Geographic potential of the world's largest hornet, *Vespa mandarinia* Smith (Hymenoptera: Vespidae), worldwide and particularly in North America

Claudia Nuñez-Penichet [1,2], Luis Osorio-Olvera [2,3], Victor H. Gonzalez [1,4], Marlon E Cobos [1,2], Laura Jiménez [1,2], Devon A. DeRaad [1,2], Abdelghafar Alkishe [1,2], Rusby G. Contreras-Díaz [3,5], Angela Nava-Bolaños [2], Kaera Utsumi [1], Uzma Ashraf [6], Adeola Adeboje [1], A. Townsend Peterson [1,2], Jorge Soberón [Corresp. 1,2]

1 Department of Ecology & Evolutionary Biology, University of Kansas, Lawrence, KS, United States

2 Biodiversity Institute, University of Kansas, Lawrence, Kansas, United States

3 Departamento de Matemáticas, Facultad de Ciencias, Universidad Nacional Autónoma de México, Ciudad de México, Ciudad de México, Mexico

4 Undergraduate Biology Program, University of Kansas, Lawrence, KS, United States

5 Posgrado en Ciencias Biológicas. Unidad de Posgrado, Universidad Nacional Autónoma de México, Ciudad de México, Ciudad de México, México

6 Department of Environmental Sciences and Policy, Lahore School of Economics, Lahore, Pakistan

Corresponding Author: Jorge Soberón
Email address: jsoberon@ku.edu

The Asian giant hornet (AGH, *Vespa mandarinia*) is the world's largest hornet, occurring naturally in the Indomalayan region, where it is a voracious predator of pollinating insects including honey bees. In September 2019, a nest of Asian giant hornets was detected outside of Vancouver, British Columbia and in May 2020 an individual was detected nearby in Washington state, indicating that the AGH successfully overwintered in North America. Because hornets tend to spread rapidly and become pests, reliable estimates of the potential invasive range of *V. mandarinia* in North America are needed to assess likely human and economic impacts, and to guide future eradication attempts. Here, we assess climatic suitability for AGH in North America, and suggest that, without control, this species could establish populations across the Pacific Northwest and much of eastern North America. Predicted suitable areas for AGH in North America overlap broadly with areas where honey production is highest, as well as with species-rich areas for native bumble bees and stingless bees of the genus *Melipona* in Mexico, highlighting the economic and environmental necessity of controlling this nascent invasion.

1   **Geographic potential of the world's largest hornet, *Vespa mandarinia* Smith**

2   **(Hymenoptera: Vespidae), worldwide and particularly in North America**

3   Claudia Nuñez-Penichet[1, 2], Luis Osorio-Olvera[2, 3], Victor H. Gonzalez[1, 4], Marlon E. Cobos[1, 2],

4   Laura Jiménez[1, 2], Devon A. DeRaad[1, 2], Abdelghafar Alkishe[1, 2], Rusby G. Contreras-Díaz[3, 5],

5   Angela Nava-Bolaños[2], Kaera Utsumi[1], Uzma Ashraf[6], Adeola Adeboje[1], A. Townsend

6   Peterson[1, 2], Jorge Soberón[1, 2*]

7   [1]Department of Ecology & Evolutionary Biology, University of Kansas, Lawrence, KS 66045

8   USA.

9   [2]Biodiversity Institute, University of Kansas, Lawrence, KS 66045 USA.

10   [3]Departamento de Matemáticas, Facultad de Ciencias, Universidad Nacional Autónoma de

11   México, Circuito Exterior s/n, Cd. Universitaria, 04510 Ciudad de México, México.

12   [4]Undergraduate Biology Program, Haworth Hall, 1200 Sunnyside Ave. University of Kansas,

13   Lawrence, KS 66045 USA.

14   [5]Posgrado en Ciencias Biológicas. Unidad de Posgrado, Edificio A, 1er Piso. Circuito de

15   Posgrados, Ciudad Universitaria, Delegación Coyoacán. C.P. 04510. Ciudad de México, México.

16   [6]Department of Environmental Sciences and Policy, Lahore School of Economics, Lahore,

17   Pakistan.

19   *Corresponding Author: Jorge Soberón

20   Department of Ecology & Evolutionary Biology and Biodiversity Institute, University of Kansas,

21   1345 Jayhawk Blvd., Lawrence, Kansas 66045, USA. jsoberon@ku.edu

23   **ORCID:**

24   CNP: https://orcid.org/0000-0001-7442-8593

25   LOO: https://orcid.org/0000-0003-0701-5398

26   VHG: https://orcid.org/0000-0002-4146-1634

27   MEC: https://orcid.org/0000-0002-2611-1767

28   LJ: https://orcid.org/0000-0002-6683-9576

29   DAD: https://orcid.org/0000-0003-3105-985X

30   AbA: https://orcid.org/0000-0003-2927-514X

31   RGCD: https://orcid.org/0000-0002-0569-8984

32   ANB: https://orcid.org/0000-0002-4371-5415

33   KU: https://orcid.org/0000-0001-5935-7299

34   UA: https://orcid.org/0000-0003-4319-9315

35   AdA: https://orcid.org/0000-0001-8513-7804

36   ATP: http://orcid.org/0000-0003-0243-2379

37   JS: https://orcid.org/0000-0003-2160-4148

38   **ABSTRACT**

39   The Asian giant hornet (AGH, *Vespa mandarinia*) is the world's largest hornet, occurring

40   naturally in the Indomalayan region, where it is a voracious predator of pollinating insects

41   including honey bees. In September 2019, a nest of Asian giant hornets was detected outside of

42   Vancouver, British Columbia and in May 2020 an individual was detected nearby in Washington

43   state, indicating that the AGH successfully overwintered in North America. Because hornets tend

44   to spread rapidly and become pests, reliable estimates of the potential invasive range of *V.*

45   *mandarinia* in North America are needed to assess likely human and economic impacts, and to

46   guide future eradication attempts. Here, we assess climatic suitability for AGH in North

47   America, and suggest that, without control, this species could establish populations across the

48   Pacific Northwest and much of eastern North America. Predicted suitable areas for AGH in

49   North America overlap broadly with areas where honey production is highest, as well as with

50   species-rich areas for native bumble bees and stingless bees of the genus *Melipona* in Mexico,

51   highlighting the economic and environmental necessity of controlling this nascent invasion.

52

53   **Keywords:** Asian giant hornet, dispersal simulation, ecological niche modeling, invasive

54   species, pollinator threats

Introduction

55

56   Invasive species represent major threats to biodiversity, as they can alter ecosystem

57 processes and functions (Pyšek & Richardson, 2010; Vilà et al., 2011), and often contribute to

58 the decline of imperiled species (e.g., Wilcove et al., 1998; Dueñas et al., 2018). The economic

59 damage to agriculture, forestry, and public health, resulting from invasive species totals nearly

60 $120 billion annually in the United States alone (Pimentel, Zuniga & Morrison, 2005).

61   Even in the midst of the global uncertainty and socio-economic distress resulting from

62 the COVID-19 pandemic, the recent detection of the Asian Giant Hornet (AGH, *Vespa*

63 *mandarinia* Smith, Hymenoptera: Vespidae), in North America (Bérubé, 2020), received

64 significant public attention. This social insect is the world's largest hornet (2.5–4.5 cm body

65 length), and occurs naturally across Asia, including in India, Nepal, Sri Lanka, Vietnam, Taiwan,

66 and Japan, at elevations ranging between 850 and 1900 (Matsuura & Sakagami, 1973; Archer,

67 2008; Smith-Pardo, Carpenter & Kimsey, 2020). As in other temperate-zone social species,

68 annual colonies of the AGH, which may contain up to 500 workers, die at the onset of winter and

69 mated queens overwinter in underground cavities. After emerging in the spring, each queen starts

70 a new colony in a pre-existing cavity, typically in tree roots or an abandoned rodent nest (Archer,

71 2008). Like other species of *Vespa*, AGH is a voracious predator of insects, particularly honey

72 bees and other social wasps. Attacks on honey bee hives occur late in the development of the

73 hornet colony and prior to the emergence of reproductive individuals (males and new queens),

74 the timing of which depends on location (e.g., Matsuura & Sakagami, 1973; Matsuura, 1988;

75 Archer, 2008).

76   In its native range, AGH attacks several species of bees, some of which have developed

77 sophisticated defense mechanisms against attacks (Ono et al., 1995; Kastberger, Schmelzer &

78   Kranner, 2008; Fujiwara, Sasaki & Washitani, 2016). The best documented, colony-level defense

79   mechanism is in the Asiatic honey bee, *Apis cerana* Fabricius, which can detect site-marking

80   pheromones released by AGH scouts, and responds by engulfing a single hornet in a ball

81   consisting of up to 500 bees. The heat generated by the vibration of the bees' flight muscles, and

82   the resulting high levels of CO_2 from respiration effectively kill the hornet (Ono et al., 1995;

83   Sugahara, Nishimura & Sakamoto, 2012). In contrast, European honey bees (*A. mellifera* L.)

84   cannot detect and respond to AGH marking pheromones and colonies are more or less

85   defenseless against AGH attacks (McClenaghan et al., 2019). As few as a dozen AGH can

86   destroy a European honey bee colony of up to 30,000 individuals (Matsuura & Sakagami, 1973).

87   In addition to the threat to the beekeeping industry, the introduction of AGH in North

88   America is also concerning for public health. Their powerful stings can induce severe allergic

89   reactions or even death in hypersensitive individuals (Schmidt et al., 1986; Yanagawa et al.,

90   2007). Annually, 30-40 people die from AGH stings in Japan, most as a result of anaphylaxis or

91   sudden cardiac arrest (Matsuura & Sakagami, 1973); similar deadly cases have been reported

92   from China (Li et al., 2015).

93   Although invasive species are typically limited by dispersal ability and suitability of

94   novel environments, vespid hornets are well known for their invasive success and excellent

95   dispersal capacity (Beggs et al., 2011; Monceau, Bonnard & Thiéry, 2014). As such, the

96   introduction of AGH in the Pacific Northwest poses a potentially serious ecological and socio-

97   economic risk in North America. Here, we use ecological niche modeling (ENM) to detect areas

98   of suitable environments for this species worldwide, with particular emphasis on North America.

99   We also use a dispersal simulation approach to detect potential invasion paths of this species

100 within North America. A similar methodology for projecting AGH invasion potential has been

101 implemented by Zhu et al. (2020); we build upon this framework by introducing several

102 modifications to the modelling approach, and investigating further the potential ecological and

103 economic impacts of an AGH invasion in North America.

