

Research



Cite this article: Menchetti M, Guéguen M, Talavera G. 2019 Spatio-temporal ecological niche modelling of multigenerational insect migrations. *Proc. R. Soc. B* **286**: 20191583.
<http://dx.doi.org/10.1098/rspb.2019.1583>

Received: 4 July 2019

Accepted: 14 August 2019

Subject Category:

Ecology

Subject Areas:

behaviour, ecology, environmental science

Keywords:

animal movement, citizen science, insect migration, Painted Lady butterfly, species distribution modelling, *Vanessa cardui*

Author for correspondence:

Gerard Talavera

e-mail: gerard.talavera@csic.es

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.4638230>.

Spatio-temporal ecological niche modelling of multigenerational insect migrations

Mattia Menchetti¹, Maya Guéguen² and Gerard Talavera^{1,3}

¹Institut de Biologia Evolutiva (CSIC-UPF), Passeig Marítim de la Barceloneta 37, 08003 Barcelona, Catalonia, Spain

²Laboratoire d'Écologie Alpine, Univ. Grenoble Alpes, CNRS, Univ. Savoie Mont Blanc, CNRS, LECA, Grenoble 38000, France

³Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA

id MM, 0000-0002-0707-7495; GT, 0000-0003-1112-1345

Modelling ecological niches of migratory animals requires incorporating a temporal dimension, in addition to space. Here, we introduce an approach to model multigenerational migratory insects using time-partitioned environmental variables (by months and years) and time- and behaviour-partitioned records (breeding records to model reproductive habitat). We apply this methodology to modelling the Palearctic-African migratory cycle of the Painted Lady butterfly (*Vanessa cardui*), based on data encompassing 36 years (646 breeding sites from 30 countries). Each breeding record is linked to a particular time (month and year), and the associated values of the bioclimatic variables are used for an ensemble modelling strategy, to finally obtain monthly projections. The results show obligated movements, mostly latitudinal, for the species' successive generations across the overall range, and only scattered locations show high probabilities of reproduction year-round. The southernmost reproductive areas estimated for the Palearctic-African migratory pool reach equatorial latitudes from December to February. We thus propose a potential distribution for the winter 'missing generations' that would expand the *V. cardui* migration cycle to encompass about 15 000 km in latitude, from northernmost Europe to equatorial Africa. In summer, Europe represents the major temporary resource for *V. cardui*, while January and February show the lowest overall suitability values, and they are potentially the most vulnerable period for the species to suffer yearly bottlenecks. In summary, we demonstrate the potential of the proposed niche modelling strategy to investigate migratory movements of insects.

1. Background

Animals frequently partition their time to perform different behaviours and this is often reflected in a spatial partition as well. While this phenomenon may take place at the microhabitat level for most species, it is extreme in the case of migratory species. For example, during the migratory phase, animals can be observed in habitats that are unsuitable for breeding, and environmental requirements of overwintering sites may be widely different from those of the reproductive phase. Thus, for ecological modellers, it is important to decide which behavioural phase is to be modelled and to filter records accordingly. On the other hand, migratory animals often track ephemeral resources that are variable throughout the year. As a consequence, models need to account for, on top of environmental regularities, stochasticity and particularities of each year, and to associate data records to the specific conditions at the time of the observation. Modelling of migratory species is thus particularly challenging, and a higher amount of data is needed in the form of: (1) time-partitioned environmental layers and (2) time-partitioned and behaviour-partitioned presence records.

Sampling presence records for migratory insect species have some additional particularities. As migratory insects involve continuous movements along the year through multiple generations, each individual completes only one part of the cycle, and the entire migratory range generally involves large, frequently continuous distributions. Each generation may use only part of the overall niche breadth of the species, and sampling needs to be designed to capture such spatio-temporal variability.

Ecologists have successfully modelled seasonal variation in distributions of highly mobile marine animals [1–3] and birds [4–6], where niches are usually not equivalent for different behaviours, i.e. breeding, foraging, migrating, or overwintering. In the Monarch butterfly (*Danaus plexippus*), monthly requirements during the breeding and overwintering seasons have also been modelled [7]. However, modelling has not been used to improve our knowledge of insect migratory species with poorly known breeding requirements and migratory routes, or at scales involving transcontinental migrations.

In this paper, we introduce an approach to build spatio-temporal niche models for migratory insect species with multiple generations per year, based on annual and monthly time-series variables. We use the Palearctic-African migratory range of the Painted Lady butterfly (*Vanessa cardui*) as a model system, and model niche requirements for reproductive behaviour using immature or oviposition records for Europe and Africa. *Vanessa cardui* performs an annual migration exploiting temperate regions in Europe and North Africa from March to November, and tropical regions in sub-Saharan Africa from September to March [8–10]. Up to 10 generations are thought to complete this migratory loop every year. Unlike most temperate species of butterflies, *V. cardui* does not have the capacity to diapause during winter, and has instead evolved the strategy of continuous migration and breeding throughout the year. After hatching from the chrysalis, each specimen undergoes first a migratory phase (more or less intense) followed by a reproductive phase [11]. Thus, this species only partitions time and space at the macrohabitat scale between two main behaviours: migration and breeding (unlike *D. plexippus*, which also overwinters). Each phase is composed of several steps with slightly different microhabitat requirements (e.g. resting at night during the migratory phase, or hilltopping behaviour during the reproductive phase). Although recent evidence suggests that most populations spend the winter months (December to February) in the Afrotropics [8,10], the location and characteristics of the breeding habitat, as well as the regional movements of the population core in such an extensive area, have not yet been precisely studied. There is thus an unresolved spatio-temporal gap of knowledge within their annual migratory range, even involving ‘missing’ winter generations, from December to February, a period for which there are virtually no records.

By projecting the ecological niche model obtained for *V. cardui* onto monthly climatic layers we estimate potential breeding grounds, and discuss their geographical shifts along the year, thus improving our knowledge of the Palearctic-African migratory movements of this species. In particular, we aim to shed light on the following questions: (1) is *V. cardui* obligated to migrate to reach suitable breeding areas in successive generations, or are there possibilities for permanent residency? (2) Are Afrotropical grounds suitable for reproduction during the European winter, and if yes, where might the

‘missing’ generations breed? (3) Are resources unbalanced along the year, and if yes, where might population numbers increase or decrease?

2. Methods

(a) Study area: the Palearctic-African migratory range

Environmental data and presence/absence records were geographically delimited between latitude -5° and 70° , and longitude -20° and 60° . This area widely includes the recognized Palearctic-African migratory system of *V. cardui* [8–10].

