*Bulletin of Entomological Research*

**cambridge.org/ber**

# Risk of spread of the Asian citrus psyllid *Diaphorina citri* Kuwayama (Hemiptera: Liviidae) in Ghana

Kodwo Dadzie Ninsin[1], Philipe Guilherme Corcino Souza[2], George Correa Amaro[3], Owusu Fordjour Aidoo[1,4], Edmond Joseph Djibril Victor Barry[5], Ricardo Siqueira da Silva[5], Jonathan Osei-Owusu[6], Aboagye Kwarteng Dofuor[1], Fred Kormla Ablormeti[7], William K. Heve[1], George Edusei[6], Lakpo Koku Agboyi[8], Patrick Beseh[9], Hettie Arwoh Boafo[8], Christian Borgemeister[10] and Mamoudou Sétamou[11]

[1]Department of Biological Sciences, School of Natural and Environmental Sciences, University of Environment and Sustainable Development, PMB, Somanya, E/R, Ghana; [2]Department of Agronomy, Instituto Federal de Ciência e Tecnologia do Triângulo Mineiro (IFTM *Campus* Uberlândia), Uberlândia, MG 38400-970, Brazil; [3]Embrapa Roraima, Boa Vista, Roraima 69301-970, Brazil; [4]Department of Entomology, College of Agricultural, Human, and Natural Resource Sciences, Washington State University, Pullman, WA 99164, USA; [5]Department of Agronomy, Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Diamantina, MG 39100-000, Brazil; [6]Department of Physical and Mathematical Sciences, School of Natural and Environmental Sciences, University of Environment and Sustainable Development, PMB, Somanya, E/R, Ghana; [7]Council for Scientific and Industrial Research (CSIR), P. O. Box 245, Sekondi, W/R, Ghana; [8]Centre for Agriculture and Biosciences International (CABI), CSIR Campus, No. 6 Agostino Neto Road, Airport Residential Area, P. O. Box CT 8630, Cantonments, Ghana; [9]Plant Protection and Regulatory Services Directorate. P. O. Box M37, Accra, Ghana; [10]Centre for Development Research (ZEF), University of Bonn, Genscherallee 3, 53113 Bonn, Germany and [11]Citrus Center, Texas A & M University-Kingsville, 312 N. International Blvd., Weslaco, TX 78599, USA

**Abstract**

The impact of invasive species on biodiversity, food security and economy is increasingly noticeable in various regions of the globe as a consequence of climate change. Yet, there is limited research on how climate change affects the distribution of the invasive Asian citrus psyllid *Diaphorina citri* Kuwayama (Hemiptera:Liviidae) in Ghana. Using maxnet package to fit the Maxent model in R software, we answered the following questions; (i) what are the main drivers for *D. citri* distribution, (ii) what are the *D. citri*-specific habitat requirements and (iii) how well do the risk maps fit with what we know to be correctly based on the available evidence?. We found that temperature seasonality (Bio04), mean temperature of warmest quarter (Bio10), precipitation of driest quarter (Bio17), moderate resolution imaging spectroradiometer land cover and precipitation seasonality (Bio15), were the most important drivers of *D. citri* distribution. The results follow the known distribution records of the pest with potential expansion of habitat suitability in the future. Because many invasive species, including *D. citri*, can adapt to the changing climates, our findings can serve as a guide for surveillance, tracking and prevention of *D. citri* spread in Ghana.

## Introduction

Species from one place can move to another due to anthropogenically driven activity and natural dispersal (Padayachee *et al.*, 2017). Alien species have the potential to become established and expand after being introduced to a new area, leading to ecological destruction and economic losses (Nahrung *et al.*, 2023; Uden *et al.*, 2023). There has been a dramatic rise in the prevalence of invasive species during the nineteenth century (Simberloff and Gibbons, 2004). The invasiveness of these species is influenced by many factors, including climatic conditions, host plant resistance, propagule pressure and natural enemies in the invaded area (Skendžić *et al.*, 2021).

Invasive pest population sizes, survival rates, geographic distributions, disease prevalence and proliferation can all be regulated by climatic factors (Aidoo *et al.*, 2021, 2022, 2023a). Their populations, mainly in temperate regions, are expected to increase worldwide due to climate change (Schneider *et al.*, 2022). However, it still needs to be determined how much of a role climate change plays in determining invasive species like the invasive Asian citrus psyllid *Diaphorina citri* Kuwayama (Hemiptera: Liviidae) in Ghana.

*Diaphorina citri* is a sap-sucking hemipteran pest of citrus and its close relatives. The psyllid undergoes five nymphal stages, which may take about 49 days, depending on the thermal

## Research Paper

**Cite this article:** Ninsin KD *et al* (2024). Risk of spread of the Asian citrus psyllid *Diaphorina citri* Kuwayama (Hemiptera: Liviidae) in Ghana. *Bulletin of Entomological Research* 1–20. https://doi.org/10.1017/S0007485324000105

Received: 18 April 2023
Revised: 26 November 2023
Accepted: 13 February 2024

**Keywords:**
Asian citrus psyllid; climate change; *Diaphorina citri*; maxnet; species distribution modeling

**Corresponding author:**
Owusu Fordjour Aidoo;
Email: ofaidoo@uesd.edu.gh; owusufordjour.aidoo@wsu.edu

conditions (Tsai and Liu, 2000). Apart from temperature, other factors that can influence the development and survival of the psyllid include availability of host plants and abundant flush growth. *Diaphorina citri* prefers warm and dry climates where conditions are suitable for developing all life stages (Hall *et al.*, 2013).

The physical damage caused by the psyllid through feeding leads to the formation of sooty molds, and leaf distortion. Indirectly, the psyllid transmits the invasive pathogens; '*Candidatus* Liberibacter asiaticus' (*C*Las) and '*Ca.* L. americanus' (*C*Lam) (Bové, 2006). '*Candidatus* Liberibacter asiaticus' occurs in Africa, Asia and the Americas, whereas *C*Lam is distributed in Brazil. These phloem-limited pathogens have been associated with the deadliest disease of citrus called citrus greening disease or Huanglongbing (HLB). The interplay between the pathogens and the local accumulation of callose affects the movement of phloem in the plant (Welker *et al.*, 2022).

There is no effective method for preventing or curing HLB. Symptoms of HLB infection may not appear for months or even years, making it difficult to detect in affected trees (Lee *et al.*, 2015). Producing seedlings from *C*Las-free nursery stock, removing affected trees and applying insecticides to curb vectors are all part of HLB management. However, if pesticides are used extensively, pest resistance to pesticides may develop (Naeem *et al.*, 2016).

In 2023, for the first time, Ghana reported *D. citri* (Hemiptera: Liviidae) (Aidoo *et al.*, 2023b). The psyllids were obtained from orange jasmine *Murraya paniculata* (L.) in one of Ghana's 16 regions. However, none of the samples tested positive for the phloem-limited bacteria; *C*Las, *C*Lam and *C*Laf. Risk maps of *D. citri*, which can serve as a guide for sampling and tracking, are urgently needed to develop quarantine and preventive measures. *D. citri* has been the subject of many species' distribution models. For example, Aidoo *et al.* (2022) used the Maxent model to project where *D. citri* might develop in the world given the current and future climates. The potential distribution of *D. citri* was estimated in China using the Maxent and CLIMEX models (Wang *et al.*, 2015, 2020). The Maxent model has a friendly user interface and works well when only presence and fewer occurrence records are available (Aidoo *et al.*, 2019, 2022).

The newly developed package maxnet in R is used to estimate Maxent models for regularised generalised linear models (Friedman *et al.*, 2010). Maxnet employs the R package glmnet instead of relying on external sources such as the Maxent Java programme. It incorporates feature classes and regularization settings to effectively fit models comparable to those generated by the Maxent Java application (Phillips *et al.*, 2017). This enables enhanced integration of Maxent modelling with R's diverse range of visualization and analysis tools. The package helps to incorporate all the derived feature classes, with a particular emphasis on hinge features, as well as the default configured regularization settings from the Maxent Java programme into the maxnet package, thereby enabling the seamless and straightforward fitting of Maxent models within the R environment (Phillips *et al.*, 2017). The performance of maxnet is similar to that of the Maxent Java programmes (Phillips and Phillips, 2021).

Understanding areas suitable for an invasive insect pest is a prerequisite for surveillance and tracking programmes. Using local distribution information of a species to examine its geographic distribution offers excellent accuracy when operating under the niche conservation tenet (Phillips, 2017). However, limiting occurrence records to only local records seems problematic given that *D. citri* has only recently become established in Ghana (Aidoo *et al.*, 2023b). These local records may not represent its ultimate niche in the country for accurate model predictions. Herein, we used the maxnet package in R to estimate the future spread risk for two shared socioeconomic pathways (SSPs) based on local and global occurrence records and climate datasets, and determine the climatic factors influencing the distribution of the pest. Also, we predicted the current suitable regions using local and global records to help policymakers develop effective responses to *D. citri*-borne disease.

## Material and methods

### Occurrence data

Four hundred and seventy (470) global occurrence records were sourced from Aidoo *et al.* (2019), Aidoo *et al.* (2023b) and Sétamou *et al.* (2023). These records were supplemented by new field surveys in Ghana, and by data obtained from the Global Biodiversity Information Facility (GBIF). For the former, citrus and non-citrus host plants in the Western, Central, Eastern, Greater Accra, Ashanti and Volta Regions of Ghana were surveyed for the presence of the pest. If *D. citri* was found, Garmin eTrex® 32 × was used to obtain the Geographical Processing System (GPS) coordinates. The survey yielded 28 occurrence points. The occurrence records from GBIF were obtained using the {rgbif} package version 3.7.7 (Chamberlain *et al.*, 2023). Only records with a spatial resolution $\leq 1\ km^2$ were retained for analysis. When only the names of locations were identified in the publications, their geographic coordinates were obtained online with Google Maps. The dataset was then formed by 498 occurrence points compiled from literature and field data, and 516 acquired from GBIF.

Occurrence records with a radius of 1 km around the capitals and centres of countries, equal absolute longitude and latitude, a radius of one degree around the GBIF headquarters, duplicated coordinates and zero values were removed, in addition to verifying that the informed coordinates corresponded to a valid coordinate reference system, using the clean_coordinate function of the {CoordinateCleaner} package version 2.0-20 (Zizka *et al.*, 2019) (table S1). After these procedures, of the total 1014 points, 707 points of occurrence remained. After extracting environmental data in their coordinates and eliminating those in which at least one of the selected variables did not present information, the final set of occurrences was 704 points, of which 628 corresponded to invaded areas around the world, and others referred to *D. citri* native areas in South and Southeast Asia (e.g. Bangladesh, Bhutan, India, Nepal, Pakistan and southern China) (Bayles *et al.*, 2017; Bové, 2014; Luo and Agnarsson, 2018).

To further mitigate the bias in occurrence data sampling, for the variable selection process, a spatial filter based on a 5 km distance was applied after data cleaning, considering the $2.5\ km^2$ resolution of the used variables. For the final model, the 704 points of presence were filtered by applying an environmental filter (Varela *et al.*, 2014; Castellanos *et al.*, 2019; Velazco *et al.*, 2020), using the function occfilt_env from the {flexsdm} package (Velazco *et al.*, 2022). As the environmental filters are sensitive to the number of intervals (bins), filters with four, five, six, eight and ten bins were made, using the selected variables to build the distribution models. Spatial autocorrelation, based on the Moran Index, was calculated for each resulting dataset and the number of bins that showed the lowest autocorrelation (5 bins, with $I = 0.352$) was selected. After all these steps, a final set of 315 unique

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

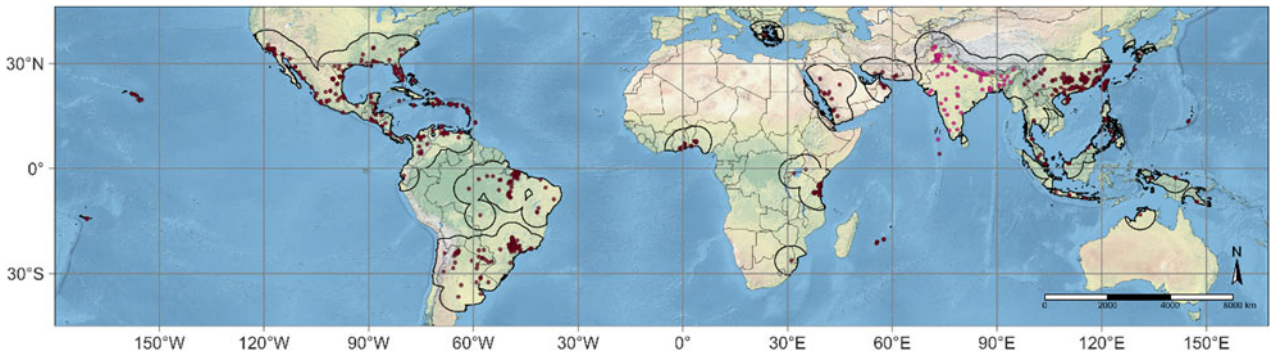


**Figure 1.** Distribution of *D. citri*: points of occurrence (red, native areas; magenta, invaded areas) and study area (areas with darker contour).

occurrence records was obtained, cleaned and filtered, for the modelling process (fig. 1).

As our intention was to estimate the areas susceptible to invasion by *D. citri* in Ghana, its probability of establishment was evaluated, and the results were interpreted without considering limitations to the future dispersal of the species, using native and invaded regions since this is the best option, especially if considering the modelling goal for invasive species (Beaumont *et al.*, 2009; Broennimann and Guisan, 2008; Zhang *et al.*, 2020).

### Environmental data

The bioclimatic variables (table 1) utilised in this study were obtained from the Worldclim database version 2.1 (Fick and Hijmans, 2017). These variables have an average spatial resolution of 2.5 arc-min, which corresponds to about 4 km at the equator. The data cover the period from 1970 to 2000. The {geodata} package version 0.5-8 (Hijmans *et al.*, 2023) was used to access and process the data. These variables were chosen as predictor variables due to their ability to capture annual climate variations and limiting factors known to affect the geographic distribution of species (O'Donnell and Ignizio, 2012). In addition, Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover was sourced from Friedl and Sulla-Menashe (2022).

Variables that exhibited a Pearson's correlation coefficient ($r$) greater than the absolute value of 0.70, which is considered significant at an $\alpha$ level of 0.05, were categorised based on the results of a hierarchical cluster analysis (fig. 2). The cluster analysis was conducted using the {corrplot} package version 0.92 (Wei and Simko, 2021). The initial set of all environmental variables was reduced using an algorithm that: (1) ranks variables based on permutation importance, (2) checks if the top-ranked variable is correlated with others (Pearson, $r > |0.7|$), and (3) if correlated, performs a Jackknife test between correlated variables, removing the variable that minimally impacts model performance when removed, based on the true skill statistics (TSS) metric (Allouche *et al.*, 2006; Lawson *et al.*, 2014; Shabani *et al.*, 2018). This process was repeated until all variables were tested, and all correlated variables were removed. Variables with low permutation importance (<3%) were then removed using Jackknife, ensuring that removal did not decrease model quality (based on estimated TSS value) using independent test data in ten permutations. We presented the contribution of the environmental variables to the base model in table S2. The resulting variables, and their descriptive statistics considering the points of presence, are presented in table 2. These variables were retained because findings from earlier studies indicate that, when developing models under various conditions, it is more useful to focus on a few variables that align with well-defined biological expectations rather than incorporating many variables with uncertain impacts on species distribution (Araújo and Guisan, 2006; Austin and Van Niel, 2011; Santini *et al.*, 2021).

