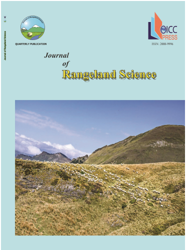

Volume 13, Issue 3, 132319 (1-11)

Journal of Rangeland Science (JRS)

https://dx.doi.org/10.57647/j.jrs.2023.1303.1577

Modeling potential habitats for Gymnocarpus decander using multivariate statistical methods and logistic regression (Case study: Sistan and Baluchestan Province)

Masome Narouei[1], Seyed Akbar Javadi[1]*, Morteza Khodagholi[2], Mohammad Jafari[3], Reza Azizinezhad[4]

[1]Rangeland Department, Faculty of Natural Resources and Environment, Science and Research Branch, Islamic Azad University, Tehran, Iran.
[2]Rangeland Research Division, Rangelands and Forests of Institute Research, Agricultural Research Extension Education Organization (AREEO), Tehran, Iran.
[3]Department of Reclamation of Arid and Mountainous Regions, Natural Resources Faculty, University of Tehran, Karaj, Iran.
[4]Department of Biotechnology and Plant Breeding, Faculty of Agricultural Sciences and Food Industries, Science and Research Branch, Islamic Azad University, Tehran, Iran.

***Corresponding author**: a.javadi@srbiau.ac.ir

Received 15 August 2021; Accepted 14 March 2022; Published online 4 July 2023

Abstract:
Identification of habitats and suitable environmental conditions for the presence of various animal and plant species is one of the key issues in ecology and environment modeling. In Iran, Gymnocarpus decander species is the main and most abundant plant of the Baluchi flora, especially in the rangeland areas. Due to its high resistance to harsh environmental conditions and salinity, this species can be useful for rangeland reclamation and restoration. In order to assess potential habitats of G. decander in Sistan and Baluchestan province, Iran, this research was conducted in May 2020 using multivariate statistical methods including factor analysis, clustering and logistic regression. To do so, the map of real habitats of G. decander, was modeled using altitude and 62 climatic variables which were prepared using 24 meteorological stations across the province and neighboring. Climatic variables were reduced to 5 factors using PCA, including temperature, precipitation, radiation, dust and maximum temperature, which accounted for 88.3% of the climate change in the studied region. According to the regression models, the precipitation, temperature and altitude variables were the most significant independent variables ($p < 0.05$) which influence the potential habitats of G. decander in the study area. Also, the study area was divided into 5 groups (G1:G5). The results suggest that models are more accurate in homogeneous groups than in the whole region. The Kappa coefficient and perdition accuracy had the highest and lowest values for G1 (68% and 88%) and G5 (19% and 59%), respectively. The results also showed that currently, 48.2% of the studied region has actual habitats. While the modeling results based on the model of the whole region showed that the potential habitat area of this species was 52.3% and based on the grouping models, it was 53.7%. Therefore, based on differences between actual data and modeling result, 4.8% of potential habitats have no G. decander. Also, results indicate that the G4 region with altitude range between 800 – 1250 m, growth season mean temperature was about 16° C and annual precipitation about 170 mm was the main habitat for G. decander in Sistan and Baluchestan. According to the results of this study, G. decander is an endangered species and it is necessary to plan long-term management to preserve and restore its habitats in the arid regions of southeastern Iran.

Keywords: Habitat distribution; Rangeland management; Climate change; Species modeling

1. Introduction

Rangelands are the most extent terrestrial ecosystems in Iran and play a key role in the production of forage, medicinal plants, soil protection and water storage. Therefore, the sustainable management and optimum utilization of the functioning and services of these ecosystems require the identification of their main characteristics and species [1]. On the other hand, in the selection and expansion of plant and animal species in each region, climatic parameters will be considered as one of the most important ecological factors of rangeland characteristics [2].

According to Kargar-Chigani et al. reports, a considerable proportion of Iran's territory is covered with arid and semi-arid rangelands and mismanagement and overexploitation of those rangelands have resulted in serious ecological degradation [3]. Therefore, the need to study the composition of existing species and relationships with environmental factors is essential to provide scientific references for species conservation and ecological restoration [3]. Based on their results, among environmental factors, altitude was the most effective factor to separate the plant groups while animal grazing was determined as the major cause of rangeland degradation. The ecological factors coupled with accessibility to moisture and climatic conditions are among the most important ecological factors affecting vegetation; therefore, the spatial distribution of plant communities have high correlation with climatic conditions [4]. In general, climatic factors directly or indirectly influence all factors that control the growth of plants [5]. From this perspective, climatic factors such as precipitation and number of rainy days, temporal and spatial distribution of precipitation, mean annual precipitation, precipitation type, mean annual temperature, temperature regime and topography will create interactions resulting in the formation of specific flora and fauna in each region [6]. On the other hand, in different regions, the contribution of climatic factors in the formation of species can be different. So far, several studies have been conducted to identify the role of climatic factors in the dominant species. For instance, Yaghmaei et al. determined the effects of climatic factors on Iranian range distribution of four most important vegetative types of Artemisia aucheri and Artemisia sieberi in Isfahan province [7]. The effect of factor analysis on these types shows that precipitation and temperature were the most influential factors in spatial distribution of Artemisia aucheri and Artemisia sieberi, respectively.

Saboohi and Khodagholi studied the climatic requirements of Bromus tomentellus and concluded that precipitation and cold season temperature were the most important factors affecting the distribution of this species in the rangelands of Isfahan province [8]. Indeed, these researcher's results showed that two factors of precipitation and cooling temperature were the most important factors for the presence of this species as dominant and its associated species. While the factors of wind and sunny hours had the least role for the presence or absence of this species. Also, the results of Khodagholi et al. showed that the six factors including heating temperature, spring and summer precipitation, wind, autumn-winter precipitation, and dusty and cloudiness days explained 37.32%, 22.54%, 7.18%, 6.6%, 4.22%, and 4.15% of data variation, respectively [9]. Together, these six factors account for 82% of data variation. The autumn-winter precipitation and heating temperature had the greatest impact on the presence of Artemisia sieberi and Artemisia aucheri, respectively so that the autumn-winter precipitation was negative in areas where Ar. sieberi was observed. The heating temperature factor is negative in areas where Ar. aucheri is present while it was positive in areas lacking Ar. aucheri. Pakzad et al. showed that rangelands of Isfahan province with an altitude of 2400 m which had approximately 50% relative humidity, with more than 400 mm of precipitation per year and relative radiation were considered as the main habitat for Astragalus adscendens [10].

As mentioned above, identifying the role of climatic factors in species distribution is important for many ecological researchers [11, 12]. One of the most effective techniques for predicting the geographical distribution of species is the use of multivariate statistical techniques such as generalized linear models (GLMs) and generalized additive models (GAM) [13]. In this regard, in a comprehensive review, Meynard and Quinn also reported that using GLM and GAM models could produce more efficient results than other models such as the classification tree [14].

