

UC Davis

UC Davis Previously Published Works

Title

A winner in the Anthropocene: changing host plant distribution explains geographical range expansion in the gulf fritillary butterfly

Permalink

<https://escholarship.org/uc/item/6cg988f5>

Journal

Ecological Entomology, 45(3)

ISSN

0307-6946

Authors

Halsch, Christopher A
Shapiro, Arthur M
Thorne, James H
[et al.](#)

Publication Date

2020-06-01

DOI

10.1111/een.12845

Peer reviewed

1 **A winner in the Anthropocene: changing host plant distribution**
2 **explains geographic range expansion in the gulf fritillary butterfly**

3

4 Christopher A. Halsch¹, Arthur M. Shapiro², James H. Thorne³, David P.
5 Waetjen³, and Matthew L. Forister¹

6

7 ¹ Department of Biology, Program in Ecology, Evolution and Conservation
8 Biology, University of Nevada, Reno, NV, U.S.A.

9 ² Department of Evolution and Ecology, Center for Population Biology,
10 University of California, Davis, CA, U.S.A.

11 ³ Department of Environmental Science and Policy, University of California,
12 Davis, CA, U.S.A.

13

14 Corresponding author

15 Christopher Halsch

16 (415) 246-9157

17 Mail Stop 314, University of Nevada Reno, 1664 N Virginia Street, Reno, NV

18 89557

19 cahalsch@nevada.unr.edu

20 **Abstract:**

- 21 1. The changing climate is altering the geographic distributions of species
22 around the world with consequences for population dynamics, resulting in
23 winners and losers in the Anthropocene.
- 24 2. *Agraulis vanillae*, the gulf fritillary butterfly, has expanded its range in the
25 past one hundred years in the western United States. We combined time
26 series analysis with species distribution modeling to investigate factors
27 limiting the distribution of *A. vanillae* and to predict future shifts under
28 warming scenarios.
- 29 3. In the western US, where we have time series and geographic data, urban
30 development has a positive association with year of colonization (the host
31 plant *Passiflora* is an ornamental in gardens). Colonization was also
32 associated to a lesser extent with winter maximum temperatures, while a
33 negative impact of minimum temperatures and precipitation was apparent
34 on population growth rates. In the eastern US, urban environments play
35 less of a role and the butterfly is primarily limited by minimum
36 temperatures in the winter and host availability later in the season.
- 37 4. Models shows different projections based on region. Eastern U.S.
38 expansion broadly follows the expectation of poleward distributional shifts,
39 especially for the butterfly's maximum distributional extent. Western U.S.
40 expansion is not limited to a single direction and is driven by urban
41 centers becoming more suitable for the host plant.
- 42 5. These results demonstrate the value of combining time series with spatial
43 modeling and incorporating biotic interactions to understand and predict
44 shifting geographic ranges in the Anthropocene.

45 **Keywords:** *Agraulis vanillae*, *Passiflora*, climate change, expansion,
46 distribution, [model fusion](#)

47 **Introduction**

48 The influences of global change, which include invasive species,
49 overexploitation, and climate change, are impacting species around the world
50 (Butchart et al., 2010). We can expect these factors will have varying effects
51 on different species, and that some species will be "winners" under altered
52 conditions (McKinney & Lockwood, 1999). Identifying successful species and
53 the reason for their success in the face of environmental change is important
54 for understanding the potential of individual species and ecosystems to
55 persist and thrive in future climates. In particular, understanding how
56 aspects of global change negatively impact some species, while benefiting
57 others, will improve our ability to predict future species assemblages. One
58 broad method for assessing "winning" and "losing" is by measuring species
59 distributions, which are already shifting in response to recent change (Chen
60 et al., 2011). For some species, ranges are expanding, while for many others
61 ranges are shifting or contracting (Parmesan, 2006). In the context of
62 warming temperatures, distributional change can be caused by direct effects
63 on development and survival (Crozier, 2004) or by indirect effects mediated
64 by biotic interactions (Gutierrez & Thomas, 2001). Ectotherms, including
65 butterflies and other insects, are particularly sensitive to changes in the
66 climate and are exemplar species for the study of these issues (Parmesan et
67 al., 1999; Warren et al., 2001). Here we investigate the gulf fritillary butterfly
68 (*Agraulis vanillae*), which appears to be benefitting from anthropogenic
69 influence and has recently expanded its range in the western United States
70 (Shapiro, 2007). In this study we seek to better understand the drivers
71 underlying this expansion using a combination of spatial occurrence data and
72 long-term population records.

73 *Agraulis vanillae* is a neotropical butterfly associated with riparian and
74 weedy or disturbed habitats (Shapiro, 2009). Over its entire distribution,
75 from temperate North America to temperate South America, there are eight
76 identified sub-species. Previous work has demonstrated genetic divergence
77 between North American and South American lineages (Runquist et al.,
78 2012). In the United States, *A. vanillae* is multi-voltine and in warmer
79 southern regions flies almost all year (Sourakov, 2008). Eastern populations
80 are known to undergo poleward movement each year (Walker, 1991), with
81 sightings as far north as North Dakota and New York (Scott, 1986). The
82 butterfly has a known sensitivity to frost, which can be lethal to all life stages
83 (Shapiro, 2007) and may limit its permanent overwintering distribution. This
84 raises the possibility that the recent expansion of this butterfly is from the
85 direct effect of rising temperature reducing the risk of extinction along the
86 northern range margins in the winter.

87 *Agraulis vanillae* utilizes most plants from the genus *Passiflora* as hosts
88 (May, 1992). The two most common species in the United States are
89 *Passiflora incarnata* and *Passiflora lutea*, both of which grow naturally across
90 much of the southeastern United States (Gremillion, 1989). *Passiflora* prefers
91 well-drained soils and is often found in disturbed sites. In the western United
92 States, *Passiflora* is not present in natural areas and is restricted to modified
93 landscapes and gardens, as various species have been introduced to urban
94 areas as ornamentals (Graves & Shapiro, 2003). We are not aware of any
95 instances where the plant has escaped urban confines and established large
96 self-sustaining populations. Winter freezing temperatures likely limit the
97 distribution of the plant in the wild, however survival can be improved by
98 active management in cultivated populations (McGuire, 1999). *A. vanillae*

99 was first reported in Southern California in 1875 and in San Francisco as early
100 as 1908. It did not permanently establish in San Francisco until 1955, where
101 it used *Passiflora* (Powell, 2000). In the 1960's and 1970's the butterfly
102 briefly established in Sacramento but was extirpated and has only recently
103 reestablished in the region. The human-propagated expansion of *Passiflora*
104 in urban centers offers an alternative biotic explanation for the expansion of
105 the gulf fritillary.

106 In this study, we utilize time series analysis and species distribution
107 modeling to address the following questions. First, using data from a long-
108 term observational study, we ask if climate or urban development better
109 explain the establishment and success of the butterfly in recent years in the
110 Sacramento Valley. Second, using citizen science occurrence data and
111 species distribution modeling, we ask if the current distribution of the
112 butterfly in the continental United States is better explained by host plant or
113 climate limitation and how this varies by region. Finally, we ask if the
114 butterfly is likely to continue to expand its distribution under different climate
115 change scenarios.

