Developing tools to recognise constraints and opportunities for breeding birds associated with forest expansion: Pilot Forest Areas in the Scottish Borders.

A report to Scottish Forestry

John Calladine, Mark Wilson

BTO Scotland, Beta Centre (Unit 15), Stirling University Innovation Park, FK9 4NF

1. INTRODUCTION

The Scottish Forestry Strategy sets out ambitious goals for woodland expansion and aims to achieve 21% land cover by 2032 through increasing annual afforestation targets to 150 km^2 by 2025 (Scottish Government 2019). New woodland creation and regeneration provide opportunities for biodiversity (including birds) but also risks to some bird species and assemblages. For example, breeding waders have declined markedly over recent decades (Balmer *et al.* 2013). Waders have a high profile among the species of conservation concern that can be negatively affected by forest expansion through replacement and fragmentation of their required open habitats and associated changes in predation risk (Calladine *et al.* 2018). However, there is also increasing evidence that forest expansion in Scotland is providing opportunities for other birds which are also conservation priorities, notably some migrant passerines that rely on woodlands and shrublands (Gillings et al. 2000, Calladine *et al.* 2019).

As forest expansion can provide both constraints and opportunities to species of conservation concern, there is a need to provide tools for use by those who are considering creating new woodlands and by those who assess and grant permissions for such proposals. Such tools would indicate areas where:

- Risks to birds of open habitats and landscapes are greatest;
- Risks to birds of open habitats and landscapes are low;
- Opportunities to enhance conservation of forest and woodland birds are greatest.

Such tools should enable new forest proposals to more effectively target areas where conservation conflicts would be minimised and net benefit could be maximised.

A tool to model species-specific risks and opportunities requires detailed spatial knowledge of species distributions. This can be achieved through modelling the relationship between a species' abundance (from appropriate sample-based surveying) and various environmental variables that determine their distribution (e.g. Brambilla *et al.* 2009, Maleki *et al.* 2016). As well as generating information on previous or existing spatial patters of distribution and abundance, this approach can allow us to understand on how a species is likely to respond to changes in its environment. This understanding can be useful to decision makers aiming to conserve the species, as well as those wishing to make changes to existing landscapes for any purposes that have the potential to compete with conservation management.

This report describes a case study in the Scottish Borders that uses existing survey data on birds and environmental information to inform decision making around forest expansion. The most developed cases presented are for breeding waders for which three mapped strata of importance for each species are proposed to inform and mitigate potential risks associated with new forest creation. Also presented are some less developed case studies for species that could potentially benefit from forest expansion.

2. METHODS

Two Pilot Forest Areas (PFAs) in the Scottish Borders region, near Etterickbridge (*ca.* 154 km^2) and around Teviothead (*ca.* 122 km^2) offer an opportunity to test the effectiveness of sensitivity mapping to represent bird conservation interests in a way that is suited for informing forest expansion planning and procedures (Fig. 1). The relative importance of areas within the PFAs are considered for each of seven species of breeding wader (Oystercatcher, Lapwing, Curlew, Snipe, Golden Plover, Redshank and Common Sandpiper) that are likely to be affected negatively by forest expansion. Simple, modelled scenarios for possible responses by three species of woodland bird (Spotted Flycatcher, Wood Warbler and Redstart) are also presented.

Bird data were sourced from extensive national atlases of breeding bird distribution (Section 2.1) and modelled with a range of environmental variables (Section 2.2) to produce indices of expected abundance at 1-km square resolution (Section 2.3). The resulting indices were then used to identify strata of relative importance for breeding waders (Section 2.3).

2.1 Bird data

Data on breeding distribution and abundance were derived from the bird atlas fieldwork conducted in 1988-91 (Gibbons *et al.* 1997, hereafter *BA1990*) and 2008-11 (Balmer *et al.* 2013, hereafter *BA2010*). For purposes here, data were restricted to those from mainland Britain only. Some islands support particularly high densities of breeding waders, in part as a result of restricted suites of predators because of natural insularity (e.g. Calladine *et al.* 2017). Such islands are concentrated in the north and west, which could potentially have confounded the apparent influences of latitude and longitude in statistical models (see Section 2.3) and so were excluded. Volunteers surveyed a sample of tetrads (2 km by 2 km), making two 1-hour visits to each tetrad (a timed-tetrad-visit or TTV); one in the early part of the breeding season (April – May) and one in the late part (June – July). Within each hectad (10 km by 10 km square), at least eight tetrads were sampled. In some cases, two-hour visits were made to each tetrad but counts were recorded for each hour separately. To assess and map recent abundance recent abundance, we used the *BA2010* atlas data and extracted the maximum count of the early and late season visits per tetrad. Where a two-hour visit was made, the mean of the two hour-long counts was used as the count contributing to the maximum between the two seasonally separated visits.

Count data were not available for waders in *BA1990*. Therefore, we used occupancy (presence/absence data) to model probabilities of occurrence, with a range of 0 – 1, as proxies for relative abundance. In order to ensure that apparent occupancy in individual survey tetrads was not affected by survey effort, we only used data from the first hour of surveys, regardless of whether survey visits were one or two-hour in duration.

It should be noted that observers recorded all species or individuals seen regardless of whether they were actively breeding. This means that observers sometimes encountered non-breeding flocks of waders. To limit the influence of these records on modelled outputs, any counts or presence information outside the known range (judged from recorded breeding evidence as less than ‘probable’ within the hectad including the tetrad count) were turned to zeroes/absences. Despite this, a small proportion of wader counts within the national data set were very large (for example, up to 400 Curlew per tetrad) indicating that some non-breeding flocks had been included within the breeding range. Prior to modelling, for all waders apart from Snipe and Common Sandpiper, large counts were capped (by changing them to the median count across mainland Britain), in order to minimise the influence of large, non-breeding flocks on models aimed at estimating breeding abundance. For all of these species, reporting frequency declined with count size (i.e. tetrads with small counts were more frequent than tetrads with large counts). However, at low numbers (less than 20 birds) even numbers (e.g. n) were reported more frequently than adjacent smaller odd numbers (e.g. n – 1), probably due to the fact that breeding waders are often detected in pairs. For each species, the lowest even count whose frequency was not higher than the succeeding odd count therefore indicated that counts of this size or above may have been associated with non-breeding flocks. This value varied from 14 (for Golden Plover) to 24 (for Lapwing, Oystercatcher and Redshank), suggesting that 20 might be a sensible value at which to cap breeding counts for these species. Additional evidence for this came from the fact that the negative relationship between frequency and count size (for even counts) broke down between counts of 18 and 20. For all species, counts of 20 were noticeably more frequent than counts of 18. This is likely due to the tendency of counters to round flock sizes to the nearest ten birds (so that some flocks of 19 or 21 birds are reported as 20). Thus, at counts of 20, there is further evidence for an influence of (presumably non-breeding) flocks on count frequencies. Assuming that counts of 20 or higher are frequently of non-breeding flocks, we capped count values at 18 for all species. For each species, counts greater than the capped value in areas where breeding evidence was ‘probable’ or better were changed to the median count for the species (Table 1).

2.2 Environmental variables

A range of environmental data was used to represent the available variables that were deemed most likely to affect breeding waders and woodland birds. Wherever possible, data were selected that were well matched, temporally and spatially, to the Bird Atlas data.

2.2.1 Climate (temperature and precipitation)

Climate data was available at the 5-km resolution from the Met Office’s UK climate projections for 2009 (available at: http://www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09/download/index.html). The data are generated for a regular 5-km grid via regression and interpolation of raw data derived from the irregular weather station network, taking into account longitude, latitude, elevation, terrain shape, coastal influence, and urban land use (Perry & Hollis 2005). To encompass conditions when the birds were breeding we used the mean of mean monthly temperatures and the mean of total monthly rainfall from the months April, May, June and July. For winter conditions, the mean of mean winter temperatures and the mean of total rainfall from the months of December, January and February before the breeding season of interest (i.e. December 2007 for 2008 survey) were calculated. The mean of these variables for each 5 km square was then calculated for each 5-km square for the years 2008 to 2011.

2.2.2 Topography (elevation and slope)

Elevation (in meters above sea level) was extracted from the GGIAR-SRTM 90m raster (Jarvis *et al.* 2008, available at http://srtm.csi.cgiar.org) taking the mean elevation over each tetrad or hectad (depending on the scale of the analysis). Slope was calculated from elevation in ARCGIS (ESRI 2011). The slope of each elevation raster cell is the maximum rate of change in elevation in one raster cell compared to its eight neighbours. The lower slope values indicate flatter areas, higher values indicate steeper areas. The median slope was taken for each 1-km square or tetrad, as this represents mostly flat areas more effectively than mean slope.

2.2.3 Habitat

Land cover data came from the 1-km square percentage cover summary of the 2015 Land Cover Map (LCM) from the Centre for Ecology and Hydrology (Rowland *et al.* 2017). Seven land cover categories were derived from the LCM dataset: (i) semi-natural unimproved grassland (including rough grassland, neutral grassland, calcareous grassland, acid grassland and fen, marsh and swamp); (ii) mountain, heath and bog (including heather, heather grassland, montane habitats and inland rock); (iii) intensively managed arable land; (iv) intensively managed improved grassland; (v) urban and suburban habitats (including built land, and suburban land); and (vi) broadleaved woodland and (vii) coniferous woodland. The mean percentage of organic carbon in topsoil from the European soil data centre (Jones *et al.* 2003) was used as an indication of the peat content of soils.

For more in-depth data on the availability and configuration of woodland we used the Forestry Commission's National Forest Inventory for 2011 (https://www.forestry.gov.uk/inventory). The relevant variables were extracted at 1km square level by intersection of the relevant forestry shape files and a polygon layer of 1km squares in ARCGIS. Percentage cover was calculated by summing over each 1km. As a measure of patchiness, the perimeter length of mature woodland was calculated for each 1km square and a measure of heterogeneity (to represent both structural diversity and patchiness) was also calculated by summing the number of separate woodland habitat parcels in each 1km square. Here, we combined regenerating shrubby understory, young growth stage plantations and shrub woodland into 'young growth stage trees' and bare ground, rock, forest tracks and grass into 'clearings'. The data available on forestry structure for the whole of Britain does not extend to types of trees beyond conifer or deciduous. Data on tree age and density of woodland is also limited to broad categories such as low density, scrub, clearings and young trees.

2.2.4 Wind farms and roads

Data on wind farm developments came from the Renewable Energy Planning Database (REPD) (Available from: https://www.gov.uk/government/publications/renewable-energy-planning-database-monthly-extract). Only developments that were in the construction or operational phrase during one of the Atlas survey periods were included (though no large-scale wind farms had been built before the 1988-91 Bird Atlas). The dataset gave the energy output expected for the wind farm (MW) and a central coordinate for the location of the wind farm but not an area. Therefore we estimated the area from a formula derived by Bright *et al.* (2008): footprint (km^2) = ($7E^{-5} \times$ Output2) + ($0.0505 \times$ Output) + 0.0295. Then, because a recent paper found breeding bird densities may be reduced within a 500-m radius of wind turbines, we added 1 km^2 to the total footprint area (Pearce-Higgins *et al.* 2009). Then we converted this area into a circular buffer centred on the coordinates given for the wind farm to use as an approximation of the footprint of impact of a wind farm. Lastly we intersected these buffers and a shapefile of 1km grid squares to determine the percentage of each tetrad/1km square was likely to be effected by wind developments.

We used the Annual Average Daily Flows (AADF) data from the department of transport for major roads (A roads and motorways) as an indication of heavy road traffic. The AADFs are calculated from around 10,000 manual point counts and automatic traffic counters, the observed data was roads adjusted to compensate for road length (Available from: https://www.dft.gov.uk/traffic-counts/).

2.2.5 Predation risk

There is not currently robust tetrad/hectad level data on mammal abundance for the whole of Britain and even presence-absence data is in short supply. Due to the lack of detailed distribution data, mammalian predators were assumed present throughout mainland Britain, which meant the resulting variable was highly skewed, predominantly consisting of presences. Fox trends generated from the Breeding Birds Survey (Massimino *et al.* 2018) were also used to calculate an index of fox abundance in 2007.

Better data are available for avian predators, from the Atlas. In order to generate tetrad-level estimates of abundance for individual avian predator species, we carried out random forest regression tree modelling, following the methods outlined below. All of the above environmental variables, including those pertaining to mammalian predators, were included in these models. We used these models to predict relative abundance (average TTV count) of Buzzard, Ravens and Carrion/Hooded Crows (combined) in every tetrad. The abundance of these predatory species were included in our models of wader abundance as surrogates for risk of avian predation on the eggs and chicks of breeding waders.