In Ghazimoradi et al. study, the ability of GAM for mapping potential distribution of Ferula ovina Boiss, and the description of species response curves to environmental variables were examined in Fereidonshar region with area of 1000 square kilometers located in west of Isfahan province [15]. According to their results, the presence of F. ovina was negatively correlated with some environmental variables including silt and clay contents of soil and also, its distribution was positively correlated with slope, altitude, organic matter content, soil saturation percentage and annual temperature. According to their results, the presence of F. ovina is more likely in the habitats with mean annual temperature: $9-11$ C, slop: $25-50\%$, altitude: $1950-3000$ m (Above sea level), $CaCO_3$ content of soil: $10-30\%$, organic matter: $4-6\%$, silt: $10-30\%$ and soil saturation percentage: $45-60\%$. Therefore, using multivariate statistical techniques such as GAM could enable managers to identify appropriate areas for rangelands rehabilitation and protection practices. [15]. Sahragard et al. modeled the habitat suitability of wild almond (Amygdalus scoparia Spach) using three individual species distribution models (SDMs), i.e., back propagation artificial neural network (BP-ANN), maximum entropy (MaxEnt), and Generalized Linear Model (GLM) as well as the ensemble technique along with measuring the landscape metrics and analyzing the relationship between the distribution of the suitable habitat of the species in different landform classes in Fars Province, southern Iran [16]. Their results also showed the good performance of all three models on independent bioclimatic variables. Moreover, evaluation of the variable importance showed that the occurrence of A. scoparia strongly depend on the climatic variables, particularly isothermality (Bio 3), temperature seasonality (Bio 4), and precipitation of the driest quarter (Bio 17).

The present study aims to identify potential habitats of

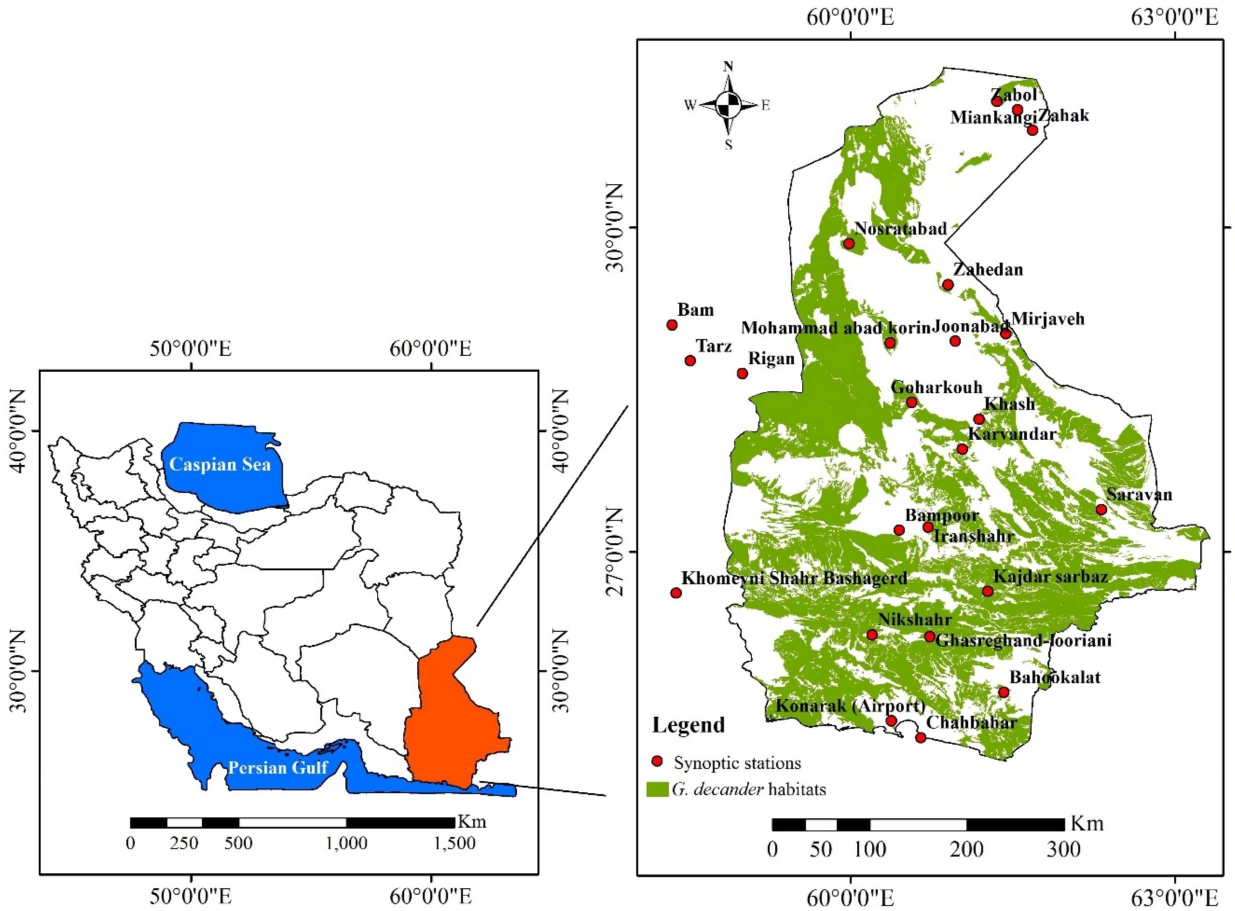

Figure 1. Distribution map of G. decander (green area) and the location of synoptic stations (red point).

the Gymnocarpus decander as one of the main rangeland species in southeastern Iran and to identify climatic factors affecting the distribution of this species. This species is one of the most important rangeland species in terms of livestock breeding and soil protection and also has an effective role in the life of local communities [17]. In this regard, identifying favorable climatic conditions and potential habitats of this species can be of great help to conservation planning and rangeland restoration plans not only in Sistan and Baluchestan rangelands but also in other arid and semi-arid regions. In addition, multivariate statistical techniques are used and the modeling method used in the present study was mostly used to identify the distribution of animal and tree species while the present study emphasizes the importance of using these techniques on rangeland species, especially in areas without data and arid and semi-arid regions. The aim of this study was to model the potential habitats of G. decander in Sistan and Baluchestan as well as modeling the distribution of this plant species in the province by considering climate variable and altitude factor. Advanced statistical methods including GLM, dimension reduction (i.e., PCA) and clustering analyses were used to identify potential areas and the optimal amount of climatic variables for the presence of this plant species.

2. Materials and methods

The study area covers an arid region located in the southeastern Iran with an area of 18.75 m hectares. It is located between latitudes 25°40′ N and longitudes 44°64′ E. This province is one of the most arid provinces in Iran in terms of low precipitation and dry climate. The rangeland areas of this province are not in a good condition due to low precipitation and limited soil.

The present ranges for G. decander and the location of the meteorological stations used in this study are presented in Fig. 1. The Gymnocarpus decander Forssk belongs to the subfamily of Paronychioidae and the family Caryophyllaceae [18].

In Iran, this species is the main and most abundant plant of the Baluchiflora of the country, especially in the provinces of Hormozgan, Bushehr, Sistan and Baluchestan and south of Kerman [19]. Also, due to the high resistance of G. decander to harsh environmental conditions such as high levels of soil salinity, drought, physical and chemical stresses in a wide range of natural habitats and salinity, this species can also be useful for land reclamation degraded rangelands. In addition, this species has an important role in terms of soil protection and production of palatable forage [20, 21].

