

Natural Environment Valuation Online Tool

Technical Documentation

Version 1.0, June 2019

Chapter 5a: Biodiversity Model

Land, Environment, Economics and Policy (LEEP) Institute

University of Exeter

Corresponding authors

Emma Wright Emma.Wright@jncc.gov.uk

Authors

Brett Day

Amy Binner

Nathan Owen

Ian Bateman

Greg Smith

Patrick Collings

Louis Haddrell

Lorena Liuzzo

Carlo Fezzi

Collaborators

Forestry Commission / Forest Research

UCL

JNCC

University of Aberdeen

The Biodiversity module

NEVO's biodiversity module uses species presence data along with environmental variables and a comprehensive modelling framework to produce probabilistic species distribution maps for 100 species of conservation priority in the UK. The module is based on the species distribution modelling framework produced by JNCC¹. The framework is designed to support data from as wide a range of taxa as possible, thereby improving species representation beyond only highly recorded species such as birds, and providing a better indication of biodiversity changes than using a single taxonomic group². The probabilities of each species are summed to give an indicative species richness for any given unit of interest. The models produced can then be used to predict species distribution based on any alternate states in NEVO, and the subsequent impacts land-use decisions may have on species richness.

Species selection

The module aims to capture taxonomic diversity across a range of broad habitat types. It is acknowledged that representing biodiversity with a limited subset of species will retain bias, but 100 species have been included to balance species representation without prohibitively reducing the interactive response time of NEVO due to computational capacity.

Species selection followed these general principles:

- Selection should capture not only taxonomic diversity within families/genera, but also conservation priority.
- Species with habitat specificity were favoured rather than generalists, as a major component of NEVO is evaluating the effect of land cover/habitat-type changes on species presence.
- Species were selected from across the range of broad habitat types present in the UK, but with emphasis on habitats of particular interest to NEVO.
- Plants make up a significant proportion of the species list due to the fact they are stationary and frequently closely associated with habitat type, and therefore may be a good indicator of biodiversity³.
- The taxonomic groupings considered were: mammals; birds; vascular plants; invertebrates; lichens, and; herptiles.
- Within each taxonomic group, species with comparatively high numbers of records were favoured due to more complete datasets being available. This does not necessarily mean the species is widespread but that effort for that taxonomic group is good across the UK, implying greater certainty for species absences than for poorly recorded species.

The first stage was to filter each taxonomic group by the 2007 Biodiversity Action Plan (BAP) list⁴ to select species of conservation priority within the UK. Since 2007, the successor to the BAP prioritises species for each of the four UK countries⁵, however using this list results in a regional bias as a result of numerous species that were only of concern in one UK country. Species that have been recorded in fewer than 20 GB hectads were removed as these were deemed too restricted in range to be informative. The list for each taxon was then divided by habitat type association, and widespread generalists were excluded. Where possible, species whose presence is heavily linked to water quality were excluded as the environmental parameters currently included in the model do not capture variation in this. Final species within each group were then selected based on the number of records available recorded in the last 20 years at a 2 km or greater precision. Partial preference was also given to charismatic animals to make it easier for a general audience to engage with the final species list, and therefore there is higher representation from mammals, birds and butterflies than other

taxonomic groups. In turn this also means these species are likely to be well recorded with good coverage across their range. The final species list is represented by:

- Terrestrial mammals (14)
- Birds (17)
- Vascular plants (38)
- Invertebrates (25, of which 10 are butterflies)
- Lichens (5)
- Herptile (1)

The full list of included species is presented in Table 1.

