Article (refereed)

Isaac, Nick J.B.; Cruickshanks, Katie L.; Weddle, Ann M.; Rowcliffe, J. Marcus; Brereton, Tom M.; **Dennis, Roger L.H.**; Shuker, David M.; Thomas, Chris D.. 2011 Distance sampling and the challenge of monitoring butterfly populations. *Methods in Ecology and Evolution*, 2 (6). 585-594. 10.1111/j.2041-210X.2011.00109.x

Distance sampling and the challenge of monitoring butterfly populations

Nick J.B. Isaac[1,2], Katie L. Cruickshanks[3], Ann M. Weddle[2,4], J. Marcus Rowcliffe[2], Tom M. Brereton[3], Roger L.H. Dennis[1,5], David M. Shuker[6] and Chris D. Thomas[4]

[1]Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB, UK

[2]Institute of Zoology, Zoological Society of London, Regent's Park, London, NW1 4RY, UK

[3]Butterfly Conservation, Manor Yard, East Lulworth, Wareham, Dorset, BH20 5QP, UK

[4]Department of Biology, University of York, Heslington, York, YO10 5DD, UK

[5]Institute for Environment, Sustainability and Regeneration, Mellor Building, Staffordshire University, College Road, Stoke on Trent, ST4 2DE, UK

[6]School of Biology, University of St Andrews, St Andrews, Fife, KY16 9TH, UK

Correspondence: *Nick Isaac, Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB. F: 01491 692424, E: njbi@ceh.ac.uk*

Running title: Distance sampling and monitoring butterflies

Keywords: apparency, bias, butterfly monitoring scheme, detectability, effective strip width, habitat, mixed models, Pollard Walk, sampling, transect.

Word Count: 7617 (including references, figures & tables)

Summary

1. Abundance indices generated by the UK Butterfly Monitoring Scheme (UKBMS) have been influential in informing our understanding of environmental change and highlighting UK conservation priorities. Here we critically evaluate the standard 'Pollard Walk' methodology employed by the UKBMS.

2. We consider the systematic sampling biases among different butterfly species and biotopes using distance sampling. We collected over 5000 observations on 17 species using distance sampling at 13 study sites in England and Wales. We fitted detection functions to explore variation in detectability among species and sites.

3. Our results suggest that around one third of individual butterflies in the Pollard Walk box were missed. However, detectability varies markedly among species and sites. We provide the first species-specific estimates of detectability for converting Pollard Walk data into population densities. A few species show no drop-off in detectability and most require only a modest correction factor, but for the least detectable species we estimate that 3/4 of individuals are not recorded.

4. Much of the variation among sites is explained by substantially higher detectability among sites in England than in Wales, which had different recorders. Biological traits have only limited explanatory power in distinguishing detectable vs undetectable species.

5. The variation in detectability is small compared with the variation in true abundance, such that population density estimates from the Pollard Walk are highly correlated with those derived from distance sampling.

Synthesis. These results are used to evaluate the robustness of the Pollard Walk for comparisons of abundance across species, across sites and over time. UKBMS data provide a good reflection of relative abundance for most species, and of large-scale trends in abundance. We also consider the practicalities of applying distance sampling to

butterfly monitoring in general. Distance sampling is a valuable tool for quantifying bias and imprecision, and has a role in surveying species of conservation concern, but is not viable as a wholesale replacement for simpler methods for large-scale monitoring of multispecies butterfly communities by volunteer recorders.

Introduction

Population abundance is a critical variable in ecology (McGill, 2006): abundance data are required to understand the basic population dynamics of species, as well as to provide information on the state of biodiversity (Loh et al., 2005). One of the largest datasets on non-pest insect population dynamics comes from the UK Butterfly Monitoring Scheme (UKBMS, Pollard & Yates, 1993; Rothery & Roy, 2001; Fox et al., 2006). The UKBMS has provided data on the abundance of butterfly populations for over three decades, and over 850 sites are now monitored annually (Botham et al., 2008). The methods developed for the UKBMS have been adopted by monitoring schemes in several other countries (van Swaay et al., 2008). Data from the UKBMS have provided valuable insights into the population-level effects of land-use and climate change (e.g. Roy & Sparks, 2000; Roy et al., 2001; Warren et al., 2001; Brereton et al., 2007; Oliver et al., 2009; Isaac et al., 2011). These findings, allied with certain aspects of butterfly biology (rapid life-cycle, microhabitat requirements), make butterflies a key indicator of environmental change (Thomas et al., 2004; Thomas, 2005).

The majority of UKBMS data are collected using a fixed-width transect count method, in which recorders count individual adult butterflies along set routes that are sub-divided into sections (Pollard et al., 1975; Pollard, 1977; Pollard & Yates, 1993). The method is known as the butterfly transect method or Pollard Walk: we use the latter to distinguish it from other transect-based methods. A key feature of the Pollard Walk is the imaginary moving box of 5 metres each side (250cm on both sides of the transect line): individuals observed within the box are counted whilst those outside are ignored. The method allows large quantities of data to be collected on butterfly communities, using simple rules that can be learned and adopted quickly. Intensive field studies have shown that counts from Pollard walks are closely correlated with absolute numbers of butterflies present, when the survey design representatively samples the site (e.g. Pollard, 1977; Thomas, 1983; Warren et al., 1984; Warren et al., 1986 ; Thomas, 1991; Sutcliffe et al., 1996; Haddad et al., 2008), but see Harker & Shreeve (2008).

