*Bulletin of Entomological Research*

**cambridge.org/ber**

## Research Paper

*These authors contributed equally to this work and should be regarded as co-first authors.

**Cite this article:** Wang Z, Zhuo Z, Peng Y, Xu D (2024). Predicting the geographical potential distribution of species *Opisina arenosella* Walker in China under different climate scenarios based on the MaxEnt model. *Bulletin of Entomological Research* 1–9. https://doi.org/10.1017/S0007485324000464

Received: 27 May 2024
Revised: 4 July 2024
Accepted: 18 August 2024

**Keywords:**
geographic distribution; invasive species; MaxEnt model; *Opisina arenosella* Walker

**Corresponding author:**
Danping Xu;
Email: xudanping@cwnu.edu.cn



# Predicting the geographical potential distribution of species *Opisina arenosella* Walker in China under different climate scenarios based on the MaxEnt model

Zhiling Wang*, Zhihang Zhuo*, Yaqin Peng and Danping Xu

College of Life Science, China West Normal University, Nanchong, China

**Abstract**

As global warming increases with the frequency of extreme weather, the distribution of species is inevitably affected. Among them, highly damaging invasive species are of particular concern. Being able to effectively predict the geographic distribution of invasive species and future distribution trends is a key entry point for their control. *Opisina arenosella* Walker is an invasive species, and its ability to live on the backs of foliage and generate canals to hide adds to the difficulty of control. In this paper, the current and future distributions of *O. arenosella* under three typical emission scenarios in 2050 and 2090 are projected based on the MaxEnt model combining 19 bioclimatic variables. Filter through the variables to find the four key environment variables: BIO 1, BIO 6, BIO 11 and BIO 4. The results show that *O. arenosella* is distributed only in the eight provinces of Tibet, Yunnan, Fujian, Guangxi, Taiwan, Guangdong, Hong Kong and Hainan in the southeastern region. Its high suitability area is concentrated in Taiwan and Hainan. In the long run, highly suitable areas will continue to increase in size, while moderately suitable areas and poorly suitable areas will decrease to varying degrees. This paper aims to provide theoretical references for the control of *O. arenosella*.

## Introduction

*Opisina arenosella* Walker (Lepidoptera: Oecophoridae) is also known as the coconut black-headed caterpillar (BHC). BHC is an essential quarantine leaf-feeding pest native to Sri Lanka and India in South Asia (Lu *et al.*, 2023). BHC was first detected and recorded in 1909 in the coastal cities of southern India, and since 1920 it has been reported as an economically relevant pest in Sri Lanka, Hindustan, Bengal Republic, Indonesia, Thailand and Malaysia (Perera *et al.*, 2009). By August 2013, BHC damage was first detected in China in Wanning City, Hainan Province. Subsequently, BHC damage to palms was also reported in Fangchenggang City, Guangxi Province and Foshan City, Guangdong Province (Tang, 2017). According to current research, BHC can harm about 34 species of palms (Li *et al.*, 2021) such as coconut (*Cocos nucifera*), royal palm (*Roystonea regia* (HBK.) O.F. Cook), Chinese fan palm (*Livistona chinensis*), date palm (*Phoenix dactylifera* L.), *Oredoxa oleracea* Kurth., *Metroxylon sagu* Rottboell., *Caryota urens* L., *Hyphaene thebacia* L., *Elaeis guineensis* Jacq. and others (Jin *et al.*, 2019). In Hainan Province, China, it mainly includes 21 species of plants that are affected by this pest, such as *R. regia*, *Phoenix sylvestris*, *Wodyetia bifurcata*, *L. chinensi*, *Butia capitat*, *Hyophore lagenicauli*, *Washingtonia robusta*, *Latania lontaroide*, *Prithchardia pacifica*, *Bismarckia nobilis*, *Corypha umbaculifera*, *Chrysalidocarpus lutescen*, *Arenga catechu*, *Archontophoenix alexandrae*, *Latania verschaffeltii*, *Arenga pinnat*, *Licuala grandis*, *Borassus flabellife*, *Saccharum sinensis* and *Musa paradisiaca*. However, it is the palms that it mainly harms (Tang, 2017). Regarding the BHC, the damage area is mainly concentrated on the leaves of the plant. The females lay their eggs on the underside of the leaf blades and the eggs successfully hatch into larvae, which hide on the underside. The larvae feed on the thin-walled tissues of the leaves and leave behind leaf debris and excreta, and use their excreta and spider threads on the underside of the coconut leaves to construct galleries (Kumar, 2002; Perera *et al.*, 2009). This causes mechanical damage to the leaves, while shading by leaf litter and excreta causes a reduction in photosynthetic area (Nasser and Abdurahiman, 2001). Not only that, the insect will damage the leaves 2–3 days later, and the leaves will automatically roll down and fold closed, providing a hiding place for the insect (Li *et al.*, 2021). Also, because the BHC feeds only in the canopy, often at heights of 10 m or more, it is difficult to detect in the early stages of the infestation (Lu *et al.*, 2023). It is worth mentioning that only the upper epidermis is left after feeding by the larvae, which, together with the fact that it is a group feeding, results in a large damaged area that externally looks burned. This not only affects the appearance but also leads to a reduction in fruit yield, even early fruit drop and slow growth (Lever, 1969;

Mao and Kunimi, 1994). For coconuts in particular, coconut yields are reduced by 45.4% in coconuts infested with the insect, and 13.8–21% of coconut leaves are damaged. Studies have shown that, in cases of severe damage, it can take up to 4 years to recover normal yields (Mohan *et al.*, 2010). With regard to the spread of the BHC, in addition to its own superior flight ability (at night), the trade transportation of the host plant (palm) and the development of the coconut chain are the main reasons for it (Jingjing *et al.*, 2023). With regard to the control of the BHC, tall palms make it difficult to spray pesticides, as the leaves are not easy to spray comprehensively, and it also increases labour costs. Also, due to the pest's characteristic of hiding behind the leaves to feed, most of the pesticides fall on the front side of the leaves and fall by gravity. This makes spraying difficult and at the same time, greatly reduces the contact rate between the BHC and the pesticide (Jin *et al.*, 2019). Nowadays, all kinds of science and technology are developed interactively, especially the combination of the distribution of various types of pests with modelling software and geographic software. It can predict the future distribution trend and geographic direction of pests, which plays a good role in pointing out control work.

