Appendix B: Ecological Modelling

To determine the ecological value of each land parcel within the case study area we developed a Species Abundance Model (SAM) for each of the three target species: the Eurasian oystercatcher (*Haematopus ostralegus*), Eurasian curlew (*Numenius arquata*) and the northern lapwing (*Vanellus vanellus*). SAMs generate predictions of abundance for unsampled locations in a study area using existing data from other sample areas, extrapolating this to new areas based on environmental characteristics (Barker et al, 2014). The development of these ecological models allowed us to obtain predictions for the abundance of the three target species based on current land use and any land-use change as a result of biodiversity offsetting.

Species data

The sampling was undertaken by the British Trust of Ornithology (BTO) as part of their Breeding Bird Survey (BBS). The BBS is undertaken on a stratified random sample of 1 by 1km squares in the UK National Grid (covering the UK, Channel Islands and the Isle of Man), where squares are stratified regionally to obtain representative coverage of regions and habitats whilst making the most of available volunteer resources. As BBS squares are randomly selected, they can turn up on any area of land in the UK. However, temporarily inaccessible squares, or which are not taken up due to their remote location, are retained to maintain the integrity of the sampling design. Therefore, there is no restrictive sampling regarding developed areas.

The data is collected by volunteers who conduct three field visits to each survey square: a reconnaissance visit and two bird recording visits. Surveys take place during the early breeding season (April – mid-May) and a second visit at least for weeks later (mid-May to the end of June). Counts begin early in the early morning to maximise bird activity and last approximately

90 minutes. Volunteers record all the birds they see or hear as they walk methodically along the transect routes (Harris et al, 2019).

For our analysis, we use the maximum count from the two visits which is in line with the standard BBS methodology (see Harris et al, 2019). Species that were not recorded are assigned a count value of zero. We requested data from the eastern region of the UK only giving a total of 1798 squares for 2016 and 1834 squares for 2017 (Figure B1). We requested data for the eastern region to ensure that data were more directly relevant to our study system, given the known regional differences across the UK for bird distributions.

The data collected by the BBS survey has been widely used in ecologically modelling to predict species distribution and abundance nationally (Balmer et al, 2013) and for regional and landscape-scale studies. Given that it is repeated through time, it also allows for a robust analysis of how bird distributions and abundances change through time (for examples see Renwick et al, 2011; Hewson et al, 2016). It is widely considered one of the most robust biodiversity monitoring datasets globally (Magurran et al, 2010). Further, we take a conservative approach to its use by only basing our models on data collected across the eastern portion of the UK to align with the location of our case study region.