(b) Breeding records dataset

Records for immature stages were gathered, including eggs and ovipositions, and adult emergences from pupae, although the vast majority were observations of larvae. Data sources based on exhaustive data mining included online databases for monitoring schemes, curated citizen science projects (e.g. iNaturalist), records shared on social networks (e.g. Instagram, Flickr, and Facebook), literature, personal communications, and records by the authors. In total, 960 occurrences of breeding sites recorded between 1971 and 2017 were databased along with specific observation dates, geographical coordinates, stages observed, and number of immatures. Breeding records included sites in 30 countries within the study area and represented 140 months for 36 years (figure 1; electronic supplementary material, table S1). Coinciding with a memorable outbreak of Painted Lady butterflies in Europe in summer 2009, records retrieved from the Waarneming.nl and Waarneming.be databases in the Netherlands and Belgium were over-represented, and these were randomly thinned to $\frac{1}{4}$. Observations within the timeframe of 5 days before or after the end of one month were considered as a presence point for both months, given that the development time of the species is longer than two weeks. Because *V. cardui* may congregate in large numbers to breed in highly suitable grounds, we wanted to reflect relative abundance in the models. Thus, records for each site were replicated according to the number of observed immatures. For records with no specific data on abundance, a value of one was used. The resulting number of records used for niche modelling was 11 762 data points within 646 breeding sites.

Documenting true absences is usually a source of contention, given the difficulty to distinguish between unsuitable or unoccupied habitat. To build our model, we used 400 randomly selected absence points for northern European latitudes (greater than 50°) corresponding to the months of November, December, January, and February (100 points per each of the four months), a period and region where *V. cardui* breeding has never been recorded. However, it is well known that the species disappears from Central and Northern Europe during the winter months, due to their migratory strategy. These absence points were randomly assigned to any year within the studied temporal frame. In addition to these 400 spatio-temporal points we also included 10 spatio-temporal points in Africa with confirmed absence of both breeding and host plant, after dedicated field searches in the region [8].

(c) Predictor variables

A suite of 10 climatic variables was retrieved from the Climatic Research Unit time-series CRU TS v. 3.24 at 0.5° resolution [12–16] as a source of time-partitioned variables. Monthly layers for a time-series between 1971 and 2017 were used for the modelling procedure. Principal component analyses were run on a monthly based series of variables to select those most independent within months. We kept those variables that were most ecologically

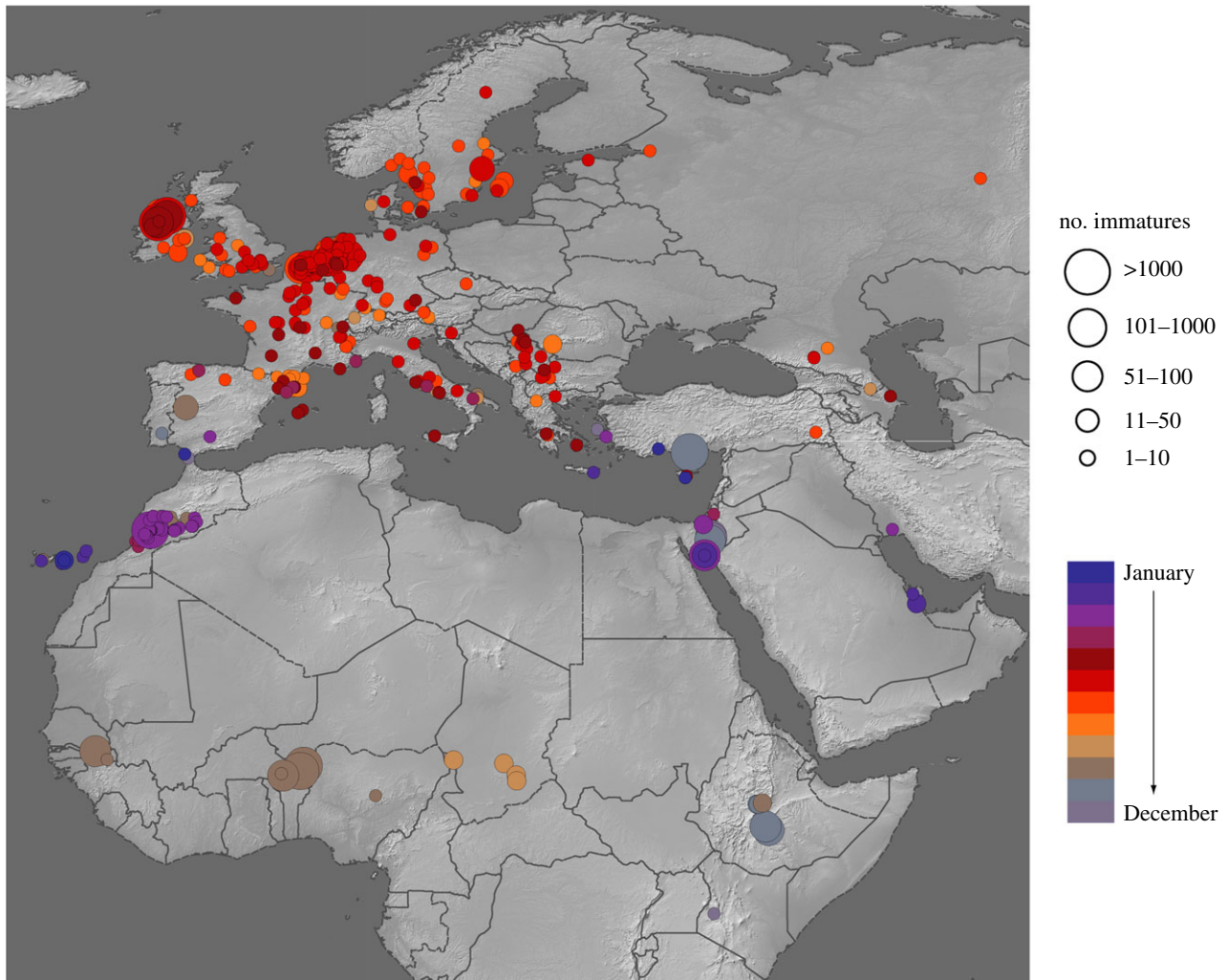


Figure 1. Map of breeding records used within the study area. A total of 646 breeding sites were recorded. Colour indicates the month when breeding was recorded (January–December, see scale). Circle sizes denote relative numbers of immatures per breeding site and date. (Online version in colour.)

relevant for the biology of *V. cardui* and also correlated less than 80% to the other variables across all months. Using these criteria we selected: frost day frequency (*frs*), precipitation (*pre*), daily mean temperature (*tmp*), vapour pressure (*vap*), and wet day frequency (*wet*). In addition to climatic data, land cover layers were obtained from Globcover v. 2.4 [17] at 30 arc-seconds resolution. Layers were rescaled to match the resolution of the climatic variables. With a total of 22 categories, land cover layers were re-categorized as: *open*, *closed*, and *bareland* (electronic supplementary material, table S2).

