Optimising trait and source selection for explaining occurrence and abundance changes: a case study using British Butterflies

Joe Middleton-Welling[1], Rachel A.Wade[1], Roger L. H. Dennis[1,2], Leonardo Dapporto[3], Tim G. Shreeve[†1]

1. Centre for Ecology, Environment and Conservation, Faculty of Health and Life Sciences, Oxford Brookes University, Headington, Oxford OX3 0BP, U.K.

2. School of Life Sciences and Education, Staffordshire University, Science Centre, Leek Road, Stoke-on-Trent ST4 2DF, U.K.

3. Dipartimento di Biologia, Via Madonna del Piano, 6, 50019 Sesto Fiorentino, Florence, Italy.

† Correspondence author e-mail: tgshreeve@brookes.ac.uk

Running head: Pitfalls of trait-based analyses

ABSTRACT

1. Functional traits that define the ecological role of an organism are increasingly being used to determine and predict responses to environmental change. Functional trait analyses of butterflies remain underexplored compared with other taxa, such as plants. Previous works using butterfly functional traits have not comprehensively addressed issues about the quality of trait data sets used and the relative predictive power of different trait types.
2. We compare the consistency of trait descriptions between six widely used trait sources for the British butterfly fauna. We analysed consistency of trait sources using Fleiss's kappa and ICC. PCA was used to produce species ordinations, comparing outputs to examine which trait sets were better at explaining recent species range and abundance changes within the UK.
3. There was a large range in congruence values for specific traits between sources. No single source can be relied upon to produce accurate trait information for British butterflies. Most trait sets are poor predictors of abundance and occurrence changes but are better at predicting current occurrence. An extensive trait set, supplementing biotope-related traits with explicit resource-based information recovers more informative ecological classifications and models than those primarily based on life-history traits or biotope descriptors. Smaller trait sets do, however, recover the specialist-generalist continuum.
4. We conclude that analyses of distribution and abundance changes that rely on traits are highly dependent on trait source and trait type. For butterflies, traits that are based on measures of biotope occupancy should be avoided in explaining changes of abundance and distribution. Including trait information that describes their resource requirements is essential for such analyses.

Key-words: butterfly biogeography, functional-traits, habitat, occurrence and abundance change, resource-use

INTRODUCTION

Functional traits are characteristics that define the ecological roles of organisms, including their interactions with other species (Díaz & Cabido, 2001). Functional trait-based studies of plant communities have revealed insights into plant strategies and tolerances of environmental conditions (e.g. Grime, 1977), invasion dynamics (e.g. Pyšek et al., 2015) and distribution changes (e.g. Powney, Giovanni, Preston, Purvis & Roy, 2014). Implicit to trait-based analyses is the understanding that identifying functional groups from shared traits will elucidate sets of species with shared responses to past, current and future climate and land-use changes and reveal the development of novel communities with environmental change. Trait-based analyses of animals are becoming common (e.g. Kotiaho, Kaitala, Komonin & Päivinen, 2005; Betzholtz, Pettersson, Ryrholm & Franzén, 2013; Luck, Carter & Smallbone, 2013), but tend to rely on small trait sets and do not consider intraspecific trait variation to the same extent as plant-based studies (e.g. Kichenin, Wardle, Peltzer, Morse, & Freschet, 2013; Celis, Halpen & Jones, 2017). Theoretical work has suggested that including intraspecific variation is essential for addressing questions of community assembly (Siefert, et al., 2015) and such studies tend to use botanical systems as examples. The same theoretical framework is not applied to animal studies to the same extent, and not at all in those that use insects as model systems.

Butterflies respond rapidly to environmental change, being representative of wider trends in insect abundance (Thomas, 2005). Trait-based analyses of butterflies are relatively well established (e.g. Mattila, Kotiaho, Kaitala, Komonen & Päivinen, 2009) with commonly chosen traits including larval hostplant specificity, habitat/biotope breadth, flight period, body size and voltinism (e.g. Dennis et al., 2004; Kotiaho et al., 2005; Pöyry, Luoto, Heikkinen, Kuusaari & Saarinen, 2009; Betzholtz et al., 2013; Gallien, Altermatt, Wiemers, Schweiger & Zimmermann, 2017). Other analyses have used various measures of mobility as predictors of responses to environmental changes, such as mobility ranking based on mark-release-recapture studies (Warren et al., 2001) and expert opinion (Nilsson, Franzén, & Jönsson, 2008). Analyses that focus on resources and behaviours generally include many more parameters than those that focus on morphology, physiology or voltinism. For example, Shreeve, Dennis, Roy and Moss (2001) used 145 character states (traits), independent of biotope occupancy, to identify relationships between trait sets and mobility, conservation status and distributions of UK butterflies; Pavlikova and Konvicka (2012) later expanded this approach to European macromoths.

Recently Moretti et al. (2016) provided a list of traits for terrestrial invertebrates applicable across taxa and measurable in a standardized way for individuals. In fact, it has been suggested that functional traits should be measured at the individual rather than population, community or ecosystem levels to avoid ambiguity about the mechanisms underlying community structure and ecosystem functioning (Violle et al., 2007). From this perspective, a series of traits commonly used for butterflies obtained from their distribution (e.g. range size, climatic niche characteristics; see Schweiger, Harpke, Wiemers & Settele, 2014)) or by defining the biotope(s) in which the species occur, cannot be considered as functional traits.

Reliable trait-based analyses require two main assumptions to be met. The first is that trait information is accurate and fully encompasses variation, both within and among species. Dennis, Hardy & Dapporto (2012) argued that trait data should be explicitly labelled temporally and spatially in order to account for variation;despite this, such explicit labelling is not common. Fitzsimmons (2013), working on Canadian butterfly species, showed that commonly used traits were not reported consistently across different literature sources. Traits with a generally low level of congruence between sources included overall habitat breadth, presence/absence in different habitat types and behavioural traits. This may indicate that individual trait sources do not report the full range of variation for these traits. European Lepidoptera are comparatively better studied than those of Canada; the UK in particular has an especially long history of field guides, reference works and, more recently, online resources for Lepidoptera (e.g. South, 1906; Frohawk, 193; Emmet & Heath, 1989; Asher et al., 2001, Eeles, 2016). Recent trait-based analyses (Curtis, Brereton, Dennis, Carbone & Isaacs, 2015) conducted on the UK butterfly fauna have used works such as Dennis (2010) but consistency between trait sources, and thus the reliability of trait-based conclusions, remains unclear.

The second assumption of reliable trait-based analyses is that the adaptive response of species' traits to land-use and climate changes are fully understood. This depends on the relative plasticity of each trait, which is not commonly considered in trait-based studies. This second assumption is more intractable than the first, carrying a risk of circular reasoning. In Lepidopteran trait-based analyses of distribution changes, the main traits used are some measure of dispersal (often approximated by wingspan or more rarely by ratio of thorax width to forewing length) (Sekar, 2012; Gallien et al., 2017), larval hostplant specificity or number of hostplants used (Öckinger et al., 2010) and overwintering stage (Mattila et al., 2009). Some traits are logical descriptors of occurrence; Dennis et al. (2005) found a positive correlation between larval hostplant breadth and overall distribution for UK butterflies. Other traits can be more problematic. 'Habitat' is also a commonly used trait in many analyses; despite often lacking a precise definition in many studies (cf., Dennis, Shreeve & Van Dyck, 2003; Dennis, Dapporto & Dover, 2014); there is a tendency to associate 'habitat' with loosely defined biotopes. If defined as a vegetation association, 'habitat' is not consistent with resource requirements, which are frequently more precise (e.g. woodland vegetation structures could include several overlapping resources). The association of species with habitats/biotopes becomes increasingly uncertain the more structurally complex the biotope and the greater the range of resources a species uses (Shreeve et al., 2001; Dennis et al., 2005; Vanreusel & Van Dyck, 2007). Resource distributions are not fixed in time and any measure of habitat used in trait-based studies should take this into account as species are expected to shift their vegetation associations in response to climate warming (Oliver, Hill, Thomas & Roy, 2009). Observed dietary breadth has also been known to shift in response to climate warming (Pateman, Hill, Roy, Fox, & Thomas, 2012) and mobility has equally been shown to be a variable trait. For instance, measured mobility potential has rapidly shifted in response to habitat fragmentation (Dennis, 1993; Shreeve, 1995; Hill, Thomas & Lewis, 1999) and is a selected response to previous and current landscape configurations (Dennis, 1977; Shreeve & Dennis 2011; Dennis, Dapporto, Dover & Shreeve, 2013) rather than being invariant. As these traits have been used as predictors of responses to environmental change, whilst also being affected by these changes, circular reasoning is a legitimate concern at least when trait plasticity is not considered.

To address these issues of trait reliability and their resulting predictive power we (i) examine the variability in trait information for British butterflies, (ii) determine the sensitivity of relationships between species occurrence and abundance and traits to the source of trait data, (iii) compare these relationships with those generated by randomly selected trait sets (iv) compare the sensitivity of trait-based analyses to the inclusion of different trait types, and (v) recommend how trait-based analyses of Lepidoptera could be improved to predict community responses more reliably to current and predicted environmental changes.

