Ecological Entomology

Ecological Entomology (2020), DOI: 10.1111/een.12845

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

CHRISTOPHER A. HALSCH,[1] ARTHUR M. SHAPIRO,[2] JAMES H. THORNE,[3] DAVID P. WAETJEN[3] and MATTHEW L. FORISTER[1] [1]Department of Biology, Program in Ecology, Evolution and Conservation Biology, University of Nevada, Reno, Nevada, U.S.A., [2]Department of Evolution and Ecology, Center for Population Biology, University of California, Davis, California, U.S.A. and [3]Department of Environmental Science and Policy, University of California, Davis, California, U.S.A.

Abstract. 1. The changing climate is altering species distributions with consequences for population dynamics, resulting in winners and losers in the Anthropocene.

2. *Agraulis vanillae*, the gulf fritillary butterfly, has expanded its range in the past 100 years in the western U.S.A. Time series analysis is combined with species distribution modelling to investigate factors limiting the distribution of *A. vanillae* and to predict future shifts under warming scenarios.

3. Time series analyses from the western U.S.A. show that urban development has a positive association with year of colonisation (the host plant *Passiflora* is an ornamental in gardens). Colonisation was also associated positively and to a lesser extent with winter maximum temperatures, whereas a negative impact of minimum temperatures and precipitation was apparent on population growth rates after establishment.

4. Species distribution models vary by region. In the eastern U.S.A., the butterfly is primarily limited by minimum temperatures in the winter and host availability later in the season. Eastern U.S. projected expansion broadly follows the expectation of poleward distributional shifts, especially for the butterfly's maximum annual extent. Western U.S. distributions are limited by the host plant, which in turn is dependent on urban centres. Projected western U.S. expansion is not limited to a single direction and is driven by urban centres becoming more suitable for the host plant.

5. These results demonstrate the value of combining time series with spatial modelling, at the same time as incorporating biotic interactions, aiming to understand and predict shifting geographical ranges in the Anthropocene.

Key words. *Agraulis vanillae*, climate change, expansion, *Passiflora*.

Introduction

The influences of global change, which include invasive species, overexploitation and climate change, are impacting species around the world (Butchart *et al.*, 2010). We can expect these factors to have varying effects on different species, and that some species will be 'winners' under altered conditions (McKinney & Lockwood, 1999). Identifying successful species and the reason for their success in the face of environmental change is important for understanding the potential of individual species and ecosystems to persist and thrive in future climates. In particular, understanding how different aspects of global change can negatively impact some species, at the same time as benefiting others, will improve our ability to predict future species assemblages. One broad method for assessing 'winning' and 'losing' is by measuring species distributions, which are already shifting in response to recent change (Chen *et al.*, 2011). For some species, ranges are expanding, whereas, for many others, ranges are shifting or contracting (Parmesan, 2006). In the

Correspondence: Christopher Halsch, Department of Biology, Program in Ecology, Evolution and Conservation Biology, University of Nevada, Mail Stop 314, Reno, 1664 N Virginia Street, Reno, NV 89557, U.S.A. E-mail: cahalsch@nevada.unr.edu

context of warming temperatures, distributional change can by caused by direct effects on development and survival (Crozier, 2004) or by indirect effects mediated by biotic interactions (Gutierrez & Thomas, 2001). Ectotherms, including butterflies and other insects, are particularly sensitive to changes in the climate and are exemplar species for the study of these issues (Parmesan *et al.*, 1999; Warren *et al.*, 2001). In the present study, we investigate the gulf fritillary butterfly (*Agraulis vanillae)*, which appears to be benefitting from anthropogenic influence and has recently expanded its range in the western U.S.A. (Shapiro & Manolis, 2007). We seek to better understand the drivers underlying this expansion using a combination of spatial occurrence data and long-term population records.

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

Agraulis vanillae utilises most plants from the genus *Passiflora* as hosts (May, 1992). The two most common species in the U.S.A. are *Passiflora incarnata* and *Passiflora lutea*, both of which grow naturally across much of the southeastern U.S.A. (Gremillion, 1989). *Passiflora* prefers well-drained soils and is often found in disturbed sites. In the western U.S.A., *Passiflora* is not present in natural areas and is restricted to modified landscapes and gardens because various species have been introduced to urban areas as ornamentals (Graves & Shapiro, 2003). We are not aware of any instances where the plant has escaped urban confines and established large self-sustaining populations. Winter freezing temperatures likely limit the distribution of the plant in the wild, however survival can be improved by active management in cultivated populations (McGuire, 1999). *Agraulis vanillae* was first reported in Southern California in 1875 and in San Francisco as early as 1908. It did not permanently establish in San Francisco until 1955, where it used *Passiflora* (Powell *et al.*, 2000). In the 1960s and 1970s, the butterfly briefly established in Sacramento but was extirpated and has only recently reestablished in the region. The human-propagated expansion of *Passiflora* in urban centres offers an alternative biotic explanation for the expansion of the gulf fritillary.

In the present study, we utilise time series analysis and species distribution modelling to address three specific questions. First, using data from a long-term observational study, we investigate whether climate or urban development better explain the establishment and success of the butterfly in recent years in the Sacramento Valley. Second, using citizen science occurrence data and species distribution modelling, we investigate whether the current distribution of the butterfly in the continental U.S.A. is better explained by host plant or climate limitation and how this varies by region. Finally, we investigate whether the butterfly is likely to continue to expand its distribution under different climate change scenarios.

Materials and methods

Sacramento Valley time series data

Long-term observational data were collected every other week by a single observer (AMS) across five sites in the Sacramento Valley. Count data of individual butterflies at these five sites have been collected subsequent to 1999 and presence/absence data have been collected from the 1970s or 1980s onward, depending on the site. Site descriptions and additional details have been reported elsewhere (Forister *et al.*, 2010). *Agraulis vanillae* did not consistently appear at any of these five sites until 2001 and did not appear at every site until 2012. Climate data in California were derived from 270-m grid climate maps of monthly and annual values for minimum and maximum temperature and precipitation (Flint & Flint, 2012; Flint *et al.*, 2013; Thorne *et al.*, 2015). We extracted the values for grid cells that overlapped with each of the sample sites in the Sacramento Valley and averaged the values for each monthly variable for each year. We calculated seasonal variables by further averaging monthly values to season and converting to water year (the start of September through the end of August).

