

Abstract

Species distribution, spatial, temperature, climate, land cover, carabid.

Introduction

 Climatic variables are widely used in models to predict species' distributions at relatively coarse resolutions and over large spatial extents, but no consensus yet exists about the role of climate in determining finer-resolution distributions approaching the scales that individuals actually experience (e.g. Pearson and Dawson 2003, Elith and Leathwick 2009). This uncertainty arises, in part, because there is a large difference between the spatial resolutions at which individuals experience the environment and those at which distributions and climate are usually represented and modelled. Most terrestrial invertebrates spend the majority of the non-dispersing phases of their lives in areas of less than a square metre to a square kilometre (see Brouwers and Newton 2009 for example dispersal rates). Even the most mobile species usually experience a potentially selective series of local environments, rather than the "average" conditions of a much larger area. In contrast, the geographic ranges of species are typically represented on distribution maps at resolutions that are three to fourteen orders of magnitude coarser than this. For example, grid resolutions of 10 km x 10 km to 100 km x 100 km are commonly used for the production of distribution atlases (e.g. Luff 1998), to help quantify recent distribution changes in response to climate change (e.g. Hickling et al. 2006), and for the purposes of modelling species' distributions and projecting the responses of species to climate change (e.g. Huntley et al. 2007, McKenney et al. 2007). Different resolution bioclimatic models give rise to quite different expectations of species' distributions and abilities to survive climatic change (e.g. Randin et al. 2009). Given this mismatch, it is unclear whether coarse- resolution analyses are sufficient to understand current distributions, or to project future potential changes and conservation priorities.

 Within a particular landscape, the background climate may be similar from place to place, and fine- resolution spatial variation in the distributions of species could be determined predominantly by non-climatic factors. These could include differences in vegetation cover and management (e.g.

 Eyre et al*.* 2003, 2005a), or the geological substrate and soil type (e.g. Blake et al. 2003). In reality, climatic and non-climatic factors are likely to combine to determine fine-resolution distributions. However, the relative importance of climate and vegetation is difficult to deduce, partly because the vegetation is itself often related to climate and partly because species can become restricted to favoured locations or habitats in regions where the background climate is marginal (Lennon et al. 2002). Species may select local conditions (e.g. sparse vegetation on south-facing slopes at northern range boundaries) that provide suitable microclimates for population growth (Thomas et al. 1999), or they may thermoregulate, enabling species to inhabit regions far outside the geographical limits that might otherwise be expected (Bryant et al. 2002). These behaviours can give rise to temperature-related geographic gradients of habitat associations (e.g. Oliver et al. 2009), and underlie shifts in the habitat affiliations of species following climatic warming (Thomas et al. 2001, Davies et al. 2006). Hence, even when non-climatic factors appear to be the most important determinants of the local distributions of species, it is difficult to evaluate whether it is the microclimatic conditions that these locations provide that are important, or whether other non- climatic differences are more important (Gottfried 1999, Gutiérrez Illán et al. 2010).

 Assessment of the relative importance of climatic *vs* non-climatic factors at different spatial scales requires the availability of both distributional and environmental data at equivalent resolutions. However, fine-resolution temperature surfaces have yet to be combined in the same analyses as land cover or habitat variables. This may in part be due to a lack of availability of fine-resolution climatic data for use in local scale studies, but the recent publication of a microclimate model (Bennie et al. 2008, 2010) enables us to investigate the relative importance of climatic and habitat variables at a finer resolution than has previously been explored.

 We utilise this microclimate model in relation to the distributions of a group of mainly predatory ground beetles (Carabidae) that should not be intrinsically limited to a particular type of vegetation.

 Habitat may still be limiting in the prey it provides and the manner in which its structure modifies the microclimate experienced (Cernusca and Seeber 1981). Carabids have previously been shown to have preferences for particular temperatures, both at a national scale (Eyre 2006) and within regions (Martinez et al*.* 2009). Other important determinants of their distributions will vary depending on species, but may include light levels, humidity, substrate types and possibly pH and salt levels (Thiele 1977), as well as soil moisture (Gardner et al*.* 1997, Eyre 2006). 350 species have been recorded from Britain and Ireland, with several species having Biodiversity Action Plans in Britain, and many more appearing on Red Data Book lists (Luff 1998).