104

105 **Methods**

106 *Occurrence and environmental data*

107 We downloaded occurrence data for *V. mandarinia* from the Global Biodiversity

108 Information Facility database (GBIF; https://www.gbif.org/). We kept records from the species'

109 native range (Fig. 1) separate from non-native occurrences facilitated by human introduction. We

110 cleaned occurrences from the native distribution following Cobos et al. (2018) by removing

111 duplicates and records with doubtful or missing coordinates. To avoid model overfitting derived

112 from spatial autocorrelation and overdominance of specific regions due to sampling bias, we

113 thinned these records spatially in two ways: by geographic distance and by density of records per

114 country (Fig. 2). In the first case (distance-based thinning), we excluded occurrences that were

115 <50 km away from another locality (Anderson, 2012). In the second thinning approach (country-

116 density thinning), we randomly reduced numbers of occurrences in countries with the densest

117 sampling, namely Japan, Taiwan, and South Korea (from 30, 6, and 5, to 6, 2, and 2 occurrences,

118 respectively), to match an approximate reference density of India, Nepal, and China. We used the

119 package ellipsenm (Cobos et al., 2020; available at https://github.com/marlonecobos/ellipsenm)

120 in R 3.6.2 (R Core Team, 2019) to clean and thin the data. We then treated both data sets

121 independently in all subsequent analysis steps.

122   For environmental predictors, we used bioclimatic variables at 10' resolution (~18 km at

123   the Equator) from the MERRAclim database (Vega, Pertierra & Olalla-Tárraga, 2018). We

124   excluded four variables because they are known to contain spatial artifacts as a result of

125   combining temperature and humidity information (Escobar et al., 2014): mean temperature of

126   most humid quarter, mean temperature of least humid quarter, specific humidity mean of

127   warmest quarter, and specific humidity mean of coldest quarter. The 15 variables remaining were

128   masked to an area for model calibration (**M**, see Ecological niche modeling).

129   These 15 variables were submitted to a principal component analysis (PCA) to reduce

130   dimensionality and multicollinearity. Raw variables and principal components (PCs) were

131   considered separately in all subsequent analyses. To select a set of raw variables, we reduced

132   them to a subset with Pearson's correlation coefficients (r) ≤ 0.85, choosing the most

133   biologically relevant or interpretable variables based on our knowledge of AGH natural history

134   (Simões et al., 2020). The PCA was calibrated using environmental variation across the **M** area,

135   and transferred to the whole world. All analyses were done in R; specifically, raster processing

136   was done using the packages raster (Hijmans et al., 2020), rgeos (Bivand et al., 2020b), and rgdal

137   (Bivand et al., 2020a); PCA was done using the ntbox package (Osorio-Olvera et al., 2020;

138   available in https://github.com/luismurao/ntbox).

139

140   *Ecological niche modeling*

141   To identify a calibration area (ostensibly equivalent to **M**; Owens et al., 2013) for our

142   models, we considered a region contained within a buffer of 500 km around the known

143   occurrence records after the 50 km thinning process (Fig. 1). This distance was selected

144   considering the species' dispersal ability (Matsuura & Sakagami, 1973; APHIS, 2020). We used

145   all pixels in **M** (15,411) as the background across which to calibrate the models.

146   Given uncertainty deriving from specific treatments of occurrence records and

147   environmental predictors in ecological niche modeling (Alkishe et al., 2020), we calibrated

148   models via four distinct schemes: (1) using raw variables and distance-based thinned

149   occurrences, (2) using PCs and distance-based thinned occurrences, (3) using raw environmental

150   variables and country-density thinned occurrences, and (4) using PCs and country-density

151   thinned occurrences (Fig. 2). For each scheme, we calibrated models five times, each time

152   randomly selecting 50% of the occurrences for calibrating models, and using the remaining

153   records for testing (Cobos et al., 2019a).

154   Each process of model calibration consisted of creating and evaluating candidate models

155   using Maxent (Phillips, Anderson & Schapire, 2006; Phillips et al., 2017) based on distinct

156   parameter settings: 10 regularization multiplier values (0.10, 0.25, 0.50, 0.75, 1, 2, 3, 4, 5, 6),

157   eight feature classes (lq, lp, lqp, qp, q, lqpt, lqpth, lqph, where l is linear, q is quadratic, p is

158   product, t is threshold, and h is hinge), and all combinations of more than two predictor variables

159   (Cobos et al., 2019b; Table S1-S2). We tested 4560 models using raw variables and 880 using

160   PCs (see Data preprocessing and model calibration), in tandem with the two methods of reducing

161   occurrence data described above. We assessed model performance using partial ROC (for

162   statistical significance; Peterson, Papeş & Soberón, 2008), omission rates ($E = 5\%$, for predictive

163   ability; Anderson, Lew & Peterson, 2003), and Akaike Information Criterion corrected for small

164   sample sizes (AICc, for model complexity; Warren & Seifert, 2011). We selected models with

165   delta AICc ≤ 2 (Cobos et al., 2019a) from those that were statistically significant and had

166   omission rates below 5%.

167   After model calibration, we created models with the selected parameter values, using all

168 occurrences after the corresponding thinning process, with 10 bootstrap replicates, cloglog

169 output, and model transfers using three types of extrapolation (free extrapolation, extrapolation

170 and clamping, no extrapolation; Owens et al., 2013). Not all calibration processes identified

171 models that met all three criteria of model selection; we did not consider those models in further

172 analyses (Fig. 2; Table 1). As a final evaluation step, we tested whether each replicate of the

173 selected models was able to anticipate the known invasive records of the species in the Americas

174 (British Columbia, Canada; Washington, USA). For each scheme, using only those model

175 replicates that met the selection criteria and correctly predicted independent occurrences, we

176 created two types of consensus: (1) a median of the medians obtained for each parameterization,

177 and (2) the sum of all suitable areas derived from binarizing each replicate using a modified least

178 presence (5% omission) threshold (Fig. 2).

179   As we transferred models to the entire world, we used the mobility-oriented parity metric

180 (MOP; Owens et al., 2013) to detect areas where strict or combinational extrapolation risks could

181 be expected, given the presence of non-analogous conditions with respect to the environments

182 manifested across the calibration area. We used areas where extrapolation risks were detected

183 using MOP to trim our binary results (suitable areas) to avoid potentially problematic

184 interpretations based on extrapolative situations. Model calibration, production of selected

185 models with replicates, and MOP analyses were done in R using the package kuenm (Cobos et

186 al., 2019a); raster processing and independent testing of models were done using the package

187 raster and other base functions in R.

188

189 *Dispersal simulations*

190   We used the binary outputs from the final consensus models (suitable and unsuitable

191 areas, without areas of strict extrapolation) to simulate invasion dynamics of the AGH. All

192 simulations were started from the Pacific Northwest, from sites already known to be occupied by

193 the AGH. The simulations were performed using the cellular automaton dynamic model included

194 in the bam R package (Osorio-Olvera & Soberón, 2020; available at

195 https://github.com/luismurao/bam). Under this discrete model, given an occupied area at time t,

196 two layers of information are needed to obtain the occupied area at time $t + 1$: (i) the binary layer

197 of suitability for the species, and (ii) a connectivity matrix determined by the species' ability to

198 reach neighboring cells in one time unit (known as "Moore's neighborhood"; Gray et al., 2003,

199 that defines patches that are connected by dispersal). At each step, each of the suitable cells can

200 be either occupied or not by the species. If a cell is occupied, adjacent cells can be visited by the

201 species, and if suitable, they become occupied. This method is similar to the one implemented in

202 the MigClim R package (Engler, Hordijk & Guisan, 2012), but uses a simpler dispersal kernel

203 and parameterization.

204   For each of the schemes followed to obtain ecological niche models for *V. mandarinia*,

205 we performed a set of simulations in which we explored different degrees of connectivity (1–8,

206 10, and 12 neighbor cells) and different suitability thresholds (10 equidistant levels from 3–10%

207 of the presence points) to create the binary maps. All simulations were done with 200 steps. In

208 the end, we visualized the simulation results by summing the occupied distribution layers

209 obtained from each set of simulations. A value of 100 in these final layers means that the species

210 reached that cell in 100% of the simulations, whereas a value of 0 means that the species never

211   reached that cell. Further details regarding the simulation processes can be found in the

212   Supplementary Information.

213

214   *Honey production and native bee richness in North America*

215   To explore potential ecological and economic impacts of the invasion of the AGH in

216   North America, we explored annual, state-level production of honey as well as species richness

217   of bumble bees (*Bombus* Latreille) and stingless bees (*Melipona* Illiger) in Mexico and the

218   United States. We extracted data on 2016 honey production (in US dollars) for the United States

219   from the U.S. Department of Agriculture (USDA; available at

220   https://quickstats.nass.usda.gov/#4A0314DA-F3E5-3B06-ADD1-CA8032FBD937) and from the

221   Instituto Nacional de Estadística, Geografía e Informática (INEGI) for Mexico

222   (https://atlasapi2019.github.io/cap4.html). For native species richness, we obtained a list of

223   species of bumble bees and stingless bees that occur in Mexico and the United States from

224   Discover Life (https://www.discoverlife.org/) and downloaded their occurrence data from GBIF.

225   We chose these bee taxa as likely targets of AGH because the species in these groups are of

226   similar body size and behavior to the typical prey of these hornets: they are social insects that

227   form annual or perennial colonies that can have a few hundreds to as many as 10,000 individuals

228   (Cueva del Castillo, Sanabria-Urbán & Serrano-Meneses, 2015; Viana et al., 2015), and store

229   honey and pollen inside their nests (Michener, 2000). To summarize species richness of these

230   two genera, we created a presence absence matrix (PAM; Arita et al., 2008) for North America,

231   based on geographic coordinates of occurrence data, with a pixel size of one degree. The PAM

232   was created in R with the package biosurvey (Nuñez-Penichet et al., 2020; available at

233   https://github.com/claununez/biosurvey).

234   To assure transparency and reproducibility of our work, we include an Overview, Data,

235   Model, Assessment, and Prediction protocol (ODMAP; Zurell et al., 2020) in our supplementary

236   materials. This metadata summary provides a detail key steps included in our analyses.

237

238   **Results**

239   *Data preprocessing and model calibration*

240   We retained 172 occurrence records for *V. mandarinia* after initial data cleaning, 49

241   records after the distance-based thinning approach, and 18 records after the country-density

242   thinning approach (Fig. 1). As environmental predictors, we selected six raw variables based on

243   correlation levels and natural history criteria: isothermality (BIO3), maximum temperature of

244   warmest month (BIO5), minimum temperature of coldest month (BIO6), temperature annual

245   range (BIO7), specific humidity of most humid month (BIO13), and specific humidity of least

246   humid month (BIO14). From the PCA, we kept the first four PC axes, as they explained 97.9%

247   of the cumulative variance (Figure S1).

