Predicted Potentially Suitable Areas for Ips Typographus in the Northwest Region of China Based on the Maxent Model

Shuo Cao [1, *], Haodong Yang [1], Duo Xu [2]

[1] School of Forestry and Landscape Architecture, Xinjiang Agricultural University, Urumqi, China, 830052

[2] School of Water Conservancy and Civil Engineering, Xinjiang Agricultural University, Urumqi, China, 830052

* Corresponding Author Email: sc87378732@163.com

Abstract. As a typical secondary pest in the family Scolytidae of Coleoptera, Ips typographus poses a serious threat to coniferous forest resources in the northwest region of China. This study forecasts the spatial distribution of potential habitats for I. typographus by using the MaxEnt model and GIS technology, based on 28 distribution points of I. typographus and 12 key environmental factor variables in the study area. Through contribution rate analysis, model validation, and superposition analysis of host plants, the driving mechanism of its habitat suitability was evaluated. The outcomes revealed that the mean value of AUC was 0.989, and the MaEent model had an extremely high prediction accuracy. Annual mean temperature (Bio1), precipitation of the driest month (Bio14), temperature seasonality (Bio4), and precipitation of the coldest quarter (Bio19) were determined as the key factors governing the distribution area of I. typographus among the 12 environmental variables used in the modeling. Their respective contribution rates were 31.5%, 26.3%, 9.8%, and 10.2%. These four variables collectively accounted for 77.8% of the total contribution. The northern mountainous areas of the study area are suitable for the spread of I. typographus. For example, Altay and Tacheng regions. The findings of this research could offer theoretical foundations for local forestry authorities to develop corresponding prevention and control measures.

Keywords: MaxEnt model, Ips typographus, Potential habitat, Environmental factors.

1. Introduction

Bark beetles belong to the family Scolytidae (Coleoptera, Insecta), a group of forest pests specialized in infesting the phloem and xylem of host plants. These pests not only cause direct physical damage to host plants but also disrupt nutrient transport systems, leading to stunted growth, debilitation, and even death of host plants [1], thereby incurring significant economic losses and ecological degradation. Among numerous bark beetle species, Ips typographus is particularly notable. This species is widely distributed in China's northwest and northeast regions, and other regions, with its damage severity and affected area increasing annually. To effectively deal with such pests, the application of Species Distribution Models (SDM) is particularly important. SDMs integrate species distribution data with environmental layers and employ specific algorithms to predict species occurrence probability, richness, and habitat suitability [2]. Common SDMs include Genetic Algorithm for Ruleset Production (GARP), Ecological-Niche Factor Analysis (ENFA), Maximum Entropy Model (MaxEnt), and Climate Matching EXtremes (CLIMEX) [3]. Among these, the MaxEt model has gained widespread use in China for predicting suitable habitats of pests in recent years due to its operational simplicity and high prediction accuracy [4].

In the current research landscape, using the MaxEt model to predict species' suitable habitats has become a trend, with many researchers adopting the model’s predicted suitable areas as a reference for formulating pest control strategies. For example, Sheng Li et al. [5] utilized the model of MaxEnt to project the spatial distribution of suitable regions for Emeia pseudosauteri are southern Zhejiang Province and riverine regions, providing a theoretical foundation for habitat conservation of E. pseudosauteri. Yi Zhao et al. [6] predicted through this model that Chinese fir (Cunninghamia lanceolata (Lamb.) Hook.) is primarily distributed in the subtropical zones of China, offering theoretical references for forest management. Xinju Wei et al. [7] predicted that Cicadella viridis is predominantly distributed across the areas of the Loess Plateau, Shandong Peninsula, and North China Plain, aiding the development of effective control strategies.

However, research on predicting suitable habitats for harmful organisms in the northwest region of China remains scarce, particularly for bark beetles. Therefore, this research uses the model of MaxEnt and integrates multi-source environmental data to comprehensively conduct a spatial assessment of potential suitable habitats for I. typographus in the northwest region of China and the driving factors behind it. By connecting climate and topography, the scope of the suitable living regions of I. typographus under the present change of climate is speculated. In addition, through superposition analysis with the host plant (P. obovata Ledeb.) in ArcGIS (Arc Geographic Information System), the forecast accuracy of the species distribution model could be further optimized, thus offering scientific support for the prevention and control of forest pests in this region.