### Model development

All procedures related to data processing, development of models and maps were performed with the R environment, version 4.3.0

**Table 1.** Environmental variables used for the initial model

| Code | Climatic variable | Unit |
|---|---|---|
| **Bio01** | Annual mean temperature | °C |
| **Bio02** | Mean diurnal range (mean of monthly (max temp–min temp)) | °C |
| **Bio03** | Isothermality (BIO2/BIO7) (× 100) | - |
| **Bio04** | Temperature seasonality (standard deviation × 100) | - |
| **Bio05** | Maximum temperature of warmest month | °C |
| Bio06 | Minimum temperature of coldest month | °C |
| Bio07 | Temperature annual range (BIO5-BIO6) | °C |
| Bio08 | Mean temperature of wettest quarter | °C |
| Bio09 | Mean temperature of driest quarter | °C |
| Bio10 | Mean temperature of warmest quarter | °C |
| Bio11 | Mean temperature of coldest quarter | °C |
| Bio12 | Annual precipitation | mm |
| Bio13 | Precipitation of wettest month | mm |
| Bio14 | Precipitation of driest month | mm |
| Bio15 | Precipitation seasonality (coefficient of variation) | mm |
| Bio16 | Precipitation of wettest quarter | mm |
| Bio17 | Precipitation of driest quarter | mm |
| Bio18 | Precipitation of warmest quarter | mm |
| Bio19 | Precipitation of coldest quarter | mm |
| MODISLC | Moderate resolution imaging spectroradiometer land cover product | - |

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

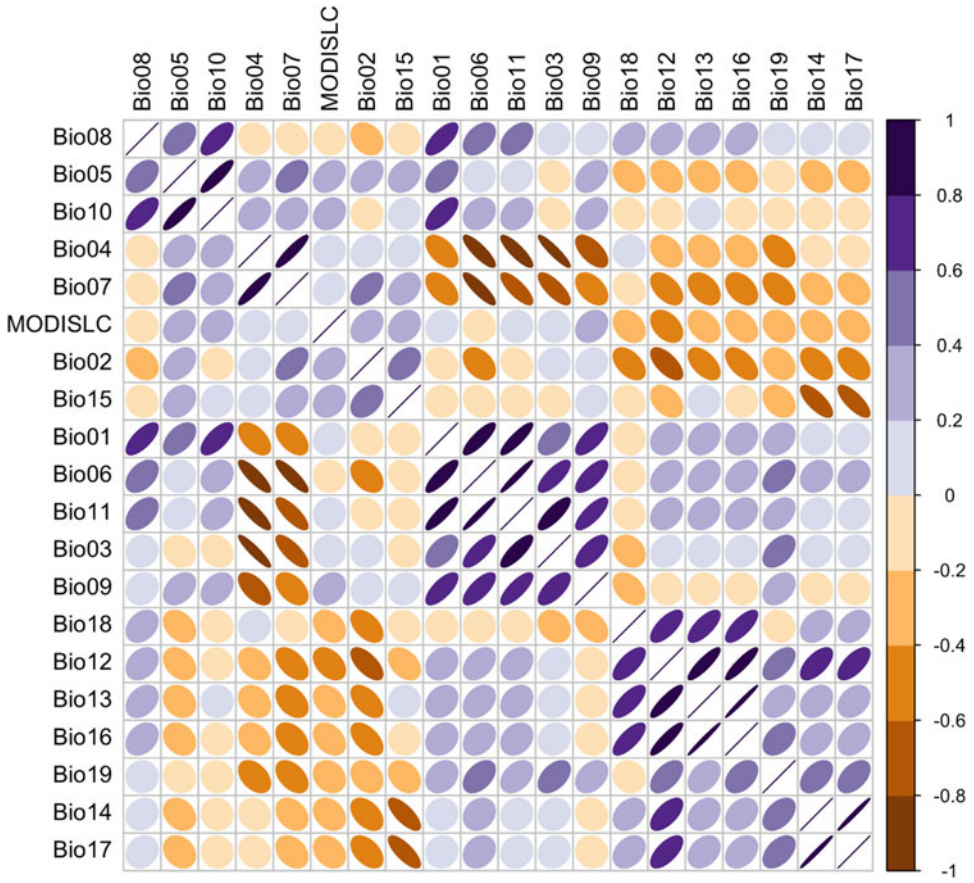


**Figure 2.** Correlation between bioclimatic variables. Blue colour sloping to the right indicates a positive correlation, while red sloping to the left indicates a negative correlation. The intensity of Pearson's correlation coefficient ($r$) increases from the circle ($r = 0$) to the ellipse ($r$ = intermediate) to the line ($r = 1$). Correlated variables, $r > 0.70$, were grouped by centroid with hierarchical cluster analysis.

'Already Tomorrow' (R Core Team, 2023) with the following packages: {rnaturalearth} version 1.0.1 (Massicotte and South, 2023), to obtain spatial data from Ghana; {terra} version 1.7-29 (Hijmans, 2023) and {sf} version 1.0-13 (Pebesma, 2018), for analysis and transformation of spatial data; {flexsdm} version 1.3.3 (Velazco *et al.*, 2022), for all species distribution modelling procedures, with resources from {maxnet} version 0.1.4 (Phillips and Phillips, 2021); {pROC} version 1.18.2 (Robin *et al.*, 2011), for graphics and estimates of the receiver operating characteristics (ROC) curve; {ade4} version 1.7-22 (Dray and Dufour, 2007) and {ecospat} version 3.5.1 (Broennimann *et al.*, 2022), for ecological analyses; {factoextra} version 1.0.7 (Kassambara and Mundt, 2022), to visualise the results of multivariate analyses; {tmap} version 3.3-3 (Tennekes, 2018), to plot all resulting maps.

The maximum entropy model (Maxent) was used because of its widespread application in species distribution modelling and its demonstrated effectiveness relative to alternative methods (Elith *et al.*, 2006, 2011; Heikkinen *et al.*, 2012; Hijmans, 2012; Venette, 2017; Helmstetter *et al.*, 2021; Valavi *et al.*, 2022). Maxent has been employed to identify suitable regions for the establishment of *D. citri* on a global scale (Aidoo *et al.*, 2022). It has been recognised as a reliable approach for species distribution modelling (Valavi *et al.*, 2022). Yet, Maxent is susceptible to sample bias and can easily result in overfitting (Zhu *et al.*, 2014).

**Table 2.** Descriptive statistics of the environmental variables considering their values at the 315 occurrence points (before the environmental filter) used to develop the models

| Variable | Minimum | Maximum | Median | Mean | SD |
|---|---|---|---|---|---|
| Bio02 | 4.93 | 19.77 | 10.62 | 10.54 | 2.69 |
| Bio04 | 28.30 | 875.18 | 380.31 | 378.16 | 246.45 |
| Bio10 | 12.69 | 34.81 | 26.97 | 26.62 | 3.12 |
| Bio15 | 10.20 | 161.74 | 67.18 | 70.04 | 26.29 |
| Bio16 | 26.00 | 3457.00 | 600.00 | 628.54 | 340.59 |
| Bio17 | 0.00 | 614.00 | 89.00 | 111.58 | 102.05 |
| Bio19 | 1.00 | 1725.00 | 159.00 | 239.95 | 283.40 |

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

To mitigate the overfitting and enhance the generalizability of a model, it is crucial to carefully select parameters. This includes determining suitable transformations for the explanatory variables and identifying the best value for the regularization multiplier. This step was carried out through constructing and validating different models, seeking the best combination of these parameters.

The process of variable transformation in Maxent involves expanding the explanatory variables to include a broader range of derived variables. The bioclimatic variables chosen for the model are represented as derived variables, which are functions of the original variables (Phillips *et al.*, 2006). This phenomenon might be interpreted as a change in the functional structure of the model specification, akin to including polynomial parts in a linear regression analysis. According to Phillips *et al.* (2017) and Phillips and Dudík (2008), the latest iterations of Maxent offer support for five types of transformations applicable to continuous independent variables. These transformations include linear, quadratic, threshold, forward hinge and reverse hinge.

A multi-step process (Perkins-Taylor and Frey, 2020; Warren *et al.*, 2020; Khan *et al.*, 2022) was applied to identify species-specific adjustments used for optimizing the model (Shcheglovitova and Anderson, 2013; Radosavljevic and Anderson, 2014). This was done to prevent excessive complexity and, consequently, low performance when projecting the model to different locations or under climate change scenarios (Elith *et al.*, 2011; Merow *et al.*, 2013; Low *et al.*, 2021). Several authors have employed a non-homogeneous Poisson process to estimate models (Renner and Warton, 2013; Renner *et al.*, 2015; Phillips *et al.*, 2017). These models incorporated linear (L), quadratic (Q) and hinge (H) features and combinations of these features such as 'LQ', 'QH' and 'LQH'. A base Maxent model was fitted using RM = 1 and FC = LQHP (default settings), employing the 'Maxnet' method and the training dataset (d), including all predictor variables. The goal was to evaluate and select the most important variables. A total of 119 models (FC = L, Q, H, LQ, QH, LQH, LQP, LQHP; RM = 1–5, with increments of 0.25) were fitted to identify the best combination of defined hyperparameters and to select the most suitable model (final model) using TSS with the selected variables (fig. S1). If results from different models are similar, the simplest one, ecologically easier to understand, was chosen. This is particularly important for the formulation of phytosanitary policies and decision-making processes.

Background points are based on occurrence distribution. In the first dataset we applied a geographic filter which is easy and faster, as we are working with all 19 variables. To run the final model, we filtered the occurrences with an environmental filter (just with the selected variables) and thus, had to create another background sample. For defining background points, the occurrence points were divided into four blocks using conventional k-fold cross-validation, allowing control over potential spatial autocorrelation between training and testing data and evaluating model transferability more appropriately (Roberts *et al.*, 2017; Santini *et al.*, 2021). Thirty grids with resolutions ranging from 0.5 to 8 degrees were generated, each with a minimum of ten occurrences (fig. S2 a, b). This method effectively addresses potential spatial autocorrelation between the training and test data, and provides a more suitable evaluation of the models' transferability compared to alternative partitioning methods (Roberts *et al.*, 2017; Santini *et al.*, 2021). In their study on effects of future climate, land use and protected areas ineffectiveness on a dark scenario for Cerrado plant species, Velazco *et al.* (2019) generated 20 grids, each with resolutions ranging from 0.5 to 5 degrees. The optimal grid size was selected considering (i) minimal spatial autocorrelation, (ii) maximum environmental similarity and (iii) minimal differences in records between training and testing data (Velazco *et al.*, 2022). In this study, we made sure each cell included at least five occurrence points. For the purpose of conducting an autocorrelation test between groups, 50% of the points of existence were used. After establishing the division of occurrence records, the allocation of 10,000 background points was carried out in a random way, making sure they were dispersed proportionally based on the number of occurrences in each partition.

Excluding product features is justified by the fact that the model is defined by the marginal response curves of each predictor variable (Phillips *et al.*, 2017). These curves are more straightforward to interpret compared to those that rely on the values of other variables. It is worth noting that some of the used variables already encompass the combination of others. Phillips *et al.* (2017) specified the use of the clog-log output format as the most appropriate method for representing the probability of occurrence or establishment of a species.

The maximum entropy approach is a statistical technique that seeks to estimate an unknown probability distribution by leveraging the observed frequencies and background information within a defined study area. The goal of this strategy is to identify the distribution that shows the highest degree of geographic uniformity, often known as maximum entropy. This is achieved by incorporating the restrictions resulting from the environmental conditions seen in the known occurrence locations. Maxent often uses randomly generated background points distributed throughout the spatial extent of the study area (Phillips *et al.*, 2009). The purpose of this is to determine the extent of environmental variability surrounding the best conditions for the survival of a species. The number of background points significantly influences the model's predictions, and it is recommended to incorporate a considerable quantity of data to effectively represent the environmental conditions in which the species is found (Barbet-Massin *et al.*, 2012; Northrup *et al.*, 2013). It is crucial to recognise that the results depict a model of habitat suitability relative in nature. Put otherwise, such findings indicate that one location is more suitable than another, without indicating the actual presence or inhabitation of the species.

The model development requires the spatial extent to encompass the accessible area for the species of interest during the relevant period (Araújo *et al.*, 2019). Additionally, the background data must be constrained to this same extent, and historical dispersion methods should be employed (Barve *et al.*, 2011; Merow *et al.*, 2013; Cooper and Soberón, 2018). The choice of a proper geographic area for sampling background points varies depending on the species and the goals of the study (Santini *et al.*, 2021). This is important for the development of niche models that rely on presence-only data, such as Maxent models (VanDerWal *et al.*, 2009; Anderson and Raza, 2010; Barbet-Massin *et al.*, 2012; Khosravi *et al.*, 2016; Cooper and Soberón, 2018; Machado-Stredel *et al.*, 2021; Amaro *et al.*, 2023). It is also crucial to make sure the sample size is large enough to adequately represent all environments (Renner *et al.*, 2015). In our study, the calibration area for the model was defined by considering the Köeppen–Geiger zones occupied by the species (Webber *et al.*, 2011; Hill *et al.*, 2017), as depicted in fig. 3. The resulting calibration area encompassed 68,026,420 $km^2$.

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

Applying this approach is pertinent to models intended for extrapolation to different geographic regions beyond the calibration area or for alternative temporal intervals (Velazco *et al.*, 2022). The best grid size was determined using the part_sblock function from the {flexsdm} package. This function systematically explores various block sizes and selects the most suitable size based on a multi-dimensional optimization procedure. The optimization process considers three dimensions: spatial autocorrelation (measured by Moran's *I*), environmental similarity (measured by Euclidean distance) and variation in data quantity among partition groups (measured by standard deviation – SD) (Velazco *et al.*, 2022). Velazco *et al.* (2022) established four partitions, each consisting of a test set of 40 cells and the resolving cells varying from 2 to 30 pixels.

