EUROPEAN AND MAGHREB BUTTERFLY TRAIT DATABASE

Compiled by: Joseph Middleton Welling, Leonardo Dapporto, Enrique García-Barros, Martin Wiemers, Piotr Nowicki, Elisa Plazio, Simona Bonelli, Michele Zaccagno, Martina Šašić, Jana Liparova, Oliver Schweiger, Alexander Harpke, Martin Musche, Josef Settele, Reto Schmucki,Tim Shreeve

Definitions of the traits included within the European and Maghreb butterfly trait database

Individual traits were defined prior to the beginning of data collation so that coding could be unambiguous. Most traits are coded as binary variables, but a minority are continuous or categorical. Most traits included in the dataset are divided up into multiple sub-traits. For example, the trait ‘overwintering stage’ comprises four binary sub-traits each of which indicates one stage of a species’ life cycle: egg, larvae, pupa and adult. A species can have any combination of 0 and 1 for each of these sub-traits. This allows for the coding of trait plasticity across a species’ range.

Overwintering stage: the life history stage (egg, larvae, pupae or adult) in which a species hibernates during the winter in a state of suspended or highly reduced activity. In cases where a species overwinters as a caterpillar within the egg this is coded as egg stage overwintering. Species that are present all year round throughout their range and lack an overwintering stage are coded as zero for all the states of this trait.

Overwintering location: place where a species passes the winter during a normal life-cycle, and any structures that are specifically used by the species as a location for overwintering. If a species overwinters in a developmental stage that is strongly associated with a particular environmental structure then the coding for ‘overwintering location’ will generally overlap with codings relating to that particular developmental stage. For example, a species that overwinters in the pupal stage will have similar trait codings for both ‘overwintering stage’ and ‘pupal location’.

Voltinism: the number of generations a species has in a year. For the binary coding of this trait, species which have one generation in a year are coded as ‘univoltine.’ Species that complete a generation every two years are coded as ‘biennial.’ Species described as having a ‘partial second generation’ are coded as ‘univoltine+partial.’ Species that have two generations in a year are coded as ‘bivoltine’. Species with more than two generations in a year are coded as ‘multivoltine.’. Whilst individual life-cycles may take two years in sites where biennial life-cycles occur, adults appear annually, the binary data precisely describes variation in the number of annual development cycles within a species.

Forewing length: is defined as the distance from the forewing apex to the body in a set specimen. **Wingspan**: is defined as the distance between the apices of the two forewings in a set specimen. These traits are regarded as good proxies for mobility in butterflies (Sekar, 2012). Data for these traits was primarily obtained from Higgins and Riley (1980), Bink (1992), Tennent (1996), Newland et al. (2015), and Paolucci (2013) for Forewing length and from Tshikolovets (2011) for Wingspan. These sources tend to contain averages for this trait from many specimens from across Europe and because of this, it was not generally necessary to take into account the variation in forewing length within the region of study. As sexual dimorphism frequently occurs in butterflies, forewing length and wingspan are coded separately for both males and females. For each sex, the maximum, minimum, mean and range are coded for both of these traits. Forewing length and wingspan are clearly highly correlated but for the reasons discussed above can vary among sources. Moreover, no single source provides data for all the species. For this reason, we also rovide a single size measure “**Wing index**” representative for all sources and covering the widest number of taxa as follows. We imputed the missing values of the six average forewing size and wingspan obtained by four different sources (two for males and females separately and two from males and females together) by using the “mice” function of the “mice” R package. The algorithm imputes an incomplete variable by generating plausible values based on other variables in the data by Multivariate Imputations by Chained Equations (MICE) (Van Buuren & Groothuis-Oudshoorn, 2011). Then, we applied a PCA to the six measurements, thus obtaining a single factor with an eigenvalue higher than 1. The scores from this factor provided an overall size measurement for all species except three and has been indicated as “wing index”. The variation in wing size has been calculated as the difference between the maximum and the minimum size reported in a given source divided by the average reported in the same source. Similarly to size, the mice procedure and PCA has been applied to these data to obtain an overall “**Variation index**” measurement of size variability for each species.

Pupal location: the location the pupal stage is located in the environment during a normal life-cycle. When the pupa is described as ‘hanging’ from a particular structure then that structure is coded as the location, as it is assumed to be required for the pupa to successfully survive pupation.