248   The number of models that met the selection criteria was considerably smaller than the

249   total number of models tested (Table 1). The calibration schemes including raw variables had

250   fewer models selected than the those using PCs (11, 19, 6, 15 models selected for raw/distance-

251   thinned, PC/distance-thinned, raw/country-density, and PC/country-density, respectively). The

252   number of replicates of those selected models that predicted the *V. mandarinia* invaded areas in

253   North America was also small and changed among types of extrapolation (Table 1).

254

Ecological niche model predictions

255

256   In our models, areas predicted as suitable for the AGH varied among calibration schemes,

257   in both scale and geographic pattern (Fig. 3, Figures S2-S4). The differences are conspicuous

258   between the two types of thinning approaches, which resulted in models with different numbers

259   of occurrence records. Models with country-density thinning (18 records) resulted in broad

260   predicted suitable areas worldwide, with areas of higher values of suitability concentrated in

261   tropical regions (Fig. 3, Figures S2-S4). In contrast, models created with the greater number of

262   occurrences (49 records) from the geographic distance thinning predicted more patches of

263   suitable areas across large extensions of Southeast Asia, Europe, West Africa, Central America,

264   northern South America, and the Pacific Northwest and southeastern United States (Fig. 3,

265   Figures S2-S4). In the calibration area, the areas detected with high levels of suitability were

266   larger in the scheme with geographic distance thinned occurrences and the raw variables and

267   smaller in the predictions obtained with the country-density thinned occurrences and the PCs as

268   environmental predictors (Fig. 3). In all schemes, the two northernmost occurrence points of this

269   species in China were accorded relatively low levels of suitability (Fig. 3). Predicted suitable

270   areas for this hornet worldwide were also different among types of extrapolation considered in

271   this study, especially as regards distribution size rather than location (Figures S2-S4).

272   In North America, across multiple model calibration schemes, our various models agreed

273   in predicting suitable areas for AGH in the Pacific region of southwestern Canada, the Pacific

274   Northwest, the southeastern United States, and from central Mexico south to southernmost

275   Panama (Fig. 4). Our model calibration schemes also agreed in identifying the Rocky Mountains

276   and Great Plains as unsuitable for this species (Fig. 4).

277   Prevalences (proportion of suitable area) varied among the data thinning schemes. In the

278   case of models created with raw variables, prevalence values of 0.171 and 0.164 were detected

279   when using spatially rarefied and country-density rarefied records, respectively. Models created

280   with PCs had prevalences of 0.248 and 0.239 for spatially rarefied and country-density rarefied

281   records, respectively (Table S3).

282

283   *Extrapolation risks in model projections*

284   The pattern of areas detected with risk of extrapolation was similar worldwide between

285   thinning methods, but different between raw variables and PCs (Fig. 5, Figure S8). Most tropical

286   areas predicted as suitable were identified as regions with high extrapolation risk (Figure S8).

287   For raw variables, the areas with extrapolation risk in North America included most of Canada

288   and Alaska, whereas for PCs areas with extrapolation risk included large portions of Mexico and,

289   the central-southwestern United States, as well as the islands north of Hudson Bay in Canada

290   (Fig. 5).

291

292   *Simulations of potential invasion*

293   The simulations of potential sequences of colonization and dispersal of AGH in North

294   America, starting from the known invaded localities, showed agreement among calibration

295   schemes in predicting an invasion across the Pacific Northwest from southernmost Alaska to

296   southernmost California in the United States (Fig. 6). In contrast, we found that the dispersal

297   distance required to invade all the way to the East Coast of North America varied among

298   calibration schemes. In the schemes using raw variables, the route of invasion to reach the East

299   Coast goes from the Pacific Northwest down to California and Mexico, and then up the East

300   Coast of North America. A dispersal distance of 10 cells (where each cell represents ~18 km)

301   was enough to reach the East Coast (see left panels in Fig. 6). For the scheme using the 50 km

302   spatially-thinned occurrences and PCs, the invasion follows a more direct route from the Pacific

303   Northwest to the East Coast that goes through the United States, and the required dispersal

304   distance to reach the East Coast was only 4 cells (top right panel in Fig. 6). Finally, in the case of

305   country-density thinned occurrences and PCs, the invasion goes from the Pacific Northwest

306   through Canada to the Atlantic, and then down the East Coast to the United States. A distance of

307   8 cells was needed to make this invasion route possible (bottom right panel in Fig. 6).

308

Discussion

309   **Discussion**

310   The patterns of suitability that we found in North America across multiple input data

311   processing schemes (Fig. 6) are broadly concurrent with the results obtained by Zhu et al. (2020),

312   who used an ensemble modeling approach. This concordance (both among our selected models,

313   and between our models and the ensemble models), gives us confidence that the Pacific

314   Northwest and southeastern United States represent suitable areas for AGH. In contrast with the

315   results of Zhu et al. (2020), however, our dispersal simulations indicate a larger potential

316   invasion area in the United States, with the AGH potentially crossing to eastern North America

317   via a southern invasion route, through Mexico and Texas; a southeast-ward route crossing Idaho,

318   Wyoming, and Colorado; or a northern route across Canada and the Great Lakes region (Fig. 6).

319   Quantifying the probability of the AGH following any one of the individual dispersal

320   routes presented would require precise quantification of dispersal ability, and discerning the real-

321   world validity of each of the four modeling outcomes. Instead of attempting to guess, we present

322   several models that offer multiple plausible invasion scenarios. Across all scenarios presented,

323   the AGH is expected to establish populations along the coastal Pacific Northwest via short-

324   distance dispersal, and it is likely to invade the southeastern United States if it has even moderate

325   dispersal potential (Fig. 6). It is important to note that these potential invasion routes consider

326   only the natural dispersal ability of this hornet, and do not take into account the effect of

327   potential accidental human-aided dispersal through the transport of soil and wood, where

328   fertilized queen AGHs overwinter (Archer, 1995). Such unwitting human-aided dispersal is a

329   serious concern, as it could potentiate a rapid invasion of this hornet to environmentally suitable,

330   yet currently isolated places across North America. Our simulations with larger numbers for

331   neighbor cells are perhaps a good illustration of what could be expected if dispersal events to

332   very long distances occur.

333   Contrasts between our prediction of extensive invasion potential, and Zhu et al.'s (2020)

334   more conservative predictions, arise from Zhu et al.'s (2020) use of MigClim (Engler, Hordijk &

335   Guisan, 2012) to model dispersal of the AGH in western North America. MigClim is a cellular

336   automaton platform that models the state of grid cells as occupied or unoccupied. Although we

337   used the same modelling technique, our dispersal kernel is a much simpler "Moore

338   Neighborhood" (Gray et al., 2003) approach, in which cells surrounding an occupied focal cell

339   (to 1,2...d neighbors) may become occupied, depending on their suitability. MigClim instead

340   assumes a probabilistic contagion model that requires parameter estimates for number of

341   propagules, and short- and long-distance-decay rates. Given the lack of empirical data to inform

342   values for those parameters, we prefer a simpler algorithm to explore how connected clusters of

343   suitable cells are across different values of the single parameter d. Another factor resulting in

344   these differences is the number of simulation steps used in our approach (200). From a biological

345   perspective, this implies that 200 dispersal events resulting in colonization of suitable cells

346   happened. Although this number may appear excessive, it gives a view of a scenario in which no

347   action is taken to prevent AGH invasion in North America and the species builds to large local

348   populations. For a more conservative view of the expected invasion, one could concentrate in

349   areas with high values on the layers obtained from our simulations.

350   The areas in North America that our models identified as highly suitable for this hornet

351   overlap broadly with the states where honey production is highest, and species richness of

352   *Bombus* and *Melipona* are highest (Fig. 7). These results give credence to public concerns that, if

353   established, the AGH could pose a serious economic threat to the beekeeping industry in Oregon,

354   northern California, Georgia, Alabama, and Florida. In the United States alone, the European

355   honey bee provides at least $15 billion worth of pollination services and generates between $300

356   and 500 million in harvestable honey and other products each year (Calderone, 2012). In

357   Mexico, impacts on the honey bee industry are also expected, in tropical areas of the country that

358   have suitable areas for the AGH, particularly in the states of Yucatan, Campeche, and Quintana

359   Roo. Beekeepers in the United States and Mexico may have to adopt mitigation practices to

360   avoid serious losses, such as those developed by Japanese beekeepers including the use of

361   protective screens or traps at the hive entrance that can exclude AGHs based on body size

362   (Matsuura & Sakagami, 1973; Mahdi, Glaiim & Ibrahim, 2008). Potential establishment of the

363   AGH in North America adds an additional layer of environmental and economic stress to a

364   beekeeping industry already suffering from high annual hive mortality rates resulting from the

365   combined effects of pesticides, diseases, and poor nutrition (Goulson et al., 2015).

366   The ecological impact of AGH on the local bee fauna is more challenging to predict than

367   the economic impact on honey production, because it is not clear which native bee species would

368   be particularly targeted by AGH in North America. We explore *Bombus* and *Melipona* species as

369   likely prey candidates of AGH because, among the >4000 bee species occurring in this region

370   (Ascher & Pickering, 2020), these two groups of bees are social, locally abundant, and make

371   annual or perennial colonies (Michener, 2007; Cueva del Castillo, Sanabria-Urbán &

372   Serrano-Meneses, 2015; Viana et al., 2015). Thus, they may represent predictable food sources

373   for the AGH. It is crucial to consider this potential threat because both *Bombus* and *Melipona*

374   bees are important pollinators that have already experienced population losses and local

375   extirpations, reflecting changes in landscape and agricultural intensification (Brown & Albrecht,

376   2001; Cameron et al., 2011). Furthermore, these species, as well as the European honey bee, lack

377   behavioral responses to prevent predation by the AGH (Matsuura & Sakagami, 1973;

378   McClenaghan et al., 2019), because they have no shared evolutionary history with the AGH, and

379   are thus vulnerable to its predatory and antagonistic behavior. The economic and cultural

380   importance of species of *Melipona* in America is well-documented, particularly in the Yucatan

381   Peninsula in Mexico, where these bees have been traditionally raised for honey and were even

382   considered gods outright in Mayan times (Ayala, Gonzalez & Engel, 2013; Quezada-Euán et al.,

383   2018). It is important to mention, however, that the risk to *Melipona* species may be lower than

384   that to *Bombus* species because entrances to the hives of some species of *Melipona* are narrow,

385   allowing a single bee to pass at a time (Couvillon et al., 2007), unlike the entrances to the hives

386   of honey bees and many bumble bees, which are wider.

387   The AGH is not the first hornet to invade North America, and species of *Vespa* are well-

388   known for their invasive success and excellent dispersal capacity (Beggs et al., 2011; Monceau,

389   Bonnard & Thiéry, 2014). The solitary giant resin bee, *Challomegachile sculpturalis* is an Asian

390   taxon which was recently introduced in the United States. Only 15 years after its initial detection

391   near Baltimore, Maryland, this species had invaded most of the southeastern United States

392   (Hinojosa-Díaz et al., 2005). Additionally, the European hornet, *Vespa crabro* L., a Eurasian

393   species that was accidentally introduced to North America in the 1800s, occupies a similar

394   invasive range in the United States (Smith et al. 2020). These examples indicate considerable

395   precedent for hornet invasion and establishment in the southeastern United States, but the AGH

396   poses a unique biodiversity risk as a direct predator of bees. Because the Pacific Northwest is

397   consistently predicted as suitable for the AGH, preventing further establishment and spread of

398   recently detected introduced populations near Seattle and Vancouver is essential. If these

399   introduced individuals are not eradicated, they may flourish under the suitable climatic

400   conditions, establishing many more colonies that will be difficult to control. Preventing

401   establishment of the AGH in the Pacific Northwest is especially critical because an established

402   AGH population in the Pacific Northwest would provide a source population for potential long-

403   range dispersers that could use multiple potential invasion routes (Fig. 6) to reach suitable habitat

404   in the eastern United States, facilitating full-scale invasion. In light of this, we recommend

405   official monitoring protocols for the vulnerable Pacific Northwest region, and encourage citizen-

406   science monitoring efforts, which may be the fastest and most effective way to detect potential

407   range expansions.