Analyses of Pollard Walk abundance estimates generally assume that a constant proportion of the butterfly population is surveyed. This may be reasonable when comparing the same site year on year for a particular species, but may not hold true for comparisons between biotopes (e.g. Brown & Boyce, 1998; Haddad et al., 2008). In addition, vegetation changes associated with climate change or management regimes may affect the detectability of butterflies and generate apparent changes in abundance or mask real trends (Davies et al., 2006; Dennis & Sparks, 2006). Systematic changes in detectability over time would further reduce the degree to which abundance estimates are comparable, thus making it difficult to draw either theoretical or applied conclusions from such data. Moreover, potential differences exist in the visibility of different species (e.g. Thomas, 1983; Pollard & Yates, 1993). Dennis et al. (2006) found that visual apparency of British butterflies at a national scale is correlated with size, wing colour and flight behaviour. For this reason, there have been few attempts to use UKBMS data for interspecific comparisons of abundance (Cowley et al., 2001; Isaac et al., 2011). However, no methodological assessment has been conducted on the relative detectability of butterfly species at the biotope level.

Accurate population estimates with defined precision are increasingly being demanded in relation to the conservation of rare species and analyses of population viability and metapopulation dynamics. This is particularly true for species with low or fluctuating abundance and patchy or restricted distribution (Brown & Boyce, 1998; Boughton, 2000; Powell et al., 2007). One problem with the Pollard Walk is that it does not generate confidence intervals around individual estimates of abundance, so the precision of UKBMS data are unknown (see also Haddad et al., 2008). Thus, a critical evaluation of the bias and precision of the Pollard Walk is important for both fundamental and applied ecological questions.

The ecological literature contains many techniques for estimating absolute population size (e.g. Southwood & Henderson, 2000). Among the most widely used is distance sampling (Buckland et al., 2001; Thomas et al., 2010), which has been shown to give accurate and unbiased estimates of population density when not all individuals within a surveyed area are sampled. Distance sampling is usually transect-based, but can also be applied to point counts. It works by fitting a detection function to observations at known distances. The shape of this function defines the effective strip width (ESW), which provides a simple measure of detectability. ESW is the distance at which the number of individuals observed further away is estimated to equal the number of individuals closer to the line that were missed. Population density can be calculated as the number of individuals counted divided by [ESW * 2 *distance travelled]. The published literature contains few applications of distance sampling to butterflies (Brown & Boyce, 1998), and none in the context of validating monitoring data (see Newson et al., 2008 for an application to bird monitoring).

The key challenge we address here is the extent to which the relative abundance estimates derived from the Pollard Walk are comparable among species and among sites. We use distance sampling to estimate the variation in detectability of butterflies on UKBMS transects and compare abundance estimates from the two methods. We explore the limitations of the Pollard Walk and address the potential for distance sampling as a tool in monitoring butterfly populations. Our inferences are based on estimates of the detection function within the Pollard Walk box on existing UKBMS transect routes, which do not represent a random sample of the landscape. We do not address the issue of survey design, which is paramount for obtaining unbiased estimates of animal abundance (Thomas et al., 2010). Our primary focus is on how distance sampling can inform the interpretation of data collected on existing UKBMS routes.

We address four specific research questions. First, for each species, what proportion of butterflies is missed by the Pollard Walk? Second, what is the magnitude of variation in detectability among species and sites? Third, to what extent is detectability explained by butterfly biology and biotope characteristics? Finally, how well correlated are Pollard Walk and distance-based estimates of population density?

Our results have potentially wide-reaching implications for butterfly monitoring. Converting Pollard Walk data into absolute abundances would greatly enhance the value of the data already collected, providing new opportunities for analyses of the viability of populations, and make the data amenable to studies of community ecology and macroecology. This is a great opportunity, both to enhance studies of past population changes and to increase the rigor of future monitoring in Europe and elsewhere (Haddad et al., 2008; Nowicki et al., 2008). Moreover, understanding detectability will inform new techniques for monitoring rare species of particular conservation concern, and in the wider countryside (Thomas, 1983; Roy et al., 2007; Nowicki et al., 2008; van Swaay et al., 2008).

Materials and Methods

Data collection

Fieldwork was carried out at nine sites in north Wales and four in southern England (table 1). Welsh field sites fall within an area of 35 km^2 on the Creuddyn Peninsula and Anglesey, and cover a range of biotopes (Cowley et al., 2001). Transects were laid out in order to sample representative habitats (following Pollard et al., 1975). Some transects followed existing footpaths, whilst others traversed open land. English sites were all UKBMS transect routes on south-facing chalk-grassland slopes with varying degrees of scrub invasion and grazing pressure: two are situated on the Dunstable Downs in Bedfordshire and two on the North Downs in Surrey.

The Welsh study was conducted by DMS between 18th May and 22nd September 1998 (n=2256 butterflies recorded). The English study was conducted by AW between June 28th and July 25th 2006 (n=3304). Perpendicular distances were estimated by eye to the nearest 10cm in the Welsh sites (0, 5, 15, 25 etc) and to the nearest 5cm at English sites. Both studies followed the UKBMS criteria for weather and time of day (Pollard & Yates, 1993). Butterflies were identified to species level, with the exception of Small and Green-veined whites (*Pieris rapae* and *P. napi*) and, in the English study, Small and Essex skippers (*Thymelicus sylvestris* and *T. lineola*), which could not be distinguished reliably in flight.

Analytical Approach

We expect that detectability varies systematically among species and sites. The nature of this variation is of primary interest (Question 2), but also means that neither the raw observations nor the derived strip widths can be considered mutually independent. For these reasons, we fitted separate detection functions to each site-species combination, pooling the data across visits, and analysed the resulting strip widths using linear mixed-effects models. This provides an effective means for partitioning and estimating the variance in detectability, but is not optimal for robustly estimating population density (Thomas et al., 2010). Our measures of population density (Question 4) should therefore be treated with caution. To test if this approach compromised our conclusions, we repeated our analysis on the factors associated with detectability (Question 3) using the Multiple Covariate Distance Sampling (MCDS) engine (Thomas et al., 2010). These results are presented in the Supporting Information.

