

# SCIENTIFIC REPORTS



OPEN

## Modeling the potential range expansion of larger grain borer, *Prostephanus truncatus* (Coleoptera: Bostrichidae)

Frank H. Arthur<sup>1</sup>, William R. Morrison III <sup>1</sup> & Amy C. Morey<sup>2</sup>

*Prostephanus truncatus* (Horn) (Coleoptera: Bostrichidae), is a beetle that is a member of a family that is primarily comprised of wood-boring insects, including forest insect pests. It is native to Mexico and Central America, where it has adapted to become a pest of stored maize. It was accidentally introduced into Africa in late 1970s, where it quickly spread throughout the sub-Saharan region, perhaps aided by adaptation to alternate hosts and the ability to persist in non-agricultural habitats. We used the correlative modelling algorithm, MaxEnt, to identify global areas of potential high suitability based on the climate locations with documented populations. Predictions using a model trained in Mexico + Central America showed potential high climatic suitability extending north into the southern United States and southward into South America, including parts of Argentina, but predictions using a model built from African occurrences did not include those areas as highly suitable. However, there was general agreement in both models that large areas of the tropics in the Western Hemisphere and in Asia have climatic conditions that could support *P. truncatus* if it were to become established. The models also showed consistency in capturing potential suitability at sites not used to build a given model. Results can be used as an initial guide to establish surveillance programs to monitor for this insect in high risk areas where it is not currently found, and to proactively mitigate the biosecurity risk from *P. truncatus*.

The larger grain borer, *Prostephanus truncatus* (Horn) (Coleoptera: Bostrichidae), is native to Mexico and Central America, where it is a pest of stored maize though it is also found in non-agricultural habitats<sup>1,2</sup>. However, it was accidentally introduced into east Africa and first documented in Tanzania in the late 1970s<sup>1,3</sup>. It spread rapidly through Central Africa to the west coast, and major outbreaks were soon reported in Ghana and Benin<sup>4,5</sup>. In addition, *P. truncatus* adapted quickly to cassava as an alternate host, and was also able to persist for short periods in non-agricultural habitats such as woodland areas, further facilitating dispersion<sup>6</sup>.

The pest status, spread, and dispersion of *P. truncatus* have been documented in multiple review articles published since the initial introduction<sup>1,6-9</sup>. It has spread throughout the major agricultural regions of sub-Saharan Africa, but most studies involving the sampling of commodities, evaluating the extent of infestations, and determining population ecology were concentrated in Eastern and Western Africa<sup>1,10-12</sup>. A possible reason for this is the presence of international research institutes in those areas, such as branches of the International Institute for Tropical Agriculture in Idaban, Nigeria and Cotonou, Benin, or scientists from the United Kingdom and Germany with extensive research experience in Africa. There is a comparative lack of information for Central Africa, which could be related to political circumstances in that region, as well as the lack of funding and scientists conducting research on stored products and stored product insects.

The invasive *P. truncatus* appears to be more destructive in Africa compared to damage reported in its native range of Mexico and Central America. There are several possible reasons for this increased capacity for damage, including the subsistence-level crib storage systems of maize cobs in Africa, adaptation to dried stored cassava chips as an alternate host, persistence in non-agricultural habitats, and the lack of native natural enemies in the

<sup>1</sup>USDA, Agricultural Research Service, Center for Grain and Animal Health Research, 1515 College Ave., Manhattan, KS, 66502, USA. <sup>2</sup>Department of Entomology, University of Minnesota, 1980 Folwell Ave., 219 Hodson Hall, St. Paul, MN, 55108, USA. Correspondence and requests for materials should be addressed to F.H.A. (email: [frank.arthur@ars.usda.gov](mailto:frank.arthur@ars.usda.gov))

invaded area, including *Teretrius nigrescens* (Lewis) (Coleoptera: Histeridae)<sup>13–17</sup>. However, since the detection of *P. truncatus* in Africa, there have been multiple studies and repeated introductions of this predator on the continent, with varying levels of permanent establishment and success in controlling and limiting *P. truncatus* outbreaks<sup>18–20</sup>.

Given the destructive nature of *P. truncatus* in Africa, knowledge relating to the potential spread of this insect northward or southward from Mexico and Central America and Africa, and eastward from Africa into tropical Asia, as well as knowledge on climatically suitable habitats for invasion would greatly benefit efforts in limiting or eliminating future infestations by targeting detection efforts to areas where the pest is likely to establish. There have been sporadic interceptions of this insect in the United States and Europe, but no true establishment to date. The use of correlative models to estimate species distributions are increasingly viable, in part due to the rise in availability of high resolution climatic and environmental datasets and advances in modelling methods<sup>21,22</sup>. These models identify statistical relationships between locations of species occurrence and a set of environmental predictors, which are then used to identify other areas with similar environmental conditions<sup>23</sup>. The simplicity of the data inputs needed makes correlative models particularly appealing to estimate suitable habitats for invasive species, where often little more information other than records of occurrence are known. However, the extrapolation of correlative model predictions to new areas of space and time – an interest of most applications to invasive species – requires caution<sup>23,24</sup>.

Of the currently available methods, MaxEnt<sup>25</sup> has shown relatively reliable performance<sup>26</sup>. MaxEnt is a presence-background (or pseudo-absence) model that compares the environmental space at a list of occurrence locations to those at background locations that fall within a user-defined area<sup>27</sup>. The algorithm minimizes the relative entropy between a probability density of the environmental variables at the presence data and another from the background locations, which provides estimates of relative suitability based on the given variables and suitability in geographical space<sup>27,28</sup>. The purpose of the current study is to predict areas for potential range expansion of *P. truncatus* from its current documented distribution. To achieve this, we (1) developed a MaxEnt model based on the documented presence of *P. truncatus* in its native range (Mexico + Central America), (2) developed a second model based on the presence of *P. truncatus* in its introduced range (Africa), and (3) compared the similarity of the two model projections and identified areas of agreement between both models, which may suggest higher risk areas for invasion by *P. truncatus* in natural areas (e.g., areas outside of grain storage facilities).