In multiple cases of species distribution modelling, the chosen model predictions are transformed into binary maps that depict areas considered suitable or unsuitable for the species under investigation. While some scholars have advocated for the complete avoidance of this procedure unless there is a clear justification for applying the model (Guillera-Arroita *et al.*, 2015; Santini *et al.*, 2021), the information provided by a map is valuable when addressing invasive species. It aids in focusing on areas for implementing phytosanitary public policies. The determination of a threshold value that serves as a cut-off point for site classification is a critical part, since varying thresholds might provide significantly divergent estimates (Jarnevich *et al.*, 2015). In this study, the threshold selected was max_sens_spec, which has been earlier shown to provide consistent findings (Liu *et al.*, 2005, 2013, 2016; Allouche *et al.*, 2006). This threshold is equal to optimizing the vertical separation between a point on the ROC curve and the diagonal line, or maximizing the TSS. Still, it is crucial to acknowledge that when using thresholds to transform distribution models, the resulting interpretations should generally be regarded as hypothetical distributions (Liu *et al.*, 2013).

When constructing a model to estimate the possible distribution of a species, the model is adjusted to the observed data using a predetermined range of values for each environmental variable, known as the calibration range. When conducting model predictions for spatial or temporal projections, parts of the environment may show conditions that fall outside the calibration range. These conditions, under non-analogous circumstances, can lead to model extrapolation (Randin *et al.*, 2006; Williams *et al.*, 2007; Fitzpatrick and Hargrove, 2009; Owens *et al.*, 2013; Yates *et al.*, 2018). To assess which areas had the highest extrapolation, the Shape metric (pers. comm.) was used, which consists of a model-independent method, since its calculation is based exclusively on the environmental distance between the training and projection data, that is, extrapolation is independent of model parameters and predictions.

### *Model evaluation*

The scholarly literature has various perspectives on the best evaluation criteria for species distribution models. The ROC is a widely used metric for evaluating classification performance. One of its most commonly employed parts is the area under the curve (AUC) value. This metric is independent of threshold and scale (Merow *et al.*, 2013). Bradley (1997) has revealed many useful attributes linked to the use of AUC as a performance metric in classification tasks. Still, the AUC values can be affected by the size of the study region (Smith, 2013; Amaro *et al.*, 2023). Specifically, if the model encompasses a vast area and the species being studied has a limited distribution inside that area, the AUC values may be artificially inflated. Evaluating a classifier only based on total AUC may not be enough when assessing its performance specifically in areas characterised by high specificity or high sensitivity (Robin *et al.*, 2009). To address these scenarios, the idea of partial area under the curve (pAUC) was created as a

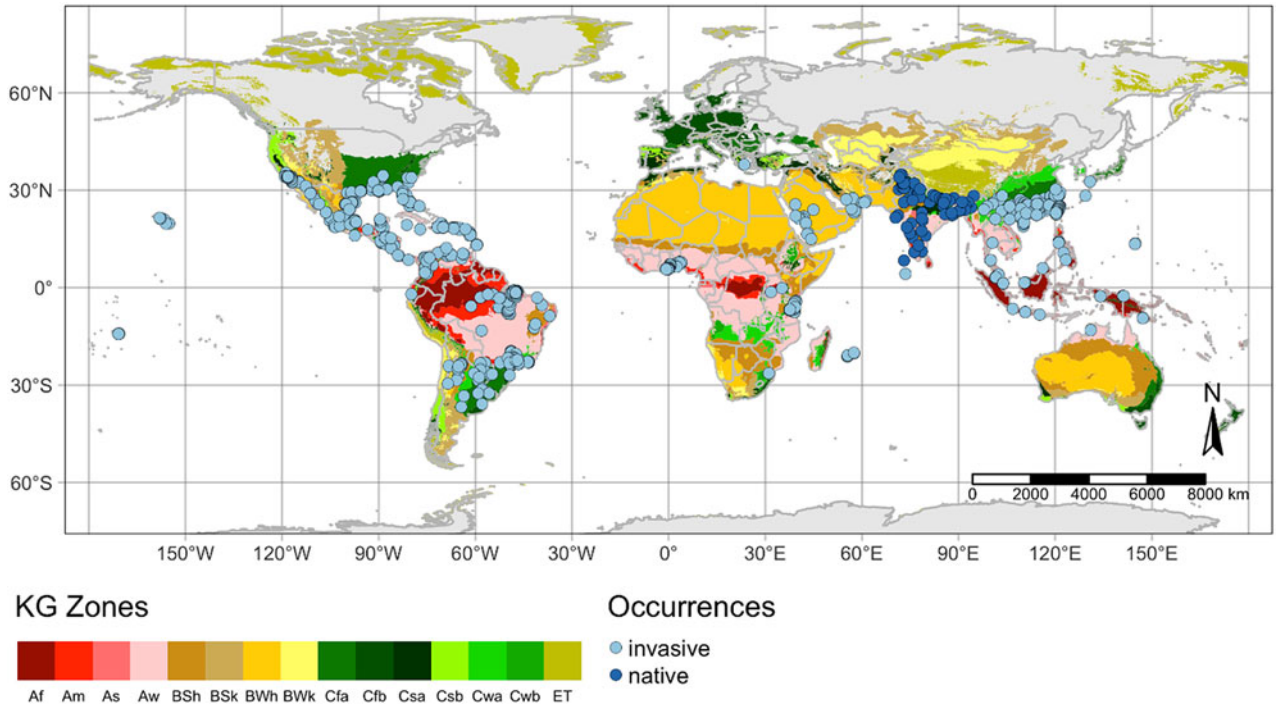


**Figure 3.** Map of the Köeppen–Geiger zones used to define the model calibration area, considering the current dispersion of *Diaphorina citri*.

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

localised comparative method that specifically examines a segment of the ROC curve (McClish, 1989; Jiang *et al.*, 1996; Streiner and Cairney, 2007).

Alternative metrics for evaluating the adequacy of a model show certain limitations (Allouche *et al.*, 2006). These limitations include sensitivity, which refers to the proportion of correctly predicted attendances or the true positive rate (TPR), specificity, which denotes the proportion of accurately predicted absences, and the TSS, a metric that remains unaffected by prevalence and is calculated using sensitivity and specificity (Allouche *et al.*, 2006). Additionally, κ, another metric unaffected by prevalence, can also be used for this purpose (Allouche *et al.*, 2006).

According to Hosmer *et al.* (2013), an ROC curve with an AUC value of 0.9 or higher indicates an excellent model fit. A value between 0.6 and 0.9 is considered good, while a value of 0.5 or lower suggests that the model performs no better than random chance. The TPR should be positioned near 1, hence indicating a heightened level of sensitivity. The TSS is a numerical measure that falls within the range of 0–1. A TSS value over 0.9 is regarded as ideal, while a value between 0.85 and 0.9 is considered exceptional. A TSS value ranging from 0.7 to 0.85 is classified as very good, while a value between 0.5 and 0.7 is considered good. Additionally, a TSS value ranging from 0.4 to 0.5 is considered decent, and any value ≤0.4 indicates a poor fit (Landis and Koch, 1977). According to Peterson (2006), to accurately represent low false-negative scores, omission values should be reduced and brought closer to zero.

### Niche change

The occurrence records of the species were divided into two sets of data: the native area and the area invaded by *D. citri*. At first, updated Köppen–Geiger climate classes (Rubel and Kottek, 2010) were extracted from a raster file obtained from the Climate Change and Infectious Diseases Group (https://koeppen-geiger.vu-wien.ac.at/present.htm) of the University of Vienna, Austria, to create histograms of the frequencies of occurrences and comparisons. To assess the ecological niches in the environment and compare the distinctions between the native and invaded regions, the COUE framework (Centroid shift, Overlap, Unfilling, and Expansion) proposed by Broennimann *et al.* (2012) was used for both datasets. The comparison of niches was conducted in a two-dimensional space using the Schoener Index (*D*), a metric that measures the extent of overlap and ranges from 0 (showing no overlap) to +1 (representing complete overlap) (Broennimann *et al.*, 2012). Subsequently, principal component analysis was used to analyse environmental data obtained from occurrence locations in both native and introduced areas to identify and assess the variability captured by the first two principal parts. The spatial extent of the environment was partitioned into a grid consisting of 100 cells, and the estimation of occurrence density within this spatial domain was conducted using a kernel density estimator, as described by Broennimann *et al.* (2012) and Parravicini *et al.* (2015).

The measured overlap (*D*) was assessed by using a niche equivalency test to determine whether it deviated significantly from a null distribution consisting of 1000 randomly generated *D*-metrics. This null distribution was created by randomly redistributing the occurrences of both niches among two datasets (Broennimann *et al.*, 2012). The null hypothesis, which posits that the niches are comparable, is considered invalid when the observed *D* value falls below the fifth percentile of the null distribution. This null distribution was generated by considering the geographic availability of environmental conditions. Specifically, one niche was randomly distributed in the background while the other remained unchanged.

In this study, we used the same metrics as Guisan *et al.* (2014) who examined various metrics related to niche dynamics in invasive species. These metrics include niche stability, which measures the proportion of environments within the introduced niche shared with the native niche. Another metric is niche expansion, which measures the proportion of environments in the introduced niche that do not overlap with the native niche. Last, the metric of unoccupied niche, also known as niche unfilling, assesses the proportion of environments that are not occupied by the invasive niche.

### Projections for future climate scenarios

For future projections under different climate change scenarios, three periods were selected (2021–2040, 2041–2060 and 2061–2080), SSPs 245 and 585, using three global climate models (GCMs) considered suitable for Ghana (Oduro *et al.*, 2021), BCC-CSM2-MR, INM-CM5-0, MRI-ESM2-0, as per the Coupled Model Intercomparison Project Phase 6 (CMIP6).

The SSPs provide different development paths, contemplating possible trends about radiative forcing (W $m^{-2}$). SSP 245 describes a society in which development follows a historical pattern without significant future deviations, with a radiative forcing in 2100 of 4.5 W $m^{-2}$, representing an increase in global temperature between 1.4 and 2.8 °C. SSP 585 assumes a society in which economy is based on fossil fuels and intensive energy use, with a projected radiative forcing of 8.5 W $m^{-2}$ in 2100 and a rise in global temperature between 3.5 and 5.5 °C.

The bioclimatic variables, SSPs and GCMs were obtained from the Worldclim. The mean model, for the periods and SSPs, was obtained by averaging the three GCMs used as a consensus model.

### Contribution of environmental variables

Maxent's first formulation is equal to maximizing the probability of a parametric exponential distribution (Phillips *et al.*, 2004), but more recently the identical maximum likelihood exponential model may be obtained through an Inhomogeneous Poisson Process (IPP) (Aarts *et al.*, 2012; Fithian and Hastie, 2013; Renner and Warton, 2013). One often used technique is the approach used by the 'varImportance' function within {fitMaxnet} package. The method used for calculating variable importance aligns with the method used in the R packages biomod2 and ecospat. The model generates predictions for every row in the combined environmental data table, created by stacking the occSWD and bkgSWD datasets. In the maxnet model object, the values of each variable undergo permutation across rows, resulting in a model prediction for each row using the permuted or shuffled data table. The permutation procedure is repeated numReplicates times for every variable. For every permutation, a Pearson correlation coefficient is calculated between the reference predictions and the predicted values obtained from the shuffled table. The significance value is calculated as the difference between 1-correlation coefficient. The outcome yields a vector of average scores for each variable, represented as a percentage relative to the sum of all average scores.

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

**Table 3.** Threshold-dependent and independent metrics (mean values and standard deviations) used to evaluate the developed model

| Metric names | Values |
|---|---|
| True-positive rate or sensitivity (TPR) | 0.63935 |
| True-negative rate or specificity (TNR) | 0.78289 |
| True skill statistic (TSS) | 0.42224 |
| Sorensen index | 0.17578 |
| Jaccard index | 0.09642 |
| F-measure on presence-background (FPB) | 0.19285 |
| Omission or false-negative rate (OR) | 0.36065 |
| Boyce index | 0.90534 |
| Area under ROC curve (AUC) | 0.75868 |
| Inverse mean absolute error (IMAE) | 0.69743 |
| Maximum sensitivity plus specificity (maxSSS) | 0.47555 |
| False-positive rate (FPR) | 0.21711 |
| Positive predictive value or precision (PPV) | 0.74650 |
| Accuracy | 0.71112 |
| F1 score | 0.68878 |
| Balanced accuracy | 0.71112 |
| Matthews correlation coefficient (MCC) | 0.42666 |
| Minimum training presence (MTP) | 0.03987 |
| 10th Percentile training presence (10TP) | 0.11588 |
| Symmetric extremal dependence index (SEDI) | 0.49300 |

## Results

Based on the TSS metric ( = 0.42224), the best combination was obtained using the linear, quadratic and hinge (LQH) classes simultaneously and regularised multiplication 0.5. The evaluation metrics of the selected model among the 50 tested models, provided by the {flexsdm} package and calculated from these, are presented in table 3, although all models have presented good performance to the random one.

The ROC curve of the final model (fig. 4a), resulting from evaluating true-positive predictions (sensitivity) and false-positive predictions (1 – specificity), showed a useful predictive capacity with an AUC value between 0.7 and 0.9. The ROC curve and partial AUC information when constraining the false- and true-positive rate ($x$ = FPR = specificity; $y$ = TPR = sensitivity) in the 90–100% range are illustrated in fig. 4b. The partial area (pAUC) can be interpreted as the mean sensitivity over the specified specificity range and the mean specificity over the specified sensitivity range. In addition, the suitability of habitats against the projection and training data are illustrated in fig. S3.

The response curves of the model are illustrated in fig. 5, which allow exploring the average marginal effect of environmental variables on the suitability of the environment for *D. citri*. The graphs show how the model response is individually influenced by each predictor variable, keeping the effects of the other variables fixed. The most suitable habitats (ideal values) for *D. citri*, as predicted by our model, are presented in supplemental information table S3. The histogram depicting the occurrence of *D. citri* against environmental variables is illustrated in fig. S4.

The analysis showed that the order of importance of the eight environmental variables is temperature seasonality (Bio04) > mean temperature of warmest quarter (Bio10) > precipitation of driest quarter (Bio17) > moderate resolution imaging spectroradiometer land cover (MODISLC) > precipitation seasonality (Bio15) > precipitation of coldest quarter (Bio19) > precipitation of wettest quarter (Bio16) > mean diurnal range (Bio02), were the most important drivers of *D. citri* distribution. Among these environmental variables, Bio04, Bio10, Bio17, MODISLC and Bio15 contributed to about 85% of the model (fig. 6). Further, the characteristics of the mean of the analysis on the *D. citri* are presented in table S4.