2. Maxent Model

2.1. The working principle of the MaxEnt model

The model of MaxEnt serves as a predictive tool for species distribution modeling based on the niche theory and a typical representative of the SDM. Its core principle is to maximize the entropy value and use the documented species distribution datasets and environmental variable datasets to construct a non-linear relationship between species distribution and environmental factors, so as to estimate where species may potentially inhabit in specific regions.

The main working principle of the MaxEnt model is very complex. This study will explain the main principles, which are as follows:

(1) The definition of entropy in the MaxEnt model.

In the case of a discrete random variable Y and its distribution of probability $P(y)$, the mathematical expression of entropy is:

$$H(Y) = -\sum_{y} P(y) \log P(y) \tag{1}$$

For the discrete random variable X and its conditional distribution of probability $P(y|x)$, the conditional expression of entropy:

$$H(Y|X) = -\sum_{x,y} P(x) P(y|x) \log P(y|x) \tag{2}$$

(2) Constraints of the MaxEnt model.

Suppose there are n eigenfunctions $f_i(x,y)$, and the expected constraint for each characteristic is:

$$E_P(f_i) = E_{\tilde{P}}(f_i) \tag{3}$$

Among them, use $E_{\tilde{P}}(f_i)$ to represent the expectation of the characteristic function $f_i(x,y)$ with respect to the empirical distribution $\tilde{P}(x,y)$:

$$E_{\tilde{P}}(f_i) = \sum_{x,y} \tilde{P}(x,y) f_i(x,y) \tag{4}$$

The model's expectation $P(y|x)$ with respect to the characteristic function $f_i(x,y)$:

$$E_P(f_i) = \sum_{x,y} \tilde{P}(x) P(y|x) f_i(x,y) \tag{5}$$

(3) In the MaxEnt model, by establishing a constrained optimization model:

$$\max_{P\in\mathcal{C}} H(P) = -\sum_{x,y}\tilde{P}(x)P(y|\ x)logP(y|\ x) \tag{6}$$

$$s.t.\ E_P(f_i) = E_{\tilde{P}}(f_i)\ i=1,2,\ldots,n \tag{7}$$

$$\sum_{y} P(y|\ x) = 1 \tag{8}$$

(4) Introducing the Lagrange multiplier λ_i enables the construction of the MaxEnt model solution:

$$L(P,\lambda) = -H(P) + \sum_{i=1}^{n}\lambda_i\left(E_P(f_i) - E_{\tilde{P}}(f_i)\right) + \lambda_0\left(\sum_{y}P(y|\ x) - 1\right) \tag{9}$$

The form of the solution can be obtained by the variational method to find the optimal solution:

$$P_\lambda(y|\ x) = \frac{1}{Z_\lambda(x)} exp\left(\sum_{i=1}^{n}\lambda_i f_i(x,y)\right) \tag{10}$$

$$Z_\lambda(x) = \sum_{y} exp\left(\sum_{i=1}^{n}\lambda_i f_i(x,y)\right) \tag{11}$$

where $Z_\lambda(x)$ is the normalization factor.

The results generated by the MaxEnt model are the optimal solutions obtained from the above steps. This is the principle of the MaxEnt model [8].

2.2. MaxEnt model building

Input the geographical distribution points of I. typographus and the relevant climate factors into the model of MaxEnt for suitability analysis. In the first step, according to the total number of the dataset, 75% of the point data on geographical distribution is employed in training dataset for model formulation and calculation of parameter, whereas the remaining 25% is employed as the testing dataset to assess the model's prediction accuracy. To make the model more stable and accurate, this study specifies the number of repetitions of the MaxEnt model as 500 times and selects the "Logistic" and "Asc" formats to facilitate subsequent analysis and visualization.