To identify the potential habitats of G. decander, 62 climatic variables from 24 meteorological stations over Sistan

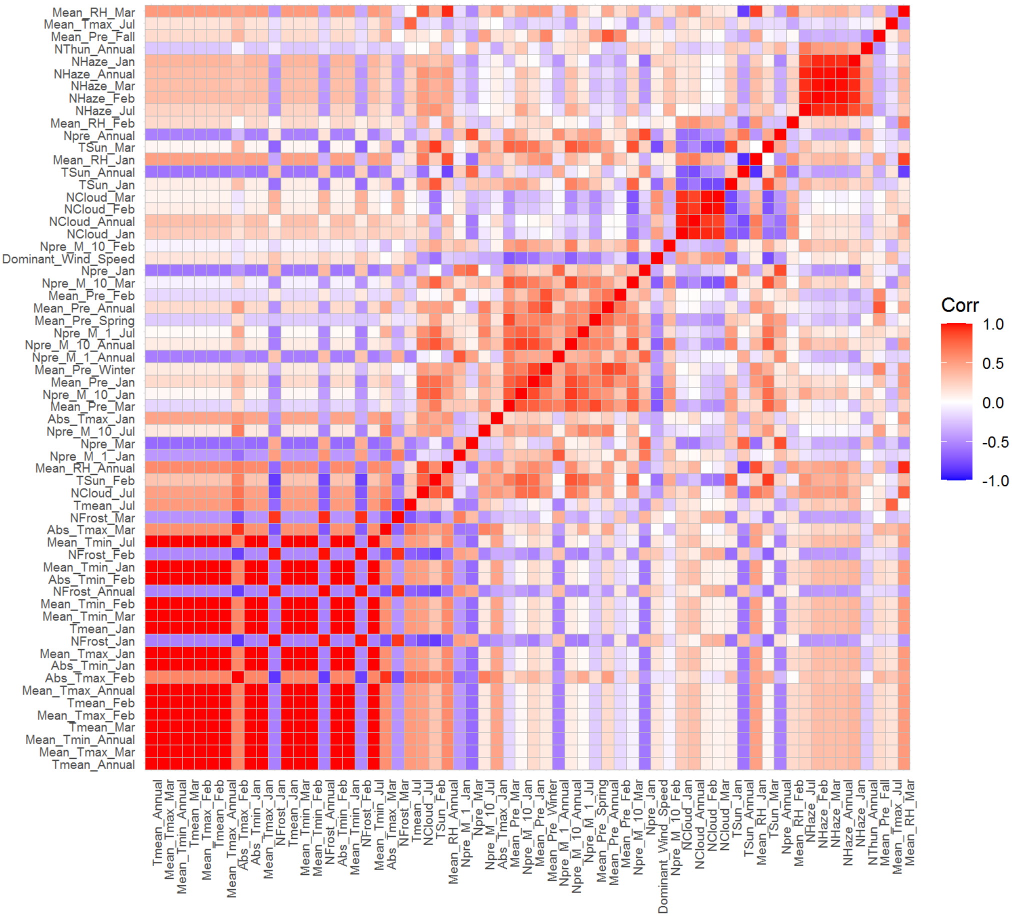

Figure 2. Correlation matrix between climatic variables.

and Baluchestan province and neighboring provinces were obtained from the Iran Meteorological Organization (Statistical period for obtaining meteorological data from the time of meteorological station formation until 2019) (Table 1). In addition, vegetation distribution maps were obtained from the National Forest and Rangeland Research Institute for the distribution of G. decander (Fig. 1). After completing and compiling climatic information, kriging and inverse weight distance (IDW) methods were used to increase the sample size and generalize the point-to-zone results; each of the

Table 1. Type and number of climatic variables.

Variable	Number
Temperature	20
Precipitation and rainy days	18
Sunny hours and cloudy days	9
Dust and thunderstorm days	6
days with frost	4
Relative humidity	4
Wind	1
Total	**64**

climatic variables was zoned in ArcMap v10.4.1. Finally, a network with 1071 cells and dimensions of 13×13 km was provided for statistical calculations. In order to reduce the dimensions of climate data, exploratory factor analysis by PCA and varimax rotation was used [22]. In this regard, to perform the factor analysis process, the first step is to prepare a correlation matrix between the variables. For this purpose, Spearman's correlation coefficients between each pair of study independent variables were computed and the correlation matrix visualized in Fig. 2 using graphical R package "ggcorrplot". Then, the preconditions of the KMO index and Bartlett test before performing the work were evaluated for valid factor analysis. The KMO index was used to evaluate the adequacy of the sample size and values greater than 0.7 were suitable for this index, but values of 0.5 to 0.7 could be used with caution [23]. The Bartlett test was also used to measure the identity of the correlation matrix of variables [22]. After performing factor analysis to increase the prediction accuracy of regression models for the presence of the studied species, k-means clustering was used to determine homogeneous regions [24]. To determine the appropriate number of clusters, the statistic of the total within sum of squares (t.wss) for different number of clusters were tested.

To predict the probability of the presence of G. decander under the condition of independent climatic variables and altitude, a multivariate logistic regression was used to identify