Sacramento Valley statistical analysis

We approached the analysis of the time series data in two phases. First, we used annual presence/absence data to examine colonisation, attempting to model the difference between years in which the butterfly was absent across our focal sites and years in which it was resident (spanning 1984 to 2018). Residency at a site was determined to be a presence in consecutive years. Random forest regression was used with presence at a site (during years of residency) in a given year as the response variable and year, percentage urban land cover (at a county level), seasonal means of minimum temperature, seasonal means of maximum temperature, and seasonal means of precipitation as covariates. In total, 500 000 trees were made with a node size of 5. Variable importance was determined by examining the increased mean squared error of the model when each variable was randomly permuted. The most influential variables identified by random forest analysis were moved forward into a Bayesian hierarchical linear regression. Although the random forest is useful for judging the potential importance of a large number of variables, including some that are highly correlated, the Bayesian model allows us to estimate coefficients and associated uncertainty in a hierarchical framework (simultaneously within and across sites). In accordance with a previous model used for data from these study sites (Nice *et al.*, 2019), presence was modelled both at the individual site level and at a higher level across all sites using a Bernoulli distribution. Vaguely informative priors were used for means and variance, with means drawn from normal distribution (mean = 0, SD = 10 000) and variances drawn from a gamma distribution (rate = 2, shape = 0.01). The Bayesian model was comprised of four chains each run for 100 000 iterations with a burn in phase of 50 000 iterations.

As a second phase, we examined annual population dynamics post-colonisation at the same focal sites, using individual survey count data summarised by year and transformed into population growth rates. Population growth was calculated as the natural log of the current year's total count divided by the previous year's total count (Sibly & Hone, 2002). To determine the most influential climate variables, population growth in a given year was then modelled using a random forest regression. Covariates in the model included year, urban development, abundance in the previous year, seasonal means of minimum monthly temperature, seasonal means of maximum monthly temperature and seasonal means of precipitation, as well as these same variables lagged by 1 year to allow in particular for effects mediated through host plants. Again, 500 000 trees with a node size of 5 were used. Variable importance was determined by examining the increased mean squared error of the model after permutation of each variable, and this was done both within and among sites. Similar to the colonisation analysis, the most influential variables identified by random forest analysis were moved forward into a Bayesian hierarchical model in which population growth was modelled both at the individual site level and at a higher level across all sites using a normal distribution. Means of covariates were drawn from a vaguely informative normal distribution (mean = 0, SD = 10 000) and variances were drawn from a gamma distribution (rate = 2, shape = 0.01). This model comprised four chains each run for 100 000 iterations with a burn in phase of 50 000 iterations. All analyses were conducted using the randomForest (Liaw & Wiener, 2018) and jagsUI (Kellner, 2019) packages in R and R STUDIO (https://rstudio.com; R Development Core Team, 2013).

National data

For U.S.-wide spatial analyses, geo-referenced data points for both *A. vanillae* and *Passiflora* were acquired from observations on iNaturalist (iNaturalist, 2019) and GBIF (GBIF.org, 2019). Additional observations of *Passiflora* were obtained from Calflora (Calflora: Information on California plants for education, research and conservation, 2014) and additional observations of *A. vanillae* from the Butterflies and Moths of North America (Lotts & Naberhaus, T coordinators, 2017) and eButterfly (Larrivee *et al*, 2018) (see Supporting information, Figs S1 and S2). Only observations made subsequent to 2000 with a spatial precision higher than 1 km were used for analysis. Both *Passiflora* and *A. vanillae* are distinct and identification is likely not a concern; however, a random subset of 100 observations with photos were checked and all were found to be correct IDs. Current climate data and future projections were obtained from WorldClim (Hijmans *et al.*, 2005). A human population density raster was obtained from the Socioeconomic Data and Applications Center, which used data from the 2010 census (Center for International Earth Science Information Network, 2018). All raster layers were cropped to include only the 48 contiguous states of the U.S.A., although *A. vanillae* is also present in Hawaii as an introduced species. Finally, *A. vanillae* points were separated based on being from the overwintering season, which was defined as between January and March, which is earlier than the earliest observed spring migrant from a study of *A. vanillae* seasonal movement in Florida (Walker, 1991).

National statistical analysis

Species distribution models were built for both *Passiflora* and *A. vanillae*. All host plant models were applied to the genus level, although *Passiflora* species known not to be hosts were excluded. The western and eastern distributions were modelled separately to allow for the possibility of different factors affecting range limits in the different regions. For all models, we used the MaxEnt algorithm, which models presence only data by comparing observations with random background points. For every model, 10 000 random background points were taken within the continental U.S.A. To account for sampling bias in the occurrence data, the random background points were spatially structured using a bias file (Phillips *et al.*, 2009). For *Passiflora*, the bias file was built from all Malpighiales observations (excluding *Passiflora)* and the bias file for *A. vanillae* was built using all Nymphalidae observations (excluding *A. vanillae*). *Passiflora* was modelled using temperature, mean precipitation and human population density as covariates. Models were built and evaluated using minimum temperature in the coldest month, mean annual temperature, maximum temperature in the warmest month, and both maximum and minimum together as temperature variables. Human population was included in the model to account for any dependence on urban cultivation, which we hypothesised to be important in the western U.S.A. The best performing host plant model was later used as a covariate for the butterfly distribution model. For *A. vanillae*, both the overwintering and maximum distributions were modelled. The overwintering distribution was modelled using the best performing *Passiflora* distribution model and temperature variables. The maximum annual distribution was similarly modelled using the *Passiflora* distribution model and temperature as covariates. As with *Passiflora* analyses, various temperature variables were used for model building and comparison, and only the highest performing model for both overwinter and dispersal distributions were used for inference and projection. The models were trained on 70% of the data and tested with the remaining 30%. Model evaluation was performed by examining the area under the curve (AUC) scores and omission error rates of both the real model and 1000 permuted null models. Methods and code for null model permutation are described by Bohl *et al.* (2019), although, in brief, observations from the real model are randomly moved around the study area and compared with the real model using the same covariates and testing data. All analyses were performed in R Studio using the dismo package (Hijmans *et al.*, 2013).