116 **Materials and methods**

117 *Sacramento Valley time series data*

118 Long-term observational data were collected every other week by a single
119 observer (AMS) across five sites in the Sacramento Valley. Count data of
120 individual butterflies at these five sites have been collected since 1999 and
121 presence/absence data have been collected since the 1970's or 1980's,
122 depending on the site. Site descriptions and additional details have been
123 reported elsewhere (Forister et al., 2010). *Agraulis vanillae* did not

124 consistently appear at any of these five sites until 2001 and did not appear at
125 every site until 2012. Climate data in California were derived from 270m grid
126 climate maps of monthly and annual values for minimum and maximum
127 temperature and precipitation (Flint & Flint 2012; Flint et al. 2013; Thorne et
128 al. 2015). We extracted the values for grid cells that overlapped with each of
129 the sample sites in the Sacramento Valley and averaged the values for each
130 monthly variable for each year. We calculated seasonal variables by further
131 averaging monthly values to season and converting to water year (the start
132 of September through the end of August).

133 *Sacramento Valley statistical analysis*

134 We approached the analysis of times series data in two phases. First, we
135 used annual presence/absence data to examine colonization, attempting to
136 model the difference between years in which the butterfly was absent across
137 our focal sites and years in which it was resident (spanning 1984 through
138 2018). Residency at a site was determined to be a presence in consecutive
139 years. Random forest regression was used with presence at a site (during
140 years of residency) in a given year as the response variable and year, percent
141 urban land cover (at a county level), seasonal means of minimum
142 temperature, seasonal means of maximum temperature, and seasonal means
143 of precipitation as covariates. A total of 500,000 trees were made with a
144 node size of 5. Variable importance was determined by examining the
145 increased mean squared error of the model when each variable was randomly
146 permuted. The most influential variables identified by random forest analysis
147 were moved forward into a Bayesian hierarchical linear regression. While the
148 random forest is useful for judging the potential importance of a large

149 number of variables, including some that are highly correlated, the Bayesian
150 model allows us to estimate coefficients and associated uncertainty in a
151 hierarchical framework (simultaneously within and across sites). Following a
152 previous model used for data from these study sites (Nice et al, 2019),
153 presence was modeled both at the individual site level and at a higher level
154 across all sites using a Bernoulli distribution. Vaguely informative priors were
155 used for means and variance, with means drawn from normal distribution
156 (mean = 0, sd = 10,000) and variances drawn from a gamma distribution (r
157 = 2, $\lambda = 0.01$). The Bayesian model was comprised of four chains
158 each run for 100,000 iterations with a burn in phase of 50,000 iterations.

159 As a second phase, we examined annual population dynamics post-
160 colonization at the same focal sites, using individual survey count data
161 summarized by year and transformed into population growth rates.
162 Population growth was calculated as the natural log of the current year's total
163 count divided by the previous year's total count (Sibly & Hone, 2002). To
164 determine the most influential climate variables, population growth in a given
165 year was then modeled using a random forest regression. Covariates in the
166 model included year, urban development, abundance in the previous year,
167 seasonal means of minimum monthly temperature, seasonal means of
168 maximum monthly temperature, seasonal means of precipitation, and these
169 same variables lagged by one year to allow in particular for effects mediated
170 through host plants. Again, a total of 500,000 trees with a node size of 5 was
171 used. Variable importance was determined by examining the increased mean
172 squared error of the model following permutation of each variable, and this
173 was done both within and among sites. Like the colonization analysis, the
174 most influential variables identified by random forest analysis were moved

175 forward into a Bayesian hierarchical model in which population growth was
176 modeled both at the individual site level and at a higher level across all sites
177 using a normal distribution. Means of covariates were drawn from a
178 vaguely-informative normal distribution (mean = 0, sd = 10,000) and
179 variances were drawn from a gamma distribution ($r = 2$, $\lambda = 0.01$).
180 This model was comprised of four chains each run for 100,000 iterations with
181 an burn in phase of 50,000 iterations. All analyses were conducted using the
182 randomForest (Liaw & Wiener, 2018) and jagsUI (Kellner, 2019) packages in R
183 Studio.

184 *National data*

185 For US-wide spatial analyses, geo-referenced data points for both *A.*
186 *vanillae* and *Passiflora* were acquired from observations on iNaturalist and
187 GBIF. Additional observations of *Passiflora* were obtained from Calflora and
188 additional observations of *A. vanillae* from the Butterflies and Moths of North
189 America and eButterfly. Only observations since 2000 with a spatial precision
190 higher than 1km were used for analysis. Both *Passiflora* and *A. vanillae* are
191 distinct and identification is likely not a concern, however a random subset of
192 100 observations with photos were checked and all were found to be correct
193 IDs. Current climate data and future projections were obtained from
194 WorldClim (Hijmans et al., 2005). A human population density raster was
195 obtained from the Socioeconomic Data and Applications Center, which used
196 data from the 2010 census (Center for International Earth Science
197 Information Network, 2018). All raster layers were cropped to include only
198 the 48 contiguous states of the USA, although *A. vanillae* is also present in
199 Hawaii as an introduced species. Finally, *A. vanillae* points were separated

200 based on being from the overwintering season, which was defined as
201 between January and March, which is earlier than the earliest observed spring
202 migrant from a study of *A. vanillae* seasonal movement in Florida (Walker,
203 1991).

204 *National statistical analysis*

205 Species distribution models were built for both *Passiflora* and *Agraulis*
206 *vanillae*. All host plant models were built at the genus level, but *Passiflora*
207 species known not to be hosts were excluded. The western and eastern
208 distributions were modeled separately, to allow for the possibility of different
209 factors affecting range limits in the different regions. For all models, we used
210 the MaxEnt algorithm, which models presence only data by comparing
211 observations with random background points. For every model, 10,000
212 random background points were taken within the continental United States.
213 To account for sampling bias in the occurrence data, the random background
214 points were spatially structured using a bias file (Phillips et al., 2009). For
215 *Passiflora*, the bias file was built from all Malpighiales observations (excluding
216 *Passiflora*) and the bias file for *A. vanillae* was built using all Nymphalidae
217 observations (excluding *A. vanillae*). *Passiflora* was modeled using
218 temperature, mean precipitation, and human population density as
219 covariates. Models were built and evaluated using minimum temperature in
220 the coldest month, mean annual temperature, maximum temperature in the
221 warmest month, and both maximum and minimum together as temperature
222 variables. Human population was included in the model to account for any
223 dependence on urban cultivation, which we hypothesized is important in the
224 western United States. The best performing host plant model was later used

225 as a covariate for the butterfly distribution model. For *A. vanillae*, both the
226 overwintering and maximum distributions were modeled. The overwintering
227 distribution was modeled using the best performing *Passiflora* distribution
228 model and temperature variables. The maximum annual distribution was
229 similarly modeled using the *Passiflora* distribution model and temperature as
230 covariates. As with *Passiflora* analyses, various temperature variables were
231 used for model building and comparison, and only the highest performing
232 model for both overwinter and dispersal distributions were used for inference
233 and projection. The models were trained on 70% of the data and tested with
234 the remaining 30%. Model evaluation was performed by examining the AUC
235 scores and omission error rates of both the real model and 1000 permuted
236 null models. Methods and code for null model permutation are described by
237 Bohl et al. (2019), but briefly, observations from the real model are randomly
238 moved around the study area and compared to the real model using the
239 same covariates and testing data. All analyses were performed in R Studio
240 using the dismo package (Hijmans et al., 2013).

