Nota Lepi. **48** 2025: 165–184 | DOI 10.3897/nl.48.144747 Research Article

Refining estimates of dry body weight from linear measurements in adult moths and butterflies (Lepidoptera)

Enrique García-Barros[1]

1 *Department of Biology - CIBC, Universidad Autónoma de Madrid, C. Darwin 2, 28049 Madrid, Spain; garcia.barros@uam.es*

https://zoobank.org/3419D122-4C65-4023-ABD7-67528E674E31

Received 17 December 2024; accepted 9 April 2025; published: 12 June 2025
Subject Editor: Zdenek Fric.

Abstract. The dry body weight of adult male Lepidoptera was estimated from thirteen linear measurements of the wings and body using multivariate regression techniques. A dataset comprising information from 2,645 species was used, significantly increasing sample size with respect to a similar former approach. Based on the logarithmically transformed values of dry body weight and several linear measurements the best single predictors for body weight are body length, thorax length and head width, none of which is among the most popular descriptors of size in this insect order (namely, forewing length and wingspan). The results show that combinations of several linear measurements lead to the most precise estimates of dry body weight. More simple models, e.g. based on wing length or wingspan and body length, may provide reasonable but suboptimal approaches. Variance partitioning of the regression residuals indicated that most of the non-explained variance is attributable to morphology rather than to phylogeny, so overall the results suggest that the best models may be stable and liable for prediction except for unusual morphologies. Alternative approaches such as a taxon-by-taxon approach or ANCOVA-based methods were tested, and the results -and problems involved- are discussed. The potential relevance of co-linearity is addressed to. Based on a limited number of species, the author attempted to estimate the female to male weight relation (which happened to be nearly isometric), as well as the percent water content (38% overall).

The dataset is made available so it can be accessed for any related research on this or related subjects.

Introduction

Body size (or body size at reproduction) is among the key life history traits because of its many -often intricate- interrelations with other life history traits and with the environment (e.g. Schmidt-Nielsen 1984; Reiss 1989; Smith and Lyons 2013; for insects, see Chown and Gaston 2010). While the meaning of size is apparently obvious in most standard languages, it represents a comparative criterion between objects of similar proportions. This ambiguity is directly translated to the ecological or evolutionary context: body size has no single scientific standard definition. There are several potential meanings depending on the biological implications investigated (e.g. the volume or surface area of an individual organism, its height or weight, speed, caloric content, or the energy required to reach reproductive maturity). Such values are represented by surrogate one-, two- or three-dimensional variables (length, area, volume or weight) of which, in the Animal Kingdom, body length (or the combinations of body length and width) or body weight are the most frequent options (e.g. Rogers et al. 1976; Blackburn and Gaston 1994; Ganihar 1997; Wardhaugh 2013; Sohlström et al. 2018). Due to the practical problems for estimating the live (fresh) body weight (but see e.g., Sohlström et al. 2018), dry body weight is generally accepted as a practical estimate of body size in arthropods (Lister and Garcia 2018; Kinsella et al. 2020). Lepidoptera, one of the hyper-diverse insect orders (with more than 160,000 described species: Kristensen et al. 2007; van Nieukerken et al. 2011) provides an interesting basis for ecological and evolutionary studies on insect body size for many reasons, including those arising from the typical (though not universal) phytophagous condition of the larvae and the flying requirements of the adults (recent examples include Davies et al. 2012; Beck et al. 2016; Chazot et al. 2016; Kivela et al. 2020; Owens et al. 2020; Hamilton et al. 2022; Foerster et al. 2024a). Moreover, recent evidence of impacts of global change on terrestrial ecosystems suggests that changes in biomass may be occurring (e.g., Hallmann et al. 2017; Lister and Garcia 2018; Harris et al. 2019; Seibold et al. 2019; Bell et al. 2020; Roth et al. 2021; Wagner et al. 2021; Dalton et al. 2023; Müller et al. 2024). This makes clear that practical, standardized means to estimate moth and butterfly body masses are needed (e.g., Kinsella et al. 2020).

The body dimensions (such as body length or width) are seldom reported for the adult Lepidoptera in monographs and field guides, where the length of the forewing or the distance between the apex of the two forewings wings in extended position (wingspan) are most often presented (e.g. Miller 1991; van Hook et al. 2012; Kinsella et al. 2020 and references therein). The length of the forewing is, in fact, known to be well correlated with body weight in these insects (Miller 1997) even if it does not stand for the best linear predictor (García-Barros 2015). The diversity of body plans in the adults of this insect order occurs not only at high taxonomic levels but also at, or below the family rank (also long acknowledged, e.g. Janzen 1984; Owens et al. 2020) making evident that bi- or multivariate approaches may be by far more precise (Foerster et al. 2024b), which imposes a balance between accuracy and practicality depending on the objectives intended.

Some years ago, the author essayed a multivariate approach to the prediction of male dry body weight in the Order Lepidoptera (García-Barros, 2015). Although based on a relatively small number of species the results suggested high reliability overall, though probably some sensitivity in the case that morphological diversity within clades was underestimated. This could only be solved by increasing morphological diversity within the taxa already present in the database. Therefore, the present study represents an extension of that former work after substantially increasing taxonomic sampling and, within reasonable limits, the intra-taxon variation in morphology. The new dataset should provide evidence to (1) validate the former results, (2) propose one or several alternative indexes to estimate dry body weight using only linear measurements and (3) estimate the relative robustness of the results in terms of the body proportions and phylogenetic position, as this might illuminate the ways for any further improvement of the results. (4) For practical reasons, research on lepidopteran allometry and related issues often concentrates on one taxon (typically, a family). For this reason, I intend to provide -and discuss the merit of- dry body size estimates based on pre-selected subtaxa (family, superfamily or informal groupings such as ‘micromoths’ or ‘macromoths’), both on a taxon-by-taxon basis or from multivariate regression approaches such as those implemented by other authors (e.g., Kinsella et al. 2020; Weiss and Linde 2022, on carabid beetles). Next, (5) possible short-term solutions to two limitations of the results for ecological studies, i.e., the restriction of the data to the male sex and the relationship between dry body weight and fresh (live) body weight. Finally (6), making accessible the set of data used in this study, which may provide a basis for related research on body size related matters in the moths and butterflies.

Materials and methods

Specimens

The methods were identical to those described in García-Barros (2015), so I refer to that publication for further detail. The sample consisted of 5,341 male adults from 2,645 species from any possible source including commercial dealers. Dry body weight in mg (DBw) was recorded after drying to constant weight at 70 °C (24–72 h). Next, the insects were re-moist and spread, then digitized together with a metric scale (on dorsal view but also ventrally whenever this was needed to appreciate the hind wing or to measure head width). For a small number of collection specimens (pinned and spread) these were photographed first, then detached from the pin and weighed. Only specimens judged to be in good condition were selected. The insects were labelled and numbered individually; the largest part of the sample was deposited in the collections of the Madrid Natural History Museum (Museo Nacional de Ciencias Naturales) except the Tortricidae (Institut Cavanilles de Biodiversitat i Biologia Evolutiva, Valencia, Spain) and a small voucher sample kept at the author’s Department. Measurements were taken on the digitized images with ImageJ (Rasband 2018). When required, complementary measures were taken under a binocular microscope with a graded scale or from scale drawings done with a camera lucida.

Identification to species level was not always possible, so in most instances reliable genus-level identification was judged sufficient for the present approach. Even when a plausible species could be attributed to such samples, these have been kept as ‘sp.’ unless confirmation by an expert or by reference to the specialized literature was possible.

A small sample of female individuals (361 species, 487 individuals) from species already represented by males (where at least the dry body weight and the forewing length were measured) was used to explore the general pattern of sexual size dimorphism of dry body weight. Another subset of field materials collected in Spanish locations (430 male individuals from 237 species) was used to explore the relationship between fresh and dry body weight. These individuals were frozen (-20 °C) shortly after capture, weighted in the laboratory within the next 24 hours and then processed as described above.

Measurements and variable abbreviations

In addition to the Dry Body weight (DBw) thirteen linear measurements (in mm) were taken from each spread specimen. These are abbreviated using a combination of 1–2 capital letters to denote the body part measured (H, T, A, FW, HW, W, B respectively for Head, Thorax, Abdomen, Forewing, Hindwing, Wing and Body) followed by small case characters to indicate the attribute measured (l = length, w = width and sp = span, with the latter followed by 1 or 2 to indicate two alternative to estimate the wingspan). Since the statistical analyses were done on the Log_{10}-transformed values, the abbreviations below (used in the text, tables and figures) refer to the log-transformed arithmetic means of such values for each species:

Hl	Log_{10}(head length): the length of the head in dorsal view. Frequently, this does not stand for the real length of the head, which can only be measured laterally. In extreme cases (e.g., the genus *Hyblaea* Fabricius) where the head is dorsally occluded by the thorax an arbitrary measure (0.01 mm) was recorded.
Hw	Log_{10}(Head width): where the head width is the distance between the outermost point of the two eyes.
Tl	Log_{10}(Thorax length): with the thorax measured dorsally; this estimate is somewhat arbitrary as in some taxa the thorax extends backwards behind the abdomen to some extent.