METHODS

Species selection and data sources

Fifty-six species of habitual current resident UK butterflies were selected for analysis. Following Fitzsimmons (2013), rare migrants and species described as ‘migrant’ were excluded; these latter comprise *Vanessa atalanta* (but see: Fox & Dennis, 2010), *Vanessa cardui* and *Colias* species. Species that may have a substantial migratory portion, but are generally considered to overwinter and breed in the UK (*Aglais* and *Pieris* species) were included. Trait sources were selected based on the following criteria: 1) containing information on all 56 resident UK species, and 2) were published recently or are currently in wide use. Five sources of non-resource-based trait information (four books: Emmet and Heath (1989); Asher et al. (2001); Thomas (2010); Newland, Still, Swash & Tomlinson (2015) and one online resource: Eeles (2016)) and one resource-based trait source (Dennis, 2010) were selected.

The recent splitting of *Leptidea sinapsis* into a species pair (*Leptidea sinapis/juvernica* (Dincă, Lukhtanov, Talavera & Vila, 2011), is not reflected in all the sources used. In this case, where trait differences were noted by locality, only traits related to locations matching the currently known locality of *Leptidea sinapis* were used for coding. Species nomenclature used in this study follows that used in Fauna Europaea (Karsholt & van Nieukerken, 2016), except *Leptidea sinapis,* where we have followed Dincă et al. (2011).

Trait coding

Data were collected and coded for 23 trait types, following Fitzsimmons (2013) and containing information on: biotope usage, behaviour, resource use, ecological niche and wingspan (Table 1). Biotope usage was coded using the European Nature Information System (EUNIS) habitat codes (Davies, Moss & Hill, 2004). When information on a trait was not provided by a source it was left blank. To minimise interpersonal bias, all coding was carried out by one author (JMW). Our species trait databases are available from the Dryad Digital Repository (Middleton Welling et al, 2018).

Data analyses

Ordinal and categorical trait agreement between sources was measured using Fleiss's kappa and continuous trait agreement was measured using a two-way intraclass correlation coefficient (ICC). Kappa and ICC scores vary from -1 (complete disagreement) to 1 (complete agreement) with 0 being equivalent to the amount of agreement expected by chance. Kappa and ICC scores were generated using the 'irr' package (Gamer, Lemon, Fellows & Singh, 2012) in R (R Core Development team, 2017)

In order to assess how trait source affected predictions of species responses to changing environmental conditions we performed PCA ordinations on data from the different trait sources , with and then without biotope information, using the 'dudi.mix' function of 'Ade4' (Dray & Dufour, 2007). Each PCA was performed on a correlation matrix scaled to unit variance. *Leptidea juvernica* was removed from the datasets (where present) so that all data sources were comparable. The number of components extracted in each PCA was determined by extracting those components with eigenvalues > 1. The 'protest' function in the 'vegan' package (Oksanen et al., 2016), returning a statistic which measured the level of correlation between two matrices, was used to test the degree of concordance between all the PCA outputs, both with and without biotope information.

Trait codings for the PCA analyses were the same as for the Kappa statistics, except when missing data would lead to biologically nonsensical results. Mate-locating method was coded as two separate traits - perching and patrolling. Myrmecophily was removed from the trait list because this information was absent in most sources, including for those species known to have some association with ants. Hilltopping, flight month information and wingspan were also removed as these variables are not present in all trait sources. Any remaining missing data within sources were assigned the average values of particular traits for that source.

Fifty random datasets (null models) were generated in R in order to compare the explanatory power of the different trait sources to randomly assembled trait sets. We first calculated the number of occurrences of each trait state using all the data sets. We then randomly sampled from these trait distributions to generate each random species x trait matrix. Random datasets were used to generate PCAs using the methods described for the source data.

To determine the explanatory power of each trait set the PCA case scores produced for each trait source were regressed against the seven metrics provided in the State of the UK's Butterflies 2015 (Fox et al. 2015). These were: the number of 10km, 1km squares and UKBMS sites occupied in 2014, and long-term (1976-2014) and short term (2005-2014) percentage occurrence and percentage abundance changes. This information is the most recent publically available assessment of short term and long term butterfly trends for the whole UK butterfly fauna. The PCA case scores were also regressed against 'colonfac', a commonly used measure of generalism (Dapporto & Dennis, 2013), for all trait sources that exclude biotope associations. The colonfac scores for all species were taken from Dapporto and Dennis (2013: Table A1). Colonfac measures species generalism by quantifying resource breadth; a higher value indicates an increased ability to exploit a hypothetical vacant patch. Dependent variables were normalised using min-max normalisation. Regressions were conducted using a generalized linear model (GLM) with quasibinomial error function and a logit link function. Model assumptions were assessed by examining the normality of the residuals and the linearity of the Q-Q plots. The first six components extracted using the PCA were initially all included as co-predictors. Two sets of models were built, one set with biotope information included and one set where it was excluded. Models were compared using log-likelihood values and the best models were retained. Factor scores underlying the principal components were compared in order to see whether the variables that were correlated with particular principle components were analogous between sources.

To test which trait types correlate most strongly with abundance and occurrence change we repeated the analysis using the dataset from Dennis (2010). We split traits into three types; ‘life-history traits’, ‘biotope traits’ and ‘resource-specific traits’. Life-history traits are those that define the life-history strategy of a particular species, by either affecting the reproductive output or relative investment in particular life-history stages. Biotope traits describe the biotope in which a species is commonly found. Resource-specific traits define consumables that a species uses either as a larva or adult, or a utility that a species uses as part of a behaviour but does not consume. We carried out the same analysis as on the more restricted trait sets; we produced PCAs for each trait set and then correlated the components with the measures of conservation status given in Fox et al. (2015). We then used the same model structure and simplification procedure as with the more restricted trait sets. The analyses were carried out on life-history traits alone, with life-history traits in combination with either biotope or resource specific traits and with the full trait set. We compared the ordinations produced by the different trait combinations using the ‘protest’ function in the ‘vegan’ package (Oksanen et al., 2016).

RESULTS

Intra-source trait consistency

Inconsistencies were found between sources for all traits. In particular, these may be caused by the lack of coverage of behavioural data in some sources and difficulty in applying EUNIS habitat criteria to sources that used unclear habitat terminology. These behavioural and habitat traits were therefore often the most inconsistent (Table 2). In comparison, basic life-history information (viz., number of generations, overwintering stage and size) was available for all species and was the most consistently described. In general the extent of agreement varied widely between both sources and traits.

Multivariate analysis

When biotope type and breadth were included in the analyses, the first component consistently explained 20-28% of the variation (see Table S1A in Supporting Information) and represented a continuum from biotope specialism to generalism; biotope total (i.e. the total number of EUNIS biotope categories that a species was recorded as utilising) was a contributing trait for all datasets. The contributions of other trait types to the first and other components were inconsistent between trait sources (see Table S2A). With biotope excluded, the first component explained 26-30% of the variation (see Table S1B) with overwintering stage, hostplant specialism and number and generation number being the most important traits. Overall, the contributions of all the underlying variables (identified from the factor scores) differed between datasets, especially for the more minor components (see Table S2B).

The Protest analysis indicates that the PCAs of each source are significantly correlated with each other, but there is evidence that each ordination is marginally different (Table 3). The random datasets were more different from the real datasets than the real datasets were from each other. Principal component analysis of data from Dennis (2010) produces an ordination that is primarily explained by resources used by all life cycle stages and adult behaviour (Table S4).

Does biotope improve correlation with change of status?

With biotope included, the best models produced by all six trait sources were highly correlated ($p < 0.001$) with all three of the measures of current occurrence from Fox et al. (2015). This is largely due to the first component being highly correlated with current occurrence (see Table S3A). Half of the sources also produced a model that was significantly correlated with long term occurrence change ($p < 0.05$) although the components that were included in the best model varied (see Table S3A) and there are differences in the contributing traits (see Table S3A). With biotope occupancy included, four sources: Emmet and Heath (1989), Asher (2001) Thomas (2010), Newland et al. (2015) and Eeles (2016) produced models that were significantly correlated with long term abundance change ($p < 0.05$). These models all included the first component as a significant predictor but otherwise varied in their composition (see Table S3B) and explanatory power (see Table 4). No trait source produced models that were significantly correlated with short-term occurrence changes. Most of the relationships produced from the restricted trait sources are either weaker or not significant when biotope is excluded and models tend to include fewer components. There are two notable instances where there is a more significant correlation when biotope is excluded: both Thomas (2010) and Newland et al. (2015) include a component related to long-term occurrence change ($p < 0.01$). Although this was more significant than when biotope was included the component involved (component 6) was minor, explaining 7.5% of the total variation (Newland et al., 2015) and 6% for Thomas (2010). For all trait sources the first component was also significantly related to the 'colonfac' scores ($p < 0.001$ in all cases) of Dapporto and Dennis (2013).