To ensure the reliability of the model results, this study runs the MaxEnt model 10 times repeatedly, each time using different random data allocations, so as to reduce the errors and fluctuations caused by data randomness [9]. Finally, import the prediction results generated by MaxEnt into ArcGIS. Through some of the previous studies conducted by others [10], the suitable habitat of I. typographus is divided into: no suitability area (0-0.05), low suitability area (0.05-0.19), moderately suitability area (0.19-0.44) and highly suitability area (0.44-1).

3. Results

The research area of this study is shown in Figure 1 and was obtained through ArcGIS. The distribution points of I. typographus used in this research are sourced from CNKI (https://www.cnki.net/) and the National Specimen Information Infrastructure of China (http://www.nsii.org.cn). WorldClim (http://worldclim.org/) served as the source of climatic factors. Data on slope, aspect, elevation, and human activity are obtained from the DEM elevation data (www.gscloud.cn), and the National Cryosphere Desert Data Center (http://www.ncdc.ac.cn) was used to acquire host data.

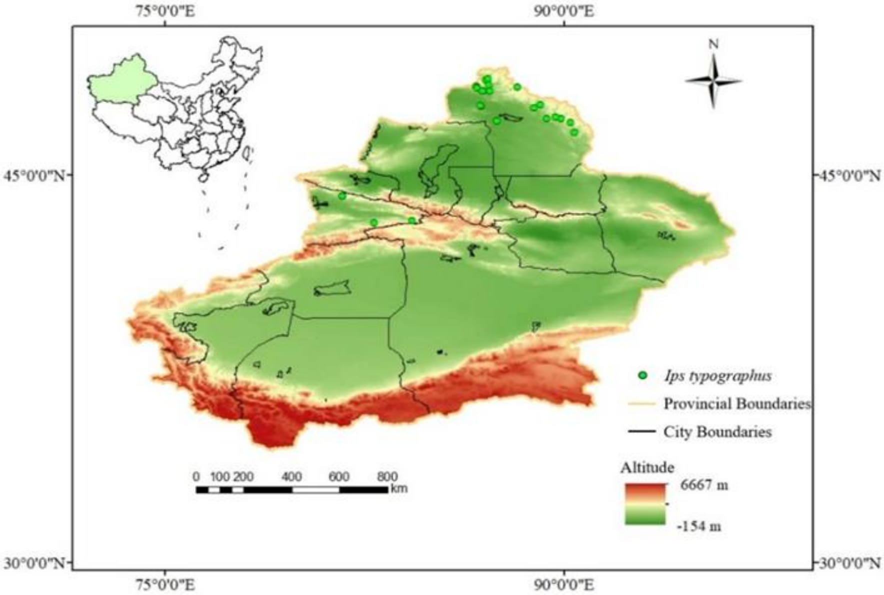

Figure 1. Distribution of Ips typographus

3.1. The MaxEnt model's prediction outcomes

The model of MaxEnt predicted an average AUC (Area Under Curve) of 0.989 for the ROC (Receiver Operating Characteristic) curve (Figure 2.). This metric suggests the model exhibits robust predictive capability and high accuracy. AUC is a key indicator: the closer its value approaches 1, the better the model's performance in discriminating species' suitable habitats. [11].

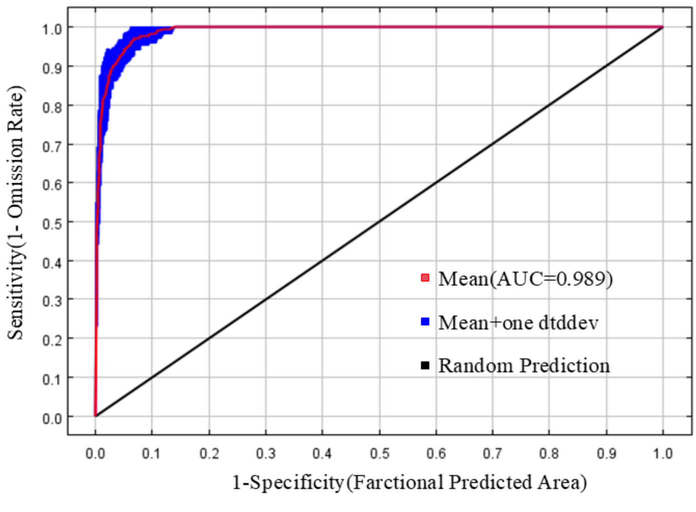

Figure 2. ROC curves and AUC averages for MaxEnt.