## Materials and Methods

**Occurrence records and background.** A necessary input for correlative species distribution models, including MaxEnt, are geographic coordinates of sites where a species has been documented to occur. An exhaustive literature search was done through various databases, including Agricola, Web of Science, and Google Scholar, to find published, peer-reviewed papers that gave specific names of states, provinces, or regions within a country, including towns or villages, where *P. truncatus* infestations were documented. In some cases, only a general location such as a village or town was listed, others gave a specific sampling area. Geographic coordinates were then obtained through a Google search by listing the location and obtaining information that gave the longitude and latitude. Locations with sporadic sightings or incidental infestations or interceptions were excluded from the list. For many of the references, the sampling sites were one-time locations, with no information on yearly persistence or regularity of infestation but were included in the list because they were in a geographic region where other studies had detected *P. truncatus*, or where *P. truncatus* could potentially survive, based on records of documented establishment in the surrounding geographic region. The literature review yielded multiple, unique geographic coordinates with documented infestations in grain storages or in the surrounding landscape, field crop sites, and forest and woodland sites (Table S1).

Limiting the extent from which MaxEnt draws background locations is recommended to reflect areas accessible to a species or the sampling bias present in occurrence records<sup>28,29</sup>. Given that we had no information on the specific nature of bias in our occurrence sites, we limited background selection to a minimum convex polygon (MCP) with a one-pixel (2.5 arc-minute) buffer drawn around the occurrence points for a given model<sup>29,30</sup>. The package ‘rgeos’<sup>31</sup> was used to generate the buffered MCPs in R version 3.5.0<sup>32</sup> and RStudio version 1.1.453<sup>33</sup>. A random sample of 10,000 locations (i.e., the default for MaxEnt) was taken from these reduced areas to serve as background locations for a given model.

**Climate data.** Characterizing current and predicted distributions of a species based on climate suitability is widely used<sup>34</sup> especially with the increasing availability of global climate data. We used 19 gridded temperature and precipitation variables from the WorldClim dataset (v. 2; [www.worldclim.org/bioclimate](http://www.worldclim.org/bioclimate))<sup>35</sup> at 2.5 arc-minute resolution. This dataset contains average monthly global climate data for 1970–2000. The climate rasters were clipped to the extent of a given buffered MCP using the R package ‘raster’ (v. 2.6–7)<sup>36</sup>. To reduce the confounding effects of potential collinearity among climate variables<sup>37</sup>, we then calculated a Pearson correlation coefficient matrix within each MCP space. Only variables with correlations of  $|r| < 0.70$  were retained, leaving the final models with 5–6 climate variables (Table 1).

**Direct climate comparisons.** For each of the documented occurrences of *P. truncatus*, the climate variable values were extracted and directly compared between populations in Mexico + Central America and Africa to understand how conditions differed between the native and introduced range. The mean value for each population was compared with a *t*-test. Climatic variables compared were those included in one or both final models. These variables included BIO3 (isothermality, percent), BIO5 (maximum temperature of the warmest month, °C), BIO6 (minimum temperature of the coldest month, °C), BIO12 (annual precipitation, mm), BIO15 (precipitation seasonality, percent), BIO18 (precipitation of warmest quarter, mm), and BIO19 (precipitation of coldest quarter, mm). BIO3, isothermality, is derived by calculating the ratio of the mean diurnal range (BIO2) to the

Model	<i>n</i>	Environmental variables <sup>a</sup>	Regularization multiplier ( $\beta$ )	Features	Data partitioning method <sup>b</sup>
Mexico + Central America (native range)	32	BIO <sub>3</sub> , BIO <sub>5</sub> , BIO <sub>6</sub> , BIO <sub>12</sub> , BIO <sub>15</sub>	1.5	hinge	$n - 1$ jackknife
Africa (invaded range)	69	BIO <sub>3</sub> , BIO <sub>5</sub> , BIO <sub>6</sub> , BIO <sub>12</sub> , BIO <sub>18</sub> , BIO <sub>19</sub>	1.0	linear, product, quadratic	block

**Table 1.** Environmental variables and parameters used in MaxEnt models of *Prostephanus truncatus*.

Environmental variables were from the WorldClim dataset ([www.worldclim.org/bioclim](http://www.worldclim.org/bioclim)) and limited to those with a  $|r| < 0.7$  within the area of model development. Regularization and feature combinations were selected for each model based on the lowest AICc value. Resampling method was selected based on sample size and study objective (spatial transferability). Model names refer to the dataset (with  $n$  occurrences) used to develop the model. <sup>a</sup>BIO<sub>3</sub> = Isothermality; BIO<sub>5</sub> = Max Temperature of Warmest Month; BIO<sub>6</sub> = Min Temperature of Coldest Month; BIO<sub>12</sub> = Annual Precipitation; BIO<sub>15</sub> = Precipitation Seasonality; BIO<sub>18</sub> = Precipitation of Warmest Quarter; BIO<sub>19</sub> = Precipitation of Coldest Quarter. <sup>b</sup>partitioning was done using the R package ‘ENMeval’<sup>43</sup>.

annual temperature range (BIO7) and then multiplying by 100. Isothermality quantifies how large the day-night temperatures oscillate relative to the summer-to-winter (annual) oscillations. Isothermality is generally useful for tropical, insular, and maritime environments<sup>38</sup>. Precipitation seasonality, BIO15, is a measure of the variation in monthly precipitation totals over the course of the year. This index is the ratio of the standard deviation of the monthly total precipitation to the mean monthly total precipitation (also known as the coefficient of variation) and is expressed as a percentage<sup>38</sup>. For all tests, significance was determined at  $P < 0.05$ .

**Model development and evaluation.** Using MaxEnt version 3.3.3k<sup>25</sup>, we built two separate models for *P. truncatus*: one based on occurrences and background from the native range (Mexico + Central America) and one based on occurrences and background from the invaded range in Africa. MaxEnt models made with default settings can result in overfit predictions, so species-specific tuning of complexity is recommended<sup>39–41</sup>. To address this, the regularization multiplier and features classes were selected for each model based on the lowest Akaike Information Criterion corrected for small sample sizes (AICc<sup>42</sup>) using the ‘ENMeval’ R package<sup>43</sup>. ‘ENMeval’ also allows additional methods to partition training and testing data, which we selected based on the recommendations summarized in Muscarella *et al.*<sup>43</sup>. For the model built using native occurrence records, ‘ $n - 1$  jackknife’ partitioning was used due to the relatively small number of occurrences ( $n = 32$ ). This method chooses one occurrence point to test a model trained with the remaining ( $n - 1$ ) occurrences, and it was repeated  $n$  times until each record has been used for testing. For the model built using invaded occurrence records from Africa, the ‘block’ method was used. This method splits approximately equal numbers of occurrence and background points into quadrants based on latitude and longitude.