 Here, we examine microclimatic variation and land cover type and their relationships with the distributions of twenty carabid species at 25 m² spatial resolution, in three areas of the British uplands (from here on local scale analyses). We compare these results with deductions based on 107 coarse-resolution (100 km² resolution) distribution models for the same species across the whole of Great Britain (from here on national scale analyses). We hypothesise that climate will be the most important determinant of species' distributions at the national scale, but that land cover will be important at the local scale. We also hypothesise that temperature will be more important to species at the edges of their range, where the background climate is likely to be marginal (e.g. Jump et al. 2010).

Methods and analysis

Field Sampling

Invertebrates were collected by pitfall trapping between the beginning of May and the end of

August 2008 at three sites: Lake Vyrnwy Royal Society for the Protection of Birds (RSPB) reserve

in Wales; Glen Finglas Woodland Trust reserve in Scotland; and the High Peak region in the Peak

121 District National Park in England (Fig. 1).

 Forty locations were sampled at Lake Vyrnwy and the Peak District, using a stratified random 143 sampling strategy based on sixteen categories defined on the basis of slope (shallow 0-10°, steep >

10°), aspect (North 315-45°, East 45-135°, South 135-225°, West 225-315°) and elevation (low <

475m, high > 475m). The sixteen categories were reduced to twelve by combining those that

covered a very small amount of the site. These twelve categories were equally represented, with the

addition of four sites in the lowest elevation areas. This strategy gave a comprehensive

 representation of the range of microclimates present at each site. Forty-eight locations were sampled at Glen Finglas, selected in the same way, the additional locations allowing better representation of the more topographically diverse terrain present. Sample locations were georeferenced in the field using a handheld Global Positioning System (Garmin GPS 60), and were at least 50 m apart to reduce spatial autocorrelation.

 At each sample location, five pitfall traps were deployed in a circle of 2 m diameter. Traps consisted of two standard plastic vending cups nested together and sunk into the soil so that the rim of the inner cup was flush with the soil surface. Each trap was filled to a 2.5 cm depth with ethylene glycol antifreeze. Traps were covered with a terracotta coloured plastic saucer suspended 11cm above the trap using galvanised wire. This served the dual purpose of limiting liquid loss by evaporation and limiting flooding from rainfall. A surround of chicken wire (mesh diameter 20 mm) aimed to exclude small mammals and livestock. Traps were emptied on a monthly basis, and the antifreeze solution was replaced as necessary. Invertebrates from all five pitfalls at each sample location were pooled then transferred to 100% ethanol for transport back to the lab. Carabids were sorted from other invertebrates and identified to species. The entire season's catch for each trapping location was then pooled for analysis to give a measure of the relative population density present at any one location; pooling data in this way should average out the effects of sample error and of extreme weather events on activity of individuals during different months of trapping (Baars 1979).

Microclimatic Modelling

 A recently published microclimatic model (Bennie et al. 2008) was used to predict the microclimate 171 across the three sites¹. Hourly radiation, wind speed and air temperature data were obtained for the

 $\frac{1}{1}$

recoded from a Bash script running under the open source GIS program GRASS to a standalone C++ program.

 meteorological stations recording radiation data that were closest to each of the study sites (Lake Vyrnwy, 52° 45′ 25″ N, 03° 38′ 45″ W, Manchester Hulme Library, 53° 28′ 01″ N, 02° 15′ 00″ W 174 and Strathallan Airfield, 56° 19′ 33″ N, 3° 43′ 44″ W). Choice of meteorological stations was limited by the requirement for radiation data, such that no other stations were close to the study sites. Digital Elevation Models (DEMs) with 25 m² horizontal resolution and 1 m vertical accuracy (from NERC's Earth Observation Data Centre, [www.neodc.rl.ac.uk\)](http://www.neodc.rl.ac.uk/) for the three sites were used to calculate slope and aspect values for each 5 m x 5 m cell using the "Slope" and "Aspect" functions in the Spatial Analyst package in ArcMap v 9.2 (ESRI 2008). The microclimate model adjusted the temperature values from the meteorological stations to take account of the elevation difference between the meteorological station and each 25 m² grid cell in the landscape, as well as differences in direct and indirect radiation associated with different slopes and aspects at different times of the day and year (including the effect of hill-shading by the surrounding landscape). Hourly predicted temperatures were generated for the top of the vegetation from the start of September 2007 until the end of August 2008. For the purposes of this study, we assume that these vegetation surfaces are correlated with the conditions experienced beneath the vegetation layer.