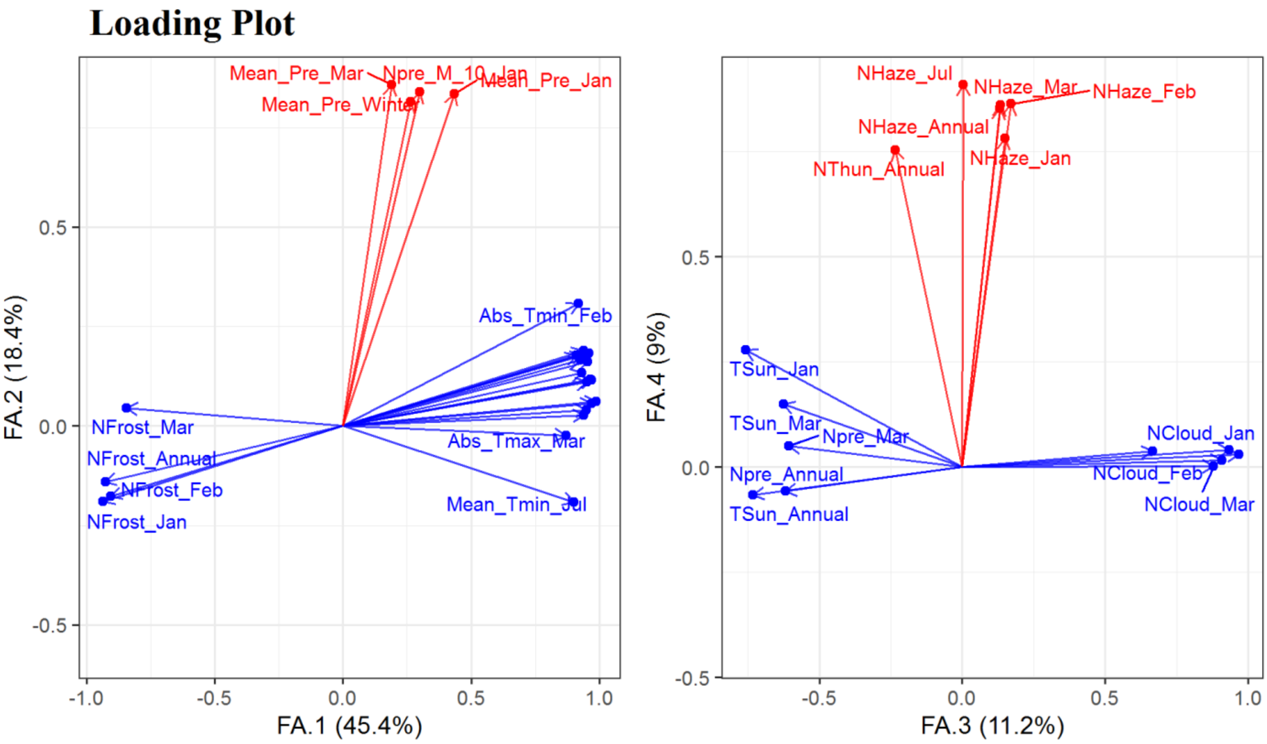

Figure 3. Displays the factor loads of the variables on the first to fourth factors. The red and blue colors represent the variables with high load factors on the factor located on the vertical and horizontal axis of each graph, respectively.

potential areas and the optimal amount of climatic variables for the presence of the plant species. In this type of regression, the dependent variable is a binary data for presence, taking values 0 for absence and 1 for presence. Thus, the probability of presence and absence of the species based on the equation (1) can be calculated under the condition of independent variables [25].

$$\mathrm{p}(Y=1|X_i) = \log it(\mathrm{p}(x)) = \frac{e^{(\alpha+\beta_1X_1+\beta_2X_2+\ldots+\beta_nX_n)}}{1+e^{(\alpha+\beta_1X_1+\beta_2X_2+\ldots+\beta_nX_n)}} \tag{1}$$

Where $\mathrm{p}(x) = \mathrm{p}(Y=1|X_i)$ is the probability of the presence of the dependent variable (here, the presence of G. decander) for the independent variables X_i, α and β_i are also the coefficients of the above equation, respectively. Also, to select the best independent and effective variables, the stepwise regression technique (Round back method) was used to delete the ineffective variables in the presence of G. decander. In addition, the goodness of fit statistics including Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were used to evaluate the performance of the models.

The logistic regression model's validation was assessed using two commonly used statistics, i.e., the kappa coefficient (κ) and prediction accuracy statistics, based on observed and fitted values in each cell of the study grid. The kappa coefficient is a statistic that is used to measure inter-rater reliability for categorical items and takes into account the possibility of the agreement occurring by chance [26]. While prediction accuracy statistics refer to the percentage of agreement between observed and fitted values. After confirming regression models, to survey potential habitats of G. decander the probability of the presence of G. decander was computed and interpolated using IDW method. Afterward, the interpolated maps were classified into two groups, presence for probability > 0.5 and absence for probability < 0.5.

3. Results

For the correlation matrix, the KMO index was equal to 0.53 and the values of chi-squared of Bartlett test were 188150 ($p < 0.01$), respectively. Therefore, the preconditions of factor analysis for climatic variables were established. Also, after factor analysis, 5 factors were identified and explained a total of 88.3% of the total variance of the data (Table 2). In this regard, the names of the factors were selected based on the high factor loads of the variables in each factor. Fig. 3 shows the variables with factor loads greater than 0.6 on the first to fourth factors. In addition, for the fifth factor, only the mean maximum temperature variable was identified in July with a factor load of -0.83, which is not shown in Fig. 3. The WSS value for different numbers of clusters is present in Fig. 4. As it is clear, the value of WSS statistic was minimized about 5 clusters and then fixed (Fig. 4). So for segregated clustering based on t.wss statistics, five clusters were found to be suitable for data grouping. The characteristics of each group are also determined from the box-plots of the factor scores. The first group is relatively mountainous areas with low hours of sunshine, high cloudy and low precipitation, the second group includes highlands with low temperatures and moderate precipitation. The third and fourth groups are the intermediate between the second

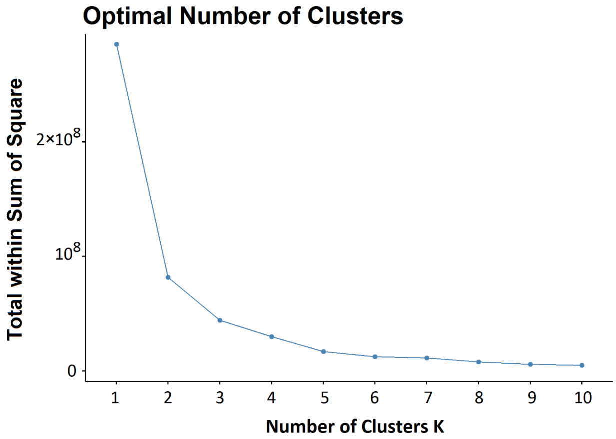

Figure 4. WSS statistics values for per number of different clusters.

and fifth groups in terms of precipitation, sunshine hours, dust and temperature, with the difference that precipitation and temperature in the fourth group was more than the third group. While the fifth group will include coastal low-lying areas with high temperatures and precipitation (Fig. 5).
The probability of the presence of G. decander relative to the changes of each climatic factor and altitude variable by the univariate logistic regression model diagram for each group and the whole region is presented in Fig. 6. For the temperature factor (FA1), in exchange for increasing the factor score, the probability of the presence of G. decander in the case of the whole region is considered (black line) and groups 1, 2 and 4 (red, blue and purple lines) increase. While in groups 3 and 5 (green and orange lines) with increasing temperature factor scores, the probability of the presence of G. decander will decrease.
For the precipitation factor (FA2), in exchange for increasing the factor score, the probability of the presence of G. decander increases in the case that the whole area (black line) and groups 1 and 4 (red and purple lines). While in groups 3 and 2 (green and blue lines) with increasing precipitation factor scores, the probability of the presence of G. decander will decrease and in group 5 (orange line) with a very small slope, increasing or decreasing precipitation factor scores has little effect on the presence of G. decander.
For the factor of the number of cloudy days (FA3), in exchange for increasing the factor score, the probability of the presence of G. decander decreases in the case of the whole region (black line) and groups 1 and 5 (red and orange lines). While in group 3 (green line) with increasing the factor scores of the number of cloudy days, the probability of the presence of G. decander will increase and in group 4 (purple line) with a very slight slope, increasing or decreasing the factor scores of the number of cloudy days has little effect on the presence of G. decander. The changes of the third factor (cloudy days and sunny hours) have no effect on the presence or absence of the G. decander in the second group (blue line with zero slopes).
For the factor of the number of dust days (FA4), in exchange for increasing the factor score, the probability of the presence of G. decander increases in the case of the whole region (black line) and groups 1 and 5 (red and orange lines). While in groups 2, 3 and 4 (blue, green and purple lines), with increasing the factor score of the number of dust days, the probability of the presence of G. decander will decrease.
For the average maximum temperature factor in the warm season (FA5), in exchange for increasing the factor score, the probability of the presence of G. decander in the whole region (black line) and groups 1, 3 and 4 (red, green and purple lines) increases. While in groups 2 and 5 (blue and orange lines) with increasing the factors of the average maximum temperature in the warm season, the probability of the presence of G. decander species will decrease. Also, in exchange for increasing altitude, the probability of the presence of G. decander in the case of the whole area (black line) and group 2 (blue line) decreases. While in exchange for increasing the altitude in groups 1, 3, 4 and 5 (red, green, purple and orange lines), the probability of the presence of G. decander in these areas will increase.
Stepwise regression was used to identify the best linear combination of climatic variables for modeling and predicting habitats of G. decander. The results of stepwise regression for each group and in the whole region without grouping along with AIC, BIC, kappa coefficient and predictive accuracy of each model are presented in Table 3. In general, the results showed that the most accurate model for predicting the habitat of G. decander is the model estimated in the first group. Kappa coefficient and prediction accuracy statistics for this model are 68.43 and 88.16%, respectively, which is more than other models. In contrast, the weakest model with the lowest values of kappa coefficient and prediction accuracy (18.8 and 59.31, respectively) belongs to the fifth group and the coastal area. For the general state of the data

