Bishop, Tom R.; Botham, Marc S.; Fox, Richard; Leather, Simon R.; Chapman, Daniel S.; Oliver, Tom H. 2013. **The utility of distribution data in predicting phenology**. *Methods in Ecology and Evolution*, 4 (11). 1024-1032. 10.1111/2041-210X.12112

**Title:** The utility of distribution data in predicting phenology

**Authors:**

Tom R. Bishop, tbish@liverpool.ac.uk, School of Environmental Sciences, University of Liverpool, Liverpool, L69 3GP

Marc S. Botham, NERC Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB

Richard Fox, Butterfly Conservation, Manor Yard, East Lulworth, Dorset, BH20 5QP

Simon R. Leather, Department of Crop and Environment Sciences, Harper Adams University College, Edgmond, Newport, Shropshire, TF10 8NB

Daniel Chapman, NERC Centre for Ecology and Hydrology, Bush Estate, Penicuik, Edinburgh, EH26 0QB, UK

Tom H. Oliver, NERC Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB

## Abstract

1. The phenology of many species has been shown to shift under climate change. However, because species respond at different rates, ecological communities may be disrupted leading to species extinctions and loss of ecosystem services. Hence, there is a need to monitor and understand phenological change.
2. Population data, gathered by standardised monitoring schemes, can be used to this end. However, such schemes require significant organisation and financial resources. Distribution data (georeferenced biological records with dates) are easier and cheaper to collect and may be an unexploited resource for phenology analyses. This would allow analysis of more taxa from more regions of the world. However, distribution data are potentially biased due to the unstandardized behaviour of biological recorders.
3. Here, the ability of distribution data record dates to accurately predict phenology is investigated by using the British butterfly fauna as a model system. We used the total number of distribution records per unit time across Great Britain as a proxy for butterfly abundance. Phenology metrics of mean flight date and flight period length were then calculated from the resulting abundance-time relationships for each year in a 15-year time series. These estimates were validated against those generated from a standardised-effort population monitoring scheme.
4. We analysed 1,078,328 records from 30 British butterflies and found that distribution data accurately predicted the mean flight date for 22 out of the 30 species tested. Flight period length was only predicted accurately for seven out of thirty species.
5. We found a non-linear but consistent positive relationship between the accuracy of mean flight date estimates and sample size (number of records) at both inter- and intraspecific scales. Our results suggest that a threshold sample size of approximately 6,500 distribution records (430 per year) is a pragmatic compromise between accuracy and recording effort,

leading to little loss of accuracy in phenology predictions (an average decrease in accuracy of 2.9 days was observed).

6. The results suggest that distribution data are a potentially useful resource for phenology research. This may allow practitioners to monitor particular regions and previously unstudied species relatively cheaply using existing mapping schemes.

**Key Words:** Distribution data; biological records; phenology, UK Butterfly Monitoring Scheme, Butterflies for the New Millenium

## Introduction

During recent decades it has become clear that climate change is having a significant effect on the phenology of many species (Parmesan 2006; Rosenzweig *et al.* 2007; Hill, Griffiths & Thomas 2011). These changes occur in the direction predicted under global warming scenarios (Parmesan 2006) and are likely to disrupt existing ecological communities as individual species respond at different rates (Walther *et al.* 2002; Root *et al.* 2003; Thackeray *et al.* 2010). Ultimately, this may lead to widespread extirpation and extinction (Thomas *et al.* 2004; Tylianakis *et al.* 2008; Willis *et al.* 2008). The potential for synergism between global change drivers, coupled with long term projections, makes predicting and monitoring the effects of climate change on phenology a key issue for 21st century biologists (Balmford *et al.* 2005; Visser 2008; Miller-Rushing *et al.* 2010).

A major challenge concerning biodiversity monitoring schemes, many of which are designed to collect phenology data, is the considerable effort required on the part of professionals and volunteers to achieve adequate levels of temporal and spatial coverage that will allow large scale or long term trends to be revealed (Thomas 2005; Fox *et al.* 2006). This problem is particularly true of invertebrates, which are often neglected by conservation biologists and funding bodies (Clark & May 2002; Leather 2009), yet is also present in a range of other taxa and geographic regions. Additionally, it is not expected that different species will adjust their phenology in the same direction or in response to the same cues (Bale *et al.* 2002; Visser & Both 2005; Doi, Gordo & Katano 2008), and so the focus on a small number of charismatic taxa or well-funded regions inevitably ignores the true scope of phenological change. Furthermore, the consequences of shifting phenologies need to be understood in the context of concurrent change in other species and environmental variables (Visser & Both 2005). This point may be particularly salient given that different trophic levels and interaction partners are known to respond to climate change at different rates (Van Nouhuys & Lei 2004; Memmott *et al.* 2007; Both *et al.* 2009; Thackeray *et al.* 2010). These unequal phenological responses are likely to influence species demography and ecosystem processes in novel ways that

may not be fully understood without data on multiple nodes within the ecological web. Consequently, we have an incomplete picture of the global phenological response to climate change.

In order to address these gaps in our knowledge, distribution data may prove useful. Distribution, or 'atlas', data are spatially and temporally explicit information on a species occurrence and are commonly used to create regional distribution atlases of specific taxa (Robertson, Cumming & Erasmus 2010). Distributions are mapped from the presence of at least one recorded occurrence within a specified grid cell (Araujo *et al.* 2005). These records may be obtained through the extraction of museum specimen data (Funk & Richardson 2002), the use of historical records (Hassall *et al.* 2007), the submission of casual species observations or through nationally coordinated surveys (Harding & Sheail 1992; Fox *et al.* 2006). Crucially, such data are a record only of species' presence and, thus, are different from more detailed presence-absence distribution data obtained from intensive standardised surveys.

Whilst they contain less information than detailed population monitoring data, distribution data are available for a greater range of taxa and geographic regions and, often for longer time periods (Thomas 2005; Robertson, Cumming & Erasmus 2010; www.gbif.org). Additionally, it may be logistically easier to collect meaningful volumes of this data type than adequately standardised and replicated population estimates. As distribution records have dates attached they can be analysed in a temporal context. In theory, one might interpret the number of distribution records available for a species throughout a time-series in an analogous fashion to population abundance data. Both data types may produce an abundance-time distribution of a species within a year from which phenology metrics may be drawn.

There has been some interest in the potential of distribution data to reveal phenological patterns (Hassall *et al.* 2007; Carroll *et al.* 2009; Altermatt 2010; Poyry *et al.* 2011), but no rigorous test of its utility in such a role. Validation tests are crucial as distribution data are likely to be highly biased in space and time. In space, data may be influenced by recorder effort (Dennis, Sparks & Hardy 1999),

the visual apparency of target species (Dennis *et al.* 2006) and the expected species richness of a site (Dennis & Thomas 2000).

There are also a number of biases specific to the application of distribution data to phenological research. These will not be apparent when using distribution data for its original purpose and are related to the behaviour of biological recorders. Within years there is often a drive amongst recorders to collect the first record of a species within a year (http://www.butterfly-conservation.org/text/853/first_sightings_2012.html). Recorders may also lose interest in a particular species as a season progresses and may have renewed interest in unusual late events. These effects may bias phenology estimates derived from distribution data.

The magnitude and influence of these biases are unknown due to the lack of data on actual recorder effort and the implementation of standardised collecting protocols. A key strength of distribution data however, is that their collection is not hampered by adhering to rigorous controls and so spatial and temporal coverage can be much greater than for standardised surveys. In the UK, for example, the standardised monitoring scheme for butterflies covers around 1000 active sites. Butterfly distribution data, on the other hand, covers 3834 unique 10km UK grid squares. The greatest potential strength of distribution data for phenology research, however, comes from its taxonomic scope. Distribution data are available for a much wider range of taxa than the Lepidoptera, birds and bats, which comprise the major population monitoring schemes. Despite the expected shortcomings in the application of distribution data to temporal research, the question remains over whether any phenological signal is strong enough to penetrate potential biases and produce reliable estimates. The British butterfly fauna provides an ideal system within which to answer this question. The UK has a spatially and temporally extensive butterfly distribution dataset generated by the Butterflies for the New Millennium (BNM) project (Fox *et al.* 2006). The aim of the BNM is to map the national distribution of species and, thus, to assess changes over time. The BNM was launched in 1995 and has run continuously with three major drives of record collection activity occurring during 1995 –

1999, 2000 – 2004 and 2005 – 2009. Over 7.5 million records have been collated. The scheme is operated through a network of volunteers and local co-ordinators who feed data to Butterfly Conservation. There is also a detailed and pioneering transect monitoring programme for butterflies: the UK Butterfly Monitoring Scheme (UKBMS) (Fox *et al.* 2006; Brereton *et al.* 2011). The UKBMS is a standardised scheme which has been in operation in some form since 1976. Volunteers undertake weekly transect walks which generate abundance measures for species on over 1000 sites across the UK (Brereton *et al.* 2011). Both these datasets have also played a large role in investigating the influence of climate change on butterflies (Parmesan *et al.* 1999; Roy & Sparks 2000; Menendez *et al.* 2006; Pateman *et al.* 2012).

We assume that phenology estimates drawn from UKBMS data will give an accurate baseline against which BNM estimates may be compared. The UKBMS is designed to detect a range of population indices, including phenology metrics, and is temporally standardised. It must be remembered however, that the UKBMS itself is not infallible. Issues concerning the visual apparency of species may apply to both the UKBMS and the BNM datasets (Dennis *et al.* 2006). Indeed, the UKBMS fails to routinely produce population trends for a number of rare or visually unapparent species (Fox *et al.* 2006). A further limitation of the UKBMS is that it is spatially and temporally restricted. Although there are c. 1000 active sites, these may not capture all of the warmest microclimates across landscapes and, therefore, very early and late individuals may be missed. This is also exacerbated by the fact that transect monitoring only starts in April and runs until the end of September. Consequently, an increasing proportion of species flight periods may occur outside of the monitoring period as phenology shifts with the warming climate. These caveats must be kept in mind when commenting on the relative accuracy of BNM phenology estimates.

