of *I. songarica, Z. eurypterum* and *P. harmala* on axis 4 showed that these species were often found on non-saline soils. The highest accuracy of prediction was related to *S. arbusculliformis* with an almost prefect accordance (kappa = 0.85) between the predicted and actual distribution (Table 3). There was a substantial agreement for *A. aphylla* (0.79), *A. aucheri* (0.76), *E. strobilacea* (0.75) and *S. tomentosa* (0.68) model, a moderate agreement for *E. ceratoides* (0.54) and *Z. eurypterum* (0.52) and a fair agreement for *I. songarica* (0.40) model. The poorest accuracy of prediction was found for *P. harmala* (0.04). Overall, the prediction accuracy for *S. tomentosa, Z. eurypterum, E.*

ceratoides, E. strobilacea and I. songarica distribution was higher by BLR than by CCA, whereas the distribution of *A. aphylla, S. arbusculiformis and A. aucheri* was more accurately predicted by CCA. The only exception was *P. harmala* with the poorest accuracy of prediction by both BLR and CCA. However, BLR provided more accurate prediction than CCA for this species. *P. harmala* appeared following the omission of plant species under an intense grazing pressure by sheep mainly in the areas close to villages in Nodushan rangelands. This was in line with the statement of Guisan *et al.* (1999) that the species for which model predictions had the poorest accuracy were under disturbance. The soil factors were the most important drivers of the distribution and occurrence of 8 vegetation types in arid rangelands except *A. aucheri* type, a semiarid vegetation type, whose distribution was affected by elevation. Elevation (Hosseini *et al.,* 2013) and precipitation (Yaghmaei *et al.,* 2008) were addressed as important factors influencing the distribution of *A. aucheri* in a local scale. Although studies showed that the species, whose distribution was well modeled by BLR, was also well modeled by CCA (Guisan *et al.*, 1999; Chahouki *et al.,* 2010), and vica versa. In the present study the vegetation types whose distribution was well modeled by BLR was not well modeled by CCA. BLR provided a prefect agreement for 1 species, a substantial agreement for 7 species and a fair agreement for 1 species, whereas CCA provided a prediction with a perfect agreement for 1 species, a substantial agreement for 4, a moderate for 2 and a fair and slight for 1 species. It was observed that BLR might give better species-specific models than CCA. Presence-absence data of species were found to be effective for an accurate modeling of vegetation distribution. This makes it needless to measure time-consuming quantitative attributes of plants such as density and cover. Determination of species-environment relations using CCA enabled us to simultaneously assess the influence of a variety of factors on the distribution of the species. Although the cumulative proportion of the variance of the species data was low, the predictions by CCA model were reasonable. Studies showed that an ordination axes that explained only a low amount of variance may still be quite informative (Gauch, 1982).

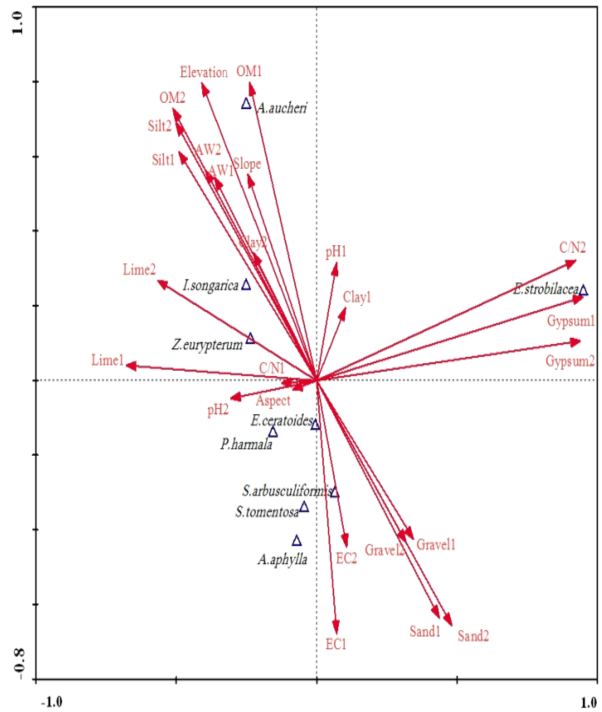

Fig 4. CCA biplots of the indicator plant species based on axis 1 versus axis 2. Species scores are linear combination scores