Ant association: describes the range of ant species that species interact with during their life cycle (myrmecophily). Many of the ‘blues’ (Lyceanidae) have close symbiotic, commensal or parasitic relationships with ant species. This is divided into three sub-traits ‘monospecific’, ‘oligospecific and ‘polyspecific’ ‘Monospecific’ species are only associated with one species of ant. ‘Oligospecific species are associated with ants within a single genus. ‘Polyspecific species are defined as those that can utilise ant species from more than one genus. In Europe, myrmecophily is confined to the Lycaenidae and the Riodinidae. Information on this trait was largely based on studies carried out by Fiedler (1991) supplemented with more recent information taken from Settele et al. (2009).

Hostplant index: calculated from the total number of hostplant species, genera and families used by a species with the following formula:

$$HPI = \frac{1}{\sqrt{(N\ Hostplant\ species\ \times N\ Hostplant\ genera\ \times N\ hostplant\ families)}}$$

This index is based on the concept of taxonomic distinctiveness presented by Freitag and Van Jaarsveld (1997). The majority of the hostplant data used to calculate this index was taken from Essens et al. (2017). Records considered to be unreliable or those based only on labrearing experiments were not used to calculate values for this trait or for any other hostplant trait. Hostplant index ranges from 0 for species which are highly polyphagous to 1 for species that are completely monophagous, providing a quantitative measure of overall hostplant specificity.

Hostplant family: the taxonomic affinity of the hostplants that a species uses with the indication of the number of genera used for each family. Hostplant information used to code this trait and all subsequent traits relating to hostplant, were taken from field guides, the primary ones used being Lafranchis (2004), Tolman and Lewington (2007), and Tshikolovets (2011). The plant families adopted were those in the APG III taxonomy (2009).

Hostplant specificity: the range of hostplants that a species can use. 'Monophagous' species are defined as species that feed on a single hostplant species. 'Narrow oligophagous' butterfly species are those which feed on multiple hostplants which are found within one genus. 'Broad oligophagous' species utilise multiple hostplants within more than one genus but within one family. 'Polyphagous' species feed on hostplants from more than one family. The level of specificity of each species was determined by comparing the recorded hostplants with the APG III taxonomy (2009).

Hostplant type: the phenology of a species' hostplants. Each plant species (gathered from the literature as part of the coding of the 'hostplant family' trait) were checked against the Flora Europea (Tutin et al., 1993), which provides information on its phenology (biennial, annual, perennial). All the various hostplants of a particular species are coded, thus if a species has annual and biennial hostplants, both are coded as being hostplant types of that species.

Hostplant growth form: the growth form of a species' hostplants. Data was taken from Flora Europea (Tutin et al., 1993). 'Short herbs' are defined as non-woody species having a primary stem of less than 1m, and 'tall herbs' (non-woody) have a primary stem of greater than 1m in length. All the various hostplants of a particular species are coded for their growth form.

Hostplant part: the part (or organ) of the hostplant that the larvae of the species consumes. In the European butterfly fauna, several hostplant parts are commonly consumed, including the leaves, buds and flowers (Munguira et al., 2009).

Hostplant age: the growth stage of the hostplant chosen by adults to lay eggs on and the larvae subsequently consume. Young hostplants are those that have not yet flowered for the first time. Mature hostplants are those that have flowered at least once. All hostplants of European butterflies are angiosperms (Munguira et al., 2009) and so this was judged to be an effective indicator of hostplant age for this taxonomic group.

Larval environment: the location and vegetative structures on which the larvae can be found throughout their various instars. The coding of this trait substantially overlaps with 'egg laying location', 'overwintering location' and 'pupal location' but may also potentially contain a wider range of states compared with each of these other traits if the larvae are highly mobile. This overlap reflects the importance of the larval stage for many species of butterfly.

Hostplant patch size: the area of continuous hostplant required for egg laying to occur under typical circumstances. A 'small patch' is defined as a single hostplant. A 'medium patch' is defined as an area of hostplant larger than a single plant but less than $1m^2$ in spatial extent. A 'large patch' is defined as an area of hostplant greater than $1m^2$.

