Do habitat association models have any generality? Predicting skylark *Alauda arvensis* abundance in different regions of southern England.

Mark J. Whittingham¹, Jeremy D. Wilson² and Paul F. Donald³

¹Farmland Bird Group, Edward Grey Institute of Field Ornithology, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, U.K.

²RSPB Scotland, Dunedin House, 25 Ravelston Terrace, Edinburgh, EH4 3TP, U.K.

³RSPB, The Lodge, Sandy, Bedfordshire SG19 2DL, U.K.

Correspondence: M.J. Whittingham

Tel: (01865) 271156

Fax: (01865) 271221

email: mark.whittingham@zoo.ox.ac.uk

Abstract

Much of conservation science is based upon determining the use by organisms of different resources. However, the field data used to construct habitat association models generally come from a small number of sites covering a fraction of the area of interest. It is important therefore to assess the generality of these models for species occurring over large geographical areas. In this paper we test the generality of models describing skylark *Alauda arvensis* abundance across farmland in southern England in relation to crop type, crop structure and field structure (i.e. height of surrounding boundaries). Skylarks responded to most predictors we measured in similar ways across three regions of differing farming practices (arable-dominated, pasture-dominated and a mixture of the two). Most of the regional differences in habitat associations could be related to differences in the speed of crop development. For example, the sowing of cereals in spring, a much lauded strategy to increase skylark populations, is likely to have less of an effect in regions where cereal development is slow than in regions where it is fast.

Most studies that explicitly test the performance of a model developed in one place elsewhere use presence/absence models. We adopt a more sensitive and novel approach by using counts. We found regression equations developed in one region performed poorly when tested as a direct predictor (i.e. a 1:1 relationship) on data from other regions. However, the skylark territories observed in any one region were positively correlated with the territory numbers predicted by models built using data from other regions; so models were good predictors of relative abundance.

The results suggest that, for this species at least, conservationists should have confidence when advocating management strategies based upon habitat-association models and extrapolating their generality to other areas. However, our results warn against using regression equations developed in one place to make absolute quantitative predictions elsewhere. Decision-makers should beware of using models in this way.

Introduction

Information generated by applied ecological studies is often used in habitat and species management plans and to make policy recommendations (Ormerod et al. 2002). Often such studies take place at a small number of sites situated in the same geographical locality (e.g. Norris et al. 1997, Perkins et al. 2000, Bradbury et al. 2000, Whittingham et al. 2000, Milsom et al. 2000) although the results are often assumed to apply elsewhere. Even when data are collected from a larger range of geographical locations, the results are often pooled into one data set (e.g. Chamberlain et al. 1999, Henderson et al. 2000, Donald et al. 2001a). Only rarely are relationships between study species and their environments tested independently across a range of geographically separated areas (e.g. Fielding and Haworth 1995, Rodríguez and Andrén 1999, Morris et al. 2001) but it is precisely these kinds of studies that are most useful to underpin conservation action for widespread species by either providing separate management solutions for each area, or should relationships prove similar, one strategy across all areas.

Skylark *Alauda arvensis* L. populations, like those of many farmland bird species, have been declining rapidly in Britain and in Western Europe (Siriwardena et al. 1998, Donald et al. 2001b). This has prompted a number of studies of the breeding ecology of skylarks on farmland (e.g. Wilson et al. 1997, Wakeham-Dawson et al. 1998, Chamberlain et al. 1999, Donald and Vickery 2000, Donald et al. 2001a). These studies have highlighted the importance of field area, field boundary structure (notably height), in-field vegetation structure and crop type on skylark breeding abundance. Skylarks favour large fields with relatively short, sparse in-field vegetation and tend to avoid the vicinity of tall, enclosing boundary structures such as trees and tall hedges. Current conservation measures for skylarks are mainly based either on studies conducted in one region (e.g. Wilson et al. 1997, Wakeham-Dawson et al. 1998) or studies conducted in many regions but pooled together (e.g. Chamberlain et al. 1999, Donald et al. 2001a). In this study we test for differences in habitat associations between different geographical regions in order to assess whether existing habitat models have generality outside the areas where data were collected. This information is important if we are to implement effective conservation recommendations across the range upon which such measures are targeted.

We used Generalized Linear Mixed Models to test the extent to which relationships between skylark abundance and habitat measures differed between lowland farmland study areas in southern England. First, we fitted interaction terms between a factor ‘region’ and habitat predictors of interest. Second, we qualitatively compared the predictors that remained in models predicting the abundance of skylarks in models fitted to each region independently. Third, we

compared quantitatively how well a regression equation developed in one region predicted skylark abundance in other regions.

Methods

Study sites

We used data available from a previous study of skylark habitat associations (Donald et al. 2001a) and also data collected from five other farms in Oxfordshire in 1998. In 1996, Donald et al. (2001a) collected data from 12 farms located in three different regions of England: East Anglia, comprising Norfolk and Suffolk (four farms); Dorset (four farms); and Oxfordshire (four farms). We pooled the data from the farms in Oxfordshire (in 1996 and 1998) to examine patterns of skylark abundance in the three different geographical regions. We excluded data from one farm in Oxfordshire used by Donald et al. (2001a). This farm was not managed on a commercial basis and was more akin to a nature reserve with a large number of organic hay meadows atypical of lowland British farmland. This farm was excluded because we were interested in examining patterns of skylark abundance on typical lowland farms to which UK agri-environment schemes (such as Countryside Stewardship would typically be targeted). In total, therefore, we examined data from 16 farms: four in East Anglia, four in Dorset and eight in Oxfordshire.

A breakdown of the crop types in each region is presented in Table 1. Dorset farms were dominated by grassland with a small amount of arable (mainly cereal) crops. Arable crops dominated the East Anglian farms, these mainly comprised cereals with some other crops, particularly sugar beet *Beta vulgaris* L. (placed in the ‘others’ category, see Table 2). All farms in the Oxford area were mixed arable (mainly winter cereal) and grassland. Thus our three ‘regions’ broadly represented the three types of farming in lowland England, namely: pastoral, arable and mixed. The field sizes were surprisingly small for East Anglia; this is difficult to explain as our study farms were chosen at random from replies to a questionnaire sent to over 200 farmers, although it is important to note that field area was controlled for in all our models (see below).