Tw	Log$_{10}$(Thorax width): the distance between the points of insertion of the forewings (not between the midpoints of the tegulae since these are movable structures covering the wing basis, e.g., Kristensen 2003).
Al	Log$_{10}$(Abdomen length).
Aw	Log$_{10}$(Abdomen width): where the width was measured at the midpoint of the abdomen length; again, this variable may not equally represent the robustness of the abdomen in all Lepidoptera, because the abdomen may be sub-cylindrical, dorso-ventrally flattened, or laterally depressed (e.g., in most Pieridae).
FWl	Log$_{10}$(Forewing length): with the wing measured from its base (see above) to the apex.
FWw	Log$_{10}$(Forewing width): the width of the wing taken at the midpoint of the FWl: axis and perpendicular to it.
HWl	Log$_{10}$(Hindwing length), the hindwing measured from its base to the most distant part of the wing edge, whatever its position.
HWw	Log$_{10}$(Hindwing width): like FWw described above.
Wsp1	(Wingspan 1, Log$_{10}$ transformed): where wingspan is the distance between the two forewing apices in the spread insect; or, when the forewing maximum length occurs posteriorly to the wing apex, between the outermost points of the two forewings while these stand in the anterior half of the wing margin.
Wsp2	Log$_{10}$(Wingspan 2): where wingspan is one half of the flying insect's maximum wingspan along a wingstroke) calculated as: wingspan 2 = forewing length + (0.5·thorax width).
Bl	Log$_{10}$(body length): where the body length is the sum of the head, thorax and abdomen lengths.

Regression

The species means for each variable were calculated as the first step, then Log10-transformed and processed through linear regression. This implies that the traditional allometric (power) equation $Y = a(X^b)$, or $\text{Log}(Y) = \text{Log}(a) + b \cdot \text{Log}(X)$ (e.g. Schmidt-Nielsen 1984) was assumed as the basic underlying relationship between the variables. Regression models were double checked with GLM (generalized linear models) fitting procedures in R (R Core Team 2023), and GRM in Statistica (StatSoft 2004); the latter was used for other standard procedures. Because of the logarithmic transformation of the data, the results should be identical to those obtained from Linear Models methods. Multicollinearity in the terms of the regression models was estimated with the variance inflation factor (VIF) (*R* package "car": Fox and Weisberg 2019). Forward stepwise selection was applied if more than one explanatory variable was entered. Fitting with RMA (Reduced Major Axis) regression, although often applied in allometric studies after the suggestions by Harvey and Pagel (1991) was discarded as its use has been criticized by several experts (Hansen and Bartoscek 2012; Kilmer and Rodríguez 2016).

Model validation

The predictive accuracy of the regression models estimated by cross-validation was given preference to evaluate the models. For that purpose, a subset of 534 species represented by at least three individuals was chosen as the validation set. Cross-validation was performed using the leave-one-out procedure (package "boot" in R: Canty et al. 2024). The Cross-validation error (CVE) measures the success of each model to predict the dry body weights for the selected species after excluding them from the fitting process.

Phylogenetic information

Comparative methods where phylogenetic relations are accounted for are the standard whenever evolutionary relationships are the aim (multiple examples during the last three decades after e.g., Felsenstein 1985 and Harvey and Pagel 1991). However, this study intends to design predictive models and explicitly avoids any evolutionary interpretations, so no comparative method was adopted. This is justified by two reasons: first, controlling' for phylogenetic effects implies that, depending on the comparative method used, predicting body weight for a given species requires that the phylogenetic position and the evolutionary distance (branch length) between that species and the nearest node in the operative phylogenetic tree are specified as a part of the input information. This is possible under phylogenetic imputation techniques (Swenson 2014; Gendre et al. 2024), but this falls beyond the scope of this study (see further comments in the discussion). The second reason is that the concern on the non-independence of the species means due to shared ancestry (e.g., Ives and Zhu 2006) may be justified or not depending on the data and the purposes of the researcher; in regression, statistical independence should apply to the error terms of the models fitted, not to the variables themselves (McNab 2003; Revell 2010; Uyeda et al. 2028). The present approach relies on the assumption (actually, a not-demonstrated expectation) that the pool of species compared covers most of the morphological range of variation at the highest taxonomic level and, within reasonable limits, within families and genera. From this point of view, phylogenetic relatedness is assumed as one of several causes behind the correlations observed, and it is accepted that the regression models selected are reliable only within the range of morphological and phylogenetic variation tested (i.e., they provide means for interpolating, not for extrapolating values).

However, phylogenetic information is relevant to this study as a means for estimating the phylogenetic signal in the data and, more relevantly, to check the potential causes of error in the models fitted to determine whether the lack of fit of the results should be attributed to phylogeny or to morphology (by variance partitioning as detailed below). The balance between these two potential sources of bias should help to improve the dataset for further work on the subject. This required an operative tree topology and an estimate of evolutionary distances.

The tree topology was built after the available information on lepidopteran taxonomy, updated to December 2022. Using the scheme used in García-Barros (2015) as the starting point, available updated hypotheses at any taxonomic level were incorporated to a single tree topology (Suppl. material 1: Documentation on phylogeny). The approach was certainly optimistic since, in the absence of phylogenetic treatment for any taxon, the formal, accepted subtaxa were kept as monophyletic entities. Polytomies were frequent (tree resolution = 68.9%), not only because of the lack of completely resolved phylogenetic hypotheses for all the families but because of the lack of specific identification of one part of the species in the data analysed. The tree topology (Suppl. material 2: tree topology) was assembled manually with Mesquite (Maddison and Maddison 2023).

Branch lengths (not the tree topology) were derived from the DNA barcode sequences from BOLD (Barcode of Life Data System: www.boldsystems.org). For 1,430 species with sequences available, one of the most complete ones was randomly chosen; for 730 species identified only to genus level, the sequence from a 'replacement species' (i.e., a species in the same genus) was adopted if available. In this way, one nucleotidic sequence was associated to 2,160 of the terminal taxa (species) (Suppl. material 3: list of species and available BOLD sequences). After aligning the sequences (MUSCLE: Edgar 2004), a distance matrix was calculated ("dist.dna" in ape: Paradis and Schliep 2018; default settings), which was later used to estimate branch lengths on the operative tree topology (with Phytools: Revell 2024).

The tests incorporating phylogenetic information were restricted to the subset of species for which branch length estimates were available. For variance partitions (see below) the residuals from that subset of species were extracted from the models fitted to the whole data set. The phylogenetic signal in each of the variables (dry body weight and the linear measurements) was estimated using Pagel's lambda (Pagel 1999; Freckleton et al. 2002; Münkemüller et al. 2012; Molina-Venegas and Rodríguez 2017) using the R package "phylosignal" (Keck et al. 2016), after ultrametrizing the tree (function "chronos" in ape: Paradis and Schliep 2018).

Variance partitioning was used to estimate the fraction of the variation not accounted for by the models which is attributable to either morphology or to phylogeny (following the ideas by Foerster et al. 2024b: phylogeny vs. body plan). The residuals from each of the regression models (the problem variable) were analysed in terms of two sets of variables designed to summarize the information on morphology and phylogeny. The method of variance partitioning (Legendre and Legendre 2012) was selected as the most effective to estimate the two types of effects with the same procedure. Morphology was not represented by the original explanatory variables but by the residuals of their linear regression on dry body weight (DBw), since DBw itself was the dependent variable in all the models. In other words, the 'predictors' for this test can be interpreted as proportions, i.e., the values of the predictor variables relative to the dry body weight, and linearly independent from it. These were submitted to a principal components analysis (PCA based in the covariances) from which ten orthogonal factors were extracted (100% quality of representation) and subsequently used to represent the morphological proportions, relative to body weight. Phylogeny was represented by a series of factors (2161) extracted from the tree topology with branch lengths, processed by Principal Coordinate Analysis (PCoA, a method suited for Euclidean distances: Legendre and Legendre 2012) as implemented in the package "ape" in the R environment (Paradis and Schliep 2018; R Core Team 2023). Of these, those with significant effects on any of the model's residuals effects (selected through backwards regression) were adopted for the variance partitions analysis (following Desdevises et al. 2003 and Legendre and Legendre 2012), which was performed using "varpart" (in the R package vegan: Oksanen 2024).

