Routledge
Taylor & Francis Group

Developing a predictive modelling capacity for a climate change-vulnerable blanket bog habitat: Assessing 1961–1990 baseline relationships

John Coll[a]*, David Bourke[b], Micheline Sheehy Skeffington[c], John Sweeney[a] and Michael Gormally[d]

[a]*Department of Geography, NUI Maynooth, Maynooth, Co Kildare, Ireland;* [b]*Applied Ecology Unit, Centre for Environmental Science, NUI Galway, Galway, Ireland. Department of Botany, NUI Galway, Galway, Ireland;* [c]*Department of Botany, NUI Galway, Galway, Ireland;* [d]*Applied Ecology Unit, Centre for Environmental Science, NUI Galway, Galway, Ireland*

Aim: Understanding the spatial distribution of high priority habitats and developing predictive models using climate and environmental variables to replicate these distributions are desirable conservation goals. The aim of this study was to model and elucidate the contributions of climate and topography to the distribution of a priority blanket bog habitat in Ireland, and to examine how this might inform the development of a climate change predictive capacity for peat-lands in Ireland.

Methods: Ten climatic and two topographic variables were recorded for grid cells with a spatial resolution of 10×10 km, covering $\sim 87\%$ of the mainland land surface of Ireland. Presence-absence data were matched to these variables and generalised linear models (GLMs) fitted to identify the main climatic and terrain predictor variables for occurrence of the habitat. Candidate predictor variables were screened for collinearity, and the accuracy of the final fitted GLM was evaluated using fourfold cross-validation based on the area under the curve (AUC) derived from a receiver operating characteristic (ROC) plot. The GLM predicted habitat occurrence probability maps were mapped against the actual distributions using GIS techniques.

Results: Despite the apparent parsimony of the initial GLM using only climatic variables, further testing indicated collinearity among temperature and precipitation variables for example. Subsequent elimination of the collinear variables and inclusion of elevation data produced an excellent performance based on the AUC scores of the final GLM. Mean annual temperature and total mean annual precipitation in combination with elevation range were the most powerful explanatory variable group among those explored for the presence of blanket bog habitat.

Main conclusions: The results confirm that this habitat distribution in general can be modelled well using the non-collinear climatic and terrain variables tested at the grid resolution used. Mapping the GLM-predicted distribution to the observed distribution produced useful results in replicating the projected occurrence of the habitat distribution over an extensive area. The methods developed will usefully inform future climate change predictive modelling for Ireland.

Keywords: active blanket bogs; conservation; climate change; adaptation; GLM; model validation; collinearity; topography

*Corresponding author. Email: john.coll@nuim.ie

22   *J. Coll* et al.

Introduction

An adaptation framework

Even the most restrictive global emissions policies proposed to date leave a sizeable chance that significant climate change will occur in the next few decades. Many of the projections suggest that this warming will surpass the 2° C warming limit adopted by the European Union (EU) that is suggested as a dangerous threshold beyond which we should not pass (Parry *et al.* 2009). This has led to suggestions that while 2° C provides a reasonable upper limit for target mitigation, it is dangerously misleading for informing the adaptation agenda (Anderson and Bows 2008). Other empirical work indicates that the emissions growth rate since 2000 has exceeded that for the Intergovernmental Panel on Climate Change (IPCC) A1FI scenario, and that no region is de-carbonising its energy supply (Raupach *et al.* 2007). These changes are coincident with a carbon cycle that is generating stronger than expected and sooner than expected climate change forcing (Canadell *et al.* 2007). A strong emissions policy must therefore be complemented with a plan to adapt to major changes with much more investment in adaptation than is currently planned (Parry *et al.* 2009, Swart *et al.* 2009).

Against this wider background, there is a growing recognition within the conservation sector that protected area networks will become increasingly important refuges for habitats and species in a warming climate, as well as being a key component in sustainable futures built on the green infrastructure which will provide vital ecosystem services for the delivery of cost effective climate change adaptation and mitigation. For Ireland specifically, adaptation initiatives in the natural resource management sector will further National Climate Change Strategy post-2012 aspirations by informing further cross-sectoral initiatives aimed at a lowering of the carbon intensity of the economy based on resilient green infrastructures (DOEHLG 2007).

Climate change and conservation

Anthropogenic impacts on global biodiversity have led to international agreements
AQ5 such as the Convention on Biological Diversity (CBD 1992) and the European Union (EU) Natura 2000 Network, established to protect endangered species and habitats in the EU based on the Birds Directive and the Habitats Directive (Donald *et al.* 2008). Member states including Ireland are bound by the Directives to monitor and report on the status of target species and habitats (Ostermann 1998); the National Parks and Wildlife Service (NPWS) has submitted a recent assessment report for Ireland to the European Commission (EC) (NPWS 2008).

Globally, there is evidence that species are shifting their ranges in response to changes in regional climates (Parmesan and Yohe 2003, Root *et al.* 2005, Walther
AQ1,2 *et al.* 2005, Lavergne *et al.* 2006); that species are altering their phenology (Menzel and Fabian 1999, Visser and Holleman 2001, White *et al.* 2003, Zavaleta *et al.* 2003, Jones *et al.* 2006, Donnelly *et al.* 2008) and that some species are facing extinction, or have become extinct (Parmesan 2006, Foden *et al.* 2007, Pauli *et al.* 2007). Therefore, developing effective conservation strategies that offset the climate change threats to species persistence will be critical in maintaining species and genetic diversity (Thuiller *et al.* 2008). The effects of climate change on biodiversity in Europe are

already observable through the changing distribution, migration and reproductive patterns of species (CBD 2007). Consequently, a major challenge is to enhance the ecological coherence of the Natura 2000 Network via extended habitat networks and linkages to increase the overall spatial coverage of natural and semi-natural habitats
95   (Gaston *et al.* 2008). Moreover, the changing relationship between biodiversity and climate will have profound implications for the economic and social well-being of Irish citizens and on the commitments to halt biodiversity loss.