In this study, two standard butterfly phenology metrics, mean flight date and flight period length (Stefanescu, Penuelas & Filella 2003), are calculated on both datasets for 30 univoltine species over a 15 year time period. We then compare the ability of BNM distribution data to predict phenology

estimates from the standardised UKBMS recording scheme. The influence of distribution record sample size on the relative accuracy of phenology estimates is also investigated and the potential application of distribution records in phenology monitoring is discussed. Due to the biases that may be present in distribution data it is expected that flight period length estimates will not be predicted well by the BNM. This metric is more likely to be sensitive to non-uniform recording effort throughout a year. Mean flight date estimates are hypothesised to be more robust to these biases and so are expected to be predicted well by the BNM.

## Methods

### DATA COLLECTION AND PREPARATION

BNM and UKBMS data were supplied by the Centre for Ecology and Hydrology and Butterfly Conservation for the years 1995 – 2009 and the 30 univoltine species given in the appendix. Analyses were restricted to univoltine species due to the problems involved in calculating the phenology metrics for multivoltine species (Botham *et al.* 2008). Multivoltine species have two or more generations per year which may overlap. This can make the chosen phenology metrics meaningless as generations cannot always be objectively separated. All UKBMS records present in the BNM were removed. For both datasets the number of days since April 1$^{st}$ was calculated for each record. The UKBMS only monitors butterfly populations between April 1$^{st}$ and September 30$^{th}$ each year (Fox *et al.* 2006) and so the BNM was also restricted to this time frame to ensure fair comparison. This filtered and restricted BNM dataset consisted of 1,078,328 records (from 30 species over 15 years). Data were then aggregated to give an abundance (UKBMS) or record count (BNM) per day for every year, species and dataset. The phenology metrics of mean flight date and flight period length were calculated for each species in each year. These correspond to the weighted mean:

$\bar{x} = \sum wx / \sum w$

and standard deviation:

$$sd = \sqrt{\sum(w(x - \bar{x})2)( \sum w/\sum(w)2 - \sum w2)}$$

respectively (Stefanescu, Penuelas & Filella 2003), where $x$ is the number of days since April 1st and $w$ is the total abundance per day recorded by the UKBMS or the number of BNM records. These metrics are used as the Gaussian phenology curve is specified by the mean and standard deviation. Both metrics are also commonly used in the study of phenology (Brakefield 1987; Roy & Sparks 2000; Stefanescu, Penuelas & Filella 2003).

We expect no systematic bias between the phenology estimates of the two datasets. Both schemes have comparable latitudinal distributions and there does not appear to be a pattern in the degree of accuracy of estimates through time (see appendix).

TESTING BNM PREDICTIONS

Observed estimates (UKBMS) were compared to predicted estimates (BNM) of each phenology metric and every species using type II major axis regression. Major axis regression is a more appropriate method than ordinary least squares regression when there is error present on both the x and y variables and the aim is to compare observed to predicted values (Legendre & Legendre 1998). Rather than minimise the sum of squares of vertical residuals as in OLS regression, MA regression minimises the sum of the squared Euclidean distances of data points to the regression line. BNM predictions were not considered to be significantly different from UKBMS estimates if 1) a significant (> 0.05) positive correlation existed between the two, 2) the 95% confidence intervals of the regression intercept encompassed zero, and 3) the 95% confidence intervals of the regression slope encompassed 1 (Mesple *et al.* 1996). Meeting these three criteria indicated that there was a good match between phenology estimates derived from the BNM and the UKBMS. Significance of the correlation coefficient was assessed using 999 permutations. Regressions were performed using the lmodel2 package in R (Legendre 2008).

## INTERSPECIFIC VARIATION IN THE MISMATCH BETWEEN BNM AND UKBMS PHENOLOGY ESTIMATES

The average absolute value of the difference between the UKBMS and BNM phenology estimates was calculated. This gave the average mismatch in days of mean flight date for each species. The 95% confidence intervals around these means were also calculated. Sample size was extracted for each species over the 15 year time period. This is defined as the total number of BNM records for a given species summed over all years. The average mismatch for each species was then regressed against their sample sizes, as were the 95% confidence intervals. The absolute value of the confidence interval was used to represent the potential error above or below the mean. Variables were log transformed to meet parametric assumptions. This analysis was not performed for the flight period length estimates due to the poor ability of the distribution data to predict this metric (see results). Model predictions were compared across species with varying total sample sizes in order to locate a threshold number of BNM records that 1) did not predict an average maximum mismatch of greater than five days, 2) was smaller than the majority of species whose BNM predictions were successful, and 3) was greater than the majority of species whose BNM predictions did not match those of the UKBMS. An accuracy of at least 5 days was chosen as reported phenological shifts over similar time periods tend to be larger than this, ensuring that distribution data could detect phenological changes if they were present (Crick *et al.* 1997; Roy & Sparks 2000; Fitter & Fitter 2002).

We tested for phylogenetic autocorrelation in model residuals using Moran I tests with Geary randomisations in the ade4 R package (Paradis 2006; Dray & Dufour 2007). We used 1000 phylogenetic trees as described in Oliver *et al.* (2012), using closely related conger species where molecular sequences were not available for four out of the 24 butterfly species. For none of the iterations for either model was significant phylogenetic autocorrelation in residuals apparent.

### INTRASPECIFIC VARIATION IN MISMATCH WITH ALTERED SAMPLE SIZE

Species that were predicted successfully by the BNM and had a sample size greater than the threshold size determined in the previous section were subsampled to further investigate the influence of decreased sample size. Subsets of decreasing size were randomly extracted from the original BNM data for each species. This procedure gave 20 levels of subsampling, decreasing from 100% in 5% increments. This randomisation was repeated 100 times to obtain the average mismatch in mean flight date and associated 95% confidence intervals for each subsampling level for each species. These mismatches and absolute confidence intervals were then regressed against the actual subsample sizes in the same way as described in the previous section. Data organisation and preparation took place in Microsoft Access and R. All analyses took place in R (R Development Core Team 2011).

## Results

### TESTING BNM PREDICTIONS

Mean flight date: 26 out of the 30 species tested had a significant linear relationship between the UKBMS and BNM estimates ($p<0.05$) for mean flight date. Of these, 22 had 95% confidence intervals of the intercept that included zero and of the slope that included one. Thus, the yearly predictions of mean flight date derived from BNM data are not significantly different from the UKBMS dataset for the majority of the British univoltine species. Figure 1 displays scatterplots and associated regressions for *Anthocharis cardamines* (Linnaeus, 1758) and *Pyronia tithonus* (Linnaeus, 1767), two example species randomly chosen from those which did not differ from a 1:1 line. Regression details and plots for all species are given in the appendix and slope estimates are displayed in figure 2.

Flight period length: Seven out of 30 species had a significant linear relationship between the UKBMS and BNM estimates ($p<0.05$) for flight period length. All seven had 95% confidence intervals

of the intercept that included zero and of the slope that included one. This indicates that the yearly predictions of flight period length from each dataset are divergent for the majority of univoltine species. Regression details for all species are given in the appendix. Figure 3 displays scatterplots and associated regressions for a successfully predicted and unsuccessfully predicted species: *A. cardamines* and *P. tithonus,* respectively. Plots for all species are presented in the appendix.

## INTERSPECIFIC VARIATION IN THE MISMATCH BETWEEN BNM AND UKBMS PHENOLOGY ESTIMATES

Mean flight date: A significant positive linear relationship was found between log average mismatch between BNM and UKBMS predictions and log species' sample size ($df = 28$, $R^2 = 0.28$, $t = -3.27$, $p<0.01$, figure 4a). The average mismatch in mean flight date estimates decreases exponentially with increasing sample size. A similar relationship is also seen between the log 95% confidence intervals of the mean mismatch and log sample size, indicating that the error about mean flight date estimates also decreases with increasing sample size ($df = 28$, $R^2 = 0.43$, $t = -4.59$, $p<0.01$, figure 4b). Combined, these results suggest a threshold sample size below which prediction accuracy rapidly deteriorates. Based on the criteria given above, a threshold of 6,500 records over 15 years predicts a maximum mismatch (mean mismatch + 95% CI) of 5.03 days. This threshold is also smaller than the sample sizes of 15 out of the 21 species predicted successfully by the BNM and is larger than six out of the eight species not predicted successfully by the BNM. The predicted mean mismatch and 95% CI at this threshold is 3.65 ± 1.38 days.

## INTRASPECIFIC VARIATION IN MISMATCH WITH ALTERED SAMPLE SIZE

Mean Flight Date: By reducing records through subsampling, an increase in the average mismatch between BNM and UKBMS predictions was observed, alongside an increase in the 95% confidence intervals. This is consistent with the results from the previous section. For example, *A. cardamines*

showed a significant negative relationship between log subsample size and log average mismatch (df=18, $R^2$=0.64, t=-5.73, p<0.01, figure 5a), as did *P. tithonus* (df=18, $R^2$=0.87, t=-10.78, p<0.01, figure 5b). The only exception, was one species, *Aphantopus hyperantus* (Linnaeus, 1758), which showed no relationship. Details for all species are presented in the appendix. Species also showed a negative relationship between the size of the 95% confidence intervals around the mismatch and subsample size – similar to the interspecific analysis. *Pyronia tithonus* illustrates the general pattern of the results (df=18, $R^2$=0.82, t=-9.2, p<0.01). Details for all species are presented in the appendix.