An important factor influencing the variation in predation risk across mainland Britain is predator control, which is particularly associated with management for Red Grouse. The control of predators (typically corvids, fox, stoat and weasel) is a key management practice for grouse moors aiming to enhance survival and breeding success of ground-nesting gamebirds (Tharme *et al.* 2001, Fletcher *et al.* 2010). Such areas arguably represent the only areas in Britain where predator control is effective at a landscape scale. Other studies (e.g. Franks et al. 2017, Douglas et al. 2014) have inferred variation in intensity of predator control from spatial patterns in muirburn (a practice closely associated with grouse management whose effects on vegetation can easily be detected in aerial photos). However, not all areas subject to predator control aimed at enhancing Red Grouse populations are intensively burned. We therefore used estimated Red Grouse densities as a more direct index of the intensity of management for that species which would be inclusive of predator control. We modelled the relative abundance of Red Grouse using Bird Atlas data, as described above for avian predators (and in more detail below for waders). However, we included avian predators as explanatory variables in the models of Red Grouse abundance.

2.3 Analyses

All data analysis and manipulations were carried out in R 3.6.1. The following packages were used: randomForest , rgdal , sf, birdatlas and stringr. The spatial extent of environmental and bird data used in these analyses were limited to the area of mainland Britain. This was done in order to exclude islands from consideration, as breeding wader populations on many island groups in Britain experience different environmental constraints (particularly related to predation risk, but also related to habitat) from waders on the mainland.

The explanatory variables, most of which were at a resolution of 1km, were rescaled to tetrad level in order to make them correspond with Bird Atlas data. Before modelling, explanatory variables were screened to reduce collinearity among them. A correlation matrix of all explanatory variables was generated and used to identify pairs of variables for which the correlation co-efficient was greater than 70%. For all such pairs, we kept only one variable in the subset used for modelling. The choice of variable to retain was based on an aim to maximise the number of variables we could include (i.e. if one variable was highly correlated to two other uncorrelated variables it was removed in preference to removing two). Beyond this, variables were selected for modelling on the basis of their likely ecological relevance. The final subset after removal of collinear variables comprised 24 variables : Easting; Northing; Mean Altitude; Land cover variables (woodland; arable; improved grassland; semi-natural grassland; mountain, heath & bog; urban & suburban); Percentage windfarm cover; Road traffic; Summer temperature; Summer rainfall; Woodland variables (broadleaf cover; conifer cover; clearings; young growth; mixed woodland; heterogeneity); Mammalian predators; Avian predators (Carrion/Hooded Crow; Raven; Buzzard) and Red Grouse.

2.3.1 Modelling wader abundance and occurrence

Random Forest modelling was used to generate models of abundance and occurrence for the nine wader species (Common Sandpiper, Curlew, Dunlin, Golden Plover, Lapwing, Oystercatcher, Redshank, Ringed Plover and Snipe). Random Forest Regression Trees (RFRTs) were used to model relative abundance (with hourly TTV counts as the response variable), and Random Forest Classification Trees (RFCTs) were used to model probability of occurrence (with presence/absence during TTVs as the response variable). The abundance models (RFRTs) were used to derive maps of relative importance for breeding waders during the time of *BA2010*. Occupancy models (RFCTs) were used to derive maps of change in the occurrence of breeding waders between the time of *BA1990* and *BA2010*. Abundance data were not available for waders in *BA1990*. Therefore, we used probabilities of occurrence, which have a range of 0 – 1, as proxies for relative abundance.

The models were built using the R package randomForest (Liaw & Weiner 2002), which is based on the random forest classifier described by Breiman (2001). A random forest is a classifier consisting of a larger number of regression or classification trees. Each tree recursively partitions a dataset; repeatedly subdividing based on thresholds values of explanatory variables that best explain variation in the dependent variable (for regression) or predict discrete outcomes (for classification). The predicted value for each terminal node (or 'leaf') of the tree is simply the sample mean of the dependent value for all datapoints in that subdivision. In a random forest, each tree is based on a bootstrapped dataset, generated by sampling the original dataset with replacement. For any given datapoint, the predicted values for each tree are averaged to yield a prediction from the whole random forest.

For both RFRTs and RFCTs the number of trees in each random forest was set at 500, and the number of variables sampled as candidates for each tree set at 5 (derived as p/3, where p is the number explanatory variables included in the full model). These models were used to predict relative abundance (from 2010 Atlas data) and probability of occurrence (from both 1990 and 2010 Atlas data) for all tetrads in Britain. Within the whole of Britain, we used iterative Chi-squared tests to determine which threshold (from 0.01 to 1) most effectively discriminated between modelled count values for tetrads in 10km squares with and without probable (or better) evidence of breeding. Values of modelled abundance or probability of occurrence that were lower than this threshold were set to zero.

The 'fit' of random forest models (i.e. how well they performed in predicting wader abundance occurrence values) was assessed as R-squared for RFRTs and as the percentage of correctly predicted presence and absence values for RFCTs. In both cases, the performance of the forest was made using out-of-the-bag (OOB) predictions. This means that our estimates of predictive power for these models are effectively based on independent data, and so should provide a good indication of their ability to model abundance and occurrence in areas where no Atlas data were collected. The importance of each variable in determining random forest model predictions was assessed according to increase in node purity (Breiman 2001). Node purity was measured as the mean decrease in sum of squares subsequent to splits based on the variable for RFRTs, and by Gini importance (based on the proportion of responses that are correctly predicted by splits based on this variable) for RFCTs.

2.3.2 Rescaling to derive model outputs at 1-km square resolution

In order to refine the scale at which we could predict and present information about abundance and distribution of birds, we took 1km resolution data for each of our explanatory variables and summarised them to tetrad level for each of four shifted tetrad grids (Figure 2). We applied the abundance and distribution models described above to all four grids to generate predicted values for each tetrad in the grid. For every 1km square in the areas of interest we extracted the predicted value from each of the four tetrads that contained the centroid of that 1km square. We used the averages of these tetrad-level estimates as measures of relative abundance and occurrence at the 1km level. It should be noted that these values, although applied to 1km squares, are at the same scales of abundance and occurrence probability as the original (tetrad-level) Atlas values. This means that values of relative abundance and probability of occurrence at the 1km scale can be directly compared with those at the tetrad scale. However, if interpreting these literally as probabilities of occurrence or number of birds likely to be counted during TTVs, or if summarising these 1km values across larger areas, it should be borne in mind that the values are tetrad means.

2.3.3 Stratification of importance for breeding waders

A principal aim of this work was to produce three strata of importance for each breeding wader species ('hot', 'warm' and 'cold'). Two alternative approaches to deriving the strata are presented: (a) Strata of national importance based on the variation in predicted abundance indices across mainland Britain; and (b) Strata of regional importance based on the variation in predicted abundance within 50 km of the PFAs in the Borders. Some provisional strata were circulated amongst ornithologists with local knowledge of the areas under consideration. Their comments and feedback were used to refine the criteria for defining strata to better match their direct experience with the area. Further confirmation of predicted outputs with tetrad-scale maps presented in the regional bird atlas for south-east Scotland (Murray *et al.* 2019).

For both approaches to developing strata (national and regional), the 99.5th percentile of the index of relative abundance was taken as a maximum marker value in order to exclude a very small number of high outlying predictions for some species. The 'hot' stratum included all squares where the predicted index of abundance was greater than 40% of the 99.5th percentile value (greater than 30% for Dunlin, Common Sandpiper and Ringed Plover). The 'warm' stratum included all squares where the predicted index of abundance lay between 10 – 40% of the 99.5th percentile value (5 – 30% for Dunlin, Common Sandpiper and Ringed Plover). The 'cool' stratum comprised all remaining squares.

Changes in relative abundance over the 20 years between the two atlas periods were calculated by subtracting the probability of occupancy in a 1-km square during *BA1990* from that during *BA2010*. Where that change was greater than 0.4 in either direction, this was categorised as a 'large' change. Where a change was between 0.1 – 0.4, that was categorised as a 'moderate' change.

Shape files to facilitate mapping of the strata for each species were created using ArcGIS (ESRI 2011). These were inclusive of a buffer of 707 m from the centre of each 1-km square to include the entire square plus a small additional area. Where there was an overlap from the small additional areas, they were included within the 'hottest' or greatest change stratum.

2.3.4 Modelling effects of woodland expansion on woodland specialists

We used RFRTs to model the abundance of three woodland specialist species: Redstart, Spotted Flycatcher and Wood Warbler, based on 2010 Atlas TTV counts. Models were as described for waders, but omitting avian predators and grouse from the suite of explanatory variables (as any influence of predation risk is likely to operate somewhat differently from species in open habitats). We used these models to predict the relative abundance of these species at a 1km scale, under four scenarios:

1. Current levels of forest cover;
2. Up to 25% increase in levels of conifer forest cover, at the expense of semi-natural grassland cover;
3. Up to 25% increase in levels of broadleaved forest cover, at the expense of arable land and improved grassland. Each of these habitats contributed to the overall decrease of farmland in proportion to their availability in the square;
4. Up to 25% increase of conifer and broadleaved forest (combined), at the expense of arable land and both improved and semi-natural grassland. As in scenario 3, each open habitat contributed to the overall decrease in proportion to their availability. The corresponding increase in forest cover was also apportioned between forest types according to the proportional availability of conifer, broadleaved and mixed forest in the square. If a square had 2% cover of broadleaves, 8% cover of conifers and no mixed forest, the 25% increase of

forest cover would be split into 5% broadleaves and 20% conifers. In squares with no pre-existing forest cover, the increase was split evenly between different types.

In scenarios 2 – 3, squares with less than 25% cover of the relevant open habitats (combined area) had correspondingly reduced increases in forest cover. In these squares, the cover of these open habitats was set to zero, and cover of the relevant forest type(s) increased by the same amount. In squares where cover of the relevant open habitat types was zero, no increase in forest cover was possible.

The predicted abundance indices for each of these woodland bird species were categorised as 'hot', 'warm' or 'cold' using the same criteria as for Curlew (see Section 2.3.3). However for the predicted abundances in response to increases in woodland cover, the 'heat' of each stratum was based on the proportions of the 99.5th percentile predicted abundance during *BA2010* at a national scale.

3. RESULTS

3.1 Predictive capabilities of breeding wader abundance models

The predictive capabilities of the models performed best for Curlew, then Oystercatcher, Golden Plover, Lapwing, Redshank, Common Sandpiper, Snipe, Ringed Plover with those for Dunlin being the least satisfactory. Predictive performance was assessed by the proportion of variation explained by the factors included in the models (R^2 values) and by the correlations of presence or absence between predicted and empirical data where this was available (Table 2).

The relative importance of variables included in the models, as indicated by their inclusive node purity, varied between species (Table 3). Factors that were important for at least some breeding waders included metrics of habitat type (e.g. grassland type), soil chemistry (soil carbon), predation risk (Red Grouse and Crows), weather (rainfall) and geographic location (latitude and longitude). The relative predictive importance of these variables models does not necessarily reflect the relative strengths of any causal effects. Each variable is inter-correlated with many others – including other model variables, as well as variation not explicitly account for in the models. In particular, it should be noted that rainfall may act as a surrogate for slope, which was omitted from the analysis due to its high correlation with rainfall.

3.2 Strata of relative importance for breeding waders

Maps depicting strata of varying importance for breeding waders were produced for seven species (Figs. 3 – 9). Dunlin and Ringed Plover were not recorded within the PFAs during *BA2010.* Strata based on national and regional importance were broadly similar within the PFAs but with a few additional squares qualifying for higher categories of importance within the regional categories for Lapwing and Redshank). Despite adopting criteria for categorising strata based on local knowledge of breeding waders, Common Sandpiper remained likely over-predicted and Snipe under-predicted.

3.3 Changes in the status of breeding waders

Based on differences in probabilities of occupancy, most breeding waders had declined in abundance between *BA1990* and *BA2010* in the PFAs and surrounding areas (Figs. 10 – 18). Note that maps of changing status should always be interpreted along with those of their most recent known status. For example, no change in abundance could indicate that birds are still present, potentially even at high densities, or equally that they never have been expected to occur in such areas.

3.4 Forest expansion scenarios and woodland birds

Scenarios of woodland expansion within the Borders PFAs were associated with increases in the abundance and distribution of the three example woodland-associated species (Figs. 19 – 21). While noting that the models on which the predicted scenarios are based may be less reliable than those for at least some of the waders (see Section 4.3), they do illustrate that woodland expansion does offer opportunities for some species, including those considered to be conservation priorities.

4. DISCUSSION

4.1 Predictive capabilities of wader models

Data collection for the bird atlases aimed to sample most species present within a representative cross-section of available habitats. They were not specifically surveys of breeding waders or of forest birds. They were also not comprehensive surveys that aimed to accurately record density or occurrence of birds at relatively fine scale resolutions. Hence there was a need to model associations of bird atlas data with environmental data sets to produce maps of likely (predicted) distribution and abundance of birds at resolutions that are likely to be of use to those planning changes of land use such as forest expansion.