<u>.</u>

 From these hourly predictions, annual mean (ANMEAN), mean over the sampling period (May to August, MAMEAN), annual maximum (ANMAX) and annual minimum (ANMIN) temperatures were calculated for each trapping location. These variables were selected as they should represent the temperatures likely to be of importance in determining the distributions of a variety of species. Some might respond to average temperatures, such that a certain number of days above a certain temperature might be necessary for growth or development (Leirikh et al. 2009), and this might be especially important during the breeding (sampling) season. Others might be killed by extremes of temperature, such that very high maxima or very low minima in an area would preclude their presence (Bayram and Luff 1993, Somero 2010). Different temperature variables have been found

C++ code available on request from PKG, programme is named "Micromaker"

 to affect different aspects of carabid life histories (Althoff et al. 1994), and genders may also have different temperature preferences (Atienza et al. 1996) due to physiological differences inherent in the breeding period. The resolution corresponds to the scale at which ground beetles might realistically be able to disperse to more favourable locations before suffering damage (see Brouwers and Newton 2009).

Local Scale Modelling

 Species were selected for modelling at the local scale according to the following criteria: They were present at five or more sampling locations in a site, with at least ten individuals across the site, and were only modelled at the sites where these criteria were met. Abundance data were transformed to presence/absence data to enable meaningful comparison with the national data, as the strength and direction of relationships between species distribution and temperature can be affected by the type of response variable (Diez & Pulliam 2007). Presence/absence was the response variable in a Generalised Linear Model (GLM) with Binomial error structure and logit link function. Predictor variables were split into two sets: temperature and habitat. The temperature set comprised 213 ANMEAN, MAMEAN, ANMAX and ANMIN in °C. There were several habitat variables, depending on the number of habitat types present in the site (seven in Glen Finglas, five in Lake Vyrnwy and six in the Peak District). Presence or absence of each habitat type at the sampled locations were considered as separate factors (extracted from the Land Cover Map 2000 dataset from the Centre for Ecology and Hydrology (CEH), which is based on the dominant habitat within 25 x 25 m polygons). In addition, the number of land-cover categories present within a 100 m 219 diameter buffer around each point (NH) was calculated. To account for damage to traps caused by grazing livestock, a trapping effort variable was included as a log-offset in all models, corresponding to the proportion of the trapping effort that was lost to such damage at each location (e.g. loss of two of the five pitfall traps at a given location in one month). Linear and quadratic

 relationships were explored, and best models were built using a backwards stepwise procedure using the Akaike Information Criterion (AIC) as the basis for rejection or inclusion within each set of variables (temperature and habitat) in the BIOMOD package for R (Thuiller 2009). Area under the receiver operating characteristic curve (AUC) was calculated using a 70/30 data split, with 70% being used for training and the remaining 30% being used for testing. Residuals from the best models for each species were checked for spatial autocorrelation using Moran's I in ArcGIS.

National Scale Modelling

 To compare the perceived effects of different study scales, we used distribution records from the National Biodiversity Network (NBN) Gateway (http://data.nbn.org.uk), at 10 km x 10 km grid 234 resolution (i.e. 100 km^2 cells) for Great Britain. Long term average climatic variables for 1961-90 were obtained at 25 km² resolution from the UK Climate Impacts Programme (UKCIP http://www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09/), and summary percentage cover of the broad habitat categories at 1 km² resolution from the CEH Land Cover Map 2000 238 (CEH www.ceh.ac.uk/sections/seo/lcm2000 home.html). The resolutions were harmonised by summarising the climatic and land cover data to 100 km² resolution in ArcGIS, to match the distributional data; the mean of four cells was used for the temperature variables, and the total area covered for the habitat categories. Species' presences were taken as 100 km² grid cells where a given species had been recorded, using data from the NBN gateway from 1961-2009 to reflect the available climate data. Because sampling of carabid beetles in Britain is not exhaustive, a species may be truly present in some cells where it has not been recorded. To reduce this concern, "absences" were taken to be only those 100 km² grid cells that were known to have been sampled for ground beetles (i.e., other carabid species had been recorded from them, Fig. 2). For each species, we generated a response variable of presences and absences from the UK. We then used a GLM with a binomial error structure and logit link function fitted to two sets of variables, habitat

- 20, p<0.02). This negative correlation shows that northerly species tended to occur in cool locations within landscapes (negative regression coefficients), whereas more southerly species tended to be found in relatively warm places (positive regression coefficients).
-
- **National scale distribution modelling**
-

 Our analysis revealed that habitat models were better than temperature models at predicting the presence/absence of all species modelled (Fig. 3). In general, model fit was poor (13/20 species AUC < 0.7 for habitat models, 16/20 species AUC < 0.7 for temperature models). However, both 284 habitat (AUC > 0.8) and temperature models (AUC > 0.7) performed best for the four most northerly-distributed species (Supplementary materials Appendix 1, Table A4), suggesting that coarse resolution habitat and temperature models both perform better on species at the edges of 287 their range.