The description and understanding of a species are inextricably linked to its distribution and drivers, which are the basis of both ecology and geography (Barker and MacIsaac, 2022). Species distribution models (SDMs) are critical modelling tools for studying the direction of species' geographic distributions, linking and combining species occurrence data with environmental predictors and utilizing them in a variety of ways (Guisan and Zimmermann, 2000). The results derived from SDM can be used to study and excavate the target species in depth in terms of response to the environment, the most critical environmental factors, the predicted probability of occurrence of the species in time and space and the predicted presence or absence of the species. From the above aspects, hypotheses related to determining ecological conditions for the distribution of target species can be formulated (Bradie and Leung, 2017). With the rapid development of SDM, many more methods have been derived, including CART (Breiman *et al.*, 1984), MARS (Friedman, 1991), GARP (Stockwell, 1999), GLMS (Guisan *et al.*, 2002) and GAMS (Guisan *et al.*, 2002). In addition to this, it is worth mentioning that the most noteworthy algorithms for machine learning include random forest (Breiman, 2001), artificial neural network (Olden *et al.*, 2008), gradient boosting (De'Ath, 2007), support vector machine (Guo *et al.*, 2005) and maximum entropy (MaxEnt) (Phillips *et al.*, 2004). In this context, MaxEnt has been widely used for species prediction and is one of the most popular methods in climate modelling studies (Elith *et al.*, 2006). It can produce relatively robust results with only a small number of presence/absence (or pseudo-absence) samples (Guisan *et al.*, 2007; Elith *et al.*, 2011). MaxEnt distinguishes between the presence and absence of environmental conditions for the classification of the target species through maximum entropy (Elith *et al.*, 2011), and then finds the probability distribution of maximum entropy under a set of constraints based on the occurrence data of the target species (Phillips *et al.*, 2006). As a purely modelling algorithm, it can be integrated with predictive variables such as climate and remote-sensing variables. These data are used to predict the relative occurrence rate of a target species in a designated landscape (Rhoden *et al.*, 2017).

Therefore, in this study, MaxEnt modelling was used to combine the occurrence data of the target species, BHC, with 19 bioclimatic variables to predict the habitat areas in China. This work was able to obtain the most important environmental variables affecting its distribution, which in turn allowed us to predict the geographic distribution of the pest under three climate scenarios for the next two periods. This provides some ideas for the control of BHC in the future, both in terms of climate and geographic distribution.

## Materials and methods

### *Sources and processing of BHC occurrence data*

In this study, data on the occurrence of BHC in China and other countries were selected. These data were mainly obtained by searching websites on the Internet, published related literature, and newspapers and books. The website is powered by the European and Mediterranean Plant Protection Organization (EPPO, https://gd.eppo.int) and Global Biodiversity Information Network (GBIF, http://www.gbif.org/) databases for BHC occurrence data. The rest of the data were obtained through CNKI and Web of Science in the relevant literature. Of these avenues, data for which precise latitude and longitude were available were recorded, and those that provided only geographic location names were obtained through Google Earth (http://ditu.google.cn), using county location descriptions for latitude and longitude. Then, all the points are filtered, duplicates and errors are removed and they are carefully checked. Finally, the geographic coordinates (latitude and longitude) were converted into an Excel spreadsheet according to the requirements of the MaxEnt model and then saved in 'CSV' format. In order to avoid overfitting, ENMTools version 1.0.4 of the R platform was used to spatially filter the data according to longitude, retaining only one point in each grid cell ($5 \times 5$ km). Finally, a total of 106 BHC distribution points were obtained.

### *Acquisition and processing of environment variable data*

MaxEnt version 3.4.4 and ArcGIS version 10.4.1 were used in this study. The input MaxEnt data required environmental variable data in addition to the geographic coordinates (latitude and longitude) of the point of occurrence. Environmental data were obtained from WorldClim (http://www.Worldclim.org), containing 19 bioclimatic variables and 48 monthly climatic variables, and the data were in ASC format for use in ArcGIS software. Current climate data (2020s) and future climate data (2050s and 2090s) were obtained from the World Climate Data website (https://www.worldclimatedata.org/), and the spatial resolution of the above data is 2.5 arc-minutes (about 4.5 $km^2$). The Sixth International Coupled Model Comparison Program (CMIP6) model proposes several shared socio-economic pathways (SSPs) designed based on different socio-economic assumptions, and in this study, three SSP emission scenarios based on the simulation of the CanESM 5 model were selected (2.6 W $m^{-2}$ (SSP1-2.6), 4.5 W $m^{-2}$ (SSP2-4.5) and 8.5 W $m^{-2}$ (SSP5-8.5)) under the climate data.

Regarding the way these data were handled, the first step was to take these environmental variables and use the Jackknife test in MaxEnt to determine the extent to which each environmental variable contributes to the construction of the model and to exclude variables with a contribution of 0. The first step was to compare the environmental variables with their Pearson correlation coefficients. To avoid overfitting and improve model accuracy, the remaining variables were compared for their Pearson correlation coefficients and the two environmental variables with absolute values greater than 0.8 were removed. This is because it indicates a strong linear relationship between them. Finally, the remaining environmental variables were used to model predictions of current and future fitness ranges.

### *MaxEnt modelling methodology and accuracy assessment*

The BHC geolocation information converted to CSV format was entered into the MaxEnt software along with the environmental variable data after harmonizing the projection and resolution. Seventy-five per cent of the distribution records were used as a randomised training dataset to build the predictive model, and 25% of the remaining distribution records were used as a test dataset. The number of iterations was then chosen as 1000, the number of background points was chosen as 10,000 and the model was repeated ten times for averaging using the cross-validation method. Finally, this work plots the receiver operating characteristic curve (ROC) and area under curve (AUC) to evaluate the model fit. The ROC curve is a graphical tool used to evaluate the performance of a binary classification model. The AUC takes values ranging from 0 to 1, with values of 0–0.6 indicating extremely poor predictive performance; values between 0.7 and 0.8 indicate fair predictive performance; values between 0.8 and 0.9 indicate good predictive performance; and values between 0.9 and 1.0 indicate excellent predictive performance. Overall, the closer the value is to 1, the better the model fits (Wang *et al*., 2023).

### *Criteria for delineating BHC potential habitat areas*

The MaxEnt model result file was imported into the ArcGIS software and reclassified by the software's reclassification function so as to derive the possible geographic distribution areas of BHC in China. The distribution values were classified according to the method of assessing probability in the IPCC report, and the suitability areas were categorised into four classes based on the suitability index $P$: high suitability areas ($P \geq 0.66$), medium suitability areas ($0.33 \leq P < 0.66$), low suitability areas ($0.05 \leq P < 0.33$) and unsuitable areas ($P < 0.05$) (Wang *et al*., 2023). The number of rasters in each area was then calculated to obtain the area of suitable habitat in each category for different time periods and climate scenarios. The comparative data and calculated differences were then graphed for more visual analysis.

## Results

### *MaxEnt modelling and accuracy testing*

In this work, the MaxEnt model successfully predicted the current and future suitable habitats for BHC. The model ran for a total of ten iterations, achieving an AUC of 0.993. The training AUC and test AUC were 0.9936 and 0.9931, respectively, with a standard deviation of 0.006. The value then performs excellently according to the criterion of the AUC mentioned in 2.3 (fig. 1). This indicates that the prediction results of the model have a very satisfactory level of reliability. It can scientifically simulate and reflect the geographical distribution of BHC in China, thus providing a basis for further research.

