



# Risk analysis of the spread of the quarantine pest mite *Schizotetranychus hindustanicus* in Brazil

George Amaro<sup>1</sup> · Elisangela Gomes Fidelis<sup>2</sup> · Cristian Madeira de Medeiros<sup>3</sup> · Ricardo Siqueira da Silva<sup>4</sup>

Received: 7 July 2022 / Accepted: 27 October 2022  
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

## Abstract

*Schizotetranychus hindustanicus* Hirst (Acari: Tetranychidae) known as the Hindustan citrus mite, is a quarantine pest present in Roraima, Brazil. In 1924 this pest was described in India. It was reported in 2002 in Venezuela and in Roraima in 2008. In 2010, the Hindustan citrus mite was reported in Colombia. It is possible that it will be introduced in other areas of Brazil, resulting in a threat to Brazilian citrus industry. Our objective was to determine the most suitable regions of Brazil for *S. hindustanicus* using a maximum entropy (Maxent) algorithm, based on native and invasive updated occurrence records from published research, field surveys and online databases. To avoid overfitting and improving transferability, we chose parameter settings of Maxent to construct and validate models by searching for the best combination of feature classes and regularization multipliers. The model obtained showed excellent performance according to all evaluation metrics used. A high potential for the establishment of *S. hindustanicus* was identified in large areas of Roraima, the extreme west of Amazonas, the entire north of the State of Pará, also in northeast, south, east, and north of the State of Amapá, and in a small portion northwest of the State of Maranhão (all states belonging to the northern region of Brazil). Our results provide information for policy making and quarantine measures, especially where *S. hindustanicus* is still absent in Brazil.

**Keywords** Hindustan citrus mite · Ecological niche modeling · Environmental modeling · Maxent

---

✉ Ricardo Siqueira da Silva  
ricardo.ufvjm@gmail.com

<sup>1</sup> Embrapa Roraima, Boa Vista, Roraima 69301-970, Brazil

<sup>2</sup> Embrapa Cerrados, Brasília, DF, Brazil

<sup>3</sup> Universidad de Salamanca, Biología e Conservación da Biodiversidade, Salamanca, Spain

<sup>4</sup> Departamento de Agronomia, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Diamantina, MG 39100-000, Brazil

## Introduction

*Schizotetranychus hindustanicus*, also known as the Hindustan citrus mite or citrus web nest mite, was described from citrus specimens in India (Hirst 1924). This mite was detected to be infesting lemon trees in northwest Venezuela, Department of Zulia, in 2002 (Quiróz and Dorado 2005). In 2008, the mite was reported in the far north region of Brazil, in Roraima (Navia and Marsaro Jr 2010), also causing severe symptoms on the lemon trees. Subsequently, *S. hindustanicus* was also found in northern Colombia (La Guajira, Magdalena; Mesa 2010). In Venezuela, *S. hindustanicus* is present in all areas where citrus is cultivated, from the northwest (Zulia) to the northeast (Sucre), in the southern and central regions (Nienstaedt and Marcano 2009). It has also been reported to infect citrus in Iran (Sheikholeslamzadeh and Sadeghi 2010).

Severe damage caused by high infestations of *S. hindustanicus* on citrus leaves and fruits has been observed in commercial plantations and small family farms, both in Venezuela and Brazil (Fantine 2011; Quiróz and Dorado 2005). Although there is no quantitative information on citrus production losses caused by *S. hindustanicus*, the infestations cause esthetic damage that reduce the economic value of fresh fruits. *Schizotetranychus hindustanicus* has been reported to infest Tahiti lemon, lime, Rangpur lime, Ponkan mandarin, Valencia orange, tangerine, lemon, sweet orange and Murcott tangor (Quiróz and Marcano 2005; Quiróz and Dorado 2005; Navia and Marsaro Jr 2010).

According to the Brazilian Association of Exporters of Fruit and Derivatives (Abrafrutas), Brazil is the third-largest fruit producer globally. The substantial potential impacts of the spread of *S. hindustanicus* in Brazil could be the imposition of phytosanitary barriers in the international trade of citrus fruits and increased production costs due to pest control measures. According to FAO data, Brazil is one of the largest producers and exporters of citrus and orange juice (<http://faostat3fao.org>).

In 2019, based on data from the Municipal Agricultural Production Survey (PAM) of the Brazilian Institute of Geography and Statistics (IBGE), the harvested area of citrus (orange, lemon, and tangerine) in Brazil was 698,901 ha, 69% of which was in the Southeast Region, with 60% of the total concentrated in São Paulo state. The value of orange production in the country was R\$ 9.5 billion (Brazilian real, ca. 1.8 billion USD), lemon production was R\$ 157 billion, and tangerine production was R\$ 999.7 million. According to data from the Foreign Trade Secretariat (Secex) compiled by Abrafrutas, lemon exports increased 10% in 2019 compared to 2018, and the total value of citrus exports by Brazil in 2019 was US\$ 1,058 million.

The anticipation of potential areas for species introduction is fundamental to prevent economic damage. Thus, predicting areas suitable for *S. hindustanicus* helps plan phytosanitary control policies. The methodology used to identify suitable areas for species invasion may be accessed with ecological niche models (De Andrade et al. 2019). Species distribution modeling (SDM) are developed based on the environmental variables in which species occur. SDM has been applied to different modeling approaches (Elith et al. 2011), and the techniques involved in the method can be adequate to estimate the invasive species potential in the region of the fundamental niche (Jiménez-Valverde et al. 2011). Also, SDM has applications in ecology, evolution, and epidemiology studies (Corsi et al. 1999; Huerta and Peterson 2008; Peterson 2001; Phillips et al. 2006).