Discussion

 By using high resolution (25 m²) surfaces that reflect local thermal environments, the local distributions of species could frequently be explained in terms of climatic variables. We interpret this result cautiously because of the slightly coarser resolution of the habitat data, but also the higher number of habitat variables. Temperature appears to be as important as habitat as a predictor of the presence of species at finer resolutions. Northerly species were most often found in cool micro-sites within landscapes. Southerly species showed the reverse pattern, being found in warm environments within the study sites. However, the effects of climatic variables were not restricted to species at the thermal edges of their geographic distributions, and many "widespread" species were statistically associated with particular thermal environments within study sites. In contrast, land cover was always more important in analyses of species' distributions at the national scale.

 Our results, and those of Gutiérrez Illán et al. (2010) appear to run counter to the commonly-held perception that habitat and vegetation variables will be most important at a local scale, whereas climatic variables will predominate in apparent importance at more extensive scales (e.g. Pearson and Dawson 2003, but see Tingley and Herman 2009). Our explanation for this apparent discrepancy relates to the spatial resolutions and extents of the analyses. Fine resolution, within- landscape analyses of the distributions and abundances of species rarely consider the importance of microclimates on species' local distributions. Microclimates are recognised as being important towards the edges of species' ranges (Thomas et al. 1999), and in determining species' survival in the context of environmental stochasticity (e.g. Kindvall 1996, Oliver et al. 2010), and there is a large body of evidence showing the importance of temperature to the growth and survival of species (e.g. Bayram and Luff 1993, Leirikh et al. 2009, Somero 2010). In the absence of high resolution microclimatic data for inclusion in distribution models, local distributions of species are typically explained in terms of the variables that are commonly available at such resolution, such as resource, habitat and vegetation variables. The distributions of these "habitat" variables are often correlated with climatic variation within a site (different vegetation types are associated with different slopes, aspects and local elevations, e.g. Lakhani and Davis 1982), so a statistical association between a species and a particular habitat type cannot always be unambiguously ascribed to habitat or to microclimate (but see Huntley and Baxter 2003).

 This is further confounded by the fact that different habitats not only *occur* predominantly in different microclimates but that they also *generate* different microclimates through variation in vegetation density and height. For example, Schneider and Eugster (2005) found that converting peat wetlands to productive agricultural land reduced the temperature range in the study site by 0.6 325 °C. Similarly, Suggitt et al. (2011) found that the minimum temperature in woodlands was 4-6 °C warmer than in nearby heathlands and grasslands, and Thomas (1983) reported that short turf can

 generate 8 ºC higher surface temperatures than tall turf, within the same grassland. It is possible that, in the absence of appropriate high-resolution microclimatic data, much of the within-landscape variation in species' abundance and distribution is being mistakenly attributed to habitat-related factors, thereby underestimating the role of climatic variables (and hence climatic change) in affecting species' local distributions. Our analyses only considered microclimatic variation associated with topography, calculated at the top of the vegetation layer. However, beetles were trapped below the vegetation, so additional sources of microclimatic variation related to habitat type were not included, and we may still have underestimated the role of thermal variation. The fine- resolution vegetation structure data required to determine this indirect effect of land cover are not yet available for most areas, although the increasing availability and use of fine resolution LiDAR data may provide a solution (Müller and Brandl 2009).

 For the national scale analysis, local climatic variation associated with different topographies and vegetation types within 100 km² cells may be as large as, or greater than, the climatic variation 341 between cells (e.g. we observed a range of over 6 $^{\circ}$ C in monthly maximum temperatures between different locations within the Peak District site). This thermal variation may partly be captured by the frequencies of different land cover types; it is possible that our national scale analysis assigns greatest importance to land cover type simply because vegetation types provide proxies for the 345 range of local microclimatic conditions within each 100 km^2 grid cell. Likewise, including elevational range within coarse grid cells can improve the fit of distribution models because this reflects the likelihood that suitable local climates will be found somewhere within the grid cell (Luoto and Heikkinen 2008), since temperatures tend to decrease with increasing elevation. For widespread species, climate may contribute to variation in presence, but such variation is only weakly correlated with the smoothed 100 km² climatic surfaces that were available to use in the 100 km² resolution analysis of Britain.