408   Although AGH is primarily found in temperate areas in its native range, some of its

409   populations reach subtropical regions like Taiwan (Archer, 2008), which indicates a broad

410   temperature tolerance. This southern part of the species' native range might explain why our

411   models predicted suitable areas in South America, Africa, and elsewhere (Figure 2S-S7).

412   Although temperature is a critical factor that determines the abundance and distribution of

413   organisms (Sunday, Bates & Dulvy, 2012), factors such as desiccation resistance may be equally

414   important for some species. For example, for ants and some bees, desiccation tolerance is a good

415   predictor of species' distributions (Bujan, Yanoviak & Kaspari, 2016; Burdine & McCluney,

416   2019). For example, humidity is important for the regulation of temperature in nests of the

417   European hornet (Klingner et al., 2005) and, in some species of stingless bees, regulation of

418   humidity appears to be more important than regulation of temperature to maintain colony health

419   (Ayton et al., 2016). Unfortunately, heat and desiccation tolerances, factors that might improve

420   predictions of this species' distributional potential, are unknown for the AGH. In other hornets,

421   subtropical populations tend to have longer population cycles than temperate populations

422   (Archer, 2008), so negative impacts of an AGH invasion may be stronger in tropical or

423   subtropical areas.

424   In summary, our modeling approach allowed us to recognize how predicted suitable

425   areas can be depending on distinct schemes of data treatment. We showed that this variability

426   can derive from crucial decisions made during the initial steps of ecological niche modeling

427   exercises. These results highlight the importance of such initial decisions, as well as the need to

428   recognize sources of variability. This point is of special importance in predicting the potential for

429   expansion of invasive species, as uncertainty increases when models are transferred to areas

430   where environmental conditions are different. Our analyses and simulations revealed the

431   potential of the AGH to invade large areas in North America and the likely paths of such an

432   invasion. We also showed that predicted suitable areas for the AGH overlap broadly with those

433   where honey production is highest in the United States and Mexico, as well as with species-rich

434   areas for bumble bees and stingless bees. These results bring light to the potential implications of

435   uncontrolled dispersal of the AGH to suitable environments in North America, and highlight the

436   need for rapid eradication actions to mitigate potential biodiversity and economic losses.

437

438   **Acknowledgments**

439   We would like to thank the members of the KUENM group for their support in the development

440   of this manuscript. We also thank Allan Smith-Pardo for letting us use the photograph of AGH in

441   lateral view (Fig. 1B). ANB would like to thank Secretaría de Educación, Ciencia, Tecnología e

442   Innovación de la Ciudad de México.

443

444   **References**

445   APHIS. 2020. New pest response guidelines: *Vespa mandarinia*. U.S. Department of

446   Agriculture, Animal and Plant Health Inspection Service, Plant Protection and

447   Quarantine.

448   Alkishe A, Cobos ME, Peterson AT, Samy AM. 2020. Recognizing sources of uncertainty in

449   disease vector ecological niche models: an example with the tick *Rhipicephalus*

450   *sanguineus* sensu lato. *Perspectives in Ecology and Conservation*. DOI:

451   10.1016/j.pecon.2020.03.002.

452   Anderson RP. 2012. Harnessing the world's biodiversity data: promise and peril in ecological

453   niche modeling of species distributions. *Annals of the New York Academy of Sciences*

454   1260:66–80. DOI: 10.1111/j.1749-6632.2011.06440.x.

455   Anderson RP, Lew D, Peterson AT. 2003. Evaluating predictive models of species' distributions:

456   criteria for selecting optimal models. *Ecological Modelling* 162:211–232. DOI:

457   10.1016/S0304-3800(02)00349-6.

458   Archer ME. 1995. Taxonomy, distribution and nesting biology of the *Vespa mandarinia* group

459   (Hym., Vespinae). *Entomologist's Monthly Magazine* 131:47–53.

460   Archer ME. 2008. Taxonomy, distribution and nesting biology of species of the genera *Provespa*

461   Ashmead and *Vespa* Linneaus (Hymenoptera, Vespidae). *Entomologist's Monthly*

462   *Magazine* 144:69–101.

463   Arita HT, Christen JA, Rodríguez P, Soberón J. 2008. Species diversity and distribution in

464   presence-absence matrices: mathematical relationships and biological implications.

465   *American Naturalist* 172:519–532. DOI: 10.1086/590954.

466   Ascher JS, Pickering J. 2020.Discover life bee species guide and world checklist (Hymenoptera:

467   Apoidea: *Anthophila*). *Available at*

468   *https://www.discoverlife.org/mp/20q?guide=Apoidea_species* (accessed July 1, 2020).

469   Ayala R, Gonzalez VH, Engel MS. 2013. Mexican stingless bees (Hymenoptera: Apidae):

470   diversity, distribution, and indigenous knowledge. In: Vit P, Pedro SRM, Roubik D eds.

471   *Pot-Honey: A Legacy of Stingless Bees*. New York, NY: Springer, 135–152. DOI:

472   10.1007/978-1-4614-4960-7_9.

473   Ayton S, Tomlinson S, Phillips RD, Dixon KW, Withers PC. 2016. Phenophysiological variation

474   of a bee that regulates hive humidity, but not hive temperature. *Journal of Experimental*

475   *Biology* 219:1552–1562. DOI: 10.1242/jeb.137588.

476   Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Soberón J,

477   Villalobos F. 2011. The crucial role of the accessible area in ecological niche modeling

478   and species distribution modeling. *Ecological Modelling* 222:1810–1819. DOI:

479   10.1016/j.ecolmodel.2011.02.011.

480   Beggs JR, Brockerhoff EG, Corley JC, Kenis M, Masciocchi M, Muller F, Rome Q, Villemant

481   C. 2011. Ecological effects and management of invasive alien Vespidae. *BioControl*

482   56:505–526. DOI: 10.1007/s10526-011-9389-z.

483   Bérubé C. 2020. Giant alien insect invasion averted Canadian beekeepers thwart apicultural

484   disaster (... or at least the zorn-bee apocalypse). *American Bee Journal*:209–214.

485   Bivand R, Keitt T, Rowlingson B, Pebesma E, Sumner M, Hijmans R, Rouault E, Warmerdam F,

486   Ooms J, Rundel C. 2020a. *rgdal: Bindings for the "geospatial" data abstraction library*.

487   https://cran.r-project.org/web/packages/rgdal/index.html.

488   Bivand R, Rundel C, Pebesma E, Stuetz R, Hufthammer KO, Giraudoux P, Davis M, Santilli S.

489   2020b. *rgeos: Interface to geometry engine - open source ('GEOS')*. https://cran.r-

490   project.org/web/packages/rgeos/index.html.

491   Brown JC, Albrecht C. 2001. The effect of tropical deforestation on stingless bees of the genus

492   *Melipona* (Insecta: Hymenoptera: Apidae: Meliponini) in central Rondonia, Brazil.

493   *Journal of Biogeography* 28:623–634. DOI: 10.1046/j.1365-2699.2001.00583.x.

494   Bujan J, Yanoviak SP, Kaspari M. 2016. Desiccation resistance in tropical insects: causes and

495   mechanisms underlying variability in a Panama ant community. *Ecology and Evolution*

496   6:6282–6291. DOI: 10.1002/ece3.2355.

497   Burdine JD, McCluney KE. 2019. Differential sensitivity of bees to urbanization-driven changes

498   in body temperature and water content. *Scientific Reports* 9:1643. DOI: 10.1038/s41598-

499   018-38338-0.

500   Cameron SA, Lozier JD, Strange JP, Koch JB, Cordes N, Solter LF, Griswold TL. 2011. Patterns

501   of widespread decline in North American bumble bees. *Proceedings of the National*

502   *Academy of Sciences USA* 108:662–667. DOI: 10.1073/pnas.1014743108.

503   Cobos ME, Jiménez L, Nuñez-Penichet C, Romero-Alvarez D, Simões M. 2018. Sample data

504   and training modules for cleaning biodiversity information. *Biodiversity Informatics*

505   13:49–50. DOI: 10.17161/bi.v13i0.7600.

506   Cobos ME, Osorio-Olvera L, Soberón J, Peterson AT, Barve V, Barve N. 2020. *ellipsenm:*

507   *Ecological niche characterizations using ellipsoids*.

508   https://github.com/marlonecobos/ellipsenm.

509   Cobos ME, Peterson AT, Barve N, Osorio-Olvera L. 2019a. kuenm: an R package for detailed

510   development of ecological niche models using Maxent. *PeerJ* 7:e6281. DOI:

511   10.7717/peerj.6281.

512   Cobos ME, Peterson AT, Osorio-Olvera L, Jiménez-García D. 2019b. An exhaustive analysis of

513   heuristic methods for variable selection in ecological niche modeling and species

514   distribution modeling. *Ecological Informatics* 53:100983. DOI:

515   10.1016/j.ecoinf.2019.100983.

516   Couvillon MJ, Wenseleers T, Imperatriz-Fonseca VL, Nogueira-Neto P, Ratnieks FLW. 2007.

517   Comparative study in stingless bees (Meliponini) demonstrates that nest entrance size

518   predicts traffic and defensivity. *Journal of Evolutionary Biology* 21:194–201. DOI:

519   10.1111/j.1420-9101.2007.01457.x.

520   Cueva del Castillo R, Sanabria-Urbán S, Serrano-Meneses MA. 2015. Trade-offs in the evolution

521   of bumblebee colony and body size: a comparative analysis. *Ecology and Evolution*

522   5:3914–3926. DOI: 10.1002/ece3.1659.

523   Dueñas M-A, Ruffhead HJ, Wakefield NH, Roberts PD, Hemming DJ, Diaz-Soltero H. 2018.

524   The role played by invasive species in interactions with endangered and threatened

525   species in the United States: a systematic review. *Biodiversity and Conservation*

526   27:3171–3183. DOI: 10.1007/s10531-018-1595-x.

527   Engler R, Hordijk W, Guisan A. 2012. The MIGCLIM R package – seamless integration of

528   dispersal constraints into projections of species distribution models. *Ecography* 35:872–

529   878. DOI: 10.1111/j.1600-0587.2012.07608.x.