Data collection

We included all fields within each study farm. Therefore fields were censused in contiguous blocks, such that every field was directly connected to at least one other (usually more than one) census field. On each of the study farms skylarks were counted every two weeks between 05.00 and 12.00 GMT in April-June. All parts of each field were visited to within 50m and Common Bird Census (CBC) activity codes used to record the number of skylarks present (Marchant et al.

1990). Censuses were not carried out in wet or windy weather, as these conditions tend to reduce skylark activity. From the census data, the number of occupied territories on each field was counted. Where a territory included parts of two fields, each was treated as having 0.5 territories. No attempt was made to distinguish between territories occupied by pairs and those occupied by unmated males because of the difficulty in proving the absence of the female.

Details of the habitat variables recorded for each field are given in Table 2. The index of the physical structure of field boundaries was calculated by dividing the perimeter of each field into sections, according to the following numerical categories: 0, no vertical structure; 1, a low (<2m) hedge, wall or bank without trees; 2, a tall (>2m) hedge, wall or bank without trees; 3, a hedge with trees or a line of trees; 4, woodland edge or a boundary with another unsuitable habitat, such as gardens, scrub or farm buildings. The length of each section (m) was multiplied by its category score, and the sum over all sections divided by the field perimeter to give a 'boundary index' for each field.

Data analysis

Generalized Linear Mixed Models (GLMMs) were used to identify significant predictors of skylark territory abundance in April-June. GLMMs allowed us to treat 'farm' as a random effect nested within region (following Milsom et al. 2000). We examined correlates of variation in skylark territories using log-linear regression (maximum number of skylarks recorded in a field over the month was specified as the response variable, assuming a Poisson error distribution and a log link, Crawley 1993). Overdispersion was automatically corrected for in our modelling procedure (using the default option in GENSTAT 4.2 which estimates the dispersion parameter: Roger Payne pers comm). We used a non-automated step-up procedure to identify those predictors (from the list given in Table 2) that accounted for a significant amount of deviance in the response variable. Field area was included as a predictor (not as an offset) because we did not wish to assume any fixed relationship between field area and skylark abundance. Models were fitted to data from all regions combined and for each of the three regions separately. We first fitted models in which the effects of the predictors were examined one at a time and noted the significance value each time. The significance of predictors was tested by the change in deviance (ΔD) in the model on addition or removal of the predictor. This change in deviance was assessed using the Wald statistic which is distributed asymptotically as χ^2 . We selected the predictor (A) with the highest significance level and then proceeded to test the effects of each of the other predictors in a model containing predictor A. The predictor with the highest significance level was then selected for inclusion in the model and so on. At each iteration the predictors already in the model were also tested for significance and removed if the significance level fell above $\alpha = 0.05$. The final model, or Minimum Adequate Model (MAM), was reached when no predictors could be either deleted or

added to the model without causing a significant change in deviance at the 5% level. Linear and quadratic relationships were explored for continuous variables by fitting simple linear terms and squared terms. The significance of a squared term was only considered when the main effect was also present in the model. Lastly, we fitted terms for the interaction between ‘region’ and each habitat variable remaining in the MAM as an initial test of the generality of the direction and strength of habitat associations across regions. Modelling was carried out using GENSTAT 4.2 (Welham 1993). The predictive power of the MAM was assessed by plotting the predicted number of skylark territories (as the x-variable) per field against the observed number of territories (as the y-variable) and fitting a regression line with an intercept of zero and a slope of 1 (i.e. $y = x$), from which an r^2 value was calculated (see Fig. 1).

We further tested the power of a model derived in a specific region ‘A’ (training data) to explain both **absolute** and **relative** skylark abundance in another region ‘B’ (test data). First, we tested whether there was some relationship between number of territories in the test data set and predicted values of the test data from a model developed from the training data (i.e. we used regression equations derived from the MAMs for the training data, in Table 4, to predict number of territories in the test data using the environmental predictors in the test data, and then regressed the predicted numbers against the actual observed numbers in the test data). We did this for each of the six combinations of training and test data set in each month. The MAMs were derived without fitting the ‘farm’ term, as it would not be meaningful to calculate parameter estimates for the test data using arbitrary codes derived from farms in the training data. In each case, we then calculated a Spearman rank correlation co-efficient for the predicted and observed skylark counts. We also used least-squares regression to calculate the linear predictor that described the most variation in the relationship between predicted and observed counts (i.e. the true ‘best fit’ line) and also the r^2 value to give a rough estimation of the power of the relationship (Fig. 1). If this linear predictor was found to have a high r^2 value, an intercept of zero and a positive slope of 1, then the MAM derived from the training data is a good predictor of **absolute** counts of skylarks on fields in the test data set (see below for explicit test). Where a linear predictor of similar power and direction has parameter estimates that differ markedly from zero and +1, respectively, then the MAM derived from the training data is a good predictor of **relative** skylarks counts in the test data set (i.e. there is a significant positive relationship but the relationship is not 1:1). Second, we further quantified the power of training data to predict **absolute** skylark counts in test data by forcing a linear predictor with an intercept of zero and a slope of one through each relationship between predicted and observed counts [i.e. a 1:1 relationship]. We calculated the total sum of squares from the observed mean from the test data and deducted the sum of squares from the 1:1 relationship to assess the fit of this line as an r^2 statistic (Fig. 1). This value indicated the proportion of variation in skylark territorial numbers that was explained by simply assuming the model from one region was directly applicable to another region. This value could be less than

zero, indicating that assigning a field a mean score from the number of territories counted across all fields in the training data was a better predictor than that derived from the test data. This method gives an indication of how well a model would perform if it were applied directly elsewhere.

Results

How similar were models from different regions?