The potential geographic distribution of *D. citri* for Ghana, resulting from our model, is shown in fig. 7a, segmenting the probability of establishment into seven classes to help with visualization and comparison between locations. The high, optimal, moderate, marginal and unsuitable probability calculated at the occurrence points was 100% (table 4, fig. 7b). Applying a threshold that maximises the sum of sensitivity and specificity (max_sens_spec = 0.3656577) resulted in the map in fig. 7b which

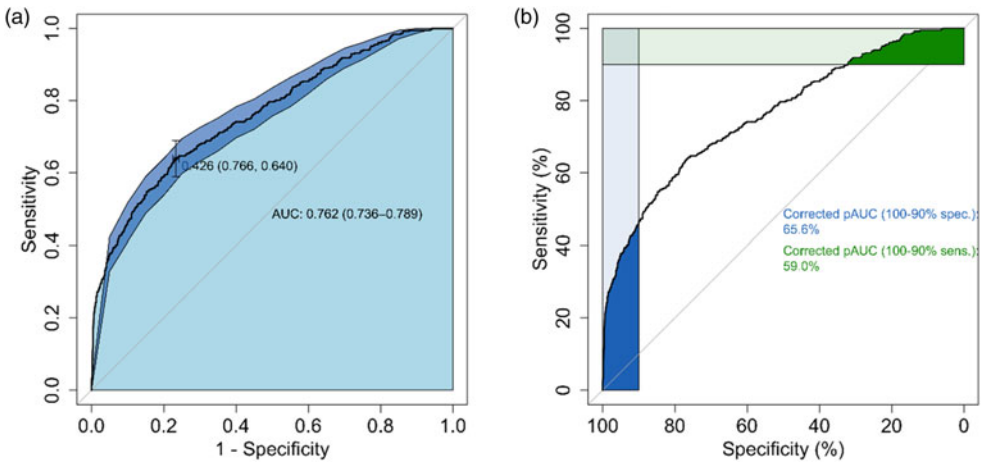


**Figure 4.** Graph of the receiver operating characteristic (ROC) curve for the potential geographic distribution model of *Diaphorina citri* in Ghana, showing the total area (AUC) under the ROC curve (a) and partial areas (pAUC) (b).

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

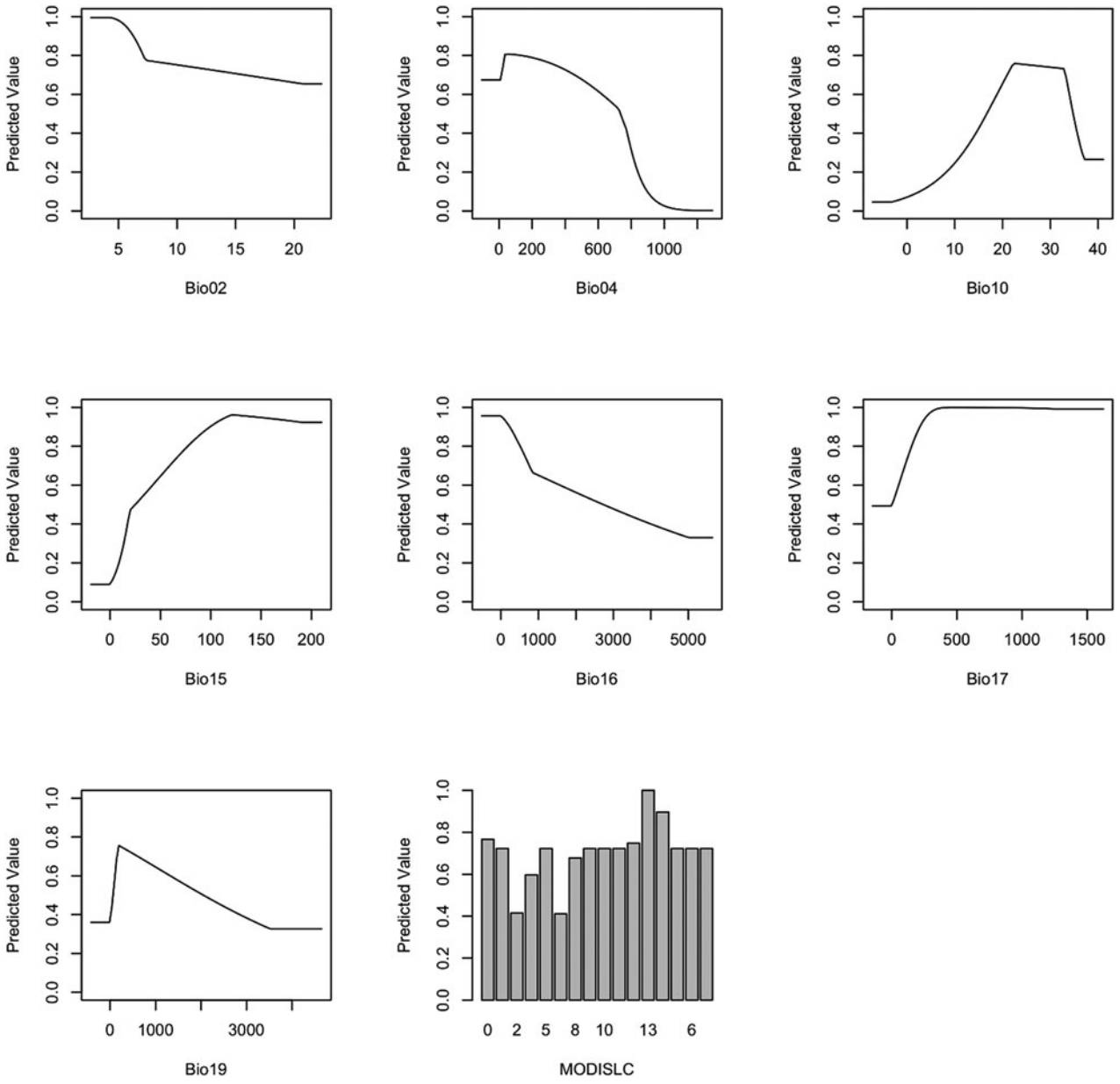


**Figure 5.** Graphs of the response curves of the variables used in the model. Mean diurnal range (Bio02), temperature seasonality (Bio04), mean temperature of warmest quarter (Bio10), precipitation seasonality (Bio15), precipitation of wettest quarter (Bio16), precipitation of driest quarter (Bio17), precipitation of coldest quarter (Bio19), Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover.

represents an area of 244,129 km$^2$ (table 4). Extrapolation of models utilizing the shape metric is contingent upon its value, whereby a larger value (shown by shades of dark blue) indicates a greater disparity in the environment. Consequently, this discrepancy leads to a diminished level of reliability in the model's predictions for the corresponding area (fig. 7c).

In its native distribution areas, *D. citri* predominantly occupies regions with hot semi-arid (steppe) climate (BSh), monsoon-influenced humid subtropical climate (Cwa) and tropical savanna, wet (Aw) climate classes according to the updated Köppen–Geiger climate classification (fig. 8a) with about 80% of the points belonging to these classes. Yet, in invaded areas, occurrences are more concentrated (about 58%) in the climate classes tropical savanna, wet (Aw), humid subtropical climate (Cfa) and tropical monsoon climate (Am) (fig. 8b). In general, we identified a greater number of climate classes occupied by *D. citri* in newly invaded compared to native areas, suggesting a change in niche.

The predicted distribution areas for *D. citri* in Ghana under the SSPs 245 and 585 and for the three periods are illustrated in fig. 9 and presented in table 4. The suitability classes for the country as predicted by the model are illustrated in fig. 10, with habitats ranging from marginal to high suitability. Yet, areas with high suitability are primarily located in the southern parts of the country. In citrus-producing regions, such as Ashanti, Central, Western, Eastern and Volta, a change in habitat suitability from the current time into the future is forecasted, with most parts showing high suitability for *D. citri* (fig. 10). The computed areas of habitat suitability for *D. citri* under the different climates are presented in table 4. We showed the optimal areas and the

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

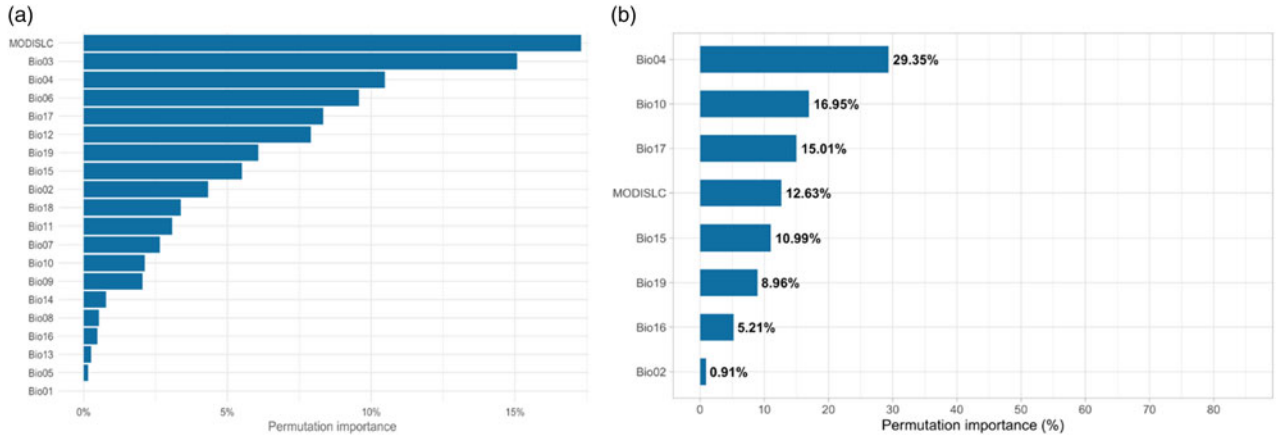


**Figure 6.** Percentage contribution of environmental variables to the base (a) and final (b) models. Annual mean temperature (Bio01), mean diurnal range (Bio02), isothermality (Bio03), temperature seasonality (Bio04), max temperature of warmest month (Bio05), min temperature of coldest month (Bio06), temperature annual range (Bio07), mean temperature of wettest quarter (Bio08), mean temperature of driest quarter (Bio09), mean temperature of warmest quarter (Bio10), mean temperature of coldest quarter (Bio11), annual precipitation (Bio12), precipitation of wettest month (Bio13), precipitation of driest month (Bio14), precipitation seasonality (Bio15), precipitation of wettest quarter (Bio16), precipitation of driest quarter (Bio17), precipitation of warmest quarter (Bio18), precipitation of coldest quarter (Bio19), Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover.

probability of establishment of the pest in the world in figs S5 and S6, respectively. Further, we illustrate Ghana's Maxent probability and classes maps in fig. S7. The probability of occurrence of *D. citri* across the time periods under climate change scenarios using a box plot are presented in fig. 11.

## Discussion

Linking climatic conditions to occurrence data is a commonly employed biogeographic strategy for characterizing species distributions and forecasting the potential effects of climate change (Guisan *et al.*, 2014; Finch *et al.*, 2021). Applying climate modelling in assessing habitat suitability has clearly demonstrated that climate change will substantially affect the distribution patterns of invasive species (Aidoo *et al.*, 2022, 2023a). Notwithstanding the limitations and uncertainties in the outcomes of species distribution models, using these models is a valuable technique for predicting potential distribution of species (Woodman *et al.*, 2019). Previous studies have established a correlation between temperature and precipitation patterns and their impacts on the spread, survival and development of *D. citri* and other invasive species (e.g. Hall *et al.*, 2012; Devi and Sharma, 2014). The present work used the CMIP6 data to predict the potential geographic distribution of *D. citri* in Ghana. Currently, the pest has a broad geographical distribution within the country, encompassing the major citrus-growing regions including, Western, Eastern, Central, Volta, Northern and the Ashanti Region (Brentu *et al.*, 2012; Asare-Bediako *et al.*, 2013). Our model now predicts a further expansion, with new areas of habitat suitability covering parts of Ahafo, Bono, Bono East, Northern, Savannah, North East, Upper East, Oti and Western North Regions. However, under the current climate areas with high and optimal suitability are restricted to southern parts of the country, including parts of Volta, Ashanti, Eastern, Western, Western North, Greater Accra and Central Regions, with moderate to marginal suitable climates in parts of the more arid Northern, North East, Upper West and Upper East Regions of Ghana. Our modelling results suggest that about 244,129 $km^2$ of Ghana are climatically suitable for *D. citri*, which corresponds well with the available occurrence data (Aidoo *et al.*, 2023b), indicating that the maxnet package can effectively predict the pest's habitats. The original formulation of Maxent equals maximizing the likelihood of a parametric exponential distribution (Phillips *et al.*, 2004). Like the Maximum Entropy java software, the maxnet package in R can generate forecasts and decrease commission mistakes when only a few occurrence records for a species are accessible.

In this study, temperature seasonality, mean temperature of warmest quarter, precipitation of driest quarter, moderate resolution imaging spectroradiometer land cover, precipitation seasonality, precipitation of coldest quarter, precipitation of wettest quarter and mean diurnal range were identified as significant environmental factors, collectively accounting for more than two-thirds of the observed geographical distribution of the pest. Earlier research on the growth and spread of *D. citri* has shown temperature and precipitation are crucial in influencing its reproductive, developmental, migratory, morphology and dispersal patterns (Antolínez *et al.*, 2022; Paris *et al.*, 2017). Further, temperature alters the flight capacity of *D. citri* (Antolínez *et al.*, 2022). New citrus flush production as well as maximum temperature, daily lowest temperature and rainfall have been positively correlated with *D. citri* infestations (Zorzenon *et al.*, 2021). The same authors found that the migration patterns of *D. citri* aligned with seasonal variations in specific climatic factors. Higher levels of humidity and daily maximum temperatures were associated with adverse effects, while increased rainfall in the preceding weeks had a favourable impact. Specifically, ideal conditions for the psyllid's spread include a mean diurnal range ranging from 2.7 to 4.2 °C, temperature seasonality of 56.5 °C and a mean temperature of warmest quarter of 22.2 °C (table S3). The changes in the climatic condition pattern resulting from climate change have the potential to exert both direct and indirect effects on the survival of invasive species. Nevertheless, it should be noted that in Brazil, the optimal environmental parameters for *D. citri* are typically observed within a temperature range of 22–28 °C (Zavala-Zapata *et al.*, 2022). However, laboratory studies predictions showed that temperatures ranging from 25 to 28 °C were optimal for the proliferation of *D. citri* and the subsequent expansion of its population (Tsai and Liu,