We quantitatively assessed model performance using the conventional metrics of area under the receiver-operator curve (AUC), which measures the discriminatory ability of each model, and omission rates, which can indicate overfitting. AUC is a threshold-independent performance measure that reflects the probability that a randomly chosen presence site will rank above a randomly chosen background site<sup>25</sup>. Values near 1.0 indicate high discriminatory ability, whereas values of 0.5 (or less) indicate discrimination no better than random<sup>42</sup>. Omission rates (OR) calculate the proportion of test locations with suitability values lower than a specified threshold. The ‘minimum training presence’ threshold omission rate (OR<sub>MTP</sub>) uses the smallest value predicted for any training location, indicating the least-suitable environment where a training point was found. Similarly, the 10% training threshold omission rate (OR<sub>10</sub>) uses the smallest value after excluding the lowest 10% of training suitability values<sup>39,44</sup>. In ideal models, the expected OR<sub>MTP</sub> and OR<sub>10</sub> are zero and 10%, respectively, and values higher than expected suggest overfitting. The OR<sub>10</sub> is less likely to be influenced by extreme/outlier occurrence locations<sup>44</sup>.

**Model projections.** Logistic MaxEnt outputs were visualized for each model in ArcMap (ESRI® ArcGIS Desktop 10.6) to show global and region-specific areas of relative suitability between 0–1.0. Models projected into areas with conditions outside of those used to build the model can be unreliable<sup>45</sup>. To identify areas of model extrapolation, multivariate environmental similarity surface (MESS) maps were produced for each MaxEnt model using the package ‘dismo’<sup>46</sup>. The MESS calculates how similar a given point is to a reference set of points for a given climate variable<sup>28</sup>. Values less than zero indicate grids where at least one variable was extrapolated; therefore, we limited each projection space to areas with MESS values of zero or greater to reduce uncertainty (Fig. S1).

To quantify how similar the final spatial projections were, we used ENMTools (v. 1.4.4) to calculate two measures of niche similarity: Schoener’s  $D$  and  $I$ , a modified Hellinger distance<sup>47,48</sup>. Both measures compare the suitability values predicted at each grid cell between two models, though  $D$  assumes the suitability scores are proportional to abundance and  $I$  does not<sup>48</sup>. Both measures are bound from 0 (no overlap) to 1 (identical), with 0.4–0.6 considered moderate overlap<sup>49</sup>. We also visually identified areas that both models forecasted as suitable, with suitability being defined for each model by the 10<sup>th</sup> percentile training presence logistic threshold (MTP<sub>10</sub>)<sup>49</sup>. Each model was converted to a binary prediction based on their respective threshold value and then regions that both models forecasted as suitable were isolated using the Intersect (Analysis) tool in ArcMap. The resulting consensus map highlighted areas of higher (suitable in both models) and lower (suitable in only one model) certainty in global suitability predictions. We also performed a Chi-Sq test to compare the suitability classes within each model compared to the null hypothesis that each class would be equally represented.

Climate Variable <sup>a</sup>	Mexico and Central America	Africa	<i>t</i>	<i>P</i>
	Mean ± SE	Mean ± SE		
BIO3	203.9 ± 4.8	220.4 ± 4.6	1.97	0.05
BIO5	229.2 ± 11.4	241.9 ± 11.1	0.64	0.52
BIO6	12.0 ± 1.8	9.8 ± 1.3	0.83	0.41
<b>BIO12</b>	<b>129.5 ± 3.2</b>	<b>118.1 ± 2.3</b>	<b>2.44</b>	<b>0.02*</b>
BIO15	332.6 ± 4.5	—	—	—
BIO18	—	245.2 ± 3.2	—	—
BIO19	—	232.4 ± 5.4	—	—

**Table 2.** Summary of climatic variables in the native (Mexico and Central America) and introduced (Africa) range of *P. truncatus*. Means (±SE) are presented for variables provided to one or both of the models. An asterisk indicates a significant ( $P < 0.05$ ) difference between the regions for a given variable if it was used in both final models. <sup>a</sup>For a definition of each variable and associated units, please see the materials and methods.

Model	AUC <sub>TEST</sub> <sup>a</sup>	AUC <sub>DIFF</sub> <sup>b</sup>	OR <sub>MTP</sub> <sup>c</sup>	OR <sub>10</sub> <sup>d</sup>	MTP <sub>10</sub> <sup>e</sup>	Schoener's <i>D</i>	<i>I</i>
Mexico + Central America (native range)	0.751	0.116	0.031	0.219	0.342	0.606	0.883
Africa (invaded range)	0.593	0.179	0.059	0.294	0.216		

**Table 3.** Evaluation and similarity metrics for MaxEnt models of *Prostephanus truncatus*. Metrics were generated in R ('ENMeval'<sup>43</sup> and ENMTools<sup>48</sup>). Similarity indices (Schoener's *D* and *I*) compared the projection areas remaining after MESS exclusion (see methods) between the two models. <sup>a</sup>AUC<sub>TEST</sub> measures the ability of the model to distinguish the occurrence points used in model testing from background points with 1.0 being perfect discrimination. <sup>b</sup>AUC<sub>DIFF</sub> is the difference between the AUC using test and training occurrences; high differences indicate model overfitting. <sup>c</sup>OR<sub>MTP</sub> is the omission rate for the proportion of test locations with suitability values lower than the smallest value predicted for any training location (minimum training presence; MTP); overfitting is indicated by deviations from the expectation of zero. <sup>d</sup>OR<sub>10</sub> is the omission rate for the proportion of test locations with suitability values lower than the smallest value after excluding the lowest 10% of training suitability values (10% training omission rate); overfitting is indicated by deviation from the expectation 0.10. <sup>e</sup>MTP<sub>10</sub> is the minimum predicted logistic value for the training sites after excluding the lowest 10% of training site values.