Table 2. Name of factors, eigenvalue, variance and cumulative variance of each factor.
(The + and - signs indicate the factor load of each variable and the type of its relationship with the factor scores of each factor.)

Factors	Factor name	Abbreviation	Eigenvalue	Variance	Cum.var
FA1	Temperature factor (+) and number of days with frost (-)	T (+)_NF (-)	28.16	45.4	45.4
FA2	Precipitation (+) and number of rainy days (+)	P (+)_NP (+)	11.42	18.4	63.8
FA3	Cloudy days (+) and sunny hours (-)	NC (+_NS (-)	6.97	11.2	75.1
FA4	Days with dust (+) and days with thunder (+)	ND (+)_NT (+)	5.6	9.0	84.1
FA5	The average maximum temperature in the hot season (-)	Tmax (-)	4.2	4.2	88.3

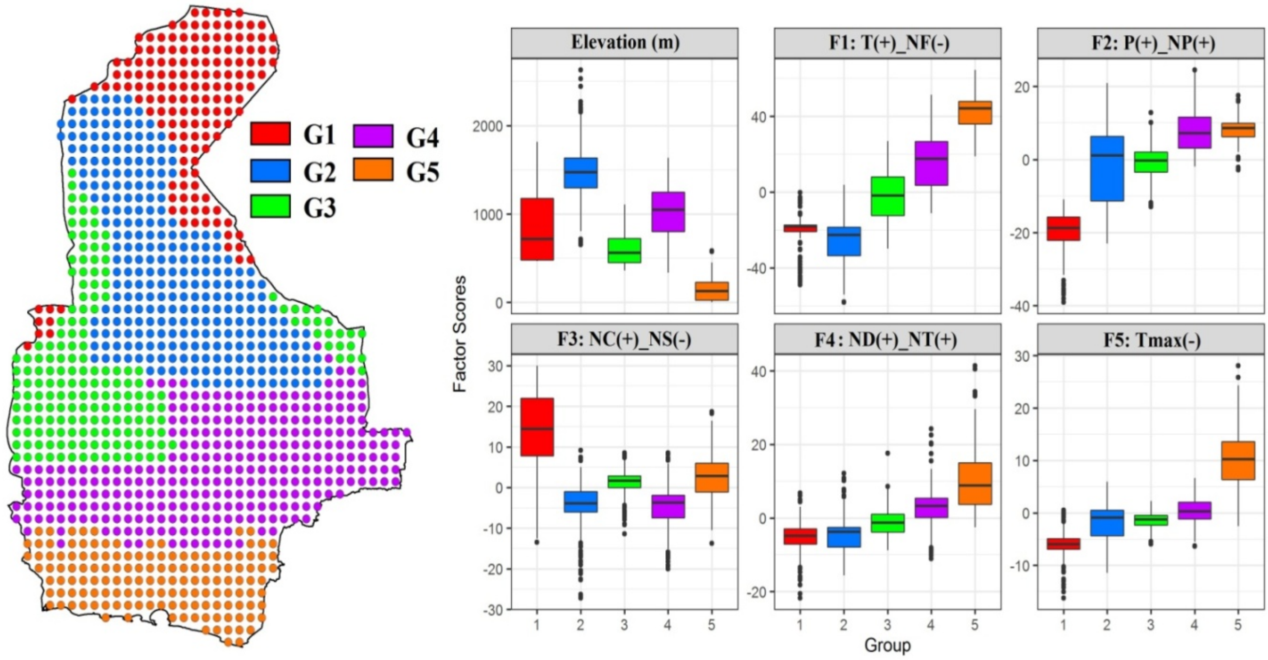

Figure 5. Displays the location of the grouped cells and the box-plots of factor scores in each group.

and regardless of grouping, the proposed model included temperature, precipitation, sunny hours and cloudy days, altitude factors and maximum temperature in the growing season, Kappa coefficient and prediction accuracy for this model were calculated to be 23.93 and 61.9%, respectively. By determining the regression models and applying them to the raster maps resulting from the zoning of the factors, the species presence probability maps, or in other words, the maps of potential habitats for the presence of the G. decander were prepared. According to the models, two types of maps were prepared for the presence of G. decander; One map for the overall data model, regardless of the grouping of the area (Fig. 7(a) and (b)) and the other for the juxtaposition of the five maps obtained from the five estimated models within each of the groups (Fig. 7(c) and (d)). Also, considering the presence of the species for a probability of more than 50% and its absence for a probability of less than 50%, the probability zoning maps were classified into two groups: the presence and absence of the species (Fig. 7(b) and (d)).

By determining the regression models and applying them to the raster maps resulting from the zoning of the factors, the species presence probability maps, or in other words, the maps of potential habitats for the presence of the G. decander were prepared. According to the models, two types of maps were prepared for the presence of G. decander; One map for the overall data model, regardless of the grouping of the area (Fig. 7(a) and (b)) and the other for the juxtaposition of the five maps obtained from the five estimated models within each of the groups (Fig. 7(c) and (d)). Also, considering the presence of the species for a probability of more than 50% and its absence for a probability of less than 50%, the probability zoning maps were classified into two groups: the presence and absence of the species (Fig. 7(b) and (d)).