As species were progressively subsampled they showed little increase in mismatch between BNM and UKBMS predictions of mean flight date whilst sample sizes were above the 6,500 record threshold. Across the 15 species that had sufficient initial sample size for this analysis the average increase in mismatch after subsampling to 6,500 records was 2.01 ± 0.28 [SE] days (Table 1). These increases in mismatch are predicted from the significant linear models described above, but their magnitude is incredibly small in terms of the number of days of mismatch that may be expected with smaller subsampling levels.

## Discussion

It appears that distribution data may be useful for predicting phenology metrics. In this study, distribution data were able to accurately predict the mean flight dates derived from a standardised-effort recording scheme for the majority of univoltine species (22 out of 30 species). Less successful however, was the prediction of flight period length. In only seven out of 30 species did distribution data accurately predict flight period. This is in accordance with our hypotheses. Furthermore, this study has shown that there is a consistent relationship between the degree of mismatch which may be expected between distribution data and a standardised-effort recording scheme, versus the number of distribution records being used. This relationship suggests a threshold sample size of

approximately 6,500 records beyond which prediction accuracy deteriorates rapidly. Below, we discuss potential reasons for the greater mismatch between the data types in certain cases and the wider applicability of distribution data in phenology research.

Firstly, the ability of distribution data to match the UKBMS predictions of mean flight date for the majority of species tested is remarkable given the temporal biases that are expected to be present within the BNM dataset. As discussed previously, these are largely related to the uneven distribution of records throughout time due to specific aspects of recorder behaviour. This study suggests that either these biases do not exist in a substantial form, or that the phenology signal is strong enough to be seen through them. Regardless, the results highlight that distribution data are an underexploited but potentially important resource for phenology research.

Of the eight species whose BNM mean flight date estimates do not match those generated from the UKBMS, six had a sample size below 6,500. In these cases, a simple explanation of poor sample size may suffice. At these low sample sizes the potential biases in recorder behaviour may have become pronounced enough to overcome the strength of the phenology signal. The issues associated with rarity may, however, equally be influencing the UKBMS estimates. If this is the case then the mismatches may be explained by the different magnitudes or directions in which the data types are influenced by rare species. For example, the UKBMS does not routinely generate population trends for *Carteorocephalus palaemon* or *Melitaea cinxia*. Both of these species are not predicted well by the BNM and have small sample sizes. Alternatively, the better geographical coverage and fewer constraints placed on recorders may mean that the BNM provides a better estimate of phenology for some species. Further testing of population monitoring methods would be able to differentiate between these two possibilities.

The BNM does not match the UKBMS estimates for *Thymelicus lineola* despite a relatively large sample size of 21,947 records. The reasons for the increased mismatch between the datasets for this species are unclear and further work investigating patterns of recording and monitoring in relation

to species traits may go some way to explaining the mismatch between the BNM and the UKBMS. However, generally speaking, the mean flight dates of widespread and relatively common species are predicted well by the BNM.

The second key finding of this study is that flight period length is not predicted accurately by the BNM. For the flight period length analyses, the majority of species (16 out of 30) displayed a regression slope less than one (see appendix). Whilst most of these slopes were not actually significant, this trend indicates that the BNM estimates of flight period length tended to be greater than those of the BMS. This trend can be explained if recorders are oversensitive to a species outside of its peak abundance period. Individuals that are seen either early or late in the season could attract a higher number of submitted records due to the novelty or unexpectedness of being sighted. The flight period length will then be overestimated due to the inflated number of records at the extreme ends of the flight period. This offers only a tentative explanation for the inability of the BNM to predict flight period length, yet it is grounded in the temporal biases likely to be present within distribution data. In addition, the marginally broader latitudinal range of BNM records (appendix figure 4) and/or a greater range of microclimates sampled by the BNM may lead to greater variance about the mean and result in greater estimates of flight period length.

Thirdly, our study finds a consistent pattern in the relationship between the average mismatch between distribution and transect data, and sample size. With decreasing sample size there is a trend for the average mismatches and their associated 95% confidence intervals to increase at both inter and intraspecific scales. This is an understandable and expected relationship. The true applied consequences of this, however, may not be appreciated without reference to the magnitude of the observed change. There is very little change in the expected accuracy as sample size is decreased above the threshold of 6,500 records. For example, at its original sample size of 132,647 records *A. cardamines* has an average mismatch of 2.26 ± 0.94 days. This increases to a predicted 2.3 ± 0.99 days at a subsample size of 6,500. This is a 95.1% reduction in sample size with an average increase

in mismatch of 0.04 days and a potential maximum increase of 1.97 days. Similar patterns are seen for the other subsampled species (Table 1). This fact not only highlights the robust nature of the methodology, but also its potential use. An average of 430 records per year (the average number of records from 6,500 records over 15 years) could be an achievable target for a range of widespread taxa and geographic regions. For example, data from the NBN Gateway suggests that around 60% of butterfly species, 20% of moths and 50% of dragonflies meet the threshold number of records (figure 6). These numbers are encouraging but emphasize the continued need for large scale citizen-science schemes, especially if we wish to understand the phenology of less well-studied taxa across the entire breadth of the ecological web.

Additionally, the possible range over which distribution data estimates may deviate from transect generated estimates for those species succeeding the 6,500 threshold is on the scale of 7.29 (*Satyrium w-album* Knoch 1782) to 0.64 days (*P. tithonus*). This margin of error tends to be below recorded long term (i.e. several decade) phenological changes for butterflies (Roy & Sparks 2000; Forister & Shapiro 2003; Stefanescu, Penuelas & Filella 2003), birds (Crick *et al.* 1997) and plants (Fitter & Fitter 2002). This suggests that distribution data could play a role in investigating long term changes. From our analyses, we suggest that species with a total sample size greater than 6,500 records over 15 years (430 per year) should be appropriate for phenology analyses.

Whilst these results are encouraging for the use of distribution data in detecting and monitoring butterfly phenology, they should not be limited to this taxon. In this study, only a single assumption has been made regarding the life history of the organisms. This is that the phenology event in question (e.g. butterfly emergence periods) follows a Gaussian distribution. Such a pattern is common for a wide range of phenological phenomena, a few individuals are relatively early or late whilst the majority fall toward the middle. This lack of complicating assumptions makes the application of distribution data to phenological research easily generalised. A caveat may be that the behaviour of butterfly recorders is fundamentally different from other biological recorders. We see

no strong reasons, however, why this should be the case. In addition, future users of distribution data in phenological research should limit their work to spatial scales no larger than that used in the validation tests performed here. We also encourage the use of distribution data at smaller spatial scales as long as the threshold sample size is reached.

In summary, this study illustrates the utility of using distribution data to predict aspects of a species phenology. Crucially, mean dates can be predicted well, whilst the range of time over which species are apparent (flight period lengths in this butterfly example) are not. These mean flight date predictions are robust and require relatively small sample sizes to achieve adequate levels of accuracy. This, coupled with the ease with which these types of data may be collected, suggest that distribution data can make a valid contribution to the continued monitoring and study of phenology in a range of taxa and locations.

## Acknowledgements

We thank the many thousands of volunteer recorders who have gathered UKBMS and BNM data over the 15 year period of this study and Graham French for supplying NBN data. We thank Albert Phillimore and anonymous reviewers for valuable comments on an earlier draft. TRB received financial support from the Grundy Educational Trust. The BNM and UKBMS projects are funded by Countryside Council for Wales, DEFRA, Environment and Heritage Service, Forestry Commission, Joint Nature Conservation Committee, Natural England, Northern Ireland Environment Agency, Redwing Trust, Scottish Executive Environment and Rural Affairs Department and Scottish Natural Heritage.

## References

Altermatt, F. (2010) Climatic warming increases voltinism in European butterflies and moths. *Proceedings of the Royal Society B-Biological Sciences,* **277,** 1281-1287.

Araujo, M.B., Thuiller, W., Williams, P.H. & Reginster, I. (2005) Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Global Ecology and Biogeography,* **14,** 17-30.

Bale, J.S., Masters, G.J., Hodkinson, I.D., Awmack, C., Bezemer, T.M., Brown, V.K., Butterfield, J., Buse, A., Coulson, J.C., Farrar, J., Good, J.E.G., Harrington, R., Hartley, S., Jones, T.H., Lindroth, R.L., Press, M.C., Symrnioudis, I., Watt, A.D. & Whittaker, J.B. (2002) Herbivory in global climate change research: direct effects of rising temperature on insect herbivores. *Global Change Biology,* **8,** 1-16.

Balmford, A., Bennun, L., ten Brink, B., Cooper, D., Cote, I.M., Crane, P., Dobson, A., Dudley, N., Dutton, I., Green, R.E., Gregory, R.D., Harrison, J., Kennedy, E.T., Kremen, C., Leader-Williams, N., Lovejoy, T.E., Mace, G., May, R., Mayaux, P., Morling, P., Phillips, J., Redford, K., Ricketts, T.H., Rodriguez, J.P., Sanjayan, M., Schei, P.J., van Jaarsveld, A.S. & Walther, B.A. (2005) The convention on biological diversity's 2010 target. *Science,* **307,** 212-213.

Both, C., van Asch, M., Bijlsma, R.G., van den Burg, A.B. & Visser, M.E. (2009) Climate change and unequal phenological changes across four trophic levels: constraints or adaptations? *Journal of Animal Ecology,* **78,** 73-83.

Botham, M.S., Brereton, T.M., Middlebrook, I., Cruickshanks, K.L., Harrower, C., Beckmann, B. & Roy, D.B. (2008) United Kingdom Butterfly Monitoring Scheme report for 2008. CEH Wallingford.