Results

Time series

For the first 25 years of the time series, *A. vanillae* only appeared as an occasional visitor; however, beginning in 2001,

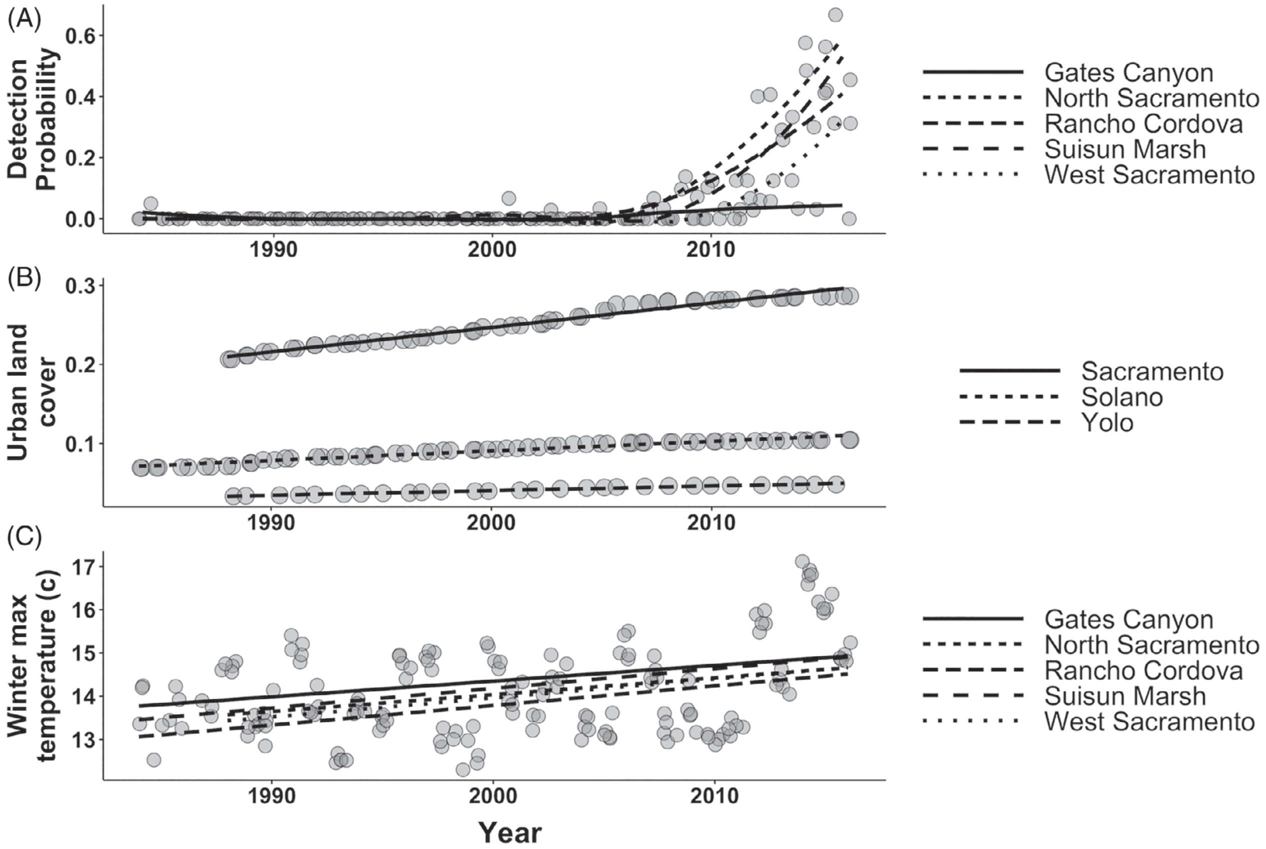

Fig. 1. (a) Change in detection probability (the ratio of days observed to total visits) over time across all sites. (b) Annual ratio of urban land cover to total land cover at a county level for the three counties containing long-term study sites: North Sacramento and Rancho Cordova are in Sacramento County; Suisun Marsh and Gates Canyon are in Solano County. (c) Mean monthly maximum winter temperature over time.

it became a frequent visitor to all sites across the Sacramento Valley. This rise in the presence of *A. vanillae* occurred during a time of rising temperature and increasing urban development in the area (Fig. 1). The random forest model attributed high importance to winter maximum temperatures, percentage urban land cover and year in predicting presence at a site (Fig. 2a). Both maximum temperature and urban land cover were increasing over time, especially land cover, which is highly correlated with year (correlation coefficients for year and land cover range from 0.973 in Solano county to 0.989 in Yolo county). In the Bayesian analysis, the model successfully converged (as judged by visual inspection of posterior probability distributions, Rhat values and effective sample size estimates) at both the individual site level and at the higher across site level. Only year was used in the model because it is highly correlated with urbanisation (precluding the inclusion of both variables). The Bayesian model confirms that both maximum winter temperatures and year are positively associated with colonisation at the higher across site level and at each individual site (Fig. 3a; see also Supporting information, Fig. S3). Specifically, the probability that maximum temperature has a greater than zero effect is 0.98 and the probability that year has a greater than zero effect is 0.92. There is a 0.98 probability that year has a stronger effect than winter, thus the positive trend of colonisation is not sufficiently explained by climate.

For annual population dynamics (represented by the natural log of the current to previous population density), the random forest analysis attributed high importance to abundance in the previous year, winter minimum temperature in the current year, winter precipitation in the current year and summer precipitation in the current year for predicting population growth (Fig. 2b; see also Supporting information, Figs S4 and S5). Although urbanisation is one of the covariates in the model, it was not found to be important for population growth rates. Coefficients in the Bayesian model for population growth converged at both the across site and individual site level. Previous year's abundance, winter minimum temperature and winter precipitation all had negative effects on population growth. The model is confident in the negative impacts of previous year's abundance, winter minimum temperature, and winter precipitation (Fig. 3b). Specifically, the probability that previous year's abundance has a negative effect is 0.84, the probability that winter minimum temperature has a negative effect is 0.80 and the probability that winter precipitation has a negative effect is 0.88. There does not appear to be a strong effect of summer precipitation in the Bayesian hierarchical regression, despite the importance attributed to it in the random forest. All three variables have approximately equal estimated effect sizes. At the individual site level, there is variation in estimated effects; however, negative density dependence is observed at all sites.