Effective Strip Widths

We estimated effective strip widths (ESW) using Distance v6.0 (Thomas et al., 2010). We stratified our analyses by each unique combination of species and study site, after removing all combinations with fewer than 20 observations. Although this is smaller than the recommended minimum sample size of 60 (Thomas et al., 2010), we feel justified in using a smaller number because our aim is to explore variation in detectability, not the precise estimation of population density. This restricted dataset consists of 5363 observations on 17 species (50 site-by-species combinations, table 2). Preliminary analysis revealed that certain combinations contained a high proportion of observations on butterflies basking on the transect path, thus violating the assumption that organisms are randomly-positioned with respect to the transect line (Thomas et al., 2010). This phenomenon tends to give distance distributions that are strongly spiked at zero, resulting in underestimated detection functions. In order to circumvent this problem, we analysed grouped data, selecting an initial bin width broad enough to remove the apparent spike (Buckland et al., 2001) and, since there were no other heaping problems, simply using this width throughout to give ten equally spaced distance bins. In practice, binning the data in this way had little effect on the estimated ESW for most combinations (table S1) and the Pearson correlation between ESWs using binned and raw distances was 0.95. Similar estimates were produced using differing numbers of bins (figure S2).

For each combination, we sought the best description of the detection function by fitting the six models suggested by Thomas et al. (2010: uniform plus cosine/polynomial adjustments, half-normal plus cosine/hermite polynomial adjustments, hazard rate plus cosine/polynomial adjustments) and selecting models in terms of goodness-of-fit statistics and AIC (Akaike's Information Criterion), following visual inspection of the data. Distance sampling data are generally truncated at some specified distance, in order to reduce the influence of outliers (Thomas et al., 2010). We generated two sets of ESWs using different truncations: one truncated at the 95th distance percentile for each combination (following Thomas et al., 2010), and one with a universal truncation distance of 250cm from the transect line (to give the width of the standard Pollard Walk box: 37% of observations were made at >250cm). The full set of ESWs is presented in the Supporting Information (table S1). We used the 250cm truncation to estimate species-specific correction factors for the UKBMS (Question 1). We used both sets of data to explore the variation in detectability (Question 2), the factors explaining detectability (Question 3) and compare estimates of population density (Question 4).

Statistical Modelling

We used linear mixed-effect models to partition the variance in ESW between sites and species and to test hypotheses about detectability. All analyses were conducted using the *lme4* package (Bates et al., 2008) in *R* (R Development Core Team, 2008). We weighted each of the 50 ESWs by the square root of the number of observations inside the truncation distance, rescaled to have a mean of 1. Weighting the data in this manner reduces the impact of combinations with small sample sizes, where ESW is likely to have been estimated imprecisely. Visual inspection of the residual distribution indicated that input variables did not require transformation, although each variable was centred on zero for modelling.

We first estimated species-specific ESWs using a model with Species as a fixed effect. These values were converted into correction factors (Question 1) by dividing them into the common truncation distance of 250cm. We then estimated the variance components (Question 2) by fitting models with random effects for Site and Species and no fixed effects. Finally, we tested a suite of hypotheses about differences in detectability among species and sites (Question 3). We used two site traits and three species traits to test these hypotheses. The site traits were Study (England vs Wales) and vegetation height measured from 1 (short grass) to 6 (high scrub, see table 1). Species traits were wingspan (in mm), bask mode (dorsal vs lateral) and colour measured from 1 (dull) to 5 (very bright), all using data presented in Dennis et al. (2006). We modelled vegetation height and colour as continuous variables. We fitted all main effects and first-order interaction terms, and then sequentially removed non-significant terms to arrive at a minimum adequate model (MAM). Significance of fixed effects was estimated by sampling 10,000 times from the posterior distribution of the fitted parameters using Markov Chain Monte Carlo methods (Bates et al., 2008).

Butterfly Population Density

We made three estimates of butterfly population density (ha^{-1}) for each site-species combination (Question 4), using a) Pollard Walk data (i.e. assuming no variation in detectability), b) distance sampling based on the 250cm truncation, and c) distance sampling based on the 95% truncation. We did not calculate confidence limits on the density estimates derived from distance sampling because our data were collected on a single transect at each site, thereby making it impossible to estimate variation in the encounter rate (Thomas et al., 2010). In addition, several combinations showed no measurable drop-off in detectability: ESW for these combinations is estimated to equal the truncation distance with zero error, in spite of the small sample sizes involved (table S1).

Results

Detection distances ranged from 0 – 1430 cm, with a median of 182 and a mean of 223 cm (figure 1). Across the 50 site-species combinations, the median ESW is 300cm for the 95% truncation and 164cm for the 250cm truncation (see Supporting Information for the full set of ESWs). These data suggest that 1-164/250 ≈ 1/3 of all individuals within the Pollard Walk box were missed.

Species-level ESWs (figure 2) range from under 60cm up to the truncation distance of 250cm, and fall into three clear groups. One group consists four highly detectable species (Brimstone, Large White and Large Skipper and Small/Essex Skippers) for whom little or no correction factor is needed (i.e. the Distance model indicates effectively no measurable drop-off in detectability within 250cm). Another group contains two species (Dingy Skipper and Brown Argus) with extremely short ESWs, suggesting that only around 25% of individuals are detected. The remaining 11 species show a moderate drop-off in detection (135cm < ESW < 210cm), and for whom a modest correction factor (1.2-1.9) would be appropriate (table 3). For nine of these intermediate species, the estimated ESW is significantly shorter than the Pollard Walk truncation of 250cm (figure 2).