Two types of taxon-by-taxon prediction of DBw were essayed using the family level as a starting point (the members of families represented by less than two species were clustered into the most closely related taxon of the same level). First, individual regressions for each family were calculated (with FWl, Wsp1 and Wsp2 as predictors). The differences between the family slopes were tested by means of ANCOVA analysis, and a T-test was used to compare the intercepts. For the main superfamilies, the phylogenetically closest families were combined when their slopes and intercepts were not significantly different. This procedure was repeated until the superfamily groups of at least 40 species were obtained. In addition, several arbitrary assemblages of potentially practical use (such as e.g., macromoths or micromoths) were evaluated. Second, an ANCOVA approach (such as that used by Weiss and Linde 2022) allowed for simultaneous testing of differences among the family intercepts and slopes (by reference to the common, shared intercept and slope). In this method family-level coefficients are assigned to the intercepts and slopes of each family. This procedure started with all the variables in model 2 (described below). Through a manually driven backwards selection process, the variables were retained only when a significant effect of the variable itself (not only the interactive term) was detected; then the interactions between the taxonomic code and the explanatory variable were checked and the families with no significant effects were merged to a common pooled 'Lepidoptera'.

Build in yourself methods

Finally, a regression model based on the factors extracted from a PCA analysis of the morphological variables and phylogenetic information was performed. 10 PCA factors were extracted from the complete set of variables (99.9% quality representation), which were combined with the

2161 PCoA factors outlining the phylogenetic relations (described above for the variance partitions protocol). The non-significant factors were excluded from the model along a manually driven backwards selection process until a final solution was selected. This method has the advantage that the factors are linearly independent from each other so co-linearity is avoided, though it has practical drawbacks (addressed to in the see discussion).

Data resources

The data underpinning the analysis reported in this paper are deposited in the Dryad Data Repository at https://doi.org/10.5061/dryad.bk3j9kdnw.

Results

The lowest range of body weights did not differ from that presented previously (García-Barros 2015) with some Nepticulidae and Heliozelidae as the smallest and lightest species (0.04–0.05 mg). From the results below, the weight of the reputedly smallest known Nepticulidae or Gracillariidae with forewing lengths of 1.2–1.3 mm (see Stonis et al. 2021) would be expected to be in the range 0.017–0.020 mg. The upper limits of body weight are of course occupied by large bombycoids and the hawk moths (Bombycoidea and Sphingidae), but in this sample these were exceeded by the dry weight of some tropical Cossidae like *Xyleutes persona* (Le Guillou, 1841) with DBw above three grams (3000 mg). On this basis the male dry body weight of known Lepidoptera covers five orders of magnitude, three in the case of the forewing length. Almost certainly, this range is underestimated since the male dry weight of the large Asuatralian Cossidae *Endoxyla cinereus* (Tepper, 1890) should at least duplicate that weight (the female fresh weight of this species reaches 30 grams: Beccaloni 2010; also see Thurman 2022).

On the logarithmically transformed values all the linear measurements were highly correlated with dry body weight ($R > 0.97$) and showed a clear phylogenetic structure, with lambda values close to 1.0 (Table 1). The slopes of the lines (the exponents in the traditional allometric equation) varied between 1.92 and 2.96, as expected from length / volume relations from the allometric equation. Bivariate correlations with DBw exceeded 0.9 except for FWw, HWl, HWw ($R > 0.8$) and, remarkably, the head length (Hl, $R = 0.77$). The covariation between the variables (correlation matrix in Suppl. material 4: table S1) suggests three groups of them (Fig. 1): (1) the best single linear predictors of body size (Tl followed by Bl and Hw, all $R > 0.98$), (2) the wing metrics and (3) the body width measures. The comparatively low correlations of Hl with the remaining variables probably reflect the low accuracy of this measurement (see methods).

Among the formerly published models, the closest one to those presented in Table 2 (i.e., based on linear measurements and not wing areas) was García-Barros 2015: supplementary material 6 (Model 4, $R^2_{adj} = 0.979$ from 375 species). This model worked remarkably well with the present data set ($R = 0.989$, adjusted $R^2 = 0.979$).

Eight alternative multivariate models were fitted, reflecting choices that would be expected to be preferred most often depending e.g., on the type of information available to predict DBw. Redundant combinations of composed measures with their components were avoided (e.g., Bl and Hl, Tl and Al or Wingspan with FWl and Tw). The statistics and coefficients are summarized in Table 2. Model 1 was obtained by stepwise selection of all the variables. Hl was ignored in the remaining models (e.g., Model 2: Fig. 2). In Model 3 Bl replaced (Hl, Tl and Al). Models 6 and 7

Table 1. Phylogenetic signal for each of the variables in the study (Lambda), and summary results of the bivariate regression models of Log_{10}(Dry body weight) (DBw) on each of the predictor variables, all values Log_{10} transformed. Intercept, coefficient and adjusted R^2 of each model, together with the cross-validation error (CVE) and the 'pure' and shared explanation of the residuals attributable to morphology (Res.: morph.), phylogeny (Res.: phylo.) and shared by morphology and phylogeny (Res.: shared) (R^2 values). For the first four columns, all but one of the values (intercept for the Tw regression) were significant at the level P > 0.001.

Variable	Lambda	Intercept	Coefficient	$R^2_{adj.}$	CVE	Res.: morph.	Res.: phylo.	Res.: shared
DBw	0.998	–	–	–	–	–	–	–
Hl	0.965	1.418	2.311	0.599	0.259	0.031	0.015	0.008
Hw	0.997	0.250	2.928	0.916	0.053	0.329	0.017	0.055
Tl	0.999	-0.434	2.708	0.948	0.033	0.042	0.001	0.008
Tw	0.984	-0.016^{ns}	2.864	0.910	0.058	0.077	0.004	0.009
Al	0.986	-1.055	2.695	0.881	0.077	0.097	0.007	0.015
Aw	0.985	0.657	2.381	0.867	0.086	0.107	0.009	0.017
FWl	0.999	-2.053	2.727	0.862	0.089	0.106	0.009	0.023
FWw	0.998	-0.340	1.953	0.780	0.142	0.172	0.024	0.025
HWl	0.985	-1.728	2.714	0.782	0.140	0.166	0.022	0.031
HWw	0.984	-0.732	2.238	0.763	0.153	0.183	0.026	0.028
Bl	0.998	-1.832	2.823	0.929	0.046	0.059	0.002	0.010
Wsp1	0.985	-3.211	2.962	0.889	0.072	0.087	0.007	0.017
Wsp2	0.987	-3.078	2.729	0.885	0.074	0.089	0.006	0.019

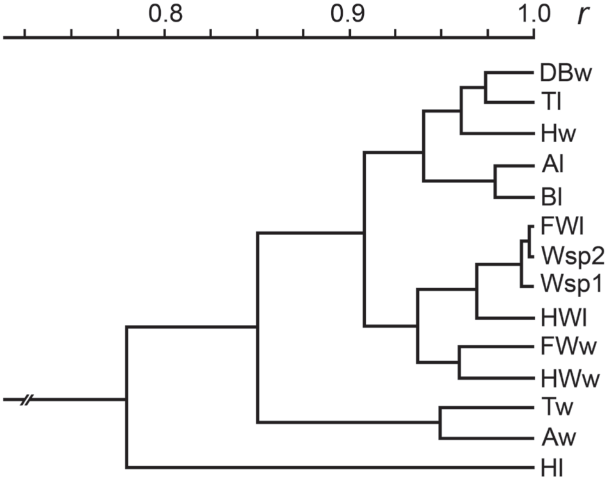

Figure 1. Patterns of covariation between all the variables in the study included dry body weight (DBW) as illustrated by a cluster based in the square correlation (r) matrix of their log-transformed values (UPGMA method).

were restricted to FWl and the thorax measures. The combinations of wingspan and body length were represented in models 4 and 5. Model 8 was fitted using FWl as the primary explanatory variable, and the residuals of the remaining variables on FWl as additional predictors; the coefficients and fitting statistics were recalculated by reference to the original variables. This presented the advantage of a high linear independence of the variables in the regression.