Species distribution modeling explores the geographical space that has the most suitable conditions for a species to survive, grow, and reproduce. Thus, SDM results could

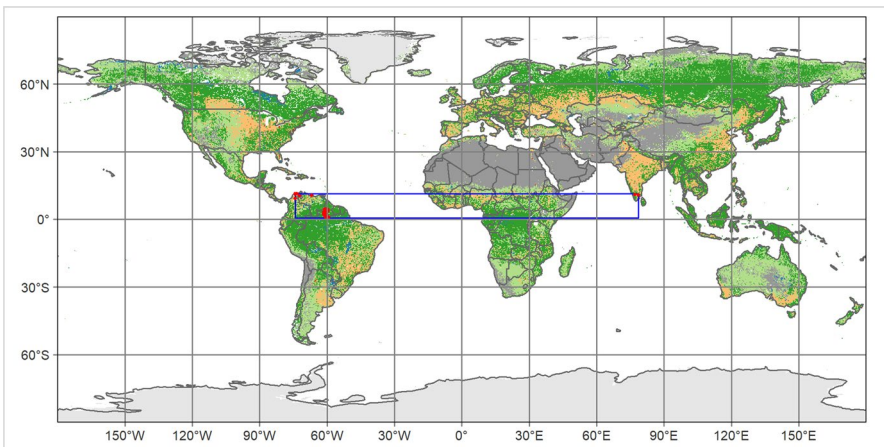
be used to predict areas of species occurrence. The model results may also be used to determine the main potential routes of accessibility for the invasion of species (Barve et al. 2011).

Maxent has been widely used and allows for modelers to use only presence data and categorical predictors (Phillips et al. 2006). Maxent can be efficient for a small sample of reported species (Benito et al. 2009; Hernandez et al. 2006; Papeş and Gaubert 2007; Pearson et al. 2007; Wisz et al. 2008). Results of Maxent are robust and maintain maximum accuracy levels for smaller samples (Hernandez et al. 2006). Therefore, we modeled and mapped Brazil's areas most suitable for the occurrence of *S. hindustanicus* using Maxent with recently registered occurrence data and bioclimatic variables.

## Material and methods

### Occurrence data

Occurrence data for *S. hindustanicus* were obtained from databases CABI and EPPO (CABI/EPPO, 2020), published research (Figueirêdo et al. 2019; Hirst 1924; Mesa 2010; Navia and Marsaro Jr 2010; Nienstaedt and Marcano 2009; Poorani 2017; Quiróz and Dorado 2005), and field sample data from the of Roraima state. 144 occurrence points were registered. Spatial filtering of the data using the R software package *spThin* (Aiello-Lammens et al. 2015) was used to reduce spatial autocorrelation, which is the degree of spatial association present in the datasets that may prevent the separation of the points used for testing. The occurrence data used for the development of the model are show in Fig. 1.



**Fig. 1** Distribution of *Schizotetranychus hindustanicus*: occurrence points (144 red dots) and study area (blue rectangle delimiting a rectangular bounding box buffer around the species extent of occurrence records)

**Table 1** Bioclimatic variables (WorldClim, v.2) that were used for modeling the species distribution of *Schizotetranychus hindustanicus* in Brazil. The variable with codes in bold were used to model fit. The ones with permutation importance in bold were bioclimatic variables that most influenced the potential geographical distribution, as stated in results

<i>Codes</i>	Description	VIF <sup>c</sup>	Contribution %	Permutation importance
<i>Bio01</i>	Annual mean temperature			
<b><i>Bio02</i></b>	Mean diurnal range	3.0622	11.05	<b>49.5700</b>
<b><i>Bio03</i></b>	Isothermality <sup>a</sup>	5.2322	14.85	<b>27.2400</b>
<b><i>Bio04</i></b>	Temperature seasonality <sup>b</sup>	5.5464	4.82	3.1860
<i>Bio05</i>	Max temperature of warmest month			
<i>Bio06</i>	Min temperature of coldest month			
<i>Bio07</i>	Temperature annual range <sup>c</sup>			
<i>Bio08</i>	Mean temperature of wettest quarter			
<b><i>Bio09</i></b>	Mean temperature of driest quarter	1.5038	3.58	1.4190
<i>Bio10</i>	Mean temperature of warmest quarter			
<i>Bio11</i>	Mean temperature of coldest quarter			
<i>Bio12</i>	Annual precipitation			
<i>Bio13</i>	Precipitation of wettest month			
<b><i>Bio14</i></b>	Precipitation of driest month	2.9120	27.05	4.8660
<i>Bio15</i>	Precipitation seasonality <sup>d</sup>			
<i>Bio16</i>	Precipitation of wettest quarter			
<i>Bio17</i>	Precipitation of driest quarter			
<b><i>Bio18</i></b>	Precipitation of warmest quarter	2.3345	17.76	<b>12.2200</b>
<b><i>Bio19</i></b>	Precipitation of coldest quarter	3.1895	20.89	1.4910

<sup>a</sup>(Bio2 / Bio7) × 100

<sup>b</sup>Standard deviation × 100

<sup>c</sup>(Bio5–Bio6)