Measures proposed to increase resilience include plans to expand protected areas, maintain varied and functional ecosystems and preserve good habitat quality
100   (Hopkins *et al.* 2007, Mitchell *et al.* 2007), as well as planning on the basis of the functional connectivity of habitats rather than simple structural connectivity. Ireland's mountains are largely coastal, and as with other maritime upland regions the steep climatic and environmental gradients result in a compression of broad latitudinal ecotones into a relatively small area (Coll *et al.* 2005), and hence add a
105   vertical component to adaptive landscape planning. Despite the need for site-specific impact assessments for priority habitats and species, scale-dependant controls on local topo-climates are not adequately captured in the present generation of global climate models (GCMs) and regional climate models (RCMs). Therefore until advances in climate system modelling improve the representation of local-scale
110   climate processes, there is a need to develop methods to better represent local projections of future changes to key climatic variables. Similarly, in order to buffer protected areas and identify migration corridors, GIS-based models that incorporate landscape ecology alongside other environmental variables are needed.

Blanket bogs and climate relationships

115   Globally, blanket bogs are rare, accounting for $\sim 3\%$ of the total peat-land area, and their distribution is restricted to temperate maritime regions typified by cool summers, mild winters and year-round rainfall (Lindsay *et al.* 1988, Warburton *et al.* 2004, Boylan *et al.* 2008, Kurbatova *et al.* 2009). In Europe, Atlantic blanket bogs are common only in Scotland and Ireland and constitute a significant global
120   component of this ecosystem (Douglas 1998, Sheehy Skeffington and O'Connell 1998). In Ireland, they cover about 6% of the land area and contain ca.19% of the nation's soil carbon stock (Tomlinson 2005). Atlantic blanket bog formation requires an annual rainfall $>1,200$mm (Crushell 2000, Boylan *et al.* 2008), although the amount of rainfall is probably less important than its distribution throughout
125   **AQ2**   the year (Lindsay *et al.* 1998, Charman 2002). Irish blanket bogs are divided into three sub-categories delineated by climatic and terrain controls:

(1) Lowland Atlantic Blanket Bog (*sensu* Moore 1962) or Oceanic Blanket Bog is confined to altitudes less than 150m (Schouten 1984) and areas with more than 250 rain days per year (Hammond 1984).
130   (2) Highland Blanket Bog (*sensu* Schouten 1984) is associated with an altitude range of 150–300m (Schouten 1984) and rainfall amounts of 1200–1250 mm/ year (Hammond 1984).
(3) Mountain or Upland Blanket Bog *(sensu* Moore 1962) occurs in areas above 300m where annual precipitation exceeds 1250 mm/year (Schouten 1984).

Although of high conservation value, only 21% of blanket bogs in Ireland remain in a relatively intact condition (Foss *et al.* 2001) due to peat extraction, drainage and forest plantation. Ireland has 50% of the remaining blanket bogs of conservation importance within the Atlantic Biogeographic Region of Europe (Foss *et al.* 2001), and while blanket bog is listed in Annex 1 of the EU Habitats Directive, only Active Blanket Bog, including areas capable of being restored to active status, is listed as a priority habitat (NPWS 2008). Therefore current conservation strategies need to be informed by a better understanding of these ecosystems (Douglas 1998), and the impacts of climate change must be considered. Climate change is expected to result in a decrease in the summer water table in peatlands through drier summers and alteration of pH, while modification of the nutrient cycle may lead to bogs becoming net emitters of carbon (Kurbatova *et al.* 2009). Most bog burst and peat slide events are triggered by high magnitude rainfall events (Crisp *et al.* 1964, Carling 1986, Dykes and Kirk 2001, Warburton *et al.* 2004, Dykes *et al.* 2008). UK and Irish data indicate that roughly half of all slippage events at present occur in the late summer months in relation to convective storm activity (Warburton *et al.* 2004). Therefore, associated with an increase in the intensity of convective activity more slippage events could be expected with climate change in the summer months, particularly if antecedent hotter and drier conditions have resulted in increased surface cracking (Sweeney *et al.* 2008). Other concerns relating to prospective seasonal changes include increases in winter rainfall leading to enhanced erosion. Overall therefore, bog habitats in Ireland are considered particularly vulnerable to the effects of a changing climate (Byrne *et al.* 2003, Jones *et al.* 2006, Donnelly *et al.* 2008).

A recent Principal Components Analysis (PCA) of grid-based climate data surfaces for Ireland confirmed these climatic controls on the distribution of Irish peat-lands and identified an east-west gradient primarily related to precipitation, and a north-south gradient related to temperature (Jones *et al.* 2006, Donnelly *et al.* 2008). Here the analysis of Donnelly *et al.* (2008) is extended by testing some of their data in an initial set of predictive models. The Donnelly *et al.* (2008) data were derived from 10×10 km climate surface interpolations (Sweeney and Fealy 2002Sweeney and Fealy 2003); although the GIS and data matching routines used in the analysis here indicate that the interpolated climate surfaces do not extend to islands and coastal boundary cells. There are sound climatologically-based reasons for this, but only values for the available cells are used here. In the analysis utilising the Donnelly *et al.* (2008) dataset, 848 10×10 km cells of the derived bioclimatic dataset are used in the predictive models.

The climate and habitat relationships based on a 1961–1990 baseline climatology are explored for each of the 10×10 km grid cells. A 30-year climatic average is useful in such an application as it smoothes out a lot of noise for use in an impact application (Carter *et al.* 1999). The use of this baseline also complies with the IPCC Task Group on Scenarios for Climate Impact Assessment (TGCIA) guidelines (Carter *et al.* 1999), and provides a common reference period in line with the World Meteorological Organisation (WMO) defined 'normal' baseline against which projected changes in climate can be assessed.

Aims and objectives

Until methods using more locally representative predictors of future change for key habitats and species can be developed, projections of climate change effects on such ecosystems cannot be made. This needs to be addressed in relatively small regions, such as Ireland, and projections should be based on an understanding of important current climatic controls on a case-by-case basis for both species and habitats. The way this might be achieved using climate and environmental variables as predictors in a regression-based approach for a selected vulnerable priority habitat is explored further here. While a number of automated model selection approaches are available, critical steps in model building are not as transparent or controlled as in manually conducted modelling exercises (Heikkinen *et al*. 2006). Multicollinearity among predictors is a known problem which hampers the analysis of species and habitat relationships in multiple regression settings (Heikkinen *et al*. 2004, 2006), and is an issue that has been poorly explored at finer spatial resolutions. The results reported here will be used to inform subsequent work using a mix of manual and automated modelling approaches for other candidate habitats and species.