Comparing trait types

The models produced by the various permutations of the Dennis (2010) dataset generally mirrored those produced by the trait sources that used a more restricted set of traits. All of the trait subsets derived from Dennis (2010) produced a model that was significantly correlated ($p<0.001$) with all three measures of current occurrence (Table 4) from Fox et al. (2015). These models included either the first (for life-history alone and life-history and biotope traits) or the second component (for life-history and resource based traits and all traits) being significantly correlated with current abundance. Unlike the trait sources based on more restricted traits, most of the trait subsets of Dennis (2010) failed to produce a model that significantly correlated with long-term occurrence changes. The only subset to do so comprised life-history and biotope traits (Table 4). In this case the third and the fourth components were significantly correlated ($p < 0.01$ to $p < 0.05$) with long-term occurrence changes. All subsets of the traits from Dennis (2010) produced models that contained a component that was significantly correlated ($p < 0.05$) with short term abundance changes. For the trait set containing only life history traits this was the seventh component but for all other subsets this was the first component. None of the subsets of the trait data in Dennis (2010) produced models that were significantly correlated with long-term abundance change. With all traits included the first component separated species on the basis of voltinism (univoltine vs multivoltine) and overwintering stage (Table S4 & Figure 1A). With life-history and just biotope it becomes more challenging to see ecologically relevant relationships (see Table S4 & Figure 1B). With life-history and resource traits only, the first components separate species depending on whether they use arboreal structures (shrubs and trees) extensively during their life cycle vs species that primarily use grass and herbaceous structures (Table S4 and compare Figures 1A and 1C). The second component separated species of xeric grassland (typically lycaenids) from generalist grassland species and those that are adapted to ruderal hostplants (typically Pieridae and *Aglais* species).

DISCUSSION

Key results

There is a significant disagreement in trait values between trait sources. The level of disagreement is similar to that reported for Canada which has a less studied butterfly fauna. The trait sources are more similar to each other than randomly generated trait sets and in general are highly co-correlated. Disagreement between trait sources is especially apparent for biotope descriptions. In general, most sources provide trait values that are adequate at explaining current occurrence but are not effective at explaining occurrence and abundance change over time. When comparing different trait types it is apparent that models are generally improved by including either biotope or resource based traits in addition to those that explain life-history. We recommend using more than one trait source to capture the full range of trait variation and to include as many traits as possible, at least during any initial analysis.

Congruence and variation within traits

A large range in congruence values for specific traits occurs between data sources. These were highest for traits relating to wingspan, overwintering stage and voltinism and lowest for traits relating to habitat preference and the adult behaviours of mudpuddling and hilltopping. All data sets tended to agree on perching and patrolling mate-locating strategy when this was present. Despite high levels of discordance being present for some traits, the different sources were more similar to each other than to randomly generated trait sets (Table 3). This indicates that the trait discordance does not obscure the broad differences between species present in the UK fauna.

The low levels of agreement found for traits describing habitat type are primarily due to inconsistent or imprecise descriptions in some trait sources. Some species have well characterised vegetation type associations (e.g. *Erebia epiphron* for *Nardus* grassland) (Emmet & Heath, 1989; Newland et al., 2015), but species occurring in a wide range of biotopes (e.g. *Pieris* and *Aglais* species) may be described as simply 'widespread' (e.g. Emmet & Heath, 1989). In such cases the level of detail is not suitable for quantitative analysis and subjective wording can lead to divergent coding and poor congruence amongst sources.

Adult behavioural traits such as mudpuddling and hilltopping had low congruence values and contained large amounts of missing data. Behavioural trait data is only currently being aggregated (e.g. Dennis et al., 2014) with Dennis (2010) giving the most comprehensive set of behavioural information. Highly congruent data types - wingspan, voltinism and overwintering stage - are either traits that have been well-documented from the early 20th century (e.g. South, 1906; Frohawk 1934) or the data have been repeated from one or more sources, most likely from Emmet and Heath (1989). Whilst this has led to consistency it does not follow that they are accurate, especially if the traits are spatially and temporally variable.

Unlike Fitzsimmons (2013) we chose to not weight some characteristics, making no *a priori* decisions about the relative importance of particular traits or the reliability of the information. Despite this, our conclusions are similar; the lowest congruence occurs in those traits that need direct field observation (e.g. adult behavioural characteristics). It is surprising that data from Britain and Canada share this trend, given the long history of butterfly ecology literature in the British Isles, the limited size of its fauna (c.56 species cf. 263 for Canada) and publication of the first guide to Canadian butterflies being relatively recent (Layberry, Hall & Lafontaine, 1998).

Disagreement between trait sources suggests that variation in trait values within species is not adequately described by any of the sources (e.g. voltinism may be more geographically variable than previously assumed). Although the expectation is that trait variation between species will increase with the number of species, our analysis produced similar Kappa values to Fitzsimmons (2013), which used a much larger number of species. Our results support those of Fitzsimmons (2013), who suggested that measures of inter-source concordance should be included in analyses and reliance on single sources should be avoided. Further work into the variability *within* traits, both temporally and spatially, will help determine how much is caused by real error rather than an underestimate of underlying variation.

The relationships between different traits

For each data source the first principal component separated species that were found across many habitats (habitat generalists) from those found in few (habitat specialists). The relationship between this first principal component and the ‘colonfac’ measure (Dapporto & Dennis, 2013) indicates that this component represents some aspect of the specialist-generalist continuum. The first or third components derived from each trait source are primarily related to the degree of hostplant specialism. Hostplant specificity has long been regarded as an important indicator of overall 'generalism' (Wiklund, 1981; Carnicer, Stefanescu, Vila, Dincă, Font & Peñuelas, 2013; Eskildsen, Carvalhiero, Kissling, Biesmeijer, Schweiger & Høye, 2015), and relates to distribution status (Dennis et al., 2004, 2005; Carnicer et al., 2013), although there are widespread ubiquitous species that use single host-plant species (e.g. *Aglais urticae* and *A. io*). Also for some species, hostplant specificity varies within geographic range (e.g. *Callophrys rubi* (Dennis, 1992)). The second extracted component is generally related to voltinism, also proposed as another important explanatory trait (Carnicer et al., 2013; Dapporto & Dennis, 2013; Dennis, Hardy & Dapporto, 2015) as it can be used as a measure of mobility, being a proxy for flight period duration. However, it is not commonly included as part of the specialism/generalism continuum (Bartonova, Benes & Konvicka, 2014).

Selecting 'good' traits - moving away from biotopes and towards resources

In comparing the models produced using different trait types it is apparent that adding either biotope or resource traits (either separately or in combination) helps to simplify the models produced but does not increase their explanatory power; fewer components are returned in the best models as the number of traits included is increased. For example, when only life-history traits are included, there are seven components that each explain more than 5% of total variation; this is reduced to five components when all the traits are included. Our results closely match those of Shreeve et al. (2001) who used an earlier version of Dennis' (2010) database employed in this study. Like Shreeve et al. (2001) we found that when we included the full dataset the first component primarily separated species into a continuum from species associated with arboreal structures, typical of forest and shrubland biotopes to species that occupied what would be typically described as grassland biotopes. This relationship was recoverable without including biotope as an original variable, which shows that ecologically relevant information can be revealed without the risk of circular reasoning that biotope traits inevitably introduce into any trait-based analysis. The overall results indicate that if information about complex traits is collected, the explained variation becomes greater and potentially of more value for practical conservation. One common criticism of trait-based studies is that their conclusions are often ambiguous and difficult to translate into practical conservation policy (Cardillo & Meijaard, 2012). Our results indicate that it is possible to link species responses to environmental change using detailed traits and for these traits to produce components that explain ecologically meaningful groupings. Resource-based traits as used here and by Shreeve et al. (2001), can be used to determine how groups of species are likely to respond to management practices, whereas imprecise descriptors such as 'woodland' do not indicate which components of that biotope are important. Thus biotope occupancy, often used as a trait to explain changes of butterfly abundance and distribution, is only of marginal use in explaining changes of status. It primarily explains this status change as a function of the status of the biotope without identifying interactions of species with resource distributions and abundance. The responses of species within any biotope to management, landscape and climate changes will therefore be unique as the occurrence and quality of individual species' resource sets will respond individually to changing conditions.

Comparisons with previous work

Recent works, (e.g. Pöyry et al., 2008; Diamond, Frame, Martin & Buckley, 2011; Eskildsen et al., 2015) generally use a narrower range of traits than either our study or the studies of Dennis and colleagues. Whilst simple relationships between traits can be extracted from limited trait sets there is a risk of oversimplification, especially if traits are treated as being invariant. We have found that trait information varies considerably between sources, so it is unwise to select individual traits and treat them as fixed values especially over the long time scales that are necessary for effective conservation (e.g. Eskildsen et al., 2015).

Trait selection for conservation

Although limited trait sets were reasonably good at explaining some recent changes of abundance and occurrence, the reliance on biotope or vegetation association descriptors in trait-based analyses has an inherent risk despite these traits increasing the overall explanatory power of the analyses. Recent increases in global temperatures are relatively modest in comparison to future projections (IPCC, 2014). Vegetation associations are expected to change in response, with potentially more fundamental changes in microhabitat structure and microclimate. It is unlikely that current vegetation associations are analogous to those from the early Holocene or earlier (Dennis & Shreeve, 1988) and equally no assumption should be made about future associations (Vera, 2000). Previous studies have shown that if the right traits are chosen, biotope traits are not necessary to predict meaningful relationships with environmental change (WallisDeVries, 2014). Additionally, biotope occupancy-based traits are likely to have decreasing predictive power over long time scales and trait-based analyses should focus on traits that govern specific vegetation and physical structure usage and describe resource use, rather than relying on the interaction of these resource requirements with current vegetation structures, i.e. current biotope associations.