Table 1 Species included in NEVO's biodiversity module

Taxonomic group	Species	English common name
Bird	<i>Acanthis cabaret</i>	Lesser Redpoll
Bird	<i>Alauda arvensis</i>	Skylark
Bird	<i>Anthus trivialis</i>	Tree Pipit
Bird	<i>Botaurus stellaris</i>	Bittern
Bird	<i>Caprimulgus europaeus</i>	Nightjar
Bird	<i>Emberiza citrinella</i>	Yellowhammer
Bird	<i>Emberiza schoeniclus</i>	Reed Bunting
Bird	<i>Locustella naevia</i>	Grasshopper Warbler
Bird	<i>Lullula arborea</i>	Woodlark
Bird	<i>Motacilla flava subsp. flavissima</i>	Yellow Wagtail
Bird	<i>Muscicapa striata</i>	Spotted Flycatcher
Bird	<i>Numenius arquata</i>	Curlew
Bird	<i>Perdix perdix</i>	Grey Partridge
Bird	<i>Phylloscopus sibilatrix</i>	Wood Warbler
Bird	<i>Streptopelia turtur</i>	Turtle Dove
Bird	<i>Turdus torquatus</i>	Ring Ouzel
Bird	<i>Vanellus vanellus</i>	Lapwing
Herptile	<i>Vipera berus</i>	Adder
Invertebrate - bee	<i>Bombus monticola</i>	Mountain Bumblebee
Invertebrate - beetle	<i>Cercyon convexiusculus</i>	Water Beetle sp.
Invertebrate - beetle	<i>Gnorimus nobilis</i>	Noble Chafer
Invertebrate - beetle	<i>Liopterus haemorrhoidalis</i>	Water Beetle sp.
Invertebrate - beetle	<i>Lucanus cervus</i>	Stag Beetle
Invertebrate - butterfly	<i>Boloria selene</i>	Small Pearl-Bordered Fritillary
Invertebrate - butterfly	<i>Coenonympha pamphilus</i>	Small Heath
Invertebrate - butterfly	<i>Coenonympha tullia</i>	Large Heath
Invertebrate - butterfly	<i>Cupido minimus</i>	Small Blue
Invertebrate - butterfly	<i>Erebia epiphron</i>	Mountain Ringlet
Invertebrate - butterfly	<i>Erynnis tages</i>	Dingy Skipper
Invertebrate - butterfly	<i>Hipparchia semele</i>	Grayling
Invertebrate - butterfly	<i>Lasiommata megera</i>	Wall
Invertebrate - butterfly	<i>Limenitis camilla</i>	White Admiral

Invertebrate - butterfly	<i>Satyrium w-album</i>	White-Letter Hairstreak
Invertebrate - cricket	<i>Leptophyes punctatissima</i>	Speckled Bush Cricket
Invertebrate - cricket	<i>Metrioptera brachyptera</i>	Bog Bush Cricket
Invertebrate - moth	<i>Cossus cossus</i>	Goat Moth
Invertebrate - moth	<i>Acronicta psi</i>	Grey Dagger
Invertebrate - moth	<i>Allophyes oxyacanthae</i>	Green-brindled Crescent
Invertebrate - moth	<i>Dasypolia templi</i>	Brindled Ochre
Invertebrate - moth	<i>Xanthorhoe decoloraria</i>	Red Carpet
Invertebrate - snail	<i>Cochlodina laminata</i>	Plaited Door Snail
Invertebrate - snail	<i>Monacha cantiana</i>	Kentish Snail
Invertebrate - snail	<i>Zonitoides excavatus</i>	Hollowed Glass Snail
Lichen	<i>Anaptychia ciliaris ciliaris</i>	Lichen subsp.
Lichen	<i>Leptogium brebissonii</i>	Lichen sp.
Lichen	<i>Parmeliella testacea</i>	Lichen sp.
Lichen	<i>Pseudocyphellaria intricata</i>	Lichen sp.
Lichen	<i>Usnea articulata</i>	String-Of-Sausage Lichen
Mammal	<i>Barbastella barbastellus</i>	Barbastelle bat
Mammal	<i>Felis silvestris</i>	Wildcat
Mammal	<i>Lepus europaeus</i>	European hare
Mammal	<i>Lepus timidus</i>	Mountain Hare
Mammal	<i>Martes martes</i>	Pine Marten
Mammal	<i>Micromys minutus</i>	Harvest Mouse
Mammal	<i>Muscardinus avellanarius</i>	Hazel Dormouse
Mammal	<i>Mustela putorius</i>	Polecat
Mammal	<i>Myotis bechsteinii</i>	Bechstein's bat
Mammal	<i>Nyctalus noctula</i>	Noctule Bat
Mammal	<i>Plecotus auritus</i>	Brown Long-eared Bat
Mammal	<i>Rhinolophus ferrumequinum</i>	Greater Horseshoe Bat
Mammal	<i>Rhinolophus hipposideros</i>	Lesser Horseshoe Bat
Mammal	<i>Sciurus vulgaris</i>	Eurasian red squirrel
Vascular plant	<i>Anchusa arvensis</i>	Field bugloss
Vascular plant	<i>Andromeda polifolia</i>	Bog Rosemary
Vascular plant	<i>Arctostaphylos alpinus</i>	Mountain bearberry
Vascular plant	<i>Asplenium viride</i>	Green spleenwort
Vascular plant	<i>Atriplex laciniata</i>	Frosted Orache
Vascular plant	<i>Blysmus rufus</i>	Saltmarsh Flat-Sedge
Vascular plant	<i>Cakile maritima</i>	Sea Rocket
Vascular plant	<i>Campanula glomerata</i>	Clustered Bellflower
Vascular plant	<i>Carex extensa</i>	Long-Bracted Sedge
Vascular plant	<i>Carex magellanica</i>	Tall Bog-Sedge
Vascular plant	<i>Centaurium pulchellum</i>	Lesser Centaury
Vascular plant	<i>Cerastium arvense</i>	Field Mouse-Ear
Vascular plant	<i>Cirsium eriophorum</i>	Woolly Thistle
Vascular plant	<i>Daphne laureola</i>	Spurge-Laurel