4. Discussion

The current study showed that the presence of G. decander could be reasonably justified by altitude and appropriate climatic variables. For example, zoned maps of the presence of G. decander (Fig. 7) for both modeling modes showed that in general, the presence of G. decander in the southern half of Sistan and Baluchestan province, except coastal areas was more than other parts of the province. According to Fig. 6, 48.17% of the entire Sistan and Baluchestan province has G. decander, with the highest share of G. decander species in the fourth group (18.39%). The average scores of temperature and precipitation factors for this range are 17.72 and 7.21, respectively (Fig. 5). In terms of real climatic variables, the average temperature of the growing season (February and March) and annual precipitation in this range are about 16 C and 170 mm, respectively. After that, the highest presence of G. decander belongs to the second group with 9.71% coverage. By comparing the range of the fourth group with the high range of the second group (Fig. 5), it can be seen that with increasing altitude and decreasing temperature, the probability of the presence of G. decander will be greatly reduced and low temperature will be a climatic limiting factor. In the third group, the presence of G. decander is 9.43% which does not decrease significantly as compared to the second group and this small difference is due to the increase in the number of cloudy days as compared to the second group (Fig. 5). In the fifth groups, the presence of 7% species is observed (Fig. 7). According to Fig. 5, it can be seen that there was a slight difference in terms of precipitation between the fourth groups and the coastal group (G5); therefore, temperature and altitude distinguished these two areas in terms of potential conditions for the presence of the species. In the coastal

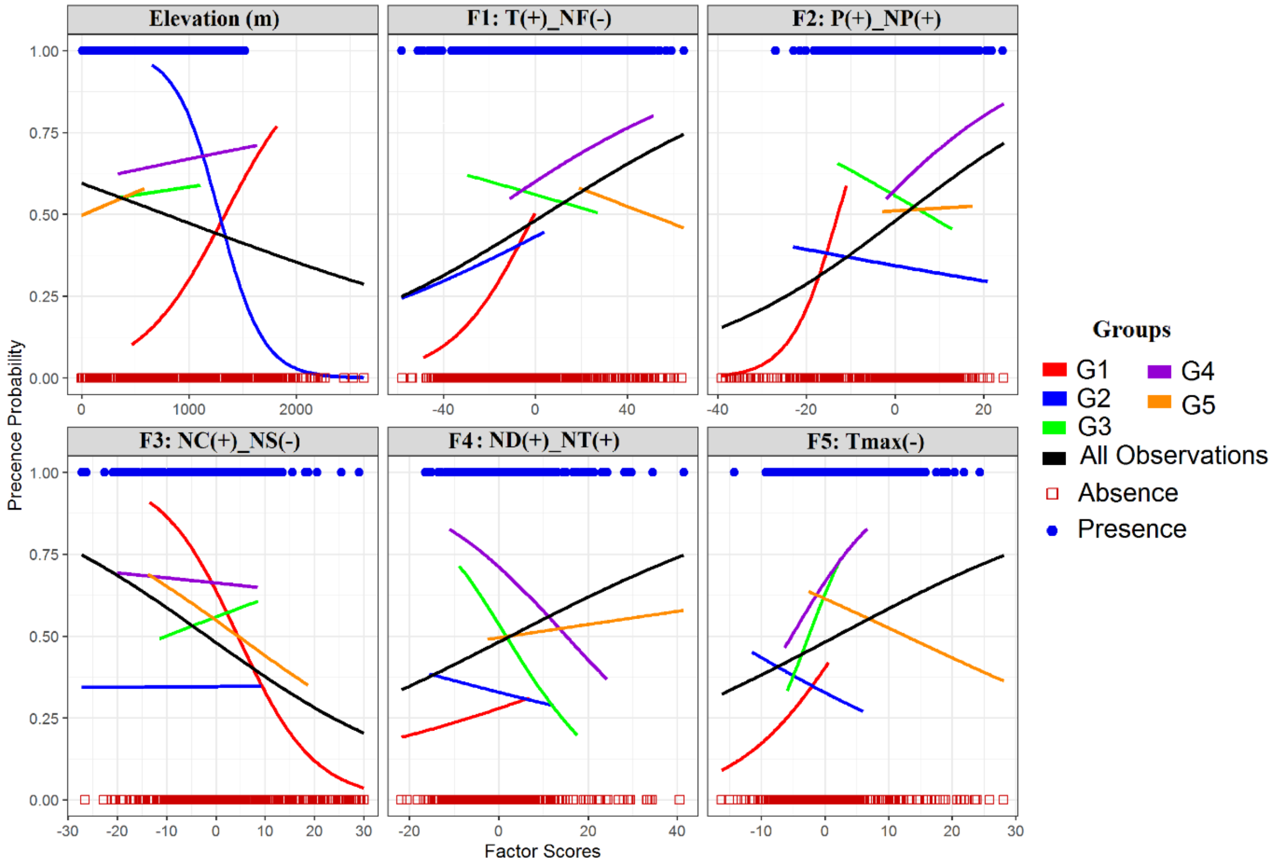

Figure 6. Univariate logistic regression of the presence of G. decander for calculated factors. The black line denotes logistic model for whole region and red, blue, green, purple and orange lines are referring to logistic models for each of G1, G2, G3, G4 and G5 regions, respectively.

area, despite the heavy precipitation, the temperature factor scores for this area are very high and high temperature will be considered as a limiting factor. In the first group, we see the lowest presence of the species with a coverage of 3.6% because as compared to the other groups, the average factor scores of temperature and precipitation are much lower and increasing the factor score of the third factor (number of cloudy days) significantly reduces the presence of G. decander (Fig. 5). The reason that the temperature factor is a limiting factor for G. decander can be related to the physiological characteristics of this species because this species is resistant against drought and water shortage conditions (Vali and Ghazavi); Therefore, the response of this species to temperature changes is more severe than precipitation changes [20]. In general, rangeland plant species in arid and semi-arid regions due to high tolerance range and adaptation

Table 3. The best fitted regression models for presence G. decander.

Groups	Stepwise logistic regression equation	AIC, BIC	Kappa Coefficient	Prediction accuracy %
G1	$y=\frac{1}{1+\exp[-(0.174^{***}(F_1)+0.00525^{***}(Elev)-2.394^{***})]}$	118.0, 127.1	68.43	88.16
G2	$y=\frac{1}{1+\exp[-(0.035^{*}(F_2)-0.0053^{***}(Elev)+6.922^{***})]}$	292.4, 303.5	36.00	73.09
G3	$y=\frac{1}{1+\exp[-(-0.2187^{***}(F_2)+0.7813^{***}(F_5)+1.076^{***})]}$	231.1, 240.7	28.72	65.92
G4	$y=\frac{1}{1+\exp[-(0.200^{***}(F_5)+0.0012^{*}(Elev)-0.5474)]}$	366.9, 377.9	28.72	68.71
G5	$y=\frac{1}{1+\exp[-(-0.186^{*}(F_2)-0.133^{**}(F_3)+1.944^{*})]}$	199.33, 208.27	18.80	59.31
whole area	$y=\frac{1}{1+\exp[-(0.011^{**}(F_1)+0.029^{**}(F_2)-0.037^{***}(F_3)-0.047^{*}(F_5)-0.00056^{**}(Elev)+0.0434^{*})]}$	1390.1, 1419.9	23.93	61.90