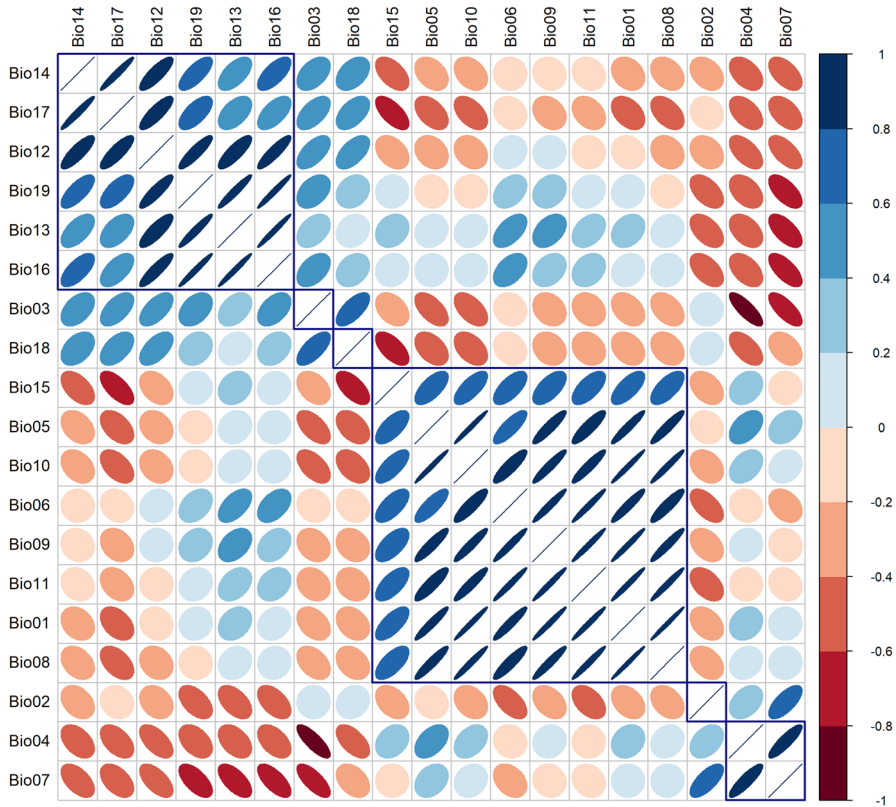
<sup>d</sup>Coefficient of variation

<sup>e</sup>Variance Inflation Factor indicates the extent of the variance of a regression coefficient estimated by one variable that is inflated due to the collinearity with the others

## Bioclimatic variables

Data from 19 bioclimatic variables were considered potential predictors of the potential distribution of *S. hindustanicus* (Table 1). The data were downloaded as layers of matrix data (raster) from the WorldClim database v.2, which contains annual averages from 1970 to 2000 (Fick and Hijmans 2017), represented in a generic grid of 5 arcmin.

The variables with Pearson's correlation coefficient ( $r$ ) > 0.801 (significant at  $\alpha = 0.05$ ) were grouped according to hierarchical cluster analysis (Fig. 2). The values of the variables were tested for collinearity using the variance inflation factor (VIF), which indicates the extent of the variance of a regression coefficient estimated by one variable that is inflated due to the collinearity with the others (Naimi et al. 2014). The procedures for selecting the variables were performed using `vif.corr` and `vif.step` functions from the `usdm` package (Naimi et al. 2014) in the R environment (R Core Team 2013). Collinearity reduces efficiency and increases uncertainty in distribution models for species (De Marco and Nóbrega 2018).



**Fig. 2** Correlation between bioclimatic variables described in Table 1. Blue color with right slope indicates positive correlation, whereas red color with left slope indicates negative correlation. The intensity of Pearson’s correlation coefficient ( $r$ ) increases from the circle ( $r=0$ ) via the ellipse ( $r$ =intermediate) to the line ( $r=1$ ). Correlated variables with  $r > 0.80$  were grouped with hierarchical cluster analysis

The use of all environmental variables was scrutinized based on collinearity during the construction of the species distribution model. Variables with a VIF > 10 (Naimi et al. 2014) were removed after evaluating their probable biological significance (Fitzpatrick et al. 2013). The bioclimatic variables used in the models are highlighted in Table 1: average variation of daytime temperature (Bio02), isothermality (Bio03), temperature seasonality (Bio04), average temperature of the driest quarter (Bio09), precipitation of the driest month (Bio14), precipitation of the coldest quarter (Bio18), and warmest quarter (Bio19).

**Model and validation**

The Maxent software implements features, which are an expanded set of transformations of the original predictor covariates: product (P), linear (L), quadratic (Q), threshold (T), and hinge (H). These predictor covariates restrict the means, variances, and covariance of the variables to correspond to their experimental values (Phillips et al. 2006). Maxent maximizes the gain function, a maximum likelihood function with a penalty term, to reduce the model’s over-parameterization (Merow et al. 2013; Phillips et al. 2006). Regularization

coefficients are adjusted by multiplying with a user-defined constant, the regularization multiplier (RM), to vary the complexity of the models (Elith et al. 2011).