The predictive capabilities of the models trained on data sets for the whole of mainland Britain for breeding waders were greatest for Curlew and Oystercatcher and least for Ringed Plover and Dunlin. The performance of the models will, in part, depend on the reliability of an observer detecting the species in question during one-hour survey visits to a tetrad. Curlew and Oystercatcher are likely to have been among the species most readily detected where present because they are relatively conspicuous both visually and aurally (Grant *et al.* 2000, Wilson & Browne 1999). Species for which the models performed least well included species that are known to be challenging to survey (e.g. Snipe; Green 1985, Hoodless *et al.* 2006), are associated with restricted habitats (e.g. Common Sandpiper and Ringed Plover; Dougall *et al.* 2010, Conway *et al.* 2019) or both (e.g. Dunlin; Grant & Pearce-Higgins 2012). Therefore, when using the outputs of the models, greater emphasis should be given to those deriving from better performing models. In considering lowland or enclosed farmland, the heat maps for Curlew and Oystercatcher, potentially supplemented with those for Lapwing and Redshank, are likely to best represent the relative importance of areas for breeding waders. For more upland and unenclosed areas the heat maps for Curlew and Golden Plover potentially provide the most reliable information.

4.2 Use of wader model outputs

The model outputs for waders are presented as heat maps representing three strata of relative importance, with alternatives based on national importance (as assessed for all of mainland Britain) and regional importance (assessed for an area including a 50 km radius around the Borders PFAs). These maps can be used to guide where positive management for breeding waders may be most beneficial, and to inform the planning and assessment of alternative land uses that would potentially be detrimental to breeding waders, notably the creation of new forests. As well as outputs based on the most recent atlas period (centred around 2010), modelled change in wader distribution and abundance in the 20 year period since 1990 are also presented. The maps of change could also be used to target management for waders by identifying areas that likely supported breeding waders relatively recently and could potentially do so with appropriate treatment.

To guide plans and applications for forest expansion, potential heat-based interpretations of the three categories of relative importance could be:

> *Hot areas* – Areas within which plans for afforestation should be accompanied with detailed surveys of breeding waders to confirm, or otherwise, their importance;

Warm areas – Areas within which plans for afforestation should be accompanied with some surveys of breeding waders to confirm, or otherwise, their importance;

Cool areas – Areas within which plans for afforestation may require minimal additional supporting information of breeding waders.

It is important to note that the above three categories and possible implications associated with them are offered here as examples only. Guidance will likely be required on any implications for planning forests in areas that supported breeding waders recently, as suggested by change maps or by new survey data. When considering plans for forests, the area of influence will likely exceed that of the actual footprint of the affected area (Wilson *et al.* 2013) and there are also likely to be cumulative effects of habitat fragmentation through multiple planned forests (Douglas *et al.* 2013). However, the scale at which such effects operate is poorly understood. While acknowledging that the applications of the wader heat maps requires further clarification and potential refinement, consideration should also be given to how this information is integrated with other data on an area's suitability for afforestation.

A number of concurrent and related modelling studies are exploring the use of bird atlas and environmental data sets to inform wader conservation (and relatedly inform forest expansion) elsewhere in Scotland and northern England. Each would benefit from feedback from all areas before release of any maps and associated shapefiles to be used widely in guiding forest planning or other land-use and management applications.

The aim is to repeat the Bird Atlases at 20-year intervals with fieldwork for next one centred around 2030. At that point these maps will be able to be refreshed and updated, ensuring their continued relevance and utility.

4.3 Use of forest bird model outputs

As well as constraints for biodiversity reliant on open habitats, forest expansion offers opportunities for species and assemblages reliant on woodlands. Considering the outcomes of forest expansion as a constraint on breeding waders through the replacement of habitats with others that are broadly unsuitable is relatively straightforward, though could benefit from a better understanding of how some constraints operate (see above). However, understanding the opportunities for birds presented by forest expansion requires a much better understanding of the influences of tree (including crop) species composition and silvicultural treatments (Calladine *et al.* 2018, Fuller & Robles 2018). The examples of predicted changes arising from modelled afforestation scenarios are presented as examples only and are unlikely to accurately predict the real-world consequences of afforestation. To be more reliable, such models will require better information on species-specific associations with woodland types and structures and, crucially, more detailed information on woodland structure to include in the models. Nevertheless, the maps of predicted changes do illustrate the potential for forest expansion to provide opportunities for some conservation priority species. Further work to develop these models into tools that could be used to inform forest planning and ongoing management plans will likely require the collection and/or preparation of additional data to supplement those already available in bird atlas and environmental data sets.

ACKNOWLEDGEMENTS

The work was funded by Scottish Forestry with Colin Edwards the nominated officer who along with David Douglas (RSPB), Jeanette Hall (SNH) and Andy Tharme (Borders Council) formed an advisory group. We are grateful to Harry Dott and Martin Moncrieff for sharing local knowledge of the areas and to Mark Holling for facilitating that process. The analyses draw heavily on earlier work undertaken by Jenni Border (BTO). Bird Atlases were a partnership between BTO, the SOC and Birdwatch Ireland and we indebted to the many volunteers who contributed to the fieldwork.

REFERENCES

Balmer, D.E., Gillings, S., Caffrey, B.J., Swann, R.L., Downie, I.S. & Fuller, R.J. (2013). *Bird Atlas 2007–11: the Breeding and Wintering Birds of Britain and Ireland*. BTO, Thetford.

Brambilla, M., Casale, F., Bergero, V., Crovetto, G. M., Falco, R., Negri, I. & Bogliani, G. (2009) GIS-models work well, but are not enough: Habitat preferences of *Lanius collurio* at multiple levels and conservation implications. *Biological Conservation* **142**: 2033-2042.

Breiman, L. (2001) Machine Learning. 2001. 45: 5. https://doi.org/10.1023/A:1010933404324

Calladine, J., Humphreys, E.M., Gilbert, L., Furness, R.W., Robinson, R.A., Fuller, R.J., Littlewood, N.A., Pakeman, R.J., Ferguson, J. & Thompson, C. (2017). Continuing influences of introduced hedgehogs *Erinaceus europaeus* as a predator of wader (Charadrii) eggs four decades after their release on the Outer Hebrides, Scotland. *Biological Invasions* **19**: 1981-1987

Calladine, J., Díaz, M., Reino, L., Jardine, D. & Wilson, M. (2018). Plantations of non-native tree species: In: Mikusinski, G., Roberge, J.M. & Fuller, R.J. (eds) *Ecology and Conservation of Forest Birds*. Cambridge University Press, Cambridge. Pp 350 – 386.

Calladine, J., Díaz, M., Reino, L., Jardine, D. & Wilson, M. (2018). Plantations of non-native tree species: In: Mikusinski, G., Roberge, J.M. & Fuller, R.J. (eds) *Ecology and Conservation of Forest Birds*. Cambridge University Press, Cambridge. Pp 350 – 386.

Calladine, J., Jarrett, D. & Wilson, M. (2019). Breeding bird assemblages supported by developing upland shrub woodland are influenced by micro-climate and habitat structure. *Bird Study* **66**: 178-186.

Conway G.J., Austin G.E., Handschuh M., Drewitt A.L. & Burton N.H.K. (2019) Breeding populations of Little Ringed Plover *Charadrius dubius* and Ringed Plover *Charadrius hiaticula* in the United Kingdom in 2007. *Bird Study* **66**: 22-31.

Dougall, T.W., Holland, P. K. & Yalden, D.W. (2010) The population biology of Common Sandpipers in Britain. *British Birds* **103**: 100-114.

Douglas, D.J.T., Bellamy, P.E., Stephen, L.S., Pearce-Higgins, J.W., Wilson, J.D. & Grant, M.C. (2013). Upland land use predicts population decline in a globally near-threatened wader. *Journal of Applied Ecology*, **51:** 194-203.

Fletcher, K., Aebischer, N.J., Baines, D., Foster, R. & Hoodless, A.N. (2010). Changes in breeding success and abundance of ground-nesting moorland birds in relation to the experimental deployment of legal predator control. *Journal of Applied Ecology* **47**: 263–272.

Franks, S.E., Douglas, D.J.T., Gillings, S. & Pearce-Higgins, J.W. 2017. Environmental correlates of breeding abundance and population change of Eurasian Curlew *Numenius arquata* in Britain. *Bird Study* **64**: 393-409.

Fuller, R.J. & Robles, H. (2019). Conservation strategies and habitat management for European forest birds. In: Mikusinski, G., Roberge, J.M. & Fuller, R.J. (eds) *Ecology and Conservation of Forest Birds*. Cambridge University Press, Cambridge. Pp 454 – 507.

Gibbons, D.W., Reid, J.B. & Chapman, R.A. (1993) *The New Atlas of Breeding Birds in Britain and Ireland: 1988–1991.* T. & A.D. Poyser, London.

Gillings, S., Fuller, R.J. & Balmer, D.E. (2000) Breeding birds in scrub in the Scottish Highlands: Variation in community composition between scrub type and successional stage. *Scottish Forestry,* **54:** 73-85.

Grant, M.C., Lodge, C., Moore, N., Easton, J., Orsman, C. & Smith, M. (2000) Estimating the abundance and hatching success of breeding Curlew *Numenius arquata* using survey data. *Bird Study* **47:** 41–51.

Grant, M.C. & Pearce-Higgins, J.W. (2012) Spatial variation and habitat relationships in moorland bird assemblages: a British perspective. Pp 207-236 in: Fuller. R.J. ed. *Birds and Habitat: relationships in Changing Landscapes.* Cambridge University Press, Cambridge.

Green, R.E. (1985) Estimating the abundance of breeding snipe. *Bird Study*, **32**: 141-149.

Hoodless, A.N., Inglis, J.G., & Baines, D. (2006) Effects of weather and timing on counts of breeding snipe *Gallinago gallinago*. *Bird Study,* **53**: 205-212.

Jarvis A., Reuter, H.I., Nelson, A., Guevara, E. 2008. *Hole-filled seamless SRTM data V4*, International Centre for Tropical Agriculture (CIAT), available from http://srtm.csi.cgiar.org

Jones, R.J., Hiederer, R., Rusco, E., Loveland, P.J. & Montanarella, L. (2003). *The map of organic carbon content in topsoils In Europe: v 1.2 September - 2003*. Luxembourg: Office for Official Publications of the European Communities, European Soil Bureau Research Report No. 15.

Liaw, A. & Wiener, M. 2002. Classification and Regression by randomForest. *R News* **2**: 18—22.

Maleki, S., Soffianian, A. R., Koupaei, S. S., Saatchi, S., Pourmanafi, S., & Sheikholeslam, F. 2016. Habitat mapping as a tool for water birds conservation planning in an arid zone wetland: The case study Hamun wetland. *Ecological Engineering*, **95**, 594-603.

Massimino D., Harris S.J. & Gillings, S. 2018. Evaluating spatiotemporal trends in terrestrial mammal abundance using data collected during bird surveys. *Biological Conservation* **226**: 153-167.

Murray, R.D., Andrews, I.J. & Holling, M. 2019. *Birds in South-east Scotland 2007-13: a tetrad atlas of the birds of Lothian and Borders.* The Scottish Ornithologists' Club, Aberlady.

Perry, M. & Hollis, D. 2005. The generation of monthly gridded datasets for a range of climatic variables over the UK. International Journal of Climatology, **25**: 1041–1054.

Rowland, C.S., Morton, R.D., Carrasco, L., McShane, G., O'Neil, A.W., and Wood, C.M. 2017. *Land Cover Map 2015 (1km percentage aggregate class, GB).* NERC Environmental Information Data Centre.

Scottish Government 2019 *Scotland's Forest Strategy 2019 – 2029.* Scottsih Government, Edinburgh.

Tharme, A.P., Green, R.E., Baines, D., Bainbridge, I.P. and O'Brien, M. 2001. The effect of management for Red Grouse shooting on the population density of breeding birds on heather-dominated moorland. *Journal of Applied Ecology* **38**: 439-457.

Wilson, A.M. & Browne, S.J. (1999). Breeding population estimates for Lapwing, Oystercatcher and Curlew in Scotland: results of the 1998 BTO Lapwing Survey. *Scottish Birds*, **20**: 73-80.

Wilson, J.D., Anderson, R., Bailey, S., Chetcuti, J., Cowie, N.R., Hancock, M.H, Quine, C.P., Russel, N., Stephen, L. & Thompson, D.B.A. (2013). Modelling edge effects of mature forest plantations on peatland waders informs landscape-scale conservation. *Journal of Applied Ecology* **51:** 204-213.

TABLES

Table 1. The proportion of TTV counts across mainland Britain that were greater than 18 (and therefore capped to reduce the influence of non-breeding flocks on modelled outputs) and the median counts with which those higher counts were replaced.

Common name	**Scientific name**	**Proportion of TTV counts > 18**	**Median count**
Curlew	*Numenius arquata*	97%	2
Lapwing	*Vanellus vanellus*	94%	2
Oystercatcher	*Haemotopus ostralegus*	93%	2
Redshank	*Tringa totanus*	95%	2
Golden Plover	*Pluvialis apricaria*	93%	2
Dunlin	*Calidris alpina*	88%	2
Ringed Plover	*Charadrius hiaticula*	95%	2

Table 2. Measures of the predicative capabilities of models trained on *BA2010* data for mainland Britain. R^2 indicates the proportion of variation in indices of abundance explained by the models variables. Correlation shows

Species	R^2	Correlation
Curlew	0.41	0.64
Lapwing	0.22	0.47
Oystercatcher	0.28	0.53
Redshank	0.19	0.43
Snipe	0.14	0.37
Golden Plover	0.25	0.50
Dunlin	0.04	0.26
Common Sandpiper	0.16	0.41
Ringed Plover	0.07	0.29

Table 3. The inclusive node purity for variables included within models to predict breeding wader abundance. The greater the value of node purity is an indicator for greater importance in the model. The three most important variables for predicted abundance of each species are presented in bold.