3.2. Main influencing factors of the suitable habitats for I. *typographus*

The MaxEnt model prediction shows that the geographical distribution of I. typographus is affected by many environmental factors, with temperature and precipitation-related variables being the most crucial. The contribution rates of these factors in the model are shown in Table.1. Among them, the annual mean temperature (Bio1) is the highest, reaching 31.5%, which reflects that temperature plays a dominant role in the suitable living region of I. typographus. The precipitation of driest month (Bio14) ranks second with a contribution rate of 26.3%, indicating that the water conditions in the dry season are very important for the survival of I. typographus. The temperature seasonality (Bio4) and the precipitation of coldest quarter (Bio19) contribute 9.8% and 10.2%

respectively. These four variables account for 77.8% of the total contribution, which has a great influence on the model and affects the potential distribution range of I. typographus. In contrast, variables such as elevation (1.6%), human activities (1.5%), and precipitation seasonality (Bio15, 0.8%) have a relatively small impact on the distribution of I. typographus, indicating that its distribution is more dependent on climatic factors.

Table.1. Contribution rate of environmental factors to I. typographus

Variable	Contribution (%)
Bio1 (Annual mean temperature)	31.5
Bio14 (Precipitation of driest month)	26.3
Bio19 (Precipitation of coldest quarter)	10.2
Bio4 (Temperature seasonality)	9.8
Bio3 (Isothermality)	6.7
Slope	6.0
Bio2 (Mean diurnal range)	2.0
Aspect	1.9
Bio11 (Mean temperature of coldest quarter)	1.7
Elevation	1.6
Human activities	1.5
Bio15 (Precipitation seasonality)	0.8

From the results obtained by analyzing the response curves of the main factors, the most suitable habitat climate conditions for I. typographus are as follows: the annual mean temperature is between -4°C and 3°C, the temperature seasonality is in the range of 1400-1500, the precipitation of driest month is 8-12mm, and the precipitation of coldest quarter reaches 70-77mm. These specific climate threshold values reflect that this species has stronger adaptability and survival probability in a relatively cold and humid climate environment (Figure 3.).