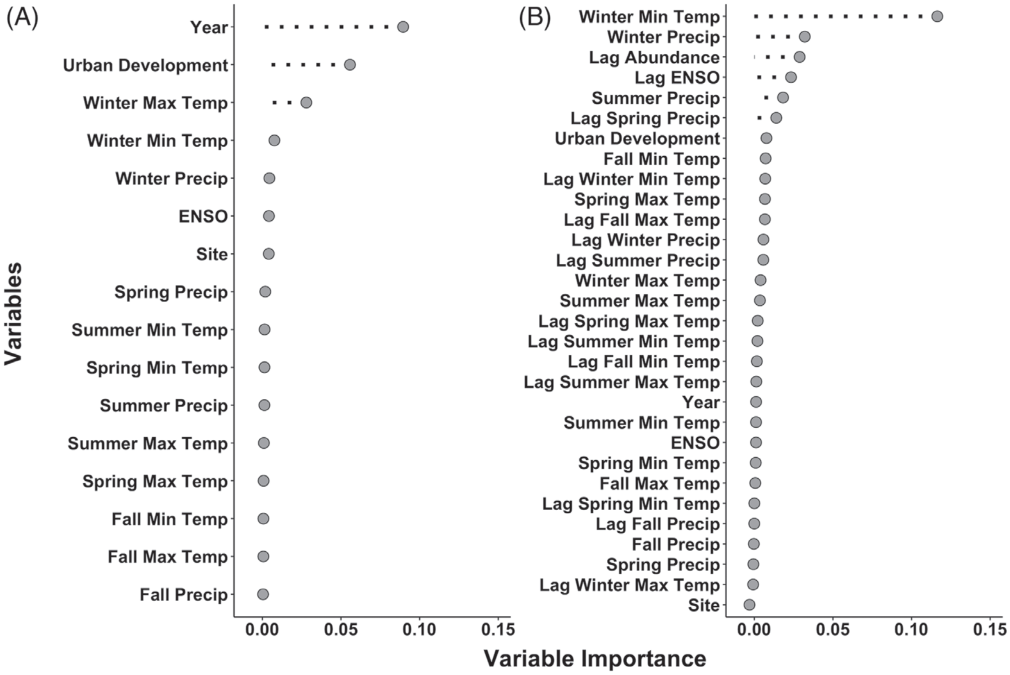

Fig. 2. (a) Variable importance of model covariates in predicting the presence of *A. vanillae* at a site in the Sacramento Valley over time. (b) Variable importance of model covariates in predicting the annual population growth after establishment.

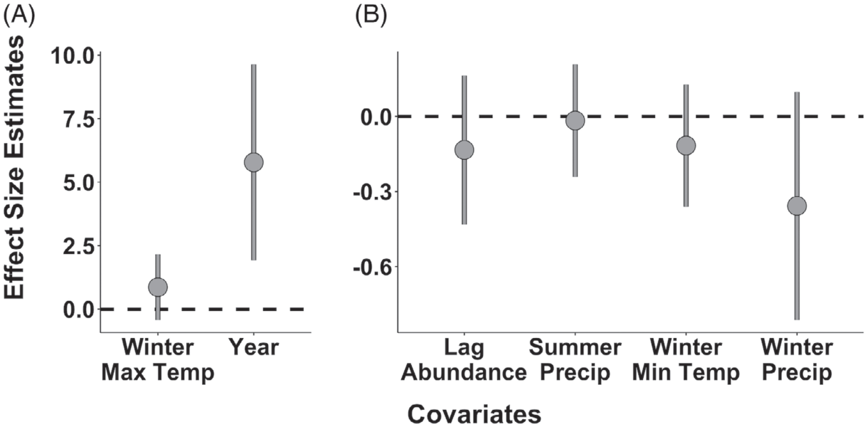

Fig. 3. 95% credible intervals for important coefficients (as determined by random forest). The *y*-axis shows scaled coefficient estimates. (a) Estimates of coefficients for establishment. (b) Estimates of coefficients for population growth.

Winter climate is also important at all sites; however, some sites have higher estimated impacts of winter precipitation, whereas others more heavily weight winter minimum temperatures (see Supporting information, Fig. S6).

Species distribution models

The predictors of highest importance of the geographical distribution of *Passiflora* vary between the eastern and western U.S.A. In the East, *Passiflora* is best predicted by winter minimum temperatures and precipitation, whereas, in the West, urban population and maximum summer temperatures are the best predictors (Table 1; see also Supporting information, Table S1). All models achieved high AUC values and performed exceptionally well when compared with permuted null models (Table 1; see also Supporting information, Fig. S7). Under the Representative Concentration Pathway (RCP) 4.5, suitable habitat in the eastern U.S.A. is predicted to increase along the northern range boundary for *Passiflora*. Habitat is also predicted to become slightly less suitable along the southern range boundary; however, the magnitude of this change in suitability is not comparable to the increase on the poleward margin (Figs 4 and 5). In the western U.S.A., current areas of suitability are predicted to expand, although not in a clear poleward direction.

The current overwintering ranges of *A. vanillae* in the eastern and western U.S.A. are best explained by both host plant and

Table 1. Variable importance and model fit of host plant distribution model.

Host plant distribution model							
Region	Maximum temperature	Minimum temperature	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**

Rows represent different regional models and columns are the different variables in the model. The area under the curve (AUC) is the performance metric of model fit. OR, omission rate.