Brakefield, P.M. (1987) Geographical variability in, and temperature effects on, the phenology of *Maniola jurtina* and *Pyronia tithonus* (Lepidoptera, Satyrinae) in England and Wales. *Ecological Entomology,* **12,** 139-148.

Brereton, T., Roy, D.B., Middlebrook, I., Botham, M. & Warren, M. (2011) The development of butterfly indicators in the United Kingdom and assessments in 2010. *Journal of Insect Conservation,* **15,** 139-151.

Carroll, E.A., Sparks, T.H., Collinson, N. & Beebee, T.J.C. (2009) Influence of temperature on the spatial distribution of first spawning dates of the common frog (Rana temporaria) in the UK. *Global Change Biology,* **15,** 467-473.

Clark, J.A. & May, R.M. (2002) Taxonomic bias in conservation research. *Science,* **297,** 191-192.

Crick, H.Q.P., Dudley, C., Glue, D.E. & Thomson, D.L. (1997) UK birds are laying eggs earlier. *Nature,* **388,** 526-526.

Dennis, R.L.H., Shreeve, T.G., Isaac, N.J.B., Roy, D.B., Hardy, P.B., Fox, R. & Asher, J. (2006) The effects of visual apparency on bias in butterfly recording and monitoring. *Biological Conservation,* **128,** 486-492.

Dennis, R.L.H., Sparks, T.H. & Hardy, P.B. (1999) Bias in butterfly distribution maps: The effects of sampling effort. *Journal of Insect Conservation,* **3,** 33-42.

Dennis, R.L.H. & Thomas, C.D. (2000) Bias in butterfly distribution maps: The influence of hot spots and recorder's home range. *Journal of Insect Conservation,* **4,** 73-77.

Doi, H., Gordo, O. & Katano, I. (2008) Heterogeneous intra-annual climatic changes drive different phenological responses at two trophic levels. *Climate Research,* **36,** 181-190.

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

Fitter, A.H. & Fitter, R.S.R. (2002) Rapid changes in flowering time in British plants. *Science,* **296,** 1689-1691.

Forister, M.L. & Shapiro, A.M. (2003) Climatic trends and advancing spring flight of butterflies in lowland California. *Global Change Biology,* **9,** 1130-1135.

Fox, R., Asher, J., Brereton, T., Roy, D. & Warren, M. (2006) *The State of Butterflies in Britain and Ireland*. Pisces Publications.

Funk, V.A. & Richardson, K.S. (2002) Systematic data in biodiversity studies: Use it or lose it. *Systematic Biology,* **51,** 303-316.

Harding, P.T. & Sheail, J. (1992) The Biological Records Centre: a pioneer in data gathering and retrieval. *Biological recording of changes in British wildlife* (ed. P.T. Harding), pp. 5-19. HMSO, London.
Hassall, C., Thompson, D.J., French, G.C. & Harvey, I.F. (2007) Historical changes in the phenology of British Odonata are related to climate. *Global Change Biology,* **13,** 933-941.
Hill, J.K., Griffiths, H.M. & Thomas, C.D. (2011) Climate Change and Evolutionary Adaptations at Species' Range Margins. *Annual Review of Entomology, Vol 56,* **56,** 143-159.
Leather, S.R. (2009) Institutional vertebratism threatens UK food security. *Trends in Ecology & Evolution,* **24,** 413-414.
Legendre, P. (2008) lmodel2: Model II Regression. www.cran.r-project.org.
Legendre, P. & Legendre, L. (1998) *Numerical Ecology,* 2nd edn. Elsevier, Amsterdam.
Memmott, J., Craze, P.G., Waser, N.M. & Price, M.V. (2007) Global warming and the disruption of plant-pollinator interactions. *Ecology Letters,* **10,** 710-717.
Menendez, R., Megias, A.G., Hill, J.K., Braschler, B., Willis, S.G., Collingham, Y., Fox, R., Roy, D.B. & Thomas, C.D. (2006) Species richness changes lag behind climate change. *Proceedings of the Royal Society B-Biological Sciences,* **273,** 1465-1470.
Mesple, F., Troussellier, M., Casellas, C. & Legendre, P. (1996) Evaluation of simple statistical criteria to qualify a simulation. *Ecological Modelling,* **88,** 9-18.
Miller-Rushing, A.J., Hoye, T.T., Inouye, D.W. & Post, E. (2010) The effects of phenological mismatches on demography. *Philosophical Transactions of the Royal Society B-Biological Sciences,* **365,** 3177-3186.
Oliver, T.H., Roy, D.B., Brereton, T. & Thomas, J.A. (2012) Reduced variability in range-edge butterfly populations over three decades of climate warming. *Global Change Biology,* **18,** 1531-1539.
Paradis, E. (2006) *Analysis of Phylogenetics and Evolution with R,* 2nd edn. Springer, Baltimore.
Parmesan, C. (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics,* **37,** 637-669.
Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J.K., Thomas, C.D., Descimon, H., Huntley, B., Kaila, L., Kullberg, J., Tammaru, T., Tennent, W.J., Thomas, J.A. & Warren, M. (1999) Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature,* **399,** 579-583.
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.
Poyry, J., Leinonen, R., Soderman, G., Nieminen, M., Heikkinen, R.K. & Carter, T.R. (2011) Climate-induced increase of moth multivoltinism in boreal regions. *Global Ecology and Biogeography,* **20,** 289-298.
R Development Core Team (2011) R: A language and environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Robertson, M.P., Cumming, G.S. & Erasmus, B.F.N. (2010) Getting the most out of atlas data. *Diversity and Distributions,* **16,** 363-375.
Root, T.L., Price, J.T., Hall, K.R., Schneider, S.H., Rosenzweig, C. & Pounds, J.A. (2003) Fingerprints of global warming on wild animals and plants. *Nature,* **421,** 57-60.
Rosenzweig, C., Casassa, G., Karoly, D.J., Imeson, A., Liu, C., Menzel, A., Rawlins, S., Root, T.L., Seguin, B. & Tryjanowski, P. (2007) Assessment of observed changes and responses in natural and managed systems *Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden & C.E. Hanson), pp. 79-131. Cambridge University Press, Cambridge, UK.
Roy, D.B. & Sparks, T.H. (2000) Phenology of British butterflies and climate change. *Global Change Biology,* **6,** 407-416.
Stefanescu, C., Penuelas, J. & Filella, I. (2003) Effects of climatic change on the phenology of butterflies in the northwest Mediterranean Basin. *Global Change Biology,* **9,** 1494-1506.

Thackeray, S.J., Sparks, T.H., Frederiksen, M., Burthe, S., Bacon, P.J., Bell, J.R., Botham, M.S., Brereton, T.M., Bright, P.W., Carvalho, L., Clutton-Brock, T., Dawson, A., Edwards, M., Elliott, J.M., Harrington, R., Johns, D., Jones, I.D., Jones, J.T., Leech, D.I., Roy, D.B., Scott, W.A., Smith, M., Smithers, R.J., Winfield, I.J. & Wanless, S. (2010) Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments. *Global Change Biology,* **16,** 3304-3313.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L. & Williams, S.E. (2004) Extinction risk from climate change. *Nature,* **427,** 145-148.

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 B-Biological Sciences,* **360,** 339-357.

Tylianakis, J.M., Didham, R.K., Bascompte, J. & Wardle, D.A. (2008) Global change and species interactions in terrestrial ecosystems. *Ecology Letters,* **11,** 1351-1363.

Van Nouhuys, S. & Lei, G.C. (2004) Parasitoid-host metapopulation dynamics: the causes and consequences of phenological asynchrony. *Journal of Animal Ecology,* **73,** 526-535.

Visser, M.E. (2008) Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society B-Biological Sciences,* **275,** 649-659.

Visser, M.E. & Both, C. (2005) Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society B-Biological Sciences,* **272,** 2561-2569.

Walther, G.R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.M., Hoegh-Guldberg, O. & Bairlein, F. (2002) Ecological responses to recent climate change. *Nature,* **416,** 389-395.

Willis, C.G., Ruhfel, B., Primack, R.B., Miller-Rushing, A.J. & Davis, C.C. (2008) Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. *Proceedings of the National Academy of Sciences of the United States of America,* **105,** 17029-17033.