Vascular plant	<i>Eriophorum latifolium</i>	Broad-Leaved Cottongrass
Vascular plant	<i>Fumaria muralis</i>	Common Ramping-Fumitory
Vascular plant	<i>Genista anglica</i>	Petty Whin
Vascular plant	<i>Genista tinctoria</i>	Dyer's Greenweed
Vascular plant	<i>Gnaphalium supinum</i>	Dwarf Cudweed
Vascular plant	<i>Goodyera repens</i>	Creeping Lady's-Tresses
Vascular plant	<i>Hypericum elodes</i>	Marsh St John's-Wort
Vascular plant	<i>Lamium hybridum</i>	Cut-Leaved Dead-Nettle
Vascular plant	<i>Leymus arenarius</i>	Lyme Grass
Vascular plant	<i>Lycopodium clavatum</i>	Stag's-Horn Clubmoss
Vascular plant	<i>Neottia nidus-avis</i>	Bird's-Nest Orchid
Vascular plant	<i>Ornithopus perpusillus</i>	Bird's-Foot
Vascular plant	<i>Orthilia secunda</i>	Serrated Wintergreen
Vascular plant	<i>Oxyria digyna</i>	Mountain Sorrel
Vascular plant	<i>Pyrola media</i>	Intermediate Wintergreen
Vascular plant	<i>Radiola linoides</i>	Allseed
Vascular plant	<i>Ranunculus omiophyllus</i>	Round-Leaved Crowfoot
Vascular plant	<i>Saxifraga tridactylites</i>	Rue-Leaved Saxifrage
Vascular plant	<i>Silaum silaus</i>	Pepper-Saxifrage
Vascular plant	<i>Thymus pulegioides</i>	Large Thyme
Vascular plant	<i>Tilia cordata</i>	Small-Leaved Lime
Vascular plant	<i>Trifolium fragiferum</i>	Strawberry Clover
Vascular plant	<i>Trifolium striatum</i>	Knotted Clover
Vascular plant	<i>Vaccinium microcarpum</i>	Small Cranberry

Species presence data were downloaded from NBN Gateway (and its successor, NBN Atlas⁶), the largest repository for UK species records. These include species presence data recorded as part of a dedicated taxon survey (e.g. ^{7,8,9,10,11}), as well as records from other environmental surveys, or reported through local or national records centres.

The modelling framework

Species distribution models use spatial data on environmental variables such as climate, elevation and land use to model against geographically explicit observation records of a species to predict the potential distribution of the species across a region based on the suitability of the environmental variables for that species. There is now a large range of modelling methods available with differing strengths and weaknesses for predicting the geographic presence of a species, and the suitability of a model may vary between species based both on the true distribution of the species and how occurrence data has been collected¹². The JNCC biodiversity module uses an ensemble modelling approach to evaluate the performance of a suite of seven modelling methodologies, combining the best performing of the resulting probabilistic maps from multiple iterations¹³.