(d) Modelling strategies

Monthly *V. cardui* breeding distributions were predicted using bioclimatic niche modelling. Each spatial data point (breeding occurrence) was linked to a particular time (month and year), thus extracting the values of the predictor variables detailed above (climate and land cover) for each of these locations and times. An ensemble of forecasts of species distribution models was then obtained, including projections from five statistical models, namely Generalized Linear Models (GLM), Generalized Additive Models (GAM), Multivariate Adaptive Regression Splines (MARS), Random Forests (RF), and Generalized Boosted Models (GBM). Models were calibrated using 70% randomly selected occurrence data, and accuracy was evaluated against the remaining 30% of the data, using the True Skill Statistic (TSS) [18] and the area under the curve (ROC) [19]. The analysis was replicated 10 times, thus providing a 10-fold internal cross-validation of the models. All calibrated models were further

projected to a monthly set of variables resulting from averaging the time-series (1971–2017), thus obtaining one prediction per month. To summarize all projections into a meaningful integrated projection per month, we used an ensemble strategy using committee averaging based on the TSS values (greater than 0.35) of each model. Models and the ensemble forecasting procedure were performed within the R package *biomod2* [20,21]. The *niche.overlap* function in the R package *phyloclim* [22] was used to quantify pairwise degrees of niche overlap between successive months, as proposed by Warren *et al.* [23].

3. Results

(a) Ecological niche modelling predictive accuracy and variable importance

Predictive accuracy across the individual models was consistently high for the two evaluation methods used (TSS and ROC). Mean score values for each modelling method exceeded 0.943 for TSS and 0.971 for ROC (electronic supplementary material, table S3), indicating that all individual models were acceptable.

The most important predictor variable was daily mean temperature (*tmp*), accounting for 30.06% of the sum of importance scores for all variables (electronic supplementary material, table S4). This was followed by frost day frequency (*frs*, 20.41%) and vapour pressure (*vap*, 17.38%). Bareland

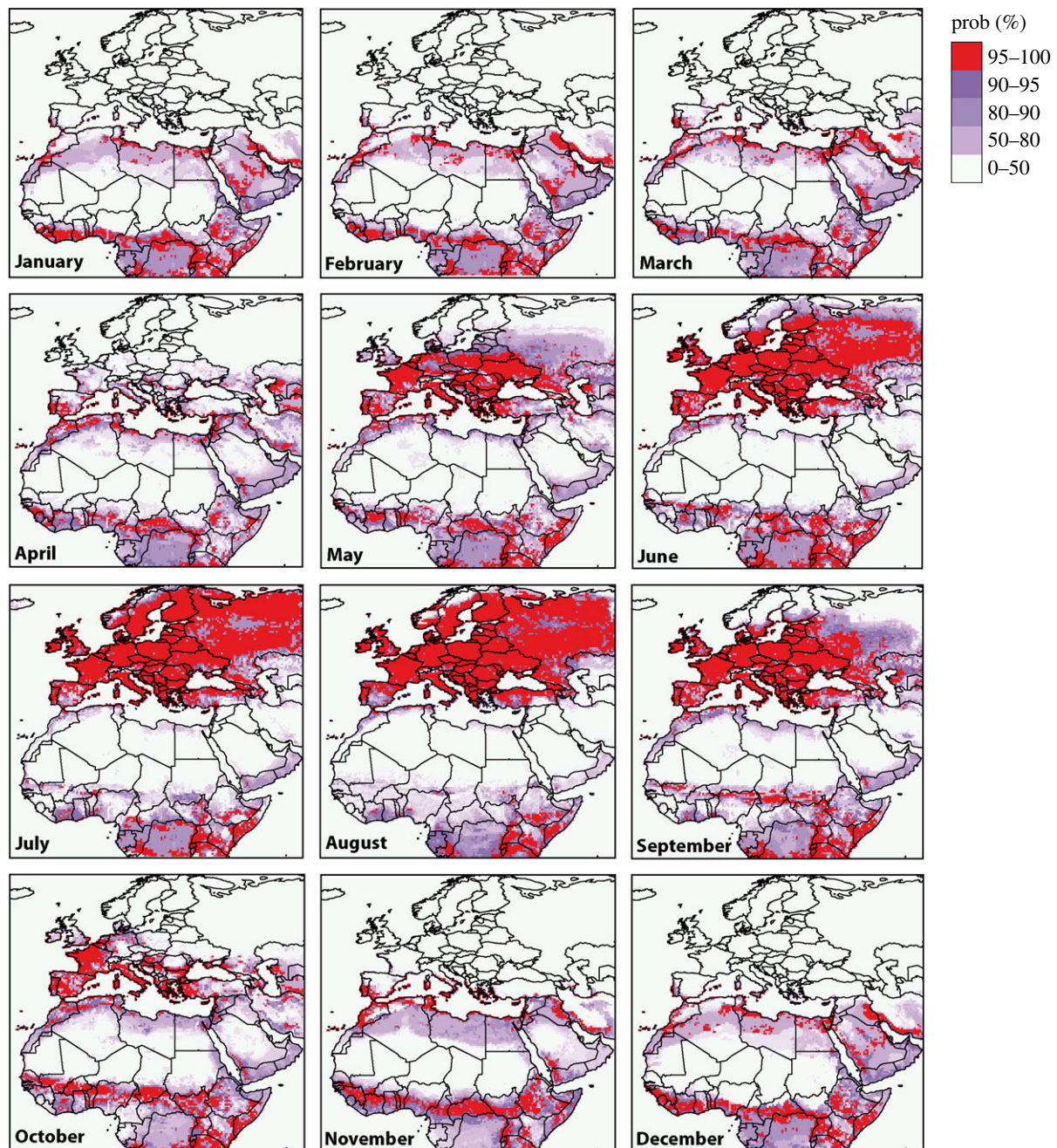


Figure 2. Monthly projections of the ensembled species distribution models. Colour denotes probabilities from 0 (white) to 100 (red). (Online version in colour.)

was the most important among the non-climatic variables (10.38%). The density plot curve for *t_{mp}* suggested that *V. cardui* prefers breeding areas with daily temperatures ranging from approximately 11° to 23°, and the *frs* curve suggested that no areas with high frosting frequency (greater than 4 days per month) are suitable for the species. Suitable water vapour pressure conditions are reached at values higher than 10 kPa, which may relate to two limiting factors for *V. cardui*: (1) cold temperatures and (2) the lack of available fresh host plant in the driest areas and periods. The density plot curve for *bareland* revealed that the species is highly adapted to semi-deserts and even deserts, which can probably be exploited when highly ephemeral plant resources are present in these areas (electronic supplementary material, figure S1).