241 **Results**

242 *Time Series*

243 For the first twenty-five years of the time series, *Agraulis vanillae* only
244 appeared as an occasional visitor, however beginning in 2001 it became a
245 frequent visitor to all sites across the Sacramento Valley. This rise in the
246 presence of *A. vanillae* occurred during a time of rising temperature and
247 increasing urban development in the area (fig. 1). The random forest model
248 attributed high importance to winter maximum temperatures, percent urban
249 land cover, and year in predicting presence at a site (fig. 2a). Both maximum

250 temperature and urban land cover were increasing over time, especially land
251 cover, which is highly correlated with year (correlation coefficients for year
252 and land cover range from 0.973 in Solano county to 0.989 in Yolo county). In
253 the Bayesian analysis, the model successfully converged (as judged by visual
254 inspection of posterior probability distributions, R_{hat} values, and effective
255 sample size estimates) at both the individual site level and at the higher
256 across site level. Only year was used in the model as it is highly correlated
257 with urbanization (precluding the inclusion of both variables). The Bayesian
258 model confirms that both maximum winter temperatures and year are
259 positively associated with colonization at the higher across site level (fig. 3a).
260 Specifically, the probability that maximum temperature has a greater than
261 zero effect is 0.98 and the probability that year has a greater than zero effect
262 is 0.92. There is a 0.98 probability that year has a stronger effect than
263 winter, thus the positive trend of colonization is not sufficiently explained by
264 climate.

265 For annual population dynamics (represented by the natural log of the
266 current to previous population density), the random forest analysis attributed
267 high importance to abundance in the previous year, winter minimum
268 temperature in the current year, winter precipitation in the current year, and
269 summer precipitation in the current year for predicting population growth (fig.
270 2b, fig. S1). Urbanization, while one of the covariates in the model, was not
271 found to be important for population growth rates. Coefficients in the
272 Bayesian model for population growth converged at both the across site and
273 individual site level. Previous year's abundance, winter minimum
274 temperature, and winter precipitation all had negative effects on population
275 growth. The model is confident in the negative impacts of previous year's

276 abundance, winter minimum temperature, and winter precipitation (fig. 3b).
277 Specifically, the probability that previous year's abundance has a negative
278 effect is 0.84, the probability that winter minimum temperature has a
279 negative effect is 0.80, and the probability that winter precipitation has an
280 effect is 0.88. There does not appear to be a strong effect of summer
281 precipitation in the Bayesian hierarchical regression, despite the importance
282 attributed to it in the random forest. All three variables have approximately
283 equal estimated effect sizes. At the individual site level, there is variation in
284 estimated effects, however negative density dependence is observed at all
285 sites. Winter climate is also important at all sites, however some sites have
286 higher estimated impacts of winter precipitation while others more heavily
287 weight winter minimum temperatures (fig. S4).

288 *Species Distribution Models*

289 The predictors of highest importance of geographic distribution of
290 *Passiflora* vary between the eastern and western United States. In the East,
291 *Passiflora* is best predicted winter minimum temperatures and precipitation
292 while in the West urban population and maximum summer temperatures are
293 the best predictors (Table S1, Table 1). All models achieved high AUC values
294 and performed exceptionally well when compared to permuted null models
295 (Table 1, fig. S5). Under the RCP 4.5 scenario, suitable habitat in the eastern
296 US is predicted to increase along *Passiflora*'s, northern range boundary.
297 Habitat is also predicted to become slightly less suitable along the southern
298 range boundary, however the magnitude of this change in suitability is not
299 comparable to the increase on the poleward margin (fig. 4; fig. 5). In the

300 western U.S.A., current areas of suitability are predicted to expand, but not in
301 a clear poleward direction.

302 The current overwintering ranges of *A. vanillae* in the eastern and western
303 U.S.A. are best explained by both host plant and winter minimum
304 temperatures (Table S1). Like the host plant model, all models performed
305 well in regard to AUC scores and in comparison with permuted null models
306 (Table 1, fig. S5). The variable importance of minimum temperature in the
307 East is slightly greater, however it is not clear if these slight differences in
308 variable importance are meaningful (Table 1). Future climate scenarios
309 project a slight increase in the suitability of some areas in the southeast for
310 overwintering, but not a major expansion (fig.4; fig. 5). The models of
311 maximum annual distribution tell a different story. Models for maximum
312 annual distribution performed best using average temperature, however
313 greater importance in both regions was given to host plant distribution (Table
314 1). Again, models performed well using both the AUC metric and permuted
315 null model comparison (Table 1, fig. S5). This greater importance of the host
316 plant is reflected in the future model predictions, which shows *A. vanillae*
317 expansion into areas that also predict *Passiflora* expansion (fig. 4, fig. 5).
318 Thus, while overwintering gains appear marginal under future warming,
319 expansion of the range during the spring and summer is potentially
320 substantial. Across all models, projections under RCP 8.5 show a slightly
321 greater expansion but do not dramatically vary from RCP 4.5 predictions (fig.
322 S6).

323 **Discussion**

324 Species are currently encountering novel biotic and abiotic conditions, which
325 can positively or negatively impact population dynamics and geographic
326 distributions (McKinney & Lockwood, 1999). Building models that parse these
327 various stressors furthers our understanding of these impacts and allows for
328 better prediction of future assemblages. In this study, we found that years in
329 which the butterfly had colonized our focal sites were characterized by
330 warmer winter maximum monthly temperatures, while winter minimum
331 temperatures had a negative impact on population growth rates in the years
332 after colonization. In particular, if the previous winter was cooler and drier,
333 the butterfly was found in higher abundance the next year. It is possible that
334 the negative impact of winter climate on *A. vanillae* that we have observed is
335 mediated through interactions with host plants or other insects. It could be
336 the case that warmer and wetter winters negatively impact *Passiflora*, but
337 another and perhaps more likely explanation is that wetter and warmer
338 winters increases parasitoid pressure and/or disease leading to reduced adult
339 emergence the following year (Harvell et al., 2002; Stireman et al., 2005). *A.*
340 *vanillae* is known to host nucleopolyhedrovirus (Rodriguez et al., 2011), which
341 could be one mechanism that generated the observed negative density
342 dependence (fig. S1), however this is not known to impact California
343 populations. Finally, at our focal sites there is a slight positive trend over
344 time in winter precipitation and winter minimum temperature (fig. S7),
345 suggesting that if anything the butterfly is persisting and expanding in the
346 Sacramento Valley despite climate, not because of it.