530   Escobar LE, Lira-Noriega A, Medina-Vogel G, Peterson AT. 2014. Potential for spread of the

531   white-nose fungus (*Pseudogymnoascus destructans*) in the Americas: use of Maxent and

532   NicheA to assure strict model transference. *Geospatial Health* 9:221–229. DOI:

533   10.4081/gh.2014.19.

534   Fujiwara A, Sasaki M, Washitani I. 2016. A scientific note on hive entrance smearing in

535   Japanese *Apis cerana* induced by pre-mass attack scouting by the Asian giant hornet

536   *Vespa mandarinia*. *Apidologie* 47:789–791. DOI: 10.1007/s13592-016-0432-z.

537   GBIF.org (07 May 2020) GBIF occurrence download. https://doi.org/10.15468/dl.kzcgc2.

538   Gray L, New A, Science K, Wolfram S. 2003. A mathematician looks at Wolfram's new kind of

539   science. *Notices of the American Mathematical Society 50 (2) (2003) 200–211, URL*

540   *http://www.ams.org/notices/200302/fea-gray.pdf. URL*

541   *http://www.ams.org/notices/200302/fea-gray.pdf* 50:200–211.

542   Hijmans RJ, Etten J van, Sumner M, Cheng J, Bevan A, Bivand R, Busetto L, Canty M, Forrest

543   D, Ghosh A, Golicher D, Gray J, Greenberg JA, Hiemstra P, Hingee K, Geosciences I for

544   MA, Karney C, Mattiuzzi M, Mosher S, Nowosad J, Pebesma E, Lamigueiro OP, Racine

545   EB, Rowlingson B, Shortridge A, Venables B, Wueest R. 2020. *raster: Geographic data*

546   *analysis and modeling*. https://cran.r-project.org/web/packages/raster/index.html.

547   Hinojosa-Díaz IA, Yáñez-Ordóñez O, Chen G, Peterson AT, Engel MS. 2005. The North

548   American invasion of the giant resin bee (Hymenoptera: Megachilidae). *Journal of*

549   *Hymenoptera Research* 14:69–77.

550   Kastberger G, Schmelzer E, Kranner I. 2008. Social waves in Giant Honeybees repel hornets.

551   *PLoS ONE* 3:e3141. DOI: 10.1371/journal.pone.0003141.

552   Klingner R, Richter K, Schmolz E, Keller B. 2005. The role of moisture in the nest

553   thermoregulation of social wasps. *Naturwissenschaften* 92:427–430. DOI:

554   10.1007/s00114-005-0012-y.

555   Li X-D, Liu Z, Zhai Y, Zhao M, Shen H-Y, Li Y, Zhang B, Liu T. 2015. Acute interstitial

556   nephritis following multiple Asian Giant Hornet stings. *American Journal of Case*

557   *Reports* 16:371–373. DOI: 10.12659/AJCR.893734.

558   Matsuura M. 1988. Ecological study on Vespine wasps (Hymenoptera: Vespidae) attacking

559   honeybee colonies: 1. Seasonal changes in the frequency of visits to apiaries by Vespine

560   wasps and damage inflicted, especially in the absence of artificial protection. *Applied*

561   *Entomology and Zoology* 23:428–440.

562   Matsuura M, Sakagami SF. 1973. A bionomic sketch of the Giant Hornet, *Vespa mandarinia*, a

563   serious pest for Japanese apiculture. *Journal of the Faculty of Science, Hokkaido*

564   *University: Series 6. Zoology* 19:125–162.

565   McClenaghan B, Schlaf M, Geddes M, Mazza J, Pitman G, McCallum K, Rawluk S, Hand K,

566   Otis GW. 2019. Behavioral responses of honey bees, *Apis cerana* and *Apis mellifera*, to

567   *Vespa mandarinia* marking and alarm pheromones. *Journal of Apicultural Research*

568   58:141–148. DOI: 10.1080/00218839.2018.1494917.

569   Michener CD. 2000. *The Bees of the World*. Baltimore: Johns Hopkins University Press.

570   Michener CD. 2007. *The Bees of the World*. Baltimore: Johns Hopkins University Press.

571   Monceau K, Bonnard O, Thiéry D. 2014. *Vespa velutina*: a new invasive predator of honeybees

572 in Europe. *Journal of Pest Science* 87:1–16. DOI: 10.1007/s10340-013-0537-3.

573 Nuñez-Penichet C, Cobos ME, Peterson AT, Barve N, Barve V, Gueta T, Soberón J. 2020.

574 *biosurvey: Tools for biological survey planning*. https://github.com/claun02/biosurvey.

575 Ono M, Igarashi T, Ohno E, Sasaki M. 1995. Unusual thermal defense by a honeybee against

576 mass attack by hornets. *Nature* 377:334–336. DOI: 10.1038/377334a0.

577 Osorio-Olvera L, Soberón J. 2020. *bam: Species distribution models in the light of the BAM*

578 *theory*. https://github.com/luismurao/bam.

579 Osorio-Olvera L, Soberón J, Barve V, Barve N, Falconi M. 2020. *ntbox: From getting*

580 *biodiversity data to evaluating species distribution models in a friendly GUI*

581 *environment*. https://github.com/luismurao/ntbox.

582 Owens HL, Campbell LP, Dornak LL, Saupe EE, Barve N, Soberón J, Ingenloff K, Lira-Noriega

583 A, Hensz CM, Myers CE, Peterson AT. 2013. Constraints on interpretation of ecological

584 niche models by limited environmental ranges on calibration areas. *Ecological Modelling*

585 263:10–18. DOI: 10.1016/j.ecolmodel.2013.04.011.

586 Peterson AT, Papeş M, Soberón J. 2008. Rethinking receiver operating characteristic analysis

587 applications in ecological niche modeling. *Ecological Modelling* 213:63–72. DOI:

588 10.1016/j.ecolmodel.2007.11.008.

589 Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME. 2017. Opening the black box: an

590 open-source release of Maxent. *Ecography* 40:887–893. DOI: 10.1111/ecog.03049.

591 Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species

592 geographic distributions. *Ecological Modelling* 190:231–259. DOI:

593 10.1016/j.ecolmodel.2005.03.026.

594 Pimentel D, Zuniga R, Morrison D. 2005. Update on the environmental and economic costs

595   associated with alien-invasive species in the United States. *Ecological Economics*

596   52:273–288. DOI: 10.1016/j.ecolecon.2004.10.002.

597   Pyšek P, Richardson DM. 2010. Invasive species, environmental change and management, and

598   health. *Annual Review of Environment and Resources* 35:25–55. DOI: 10.1146/annurev-

599   environ-033009-095548.

600   Quezada-Euán JJG, Nates-Parra G, Maués MM, Roubik DW, Imperatriz-Fonseca VL. 2018. The

601   economic and cultural values of stingless bees (Hymenoptera: Meliponini) among ethnic

602   groups of tropical America. *Sociobiology* 65:534–557. DOI:

603   10.13102/sociobiology.v65i4.3447.

604   R Core Team. 2019. *R: A language and environment for statistical computing*. Vienna, Austria:

605   R Foundation for Statistical Computing.

606   Schmidt JO, Yamane S, Matsuura M, Starr CK. 1986. Hornet venoms: lethalities and lethal

607   capacities. *Toxicon* 24:950–954. DOI: 10.1016/0041-0101(86)90096-6.

608   Simões M, Romero-Alvarez D, Nuñez-Penichet C, Jiménez L, Cobos ME. 2020. General theory

609   and good practices in ecological niche modeling: a basic guide. *Biodiversity Informatics*

610   15:67–68.

611   Smith-Pardo AH, Carpenter JM, Kimsey L. 2020. The diversity of hornets in the genus *Vespa*

612   (Hymenoptera: Vespidae; Vespinae), their importance and interceptions in the United

613   States. *Insect Systematics and Diversity* 4:1–27.

614   Sugahara M, Nishimura Y, Sakamoto F. 2012. Differences in heat sensitivity between Japanese

615   honeybees and hornets under high carbon dioxide and humidity conditions inside bee

616   balls. *Zoological Science* 29:30–36. DOI: 10.2108/zsj.29.30.

617   Sunday JM, Bates AE, Dulvy NK. 2012. Thermal tolerance and the global redistribution of

618   animals. *Nature Climate Change* 2:686–690. DOI: 10.1038/nclimate1539.

619   Vega GC, Pertierra LR, Olalla-Tárraga MÁ. 2018. MERRAclim, a high-resolution global dataset

620   of remotely sensed bioclimatic variables for ecological modelling. *Scientific Data*

621   4:170078. DOI: 10.1038/sdata.2017.78.

622   Viana JL, Sousa H de AC, Alves RM de O, Pereira DG, Silva Jr. JC, Paixão JF da, Waldschmidt

623   AM, Viana JL, Sousa H de AC, Alves RM de O, Pereira DG, Silva Jr. JC, Paixão JF da,

624   Waldschmidt AM. 2015. Bionomics of *Melipona mondury* Smith 1863 (Hymenoptera:

625   Apidae, Meliponini) in relation to its nesting behavior. *Biota Neotropica* 15. DOI:

626   10.1590/1676-06032015009714.

627   Vilà M, Espinar JL, Hejda M, Hulme PE, Jarošík V, Maron JL, Pergl J, Schaffner U, Sun Y,

628   Pyšek P. 2011. Ecological impacts of invasive alien plants: a meta-analysis of their

629   effects on species, communities and ecosystems. *Ecology Letters* 14:702–708. DOI:

630   10.1111/j.1461-0248.2011.01628.x.

631   Warren DL, Seifert SN. 2011. Ecological niche modeling in Maxent: the importance of model

632   complexity and the performance of model selection criteria. *Ecological Applications*

633   21:335–342. DOI: 10.1890/10-1171.1.

634   Wilcove DS, Rothstein D, Dubow J, Phillips A, Losos E. 1998. Quantifying threats to imperiled

635   species in the United States. *BioScience* 48:607–615. DOI: 10.2307/1313420.

636   Yanagawa Y, Morita K, Sugiura T, Okada Y. 2007. Cutaneous hemorrhage or necrosis findings

637   after *Vespa mandarinia* (wasp) stings may predict the occurrence of multiple organ

638   injury: a case report and review of literature. *Clinical Toxicology* 45:803–807. DOI:

639   10.1080/15563650701664871.

640   Zhu G, Illan JG, Looney C, Crowder DW. 2020. *Assessing the ecological niche and invasion*

641   *potential of the Asian giant hornet. bioRxiv.* DOI: 10.1101/2020.05.25.115311.

642   Zurell D, Franklin J, König C, Bouchet PJ, Dormann CF, Elith J, Fandos G, Feng X,

643   Guillera-Arroita G, Guisan A, Lahoz-Monfort JJ, Leitão PJ, Park DS, Peterson AT,

644   Rapacciuolo G, Schmatz DR, Schröder B, Serra-Diaz JM, Thuiller W, Yates KL,

645   Zimmermann NE, Merow C. 2020. A standard protocol for reporting species distribution

646   models. *Ecography.* DOI: 10.1111/ecog.04960.