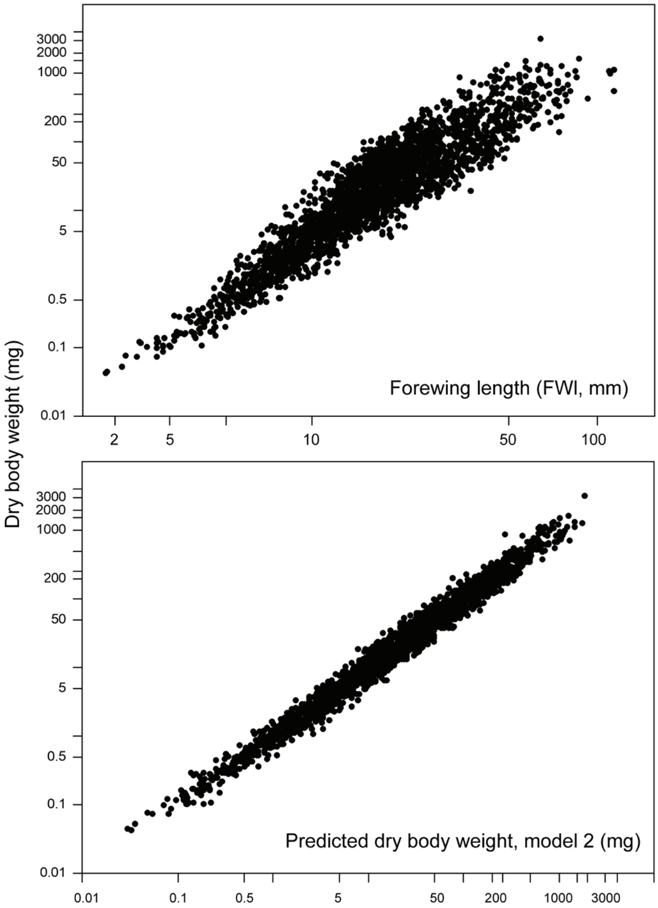

Figure 2. Relationship between the Log_{10}-transformed species mean dry body weights (DBW) and one of the most widely used descriptors of body size in the Lepidoptera, forewing length (FWl), and with the predicted values from multivariate model 2 (details in Tables 1, 2). Notice Log-scale in both axes.

All the alternatives proved high fit (adjusted R^2 above 0.93) with models 1, 2 and 8 rendering the best predictive performance (lowest cross validation errors).

Regarding the 'taxon-by-taxon approach' the explanation of DBw in terms of FWl, WS1 and WS2 were acceptable ($R^2 > 0.73$, $P < 0.001$), on average respectively 0.83, 0.84 and 0.85 (Suppl. material 4: table S2). Detailed comparisons between closely related families proved useful sometimes (as for the Papilionoidea, Fig. 3). Most often than not, significantly different slopes or intercepts between sister groups precluded phylogeny-oriented clustering, while it was clear that moderately distant families might have been combined. The ANCOVA approach initially had high fitting, but with most of the interactions being non-significant. Only four to nine taxa remained after dropping the non-significant contributions, with the performance of the model not exceeding that of the former alternatives. Depending on whether FWl (model $R^2 = 0.87$) or Wsp1 (model $R^2 = 0.91$) were the variables retained, the taxa with significant coefficients included the Gelechioidea, Pyraloidea (or the family Pyralidae), Tineoidea (or just the Tineidae) and the Yponomeutoidea (Suppl. material 4: table S3).

Table 2. Summary of the results of the multivariate linear regression models of Log$_{10}$(dry body weight), DBW, on different subsets of the potential explanatory variables, based on the Log10-transformed values (all the values reported for the intercepts, coefficients, model R^2, F tests and AIC scores were significant, $P < 0.001$). The four lower rows contain the cross-validation errors and the explanation of the model residuals attributable to morphology alone, phylogeny alone, and to the shared effects of morphology and phylogeny (Resids., respectively: morph., phylo., shared).

Parameters	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	-0.5206	-0.4474	-0.6883	-2.2383	-2.1903	-1.0617	-0.7148	-0.5329
Hl	-0.1107	–	–	–	–	–	–	–
Hw	0.1994	0.0847	0.4516	–	–	–	–	–
Tl	1.1379	1.1281	–	–	–	–	1.1757	1.2221
Tw	0.4584	0.5020	0.7435	–	–	1.7776	1.1910	0.5139
Al	0.2802	0.2738	–	–	–	–	–	0.2450
Aw	0.5111	0.5275	0.5067	–	–	–	–	0.5200
FWl	0.2859	0.2454	0.3828	–	–	1.2586	0.5829	0.3760
FWw	0.0858	0.0992	–	–	–	–	–	–
Bl	–	–	0.6431	2.1421	2.1751	–	–	–
Wsp1	–	–	–	0.7630	–	–	–	–
Wsp2	–	–	–	–	0.6873	–	–	–
Model statistics								
R^2_{adj}	0.980	0.980	0.974	0.935	0.933	0.963	0.974	0.979
F	16242.6	18250.5	16499.9	18794.3	18716.6	34146.0	33640.9	25246.2
AIC	-4002.8	-3959.7	-3300.8	-849.94	-839.68	-2352.1	-3353.6	-3847.6
CVE	0.0129	0.0131	0.0168	0.0425	0.0427	0.0241	0.0165	0.0137
Minimum VIF	4.04	11.83	10.50	11.86	11.36	3.49	6.73	1.00
Min. VIF variable	Hl	Al	Aw	Wsp1	Wsp2	FWl	Tw	FWl
Resids., R^2: morph.	0.017	0.017	0.022	0.052	0.054	0.033	0.023	0.020
Resids., R^2: phylo.	<0.001	<0.001	<0.001	0.002	0.002	0.001	<0.001	<0.001
Resids., R^2: shared	0.002	0.003	0.004	0.011	0.011	0.004	0.003	0.001

As shown in Table 2 (and Suppl. material 4: table S3 for the ANCOVA-type approaches) the variance inflation factor (VIF) was high whenever more than one variable was included in the models, except in Model 8. With no exception, the regression residuals primarily reflected differences in the body plan (morphology) and to a lesser extent (by a 10-fold proportion) a shared effect of morphology and phylogeny, while the ‘pure’ effect of phylogeny was virtually negligible. Cross-validation testing (CVE) on the 534 species represented by at least three individuals largely parallelled the R^2 values.

The models based in the linearly independent factors representing morphology and phylogeny were slightly superior to any of the alternatives in Table 2: adjusted $R^2 = 0.982$, $F_{[27, 2099]} = 4218$ ($P > 0.001$), CVE = 0.012 and VIF values in the range 1.01–1.13. Using only the components derived from morphology would still render reliable results ($R^2 = 0.981$). The details of this model are not presented in length as it cannot be applied for predicting body weights for species not represented in the data set (see discussion).

Regarding fresh (live) body weight, on average and on the non-logarithmically transformed values the water content was slightly below 40% (38.64%, s.d.= 7.11, Fig. 4). Through regression on the log-transformed values, the slope for the female sex (n = 53) differed slightly but significantly from that of the males ($F = 7.25$, $P = 0.008$). Accordingly, fresh body weight was estimated as a function of BDw only from the male data, as: Log$_{10}$(fresh body weight) = $0.4829 + 0.9841 \cdot$Log$_{10}$(dry

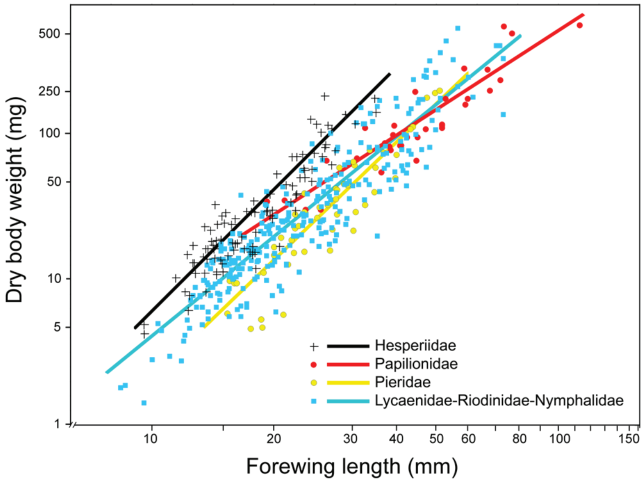

Figure 3. The butterflies (Papilionoidea) as an example of the potential problems involved in pre-selected taxon-specific approaches the prediction of body weight (Log_{10}-transformed, DBw) from forewing length (Log_{10}-transformed, FWl). Although three families (Lycaenidae, Riodinidae and Nymphalidae) were found to be homogeneous for the slope and intercept of the linear relationship, these and the three remaining families differed significantly from each other in the intercept, the slope, or both. Notice the logarithmic scale in both axes.