### *Key environmental factors affecting the geographic distribution of BHC*

In MaxEnt's modelling, for key environmental factors, the mean percentage contribution (PC) is usually used to describe their

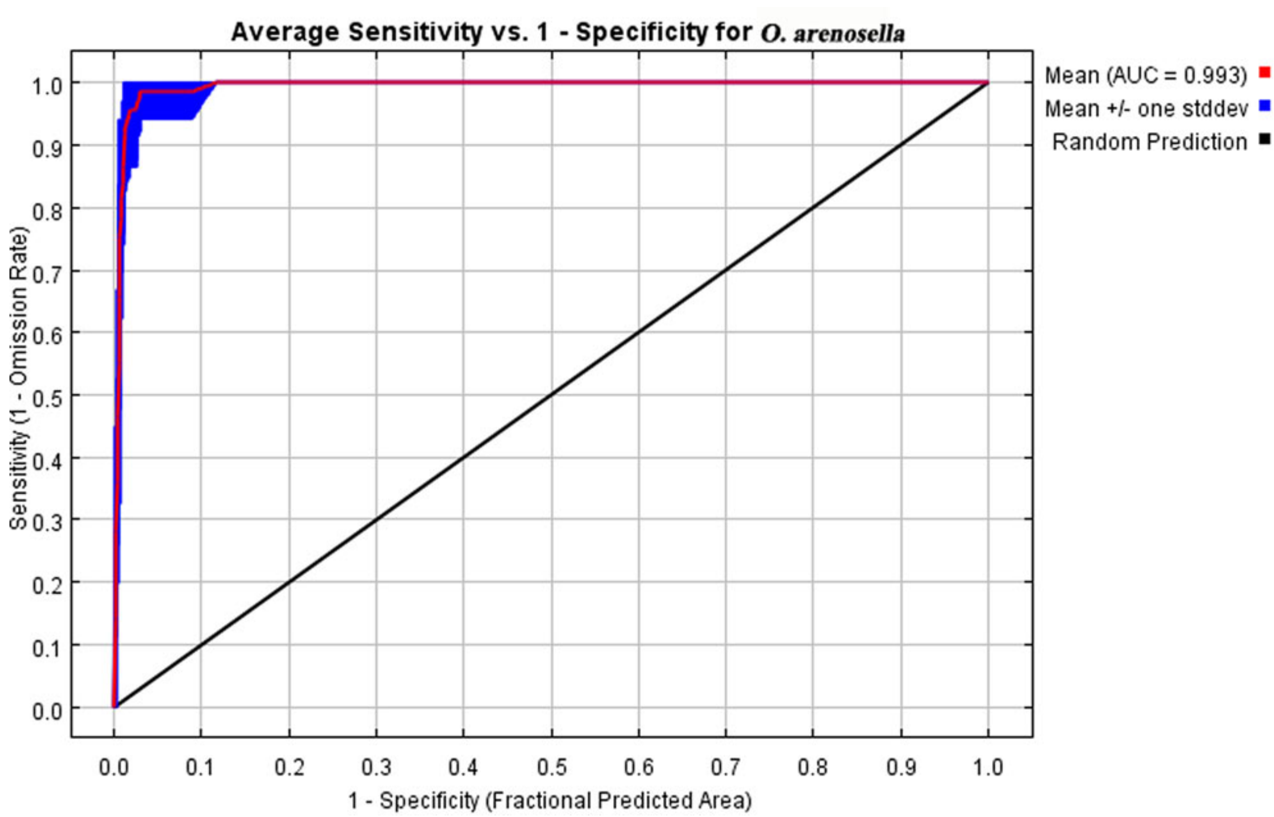


**Figure 1.** Receiver operating characteristic curve and AUC result of MaxEnt modelling.

https://doi.org/10.1017/S0007485324000464 Published online by Cambridge University Press

**Table 1.** Per cent contribution (%) and permutation importance (%) of environmental variables in predicting the occurrence of *O. arenosella* in the MaxEnt model.

| Code | Per cent contribution (%) |
|---|---|
| BIO 1 (annual mean temperature (°C)) | 88 |
| BIO 6 (minimum temperature of coldest month (°C)) | 4.9 |
| BIO 11 (mean temperature of coldest quarter (°C)) | 3.7 |
| BIO 4 (temperature seasonality (standard deviation × 100)) | 3.4 |

importance. According to the screening method mentioned in section 'Acquisition and processing of environment variable data', screening was performed using the Jackknife test and Pearson correlation coefficient. Table 1 presents the PC values of each environmental factor after screening. Among them are the values of four environmental variables, BIO 1 (annual mean temperature), BIO 6 (minimum temperature of the coldest month (°C)), BIO 11 (mean temperature of the coldest quarter (°C)) and BIO 4 (temperature seasonality (standard deviation × 100)), which are four environmental variables with values of 88, 4.9, 3.7 and 3.4%, respectively. It can be seen that they are all closely related to temperature. It is interesting to note that BIO 1 has the highest PC value, 88% for this one alone. This is a huge difference from the other three environmental factors. Figure 2 shows the Jackknife test plot for the MaxEnt model. When used independently, the gain values provided by BIO 1, BIO 11 and BIO 6 all show high values, and only BIO 4 has a large gap in value. It can be concluded that the geographic distribution of BHC in China is largely influenced by these four temperature-related factors, especially BIO 1.

### *Analysis of response curves for environmental factors*

Based on the identification of key environmental factors, this work further investigates them. Response curves were used for each key environmental factor to evaluate its relative impact on MaxEnt predictions. Figure 3 shows the response plots of the MaxEnt model with respect to the key environmental factors. These curves give a good picture of how the different key variables affect the course of the probability of BHC existence while maintaining the average sample values of all other variables. The values of the response curves in the figure are averaged over ten repeated runs of MaxEnt. The response curve of BIO 1 (annual mean temperature) can be seen in fig. 3A. The curve starts to show an upward trend when the temperature rises to about 17.0°C. Above 20.0°C the curve rises sharply until it reaches about 25.0°C, where it remains steady and its output value stabilises at about 0.96. The output value is more informative when it is greater than 0.6. That is to say, when the temperature is greater than 23.6°C and less than 25.0°C, the interval is suitable for the survival of BHC. Moreover, its survival probability is positively correlated with the increase in temperature in this interval, and the survival of the insect is most suitable when the annual mean temperature is 25.0°C. For BIO 4 (temperature seasonality (standard deviation × 100)) (fig. 3B), after the temperature seasonality value of 300, the output value starts to drop off a cliff. Until the temperature seasonality value stabilises around 750, the output value will be 0. This indicates that the temperature seasonality value between 300 and 400 is the most suitable for the survival of the insect, but with the rise of this value, there is a negative correlation with the output value. However, minimum temperature of the coldest month (fig. 3C) is different. As can be seen from the figure, the whole curve shows an increasing trend, with a positive correlation between the higher temperature and the higher output value. The temperature starts to curve up from 5.0°C and reaches a plateau around 17.5°C. When the temperature rises to 13.6°C, the output value is 0.6. It shows that minimum temperature of the coldest month is suitable for the survival of this worm in the interval of 13.6–17.5°C. Similarly, BIO 11 (mean temperature of the coldest quarter) (fig. 3D) showed a positive correlation between the increase in temperature and output values. The interval between 10.0 and 21.8°C was suitable for the survival of BHC.