In conclusion, we recommend that wherever possible, trait-based analyses should be focused explicitly on a resource-based view of traits, especially when aiming to investigate complex inter-species and species-environment relationships. Trait over-splitting should be avoided, and a rigorous analysis of inter-trait correlations should be performed prior to any analyses. Conservation actions and biogeographic appraisals should rely on finely audited resource usage traits that are ideally spatially and temporally explicit. For many European taxa, this level of trait information is not collated but much exist in the literature. A promising amount of data has been gathered for UK Lepidoptera using UKBMS transects (Curtis et al., 2015), with which the role of traits in species distributions and abundances can be tested. In recent years, other European countries have started similar projects and we argue that for butterflies the construction of a Europe-wide trait-database should be a priority for future research. Our analyses are based on the UK butterflies which may have more restricted niche dimensions and trait states in comparison to their potential trait states in mainland Europe. A European trait-database has the potential to address the issue of the effects of trait plasticity on species responses to environmental changes. Outside of Europe and North America detailed trait information may not currently exist, although we believe that information on the Lepidoptera may not be as sparse as for other groups.

ACKNOWLEDGEMENTS

Nora Braak provided helpful discussions. Roger Vila assisted with proofreading the manuscript. RAW is funded by a Daphne Jackson Research Fellowship via NERC and by Oxford Brookes University ($p < 0.05$) and JMW by a Nigel Groome Studentship.

AUTHORS’ CONTRIBUTIONS

The concept was designed by JMW, TGS and RAW and refined by RLHD and LD. All authors refined the analyses and contributed equally to the drafting and revisions of the article.

DATA ACCESSIBILITY

Data deposited in the Dryad Digital Repository: http://doi.org/10.5061/dryad.017vp2q, (Middleton Welling et al. 2018)

REFERENCES

Asher, J., Warren, M., Fox, R., Harding, P., Jeffcoate, G. & Jeffcoate, S. (2001). The Millennium Atlas of Butterflies in Britain and Ireland. Oxford: Oxford University Press.

Bartonova, A., Benes, J. & Konvicka, M. (2014). Generalist-specialist continuum and life history traits of Central European butterflies (Lepidoptera) - are we missing a part of the picture? European Journal of Entomology, 111, 543-553. doi: 10.14411/eje.2014.060

Betzholtz, P.E., Pettersson, L.B., Ryrholm, N. & Franzén, M. (2013). With that diet, you will go far: trait-based analysis reveals a link between rapid range expansion and a nitrogen-favoured diet. Proceedings of the Royal Society B. Biological Sciences. 280, 20122305. doi: 10.1098/rspb.2012.2305

Carnicer, J., Stefanescu, C., Vila, R., Dincă, V., Font, X. and Peñuelas, J., 2013. A unified framework for diversity gradients: the adaptive trait continuum. Global Ecology and Biogeography, 22, 6-18. Doi: 10.1111/j.1466-8238.2012.00762.x

Cardillo, M. & Meijaard, E. (2012). Are comparative studies of extinction risk useful for conservation? Trends in Ecology & Evolution, 27, 167-171. doi: 10.1016/j.tree.2011.09.013

Celis, J. Halpern, C.B. & Jones, F.A. (2017). Intraspecific trait variation and the differential decline of meadow species during conifer encroachment. Plant Ecology, 218, 565-578.

Curtis, R.J., Brereton, T.M., Dennis, R.L.H., Carbone, C. & Isaacs, N.J.B. (2015). Butterfly abundance is determined by food availability and is mediated by species traits. Journal of Applied Ecology, 52, 1676-1684. doi: 10.1111/1365-2664.12523.

Dapporto, L. & Dennis, R.L.H. (2013). The generalist–specialist continuum: testing predictions for distribution and trends in British butterflies. Biological Conservation, 157, 229-236. doi: 10.1016/j.biocon.2012.09.016

Davies, C.E., Moss, D & Hill M.O., (2004). EUNIS habitat classification revised 2004. European Environment Agency & European Topic Centre on Nature Protection and Biodiversity.

Dennis, R.L.H. (1977). The British Butterflies. Their Origin and Establishment. Faringdon, Oxford: E.W. Classey Ltd.

Dennis, R.L.H. (ed.) (1992). The Ecology of Butterflies in Britain. Oxford: Oxford University Press.

Dennis, R.L.H. (1993). Butterflies and climate change. Manchester: Manchester University Press.

Dennis, R.L.H. (2010). A resource-based habitat view for conservation: butterflies in the British landscape. Oxford: John Wiley & Sons.

Dennis, R.L.H., Dapporto, L. & Dover, J.W. (2014). Ten years of the resource-based habitat paradigm: the biotope-habitat issue and implications for conserving butterfly diversity. Journal of Insect Biodiversity, 2, 1-32.

Dennis, R.L.H., Dapporto, L., Dover, J.W. & Shreeve, T.G. (2013). Corridors and barriers in biodiversity conservation: a novel resource-based habitat perspective for butterflies. Biodiversity & Conservation, 22, 2709-2734. doi: 10.1007/s10531-013-0540-2

Dennis, R.L.H., Hardy P. & Dapporto, L. (2012). Nestedness in island faunas: novel insights into island biogeography through butterfly community profiles of colonization ability and migration capacity. Journal of Biogeography, 39, 1412-1426. doi: 10.1111/j.1365-2699.2012.02698.x

Dennis, R.L.H., Hodgson, J.G., Greyner, R., Shreeve, T. G. & Roy, D. B. (2004). Host plants and butterfly biology. Do hostplant strategies drive butterfly status? Ecological Entomology, 29, 12-26. doi: 10.1111/j.1365-2311.2004.00572.x

Dennis, R.L.H. & Shreeve, T.G. (1988). Hostplant-habitat structure and the evolution of butterfly mate-location behaviour. Zoological Journal of the Linnean Society, 94, 301-318. doi: 10.1111/j.1096-3642.1988.tb01198.x

Dennis, R.LH., Shreeve, T.G., Arnold, H.R. & Roy, D.B. (2005). Does diet breadth control herbivorous insect distribution size? Life history and resource outlets for specialist butterflies. Journal of Insect Conservation, 9, 187-200. doi:10.1007/s10841-005-5660-x

Dennis, R.L.H., Shreeve, T.G. & Van Dyck, H. (2003). Towards a functional resource-based concept for habitat: a butterfly biology viewpoint. Oikos, 102, 417-426. doi: 10.1034/j.1600-0579.2003.12492.x

Diamond, S.E., Frame, A.M., Martin, R.A. & Buckley, L.B. (2011). Species' traits predict phenological responses to climate change in butterflies. Ecology, 92, 1005-1012. doi: 10.1890/10-1594.1

Díaz, S. & Cabido, M. 2001. Vive la difference: plant functional diversity matters to ecosystem processes. Trends in Ecology and Evolution, 16, 646-655. doi: 10.1016/S0169-5347(01)02283-2

Dincă, V., Lukhtanov, V.A., Talavera, G. & Vila, R. (2011). Unexpected layers of cryptic diversity in wood white *Leptidea* butterflies. Nature Communications, 2, 324. doi: 10.1038/ncomms1329

Dray, S. & Dufour, A.B. (2007). The ade4 package: implementing the duality diagram for ecologists. Journal of Statistical Software, 22, 1-20.

Eeles, P. (2016). UK Butterflies [online] Available at: http://www.ukbutterflies.co.uk

Emmet, A. & Heath, J. eds., (1989). The Moths and Butterflies of Great Britain. Colchester: Harley Books.

Eskildsen, A., Carvalhiero, L.G, Kissling, W.G., Biesmeijer, J.C., Schweiger, O. & Høye, T.T. (2015). Ecological specialization matters: long-term trends in butterfly species richness and assemblage composition depend on multiple functional traits. Diversity and Distributions, 21, 792-802. doi: 10.1111/ddi.12340

Fitzsimmons, J. M. (2013). How consistent are trait data between sources? A quantitative assessment. Oikos, 122,1350-1356. doi: 0.1111/j.1600-0706.2013.00207.x

Fox, R., Brereton, T. M., Asher, J., August, T.A., Botham, M. S., Bourn, N. A.D., Cruickshanks, … Roy, D. B. (2015). The State of the UK's Butterflies 2015. Wareham: Butterfly Conservation and the Centre for Ecology & Hydrology.

Fox, R. & Dennis, R.L.H. (2010). Winter survival of *Vanessa atalanta* (Linnaeus, 1758) (Lepidoptera: Nymphalidae): A new resident butterfly for Britain and Ireland? Entomologist's Gazette, 61, 94-103.

Frohawk, F.W. (1934). The complete book of British butterflies. London: Ward, Lock & Co.

Gallien, L., Altermatt, F., Wiemers, M., Schweiger, O. & Zimmermann, N.E., 2017. Invasive plants threaten the least mobile butterflies in Switzerland. Diversity and Distributions, 23, 185-195. doi; 10.1111/ddi.12513

Gamer, M. Lemon, J., Fellows, I & Singh, P. (2012). irr: Various coefficients of interrater reliability and agreement. R package version 0.84. Available at https://cran.r-project.org/web/packages/irr/

Grime, J. P. (1977). Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. American Naturalist, 111, 1169-1194. doi: 10.1086/283244

Hill, J.K., Thomas, C.D. & Lewis, O.T. (1999). Flight morphology in fragmented populations of a rare British butterfly, *Hesperia comma*. Biological Conservation, 87, 277-283. doi: 10.1016/S0006-3207(98)00091-3

IPCC (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)). Geneva: IPCC.