Variables	Common Sandpiper	Curlew	Dunlin	Golden Plover	Lapwing	Oystercatcher	Redshank	Ringed Plover	Snipe
Buzzard	451	3250	76	937	8983	4491	1085	398	449
Built-up area	86	1342	10	92	4745	2276	641	208	99
Crow	**559**	3826	**149**	995	8429	4214	916	479	342
Fox	438	3450	**198**	847	8807	3715	882	342	**466**
Improved grassland	204	3304	18	410	7740	3875	606	318	270
Mountain, heath and bog	449	2986	69	**1187**	3903	1387	177	109	295
Organic soil carbon	474	3303	89	1063	7786	3860	806	524	458
Red Grouse	**615**	**10398**	78	**2505**	**9849**	4112	**1157**	**553**	710
Raven	459	3010	67	981	8953	3721	1052	377	429
Road	56	282	1	48	1995	816	116	57	33
Semi-natural grassland	282	2494	19	584	5849	2188	600	262	**482**
Wind-farms	4	134	2	69	298	86	7	14	22
Arable	192	1464	6	28	5046	2040	514	216	147
Broadleaved woodland	234	1571	8	252	6011	2328	525	251	235
Woodland clearings	78	308	1	40	914	666	109	13	51
Conifer woodland	160	1118	2	99	3043	1594	186	51	160
Easting	454	**5241**	95	875	**9423**	4158	**1374**	414	297
Forest heterogeneity	166	1297	12	146	3944	1779	364	133	157
Mixed woodland	133	747	1	122	2385	1222	309	34	155
Northing	**740**	**8665**	**129**	**1129**	**13202**	**7903**	854	432	341
Forest edge	193	505	3	43	1075	1905	124	26	252
Forest understorey	133	758	1	83	2534	1104	210	72	111
Summer rainfall	435	3564	65	716	8487	**7461**	**1411**	**580**	**535**
Summer temperature	437	3891	87	779	7739	**7744**	977	**560**	343

FIGURES

Figure 1. Location of the Borders Pilot Forest Areas (PFAs) in southern Scotland

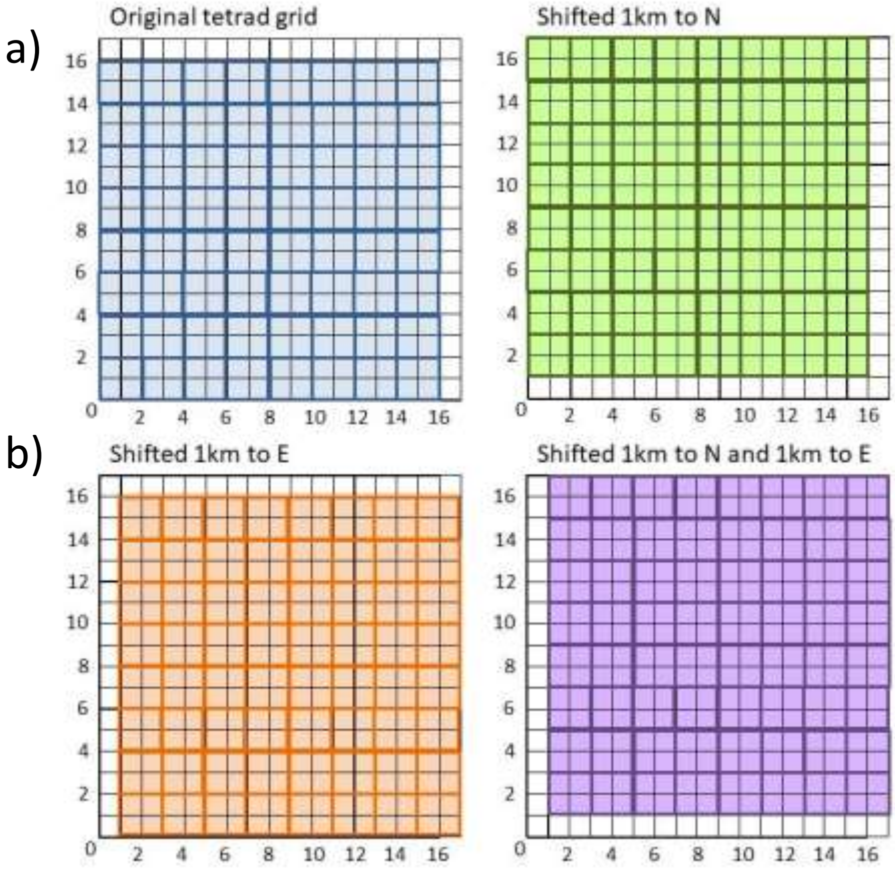


Figure 2. In order to generate 1km resolution estimates from tetrad level models, a) model-based predictions were generated using not only the original tetrad-based prediction data, but also using explanatory datasets summarised on shifted tetrad grids. The predicted abundance in each 1km square was then calculated as b) the mean abundance of the four tetrads overlapping the square.

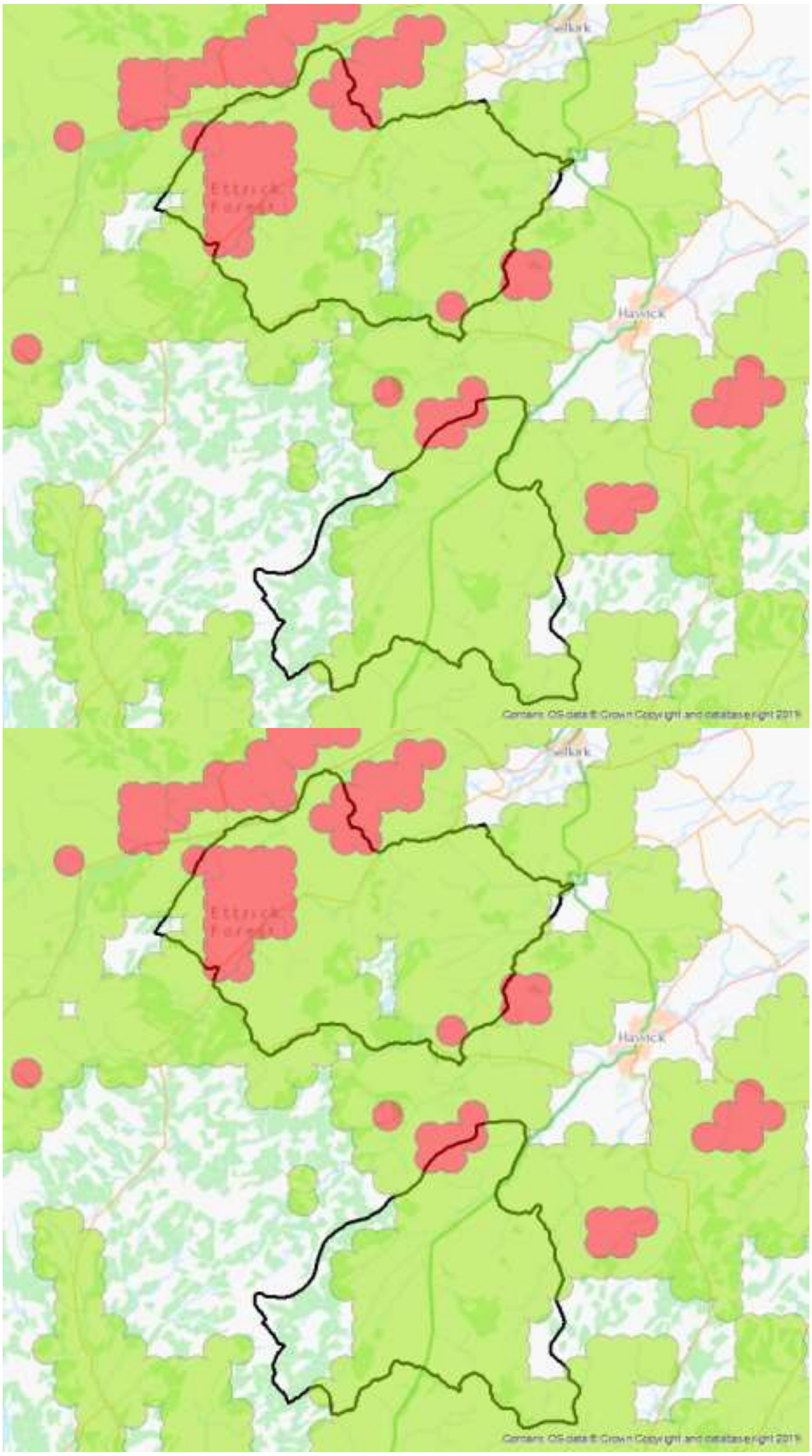


Figure 3. Modelled heat maps for breeding **Curlew** in the area of the Borders PFAs. Red areas encompass 1-km squares predicted to include the most important ('hot') areas. Green areas encompass 1-km squares predicted to include fairly important ('warm') areas. The top map is based on relative importance across mainland Britain. The lower map is based on relative importance within 50 km of the PFAs.

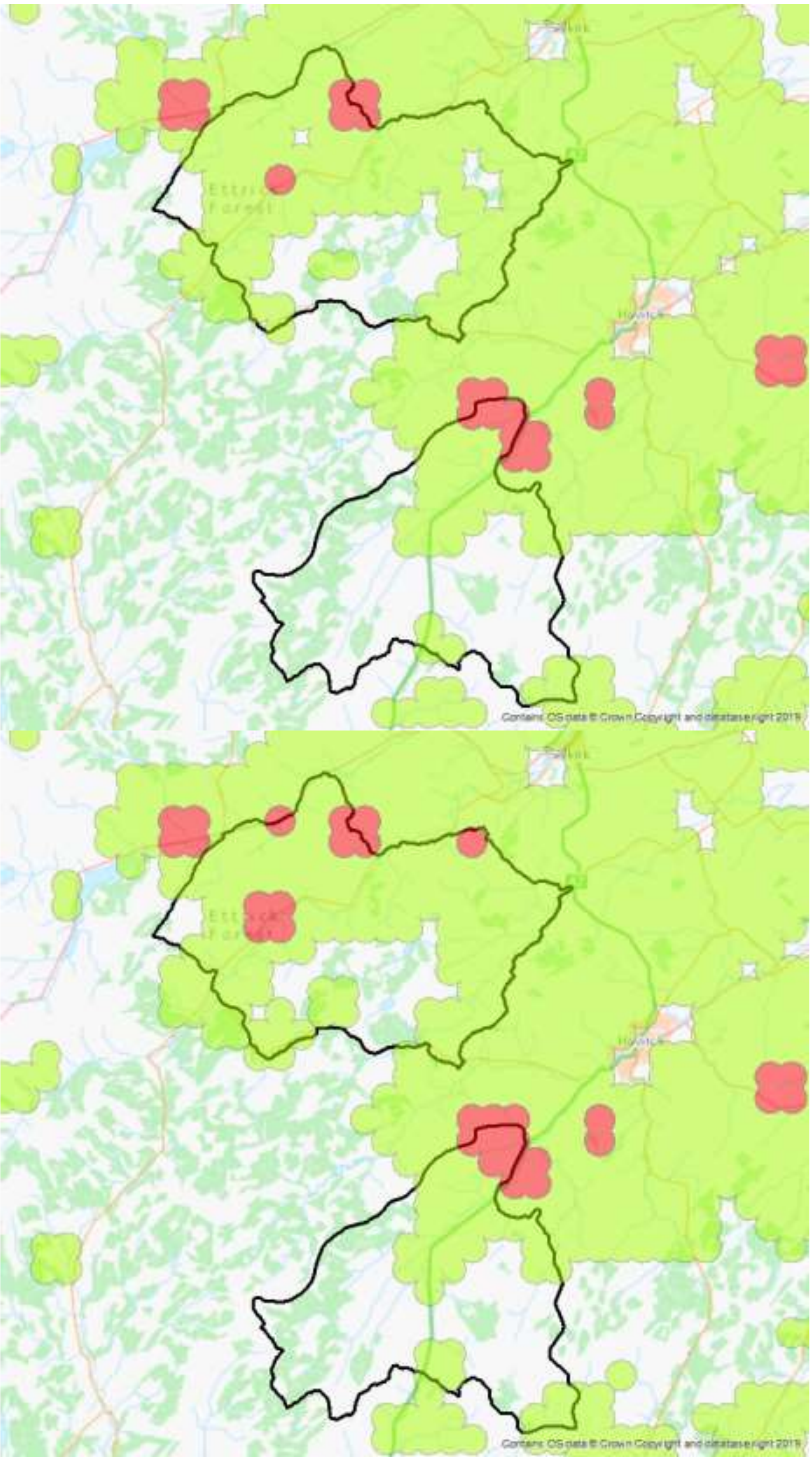


Figure 4. Modelled heat maps for breeding **Lapwing** in the area of the Borders PFAs. Red areas encompass 1-km squares predicted to include the most important ('hot') areas. Green areas encompass 1-km squares predicted to include fairly important ('warm') areas. The top map is based on relative importance across mainland Britain. The lower map is based on relative importance within 50 km of the PFAs.

