Latitude, longitude, and the evolution of Iberian butterfly faunistics (Lepidoptera). A preliminary test for shifts in distribution areas in the Western Mediterranean

Enrique García-Barros & Helena Romo

Department of Biology, Universidad Autónoma de Madrid, E-28049 Madrid, Spain
Contact: garcia.barros@uam.es; helena.romo@uam.es

There is increasing evidence that human-induced climate change is playing a role in recent shifts in the geographic distribution of organisms, including butterflies. Tracing back distributional trends along the last two centuries requires a relatively exhaustive and unbiased faunistic documentation, and hence the kind of information that might not be available from all European countries. Further, it may prove difficult to ‘dissect’ any apparent historical patterns to discriminate between latitudinal trends, historical sampling biases, and direct effects of changing land use.

A relatively exhaustive database was recently compiled from collection and published data as a basis for the provisional distribution atlas of Iberian butterflies (that is, the Papilionoidea and Hesperioidea of the continental territories of Spain and Portugal). This compilation contained ca. 285,000 records with reliable date and year data, or that were at least attributable to a precise decade. Each record refers to one or more individuals of any species associated to a different combination of site, date, and collector, from the last 210 years. All the data were geo-coded with reference to the UTM projection system, and the 10 km squares were used as operative geographic units. These data were explored in order to discover possible geographic trends in the distributions of the butterflies within the Iberian Peninsula.

This preliminary approach was based in standard regression techniques; these were applied to seek for any correlation between the recorded dates and the associated geographic location data (latitude and longitude), both within each species and in the whole data set. The basic procedure was to seek significant correlations between latitude or longitude and the year reported, using as much information as possible from the original data. The effect of the two geographic variables on the variable ‘year’ were tested simultaneously by means of multiple regression; this was necessary because, given the profile-shape of the study area, shifts in the average value of one of the two variables often imply a correlated change in the other. The species that occupy less than twenty 10 km squares were excluded from the analyses, as were all the skippers (family Hesperiidae) due to uncertainties about the reliability of the oldest records. However, all those data were retained for the global analyses where an estimate of the amount of data per square, and not the species, were required. Specific comparisons were feasible for 180 species.
Applying a simple multiple regression model yielded significant results for most species (155 out of 180 showed a significant correlation between year and either latitude, longitude, or both). This, however, was very obviously a consequence of a general trend towards a SW to NE shift in the centre of gravity of the faunistic efforts along the last two centuries: across the whole data set, both latitude and longitude were slightly (but highly significantly, \( P < 0.0001 \)) correlated to the year of observation (respectively \( r = 0.08 \) and \( r = 0.17 \)), and the two geographic position variables were inter-correlated (\( r = 0.37, P < 0.0001 \)). Four methods were applied to statistically control these general effects. (a) First, the raw individual positional data were divided by the global year-specific mean latitude or longitude. (b) Second, the residuals (standardised residual values) from the regressions of latitude and longitude on year (calculated from the whole data set) were estimated for each record, and used as the input for new regressions at the species level. (c) Third, the General Lineal Models procedure was applied; two additional variables were introduced, i.e., ‘period’ (four different periods of time) and ‘large region’ (eight large subareal units intended to account for possible interregional biases). The residuals were retained for further analysis. (d) And last (and partly supported by the results of the former), a nested ANCOVA model with the variables described formerly, with ‘period’ nested within ‘large region’, and ‘year’ as a continuous predictor. Again, the residuals were saved.

The overall model results from the analyses (b-d) were able to explain important amounts of the variance. However, only the residual values from the two last treatments (c and d) demonstrated full temporal independence. In other words, the historic overall trends in the geographic position variables almost disappeared (\( r = 0.00 \) in GLM procedure, \( r = 0.01 \) in nested ANCOVA). The correlation between latitude and longitude decreased without completely disappearing in the last two analyses (\( r = 0.13, P < 0.0001 \) in both). Based on the residuals from these two treatments, the species-specific correlations were more conservative than after the regressions done on the raw data, or on the residuals from procedures a and b. Even so, significant results still arose from the data of 104 species (with slight differences between the data from methods c and d).

The results certainly represent evidence for changes in the size and shape of the species’ geographic distributions in Spain and Portugal. However, any interpretation must be cautious. It is evident that any potential latitudinal (and/or longitudinal) shift in the distribution of any of the Iberian species must have progressed along with the development of faunistic studies in the area; in fact, the last is still far from complete, at least in terms of the percentage of territory that has been thoroughly prospected. A vast majority of the records (67%) was made after 1976. Most of the oldest data are bibliographic and lack detailed specimen counts, which implies that the volume of effort carried out in past decades becomes progressively underestimated as the original material disappears. In addition, important changes in the patterns of land use have prevailed across most of the study area in relation with the process of industrial development along the XXth Century, and hence ---probably--- slightly predating the spread of the bulk of collection activities. Finally, the patterns of geographic distribution of faunistic efforts seem to have followed a roughly SW to NE orientation. This is partly attributable to an important development of extensive work in the northern part of Spain in the decade of 1980. Although in theory the residuals used in the calculations (methods c and d above) were free of geographic bias, more detailed analyses, together with a species-by-species intuitive reconsideration, are necessary to ensure that any patterns detected do not just reflect the rhythm of discovery of new sites (thus, the process of ‘filling’ the still imperfectly known geographic ranges) for the species involved.
To mention a few specific cases, we concentrate on those species for which a significant, and relatively high ($r>0.3$) partial correlation exists between the year of observation and latitude or longitude, and with similar results in the two approaches derived from methods c and d.

**Trend** (spreading towards)/ Species

**N:** *Classiana eunomia, Erebia pronoe, Pseudochezara hyppolite, Minois dryas, Maculinea aleon, Agriades pyrenaicus, Plebejus hespericus.*

**NE:** *Erebia zapateri.*

**E:** *Colotis daira, Vanessa virginiensis, Tarucus teophrastus, Polyommatus nivescens.*

**SE:** *Satyrs furia, Danaus chrysippus.*

**S:** *Brenthis ino.*

**SW:** *Erebia pandrose, Danaus plexippus.*

**W:** *Erebia palarica, Cacyreus marshalli.*