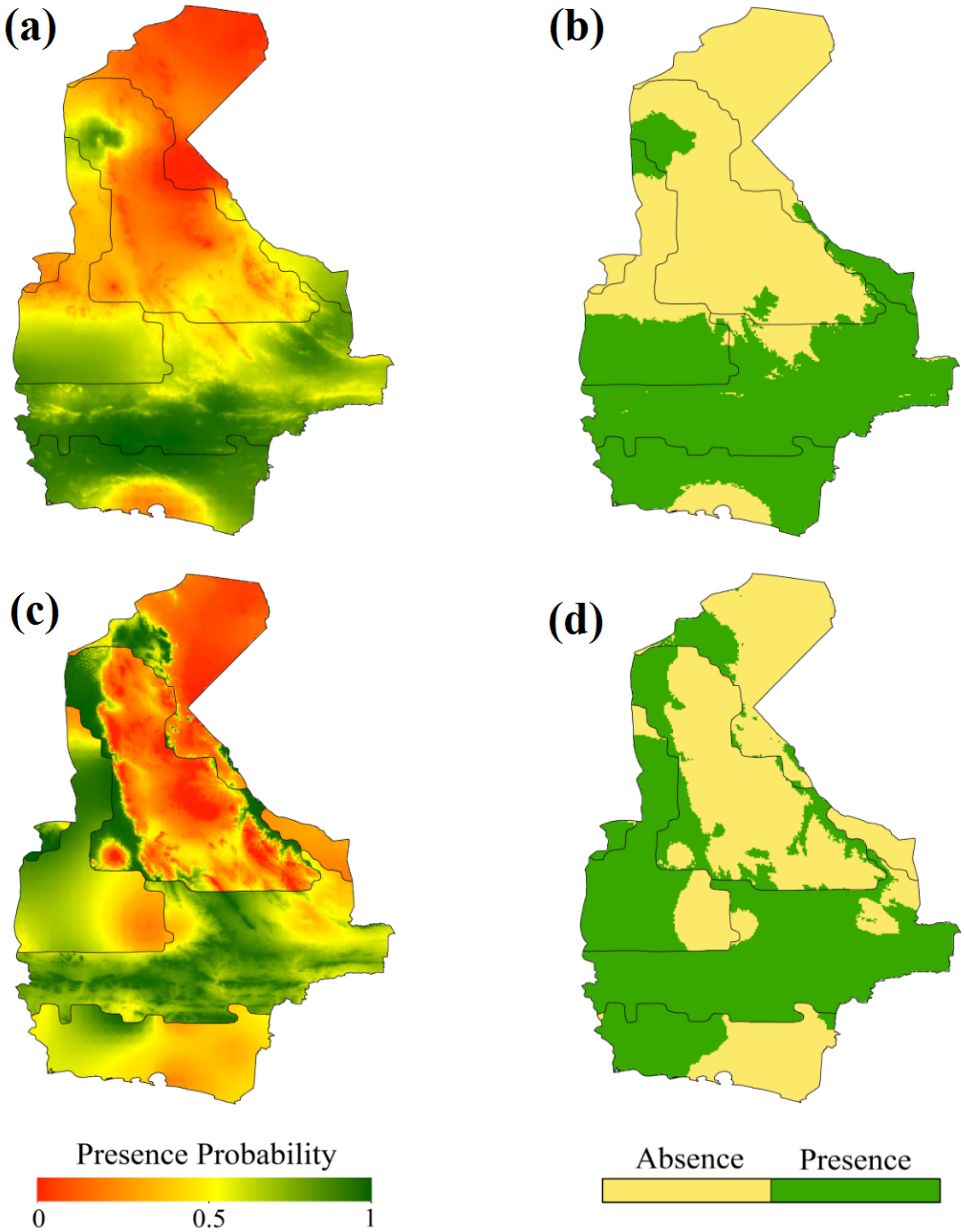

Figure 7. (a) and (c) are the zoning maps of the probability of the presence of G. decander under the condition of bioclimatic variables in the base period, respectively. (b) and (d), classify (a) and (c) maps for a probability of more than 50% or presence (green) and a probability of less than 50% or absence (cream), respectively. The lines on the map indicate grouping ranges.

to drought conditions are mainly more restricted by temperature factors. Results of some similar studies on other rangeland species such as Yaghmaei et al. for Artemisia sieberi, Arvin et al. for Thymus daenensis and Samadi et al. for Nepetacataria also indicate the predominance of temperature factors to precipitation on the distribution of rangeland species in arid and semi-arid regions [7, 27, 28]. In addition, modeling results (Table 3) showed that the fourth factor or variable of dust and thunderstorm days did not have much effect on the presence of G. decander so that in none of the models, the fourth factor of the present study was not significant. Also, in Fig. 6, the slope of univariate graphs for dust factor was severe and inverse only in the third and fourth groups (purple and green lines), which can be partly due to the effect of sandstorms on the distribution and presence of G. decander in the southwestern desert area knowing these two groups.
On the other hand, the altitude and climatic characteristics of the fourth group can be considered as the optimal range in terms of habitat suitability for the G. decander. This range has a height between 800 to 1250 m with an average of 1050 m. The average scores of temperature and precipitation factors for this range are 17.72 and 7.21, respectively. In terms of real climatic variables, the average temperature

of the growing season (February and March) and annual precipitation in this range are about 16 C and 170 mm, respectively. These characteristics are almost consistent with the reported habitat conditions in Karampooret al. [21].
In general, the results showed that currently, 48.2% of the region has actual habitats (area of green poligons in Fig. 1). While the modeling results show that the percentage of the potential habitat area is 52.3% based on the model of the whole region (area of green poligons in Fig. 7(b)) and 53.7% based on the grouping models (area of green poligons in Fig. 7(d)). Therefore, based on differences between aforementioned numbers on average around 4.8% of potential habitats have no G. decander. It was generally normal to differentiate between potential and actual habitats of species. However, Yousefi et al. reported that climate change has adverse effects on the potential habitats of 81% of the animal and plant species studied in Iran since 2014 [29]. Therefore, due to climate change conditions, the difference between potential and actual habitats estimated in the present study for the G. decander may not be fixed. In this respect, Abolmaali et al. suggest the use of models prepared from species distribution to predict potential habitat habitats under different climate change models and scenarios [30].
From the point of view of the technique used to model the presence of G. decander, the results showed that the use of clustering and modeling process within clusters will greatly increase the accuracy of the models; As shown in Fig. 7 zoning maps, the probability of the presence of G. decander in the grouping mode was well able to obtain the pattern of changes in vegetation maps and field data compared to modeling in general.

5. Conclusion

In the present study, due to the great importance of G. decander in terms of high protein forage production, its role in the life of local communities as well as the ability to form plant types in harsh climatic conditions and the role of soil protection, this species was selected as the target species. Unfortunately, the present study showed that about 4.8% of potential habitats in the region do not have this species. This may be due in part to climate change, human activities, and species interactions that require more detailed study. It is normal for differences between potential and actual habitats of species. But when this difference is dynamic and increasing, there is a risk of limiting the extinction of the species. Therefore, continuous monitoring of factors affecting the distribution of species as well as determining their potential and actual habitats over time for long-term planning for the conservation and development of rangeland species, especially in sensitive arid and semi-arid ecosystems will be critical. Finally, according to what was expressed, it can be stated that the use of multivariate scientific and statistical techniques and modeling to identify potential habitats of rangeland species, especially in rangeland ecosystems of arid and semi-arid regions where the distribution of species is strongly influenced by climatic factors is accurate. It is suitable. Therefore, the use of these techniques and methods of modeling the distribution of species can provide useful information for rangeland improvement and rehabilitation projects by identifying and introducing areas with high potential.

Acknowledgement:
The data of current paper has been extracted from the study area in Iran; and the study held under the auspices of the respected council; in this regard, the authors are truly grateful.

Conflict of interest statement:
The authors declare that they have no conflict of interest.