(b) Monthly predictions

Predicted monthly distributions were variable within the entire migratory range, both in geographical distribution

and in total range extension (figure 2). The model output generally matched the expected distributions for the best-known regions and months (Europe from April to August). Predicted monthly distributions from September to March agreed with the working hypothesis scenario [8,10]: (1) a latitudinal shift southwards where distributions are initially on both sides of the Sahara but are restricted to equatorial latitudes in December and January and (2) that reverses from February to March to expand again in the north. The curve of monthly probability summation shows higher values from May to September (maximum peak in July) and lower values from December to April (minimum in January) (figure 3c). When considering monthly distributions in three latitudinal slices independently (Sahara and Afrotropics, Mediterranean and Middle East, and Central and North Europe), most variability was due to a massive peak in Central and North Europe (figure 3b) from May until September. The occurrence probability at this latitudinal slice becomes zero from December to March. Occurrence probabilities in the Sahara and Afrotropics remain fairly

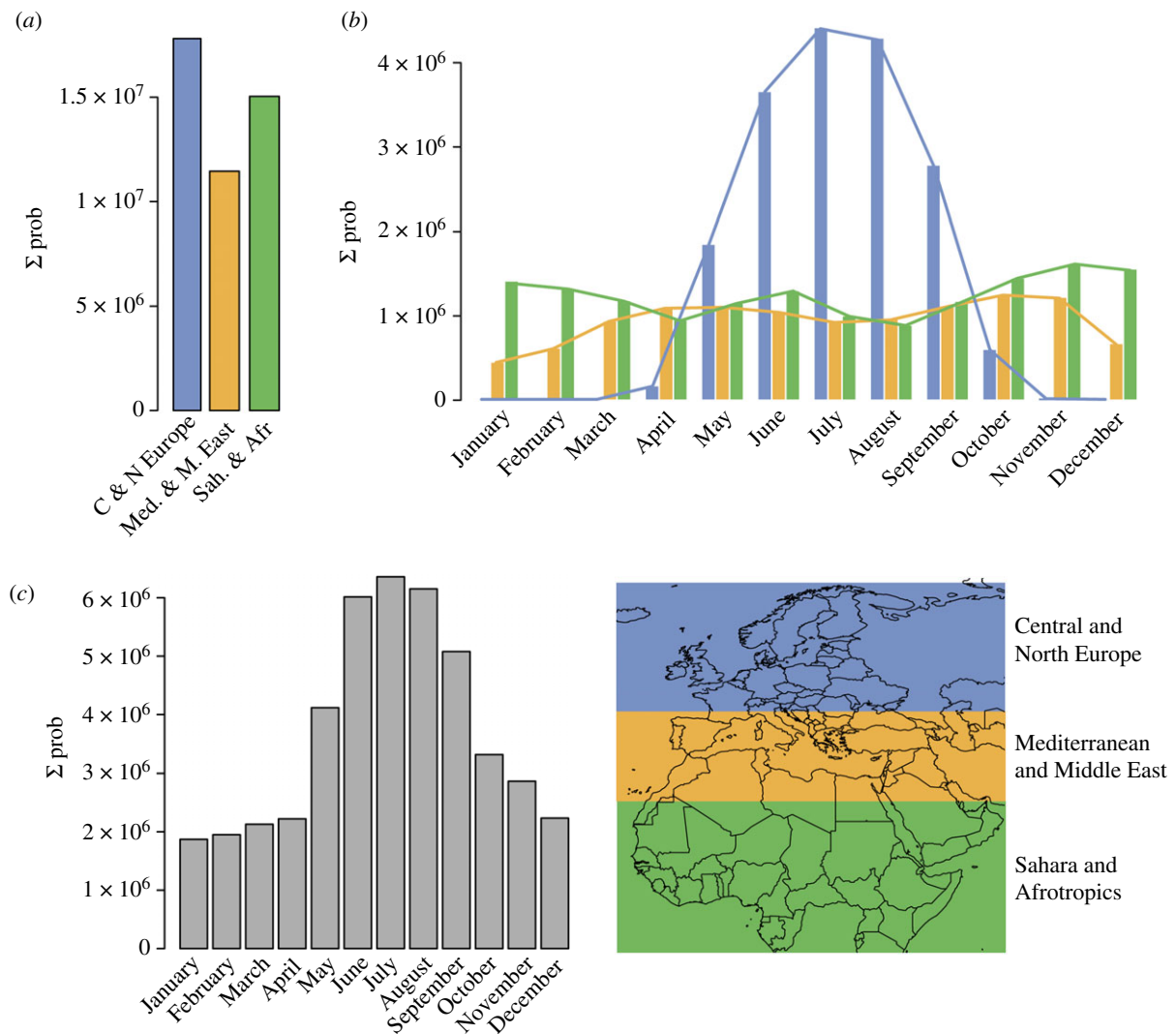


Figure 3. Variation of predicted suitable habitat (presence probability summations) for *Vanessa cardui* in the Palearctic-African migratory range: (a) yearly overall summations along latitudinal divisions, (b) monthly overall summations along latitudinal divisions, and (c) monthly overall summations for the entire range. Colours represent three latitudinal divisions as represented in the map: blue (Central and North Europe), yellow (Mediterranean and Middle East), and green (Sahara and Afrotropics). (Online version in colour.)

constant along the year, but with minor fluctuations showing the lowest values in April and August. The Mediterranean and the Middle East also show progressive but minor fluctuations, indicating two peaks of occurrence in April and October, and the lowest values from December to February (figure 3b).

4. Discussion

(a) Predicted breeding areas and migratory cycle for *V. cardui*

The resulting model predicts monthly transitions of breeding habitat, mostly latitudinal, which support the view that the successive generations of *V. cardui* are generally forced to move (figure 2; see movie in electronic supplementary material, File S2). For the majority of the year, suitable breeding habitat is geographically connected from one month to the next, suggesting that progressive latitudinal short-range migratory movements are likely to occur. By contrast, two major breakdowns in such spatio-temporal progression occur in autumn and spring, which are reflected in small niche overlap (figure 4), with gains of suitable breeding habitat in the Afrotropics and the Palearctic, respectively (figure 3).