### *Projections of BHC suitable habitat (current/future)*

Figures 4 and 5 show the contemporary and future distribution of suitable habitat at BHC, respectively. The colouring patterns are mainly in red, orange, yellow and white to indicate highly suitable areas, moderately suitable areas, poorly suitable areas and unsuitable areas, respectively. In general, at any given time, the highly suitable area is mainly found in the island provinces of Hainan and Taiwan. Almost the entire island of Hainan Province is covered by highly suitable areas, especially the sea-rimmed areas, except for a small amount of moderately suitable and poorly suitable areas, of which the poorly suitable area is the smallest.

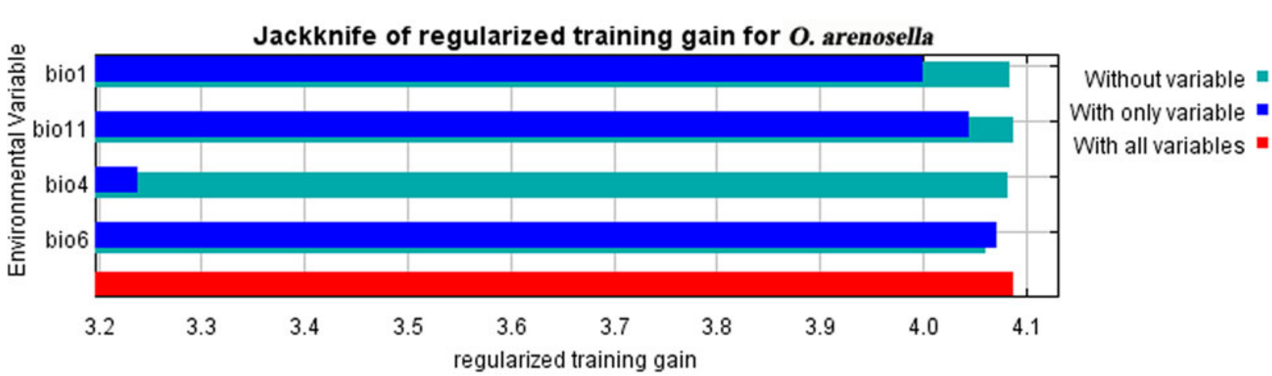


**Figure 2.** Importance of environmental variables to *O. arenosella* by jackknife test.

https://doi.org/10.1017/S0007485324000464 Published online by Cambridge University Press

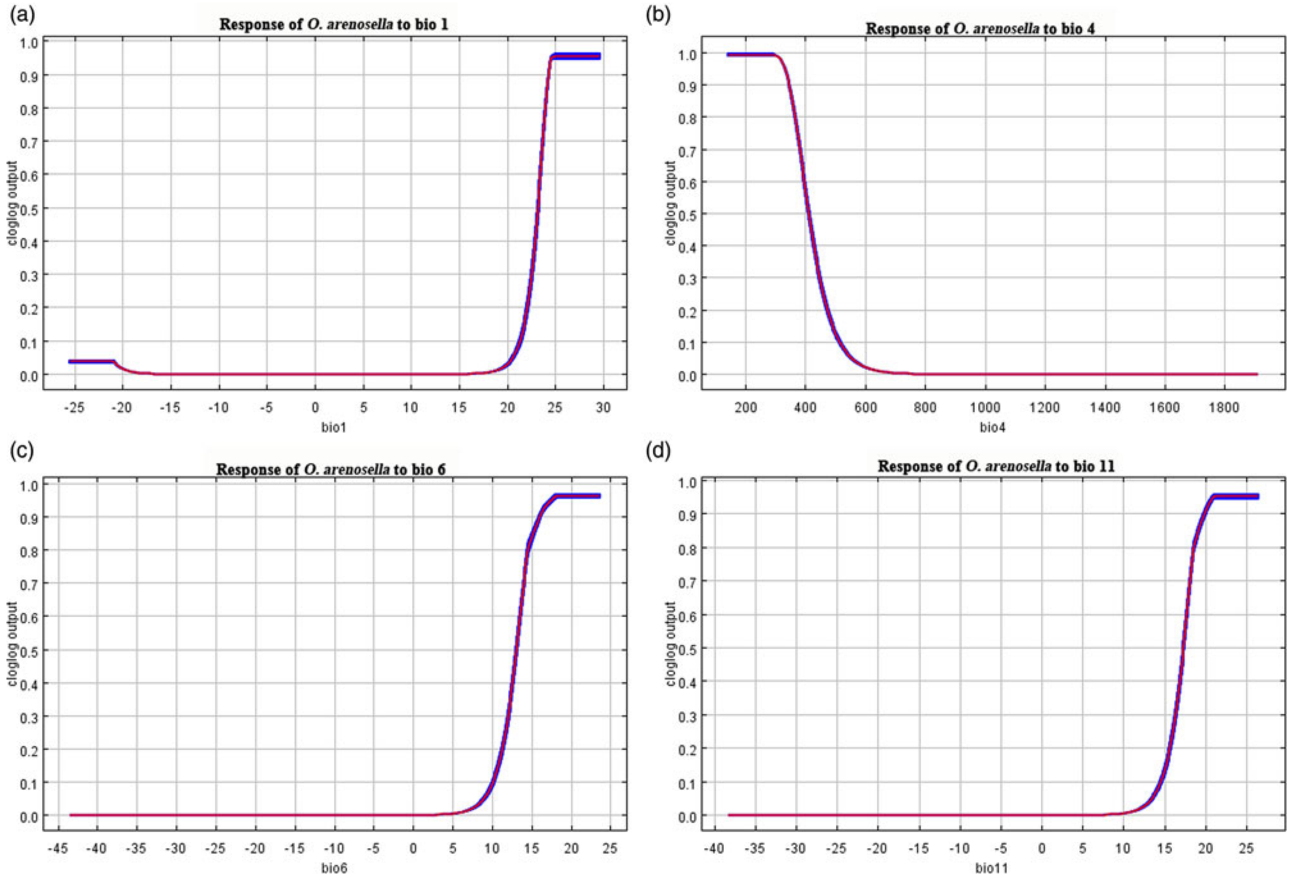


**Figure 3.** Response curves of the environmental variables that contributed most to the MaxEnt models. (A) Annual mean temperature (Bio 1); (B) temperature seasonality (standard deviation × 100) (Bio 4); (C) coldest month minimum temperature (Bio 6); (D) mean temperature of coldest quarter (°C) (Bio 11).