Kichenin, E. Wardle, D.A. Peltzer, D.A., Morse, C.W. & Freschet, G.T. (2013). Contrasting effects of plant inter‐and intraspecific variation on community‐level trait measures along an environmental gradient. Functional Ecology, 27, 1254-1261.

Layberry, R.A., Hall, P.W. & Lafontaine, J.D. (1998). The butterflies of Canada. Toronto: University of Toronto Press.

Luck, G.W., Carter, A. & Smallbone, L. (2013). Changes in bird functional diversity across multiple land uses: interpretations of functional redundancy depend on functional group identity. PLOS ONE. 8: e63671. doi: 0.1371/journal.pone.006367

Karsholt, K. & van Nieukerken, E. (2016.) Fauna Europaea: Papilionoidea Fauna Europaea version 2.6 Available at: http://www.faunaeur.org

Kotiaho, J.S., Kaitala, V., Komonin, A. & Päivinen, J. (2005). Predicting the risk of extinction from shared ecological characteristics. Proceedings of the National Academy of Sciences, 102, 1963-1967. doi: 10.1073/pnas.0406718102

Mattila, N., Kotiaho, J.S., Kaitala, V., Komonen, A. & Päivinen, J. (2009). Interactions between ecological traits and host plant type explain distribution change in noctuid moths. Conservation Biology, 23, 703-709. doi: 10.1111/j.1523-1739.2008.01138.x

Middleton-Welling, J., Wade, R.A., Dennis, R.L.H., Dapporto, L., Shreeve T.G. (2018). Data from: Optimising trait and source selection for explaining occurrence and abundance changes: a case study using British Butterflies. Dryad Digital Repository. http://doi.org/10.5061/dryad.017vp2q

Moretti, M., Dias, A.T.C., de Bello, F., Altermatt, F., Chown, S. L., Azcárate, F.M., … Berg, M. P. (2016). Handbook of protocols for standardized measurement of terrestrial invertebrate functional traits. Functional Ecology, 31, 558-576. doi: 10.1111/1365-2435.12776

Newland, D., Still, R., Swash, A. & Tomlinson, D. (2015). Britain's butterflies: a field guide to the butterflies of Britain and Ireland. New Jersey: Princeton University Press.

Nilsson, S.G., Franzén, M., & Jönsson, E. (2008). Long term land-use changes and extinction of specialised butterflies. Insect Conservation and Diversity, 1, 197-207. doi: 10.1111/j.1752-4598.2008.00027.x

Öckinger, E., Schweiger, O., Crist, T.O., Debinski, D. M., Krauss, J., Kuussaari, M. …Bommarco, R. (2010). Life-history traits predict species responses to habitat area and isolation: a cross-continental synthesis. Ecology Letters, 13, 969-979. doi: 0.1111/j.1461-0248.2010.01487.x

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, McGlinn, P.D., … Wagner, H. (2016). Vegan: an R package for community ecologists. Available from: https://github.com/vegandevs/vegan

Oliver, T., Hill, J.K., Thomas, C.D. & Roy, D.B. (2009). Changes in habitat specificity of species at their climatic range boundaries. Ecology Letters, 12, 1091-1102. doi: 10.1111/j.1461-0248.2009.01367.x

Pateman, R. M., Hill, J. K., Roy, D.B., Fox, R. & Thomas, C.D. (2012). Temperature-dependent alterations in host use drive rapid range expansion in a butterfly. Science, 336, 1028-1030. doi: 10.1126/science.1216980

Pavlikova, A. & Konvicka, M. (2012). An ecological classification of Central European macromoths: habitat associations and conservation status returned from life history attributes. Journal of Insect Conservation, 16, 187-206. doi: 10.1007/s10841-011-9405-8

Powney, G. D., Giovanni, R., Preston, C.D., Purvis, A. & Roy, D.B. (2014). A phylogenetically-informed trait-based analysis of range change in the vascular plant flora of Britain. Biodiversity and Conservation. 23, 171-185. doi: 10.1007/s10531-013-0590-5

Pöyry, J., Luoto, M., Heikkinen, R.K., Kuusaari, M. & Saarinen, K. (2009). Species traits explain recent range shifts of Finnish butterflies. Global Change Biology, 15, 732-743.doi: 10.1111/j.1365-2486.2008.01789.x

Pyšek, P., Manceur, A.M., Alba, C., McGregor, K. F., Pergl, J., Štajerová, K, … Kühn, I. (2015). Naturalization of central European plants in North America: species traits, habitats, propagule pressure, residence time. Ecology, 96, 762-774. doi: 10.1890/14-1005.1

R Core Development Team (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available from: https://www.R-project.org/.

Schweiger, O., Harpke, A., Wiemers, M. & Settele, J. 2014. CLIMBER: Climatic niche characteristics of the butterflies in Europe. ZooKeys, 367, 65-84. doi: 10.3897/zookeys.367.6185

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. doi: 10.1111/j.1365-2656.2011.01909.x

Shreeve, T. (1995). Butterfly mobility. In: Pullin, A. S. (ed). Ecology and Conservation of Butterflies. Dordrecht: Springer. pp.37-45

Shreeve, T.G. & Dennis, R.L.H. 2011. Landscape scale conservation: resources, behaviour, the matrix and opportunities. Journal of Insect Conservation, 15, 175-188. doi: 10.1007/s10841-010-9336-9

Shreeve, T.G., Dennis, R.L.H., Roy, D.B. & Moss, D. (2001). An ecological classification of British butterflies: ecological attributes and biotope occupancy. Journal of Insect Conservation, 5, 145-161. doi: 10.1023/A:1017556113534

Siefert, A., Violle, C., Chalmandrier, L., Albert, C. H., Taudiere, A., Fajardo, A., …Wardle, D. A. (2015). A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. Ecology Letters, 18, 1406–1419. doi:10.1111/ele.12508

South, R. (1906). The Butterflies of the British Isles. London: Frederick Warne and Co.

Thomas, J.A. (2010). Butterflies of Britain and Ireland- 2nd Edition. British Wildlife Publishing.

Thomas, J.A. (2005). Monitoring change in the abundance and distribution of insects using butterflies and other indicator groups. Philosophical Transactions of the Royal Society Biological Science, 360, 339-357. doi:10.1098/rstb.2004.1585

Vanreusel, W. & Van Dyck, H. (2007). When functional habitat does not match vegetation types: a resource-based approach to map butterfly habitat. Biological Conservation, 135, 202-211.

Vera, F. W. M. (2000). Grazing, Ecology and Forest History. Wallingford: CABI publishing.

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. (2007) Let the concept of trait be functional! Oikos, 116, 882-892. doi: 10.1111/j.0030-1299.2007.15559.x

WallisDeVries, M.F., (2014) Linking species assemblages to environmental change: Moving beyond the specialist-generalist dichotomy. Basic and Applied Ecology, 15, 279-287. doi:10.1016/j.baae.2014.05.001

Warren, M.S., Hill, J.K., Thomas, J.A., Asher, J., Fox, R., Huntley, B., … Thomas, C. D. (2001). Rapid responses of British butterflies to opposing forces of climate and habitat change. Nature, 414, 65-69. doi: 10.1038/35102054

Wiklund, C. (1981). Generalist vs. specialist oviposition behaviour in *Papilio machaon* (Lepidoptera) and functional aspects on the hierarchy of oviposition preferences. Oikos, 36, 163-170. doi: 10.2307/3544441

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.

Table S1. Eigenvalues and explained variance of PCAs using species x trait values from six trait sources A) with and B) without biotope information