 At yet coarser resolutions, land cover may appear to be less important (Luoto et al. 2007) because all grid cells contain a diversity of topographies (south- and north-facing slopes) and vegetation types (which vary in shadiness, for example), providing a wide range of local climates. Such coarse- resolution grids give the impression of solid distributions within the core parts of species' ranges, with a relatively "sharp" climatic margin, giving rise to the conclusion that climatic variables dominate distributions at very coarse resolution, particularly when analysed at continental extents. We are left with the somewhat uncomfortable conclusion that the perceived relative importance of climatic and non-climatic determinants of species' distributions may have as much to do with the nature of the data available, and the resolution and extent of the analysis, as with the real separate and combined effects of climatic versus non-climatic variables.

 More positively, finer-resolution microclimatic approaches, such as those adopted here, may begin to provide projections of species' responses to climatic change at a scale that is relevant to conservation planning. Thermally-driven shifts in species' vegetation associations (Davies et al. 2006) may enable us to identify ways of adapting management regimes to maintain suitable microclimates for species within sites. Microclimate models such as that used here can help identify heterogeneous landscapes and climatic refugia where species may be most buffered against extinction (e.g. Kindvall 1996, Oliver et al. 2010). Microclimatic modelling could also be used to improve assessments of conservation status by refining estimates of the area of land occupied by species, especially in mountainous regions, where many species of conservation concern occur. The conservation status of species is often determined, in part, by estimates of the extent of occurrence (from coarse resolution maps) or area of occupancy (the area actually occupied within the extent of occurrence, from finer resolution maps) (IUCN 2010). The coarse resolution extent of occurrence may include large areas that are unsuitable for a species, hence it is preferable to use area of occupancy criteria. However, this more accurate measure is unknown for 98% of all birds (Sekercioglu et al. 2008), which are a relatively well studied group. Fine-resolution models

ACKNOWLEDGEMENTS

in Britain. - Journal of Applied Ecology 20: 59-83.

- 579 Table 1: The abundance of the twenty selected species at each field site. Species present at a site,
- 580 but not fulfilling the modelling criteria, are designated by *.
- 581

Figure legends:

 Figure 1: a) Location of the three field sites (white circles) within Great Britain in relation to annual mean temperature (ºC) during the period 1961-90 at 5 km resolution; (b,c,d) Examples of the presence (black triangles) and absence (white circles) of individual species at the three study sites in relation to modelled annual mean temperature (ºC) at 5 m resolution. All temperature layers are shown on the same scale.

 Figure 2: The distribution of *Carabus glabratus* within Great Britain. Black squares are where the species has been recorded as present, white areas are where the species has not been recorded, but a record of at least one other carabid exists and grey squares are where no carabids of any species have been recorded. White circles show the locations of the three field sites, and distances are in kilometres.

 Figure 3: The relative importance of habitat and temperature variables to the twenty carabid species at the three sites and across Great Britain.

Figure 4: The relationship between the mean regression coefficient (from up to three sites) of

species presence/absence on modelled annual mean temperature (ANMEAN) and the mean latitude

of the GB distribution of the 20 selected Carabid species. For the purposes of illustration, all models

were constrained to use ANMEAN (the mean annual temperature in °C), despite other temperature

variables often being better predictors of the distributions of species. Labels denote codes for each

species; Af = *Agonum fuliginosum*, Al = *Amara lunicollis*, Ap = *Abax paralllelepipedus*, Ca =

Carabus arvensis, Cg = *C. glabratus*, Cp = *C. problematicus*, Cv = *C. violaceus*, Cc = *Cychrus*

caraboides, Lp = *Loricera pilicornis*, Neb = *Nebria brevicollis*, Nob = *Notiophilus biguttatus*, Pas =

Patrobus assimilis, Pad = *Pterostichus adstrictus*, Pae = *P. aethiops*, Pd = *P. diligens*, Pma = *P.*

madidus, Pme = *P. melanarius*, Pn = *P. nigrita*, PN = *P. niger*, Pv = *Poecilus versicolor*.

Lake Vyrnwy

Peak District

616 **Figure 3:**

