Potential Habitat Prediction of Invasive Species Flaveria Bidentis in China Based on ArcGIS and MAXENT

Xuanwei Liang

Chu Kochen Honors College Zhejiang University Beijing, China

3220101840@zju.edu.cn

Abstract. The flaveria bidentis, an alien invasive plant that poses a threat to ecological stability and agricultural output, has invaded China since the early 1990s and may have the potential to spread further. An accurate assessment of its potential spread can provide a basis for early warning and prevention. The flaveria bidentis's probable habitat in China under present ecological circumstances was predicted via a maxent model that was constructed based on dispersal data and environmental factors. The result showed that the suitable habitat of the flaveria bidentis is currently widely distributed in China, making up around 3.98% of China's total land area, having a total size of about 381,795km^2. The Jackknife method showed that the main factors that affected the distribution of the flaveria bidentis were the precipitation seasonality and the minimum temperature of the most frigid month. The model's AUC was 0.987, which means the precision of the forecast was quite high. The flaveria bidentis will continue to spread, therefore the pertinent departments ought to pay particular scrutiny and react promptly to set up protective measures in place.

Keywords: Flaveria bidentis, potential distribution, early monitoring and warning, MaxEnt model.

1. Introduction

Invasive alien species, which include insects, plants, and other organisms that affect the new ecosystem, worsen human health and cause economic and environmental devastation [1]. Numerous species worldwide experience ecological and habitat changes as a result of biological invasion. Once the invasive alien species are successfully colonized, they will be extremely difficult to eradicate, and predicting their potential habitat in advance is an important basis for formulating control strategies to prevent them from spreading [2]. Total economic damages worldwide between 1970 and 2017 brought on by invading alien species exceeded no less than $ 1.288 trillion [3]. For instance, the Flaveria bidentis not only inhibits other crops to grow, but can also devastate the existing ecosystems, posing a major threat to the invaded area's agricultural production and ecological security.

The flaveria bidentis, belonging to the Asteraceae family, is an annual herbaceous plant native to Brazil, Bolivia, Ecuador, Peru, Argentina, Chile, Paraguay and other South American countries and regions. It was then spread to Central America, southern North America and the West Indies and spread to Egypt, South Africa, the United Kingdom, France, Australia and other regions in Asia later [4]. In China, the Flaveria bidentis first appeared in Hebei Province in 1994, and was then spread to other surrounding regions centered on the central and southern parts of Hebei Province and since then has been one of the most hazardous invasive weed species distributed in China [5].

Ecological niche models have been very popular in recent years for analyzing the possible dispersion of invasive alien species, and the basic method is to use the origin distribution data of invasive alien species and the environmental data to construct models that can analyse the environmental conditions required for the growth of the invasive alien species, and then project them to predict the potential habitat [6]. The maximum entropy model is a machine-learning model that estimates the probability of habitat by calculating the probability distribution of maximum entropy. Compared with other models, the MaxEnt model can still perform accurate prediction with just the distribution data, and meets the needs of this research. Therefore, in order to provide a certain reference for the control of Flaveria bidentis's invasion, this research uses the MaxEnt model to predict the suitable habitat of the Flaveria bidentis in China and divides the potential habitat into 4 suitable grades according to certain criteria. This paper aims to stop the Flaveria bidentis from spreading over China.

2. Methods

2.1. Distribution Data

The distribution data of the Flaveria bidentis involved in this study are derived from the website of the Global Biodiversity Information Facility (GBIF, https://www.gbif.org) and Chinese Virtual Herbarium(https://www.cvh.ac.cn), as well as from literature and publications [6]. The author gained the corresponding latitude and longitude coordinates from Latitude query platform (http://www.gpsspg.com). In order to dwindle sampling bias, the Geographical scope of the study was divided by a 1 km×1 km grid and only one distribution point is randomly selected in each grid. After deleting the repetitive and null distribution points, 27 effective distribution points were retained. The specific coordinates of the Flaveria bidentis gained data were listed as.csv format, in which the longtitude and latitude data is saved as decimal. By importing the distribution data into arcgis system, distribution of Flaveria bidentis could be visualized. Through cursory observation, this research found that the distribution points are mainly concentrated in the southern part of northern China, the northern part of central China and the northern part of eastern China, as well as in Taiwan Island.