Despite these differences, species identity contributes only a small portion of the variance in detectability *within* the Pollard Walk box. Just 7% of the variance is among species, compared with 29% among sites and 65% residual error. However, the picture is quite different when observations beyond 250cm are considered (i.e. using the 95% distance truncation): variance among species in detectability contributes 52% of the total, with 35% among sites and 14% due to residual error. This difference between our two sets of ESWs reflects the fact that strip widths cannot be larger than the truncation distance, and that some species with large ESWs (notably the Large White) have few observations within 250cm of the transect line (and therefore low weight). The total variance among the 50 ESWs is six times greater using the 95% set than using the 250cm truncation.

The minimum adequate models of detectability (table 4) reveal that much of the variation among sites is attributable to study: ESWs in the Welsh study were much shorter than for sites in England. Other correlates depend on the choice of truncation distance used. Within the Pollard Walk box, the only other significant correlate of detectability is the interaction between study and wingspan: each millimetre increase in butterfly wingspan leads to an increase in ESW of around 4cm in Wales, but had no significant effect among English sites. Using the 95% truncation, we find that colourful species are easier to detect: the fitted difference in ESW between the dullest species (colour=1) and the brightest (colour=5) is nearly three metres. We found small but non-significant positive relationships between size and detectability ($p \sim 0.07$) and the interaction between colour and wingspan ($p \sim 0.06$): each millimetre increase in butterfly wingspan leads to an increase of 8.5cm in ESW for the brightest species but no increase for dull species. Detectability does not differ between species that bask dorsally versus those basking laterally, nor does it correlate with our index of vegetation height. Broadly similar results were obtained using the MCDS engine (table S2), which suggests the 'proportion missed' within the Pollard Walk box is 33% in Welsh sites, compared with just 11% in English ones.

In spite of the differences in detectability we have observed, the population density estimates derived from distance sampling and the Pollard Walk are broadly similar (figure 3). The Pollard Walk densities tend to be under-estimates because they do not take into account any drop-off in detectability. Pollard Walk densities are more tightly correlated with density estimated from the 95% truncation (r^2=0.933) than from the 250cm truncation (r^2=0.789). This is surprising, because of the higher variance in ESW for the 95% truncation, and because the densities based on Pollard Walk and the 250cm truncation use the same numerator (number butterflies) in the density estimate. Although the overall correlation is high, the degree of under-estimation is extreme in a minority of cases: around 10% are underestimated by a factor of 3 or worse (dotted line on figure 3). The relationship between Pollard Walk and distance estimates of density is somewhat triangular: the mean discrepancy between the two estimates is greater at high density. Naïve interpretation might suggest that populations occurring at low density tend on average to be more detectable, and that the Pollard Walk is therefore less biased for rare than common populations. However, this phenomenon is almost certainly an artefact of excluding combinations with small sample sizes: low density populations that are difficult to detect would not generate enough data to be considered, whereas high density populations with similarly low detectability would show up as poorly-estimated by the Pollard Walk.

Discussion

Our results reveal that a sizeable proportion of butterflies are missed by the Pollard Walk, and that detectability (the proportion missed) varies substantially among species and sites. Whilst previous studies have reported variation in detectability of butterflies among species (Kery & Plattner, 2007) and biotopes (Brown & Boyce, 1998), ours is the first to quantify, compare and model them.

Overall, the variability among species in detectability is large. However, most of this disappears if observations outside the Pollard Walk box are excluded. This means that UKBMS data provide a good reflection of relative abundance for most species. Our species-specific correction factors (table 3) estimate the degree to which different species are under-recorded at the 'average' UKBMS site (but see below). We stress these are preliminary estimates based on relatively few sites and, in some cases, on small sample sizes. Notwithstanding these caveats, the numbers suggest that several species of UK conservation concern are being systematically under-recorded: the Dingy skipper, Grayling and Silver-studded blue (see also Dennis & Sparks, 2006) are all priorities on the UK Biodiversity Action Plan and are among the least detectable of the 17 species studied here (figure 2). Among species, both colour and size have limited power in explaining detectability, although the relative position of most species on this gradient of detectability is not surprising: the Dingy Skipper and Grayling are both well-camouflaged and known to be difficult to spot, whilst the three pierid species are all highly conspicuous.