Central and North Europe represent an area with almost geographically continuous resources for the species to breed between May and September (figure 3b). This temporal habitat triples the size of the hosting capacity of the Afrotropics or the Mediterranean during their annual maximum peaks. Contrasting to the areas of the migratory range at southern latitudes, Central and North Europe becomes unsuitable during part of the year (November to March). Other migratory insects have also been shown to undergo significant population increases during the summer months in Europe, such as the silver Y moth [24] and some migratory hoverflies [25]. Thus this region becomes a time-constrained but extraordinary resource for the Painted Lady butterfly. Relevant gains in the Afrotropics are also observed starting in September, with resource pulses suitable to the species appearing in the Sahel and moving south until a decrease in January.

Spatio-temporal fluctuations of gains and losses in overall suitable habitat (figure 3c) may affect species population size dynamics. In extreme cases, these could lead to both population explosions and bottlenecks. The extraordinary window of resources opening in Europe during the summer months could provide the opportunity for the species to yearly increase population size. Because this large habitat suitable for

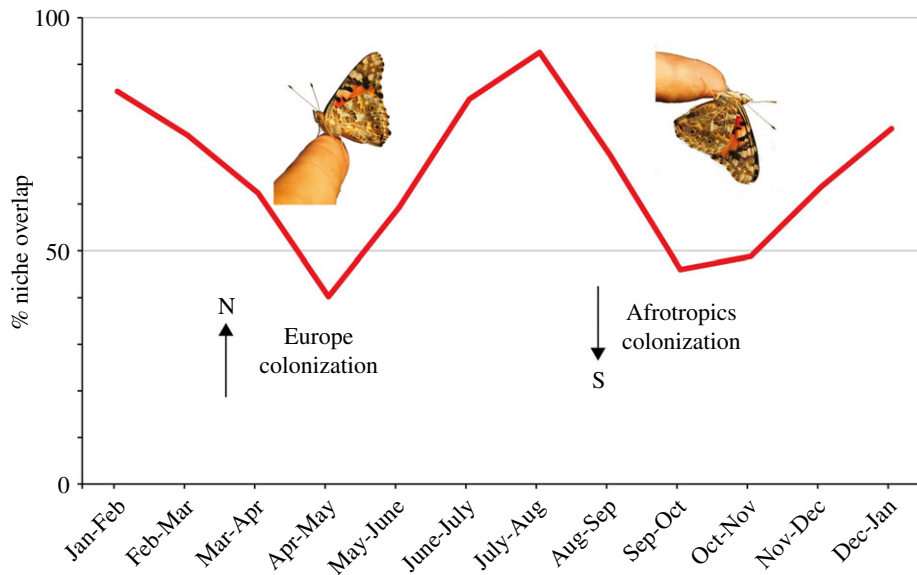


Figure 4. Percentage of niche overlap between successive months. Timing of the main northbound and southbound intercontinental migration events are indicated. (Online version in colour.)

reproduction is relatively sustained in time (from May to September) and involves multiple generations (possibly 3–4), late summer generations could potentially build up their numbers. This yearly cycle of potential population increase during summer is ecologically reasonable taking into account that these individuals are about to embark on a long-distance migration of up to 4000 km [8,9], which most likely will have a heavy toll on population numbers. On the other hand, the original population numbers colonizing Europe are necessarily dependent on the numbers of previous generations at lower latitudinal breeding grounds, which are more likely to vary from year to year depending on climatic conditions. Indeed, annual systematic counts for Europe show important fluctuations in population densities [26]. Thus, despite the increase of habitat availability during the summer, population numbers in Europe may not become notoriously large every year, but it could function as a recovery from previous bottlenecks. Moreover, habitat quality in Europe at present may not be as good due to anthropogenic alterations, like intensive agriculture and pesticide use. Population bottlenecks in migratory species can occur during the breeding phase because of restricted availability of breeding habitat, but also during the migratory phase due to harsh conditions along migratory movements. In the case of *V. cardui*, the timeframe between December and April shows the lowest values of overall suitable breeding habitat (figure 3c), which suggests that this period is the most vulnerable for the species to suffer yearly reproductive bottlenecks. Moreover, crossing the Sahara northwards may represent additional pressure, considering that the prevalent winds in this period are the southbound Harmattan.

The predicted pattern of successive losses and gains in suitable breeding areas matches migratory events that are known to recurrently occur between these areas: (1) European migrations southwards in the sub-Saharan during the autumn (September to November) [8] and in NW Africa (October to November) [9,27], (2) early spring northwards trans-Saharan migrations in North Africa and the Mediterranean region (February to March) [10,28], (3) further spring colonizations of Europe from North Africa (April–May) [26,28], (4) progressive northward shifts within Europe during spring and

summer [26]. The Sahara Desert, predicted by the model as non-suitable habitat, functions as a large natural barrier forcing migratory movements to be much longer than movements occurring at other latitudes. Main trans-Saharan crossings coincide with large resource pulses emerging in both sides of the desert. Indeed, the lowest percentages of niche overlap between consecutive months occur during the periods involving trans-Saharan migrations (figure 4).

(b) Breeding dynamics in the Afrotropics: the missing generations

There is no substantial breeding data available for the species in the Afrotropics from December to February. The current conceptual model for the Palearctic-African migratory system establishes that the Afrotropics may host the core populations from October to February, approximately, before the Mediterranean region is replenished from the south [8,10]. In addition, northwest Africa (Maghreb) also harbours massive populations during October/November and March/April [26,27]. Breeding areas have been identified in the Sahelian belt, tropical savannah, and Ethiopian highlands in October and November. The offspring of these generations have been hypothesized to progressively migrate farther south until a reverse of their migration direction in February [8]. This last step has been recently demonstrated by means of stable isotope evidence [10]. The spatio-temporal model presented here allows the prediction of solid hypotheses about the most suitable areas for the species to breed during the European winter months and, as a consequence, to propose the potential southernmost areas that the Palearctic-African cycle may reach.