2.2. Environmental Variables

The environmental variables used in this study is the World Clim world environment database. 19 environmental climate parameters with a spatial resolution of 30s were extracted from the WorldClim environmental database (http://www.worldclim.org/). These variables were derived from the temperature and precipitation data to form more biologically valid factors, often showing a significance in ecological niche modeling. The detailed information of these variables is shown in the coming table.

Table 1. Environmental Variables

Environmental Variables	Information
Bio-1	mean temperature of the year
Bio-2	mean diurnal various scope
Bio-3	isothermal characteristic
Bio-4	temperature seasonal characteristic
Bio-5	maximum temperature of the hottest month
Bio-6	minimum temperature of the most frigid month
Bio-7	temperature range of the year
Bio-8	mean temperature of the most humid season
Bio-9	mean temperature of the most arid season
Bio-10	mean temperature of the hottest season
Bio-11	mean temperature of the most frigid season
Bio-12	precipitation of the year
Bio-13	precipitation of the most humid month
Bio-14	precipitation of the most arid month
Bio-15	precipitation Seasonal characteristic
Bio-16	precipitation of the most humid season
Bio-17	precipitation of the most arid season
Bio-18	precipitation of the hottest season
Bio-19	precipitation of he most frigid season

2.3. Map Data and Software

The China vector map data was gained from the website of the National Geomatic Center of China(www.ngcc.cn/ngcc/). The niche modeling tool used in this research is the Maxent software gained from the BurdGIS website(https://www.burdgis.com/blog/view/maxent-is-open-source/).

ArcGIS spatial platform used in this research is a complete set of GIS tools and methods gained from the ArcGIS Online platform(https://www.arcgis.com/index.html).

2.4. Model and Methods

This research would import distribution data and environment data into Maxent model,a season of the original distribution data was chosen as test data set, and the remaining portion was left as training data. The model should then be trained 10 times. In the process of parameter setting, this research relies Jackknife method for accuracy test, the other parameters are based on software defaults. The analysis results gained from the maxent model are output in Logistic format and ASCII type files which could be further processed with Arcgis system. The Jackknife method is often used to analyze the extent to which individual environmental variables affect predicted outcomes, or to identify key environmental factors that affect species distribution. The correlation and restriction of the occurrence of Flaveria bidentis were sorted according to environmental variables by knife cutting method, and the main influencing factors on the occurrence of Flaveria bidentis were screened out. In Arcgis software, raster tools could be used to estimate the size and proportion of the habitat of Flaveria bidentis as well as to plot an intuitive habitat distribution map.

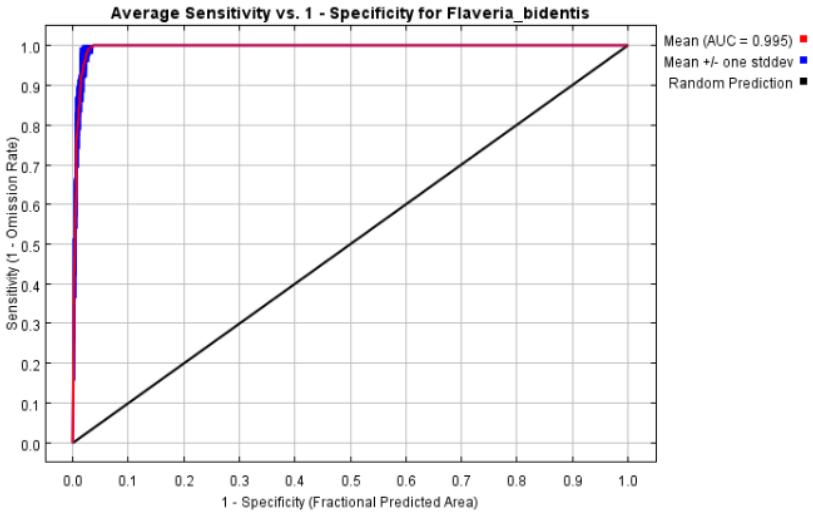


Figure 1. ROC curve of the prediction (Photo/Picture credit :Original).