Field area, crop type and boundary index not only explained a large amount of the variation in skylark territorial numbers in all months but their effects were similar in each region (Tables 3 and 4). This was supported by the low number of interaction terms between region and these predictors (Table 3) and the consistency of direction of the effect of these predictors (Table 4). However, there were two exceptions. (1) In April the effect of crop type on skylark territory abundance differed significantly between regions and there were also other differences in later months (Tables 3 and 5). Oilseed rape was generally little used in all regions except East Anglia where it was the crop type on which the greatest number of skylark territories were recorded in two of the three months. Spring-sown cereal fields were preferred to winter-sown cereal fields in all months in East Anglia, though only later in the season in Oxfordshire and Dorset. One reason to explain the consistent preference of spring over winter-sown cereal crops in East Anglia was that it was here that the greatest difference in height between the two sowing times was found (see discussion). (2) In East Anglia in April, the relationship between boundary index and skylark abundance was weak in contrast to Oxfordshire and Dorset, which showed a strong negative effect of boundary index. This difference is highlighted by the significant interaction term between boundary index and region (Table 3).

The negative relationship between increasing boundary height and skylark territorial numbers was generally consistent across different field sizes (Fig. 2). However, there was some indication that this relationship was not so straightforward on larger fields (>15 hectares), although sample sizes were small ($n = 6$ fields boundaries categorised as 3 or 4). This latter statement was supported by the significance of the interaction term of area and boundary index when added to the MAMs for May and June (see Table 4) in each region ($p < 0.001$ in each case), suggesting a less strong negative relationship between skylark territory numbers and boundary index on the largest fields.

The amount of variation explained by the remaining three predictors (vegetation height, vegetation cover and field shape) was smaller than that explained by the first three described above (Tables 3 and 4), although crop type and vegetation measures are confounded (i.e. vegetation measures had

stronger effects when crop type was removed from models). Generally the additive effects of both vegetation height and vegetation cover were similar in different regions, although there were some regional differences in the effect of field shape. Although the interaction term between vegetation height and region was not significant in any month (Table 3), there were some minor regional differences. Later in the season vegetation height had a significant additive effect in explaining variation in skylark territory numbers; in May in East Anglia, and in June in East Anglia and Oxfordshire when the crops became taller (Tables 4b-c). Vegetation cover was not significantly related to numbers of skylark territories in any region or month except Oxfordshire in April and showed a negative relationship with skylark territories (Table 4a). A weak curvilinear relationship was found for data from all regions in May. Overall, vegetation density had only a very weak additive influence on skylark abundance. The interaction between field shape and region was significant in two months (Table 3). The negative effect of field shape (Tables 4a and c) indicates that fields with higher ratios of field perimeter to area were avoided and so perhaps this effect was stronger in Dorset (see Table 1) where more of such fields were present.

In summary, in models built using data from all regions, the interaction term between the environmental predictors of interest and 'region' were significant in only four out of 18 cases. The same predictors tended to be present in models derived for each region separately and the effect of these predictors tended to be similar (i.e. in the same direction). Models tended to explain a large amount of the variation suggesting they provided a good fit to the data.

Regional model performance in other regions

In general the number of skylark territories predicted by a model developed in one region and then applied to other regions were positively correlated with observed territory numbers in those other regions (i.e. the slope of the relationship was positive in all cases). The amount of variation explained in other regions varied between months and regions from 1% to 48%, although on average when models were applied to other regions they only explained 16%, 15% and 36% of the variation in the three regions (Table 6). Models explaining variation in number of birds in April tended to have lower r^2 values than those in May or June. In summary skylark models developed in one place performed well in predicting relative numbers of skylark territories in other regions, but poorly in predicting the exact (absolute) number of territories (Table 7): none of the variation was explained in 12 out of 18 cases (when one model was tested in another region). Assuming that the number of skylarks found in a field was simply the average number counted across all fields on a specific farm frequently produced better estimates than using predictive equations developed elsewhere.

Discussion

How similar were models from different regions?

Habitat variables identified by previous studies (Wilson et al. 1997, Wakeham-Dawson et al. 1998, Chamberlain et al. 1999) as key predictors of skylark abundance (namely crop type, boundary index and crop height) were repeatedly found in models describing skylark territory numbers in three agriculturally contrasting regions. Furthermore, the effects of these predictors (or any other of our predictors, except field shape which explained only a small amount of variation) only rarely differed between regions. We explore the effect of each habitat variable, highlighting minor differences between regions, below.

Oilseed rape fields in East Anglia were used more than any other crop in May and June and accounted for the second highest counts in April. In contrast they were generally one of the least used crops in the other three regions. Although the number of oilseed rape fields was small (see Table 1) it is interesting to explore the reasons behind this pattern. Two of the three oilseed rape fields in East Anglia were from the same farm. Although these were sown in the winter, drilling was carried out directly on to a stubble field (which held very high densities of granivorous birds, including skylarks in the winter). Establishment of the crop was very slow and even when it became established skylarks were able to nest in bare patches which persisted throughout the crop. Such late and patchy crop establishment was atypical of oilseed rape fields elsewhere in the study.

Skylark territories were more numerous on winter-sown cereal fields than on spring-sown cereal and linseed fields for a greater part of the season in Dorset than in either East Anglia (where spring-sown cereals were always preferred) or Oxfordshire. This is probably explained by the slower growing winter-sown cereals in Dorset (e.g. mean height of winter cereal in May in Dorset was 28 cm), whereas in Oxfordshire and East Anglia this crop was taller (e.g. mean height of winter cereal in May in Oxfordshire and East Anglia was 61 cm). The crops need to reach a certain height (see below) before they have a negative effect on the number of skylark territories and only in June do winter cereals in Dorset reach this height. Spring-sown cereals do indeed hold higher skylark densities than winter-sown cereals in Dorset in June. Overall the regional differences in skylark territorial abundance between fields sown at different dates can probably be explained by differences in crop growth rate between regions.

Increasing boundary structure (e.g. taller surrounding hedges) had a consistently negative effect on skylark territorial numbers in all regions and in all months (except East Anglia in April). It is interesting to add here that even on the largest fields (>20 hectares) higher skylark densities were

found on fields with low boundaries, though the relationship with boundary index was not a straightforward one (Fig. 2).