## Results

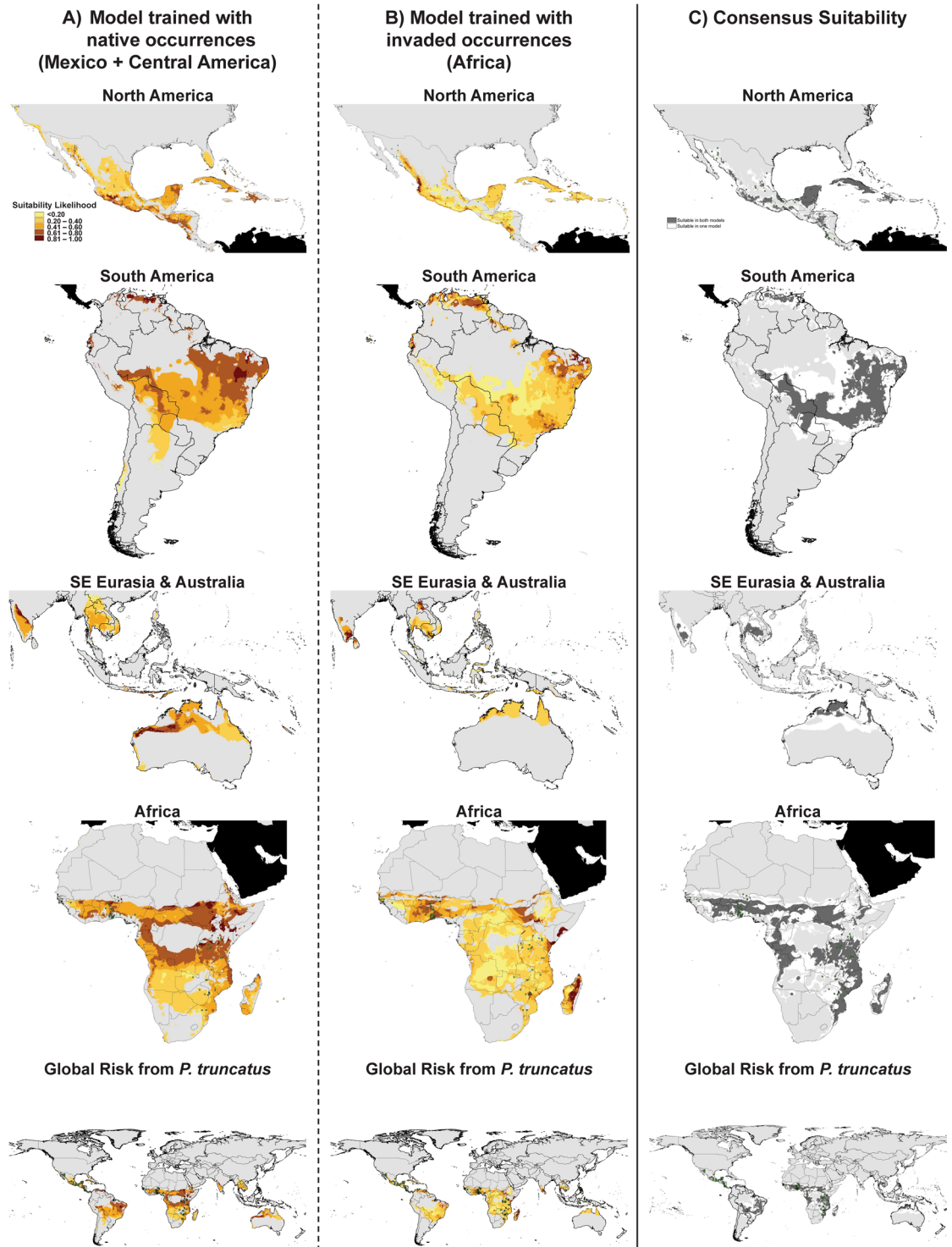
**Direct climate comparisons.** The climate conditions for recorded occurrences of *P. truncatus* in Mexico + Central America, and Africa were similar for variables BIO3, BIO5, BIO6, and BIO12 (Table 2). BIO15 did not ultimately contribute to the final Mexico + Central America model (Table S2). In the final Africa model, BIO19 and BIO5 showed no permutation importance, but did show some percent contribution (Table S2).

**Assessment of Mexico + Central America model and global predictions.** The model trained in Mexico + Central America showed a high degree of discriminatory ability based on the AUC, and closely met the expected value for the omission rate (AUC and OR<sub>MTP</sub>, see Table 3). The AUC<sub>diff</sub> and OR<sub>10</sub> values for the model suggests that there is some degree of overfitting for the model (Table 3). The Mexico + Central America-trained model ultimately included four uncorrelated variables from the WorldClim dataset, namely BIO3, 5, 6, and 12.

As expected, the Mexico + Central America-trained model showed that all the recorded occurrences of *P. truncatus* in its native range were highly suitable for potential establishment and population development (Fig. 1A). In addition, there are areas in Mid- and Southern California, most of Florida, and large sections of the Caribbean that appear climatically highly suitable for colonization by *P. truncatus* according to this model.

In South America, large areas of Brazil and the northern regions of Argentina, Venezuela, much of Bolivia, and western Paraguay appear highly suitable (>0.80) for this species (Fig. 1A). In Southeast Asia, the most probable areas of suitability are western and southern India, large sections of Myanmar, Thailand, Cambodia, Vietnam, Indonesia, and in sporadic locations in the Philippines. Finally, large sections of northern Australia in a belt stretching from the eastern to western coasts of the continent also appear highly suitable for establishment of *P. truncatus*. Much of equatorial Africa from 12°N to -20°S, except for a large area of central Democratic Republic of Congo is considered a climatically highly-suitable habitat for *P. truncatus* according to the Mexico + Central America-trained model. The majority of the documented occurrences in Africa (55%) were predicted by the Mexico + Central America-trained model.

**Assessment of the African model and global predictions.** The second model was trained on the recorded occurrences of *P. truncatus* from the invaded range in Africa. This model showed a moderate degree of discriminatory ability based on the AUC (Table 3). Notably, there were several recorded occurrences in the native range that the African-trained model did not predict as potential suitable areas, but 68% of the occurrences were still predicted by the Africa-trained model (Table 5). Further, this model had a higher AUC<sub>diff</sub> value and omission



**Figure 1.** Forecasted geographic suitability of *Prostephanus truncatus* based on temperature and moisture. The first two columns show the predicted suitability projected in various regions from two different MaxEnt models: (A) a model trained using occurrence and background locations from the native range of *P. truncatus* in Mexico + Central America and, (B) a model trained using occurrence and background locations from the invaded range of *P. truncatus* in Africa. The third column (far right) shows regions where the two models overlap (dark grey) in projected suitability based on their respective 10% minimum training thresholds (Table 3). Light grey areas in all maps indicate areas removed from model projection based on MESS (see Materials and Methods and Fig. S1). Green dots are documented records of occurrence (see Table S1).