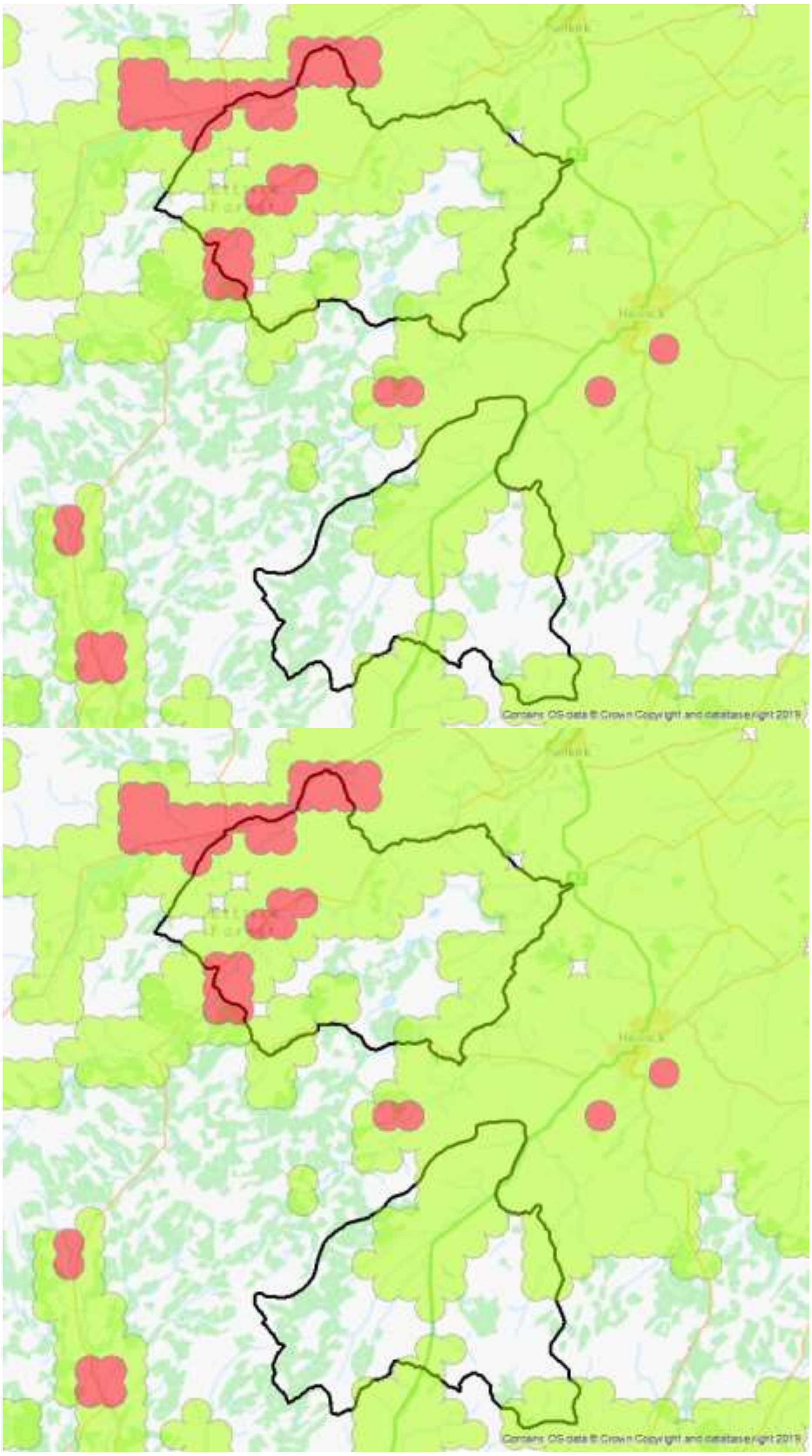


Figure 5. Modelled heat maps for breeding **Oystercatcher** in the area of the Borders PFAs. Red areas encompass 1-km squares predicted to include the most important ('hot') areas. Green areas encompass 1-km squares predicted to include fairly important ('warm') areas. The top map is based on relative importance across mainland Britain. The lower map is based on relative importance within 50 km of the PFAs.

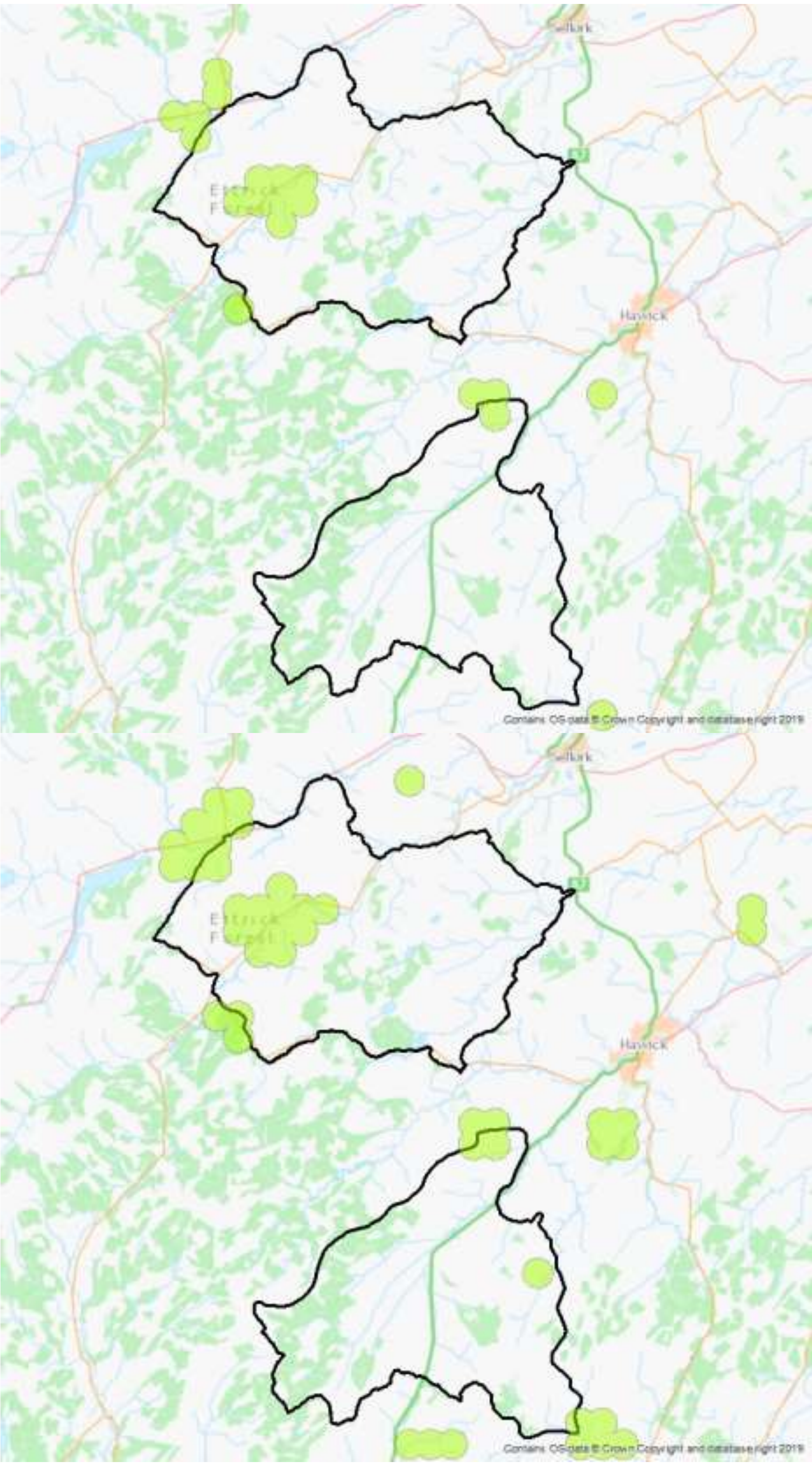
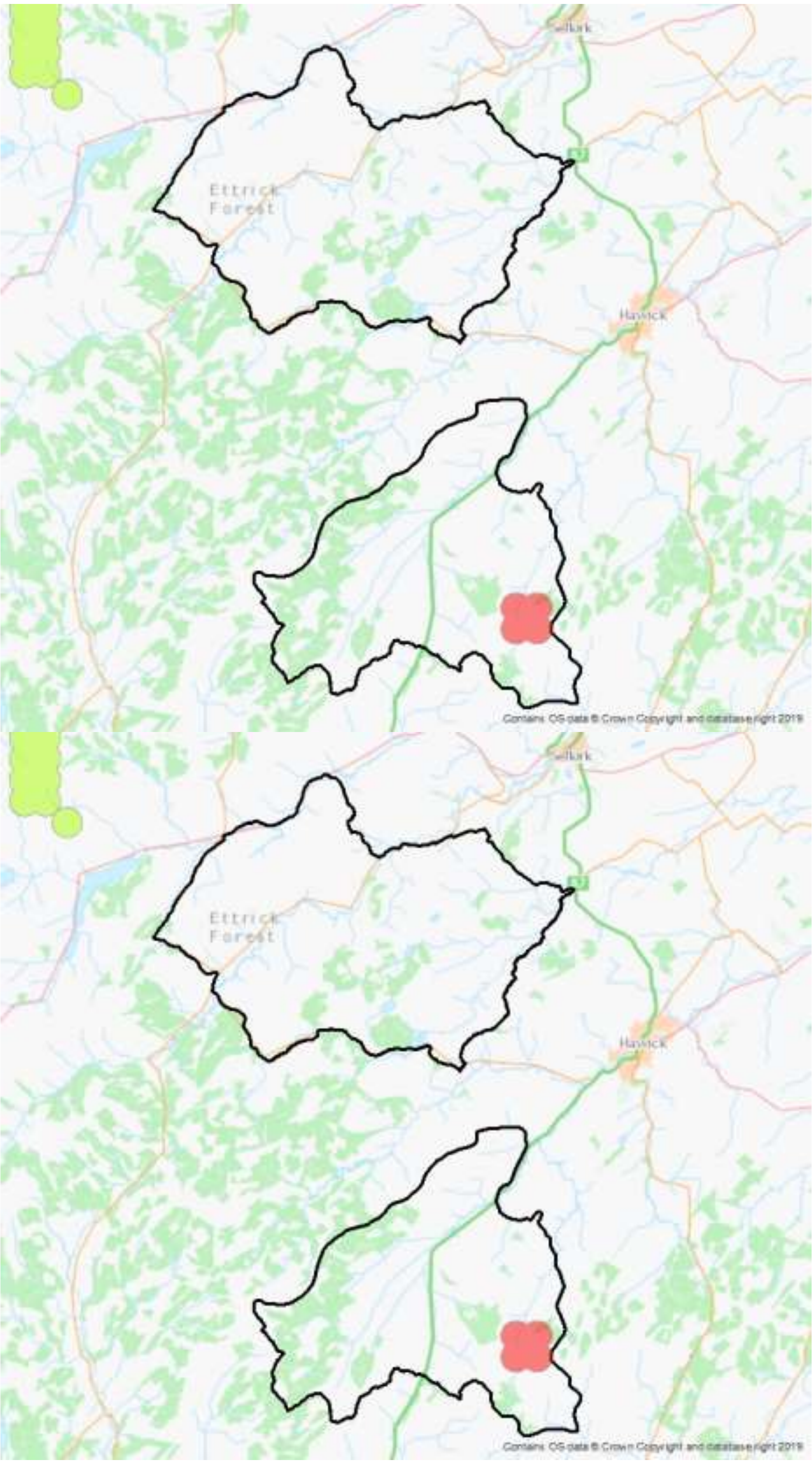


Figure 6. Modelled heat maps for breeding **Redshank** in the area of the Borders PFAs. Red areas encompass 1-km squares predicted to include the most important ('hot') areas. Green areas encompass 1-km squares predicted to include fairly important ('warm') areas. The top map is based on relative importance across mainland Britain. The lower map is based on relative importance within 50 km of the PFAs.



Contains OS data © Crown Copyright and database right 2019

Contains OS data © Crown Copyright and database right 2019

Figure 7. Modelled heat maps for breeding **Golden Plover** in the area of the Borders PFAs. Red areas encompass 1-km squares predicted to include the most important ('hot') areas. Green areas encompass 1-km squares predicted to include fairly important ('warm') areas. The top map is based on relative importance across mainland Britain. The lower map is based on relative importance within 50 km of the PFAs.