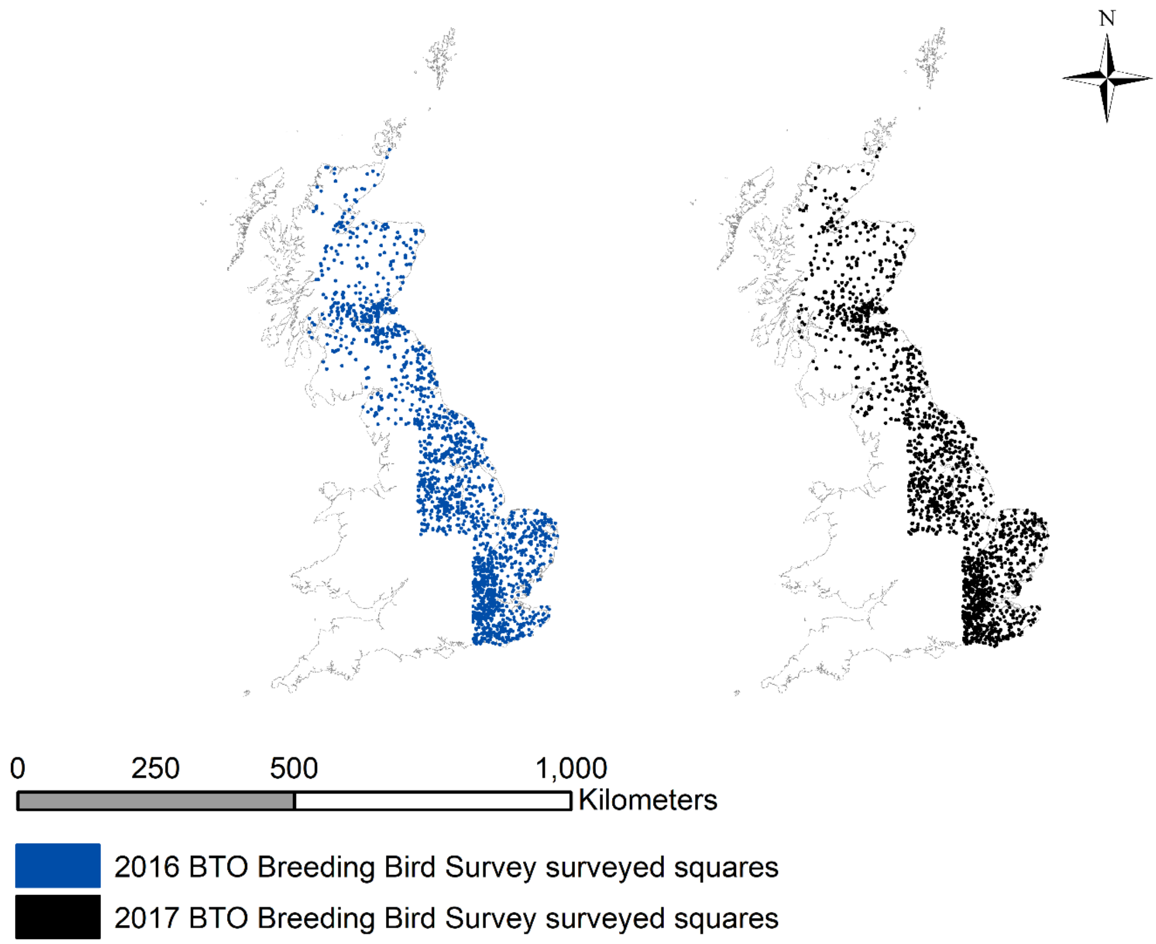

Figure B1: 2016 and 2017 BBS sampled grid squares for eastern region (1km by 1km). 2016 n = 1798, 2017 n = 1835

Explanatory variables

Explanatory variables for the ecological modelling were chosen based on the existing literature for predicting bird habitat suitability (Brotons et al, 2004; Tattoni, Rizzolli, and Pedrini, 2012) and also to link with the economic models. The explanatory variables were derived from their source dataset and sampled to the 1 km grid using ESRI ArcGIS v10.4 (ESRI, Redlands, CA).

Habitat data for each land parcel was derived from the Centre of Ecology and Hydrology (CEH) Land Cover Map (LCM2015) (Rowland et al, 2015). This produces 21 broad classifications following the UK Biodiversity Action Plan Broad Habitats classes including urban, improved grassland, arable and horticulture. Using the CEH Land Cover plus Crops map, arable and horticulture is further sub-divided into 11 crop types (beet, field beans, grassland, maize, oilseed rape, potatoes, spring barley, winter barley, spring wheat, winter wheat (including oats) and other). For our analysis "improved grassland" from the Land Cover Map 2015 and "grassland" from the Land Cover plus Crops map were combined into one category, thus we have 30 land classifications in total.

LCM2015 is created by classifying two-date composite images and is based mainly on data from Landsat-8 (30m resolution) supplemented with AWIFS data (60m resolution) as required (NERC CEH). A limitation for our work is that the data is periodically updated (the previous iteration being LCM2007) and as such there may be differences between the actual habitats in the 2016 and 2017 BBS data.

To derive crop coverage we used the 2016 and 2017 CEH Land Cover Plus: Crops map (NERC CEH). This is the first high-resolution map of UK crops and classifies two million land parcels based on the LCM2015. Crop data is derived from a combination of Copernicus Sentinel-1 C-band SAR (Synthetic Aperture Radar) and Sentinel-2 optical data. Accuracy of the data is assessed using crop type data collated by the Rural Payments Agency, Rural Payments Wales and Scottish Government Rural Payments and Services which detail all land registered for agricultural use and are submitted annually by farmers.