Data and methods

Habitat data

Data on the distribution of the habitat on a 10×10 km grid was provided by the Irish National Parks and Wildlife Service (NPWS) for use in GIS analysis (ArcView GIS 9.3) alongside the climatic and environmental data. The NPWS maps are based on a combination of habitat maps, e.g. the Peatland Map (Hammond 1979), the Corine 2000 Land Cover Map and the Teagasc Soils Map (2006) and NPWS surveys (NPWS 2008). These have also been coupled with maps of important controls such as rainfall and altitude. Though these sources are to some degree incomplete and none fully depict the national resource of blanket bog, as with elsewhere across the EU (Evans 2006), the data used here are considered adequate (NPWS 2008). Here the Active Blanket Bog (Habitats Directive Annex I code 7130) community distribution is also mapped for comparison against the Natura 2000 (N2000) network of sites for both Ireland and Northern Ireland (NI) (Figure 1) and is the focus of the predictive distribution modelling. For the extended spatial analysis

AQ3 described in Section 2.3.3, spatial data for the NI Active Blanket Bog (code 7130) habitats from the Joint Nature Conservancy Council (JNCC) database are also incorporated in the models (JNCC 2007).

Climate and elevation data

Irish Grid (IG) co-ordinates for the 10×10 km grid cells coincident with mapped presences for the Irish Active Blanket Bog (code 7130) habitat were extracted from the NPWS GIS data alongside the 10×10 km climate data for the available cells. These 10×10 km resolution data are derived from monthly climate data for 560 precipitation stations and 70 stations for temperature spatially interpolated using a polynomial regression method with an inbuilt adjustment for elevation (Sweeney and Fealy 2002Sweeney and Fealy 2003). The Donnelly *et al*. (2008) derived bioclimatic

26   *J. Coll* et al.
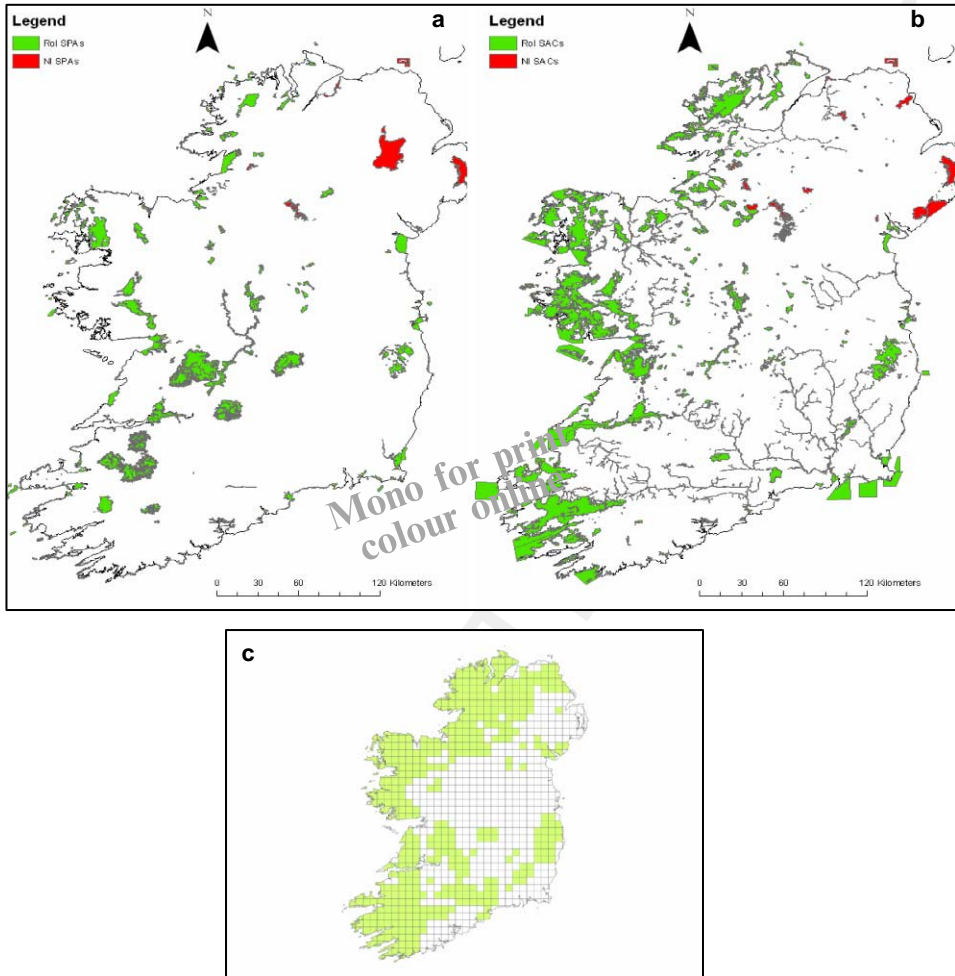
AQ1 Figure 1. The Irish and Northern Ireland Natura 2000 protected area network. a. Special Protection Areas (SPAs). b. Special Areas of Conservation (SACs). c. 10 x 10 km distribution of Active Blanket Bog (code 7130) habitat mapped from National Parks and Wildlife Service (NPWS) and Joint Nature Conservancy Council (JNCC) data respectively. See text for explanation.

data (Table 1) was also matched to the IG projections in order to test the variables as candidate predictors.

For the topographic data, the range in elevation variables (highest elevation – lowest elevation in the focal cell) was calculated in each grid cell using ArcGIS 9.3 AQ4 software (ESRI, Redlands, CA, USA) from a global digital elevation model (DEM) (GTOPO30) with a horizontal grid spacing of 30 arcs (approximately 1km). Mean elevation for each cell was also derived from the DEM and the data referenced to the climatic datasets. Presence (1) and absence (0) matrices were constructed for the Active Blanket Bog (code 7130) habitat using the extracted NPWS GIS data and referenced to the corresponding climatic and elevation data. Although there are 987

Table 1. Environmental predictors used in the distribution modelling of Active Blanket Bog (code 7130) habitat (n = 848).