Model training region	Suitability Class	Area (km <sup>2</sup> )	Proportion of total projection area <sup>a</sup>
Mexico + Central America	<21	376,379	0.013*
	21–40	8,426,985	0.281
	41–60	12,285,599	0.410*
	61–80	7,803,100	0.260
	81–100	1,066,489	0.035
	Total	29,958,552	
Africa	<21	4,505,269	0.170
	21–40	15,324,383	0.577*
	41–60	4,389,318	0.165
	61–80	1,844,947	0.069*
	81–100	513,238	0.019*
	Total	26,577,154	

**Table 4.** Areas of suitability for *P. truncatus* as predicted by the Mexico + Central America and African models. Suitability class, area, and proportion of total global projection area are listed for each model. Expected proportion of total projection area was assumed to be 0.2 for each class within each model under the null hypothesis. Significant deviations from the null hypothesis are denoted by an asterisk ( $\chi^2$ -test, Bonferroni correction). For both models, there were no areas predicted as 0 suitability within our areas of projection. <sup>a</sup>Total area where overlap between the two models forecasted suitability >0.0 was 18,509,504 km<sup>2</sup>. The proportion of the Africa model in agreement with the Central America model was 0.70. The proportion of the Central America model in agreement with the Africa model was 0.62.

Model	Occurrence record dataset													
	Mexico/Central America (n=32)							Africa (n=69)						
	Within modeled suitability space*	Outside modeled suitability space	<0.21	0.21–0.40	0.41–0.60	0.61–0.80	0.80–1.00	Within modeled suitability space	Outside modeled suitability space	<0.21	0.21–0.40	0.41–0.60	0.61–0.80	0.80–1.00
Africa	22	10	6	14	1	1	0	69	0	7	28	13	14	7
Mexico + C. America	32	0	0	7	9	15	1	38	31	1	13	9	14	1
Consensus	16	16	n/a	n/a	n/a	n/a	n/a	24	45	n/a	n/a	n/a	n/a	n/a

**Table 5.** Counts of occurrence records for *Prostephanus truncatus* (see Table S1) in relation to various categories of predicted suitability in each model. See Material and Methods for full description of each model. \*The suitability space for the binary Consensus model was considered here as those areas that both models predicted suitable, based on their respective 10 percent minimum training presence logistic threshold (i.e., dark grey areas in Fig. 1c). The suitability space for the remaining two models were those areas with a continuous suitability value >0 (i.e., colored areas in Fig. 1a–b).

rate than the Mexico + Central America-trained model, with a  $OR_{MTP}$  roughly near zero, but an  $OR_{10}$  rate near 30% (Table 3), indicating some degree of overfitting. The Africa-trained model included six variables instead of four, namely BIO3, 5, 6, 12, 18, and 19 (Table 1).

Overall, fewer areas within our projection space were predicted as highly suitable in the African-trained model (>0.80 suitability; Table 4, Fig. 1B). In addition, overall there was also more land area considered to have low suitability (<0.21 suitability) for *P. truncatus* in the African model compared to the Mexican + Central American model.

With regards to North America, most areas in the United States, were outside the range of our data, but Mexico the Caribbean Cuba, most of Haiti and the Dominican Republic, and Jamaica and Puerto Rico were considered potential areas of establishment (Fig. 1B). Similar to the Mexico + Central American-trained model, large areas of Brazil, Bolivia, Paraguay, and northern Venezuela were considered climatically highly-suitable for establishment of *P. truncatus* by the African-trained model, though not northern Argentina. In southeastern Eurasia and Oceania, Thailand, Cambodia, Vietnam, southern India, and northern Laos were considered highly suitable by the model, with sporadic high climatic suitability in the Philippines and Indonesia. Similar to the Mexican + Central American-trained model, the African-trained model considered regions of northern Australia climatically suitable for *P. truncatus*. Much of equatorial Africa was also considered to be climatically highly-suitable by the African-trained model, and included all the recorded occurrences of the species, but also similarly predicted less climatic suitability in the central Democratic Republic of Congo.

**Assessment of dual model consensus and global predictions.** There was moderate to strong overlap in areas considered climatically highly-suitable by the models depending on whether Schoener's *D* or *I* is considered. Both models assessed much of Western and Southern Mexico, as well as much of Central America, and the larger Caribbean islands as potential habitats for establishment and dispersal of *P. truncatus*, though some of the

recorded occurrences fell outside the consensus zones in the native range (Table 5). In South America, consensus zones of climatic suitability included large areas of Brazil, Bolivia, western Paraguay, and coastal Venezuela. Both models overall predicted that much of tropical Africa was highly suitable for *P. truncatus*. In Eurasia and Oceania, both models agreed that most of Cambodia, northern Laos, large areas of Thailand, southern India, sporadic locations in Indonesia, and northern Australia were moderately suitable (0.40 to 0.80) areas for *P. truncatus*. Both models overall predicted that much of tropical Africa was highly suitable for *P. truncatus*.

## Discussion

Our MaxEnt models utilized four to six different climatic variables to predict areas where *P. truncatus* could potentially establish outside of storage facilities, based on either the current distribution in Mexico + Central America or in Africa. The two models ultimately were based on similar climate variables, with the Africa model using two additional variables (i.e., BIO18 and BIO19). Of the shared variables, only the amount of annual precipitation (i.e., BIO12) differed significantly between the two modelled regions. However, there were some other differences between the two models, in that the Mexican + Central American model identified sites in the southern US that were potential areas of high suitability (i.e., >0.8), while the African model did not. In the southern US, maize is typically harvested in late July and early August, and population prediction models coupled with historical weather data show that the warm temperatures will support rapid population growth of other stored product species, such as the maize weevil, *Sitophilus zeamais* (Motschulsky), even with the use of aeration to cool the grain mass and modify storage bin temperatures<sup>50</sup>.

The African model was perhaps more conservative, but both models identified areas in the Caribbean and areas in tropical and sub-tropical South America as areas of high climatic suitability. In general, though there were some zones of disagreement, both models showed significant overlap with each other in areas designated as climatically suitable (Fig. 1C). Importantly, for both models, we conservatively chose to limit the areas of prediction to those regions that did not require extrapolation beyond the range of our data (i.e., light grey areas in Fig. 1). This decision was intended to reduce uncertainty associated with extrapolation<sup>23</sup> and does not indicate that these areas are not at risk. There are three possible explanations for what may be happening in the grey areas of our projections: (1) they could truly be outside the climatic tolerances for this species, (2) they could be outside the current realized niche of the species, and/or (3) our occurrence datasets may be incomplete in representing the current niche of the species and additional records may change the extent of these areas. It is possible that *P. truncatus* may be able to survive inside a buffered anthropogenic structure such as a grain bin or food facility where temperature effects are mitigated.