The Tabulate Intersection, spatial statistics analysis tool in ArcGIS was used to derive the total area for each habitat and crop type with the land parcel. For the majority of grid squares, LCM Crops Plus data was overlaid with the LCM2015 Arable and Horticulture classification. Where differences were found, preference was given to the Land Cover Plus: Crop classification.

A further set of environmental variables were derived using OS Open Map products. These variables were chosen based on the existing literature for predicting bird habitat suitability (Brotons et al, 2004; Tattoni et al, 2012). The full list of the environmental variables derived is available in Table B1.

The average altitude within the grid square was derived using OS Terrain 5 digital terrain model (DTM). OS Open Map local (vector) was used to derive the distance from the nearest railway line, overhead electricity pylons, road (divided into Motorway, A road and B road) watercourse and tidal boundary. The inclusion of variables for the road network, railway line and overhead electricity pylons were included based on previous studies which have shown some birds are sensitive to these human structures in the landscape. For each of these measures, we calculated the distance from the centroid of each grid square to the polygon/polyline feature were calculated using the Arc GIS analysis tool 'Near' within the 'Proximity' toolbox. We also calculated the density for the road network, overhead electricity pylons and railway lines within the parcel. We use a measure of density, rather than distance for the road, pylon and railway variables as density provides a more suitable measure as it is more variable. Density was calculated using the 'Line Density Tool' within the Spatial Analyst toolbox.

Table B1: Descriptive statistics for the environmental variables derived from the open data products for model years 2016 and 2017

VARIABLES		2016: *n =1798*			2017: *n = 1835*		
		Mean	Min	Max	Mean	Min	Max
UK Broad Habitat Classifications	Acid Grassland (area hectares)	5.38	0	100	4.82	0	100
	Bog (area hectares)	2.43	0	100	2.19	0	100
	Broadleaf Woodland (area hectares)	7.48	0	92.66	7.67	0	92.66
	Calcareous Grassland (area hectares)	0.28	0	59.22	0.28	0	59.22
	Coniferous Woodland (area hectares)	4.22	0	100	4.34	0	100
	Fen, Marsh and Swamp (area hectares)	0.07	0	57.54	0.07	0	57.54
	Freshwater (area hectares)	0.91	0	90.47	0.96	0	90.47
	Improved Grassland (area hectares)	28.7	0	99.48	28.79	0	99.41
	Heather (area hectares)	2.86	0	100	3.13	0	100
	Heather Grassland (area hectares)	1.73	0	100	1.79	0	100
	Inland Rock (area hectares)	0.3	0	100	0.39	0	100
	Littoral Rock (area hectares)	0.04	0	10.76	0.03	0	10.76
	Littoral Sediment (area hectares)	0.22	0	69.65	0.2	0	69.65
	Neutral Grassland (area hectares)	0.1	0	20.01	0.1	0	20.01
	Saltmarsh (area hectares)	0.21	0	65.7	0.21	0	65.7
	Saltwater (area hectares)	0.31	0	66.19	0.3	0	66.19
	Suburban (area hectares)	9.33	0	97.76	9.76	0	97.76
	Supralittoral Rock (area hectares)	0	0	2.46	0	0	2.46
	Supralittoral Sediment (area hectares)	0.17	0	93.23	0.19	0	93.23
	Urban (area hectares)	3.62	0	100	3.89	0	100
Crop classifications	Beets (area hectares)	0.47	0	60.5	1.05	0	57.52
	Field Beans (area hectares)	1.79	0	61.06	1.54	0	69.93
	Maize (area hectares)	0.6	0	39.59	0.99	0	44.52
	Oil Seed Rape (area hectares)	3.45	0	89.29	3.09	0	72.12
	Other crop type (area hectares)	6.9	0	71.28	4.86	0	64.22
	Potato (area hectares)	0.63	0	35.63	0.92	0	43.37
	Spring Barley (area hectares)	1.76	0	54.69	2.49	0	63.18
	Spring Wheat (area hectares)	2.48	0	60.27	2.16	0	56.07
	Winter Barley (area hectares)	2.92	0	61.56	2.97	0	60.21
	Winter Wheat (area hectares)	10.28	0	96.41	10.41	0	88.44
Other environmental variables							
	Distance from nearest A road (km)	2.68	0.00	67.85	2.57	0	66.87
	Distance from nearest Motorway (km)	31.96	0.02	278.25	31.88	0.02	278.25
	Distance from nearest B Road (km)	3.1	0.00	70.64	3.04	0	69.11
	Distance from nearest Any road (km)	0.66	0.00	66.43	0.54	0	66.04
	Distance from nearest overhead electricity pylons (km)	5.05	0.00	48.68	5	0	48.68
	Distance from nearest railway line (km)	5.13	0.00	59.83	5.02	0	55.32
	Road Density within the land parcel (weighted)	4.24	0.00	30.68	4.32	0	27.24
	Road Density within the land parcel (unweighted)	3.33	0.00	24.47	3.4	0	24.47