Name	Definition	Unit	Mean	Minimum	Maximum
Climate					
T	Mean annual temperature	°C	8.9	6.1	10.4
Tmin	Mean temperature-coldest month	°C	4.2	0.95	6.3
Tmax	Mean temperature-warmest month	°C	12.8	10.4	13.9
P	Total mean annual precipitation	mm	1205.3	704.5	2025.2
Ps	Mean summer precipitation (June-August)	mm	227.9	148.2	342.3
Pw	Mean winter precipitation (December-February)	mm	383.6	208.5	693.9
Eto	Mean annual potential evapotranspiration[1]	mm	615.4	548.5	646.8
Pn	Mean annual net precipitation	mm	589.9	70.8	1426.0
Hu	Mean annual humidity index	mm	38.7	20.7	86.6
Ic	Continentality Index[2]		17.1	14.1	18.8
Topography					
ELmean	Mean elevation[3]	m	115.4	1.00	495.03
ELrng	Elevation range[3]	m	177.69	0.00	895.00

AQL2 Notes: (1) after Thornthwaite; (2) mean temperature range warmest and coldest months of the year; (3) based on the 841 10km grid cells used for this part of the analysis

10×10 km resolution terrestrial grid cells associated with the extraction routines from the DEM, climate and climate change data is not available for all these grids as explained above.

Model variable selection and testing

Logistic regression (LR), with a logit link function and a binomial error distribution, is a generalised linear modelling approach and is a popular method for characterising species or community–environment relationships (Guisan and Zimmermann 2000). The LR model with a logit link function can be written as:

$$\ln\left(\frac{P}{1-P}\right) = B_o + B_1 \ x_1 + \ldots B_{n+1} \ x_{n+1}$$

This can be rewritten in terms of odds rather than log odds as:

$$\frac{P}{1-P} = e^{B_o + B_1 \ x_1 + \ldots B_{n+1} \ x_{n+1}}$$

$$P = \text{probability of an event}$$
$$E = \text{base of the natural logarithms}$$
$$X = \text{independent variable}$$
$$B_0, \ B_1 = \text{coefficients estimated from the data}$$

28   *J. Coll* et al.

245   This special type of generalised linear model (GLM) for binary response variables
compares favourably with more complex machine-based approaches, yielding good
AQ5 results (Elith ~~*et al.*~~ 2002, 2006) and is one of a suite of modelling approaches applied
in bioclimatic envelope modelling (Table 2). GLMs have the added advantage of
fitting probability distributions to the variable being modelled, thereby offering an
250   extension to normal linear regression techniques which only model the mean of the
distribution (Fealy and Sweeney 2007). Thus in GLMs, data may be assumed to be
from several families of probability distributions, including the normal, binomial,
poisson, negative binomial, or gamma distribution, many of which better fit the non-
normal error structures of most ecological data (Austin 1987, Guisan *et al.* 2002).
255   While more sophisticated machine-based modelling methods including Artificial

Table 2. Summary of the statistical techniques, and their abbreviations, applied in bioclimatic
envelope modeling (Heikinnen, *et al.*, 2006).

Study	Modelling methods
Brereton *et al.* 1995, Beaumont and Hughes 2002, Kadmon *et al.* 2003, Meynecke 2004, Beaumont *et al.* 2005	BIOCLIM
Box *et al.* 1993, 1999; Crumpacker *et al.* 2001	'The Florida Model'
Walker and Cocks 1991	HABITAT
Carpenter *et al.* 1993	DOMAIN
Baker *et al.* 2000	CLIMEX
Skov and Svenning 2004; Svenning and Skov 2004	Fuzzy minimal rectilinear envelope Modeling
Sykes *et al.* 1996; Walther *et al.* 2005	STASH
Iverson and Prasad 1998, 2001, 2002	Classification and regression tree analysis (CTA/CART/RTA)
Guisan and Theurillat, 2000, Price, 2000 Bakkenes *et al.* 2002, Burns *et al.*, 2003	Logistic regression/binomial GLM
AQ5 thwick *et al.* 1996, Midgley *et al.* 2003 Araújo *et al.* 2004, Luoto *et al.* 2005	GAM
Beerling *et al.* 1995, Huntley *et al.* 1995, 2004 Hill *et al.* 1999, 2002	Locally weighted regression (local regression/loess)
Berry *et al.* 2002, Pearson *et al.* 2002, 2004	ANN
AQ5 erson 2001, Anderson *et al.* 2002a, 2002b, Peterson *et al.* 2002a, 2002b, 2004	GARP
Prasad and Iverson 2000	MARS

Notes: ANN = artificial neural networks; GAM = generalized additive models; GARP = genetic
AQ12 orithm for rule-set prediction; GLM = generalized linear models; GM-SMAP = Gaussian mixture
distributions and multiscale segmentation; MARS = multivariate adaptive regression splines

Neural Nets (ANNs) are available, GLMs in certain applications have been found to perform comparably well (Marmion *et al.* 2009). In the modelling approach described here, GLMs were used on the basis that a manual model building process would provide a greater insight into the spatial structures and dependencies within the data.

Once the model has been built and predictions produced, it is necessary to determine how effective that model is at predicting the dependent variable, in this case the presence or absence of the Active Blanket Bog (code 7130) habitat. Prior to model building a correlation matrix for all ten of the candidate predictor climatic variables was derived. Unsurprisingly, there are strong and positive correlations between the various temperature and precipitation variables (Appendix Table A1). Given the importance of precipitation and temperature gradients as controls on the Active Blanket Bog (code 7130) habitat distribution noted above, it would be expected that these should contribute component variables to a predictive model.