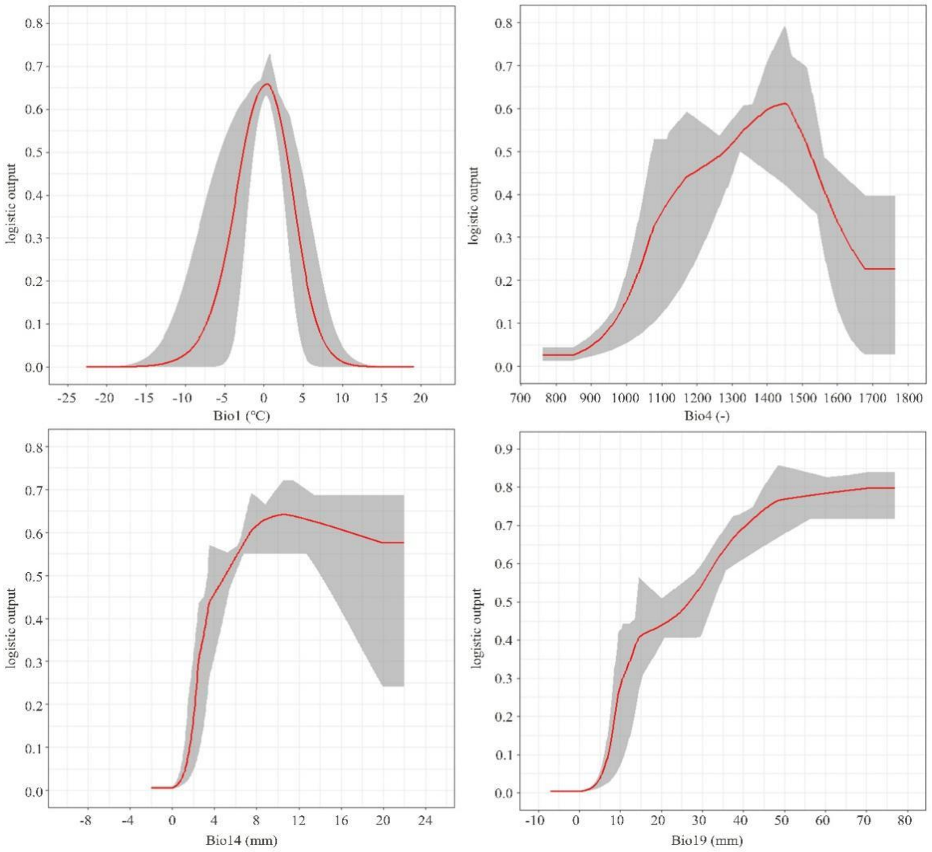

Figure 3. The response curves for variables Bio1, Bio4, Bio14, and Bio19. (The red line represents the curve trend, and the gray area represents the curve range).

3.3. Prediction of the suitable regions of I. typographus in the northwest region of China

The prediction results obtained from the model of MaxEnt show that the total region of suitable areas for I. typographus in the study region reaches $18.54\times10^4 km^2$ (Figure 4.), constituting 11.14% of the total study area. This result indicates that there is a large potential distribution area of I. typographus in the study area. According to different suitability levels, the levels are divided into low suitability area, moderately suitability area, and highly suitability area according to the standards given above. The distribution of each level is as follows:

The region of the highly suitability area is $1.35\times10^4 km^2$, accounting for 0.81% of the total area of study area. The high-suitability areas are mostly concentrated in Altay region and Tacheng region. These places have relatively stable climatic conditions and suitable environmental factors, thus providing a good living and breeding environment, which is suitable for the growth and spread of I. typographus.

The area of the moderately suitability area is $4.21\times10^4 km^2$, accounting for 2.53% of the total area of study area, and is mostly distributed in Tacheng region, Altay region, and Ili Kazak Autonomous Prefecture. Although the environment here can support the survival of I. typographus, there are still some restrictive factors compared with the high-suitability areas, such as large temperature fluctuations or uneven precipitation distribution.

The area of the low suitability area reaches $12.98\times10^4 km^2$, accounting for 7.8% of the total area of study area. The environmental conditions in this area are relatively harsh, but I. typographus can still survive here. The low-suitability areas are mainly concentrated in Ili Kazak Autonomous Prefecture, Hami region, Changji Hui Autonomous Prefecture, Tacheng region, and Altay region. The climate in these places may have obvious seasonal changes or low precipitation, thus limiting the distribution range of I. typographus.