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

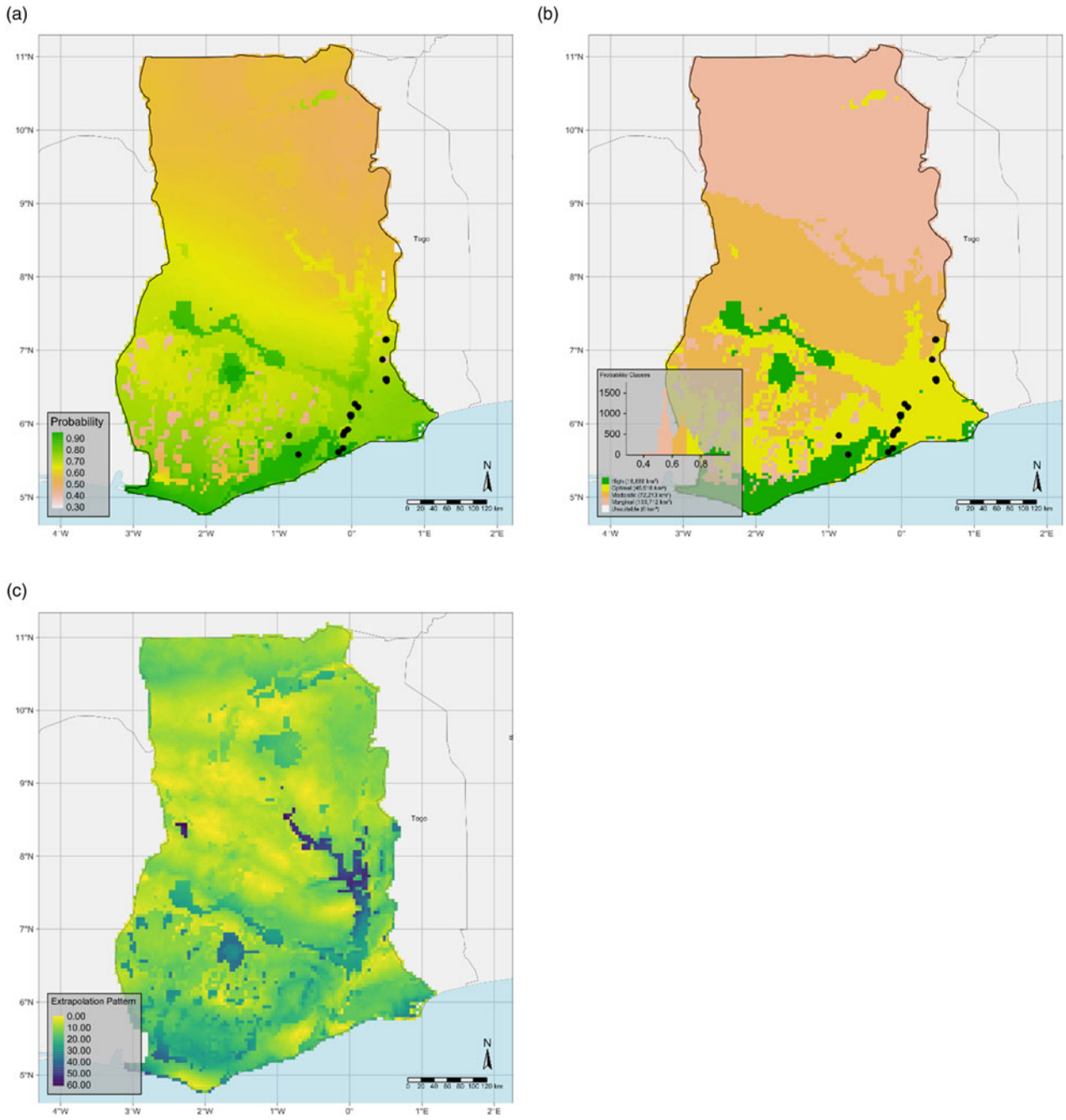


**Figure 7.** The extrapolation map of the model, considering the presence and background points, elaborated from the extra_eval function of the {flexsdm} package, using the Shape metric. (A) Predicted probability of establishment of *Diaphorina citri* in Ghana. Warmer colours (closer to red) indicate environments more suitable for establishing the species. Small black dots refer to identified occurrences. (B) Regions considered optimal for the occurrence of *D. citri* in Ghana (area in red), considering max_sens_spec = 0.3656577, and (C) model extrapolation based on the Shape metric: the higher the value (shades of dark blue), the greater the environmental difference and, consequently, the lower the degree of reliability of a model prediction for the area.

2000). Previous research conducted under controlled and standardised thermal conditions demonstrated that *D. citri* is unable to complete its life cycle above a temperature barrier of 33 °C (Tsai and Liu, 2000). Yet, a more recent investigation by Milosavljević *et al.* (2020) has provided empirical evidence that *D. ciri* has the ability to undergo developmental processes at temperatures exceeding 35 °C, under both constant and fluctuating temperature regimes. Furthermore, Antolínez *et al.* (2022) demonstrated an effective completion of the psyllid's life cycle throughout a range of temperature treatments spanning from 28 to 43 °C. However, it is important to note that the daily cycles with a temperature of 43 °C sustained for 6 h posed an exception to this pattern. Moreover, these authors observed a significant decline in the rate of adult emergence as the temperature surpassed 40 °C. In their study, Hall *et al.* (2011) conducted an estimation of temperature thresholds pertaining to the oviposition behaviour of *D. citri*. The results indicated that the lower and higher thresholds for oviposition were 16 and 41.6 °C, respectively. Comparable ranges for

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

**Table 4.** Changes in habitat suitability from the current time to future with unsuitability, marginal, moderate, optimal and high percentages.

| Climate | Unsuitable ($km^2$) | Marginal ($km^2$) | Moderate ($km^2$) | Optimal ($km^2$) | High ($km^2$) | % Unsuitable | % Marginal | % Moderate | % Optimal | % High |
|---|---|---|---|---|---|---|---|---|---|---|
| Currently (historic 1970-2000) | 0 | 103,712 | 72,213 | 49,516 | 18,688 | - | 100.00 | 100.00 | 100.00 | 100.00 |
| CMIP6 BCC-CSM2-MR SSP245 (2021-2040) | 0 | 54,361 | 46,112 | 55,584 | 88,072 | - | 52.42 | 63.86 | 112.25 | 471.28 |
| CMIP6 INM-CM5-0 SSP245 (2021-2040) | 0 | 11,943 | 67,572 | 50,932 | 113,683 | - | 11.52 | 93.57 | 102.86 | 608.32 |
| CMIP6 MRI-ESM2-0 SSP245 (2021-2040) | 0 | 126 | 73,426 | 52,109 | 118,468 | - | 0.12 | 101.68 | 105.24 | 633.93 |
| CMIP6 Mean SSP245 (2021-2040) | 0 | 15,039 | 70,737 | 52,216 | 106,137 | - | 14.50 | 97.96 | 105.45 | 567.94 |
| CMIP6 BCC-CSM2-MR SSP245 (2041-2060) | 0 | 30,783 | 46,084 | 60,045 | 107,218 | - | 29.68 | 63.82 | 121.26 | 573.73 |
| CMIP6 INM-CM5-0 SSP245 (2041-2060) | 0 | 3411 | 53,704 | 64,967 | 122,048 | - | 3.29 | 74.37 | 131.20 | 653.08 |
| CMIP6 MRI-ESM2-0 SSP245 (2041-2060) | 0 | 1932 | 66,883 | 61,571 | 113,743 | - | 1.86 | 92.62 | 124.35 | 608.64 |
| CMIP6 Mean SSP245 (2041-2060) | 0 | 8707 | 60,606 | 59,315 | 115,502 | - | 8.40 | 83.93 | 119.79 | 618.05 |
| CMIP6 BCC-CSM2-MR SSP245 (2061-2080) | 0 | 40,351 | 38,326 | 63,832 | 101,621 | - | 38.91 | 53.07 | 128.91 | 543.78 |
| CMIP6 INM-CM5-0 SSP245 (2061-2080) | 0 | 30,593 | 57,819 | 45,847 | 109,870 | - | 29.50 | 80.07 | 92.59 | 587.92 |
| CMIP6 MRI-ESM2-0 SSP245 (2061-2080) | 0 | 17,436 | 39,404 | 67,256 | 120,034 | - | 16.81 | 54.57 | 135.83 | 642.31 |
| CMIP6 Mean SSP245 (2061-2080) | 0 | 27,579 | 51,211 | 54,538 | 110,801 | - | 26.59 | 70.92 | 110.14 | 592.90 |
| CMIP6 BCC-CSM2-MR SSP585 (2021-2040) | 0 | 24,850 | 49,702 | 78,578 | 90,999 | - | 23.96 | 68.83 | 158.69 | 486.94 |
| CMIP6 INM-CM5-0 SSP585 (2021-2040) | 0 | 57,515 | 43,959 | 39,990 | 102,666 | - | 55.46 | 60.87 | 80.76 | 549.37 |
| CMIP6 MRI-ESM2-0 SSP585 (2021-2040) | 0 | 11,437 | 73,112 | 45,157 | 114,423 | - | 11.03 | 101.24 | 91.20 | 612.28 |
| CMIP6 Mean SSP585 (2021-2040) | 0 | 24,418 | 63,365 | 53,152 | 103,194 | - | 23.54 | 87.75 | 107.34 | 552.19 |
| CMIP6 BCC-CSM2-MR SSP585 (2041-2060) | 0 | 34,729 | 35,867 | 73,678 | 99,856 | - | 33.49 | 49.67 | 148.80 | 534.33 |
| CMIP6 INM-CM5-0 SSP585 (2041-2060) | 0 | 61,704 | 37,910 | 40,876 | 103,640 | - | 59.50 | 52.50 | 82.55 | 554.58 |
| CMIP6 MRI-ESM2-0 SSP585 (2041-2060) | 0 | 17,727 | 35,284 | 62,742 | 128,378 | - | 17.09 | 48.86 | 126.71 | 686.95 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| CMIP6 Mean SSP585 (2041-2060) | 0 | 30,168 | 50,750 | 54,298 | 108,914 | - | 29.09 | 70.28 | 109.66 | 582.80 |
| CMIP6 BCC-CSM2-MR SSP585 (2061-2080) | 0 | 50,728 | 22,790 | 40,498 | 130,114 | - | 48.91 | 31.56 | 81.79 | 696.24 |
| CMIP6 INM-CM5-0 SSP585 (2061-2080) | 2058 | 77,203 | 25,363 | 38,259 | 101,246 | - | 74.44 | 35.12 | 77.27 | 541.77 |
| CMIP6 MRI-ESM2-0 SSP585 (2061-2080) | 2037 | 74,077 | 21,273 | 19,862 | 126,881 | - | 71.43 | 29.46 | 40.11 | 678.94 |
| CMIP6 Mean SSP585 (2061-2080) | 105 | 66,449 | 26,077 | 33,514 | 117,985 | - | 64.07 | 36.11 | 67.68 | 631.34 |

*D. citri* response curves were found for the warmest quarter's mean temperature, warmest month's maximum temperature and temperature seasonality: 24.46–34.27, 28.6–40.91 and 56.83–818.03 °C, respectively (Wang *et al.*, 2019). According to Skendžić *et al.* (2021), global climate warming has the potential to affect insects, including the expansion of their geographic ranges, enhanced survival during winter months, an increase in the number of generations, heightened susceptibility to invasive insect species and insect-borne plant diseases and alterations in the dynamics of their interactions with host plants and natural predators.

Our predictions showed that precipitation seasonality of 121.7 mm, precipitation of wettest quarter of 25.9–317.1 mm, precipitation of driest quarter of 459.5 mm and precipitation of coldest quarter of 183.8 mm were ideal for the pest's proliferation in Ghana. Smaller insects such as *D. citri* are very vulnerable to severe precipitation which can dislodge or wash them off from their hosts. For example, Beattie (2020) showed that heavy rains washed off eggs and nymphs, thereby considerably affecting *D. citri* populations. The forecast conducted in China indicated that the optimal range of precipitation during the wettest quarter for *D. citri* was found to be between 562.89 and 1189.75 mm. Similarly, the range of precipitation during the warmest quarter was seen to vary from 503.73 to 1533.58 mm (Wang *et al.*, 2019). Several studies have shown that the primary constraint on the population size, geographical range and possible spread of *D. citri* is the cold temperatures experienced during winter months (e.g. Hall *et al.*, 2011; López-Collado *et al.*, 2013). This suggests that the survival and distribution of *D. citri* are primarily influenced by climatic factors such as temperature and precipitation.

Results from the maxnet package indicated high-suitability habitats for *D. citri*, which overlap with the major citrus-producing regions in Ghana. Hence, the model showed a high level of reliability in forecasting the distribution of *D. citri* confirming previous findings (Aidoo *et al.*, 2023b). Based on the model outputs for the projected future climatic conditions of the SSPs245 and 585 from the 2040s to the 2080s, Ghana's regions including Volta, Greater Accra, Central, Western and Eastern constitute areas that harbour highly suitable climatic conditions for *D. citri*. In contrast, some of Ghana's northern regions, including Savannah, Northern, North East, Upper East and Upper West, are projected to be less suitable for the psyllid. The climate suitable areas will increase from the current time until the 2080s for both climate change scenarios, with a few areas in the future showing unsuitability for the pest.

A study by Aidoo *et al.* (2023b) reported for the first time the presence of *D. citri* in the Volta Region of Ghana and suggested regular tracking and surveillance of the pest. Our risk maps can serve as a guide for the future development of control measures to successfully manage the pest under current and future climate change conditions. Such pest management measures are urgently needed to prevent the further spread of *D. citri* in Ghana and beyond because of the expected rising temperatures expanding suitability for the pest in most parts of the country. However, even a reduced distribution of *D. citri* could still threaten citrus production in Ghana. Moreover, invasive species can often adapt to changing climates (Barrett, 2000). Hence, the national phytosanitary authorities in Ghana should continue their efforts in preventing the spread of *D. citri* in the country as well as potential spillovers into the region.