In terms of the contemporary area of suitable distribution areas, table 2 shows us the area of different suitable areas and the percentage of each province. Table 2 shows that not many provinces are suitable for BHC distribution, mainly Tibet, Yunnan, Fujian, Guangxi, Guangdong, Hong Kong, Taiwan and Hainan, of which only a small amount of moderately suitable area, poorly suitable area and poorly suitable area exist in Tibet, Yunnan and Guangdong, while only a small amount of moderately suitable area, poorly suitable area and poorly suitable area exist in Fujian, Guangxi and Hong Kong. Among them, only a few moderately suitable and poorly suitable areas exist in Tibet, Yunnan and Guangdong, while only a few poorly suitable areas exist in Fujian, Guangxi and Hong Kong. The highly suitable areas coincide with those presented in fig. 4, with highly suitable areas existing only in Hainan and Taiwan Provinces. The percentages of highly suitable areas in Hainan and Taiwan Provinces were 59.6 and 5.8%, respectively, of the province, with the highest percentage in Taiwan Province. In terms of the future size of the suitable distribution area, table 3 shows the distribution of the area for two future periods and three climate scenarios, as well as the increase or decrease in comparison with contemporary data. Notably, the highly suitable area of BHC showed an increase under all three climate scenarios in the 2050s (SSP1-2.6, SSP2-4.5 and SSP5-8.5), while the other two suitable areas showed an overall decrease in area. Interestingly, the overall trend was consistent in the 2090s compared to the 2050s, although the area did not increase or decrease as consistently across climate scenarios as it did in the 2050s. Of these, the highly suitable area has the largest increase in area under the 2090s SSP5-8.5 scenario, with an increase of 24.03%. The largest decrease in area was in moderately suitable area under the 2090s SSP2-4.5 scenario, with a decrease of 49.40%. Overall, the highly suitable area will continue to increase in size in the future, while the moderately suitable area and poorly suitable area will decrease to varying degrees.

## Discussion

In today's scientific research, SDM has been applied in a variety of scientific fields, such as ecology, biology and geography (Elith and Leathwick, 2009). It provides a key approach to the control of harmful invasive species (Bowen and Stevens, 2020) and the conservation of rare and endangered species (Escalera *et al.*, 2018), and can also be used to predict potential impacts on target species by assessing changes in climate (Nneji *et al.*, 2019). Most SDM methods can be used to form statistical relationships between species occurrence data and environmental variables and for prediction. However, in this study, the MaxEnt model was selected as a prediction tool for the invasive species BHC among many SDMs. MaxEnt is an integration of environmental variables as a simulation of the general environment, and the presence and location of the target species can be predicted using very little species occurrence data. It includes the density probability of presence location and further derives the proportion of area present in different provinces and survival suitability (Elith *et al.*, 2011). Because of these superiorities of MaxEnt as a predictive model, the model was chosen to predict its geographic distribution for BHC.

https://doi.org/10.1017/S0007485324000464 Published online by Cambridge University Press

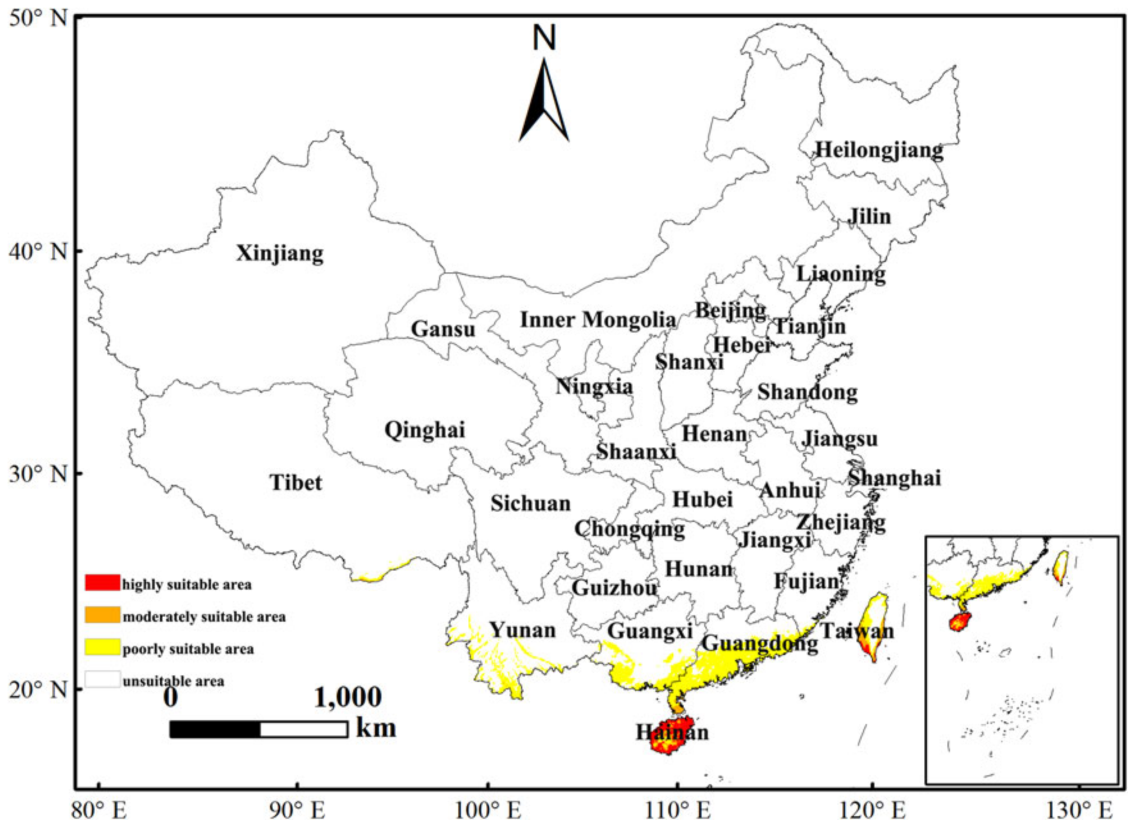


**Figure 4.** Current suitable climatic distribution of *O. arenosella* in China. The probability of *O. arenosella* is shown by the colour scale in the area. Red indicates a highly suitable area with a probability of higher than 0.66, orange indicates a moderately suitable area with a probability of 0.33–0.66, yellow indicates a poorly suitable area with a probability ranging from 0.05 to 0.33, and white represents unsuitable areas.