347 The local impact of climate on the population dynamics of *Agraulis*
348 *vanillae* in the Sacramento Valley also has implications for explaining the
349 limiting factors of its current distribution in the western United States.

350 Distribution models of *A. vanillae* in the east and west place high importance
351 on the distribution of the host plant, however only the western host plant
352 model identified human population density as an important predictor (after
353 accounting for sampling bias in the data). One explanation for the recent
354 colonization of the area by the butterfly is thus the increasing urbanization of
355 the Sacramento Valley. Over the past twenty years the suburbs of
356 Sacramento have expanded at a steady rate (Forister et al., 2010), which has
357 likely resulted in an increase in *Passiflora* in the region. Random forest
358 analysis ranked urban land cover over any climate variable when predicting
359 colonization and the Bayesian model found a much greater effect of year
360 (which is highly correlated with urbanization). In the eastern United States,
361 the impacts of temperature, specifically minimum temperatures, are
362 apparent in geographic distribution models. In the east, the distribution of
363 *Passiflora* extends further north in the winter compared to *A. vanillae*, while
364 in the west the overwintering distribution closely resembles that of *Passiflora*.
365 Once the weather warms in the east, the butterflies can then expand to cover
366 the distribution of the host plant. Thus, while minimum temperature plays an
367 important role in the overwintering locations of the eastern gulf fritillary, its
368 maximum extent appears to be host plant limited.

369 Although all analyses involve a single focal species, an interesting result of
370 our work is the discovery that variation in limiting factors between the east
371 and west result in quite different predictions for distributional change under
372 future climates by season and by region. In the eastern U.S., models using
373 the RCP 4.5 and 8.5 climate scenarios broadly follow the expectation of
374 poleward movement, with more suitable habitat along the northern range
375 margin and a slight reduction in habitat suitability in south. In the winter the

376 butterfly is limited by temperature and predicted expansion during this time
377 will largely be due to increasing temperatures. Later in the season the
378 butterfly is primarily limited by the distribution of the host plant and this
379 expansion would be better explained by an indirect effect of temperature
380 mediated the distribution of the host plant. In the western U.S., expansion is
381 also predicted, but not in a single direction. This region is much more
382 climactically and topographically complex and this this result is perhaps not
383 surprising. A recognizable pattern is the importance of population centers,
384 especially in the expansion of the maximum annual distribution of the
385 butterfly. It is important to note that our future projections were created
386 using climate forecasts, but not human population forecasts. This means that
387 there is an underlying assumption in the projection that population density
388 will remain the same, which almost certainly will not be met. Given the
389 predictive power of population in the U.S. west in our models, we suggest
390 that this that these projections are conservative. Overall, newly suitable
391 areas for the butterfly closely follow the newly suitable areas for the host
392 plant, thus we infer that expansion in the West is more closely tied to the
393 indirect effect of host plant expansion.

394 These findings add to the literature stressing the utility of accounting for
395 biotic interactions species distribution modeling and forecasting. Biotic
396 interactions are an important factor in shaping the distributions of species but
397 have been incorporated into few studies examining climate change (Araujo &
398 Luoto, 2007; Heikkinen et al., 2007; Preston et al., 2008; Schweiger et al.,
399 2008), at least relative to abiotic-only distribution models. Many of the
400 studies that do incorporate biotic information demonstrate that, whether the
401 biotic element be a host plant or a mutualist, model performance is

402 improved. Similarly, we show that the host plant has high predictive
403 importance and allows for a better understanding of the current distributional
404 limits of the butterfly. Another important component of these results is the
405 observed within-distribution variation, as we show the importance of host
406 plant varying by season and region. Recognizing and accounting for this
407 variation is critical in order to better predict future responses to change,
408 especially for species with large spatial distributions (Murphy & Lovett-Doust,
409 2006; O'Neill et al., 2008). By incorporating both a key host plant interaction
410 and allowing it to vary by region, we have a more complete understanding of
411 this observed expansion.

412 The gulf fritillary is a notable example of a “winner” in the Anthropocene.
413 While insects are declining on a large scale (Hallmann et al., 2017; Lister &
414 Garcia, 2018; Salcido et al., 2019; Sanchez-Bayo & Wyckhuys, 2019;
415 Wepprich et al., 2019), altered conditions create opportunities for some to
416 prevail. The nuances of each success story are different; but it is clear that
417 increasing temperature is playing a vital role in facilitating the distributional
418 expansion of many of these insect winners. Other studies have shown that
419 rising temperature can impact insect distributions by increasing
420 overwintering survival along a northern range margin (Crozier, 2004), by
421 increasing access to food resources (Raffa et al., 2013), or by increasing diet
422 breadth (Pateman et al., 2012). As temperatures continue to warm, insects
423 will continue to be prime candidates for temperature-driven distributional
424 change, for better or for worse. Continuing to observe these phenomena and
425 developing methods by which to understand them is critical. Here the
426 combination of long-term time series data and large-scale citizen science
427 spatial data allowed for a detailed examination of the underlying causes for

428 such an expansion. As these types of data continue to become more widely
429 accessible, the common themes behind insect distributional change in the
430 Anthropocene will continue to become more apparent.

431

432 **Acknowledgements**

433 We thank Ken Nussear for discussion about the distribution models. Data
434 were provided by the Butterfly and Moth Information Network and the many
435 participants who contribute to its Butterflies and Moths of North America
436 project. Data were also provided by iNaturalist, GBIF, eButterfly, and Calflora.
437 MLF was supported by a Trevor James McMinn professorship.

438

439 **Contribution of authors**

440 A.M.S. collected the Sacramento Valley observational data. J.H.T. and D.P.W.
441 provided the climate data. C.A.H. conducted the species distribution
442 analyses. C.A.H. and M.L.F. conducted time series analysis. C.A.H. and M.L.F.
443 wrote the manuscript with input from co-authors.

444

445 **References**

446 Araujo M.B & Luoto, M. (2007) The importance of biotic interactions for
447 modelling species distributions under climate change. *Journal of Global*
448 *Ecology and Biogeography*, **16**, 743-753.

449 Bohl, C.L., Kass, J.M. & Anderson, R.P. (2019) A new null model approach to
450 quantify performance and significance for ecological niche models of species
451 distributions. *Journal of Biogeography*, **46**, 1101-1111.

452 Butchart, S.H.M., Walpole, M., Collen, B., Strien, A.V., Scharlemann, J.P.W.,
453 Almond, R.E.A., Baillie, J.E.M., Bomhard, B., Browxn, C., Bruno, J., Carpenter,
454 K.E., Carr, G.M., Chanson, J., Chenery, A.M., Csirke, J., Davidson, N.C.,
455 Dentener, F., Foster, M., Galli, A., Galloway, J.N., Genovesi, P., Gregory, R.D.,
456 Hockings, M., Kapos, V., Lamarque, J., Leverington, F., Loh, J., McGeoch, M.A.,
457 McRae, L., Minasyan, A., Hernández Morcillo, M., Oldfield, T.E.E., Pauly, D.,
458 Quader, S., Revenga, C., Sauer, J.R., Skolnik, B., Spear, D., Stanwell-Smith, D.,
459 Stuart, S.N., Symes, A., Tierney, M., Tyrrell, T.D., Vié, J. and R. Watson. 2010.
460 Global Biodiversity: Indicators of Recent Declines. *Science*, **328**, 1164-1168.