To refine the candidate variable selection, a backward stepwise selection procedure, based on the Wald Statistic was undertaken in SPSS to successively select the most relevant predictors among the ten climatic variables incorporated (McCullagh and Nelder 1989). Binary LR with a backwards, stepwise-variable selection is a common statistical method used in conservation biology to estimate occurrence probabilities in relation to predictors (Kleyer *et al.* 1999, Cowley *et al.* 2000, Bässler *et al.* 2010). Backward elimination over forward selection is frequently preferable on the basis that it is safer to delete terms from an overly complex model than to add terms to an overly simple one (Agresti 2002). It was also considered that a stepwise selection procedure would enable the production of a parsimonious and accurate predictive model based on the successive elimination of non-explanatory predictor variables. Stepwise selection of variables is a useful and effective data analysis tool when the outcome being studied is relatively new (as was the case here), the importance of individual covariates may not be known, and associations with the outcome are not well understood (Hosmer and Lemeshow 1989). Additionally, when aiming at prediction with regression analysis, valuable insights can be developed by undertaking a sequential regression (Graham 2003). There is also a recognition that classification success using LR is sensitive to the relative proportion of presences and absences in the sample, independently of the fit of the model (Hosmer and Lemeshow 1989).

By definition, the logistic function is symmetric and its inflection point corresponds to a probability (p) value of 0.5. This value is commonly used as a default threshold above which it is assumed the model predicts presence, and is the classification cut-off point reported on here. However, when the proportions of presences and absences are not equal within the sample, the LR output within the logit function's domain is not symmetrical, and deviates towards the extreme with a greater number of cases (Real *et al.* 2006). LR is noted for producing scores biased towards the larger group (Hosmer and Lemeshaw 1989), therefore unequal group sizes (prevalence) can influence the scores (Fielding and Bell 1997). However, given the balance of the presence versus absence data for both sets of analysis here it is not heavily skewed, for this analysis the binomial distribution is considered to be appropriate and the 0.5 threshold suitable. This lack of zero-inflation in the binary count data for the habitat also avoided the need to apply more complicated zero-inflated generalised poisson (ZIGP) models (Czado *et al.* 2007).

30   *J. Coll* et al.

Statistical software applications

305   The initial stepwise selection procedures were undertaken in SPSS, with data exploration also undertaken in the S+ and R-based Brodgar (Zuur *et al*. 2007) statistical computing environments (Appendix Table A2). Brodgar was used to cross-check the decrease in residual deviance associated with each step in the backwards selection procedure to identify the relevant predictors (McCullagh and Nelder 1989).

310   The observed and fitted graph for the LR model are provided (Appendix Figure A1), together with the S+ derived residuals plots, alongside the normal Q-Q, scale location and observed and fitted plots (Appendix Figure A2a-d). Spatial Analysis in Macroecology (SAM) freeware (Rangel *et al*. 2006) was also used in order to explore alternative output measures for model performance (Appendix Table A3). The

315   climate data was further explored in a parallel multivariate (MV) environment (CANOCO) to complement the analysis in the univariate (UV) environments (S+, SPSS, SAM, Brodgar). It was considered that this two stage approach would inform future work by addressing a key challenge in bioclimatic modelling, i.e. to arrive at an understanding of the limitations of models by calibrating and validating models

320   with data that are distinctively independent of each other (Heikkinen *et al*. 2006).

Results

Specifying an initial model

AQ6   Although the residuals versus fitted plot for the model shows a distinct pattern (Appendix Figure A2), this is because of the presence-absence nature of the data and

325   does not indicate a lack of fit (Zuur *et al*. 2007). However, the graph is useful for checking that there are no samples with a score of 1 in the string of points for the 0s and vice versa, whereas the QQ-plot (Appendix Figure A2b) can be ignored for presence-absence data as the residuals will never be normally distributed (Zuur *et al*. 2007). Nagelkerke's R-square (R_{2N}) (Nagelkerke 1991) values from the SPSS outputs

330   were used to further evaluate model calibration as this measure quantifies the proportion of variance explained by the model, with values exceeding 0.4 indicating a good calibration (Harrell 2001, Steyerberg *et al*. 2001, Reineking and Schroder 2006). Typically R_{2N} varies between 0 where independent variables fail entirely to predict the dependent variable, and 1 where the independent variables in the model

335   predict the dependent variable perfectly. Conversion of the GLM R_{2N} ratio of 0.592 to a percentage indicates that 59.20% of variance in the dependent variable is accounted for by the predictors ($p < 0.001$) used in the LR equation.

Applying these routines between the various statistical applications reduced the predictor variables to four from the original 10 with a predictive accuracy of 74.0%

340   and 86.8% for the presence and absence data respectively. Some authors have criticised the use of the Wald Statistic; for example Hauck and Donner (1977) found that it behaved in an aberrant manner, often failing to reject when the coefficient was significant. This has led some authors to advocate the use of the likelihood ratio test (Hosmer and Lemeshaw 1989, Fitzmaurice *at al*. 2004, Zuur *et al*. 2007), although

345   others argue that for large samples the two tests usually give similar results (Agresti 2002). However, running the four predictors on a backward stepwise routine had no effect on the model output scores or the values of the variables obtained in the equation. Global measures of model performance are preferred by researchers, with

overall accuracy (OA) being the most commonly used in ecology (Fielding and Bell
1997), i.e. the probability that a site (either presence or absence) is correctly predicted
(Liu *et al.* 2009). In this context the percentage prediction outputs from SPSS are
interpreted as the OA score of the fitted model.

The model was cross-validated against different indicator scores for the same four
predictor variables in SAM; thus for example, SAM returned a McFadden's Rho-
Squared value of 0.428 for the same Chi Square value (489.04) obtained in the SPSS
analysis. A transformation of the likelihood ratio statistic, McFadden's Rho-Squared
AQ7 value is intended to mimic the R-squared values associated with Ordinary Least
Squares (OLS) regression in LR. Low Rho-squared values do not imply poor model
fit since these are usually much lower than R-squared values and a rho-squared value
of 0.428 is considered very satisfactory (Hensher and Johnson 1981). Similarly,
Kappa a satisfactory indicator of model performance (Manel *et al.* 1999) is recorded
at 0.534 in the SAM outputs, indicating a good predictive power (Monserud and
Leemans 1992). SAM also returned a True Skill Statistic (TSS) of 0.497 for the
model, with some authors (e.g. Allouche *et al.* 2006) arguing that TSS also known as
the Hanssen–Kuipers discriminant compensates for the shortcomings of Kappa
while preserving its advantages, and providing results closely correlated with those of
AQ3 the threshold-independent AUC statistic (Section 3.4).