The positive effect of area was consistently found in all months across all regions, although the effect of field size did reach a plateau for larger field sizes in some models. Although we did not use skylark density as the response variable (because the relationship between field area and skylark abundance may well not be linear and so we controlled for field area within our models, see ‘Data analysis’ section) we were able to plot skylark density against field area (Fig. 2) and we found that small fields (<5 hectares) held the lowest densities but larger fields held similar densities suggesting a plateau as indicated by our statistical models. Our findings contrast with those in Switzerland that reported that skylark densities were higher in small fields (Jenny 1990, Schläpfer 1988). However, those authors worked in continental strip farming systems with no vertical structures between fields. This contrast emphasises the important role that tall boundaries play in skylark territorial abundance and suggests that targeting small fields for ‘skylark-friendly’ in-field management may be less effective than focussing on larger fields when vertical structures are present between fields. This should be considered in studies of landscape design (e.g. Swetnam et al. 2001).

After controlling for crop type, vegetation height did not explain a significant amount of variation in skylark counts in any region in April when the crops were short. However, it explained a significant amount of variation in skylark territory numbers in May and June when the crops became taller in both Oxfordshire and East Anglia, though not in Dorset. Not surprisingly, a large amount of the variation in vegetation height could be explained by differences between crop types ($F_{6,380} = 73.4$, $p < 0.001$, $r^2 = 54\%$). However, most of the variation could be explained by differences between winter-sown cereals and oilseed rape (e.g. mean heights of 52 and 111 cm respectively in May) and the other five crop types (e.g. mean heights in May ranging from 12-32 cm) ($F_{2,385} = 181.7$, $p < 0.001$, $r^2 = 49\%$). Both winter-sown cereals and oilseed rape crops were much taller in May in Oxfordshire and East Anglia (winter-sown cereal mean = 61 cm, oilseed rape mean = 138 cm) than in Dorset (winter cereal mean = 43 cm, oilseed rape mean = 67 cm). The relationships derived from the MAMs for Oxfordshire and East Anglia in Tables 4b and 4c suggest that it is at heights of >60 cm that numbers of skylark territories are negatively associated with height (see also Donald et al. 2001a). Crops may simply not have grown sufficiently in Dorset (the thin chalky soils which characterise the region do not promote fast growth) to have a noticeable negative effect on skylark abundance. Our results have implications for the use of spring-sown cereals as a conservation measure for skylarks. Policy-makers and land managers should bear in mind that it is likely to be a more effective strategy in parts of the skylark’s range where crops grow more quickly.

The effect of field shape was found to differ between regions. Fields in Dorset tended to have higher ratios of perimeter to area than those in the other two regions, thus accentuating the effect of boundaries. When planning landscapes for farmland birds that select open landscapes, such as skylarks, it would be better to choose fields with lower ratios of perimeter to area (i.e. approximating a square) to implement in-field management measures. As we have shown some groups of farms, which may be representative of the regions from which they have come, have more suitable field shapes for skylarks. However, it is important to note that field shape is only a minor consideration when implementing ‘skylark-friendly’ management compared with targeting large, open fields and managing vegetation structure.

One potential criticism of our study was that we did not include information on occupancy of surrounding fields as predictors within our models. However we feel that including such information would add little to our study. Models developed within a region explained a large amount of variation (Table 6) suggesting that the amount of noise generated by the spatial structure of the skylark data did little to mask the clear habitat-association patterns. Furthermore, multiple skylark territories were found across many fields and so the scale of the territory was often far smaller than the size of the sampling unit. In addition, skylarks on any study farm could easily sample all the fields across that farm; a skylark could have flown across any study farm in less than one minute.

Our results apply only to skylarks and it is important to consider that other farmland birds may have different habitat preferences. In particular many other declining species use hedgerows for nesting, feed within non-cropped habitats, such as grass margins [e.g. linnet *Carduelis cannabina* (Moorcroft 2000); yellowhammer *Emberiza citrinella* (Morris et al. 2001; Perkins et al. 2002)] and, in contrast to skylarks, prefer landscapes with taller hedgerows. Therefore hedgerow trimming (which may benefit skylarks) is detrimental to some other farmland bird species (Swetnam et al. in press). Some species prefer trees (either within hedges or in woodland blocks) within farmed landscapes, e.g. chaffinch *Fringilla coelebs* (Whittingham et al. 2001), creating another conflict with skylark habitat preferences. Therefore our results should be used with caution if the aim is to enhance farmland bird populations in general. It is also worth noting that although skylark models proved to be fairly ‘portable’ between regions, this may not be the case for other farmland bird species, whose habitat preferences may be less easy to characterise, e.g. linnet (Moorcroft 2000). However, only time will tell how well models for other species perform across a range of different regions.

Regional model performance in other regions

Skylark abundances predicted from models developed in one region and applied to data in a second region were generally positively correlated with observed territorial numbers in the second region. However, the amount of variation explained varied considerably between regions and months. Models in April tended to have lower r^2 values than those in May or June (Table 6), this may have been because crops are more homogeneous between regions as they approach maturity than earlier in the season when local conditions and sowing times are more influential. One factor potentially contributing to the error in our models was the difference in the range of predictor variable values between regions (Table 1). Extrapolating a regression equation beyond the range of data from which it was constructed is an act of faith, which is likely to yield inaccuracies (Sokal and Rohlf 1995). However, the habitat characteristics in one region seem likely, in some cases at least, to lie outside the range found in other regions which may be inhabited by the same species (e.g. in our study it would seem strange if crop patterns and growth were uniform despite the differences in climatic conditions between regions which are known to affect growth and favour certain types of crops). Therefore the problem we highlight seems likely to be typical for studies of an animal over large geographical scales. However, despite these problems our models still found strong positive relationships in nearly all cases (16 out of 18) between predicted values from models built from training data and observed values from test data.

Another factor that could contribute to the lack of fit between predicted and observed skylark counts is the overall population density in different regions. Generally, the region with the lowest skylark densities (Dorset, Table 1) tended to under-predict the territorial numbers found in Oxfordshire (the region with the highest densities, Table 1), whilst models developed in Oxfordshire tended to over-predict the number of territories found in Dorset and East Anglia (Table 6). It is likely that landscape scale effects, for example the amount of arable land within the landscape surrounding each individual farm (Robinson et al. 2001), or regional differences in soils, weather or spatial configuration of crop types may have also contributed to the unexplained error in our models. These factors may well explain why, although models could predict relative densities in other regions quite well, they were very poor at predicting absolute densities (Table 7).