461

462 Calflora: Information on California plants for education, research and
463 conservation.

464 [web application] (2014) Berkeley, California: The Calflora Database [a non-
465 profit organization].

466 Available: <https://www.calflora.org/> (Accessed: November 5, 2019).

467

468 Center for International Earth Science Information Network – CIESIN –
469 Columbia University. (2018) Gridded Population of the World, Version 4
470 (GPWv4): Population Density, Revision 11. Palisades, NY: NASA

471 Socioeconomic Data and Application Center (SEDAC). [https://doi.org/10.7927/](https://doi.org/10.7927/H49C6VHW)
472 [H49C6VHW](https://doi.org/10.7927/H49C6VHW). (Accessed May 5, 2019),

473

- 474 Chen, I.C., Hill, J.K., Ohlemuller, R., Roy, D.B. & Thomas, C.D. (2011) Rapid
475 Range Shifts of Species Associated with High Levels of Climate Warming.
476 *Science*, **333**, 1024-1026.
- 477 Crozier, L. (2004) Warmer winters drive butterfly range expansion by
478 increasing survivorship. *Ecology*, **85**, 231-241.
- 479 Flint, L. E., & Flint, A.L. (2012) Downscaling future climate scenarios to fine
480 scales for hydrologic and ecological modeling and analysis. *Ecological*
481 *Processes* 1:1-15.
- 482 Flint, L. E., Flint A.L., Thorne J. H. & Boynton, R. (2013) Fine-scale hydrologic
483 modeling for regional landscape applications: the California Basin
484 Characterization Model development and performance. *Ecological Processes*
485 2:1-21.
- 486 Forister, M.L., McCall, A.C., Sanders, N.J., Fordyce, J.A., Thorne, J.H., O'Brien, J.,
487 Waetjen, D.P. & Shapiro, A.M. (2010) Compounded effects of climate change
488 and habitat alteration shift patterns of butterfly diversity. *Proceedings of the*
489 *National Academy of Sciences*, **107**, 2088-2092.
- 490 GBIF.org (3rd May 2019) GBIF Occurrence Download [https://doi.org/10.15468/](https://doi.org/10.15468/dl.twov45)
491 [dl.twov45](https://doi.org/10.15468/dl.twov45)
- 492 Graves, S.D. & Shapiro, A.M. (2003) Exotics as host plants of the California
493 butterfly fauna. *Biological Conservation*, **110**, 413-433.
- 494 Gremillion, K.J. The Development of a mutualistic relationship between
495 humans and maypops (*Passiflora incarnata* L.) in the Southeastern United
496 States. *Journal of Ethnobiology*, **9**, 135-155.
- 497 Gutierrez D. & Thomas, C.D. (2001) Marginal range expansion in a host-
498 limited butterfly species *Gonepterys rhamni*. *Ecological Entomology*, **25**, 165-
499 170.
- 500 Hallmann, C.A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H.,
501 Stenmans, W., Muller, A., Surnser, H., Horren, T., Goulson, D. & de Kroon, H.
502 (2017) More than 75 percent decline over 27 years in total flying insect
503 biomass in protected areas. *PLoS One*, **12**, e0185809.
- 504
- 505 Harvell, C.D., Mitchell, C.E., Ward, J.R., Altizer, S., Dobson, A.P., Ostfeld, R.S.
506 & Samuel, M.D. (2002) Climate Warming and Disease Risks for Terrestrial and
507 Marine Biota. *Science*, **296**, 2158 - 2162.
- 508
- 509 Heikkinen R.K., Luoto, M., Virkkala, R., Pearson, R.G. & Korber J. (2007) Biotic
510 interactions improve prediction of boreal bird distributions at macro-scales.
511 *Global Ecology and Biogeography*, **16**, 754-763.
- 512
- 513 Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very
514 high resolution interpolated climate surfaces for global land areas.
515 *International Journal of Climatology* 25: 1965-1978.
- 516
- 517 Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2013) dismo: Species
518 distribution modeling. *R package version 0.8-17*.
- 519

- 520 iNaturalist. Available from <https://www.inaturalist.org>. Accessed [2019-11-
521 05].
- 522 Kellner, K. (2019) jagsUI: A Wrapper Around 'rjags' to Streamline 'JAGS'
523 Analyses. *R package version 1.5.1*.
- 524
525 Larrivee, M., Prudic, K.L., McFarland, K.P., & Kerr, J. (2018) eButterfly: a
526 citizen-based butterfly database in the biological sciences. [http://www.e-
527 butterfly.org](http://www.e-butterfly.org)
- 528 Liaw, M.A. & Wiener, M. (2018) randomForest: Breiman and Cutler's Random
529 Forests for Classification and Regression. *R package version 4.6-14*.
- 530
531 Lister, B.C. & Garcia, A. (2018) Climate-driven declines in arthropod
532 abundance restructure a rainforest food web. *Proceedings of the National
533 Academy of Sciences*, **115**, 10397-10406.
- 534
535 Lotts, K. & Naberhaus, T coordinators. (2017) Butterflies and Moths of North
536 America. Data set accessed (or exported) 2019-06-13
537 at <http://www.butterfliesandmoths.org/>.
- 538
539 May, P.G. (1992) Flower Selection and the Dynamics of Lipid Reserve in Two
540 Nectarivorous Butterflies. *Ecology*, **73**, 2181-2191.
- 541 McGuire, M. (1999) *Passiflora incarnata* (Passifloraceae): A New Fruit Crop.
542 *Economic Botany*, **53**, 161-167.
- 543 McKinney M.L. & Lockwood J.L. (1999) Biotic homogenization: a few winners
544 replacing many losers in the next mass extinction. *Trends in Ecology and
545 Evolution*, **14**, 450-453.
- 546 Murphy, H.T. & Lovett-Doust, J. (2007) Accounting for regional niche variation
547 in habitat suitability models. *Oikos* **116**, 99-110.
- 548 Nice, C.C., Forister, M.L., Harrison, J.G., Gompert, Z., Fordyce, J.A., Thorne,
549 J.H., Waetjen, D.P. & Shapiro, A.M. (2019) Extreme heterogeneity of
550 population response to climatic variation and the limits of prediction. *Global
551 Change Biology*, **2019**,1-10.
- 552 O'Neill, G.A., Hamann, A. & Wang, T. (2008) Accounting for population
553 variation improves estimated of the impact of climate change on species'
554 growth and distribution. *Journal of Applied Ecology*, **45**, 1040-1049.
- 555 Parmesan, C. (2006) Ecological and Evolutionary Responses to Recent
556 Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, **37**,
557 637-669.
- 558 Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J.K., Thomas, C.D., Descimon,
559 H., Huntley, B., Kaila, L., Kullberg, J., Tammaru, T., Tennent, W.J., Thomas, J.A.
560 & Warren, M. (1999) Poleward shifts in geographical ranges of butterfly
561 species associated with regional warming. *Nature*, **339**, 579-583.
- 562 Pateman, R.M., Hill, J.K., Roy, D.B., Fox, R. & Thomas, C.D. (2012)
563 Temperature-Dependent Alterations in Host Use Drive Rapid Range
564 Expansion in a Butterfly. *Science*, **336**, 1028-1030.