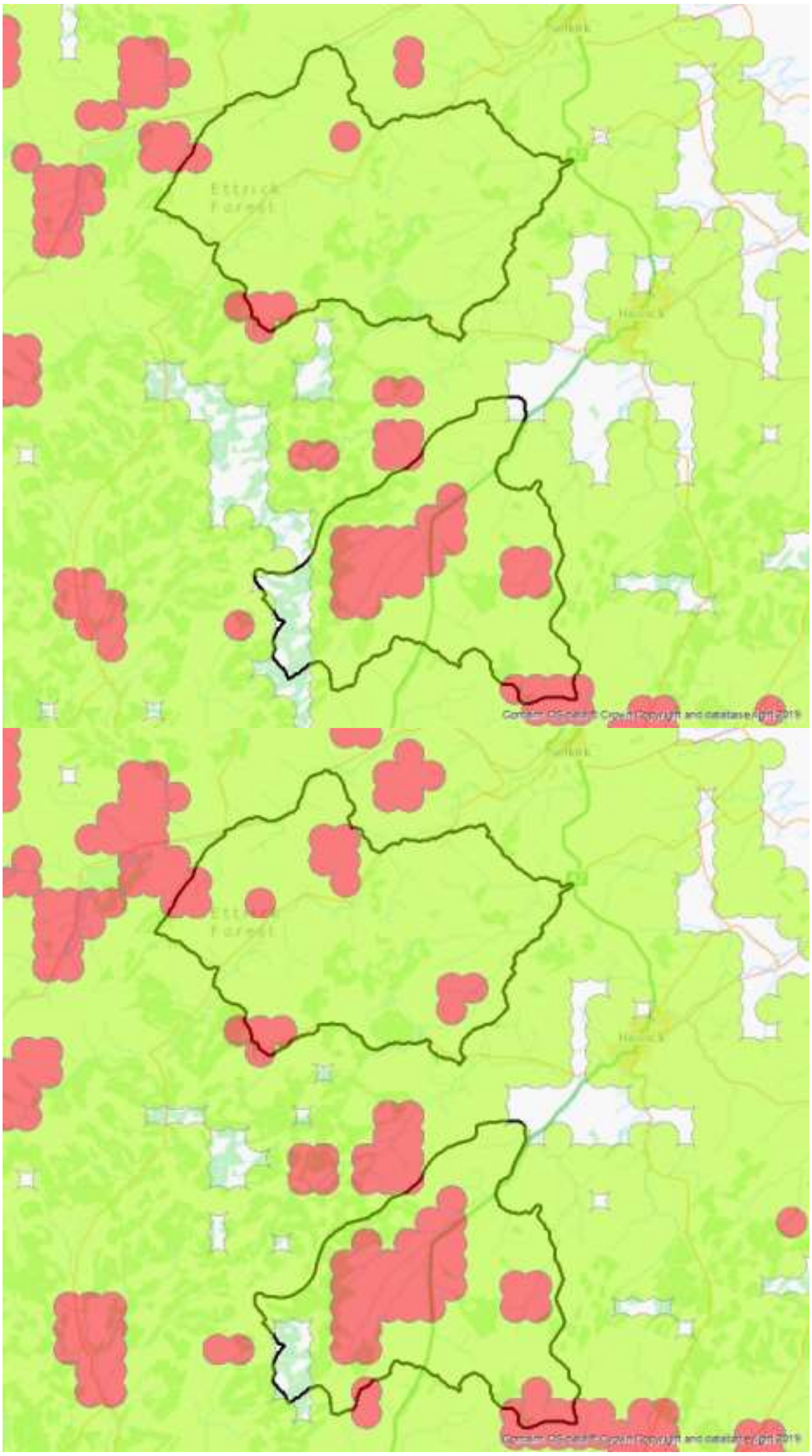


Figure 8. Modelled heat maps for breeding **Common Sandpiper** in the area of the Borders PFAs. Red areas encompass 1-km squares predicted to include the most important ('hot') areas. Green areas encompass 1-km squares predicted to include fairly important ('warm') areas. The top map is based on relative importance across mainland Britain. The lower map is based on relative importance within 50 km of the PFAs.

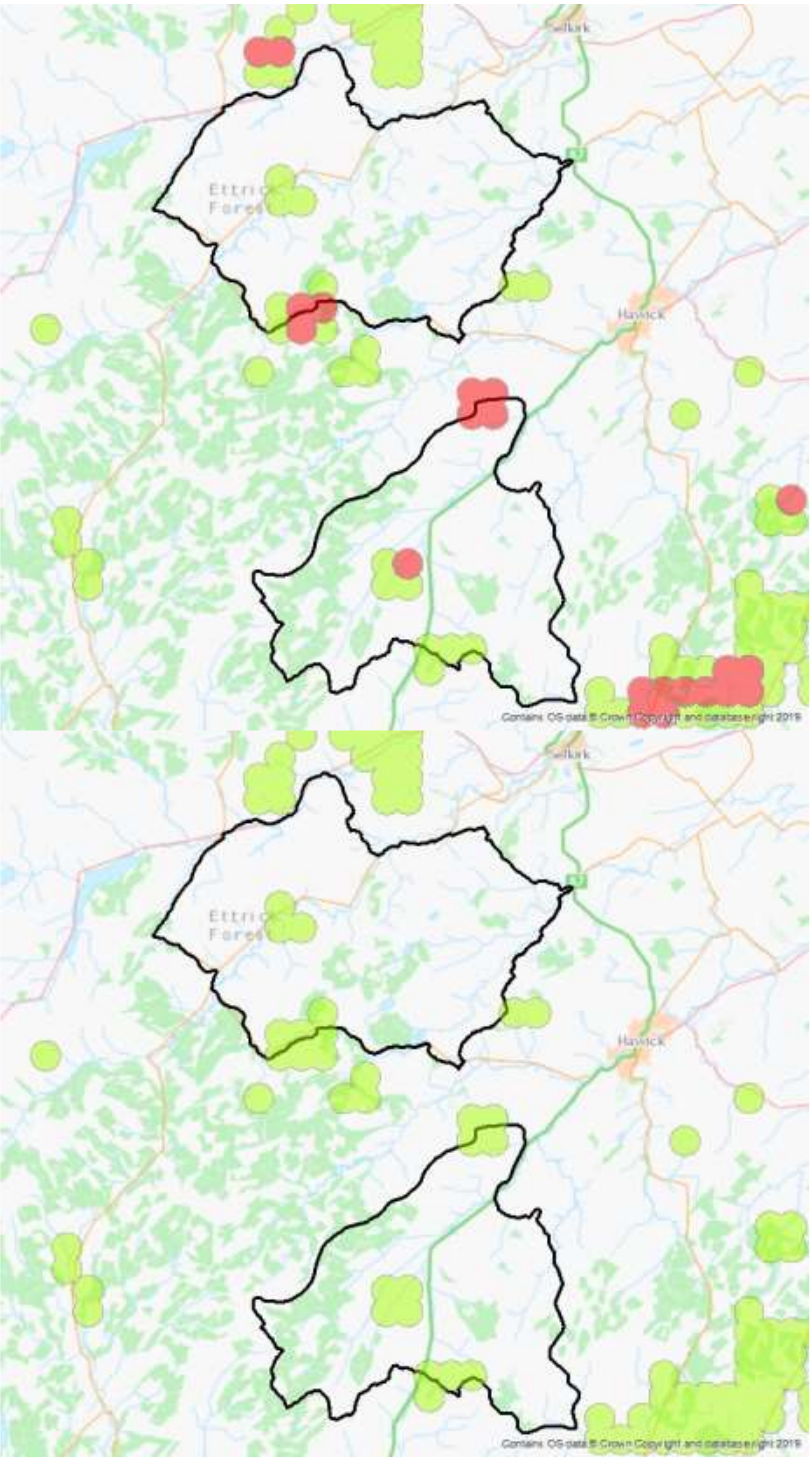


Figure 9. Modelled heat maps for breeding **Snipe** in the area of the Borders PFAs. Red areas encompass 1-km squares predicted to include the most important ('hot') areas. Green areas encompass 1-km squares predicted to include fairly important ('warm') areas. The top map is based on relative importance across mainland Britain. The lower map is based on relative importance within 50 km of the PFAs.

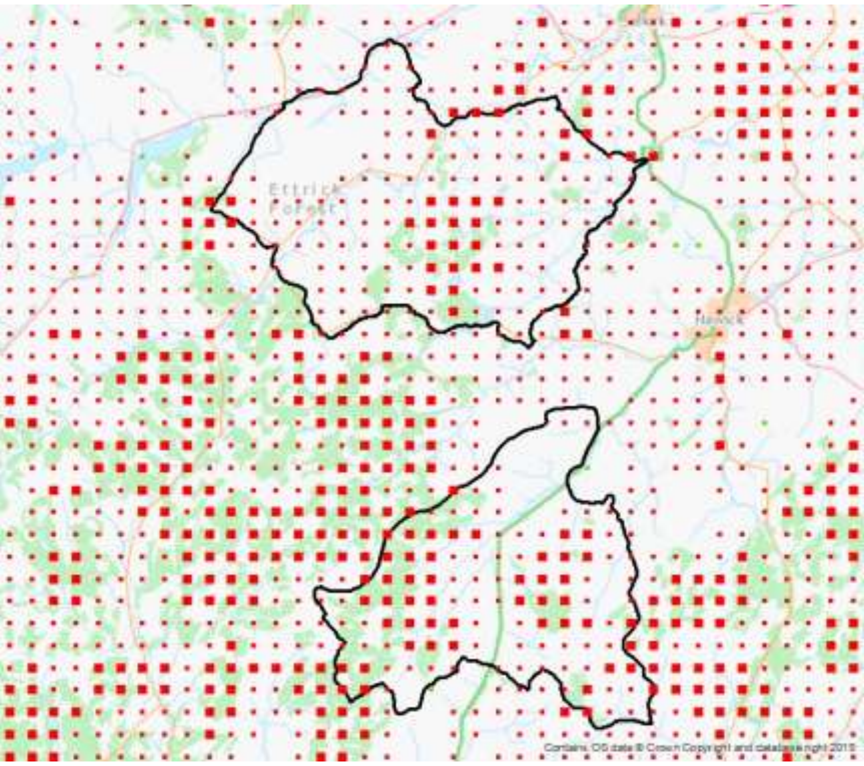


Figure 10. Predicted changes in abundance by breeding **Curlew** between *BA1990* and *BA2010* at 1-km resolution. *Red* indicates where the probability of occurrence decreased. *Green* indicates where the probability of occurrence increased. Larger symbols represent predicted changes > 40%. Smaller symbols represent predicted changes > 10% but < 40%.

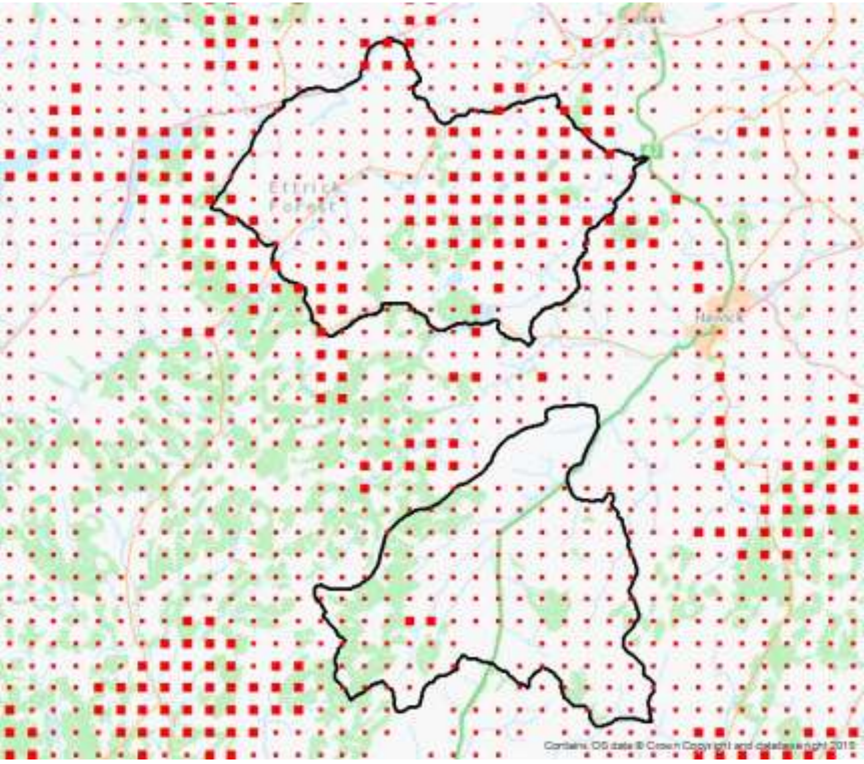


Figure 10. Predicted changes in abundance by breeding **Lapwing** between *BA1990* and *BA2010* at 1-km resolution. *Red* indicates where the probability of occurrence decreased. *Green* indicates where the probability of occurrence increased. Larger symbols represent predicted changes > 40%. Smaller symbols represent predicted changes > 10% but < 40%.