It is important to note that our models were developed and tested in lowland agricultural landscapes that were fairly homogeneous in nature. It was beyond the scope of this study to predict how models would perform in more diverse landscapes, such as are found in upland areas (>300 above sea level). Although it is important to note that 71% of the UK skylark population is associated with lowland agriculture (Donald and Vickery 2000) and so our study sites are likely to be representative of ‘typical’ skylark habitats, at least within the UK.

Comparison with previous habitat-association studies across different geographical areas

Some studies have failed to find similar habitat variables predicting bird abundance in different regions (e.g. Fielding and Haworth 1995), though others have found such consistency (e.g. Rodríguez and Andrén 1999, Morris et al. 2001). Our study joins the cohort of latter studies in that we have demonstrated a list of environmental variables that have consistent effects (at least in the same direction) across different regions.

This study is novel in predicting counts, rather than simply presence/absence in spatial units (e.g. Fielding and Haworth 1995, Rodríguez and Andrén 1999, Collingham et al. 2000, Ormerod et al. 2001). This step-up in resolution has revealed the dangers of direct extrapolation, even though the models from one region were generally correlated with numbers elsewhere.

Conclusions

Our study suggests policy makers and conservationists should have confidence when applying conservation measures for skylarks based on habitat-association models developed in any of the regions included in our study. These include encouraging the implementation of set-aside fields, advocating spring-sowing of cereals and suggesting to individual land-owners that they target ‘skylark-friendly’ measures on large fields with low surrounding boundaries (and also in areas with lower ratios of field perimeter to area, e.g. square rather than rectangular fields). However, our study suggests that direct extrapolation of results from one region to another does not yield accurate quantitative predictions. Instead we found a general positive correlation between skylarks predicted from one region and those observed elsewhere, suggesting that ‘skylark friendly’ habitat modifications based on the findings in a single region are likely to yield gains elsewhere. It may be difficult to determine the extent of habitat modification necessary to ensure that the targets set by the UK government for population increases are met (Aebischer et al. 2002), given the lack of direct extrapolation found in this study for a species which responds consistently to the same environmental cues. More refined studies are possible, such as spatially-explicit population models (e.g. Rushton et al. 1997) or individually-based population models (e.g. Sutherland 1996), but the level of detail necessary to parameterise such models is so time-consuming and difficult to collect that it may prohibit such approaches.

Acknowledgements

We would like to thank the many landowners who allowed us access to their land and provided much additional help. MJW was supported by BBSRC grant no. 43/D13408. For help with

fieldwork we thank Guy Anderson, Lys Muirhead, Dave Buckingham, Will Kirby, John Calladine, Julianne Evans, Tony Balbi, Tom Jenner and Jenny Gill. The RSPB Skylark research project was funded by Tesco Stores Ltd. in their role of champions of the Skylark Biodiversity Action Plan. Professor Roger Payne provided statistical advice and we are grateful to him and to Alan Fielding, Gavin Siriwardena, Professor Mönkkönen and one anonymous referee for improving early drafts of this paper.

References

- Aebischer, N. J. et al. 2002. Predicting the response of farmland bird populations to agricultural change. - British Trust for Ornithology report no. 289, Tring, UK.
- Bradbury, R. B. et al. 2000. Habitat associations and breeding success of yellowhammers on lowland farmland. - *Journal of Applied Ecology* 37: 789-805.
- Chamberlain, D. E. et al. 1999. Effects of habitat type and management on the abundance of skylarks in the breeding season. - *Journal of Applied Ecology* 36: 856-870.
- Collingham, Y. C. et al. 2000. Predicting the spatial distribution of non-indigenous riparian weeds: issues of spatial scale and extent. - *Journal of Applied Ecology* 37 (Suppl. 1): 13-27.
- Crawley, M. J. 1993. GLIM for Ecologists. - Blackwell Scientific Publications, Oxford, UK.
- Donald, P. F. et al. 2001a. Factors affecting the territory distribution of skylarks *Alauda arvensis* breeding on lowland farmland. - *Bird Study* 48: 271-278.
- Donald, P. F., Green, R. E. and Heath, M. F. 2001b. Agricultural intensification and the collapse of Europe's farmland bird populations. - *Proc. Roy. Soc. Lond. B.* 268: 25-29.
- Donald, P. F. and Vickery, J. A. 2000. The importance of cereal fields to breeding and wintering skylarks *Alauda arvensis* in the UK. In: Aebischer, N. J. et al. (eds), *Ecology and conservation of lowland farmland birds*. British Ornithologists Union, pp 140 – 151.
- Fielding, A. and Haworth, P. F. 1995. Testing the generality of bird-habitat models. - *Conservation Biology* 9: 1466-1481.

Henderson, I. G. et al. 2000. The relative abundance of birds on set-aside and neighbouring fields in summer. - Journal of Applied Ecology 37: 335-347.

Jenny, M. 1990. Territorialität und Brutbiologie der Feldlerche *Alauda arvensis* in einer intensive genutzten Agrarlandschaft. - Journal für Ornithologie 131: 241-265.

Manel, S., Williams, H. C. and Ormerod, S. J. 2001. Evaluating presence/absence models in ecology: the need to account for prevalence. - Journal of Applied Ecology 38: 921-932.

Marchant, J. H. et al. 1990. Population trends in British breeding birds. - British Trust for Ornithology, Tring, UK.

Milsom, T. P. et al. 2000. Habitat models of bird species' distribution: an aid to the management of coastal grazing marshes. - Journal of Applied Ecology 37: 706-728.

Moorcroft, D. 2000. The causes of the decline in the linnet *Carduelis cannabina* within the agricultural landscape. Unpublished DPhil thesis, Oxford University.

Morris, A.J. et al. 2001. Foraging habitat selection by yellowhammers *Emberiza citrinella* nesting in agriculturally contrasting regions in lowland England. - Biological Conservation 101: 197-201.

Norris, K. et al. 1997. The density of redshank *Tringa totanus* breeding on the salt marshes of the Wash in relation to habitat and its grazing management. - Journal of Applied Ecology 34: 999-1013.