The model shows, from September to January, a latitudinal belt of high probabilities progressively shifting south from the Sahel to the equator (figure 2). This transition coincides with a retreat of the Intertropical Convergence Zone (ITCZ), and the associated rains responsible for seasonality and vegetation growth in Africa [29]. The model predicts that the species could breed as far south as the equatorial latitudes between December and February, particularly in East Africa. High-probability predictions also exist for this period in West

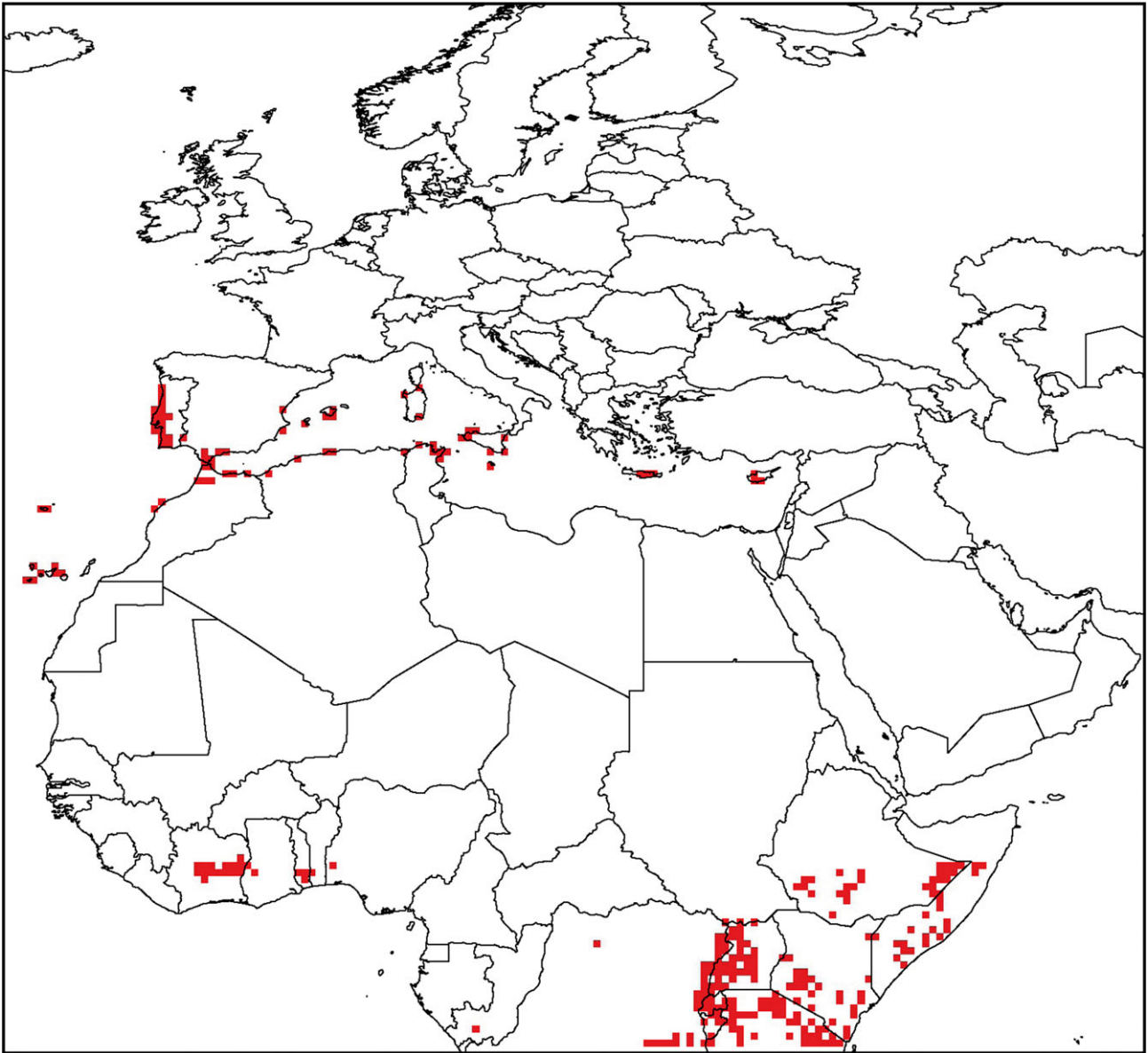


Figure 5. Regions (in red) showing year-round permanent high probabilities (greater than 90%) of suitable breeding habitat for *Vanessa cardui*. (Online version in colour.)

Africa, following the coast from Guinea to Nigeria. The highlands in Cameroon and the savannah in the Central African Republic and northern Democratic Republic of the Congo also reach high probabilities of occurrence. This scenario agrees with the hypothesis that the overall migratory circuit undertaken by the annual successive generations may encompass up to 15 000 km, from the equator (e.g. Kenyan or Cameroon highlands) to northern Scandinavia.

The model also predicts suitable breeding habitat during the winter months in certain areas of Northern Africa, and in a minor proportion in the circum-Mediterranean coastal areas. Some populations are suspected to spend the entire winter months in these areas [10,27,30]. However, predicted suitable breeding habitat is higher in the Afrotropics than in the Mediterranean region between December and February, supporting the view that the population core of the species breeds in winter in the sub-Saharan in higher numbers than known sites in North Africa or Mediterranean Europe. The Middle East may also host important areas suitable for breeding during the winter, and may contribute with population stock replenishing Eastern Europe in spring.

As *V. cardui* occurs in the Southern African Hemisphere too, other migratory pools may widely circulate within this region only, and due to an inverted seasonality regime, remain spatio-temporally unlinked to the populations transitioning between Europe and the Northern Hemisphere of Africa (electronic supplementary material, figure S2).

(c) Suitability and occupancy

Even though *V. cardui* is a polyphagous species, feeding on multiple plants of several families, the species might not find abundant host plants in some regions that are identified in the model to be highly suitable climatically, thus the model prediction may over-represent probabilities in certain areas. For example, widespread predicted high probabilities in desert areas in Western Sahara or in Arabia during the winter months seem, *a priori*, unrealistic. However, we can not discard the possibility that the species breed in localized spots in these regions (e.g. oasis, irrigated fields), but the availability of fresh plants must be low and fragmented so as to support the extensive areas that the model predicts with

high likelihood. One possibility to consider is the desert blooming phenomena that can occur in exceptional years, an ephemeral situation that *V. cardui* can take strong advantage of. Unusual heavy rains in desert areas have been shown to correlate with the observation of adult outbreaks in the Mediterranean [26,30,31]. Pollen metabarcoding analyses of *V. cardui* migrants in the Mediterranean coast of the Iberian Peninsula have found pollen load from desert-adapted plants of Saharo-Arabian distribution [28]. However, these occasional resources are probably not able to sustain the basal population size of the species every winter.

Regions at northern latitudes could be over-represented during the summer months as well, reaching latitudes in our model as north as 70° in July, in northern Finland and Norway. It is important to note that developmental time at those latitudes might be significantly slowed down with respect to those at southern European latitudes, because of lower temperatures. Developmental times at temperatures ranging from 10 to 20°C (night-day) may extend the completion of the entire cycle by to up to two months (G Talavera 2016, unpublished data), likely too long because suitable conditions are virtually limited to July in these northern latitudes.