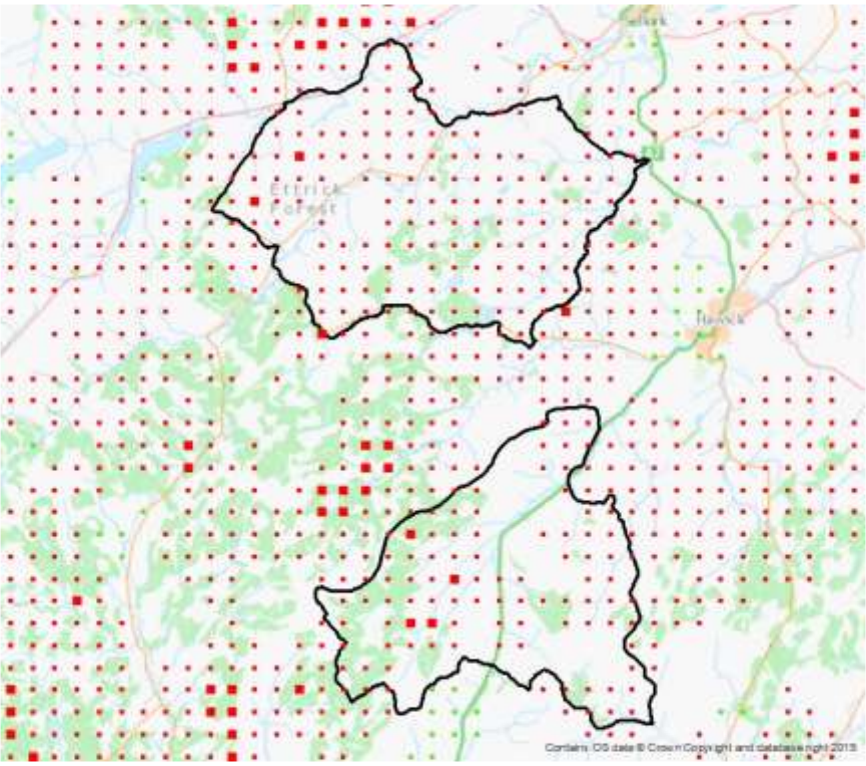


Figure 12. Predicted changes in abundance by breeding **Oystercatcher** between *BA1990* and *BA2010* at 1-km resolution. *Red* indicates where the probability of occurrence decreased. *Green* indicates where the probability of occurrence increased. Larger symbols represent predicted changes > 40%. Smaller symbols represent predicted changes > 10% but < 40%.

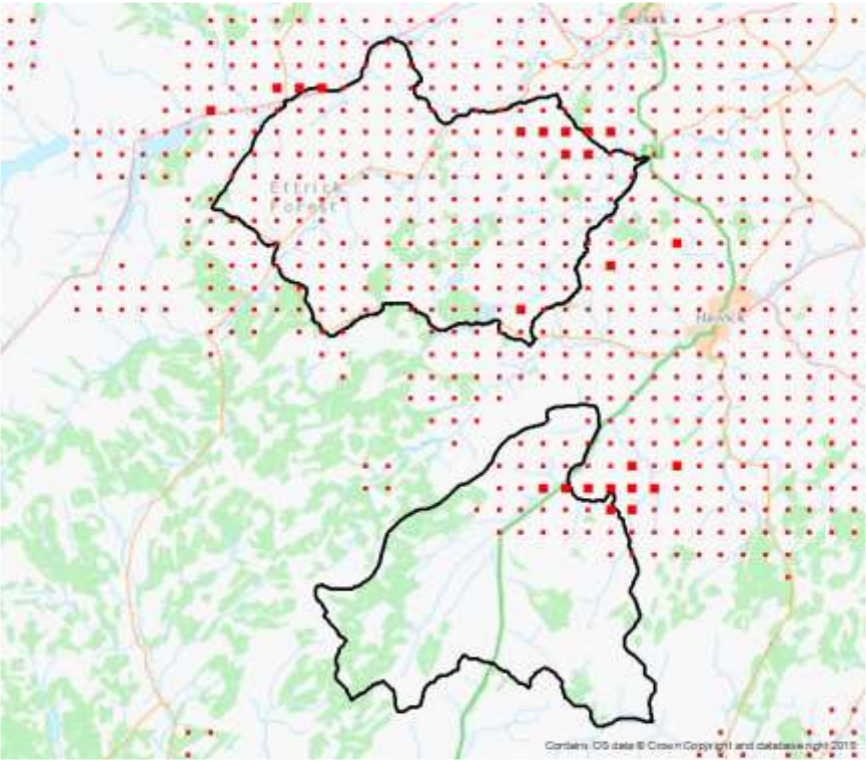


Figure 13. Predicted changes in abundance by breeding **Redshank** between *BA1990* and *BA2010* at 1-km resolution. *Red* indicates where the probability of occurrence decreased. *Green* indicates where the probability of occurrence increased. Larger symbols represent predicted changes > 40%. Smaller symbols represent predicted changes > 10% but < 40%.

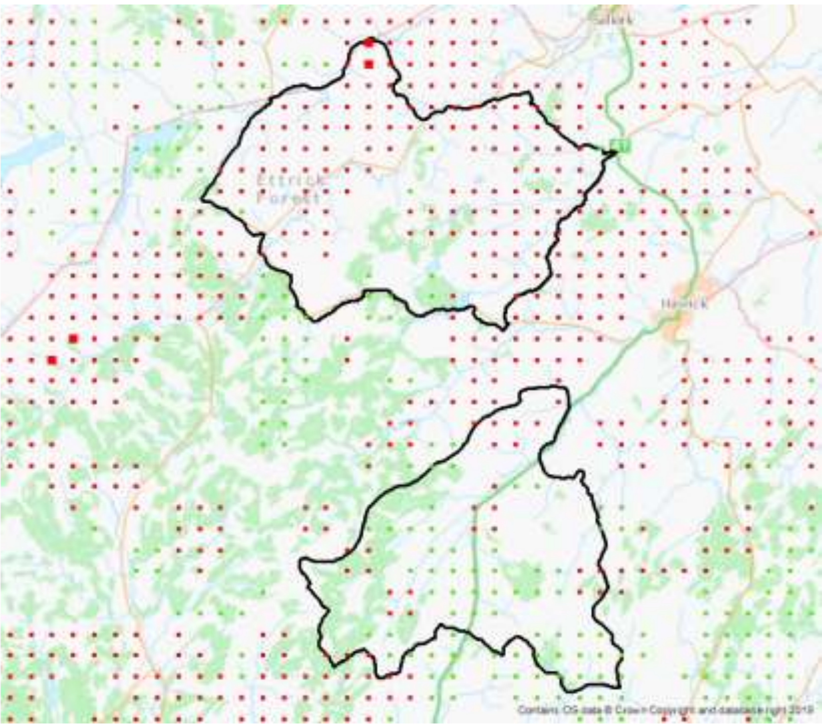


Figure 14. Predicted changes in abundance by breeding **Snipe** between *BA1990* and *BA2010* at 1-km resolution. *Red* indicates where the probability of occurrence decreased. *Green* indicates where the probability of occurrence increased. Larger symbols represent predicted changes > 40%. Smaller s symbols represent predicted changes > 10% but < 40%.

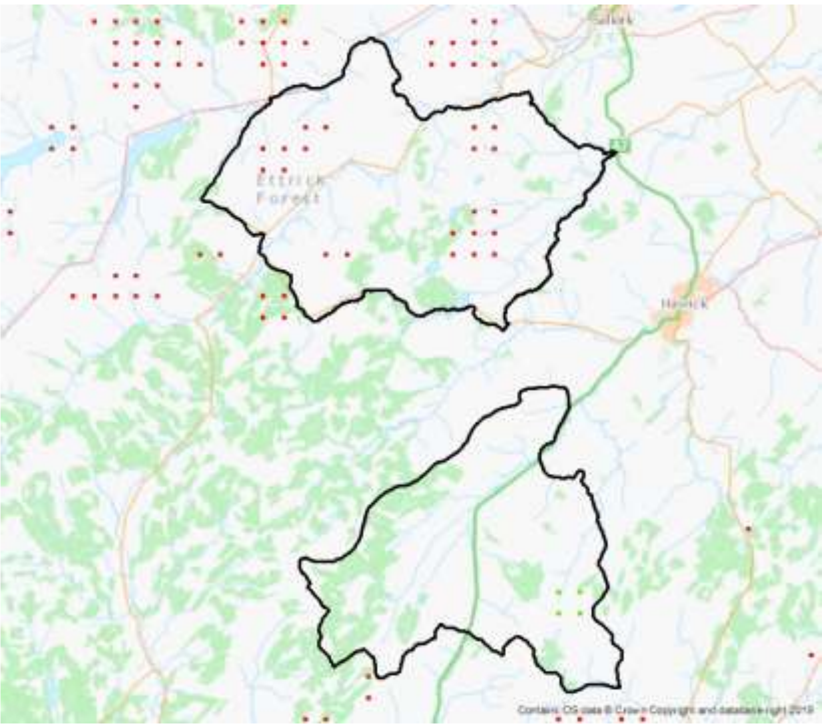


Figure 15. Predicted changes in abundance by breeding **Golden Plover** between *BA1990* and *BA2010* at 1-km resolution. *Red* indicates where the probability of occurrence decreased. *Green* indicates where the probability of occurrence increased. Larger symbols represent predicted changes > 40%. Smaller symbols represent predicted changes > 10% but < 40%.

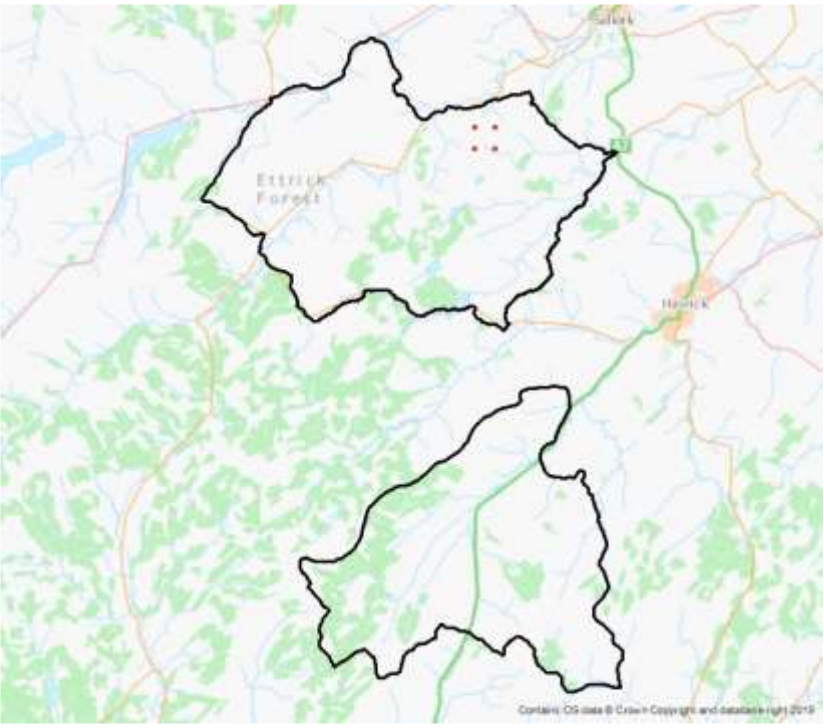


Figure 16. Predicted changes in abundance by breeding **Dunlin** between *BA1990* and *BA2010* at 1-km resolution. *Red* indicates where the probability of occurrence decreased. *Green* indicates where the probability of occurrence increased. Larger symbols represent predicted changes > 40%. Smaller symbols represent predicted changes > 10% but < 40%.

Figure 17. Predicted changes in abundance by breeding **Ringed Plover** between *BA1990* and *BA2010* at 1-km resolution. *Red* indicates where the probability of occurrence decreased. *Green* indicates where the probability of occurrence increased. Larger symbols represent predicted changes > 40%. Smaller symbols represent predicted changes > 10% but < 40%.

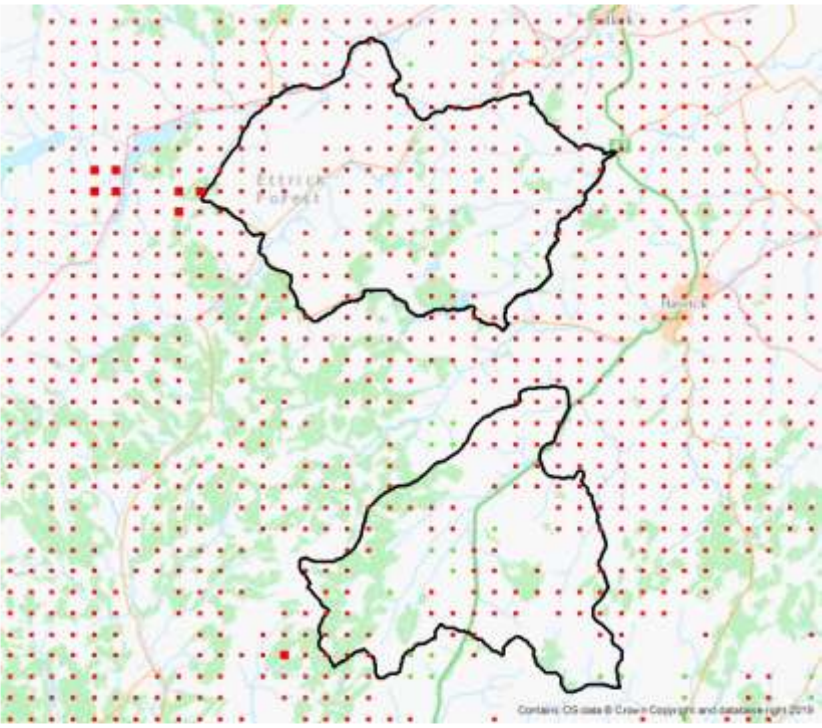


Figure 18. Predicted changes in abundance by breeding **Common Sandpiper** between *BA1990* and *BA2010* at 1-km resolution. *Red* indicates where the probability of occurrence decreased. *Green* indicates where the probability of occurrence increased. Larger symbols represent predicted changes > 40%. Smaller symbols represent predicted changes > 10% but < 40%.