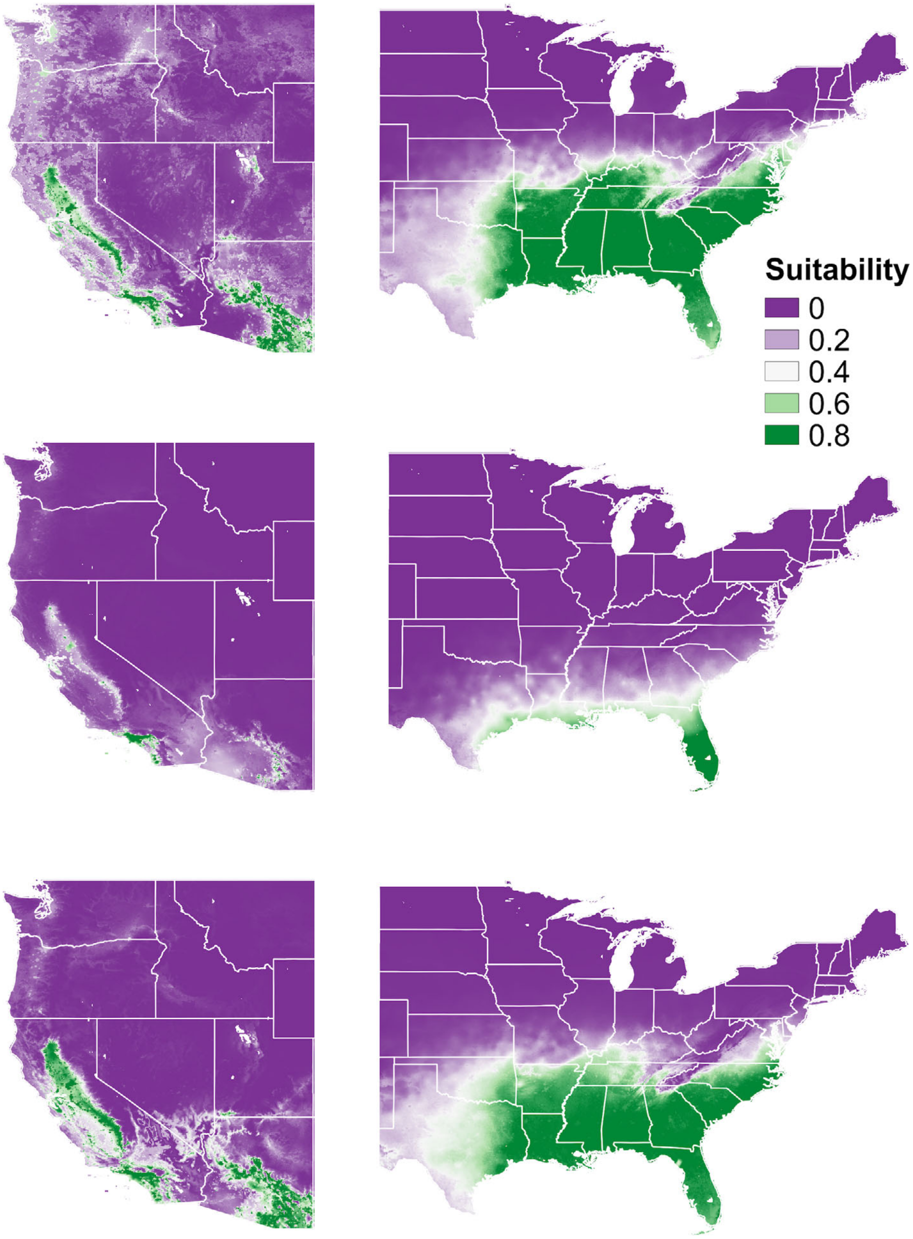

Fig. 4. Current distribution of suitability for *Passiflora* in the West (top left), *Passiflora* in the East (top right), overwintering *A. vanillae* in the West (middle left), overwintering *A. vanillae* in the East (middle right), maximum annual *A. vanillae* in the West (bottom left) and maximum annual *A. vanillae* in the East (bottom right). [Colour figure can be viewed at wileyonlinelibrary.com].

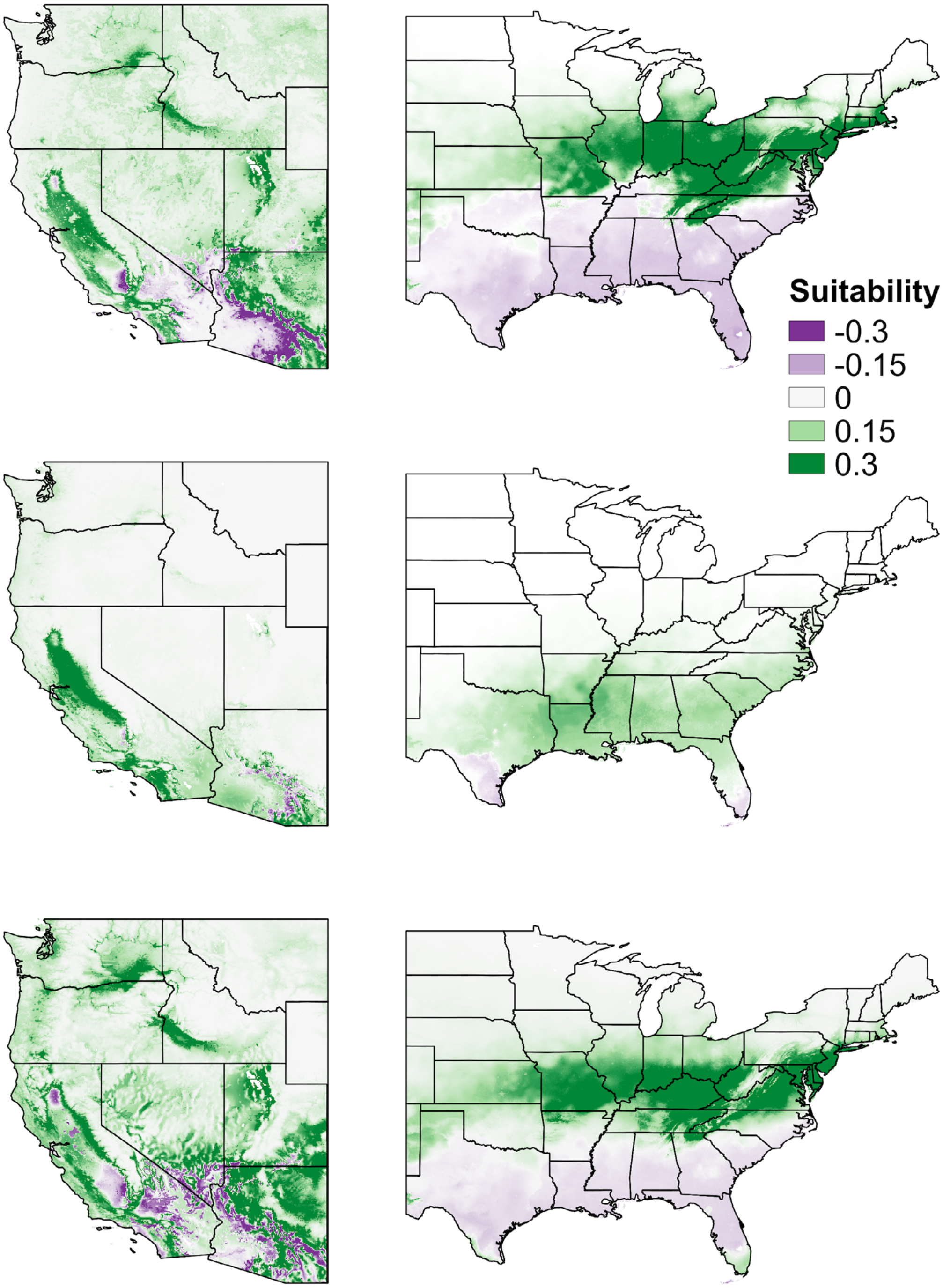

Fig. 5. The expanding gulf fritness landscape. Predicted change in suitability in 2050 under RCP 4.5 for *Passiflora* in the West (top left), *Passiflora* in the East (top right), overwintering *A. vanillae* in the West (middle left), overwintering *A. vanillae* in the East (middle right), maximum annual *A. vanillae* in the West (bottom left) and maximum annual *A. vanillae* in the East (bottom right). [Colour figure can be viewed at wileyonlinelibrary.com].