# Tables

Table 1. Estimates for the mismatch in mean flight date between BNM and UKBMS datasets using either the original BNM number of records or subsampled to 6,500 records. For the subsampled data, estimates of mismatch and 95% confidence intervals in days are calculated from log-linear models of mismatch and 95% confidence interval versus subsample size over 100 iterations.

| | Average Mismatch ± 95% CI | | |
|---|---|---|---|
| Species | Original record number | Threshold record number | Percent reduction in record number |
| *Anthocharis cardamines* | 2.26 ± 0.94 | 2.3 ± 0.99 | 95.1 |
| *Aphantopus hyperantus* | 2.52 ± 0.47 | 2.52 ± 0.57 | 93.72 |
| *Argynnis aglaja* | 4.31 ± 1.22 | 4.34 ± 1.3 | 57.74 |
| *Argynnis paphia* | 1.65 ± 0.52 | 1.74 ± 0.63 | 61.6 |
| *Callophrys rubi* | 3.4 ± 1.55 | 3.48 ± 1.57 | 67.17 |
| *Hipparchia semele* | 2.45 ± 0.93 | 2.52 ± 0.96 | 48.31 |
| *Limenitis camilla* | 2.13 ± 0.94 | 2.16 ± 0.96 | 40.63 |
| *Polyommatus coridon* | 2.94 ± 1.41 | 2.94 ± 1.43 | 41.18 |
| *Maniola jurtina* | 3.59 ± 1.15 | 3.61 ± 1.19 | 97.77 |
| *Melanargia galathea* | 2.45 ± 0.67 | 2.47 ± 0.77 | 82.75 |
| *Favonius quercus* | 2.25 ± 0.91 | 2.37 ± 0.96 | 66.19 |
| *Ochlodes sylvanus* | 1.24 ± 0.55 | 1.37 ± 0.61 | 91.75 |
| *Plebejus argus* | 6.26 ± 2.49 | 6.21 ± 2.47 | 7.3 |
| *Pyronia tithonus* | 0.55 ± 0.23 | 0.68 ± 0.29 | 96.09 |
| *Thymelicus sylvestris* | 1.21 ± 0.35 | 1.26 ± 0.47 | 92.55 |

# Figures

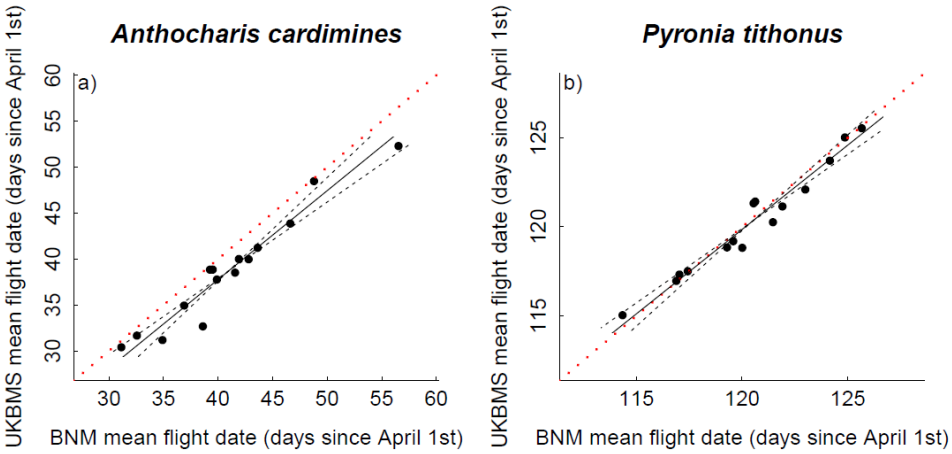



Figure 1. Mean flight date predictions of the BMS plotted against those of the BNM for a) A*nthocharis. cardamines* and b) *Pyronia tithonus*. Dotted red line marks the 1:1 line. Solid black line is the major axis regression line for each species. Dashed lines are 95% confidence intervals for the regression slope.

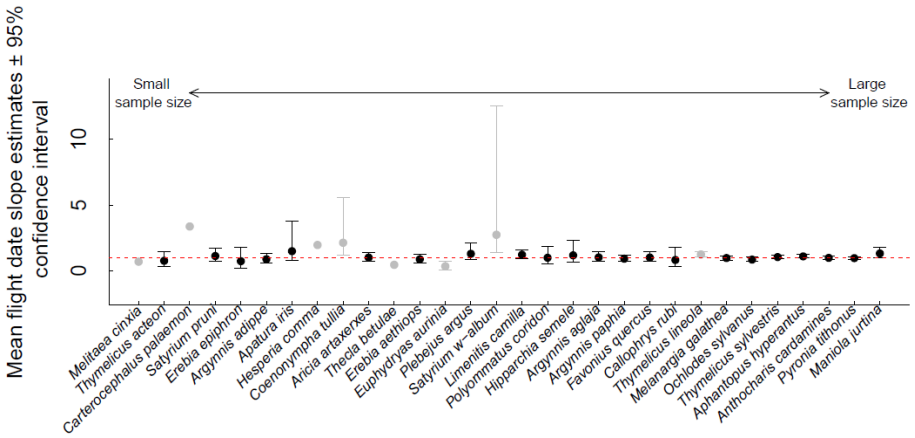



Figure 2. Slope estimates ± 95% confidence interval for mean flight date regressions. Black points indicate species that conformed to a 1:1 line, indicating that distribution data (BNM) estimates matched those of the transect data (UKBMS). Grey points indicate species which did not conform to a 1:1 line. Species ordered from small to large sample sizes (total number of records). Red dashed lines marks a slope value of 1. Confidence intervals not included for those species whose mean flight date estimates were not significantly correlated.

1

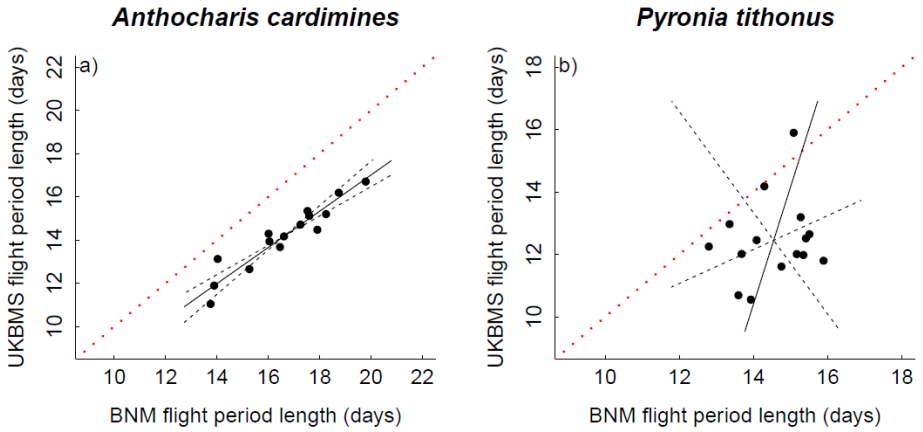



Figure 3. Flight period length predictions of the BMS plotted against those of the BNM for a) *Anthocharis cardamines* and b) *Pyronia tithonus*. Dotted red line marks the 1:1 line. Solid black line is the major axis regression line for each species. Dashed lines are 95% confidence intervals for the regression slope.

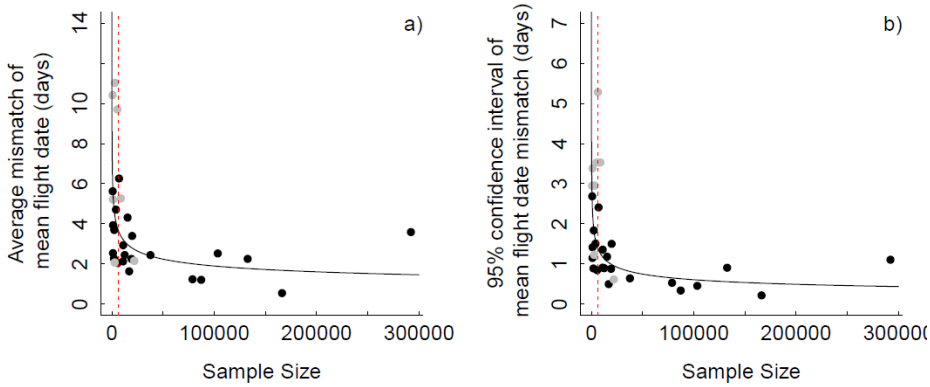



Figure 4. a) Plot of average mismatch in mean flight date estimates between BNM and UKBMS datasets in days (m) against sample size (s): The fitted curve is the equation: log(m) = 3.38 – 0.24*log(s). b) plot of 95% confidence intervals (CI) of the average mismatch against sample size: log(CI) = 2.94 – 0.30*log (s). In both plots, each point represents a single species and the vertical red dashed line indicates the selected 6,500 record threshold, below which prediction accuracy rapidly declines.

1

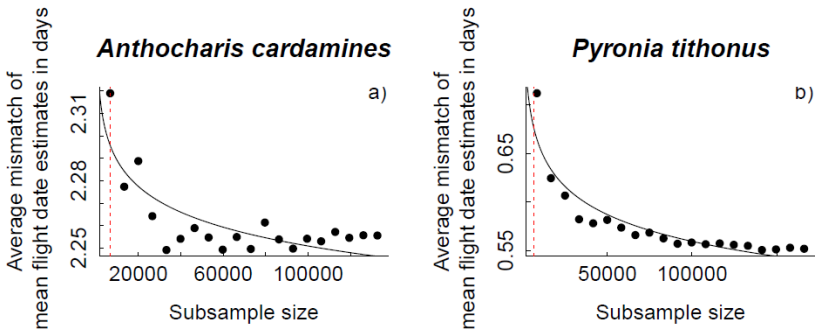



Figure 5. Plots of average mismatch in mean flight date estimates between BNM and UKBMS datasets in days (m) against subsample size (s) for *A. cardamines* (panel a; curve equation: log(m) = 0.89 – 0.01*log(s); and *P. tithonus* (panel b; curve equation: log(m) = 0.21 – 0.07*log(s)). The dashed line refers to the 6,500 record threshold in both plots.

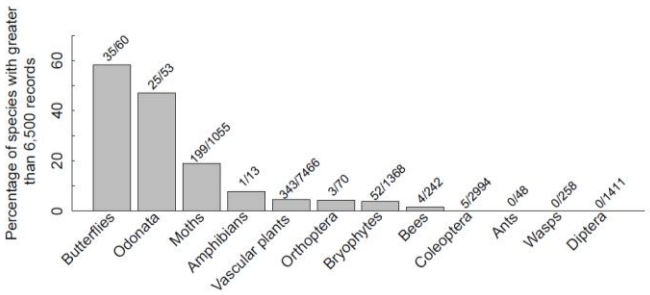



Figure 6. Barplot to show the percentage of species from various taxa that reach the 6,500 record threshold. Data obtained from the NBN Gateway (data.nbn.org.uk, accessed January 2013).