2.5. Model Accuracy Test

The receiver operating characteristic (ROC) is the curve with the rate of true positives rate as the horizontal axis and the rate of false positives as the vertical axis. The AUC (are under the curve) value, which is typically used to determine the accuracy of the model's prediction outcomes, refers to the region circumscribed by the contour of the curve and the abscissa [7]. The more closely the AUC number resembles 1, the further the forecast is from the random distribution, the more strongly the environmental variables are correlated with the anticipated species' geographic distribution, and the more precisely the model predicts the species [8]. In general, for the maxent model, prediction with an AUC value higher than 0.9 is generally regarded as an excellent performance, good between 0.8 and 0.9, fair between 0.7 and 0.8, and poor below 0.7 [9].

3. Results and Analysis

3.1. Accuracy Test

It can be derived from Fig. 1 that the ROC curves of the two sets of training and testing datasets are quite different from the ROC curves of the random distribution model. The standard deviation is 0.001 and the average training AUC is 0.995. which were significantly higher than the AUC values of the random distribution (0.5). The conclusion is that the MaxEnt presented overall “excellent”

performance with a high accuracy and the prediction performed by the Maxent model is accurate and reliable.

3.2. Grouping of the Results

Each grid in China (the study area) is given a suitability value from 0 to 1 as part of the prediction output; the nearer the suitability value gets to 1, higher the likelihood that the grid is fit for the occurrence of a species. This data is then fed into the ArcGIS system for reclassification. The habitat prediction of the Flaveria bidentis could be divided into four levels by referring to others publications [10], namely unsuitable area (0≤suitability value<0.12), low suitable habitat (0.12≤suitability value<0.25), suitable habitat (0.25≤suitability value<0.50), highly suitable habitat (0.5≤suitability value<1). After the reclassification of the prediction data, the overall habitat prediction of the Flaveria bidentis in China could be observed in Fig. 2. The distribution points of the Flaveria bidentis were superimposed with the prediction map, and more than 98% of the distribution points are within the prediction habitat, which further verified the accuracy and reliability of the model.

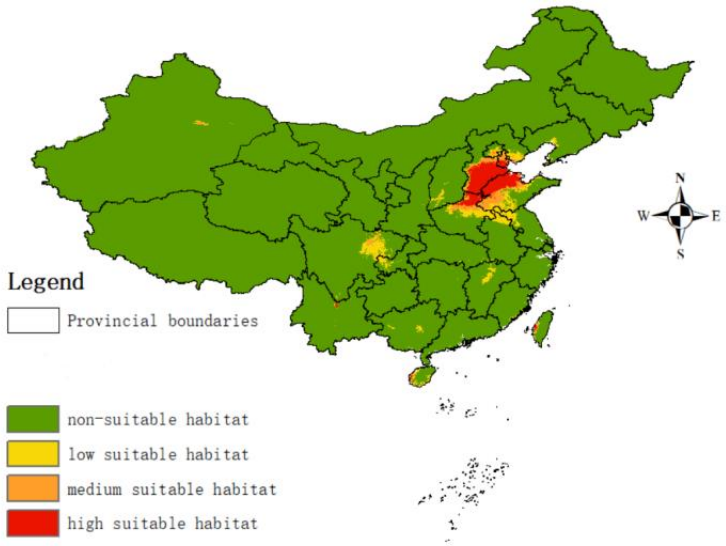


Figure 2. Habitat prediction map (Photo/Picture credit :Original)

Derived from Fig. 2, the suitable habitat of the Flaveria bidentis is mainly located in North China at 110°~ 120°E, 34°~ 43°N, and has a relatively large area distribution in Hebei, Henan, Tianjin, Beijing, Hainan and Shandong in China, but merely has a small area distribution in Jiangsu, Anhui, Sichuan, Chongqing, Jiangxi, Taiwan, Yunnan, Guangxi and Xinjiang. With a total area of about 381,795 km^2, or 3.98% of China's land area, a potential habitat area for Flaveria bidentis is currently widely dispersed throughout the country. Of this total, highly suitable habitat makes up 34.6% of the area, suitable habitat makes up 19.6%, and low suitable habitat makes up around 45.8%. A frequency table which is shown below is compiled for 13809840 raster data (suitability value) according to the four levels of habitat that have been divided, with frequencies rounded up to three decimal places.