winter minimum temperatures (see Supporting information, Table S1). Similar to the host plant model, all models performed well in regard to AUC scores and when compared with permuted null models (Table 2; see also Supporting information, Fig. S7). The variable importance of minimum temperature in the East is slightly greater, however it is not clear if these slight differences in variable importance are meaningful (Table 2). Future climate scenarios project a slight increase in the suitability of some areas in the southeast for overwintering, although not a major expansion (Figs 4 and 5). The models of maximum annual distribution tell a different story. Models for maximum annual distribution performed best using average temperature; however, greater

Table 2. Variable importance and model fit of overwintering distribution model.

Overwintering distribution model					
Region	Maximum 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**

Rows represent different regional models and columns are the different variables in the model. The area under the curve (AUC) is the performance metric of model fit. OR, omission rate.

Table 3. Variable importance and model fit of maximum annual distribution model.

Maximum annual distribution model					
Region	Average temperature	Host plant	AUC	OR	*P*-value
East	11.3	88.7	0.843	0.102	**< 0.05**
West	34.0	66.0	0.821	0.073	**< 0.05**

Rows represent different regional models and columns are the different variables in the model. The area under the curve (AUC) is the performance metric of model fit. OR, omission rate.

importance in both regions was given to host plant distribution (Table 3). Again, models performed well using both the AUC metric and permuted null model comparison (Table 3; see also Supporting information, Fig. S7). This greater importance of the host plant is reflected in the future model predictions, which show *A. vanillae* expansion into areas that also predict *Passiflora* expansion (Figs 4 and 5). Thus, although overwintering gains appear marginal under future warming, expansion of the range during the spring and summer is potentially substantial. Across all models, projections under RCP 8.5 show a slightly greater expansion but do not dramatically vary from predictions under RCP 4.5 (see Supporting information, Fig. S8).

Discussion

Species are currently encountering novel biotic and abiotic conditions, which can positively or negatively impact population dynamics and geographical distributions (McKinney & Lockwood, 1999). Building models that parse these various stressors furthers our understanding of these impacts and allows for better prediction of future assemblages. In the present study, we found that years in which the butterfly had colonised our focal sites were characterised by warmer winter maximum temperatures, whereas winter minimum temperatures had a negative impact on population growth rates in the years after colonisation. In particular, if the previous winter was cooler and drier, the butterfly was found in higher abundance the next year. It is possible that the negative impact of winter climate on *A. vanillae* that we observed is mediated through interactions with host plants or other insects. It could be the case that warmer and wetter winters negatively impact *Passiflora*, although another and perhaps more likely explanation is that wetter and warmer winters increase parasitoid pressure and/or disease leading to reduced adult emergence the next year (Harvell *et al.*, 2002; Stireman III *et al.*, 2005). *Agraulis vanillae* is known to host nucleopolyhedrovirus (Rodriguez *et al.*, 2011), which could be one mechanism that generated the observed negative density dependence (see Supporting information, Fig. S4); however, this is not known to impact California populations. Finally, at our focal sites, there is a slight positive trend over time in winter precipitation and winter minimum temperature (see Supporting information, Fig. S9), suggesting that, if anything, the butterfly is persisting and expanding in the Sacramento Valley despite climate and not because of it.

The local impact of climate on the population dynamics of *A. vanillae* in the Sacramento Valley also has implications for explaining the limiting factors of its current distribution in the western U.S.A. Distribution models of *A. vanillae* in the east and west place high importance on the distribution of the host plant; however, only the western host plant model identified human population density as an important predictor (after accounting for sampling bias in the data). One explanation for the recent colonisation of the area by the butterfly is the increasing urbanisation of the Sacramento Valley. Over the past 20 years, the suburbs of Sacramento have expanded at a steady rate (Forister *et al.*, 2010), which has likely resulted in an increase in *Passiflora* in the region. Random forest analysis ranked urban land cover over any climate variable when predicting colonisation and the Bayesian model found a much greater effect of year (which is highly correlated with urbanisation). In the eastern U.S.A., the impacts of temperature, specifically minimum temperatures, are apparent in geographical distribution models. In the east, the distribution of *Passiflora* extends further north in the winter compared with *A. vanillae*, whereas, in the west, the overwintering distribution closely resembles that of *Passiflora*. Once the weather warms in the east, the butterflies can then expand to cover the distribution of the host plant. Thus, although minimum temperature plays an important role in the overwintering locations of the eastern gulf fritillary, its maximum extent appears to be host plant limited.

Although all analyses involve a single focal species, an interesting result of the present study is the discovery that variation in limiting factors between the east and west result in quite different predictions for distributional change under future climates by season and by region. In the eastern U.S., models using the RCP 4.5 and 8.5 climate scenarios broadly follow the expectation of poleward movement, with more suitable habitat along the northern range margin and a slight reduction in habitat suitability in south. In the winter the butterfly is limited by temperature and predicted expansion during this time will largely be as a result of increasing temperatures. Later in the season, the butterfly is primarily limited by the distribution of the host plant and this expansion would be better explained by an indirect effect of temperature mediated the distribution of the host plant. In the western U.S., expansion is also predicted, although not in a single direction. This region is much more climactically and topographically complex and this result is perhaps not surprising. A recognisable pattern is the importance of population centres, especially in the expansion of the maximum annual distribution of the butterfly. It is important to note that our future projections were created using climate forecasts and not human population forecasts. This means that there is an underlying assumption in the projection that population density will remain the same, which almost certainly will not be met. Given the predictive power of population in the western U.S.A. models, we suggest that these projections are conservative. Overall, newly suitable areas for the butterfly closely follow the newly suitable areas for the host plant, thus we infer that expansion in the West is most closely tied to the indirect effect of host plant expansion.