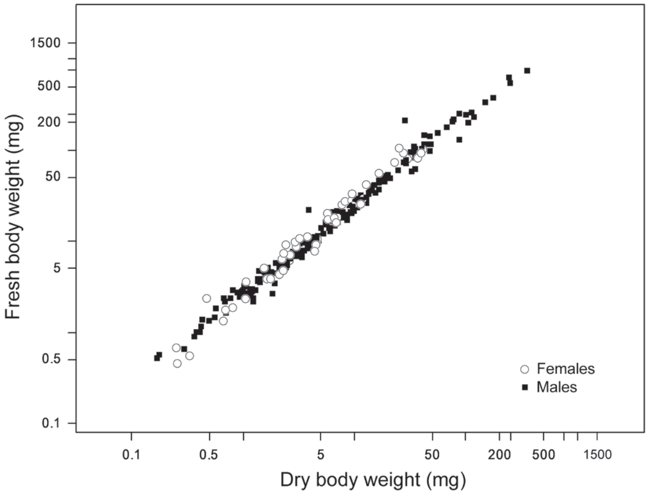

Figure 4. Fresh body weight as predicted by the Log10-transformed dry body weight values (DBW) of males and females from 237 species. The overall relationship was significant (adjusted $R^2 = 0.985$, $p < 0.001$) with a slope close to isometry: Log_{10}(Fresh body weight) = 0.428 + 0.992(Log_{10}DBW). The difference between the slope of the two sexes was small, but significant (details in the text). Notice the logarithmic scale in both axes.

body weight), $R^2 = 0.987$, $P < 0.001$ (n = 183 species). From this equation, the dry body weight accounts for 38.02% of the live weight for the male sex.

For the 362 species with weight data from the two sexes (Fig. 5), sexual dimorphism was described by the equation $DBw_{female} = 0.1574 + 0.9770(DBw_{male})$, $R^2_{adj} = 0.953$, $P < 0.001$. This suggests a nearly isometric relationship between the body mass of the two sexes, which was not expected.

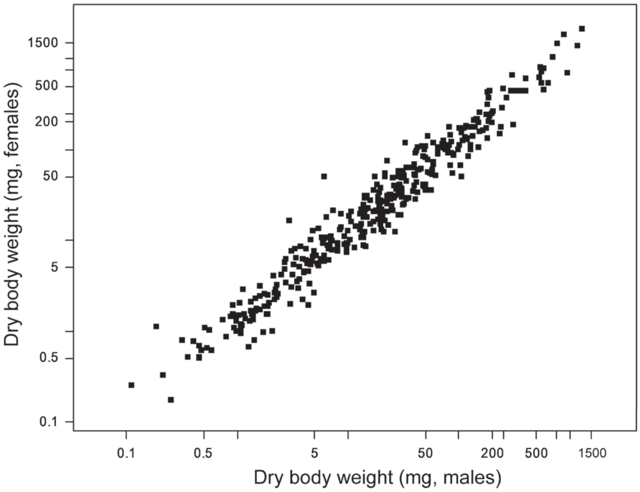

Figure 5. Male dry body weight as a predictor of the female dry body weight, from a sample of 362 species. Overall, on the log-transformed values the female dry body weight can be approached as female DBW = 0.1574 + 0.9770(male BDW) (adjusted $R^2 = 0.953$, $P < 0.0001$). Notice the logarithmic scale in both axes.

Discussion

The interpretation of the results allows for two opposed perspectives — both realistic — a pessimistic and an optimistic one.

As far as body weight prediction is concerned, the pessimistic point of view derives from the putatively low taxonomic and geographic representativity of the species used (hardly more than 2% of the described species of Lepidoptera: Kristensen et al. 2007; van Nieukerken et al. 2011), the small sample size (most often 1–3 individuals per species), the scarcity of data from the female sex and the frequently rough taxonomic identification. The amount of statistical error imposed by these shortfalls may exceed that caused by the technical details of the analytical methods adopted (Hansen and Bartoszek 2012). Moreover (as stated by Foerster et al. 2024b) the weights recorded here probably differ from the weight on emergence and this may partly obscure the life historical interpretation of the data.

Missing information from one part of the species is a recurrent problem in interspecific comparative studies (Nakagawa and Freckleton 2008; Penone et al. 2014). However, unlike for e.g., plants and the vertebrate animal taxa, the problem with hyper-diverse taxa such as the Lepidoptera is not one on interpolating the missing values for one part of the species but on the need to extrapolate them from a minority of species. Phylogenetic imputation (Young et al. 2014; Molina-Venegas et al. 2018) is an alternative to be explored, especially because such methods might render acceptable results if additional correlated variables (e.g. the body dimensions) can be coded for (Penone et al. 2014; Debastiani et al. 2021). However, while the broad lines of Lepidopteran phylogeny are progressively being fixed (e.g., Mayer et al. 2012; Heikkilä et al. 2015; Kawahara et al. 2019; other references and discussion in Rota et al. 2022), a comprehensive reference phylogenetic scheme may take years to come. Of course, some taxa in specific geographic regions are exceptions, such as the European Papilionoidea (Wiemers et al. 2020).

However, an optimistic interpretation is straightforward from the results. First, a seven-fold increase of sample size in comparison to García-Barros (2015) (375 to 2645 species) raised the fit of the best regression model by merely 1%. This, together with the low pure contribution of phylogeny in the non-explained variances, suggests that the variation of the broad lepidopteran morphology across taxa is redundant enough for dry body weight to be predicted with reasonable accuracy on a multivariate basis. Yet one or two variables cannot cover the morphological heterogeneity of the Lepidoptera so a multivariate model will generally perform better. This may impose costs in terms of collinearity, which in fact holds for any model with more than one predictor. A VIF value above 4–10 is generally identified as excessive (depending on authors: Neter et al. 1983; Legendre and Legendre 2012 and references therein). From this point of view only model 8 (Table 2) would be safe. The main risk of collinearity in the predictor variables is ‘overprediction’, meaning that unusual morphologies (body plans not covered by the dataset) would lead to inaccurately predicted body weights. This would represent a lesser concern under the assumption that the main morphological gradients were covered by the dataset. However, caution is required whenever any such unusual shaped species is involved. Alternatively, co-linearity may be avoided by a model constructed with the PCA factors derived from all the potential independent variables. Unfortunately, predictions from this model are possible only if the measurements of the candidate species are added to the data table before the PCA factors are extracted. After fitting the regression model (on the species with known DBw) the values can be interpolated for the problem species. The same holds for phylogenetic relations, the species to be predicted must be given a position and a branch length in the reference phylogeny before the PCoA factors describing phylogeny are extracted from the tree. All this makes these methods far less convenient for prediction than a ready to use equation (however, the PCoA alternative is feasible using the dataset provided in this study).

A taxon-specific approach (e.g., predictions for one single family or superfamily using only FWl or Wingspan) may eventually be justified for practical reasons despite the generally suboptimal results. Such approach implicitly assumes a high degree of phylogenetic conservatism within the taxon studied, which may or may not be the case. As found in the Papilionoidea, there is a risk that the weights of part of the species are systematically overestimated or underestimated. The low weight of phylogenetic structure in the regression residuals suggests that species clustering on morphological grounds rather than on taxonomic ones might by far be preferable. This applies equally to the ANCOVA-like approaches where the initial results are of high fit overall, but with many non-significant terms which ultimately lead to comparatively poor results.

In summary, the most complex regression models (based in five to seven body measurements) predicted dry body weight most accurately. The author would suggest that models 2 and 8 (in Table 2) are given preference. Estimates from a smaller number of variables (FWl, Wsp1, Wsp2 or combinations of these with Bl) should generally represent suboptimal solutions but may still be acceptable (as suggested by Foerster et al. 2024b) while it is acknowledged that high R^2 scores from this and similar datasets are not surprising given the wide range in body sizes and the possible underlying functional constraints on the body proportions. In any case, given the small sample size for each species in the dataset used, the species means should be interpreted as rough approaches of the true values. For similar reasons, ideally, reliable predictions for a species should not be done on the data from a single specimen but on the average from a) the measures from several individuals or b) the average of the individual predictions for those specimens.

It is not the author's intention to negate the utility of traditional size standards used for the Lepidoptera such as wingspan and forewing length. As shown by the results these measures retain a good correlation with body weight and may represent the only metric available for most species, besides their interest as descriptive standards or as the basis for the study of e.g. intra-specific variation of moth and butterfly body size.

The fact that only male insects were analysed imposes severe limitations for implementation of the present results in field studies. In addition to the differences between live insects and preserved ones due to tissue contraction after desiccation (Van Hook et al. 2012), the female body weights or the fresh body weight may eventually be interesting. These shortfalls can tentatively be overcome by correcting the predicted body masses for female weight (to average the male and female values), and by estimating fresh body weights from the results described above. However, while the female to male body size relationship was clear-cut in the results, there is ample evidence that such relationship is not uniform across all the moth and butterfly families (Stilwell et al. 2009; Allen et al. 2011; Teder 2014). The available sample points out an approximate 38% water content, which falls within the typical range for adult insects (e.g., Hutchinson et al. 1997; Barker et al. 1998). Again, there is no reason to believe that this proportion is constant across all taxa. It is worth noting that the range of sizes used for those comparisons is markedly shorter than that of the general sample.