Maxent compares the locations where a species has been found in all study area environments. It defines these available environments by sampling many points throughout this area, referred to as background points. Because background points can include locations where the species is known to occur, background points are not the same as pseudo-absence points because they define the available environment. Using this method, 10,000 (Maxent standard) random points were generated within the species extent area of occurrence (red rectangle in Fig. 1), according to Phillips (2008), using the random points function of the *dismo* package (Hijmans and Elith 2013) in R.

Maxent default settings are based on an extensive study of empirical tuning (Phillips and Dudík 2008), a study has shown that its use may result in poor performance models (Radosavljevic and Anderson 2014; Shcheglovitova and Anderson 2013). In addition, artificial spatial autocorrelation between training and test data partitions (e.g., due to sampling bias) may increase the number of metrics used to assess model performance (Radosavljevic and Anderson 2014; Veloz 2009; Wenger and Olden 2012).

It is important to analyze combinations of parameters to select the best model (Morales et al. 2017); thus, the ideal parameters of the model were adjusted using the ENMevaluate function of the *ENMeval* package (Muscarella et al. 2014) in R, and the following combinations of features were evaluated: L, Q, H, LQ, LQH, LQHP, and LQHPT. The regularization multipliers ranged from 0.50 to 200 in increments of 0.25 (half and double of Maxent's default = 100).

The k-fold cross-validation method was used to partition the dataset into 10 groups for training and testing the model (Burman 1989; Fielding and Bell 1997; Peterson 2001). This method is helpful for modeling studies that involve transferring models in time or space, including the possibility of finding non-analogous conditions (Wenger and Olden 2012).

Groups of 5, 10, or 20 are recommended to be statistically stable, and the choice of group size is determined by the size of the study population (Kohavi 1995; Salzberg 1997). Thus, considering the complete set of features ( $s=7+7=14$  parameters), including the regularization multipliers, and the data partitioning ( $k=10$ ), according to Muscarella et al. (2014), a total of  $s(k+1)=154$  models were compared to determine the best model. The models were classified using the corrected Akaike information criterion (AICc) (Hurvich and Tsai 1993), the best model is the model with the lowest AICc value (Morales et al. 2017).

AICc is a comparison of the model's adjustment and complexity. It is a variation of the complete AIC metric, with adjustments to the formula to consider small sample sizes (Burnham and Anderson 2002). Under the statistical principle of parsimony, the AICc seeks to balance excess of simplicity and complexity in the model while considering the model's general adjustment. Models with lower AICc scores are desirable, as they indicate less complexity and a better fit.

In addition to these more traditional metrics, the *ENMeval* package results include additional metrics used to evaluate SDMs. The *avg.test.orMTP* (Minimum Training Presence Omission Rate), *avg.test.or10pct* (10% Training Omission Rate), and *var.diff.AUC*, which is the difference between the area under the curve (AUC) of the training data vs. the test data (Warren and Seifert 2011), indicates potential over-adjustment (by high values). The AUC is a model performance measure and ranges between 0 and 1, and perfect discrimination shows a value of 1 (Fielding and Bell 1997). The AUC was an efficient autonomous threshold index capable of evaluating the ability of a model to discriminate presence from absence. AUC scores could be divided into five categories (Swets 1988):  $AUC < 0.5$

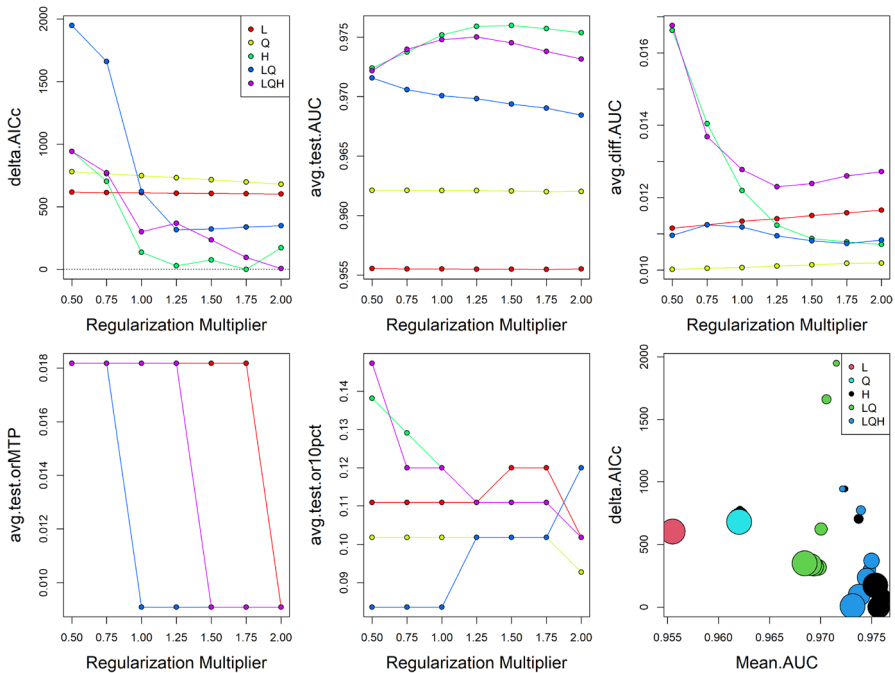
describes models that are less than chance and rarely occur; an AUC of 0.5 represents a model that is no better than random; model performance is classified as failing (0.5–0.6), bad (0.6–0.7), reasonable (0.7–0.8), good (0.8–0.9), or great (0.9–1).