647   **Tables**

648   Table 1. Summary of results of ecological niche modeling for *Vespa mandarinia*, including

649   model calibration, model evaluation, and features for models selected after independent testing.

650   The variables included in the sets mentioned on this table can be found in Table 1S-2S. E: free

651   extrapolation; EC: extrapolation and clamping; NE: no extrapolation.

652 **Figures**

653 Figure 1. Hypothesis of accessible areas (**M**) and occurrence records of *Vespa mandarinia* across

654 its native distribution. The three panels represent the occurrences left after cleaning (A) and after

655 applying the two thinning approaches (B and C).

656 Figure 2. Schematic representation of methods used to obtain ecological niche models for *Vespa*

657 *mandarinia*, considering the uncertainty coming from the distinct treatments applied to the data

658 and the variability resulting from different procedures and methodological decisions made during

659 model calibration.

660 Figure 3. Median of potentially suitable areas for *Vespa mandarinia* predicted with free

661 extrapolation for different calibration schemes in the calibration area (left panels) and in North

662 America (right panels).

663 Figure 4. Sum of all suitable areas for *Vespa mandarinia* in North America derived from

664 binarizing (using a 5% threshold) each replicate of selected models (model transfers done with

665 extrapolation) that predicted the known invaded localities of this hornet.

666 Figure 5. Agreement of areas with extrapolation risk for models of *Vespa mandarinia* in North

667 America, separated by calibration schemes.

668 Figure 6. Results from simulations of the potential dynamics of invasion of *Vespa mandarinia* in

669 North America. Dark shades of green show areas that the species reached in a high percentage of

670 scenarios, while light shades of green represent areas reached only rarely by the species. Arrows

671 represent the general path of potential invasion.

672 Figure 7. Representation of potential ecological and economic impacts of an invasion of *Vespa*

673   *mandarinia*. Top panel: honey production (in US dollars) in Mexico and the United States in

674   2016. Bottom panel: species richness of the genera *Bombus* (bumble bees) and *Melipona*

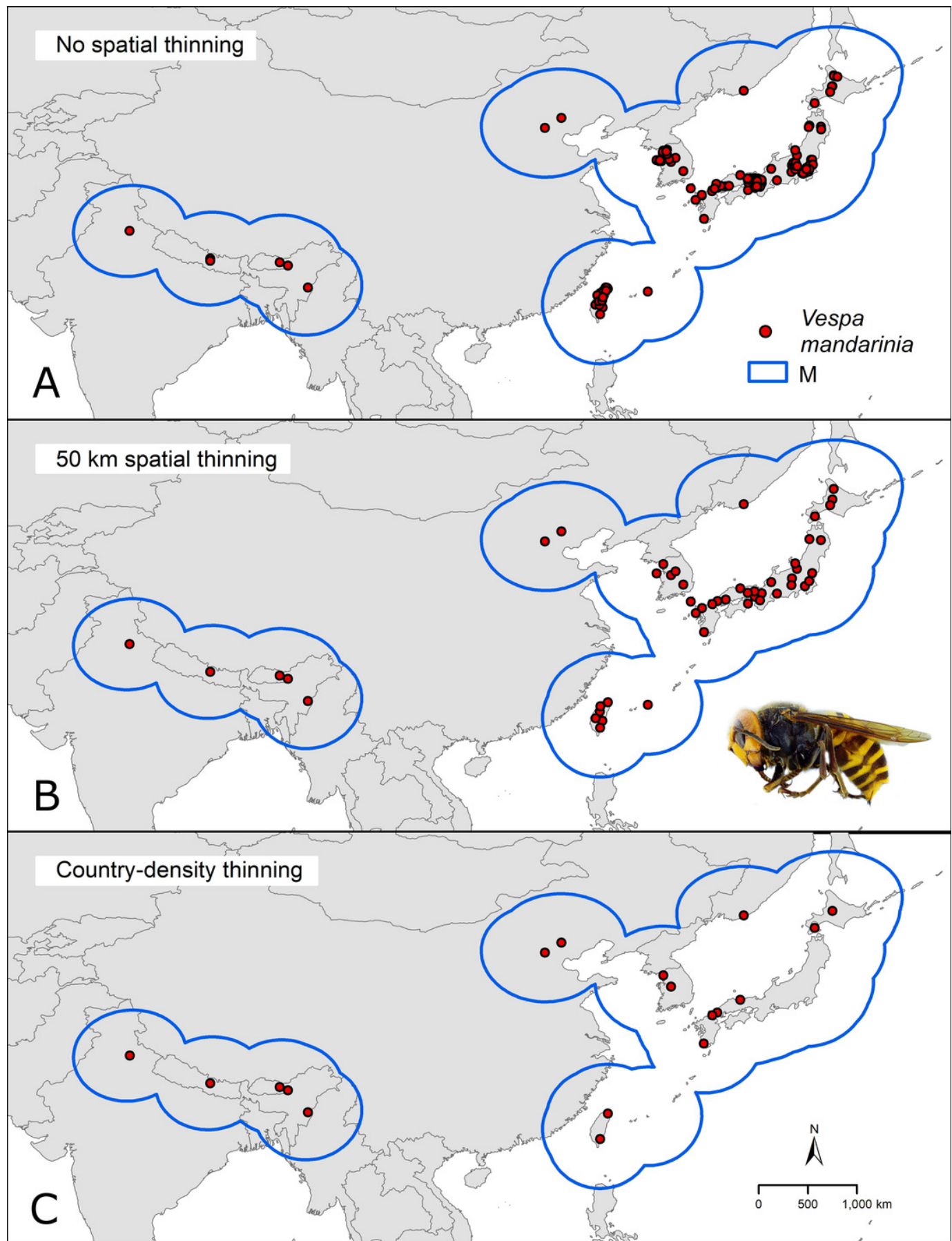
675   (stingless bees) in North America. The area shaded in gray represents the simulated potential

676   invaded area of *Vespa mandarinia* in North America obtained with the 50 km spatial thinning

677   occurrences and PCs as environmental predictors.

Figure 1

Hypothesis of accessible areas (M) and occurrence records of *Vespa mandarinia* across its native distribution.

The three panels represent the occurrences left after cleaning (A) and after applying the two thinning approaches (B and C).

A **No spatial thinning**

● *Vespa mandarinia*
▭ M

B **50 km spatial thinning**

C **Country-density thinning**

N

0   500   1,000 km

Figure 2

Schematic representation of methods used to obtain ecological niche models for *Vespa mandarinia*.

Uncertainty coming from the distinct treatments applied to the data and the variability resulting from different procedures and methodological decisions made during model calibration are considered.

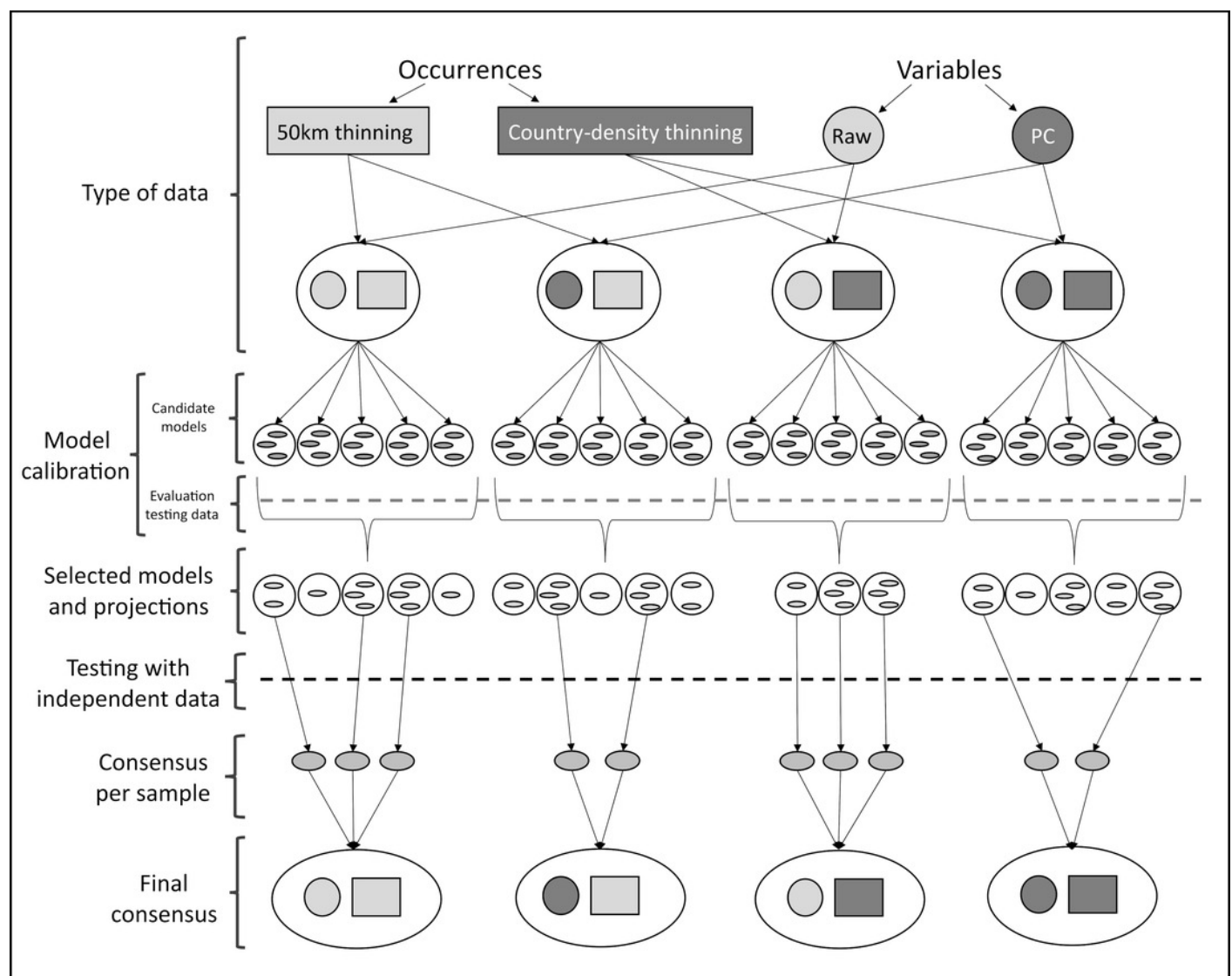



Figure 3

Median of potentially suitable areas for *Vespa mandarinia* predicted with free extrapolation for different calibration schemes in the calibration area (left panels) and in North America (right panels).