Table 2. Frequency Table of Potential Habitat

Group	Value	Occurrence	Frequency
unsuitable area	0≤suitability value<0.12	13260618	0.96
low suitable habitat	0.12≤suitability value<0.25	251418	0.018
suitable habitat	0.25≤suitability value<0.50	107675	0.008
highly suitable habitat	0.5≤suitability value<1	190067	0.014

1) Highly suitable habitat: The majority of northern Shanxi, northern Henan, the majority of southern Hebei, and the majority of southern Tianjin make up North China's highly suitable habitat for

Flaveria bidentis. A small proportion of them are distributed in the western coastal areas of Taiwan Island and Hainan Island, moreover a small area of the north-central part of Yunnan.

2) Suitable habitat: The medium-suitable habitat of Flaveria bidentis (in China) is primarily distributed on the north and south sides of the extremely highly suitable distribution area in North China, including portions of central Hebei, the majority of central Tianjin, central Beijing, southern Shanxi, north-central Shanxi, and north-central Henan. A small proportion of them are distributed in the center of Xinjiang's Turpan Basin, the northeast coast of Liaodong Bay and adjacent inland areas, the sporadic area in the northwest of the Sichuan Basin, the sporadic areas of southern Sichuan and the western coastal areas of Taiwan Island and Hainan Island.

3) Low suitable habitat: The northern coast and adjacent inland areas of Bohai Bay, the northeastern coast and adjacent inland areas of Liaodong Bay, the central, southern, and eastern parts of Beijing, the southern and central parts of Shandong, the majority of the Sichuan Basin and Turpan Basin, the north-western part of Guizhou, the western, southern, and eastern part of the coastal areas of Hainan Island, and the sporadic locations are the main distribution areas for the lowsuitable habitat of the Flaveria bidentis.

For the prevention and management of this invasive alien species, it is obviously important to analyze the varied levels of Flaveria bidentis's suitable habitat in each province with prospective habitat.

The proportion of area occupied by potential habitat at each level was calculated for 17 of the 34 provincial-level administrative divisions in China and shown in the table below. The average suitability value AP in a certain region could be regarded as the comprehensive suitability index, showing the overall suitability of the Flaveria bidentis in one particular region [11]. For grouped data with known frequency distributions, the value of average suitability value in a certain region can be calculated by the following formula:

$$AP = \sum_{i=1}^{4} X_i \times P_i \tag{1}$$

Noted that X_i is the median suitability value of the habitat level i, and P_i is the area percentage of the habitat level.

Table 3. Potential Habitat Area Proportion

Region	Area Percentage				Comprehensive Suitabil-ity Index
	unsuita-ble area	***low suitable habitat***	***suitable habitat***	***highly suitable habitat***	
Tianjin	3.7	14.0	42.8	39.4	0.458
Shandong	35.7	17.6	19.4	27.8	0.317
Hebei	52.6	7.3	7.4	32.7	0.298
Beijing	54.8	28.7	16.5	0.2	0.147
Henan	53.4	12.2	6.8	9.5	0.145
Hainan	70.3	20.5	6.9	2.4	0.121
Taiwan	89.7	3.1	3.6	3.6	0.098
Jiangsu	84.5	15.2	0.3	0.0	0.098
Sichuan	91.6	7.5	0.9	0.0	0.072
Chongqing	89.8	10.1	0.1	0.0	0.072
Anhui	77.1	10.6	0.0	0.0	0.065
Jiangxi	95.5	4.5	0.0	0.0	0.065
Shanxi	95.9	3.9	0.2	0.0	0.065
Liaoning	97.5	2.4	0.1	0.0	0.063
Fujian	99.1	0.8	0.1	0.0	0.061
Guangxi	99.1	0.9	0.0	0.0	0.061
Yunnan	99.5	0.2	0.1	0.1	0.061

The result shows that Tianjin, Shandong and Hebei are covered with suitable and highly suitable habitats accounting for more than 40% of its area meanwhile have a comprehensive suitability index greater than or close to 0.3. Beijing, Taiwan, Henan and Hainan have suitable and highly suitable habitats covering more than or close to 10% of its area.