Based on thresholds, the *avg.test.orMTP* and *avg.test.or10pct* metrics are measured (Radosavljevic and Anderson 2014), with the *avg.test.orMTP* indicating the proportion of sites with species presence in the test data below the lowest training sites. Likewise, *avg.test.or10pct* sets the test limit at a level of 10% of the training data.

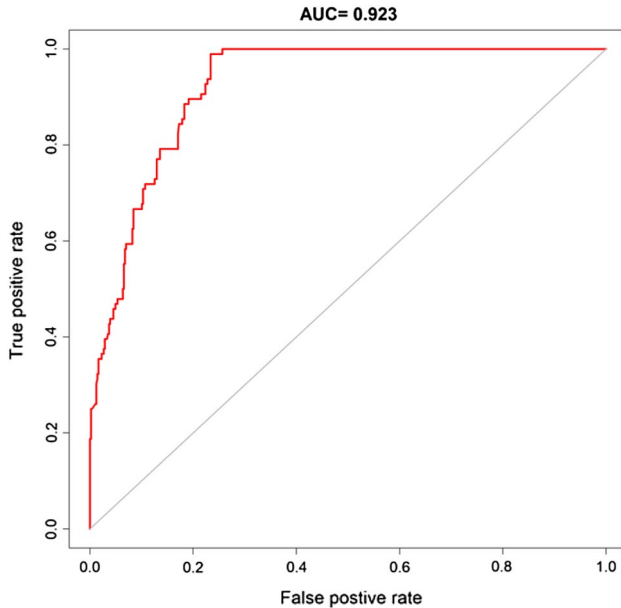
## Results

The metrics of combinations of feature classes (FC) with regularization multipliers (RMs) are shown in Fig. 3. The best combination was obtained using hinge feature (H), with a regularization multiplier of 1.75. The metrics (ENMEval) of evaluation for the test data with this model were:  $\Delta AICc=0$ ,  $AICc=2,065$ ,  $train.AUC=0.9784$ ,  $avg.test.AUC=0.9757$ ,  $avg.diff.AUC=0.01078$ ,  $avg.test.orMTP=0.009091$ ,  $avg.test.or10pct=0.1119$ , and  $TSS=0.4327$ . The ROC curve of the model is shown in Fig. 4, with an  $AUC=0.923$ .

The bioclimatic variables that most influenced the potential distribution of *S. hindustanicus* were the average variation of daytime temperature (Bio02), isothermality (Bio03), and precipitation of the coldest quarter (Bio18) (Table 1).



**Fig. 3** Metrics for selection of the feature sets and regularization multiplier value for the final model. The plots show how the five metrics used to evaluate the model vary according to changes made to the regularization multiplier. The last graph allows to evaluate the best feature combination according to the difference in  $AICc$  ( $\Delta AICc$ ) and the mean AUC



**Fig. 4** Receiver Operating Characteristic (ROC) graph for the potential geographical model for *Schizotetranychus hindustanicus* in Brazil, with area under the curve (AUC) calculation. An AUC value  $\geq 0.9$  was considered an excellent model fit

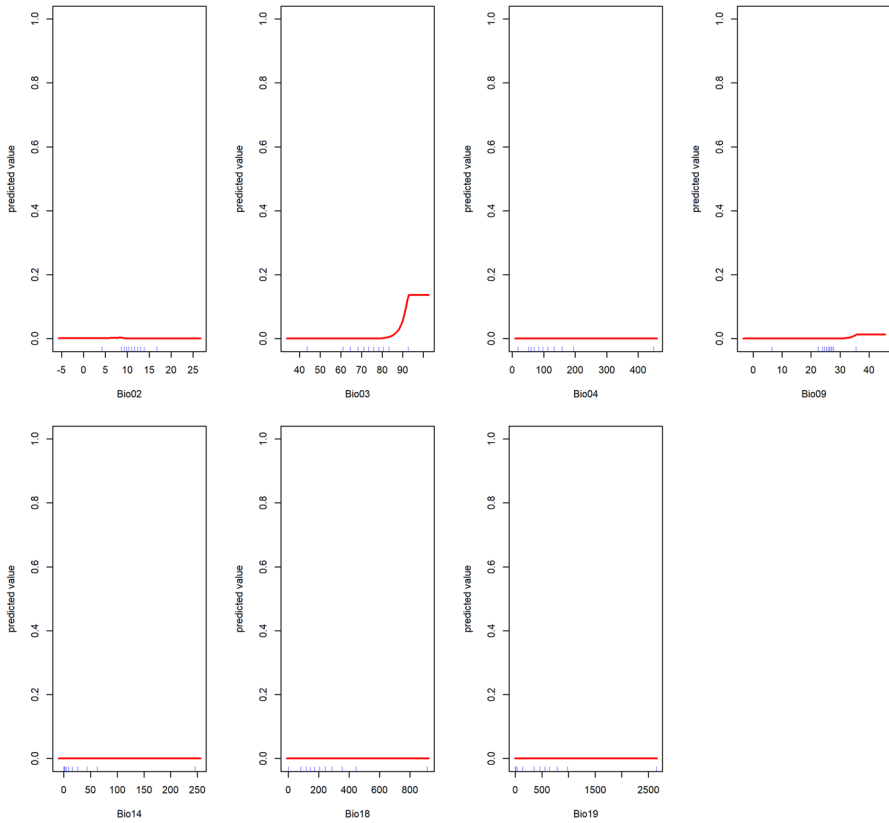
The response curves for bioclimatic variables used are shown in Fig. 5. The values on the y-axis are the predicted probabilities of suitable conditions, as given by the logistic output format, with only the predictor variables used to develop the Maxent. This result presents the response curves showing the relationship between the environmental variables and the predicted probability of occurrence. According to the response curves, the following trends were observed in the predicted suitability: suitability has small variation with oscillations of the average daytime temperature variation (Bio2) between 9 and 14 °C, it increases with isothermality (Bio3) above 80, and it presents a slight increase with an average temperature of the driest quarter (Bio09) above 36 °C.