	Density of railway lines within the land parcel	0.14	0.00	2.97	0.14	0	4.47
	Density of overhead pylons within the land parcel	0.16	0.00	2.79	0.17	0	2.79
	Contour (mean elevation m)	120.86	-1.24	996.09	122.11	-1.24	1,052.58
	Distance from tidal boundary (km)	19.09	0.00	89.33	19.17	0	89.33
	Distance from nearest watercourse (km)	0.8	0.00	8.15	0.79	0	7.16

Modelling framework

The first stage of the modelling process was to produce a multiple regression model to identify which parameters (environmental characteristics) predict current distributions for each of the target species (curlew, oystercatcher, lapwing) across the eastern region of the UK. This allowed us to identify which land management practices are most beneficial to the target species (related to the parameters used in the modelling process) and thus serve as our "preferred" land management practice within the biodiversity offset market. The SAM could then be used to predict how the distribution and abundance of the target species could alter as a result of land-use change at the 1 km land parcel.

The choice of a specific functional form for SAMS is governed by the knowledge of the species and characteristics of the available data. The majority of SAMS sit within the generalised linear model framework (GLM). Such models are typically described in terms of their systematic component in which the response is linked to the environmental data, and their stochastic structure that describes the error distribution (Venables and Springer 2002).

For our analysis, we employed the negative binomial model using pooled count data from 2016 and 2017. This is a count data model, derived from the Poisson model but accounts for overdispersion within the data (Potts and Elith 2006). Let us denote as a random variable the single species count, Y_{ij}, from the ith site and the jth count. We assume:

$$E(Y_{ij}) = \mu_{ij} \qquad [B1]$$

We allow the mean to be a positive function of covariates

$$\mu_{ij} = \exp\left(\boldsymbol{x}'_{ij}\boldsymbol{\beta}\right), \qquad \text{[B2]}$$

where $\boldsymbol{x}_{ij}$ is a vector of measured covariates for the *j*th count of the *i*th site, and $\boldsymbol{\beta}$ a vector of parameters. We fit the following model for each regression method:

$$x'_{ij}\beta = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in} + \varepsilon_t \qquad \text{[B3]}$$

The models were fit by minimizing the negative log-likelihood using the minimization procedure (glm2) in the software package R (Version 3.6.2). The best model for each species was selected by comparing the AIC values (Burnham and Anderson 1998). Models with the smallest AIC value are considered the best model. The model was also evaluated for a sub-sample of the 2016 data set which was independent of the dataset used for the primary modelling.

Ecological Modelling Results

The results highlight that none of the crop types is positively associated with increased abundance of any of the target species. However, switching from the crop types most negatively associated with each species to improved grassland could benefit each species, dependent on other environmental factors within the 1 by 1 km land parcel (Table B2).

Table B2: Results of the negative binomial model for predicting individual species abundance