The response curves from our model illustrate the manner in which the predicted probability of the species' habitat alters in

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

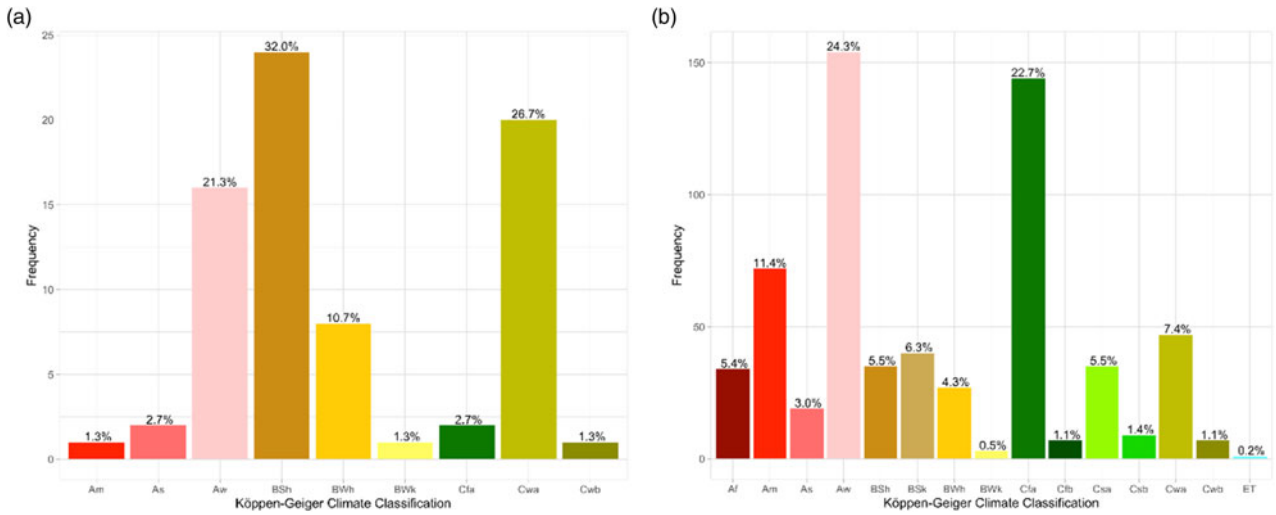


**Figure 8.** Frequency histograms of the climatic classes occupied by the recorded occurrences of *D. citri* (native (A) and none-native (B) ranges, tropical monsoon climate (Am), tropical dry savanna climate (As), tropical savanna, wet (Aw), hot semi-arid (steppe) climate (Bsh), hot deserts climate (BWh), cold desert climate (BWk), humid subtropical climate (Cfa), monsoon-influenced humid subtropical climate (Cwa), subtropical highland climate or temperate oceanic climate with dry winters (Cwb), tropical rainforest climate (Af), cold semi-arid (steppe) climate (BSk), temperate oceanic climate (Cfb), hot-summer Mediterranean climate (Csa) and warm-summer Mediterranean climate (Csb).

response to fluctuations in each predictor, while holding all other variables at their respective average sample values. They offer insights into the environmental factors correlated with varying species occurrence probabilities (Merow *et al.*, 2013; Tesfamariam *et al.*, 2022). These parameters may correspond to the species' established ecological preferences or tolerances, explaining certain features of its biology. However, species' biology is frequently shaped by many interconnected elements,

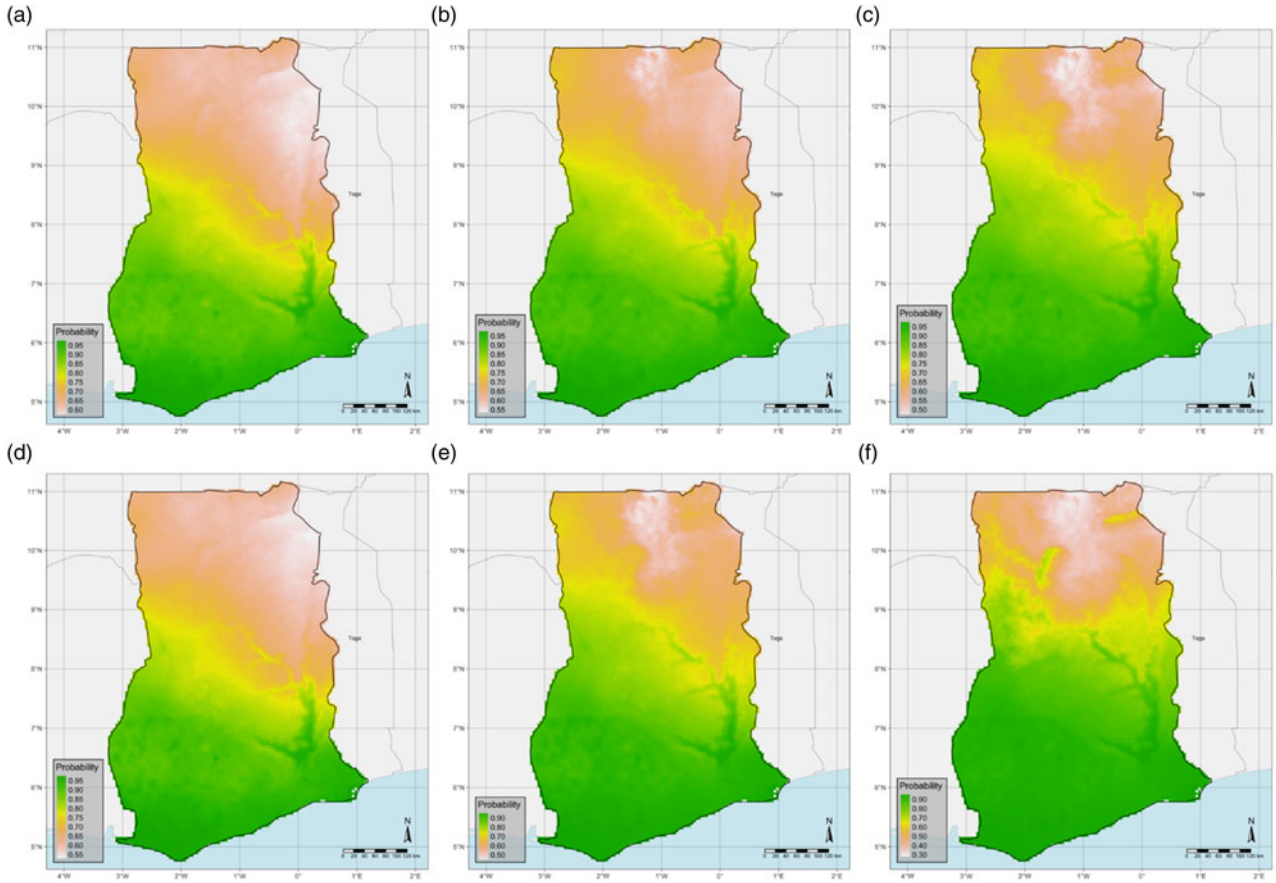


**Figure 9.** Mean future predictions of suitable areas for *Diaphorina citri*. (a) SSP 245 (2021-2040), (b) SSP 245 2041-2060, (c) SSP 245 2061-2080, (d) SSP 585 2021-2040, (e) SSP 585 2041-2060 and (f) SSP 585 2061-2080.

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

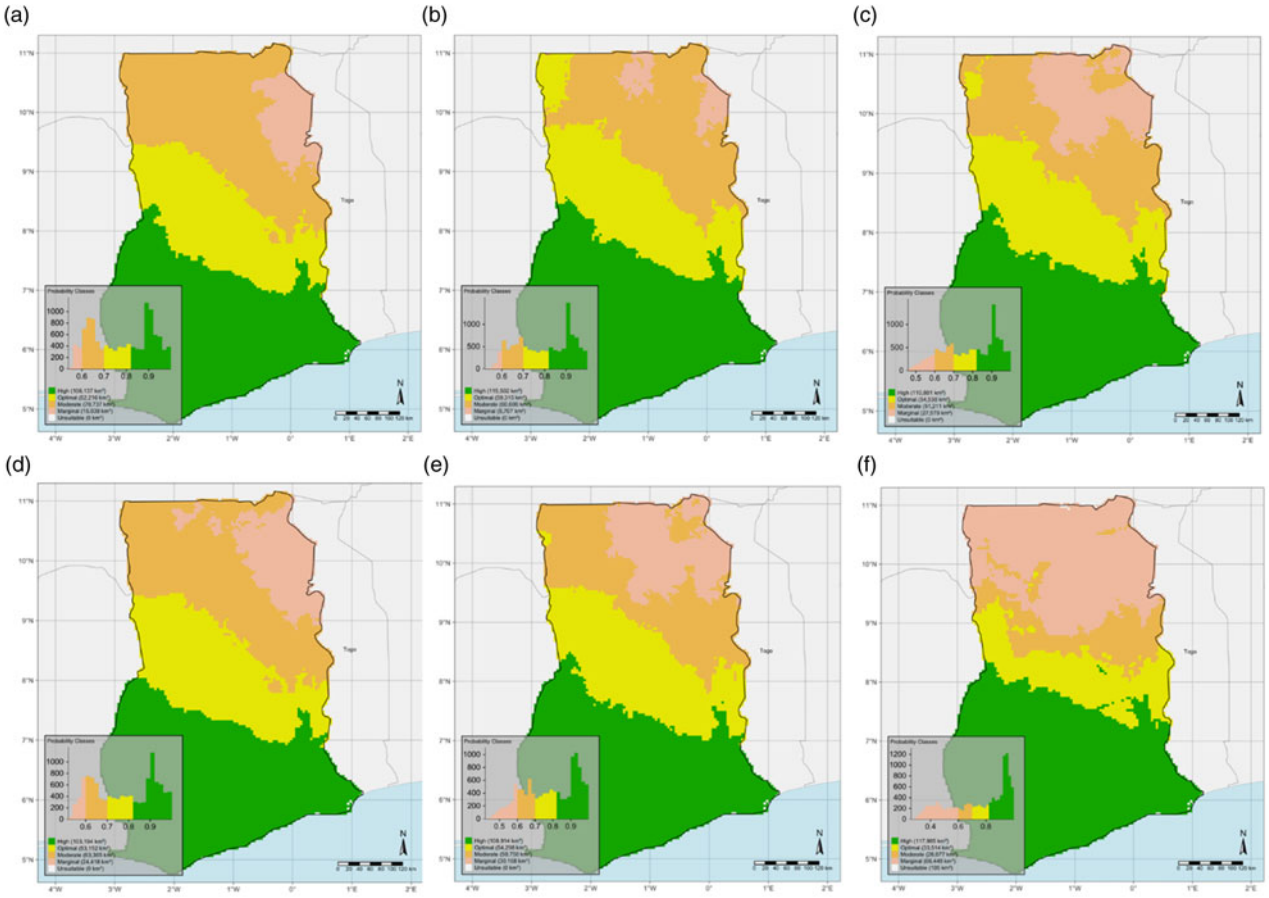


**Figure 10.** Classes of future predictions of suitable areas for *Diaphorina citri*. (a) SSP 245 (2021-2040), (b) SSP 245 2041-2060, (c) SSP 245 2061-2080, (d) SSP 585 2021-2040, (e) SSP 585 2041-2060 and (f) SSP 585 2061-2080.

encompassing biotic interactions, reproductive strategies and other relevant characteristics. The ecological significance of response curves may vary depending on the particular context and size of the investigation. Varying reaction patterns may arise due to different life phases, geographical regions or year's seasons. Maxent models can capture intricate interactions; however, it is essential to note that overfitting, which entails fitting noise present in the data, can lead to unrealistic response curves that then need to be revised (Phillips *et al.*, 2009). Thorough model validation and careful evaluation of ecological plausibility are crucial aspects to be considered when modelling species distribution. Nonetheless, to thoroughly comprehend species biology, it is necessary to incorporate broader ecological information, field observations and domain experience alongside the findings obtained using maxnet in future investigations and analysis.

SDMs, like the maxnet package in R, can have some important drawbacks. We did not consider the potential effects of socioeconomic development, *D. citri* evolution and adaptation, the introduction of new plant protection and regulatory services policies, and changed farm-level management practices, all of which could considerably change the spread risk, in addition to human movement and interventions. Our predictions are based on SSPs, and the chosen combinations of emission and socioeconomic scenarios can influence our projections. Thus, different modelling methods and their combinations should be used in future studies to provide a better understanding of the uncertainty around our estimations. Biotic and abiotic factors such as pest pressure, predators, parasitoids and elevation that were not included in our model should be considered in future predictions. Notwithstanding these drawbacks, our results offer valuable data for developing surveillance and preventive policies against a further spread of *D. citri* in Ghana and the pest's impact on the country's citrus sector.

## Conclusion

The present investigation examined the suitable ecological environment for *D. citri* in Ghana, including both present and projected climate change conditions represented by the SSP245 and SSP585 scenarios. Such an analysis is crucial for the formulation of effective strategies and policies aimed at an effective control of *D. citri*. The potential habitats of the pest was predicted using the maxnet in R. Our findings indicate that temperature and rainfall conditions are significant predictors for the possible distribution of *D. citri* in Ghana. Thus, the current climatic regions in Ghana are highly conducive for *D. citri* and are expected to expand until the 2080s. Still, it is projected that certain crucial citrus-producing regions in southern Ghana will continue to exhibit a high level of suitability for *D. citri*. Our findings can thus assist researchers and policymakers in developing effective and well-targeted pest management strategies for *D. citri* in a changing climate.

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

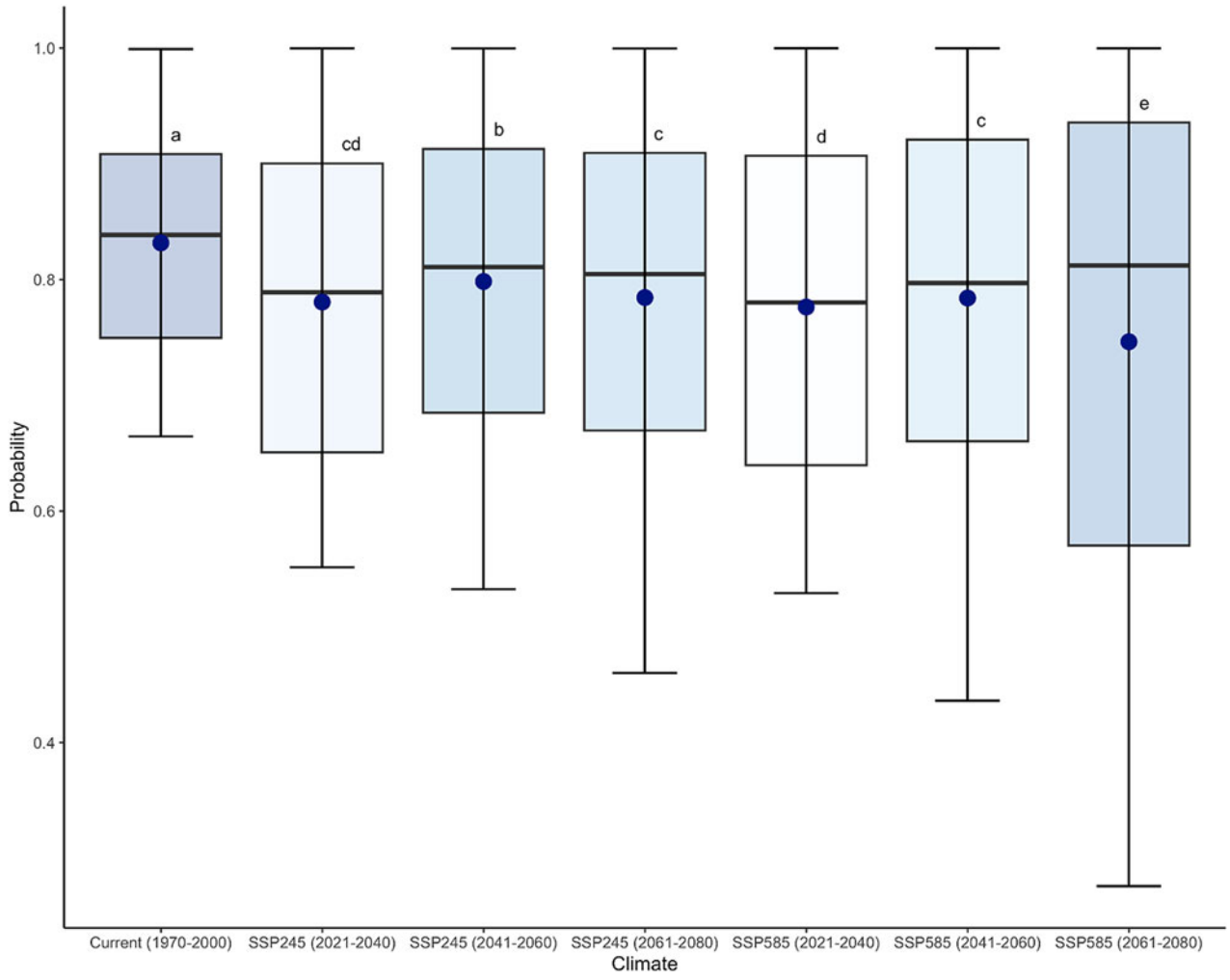


**Figure 11.** Tukey climates from the current time (1970–2000) to the future (2021–2040; 2021–2060; 2061–2080) under SSPs 245 and 585.