Collinearity tests for identified predictors

While the selected variables made sense climatically in terms of the study habitat, it
was decided to test the statistical basis of the initial GLM further before attempting
to refine it by the incorporation of elevation data as an additional predictor. Having
re-examined the correlation matrix scores, further tests for collinear predictor
variables contributing to multicollinearity in the results were undertaken. Multi-
collinearity occurs when one or more variables are exact or near exact linear
functions of other variables in the data set (Munoz and Felicisimo 2004) and is a
common problem associated with organism and community modelling (Brown 1994,
De Veaux and Ungar 1994). Hence multicollinearity among predictors may hamper
the accurate analysis of species and habitat relationships in multiple regression
settings (MacNally 2000, Luoto *et al.* 2002, Heikkinen *et al.* 2004, Heikkinen *et al.*
2006); and despite these known problems, relatively little attention has been paid to
multicollinearity in bioclimatic modelling studies (Guisan and Thuiller 2005, Luoto
AQ6 *et al.* 2006). This is a significant issue, since multicollinearity (i.e. intercorrelation)
among *X*s means that causal *X*s may be lost from final application models because
other, non-causal *X*s are correlated with those causal variables and may be retained
in models at their expense (MacNally 2000).

Since the MV scaling applied in the Principal Components Analysis (PCA)
focused on variable correlations, the relative direction of the arrows (Appendix
Figure A3) approximates the linear correlation coefficients among the variables
represented in Appendix Table A1 (Leps and Smilaur 2003). These findings support
the use of data reduction techniques such as PCA to reduce the dimensions of
predictor data sets (Gates and Donald 2000, Suarez-Seoane *et al.* 2002, Munoz and
Felicisimo 2004); with the MV approach indicating strong collinearity among some
of the variables analysed here. This applies to both the full range of variables, and to

the temperature and precipitation variables identified in the initial model specifica-
tion.

The UV approach confirms this assessment, particularly the scatter associated
with the Pairplots for the temperature and precipitation variables (Appendix Figure
A4a, b). With the PCA also indicating that mean annual potential evapotranspira-
tion (Eto) and the mean annual humidity index (HU) are collinear with temperature
and precipitation respectively, this was tested further using a standard OLS
regression in both SPSS and SAM. The high r^2 values obtained using T as a
predictor for Eto ($r^2 = 0.992$; $p < 0.001$) and P as a predictor for HU ($r^2 = 0.902$;
$p < 0.001$) alongside the high Akaike's Information Criterion (AIC) scores in SAM
provide further indication of this.

For the refined model a reduced set of predictor variables were tested further.
These comprised T, P and Ic from the climate data and ElRng from the elevation
variables, with further PCA analysis indicating collinearity between the elevation
variables (Appendix Figure A5). A mix of further stepwise selection routines in UV
OLS regression confirmed that mean elevation was less significant as a predictor
than elevation range. CANOCO-based selection routines also detected collinearity of
Ic with T and P as indicated by high variance inflation factors (VIFs). However,
elimination of Ic from the model substantially reduced the VIFs for the three
remaining variables.

These procedures support the elimination of variables based on VIFs as
indicators of correlation between variables (Brauner and Schacham 1998, Cawsey
et al. 2002, Elith and Burgman 2002, Guisan *et al*. 2002) at the model building stage.
Further testing of the four variable (T, P, Ic, ElRng) model versus the three variable
(T, P, ElRng) reduced model in SAM also supported dropping Ic from the final
model based on test score outputs. For the four variable model a McFadden's Rho-
Squared score of 0.599 was returned with a Kappa and TSS of 0.695 and 0.694
respectively, but with statistically insignificant p-values for T ($p = 0.557$) and Ic
($p = 0.172$). Whereas the three variable fitted model (Figure 2) returned Kappa and
TSS scores of 0.755 and 0.755 respectively against a McFadden's Rho-Square value
of 0.599 with all p-values < 0.05. On all other performance measures, the three
variable model offered an improved predictive skill, thus for example, there is a
reduced AIC score (477.67). The reduced predictor model also provided an increased
Nagelkerke's R_{2N} score (0.751), together with an improved predictive accuracy for
both presence (86.1%) and absence (89.6%) data (Appendix Table A5) as a measure
of OA (Fielding and Bell 1997).

Combining climate and topographic variables to specify a refined model

To be able to correctly specify a model in an ideal world, the factors that determine
the exact nature of the relationship between the presence-absence data and each
predictor variable must be known (i.e. the appropriate functional form of the
variable) (Pearce and Ferrier 2000a). However, as this level of information is rarely
available in regional studies of species and habitat distributions, more exploratory
approaches must be adopted in which the data available determine the model (Pearce
and Ferrier 2000a).

However, elevation is *a priori* a known control on the distribution of blanket bog
via its influence on temperature and precipitation, therefore elevation data was
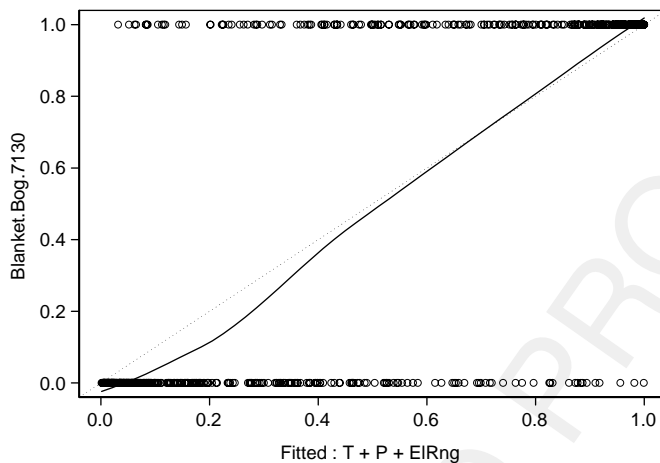
Figure 2. Observed and fitted probability values for the Active Blanket Bog (code 7130) **AQ12** ꞊itat. Presence-absence data (0s and 1s denoted as circles). Hypothetical fitted probability (dotted line) and fitted refined logistic regression model (solid line).