- 565
566 Phillips, S.J., Dudik, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. &
567 Ferrier, S. (2009) Sample selection bias and presence-only distribution
568 models: implications for background and pseudo-absence data. *Ecological*
569 *Applications*, **19**, 181-197.
- 570 Powell, J.A., Russell, P., Russell, S. & Sperling, F.A.H. (2000) Northward
571 expansion of two mint-feeding species of *Pyrausta* in California (Lepidoptera:
572 Pyraloidea: Crambidae). *Holarctic Lepidoptera*, **7**, 55-58.
- 573 Preston, K.L., Rotenberry, J.T., Redak, R.A. & Allen, M.F. (2008) Habitat shifts of
574 endangered species under altered climate conditions: importance of biotic
575 interactions. *Global Change Biology*, **14**, 2501-2515.
- 576 R Development Core Team. (2013) R: A language and environment for
577 statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
578 [WWW document]. URL <http://www.R-project.org/> [accessed on 05 November
579 2019].
580
- 581 Raffa, K.F., Powell, E.N. & Townsend, P.A. (2013) Temperature-driven range
582 expansion of an irruptive insect heightened by weakly coevolved plant
583 defenses. *Proceedings of the National Academy of Sciences*, **110**, 2193-2198.
- 584 Rodriguez, V.A., Belaich, M.N., Gomez, D.L.M., Sciocco-Cap, A. & Ghiringhelli,
585 P.D. (2011) Identification of nucleopolyhedrovirus that infect Nymphalid
586 butterflies *Agraulis vanillae* and *Dione juno*. *Journal of Invertebrate*
587 *Pathology*, **106**, 255-262.
- 588 Runquist, E.B., Forister, M.L. & Shapiro, A.M. (2012) Phylogeography at large
589 spatial scales: incongruent patterns of population structure and demography
590 of Pan-American butterflies associated with weedy habitats. *Journal of*
591 *Biogeography*, **39**, 382-396.
- 592 Salcido, D. M., Forister, M., Lopez, H. G. & Dyer, L. A. (2019) Ecosystem
593 services at risk from declining taxonomic and interaction diversity in a
594 tropical forest. bioRxiv, 631028.
- 595 Sanchez-Bayo, F. & Wyckhuys, K.A.G. (2019) Worldwide decline of the
596 entomofauna: A review of its drivers. *Biological Conservation*, **232**, 8-27.
- 597 Scott, J.A. **1986**. The butterflies of North America: A natural history and field
598 guide.
- 599 Shapiro, A.M. (2009) The Neo-Riparian butterfly fauna of western Argentina.
600 *Journal of Research on the Lepidoptera*, **41**, 24-30.
601
- 602 Shapiro, A.M. & Manolis, T.D. (2007) Field Guide to Butterflies of the San
603 Francisco Bay and Sacramento Valley Regions.
604
- 605 Schweiger, O., Settele, J., Kudrna, O., Klotz, S. & Kuhn, I. (2008) Climate
606 change can cause spatial mismatch of trophically interacting species.
607 *Ecology*, **89**, 3472-3479.
608
- 609 Sibly, R.M. & Hone, J. (2002) Population growth rate and its determinants: an
610 overview. *Philosophical Transactions of the Royal Society B*, **357**, 1153-1170.
611

- 612 Sourakov, A. (2008) Notes on the biology of the gulf fritillary *Agraulis vanillae*
613 (Lepidoptera: Nymphalidae), in North-Central Florida. *Journal of the*
614 *Lepidopterists' Society*, **63**, 127.
615
- 616 Stireman III, J.O., Dyer, L.A., Janzen, D.H., Singer, M.S., Lill, J.T., Marquis, R.J.,
617 Ricklefs, R.E., Gentry, G.L., Hallwachs, W., Coley, P.D., Barone, J.A., Greeney,
618 H.F., Connahs, H., Barbosa, P., Morais, H.C. & Diniz, I.R. (2005) Climatic
619 unpredictability and parasitism of caterpillars: Implications of global warming.
620 *Proceedings of the National Academy of Sciences*, **102**, 17384-17387.
621
- 622 Thorne, J.H., Boynton, R.M., Flint, L.E. & A.L. Flint (2015) Comparing historic
623 and future climate and hydrology for California's watersheds using the Basin
624 Characterization Model. *Ecosphere* 6(2). Online.
625
- 626 Walker, T.J. (1991) Butterfly migration from and to peninsular Florida.
627 *Ecological Entomology*, **16**, 241-252.
628
- 629 Warren, M.S., Hill, J.K., Thomas, J.A., Asher, J., Fox, R., Huntley, B., Roy, D.B.,
630 Telfer, M.G., Jeffcoate, S., Harding, P., Willis, S.G., Greatorex-Davies, J.N.,
631 Moss, D., Thomas, C.D. (2001) Rapid responses of British butterflies to
632 opposing forces of climate and habitat change. *Nature*, **414**. 65-69.
633
- 634 Wepprich, T., Adrion, J.R., Ries, L., Wiedmann, J. & Haddad, N.M. (2019)
635 Butterfly abundance declines over 20 years of systematic monitoring in Ohio,
636 USA. *Plos ONE*, **14**, e0216270.
637

- 638 Figure 1. (a) Change in detection probability (the ratio of days observed to
639 total visits) over time across all sites. (b) Annual ratio of urban land cover to
640 total land cover at a county level for the three counties containing long-term
641 study sites: North Sacramento and Rancho Cordova are in Sacramento
642 County; Suisun Marsh and Gates Canyon are in Solano County. (c) Mean
643 monthly maximum winter temperature over time.
- 644 Figure 2. (a) Variable importance of model covariates in predicting the
645 presence of *A. vanillae* at a site in the Sacramento Valley over time. (b)
646 Variable importance of model covariates in predicting the annual population
647 growth after establishment.
- 648 Figure 3. Bayesian posterior distributions for important coefficients (as
649 determined by random forest). Y-axis shows scaled coefficient estimates. (a)
650 Estimates of coefficients for establishment. (b) Estimates of coefficients for
651 population growth.
- 652 Figure 4. Current distribution of suitability for (a) *Passiflora* in the West. (b)
653 *Passiflora* in the East. (c) Overwintering *A. vanillae* in the West. (d)
654 Overwintering *A. vanillae* in the East. (e) Seasonal *A. vanillae* in the West. (f)
655 Seasonal *A. vanillae* in the East.
- 656 Figure 5. The expanding gulf fritness landscape. Predicted change in
657 suitability in 2050 under RCP 4.5 for (a) *Passiflora* in the West. (b) *Passiflora*
658 in the East. (c) Overwintering *A. vanillae* in the West. (d) Overwintering *A.*
659 *vanillae* in the East. (e) Seasonal *A. vanillae* in the West. (f) Seasonal *A.*
660 *vanillae* in the East.

