ssibilities of false entries. A set of reports helps to produce routine inquiries. All documents, including correspondence, are suitably archived at our institute, so that they can be found immediately if another detailed check seems appropriate. For the examination of earlier entries it is useful to have the data within easy reach.

At the end of the year, when all files have been checked and stored, some routine checks are made on the whole dataset. Subsequently, for each kilometre square two identical lists are produced. They contain the general information concerning the square, the name and address of the responsible fieldworker, the dates of the visits and the number of territories per species for the last 10 years. We send these lists to the fieldworkers and ask them to make a thorough check. Thanks to this procedure, erroneous or missing entries can be detected. At the same time the fieldworker is asked whether he or she will be able to survey the square again the following year. The fieldworker signs one of these lists, sends it back to the institute and keeps the other one.

4.3.2 Bird monitoring and spatial modelling of species distribution: an example from the Catalan common bird monitoring scheme

Lluís Brotons and Sergi Herrando

Long-term bird monitoring schemes (LTM) provide us with a great deal of spatial data that have the potential to be used to create maps showing changes in species distribution and abundance. One possibility is to obtain distribution maps from sample locations, by using surrogate environmental data and estimating species habitat preferences in spatial terms. Recent developments in numerical methods, and the increasing availability of remote census data sources have boosted the application of habitat-based models in ecology. This approach is based on the hypothesis that if species environmental associations can be robustly established, one may use them to estimate species distributions through the identification of suitable habitat in areas from which faunal data has not been recorded, but where environmental information is available.

Habitat modelling is being progressively extended but has not yet been widely applied to LTM data. Here, we briefly describe how to use data from a bird LTM scheme to obtain species distribution maps. We use data from the Catalan common bird survey as an example to produce habitat based maps and assess its predictive accuracy using independent data from recently completed atlas work (Brotons *et al.* 2007).

How to conduct habitat based models: the bird data

We used LTM data from the Catalan common bird survey (SOCC, from the Catalan "Seguiment d'Ocells Comuns a Catalunya"). The SOCC scheme started in

2002 and is based on a line transect approach, in which observers record all individuals of all bird species seen or heard on a 3-km transect divided into three 1-km sections. We used these 1-km sections as sampling units for modelling and mapping purposes (see www.ornitologia.org).

At present, 226 SOCC transects have been conducted for at least one year during the period 2002-2005 (Fig. 4.2). The mean number of available years (maximum of 4 years) per transect during this period was 3.01. Since the SOCC scheme is essentially based on volunteer ob- servers, the survey is constrained by the number of available sampling transects and is prone to poor spatial cover of remote areas, resulting in a regionally biased sampling distribution. We took into account heterogeneity in sampling effort by weighting sample locations according to their abundance in different subregions.

In the end, we were able include 99 species that appeared in at least 10 different transects during the 2002-2005 period.

In order to evaluate the predictive ability of habitat models conducted using SOCC data, we used species occurrence from the Catalan Breeding Bird Atlas (CBBA, Estrada *et al.* 2004). The CBBA is a large-scale survey that covered the whole of the Catalonia between 1999 and 2002, using a grid based 10-km Universal Transverse Mercator (UTM) squares. A sub-sample of a total of 3,077 1-km squares (approximately 9% of the total area) was selected to conduct standardised intensive surveys of species presence in a stratified fashion to cover the main habitat types present within each of the 10-km UTM squares.

How to conduct habitat based models: the environmental data

We used 39 environmental variables to build the model, which were generated from available digital layers such as land use and climatic maps or digital elevation models. All environmental variables were generated for each 1-km UTM square in Catalonia and for 1-km square buffers around the central point of each SOCC section. If possible, the environmental variables (i.e. land use maps), were estimated from different data sources so that they better matched the sampling periods of each of the surveys.

How to conduct habitat based models: the model

We conducted occupancy models using presence/absence data over the 2002-2005 period from SOCC transect sections (SOCC models), by means of generalised linear modelling with binomial error distribution (GLM), including as potential predictors in the model all linear and quadratic terms, and selected the most parsimonious model using the Akaike Information Criteria (AIC). Other modelling methods are currently available and easily implemented in customised software, depending on the type of bird data available and the complexity of the database (Maxent, Phillips *et al.* 2006, BIOMOD, Thuiller 2003).