Incidentally, significant between-taxa differences (such as those within the Papilionoidea already mentioned, and probably other groups) may reflect differences in the broad body plans with ecological, life-historical or biomechanical implications which deserve further attention. On similar grounds, the bivariate slopes (form the regressions of dry body weight on the individual variables) revealed allometric slopes within the range 2.0–3.0, where 3.0 represents isometry. Interestingly, the steepest slopes (2.8–3.0) are among those involving body measurements explicitly (thorax width, total body length) or implicitly (such as wingspan 1) together with head width (as already documented by García-Barros 2015). Wing length values scale to dry body weight with slopes of ca. 2.7 while wing widths do so with lower values (1.9–2.2), so overall one might envisage a pattern of progressive wing narrowing along the positive gradient of body size, probably encompassed by comparatively higher wing loadings. While suggestive, these observations require corroboration from further research done under comparative methods to incorporate the effect of phylogenetic relatedness.

To conclude, the dry body weight of male Lepidoptera can be predicted with reasonable reliability from a combination of linear measures not difficult to obtain. Rough estimates of the live weight or of the average male-female dry weight may be possible at the cost of accuracy. For decades, work devoted to the estimation of body weight in insects has relied in the objective of finding reference equations whose results could be used for extrapolation to other species (to quote some examples: Rogers et al. 1976; Sage 1982; Ganihar 1997; Benke et al. 1999; Sohlström et al. 2018). While a single reference model (e.g. at the family or order levels) would of course represent a useful tool for standardizing results, the present facilities for statistical computing combined with the widespread tendency to make the research data available lead to new alternatives such as the ad-hoc assemblage of data sets for specific uses. Facing the urgency of detecting and understanding the possible effects of Global Change there is a renewed interest in collecting and compiling life history data under reasonable quality standards (Keller et al. 2022). Needless to say—and as argued in the introduction—the body size of moths and butterflies deserves a role in this research program, therefore the explicit publication of more data on this subject (e.g., Kinsella et al. 2020) should be encouraged.

Acknowledgements

Over the last few years, the author received help from several colleagues who either provided specimens, facilitated access to them or to documentation, assisted with identification or solved methodological issues. While discharging them from for any accuracies in the results, I am indebted to (in alphabetical order) N. Agustí, J. Baixeras, U. Benardo, J.W. Brown, J.P. Cancela, M. Corley, M. Costa, E. Drouet, J. de Freina, O. Karsholt, M. Laguerre, W. Mey, M.L. Munguira, J.E. Murria Beltrán, A. Orellana, L. Przybylowicz, F. Pühringer, J. Razowski, L. Ronkay, H. Romo, P. Sihvonen, H. Staude, J. Sumpich, G. Tarmann, E. Toro Delgado, A. Vives Moreno, T.J. Witt (†) and J.L. Yela.

References

Allen CE, Zwaan BJ, Brakefield PM (2011) Evolution of sexual dimorphism in the Lepidoptera. Annual Review of Entomology 56: 445–464. https://doi.org/10.1146/annurev-ento-120709-144828

Barker D, Fitzpatrick MP, Dierenfeld ES (1998) Nutrient composition of selected whole invertebrates. Zoo Biology 17 (2): 123–134. https://doi.org/10.1002/(SICI)1098-2361(1998)17:2%3C123::AID-ZOO7%3E3.0.CO;2-B

Beccaloni G (2010) Big bugs life-size. Natural History Museum, London, 84 pp.

Beck J, Liedtke HC, Widler S, Altermatt F, Loader SP, Hagmann R, Lang S, Fiedler K (2016) Patterns or mechanisms? Bergmann's and Rapoport's rule in moths along an elevational gradient. Community Ecology 17(2): 137–148. https://doi.org/10.1556/168.2016.17.2.2

Bell JR, Blumgart D, Shortall CR (2020) Are insects declining and at what rate? An analysis of standardised, systematic catches of aphid and moth abundances across Great Britain. Insect Conservation and Diversity 13: 115–126. https://doi.org/10.1111/icad.12412

Benke AC, Huryn AD, Smock LA, Wallace JB (1999) Length-mass relationships for freshwater macroinvertebrates in North America with particular reference to the Southeastern United States. Journal of the North American Benthological Society 18: 308–343. https://doi.org/10.2307/1468447

Blackburn TM, Gaston KJ (1994) Animal body size distributions: patterns, mechanisms and implications. Trends in Ecology and Evolution 9(12): 471–474. https://doi.org/10.1016/0169-5347(94)90311-5

Canty A, Ripley BD, Brazzale AR (2024) boot: Bootstrap R (S-Plus) Functions. R package version 1.3-31. https://cran.r-project.org/web/packages/boot/

Chazot N, Panara S, Zilbermann N, Blandin P, Le Poul Y, Cornette R, Elias M, Debat V (2016) Morpho morphometrics: shared ancestry and selection drive the evolution of wing size and shape in Morpho butterflies. Evolution 70:181–194. https://doi.org/10.1111/evo.12842

Chown SL, Gaston KJ (2010) Body size variation in insects: a macroecological perspective. Biological Reviews 85(1): 139–169. https://doi.org/10.1111/j.1469-185X.2009.00097.x

Dalton RM, Underwood NC, Inouye DW, Soulé ME, Inouye BD (2023) Long-term declines in insect abundance and biomass in a subalpine habitat. Ecosphere 14(8): e4620. https://doi.org/10.1002/ecs2.4620

Davies RB, Ounap E, Javois J, Gerhold P, Tammaru T (2012) Degree of specialization is related to body size in herbivorous insects: a phylogenetic confirmation. Evolution 67: 583–589. https://doi.org/10.1111/j.1558-5646.2012.01776.x

Debastiani V J, Bastazini VAG, Pillar VD (2021) Using phylogenetic information to impute missing functional trait values in ecological databases. Ecological Informatics 63: 101315. https://doi.org/10.1016/j.ecoinf.2021.101315

Desdevises Y, Legendre P, Azouzi L, Morand S (2003) Quantifying phylogenetically structured environmental variation. Evolution 57(11): 2647–2652. https://doi.org/10.1111/j.0014-3820.2003.tb01508.x

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32(5): 1792–1797. https://doi.org/10.1093/nar/gkh340

Felsenstein J (1985) Phylogenies and the Comparative Method. The American Naturalist 125: 1–15. https://doi.org/10.1086/284325

Foerster SIA, Clarke JT, Õunap E, Teder T, Tammaru T (2024a) A comparative study of body size evolution in moths: evidence of correlated evolution with feeding and phenology-related traits. Journal of Evolutionary Biology 37(8): 891–904. https://doi.org/10.1093/jeb/voae072

Foerster SIA, Javoiš J, Holm S, Tammaru T (2024b) Predicting insect body masses based on linear measurements: a phylogenetic case study on geometrid moths. Biological Journal of the Linnean Society 141: 71–86. https://10.1093/biolinnean/blad069

Fox J, Weisberg S (2019) An R Companion to Applied Regression, third edition. Sage, Thousand Oaks CA. https://www.john-fox.ca/Companion/

Freckleton RP, Harvey PH, Pagel M (2002) Phylogenetic analysis and comparative data: a test and review of evidence. The American Naturalist 160(6): 712–726. https://doi.org/10.1086/343873

Ganihar SR (1997) Biomass estimates of terrestrial arthropods based on body length. Journal of Biosciences 22: 219–224. https://doi.org/10.1007/BF02704734

García-Barros E (2015) Multivariate indices as estimates of dry body weight for comparative study of body size in Lepidoptera. Nota lepidopterologica 38(1): 59–74. https://doi.org/10.3897/nl.38.8957

Gendre M, Hauffe T, Pimiento C, Silvestro D (2024) Benchmarking imputation methods for categorical biological data. Methods in Ecology and Evolution 15(9): 1624–1638. https://doi.org/10.1111/2041-210X.14339

Hallmann CA, Sorg M, Jongejans E, Siepel H, Hofland N, Schwan H, Stenmans W, Müller A, Sumser H, Hörren T, Goulson D, de Kroon H (2017) More than 75% decline over 27 years in total flying insect biomass in protected areas. PLoS ONE 12: e0185809. https://doi.org/10.1371/journal.pone.0185809