440   included in the revised LR GLM. For this analysis, sample n = 841 10×10 km grid squares in total, due to the spatial mismatches between the available climate and habitat presence data indicated previously, the number of records was extended by inclusion of the NI Active Blanket Bog (code 7130) habitat presence-absence data. Thus the refined model was tested against all the available habitat records for
445   Ireland, with presence (n = 417) and absence (n = 424) for a total of 841 grids comparing to presence (n = 342) and absence (n = 506) for a total of 848 grids in the preceding model. The addition of the NI data also usefully balanced the group scores for the refined model given the prevalence biases associated with LR.

Model performance evaluation

450   The various approaches to model building explored here indicate the sort of problems associated with classical regression analysis that arise when many predictors are used. An increase in the number of predictors implies a greater increase in the number of possible regression structures, and the almost inevitable problem with multicollinearity (Munoz *et al.* 2004). Having dealt with multi-
455   collinearity, testing the refined model is a vital step in model development (Pearce and Ferrier 2000b). Despite the increasing use of statistical modelling techniques such as LR, relatively little attention has been paid to the development and application of appropriate evaluation techniques for assessing the predictive performance of habitat models (Pearce and Ferrier 2000b).
460   However, both Metz (1986), and Fielding and Bell (1997) have identified an index that represents an unbiased discrimination index derived from the area under a relative operating characteristic curve (ROC) (Pearce and Ferrier 2000b). The refined model's performance was therefore assessed using a threshold independent method: the area under the ROC curve, commonly termed AUC (Munoz *et al*. 2004). The
465   ROC curve is recommended for comparing two class classifiers as it does not merely summarise performance at a single arbitrarily selected decision threshold, but across

34   *J. Coll* et al.

all possible decision thresholds (Fielding and Bell 1997, Harrell 2001, Bässler *et al.* 2010, Liu *et al.* 2009). It plots the sensitivity (i.e. true positives) versus (1– specificity)

AQ8   (i.e. false positives) (Pearce and Ferrier 2000, Munoz *et al.* 2004). Thus an ideal classifier hugs the left side and top side of the graph and the area under the curve equals one.

A random classifier should achieve *ca*. 0.5, and whilst most analyses use 0.5 as the decision threshold to consider a case as present or absent, this value is arbitrary and it does not necessarily give a more accurate model (Fielding and Bell 1997, Manel *et al*. 1999). In addition, ROC curves are invariant under changing distributions of the binary classes (presence-absence), as they actually plot the 'percentage of class-1 observations' vs 'percentage of class-0 observations' and are therefore independent of the balance between the two classes (Munoz *et al*. 2004).

AUC, traditionally used for assessing the accuracy of weather forecasts (Saseendran *et al*. 2002, Elmore *et al*. 2003, Accadia *et al*. 2005), compares the number of correct forecasts, minus those attributable to random guessing, to that of a hypothetical set of perfect forecasts, and has a long history in this field (Finley 1884). Hanley and McNeil (1983) have shown that when dealing with a single scoring model, the AUC is equal to the empirically observed probability of a class-1 observation attaining a higher score than a class-0 observation. It has also been shown that the AUC is equivalent to the normalised Mann-Whitney two-sample statistic, making it equivalent to the Wilcoxon statistic (Bambar 1975, Sokal and Rohlf 1981, Pearce and Ferrier 2000b). AUC is a measure of model accuracy, but it does not provide a rule for the classification of cases (Fielding and Bell 1997, Liu *et al*. 2005) and has recently received some criticism (Lobo *et al*. 2008). However, rates higher than 0.9 indicate very good discrimination because the sensitivity rate is high relative to the false positive rate (Swets 1988). Therefore the refined model developed here demonstrates a very good discrimination ability, with an ROC area of 0.941; SE $\pm$ 0.006 (Figure 3).

Using the casewise listing outputs from SPSS, model predicted outputs for each grid were imported back into GIS to visually compare the model performance with the mapped presence and absence data for the habitat (Figure 4 and 5). This enabled the mapping of outliers (SE $\pm$ 2.0) in the GLM-predicted values for wrongly called false presences (FPs) and false absences (FAs) against the observed habitat distribution (Figure 4). It has been suggested that FPs close to real positives may be less serious errors than FPs distant from a real positive (Fielding and Bell 1997) as is the case with results here. This also enabled all the GLM-predicted presence/absence calls to be mapped against the habitat distribution maps to visually represent GLM predictive performance overall. Figure 5 visually reinforces the good predictive skill of the refined GLM as would be expected from the statistical performance measures (Appendix Tables A4 and A5), and when considered alongside the AUC score from the ROC. While not mapped here, many of the GLM-predicted FP and FA calls appear to be associated with the influence of topographic variation on orography at a local scale. Specifically, the FPs and FAs seem to be associated with lee areas in the rain shadow of uplands, although this is not uniformly the case. These findings may indicate a need for data splitting in subsequent work in order to analyse the spatial pattern in and around upland regions separately. Alternatively there is a likely need to incorporate a predictive component for slope aspect, for example, in order to refine the predictive capacity of subsequently developed models.
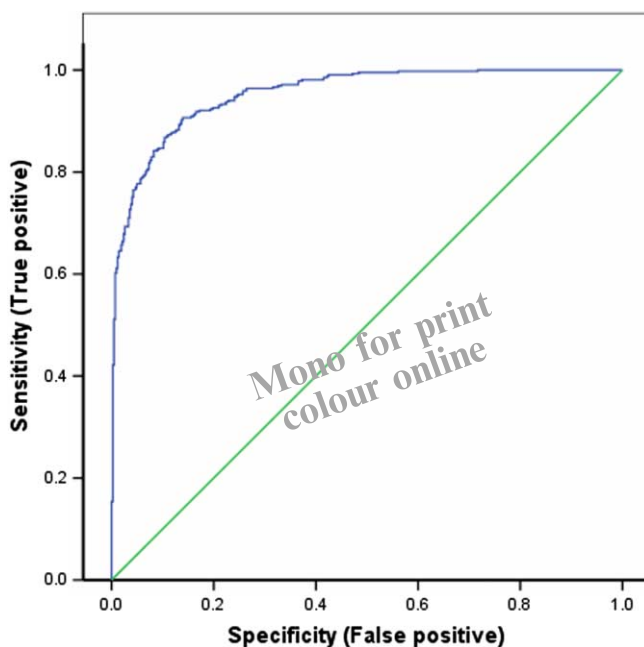
AQ12 Figure 3. The ROC curve for the climate and elevation variables model (AUC = 0.941; SE ± 0.006).