Long term monitoring programs and spatial modelling: perspectives and applications

Overall, the accuracy of models estimated with the SOCC data performed better than random for all the species analysed. Furthermore, the evaluation of their predictive accuracy on independent atlas (CBBA) field data provided acceptable to excellent results for most species and were, in general, highly comparable to the maps produced by the Catalan atlas (Figure 4.2).

Given the number of LTM schemes currently running in many countries, application of spatial modelling techniques to these data may prove a major contributor to conservation and land use planning in many areas. Spatial mapping of LTM data may substantially enhance the general efficiency of large-scale biodiversity assessments by adding a potentially useful spatially explicit component allowing accurate representation of species distributions. Furthermore, spatial mapping of LTM data may be integrated in current projects specifically aimed at mapping species distributions at large spatial scales. For instance, during Atlas work periods, spatial mapping derived from LTM data may become an integral part of Atlas methodology covering more common species.

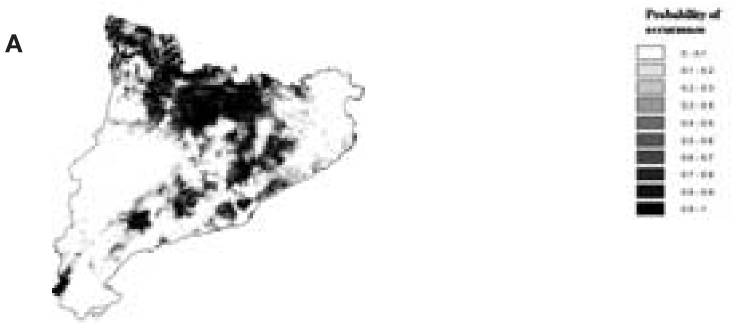
Some limitations of habitat modelling, however, such as the difficulty of accounting for fine-scale habitat structure, should be kept in mind to enhance proper use of distributions maps derived from LTM data. For instance, many authors consider maps generated by habitat or niche modelling as equivalent to potential distribution maps. Although our models predicted the occurrence of most species with high accuracy, some additional steps may be added to ensure that final relative abundance maps corresponded as accurately as possible to real rather than potential distribution maps. A possibility is to filter out hypothetical occurrence areas for each species from the known distribution of the species, gathered either from expert knowledge or coarse resolution field atlas data.

How to improve monitoring schemes to obtain more reliable distribution maps?

We think that if used for mapping purposes based on habitat modelling, LTM schemes should benefit from an effort to increase sample size. Such an increase in sampling effort is also likely to benefit trend estimation, which is the main aim of most LTM schemes. There is, however, a trade-off between the number of locations that could potentially be sampled and the distance volunteer surveyor have to travel to cover them. We suggest that LTM data based on long transects, or possibly other methods (e.g. point counts), may be disaggregated into smaller sampling units (i.e. 1-km transect sections in the case of the SOCC), leading to significant increases in the predictive accuracy of habitat models. The optimal degree of disaggregation to develop accurate habitat models from LTM data should be further investigated and is likely to depend on factors such as minimum

unit size, species ecology and spatial distribution of the sampling locations. The spatial coverage of the sampling scheme is also likely to be an important factor in many cases and therefore, improving this feature should be also favoured for mapping attempts.

Finally, we have shown that LTM data is ideally suited for occurrence data, which has been often found to be a good surrogate of abundance. However, since LTM schemes often collect count or density data, they have the potential to be used for more informative modelling of abundance data. It is expected that combining presence/absence modelling and abundance models will better fit the data when factors determining occurrence are different from those determining abundance.



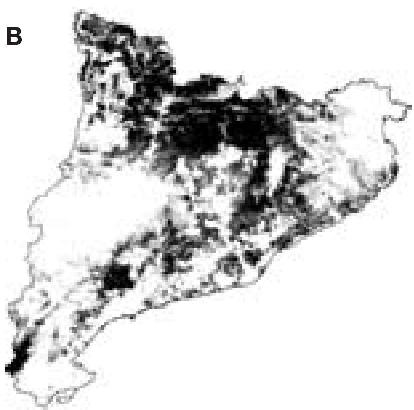


Figure 4.2. *Maps showing the predicted distributions for one of the species analysed, the Coal tit (Parus ater) in Catalonia as generated by (A) the Atlas occupancy model (CBBA, AUC = 0.91), (B) the long term monito- ring data occupancy model (SOCC, AUC= 0.85).*