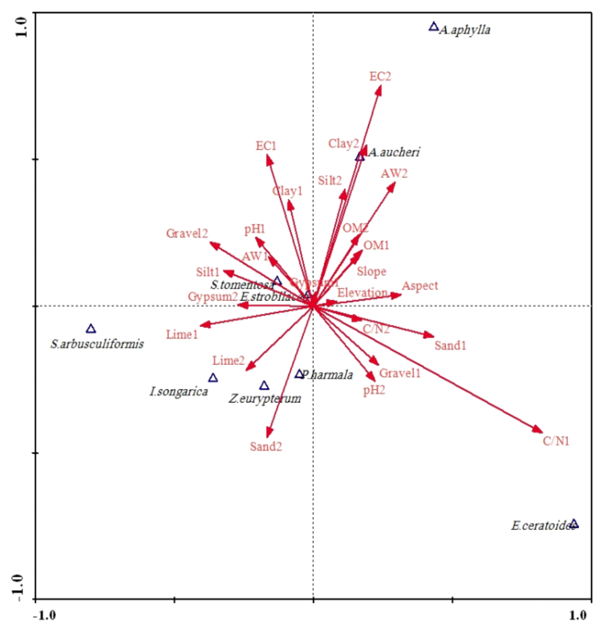

Fig 5. CCA biplots of the indicator plant species based on axis 3 versus axis 4. Species scores are linear combination scores

Conclusion

It was concluded that BLR models gave better predictions than CCA models for most of the study vegetation types because a species-specific subset of environmental predictors can be selected in BLR. Although a number of vegetation types were better modeled by CCA than by BLR, these vegetation types were well modeled by BLR too. Both BLR and CCA have limitations for modeling distribution of the vegetation types that are related to the grazing disturbance. However in such case, BLR is preferred over CCA. It is recommended that the BLR be used for modeling the distribution of vegetation types based on indicator species in arid and semiarid climates. Accurately prediction of species distribution enables an effective degraded land restoration planning and rangelands rehabilitation measures.

Acknowledgement

The author wishes to thank Yazd Branch, Islamic Azad University for its support of this research project

References

Aspinall, R.J. 2002. Use of logistic regression for validation of maps of the spatial distribution of vegetation species derived from high spatial resolution hyperspectral remotely sensed data. *Ecological Modelling* 157: 301-312.

Brunetto, G., M. Ventura, F. Scandellari, C.A. Ceretta, J. Kaminski, G.W. de Melo and M. Tagliavini. 2011. Nutrient release during the decomposition of mowed perennial ryegrass and white clover and its contribution to nitrogen nutrition of grapevine. *Nutrient Cycling in Agroecosystems* 90: 299-308.

Calef, M.P., A..D., McGuire, H.E. Epstein, T. Scott Rupp and H.H. Shugart. 2005. Analysis of vegetation distribution in Interior Alaska and sensitivity to climate change using a logistic regression approach. *Journal of Biogeography* 32: 863-878.

Chahouki, M.A.Z. and L.K. Ahvazi. 2012. Predicting potential distributions of Zygophyllum eurypterum by three modeling techniques (ENFA, ANN and logistic) in North East of Semnan, Iran. *Range Management and Agroforestry* 33: 123-128.

Chahouki, M.A.Z., L.K Ahvazi and H. Azarnivand. 2012. Comparison of three modeling approaches for predicting plant species distribution in mountainous scrub vegetation (Semnan Rangelands, Iran). *Polish Journal of Ecology* 60: 277-289.