**Supplementary material.** The supplementary material for this article can be found at https://doi.org/10.1017/S0007485324000105.

**Data availability.** The datasets generated in this study are available from the corresponding author upon request.

**Acknowledgements.** We thank the Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM) and the University of Environment and Sustainable Development (UESD) for access to instruments. We would like to thank the National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq) and the Minas Gerais State Research Foundation (Fundação de Amparo a Pesquisa do Estado de Minas Gerais – FAPEMIG) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

**Author contributions.** The manuscript was conceived, designed and written by K. D. N. and O. F. A. Material preparation and data collection were conducted by K. D. N., G. C. A., J. O.-O., O. F. A., F. K. A., A. K. D., W. K. H., G. E., L. K. A., P. B. and H. A. B. Analysis were performed by P. G. C. S., G. C. A., E. J. D. V. B. and O. F. A. The manuscript was reviewed and edited by R. S. d. S. C. B. and M. S. contributed to review and editing of the article. All authors read and approved the final manuscript.

**Financial support.** Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM).

**Competing interests.** None.

## References

**Aarts G, Fieberg J and Matthiopoulos J** (2012) Comparative interpretation of count, presence–absence and point methods for species distribution models. *Methods in Ecology and Evolution* **3**, 177–187.

**Aidoo OF, Tanga CM, Mohamed SA, Rasowo BA, Khamis FM, Rwomushana I, Kimani J, Agyakwa AK, Daisy S, Sétamou M, Ekesi S and Borgemeister C** (2019) Distribution, degree of damage and risk of spread of *Trioza erytreae* (Hemiptera: Triozidae) in Kenya. *Journal of Applied Entomology* **143**, 822–833.

**Aidoo OF, Cunze S, Guimapi RA, Arhin L, Ablormeti FK, Tettey E, Dampare F, Afram Y, Bonsu O, Obeng J, Lutuf H, Dickinson M and Yankey N** (2021) Lethal yellowing disease: insights from predicting potential distribution under different climate change scenarios. *Journal of Plant Diseases and Protection* **128**, 1313–1325.

**Aidoo OF, Souza PG, da Silva RS, Santana PA Jr., Picanço MC, Kyerematen R, Sètamou M, Ekesi S and Borgemeister C** (2022) Climate-induced range shifts of invasive species (*Diaphorina citri* Kuwayama). *Pest Management Science* **78**, 2534–2549.

**Aidoo OF, Souza PG, Silva RS, Júnior PA, Picanço MC, Heve WK, Duker RQ, Ablormeti FK, Sétamou M and Borgemeister C** (2023a) Modeling climate change impacts on potential global distribution of *Tamarixia radiata* Waterston (Hymenoptera: Eulophidae). *Science of The Total Environment* **864**, 160962.

**Aidoo OF, Ablormeti FK, Ninsin KD, Antwi-Agyakwa AK, Osei-Owusu J, Heve WK, Dofuor AK, Soto YL, Edusei G, Osabutey AF, Sossah FL, Aryee CO, Alabi OJ and Sétamou M** (2023b) First report on the presence of huanglongbing vectors (*Diaphorina citri* and *Trioza erytreae*) in Ghana. *Scientific Reports* **13**, 11366.

**Allouche O, Tsoar A and Kadmon R** (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal Applied Ecology* **43**, 1223–1232.

**Amaro G, Fidelis EG, da Silva RS and Marchioro CA** (2023) Effect of study area extent on the potential distribution of species: a case study with models for *Raoiella indica* Hirst (Acari: Tenuipalpidae). *Ecological Modelling* **483**, 110454.

**Anderson RP and Raza A** (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela: effect of study region on models of distributions. *Journal of Biogeography* **37**, 1378–1393.

**Antolínez CA, Olarte-Castillo XA, Martini X and Rivera MJ** (2022) Influence of daily temperature maximums on the development and short-distance movement of the Asian citrus psyllid. *Journal of Thermal Biology* **110**, 103354.

**Araújo MB and Guisan A** (2006) Five (or so) challenges for species distribution modelling. *Journal Biogeography* **33**, 1677–1688.

**Araújo MB, Anderson RP, Márcia Barbosa A, Beale CM, Dormann CF, Early R, Garcia RA, Guisan A, Maiorano L, Naimi B, O'Hara RB, Zimmermann NE and Rahbek C** (2019) Standards for distribution models in biodiversity assessments. *Science Advances* **5**, eaat4858.

**Asare-Bediako E, Addo-Quaye AA, Tetteh JP, Buah JN, Van Der Puije GC and Acheampong RA** (2013) Prevalence of mistletoe on citrus trees in the Abura-Asebu-Kwamankese district of the Central Region of Ghana. *International Journal of Scientific & Technology Research* **2**, 122–127.

**Austin MP and Van Niel KP** (2011) Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography* **38**, 1–8.

**Barbet-Massin M, Jiguet F, Albert CH and Thuiller W** (2012) Selecting pseudo-absences for species distribution models: how, where and how many?: how to use pseudo-absences in niche modelling? *Methods in Ecology and Evolution* **3**, 327–338.

**Barrett SCH** (2000) Microevolutionary influences of global change on plant invasions. In Mooney HA and Hobbs RK (eds), *The Impact of Global Change on Invasive Species*. Covelo, CA: Island Press, pp. 115–139.

**Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Soberón J and Villalobos F** (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling* **222**, 1810–1819.

**Bayles BR, Thomas SM, Simmons GS, Grafton-Cardwell EE and Daugherty MP** (2017) Spatiotemporal dynamics of the Southern California Asian citrus psyllid (*Diaphorina citri*) invasion. *PLoS ONE* **12**, e0173226.

**Beattie GA** (2020) 12 Management of the Asian Citrus Psyllid in Asia. In *Asian Citrus Psyllid: Biology, Ecology and Management of the Huanglongbing Vector*, p. 179.

**Beaumont LJ, Gallagher RV, Thuiller W, Downey PO, Leishman MR and Hughes L** (2009) Different climatic envelopes among invasive populations may lead to underestimations of current and future bio- logical invasions. *Diversity and Distribution*, **15**, 409–420.

**Bové JM** (2006) Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. *Journal of Plant Pathology* **88**, 7–37.

**Bové JM** (2014) Huanglongbing or yellow shoot, a disease of Gondwanan origin: will it destroy citrus worldwide? *Phytoparasitica* **42**, 579–583.

**Bradley AP** (1997) The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition* **30**, 1145–1159.

**Brentu FC, Oduro KA, Offei SK, Odamtten GT, Vicent A, Peres NA and Timmer LW** (2012) Crop loss, aetiology, and epidemiology of citrus black spot in Ghana. *European Journal of Plant Pathology* **133**, 657–670.

**Broennimann O and Guisan A** (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters*, **4**, 585–589.

**Broennimann O, Fitzpatrick MC, Pearman PB, Petitpierre B, Pellissier L, Yoccoz NG, Thuiller W, Fortin M-J, Randin C, Zimmermann NE, Graham CH and Guisan A** (2012) Measuring ecological niche overlap from occurrence and spatial environmental data: measuring niche overlap. *Global Ecology and Biogeography* **21**, 481–497.

**Broennimann O, Di Cola V and Guisan A** (2022) ecospat: spatial ecology miscellaneous methods, R package ver. 3.2.1. https://CRAN.R-project.org/package=ecospat

**Castellanos AA, Huntley JW, Voelker G and Lawing AM** (2019) Environmental filtering improves ecological niche models across multiple scales. *Methods in Ecology and Evolution*, **10**, 481–492.

**Chamberlain S, Barve V, Mcglinn D, Oldoni D, Desmet P, Geffert L and Ram K** (2023) rgbif: Interface to the Global Biodiversity Information Facility API. Website: https://CRAN.R-project.org/package=rgbif (accessed 13 March 2024).

**Cooper JC and Soberón J** (2018) Creating individual accessible area hypotheses improves stacked species distribution model performance. *Global Ecology and Biogeography* **27**, 156–165.

**Devi HS and Sharma DR** (2014) Impact of abiotic factors on build-up of citrus psylla, *Diaphorina citri* Kuwayama population in Punjab, India. *Journal of Applied and Natural Science* **6**, 371–376.

**Dray S and Dufour A-B** (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* **22**, 1–20.

**Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS and Zimmermann NE** (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151.

**Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE and Yates CJ** (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* **17**, 43–57.

**Fick SE and Hijmans RJ** (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* **37**, 4302–4315.

**Finch DM, Butler JL, Runyon JB, Fettig CJ, Kilkenny FF, Jose S, Frankel SJ, Cushman SA, Cobb RC, Dukes JS, Hicke JA and Amelon SK** (2021) Effects of climate change on invasive species. Invasive species in forests and rangelands of the United States: a comprehensive science synthesis for the United States forest sector, 57-83.

**Fithian W and Hastie T** (2013) Finite-sample equivalence in statistical models for presence-only data. *The Annals of Applied Statistics* **7**, 1917.

**Fitzpatrick MC and Hargrove WW** (2009) The projection of species distribution models and the problem of non-analog climate. *Biodiversity and Conservation* **18**, 2255–2261.

**Friedl M and Sulla-Menashe D** (2022) MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 0.05Deg CMG V061 [Data set]. NASA EOSDIS Land Processes Distributed Active Archive Center. Accessed 2023-10-11 from https://doi.org/10.5067/MODIS/MCD12C1.061

**Friedman J, Hastie T and Tibshirani R** (2010) Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software* **33**, 1.

**Guillera-Arroita G, Lahoz-Monfort JJ, Elith J, Gordon A, Kujala H, Lentini PE, McCarthy MA, Tingley R and Wintle BA** (2015) Is my species distribution model fit for purpose? Matching data and models to applications: matching distribution models to applications. *Global Ecology and Biogeography* **24**, 276–292.

**Guisan A, Petitpierre B, Broennimann O, Daehler C and Kueffer C** (2014) Unifying niche shift studies: insights from biological invasions. *Trends in Ecology and Evolution* **29**, 260–269.

**Hall DG, Wenninger EJ and Hentz MG** (2011) Temperature studies with the Asian citrus psyllid, *Diaphorina citri*: cold hardiness and temperature thresholds for oviposition. *Journal of Insect Science* **11**, 83.

**Hall DG, Hentz MG, Meyer JM, Kriss AB, Gottwald TR and Boucias DG** (2012) Observations on the entomopathogenic fungus *Hirsutella citriformis* attacking adult *Diaphorina citri* (Hemiptera: Psyllidae) in a managed citrus grove. *BioControl* **57**, 663–675.

**Hall DG, Richardson ML, Ammar ED and Halbert SE** (2013) Asian citrus psyllid, Diaphorina citri, vector of citrus huanglongbing disease. *Entomologia Experimentalis et Applicata* **146**, 207–223.

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

**Heikkinen RK, Marmion M and Luoto M** (2012) Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography (Cop.)* **35**, 276–288.

**Helmstetter NA, Conway CJ, Stevens BS and Goldberg AR** (2021) Balancing transferability and complexity of species distribution models for rare species conservation. *Diversity and Distributions* **27**, 95–108.

**Hijmans RJ** (2012) Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology* **93**, 679–688.

**Hijmans RJ** (2023) terra: Spatial data analysis. https://CRAN.R-project.org/package=terra

**Hijmans RJ, Barbosa M, Ghosh A and Mandel A** (2023a) geodata: Download geographic data.

**Hill MP, Gallardo B, Terblanche JS** (2017) A global assessment of climatic niche shifts and human influence in insect invasions: HILL et al. *Global Ecology Biogeography* **26**, 679–689.

**Hosmer DW, Lemeshow S and Sturdivant RX** (2013) *Applied Logistic Regression, Applied Logistic Regression*, 3rd Edn. Wiley. https://doi.org/10.1002/9781118548387. Wiley series in probability and statistics.

**Jarnevich CS, Stohlgren TJ, Kumar S, Morisette JT and Holcombe TR** (2015) Caveats for correlative species distribution modeling. *Ecological Informatics* **29**, 6–15.

**Jiang Y, Metz CE and Nishikawa RM** (1996) A receiver operating characteristic partial area index for highly sensitive diagnostic tests. *Radiology* **201**, 745–750.

**Kassambara A and Mundt F** (2022) factoextra: Extract and visualize the results of multivariate data analyses (1.0.7) [R]. https://cran.rproject.org/web/packages/factoextra/index.html

**Khan AM, Li Q, Saqib Z, Khan N, Habib T, Khalid N, Majeed M and Tariq A** (2022) MaxEnt modelling and impact of climate change on habitat suitability variations of economically important chilgoza pine (*Pinus gerardiana* wall.) in South Asia. *Forests* **13**, 715.

**Khosravi R, Hemani MR, Malekian M, Flint A and Flint L** (2016) Maxent modeling for predicting potential distribution of goitered gazelle in central Iran: the effect of extent and grain size on performance of the model. *Turkish Journal of Zoology* **40**, 574–585.

**Landis JR and Koch GG** (1977) The measurement of observer agreement for categorical data. *Biometrics* **33**, 159.

**Lawson CR, Hodgson JA, Wilson RJ and Richards SA** (2014) Prevalence, thresholds and the performance of presence–absence models. *Methods in Ecology and Evolution* **5**, 54–64.

**Lee JA, Halbert SE, Dawson WO, Robertson CJ, Keesling JE and Singer BH** (2015) Asymptomatic spread of huanglongbing and implications for disease control. *Proceedings of the National Academy of Sciences*, **112**, 7605–7610.

**Liu C, Berry PM, Dawson TP and Pearson RG** (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28**, 385–393.

**Liu C, White M and Newell G** (2013) Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography* **40**, 778–789.

**Liu C, Newell G and White M** (2016) On the selection of thresholds for predicting species occurrence with presence-only data. *Ecology and Evolution* **6**, 337–348.

**López-Collado J, López-Arroyo JI, Robles-García PL and Márquez-Santos M** (2013) Geographic distribution of habitat, development, and population growth rates of the Asian citrus psyllid, *Diaphorina citri*, in Mexico. *Journal of Insect Science* **13**.

**Low BW, Zeng Y, Tan HH and Yeo DC** (2021) Predictor complexity and feature selection affect Maxent model transferability: evidence from global freshwater invasive species. *Diversity and Distributions* **27**, 497–511.