Network mapping showing routes of trade and their relative frequency with ports in the native and invaded ranges of *P. truncatus* may help illuminate relative risk in these grey areas. For example, Paini and Yemshanov<sup>51</sup> used the international marine shipping network centered around Australia to predict the relative risk of entry by the quarantine threat posed by the, khapra beetle, *Trogoderma granarium* Everts (Coleoptera: Dermestidae), and found ships arriving from Taiwan and the Republic of Korea were the most likely sources for entry of the species into the country. Moreover, it is also possible that global climate change may shift the range of suitable areas northward and southward towards the poles for *P. truncatus*, as has been reported for other insect species<sup>52,53</sup>. This may have repercussions for the stored product species community ecology in newly invaded areas, with *P. truncatus* likely dominating on maize, and causing correspondingly more damage<sup>54</sup>.

Both models consistently identified tropical and sub-tropical areas as climatically highly-suitable for establishment of *P. truncatus*. This is important because cassava is a primary crop in the tropics and subtropics, and many of the publications documenting the rapid spread and dispersal of *P. truncatus* in Africa cite the species' quick adaptation to cassava as an alternate agricultural host as a primary contributing factor<sup>15,55</sup>. In Africa, infestations often occur when cassava is stored as dried chips. Cassava is native to South America and is now grown throughout Brazil, in tropical Africa and Asia, in India and China, and parts of tropical Australia<sup>56</sup>. It is used as a staple food because of its high starch content in the roots, as well as in industrial processes<sup>57</sup>. Production for human food is increasing and some researchers have stated that cassava now ranks behind wheat, corn, and rice in global production and utilization<sup>58</sup>. Although potential spread of *P. truncatus* might be limited because cassava is cultivated in large plantations in South America and tropical Asia, and may not be stored as dried chips, it is still reasonable to assume that if *P. truncatus* was introduced into areas where cassava is grown and where *P. truncatus* has not been established, some adaptation and dispersal could still occur.

Many studies in Africa also document survival and reproduction of *P. truncatus* on woody host material, including live and dead branches, branches girdled by attacking wood-boring insects, or on seeds<sup>6,59,60</sup>. Though listings for specific species can be inconsistent and contradictory, there are obviously some woody species that will support development and reproduction in addition to incidental or low-level attack and survival. Nansen *et al.*<sup>59</sup> reported reproduction on teak (*Tectona grandis* L.) seeds. The natural range of teak is the Indian peninsula and Southeast Asia, but it is also planted in managed plantations throughout the tropics<sup>61</sup>. While there is no evidence that *P. truncatus* would attack mature trees, it seems that it could indeed survive and reproduce on dead branches and seeds and serve as an alternate host as well. Since *P. truncatus* is a member of the family Bostrichidae, and most members of this family are wood-boring insects, it is not surprising that it can potentially survive and reproduce on woody material. Both the Mexican + Central American and African models predict that *P. truncatus* could survive in tropical areas where teak and cassava are both present. It is also possible that other woody species could serve as alternate hosts as well.

The MaxEnt modeling approach performs particularly well with small samples size<sup>59</sup> and has been utilized for a variety of invasive species, including the brown marmorated stink bug (*Halyomorpha halys* [Stål])<sup>62</sup>, the emerald ash borer (*Agrilus planipennis* Fairmaire)<sup>63</sup>, and spotted wing drosophila (*Drosophila suzukii*) Matsumura<sup>64</sup>. There were some discrepancies between our two models but predicted climatic suitability can differ between models based on native versus invaded populations for multiple reasons, such as shifts in climatic niche with invasion<sup>65,66</sup>, other

biotic sieves such as interspecific competition that limit the realized niche<sup>67</sup> or an incomplete description of climatic suitability in the invaded range due to lack of equilibrium with the environment<sup>23</sup>. We compared the two models using two indices of similarity, Schoener's  $D$  and  $I$ , a modified Hellinger distance<sup>45,47</sup>. Others have found that  $I$  tends to estimate greater similarity than Schoener's  $D$ <sup>48,68</sup>, and that  $D$  may be better suited for computing niche overlap<sup>48</sup>.

In conclusion, our modeling results provide guidelines for predicting potential geographic regions where *P. truncatus* could survive outside of storage facilities or in unbuffered environments, and consequently, the increased risk of invasion into storage facilities due to populations sustained outside these facilities. Our models suggest grains stored in tropical Asia may be especially at risk if *P. truncatus* were to become established in the surrounding environment. This modeling approach could be used to guide directed studies on temperature tolerance for management in storage facilities. These results emphasize the need for monitoring of this pest in tropical areas where it is not currently present or established, and for governments and research institutions to be proactive in their surveillance and detection efforts<sup>69</sup>.

## Data Availability

All data on which the model in this paper was based are available as Supplemental Files.