## Appendix

**Appendix Table 1.** Mean flight date BMS versus BNM regression details for all species. Prediction accuracy indicates whether a 1:1 line is achieved in the regression. Prediction accuracy is assessed on the p value of the correlation coeffiecient *r* ($p<0.05$), whether the intercept confidence intervals encompass 0 and whether the slope confidence intervals encompass 1. Sample size is the total number of BNM records between 1995 and 2009.

| Species | Intercept | 2.5% CI Intercept | 97.5% CI Intercept | Slope | 2.5% CI Slope | 97.5% CI Slope | r | p | Prediction Accuracy | Sample Size |
|---|---|---|---|---|---|---|---|---|---|---|
| *Anthocharis cardamines* | -0.89 | -7.39 | 4.72 | 0.97 | 0.83 | 1.13 | 0.97 | <0.01 | 1:1 | 132647 |
| *Apatura iris* | -51.13 | -293.86 | 26.32 | 1.48 | 0.74 | 3.78 | 0.64 | 0.01 | 1:1 | 2353 |
| *Aphantopus hyperantus* | -10.98 | -25.46 | 1.74 | 1.08 | 0.96 | 1.22 | 0.98 | <0.01 | 1:1 | 103442 |
| *Argynnis adippe* | 13.83 | -30.11 | 44.93 | 0.86 | 0.57 | 1.28 | 0.84 | <0.01 | 1:1 | 2042 |
| *Argynnis aglaja* | -4.9 | -49.5 | 26.61 | 1.01 | 0.71 | 1.42 | 0.87 | <0.01 | 1:1 | 15382 |
| *Argynnis paphia* | 9.99 | -19.37 | 33.41 | 0.9 | 0.7 | 1.16 | 0.92 | <0.01 | 1:1 | 16926 |
| *Aricia artaxerxes* | -3.85 | -44.26 | 24.75 | 0.99 | 0.7 | 1.4 | 0.87 | <0.01 | 1:1 | 4002 |
| *Callophrys rubi* | 9.97 | -37.89 | 34.88 | 0.81 | 0.32 | 1.75 | 0.63 | 0.01 | 1:1 | 19796 |
| *Carterocephalus palaemon* | -155.86 | 299.49 | 2.72 | 3.36 | 0.89 | -3.72 | 0.31 | 0.28 | - | 1100 |
| *Coenonympha tullia* | -119.61 | -453.72 | -28.26 | 2.12 | 1.18 | 5.56 | 0.66 | 0.01 | - | 3158 |
| *Erebia aethiops* | 20.21 | -28.44 | 55.71 | 0.86 | 0.58 | 1.23 | 0.86 | <0.01 | 1:1 | 5642 |
| *Erebia epiphron* | 31.14 | -77.12 | 85.28 | 0.71 | 0.18 | 1.77 | 0.56 | 0.03 | 1:1 | 1151 |
| *Euphydryas aurinia* | 34.21 | 3.1 | 59.95 | 0.33 | 0.03 | 0.68 | 0.54 | 0.04 | - | 6372 |
| *Favonius quercus* | -0.28 | -51.91 | 35.42 | 0.99 | 0.68 | 1.44 | 0.86 | <0.01 | 1:1 | 19224 |
| *Hesperia comma* | -130.75 | -26654.4 | 36.83 | 1.95 | 0.73 | 195.67 | 0.47 | 0.07 | - | 2813 |
| *Hipparchia semele* | -18.49 | -156.07 | 46.81 | 1.16 | 0.63 | 2.28 | 0.71 | <0.01 | 1:1 | 12575 |
| *Limenitis camilla* | -23.08 | -61.96 | 5.34 | 1.21 | 0.93 | 1.58 | 0.92 | <0.01 | 1:1 | 10948 |
| *Maniola jurtina* | -30.62 | -85.57 | 6.69 | 1.3 | 0.97 | 1.79 | 0.89 | <0.01 | 1:1 | 292134 |
| *Melanargia galathea* | 2.27 | -15.94 | 17.77 | 0.95 | 0.81 | 1.13 | 0.96 | <0.01 | 1:1 | 37688 |
| *Melitaea cinxia* | 23.54 | -253.22 | 77.89 | 0.68 | -0.19 | 5.08 | 0.44 | 0.17 | - | 716 |
| *Ochlodes sylvanus* | 15.09 | -1.94 | 29.55 | 0.83 | 0.68 | 1.02 | 0.95 | <0.01 | 1:1 | 78797 |
| *Plebejus argus* | -33.71 | -114.3 | 11.29 | 1.28 | 0.82 | 2.09 | 0.8 | <0.01 | 1:1 | 7012 |
| *Polyommatus coridon* | 0.41 | -111.42 | 60.5 | 0.98 | 0.51 | 1.84 | 0.71 | <0.01 | 1:1 | 11050 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Pyronia tithonus* | 6.19 | -9.12 | 19.76 | 0.95 | 0.83 | 1.07 | 0.98 | <0.01 | 1:1 | 166194 |
| *Satyrium pruni* | -9.16 | -60.01 | 22.24 | 1.11 | 0.72 | 1.73 | 0.85 | <0.01 | 1:1 | 1105 |
| *Satyrium w-album* | -180.54 | -1210.19 | -40.39 | 2.73 | 1.39 | 12.54 | 0.58 | 0.02 | - | 8715 |
| *Thecla betulae* | 88.49 | -0.56 | 149.34 | 0.44 | 0.01 | 1.07 | 0.49 | 0.07 | - | 5322 |
| *Thymelicus acteon* | 34.67 | -50.18 | 84.77 | 0.74 | 0.32 | 1.47 | 0.66 | 0.01 | 1:1 | 809 |
| *Thymelicus lineola* | -27.51 | -54.48 | -5.64 | 1.23 | 1.03 | 1.46 | 0.96 | <0.01 | - | 21947 |
| *Thymelicus sylvestris* | -4.22 | -20.11 | 9.69 | 1.03 | 0.9 | 1.17 | 0.98 | <0.01 | 1:1 | 87266 |

**Appendix Table 2.** Flight period length BMS versus BNM regression details for all species. Prediction accuracy indicates whether a 1:1 line is achieved in the regression. Prediction accuracy is assessed on the p value of the correlation coeffiecient r ($p<0.05$), whether the intercept confidence intervals encompass 0 and whether the slope confidence intervals encompass 1. Sample size is the total number of BNM records between 1995 and 2009.

| Species | Intercept | 2.5% CI Intercept | 97.5% CI Intercept | Slope | 2.5% CI Slope | 97.5% CI Slope | r | p | Prediction Accuracy | Sample Size |
|---|---|---|---|---|---|---|---|---|---|---|
| *Anthocharis cardamines* | 0.26 | -2.86 | 2.9 | 0.84 | 0.68 | 1.03 | 0.95 | <0.01 | 1:1 | 132647 |
| *Apatura iris* | 707.85 | 70.66 | -55.17 | -66.34 | 6.57 | -5.46 | -0.05 | 0.85 | - | 2353 |
| *Aphantopus hyperantus* | -21.85 | 54.39 | 3.88 | 2.42 | 0.52 | -3.19 | 0.31 | 0.26 | - | 103442 |
| *Argynnis adippe* | -215.96 | 99.86 | -31.37 | 13.8 | 2.81 | -5 | 0.16 | 0.57 | - | 2042 |
| *Argynnis aglaja* | -2.15 | -36.7 | 9.58 | 0.93 | 0.27 | 2.88 | 0.54 | 0.04 | 1:1 | 15382 |
| *Argynnis paphia* | -7.48 | -97.77 | 7.48 | 1.31 | 0.44 | 6.52 | 0.52 | 0.05 | 1:1 | 16926 |
| *Aricia artaxerxes* | -44.22 | 94.62 | -0.09 | 3.05 | 0.82 | -3.95 | 0.31 | 0.25 | - | 4002 |
| *Callophrys rubi* | -133.47 | 107.18 | -19.43 | 8.43 | 2.13 | -4.87 | 0.22 | 0.44 | - | 19796 |
| *Carterocephalus palaemon* | 7.5 | -1 | 14.38 | 0.18 | -0.38 | 0.86 | 0.21 | 0.48 | - | 1100 |
| *Coenonympha tullia* | -46.86 | NA | NA | 4.01 | NA | NA | 0.06 | 0.82 | - | 3158 |
| *Erebia aethiops* | 4.31 | -5.16 | 10.55 | 0.38 | -0.13 | 1.16 | 0.39 | 0.15 | - | 5642 |
| *Erebia epiphron* | 7.61 | -0.33 | 15.02 | 0.05 | -0.6 | 0.75 | 0.05 | 0.85 | - | 1151 |
| *Euphydryas aurinia* | 12.09 | 3.48 | 20.82 | -0.04 | -0.24 | 0.16 | -0.12 | 0.68 | - | 6372 |
| *Favonius quercus* | -33.42 | 69.26 | 3.11 | 2.71 | 0.62 | -3.16 | 0.3 | 0.27 | - | 19224 |
| *Hesperia comma* | 1.94 | -3.81 | 5.85 | 0.67 | 0.32 | 1.18 | 0.71 | <0.01 | 1:1 | 2813 |
| *Hipparchia semele* | -298.06 | 329.05 | -86.27 | 16.85 | 5.54 | -16.64 | 0.29 | 0.3 | - | 12575 |
| *Limenitis camilla* | 7.6 | -2.82 | 15.01 | 0.3 | -0.22 | 1.03 | 0.32 | 0.24 | - | 10948 |
| *Maniola jurtina* | -1.36 | -9.97 | 4.76 | 1.05 | 0.76 | 1.45 | 0.88 | <0.01 | 1:1 | 292134 |
| *Melanargia galathea* | -36.22 | 19.29 | 14.03 | 3.39 | -0.12 | -0.49 | 0.16 | 0.56 | - | 37688 |
| *Melitaea cinxia* | 5.31 | -5.87 | 13.72 | 0.3 | -0.16 | 0.9 | 0.46 | 0.18 | - | 716 |
| *Ochlodes sylvanus* | -0.52 | -8.46 | 4.88 | 0.94 | 0.62 | 1.41 | 0.83 | <0.01 | 1:1 | 78797 |
| *Plebejus argus* | -23.46 | 976.82 | -0.74 | 2.23 | 0.86 | -58.04 | 0.47 | 0.08 | - | 7012 |
| *Polyommatus coridon* | 271.98 | 90.18 | -170.46 | -16.29 | 11.78 | -4.75 | -0.24 | 0.38 | - | 11050 |