661
 662
 663
 664
 665
 666
 667

668 Table 1. Variable importance and model fit of all species distribution models.
 669 Rows represent different regional models and columns are the different variables
 670 in the model. AUC (area under the curve) is the performance metric of model
 671 fit.

Host Plant Distribution Model							
Region	Max. Temp.	Min. Temp.	Population	Precipitation	AUC	OR	P-value
East	10.3	62.9	3.1	23.7	0.822	0.076	<< 0.05
West	25.2	16.8	42.2	15.8	0.830	0.125	<< 0.05

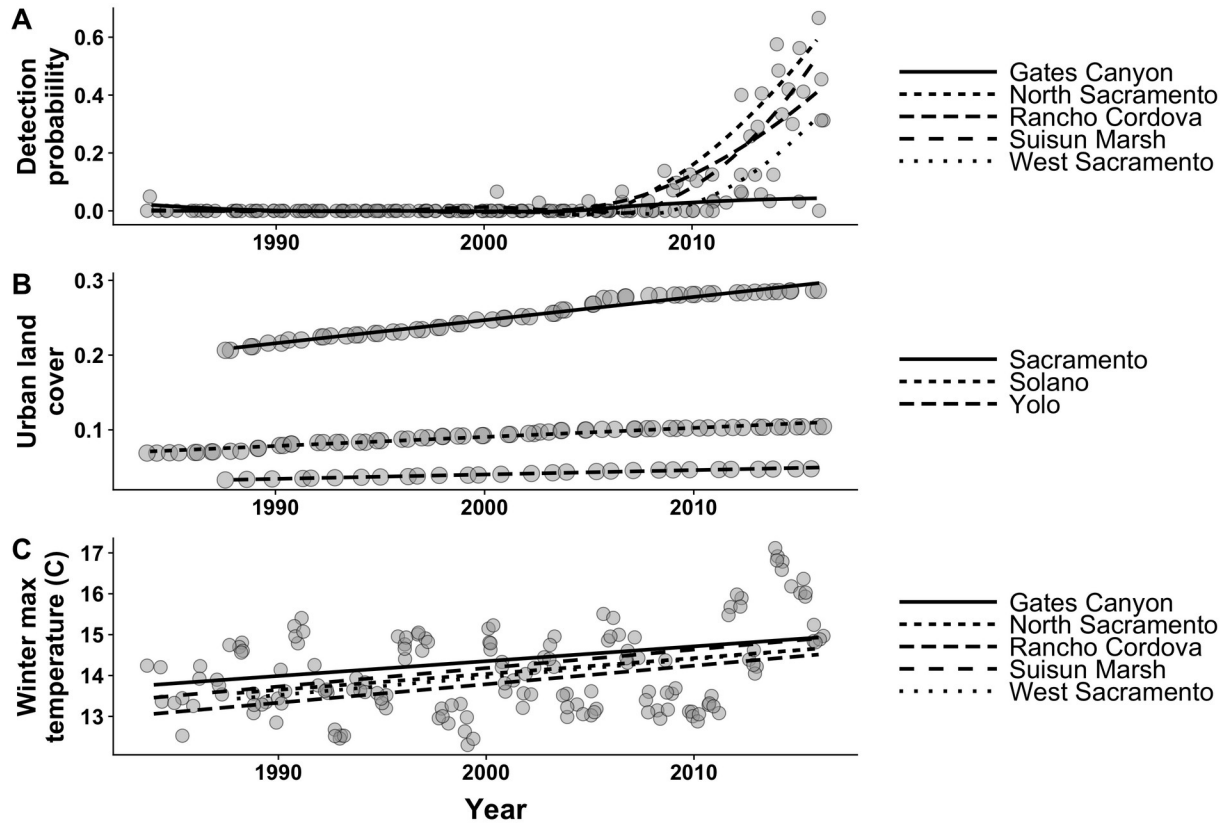
Overwintering Distribution Model						
Region	Min. Temperature	Host Plant	AUC	OR	P-value	
East	65.7	34.3	0.931	0.095	<< 0.05	
West	46.9	53.1	0.855	0.140	<< 0.05	

Maximum Annual Distribution Model					
Region	Av. Temperature	Host Plant	AUC	OR	P-value

East	11.3	88.7	0.843	$\frac{0.10}{2}$	\ll 0.05
West	34.0	66.0	0.821	$\frac{0.07}{3}$	\ll 0.05