## References

- Hodges, R. J., Dunstan, W. R., Magazini, I. & Golob, P. An outbreak of *Prostephanus truncatus* (Horn) (Coleoptera: Bostrichidae) in east Africa (Tabora, Tanzania). *Prot. Ecol.* **5**, 183–194 (1983).
- Tigar, B. J., Osborne, P. E., Key, G. E., Flores, S. E. & Vazque, A. Insect pests associated with rural maize stores in Mexico with particular reference to *Prostephanus truncatus* (Coleoptera: Bostrichidae). *J. Stored Prod. Res.* **30**, 267–281 (1994).
- Golob, P. & Hodges, R. J. Study of an outbreak of *Prostephanus truncatus* (Horn) in Tanzania. Report G. 164. *Tropical Products Institute*, London (1982).
- Addo, S., Birkinsaw, L. A. & Hodges, R. J. Ten years after the arrival in Ghana of Larger Grain Borer: Farmer's responses and adoption of IPM strategies. *Int. J. Pest Manage.* **48**, 315–325 (2002).
- Meikle, W. G., Holst, N., Cholz, D. & Markham, R. H. Simulation model of *Prostephanus truncatus* (Coleoptera: Bostrichidae) in rural maize stores in the Republic of Benin. *Environ. Entomol.* **27**, 59–69 (1998).
- Nansen, C., Meikle, W. G. & Korie, S. Spatial analysis of *Prostephanus truncatus* (Horn) (Bostrichidae: Coleoptera) flight activity near maize stores and in different forest types in southern Benin, West Africa. *Ann. Entomol. Soc. Am.* **95**, 66–74 (2002).
- Golob, P. Current status of the larger grain borer *Prostephanus truncatus* (Horn) in Africa. *Int. J. Trop. Insect Sci.* **9**, 737–745 (1988).
- Nansen, C. & Meikle, W. G. The biology of the larger grain borer, *Prostephanus truncatus* (Horn) (Coleoptera: Bostrichidae). *Integr. Pest Manag. Rev.* **7**, 91–104 (2002).
- Muatinte, B. L., Van Den Berg, J. & Santos, L. A. *Prostephanus truncatus* in Africa: a review of biological trends and perspectives on future pest management strategies. *Afr. Crop Sci. J.* **22**, 237–256 (2014).
- Dendy, J., Dobie, P., Saidi, J. A., Smith, J. L. & Uronu, B. Trapping the larger grain borer *Prostephanus truncatus* in maize fields using synthetic pheromones. *Entomol. Exp. App.* **50**, 241–244 (1989).
- Gueye, M. T., Goergen, G., Badiane, D., Hell, K. & Lamboni, L. First report on occurrence of the larger grain borer *Prostephanus truncatus* (Horn) (Coleoptera: Bostrichidae) in Senegal. *African Entomol.* **16**, 309–311 (2008).
- Nang'ayo, F. L. O. *et al.* The natural environment as a reservoir for larger grain borer *Prostephanus truncatus* (Horn) Coleoptera: Bostrichidae in Kenya. *Afr. Crop Sci. J.* **1**, 39–47 (1993).
- Rees, D. P., Rodriguez, R. & Herrera, R. Observations on the ecology of *Teretriosoma nigrescens* Lewis (Col.: Histeridae) and its prey *Prostephanus truncatus* (Horn) (Col.: Bostrichidae) in the Yucatan. *Trop. Sci.* **30**, (153–165 (1990).
- Boxall, R. A. Damage and loss caused by the larger grain borer *Prostephanus truncatus*. *Integrated Pest. Manag. Rev.* **7**, 105–121 (2002).
- Ospitan, A. A., Sangowusi, V., Lawal, O. I. & Popoola, K. O. Correlation of chemical compositions of cassava varieties to their resistance to *Prostephanus truncatus* Horn (Coleoptera: Bostrichidae). *J. Ins. Sci.* **15**, 1–7 (2015).
- Hodges, R. J. The biology and control of *Prostephanus truncatus* (Horn) (Coleoptera: Bostrichidae) a destructive storage pest with increasing range. *J. Stored Prod. Res.* **22**, 1–14 (1986).
- Hodges, R. J. Recent advances in the biology and control of *Prostephanus truncatus* (Coleoptera: Bostrichidae). Proceedings of the 6<sup>th</sup> International Working Conference on Stored Product Production. Highley, E., Wright, E. J., Banks, H. J. & Champ, B. R. (eds), pp. 029–934, (CABI, Canberra, Australia, 1994).
- Hell, K., Lamboni, Y., Houndekon, T. & Aildou, G. M. Augmented release of *Teretrius nigrescens* Lewis (Coleoptera: Histeridae) for the control of *Prostephanus truncatus* (Horn) (Coleoptera: Bostrichidae) in stored cassava chips. *J. Stored Prod. Res.* **42**, 367–376 (2006).
- Omondi, B. A., Jiang, N., van den Berg, J. & Schulthess, F. The flight activity of *Prostephanus truncatus* (Horn) (Coleoptera: Bostrichidae) and *Teretrius nigrescens* Lewis (Coleoptera: Histeridae) in Kenya. *J. Stored Prod. Res.* **47**, 13–19 (2011).
- Muatinte, B. L. & Cugala, D. R. Monitoring the establishment and dispersal of *Teretrius nigrescens* Lewis (Coleoptera: Histeridae), a predator of *Prostephanus truncatus* Horn (Coleoptera: Bostrichidae) in Manica Province, Mozambique. *Afr. Entomol.* **23**, 250–254 (2015).
- Elith, J. & Leithwick, J. R. Species distribution models: ecological explanations and predictions across space and time. *Ann. Rev. Ecol. Evol. System.* **40**, 677–697 (2009).
- Peterson, A. T. *et al.* Ecological niches and geographical distribution (Princeton University Press, Princeton, NJ, USA, 2011).
- Elith, J. Predicting distributions of invasive species. School of Biosciences, University of Melbourne, Australia, <https://arxiv.org/ftp/arxiv/papers/1312/1312.0851.pdf> Accessed 12 Feb. 2019 (2013).
- Thuiller, W. *et al.* Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biol.* **11**, 2234–2250 (2005).
- Phillips, S. J., Anderson, R. P. & Schapire, R. E. Maximum entropy modeling of species geographic distributions. *Ecol. Modell.* **190**, 231–259 (2006).
- Elith, J. *et al.* Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151 (2006).
- Merow, C., Smith, M. J. & Silander, J. A. A practical guide to MaxEnt for modeling species' distributions: What it does, and why inputs and settings matter. *Ecography* **36**, 1058–1069 (2013).
- Elith, J. *et al.* A statistical explanation of MaxEnt for ecologists. *Diversity Distrib.* **17**, 43–57 (2011).
- Barve, N. *et al.* The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecol. Modeling* **222**, 1810–1819 (2011).
- Jarnevich, C. S., Stohlgren, T. J., Kumar, S., Morissette, J. T. & Holcombe, T. R. Caveats for correlative species distribution modeling. *Eco. Inform.* **29**, 6–15 (2015).
- Bivand, R. & Rundel, C. RGeos: Interface to Geometry Engine - Open Source ('GEOS'). R package version 0.3–26, <https://CRAN.R-project.org/package=rgeos> (2017).



32. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/> (Accessed 12 February 2019).
33. RStudio. Integrated development environment for R, Version 1.1.453, <http://www.rstudio.org/> (Accessed 12 February 2019).
34. Beaumont, L. J., Hughes, L. & Poulson, M. Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecol. Modelling* **186**, 250–269 (2005).
35. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
36. Hijmans, R. J. Raster: Geographic data analysis and modeling. R package version 2.6-7 (2017).
37. Dormann, C. F. *et al.* Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 027–046 (2013).
38. Nix, H. A biogeographic analysis of Australian Elapid snakes. Longmore, R. (ed.), *Snakes: Atlas of Elapid snakes of Australia*, pp. 4–10, (Bureau of Flora and Fauna, Canberra, AU 1986).
39. Anderson, R. P. & Gonzalez, I. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecol. Modelling* **222**, 2796–2811 (2011).
40. Shcheglovitova, M. & Anderson, R. P. Estimating optimal complexity for ecological niche models: A jackknife approach for species with small sample sizes. *Ecol. Modelling* **269**, 9–17 (2013).
41. Warren, D. L. & Seifert, S. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecol. Soc. Am.* **21**, 335–342 (2011).
42. Elith, J. Quantitative methods for modeling species habitat: Comparative performance and an application to Australian plants. Ferson, S. & Burgman, M. (Eds), *Quantitative methods for conservation biology*, pp. 39–58, (Springer-Verlag, NY, NY, USA, 2000).
43. Muscarella, R. *et al.* ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods Ecol. Evol.* **5**, 1198–1205 (2014).
44. Radosavljevic, A. & Anderson, R. P. Making better MaxEnt models of species distributions: complexity, overfitting, and evaluation. *J. Biogeogr.* **41**, 629–643 (2014).
45. Peterson, A. T., Williams, R. & Chen, G. Modeled global invasive potential of Asian gypsy moths, *Lymantria dispar*. *Entomol. Exp. Appl.* **125**, 39–44 (2007).
46. Hijmans, R. J., Phillips, S., Leathwick, J. & Elith, J. dismo: Species Distribution Modeling. R package version 1.1-4 Available online at: <http://cran.r-project.org/web/packages/dismo/dismo.html> (Accessed 12 Feb. 2019).
47. Warren, D. L., Glor, R. E. & Turelli, M. Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* **62**, 2868–2883 (2008).
48. Warren, D. L., Glor, R. E. & Turelli, M. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* **33**, 607–611 (2010).
49. Rödder, D. & Engler, J. O. Quantitative metrics of overlaps in Grinnellian niches: Advances and possible drawbacks. *Glob. Ecol. Biogeogr.* **20**, 915–927 (2011).
50. Arthur, F. H., Throne, J. E., Maier, D. E. & Montross, M. D. Initial cooling cycles for corn stored in the southern United States: Aeration strategies based on weather data. *Am. Entomol.* **44**, 118–123 (1998).
51. Paini, D. R. & Yemshanov, D. Correction: Modelling the arrival of invasive organisms via the international marine shipping network: A khapra beetle study. *PLoS ONE* **7**, 10.137 (2012).
52. Parmesan, C. *et al.* Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature* **399**, 579–583 (1999).
53. Walther, G.-R. *et al.* Ecological responses to recent climate change. *Nature* **416**, 389–395 (2002).
54. Sakka, M. K. & Athanassiou, C. G. Competition of three stored-product bostrychids on different temperatures and commodities. *J. Stored Prod. Res.* **79**, 24–39 (2018).
55. Osipitan, A. A., Akintokun, K., Odeyemi, S. & Bankole, S. O. Evaluation of damage of some food commodities by larger grain borer – *Prostephanus truncatus* (Horn) (Coleoptera: Bostrichidae) and microbial composition of frass induced by the insect. *Arch. Phytopathol. Plant Prot.* **44**, 537–546 (2011).
56. Blagbrough, I. S., Bayoumi, S. A. L., Rowan, M. G. & Beeching, J. R. Cassava: an appraisal of its phytochemistry and its biotechnological prospects. *Phytochem.* **71**, 1940–1951 (2010).
57. Mombo, S., Dumat, C., Shahid, M. & Schreck, E. A socio-scientific analysis of the environmental and health benefits as well as potential risks of cassava production and consumption. *Environ. Sci. Poll. Res. Int.* **24**, 5207–5221 (2016).
58. Latif, S. & Müller, J. Potential of cassava leaves in human nutrition: A review. *Trends Food Sci. Technol.* **44**, 147–158 (2015).
59. Nansen, C., Meikle, W. G., Tigar, B., Harding, S. & Tchabi, A. Nonagricultural hosts of *Prostephanus truncatus* (Coleoptera: Bostrichidae) in a West African forest. *Ann. Entomol. Soc. Amer.* **97**, 481–491 (2004).
60. Mutainté, B. L. Population dynamics and integrated management of *Prostephanus truncatus* (Coleoptera: Bostrichidae) in Manica Province, Mozambique. Ph.D. thesis, (Environmental Sciences and Management at the North-West University (Potchefstroom Campus, 2017). Accessed 12 Feb. 2018).
61. Deb, J. C., Phinn, S., Butt, N. & McAlpine, C. A. Climatic-induced shifts in the distribution of teak (*Tectona grandis*) in tropical Asia: implications for forest management and planning. *Environ. Manag.* **60**, 422–435 (2017).
62. Zhu, G., Bu, W., Gao, Y. & Liu, G. Potential geographic distribution of brown marmorated stink bug invasion (*Halyomorpha halys*). *PLoS ONE* **7**, e31246 (2012).
63. Sobek-Swank, S., Kluza, D. A., Cuddington, K. & Byons, D. B. Potential distribution of emerald ash borer: What can we learn from ecological niche models using Maxent and GARP? *Forest Ecol. Manag.* **281**, 23–31 (2012).
64. Dos Santos, L. A. *et al.* Global potential distribution of *Drosophila suzukii* (Diptera, Drosophilidae). *PLoS ONE* **12**, e0174318 (2017).
65. Broennimann, A. *et al.* Evidence of climatic niche shift during biological invasion. *Ecol. Lett.* **10**, 701–709 (2007).
66. Hill, M. P., Elith, J., Macfadyen, S., Umina, P. & Hoffmann, A. A. Understanding niche shifts: Using current and historical data to model the invasive redlegged earth mite, *Halotydeus destructor*. *Divers. Distrib.* **18**, 191–203 (2012).
67. Duyck, P.-F. & Quilici, D. P. A review of relationships between interspecific competition and invasions in fruit flies (Diptera: Tephritidae). *Ecol. Entomol.* **29**, 511–520 (2004).
68. Bentlage, B. & Shcheglovitova, M. NichePy: Modular tools for estimating the similarity of ecological niche and species distribution models. *Methods Ecol. Evol.* **3**, 484–489 (2012).
69. Li, H., Liu, Y., Gao, X. & Li, X. Preparation and characterization of cassava starch-based adsorbents for separating of azeotropic ethanol-water in biofuels ethanol production. *J. Chem. Technol. Biotechnol.* **91**, 977–984 (2016).

## Acknowledgements

This paper reports the results of research only. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA or the University of Minnesota. The USDA and the University of Minnesota are equal opportunity employers and providers.

### Author Contributions

F.H.A. and W.R.M. conceived of the project. F.H.A. compiled documented occurrences for the species. A.C.M. added further ideas and performed the statistical and computational modeling for the paper. W.R.M. compiled the Figures. All authors contributed equally to the drafting, writing, and editing of the manuscript, and all approved the final version for publication.

### Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-42974-5>.

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

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019