50 km spatial thinning - raw variables

Vespa mandarinia
○ native
◉ invaded

Suitability
High
Low

50 km spatial thinning - PCs

Country-density thinning - raw variables

Country-density thinning - PCs

0   1,500 km

0   1,500 km

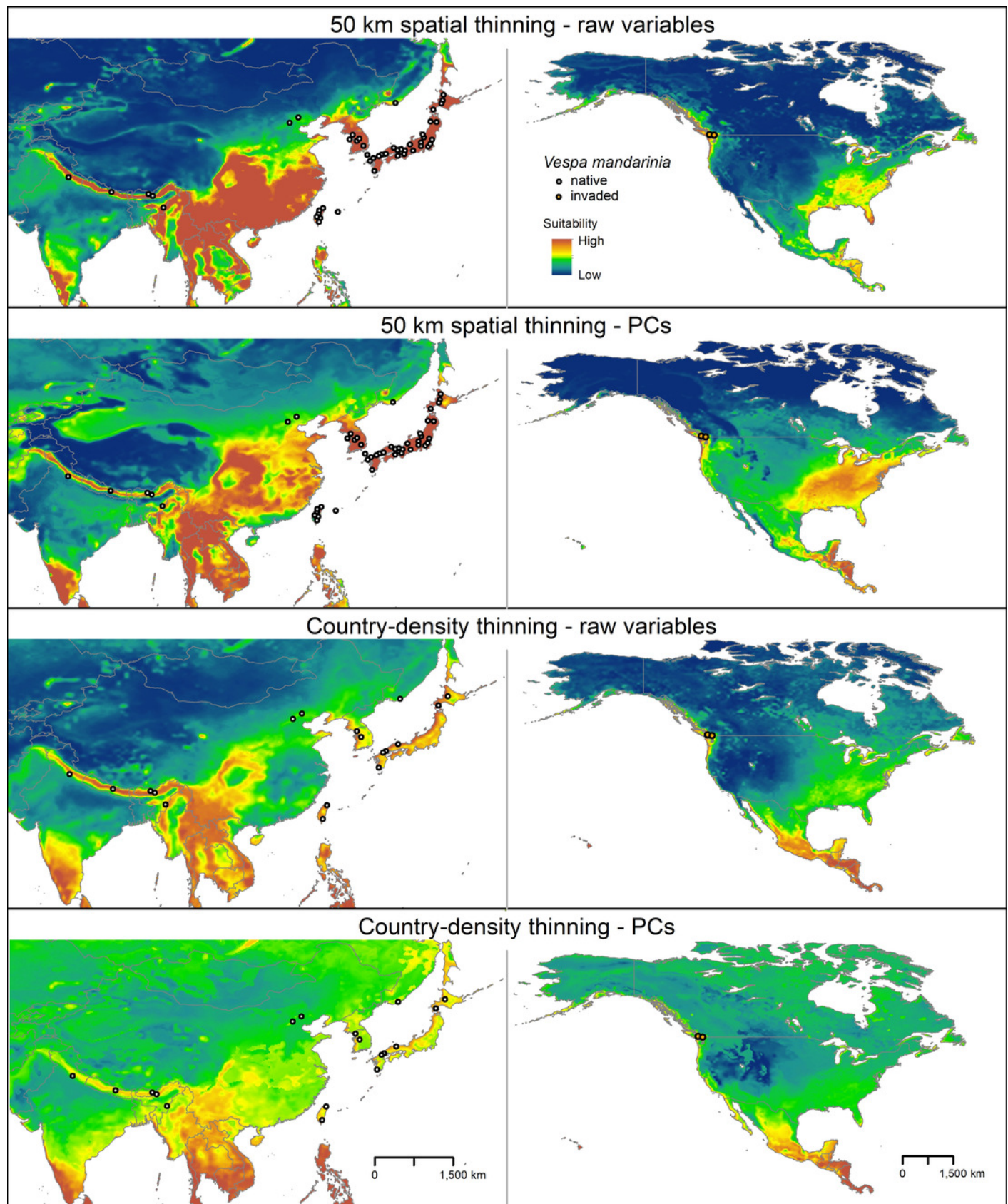



Figure 4

Sum of all suitable areas for *Vespa mandarinia* in North America derived from binarizing (using a 5% threshold) each replicate of selected models (model transfers done with extrapolation) that predicted the known invaded localities of this hornet.

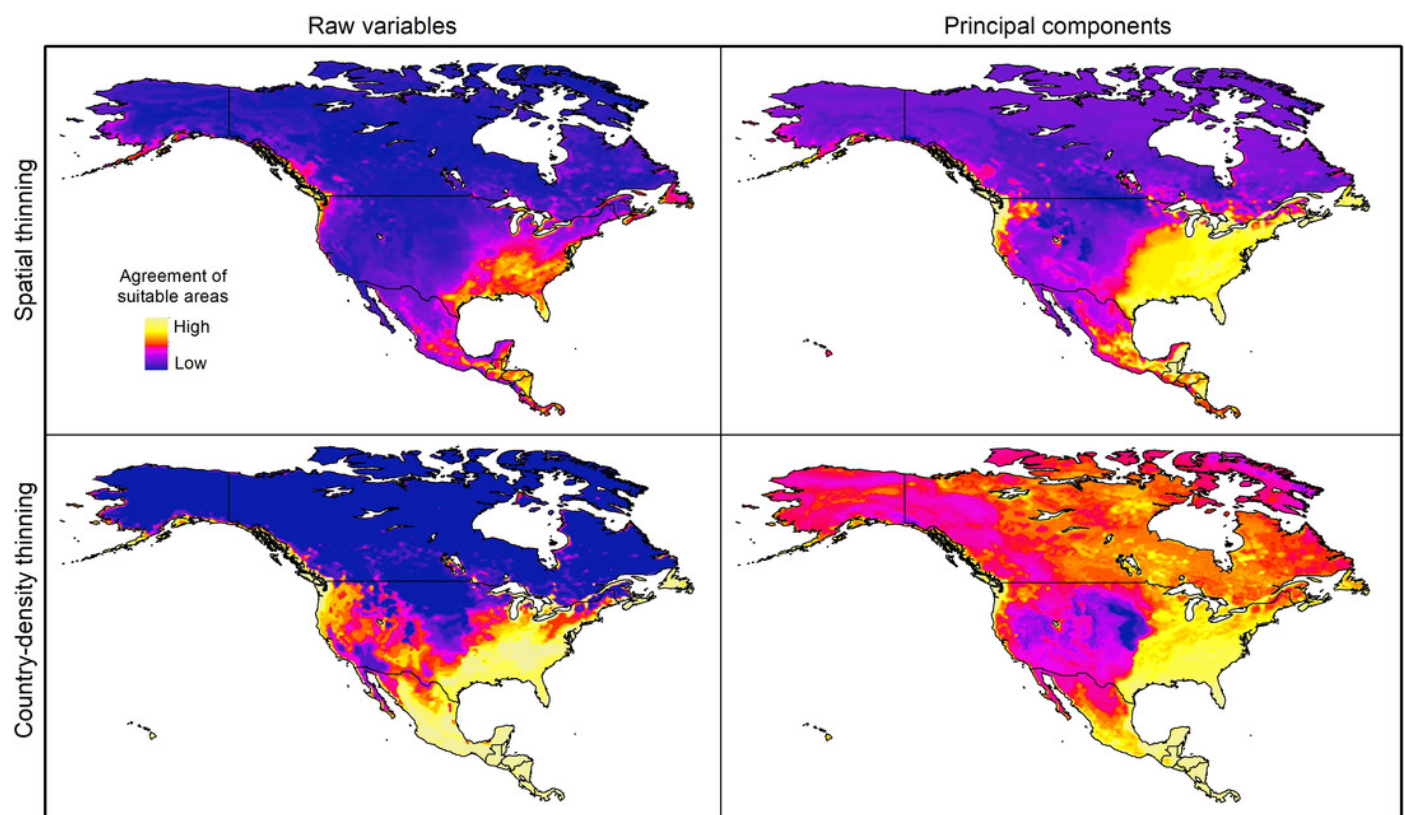



Figure 5

Agreement of areas with extrapolation risk for models of *Vespa mandarinia* in North America, separated by calibration schemes.

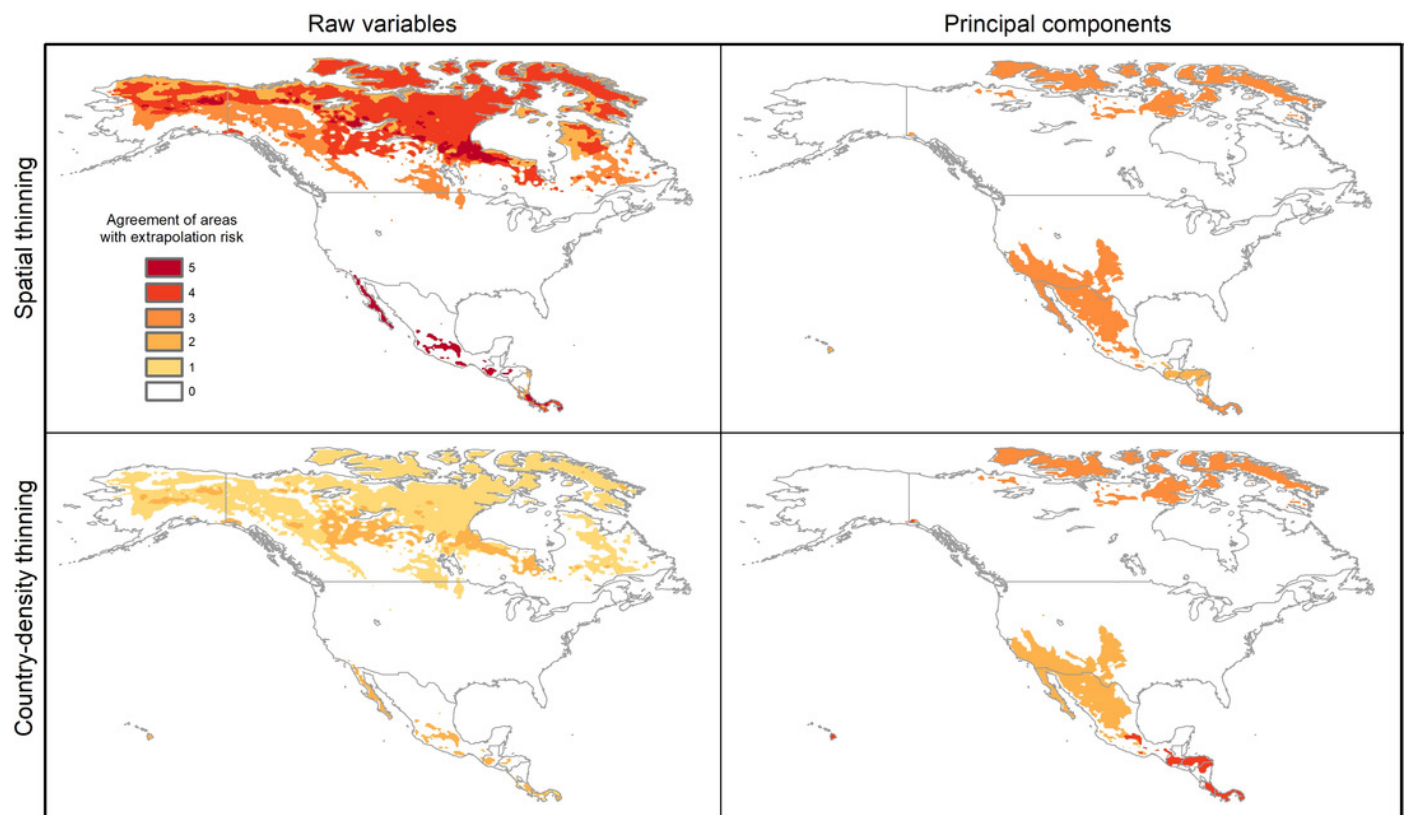



Figure 6

Results from simulations of the potential dynamics of invasion of *Vespa mandarinia* in North America.