Egg laying type: the number of eggs that are laid by a particular species in a discrete location. This trait was divided into three categories ''Single' eggs are laid one per hostplant or one per hostplant part or are dropped in flight by the butterfly. In the field 'single' eggs may approximate small batches if different females deposit eggs in the same locations, thus the trait is defined by female laying strategy at the individual level. ‘Small batches’ are when eggs are laid together in groups of 2-8 eggs on one part of a hostplant in close proximity. ‘Large batches’ are groups >8 eggs laid at one time on the same part of the hostplant in close proximity. Typically, eggs in small or large batches are found in a very small area and the eggs are found touching in a contiguous cluster. The trait is coded as binary as individuals of some species may lay single eggs or small batches.

Egg laying location: the structures that eggs are laid on by a particular species and the environmental conditions under which the eggs are preferentially laid.

Adult feeding: the food sources used by adult butterflies. In general, types of adult feeding are divided into three broad groups by plant type, and also include a number of non-plant resources that adult butterflies can also feed on. The three plant types that are included in this trait are ‘herbs and herbaceous flowers,’ ‘ergot’ and ‘shrubs and trees.’ ‘Herbs’ in this instance includes all plants without a woody stem. ‘Shrubs and trees’ in this instance refer to generally larger plants with a woody stem. Sap feeding behaviour, feeding on decaying plant material and feeding on honeydew were coded as separate binary sub-traits. ‘Animal’ feeding primarily refers to feeding on animal dung but also may indicate feeding on animal carcasses. ‘Mineral’ feeding includes mud puddling behaviour and salt licks where adult butterflies will feed on concentrations of minerals or bodies of water and wet microhabitats.

Adult roosting: the structures a butterfly uses during the night and during periods of inclement weather or other unsuitable environmental conditions. Therefore, the most reliable way to get information on this trait was either to observe this in the field or take it from the literature (e.g. Fric, 2000).

Mate locating type: the type of behaviour the males of a species use to locate females. In butterflies three main types of mate locating behaviour are recognised, patrolling, perching and lekking (Dennis and Shreeve, 1988). ‘Patrolling’ behaviour involves the male performing exploratory flights in search of females. ‘Perching’ behaviour comes in two types: ‘Perching’ and ‘Territorial perching’ both of which involve males sitting on a physical structure or other resource and waiting for females to pass by, whereupon the males will generally fly up from the perch and investigate. In ‘territorial perching’ males will attempt to chase off other males and will not perch on the same structure as other males. In non-territorial perching, several males may share the same perching location. ‘Lekking’ is a behaviour where males aggregate at a specific location and compete for females at this location, the presence of other males being the cue for aggregation and competitive behaviours. Once at the lek they may perch or patrol and therefore ‘lekking’ will overlap with other mate locating behaviours and these behaviours are also coded when ‘lekking’ is recorded.

Mate locating location: the physical location where mate locating behaviour occurs including what (if any) resources are either required or preferred by males attempting mate locating behaviour. This trait describes the structures used as perching sites, and the environmental features and structures which are used as guidelines for patrolling flights as well as the over what resources patrolling flights take place. This trait also describes the location of putative leks.

Basking type: the wing orientation that the butterflies use relative to the sun when attempting to raise their body temperatures (i.e. during basking behaviour). ‘Dorsal absorption’ basking is defined as a basking behaviour where the wings are held open and the upper wing surfaces directly absorb solar radiation. The areas of the wings in these species are generally dark in colour to facilitate absorption of high levels of solar radiation (Clench, 1966). ‘Dorsal reflectance’ basking is a behaviour with similar wing positioning, but the wings are opened at a slightly more acute angle so that solar radiation is reflected from the wings onto the body, rather than being directly absorbed by the wings themselves. Species that use this strategy may have lighter coloured wings to facilitate reflectance, and there is often a darker coloured area either near or on the thorax where absorption of the reflected radiation takes place (Kingsolver, 1985). ‘Lateral basking’ is a behaviour pattern where the wings remain closed and the whole wing area is angled to the sun, and the undersides of the wings absorb solar radiation. In general, dorsal absorption is observable from photographs as the behaviour produces a very obvious 'open' winged posture. Other types of basking are not observable from photographs as they're generally indistinguishable from the normal resting position of many species. Therefore, these behaviours were inferred from observation in the field, taking into account sun position relative to the butterfly or taken from the literature if basking behaviour has been described.