These findings add to the literature stressing the utility of accounting for biotic interactions in species distribution modelling and forecasting. Biotic interactions are an important factor in shaping the distributions of species but have been incorporated into few studies examining climate change (Araujo & Luoto, 2007; Heikkinen *et al.*, 2007; Preston *et al.*, 2008; Schweiger *et al.*, 2008), at least relative to abiotic-only distribution models. Many of the studies that do incorporate biotic information demonstrate that, whether the biotic element be a host plant or a mutualist, model performance is improved. Similarly, we show that the host plant has high predictive importance and allows for a better understanding of the current distributional limits of the butterfly. Another important component of these results is the observed within-distribution variation because we show the importance of host plant varying by season and region. Recognising and accounting for this variation is critical to better predict future responses to change, especially for species with large spatial distributions (Murphy & Lovett-Doust, 2007; O'Neill *et al.*, 2008). By incorporating both a key host plant interaction and allowing it to vary by region, we have a more complete understanding of this observed expansion.

The gulf fritillary is a notable example of a 'winner' in the Anthropocene. Although insects are declining on a large scale (Hallmann *et al.*, 2017; Lister & Garcia, 2018; Salcido *et al.*, 2020; Sanchez-Bayo & Wyckhuys, 2019; Wepprich *et al.*, 2019), altered conditions create opportunities for some to prevail. The nuances of each success story are different, although it is clear that increasing temperature is playing a vital role in facilitating the distributional expansion of many of these insect winners. Other studies have shown that rising temperature can impact insect distributions by increasing overwintering survival along a northern range margin (Crozier, 2004), by increasing access to food resources (Raffa *et al.*, 2013) or by increasing diet breadth (Pateman *et al.*, 2012). As temperatures continue to warm, insects will continue to be prime candidates for temperature-driven distributional change, for better or for worse. Continuing to observe these phenomena and developing methods by which to understand them is critical. In the present study, the combination of long-term time series data and large-scale citizen science spatial data allowed for a detailed examination of the underlying causes for such an expansion. As these types of data continue to become more widely accessible, the common themes behind insect distributional change in the Anthropocene will continue to become more apparent.

Acknowledgements

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

AMS collected the Sacramento Valley observational data. JHT and DPW provided the climate data. CAH conducted the species distribution analyses. CAH and MLF conducted time series analysis. CAH and MLF wrote the manuscript with input from the co-authors.

The authors declare that they have no conflicts of interest.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Spatial model scores for various temperature variables.

Figure S1. Raw occurrence data used in analysis for *Passiflora*.

Figure S2. Raw occurrence data used in analysis for *Agraulis vanillae*.

Figure S3. Individual site level Bayesian posterior distributions from the colonisation model

Figure S4. Relationship between population growth and covariates

Figure S5. Correlation matrix of all model covariates

Figure S6. Individual site level Bayesian posterior distributions from the population growth model

Figure S7. Real SDM comparison with permuted null models.

Figure S8. *Passiflora*, *A. vanillae* overwintering and *A. vanillae* maximum distributions SDMs under RCP 8.5.

Figure S9. Change in winter minimum temperature and winter precipitation over time.

References

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

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

Butchart, S.H.M., Walpole, M., Collen, B., Strien, A.V., Scharlemann, J.P.W., Almond, R.E.A. *et al.* (2010) Global biodiversity: indicators of recent declines. *Science*, **328**, 1164–1168.

Calflora: Information on California plants for education, research and conservation. [web application] (2014) Berkeley, California: The

Calflora Database [a non-profit organization]. Available: https://www.calflora.org/ [Accessed: November 5, 2019].

Center for International Earth Science Information Network – CIESIN – Columbia University. (2018) Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11. Palisades, NY: NASA Socioeconomic Data and Application Center (SEDAC). https://doi.org/10.7927/H49C6VHW. (Accessed May 5, 2019).

Chen, I.C., Hill, J.K., Ohlemuller, R., Roy, D.B. & Thomas, C.D. (2011) Rapid range shifts of species associated with high levels of climate warming. *Science*, **333**, 1024–1026.

Crozier, L. (2004) Warmer winters drive butterfly range expansion by increasing survivorship. *Ecology*, **85**, 231–241.

Flint, L.E. & Flint, A.L. (2012) Downscaling future climate scenarios to fine scales for hydrologic and ecological modeling and analysis. *Ecological Processes*, **1**, 1–15.

Flint, L.E., Flint, A.L., Thorne, J.H. & Boynton, R. (2013) Fine-scale hydrologic modeling for regional landscape applications: the California Basin characterization model development and performance. *Ecological Processes*, **2**, 1–21.

Forister, M.L., McCall, A.C., Sanders, N.J., Fordyce, J.A., Thorne, J.H., Obrien, J. *et al.* (2010) Compounded effects of climate change and habitat alteration shift patterns of butterfly diversity. *Proceedings of the National Academy of Sciences*, **107**, 2088–2092.

GBIF.org (3rd May 2019) GBIF Occurrence Download https://doi.org/10.15468/dl.twov45

Graves, S.D. & Shapiro, A.M. (2003) Exotics as host plants of the California butterfly fauna. *Biological Conservation*, **110**, 413–433.

Gremillion, K.J. (1989) The development of a mutualistic relationship between humans and maypops (*Passiflora incarntata L.*) in the southeastern United States. *Journal of Enthnobiology*, **9**, 135–155.

Gutierrez, D. & Thomas, C.D. (2001) Marginal range expansion in a host-limited butterfly species *Gonepterys rhamni*. *Ecological Entomology*, **25**, 165–170.

Hallmann, C.A., Sorg, M., Jongejans, E., Siepel, H., Hofland, N., Schwan, H. *et al.* (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS One*, **12**, e0185809.

Harvell, C.D., Mitchell, C.E., Ward, J.R., Altizer, S., Dobson, A.P., Ostfeld, R.S. *et al.* (2002) Climate warming and disease risks for terrestrial and marine biota. *Science*, **296**, 2158–2162.

Heikkinen, R.K., Luoto, M., Virkkala, R., Pearson, R.G. & Korber, J. (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, **16**, 754–763.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2013) dismo: Species distribution modeling. *R package version 0*.8–17.

iNaturalist. Available from https://www.inaturalist.org. Accessed [2019-11-05].

Kellner, K. (2019) jagsUI: A Wrapper Around 'rjags' to Streamline 'JAGS' Analyses. *R package version 1.5.1*.