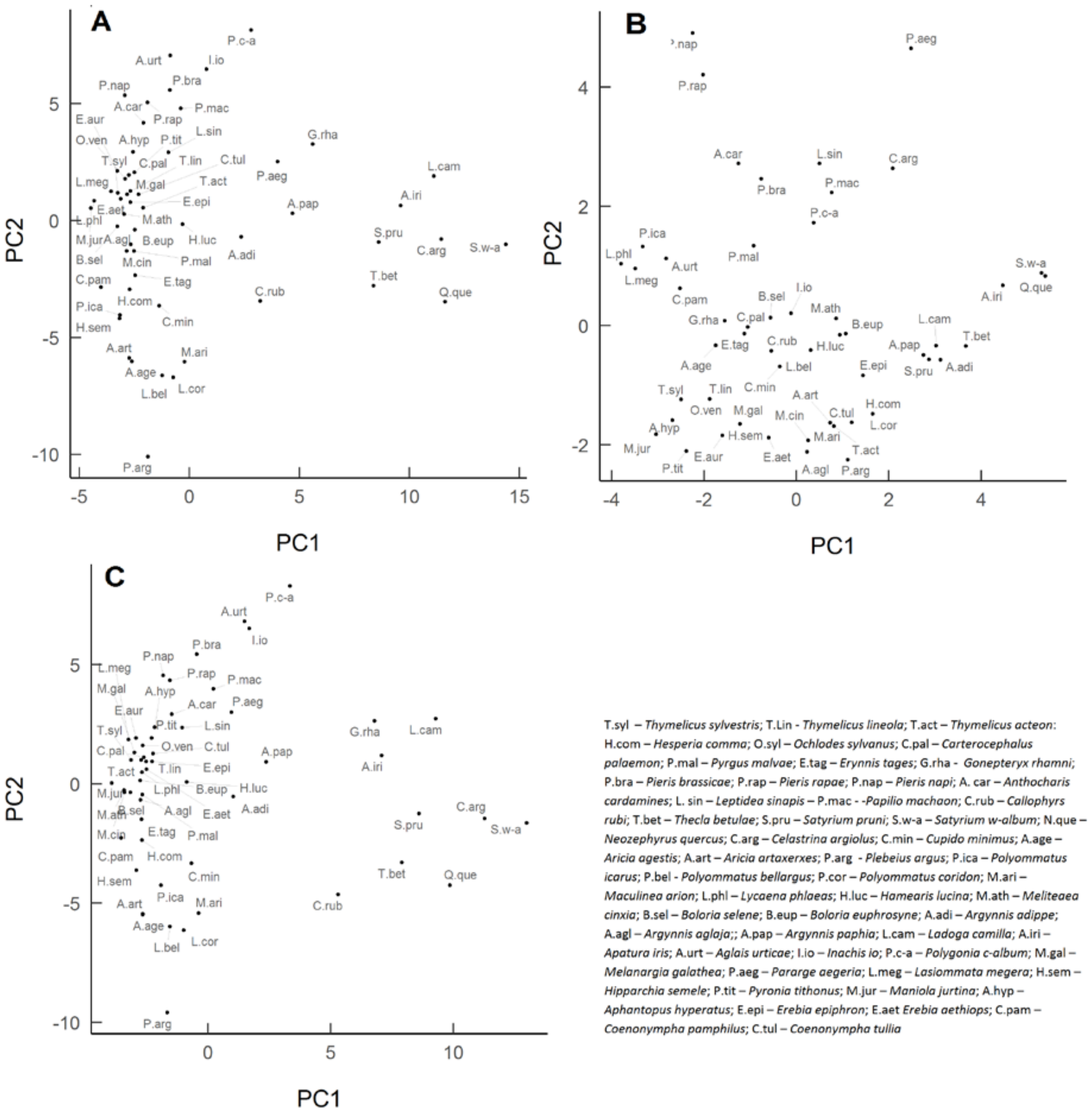
Table S2 Beta coefficients of PCA loadings from six trait sources, A) with and B) without biotope information

Table S3 Summary of the results of the best generalised linear models relating principal components to changes of abundance and distribution from six data sources.

Table S4 The major traits contributing to the PCA ordination of UK butterflies, using different combinations of life-history, resource-use and biotope occupancy, from data in Dennis (2010).

Figure Legend

Figure 1 PCA plots produced from data in Dennis (2010) comparing different trait sets A) with life-history , biotope and resource traits, B) with life-history and biotope and C) with life-history and resource traits.



T.syl – *Thymelicus sylvestris*; T.Lin - *Thymelicus lineola*; T.act – *Thymelicus acteon*: H.com – *Hesperia comma*; O.syl – *Ochlodes sylvanus*; C.pal – *Carterocephalus palaemon*; P.mal – *Pyrgus malvae*; E.tag – *Erynnis tages*; G.rha - *Gonepteryx rhamni*; P.bra – *Pieris brassicae*; P.rap – *Pieris rapae*; P.nap – *Pieris napi*; A. car – *Anthocharis cardamines*; L. sin – *Leptidea sinapis* – P.mac - -*Papilio machaon*; C.rub – *Callophyrs rubi*; T.bet – *Thecla betulae*; S.pru – *Satyrium pruni*; S.w-a – *Satyrium w-album*; N.que – *Neozephyrus quercus*; C.arg – *Celastrina argiolus*; C.min – *Cupido minimus*; A.age – *Aricia agestis*; A.art – *Aricia artaxerxes*; P.arg - *Plebeius argus*; P.ica – *Polyommatus icarus*; P.bel - *Polyommatus bellargus*; P.cor – *Polyommatus coridon*; M.ari – *Maculinea arion*; L.phl – *Lycaena phlaeas*; H.luc – *Hamearis lucina*; M.ath – *Meliteaea cinxia*; B.sel – *Boloria selene*; B.eup – *Boloria euphrosyne*; A.adi – *Argynnis adippe*; A.agl – *Argynnis aglaja*;; A.pap – *Argynnis paphia*; L.cam – *Ladoga camilla*; A.iri – *Apatura iris*; A.urt – *Aglais urticae*; I.io – *Inachis io*; P.c-a – *Polygonia c-album*; M.gal – *Melanargia galathea*; P.aeg – *Pararge aegeria*; L.meg – *Lasiommata megera*; H.sem – *Hipparchia semele*; P.tit – *Pyronia tithonus*; M.jur – *Maniola jurtina*; A.hyp – *Aphantopus hyperatus*; E.epi – *Erebia epiphron*; E.aet *Erebia aethiops*; C.pam – *Coenonympha pamphilus*; C.tul – *Coenonympha tullia*

Table 1. Traits used in analysis of congruence between UK butterfly trait source data and in the ordination of species.

Trait type[1]	Data type	Source availability[2]	Number of species compared	Possible trait states
Myrmecophily	Categorical	All	16	0/1
Biotope use (Nine variables)	Categorical	All	55	0/1 for each biotope
Hilltopping	Categorical	All	25	0/1
Mate choice strategy	Categorical	All	25	1/2/3
Mudpuddling	Categorical	All	33	0/1
Overwintering stage	Categorical	All	53	1/2/3/4
Patrolling	Categorical	All	25	0/1
Perching	Categorical	All	25	0/1
Hostplant number	Continuous	All	55	Continuous
Wingspan (average)	Continuous	1,4,5	55	Continuous
Wingspan (max)	Continuous	1,4,5	55	Continuous
Wingspan (min)	Continuous	1,4,5	55	Continuous
Flight period start month	Ordinal	1-5	53	1-12
Flight period end month	Ordinal	1-5	53	1-12
Max. no. of generations	Ordinal	All	55	1-3
Min. no. of generations	Ordinal	All	55	0.5-3
Number of biotopes occupied	Ordinal	All	55	1-8
Hostplant specificity	Ordinal	All	55	1-5

1.Trait type: *Myrmecophily*: larval and/or pupal association with ants. *Hilltopping*: assembling at hilltops as part of mating strategy. *Mate searching strategy*: 1 - perching, 2 - patrolling and 3 - perching and patrolling. For PCA analysis only, mate location coded as two binary traits: *Perching* and *Patrolling*. *Mudpuddling:* adults feeding from puddles/ damp patches. *Overwintering stage*: 1 – egg, 2 – larva, 3 – pupa, 4 - adult (with multiple overwintering stages the earliest stage was scored). *Hostplant number*: maximum number of plant species used. *Wingspan*: wing length from base to tip (mm), coded as mean, maximum and minimum. *Flight period*: start and end month number. *Number of generations*: maximum and minimum number of generations per year. Partial generations were rounded up. *Number of biotopes occupied*: sum of biotopes occupied. *Hostplant specificity*: 1 - single hostplant, 2 - multiple hostplants within the same genus, 3 - multiple hostplants within the same family, 4 - multiple hostplants within the same order, 5 - multiple hostplants in multiple orders.

2. Trait source: 1 - Emmet and Heath (1989), 2 - Asher et al. (2001), 3 - Thomas (2010), 4 - Newland et al. (2015), 5 - Eeles (2016), 6 - Dennis (2010).

Table 2. Individual trait agreement values between six trait sources for UK butterflies. Ordinal and categorical traits are compared with Fleiss's Kappa and continuous traits with a two-way intra-class correlation coefficient (ICC). For trait definitions see Table 1.

Categorical Traits	**Kappa**	**Trait sources**
Overwintering	0.973	All
Perching	0.778	All
Patrolling	0.745	All
Mate choice strategy	0.678	All
Biotope G	0.654	All
Biotope E	0.637	All
Biotope B	0.524	All
Biotope I	0.501	All
Biotope D	0.443	All
Biotope C	0.428	All
Myrmecophily	0.422	All
Biotope F	0.394	All
Biotope J	0.288	All
Mudpuddling	0.273	All
Biotope H	0.207	All
Hilltopping	0.173	All

Ordinal Traits	**Kappa**	**Trait sources**
Generations (max)	0.812	All
Hostplant specificity	0.66	All
Flight Period (start)	0.512	1-5
Generations (min)	0.475	All
Flight period (end)	0.457	1-5
Biotope total	0.236	All

Continuous traits	**ICC**	**Trait sources**
Wingspan (av.)	0.998	1,4,5
Wingspan (max)	0.995	1,4,5
Wingspan (min)	0.993	1,4,5
Hostplant number	0.355	All

Trait source: 1 - Emmet and Heath (1989), 2 - Asher et al. (2001), 3 - Thomas (2010), 4 - Newland et al. (2015), 5 - Eeles (2016), 6 - Dennis (2010).