Site effects are at least as important as species identity in determining detectability. Within the Pollard box, variance in detectability is much greater among sites than among species, which suggests that any correction factor applied to UKBMS data should be biotope-specific as well as species-specific. Our variance components model predicts the correction factor for the 'average' species to be in the range 1.1 – 2.5 for 95% of sites; comparable prediction intervals for species at the 'average' site are 1.3 – 1.9. This suggests that UKBMS data might not be especially reliable for comparing butterfly abundance between sites in individual years. However, the 2.5-fold variation in detectability remains small compared with the 100-fold variation in estimated abundance that is typical of species on the UKBMS (Thomas et al., 2011). The importance of the site effect is evident in the left-hand panel of figure 3: most of the severely under-estimated population densities are found at just a few sites (principally the Dulas Valley sites). The lack of significant relationship between vegetation height and detectability suggests that sites differ in ways that are not captured by our simple index, especially since butterfly behaviour varies among biotopes (Dennis, 2004) in ways that have unpredictable consequences for detectability. The strongest pattern in detectability is that detection distances in at Welsh sites were substantially shorter than in English ones. This could be explained by the coastal location and therefore higher wind speeds in Wales (wind makes it difficult to identify butterflies, especially in flight). However, the studies were conducted on different butterfly species at different times and by different observers. We can reject the effect of species composition, since the regional difference is pronounced among four of the five species shared between English and Welsh study sites (figure 4). The survey year is potentially confounding, because the English data were collected during an extremely hot summer (2006), whilst the Welsh study was conducted during a relatively poor year for butterflies (1998). The UKBMS minimum weather conditions (Pollard & Yates, 1993) were observed during both studies presented here, but it is likely that variation in weather above these minima exert a strong influence on butterfly behaviour that have knock-on effects for detectability (Dennis & Sparks, 2006; Wikstrom et al., 2009). The final complication is that two different observers collected the data. Both observers received suitable training, and it seems unlikely that differences in their ability to identify butterflies and estimate distances can account for the much larger ESWs at sites in England. Variation among observers presents greater problems for extrapolating our results to the wider question of detectability. Both our observers were relatively naïve: more experienced recorders might have an established search image of species of particular conservation concern, even if they are difficult to see. Such experience almost certainly increases the detectability of species with distinctive flight patterns (e.g. dingy skipper), but also presents an extra source of variation. Variation among recorders therefore deserves further consideration (Kery & Plattner, 2007; Nowicki et al., 2008), possibly by observing a range of recorders surveying the same sites. The importance of intraspecific variation in detectability means that untangling these multiple causal factors is a priority for future research in this area.

We found tight correlations between densities estimated using the Pollard Walk and distance sampling (figure 3). This is because variation in detectability, whilst substantial, is small compared to the huge variation in population density across sites and species (c.f. Thomas et al., 2011). However, for some populations the Pollard Walk gives a substantial underestimate. Thus, it would be unwise to treat Pollard Walk data as absolute estimates of abundance without considering the factors correlated with detectability. Most existing applications of butterfly monitoring scheme data are based on trends over time within populations (Roy & Sparks, 2000; Roy et al., 2001; Warren et al., 2001; Brereton et al., 2007): the key question here is whether the variation in detectability within populations is of comparable magnitude to real changes in population size. The widely-reported trends in butterfly abundance (e.g. Fox et al., 2006) might be compromised if biotopes themselves had changed in a consistent way across the country over the period of monitoring, thus leading to a systematic trend in detectability. National scale trends are probably quite robust, given that species declining on the UKBMS tend also to have shrinking distributions (Warren et al., 2001; Thomas, 2005; Fox et al., 2006), but individual site-level trends might be less precise. Long-term vegetation change might conceivably increase detectability (making it harder to detect declines in abundance) or decrease it (making it appear that stable populations are in fact declining). We suspect that the observed inter-site variation is far greater than the likely range for any one site, even under the combined effects of ecological succession, management, weather and climate change. However, this unanswered question could be addressed by a combination of monitoring detectability at reference sites and controlled experiments that manipulate biotope structure in realistic ways. Such focussed research should use MCDS (Thomas et al., 2010) rather than the stratified approach employed in this study.

Although we have demonstrated the value and potential of distance sampling in butterfly monitoring, there are both practical and theoretical considerations that make distance sampling unviable as an alternative to the Pollard Walk for large-scale multi-species monitoring. The practical issue is the potentially large number of butterflies occurring in peak season, when it is commonplace to record a butterfly every second. The effort of keeping separate counts for each species is so intense that it would be impossible to record distance estimates for each observation, even in the wider countryside. More fundamentally, most animals tend to be observed in flight, which violates one of the key assumptions of distance sampling (but see Buckland et al., 2001 p198). In addition, UKBMS routes do not sample habitat randomly, either at small spatial scales (many routes follow linear features or public rights of way) or large (sites tend to be selected because they contain abundant populations), leading to biased estimates of population density (either from distance sampling or the Pollard Walk). Our detection function models were hampered by the fact that several transects followed paths, which provide warm microclimates that attract aggregations of basking butterflies, thus violating another key assumption of distance sampling. Unfortunately, it would be impractical and undesirable to relocate traditional UKBMS transects to be more representative without breaking the continuity of >3 decades continuous monitoring that is the major strength of the scheme. The UK monitoring has recently been extended through a complementary scheme, the Wider Countryside Butterfly Survey, that samples a stratified-random selection of survey locations (1km grid squares) across the UK (Roy et al., 2007; Brereton et al., 2011). Although the wider countryside scheme addresses the bias towards sampling high abundance sites, it still involves routes that follow linear features or public rights of way. In spite of these reservations, we suggest that distance sampling, particularly MCDS, has two important roles in butterfly monitoring. One is to conduct intensive studies on a species-by-species basis, in order to refine our estimates of detectability and quantify the importance of variation due to biotope, management conditions, weather conditions, observer, butterfly behaviour (perched versus flying) and sex (across the three species of Blue butterflies, 90% of observations were on males). The second is to conduct targeted surveys and monitoring in relatively open biotopes, where trained observers can collect data outside the 250cm of the Pollard Walk box. This approach would be especially suitable for species of high conservation concern (e.g. Large Blue and High Brown Fritillary), where absolute abundance estimates may be important for conservation and research.

The work described here is not the final word on detectability of butterflies on transects, but provides an important step in testing the robustness of Pollard Walk data (see also Haddad et al., 2008; Nowicki et al., 2008). Monitoring schemes like the UKBMS are increasingly being used to address questions about global change (de Heer et al., 2005). Validation of these data, using well-established ecological methodology, is therefore essential for delivering policy objectives for biodiversity, both nationally (Sutherland et al., 2006) and internationally (Dobson, 2005). With this in mind, we hope that our work will provoke new enquiry into methodological questions about biodiversity change and contribute to the development of more rigorous standards in applied ecology and conservation (Sutherland et al., 2004; Stewart, 2010).