Hamilton CA, Winiger N, Rubin JJ, Breinholt J, Rougerie R, Kitching IJ, Barber JR, Kawahara AY (2022) Hidden phylogenomic signal helps elucidate Arsenurine silkmoth phylogeny and the evolution of body size and wing shape trade-offs. Systematic Biology 71(4): 859–874. https://doi.org/10.1093/sysbio/syab090

Hansen TF, Bartoszek K (2012) Interpreting the Evolutionary Regression: The Interplay Between Observational and Biological Errors in Phylogenetic Comparative Studies. Systematic Biology 61(3): 413–425. https://doi.org/10.1093/sysbio/syr122

Harvey PH, Pagel MD (1991) The comparative method in evolutionary biology. Oxford University Press, Oxford, 239 pp. https://doi.org/10.1093/oso/9780198546412.001.0001

Harris JE, Rodenhouse NL, Holmes RT (2019) Decline in beetle abundance and diversity in an intact temperate forest linked to climate warming. Biological Conservation 240: 108219. https://doi.org/10.1016/j.biocon.2019.108219

Heikkilä M, Mutanen M, Wahlberg N, Sihvonen P, Kaila L (2015) Elusive ditrysian phylogeny: an account of combining systematized morphology with molecular data (Lepidoptera). BMC Evolutionary Biology 15: 260. https://doi.org/10.1186/s12862-015-0520-0

Hutchinson JMC, McNamara JM, Houston AI, Vollrath F (1997) Dyar’s rule and the investment principle: optimal moulting strategies if feeding rate is size-dependent and growth is discontinuous. Philosophical Transactions of the Royal Society of London, B, 352: 113–138. https://doi.org/10.1098/rstb.1997.0007

Ives AR, Zhu J (2006) Statistics for correlated data: phylogenies, space, and time. Ecological Applications 16(1): 20–32. https://doi.org/10.1890/04-0702

Janzen DH (1984) Two ways to be a tropical big moth: Santa Rosa saturniids and sphingids. Oxford Surveys in Evolutionary Biology 1: 85–140. [pls. 1–5]

Kawahara AY, Plotkin D, Espeland M, Meusemann K, Toussaint EFA, Donath A. Gimnich F, Frandsen PB, Zwick A, dos Reis M, Barber JR, Peters RS, Liu S, Zhou X, Mayer C, Podsiadlowski L, Storer C, Yack JE, Misof B, Breinholt JW (2019) Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. Proceedings of the National Academy of Sciences of the United States of America 116: 22657–22663. https://doi.org/10.1073/pnas.1907847116

Keck F, Rimet F, Bouchez A, Franc A (2016) phylosignal: an R package to measure, test, and explore the phylogenetic signal. Ecology and Evolution 6(9): 2774–2780. https://doi.org/10.1002/ece3.2051

Keller A, Ankenbrand MJ, Bruelheide H, Dekeyzer S, Enquist BJ, Erfanian MB, Falster DS, Gallagher RV, Hammock J, Kattge J, Leonhardt SD, Madin JS, Maitner B, Neyret M, Onstein RE, Pearse WD, Poelen JH, Salguero-Gomez R, Schneider FD, Tóth AB, Penone C (2022) Ten (mostly) simple rules to future-proof trait data in ecological and evolutionary sciences. Methods in Ecology and Evolution 14: 444–458. https://doi.org/10.1111/2041-210X.14033

Kilmer JT, Rodríguez RL (2016) Ordinary least squares regression is indicated for studies of allometry. Journal of Evolutionary Biology 30: 4–12. hhttps://doi.org/10.1111/jeb.12986

Kinsella RS, Thomas CD, Crawford TJ, Hill JK, Mayhew PJ, Macgregor CJ (2020) Unlocking the potential of historical abundance datasets to study biomass change in flying insects. Ecology and Evolution 10: 8394–8404. https://doi.org/10.1002/ece3.6546

Kivelä SM, Davis RB, Esperk T, Gotthard K, Mutanen M, Valdma D, Tammaru T (2020) Comparative analysis of larval growth in Lepidoptera reveals instar-level constraints. Functional Ecology 34: 1391–1403. https://doi.org/10.1111/1365-2435.13556

Kristensen NP (2003) Skeleton and muscles: adults. In: Kristensen NP (Ed.) Lepidoptera, moths and butterflies, 36(2), Morphology, physiology and development. Handbook of Zoology/Handbook of Zoology 4. De Gruyter, Berlin, New York, 39–131. https://doi.org/10.1515/9783110893724.39

Kristensen NP, Scoble MJ, Karsholt O (2007) Lepidoptera phylogeny and systematics: the state of inventorying moth and butterfly diversity. Zootaxa 1668(1): 699–747. https://doi.org/10.11646/zootaxa.1668.1.30

Legendre P, Legendre L (2012) Numerical Ecology (3rd edn). Elsevier, Amsterdam, 990 pp.

Lister BC, Garcia A (2018) Climate-driven declines in arthropod abundance restructure a rainforest food web. Proceedings of the National Academy of Sciences 115(44): 201722477. https://doi.org/10.1073/pnas.1722477115

Maddison WP, Maddison DR (2023) Mesquite: a modular system for evolutionary analysis. Version 3.81. http://www.mesquiteproject.org

Mayer C, Dietz L, Call E, Kukowka S, Martin S, Espeland M (2021) Adding leaves to the Lepidoptera tree: capturing hundreds of nuclear genes from old museum specimens. Systematic Entomology 46: 649–671. https://doi.org/10.1111/syen.12481

McNab BK (2003) Standard energetics of phyllostomid bats: the inadequacies of phylogenetic-contrast analyses. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology 135(3): 357–368. https://doi.org/10.1016/S1095-6433(03)00090-4

Miller WE (1991) Body size in North American Lepidoptera as related to geography. Journal of the Lepidopterists’ Society 45(2): 158–168. https://www.biodiversitylibrary.org/part/80830

Miller WE (1997) Body weight as related to wing measure in hawkmoths (Sphingidae). Journal of the Lepidopterists’ Society 51: 91–92. https://www.biodiversitylibrary.org/part/80576

Molina-Venegas R, Rodríguez MA (2017) Revisiting phylogenetic signal; strong or negligible impacts of polytomies and branch length information? BMC Evolutionary Biology 17: 53. https://doi.org/10.1186/s12862-017-0898-y

Molina-Venegas R, Moreno-Saiz JC, Castro Parga I, Davies TJ, Peres-Neto PR, Rodríguez MA (2018) Assessing among-lineage variability in phylogenetic imputation of functional trait datasets. Ecography 41(10): 1740–1749. https://doi.org/10.1111/ecog.03480

Müller J, Hothorn T, Yuan Y, Seibold S, Mitesser O, Rothacher J, Freund J, Wild C, Wolz M, Menzel A. (2024) Weather explains the decline and rise of insect biomass over 34 years. Nature 628: 349–354. https://doi.org/10.1038/s41586-023-06402-z

Münkemüller T, Lavergne S, Bzeznik B, Dray S, Jombart T, Schiffers K, Thuiller W (2012) How to measure and test phylogenetic signal. Methods in Ecology and Evolution 3(4): 743–756. https://doi.org/10.1111/j.2041-210X.2012.00196.x

Nakagawa S, Freckleton RP (2008) Missing inaction: the dangers of ignoring missing data. Trends in Ecology and Evolution 23: 592–596. https://doi.org/10.1016/j.tree.2008.06.014

Neter J, Wasserman W, Kutner MH (1983) Applied linear regression models. RD Irwin Inc., Illinois. 547 pp. https://www.ime.unicamp.br/~dias/John%20Neter%20Applied%20linear%20regression%20models.pdf [accessed 01/12/2024]

Oksanen J [maintainer] (2024) Package 'vegan', Community Ecology Package. https://CRAN.R-project.org/package=vegan

Owens HL, Lewis DS, Condamine FL, Kawahara A, Guralnick R (2020) Comparative Phylogenetics of Papilio Butterfly Wing Shape and Size Demonstrates Independent Hindwing and Forewing Evolution. Systematic Biology 69(5): 813–819. https://doi.org/10.1093/sys-bio/syaa029.hal-03032131

Pagel M (1999) Inferring the historical patterns of biological evolution. Nature 401: 877–884. https://doi.org/10.1038/44766

Paradis E, Schliep K (2018) ape 5.3: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35: 526–528. https://doi.org/10.1093/bioinformatics/bty633

Penone C, Davidson AD, Shoemaker KT, Di Marco M, Rondinini C, Brooks TM, Young BE, Graham CH, Costa GC (2014) Imputation of missing data in life-history trait datasets: Which approach performs the best? Methods in Ecology and Evolution 5(9): 961–970. https://doi.org/10.1111/2041-210X.12232