Dark shades of green show areas that the species reached in a high percentage of scenarios, while light shades of green represent areas reached only rarely by the species. Arrows represent the general path of potential invasion.

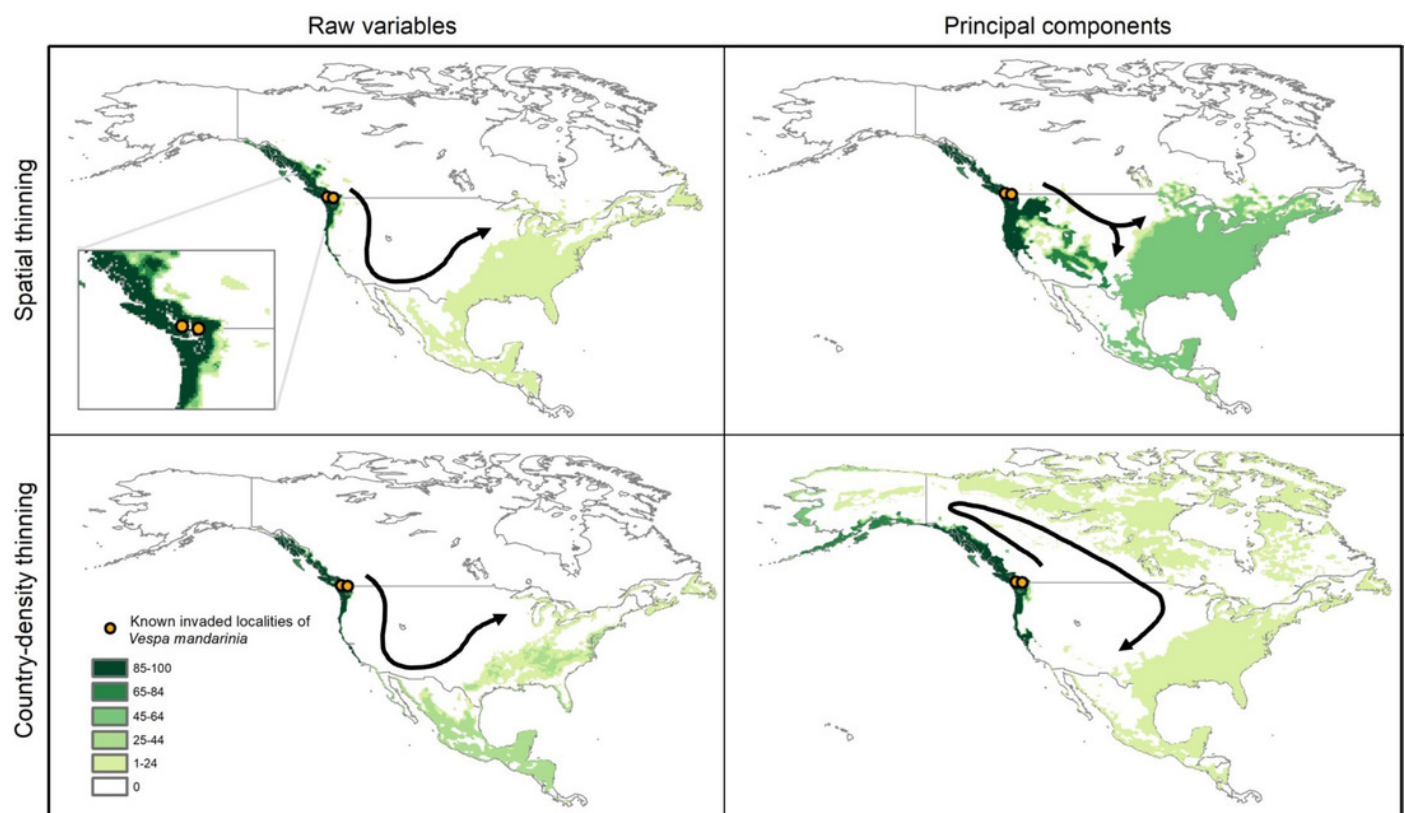
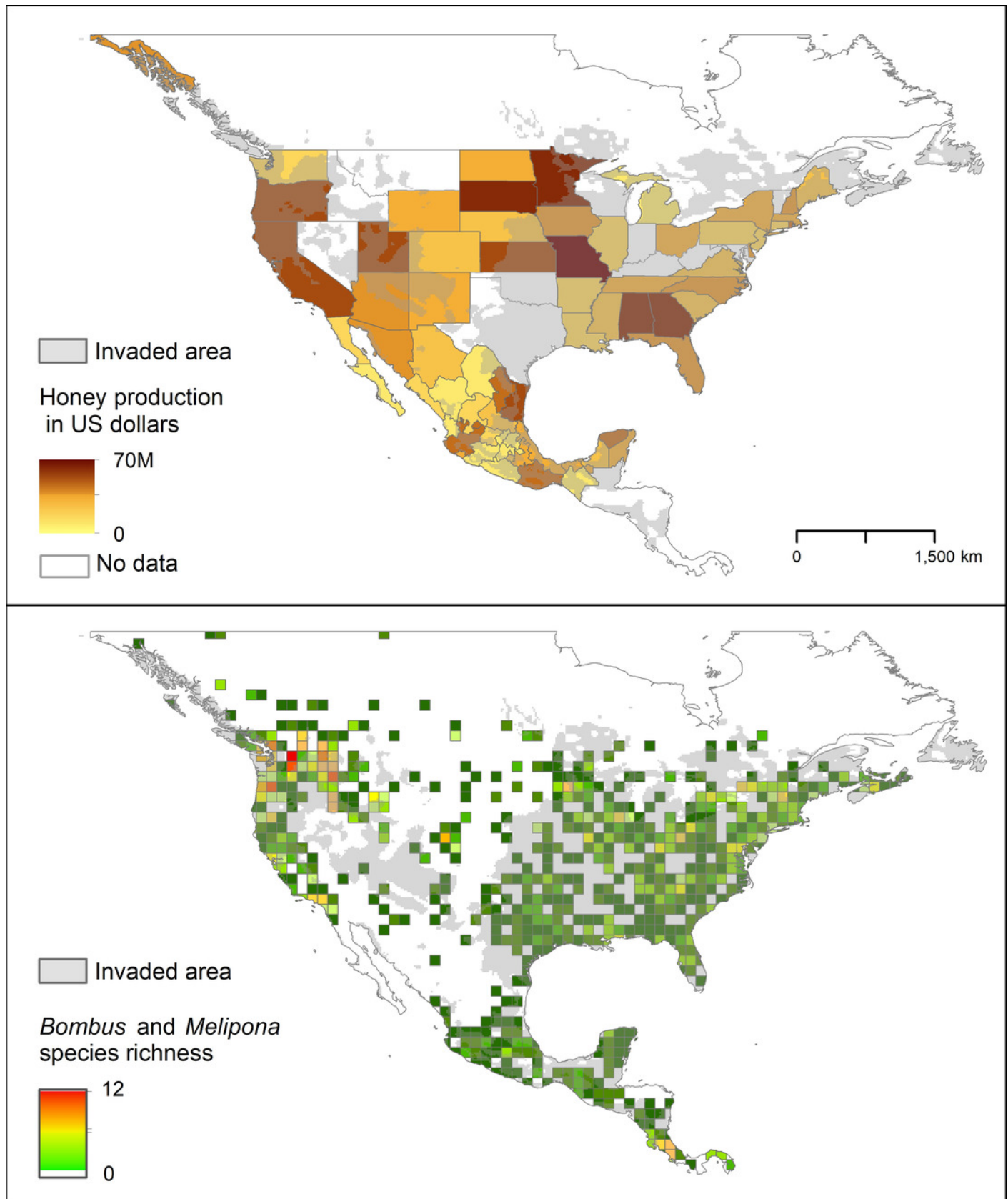



Figure 7

Representation of potential ecological and economic impacts of an invasion of *Vespa mandarinia*.

Top panel: honey production (in US dollars) in Mexico and the United States in 2016. Bottom panel: species richness of the genera *Bombus* (bumble bees) and *Melipona* (stingless bees) in North America. The area shaded in gray represents the simulated potential invaded area of *Vespa mandarinia* in North America obtained with the 50 km spatial thinning occurrences and PCs as environmental predictors.

Summary of results of ecological niche modeling for *Vespa mandarinia*, including model calibration, model evaluation, and features for models selected after independent testing.

The variables included in the sets mentioned on this table can be found in Table 1S-2S. E: free extrapolation; EC: extrapolation and clamping; NE: no extrapolation.

Calibration scheme	Calibration processes	Meeting selection criteria	Passing independent testing (E; EC; NE)	Regularization multiplier	Feature classes	Variable sets
Raw variables and distance thinned occurrences	1	6	8; 2; 10	0.25; 0.5; 0.75	lq; lqpt	42; 43; 50; 51; 57
	2	1	-	-	-	-
	3	1	6; 4; 4	0.75	lqpth	12
	4	1	2; -; 1	0.25	lq	21
	5	2	7; 7; 13	0.1; 0.25	lq	26
PCs and distance thinned occurrences	1	4	24; 18; 20	5	lqph; lqpth	7; 11
	2	2	10; 5; 5	0.25; 0.5	qp	11
	3	4	22; 23; 19	0.1; 0.25; 0.5; 0.75	lp	11
	4	3	9; 8; 11	0.25; 0.5; 0.75	qp	7
	5	6	21; 21; 26	0.1; .25; 0.5; 0.75	lqp	2; 9
Raw variables and country-density thinned occurrences	1	1	4; 4; 6	0.1	lqp	22
	2	2	4; 11; 8	0.1	lq; lqp	5; 22
	3	-	-	-	-	-
	4	3	15; 13; 16	0.1; 2	lq; lqph; lqpth	13; 32
	5	-	-	-	-	-
PCs and country-density thinned	1	7	32; 31; 32	0.25; 0.5; 0.7; 1	lp; lqpt; lqpth	2; 4; 5; 6; 8
	2	1	1; 3; 1	1	lqpth	8
	3	1	5; 6; 7	0.1	q	4
	4	4	24; 22; 18	0.1; 0.25; 0.5; 0.74	lp	1
	5	2	5; 7; 5	1	lqp	1; 6