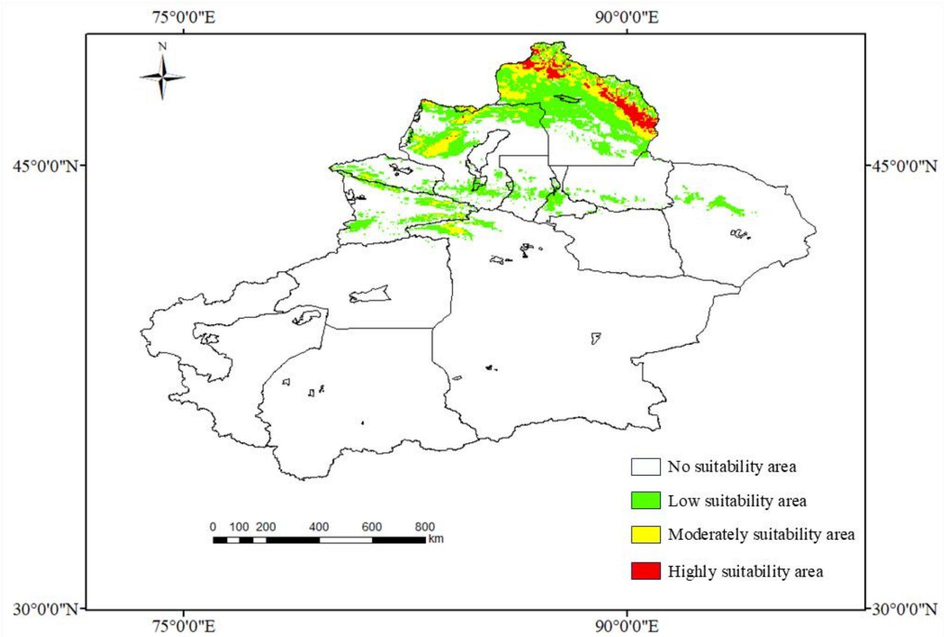

Figure 4. The suitable habitat of the I. typographus

3.4. Prediction of the final potential distribution of I. typographus in northwestern China via host distribution overlay

Use the overlay analysis tool in the ArcGIS toolbox to integrate the distribution data of the host (Picea obovata Ledeb.) and I. typographus, and assign the same weight to both to guarantee the precision of model prediction outcomes. Overlay analysis results indicate that the total potential distribution area of I. typographus is $11.8\times10^4 km^2$ (Figure 5.), constituting 7.08% of the total study area. Through the division of the potential distribution region, the region of the high potential distribution area is $0.68\times10^4 km^2$, accounting for 0.41% of the overall area of the study region. This area is primarily located in Altay Prefecture and Hami Prefecture. The area of medium potential distribution area is $1.72 \times 10^4 km^2$, constituting 1.03% of the total study area. This area is mainly concentrated in Altay Prefecture and Tacheng Prefecture. The area of the low potential distribution area reaches $9.4 \times 10^4 km^2$, constituting 5.64% of the overall study zone. This area is mainly in Altay Prefecture, Tacheng Prefecture, Ili Kazak Autonomous Prefecture, and Hami Prefecture.