Rasband WS (2018) ImageJ, 1.54. U.S. National Institutes of Health, Bethesda, Maryland, USA. https://imagej.net/ij

R Core Team (2023) R: A language and environment for statistical computing, version 4.3.1. R Foundation for Statistical Computing, Viena. https://www.R-project.org/

Reiss MJ (1989) The allometry of growth and reproduction. Cambridge University Press, Cambridge, 182 pp. https://doi.org/10.1017/CBO9780511608483

Revell LJ (2010) Phylogenetic signal and linear regression on species data. Methods in Ecology and Evolution 1: 319–329. https://doi.org/10.1111/j.2041-210X.2010.00044.x

Revell LJ (2024) phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). PeerJ 12: e16505. https://doi.org/10.7717/peerj.16505

Rogers LE, Hinds WT, Buschbom RL (1976) A General Weight vs. Length Relationship for Insects. Annals of the Entomological Society of America 69: 387–389. https://doi.org/10.1093/aesa/69.2.387

Rota J, Twort V, Chiocchio A, Peña C, Wheat CW, Kaila L, Wahlberg N (2022) The unresolved phylogenomic tree of butterflies and moths (Lepidoptera): Assessing the potential causes and consequences. Systematic Entomology 47(4): 531–550. https://doi.org/10.1111/syen.12545

Roth N, Hacker HH, Heidrich L, Friess N, García-Barros E, Habel JC, Thorn S, Müller J (2021) Host specificity and species colouration mediate the regional decline of nocturnal moths in central European forests. Ecography 44: 1–12. https://doi.org/10.1111/ecog.05522

Sage RD (1982) Wet and dry-weight estimates of insects and spiders based on length. The American Midland Naturalist, 108(2): 407–411. https://doi.org/10.2307/2425505 [Accessed: 01-12-2024]

Schmidt-Nielsen K (1984) Scaling. Why is animal size so important? Cambridge University Press, Cambridge, 241 pp. https://doi.org/10.1017/CBO9781139167826

Seibold S, Gossner MM, Simons NK, Blüthgen N, Müller J, Ambarli D, Ammer C, Bauhus J, Fischer M, Habel JC, Linsenmair KE, Nauss T, Penone C, Prati D, Schall P, Schulze E-D, Vogt J, Wöllauer S, Weisser WW (2019) Arthropod decline in grasslands and forests is associated with landscape-level drivers. Nature 574: 671–674. https://doi.org/10.1038/s41586-019-1684-3

Smith FA, Lyons SK (2013) Animal Body Size. Linking Pattern and Process across Space, Time, and Taxonomic Group. The Chicago University Press, Chicago, 280 pp. https://doi.org/10.7208/chicago/9780226012285.001.0001

Sohlström EH, Marian L, Barnes AD, Haneda NF, Scheu S, Rall BC, Brose U, Jochum M (2018) Applying generalized allometric regressions to predict live body mass of tropical and temperate arthropods. Ecology and Evolution 8(24): 12737–12749. https://doi.org/10.1002/ece3.4702

StatSoft Inc. (2004) STATISTICA, data analysis software system, version 7. Tulsa, Oklahoma.

Stillwell RC, Blanckenhorn WU, Teder T, Davidowitz G, Fox CW (2009) Sex differences in phenotypic plasticity affect variation in sexual size dimorphism in insects: from physiology to evolution. Annual Review of Entomology 55: 227–245. https://doi.org/10.1146/annurev-ento-112408-085500

Stonis JR, Remeikis A, Diškus A, Baryshnikova S, Solis MA (2021) What are the smallest moths (Lepidoptera) in the world? Zootaxa 4942(2): 269–289 https://doi.org/10.11646/zootaxa.4942.2.8

Swenson NG (2014) Phylogenetic imputation of plant functional trait databases. Ecography 37: 105–110. https://doi.org/10.1111/j.1600-0587.2013.00528.x

Teder T (2014) Sexual size dimorphism requires a corresponding sex difference in development time: a meta-analysis in insects. Functional Ecology 28: 479–486. https://doi.org/10.1111/1365-2435.12172

Thurman JH (2022) Beyond the pest: Life history, ecology and ethnoentomology of the giant wood moth (*Endoxyla cinereus*). Austral Ecology 47(4): 733–747. https://doi.org/10.1111/aec.13165

Uyeda JC, Zenil-Ferguson R, Pennell MW (2018) Rethinking phylogenetic comparative methods. Systematic Biology 67(6): 1091–1109. https://doi.org/10.1093/sysbio/syy031

Van Hook T, Williams EH, Brower LP, Borkin S, Hein J (2012) A standardized protocol for ruler-based measurement of wing length in monarch butterflies, *Danaus plexippus* L. (Nymphalidae, Danainae). Tropical Lepidoptera Research 22: 42–52.

van Nieukerken EJ, Kaila L, Kitching IJ, Kristensen NP, Lees DC, Minet J, Mitter C, Mutanen M, Regier JC, Simonsen TJ, Wahlberg N, Yen SH, Zahiri R, Adamski D, Baixeras J, Bartsch D, Bengtsson BÅ, Brown JW, Bucheli SR, Davis DR, De Prins J, De Prins W, Epstein ME, Gentili-Poole P, Gielis C, Hättenschwiler P, Hausmann A, Holloway JD, Kallies A, Karsholt O, Kawahara AY, Koster S, Kozlov MV, Lafontaine JD, Lamas G, Landry JF, Lee S, Nuss M, Park KT, Penz C, Rota J, Schintlmeister A, Schmidt BC, Sohn JC, Solis MA, Tarmann GM, Warren AD, Weller S, Yakovlev RV, Zolotuhin VV, Zwick A (2011) Order Lepidoptera Linnaeus, 1758. In: Zhang Z-Q (Ed.) Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. Zootaxa 3148: 212–221. https://doi.org/10.11646/zootaxa.3148.1.41

Wagner DL, Grames EM, Forister ML, Stopak D (2021) Insect decline in the Anthropocene: Death by a thousand cuts. Proceedings of the National Academy of Sciences 118(2): e2023989118. https://doi.org/10.1073/pnas.2023989118

Wardhaugh CW (2013) Estimation of biomass from body length and width for tropical rainforest canopy invertebrates. Australian Journal of Entomology 52: 291–298. https://doi.org/10.1111/aen.12032

Weiss F, Linde A (2022) How to estimate carabid biomass? -an evaluation of size-weight models for ground beetles (Coleoptera: Carabidae) and perspectives for further improvement. Journal of Insect Conservation 26: 537–548. https://doi.org/10.1007/s10841-022-00391-6

Wiemers M, Chazot N, Wheat CW, Schweiger O, Wahlberg N (2020) A complete time-calibrated multi-gene phylogeny of the European butterflies. ZooKeys 938: 97–124. https://doi.org/10.3897/zookeys.938.50878

Supplementary material 1

Documentation on phylogeny

Authors: Enrique García-Barros

Data type: pdf

Explanatory note: Documentation on phylogeny. Documentation used to assemble the branching pattern of the cladogram for the species included in the study.

Link: https://doi.org/10.3897/nl.48.144747.suppl1

Supplementary material 2

Tree topology

Authors: Enrique García-Barros

Data type: nwk

Explanatory note: Tree describing the topology of the phylogenetic hypothesis used in this study.

Link: https://doi.org/10.3897/nl.48.144747.suppl2

Supplementary material 3

List of species and available BOLD sequences

Authors: Enrique García-Barros

Data type: pdf

Explanatory note: List of the species included in the study, and the related COI sequences adopted to estimate between-species distances. The CODE is an arbitrary one (A001 to ZL013) used to denote each operative species-level taxon in this study. When no sequences associated to a species were found, that available from a closely related species (e.g., same genus) was adopted (listed in the column headed "Replace. / altern. SP") to get approximated between-species distances at different taxonomic levels. The information was retrieved from the BOLD data portal (https://v3.boldsystems.org/) during 2023.

Link: https://doi.org/10.3897/nl.48.144747.suppl3

Supplementary material 4

Additional statistics

Authors: Enrique García-Barros

Data type: pdf

Explanatory note: Supplementary tables describing the bivariate correlations (table S1), taxon-by-taxon regression results (table S2) and ANCOVA approaches (table S3).

Link: https://doi.org/10.3897/nl.48.144747.suppl4

Zeitschrift/Journal: Nota lepidopterologica

Jahr/Year: 2025

Band/Volume: 48

Autor(en)/Author(s): Garcia-Barros Enrique

Artikel/Article: Refining estimates of dry body weight from linear measurements in adult moths and butterflies (Lepidoptera) 165-184