Acknowledgements

We are grateful to Douglas Bates, Steve Buckland & Stephen Freeman for statistical advice, to Marc Botham, Tom Oliver and especially David Roy for comments on previous versions of the manuscript, and to four anonymous reviewers for providing insights and constructive criticism that lead to substantial improvements. Gail Jeffcoate and Emily Brennan helped with logistical arrangements for the English field sites. Christine Weddle, Peter Curnock and other Butterfly Conservation volunteers provided assistance and advice in the field. NJBI is supported by a NERC Fellowship (NE/D009448/2) and DMS is supported by a NERC Advanced Fellowship (NE/D009979/2).

References

Bates, D.M., Maechler, M., & Dai, B. (2008) lme4: Linear mixed-effects models using S4 classes, Vol. version 0.999375-28.

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*. Centre for Ecology & Hydrology, Wallingford.

Boughton, D.A. (2000) The dispersal system of a butterfly: A test of source-sink theory suggests the intermediate-scale hypothesis. *American Naturalist*, **156**, 131-144.

Brereton, T.M., Cruickshanks, K.L., Risely, K., Noble, D.G., & Roy, D.B. (2011) Developing and launching a Wider Countryside Butterfly Survey across the United Kingdom *Journal of Insect Conservation*, **15**, in press.

Brereton, T.M., Warren, M.S., Roy, D.B., & Stewart, K. (2007) The changing status of the Chalkhill Blue butterfly *Polyommatus coridon* in the UK: the impacts of conservation policies and environmental factors. *Journal of Insect Conservation*, **12**, 629-638.

Brown, J.A. & Boyce, M.S. (1998) Line transect sampling of Karner blue butterflies (*Lycaeides melissa samuelis*). *Environmental and Ecological Statistics*, **5**, 81-91.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., & Thomas, L. (2001) *Introduction to Distance Sampling* Oxford University Press, Oxford.

Cowley, M.J.R., Thomas, C.D., Roy, D.B., Wilson, R.J., Leon-Cortes, J.L., Gutierrez, D., Bulman, C.R., Quinn, R.M., Moss, D., & Gaston, K.J. (2001) Density-distribution relationships in British butterflies. I. The effect of mobility and spatial scale. *Journal of Animal Ecology*, **70**, 410-425.

Davies, Z.G., Wilson, R.J., Coles, S., & Thomas, C.D. (2006) Changing habitat associations of a thermally constrained species, the silver-spotted skipper butterfly, in response to climate warming. *Journal of Animal Ecology*, **75**, 247-256.

de Heer, M., Kapos, V., & ten Brink, B.J.E. (2005) Biodiversity trends in Europe: development and testing of a species trend indicator for evaluating progress towards the 2010 target. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **360**, 297-308.

Dennis, R.L.H. (2004) Butterfly habitats, broad scale biotope affiliations and structural exploitation of vegetation at finer scales: the matrix revisited. *Ecological Entomology*, **29**, 744-752.

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. (2006) When is a habitat not a habitat? Dramatic resource use change under differing weather conditions for the butterfly *Plebejus argus*. *Biological Conservation*, **129**, 291-301.

Dobson, A. (2005) Monitoring global rates of biodiversity change: challenges that arise in meeting the Convention on Biological Diversity (CBD) 2010 goals. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **360**, 229-241.

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

Haddad, N.M., Hudgens, B., Damiani, C., Gross, K., Kuefler, D., & Pollock, K. (2008) Determining optimal population monitoring for rare butterflies. *Conservation Biology*, **22**, 929-940.

Harker, R.J. & Shreeve, T.G. (2008) How accurate are single site transect data for monitoring butterfly trends? Spatial and temporal issues identified in monitoring *Lasiommata megera*. *Journal of Insect Conservation*, **12**, 125-133.

Isaac, N.J.B., Girardello, M., Brereton, T.M., & Roy, D.B. (2011) Butterfly abundance in a warming climate: patterns in space and time are not congruent. *Journal of Insect Conservation*, **15**, in press.

Kery, M. & Plattner, M. (2007) Species richness estimation and determinants of species detectability in butterfly monitoring programmes. *Ecological Entomology*, **32**, 53-61.

Loh, J., Green, R.E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V., & Randers, J. (2005) The Living Planet Index: using species population time series to track trends in biodiversity. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **360**, 289-295.

McGill, B.J. (2006) A renaissance in the study of abundance. *Science*, **314**, 770-772.

Newson, S.E., Evans, K.L., Noble, D.G., Greenwood, J.J.D., & Gaston, K.J. (2008) Use of distance sampling to improve estimates of national population sizes for common and widespread breeding birds in the UK. *Journal of Applied Ecology*, **45**, 1330-1338.

Nowicki, P., Settele, J., Henry, P.Y., & Woyciechowski, M. (2008) Butterfly monitoring methods: The ideal and the real world. *Israel Journal of Ecology & Evolution*, **54**, 69-88.

Oliver, T., Hill, J.K., Thomas, C.D., Brereton, T., & Roy, D.B. (2009) Changes in habitat specificity of species at their climatic range boundaries. *Ecology Letters*, **12**, 1091-1102.

Pollard, E. (1977) A method for assessing changes in the abundance of butterflies. *Biological Conservation*, **12**, 115-134.

Pollard, E., Elias, D.O., Skelton, M.J., & Thomas, J.A. (1975) A method of assessing the abundance of butterflies in Monks Wood National Nature Reserve England in 1973. *Entomologist's Gazette*, **26**, 79-88.