Interestingly, although high probabilities are obtained in an extensive region encompassing rainforest in equatorial Africa, those predicted suitable areas with dense rainforest are relatively small. The presence of rainforests, where *V. cardui* is very rarely observed, appears to be an important limiting factor for this species. Rainforest in West Africa (from Guinea to Ghana) is however shown to be highly suitable from December to February. Even so, forest fragmentation in these countries is remarkable at present, and forest alternates with agricultural land, which may be beneficial to *V. cardui* breeding requirements.

(d) Permanent habitat and potential behaviour change

Building spatio-temporal models offer the possibility to identify sites that might be permanently suitable for reproduction along the entire year. If such sites exist and are large enough, part of the migratory populations could, in theory, exhibit sedentary behaviour. Our model predicts a few spots where *V. cardui* could reproduce all the year-round (figure 5). These areas mainly include: (1) highlands and coastal areas in eastern Africa (e.g. Uganda, Rwanda, Kenya, Somalia), (2) western Africa (Cote d'Ivoire and Ghana), (3) the Atlantic coast in SW Portugal, (4) the Strait of Gibraltar, (5) western Canary Islands and Madeira, and (6) isolated spots along the Mediterranean coast (Tunisia, Morocco, Algeria, Balearic Islands, Sicily, Corsica, Crete, and Cyprus) (figure 5). All these areas, however, are relatively small and heavily recolonized once or twice every year, including archipelagos like the Canary Islands and Madeira [32]. The possibility that some butterflies remain in the area and avoid migration needs to be studied. However, these would represent a very small fraction of the total population and it is very unlikely that sedentary behaviour becomes fixed, which would only be favoured in cases of strong spatio-temporal isolation. Such a situation could eventually revert if the overall permanent suitable extent grows substantially in the future, as a consequence of global warming. In the case where the Mediterranean becomes largely suitable for overwintering, migratory movements could become shorter and restricted

to the Palearctic, and populations at both sides of the Sahara become fragmented.

5. Conclusion

By using time-partitioned variables and time- and behaviour-partitioned records it is possible to model ecological niches for migratory animals in general, and multigenerational insects in particular. The strategy here proposed allows investigating the potential breeding preferences of the Painted Lady butterfly in Africa, an unresolved spatio-temporal gap of knowledge within their annual migratory range. The model predicts that the species might breed as far south as equatorial latitudes between December and February. This scenario agrees with the view that the overall migratory circuit undertaken by the annual successive generation may encompass up to 15 000 km, from the equator (e.g. Kenyan highlands) to northern Scandinavia. Two major breakdowns in monthly gains and losses of suitable breeding habitat coincide with the main known long-range migratory movements colonizing the Afrotropics (September–October) and Europe (April–May). Predicted breeding habitat is relatively stable along the year in the Sahel-Afrotropics region, and the Mediterranean-Middle East region displays a higher fluctuation reaching a minimum in January. By contrast, Central and North Europe show high probabilities from May to September and represent a huge spatio-temporal window for suitable habitat that triples those in the south. January and February display the lowest suitability values, which suggest that these months are the most vulnerable for the species to suffer yearly population bottlenecks. This, in addition to the fact that these populations are facing imminent northwards trans-Saharan migrations at high cost, highlights the relevance of the sub-Saharan winter population stock in sustaining the migrations of the species into Europe. The model predicts few and scattered locations with high probabilities which allow reproduction all year-round, and discard the possibility of a transition of the species as a whole towards permanent residency. In addition to the Palearctic-African migratory cycle, this study also reveals the potential of ecological niche modelling to provide insight into migratory behaviour of *V. cardui* in other geographical regions, as well as of other migratory insects. Given the importance of the vast data of immatures here used, coming from multiple programmes, databases, and platforms mostly nourished by citizen scientists, this study exemplifies the value of citizen science in contributing to large-scale projects, both across geography and time.

Data accessibility. All data are accessible via the electronic supplementary material.

Authors' contributions. G.T. and M.M. conceived the study and gathered data; M.G. and G.T. performed modelling analyses; all authors contributed to writing the manuscript.

Competing interests. We declare we have no competing interests.

Funding. This work was funded by the Committee for Research and Exploration of the National Geographic Society (grant no. WW1-300R-18), the British Ecological Society (grant no. LRB16/1015), the PRIC grants of the Fundació Barcelona Zoo, and by project CGL2016-76322 (AEI/FEDER, UE) to G.T. G.T. is supported by the MINECO programme Juan de la Cierva Incorporación (IJCI-2016-29083), the Marie Curie Actions FP7-People-2013 IOF (project 622716) and 'La Caixa' Foundation (grant LCF/BQ/PR19/11700004).

Acknowledgements. We thank Marta Barbieri, Dubi Benyamini, Giuseppe Cagnetta, Mariano Cuadrado, Sylvain Cuvelier, Leonardo Dapporto, Vlad Dinca, Martin Gascoigne-Pees, Marco Gottardo, Cosmin