There are large areas with suitable conditions in Brazil where *S. hindustanicus* has not been reported to occur, such as in Amazonas, Amapá, and Pará (Fig. 6).

## Discussion

Our model results demonstrated high predictive capacity. The temperature threshold for *S. hindustanicus* is 20–30 °C with an optimal developmental temperature of 27 °C (EG Fidelis, unpubl. data), which has also been shown in population studies (Nienstaedt and Marcano 2009). The most suitable regions for invasion by this mite had higher temperatures and lower variation in climate.

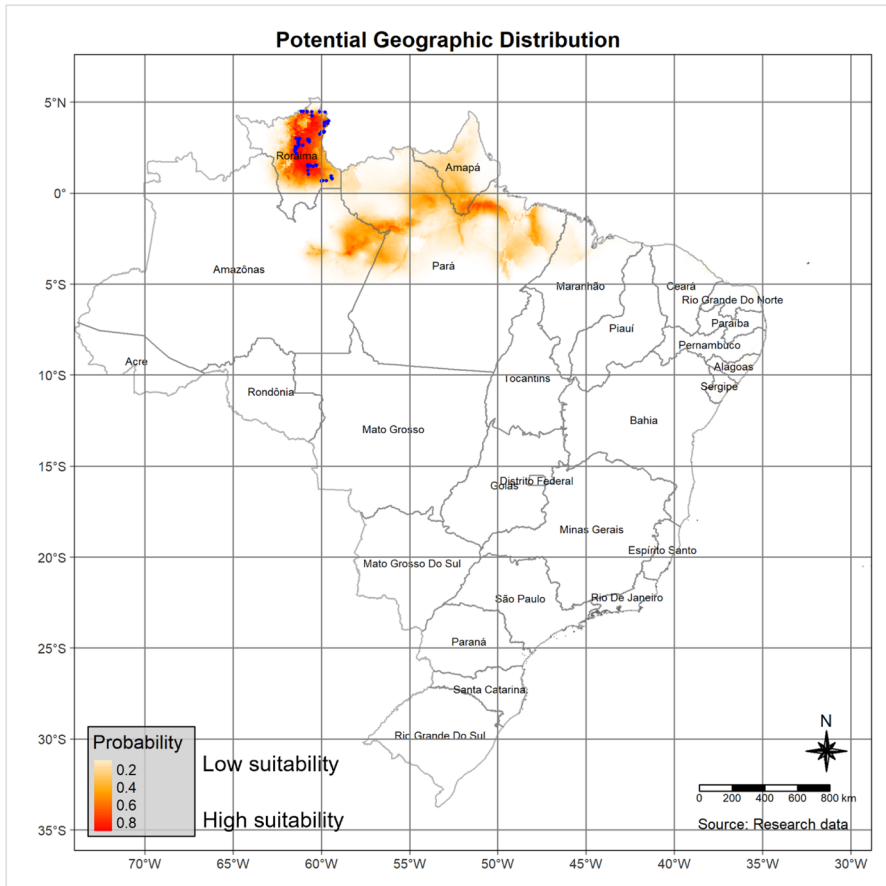




**Fig. 5** Response curves for seven bioclimatic variables (bold codes in Table 1) used as predictors in distribution model for *Schizotetranychus hindustanicus* in Brazil. For each climate variable, the predicted suitability over all values in the gradient of the variable was generated, whereas the other variables were kept constant

The model prediction shows a high probability of suitable conditions for *S. hindustanicus* in the following locations: most of Roraima, the far west of Amazonas, all of the north of Pará, especially in its northeastern region, southern, eastern, and northern parts of the State of Amapá, and a small portion to the northwest of Maranhão (all states belonging to the northern region of Brazil).

The bioclimatic variables that most influenced the potential geographical distribution of *S. hindustanicus*, were the average variation of daytime temperature (Bio02), isothermality (Bio03), and precipitation of the coldest quarter (Bio18). We observed that *S. hindustanicus* has a high potential for establishing and dispersing new regions in the north. This dispersal may make areas with a low probability of climate suitability become highly suitable over time. This is because Maxent correlates with the existing occurrence data. Thus, updating the model over time is necessary if new records and bioclimatic variables can further influence its establishment and dispersion. This model shows that areas of high suitability should be prioritized to monitor this pest in the north and avoid invasion into other states in Brazil over time.



**Fig. 6** Climatic suitability of predicted suitability for *Schizotetranychus hindustanicus* in Brazil according to the adjusted Maxent model (probability of occurrence from 0 to 1). Warmer colors (the closer to red) represent areas with better environmental conditions based on the occurrence records (blue dots)

**Acknowledgements** We thank Embrapa and the Brazilian National Council for Scientific and Technological Development (CNPq) for the financial support.