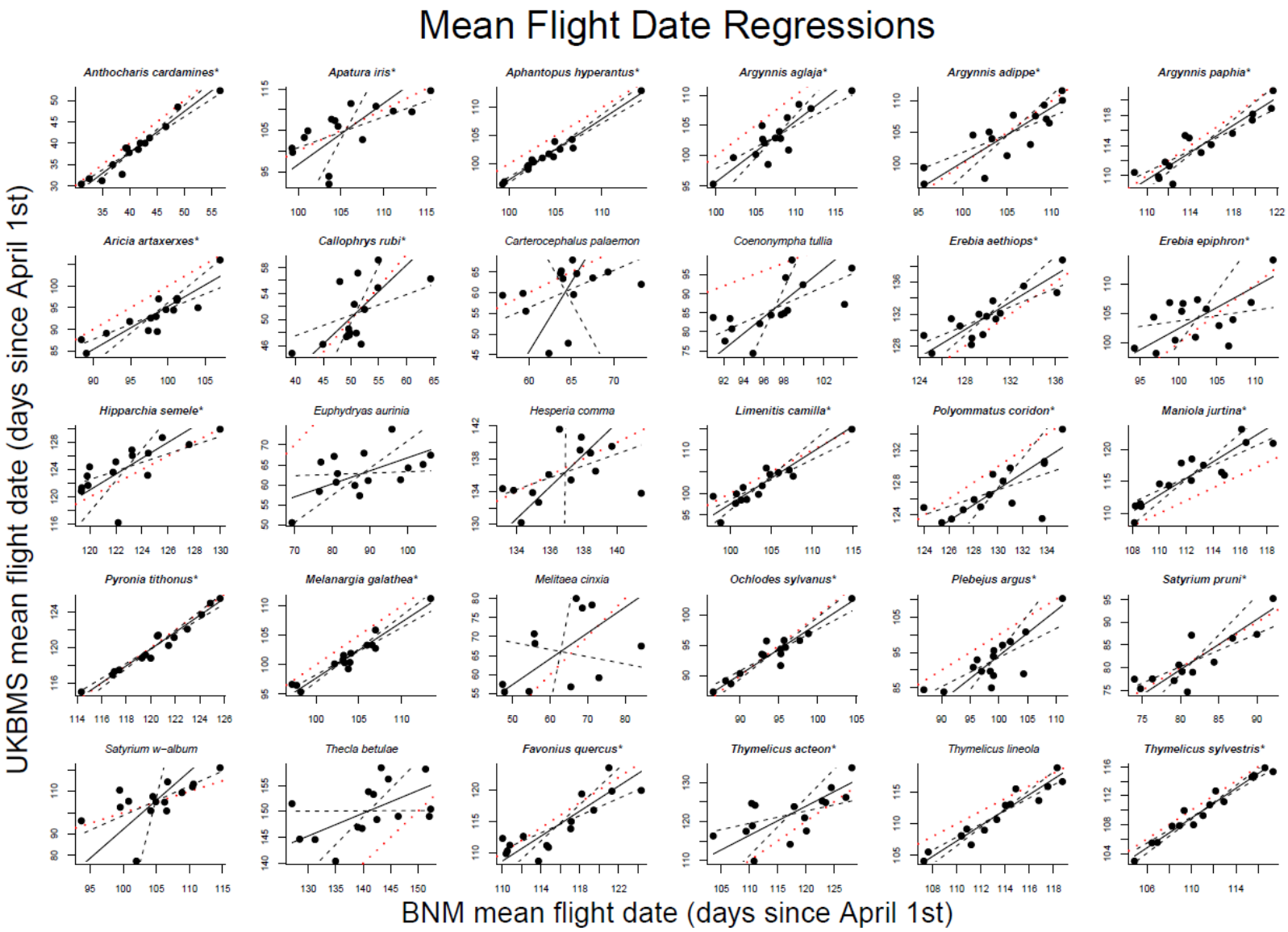
| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Pyronia tithonus* | -42.14 | 36.03 | 4.56 | 3.75 | 0.54 | -1.62 | 0.2 | 0.47 | - | 166194 |
| *Satyrium pruni* | 11.36 | 0.37 | -12.85 | -0.69 | 2.73 | 0.86 | -0.32 | 0.31 | - | 1105 |
| *Satyrium w-album* | -32.7 | 582.72 | -2.28 | 2.26 | 0.84 | -26.45 | 0.45 | 0.09 | - | 8715 |
| *Thecla betulae* | 13.69 | 8.96 | 18.38 | 0.03 | -0.09 | 0.15 | 0.17 | 0.56 | - | 5322 |
| *Thymelicus acteon* | 7.93 | -0.11 | 14.68 | 0.28 | -0.04 | 0.65 | 0.46 | 0.09 | - | 809 |
| *Thymelicus lineola* | -20.83 | 93.08 | 2.32 | 2.43 | 0.72 | -5.98 | 0.37 | 0.17 | - | 21947 |
| *Thymelicus sylvestris* | -0.97 | -13.42 | 5.96 | 0.9 | 0.45 | 1.7 | 0.7 | <0.01 | 1:1 | 87266 |

**Appendix Table 3.** Regression details for log average mismatch in days versus log subsample size. Subsamples generated through randomisation procedure.

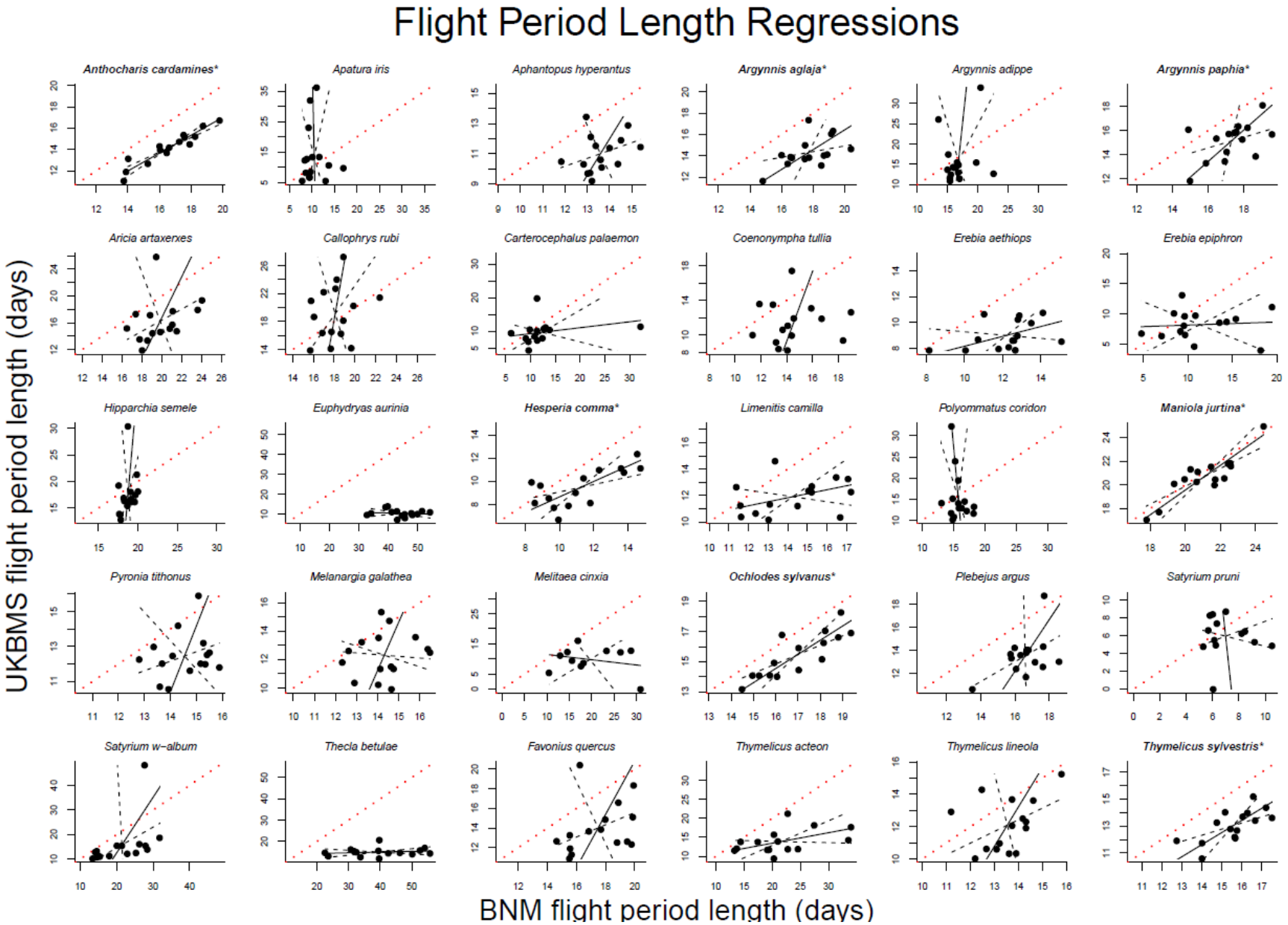
| Species | Intercept | Slope | df | R.sq | T.val.slope | Slope.p |
|---|---|---|---|---|---|---|
| *Anthocharis cardamines* | 0.89 | -0.01 | 18 | 0.65 | -5.73 | <0.01 |
| *Aphantopus hyperantus* | 0.92 | 0 | 18 | 0.05 | 1.02 | 0.32 |
| *Argynnis aglaja* | 1.59 | -0.01 | 18 | 0.73 | -6.91 | <0.01 |
| *Argynnis paphia* | 1.42 | -0.1 | 18 | 0.81 | -8.68 | <0.01 |
| *Callophrys rubi* | 1.52 | -0.03 | 18 | 0.82 | -9.2 | <0.01 |
| *Hipparchia semele* | 1.65 | -0.08 | 18 | 0.8 | -8.58 | <0.01 |
| *Limenitis camilla* | 1.45 | -0.08 | 18 | 0.9 | -12.63 | <0.01 |
| *Polyommatus coridon* | 1.47 | -0.04 | 18 | 0.81 | -8.62 | <0.01 |
| *Maniola jurtina* | 1.3 | 0 | 18 | 0.48 | -4.07 | <0.01 |
| *Melanargia galathea* | 0.99 | -0.01 | 18 | 0.45 | -3.83 | <0.01 |
| *Favonius quercus* | 1.44 | -0.07 | 18 | 0.88 | -11.37 | <0.01 |
| *Ochlodes sylvanus* | 0.7 | -0.04 | 18 | 0.89 | -12.18 | <0.01 |
| *Plebejus argus* | 1.98 | -0.02 | 18 | 0.6 | -5.2 | <0.01 |
| *Pyronia tithonus* | 0.22 | -0.07 | 18 | 0.87 | -10.78 | <0.01 |
| *Thymelicus sylvestris* | 0.39 | -0.02 | 18 | 0.73 | -7 | <0.01 |