Basking site: the resources and structures that a species uses as part of its basking behaviour (as described above). The majority of information on this trait was taken from the literature that directly describes the basking behaviour of a species rather than from photographs, but if basking could be distinguished from roosting in photographs then photographic information was also used if needed.

Flight Months: the number of months of the year a species is observed flying. This trait had both a binary element and a continuous element. The flight months for each species were recorded as 12 binary sub-traits, one for each month of the year. The maximum and minimum number of flight months are given as continuous values. The minimum and maximum value for total flight months described by different sources is recorded in the database. In addition to this the difference between the minimum and maximum values is also recorded (the range).

References

Angiosperm Phylogeny Group, (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, **61**, 105-121.

Bink, F.A. (1992). *Ecologische Atlas Van De Dagvlinders an Noordwest-Europa.* Schuyt.

Clench, H.K. (1966). Behavioural thermoregulation in butterflies. *Ecology*, **47**, 1021-1034.

Dennis, R.L.H. & Shreeve, T.G. (1988). Hostplant-habitat structure and the evolution of butterfly mate-locating behaviour. *Zoological Journal of the Linnean Society*, **94**, 301-318.

Essens, T., van Langevelde, F., Vos, R.A., Van Swaay, C.A. & WallisDeVries, M.F. Ecological determinants of butterfly vulnerability across the European continent. *Journal of. Insect Conservation,* **21**, 439-450.

Fiedler, K. (1991). European and North West African Lycaenidae (Lepidoptera) and their associations with ants. *Journal of Research on the Lepidoptera*, **28**, 239-57.

Freitag, S. & Van Jaarsveld, A.S. (1997). Relative occupancy, endemism, taxonomic distinctiveness and vulnerability: prioritizing regional conservation actions. *Biodiversity and Conservation*, **6**, 211-232.

Fric, Z. (2000). Adult population structure and behaviour of two seasonal generations of the Europian Map Butterfly, *Araschnia levana*, species with seasonal polyphenism (Nymphalidae). *Nota Lepidopterologica,* **23**, 2-25.

Higgins, L.G. & Riley, N.D. (1980). *A Field Guide to the Butterflies of Britain and Europe* 3rd edition. Collins.

Kingsolver, J. G. (1985). Thermal ecology of *Pieris* butterflies (Lepidoptera: Pieridae): a new mechanism of behavioral thermoregulation. *Oecologia*, **66**, 540-545.

Lafranchis T. (2004). *Butterflies of Europe: New Field Guide and Key*. Diatheo.

Munguira, M., García-Barros, E. & Cano, J.M. (2009). Butterfly herbivory and larval ecology. In: Settele, J., Shreeve, T., Konvička, M. & Van Dyck H. *Ecology of Butterflies in Europe.* Cambridge University Press, pp. 43-54.

Newland, D., Still, R., Swash, A. & Tomlinson, D. (2015). *Britain's Butterflies: A Field Guide to the Butterflies of Britain and Ireland-Fully Revised and Updated .3rd Edition*. Princeton University Press.

Paolucci, P. (2013). *Butterflies And Burnets Of The Alps And Their Larvae, Pupae and Cocoons* .WBA-Books.

Sekar, S. (2012). A meta-analysis of the traits affecting dispersal ability in butterflies: can wingspan be used as a proxy? *Journal of Animal Ecology*, **81**, 174-184.

Settele, J., Shreeve, T., Konvička, M. & Van Dyck H. (2009). *Ecology of Butterflies in Europe.* Cambridge University Press.

Tennent, J. (1996). *The Butterflies of Morocco, Algeria and Tunisia*. Gem Publishing.

Tolman, T. & Lewington, R. (2008). *Collins Butterfly Guide: The Most Complete Guide to the Butterflies of Britain and Europe*. Collins.

Tshikolovets, V.V., (2011*). Butterflies of Europe and the Mediterranean area. Butterflies of Europe & the Mediterranean area*. Tshikolovets Publications.

Tutin, T., Burges, N., Chater, A., Edmondson, J., Heywood, V., Moore, D., Valentine, D., Walters, S. & Webb D. (1964-1980). *Flora Europaea; Vols. 1-5.* Cambridge University Press.

Van Buuren, S., Groothuis-Oudshoorn, K. (2011) Mice: Multivariate Imputation by Chained

Equations in R. *Journal of Statistical Software*, 45, 1-67,