Table 3. Pairwise Procrustes regression values between PCA ordinations of UK butterflies constructed on the basis of traits obtained from six data sources and the average ordination of 50 randomly assembled species x trait matrices (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Upper right, with biotope information included; lower left, without biotope information.

	Emmet & Heath (1989)	**Asher et al. (2001)**	**Thomas (2010)**	**Newland et al. (2015)**	**Eeles (2016)**	**Dennis (2010)**	**Random**
Emmet & Heath (1989)		0.657***	0.601***	0.676***	0.662***	0.641***	0.256
Asher et al.(2001)	0.581***		0.646***	0.673***	0.720***	0.692***	0.288
Thomas (2010)	0.666***	0.683***		0.692***	0.729***	0.665***	0.300
Newland et al. (2015)	0.558***	0.628***	0.667***		0.683***	0.642***	0.310
Eeles (2016)	0.685***	0.733***	0.764***	0.643***		0.688***	0.290
Dennis (2010)	0.638***	0.647***	0.673	0.628***	0.713***		0.286
Random	0.336	0.307	0.373	0.326	0.339	0.339	

Table 4. Summary of the results of the best generalised linear models from data in Dennis (2010) comparing how various trait types correlate with criteria of status change in Fox et al. (2015). Components shown are those that are significantly correlated with measures of status (* p < 0.05, ** p <0.01, *** p <0.001).

		Trait combination		
Status	**Life-history**	**Life-history and biotope**	**Life-history and resources**	**All**
% change Occurrence (all)	5*,7*	3*,4**	none	none
% change Occurrence (recent)	None	none	none	none
% change Abundance (all)	None	none	none	none
% change Abundance (recent)	7**	1*,	1*	1**
UKBMS Sites occupied	1**	1***,2***,3*,6*	2**	2**
10km grid squares occupied	1***,5*	1***,2***,3**,4*,5**,6**,7**	2**,3**	2***,3*
1km grid squares occupied	1***	1***,2***,3**,4**,5*,6**	2***,3**	2***,3*

Functional Ecology

Supplementary material: Optimising trait and source selection for explaining occurrence and abundance changes: a case study using British Butterflies

Middleton-Welling, J.[1], Wade, R.A.[1], Dennis, R.L.H.[1,2], Dapporto, L.[3], Shreeve T.G.[†1]

1. Centre for Ecology, Environment and Conservation, Faculty of Health and Life Sciences, Oxford Brookes University, Headington, Oxford OX3 0BP, UK

2.

School of Life Sciences and Education, Staffordshire University, Science Centre, Leek Road, Stoke-on-Trent ST4 2DF, U.K.

3. Dipartimento di Biologia, Via Madonna del Piano, 6, 50019 Sesto Fiorentino, Florence, Italy

.

Middleton-Welling, J., Wade, R.A., Dennis, R.L.H., Dapporto, L., Shreeve T.G. Optimising trait and source selection for explaining occurrence and abundance changes: a case study using British butterflies.

Table S1. Eigenvalues and explained variance of PCA using species x trait values from six trait sources for UK butterflies A) with biotope type and breadth included and B) with biotope type and breadth excluded

	Emmet & Heath (1989)		**Asher et al. (2001)**		**Thomas (2010)**		**Newland et al. (2015)**		**Eeles (2016)**		**Dennis (2010)**	
Component	**Eigen value**	**%variation explained**	**Eigen value**	**%variation explained**	**Eigen value**	**%variation explained**	**Eigen value**	**%variation explained**	**Eigen value**	**%variation explained**	**Eigen value**	**%variation explained**
A) with biotope												
1	5.07	28.17	5.08	28.21	3.67	20.38	4.59	25.5	3.84	21.31	4.99	27.75
2	2.40	13.31	2.32	12.90	2.53	14.08	2.35	13.03	2.59	14.37	2.15	11.93
3	1.63	9.05	1.80	9.97	2.30	12.77	2.10	11.67	1.83	10.15	1.87	10.37
4	1.43	7.95	1.54	8.53	1.88	10.44	1.66	9.23	1.47	8.19	1.47	8.18
5	1.31	7.27	1.20	6.67	1.30	7.25	1.27	7.06	1.30	7.24	1.28	7.08
6	1.19	6.60	1.01	6.05	1.03	5.73	1.204	6.69	1.24	6.89	1.06	5.90
7	1.07	5.98							1.02	5.64		
B) without biotope												
1	2.15	26.86	2.39	29.82	2.32	29.05	2.57	32.10	2.41	30.06	2.21	27.66
2	1.37	17.10	1.60	20.00	1.58	19.71	1.33	16.62	1.37	17.09	1.51	18.86
3	1.10	13.76	1.27	15.90	1.14	14.21	1.22	15.30	1.22	15.27	1.26	15.78
4	1.03	12.92	1.00	12.53					1.01	12.68		

Table S2. Beta coefficients of PCA loadings from six trait sources 1-Emmet and Heath (1989), 2- Asher et al. (2001), 3 - Thomas (2010), 4 -Newland et al. (2015), 5 -Eeles (2016), 6 - Dennis (2010) regressed against measures of occurrence and changes of occurrence and abundance (from Fox et al.,2015) with biotope information included (A) and without biotope information (B).

A With biotope information included

Emmet and Heath (1989)

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
% change occurrence (1976-2014)	-0.221	-0.048	-0.099	-0.010	0.200	-0.058	0.099
% change occurrence (2005-14))	0.100	-0.071	-0.156	0.127	0.076	0.055	0.263
% change abundance (1976-2014)	0.268	0.104	0.149	0.041	0.057	0.208	0.123
% change abundance (2005-14)	-0.330*	0.280*	0.080	0.119	-0.156	0.034	0.026
UKBMS Sites	-0.588***	0.067	-0.032	0.093	0.146	-0.171	-0.166
10km grid squares	-0.676***	0.009	-0.073	0.069	0.183	-0.174	-0.094
1km grid squares	-0.680***	0.069	-0.017	0.045	0.238	-0.164	-0.059

Asher (2001)

	PC1	PC2	PC3	PC4	PC5	PC6
% change occurrence (1976-2014)	-0.186	-0.265	0.003	0.098	-0.234	-0.133
% change occurrence (2005-14))	0.120	-0.099	-0.052	0.024	0.056	0.047
% change abundance (1976-2014)	0.279*	-0.151	0.323*	-0.050	-0.029	-0.070
% change abundance (2005-14)	-0.250	0.047	0.140	-0.190	0.278	0.332*
UKBMS Sites	-0.611***	-0.055	-0.090	-0.003	-0.043	0.015
10km grid squares	-0.709***	-0.176	-0.085	0.066	-0.106	0.044
1km grid squares	-0.720***	-0.269	-0.002	0.046	-0.012	0.097

Table S2A continued

Thomas (2010)

	PC1	PC2	PC3	PC4	PC5	PC6
% change occurrence (1976-2014)	0.163	-0.242	0.191	-0.059	-0.037	-0.028
% change occurrence (2005-14))	-0.091	-0.071	-0.111	-0.204	0.221	0.142
% change abundance (1976-2014)	0.326*	0.167	-0.382	-0.005	0.040	0.063
% change abundance (2005-14)	0.295*	0.271*	-0.167	-0.040	-0.70	-0.236
UKBMS Sites	0.464***	-0.128	0.201	-0.073	-0.012	-0.226
10km grid squares	0.578***	-0.102	0.316*	-0.098	-0.091	-0.281
1km grid squares	0.505***	-0.031	0.408**	0.668	-0.134	-0.202

Newland et al. (2015)

	PC1	PC2	PC3	PC4	PC5	PC6
% change occurrence (1976-2014)	-0.177	0.264	0.151	-0.113	-0.038	-0.093
% change occurrence (2005-14))	0.097	0.003	-0.017	-0.174	-0.174	0.025
% change abundance (1976-2014)	0.298*	0.001	-0.151	0.132	-0.065	-0.069
% change abundance (2005-14)	-0.323*	-0.023	-0.249	0.060	-0.224	-0.248
UKBMS Sites	-0.633***	0.144	-0.056	0.006	-0.034	-0.195
10km grid squares	-0.729***	0.282*	0.100	-0.119	0.040	-0.212
1km grid squares	-0.716***	0.353**	-0.018	-0.071	0.055	-0.143

Table S2A continued

Eeles (2016)

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
% change occurrence (1976-2014)	0.221	-0.048	-0.099	0.010	0.200	-0.057	0.099
% change occurrence (2005-14))	0.102	-0.043	0.029	0.085	-0.167	-0.167	0.233
% change abundance (1976-2014)	0.252	0.024	0.080	0.119	-0.156	0.227	0.192
% change abundance (2005-14)	-0.330*	-0.280*	-0.906	1.745	-0.075	0.034	0.026
UKBMS Sites	-0.656***	0.174	0.020	0.092	-0.007	-0.131	-0.186
10km grid squares	-0.772***	0.240	0.020	0.056	0.022	-0.233	-0.523
1km grid squares	-0.725***	0.269*	0.196	0.140	-0.018	-0.156	0.007

Dennis (2010)