**Appendix Table 4.** Regression details for log average mismatch 95% CIs in days versus log subsample size. Subsamples generated through randomisation procedure.

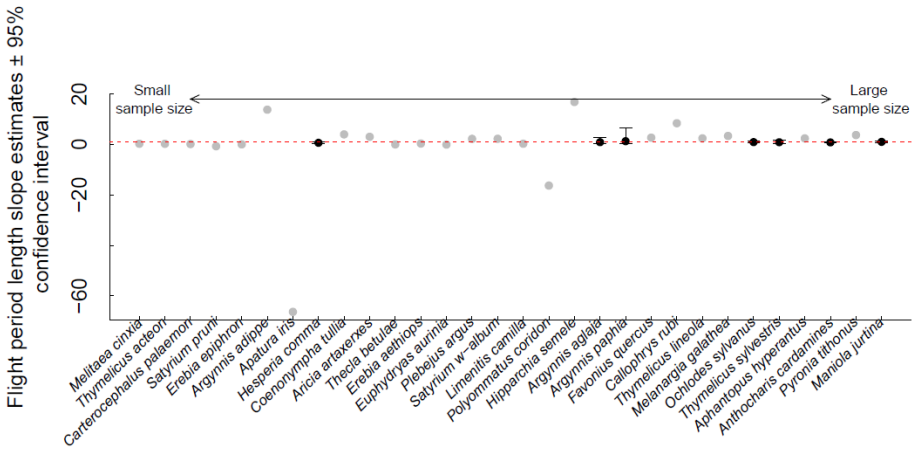
| Species | Intercept | Slope | df | R2 | t | p |
|---|---|---|---|---|---|---|
| *Anthocharis cardamines* | 0.13 | -0.02 | 18 | 0.84 | -9.78 | <0.01 |
| *Aphantopus hyperantus* | 0.1 | -0.08 | 18 | 0.89 | -12.24 | <0.01 |
| *Argynnis aglaja* | 1.09 | -0.09 | 18 | 0.93 | -15.33 | <0.01 |
| *Argynnis paphia* | 1.59 | -0.23 | 18 | 0.97 | -24.78 | <0.01 |
| *Callophrys rubi* | 0.64 | -0.02 | 18 | 0.63 | -5.5 | <0.01 |
| *Hipparchia semele* | 1.09 | -0.13 | 18 | 0.9 | -12.85 | <0.01 |
| *Limenitis camilla* | 0.73 | -0.09 | 18 | 0.86 | -10.63 | <0.01 |
| *Polyommatus coridon* | 0.82 | -0.05 | 18 | 0.94 | -16.13 | <0.01 |
| *Maniola jurtina* | 0.28 | -0.01 | 18 | 0.77 | -7.81 | <0.01 |
| *Melanargia galathea* | 0.55 | -0.09 | 18 | 0.91 | -13.39 | <0.01 |
| *Favonius quercus* | 0.56 | -0.07 | 18 | 0.82 | -9.16 | <0.01 |
| *Ochlodes sylvanus* | -0.14 | -0.04 | 18 | 0.91 | -13.85 | <0.01 |
| *Plebejus argus* | 1.29 | -0.04 | 18 | 0.9 | -12.76 | <0.01 |
| *Pyronia tithonus* | -0.55 | -0.08 | 18 | 0.82 | -9.19 | <0.01 |
| *Thymelicus sylvestris* | 0.31 | -0.12 | 18 | 0.94 | -16.67 | <0.01 |



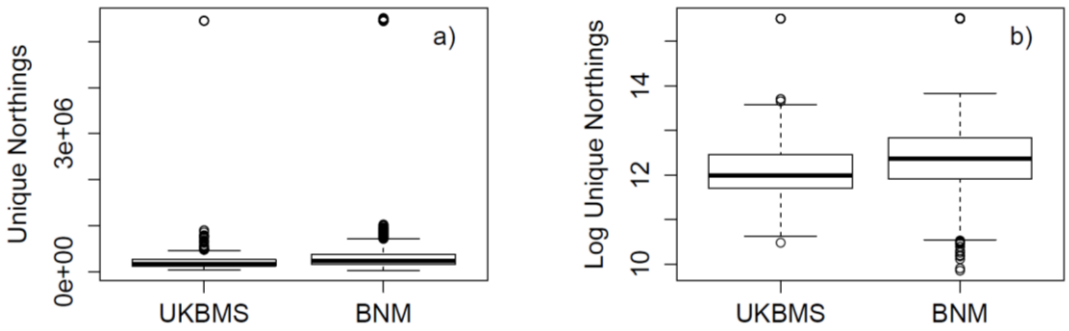

**Appendix figure 1.** Mean flight date predictions of the BMS plotted against those of the BNM for all species. Dotted red line marks the 1:1 line. Solid black line is the regression line for each species; dashed black lines are the 95% confidence intervals around slope estimates. Species in bold with an asterisk (*) fitted a 1:1 line.



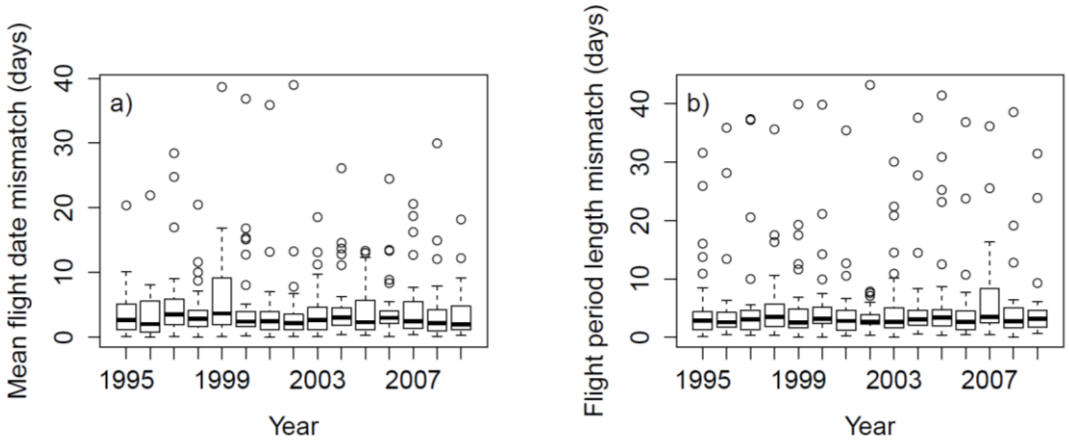

**Appendix figure 2**. Flight period length predictions of the BMS plotted against those of the BNM for all species. Dotted red line marks the 1:1 line. Solid black line is the regression line for each species; dashed black lines are the 95% confidence intervals around slope estimates. Species in bold with an asterisk (*) fitted a 1:1 line.




**Appendix figure 3**. Slope estimates ± 95% confidence interval for flight period length regressions. Black points indicate species that conformed to a 1:1 line, indicating that distribution data (BNM) estimates matched those of the transect data (UKBMS). Grey points indicate species which did not conform to a 1:1 line. Species ordered from small to large sample sizes (total number of records). Red dashed lines marks a slope value of 1. Confidence intervals not included for those species whose mean flight date estimates were not significantly correlated.




**Appendix figure 4**. Distribution of UKBMS and BNM records with unique northings. a) shows raw northings, b) shows log transformed northings.




**Appendix figure 5.** Degree of mismatch between UKBMS and BNM datasets over time. There is is no consistent change in the accuracy of mean flight date (a) or flight period length (b) estimates through time. This suggests that there is no systematic temporal bias between the two datasets.