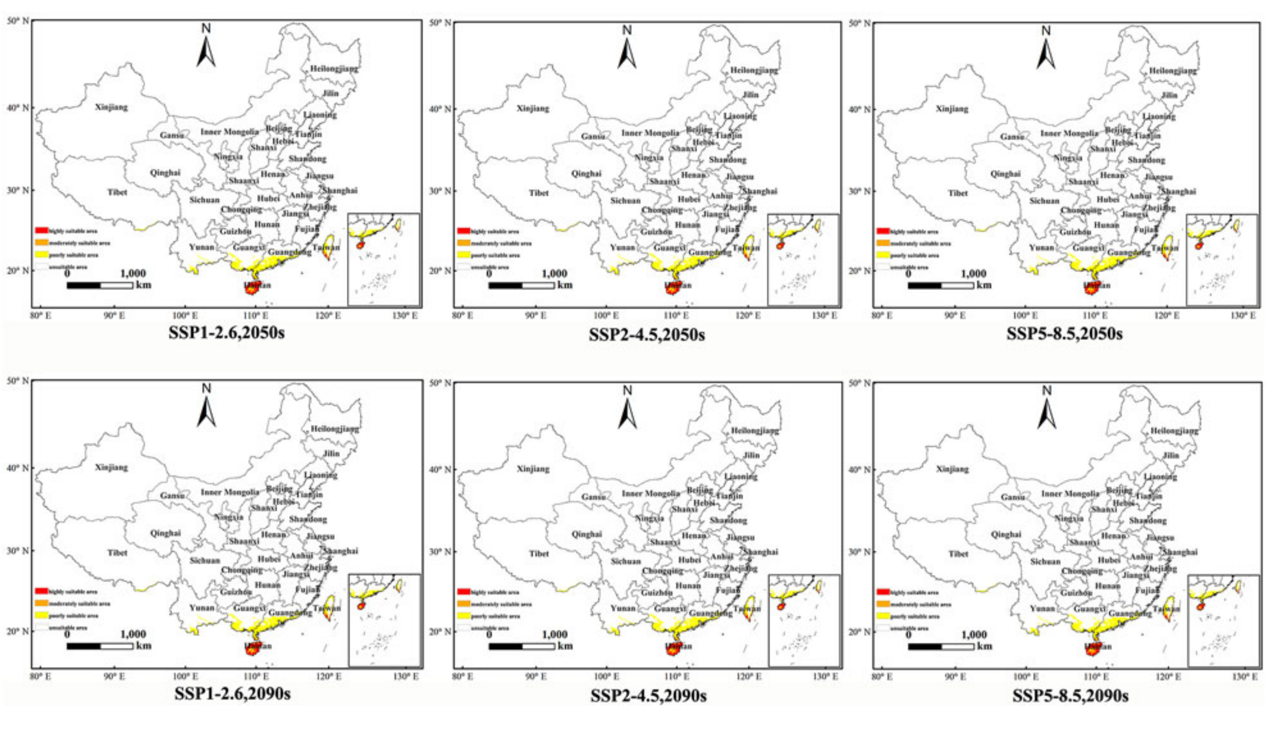


**Figure 5.** Potential distribution of suitable areas of *O. arenosella* under different climate change scenarios in China. The probability of *O. arenosella* is shown by the colour scale in the area. Red indicates a highly suitable area with a probability of higher than 0.66, orange indicates a moderately suitable area with a probability of 0.33–0.66, yellow indicates a poorly suitable area with a probability ranging from 0.05 to 0.33, and white represents unsuitable areas.

**Table 2.** Analysis of main suitable distributions of *O. arenosella* in China.

| Province | Highly suitable area ($10^4$ km$^2$) | Moderately suitable area ($10^4$ km$^2$) | Poorly suitable area ($10^4$ km$^2$) | Unsuitable area ($10^4$ km$^2$) | Total ($10^4$ km$^2$)[a] | Percentage of high-suitable area in the province (%) |
|---|---|---|---|---|---|---|
| Heilongjiang | / | / | / | 47.30 | 47.30 | 0 |
| Inner Mongolia | / | / | / | 118.30 | 118.30 | 0 |
| Xinjiang | / | / | / | 166.49 | 166.49 | 0 |
| Jilin | / | / | / | 18.74 | 18.74 | 0 |
| Liaoning | / | / | / | 14.80 | 14.80 | 0 |
| Gansu | / | / | / | 41.53 | 42.58 | 0 |
| Hebei | / | / | / | 18.88 | 18.88 | 0 |
| Beijing | / | / | / | 1.64 | 1.64 | 0 |
| Shanxi | / | / | / | 15.67 | 15.67 | 0 |
| Tianjin | / | / | / | 1.20 | 1.20 | 0 |
| Shaanxi | / | / | / | 20.38 | 20.56 | 0 |
| Ningxia | / | / | / | 5.27 | 6.64 | 0 |
| Qinghai | / | / | / | 71.34 | 72.23 | 0 |
| Shandong | / | / | / | 15.39 | 15.58 | 0 |
| Tibet | / | 0.07 | 0.35 | 113.90 | 122.84 | 0 |
| Henan | / | / | / | 16.13 | 16.70 | 0 |
| Jiangsu | / | / | / | 9.75 | 10.72 | 0 |
| Anhui | / | / | / | 13.36 | 14.01 | 0 |
| Sichuan | / | / | / | 45.53 | 48.60 | 0 |
| Hubei | / | / | / | 17.56 | 18.88 | 0 |
| Chongqing | / | / | / | 7.74 | 8.24 | 0 |
| Shanghai | / | / | / | 0.59 | 0.63 | 0 |
| Zhejiang | / | / | / | 9.46 | 10.555 | 0 |
| Hunan | / | / | / | 19.39 | 21.18 | 0 |
| Jiangxi | / | / | / | 15.27 | 16.69 | 0 |
| Yunnan | / | 0.02 | 2.89 | 31.37 | 39.41 | 0 |
| Guizhou | / | / | / | 15.96 | 17.62 | 0 |
| Fujian | / | / | 0.23 | 10.76 | 12.40 | 0 |
| Guangxi | / | / | 3.32 | 17.64 | 23.76 | 0 |
| Taiwan | 0.21 | 0.61 | 1.01 | 1.36 | 3.60 | 5.8 |
| Guangdong | / | 0.42 | 7.43 | 7.76 | 17.97 | 0 |
| Hong Kong | / | / | 0.09 | / | 0.11 | 0 |
| Hainan | 2.11 | 0.56 | 0.24 | 0.01 | 3.54 | 59.6 |

[a]Indicates the total area of the corresponding province.

However, all models have certain shortcomings based on their strengths, and MaxEnt is no exception, with certain recognised performance limitations. For example, in terms of background samples, although a small sample size is sufficient to support the completion of the modelling, the representativeness of the sample selection affects the prediction results. But the model's currently known limitations can be mitigated relatively well. For example, screening for environmental factors effectively avoids the question of whether the background sample is representative. For example, for missing species occurrence data, MaxEnt can be modelled with easily accessible presence-only datasets (Elith *et al.*, 2011).