Ormerod, S. J. et al. 2002. The uptake of applied ecology. - Journal of Applied Ecology 39: 1-7.

Perkins A. J. et al. 2000. Habitat characteristics affecting use of lowland agricultural grassland by birds in winter. - Biological Conservation 95: 279-294.

Perkins A. J. et al. 2002. Use of field margins by foraging yellowhammers *Emberiza citrinella*. - Agriculture, Ecosystems and Environment 93: 413-420.

Robinson, R. A., Wilson, J. D. and Crick, H. Q. P. 2001. The importance of arable habitat for farmland birds in grassland landscapes. - Journal of Applied Ecology 38: 1059-1069.

Rodríguez, A. and Andrén, H. 1999. A comparison of Eurasian red squirrel distribution in different fragmented landscapes. - Journal of Applied Ecology 36: 649-662.

Rushton, S. P. et al. 1997. Modelling the distribution of the red and grey squirrel at the landscape scale: a combined GIS and population dynamics approach. - Journal of Applied Ecology 34: 1137-1154.

Schläpfer, A. 1988. Populationsökologie der Feldlerche *Alauda arvensis* in der intensiv genutzten Agrarlandschaft. - Ornithologische Beobachter 85: 309-371.

Siriwardena, G. M. et al. 1998. Trends in the abundance of farmland birds: a quantitative comparison of smoothed Common Birds Census indices. - Journal of Applied Ecology 35: 24-43.

Sokal, R. R. and Rohlf, F. J. 1995. Biometry. - W.H. Freeman and Co., New York.

Sutherland, W. J. 1996. From individual behaviour to population ecology. - Oxford series in Ecology and Evolution, Oxford University Press, Oxford, UK.

Swetnam, R. D. et al. 2001. Modelling the effects of agricultural change on skylark numbers using GIS: a pilot study. In: Donald P. F. and Vickery J. A. (eds) The ecology and conservation of skylarks *Alauda arvensis*. RSPB, pp 209 – 219.

Swetnam, R. D. et al. in press. Designing lowland landscapes for farmland birds: scenario testing with GIS. - Computers, Environment and Urban Systems.

Wakeham-Dawson, A. et al. 1998. Breeding skylarks *Alauda arvensis* on Environmentally Sensitive Area arable reversion grass in southern England: survey-based and experimental determination of density. - Journal of Applied Ecology 35: 635-648.

Welham, S. 1993. Procedure GLMM. Genstat 5 Procedure Library Manual. - NAG, Oxford, UK.

Whittingham M. J., Percival S. M. and Brown A. F. 2000. Time budgets and foraging of breeding golden plover *Pluvialis apricaria*. - Journal of Applied Ecology 37: 632-646.

Whittingham, M. J. et al. 2001. Chaffinch *Fringilla coelebs* foraging patterns, nestling survival and territory distribution on lowland farmland. - Bird Study 48: 257-270.

Wilson J. D. et al. 1997. Territory distribution and breeding success of skylarks *Alauda arvensis* on organic and intensive farmland in southern England. - Journal of Applied Ecology 34: 1462-1478.

Table 1. Field characteristics in each of the three study regions. For each crop type the number of fields within each region are presented (percentages are given in parentheses). The mean scores for the environmental predictors measured on each field (see Table 2 and text) are also presented. Next to each of these predictors a symbol is given (NS $p>0.05$, *** $p<0.001$) to indicate the result of the appropriate statistical test (e.g. either one-way ANOVA or Kruskal-Wallis test) between that predictor and region as a three level factor. Note: both vegetation height and cover differed significantly ($p<0.001$) between regions, this was tested by fitting a General Linear Model between either vegetation height or cover and adding the factors 'region', 'crop type' and 'month'.

Crop type	East Anglia	Oxfordshire	Dorset
Winter cereal	33 (32%)	59 (43%)	23 (16%)
Spring cereal and linseed	9 (9%)	11 (8%)	6 (4%)
Oilseed rape	3 (3%)	4 (3%)	5 (3%)
Set-aside	10 (10%)	5 (4%)	19 (13%)
Silage	3 (3%)	25 (18%)	35 (24%)
Grazed pasture	8 (8%)	22 (16%)	47 (33%)
Other crops (see Table 2)	36 (35%)	12 (9%)	9 (6%)
Total number of fields	102	138	144
Mean boundary index (± 1 se) NS	1.90 $\pm$ 0.09 [Range: 0.4-4.0]	1.81 $\pm$ 0.06 [Range: 0.1-3.8]	1.82 $\pm$ 0.06 [Range: 0.6-3.6]
Mean field area in ha ⁻¹ (± 1 se) ***	6.41 $\pm$ 0.3 [Range: 0.8-35.6]	10.2 $\pm$ 0.6 [Range: 1.5-38.6]	8.15 $\pm$ 0.5 [Range: 0.2-17.9]
Mean field shape score (± 1 se), calculated using area in m ² NS	1.22 $\pm$ 0.02 [Range: 1-2.13]	1.28 $\pm$ 0.03 [Range: 1-4.08]	1.33 $\pm$ 0.04 [Range: 1-6.16]
Mean skylark territory density per ha ⁻¹ in April (± 1 se) ***	0.10 $\pm$ 0.01	0.15 $\pm$ 0.02	0.04 $\pm$ 0.004
Mean skylark territory density per ha ⁻¹ in May (± 1 se) ***	0.08 $\pm$ 0.01	0.12 $\pm$ 0.02	0.04 $\pm$ 0.006
Mean skylark territory density per ha ⁻¹ in June (± 1 se) ***	0.09 $\pm$ 0.02	0.11 $\pm$ 0.01	0.04 $\pm$ 0.006

Table 2. Predictors used in models to explain skylark territory abundance.