3.3. Environmental Variables Analysis

Table 4. Percentage Contribution of Variables

Variable	Percentage Contribution
Bio15	30.6
Bio6	19.8
Bio13	9
Bio9	8.8
Bio4	8.1
Bio11	7.8

In each iteration of the procedure for training, the rise in established gain is included to the contribution made by the relevant variable, or deducted from it if the alteration to the absolute amount of lambda is negative, to calculate the variable's percentage contribution. It can be seen in the following table that among the environmental parameters, factors like bio_15 (Precipitation Seasonality) contributed the maximum of 30.6%, followed by bio-6 (minimum temperature of most frigid month) 19.8% respectively. Bio-13 (precipitation of most humid month), bio-9 (mean temperature of most arid season), bio-4(temperature seasonal characteristic) and bio-11(mean temperature of most frigid season) contributed 9%, 8.8%, 8.1% and 7.8%. However other variables does not have such contribution.

The following figure (Fig. 3) shows the results of the jackknife test of variable importance [12]. This figure demonstrates the importance of each factor in affecting the Flaveria bidentis's distribution when used separately, and how the model is affected when each variable is wiped off. The environmental factor bio-6 (mean temperature of the most frigid season), which has the greatest gain when utilized alone, appears to offer the greatest amount of useful data contributions. The environmental factor bio-14 (precipitation of most arid months) seems to provide the most data that isn't accessible from all the other variables because it reduces the gain the most. Results shown in the below figure are averages over replicate runs.

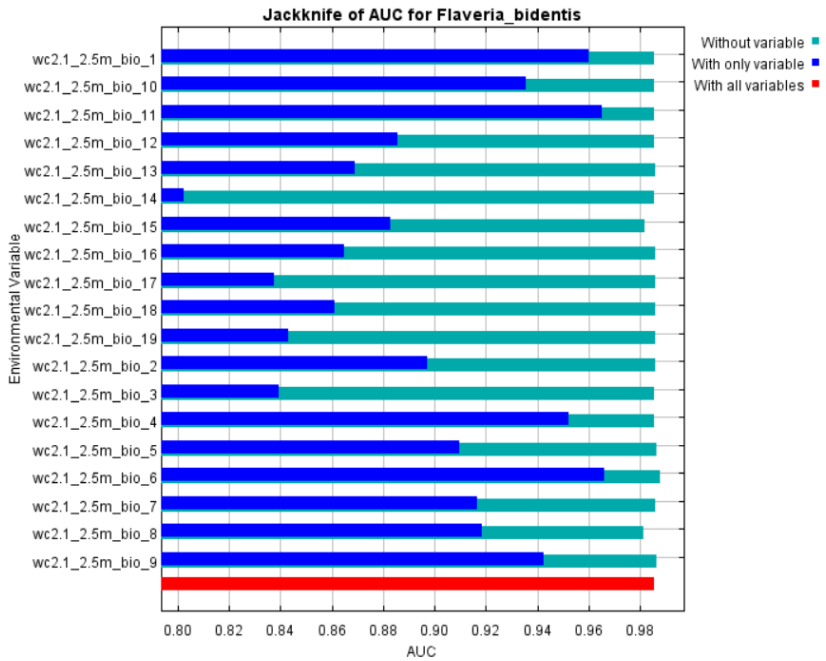


Figure 3. Results of jackknife test (Photo/Picture credit : Original)

4. Discussion

4.1. Ascription of The Results

This research constructed a reliable maxent model based on the distribution data of the Flaveria bidentis, which had successfully performed an accurate prediction of the potential suitable habitat of Flaveria bidentis in China as well as divided the potential habitat into 4 levels. With an AUC value of 0.995 corresponding to a high ability to predict habitats, the prediction in this research might be regarded as the most accurate in published researches [10, 13].