Larrivee, M., Prudic, K.L., McFarland, K.P., & Kerr, J. (2018) eButterfly: a citizen-based butterfly database in the biological sciences. http://www.e-butterfly.org

Liaw, M.A. & Wiener, M. (2018) randomForest: Breiman and Cutler's random forests for classification and regression. *R package version*, **4**, 6–14.

Lister, B.C. & Garcia, A. (2018) Climate-driven declines in arthropod abundance restructure a rainforest food web. *Proceedings of the National Academy of Sciences*, **115**, 10397–10406.

Lotts, K. & Naberhaus, T coordinators. (2017) Butterflies and Moths of North America. Data set accessed (or exported) 2019-06-13 at http://www.butterfliesandmoths.org/.

May, P.G. (1992) Flower selection and the dynamics of lipid Reserve in two Nectarivorous Butterflies. *Ecology*, **73**, 2181–2191.

McGuire, M. (1999) *Passiflora incarnata* (Passifloraceae): a new fruit crop. *Economic Botany*, **53**, 161–167.

McKinney, M.L. & Lockwood, J.L. (1999) Biotic homogenization: a few winners replacing many losers in the next mass extinction. *Trends in Ecology and Evolution*, **14**, 450–453.

Murphy, H.T. & Lovett-Doust, J. (2007) Accounting for regional niche variation in habitat suitability models. *Oikos*, **116**, 99–110.

Nice, C.C., Forister, M.L., Harrison, J.G., Gompert, Z., Fordyce, J.A., Thorne, J.H. *et al.* (2019) Extreme heterogeneity of population response to climatic variation and the limits of prediction. *Global Change Biology*, **2019**, 1–10.

O'Neill, G.A., Hamann, A. & Wang, T. (2008) Accounting for population variation improves estimated of the impact of climate change on species' growth and distribution. *Journal of Applied Ecology*, **45**, 1040–1049.

Parmesan, C. (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 637–669.

Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J.K., Thomas, C.D., Descimon, H. *et al.* (1999) Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, **339**, 579–583.

Pateman, R.M., Hill, J.K., Roy, D.B., Fox, R. & Thomas, C.D. (2012) Temperature-dependent alterations in host use drive rapid range expansion in a butterfly. *Science*, **336**, 1028–1030.

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

Powell, J.A., Russell, P., Russell, S. & Sperling, F.A.H. (2000) Northward expansion of two mint-feeding species of *Pyrausta* in California (Lepidoptera: Pyraloidea: Crambidae). *Holarctic Lepidoptera*, **7**, 55–58.

Preston, K.L., Rotenberry, J.T., Redak, R.A. & Allen, M.F. (2008) Habitat shifts of endangered species under altered climate conditions: importance of biotic interactions. *Global Change Biology*, **14**, 2501–2515.

R Development Core Team. (2013) R: A language and environment-for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [WWW document]. URL http://www.R-project.org/ [accessed on 5 November 2019].

Raffa, K.F., Powell, E.N. & Townsend, P.A. (2013) Temperature-driven range expansion of an irruptive insect heightened by weakly coevolved plant defenses. *Proceedings of the National Academy of Sciences*, **110**, 2193–2198.

Rodriguez, V.A., Belaich, M.N., Gomez, D.L.M., Sciocco-Cap, A. & Ghiringhelli, P.D. (2011) Identification of nucleopolyhedrovirus that infect Nymphalid butterflies *Agraulis vanillae* and *Dione juno*. *Journal of Invertebrate Pathology*, **106**, 255–262.

Runquist, E.B., Forister, M.L. & Shapiro, A.M. (2012) Phylogeography at large spatial scales: incongruent patterns of population structure and demography of pan-American butterflies associated with weedy habitats. *Journal of Biogeography*, **39**, 382–396.

Salcido, D.M., Forister, M.L., Garcia Lopez, H., & Dyer, L.A. (2020) Loss of dominant caterpillar genera in a protected tropical forest. *Scientific Reports*, **10**. https://doi.org/10.1038/s41598-019-57226-9.

Sanchez-Bayo, F. & Wyckhuys, K.A.G. (2019) Worldwide decline of the entomofauna: a review of its drivers. *Biological Conservation*, **232**, 8–27.

Schweiger, O., Settele, J., Kudrna, O., Klotz, S. & Kuhn, I. (2008) Climate change can cause spatial mismatch of trophically interacting species. *Ecology*, **89**, 3472–3479.

Scott, J.A. (1986) *The butterflies of North America: A Natural History and Field Guide*. Stanford University Press, Palo Alto, CA.

Shapiro, A.M. (2009) The neo-riparian butterfly fauna of western Argentina. *Journal of Research on the Lepidoptera*, **41**, 24–30.

Shapiro, A.M. & Manolis, T.D. (2007) *Field Guide to Butterflies of the San Francisco Bay and Sacramento Valley Regions*. University of California Press, Berkeley, CA.

Sibly, R.M. & Hone, J. (2002) Population growth rate and its determinants: an overview. *Philosophical Transactions of the Royal Socity B*, **357**, 1153–1170.

Sourakov, A. (2008) Notes on the biology of the gulf fritillary *Agraulis vanillae* (Lepidoptera: Nymphalidae), in north-Central Florida. *Journal of the Lepidopterists' Society*, **63**, 127.

Stireman, J.O. III, Dyer, L.A., Janzen, D.H., Singer, M.S., Lill, J.T., Marquis, R.J. *et al.* (2005) Climatic unpredictability and parasitism of caterpillars: implications of global warming. *Proceedings of the National Academy of Sciences*, **102**, 17384–17387.

Thorne, J.H., Boynton, R.M., Flint, L.E. & Flint, A.L. (2015) The magnitude and spatial patterns of historical and futurehydrologic change in California's watersheds. *Ecosphere*, **6**, 1–30.

Walker, T.J. (1991) Butterfly migration from and to peninsular Florida. *Ecological Entomology*, **16**, 241–252.

Warren, M.S., Hill, J.K., Thomas, J.A., Asher, J., Fox, R., Huntley, B. *et al.* (2001) Rapid responses of British butterflies to opposing forces of climate and habitat change. *Nature*, **414**, 65–69.

Wepprich, T., Adrion, J.R., Ries, L., Wiedmann, J. & Haddad, N.M. (2019) Butterfly abundance declines over 20 years of systematic monitoring in Ohio, USA. *PLoS One*, **14**, e0216270.

Accepted 6 January 2020
Associate Editor: Christopher Hassall