Variables	Type	Levels	Description
Crop type	Factor	7	1 = winter cereal 2 = spring cereal and linseed 3 = oilseed rape 4 = set-aside 5 = silage 6 = grazed pasture 7 = others (including sugar beet, maize, bare till, oats, potatoes, carrots, rye, flower-rich hay meadows*).
Farm	Factor	16	Individual level declared for each farm. Farm was included as a random effect (see text).
Region	Factor	3	1 = East Anglia 2 = Oxfordshire 3 = Dorset
Vegetation height	Continuous variable	N/A	Mean vegetation height (cm), from ten measurements within a 0.5 by 0.5m quadrat, from ten randomly selected points across the field**
Vegetation density	Continuous variable	N/A	Mean % vegetation cover, estimated by eye within a 0.5 by 0.5m quadrat, from ten randomly selected points across the field**
Field area	Continuous variable	N/A	Size of field in hectares
Boundary Index	Continuous variable	N/A	Measure of physical structure of field boundary: varies from 0 (no vertical structure in field boundary) to 4 (field surrounded by woods or unsuitable habitat such as gardens, scrub or farm buildings). See data collection (methods) for further details.
Field shape index	Continuous variable	N/A	Ratio of the actual field perimeter length to the perimeter length of a circular field of the same area calculated as:

$$\text{Shape} = \frac{\text{perimeter}}{2\pi\sqrt{\frac{\text{area}}{\pi}}}$$

*Flower-rich hay meadows were only found in the Cotswolds and they were so distinct that it was not considered appropriate to lump them with permanent pasture so they were placed in the ‘others’ category.

**In Oxfordshire measurements were taken in a diagonal across the field but in other regions they were taken randomly from anywhere within the field.

Table 3. Minimum Adequate Models (MAMs) for data from all regions combined in April-June. Wald statistics are presented along with the associated degrees of freedom in parentheses. The predictors accounting for the greatest variation in the data can be approximated by looking at the relative size of the Wald statistics values for each predictor compared with other predictors in each model (although it is important to note that these values should only be used as an approximate guide to the amount of variation explained). All main effects from Table 2 and two-way interactions with ‘region’ were included in the modelling process. Interactions between the predictor and the factor ‘region’ are shown in italics, e.g. the relationship with boundary index differed in the three regions in April but not in May or June. Direction of relationship: - (negative), + (positive), $\cup$ (concave), $\cap$ (convex). No direction of relationship is shown when the interaction terms were significant. R^2 values were calculated by fitting a line with an intercept of zero and a slope of one (i.e. a 1:1 relationship).

Month	Region	Crop type	Boundary Index	Area	Vegetation height	Vegetation cover	Field Shape	Observed vs predicted (r^2)
April	6.5 (2) *	51.5 (12) ***	11.6 (2) **	160.2 (1) *** (+)	4.6 (2) NS	2 (2) NS	10.4 (2) **	83%
May	7.5 (2) *	56.9 (6) ***	20.1 (2) *** ($\cup$)	80.9 (2) *** ($\cap$)	14.8 (2) *** ($\cap$)	8.7 (2) * ($\cap$)	0.7 (2) NS	82%
June	7.7 (2) *	74.6 (6) ***	38.9 (1) *** (-)	78.5 (2) *** ($\cap$)	17.1 (2) *** ($\cap$)	0.2 (2) NS	15 (2) ***	77%

Note: NS $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4a-c. Minimum Adequate Models (MAMs) explaining skylark abundance in **April-June** derived from each region separately. Wald statistics are presented along with the associated degrees of freedom in parentheses. The predictors accounting for the greatest variation in the data can be approximated by looking at the relative size of the Wald statistics values for each predictor compared with other predictors in each model (although it is important to note that these values should only be used as an approximate guide to the amount of variation explained). Having two degrees of freedom rather than one for continuous variables indicates quadratic relationships. Direction of relationship: - (negative), + (positive), $\cup$ (concave), $\cap$ (convex). R^2 values were calculated by fitting a line with an intercept of zero and a slope of one (i.e. a 1:1 relationship). Note: NS $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4a. MAMs explaining skylark abundance in **April** derived from each region separately.

	Crop type	Boundary Index	Area	Vegetation height	Vegetation cover	Field Shape	Observed vs predicted (r^2)
Oxfordshire	31.2 (6) ***	35.6 (1) *** (-)	53.6 (1) *** (+)	2.2 (2) NS	4.8 (1) * (-)	2.5 (2) NS	88%
East Anglia	65.4 (6) ***	3.2 (2) NS	40.3 (1) *** (+)	1.5 (2) NS	2.6 (2) NS	0.5 (2) NS	59%
Dorset	26.8 (6) ***	7.9 (1) ** (-)	46.3 (2) *** ($\cap$)	1.9 (2) NS	3.3 (2) NS	18.2 (1) *** (-)	74%

Table 4b. MAMs explaining skylark abundance in **May** derived from each region separately.

	Crop type	Boundary Index	Area	Vegetation height	Vegetation cover	Field Shape	Observed vs predicted (r^2)
Oxfordshire	24.7 (6) ***	12.6 (1) *** (-)	21.4 (1) *** (+)	6.3 (2) * ($\cap$)	4 (2) NS	1.8 (2) NS	80%
East Anglia	73.1 (6) ***	24.5 (2) *** ($\cup$)	102.4 (2) *** ($\cap$)	45.2 (2) *** ($\cap$)	3.2 (2) NS	2 (2) NS	67%
Dorset	41.3 (6) ***	38.3 (1) *** (-)	36.2 (2) *** ($\cap$)	1.89 (2) NS	1.63 (2) NS	2.28 (2) NS	53%

Table 4c. MAMs explaining skylark abundance in **June** derived from each region separately.

	Crop type	Boundary Index	Area	Vegetation height	Vegetation cover	Field Shape	Observed vs predicted (r^2)
Oxfordshire	12.9 (6) *	11.4 (1) *** (-)	23.1 (1) *** (+)	9.1 (2) * ($\cap$)	1.1 (2) NS	4.3 (2) NS	80%
East Anglia	33.7 (6) ***	13.6 (1) *** (-)	30.9 (2) *** ($\cap$)	10.9 (1) *** (-)	3.0 (2) NS	8.2 (1) ** (-)	59%
Dorset	67.1 (6) ***	12.7 (1) *** (-)	53.0 (2) *** ($\cap$)	1.5 (2) NS	1.0 (2) NS	5.9 (1) * (-)	76%

Table 5. Crop selection by skylarks in three different regions in April-June. Rank order of crops (1 indicates the greatest number of territories) calculated from parameter estimates derived from MAMs in Tables 4a-c. The rank orders of crop types presented here were derived after controlling for the other predictors that were significant in the MAMs in Tables 4a-c.