The localities of the species records are published at a range of precisions, typically between 100 m and 10 km. This means for low resolution data (published at 10 km precision), the actual locality of the observation may not align with the resolution of the underlying covariables, which within NEVO are 2 km². The framework therefore randomly moves the locality of low resolution records within their 10 km² 'window'. The repeat processing therefore captures the record throughout the potential area

of the ‘true’ locality along with the accompanying environmental variables, rather than excluding low resolution records or falsely assigning a static locality as a known. While this does improve the species distribution maps, particularly for species with fewer records, this does introduce additional uncertainty into the model. Therefore, if species were deemed to have sufficient high precision records (more than 5000 recorded at a 2 km or higher precision), low precision records were excluded from the dataset.

Species data are comprised of presence-only data, but most modelling methods also require absence data. 5000 pseudo-absence points were generated within a mask of cells where priority species from the same taxonomic group had been recorded but the target species had not, in order to reduce the bias of false-absence due to a lack of recorder effort¹⁴.

The covariables

Species presence was modelled against a range of spatial climatic, environmental and topographic datasets which relate to species’ presence, and are compatible with data from other modules in NEVO and the overall tool’s capabilities to project to alternate states.

The monthly minimum temperature, maximum temperature and precipitation were downloaded from the Met Office for 1961-1990 at a 2 km resolution. Using the biovars function in the R package dismo¹⁵, these were then used to generate 19 different climatic variables¹⁶. An additional climate variable of total precipitation during growing season was added, which is used in the agricultural module of NEVO.

The land cover and soil variables used by other NEVO modules were also included as potential covariables. These layers help relate the species presence to the rest of the NEVO model, and are also expected to be good predictors for plants and therefore underlying habitat.

The correlation coefficients between all variable pairs were tested. Where correlation coefficient was greater than ± 0.7 , only one variable was retained. The final list of variables used in the modelling is shown in Table 2.

Table 2 Climate, environmental and topographic parameters used in species distribution models

Layer	Description	Data source
Isothermality	Mean diurnal range (mean of monthly temperature range / annual temperature range) (averaged 1961-1990) (°C)	Met Office ¹⁷
Temperature of wettest quarter	Mean temperature of wettest quarter (averaged 1961-1990) (°C)	Met Office ¹⁷
Temperature of driest quarter	Mean temperature of driest quarter (averaged 1961-1990) (°C)	Met Office ¹⁷
Temperature annual range	Max temperature of warmest month - min temperature of coldest month (averaged 1961-1990) (°C)	Met Office ¹⁷
Temperature of warmest quarter	Mean temperature of warmest quarter (averaged 1961-1990) (°C)	Met Office ¹⁷

Temperature of coldest quarter	Mean temperature of coldest quarter (averaged 1961-1990) (°C)	Met Office ¹⁷
Total precipitation during growing season	Total precipitation during growing season (averaged 1961-1990) (mm)	See section X.X
Elevation	Average elevation for entire cell (metres)	See section X.X
Aspect	Dominant aspect of the cell	Derived from elevation data (above)
Coast	% area coverage by coastal margins	See section X.X
Freshwater	% area coverage by freshwater	See section X.X
Marine	% area coverage by sea and estuary	See section X.X
Urban	% area coverage by urban and other developed land	See section X.X
Rough grazing	% area coverage by rough grazing	See section X.X
Semi-natural habitat	% area coverage by semi-natural grass, mountains, moors and heaths where NOT used for farming	See section X.X
Potatoes	% area coverage by potatoes	See section X.X
Maize	% area coverage by maize	See section X.X
Horticulture	% area coverage by total horticulture	See section X.X
Sugar beet	% area coverage by sugar beet not for stockfeed	See section X.X
Improved grassland	Sum of % area of permanent and temporary grassland	Derived from data in section X.X
Woodland	Sum of % farm woodland and woodland not used for farming	Derived from data in section X.X
Cereals and oilseed rape	Sum of % barley, oilseed rape, wheat and other cereals	Derived from data in section X.X
Clay	% soil area classified as 'clay'	See section X.X
Loamy sand	% soil area classified as 'loamy sand'	See section X.X
Loam	% soil area classified as 'loam'	See section X.X
Sandy loam	% soil area classified as 'sandy loam'	See section X.X
Clay/loam	% soil area classified as 'clay/loam'	See section X.X
Sand	% soil area classified as 'sand'	See section X.X
Silt/loam	% soil area classified as 'silt/loam'	See section X.X