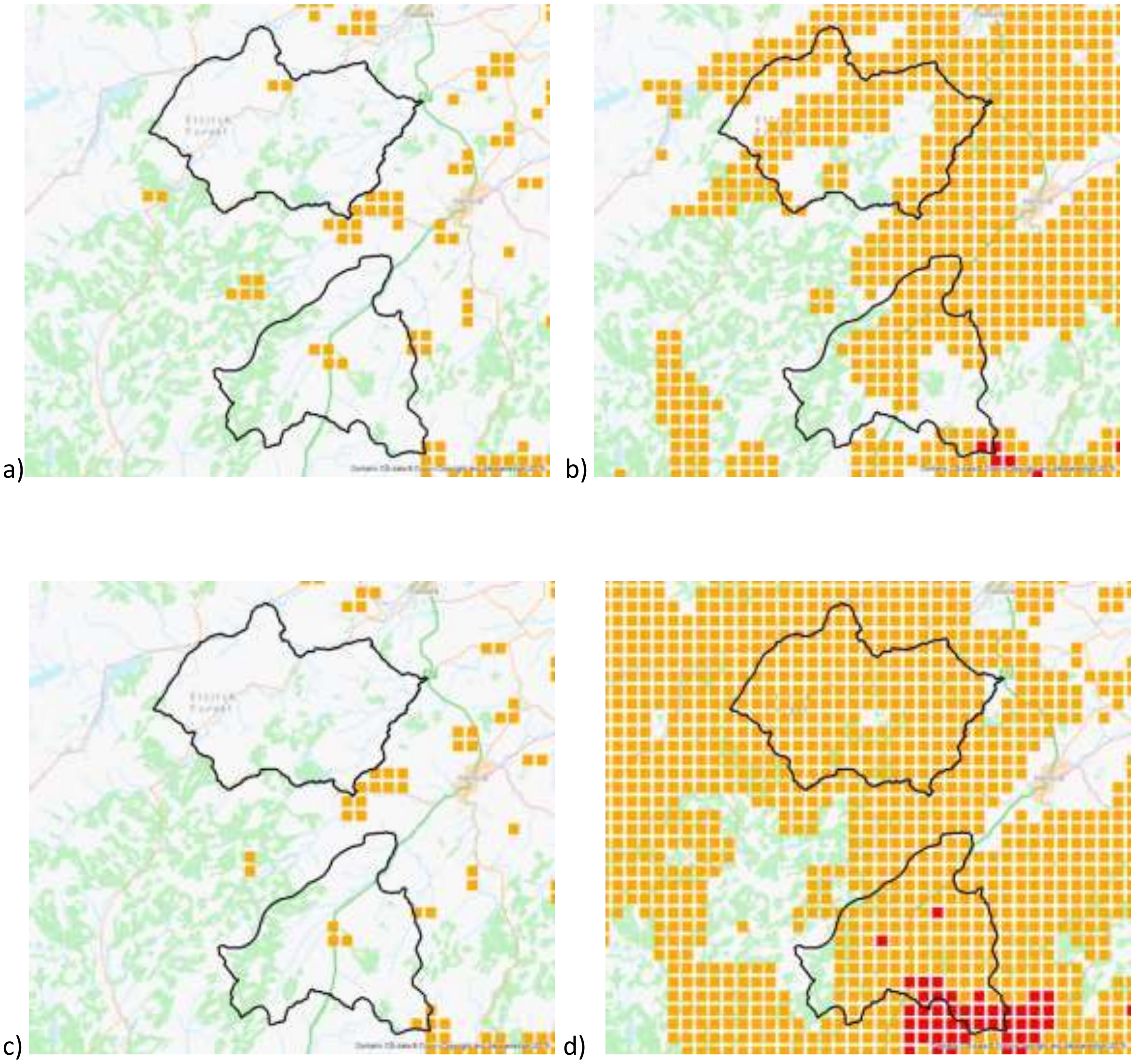


Figure 19. The predicted distributions of **Redstart:** (a) based on data collected during *BA2010*; (b) if cover with broadleaf woodland increased in lowlands; (c) if coniferous woodland increased in uplands; (d) if cover with mixed woodlands increased throughout.

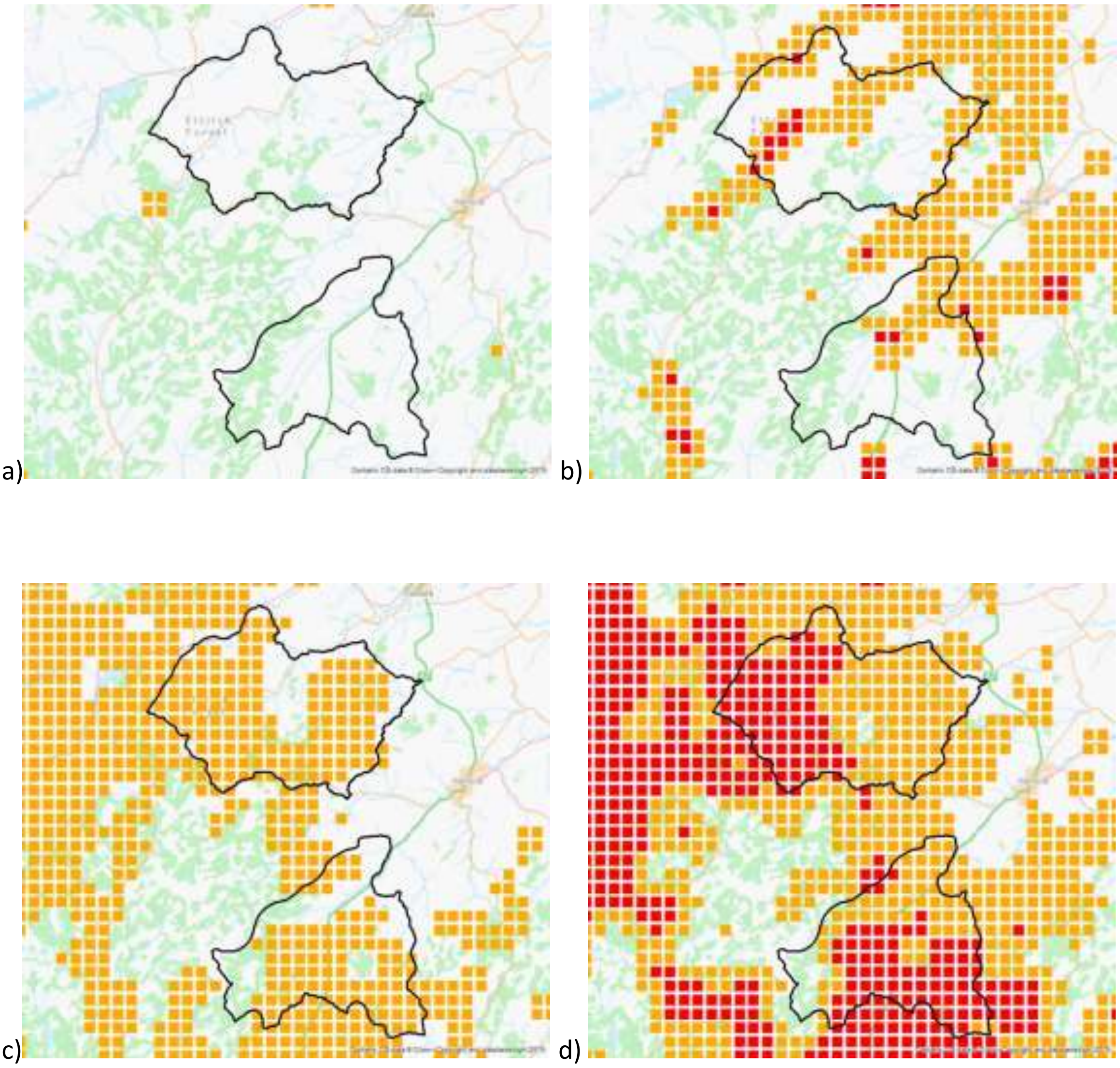
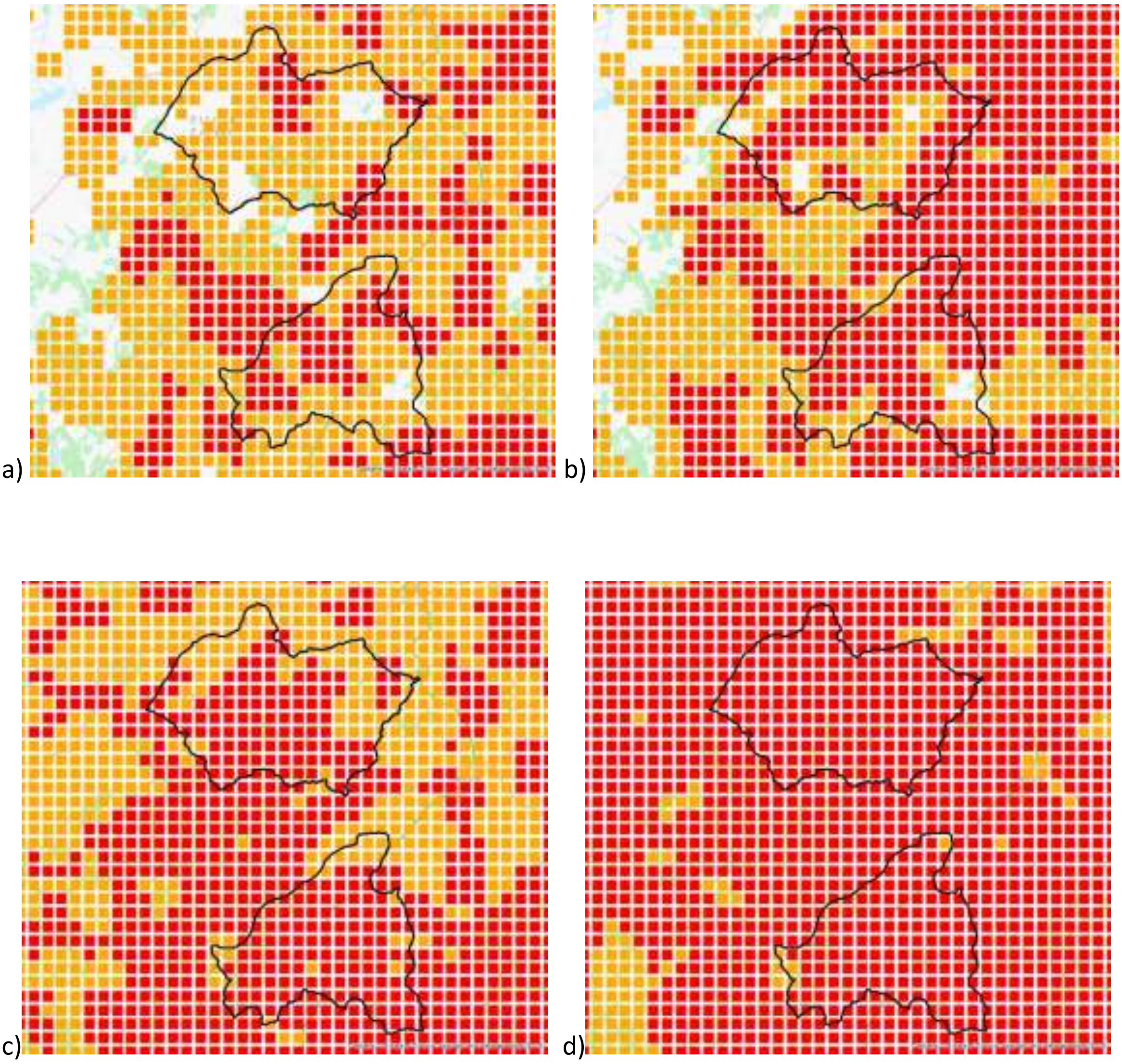


Figure 20. The predicted distributions of **Wood Warbler:** (a) based on data collected during *BA2010*; (b) if cover with broadleaf woodland increased in lowlands; (c) if coniferous woodland increased in uplands; (d) if cover with mixed woodlands increased throughout.



a) b) c) d)

Figure 21. The predicted distributions of **Spotted Flycatcher:** (a) based on data collected during *BA2010*; (b) if cover with broadleaf woodland increased in lowlands; (c) if coniferous woodland increased in uplands; (d) if cover with mixed woodlands increased throughout.