**Author contributions** GA and EGF: made substantial contributions to the conception and design of the work; the acquisition, analysis, and interpretation of data used in the work. GA, EGF, CMM, and RSS: revised the work. GA, EGF, CMM, and RSS: approved the version to be published, and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Funding** The Brazilian Agricultural Research Corporation funded this work—Embrapa, project number 131604030.00.00, developed at Embrapa Roraima.

**Data availability** The data for the present study are available upon descent request from the corresponding author.

## Declarations

**Competing interests** The authors declare no competing interests.

**Conflict of interest** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Consent for publication** Not applicable.

## References

- Aiello-Lammens ME, Boria RA, Radosavljevic A, Vilela B, Anderson RP (2015) spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* 38:541–545
- Allouche O, Tsoar A, Kadmon R (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J Appl Ecol* 43:1223–1232
- Ashraf U, Peterson AT, Chaudhry MN, Ashraf I, Saqib Z, Rashid Ahmad S, Ali H (2017) Ecological niche model comparison under different climate scenarios: a case study of *Olea* spp in Asia. *Ecosphere* 8:e01825
- Baldwin RA (2009) Use of maximum entropy modeling in wildlife research. *Entropy* 11:854–866
- Barry S, Elith J (2006) Error and uncertainty in habitat models. *J Appl Ecol* 43:413–423
- Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Villalobos F (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecol Model* 222(11):1810–1819
- Benito BM, Martínez-Ortega MM, Muñoz LM, Lorite J, Penas J (2009) Assessing extinction-risk of endangered plants using species distribution models: a case study of habitat depletion caused by the spread of greenhouses. *Biodivers Conserv* 18:2509–2520
- Burman P (1989) A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. *Biometrika* 76:503–514
- Burnham KP, Anderson DR (2002) A practical information-theoretic approach Model selection and multimodel inference, 2nd edn. Springer, New York
- CABI, Undated, (2020) CABI Compendium: Status inferred from regional distribution. CABI, Wallingford
- Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Measur* 20:37–46
- Corsi F, Duprè E, Boitani L (1999) A large-scale model of wolf distribution in Italy for conservation planning. *Conserv Biol* 13:150–159
- De Marco P, Nóbrega CC (2018) Evaluating collinearity effects on species distribution models: an approach based on virtual species simulation. *PLoS ONE* 13:e0202403
- De Andrade AFA, Velazco SJE, De Marco P (2019) Niche mismatches can impair our ability to predict potential invasions. *Biol Invasions* 21(10):3135–3150
- El-Gabbas A, Dormann CF (2018) Improved species-occurrence predictions in data-poor regions: using large-scale data and bias correction with down-weighted poisson regression and Maxent. *Ecography* 41:1161–1172
- Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Divers Distrib* 17:43–57
- Engler R, Guisan A, Rechsteiner L (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *J Appl Ecol* 41:263–274
- EPPO (2020) EPPO Global Database (available online). <https://gd.eppo.int/taxon/SCZTHI/categorization>. Accessed 20 June 2020
- Fantine AK (2011) *Schizotetranychus hindustanicus* (Hirst, 1924) (Acari: Tetranychidae): rotas de risco e potencial de impacto para a citricultura brasileira. Dissertation, Universidade Federal de Viçosa.
- Fick SE, Hijmans RJ (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int J Climatol* 37:4302–4315
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24(1):38–49. <https://doi.org/10.1017/S0376892997000088>
- Figueirêdo F, Fidelis E, Pereira R, Santos J, Negrini M, Oliveira D, Moraes G (2019) Geographical distribution of *Schizotetranychus hindustanicus* and associated mites in Roraima, Brazil. *Neotrop Entomol* 48:866–872
- Fitzpatrick MC, Gotelli NJ, Ellison AM (2013) MaxEnt versus MaxLike: empirical comparisons with ant species distributions. *Ecosphere* 4:1–15
- Hernandez PA, Graham CH, Master LL, Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29:773–785

- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25:1965–1978
- Hijmans RJ, Phillips S, Leathwick J, Elith J, Hijmans MRJ (2017) Package ‘dismo.’ *Circles* 9:1–68
- Hijmans RJ, Elith J (2013) Species distribution modelling with R. R CRAN Project. <http://www2.uaem.mx/r-mirror/web/packages/dismo/vignettes/sdm.pdf>
- Hirst S (1924) LV—On some new species of red spider. *J Nat Hist* 14:522–527
- Huerta MAO, Peterson AT (2008) Modeling ecological niches and predicting geographic distributions: a test of six presence-only methods. *Revista Mexicana De Biodiversidad* 1:205–216
- Hurvich CM, Tsai CL (1993) A corrected Akaike information criterion for vector autoregressive model selection. *J Time Ser Anal* 14:271–279
- Jaynes ET (1957) Information theory and statistical mechanics. *Phys Rev* 106:620
- Jiménez-Valverde A, Peterson AT, Soberón JOJM, Aragón P, Lobo JM (2011) Use of niche models in invasive species risk assessments. *Biol Invasions* 13(12):2785–2797
- Kohavi R (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th International Joint Conference on Artificial Intelligence, Vol. 2, Montreal, Canada, 20–25 August 1995, pp. 1137–1145.
- Leroy B, Delsol R, Hugué B, Meynard CN, Barhoumi C, Barbet-Massin M, Bellard C (2018) Without quality presence–absence data, discrimination metrics such as TSS can be misleading measures of model performance. *J Biogeogr* 45:1994–2002
- Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. *Glob Ecol Biogeogr* 17:145–151
- Merow C, Smith MJ, Silander JA Jr (2013) A practical guide to MaxEnt for modeling species’ distributions: what it does, and why inputs and settings matter. *Ecography* 36:1058–1069
- Mesa N (2010) Ácaros asociados a cítricos en Colombia. In: Primer Congreso Latinoamericano de Citricultura, 2010, p. 21.
- Morales NS, Fernández IC, Baca-González V (2017) MaxEnt’s parameter configuration and small samples: are we paying attention to recommendations? *Systematic Rev Peerj* 5:e3093
- Muscarella R, Galante PJ, Soley-Guardia M, Boria RA, Kass JM, Uriarte M, Anderson RP (2014) ENM eval: an R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods Ecol Evol* 5:1198–1205
- Naimi B, Hamm NA, Groen TA, Skidmore AK, Toxopeus AG (2014) Where is positional uncertainty a problem for species distribution modelling? *Ecography* 37:191–203
- Navia D, Marsaro AL Jr (2010) First report of the citrus hindu mite, *Schizotetranychus hindustanicus* (Hirst) (Prostigmata: Tetranychidae), in Brazil. *Neotrop Entomol* 39:140–143
- Nienstaedt B, Marcano R (2009) Fluctuación poblacional y distribución vertical del ácaro *Schizotetranychus hindustanicus* (Hirst, 1924), sobre especies de Citrus. *Entomotropica* 24:57–63
- Nix H (1986) A biogeographic analysis of the Australian elapid snakes. In: Longmore R (ed) Atlas of Elapid Snakes of Australi, Australian Flora and Fauna Series. Australian Government Publishing Service, Australia
- Papeş M, Gaubert P (2007) Modelling ecological niches from low numbers of occurrences: assessment of the conservation status of poorly known viverrids (Mammalia, Carnivora) across two continents. *Divers Distrib* 13:890–902
- Pearce JL, Boyce MS (2006) Modelling distribution and abundance with presence-only data. *J Appl Ecol* 43:405–412
- Pearson RG, Raxworthy CJ, Nakamura M, Townsend Peterson A (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *J Biogeogr* 34:102–117
- Peterson AT (2001) Predicting species’ geographic distributions based on ecological niche modeling. *The Condor* 103:599–605
- Phillips SJ (2008) Transferability, sample selection bias and background data in presence-only modelling: a response to Peterson. *Ecography* 31:272–278
- Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* 31:161–175
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Model* 190:231–259
- Poorani J (2017) *Stethorus* spp. (Coleoptera: Coccinellidae) predatory on *Schizotetranychus hindustanicus* (Hirst) (Acari: Tetranychidae) from South India, including a new species and a new synonymy in Indian *Stethorus*. *Zootaxa* 4277:591–599

- Quiróz M, Dorado I (2005) Eficiencia de tres productos comerciales en el control del ácaro hindú de las cítricas *Schizotetranychus hindustanicus* (Hirst), en el laboratorio. In: Universidad del Zulia. Facultad de Agronomía. Dpto. Fitosanitario. Resumen de Congreso de Entomología Carlos Julio Rosales, 2005.
- R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria
- Radosavljevic A, Anderson RP (2014) Making better Maxent models of species distributions: complexity, overfitting and evaluation. *J Biogeogr* 41:629–643
- Ray D, Behera MD, Jacob J (2018) Evaluating ecological niche models: a comparison between Maxent and GARP for predicting distribution of *Hevea brasiliensis* in India. *Proc Natl Acad Sci India Sec b Biol Sci* 88:1337–1343
- Salzberg SL (1997) On comparing classifiers: Pitfalls to avoid and a recommended approach. *Data Min Knowl Disc* 1:317–328
- Saupe E et al (2012) Variation in niche and distribution model performance: the need for a priori assessment of key causal factors. *Ecol Model* 237:11–22
- Shao G, Halpin PN (1995) Climatic controls of eastern North American coastal tree and shrub distributions. *J Biogeogr* 22(6):1083–1089. <https://doi.org/10.2307/2845837>
- Shcheglovitova M, Anderson RP (2013) Estimating optimal complexity for ecological niche models: a jack-knife approach for species with small sample sizes. *Ecol Model* 269:9–17
- Sheikholeslamzadeh S, Sadeghi H (2010) First records of four mite species (Acari: Tetranychidae) in Iran. *Appl Entomol Phytopathol* 78:121–125
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240:1285–1293
- Thompson I, Mackey B, McNulty S, Mosseler A (2009) Forest resilience, biodiversity, and climate change. Secretariat of the Convention on Biological Diversity. Montreal Tech Series 43:1–67
- Thuiller W, Lavorel S, Araújo MB (2005) Niche properties and geographical extent as predictors of species sensitivity to climate change. *Glob Ecol Biogeogr* 14:347–357
- Veloz SD (2009) Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *J Biogeogr* 36:2290–2299
- Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecol Appl* 21:335–342
- Wenger SJ, Olden JD (2012) Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods Ecol Evol* 3:260–267
- Wisz MS, Hijmans R, Li J, Peterson AT, Graham C, Guisan A, Group NPSDW (2008) Effects of sample size on the performance of species distribution models. *Divers Distrib* 14:763–773

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.