pH1	% soil area classified as ' pH 1' (pH <4.5)	See section X.X
pH 2	% soil area classified as ' pH 2' (pH = 4.5-5.5)	See section X.X
pH 3	% soil area classified as ' pH 3' (pH 5.50001-7.2)	See section X.X
pH 4	% soil area classified as ' pH 4' (pH 7.20001-8.5)	See section X.X
Swr1	% soil area classified as dominant annual average soil water regime class 1	See section X.X
Swr2	% soil area classified as dominant annual average soil water regime class 2	See section X.X
Swr3	% soil area classified as dominant annual average soil water regime class 3	See section X.X
Swr4	% soil area classified as dominant annual average soil water regime class 4	See section X.X

Seventy-five percent of the species occurrence (including presence and pseudo-absence) data were randomly selected as a training set for each of the distribution modelling approaches. The model methodologies included were: Bioclim, boosted regression trees, general linear model, generalised additive model, kernel support vector machine, Maxent, and random forest. The outputs of each of the models were evaluated against the remaining 25% of the occurrence data, based on the area under the Receiver Operating Characteristic curve (AUC) statistic. After 100 iterations, the probabilistic map produced by the best performing model from each iteration was averaged to produce a final mean map. If the best performing model was not considered good (AUC of 0.7 or greater) it was excluded from the ensemble. All maps were required to have at least 50 qualifying iterations in the ensemble. The mean AUC of the included models was 0.97, with a standard deviation of 0.17. Figure 1 shows examples of the localities of presence data, and the resulting ensemble probability maps of species distributions included in NEVO.

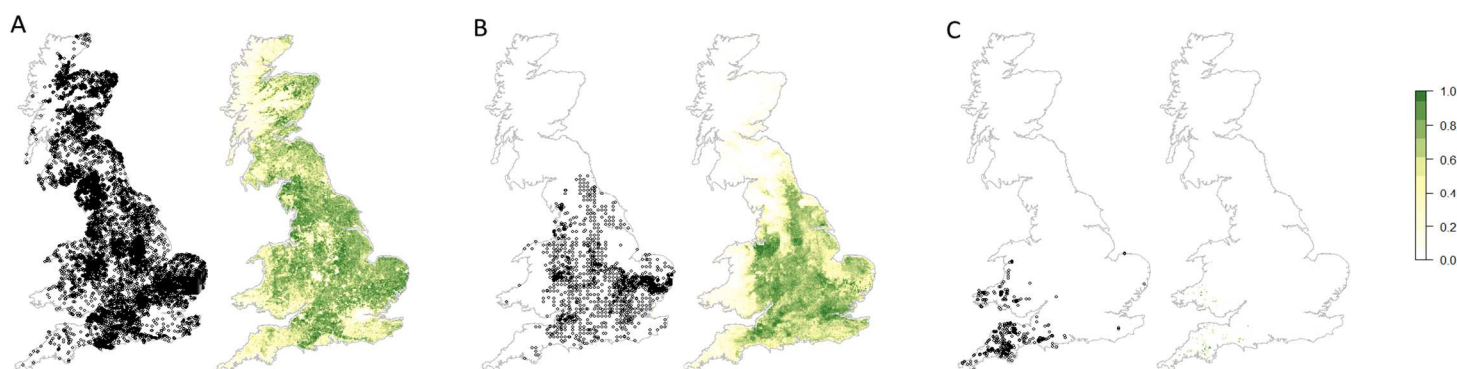


Figure 1 Input data (locality of species presence records) and resulting ensemble map of probability of species presence, for a) European hare (*Lepus europaeus*, mean ensemble AUC = 0.82), b) white-letter hairstreak butterfly (*Satyrium w-album*, mean ensemble AUC = 0.91), and c) string-of-sausage lichen (*Usnea articulata*, mean ensemble AUC = 0.95)

Species richness

The overall species richness per unit is calculated by summing the occurrence probabilities of all species for that grid square or area. In this way species with a lower probability of being found there contribute less to the overall species richness than a species with a high probability of being there. This method is mathematically justified and is less likely to overpredict species richness than using a probability threshold to define a species' predicted presence or absence¹⁸.