Pollard, E. & Yates, T.J. (1993) *Monitoring butterflies for ecology and conservation* Chapman and Hall, London.

Powell, A., Busby, W.H., & Kindscher, K. (2007) Status of the regal fritillary (Speyeria idalia) and effects of fire management on its abundance in northeastern Kansas, USA. *Journal of Insect Conservation*, **11**, 299-308.

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rothery, P. & Roy, D.B. (2001) Application of generalized additive models to butterfly transect count data. *Journal of Applied Statistics*, **28**, 897-909.

Roy, D.B., Rothery, P., & Brereton, T. (2007) Reduced-effort schemes for monitoring butterfly populations. *Journal of Applied Ecology*, **44**, 993-1000.

Roy, D.B., Rothery, P., Moss, D., Pollard, E., & Thomas, J.A. (2001) Butterfly numbers and weather: predicting historical trends in abundance and the future effects of climate change. *Journal of Animal Ecology*, **70**, 201-217.

Roy, D.B. & Sparks, T.H. (2000) Phenology of British butterflies and climate change. *Global Change Biology*, **6**, 407-416.

Southwood, T.R.E. & Henderson, P.A. (2000) *Ecological Methods*, Third edn. Blackwell Science, Oxford.

Stewart, G. (2010) Meta-analysis in applied ecology. *Biology Letters*, **6**, 78-81.

Sutcliffe, O.L., Thomas, C.D., & Moss, D. (1996) Spatial synchrony and asynchrony in butterfly population dynamics. *Journal of Animal Ecology*, **65**, 85-95.

Sutherland, W.J., Armstrong-Brown, S., Armsworth, P.R., Brereton, T., Brickland, J., Campbell, C.D., Chamberlain, D.E., Cooke, A.I., Dulvy, N.K., Dusic, N.R., Fitton, M., Freckleton, R.P., Godfray, H.C.J., Grout, N., Harvey, H.J., Hedley, C., Hopkins, J.J., Kift, N.B., Kirby, J., Kunin, W.E., Macdonald, D.W., Marker, B., Naura, M., Neale, A.R., Oliver, T., Osborn, D., Pullin, A.S., Shardlow, M.E.A., Showler, D.A., Smith, P.L., Smithers, R.J., Solandt, J.L., Spencer, J., Spray, C.J., Thomas, C.D., Thompson, J., Webb, S.E., Yalden, D.W., & Watkinson, A.R. (2006) The identification of 100 ecological questions of high policy relevance in the UK. *Journal of Applied Ecology*, **43**, 617-627.

Sutherland, W.J., Pullin, A.S., Dolman, P.M., & Knight, T.M. (2004) The need for evidence-based conservation. *Trends in Ecology and Evolution*, **19**, 305-308.

Thomas, J.A. (1983) A quick method for estimating butterfly numbers during surveys. *Biological Conservation*, **27**, 195-211.

Thomas, J.A. (1991). Rare species conservation: case studies of European butterflies. In *The scientific management of temperate communities for conservation* (eds I.F. Spellerberg, F.B. Goldsmith & M.G. Morris), pp. 149–197. Blackwell, Oxford.

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.

Thomas, J.A., Simcox, D., & Hovestadt, T. (2011) Evidence based conservation of butterflies. *Journal of Insect Conservation*, **15**, in press.

Thomas, J.A., Telfer, M.G., Roy, D.B., Preston, C.D., Greenwood, J.J.D., Asher, J., Fox, R., Clarke, R.T., & Lawton, J.H. (2004) Comparative losses of British butterflies, birds, and plants and the global extinction crisis. *Science*, **303**, 1879-1881.

Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley, S.L., Bishop, J.R.B., Marques, T.A., & Burnham, K.P. (2010) Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology*, **47**, 5-14.

van Swaay, C.A.M., Nowicki, P., Settele, J., & van Strien, A.J. (2008) Butterfly monitoring in Europe: methods, applications and perspectives. *Biodiversity and Conservation*, **17**, 3455-3469.

Warren, M.S., Hill, J.K., Thomas, J.A., Asher, J., Fox, R., Huntley, B., Roy, D.B., Telfer, M.G., Jeffcoate, S., Harding, P., Jeffcoate, G., Willis, S.G., Greatorex-Davies, J.N., Moss, D., & Thomas, C.D. (2001) Rapid responses of British butterflies to opposing forces of climate and habitat change. *Nature*, **414**, 65-69.

Warren, M.S., Pollard, E., & Bibby, T.J. (1986) Annual and long-term changes in a population of the wood white butterfly *Leptidea sinapis*. *Journal of Animal Ecology*, **55**, 707-719.

Warren, M.S., Thomas, C.D., & Thomas, J.A. (1984) The Status of the Heath Fritillary Butterfly *Mellicta athalia* Rott in Britain. *Biological Conservation*, **29**, 287-305.

Wikstrom, L., Milberg, P., & Bergman, K.O. (2009) Monitoring of butterflies in semi-natural grasslands: diurnal variation and weather effects. *Journal of Insect Conservation*, **13**, 203-211.

Table 1: Description of study sites. L is the transect length (in metres), N is the number of walks and VH is the index of vegetation height.