672
673
674
675
676
677
678

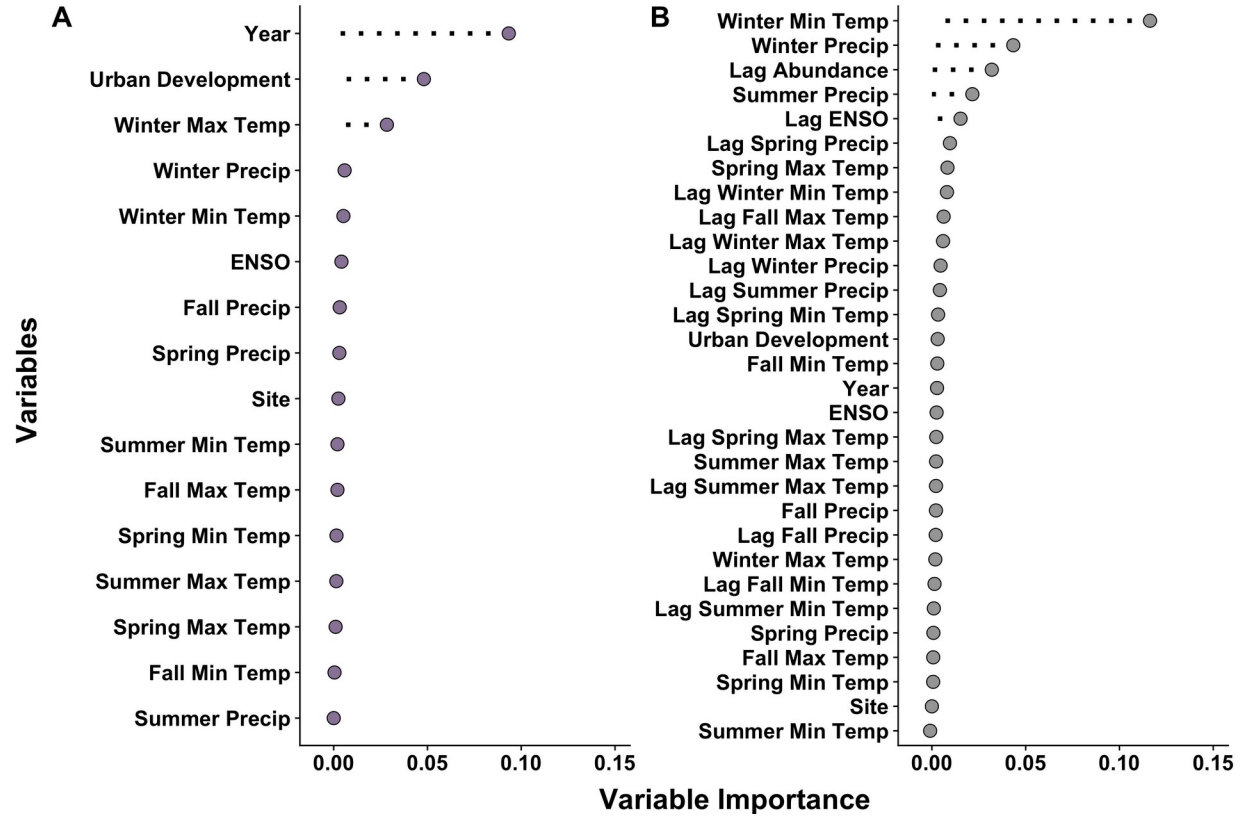
679
680
681
682
683



684
685
686
687
688
689
690
691
692

Figure 1

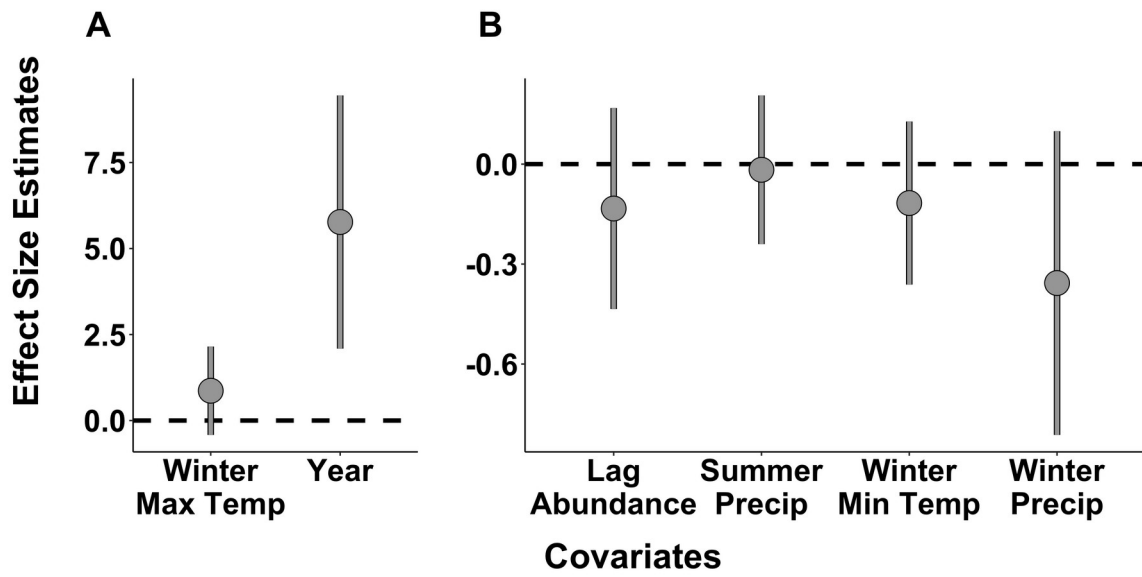
693
694
695
696
697
698



699
700
701
702
703

Figure 2

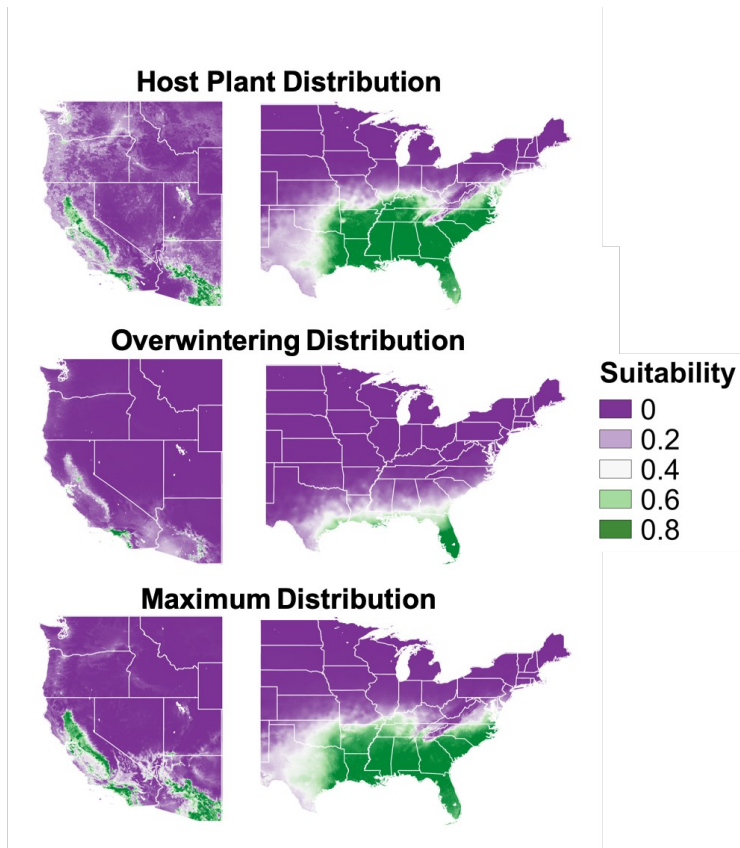
704
705
706
707
708
709
710
711



712 **Figure 3**

713
714
715
716
717
718
719
720
721

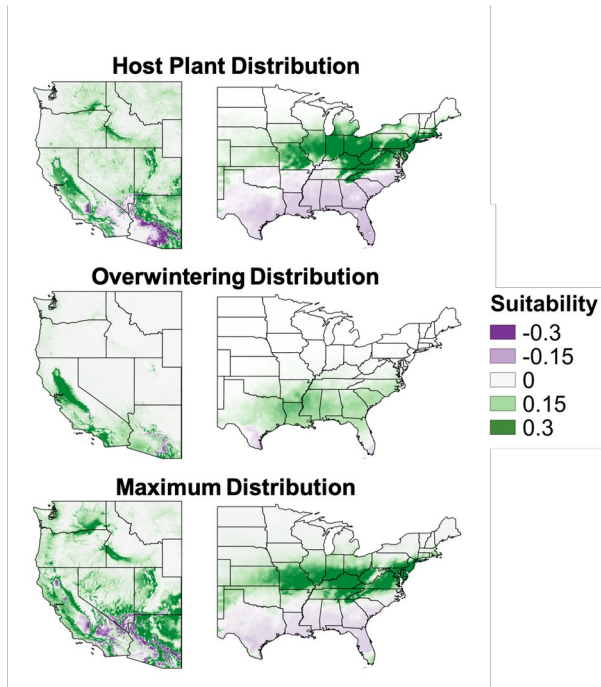
722



723

724 **Figure 4**

725



726

727 **Figure 5**

728

729

730

731

732

733

734

735

736

737

738

739

740

741