¹ Available as an R package at <https://github.com/jncc/sdms>.

² Larsen, F. W., Bladt, J. and Rahbek, C. (2009). Indicator taxa revisited: useful for conservation planning? *Diversity and Distributions*, 15: 70-79. doi:10.1111/j.1472-4642.2008.00507.x

³ Brunbjerg, A. K., Bruun, H. H., Dalby, L., Fløjgaard, C., Frøslev, T. G., Høye, T. T., Goldberg, I., Læssøe, T., Hansen, M. D. D., Brøndum, L., Skipper, L., Fog, K. and Ejrnæs, R. (2018). Vascular plant species richness and bioindication predict multi-taxon species richness. *Methods in Ecology and Evolution*, 2018;9: 2372–2382. <https://doi.org/10.1111/2041-210X.13087>

⁴ JNCC (2007). Report on the Species and Habitat Review. Report by the Biodiversity Reporting and Information Group (BRIG) to the UK Standing Committee. Available from: <http://jncc.defra.gov.uk/page-5717>

⁵ JNCC and Defra (on behalf of the Four Countries' Biodiversity Group) (2012). UK Post-2010 Biodiversity Framework. Available from: <http://jncc.defra.gov.uk/page-6189>

⁶ NBN Atlas. National Biodiversity Network <https://nbnatlas.org/>

⁷ National Bat Monitoring Programme <https://www.bats.org.uk/our-work/national-bat-monitoring-programme>

⁸ National Plant Monitoring Scheme <https://www.npms.org.uk/>

⁹ The UK Butterfly Monitoring Scheme <http://www.ukbms.org/>

¹⁰ The Breeding Bird Survey <https://www.bto.org/volunteer-surveys/bbs>

¹¹ The Wetland Bird Survey <https://www.bto.org/volunteer-surveys/webs>

¹² Elith, J. , Graham, C. H., Anderson, R. P., Dudík, M. , Ferrier, S. , Guisan, A. , Hijmans, R. J., Huettmann, F. , Leathwick, J. R., Lehmann, A. , Li, J. , Lohmann, L. G., Loiselle, B. A., Manion, G. , Moritz, C. , Nakamura, M. , Nakazawa, Y. , Overton, J. McC. M., Townsend Peterson, A. , Phillips, S. J., Richardson, K. , Scachetti-Pereira, R. , Schapire, R. E., Soberón, J. , Williams, S. , Wisz, M. S. and Zimmermann, N. E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, 29: 129-151. doi:10.1111/j.2006.0906-7590.04596.x

¹³ Croft S, Chauvenet ALM, Smith GC (2017) A systematic approach to estimate the distribution and total abundance of British mammals. *PLOS ONE* 12(6): e0176339. <https://doi.org/10.1371/journal.pone.0176339>

¹⁴ Mateo, R. G., Croat, T. B., Felicísimo, Á. M. and Muñoz, J. (2010), Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. *Diversity and Distributions*, 16: 84-94. doi:10.1111/j.1472-4642.2009.00617.x

¹⁵ Hijmans, R.J., Phillips, S., Leathwick, J. and Elith, J. (2017). dismo: Species Distribution Modeling. R package version 1.1-4. <https://CRAN.R-project.org/package=dismo>

¹⁶ <http://www.worldclim.org/bioclim>

¹⁷ Hadley Centre for Climate Prediction and Research (2017): UKCP09: Observed UK climate data (1961-1990). Centre for Environmental Data Analysis. <http://catalogue.ceda.ac.uk/uuid/87b3ab3b9bae47adab0c15d594d443b8>

¹⁸ Calabrese, J. M., Certain, G., Kraan, C. and Dormann, C. F. (2014). Stacking species distribution models. *Global Ecology and Biogeography*, 23: 99-112. doi:10.1111/geb.12102