From the prediction in this research, it could be derived that the suitable habitat of the Flaveria bidentis in China is far more than limited to Tianjin, Hebei and other places where distribution data have been collected. By referring to the ecological and environmental conditions of the origin of the Flaveria bidentis and its invasion area, a deeper reason for this invasion might be found.

Native to Brazil, Argentina, and other countries in South America, the Flaveria bidentis spread to Central America, southern North America, and the West Indies, and later spread to Egypt, South Africa, the United Kingdom, France, Australia, and Japan. The Flaveria bidentis prefers a sunlit, humid environment and has a strong absorbent and developed root system, allowing it to adapt to a wide range of suitable environments. Under natural conditions, more than 90% of the seeds are distributed in the soil layer of 0~1cm, and the seeds can germinate under the condition of 13~45℃, especially in the soil habitat with 60% moisture content and an ambient temperature of 28℃ [4].

In China, the North China Plain, Sichuan Basin, Hainan, Xinjiang, Taiwan, and Sporadic areas in South China have suitable ecological and environmental conditions for the survival of the Flaveria bidentis. Nevertheless, in addition to Henan, Hebei, Shandong, Shanxi, and Taiwan where Flaveria bidentis's distribution are already noticed, Hainan, Xinjiang, Sichuan, Chongqing, Yunnan, and Guangxi where the distribution of the Flaveria bidentis are not widely spread are also covered with its potential habitat. Therefore, at present, the distribution of the Flaveria bidentis in China has not reached its maximum potential range, and there is still a high possibility of continuing invasion on temporal and spatial scales.

From the results of this knife-cut analysis, the influence factors that affect most on the model prediction are precipitation seasonal characteristic, minimum temperature of the moat frigid month, precipitation of the most humid month, mean temperature of the most arid season, temperature seasonal characteristic and mean temperature of the most frigid season. This may reflect the competitive advantage of the Flaveria bidentis over native plants under extreme temperature and drought conditions.

4.2. Proposal for Preventing Invasion

For Henan, Hebei, Shandong, Shanxi and Taiwan where the Flaveria bidentis's distribution are already noticed, scientific and reasonable measures must be taken to prevent Flaveria bidentis from spreading to other regions. For Hainan, Xinjiang, Sichuan, Chongqing, Yunnan and Guangxi where the distribution of the Flaveria bidentis are not widely spread are also covered with its potential habitat, the inspection and monitoring of seedlings of exotic tree species, flowers, food crops and other seedlings should be strengthened to prevent seedlings of the Flaveria bidentis from entering.

All in all, invasive plants are difficult to be monitored when the population size is small. Once widely found, they are often in the phase of prompt outbreak and are difficult to eradicate. Risk assessment, early warning and rapid interception are important means to effectively prevent and control biological invasions of the Flaveria bidentis for the above regions. Meanwhile, in the context of the current global warming, many areas that are currently unsuitable for the Flaveria bidentis are might become suitable habitats gradually, indirectly promoting the rapid spread of this invasive alien species.

4.3. Innovations and Limitations

Differing from the previous research [10, 13], this research found that the Turpan Basin, Sichuan Basin and Hainan Island were also suitable habitats for the Flaveria bidentis, but also found that the Northeast Plain, Loess Plateau and Yunnan-Guizhou Plateau are not covered with suitable habitat.

Although the prediction in this research was performed a high accuracy, because of the drifts of ecological niche of invasive species, the prediction of potential habitat constructed by distribution data may be distorted [14].

The 19 environmental variables of temperature and precipitation that were included in this study's construction of the maximum model were selected because they have the potential to have a substantial impact on Flaveria bidentis plant development and seed germination. But the factors affecting the distribution of plants are far more than these 19 factors: light conditions, soil texture, interaction between species and human interference are all inextricably related to the distribution of the Flaveria bidentis. This may bring some deviations to the prediction of habitat which need to be further improved in future studies.