VARIABLES		Curlew model		Lapwing model		Oystercatcher model	
		Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
UK Broad Habitat Classifications	Acid Grassland (area ha)	0.003	-0.011	-0.036***	-0.009	-0.023***	-0.009
	Bog (area ha)	0.001	-0.011	-0.043***	-0.01	-0.024***	-0.009
	Broadleaf woodland (area ha)	-0.053***	-0.012	-0.081***	-0.011	-0.041***	-0.01
	Calcareous grassland (area ha)	0.023	-0.017	-0.036**	-0.016	0.002	-0.017
	Coniferous woodland (area ha)	-0.021**	-0.011	-0.074***	-0.01	-0.033***	-0.009
	Fen, marsh and swamp (area ha)	-0.037	-0.034	-0.005	-0.013	-0.022	-0.014
	Freshwater (area ha)	-0.025*	-0.014	-0.048***	-0.013	0.011	-0.011
	Heather (area ha)	0.004	-0.011	-0.038***	-0.01	-0.023**	-0.009
	Improved Grassland (area ha)	0.002	-0.011	0.028	-0.009	-0.01	-0.009
	Heather grassland (area ha)	0.008	-0.011	-0.035***	-0.01	-0.011	-0.009
	Inland rock (area ha)	-0.097***	-0.027	-0.130***	-0.038	-0.053***	-0.016
	Littoral rock (area ha)	0.088	-0.149	-0.646***	-0.177	0.283***	-0.093
	Littoral sediment (area ha)	0.023	-0.024	-0.168***	-0.05	0.076***	-0.022
	Neutral grassland (area ha)	0.048	-0.043	0.036	-0.035	-0.037	-0.048
	Saltmarsh (area ha)	0.063***	-0.023	0.029*	-0.017	0.047***	-0.016
	Saltwater (area ha)	0.013	-0.015	-0.086***	-0.019	0.004	-0.013
	Suburban (area ha)	-0.079***	-0.014	-0.100***	-0.012	-0.031***	-0.009
	Supralittoral rock (area ha)	-0.543*	-0.321	0.082	-0.159	-0.118	-0.152
	Supralittoral sediment (area ha)	-0.044**	-0.022	-0.048***	-0.015	0.003	-0.011
	Urban (area ha)	-0.068***	-0.019	-0.078***	-0.015	-0.064***	-0.012
Crop classifications	Beets (area ha)	-0.072	-0.046	-0.019	-0.013	0.002	-0.014

	Field Beans area ha)	-0.040**	-0.019	-0.039***	-0.012	-0.061***	-0.017
	Maize (area ha)	-0.043	-0.028	-0.027*	-0.015	-0.089***	-0.019
	Oil Seed Rape (area ha)	-0.032***	-0.012	-0.048***	-0.01	-0.054***	-0.011
	Potatoes (area ha)	0.027	-0.026	-0.032**	-0.014	0.014	-0.012
	Spring Barley (area ha)	-0.003	-0.013	-0.049***	-0.012	0.004	-0.01
	Spring Wheat (area ha)	-0.01	-0.013	-0.034***	-0.011	-0.006	-0.01
	Winter Barley (area ha)	-0.024*	-0.014	-0.040***	-0.012	-0.016	-0.012
	Winter Wheat (area ha)	-0.045***	-0.013	-0.052***	-0.01	-0.042***	-0.009
	Other crop type (area ha)	-0.033**	-0.013	-0.030***	-0.01	-0.020**	-0.01
Additional explanatory variables	Pylons Density (Density of electricity pylons within the land parcel)	-0.445***	-0.116	-0.336***	-0.117	-0.375**	-0.155
	Tidal (distance of land parcel from nearest tidal boundary km)	0.015***	-0.003	0.013***	-0.004	-0.002	-0.003
	Watercourse (distance of land parcel from nearest tidal boundary km) (distance km)	-0.232**	-0.091	-0.409***	-0.085	-0.401***	-0.087
	Model year dummy: 0 = 2016, 1= 2017	-0.02	-0.091	0.051	-0.101	0.013	-0.103
Constant		0.771	-1.053	4.350***	-0.902	1.644*	-0.853
Observations		3,626		3,620		3,610	
chi2		835.6		3.776		525.4	
*** p<0.01, ** p<0.05, * p<0.1							

Once the model was established for the eastern region data set, predictions could then be made for the target species abundance for all land parcels across the case study region based on current land use (Table B3) and changes in land use (Table B4). The predictive model allows us to calculate the abundance for each target species for each land parcel within the case study region based on the current land use. In addition, we can predict which land parcels offer the most opportunity for increases in the target species abundance which allows us to identify which land parcels could offer biodiversity offsets. In our modelling framework, this is achieved by transforming all land parcels which currently contain a level of crops into improved grassland. The predictive model is then re-run over this new landscape where all parcels containing crops have been switched to improved grassland.