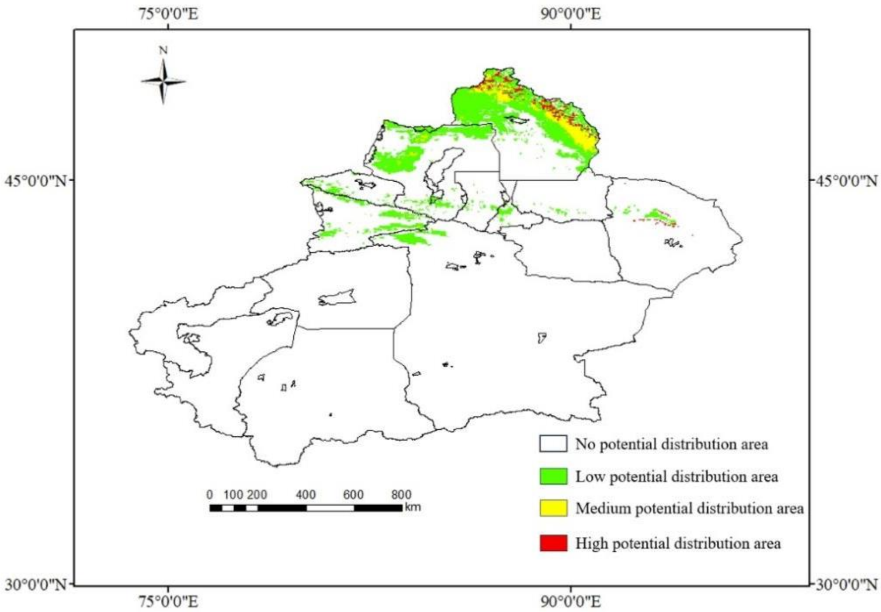

Figure 5. The potential distribution area of the I. typographus

4. Conclusions

This study predicted the habitats meeting the suitability requirements of I. typographus in the northwest region of China through the MaxEnt model. The prediction results show that the highly suitable habitats of I. typographus are mostly concentrated in Altay region, Tacheng region, and Ili Kazakh Autonomous Prefecture, with an area accounting for approximately 3.34% of the total area of the study area. Major climatic variables determining the distribution of suitable habitats for I. typographus include the annual mean temperature (Bio1), the precipitation in the driest month (Bio14), the standard deviation of temperature seasonality (Bio4), and the precipitation in the coldest quarter (Bio19), which may be closely related to the life characteristics of I. typographus and its environmental adaptation strategies. In addition, through further analysis by overlaying host data, it can be known that the overall area of the potential distribution area (Figure 5.) has become smaller, and Hami region appears in the high potential distribution area. This change indicates that there may be some environmental factors in Hami region that are conducive to the survival of I. typographus. The above research results systematically analyze the potential distribution characteristics and driving factors of I. typographus in the northwest region of China, promote the popularization of spatial information technology in the field of forest pest prevention and control, and provide multi-dimensional theoretical support for the scientific prevention and control of this pest and the protection of forest resources.

Although the MaxEnt model has a high prediction accuracy, there are still some shortcomings in this study. First, the spatial coverage of species distribution point data is lacking, which may affect the accuracy of the model output. Second, when the MaxEnt model is formed, it often relies on historical data for modeling and cannot accurately grasp the complex impacts of climate change promptly. Therefore, in the next step of the research, more distribution points of I. typographus should be included, and prediction and analysis should be carried out in combination with future climate data. Through these improvement measures, the MaxEnt model's prediction accuracy can be improved, providing a more reliable basis for future monitoring and management.

References

[1] Kozhoridze G, Korolyova N, Komarek J, et al. Direct and mediated impacts of mixed forests on Norway spruce infestation by European bark beetle Ips typographus[J]. Forest Ecology and Management, 2024, 569: 122184.

[2] Rathore M K, Sharma L K. Efficacy of species distribution models (SDMs) for ecological realms to ascertain biological conservation and practices[J]. Biodiversity and Conservation, 2023, 32(10): 3053-3087.

[3] Hayat U, Shi J, Wu Z, et al. Which SDM Model, CLIMEX vs. MaxEnt, Best Forecasts Aeolesthes sarta Distribution at a Global Scale under Climate Change Scenarios?[J]. Insects, 2024, 15(5): 324.

[4] Zhang H, Song J, Zhao H, et al. Predicting the distribution of the invasive species Leptocybe invasa: Combining MaxEnt and geodetector models[J]. Insects, 2021, 12(2): 92.

[5] Li S, Wang Z, Zhu Z, et al. Predicting the potential suitable distribution area of Emeia pseudosauteri in Zhejiang Province based on the MaxEnt model[J]. Scientific Reports, 2023, 13(1): 1806.

[6] Zhao Y, Deng X, Xiang W, et al. Predicting potential suitable habitats of Chinese fir under current and future climatic scenarios based on Maxent model[J]. Ecological Informatics, 2021, 64: 101393.

[7] Wei X, Xu D, Zhuo Z. Predicting the impact of climate change on the geographical distribution of leafhopper, Cicadella viridis in China through the MaxEnt model[J]. Insects, 2023, 14(7): 586.

[8] Merow C, Smith M J, Silander Jr J A. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter[J]. Ecography, 2013, 36(10): 1058-1069.

[9] Zhang Y, Tang J, Ren G, et al. Global potential distribution prediction of Xanthium italicum based on Maxent model[J]. Scientific reports, 2021, 11(1): 16545.

[10] Zhang H, Zhu Y, Ma Z, et al. Simulating the impact of climate change on the suitable area for cotton in Xinjiang based on SDMs model[J]. Industrial Crops and Products, 2025, 227: 120750.

[11] Shi X, Wang J, Zhang L, et al. Prediction of the potentially suitable areas of Litsea cubeba in China based on future climate change using the optimized MaxEnt model[J]. Ecological Indicators, 2023, 148: 110093.