References

[1] H. Azarnivand, M. Jafari, M. R. Moghaddam, A. Jalili, and C. M. Zare. "The effects of soil characteristics and elevation on distribution of two Artemisia species (case study: VardAvard, Garmsar and Semnan rangelands)". *Iranian Journal of Natural Resources*, **56**:93–100, 2003.

[2] A. Yousefi and D. Yarahmadi. "Evaluation of the relation between climatic elements and rangelands species Hammada salicornica". *Academic Journal of Interdisciplinary Studies*, **4**:265–265, 2015.

[3] H. Kargar-Chigani, S. A. Javadi, G. Zahedi-Amiri, S. J. Khajeddin, and M. Jafari. "Vegetation composition differentiation and species-environment relationships in the northern part of Isfahan Province Iran". *Journal of Arid Land*, **9**:161–175, 2017.

[4] B. Sugier. Soil vegetation atmosphere. :107, 1996.

[5] R. Khatibi, S. Soltani, and M. Khodagholi. "Effects of climatic factors and soil salinity on the distribution of vegetation types containing Anabasis aphylla in Iran: a multivariate factor analysis". *Arabian Journal of Geosciences*, **10**:36, 2017.

[6] J. Moghimi. *Introduction some of important rangeland species for improvement of Iranian rangelands*. Arvan press, 1th edition, 2005.

[7] L. Yaghmaei, K. S. Soltani, and G. M. Khoda. "Effect of climatic factors on distribution of Artemisia sieberi and Artemisia aucheri in Isfahan Province using multivariate statistical methods". *Journal of Soil and Water Sciences*, **12**:359–371, 2008.

[8] R. Saboohi and M. Khodagholi. "Studying the acclimation of Bromus tomentellus in Esfahan Province". *Iranian Journal of Applied Ecology*, **2**:57–72, 2013.

[9] M. Khodagholi, R. Saboohi, and E. Z. Esfahani. "A Comparison of Vegetative Climate of Artemisia sieberi Besser and Artemisia aucheri Boiss in Iran". *Theoretical and Applied Climatology*, **146**:921–939, 2021.

[10] Z. Pakzad, M. RaeiniSarjaz, and M. Khodagholi. "Evaluation the effects of climate factors on distribution of the habitats of Astragalus adscendens in Isfahan province". *Iranian Journal of Range and Desert Research*, **20**:199–212, 2013.

[11] W. Hallgren, F. Santana, S. Low-Choy, Y. Zhao, and B. Mackey. "Species distribution models can be highly sensitive to algorithm configuration". *Ecological Modeling*, **3**:108719, 2019.

[12] A. Norberg, N. Abrego, F. G. Blanchet, F. R. Adler, B. J. Anderson, J. Anttila, M. B. Araujo, T. Dallas, D. Dunson, J. Elith, S. D. Foster, R. Fox, J. Franklin, W. Godsoe, A. Guisan, B. ÓHara, N. A. Hill, R. D. Holt, F. K. C. Hui, M. Husby, J. A. Kalas, A. Lehikoinen, M. Luoto, H. K. Mod, G. Newell, I. Renner, T. Roslin, J. Soininen, W. Thuiller, J. Vanhat-alo, D. Warton, M. White, N. E. Zimmermann, D. Gravel, and O. Ovaskainen. "A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels". *Ecological Monographs*, **89**:e01370, 2019.

[13] T. J. Hastie and R. J. Tibshirani. *Generalized additive models*. CRC press, 1th edition, 1990.

[14] C. N. Meynard and J.F. Quinn. "Predicting species distributions: a critical comparison of the most common statistical models using artificial species". *Journal of Biogeography*, **34**:1455–1469, 2007.

[15] M. Ghazimoradi, M. Tarkesh, H. Bashari, and M. R. Vahabi. "Determining the potential habitat of Ferula ovina (Boiss) using Generalized Additive Model in Fereidonshar Rangelands-Isfahan". *Journal of Range and Watershed Management*, **69**:677–689, 2016.

[16] H. P. Sahragard, M. Ajorlo, and P. Karami. "Landscape structure and suitable habitat analysis for effective restoration planning in semi-arid mountain forests". *Ecological Processes*, **10**:1–13, 2021.

[17] M. A. Naseri Baziari, M. Pichand, K. Najafi, and T. Shabankareh. "Evaluation of forage quality of Gymnocarpus decander in different phenological stages in Hormozgan province (Case study: Rangelands of Rudan city)". *Journal of Plant Ecophysiology*, **10**:246–254, 2018.

[18] A. Ghahraman, A. Naghinezhad, and F. Atar. "Habitats and flora of the Chamkhaleh-Jirbagh coastline and Amirkelayeh wetland". *Journal of Environmental Science*, **30**:46–67, 2004.

[19] V. Mozaffarian. "Studies on the flora of Iran, four new species and a short note on an interesting Rubiaceae". *The Iranian Journal of Botany*, **12**:107–113, 2006.

[20] A. Vali and G. R.Ghazavi. "The relationship between plant density and soil salinity and texture in Korsiah saline area in Darab region". *Desert Magazine*, **8**:236–248, 2003.

[21] M. Karampoor, A. Yousefi, and N. Koohpaye. "Relationship between climatic elements with vegetation cover of meadows in the Hormozgan province (a case study: Gymnocarpus decander)". *Iranian Natural Ecosystems*, **6**:41–48, 2015.

[22] M. Auerswald and M. Moshagen. "How to determine the number of factors to retain in exploratory factor analysis: A comparison of extraction methods under realistic conditions". *Psychological Methods*, **24**:468, 2019.

[23] P. Kline. *An easy guide to factor analysis*. Routledge, 1th edition, 2014.

[24] J. Xu and K. Lange. "Power k-means clustering". *Proceedings of the 36th International Conference on Machine Learning, PMLR*, **97**:6921–6931, 2019.

[25] A. Agresti. *An introduction to categorical data analysis*. Wiley, 2th edition, 2007.

[26] M. L. McHugh. "Interrater reliability: the kappa statistic". *Biochemiamedica*, **22**:276–282, 2012.

[27] A. A. Arvin, M. Khodagholi, and S. Moazeni. "Investigation of the bio-climatic needs of Thymus daenensis Celak: The case of Isfahan Province". *Journal of Range and Watershed Management*, **73**:257–272, 2020.

[28] V. Samadi, M. Khodagholi, and A. Shaemi. "Evaluation the effect of climatic factors on the development of medicinal species Nepeta cataria habitats in Northwestern of Iran". *Journal of Climate Research*, **37**:57–68, 2020.

[29] M. Yousefi, A. Kafash, N. Valizadegan, S. S. Ilanloo, M. Rajabizadeh, S. Malekoutikhah, S. S. H. Yousefkhani, and S. Ashrafi. "Climate change is a major problem for biodiversity conservation: A systematic review of recent studies in Iran". *Contemporary Problems of Ecology*, **12**:394–403, 2019.

[30] S. M. R. Abolmaali, M. Tarkesha, and H. Bashari. "MaxEnt modeling for predicting suitable habitats and identifying the effects of climate change on a threatened species, Daphne mucronata, in central Iran". *Ecological Informatics*, **43**:116–123, 2018.