Table B3: Predicted species abundance across the case study region based on current land

Bird species	Total predicted abundance across case study region (95% confidence intervals)	Total number of parcels with presence	Median count of birds in a single land parcel with presence (min – max)
Lapwing	9267 (8962 - 9572)	4764	1.6 (0.1 - 62.0)
Oystercatcher	4301 (3865 - 4737)	4860	0.4 (0.1- 45.0)
Curlew	8986 (8697 - 9276)	4252	1.5 (0.1 - 40.3)

Table B4: Predicted species abundance across the case study region based and under a full offset scenario (the full offset scenario is based on all crop hectares being converted to improved grassland)

Bird species	Total predicted abundance across case study region (95% confidence intervals)	Total number of parcels with presence	Median count of birds in a single land parcel with presence (min – max)
Lapwing	12485 (12165 - 12803)	4806	2.3 (0.1 - 65.4)
Oystercatcher	5834 (5324 - 6344)	5138	0.8 (0.1 - 74.3)
Curlew	11149 (10883 - 11416)	4762	2.0 (0.1 - 40.4)

Appendix B Reference List

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

Barker, Nicole K. S., Stuart M. Slattery, Marcel Darveau, and Steve G. Cumming. 2014. "Modeling Distribution and Abundance of Multiple Species: Different Pooling Strategies Produce Similar Results." Ecosphere 5 (12): art158. https://doi.org/10.1890/ES14-00256.1

Brotons, Lluís, Wilfried Thuiller, Miguel B. Araújo, Alexandre H. Hirzel Brotons, Araú Jo, M. B. Hirzel, and W. Thuiller. 2004. "Presence-Absence versus Presence-Only Modelling Methods for Predicting Bird Habitat Suitability." Ecography 27: 4

Burnham, Kenneth P., and David R. Anderson. 1998. "Practical Use of the Information-Theoretic Approach." In Model Selection and Inference, 75–117. New York: Springer

Harris, S. J., D. Massimino, S. E. Newson, M. A. Eaton, D. E. Balmer, D. G. Noble, A. J. Musgrove, S. Gillings, D. Procter, and J. W. Pearce-Higgins. 2015. "The Breeding Bird Survey 2014." BTO Research Report 673. Thetford: BTO.

Hewson, C. M., K. Thorup, K., J. W. Pearce-Higgins, and P. W. Atkinson. 2016. "Population Decline Is Linked to Migration Route in the Common Cuckoo." Nature Communications 7(1): 1–8.

Magurran, Anne E., Stephen R. Baillie, Stephen T. Buckland, Jan M. Dick, David A. Elston, E. Marian Scott, Rognvald I. Smith, Paul J. Somerfield, and Allan D. Watt. 2010. "Long-Term Datasets in Biodiversity Research and Monitoring: Assessing Change in Ecological Communities through Time." Trends in Ecology & Evolution 25 (10): 574–82

Potts, Joanne M., and Jane Elith. 2006. "Comparing Species Abundance Models." Ecological Modelling 199 (2): 153–63. https://doi.org/10.1016/j.ecolmodel.2006.05.025.

Renwick, A. R., D. Massimino, S. E. Newson, D. E. Chamberlain, J. W. Pearce-Higgins, and A. Johnston. 2012. "Modelling Changes in Species' Abundance in Response to Projected Climate Change." Diversity and Distributions 18 (2): 121–32

Rowland, C. S., R. D. Morton, L. Carrasco, G. McShane, A. W. O'Neil, and C. M. Wood. 2017. "Land Cover Map 2015 (vector, GB)." NERC Environmental Information Data Centre. https://doi.org/10.5285/6c6c9203-7333-4d96-88ab-78925e7a4e73.

Tattoni, Clara, Franco Rizzolli, and Paolo Pedrini. 2012. "Can LiDAR Data Improve Bird Habitat Suitability Models?" Ecological Modelling 245: 103–10

Venables, W. N., and B. D. Ripley Springer. n.d. "Modern Applied Statistics with S Fourth Edition." Available at http://www.insightful.com (accessed February 3, 2020).