Region	Month	Winter cereal (1)	Spring cereal & linseed (2)	Oilseed rape (3)	Set- aside (4)	Silage (5)	Grazed Pasture (6)	Other crops (7)
Oxfordshire	April	2	4	6	1	3	7	5
	May	4	2	7	1	5	6	3
	June	3	2	5	1	4	7	6
East Anglia	April	6	4	2	3	1	7	5
	May	6	3	1	2	5	7	4
	June	4	3	1	2	5	7	6
Dorset	April	3	6	5	1	4	7	2
	May	4	5	3	1	2	7	6
	June	5	2	7	1	3	6	4

Table 6. Performance of models developed in one region in predicting the **relative** number of skylark territories in other regions. The amount of variation explained (r^2) by regressing the observed values from the column-wise region against those predicted from the MAMs in Tables 4a-c (excluding farm) is presented. # For example, the model developed from East Anglian data for April (training data) explained 7% of the variation in skylark abundance in Oxfordshire (test data). The regression equation is presented in square brackets. # For example, the mean number of skylark territories in Oxfordshire in April was equal to $1.31 + 0.20$ times the number of territories predicted from the equation from East Anglia in April (see Table 4a). The gradient of 0.2 indicates that for every 0.2 skylark territories observed in Oxfordshire the model from East Anglia predicts that 1 skylark territory should be found. Symbols next to the regression equation indicate both the level of significance (NS $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$) and the direction of the relationship (+ positive, - negative) of the Spearman rank correlation between training and test data. ## For example, the model in East Anglia (training data) was significantly positively related to the Oxfordshire data (test data) at $p < 0.001$.

	Month	East Anglia	Oxfordshire	Dorset
Variation explained by model developed in East Anglia	April	0.64	#0.07 [y = 1.31 + 0.20x] ## *** +	0.01 [y = 0.29 + 0.03x] NS
	May	0.72	0.26 [y = 0.39 + 1.31x] *** +	0.47 [y = 0.06 + 0.48x] *** +
	June	0.61	0.13 [y = 0.70 + 0.91x] *** +	0.41 [y = 0.02 + 0.95x] *** +
Mean r^2 values across all Months		0.66	0.15	0.30
Variation explained by model developed in Oxfordshire	April	0.08 [y = 0.56 + 0.19x] ** +	0.71	0.33 [y = 0.04 + 0.22x] *** +
	May	0.19 [y = 0.39 + 0.24x] *** +	0.65	0.44 [y = 0.01 + 0.28x] *** +
	June	0.17 [y = 0.31 + 0.18x] *** +	0.55	0.48 [y = -0.11 + 0.43x] *** +
Mean r^2 values across all Months		0.15	0.64	0.42
Variation explained by model developed in Dorset	April	0.01 [y = 0.72 + 0.27x] NS	0.05 [y = 1.00 + 2.16x] ** +	0.68
	May	0.10 [y = 0.51 + 0.60x] ** +	0.16 [y = 0.51 + 3.15x] *** +	0.57
	June	0.41 [y = 0.25 + 0.89x] *** +	0.21 [y = 0.49 + 2.54x] *** +	0.74
Mean r^2 values across all Months		0.17	0.14	0.66
Mean r^2 values across all months and both regions to which models derived elsewhere applied		0.16	0.15	0.36

Table 7. Performance of models developed using training data from one region and applied to test data from another region (derived from MAMs in Table 4a-c but excluding the predictor ‘farm’) and plotted against observed test data to predict **absolute** territory numbers of skylarks. A line with an intercept of zero and a slope of one [i.e. a 1:1 relationship] was forced through the plot and an r^2 value derived (i.e. we tested how well a direct extrapolation from one region to another would predict skylark abundance). Note that it is possible for an r^2 value to be less than zero (i.e. the mean value of skylarks actually observed for the series of fields on a given farm is a better predictor than that predicted by the regression equation derived from another region).

	Month	East Anglia	Oxfordshire	Dorset
Variation explained by	April	0.64	<0	<0
model developed in	May	0.72	0.18	<0
East Anglia	June	0.61	0.04	0.41
Variation explained by	April	<0	0.71	<0
model developed in	May	0.15	0.70	0.28
Oxfordshire	June	<0	0.55	<0
Variation explained by	April	<0	<0	0.68
model developed in	May	<0	<0	0.59
Dorset	June	0.35	<0	0.74

Fig. 1. Method for testing absolute and relative abundance of skylarks from one region in another region. Data presented here are from the model derived in Oxfordshire in April and applied to data from fields in East Anglia (each point on the graph represents number of territories on one field). The mean number of skylark territories recorded per field in East Anglia was 0.85, the distance between every point and this mean value was calculated and squared, and then all these values summed to give the total sum of squares. The ‘best-fit’ line was derived using a software programme (e.g. excel) to reduce the sum of squares as much as possible. This line is shown below along with the regression equation and the amount of variation explained (7.6%) (See also Table 6). Thus over 92% of the variation remains unexplained by the Oxfordshire model predictions. However, when a direct 1:1 relationship is fitted (see below) the total sum of squares (i.e. calculated by taking the distance from each point to the 1:1 line, squaring it and then summing together all the values) was greater than the total sum of squares derived from the mean value. Hence, the mean number of territories recorded per field in East Anglia gave a better prediction of each individual territory score for each East Anglian field than the absolute amount that the Oxfordshire model predicted (see also Table 7, i.e. <0 amount of variation explained).

Fig. 2. The relationship of mean observed skylark density with boundary index in May (1 = low boundaries with scores <1, 2 = boundaries with scores 1 – 1.99, 3 = boundaries with scores 2 – 2.99, 4 = tall boundaries with scores of 3 - 4, see Methods for further details) on fields of different sizes (1 = fields <5 hectares, 2 = fields 5-9.99 hectares, 3 = fields 10-14.99 hectares, 4 = fields 15-20 hectares). Data were combined across all three regions because similar patterns with boundary index and field area were found in all cases (and also in April and June).