References

- Ward EJ, Pess GR, Anlauf-Dunn K, Jordan CE. 2012 Applying time series models with spatial correlation to identify the scale of variation in habitat metrics related to threatened coho salmon (*Oncorhynchus kisutch*) in the Pacific Northwest. *Can. J. Fish. Aquat. Sci.* **69**, 1773–1782. (doi:10.1139/f2012-096)
- Ward EJ, Jannot JE, Lee Y-W, Ono K, Shelton A, Thorson JT. 2015 Using spatio-temporal species distribution models to identify temporally evolving hotspots of species co-occurrence. *Ecol. Appl.* **25**, 2198–2209. (doi:10.1890/15-0051.1)
- Mannocci L *et al.* 2017 Temporal resolutions in species distribution models of highly mobile marine animals: recommendations for ecologists and managers. *Divers. Distrib.* **23**, 1098–1109. (doi:10.1111/ddi.12609)
- Martínez-Meyer E, Peterson AT, Navarro-Sigüenza, AG. 2004 Evolution of seasonal ecological niches in the *Passerina* buntings (Aves: Cardinalidae). *Proc. R. Soc. B.* **271**, 1151–1157. (doi:10.1098/rspb.2003.2564)
- Nakazawa Y, Peterson AT, Martínez-Meyer E, Navarro-Sigüenza AG. 2004 Seasonal niches of nearctic-neotropical migratory birds: implications for the evolution of migration. *Auk* **121**, 610–618. (doi:10.1642/0004-8038(2004)121[0610:SNONMB]2.0.CO;2)
- Fink D, Hochack WM, Zuckerberg B, Kelling ST. 2011 Modeling species distribution dynamics with spatiotemporal exploratory models: discovering patterns and processes of broad-scale avian migrations. *Procedia Environ. Sci.* **7**, 50–55. (doi:10.1016/j.proenv.2011.07.010)
- Batalden RV, Oberhauser K, Peterson AT. 2007 Ecological niches in sequential generations of Eastern North American Monarch Butterflies (Lepidoptera: Danaidae): the ecology of migration and likely climate change implications. *Environ. Entomol.* **36**, 1365–1373. (doi:10.1603/0046-225X(2007)36[1365:ENISGO]2.0.CO;2)
- Talavera G, Vila R. 2016 Discovery of mass migration and breeding of the painted lady butterfly *Vanessa cardui* in the Sub-Saharan: the Europe–Africa migration revisited. *Biol. J. Linn. Soc.* **120**, 274–285. (doi:10.1111/bj.12873)
- Stefanescu C, Soto DX, Talavera G, Vila R, Hobson KA. 2016 Long-distance autumn migration across the Sahara by painted lady butterflies: exploiting resource pulses in the tropical savannah. *Biol. Lett.* **12**, 20160561. (doi:10.1098/rsbl.2016.0561)
- Talavera G, Bataille C, Benyamini D, Gascoigne-Pees M, Vila R. 2018 Round-trip across the Sahara: Afrotropical Painted Lady butterflies recolonize the Mediterranean in early spring. *Biol. Lett.* **14**, 20180274. (doi:10.1098/rsbl.2018.0274)
- Dingle H. 1972 Migration strategies of insects. *Science* **175**, 1327–1335. (doi:10.1126/science.175.4028.1327)
- New M, Hulme M, Jones PD. 1999 Representing twentieth century space-time climate variability. Part 1: development of a 1961–90 mean monthly terrestrial climatology. *J. Clim.* **12**, 829–856. (doi:10.1175/1520-0442(1999)012<0829:RTSCST>2.0.CO;2)
- New M, Hulme M, Jones PD. 2000 Representing twentieth century space-time climate variability. Part 2: development of 1901–96 monthly grids of terrestrial surface climate. *J. Clim.* **13**, 2217–2238. (doi:10.1175/1520-0442(2000)013<2217:RTSCST>2.0.CO;2)
- New M, Lister D, Hulme M, Makin I. 2002 A high-resolution data set of surface climate over global land areas. *Clim. Res.* **21**, 1–25. (doi:10.3354/cr021001)
- Mitchell TD, Jones PD. 2005 An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.* **25**, 693–712. (doi:10.1002/joc.1181)
- Harris I, Jones PD, Osborn TJ, Lister DH. 2014 Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *Int. J. Climatol.* **34**, 623–642. (doi:10.1002/joc.3711)
- Arino O, Ramos-Perez JJ, Kalogirou V, Bontemps S, Defourny P, Van Bogaert E. 2012 *Global land cover Map for 2009 (GlobCover 2009)*. Pangaia: European Space Agency (ESA) & Université Catholique de Louvain (UCL).
- Allouche O, Tsoar A, Kadmon R. 2006 Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **43**, 1223–1232. (doi:10.1111/j.1365-2664.2006.01214.x)
- Swets KA. 1988 Measuring the accuracy of diagnostic systems. *Science* **240**, 1285–1293. (doi:10.1126/science.3287615)
- Thuiller W, Lafourcade B, Engler R, Araújo MB. 2009 BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography* **32**, 369–373. (doi:10.1111/j.1600-0587.2008.05742.x)
- Thuiller W, George D, Engler R. 2014 biomod2: Ensemble platform for species distribution modelling. R package version 3.1–64. See <http://CRAN.R-project.org/package=biomod2>.
- Heibl C. 2012 phylodlim: integrating phylogenetics and climatic niche modelling. See <http://cran.r-project.org/web/packages/phylodlim/>.
- Warren DL, Glor RE, Turelli M. 2008 Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* **62**, 2868–2883. (doi:10.1111/j.1558-5646.2008.00482.x)
- Chapman JW, Bell JR, Burgin LE, Reynolds DR, Pettersson LB, Hill JK, Bonsall MB, Thomas JA. 2012 Seasonal migration to high latitudes results in major reproductive benefits in an insect. *Proc. Natl Acad. Sci. USA* **109**, 14 924–14 929. (doi:10.1073/pnas.1207255109)
- Wotton KR, Gao B, Menz MHM, Morris RKA, Ball SG, Lim KS, Reynolds DR, Hu G, Chapman JW. 2019 Mass seasonal migrations of hoverflies provide extensive pollination and crop protection services. *Curr. Biol.* **29**, 2167–2173. (doi:10.1016/j.cub.2019.05.036)
- Stefanescu C *et al.* 2013 Multi-generational long-distance migration in insects: studying the painted lady butterfly in the Western Palaearctic. *Ecography* **36**, 474–486. (doi:10.1111/j.1600-0587.2012.07738.x)
- Stefanescu C, Puig-Montserrat X, Samraoui B, Izquierdo R, Ubach A, Arrizabalaga A. 2017 Back to Africa: autumn migration of the painted lady butterfly *Vanessa cardui* is timed to coincide with an increase in resource availability. *Ecol. Entomol.* **42**, 737–747. (doi:10.1111/een.12442)
- Suchan T, Talavera G, Sáez L, Ronikier M, Vila R. 2018 Pollen metabarcoding as a tool for tracking long-distance insect migration. *Mol. Ecol. Res.* **19**, 149–162. (doi:10.1111/1755-0998.12948)
- Pedgley DE, Reynolds DR, Tatchell GM. 1995 Long-range insect migration in relation to climate and weather: Africa and Europe. In *Insect migration: tracking resources through space and time* (eds VA Drake, AG Gatehouse), pp. 3–29. Cambridge, UK: Cambridge University Press.
- Benyamini D. 2017 A swarm of millions of *Vanessa cardui* (Linnaeus, 1758) in winter–spring 2015–2016 in the south-east Mediterranean – the missing link. *Atalanta* **48**, 103–128.
- John E. 2001 Migration of *Vanessa cardui* (L.) (Lep: Nymphalidae) through Cyprus, March 2001. *Entomol. Rec. J. Var.* **113**, 269–281.
- Wiemers M. 1995 The butterflies of the Canary Islands. A survey on their distribution, biology and ecology (Lepidoptera: Papilionoidea and Hesperioidea) (second part). *Linneana Belg.* **115**, 87–118.