Location	Coordinates	Name	Description	L	N	VH
Dulas Valley, Conwy	53°16'49"N 3°38'25"W	DV1	Lots of low *Cotoneaster*, some scrub and grasses, quite open	306	18	2
		DV2	Mix of longer grasses and open turf, some scrubby vegetation	255	11	4
Glan Conwy, Conwy	53°16'33"N 3°47'51"W	GC	RSPB reserve. Open, grasses and herbs, tall in places.	640	11	3
Great Orme, Llandudno	53°19'45"N 3°51'12"W	GOI	Mix of *Rubus* scrub and grassland, generally quite scrubby	160	14	4
		GO2	Short, close cropped turf, very open	445	5	1
		GO3	Short, close cropped turf, open with scrub	515	3	1
Newborough Warren, Anglesey	53°10'37"N 4°22'40"W	NW1	Taller grasses and herbs	1375	3	5
		NW2	Mix of open turf and longer grasses	515	3	3
Llangwstenin, Conwy	53°17'46"N 3°46'17"W	LST	Quite tall woody scrub	330	3	6
Bison Hill, Dunstable	51°51'44"N 0°32'45"W	BH	SSSI. Thick grass. Ungrazed, mown annually. Sward height ~70cm	620	5	3
Whipsnade Zoo, Dunstable	51°51'07"N 0°33'05"W	WZ	Heavily grazed by wallabies and Chinese water deer. Sward height <5cm	1450	5	1
Pewley Downs, Guilford	51°13'48"N 0°33'24"W	PD	Grass with some scrub invasion. Ungrazed, but mown annually. Sward height ~75cm	630	5	4
Denbies Landbarn, Dorking	51°14'17"N 0°22'35"W	DL	Grazed by ponies. Sward height ~35cm	740	5	2

Table 2: Number of butterflies recorded for each species-site combination. Combinations with fewer than 20 observations were excluded. Site names as in table 1. *The Essex skipper does not occur in North Wales.

	Welsh Sites									English Sites			
	DV1	DV2	GC	GO1	GO2	GO3	NW1	NW2	LST	BH	WZ	PD	DL
PIERIDAE													
Brimstone												55	
Large White										28		29	
Small/Green-veined White										24		42	
LYCAENIDAE													
Brown Argus	46	32		53									
Chalkhill Blue												390	352
Common Blue			36										
Silver-studded Blue	47	62		399									
NYMPHALIDAE													
Gatekeeper	50	38	216	58						75		78	27
Grayling					71	190							
Marbled White										260		96	185
Meadow Brown		38	299	96			22	20		254	50	345	225
Ringlet							63	102		154		66	23
Small Heath		28											137
Speckled Wood									24				
HESPERIIDAE													
Dingy Skipper	62	71											
Large Skipper												30	
Small/Essex Skipper			67*							185		35	28

Table 3: Species-specific ESWs, associated standard errors (SE) and correction factors (CF) for the 250cm truncation. Figures are fitted values from a linear mixed-effects model (see text for further details).

Family	Common name	Latin name	ESW/cm	SE	CF
Pieridae	Brimstone	*Gonepteryx rhamni*	250.0	51.9	1
	Large White	*Pieris brassicae*	250.0	53.0	1
	Small/Green-veined White	*Pieris sp.*	198.2	46.2	1.26
Lycaenidae	Brown argus	*Aricia agestis*	63.7	27.6	3.92
	Chalkhill Blue	*Polyommatus coridon*	198.6	22.2	1.26
	Common Blue	*Polyommatus icarus*	141.1	51.4	1.77
	Silver-studded Blue	*Plebejus argus*	145.3	21.4	1.72
Nymphalidae	Gatekeeper	*Pyronia tithonus*	182.6	16.9	1.37
	Grayling	*Hipparchia semele*	135.8	26.4	1.84
	Marbled White	*Melanargia galathea*	199.7	23.2	1.25
	Meadow brown	*Maniola jurtina*	160.2	13.2	1.56
	Ringlet	*Aphantopus hyperantus*	206.9	20.3	1.21
	Small Heath	*Coenonympha pamphilus*	169.3	31.9	1.48
	Speckled Wood	*Pararge aegeria*	163.7	56.9	1.53
Hesperiidae	Dingy skipper	*Erynnis tages*	56.2	30.1	4.45
	Large Skipper	*Ochlodes sylvanus*	250.0	55.4	1
	Small/Essex skipper	*Thymelicus sp.*	232.8	22.1	1.07

Table 4: Parameters from the minimum adequate model of the variability in detectability among species and sites (n=50 combinations). P-values were estimated by sampling 10,000 times from the posterior distribution of the fitted parameters using Markov Chain Monte Carlo methods.

	95% truncation			250cm truncation		
	Estimate	SE	p	Estimate	SE	p
Intercept	428.8	30.2	<0.0001	209.8	10.9	<0.0001
Study (Wales)	-172.8	33.2	<0.0001	-66.2	14.5	<0.0001
Colour	73.4	19.2	0.0002			NS
Wingspan			~0.07	-1.14	1.12	0.28
Study:Wingspan			NS	3.20	1.46	0.036

Figure Legends

Figure 1. Histogram of detection distances among 5363 observations of butterflies on transects. The vertical bar indicates the edge of the Pollard Walk box, outside which butterflies are not counted.

Figure 2. Species-level strip widths (in cm) for data collected within the 250cm Pollard Walk box. Data are parameter estimates from a model of 50 site-species combinations with species as an explanatory variable. Error bars define the 95% confidence limits.

Figure 3. Comparison of population density (individuals per hectare) estimated by the Pollard Walk and distance sampling, using both a 250cm truncation (left panel) and the 95% truncation (right panel). Each symbol represents a different study site. The solid line indicates the 1:1 relationship that would be observed if populations were completely detectable. Dashed and dotted lines correspond to corrections factors of 2 and 3, respectively. Note log-log axes.

Figure 4. Box-and-whiskers plot showing variation among sites in effective strip widths (in cm) for species observed at sites in both England and Wales. Data derived from data in 10 equally-spaced bins after truncating at the 95% of observations for each site-species combination.