5. Conclusion

The MaxEnt showed an excellent performance with a high accuracy with the AUC value of 0.995. The environmental variables like the Precipitation Seasonality contributed the maximum followed by the minimum temperature of most frigid month to predict the suitable habitat. The area of the potential suitable habitat of the Flaveria bidentis is around 381,700 km^2. Currently, the Flaveria bidentis has not reached its maximum potential distribution range in China, and is still possible to spread in China. This research recommends establishing interception belts at the edge of the distribution regions and the potential habitats, taking scientific and reasonable measures to prevent Flaveria bidentis from spreading. For further improvements, future studies might be focusing on adding more environmental variables or gaining more accurate data.

References

[1] Pyšek, Petr, Richardson and M. David, Invasive species, environmental change and management, and health. Annual review of environment and resources, 2010, vol. 35, pp. 25-55.

[2] D. R. Uden, C. R. Allen, D. G. Angeler, L. Corral, K. A. Fricke, Adaptive invasive species distribution models: a framework for modeling incipient invasions. Biological invasions, 2015, vol. 17, pp. 2831-2850.

[3] C. Diagne, B. Leroy, R. Gozlan, A. C. Vaissière, C. Assailly, and L. Nuninger, InvaCost, a public database of the economic costs of biological invasions worldwide. Scientific data, 2020, vol. 7, no. 1, pp. 277.

[4] X. Gao, T. Tang, Y. Liang, T. Zheng, W. Sang and Y. Chen, A warning about the biological invasion of a new exotic plant, Flaveria bidentis, and management techniques. Biodiversity Science, 2004, vol. 12, no. 2, pp. 274.

[5] Q. Wu, C. Xu, J. Li, W. Liu, F. Wan, J. Guo, and R. Wang, A number of causes have contributed to the spread of the non-native plant Flaveria bidentis (L.) Kuntze, resulting in patchy distribution patterns. Ecology and evolution, 2022, vol. 12, no. 9, pp. 9303.

[6] G. Zhu, G. Liu, W. Bu, and Y. Gao, Ecological niche modeling and its applications in biodiversity conservation. Biodiversity Science, 2013, vol. 21, no. 1, pp. 90.

[7] A. H. Hirzel, G. L. Lay, V. Helfer, C. Randin, and A. Guisan, Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling, 2006, vol. 199, no. 2, pp. 142-152.

[8] J. H. Elith, C. P. Graham, R. Anderson, M. Dudík, S. Ferrier, A. Guisan and N. E. Zimmermann, Novel methods improve prediction of species' distributions from occurrence data, Ecography, 2006, vol. 29, no. 2, pp. 129-151.

[9] J. A. PSwets, Measuring the accuracy of diagnostic systems. Science, 1988, vol. 240, no. 4857, pp. 1285-1293.

[10] C. Xiang-feng, Q. Guo-liang, H. Bai-shi, and L. Feng-quan, Prediction of potential suitable distribution area of Flaveria bidentis in China based on niche models, Yingyong Acta Ecologica Sinica, 2010, vol. 21, no. 12.

[11] K. Ming-jian, W. Hai-bo, Y. Ze, L. Ji-can, S. Ying and L. Guo-sheng, Prediction of the potential distribution areas of brown-banded cockroach in China based on maxent model. Port Health Control, 2017.

[12] R.G. Pearson, C. J. Raxworthy, M. Nakamura, and A. T. Peterson, Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of biogeography, 2007, 34, no. 1, pp. 102-117.

[13] X. F. Cao, G. L. Qian, B. S. Hu and F. Q. Liu, Prediction of potential suitable distribution area of flaveria bidentis in China based on niche models. Yingyong Shengtai Xuebao, 2010, 21, no. 12, pp. 3063-3069.

[14] O. Broennimann, U. A. Treier, H. Müller-Schärer, W. Thuiller, A. T. Peterson and A. Guisan, Evidence of climatic niche shift during biological invasion. Ecology letters, 2007, 10, no. 8, pp. 701-709.