**Luo Y and Agnarsson I** (2018) Global mt DNA genetic structure and hypothesized invasion history of a major pest of citrus, *Diaphorina citri* (Hemiptera: Liviidae). *Ecology and Evolution*, **8**, 257–265.

**Machado-Stredel F, Cobos ME and Peterson AT** (2021) A simulation-based method for selecting calibration areas for ecological niche models and species distribution models. *Frontiers of Biogeography* **13**. https://doi.org/10.21425/F5FBG48814.

**Massicotte P and South A** (2023) rnaturalearth: World Map Data from Natural Earth. R package version 0.3.3.9000. Available from https://docs.ropensci.org/rnaturalearth/index.html (accessed April 2023).

**McClish DK** (1989) Analyzing a portion of the ROC curve. *Medical Decision Making* **9**, 190–195.

**Merow C, Smith MJ and Silander JA** (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography (Cop.)*. **36**, 1058–1069.

**Milosavljević I, McCalla KA, Morgan DJ and Hoddle MS** (2020) The effects of constant and fluctuating temperatures on development of *Diaphorina citri* (Hemiptera: Liviidae), the Asian citrus psyllid. *Journal of Economic Entomology* **113**, 633–645.

**Naeem A, Freed S, Jin FL, Akmal M and Mehmood M** (2016) Monitoring of insecticide resistance in *Diaphorina citri* Kuwayama (Hemiptera: Psyllidae) from citrus groves of Punjab, Pakistan. *Crop Protection*, **86**, 62–68.

**Nahrung HF, Liebhold AM, Brockerhoff EG and Rassati D** (2023) Forest insect biosecurity: processes, patterns, predictions, pitfalls. *Annual Review of Entomology*, **68**, 211–229.

**Northrup JM, Hooten MB, Anderson CR Jr. and Wittemyer G** (2013) Practical guidance on characterizing availability in resource selection functions under a use–availability design. *Ecology* **94**, 1456–1463.

**O'Donnell MS and Ignizio DA** (2012) Bioclimatic predictors for supporting ecological applications in the conterminous United States: U.S. *Geological Survey Data Series* **691**, 10.

**Oduro C, Shuoben B, Ayugi B, Beibei L, Babaousmail H, Sarfo I, Ullah S and Ngoma H** (2021) Observed and Coupled Model Intercomparison Project 6 multimodel simulated changes in near-surface temperature properties over Ghana during the 20th century. *International Journal of Climatology*, **42**, 1–21.

**Owens HL, Campbell LP, Dornak LL, Saupe EE, Barve N, Soberón J, Ingenloff K, Lira-Noriega A, Hensz CM, Myers CE and Peterson AT** (2013) Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecological Modelling* **263**, 10–18.

**Padayachee AL, Irlich UM, Faulkner KT, Gaertner M, Procheș Ș, Wilson JR and Rouget M** (2017) How do invasive species travel to and through urban environments?. *Biological Invasions*, **19**, 3557–3570.

**Paris TM, Allan SA, Hall DG, Hentz MG, Croxton SD, Ainpudi N and Stansly PA** (2017) Effects of temperature, photoperiod, and rainfall on morphometric variation of *Diaphorina citri* (Hemiptera: Liviidae). *Environmental Entomology* **46**, 143–158.

**Parravicini V, Azzurro E, Kulbicki M and Belmaker J** (2015) Niche shift can impair the ability to predict invasion risk in the marine realm: an illustration using Mediterranean fish invaders. *Ecology Letters* **18**, 246–253.

**Pebesma EJ** (2018) Simple features for R: standardized support for spatial vector data. *R Journal* **10**, 439.

**Perkins-Taylor IE and Frey JK** (2020) Predicting the distribution of a rare chipmunk (*Neotamias quadrivittatus oscuraensis*): comparing MaxEnt and occupancy models. *Journal of Mammalogy* **101**, 1035–1048.

**Peterson AT** (2006) Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics* **3**, 59–72.

**Phillips SJ** (2017) A brief tutorial on Maxent. Disponível: http://biodiversityinformatics.amnh.org/open_source/maxent/. Acesso em: 20-nov-2018.

**Phillips SJ and Dudík M** (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* **31**, 161–175.

**Phillips S and Phillips MS** (2021) Package 'maxnet'. Version 0.1, 4. https://github.com/mrmaxent/maxnet (Last accessed 2nd February 2024).

**Phillips SJ, Dudík M and Schapire RE** (2004) A maximum entropy approach to species distribution modeling. In *Proceedings of the Twenty-First International Conference on Machine Learning* (*ICML* 2004), Banff, Alberta, Canada, July 4–8, 2004, p. 83.

**Phillips SJ, Anderson RP and Schapire RE** (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231–259.

**Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J and Ferrier S** (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications* **19**, 181–197.

**Phillips SJ, Anderson RP, Dudík M, Schapire RE and Blair ME** (2017) Opening the black box: an open-source release of Maxent. *Ecography* **40**, 887–893.

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

**Radosavljevic A and Anderson RP** (2014) Making better Maxent models of species distributions: complexity, overfitting and evaluation. *Journal of biogeography* **41**, 629–643.

**Randin CF, Dirnböck T, Dullinger S, Zimmermann NE, Zappa M and Guisan A** (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography* **33**, 1689–1703.

**R Core Team** (2023) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Available at https://www.R-project.org/

**Renner IW and Warton DI** (2013) Equivalence of MAXENT and Poisson point process models for species distribution modeling in ecology: equivalence of MAXENT and Poisson point process models. *BIOM* **69**, 274–281.

**Renner IW, Elith J, Baddeley A, Fithian W, Hastie T, Phillips SJ, Popovic G and Warton DI** (2015) Point process models for presence-only analysis. *Methods in Ecology and Evolution* **6**, 366–379.

**Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, Hauenstein S, Lahoz-Monfort JJ, Schröder B, Thuiller W, Warton DI, Wintle BA, Hartig F and Dormann CF** (2017) Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography* **40**, 913–929.

**Robin X, Turck N, Hainard A, Lisacek F, Sanchez J-C and Müller M** (2009) Bioinformatics for protein biomarker panel classification: what is needed to bring biomarker panels into in vitro diagnostics? *Expert Review of Proteomics* **6**, 675–689.

**Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C and Müller M** (2011) pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77.

**Rubel F and Kottek M** (2010) Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen-Geiger climate classification. *metz* **19**, 135–141.

**Santini L, Benítez-López A, Maiorano L, Čengić M and Huijbregts MAJ** (2021) Assessing the reliability of species distribution projections in climate change research. *Diversity and Distribution* **27**, 1035–1050.

**Schneider L, Rebetez M and Rasmann S** (2022) The effect of climate change on invasive crop pests across biomes. *Current Opinion in Insect Science*, **50**, 100895.

**Sétamou M, Soto YL, Tachin M and Alabi OJ** (2023) Report on the first detection of Asian citrus psyllid *Diaphorina citri* Kuwayama (Hemiptera: Liviidae) in the Republic of Benin, West Africa. *Scientific Reports*, **13**, 801.

**Shabani F, Kumar L and Ahmadi M** (2018) Assessing accuracy methods of species distribution models: AUC, specificity, sensitivity and the true skill statistic. *Global Journal of Human-Social Science: B Geography, Geo-Sciences, Environmental Science & Disaster Management* **18**.

**Shcheglovitova M and Anderson RP** (2013) Estimating optimal complexity for ecological niche models: a jackknife approach for species with small sample sizes. *Ecological Modelling* **269**, 9–17.

**Simberloff D and Gibbons L** (2004) Now you see them, now you don't! – population crashes of established introduced species. *Biological Invasions*, **6**, 161–172.

**Skendžić S, Zovko M, Pajač Živković I, Lešić V and Lemić D** (2021) Effect of climate change on introduced and native agricultural invasive insect pests in Europe. *Insects* **12**, 985.

**Smith AB** (2013) On evaluating species distribution models with random background sites in place of absences when test presences disproportionately sample suitable habitat. *Diversity and Distributions* **19**, 867–872.

**Streiner DL and Cairney J** (2007) What's under the ROC? An introduction to receiver operating characteristics curves. *Canadian Journal of Psychiatry*, 121–128. doi: 10.1177/070674370705200210, PMID: 17375868.

**Tennekes M** (2018) tmap: Thematic maps in R. *Journal of Statistical Software* **84**, 1–39.

**Tesfamariam BG, Gessesse B and Melgani F** (2022) MaxEnt-based modeling of suitable habitat for rehabilitation of Podocarpus forest at landscape-scale. *Environmental Systems Research* **11**, 1–12.

**Tsai JH and Liu YH** (2000) Biology of *Diaphorina citri* (Homoptera: Psyllidae) on four host plants. *Journal of Economic Entomology* **93**, 1721–1725.

**Uden DR, Mech AM, Havill NP, Schulz AN, Ayres MP, Herms DA, Hoover AM, Gandhi KJ, Hufbauer RA, Liebhold AM and Marsico TD** (2023) Phylogenetic risk assessment is robust for forecasting the impact of European insects on North American conifers. *Ecological Applications*, **33**, e2761.

**Valavi R, Guillera-Arroita G, Lahoz-Monfort JJ and Elith J** (2022) Predictive performance of presence-only species distribution models: a benchmark study with reproducible code. *Ecological Monographs* **92**. https://doi.org/10.1002/ecm.1486.

**VanDerWal J, Shoo LP, Johnson CN and Williams SE** (2009) Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *The American Naturalist* **174**, 282–291.

**Varela S, Anderson RP, Garcia-Valdes R and Fernandez-Gonzalez F** (2014) Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography*, **37**, 1084–1091.

**Velazco SJE, Villalobos F, Galvão F and De Marco Júnior P** (2019) A dark scenario for Cerrado plant species: effects of future climate, land use and protected areas ineffectiveness. *Diversity and Distributions* **25**, 660–673.

**Velazco SJE, Svenning J-C, Ribeiro BR and Laureto LMO** (2020) On opportunities and threats to conserve the phylogenetic diversity of Neotropical palms. *Diversity and Distributions*, **27**, 512–523.

**Velazco SJE, Rose MB, de Andrade AFA, Minoli I and Franklin J** (2022) flexsdm: An R package for supporting a comprehensive and flexible species distribution modelling workflow. *Methods Ecology and Evolution* **13**, 1661–1669.

**Venette RC** (2017) Climate analyses to assess risks from invasive forest insects: simple matching to advanced models. *Current Forestry Reports* **3**, 255–268.

**Wang S, Xiao Y and Zhang H** (2015) Studies of the past, current and future potential distributions of *Diaphorina citri* Kuwayama (Homoptera: Psyllidae) in China. *Chinese Journal of Applied Entomology* **52**, 1140–1148.

**Wang R, Yang H, Luo W, Wang M, Lu X, Huang T, Zhao J and Li Q** (2019) Predicting the potential distribution of the Asian citrus psyllid, *Diaphorina citri* (Kuwayama), in China using the MaxEnt model. *PeerJ* **7**, e7323.

**Wang R, Yang H, Wang M, Zhang Z, Huang T, Wen G and Li Q** (2020) Predictions of potential geographical distribution of *Diaphorina citri (Kuwayama)* in China under climate change scenarios. *Scientific Reports* **10**, 9202.

**Warren DL, Matzke NJ and Iglesias TL** (2020) Evaluating presence-only species distribution models with discrimination accuracy is uninformative for many applications. *Journal of Biogeography* **47**, 167–180.

**Webber BL, Yates CJ, Le Maitre DC, Scott JK, Kriticos DJ, Ota N, McNeill A, Le Roux JJ, Midgley GF** (2011) Modelling horses for novel climate courses: insights from projecting potential distributions of native and alien Australian acacias with correlative and mechanistic models. *Diversity and Distributions* **17**, 978–1000.

**Wei T and Simko V** (2021) R package 'corrplot': Visualization of a Correlation Matrix. (Version 0.92). https://github.com/taiyun/corrplot

**Welker S, Pierre M, Santiago JP, Dutt M, Vincent C and Levy A** (2022) Phloem transport limitation in Huanglongbing-affected sweet orange is dependent on phloem-limited bacteria and callose. *Tree Physiology*, **42**, 379–390.

**Williams JW, Jackson ST and Kutzbach JE** (2007) Projected distributions of novel and disappearing climates by 2100 AD. *Proceedings of the National Academy of Sciences* **104**, 5738–5742.

**Woodman SM, Forney KA, Becker EA, DeAngelis ML, Hazen EL, Palacios DM and Redfern JV** (2019) Esdm: a tool for creating and exploring ensembles of predictions from species distribution and abundance models. *Methods in Ecology and Evolution* **10**, 1923–1933.

**Yates KL, Bouchet PJ, Caley MJ, Mengersen K, Randin CF, Parnell S, Fielding AH, Bamford AJ, Ban S, Barbosa AM, Dormann CF, Elith J, Embling CB, Ervin GN, Fisher R, Gould S, Graf RF, Gregr EJ, Halpin PN, Heikkinen RK, Heinänen S, Jones AR, Krishnakumar PK, Lauria V, Lozano-Montes H, Mannocci L, Mellin C, Mesgaran MB, Moreno-Amat E, Mormede S, Novaczek E, Oppel S, Ortuño Crespo G, Peterson AT, Rapacciuolo G, Roberts JJ, Ross RE, Scales KL, Schoeman D, Snelgrove P, Sundblad G, Thuiller W, Torres LG, Verbruggen H, Wang L, Wenger S, Whittingham MJ, Zharikov Y, Zurell D and Sequeira AMM** (2018) Outstanding challenges in the transferability of ecological models. *Trends in Ecology and Evolution* **33**, 790–802.

**Zavala-Zapata V, Lázaro-Dzul MO, Sánchez-Borja M, Vargas-Tovar JA, Álvarez-Ramos R and Azuara-Domínguez A** (2022) Abundance of

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press

*Diaphorina citri* Kuwayama1 associated with temperature and precipitation at Tamaulipas, Mexico. *Southwestern Entomologist* **47**, 713–722.

**Zhang Z, Mammola S, McLay CL, Capinha C and Yokota M** (2020) To invade or not to invade? Exploring the niche-based processes underlying the failure of a biological invasion using the invasive Chinese mitten crab. *Total Environment*, **728**, 1–10.

**Zhu GP, Liu Q and Gao YB** (2014) Improving ecological niche model transferability to predict the potential distribution of invasive exotic species. *Biodiversity Science* **22**, 223–230.

**Zizka A, Silvestro D, Andermann T, Azevedo J, Duarte Ritter C, Edler D, Farooq H, Herdean A, Ariza M, Scharn R, Svantesson S, Wengström N, Zizka V and Antonelli A** (2019) CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, **10**, 744–751.

**Zorzenon FP, Tomaseto AF, Daugherty MP, Lopes JR and Miranda MP** (2021) Factors associated with *Diaphorina citri* immigration into commercial citrus orchards in São Paulo State, Brazil. *Journal of Applied Entomology* **145**, 326–335.

https://doi.org/10.1017/S0007485324000105 Published online by Cambridge University Press