	PC1	PC2	PC3	PC4	PC5	PC6
% change occurrence (1976-2014)	-0.294*	-0.250	-0.067	0.044	0.031	-0.088
% change occurrence (2005-14))	0.109	-0.080	0.001	-0.010	0.030	0.117
% change abundance (1976-2014)	0.196	-0.157	0.030	0.144	0.122	0.091
% change abundance (2005-14)	-0.071	0.186	0.076	-0.252	0.213	0.102
UKBMS Sites	-0.532***	0.006	0.008	-0.057	0.042	0.142
10km grid squares	-0.655***	-0.063	0.049	-0.088	-0.013	0.047
1km grid squares	-0.643***	-0.119	0.078	0.1102	0.137	-0.039

Table S2 continued

B) without biotope

Emmet and Heath (1989)

	PC1	PC2	PC3	PC4
% change occurrence (1976-2014)	0.271*	-0.055	0.208	-0.061
% change occurrence (2005-14))	-0.069	-0.193	0.204	-0.139
% change abundance (1976-2014)	-0.210	-0.148	-0.065	-0.015
% change abundance (2005-14)	0.046	-0.218	-0.067	-0.177
UKBMS Sites	0.442***	-0.064	-0.022	0.159
10km grid squares	0.541***	-0.085	0.074	0.114
1km grid squares	0.553***	-0.098	0.087	0.170

Asher et al. (2001)

	PC1	PC2	PC3	PC4
% change occurrence (1976-2014)	-0.294*	-0.250	-0.067	0.045
% change occurrence (2005-14))	0.109	-0.080	0.001	-0.010
% change abundance (1976-2014)	0.196	-0.157	-0.030	0.144
% change abundance (2005-14)	-0.071	0.187	0.076	-0.252
UKBMS Sites	-0.532***	0.006	0.008	-0.057
10km grid squares	-0.655***	-0.063	0.049	-0.088
1km grid squares	-0.643***	-0.119	0.078	-0.110

Table S2B continued

Thomas (2010)

	PC1	PC2	PC3
% change occurrence (1976-2014)	-0.291	0.025	0.036
% change occurrence (2005-14))	0.093	0.299*	-0.057
% change abundance (1976-2014)	0.252	0.017	-0.009
% change abundance (2005-14)	0.020	0.050	-0.97
UKBMS Sites	-0.406**	0.010	-0.170
10km grid squares	-0.522***	-0.031	-0.158
1km grid squares	-0.489***	-0.078	-0.173

Newland et al. (2015)

	PC1	PC2	PC3
% change occurrence (1976-2014)	-0.124	0.184	-0.007
% change occurrence (2005-14))	0.094	0.229	0.123
% change abundance (1976-2014)	0.240	0.007	-0.104
% change abundance (2005-14)	-0.077	0.193	-0.081
UKBMS Sites	-0.393**	0.124	-0.082
10km grid squares	-0.464***	0.193	0.055
1km grid squares	-0.431**	0.255	0.044

Table S2B continued

Eeles (2016)

	PC1	PC2	PC3	PC4
% change occurrence (1976-2014)	-0.337*	0.056	-0.061	0.109
% change occurrence (2005-14))	0.107	0.114	-0.167	0.213
% change abundance (1976-2014)	0.165	-0.043	0.037	-0.016
% change abundance (2005-14)	-0.074	0.077	-0.018	-0.262
UKBMS Sites	-0.539***	-0.010	-0.065	-0.110
10km grid squares	-0.659***	0.022	-0.001	-0.095
1km grid squares	-0.640***	0.078	-0.034	-0.144

Dennis (2010)

	PC1	PC2	PC3
% change occurrence (1976-2014)	0.232	0.261	0.192
% change occurrence (2005-14))	-0.181	0.126	0.031
% change abundance (1976-2014)	-0.054	-0.277*	0.016
% change abundance (2005-14)	0.102	0.081	-0.175
UKBMS Sites	0.450***	0.259	-0.017
10km grid squares	0.497***	0.370	0.026
1km grid squares	0.512***	0.342	-0.012

P <0.05, ** P < 0.01, *** P < 0.001

Table S3: Summary of the results of the best linear models relating trait sources to criteria of status change in Fox et al. (2015)., A) with and B) without biotope information included. Models were initially produced using all the significant principle components and then simplified using AIC. Retained components are numbered and those that are significantly correlated with measures of status are indicated (* p < 0.05 , ** p < 0.01, *** p < 0.001). Trait sources- 1-Emmet and Heath (1989), 2- Asher et al. (2001), 3 - Thomas (2010), 4 -Newland et al. (2015), 5 -Eeles (2016), 6 - Dennis (2010)

A) with biotope included	Trait source					
	1	2	3	4	5	6
% change Occur (all)	1,5	1,2*,5	2,3,	1,2*,8	1*,2*,4,5,6,7,8*	1*,2
% change Occurrence (recent)	7	none	4,5	7	7	none
% change Abundance (all)	1*,6	1*,3*	1*	1*,7*	1,6,7	1
% change Abundance (recent)	1,2**,3*,4,5*,6,7	1*,4,5*,6**	1*,2*,3,6	1*,3,5,6*,8	1**,2*,7*	4,5
UKBMS Sites	1***,6,7,	1***	1***,3,6	1***,2,6,7*	1***,2,7	1***
10km grid squares	1***,5,6,	1***,2	1***,3**,6**	1***,2**,4,6*	1***,2**,6**,8	1***
1km grid squares	1***,5*,6,	1***,2,5*,6	1***,3***,6	1***,2***,6	1***,2**,3*,4,6	1***

Table S3 continued

B) with biotope included	Trait source					
	1	2	3	4	5	6
% change Occurrence (all)	1*,3	1*,2	1*,6**	2,6**	1*	1,2*,3,4*
% change Occur (recent)	2,3,6	none	2*	2,4	4	none
% change Abundance (all)	1	1	1	1	none	2*
% change Abundance (recent)	2,4,6,7**	2,4	4***	2,5*	4,5	5
UKBMS Sites	1***,7	1***	1**,3,6	1**,6*	1***	1***,2*,6*
10km grid squares	1***,7*	1***	1***,3,4*,6	1***,6*	1***	1***,2***,6***
1km grid squares	1***,3,7**	1***	1***,3,4**,6**	1***,6**	1***	1***,2**,5*,6*

Table S4. The first five factors contributing to the first three components of PCA analyses of data from 6 trait sources A) with biotope information and B) without information

Trait sources*:* 1 -Emmet and Heath (1989), 2- Asher et al. (2001), 3 - Thomas (2010), 4 -Newland et al. (2015), 5 -Eeles (2016), 6 - Dennis (2010).

Trait	Trait source																	
	1	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**	**1**	**2**	**3**	**4**	**5**	**6**
A) with biotope	PC1						PC2						PC3					
Biotope total	-0419	-0.414	0.473	-0.429	-0.457	-0.427												
Biotope .B				-0.300		-0.310		0.377										
Biotope C				-0.258		-0.312	0.545		0.477	0.347	-0.332					-0.355	0.348	
Biotope D							0.545		0.486	0.347						-0.355	0.335	
Biotope E								0.300					0.361		-0,373			0.489
Biotope F			0.309		-0.302							-0.296	-0.260	-0.504				
Biotope G									-0.234	0.339		-0.386	-0.616	-0.526				
Biotope H	-0.317	-0.315							0.361	-0.362	-0.285							
Biotope I	-0.346	-0.298			-0.268	-0.305									0.305			-0.292
Biotope J	-0.352	-0.273	0.321					0.309		-0.430							-0.404	
Overwintering			0.268	-0.296	-0.304										0.312		0.377	-0.290
Hostplant specialism			0.288	-0.284			-0.262		-0.252		0.353					0.297		
Hostplant number						-0.286	-0.283		-0.280		0.342		0.343				-0.332	
Generations (max)	-0.307	-0.317			-0.310		-0.302								0.478	0.334		-0.347
Generations (min)											0.376	0.2422			0.357	0.401		
Perch								-0.387				-0.476						
Patrol								-0.364				0.459		0.269				
Mudpuddling													-0.312	0.251				-0.520
B) without biotope																		
Overwintering	0.376	-0.391	-0.362	-0.386	-0.370	0.366	-0.401	-0.051		0.269		0.242	0.258					
Hostplant.Specialism	0.370	-0.465	-0.472	-0.473	-0.464	0.449				0.331	-0.413		0.346	-0.484	0.462	-0.162		0.490
Hostplant number	0.411	-0.446	-0.400	-0.405	-0.414	0.445	0.307	-0.129		0.419	-0.448			-0.479	0.577	-0.211		0.475
Generations (max)	0.538	-0.489	-0.454	-0.456	-0.496	0.401			-0.216	-0.479		0.356		0.477	-0.456		0.244	-0.419
Generations (min)	0.477	-0.429	-0.466	-0.425	-0.467				-0.182	-0.547	0.378		-0.319	0.532	-0.432		0.272	-0.459
Perch							0.487	-0.648	0.583		0.441	0.607	0.394			0.473	-0.458	0.322
Patrol						0.405	0.645	-0.529	-0.622			-0.435				-0.685	0.742	
Mudpuddling								0.523	0.438		0.339	0.497	0.710	-0.141		-0.429	0.199	

Middleton-Welling, J., Wade, R.A., Dennis, R.L.H., Dapporto, L., Shreeve T.G. Optimising trait and source selection for explaining occurrence and abundance changes: a case study using British butterflies.