Chahouki, M.Z., H. Azarnivand, M. Jafari and A. Tavili. 2010. Multivariate statistical methods as a tool for model-based prediction of vegetation types. *Russian Journal of Ecology* 41: 84-94.

Gauch, H.G. 1982. *Multivariate Analysis in Community Ecology*. Cambridge University Press, Cambridge.

Ghorbani, J., P. Dowlati and G. Heydari. 2015. Effects of floodwater spreading on the vegetation and soil in an arid rangeland. *Arid Land Research and Management* 29: 473-486

Guisan, A., Jr.T.C. Edwards and T. Hastie. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modeling* 157: 89-100.

Guisan, A., S.B. Weiss and A.D. Weiss. 1999. GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology* 143: 107-122.

Hosseini, S. Z., M. Kappas, M.A. Zare Chahouki, G. Gerold, S. Erasmi and A. Rafiei Emam. 2013. Modeling potential habitats for *Artemisia sieberi* and *Artemisia aucheri* in Poshtkouh area, central Iran using the maximum entropy model and geostatistics. *Ecological Informatics* 18: 61-68.

Landis J.R. and G.G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33: 159-174.

Liu, C., P.M. Berry, T.P. Dawson and R. G. Pearson. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28: 385-393.

Mousaei Sanjerehei, M. 2014. Assessment of spectral vegetation indices for estimating vegetation cover in arid and semiarid shrub lands. *Range Management and Agroforestry* 35: 91-100.

Mousaei Sanjerehei, M. 2014. Modeling the distribution of Iran life zones using CCA. *Indian Journal of Natural Sciences* 4: 1586-1596.

Mousaei Sanjerehei, M., M. Jafari, A. Mataji, N.B. Meybodi and M.R. Bihamta 2011. Facilitative and competitive interactions between plant species (an example from Nodushan rangelands, Iran). *Flora-Morphology, Distribution, Functional Ecology of Plants* 206: 631-637.

Ohmann, J.L. and M. J. Gregory. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, USA. *Canadian Journal of Forest Research* 32: 725-741.

Riordan, E.C. and P.W. Rundel. 2014. Land use compounds habitat losses under projected climate change in a threatened California ecosystem. *PLoS One* 9: 1-15.

Rodríguez-Castañeda, G., A.R. Hof, R. Jansson and L.E. Harding. 2012. Predicting the fate of biodiversity using species' distribution models: enhancing model comparability and repeatability. *PLoS One* 7: e44402.

Sahragard, H. P., M.A.Z. Chahouki and H. Gholami. 2015. Predictive distribution models for determination of optimal threshold of plant species in central Iran. *Range Management and Agroforestry* 36: 146-150.

Ter Braak, C. J. F. and P. Smilauer. 2002. *Canoco for Windows version 4.5*. Biometris–Plant Research International, Wageningen.

Ter Braak, C.J. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67: 1167-1179.

Ter Braak, C.J. 1987 The analysis of vegetation-environment relationships by canonical correspondence analysis. In: *Theory and Models in Vegetation Science*. Springer, Netherlands. pp. 69-77.

Ter Braak, C.J. and C. W. Looman. 1986. Weighted averaging, logistic regression and the Gaussian response model. *Vegetatio* 65: 3-11.

Yaghmaei, L., S. Soltani Koupaei and M. Khoda Gholami. 2008. Effect of climatic factors on distribution of *Artemisia sieberi* and *Artemisia aucheri* in isfahan province using multivariate statistical methods. *Water and Soil Science* 12: 359-371

Zarechahouki, M.A. and J. Esfanjani. 2015. Predicting potential distribution of plant species by modeling techniques in southern rangelands of Golestan, Iran. *Range Management and Agroforestry* 36: 66-71.

Zimmermann, N.E. and F. Kienast. 1999. Predictive mapping of alpine grasslands in Switzerland: species versus community approach. *Journal of Vegetation Science* 10: 469-482.