The predicted results for BHC show that temperature is an important factor influencing the distribution of this pest. The key four environmental factors (BIO 1, BIO 6, BIO 11 and BIO 4) were all related to temperature, with BIO 1 (Annual Mean Temperature) being the most influential factor. There is an effect of climate change on the population dynamics of BHC in the field (Ramkumar *et al.*, 2006; Shameer *et al.*, 2018). It has been shown that lower temperatures can have a negative effect

https://doi.org/10.1017/S0007485324000464 Published online by Cambridge University Press

**Table 3.** Prediction of the suitable areas for *O. arenosella* under current and future climatic conditions.

| | | Predicted area ($10^4$ km$^2$) | | | Comparison with current distribution (%) | | |
|---|---|---|---|---|---|---|---|
| Decade | Scenarios | Highly suitable area | Moderately suitable area | Poorly suitable area | Highly suitable area | Moderately suitable area | Poorly suitable area |
| Current | | 2.33 | 1.68 | 15.55 | | | |
| 2050s | SSP1-2.6 | 2.47 | 1.22 | 11.85 | 6.01 | −27.38 | −23.7 |
| | SSP2-4.5 | 2.62 | 1.23 | 13.97 | 12.45 | −26.79 | −10.16 |
| | SSP5-8.5 | 2.71 | 1.09 | 17.61 | 16.31 | −35.12 | 13.25 |
| 2090s | SSP1-2.6 | 2.28 | 1.90 | 16.16 | −2.15 | 13.10 | 3.92 |
| | SSP2-4.5 | 2.62 | 0.85 | 13.84 | 12.45 | −49.40 | −11.00 |
| | SSP5-8.5 | 2.89 | 0.99 | 13.69 | 24.03 | −41.07 | −11.96 |

on population growth. At lower temperatures, the spawning period of BHCs is prolonged, which reduces population growth (Muralimohan *et al*., 2013). This happens to also corroborate the response curve for the most influential BIO 1 environmental factor (PC = 88%), i.e. temperature is positively correlated with the probability of presence when the average temperature of the year is in the interval 23.6–25.0°C.

In summary, there are not many highly suitable areas for BHC in China, nor is there a trend of rapid expansion. The suitable areas are mainly located in the cities along the southeast coast, with the highly suitable areas existing mainly in Hainan Province and Taiwan Province. It is noteworthy that the whole province of Hainan and the southern part of Taiwan Province have a tropical monsoon climate. This coincides with the distribution of high suitability areas of BHC in these two provinces. Regarding areas with a tropical monsoon climate, the temperature is mainly characterised by high temperatures throughout the year, which again confirms the previous point. For invasive pests, early detection is an important method for effective containment, eradication and control of target species. Therefore, the prediction work in this study provides directions for prevention in places where invasive pest species have not yet been detected. Places that may have suitable areas for BHC in the future should be detected in time to strengthen the precaution. According to the habit of this pest, it usually provides a protective barrier by making channels on the surface under the leaves using excreta. Therefore, the use of insecticides to control this pest is not very desirable, either in terms of cost or feasibility. As early as 1924, attempts at biological control were made in India (Jayanth and Nagarkatti, 1984). In the future, rigorous testing of areas where this is likely to occur and the timely placement of natural enemies will be a more effective program.

This study combines 19 bioclimatic variables modelled using the MaxEnt model to explore both the contemporary and future geographic distribution of BHC in China. The bioclimatic variables are mainly composed of two elements: temperature and precipitation. It can well simulate the climate conditions in different situations and is also a more scientific indicator of climate nowadays. Meanwhile, the collection of data on the occurrence of BHC was screened and integrated. On this basis, the MaxEnt model was used to predict the potential distribution area of BHC in China and effectively reflect the changes in the size of the suitable area.

## Conclusion

In this study, we successfully simulated the potential geographic distribution of the BHC under three climate change scenarios (SSP1-2.6, SSP2-4.5 and SSP5-8.5) for the current and future periods (2050s and 2090s) based on the MaxEnt model. Under the current climate conditions, the BHC is distributed only in the provinces of Tibet, Yunnan, Fujian, Guangxi, Taiwan, Guangdong, Hong Kong and Hainan. Among them, high suitability zones existed only in Hainan and Taiwan. The total areas of high, medium and low suitability zones were $2.32 \times 10^4$, $1.68 \times 10^4$ and $15.56 \times 10^4$ km$^2$, respectively. The environmental factor that had the greatest impact on the geographic distribution of the bug was BIO 1 (annual mean temperature), followed by BIO 6 (minimum temperature of the coldest month), BIO 11 (mean temperature of the coldest quarter) and BIO 4 (temperature seasonality). All these factors are related to temperature. In other words, temperature has an important effect on the distribution of BHCs, which is also consistent with other related studies. The expected temperature increase is projected to lead to the geographic expansion of this pest's distribution. Also, the potential suitable distribution areas predicted by the BHC were generally consistent with the actual distribution areas. This work aims to provide new control perspectives for the invasive pest, the BHC.

## Data

The data supporting the results are available in a public repository from: Zhiling Wang. List of locations used for *Opisina arenosella*, with longitude, latitude and sources. figshare. Dataset. https://doi.org/10.6084/m9.figshare.25188527.v1. EPPO (2023) *Opisina arenosella*. EPPO datasheets on pests recommended for regulation. https://gd.eppo.int (accessed 21 May 2023), and GBIF.org (08 February 2024) GBIF Occurrence Download https://www.gbif.org/occurrence/download/0007921-240202131308920.

**Author contributions.** Conceptualisation, Zhiling Wang and Danping Xu; data curation, Yaqin Peng; formal analysis, Zhiling Wang and Yaqin Peng; investigation, Zhiling Wang and Yaqin Peng; methodology, Zhiling Wang; software, Yaqin Peng and Danping Xu; supervision, Zhihang Zhuo and Danping Xu; writing – original draft, Zhiling Wang; writing – review and editing, Zhihang Zhuo and Danping Xu. All authors have read and agreed to the published version of the manuscript.

**Financial support.** This work was funded by the Ministry of Sichuan Province Science and Technology Support Program (2022NSFSC0986) and

China West Normal University Support Program (20A007, 20E051, 21E040 and 22kA011).

**Competing interests.** None.

**Institutional review board statement.** Not applicable.

**Informed consent statement.** Not applicable.

## References

**Barker JR and MacIsaac HJ** (2022) Species distribution models: administrative boundary centroid occurrences require careful interpretation. *Ecological Modelling* **472**, 110107. doi: 10.1016/j.ecolmodel.2022.110107.

**Bowen A and Stevens M** (2020) Temperature, topography, soil characteristics, and NDVI drive habitat preferences of a shade-tolerant invasive grass. *Ecology and Evolution* **10**, 10785. doi: 10.1002/ece3.6735.

**Bradie J and Leung B** (2017) A quantitative synthesis of the importance of variables used in MaxEnt species distribution models. *Journal of Biogeography* **44**, 1344–1361. doi: 10.1111/jbi.12894.

**Breiman L** (2001) Random forests. *Machine Learning* **45**, 5–32. doi: 10.1023/A:1010933404324.

**Breiman L, Friedman J, Olshen RA and Stone CJ** (1984) *Classification and Regression Trees*. United states: Springer Nature, 368 pp.

**De'Ath G** (2007) Boosted trees for ecological modeling and prediction. *Ecology* **88**, 243–251. doi: 10.1890/0012-9658(2007)88[243:BTFEMA]2.0.CO,2.

**Elith J and Leathwick JR** (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* **40**, 677–697. doi: 10.1146/annurev.ecolsys.110308.120159.

**Elith J, Graham CH, Anderson RP, Dudík M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR and Lehmann A** (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151. doi: 10.1111/j.2006.0906-7590.04596.x.

**Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE and Yates CJ** (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* **17**, 43–57. doi: 10.1111/j.1472-4642.2010.00725.x.

**Escalera L, Hernandez R, Soto-Rojas C and Suazo-Ortuño I** (2018) Predicting *Ambystoma ordinarium* habitat in central Mexico using species distribution models. *Herpetologica* **74**, 117–126. doi: 10.1655/Herpetologica-D-16-00078.1.

**Friedman JH** (1991) Multivariate adaptive regression splines. *Annals of Statistics* **19**, 1–141. doi: 10.1214/aos/1176347963.

**Guisan A and Zimmermann NE** (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* **135**, 147–186. doi: 10.1016/S0304–3800(00)00354-9.

**Guisan Jr. A, Edwards TC and Hastie T** (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling* **157**, 89–100. doi: 10.1016/S0304-3800(02)00204-1.

**Guisan A, Zimmermann NE, Elith J, Graham CH, Phillips S and Peterson AT** (2007) What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? *Ecological Monographs* **77**, 615–630. doi: 10.1890/06-1060.1.

**Guo Q, Kelly M and Graham CH** (2005) Support vector machines for predicting distribution of Sudden Oak Death in California. *Ecological Modelling* **182**, 75–90. doi: 10.1016/j.ecolmodel.2004.07.012.

**Jayanth KP and Nagarkatti S** (1984) *Testingbessa remota* [Dip.: Tachinidae] against *Opisina arenosella* [Lep.: Cryptophasidae] and other lepidopterous hosts in India. *Entomophaga* **29**, 415–419. doi: 10.1007/BF02372163.

**Jin T, Lin Y, Han S, Ma G, Wen H and Peng Z** (2019) Host performance of Trichogramma species on *Opisina arenosella*, and evaluation of their biological control efficacy. *Journal of Asia-Pacific Entomology* **22**, 990–996. doi: 10.1016/j.aspen.2019.06.010.

**Jingjing D, Siming F, Jiaqiang Z, Juan S and Te L** (2023) Analysis of insect pests identified to species that were intercepted at Chinese ports of entry. *Management of Biological Invasions* **14**, 641–657. doi: 10.3391/mbi.2023.14.4.05.

**Kumar R** (2002) *Studies on Ecology, Reproductive Biology and Management of Opisina arenosella (Walker) (Lepidoptera: Studies on Ecology, Reproductive Biology and Management of Opisina arenosella (Walker) (Lepidoptera: Oecophoridae)* (Doctor Type). University of Agricultural Sciences, Bangalore.

**Lever RJAW** (1969) *Pests of the Coconut Palm*. Finland: FAO Agricultural Study.

**Li L, Qingzhao L, Xiongfei Y, Guangchang M, Bo L, Fan Y and Wei Y** (2021) Observation on the ultrastructures of antennal sensilla in *Opisina arenosella*. *Journal of Environmental Entomology* **43**, 1070–1078.

**Lu H, Lyu B, Tang J, Wu Q, Wyckhuys KAG, Le KH, Chongchitmate P, Qiu H and Zhang Q** (2023) Ecology, invasion history and biodiversity-driven management of the coconut black-headed caterpillar *Opisina arenosella* in Asia. *Frontiers in Plant Science* **14**. doi: 10.3389/fpls.2023.1116221.

**Mao H and Kunimi Y** (1994) Longevity and fecundity of *Brachymeria lasus* (WALKER) (Hymenoptera: Chalcididae), a pupal parasitoid of the Oriental Tea Tortrix, *Homona magnanima* DIAKONOFF (Lepidoptera: Tortricidae) under laboratory conditions. *Applied Entomology and Zoology* **29**, 237–243.

**Mohan C, Nampoothiri CK, Radhakrishnan Nair CP and Rajan P** (2010) Leaf-eating caterpillar (*Opisina arenosella*)-induced yield loss in coconut palm. *International Journal of Tropical Insect Science* **30**, 132–137. doi: 10.1017/S174275841000024X.

**Muralimohan K, Srinivasa YB and Pn S** (2013) Effect of temperature on developmental biology of *Opisina arenosella* (Walker). *Current Biotica* **7**, 228–232.

**Nasser M and Abdurahiman UC** (2001) Biological control of the coconut caterpillar *Opisina arenosella* (Lepidoptera: Xylorictidae): achievements and prospects. In Upadhyay RK, Mukerji KG and Chamola BP (eds), *Biocontrol Potential and its Exploitation in Sustainable Agriculture: Volume 2: Insect Pests*. Boston, MA: Springer US, pp. 285–305.

**Nneji LM, Salako G, Oladipo S, Ayoola A, Onadeko A, Adedeji B, Omotoso O, Ugwumba A and Adeola A** (2019) Species distribution modelling predicts habitat suitability and reduction of suitable habitat under future climatic scenario for *Sclerophrys perreti*: a critically endangered Nigerian endemic toad. *African Journal of Ecology* **58**, 481–491. doi: 10.1111/aje.12713.

**Olden JD, Lawler JJ and Poff NL** (2008) Machine learning methods without tears: a primer for ecologists. *The Quarterly Review of Biology* **83**, 171–193.

**Perera PACR, Hassell MP and Godfray HCJ** (2009) Population dynamics of the coconut caterpillar, *Opisina arenosella* Walker (Lepidoptera: Xyloryctidae), in Sri Lanka. *Bulletin of Entomological Research* **78**, 479–492. doi: 10.1017/S0007485300013237.

**Phillips SJ, Dudík M and Schapire R** (2004) A maximum entropy approach to species distribution modeling. *International Conference on Machine Learning*, 655–663. doi: 10.1145/1015330.1015412.

**Phillips SJ, Anderson RP and Schapire RE** (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231–259. doi: 10.1016/j.ecolmodel.2005.03.026.

**Ramkumar, Muralimohan K, Kiranmayi L and Srinivasa YB** (2006) Discrete generation cycles in the tropical moth *Opisina arenosella*. *Current Science India*, **91**, 811–816.

**Rhoden CM, Peterman WE and Taylor CA** (2017) Maxent-directed field surveys identify new populations of narrowly endemic habitat specialists. *Peerj* **5**, e3632. doi: 10.7717/peerj.3632.

**Shameer KS, Nasser M, Mohan C and Hardy ICW** (2018) Direct and indirect influences of intercrops on the coconut defoliator *Opisina arenosella*. *Journal of Pest Science* **91**, 259–275. doi: 10.1007/s10340-017-0904-6.

**Stockwell D** (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science* **13**, 143–158. doi: 10.1080/136588199241391.

**Tang Y** (2017) *Studies on the Suitability of Coconut Weaver Moth to Four Host Plants and its Distribution Pattern on Coconut Trees* (Master Type). Hainan University, China.

**Wang Z, Xu D, Liao W, Xu Y and Zhuo Z** (2023) Predicting the current and future distributions of *Frankliniella occidentalis* (Pergande) based on the MaxEnt species distribution model. *Insects* **14**, 458. doi: 10.3390/insects14050458.