Discussion

The study in context

The main finding of this study was that following extensive screening for collinearity, the subsequent inclusion of elevation range in the grid cells improved the predictive accuracy of the GLM developed for the Active Blanket Bog (code 7130) habitat distribution. This finding is unsurprising since climatic and topographical gradients operate at different spatial scales, with the latter nested in the former (Bruun *et al.* 2006), hence the inclusion of the elevation data provide an additional and more local component for the refined GLM. Similarly, by extending the spatial extent of the habitat presence-absence data by the inclusion of NI data, GLM performance was improved (Appendix Tables A3 and A4). It has been suggested that the limitations of bioclimatic models must be understood in order to make them useful tools for the scientific community and policy-makers (Araújo *et al.* 2005a, Heikinnen *et al.* 2006, Lawler *et al.* 2006, Beale *et al.* 2008). In this regard and based on the results here, it is suggested that the lessons learned at the model building stage are crucial. The various approaches applied to model building and variable selection emphasise the importance of understanding climate-environment relationships in the present prior to applying more sophisticated machine-based modelling methods for the prediction of future climate change effects on distributions.

The approach therefore supports the view that a manual model-building process gives a better control to the modeller than automated techniques and enables the development of ecologically plausible models (Nichols 1989, Crawley 1993, Pausas *et al.* 2003). In particular, the extensive routines aimed at identifying collinearity

36   *J. Coll* et al.
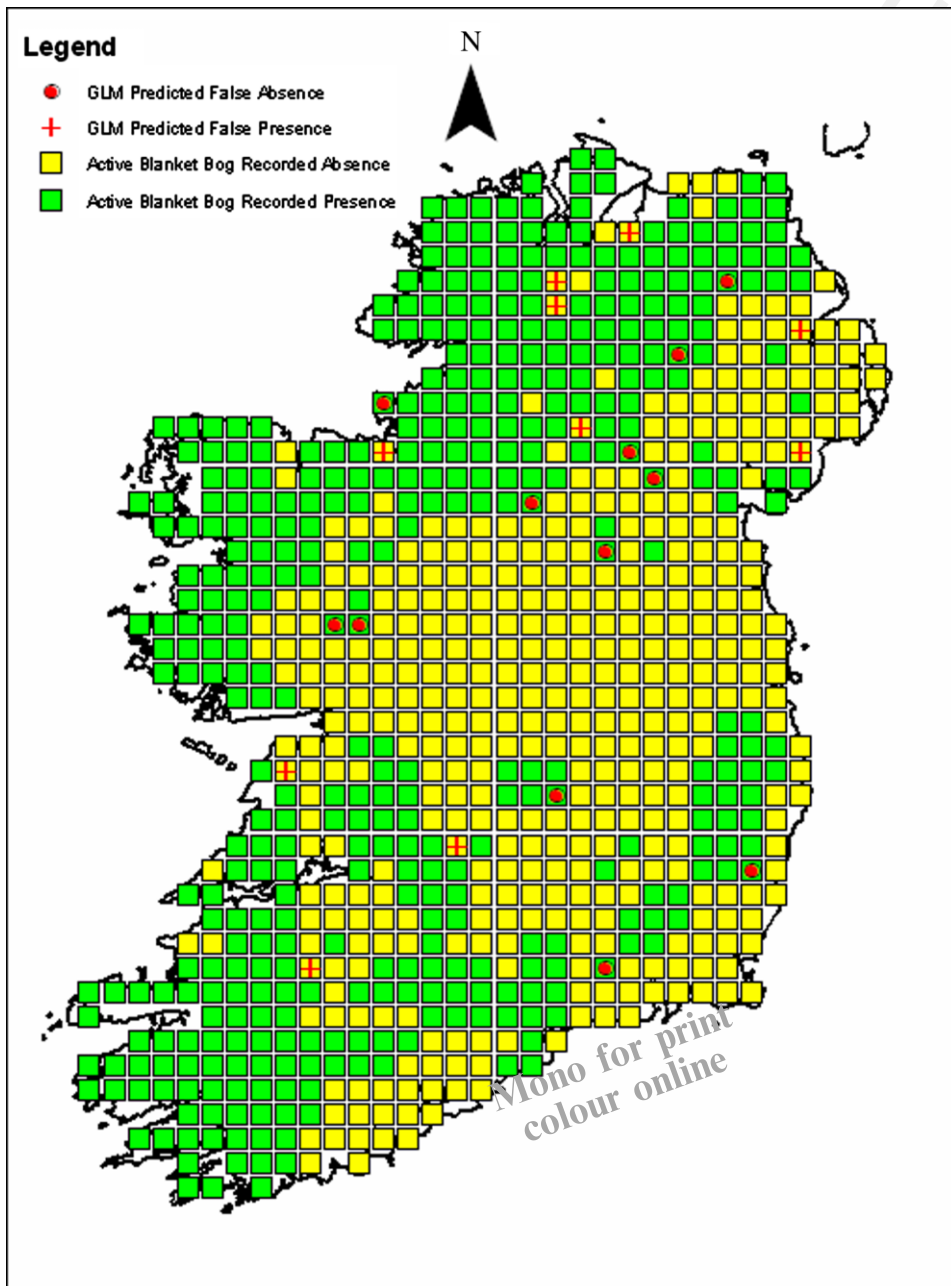
Figure 4. 10 x 10 km grid resolution recorded presence and absence for the Active Blanket Bog (code 7130) habitats (n = 841).
Note: Red circles denote GLM misclassified false absence (0 for 1; n = 12); red crosses denote GLM misclassified false presence (1 for 0; n = 10). Mapped casewise list for all outliers AQ12 ± 2.0)
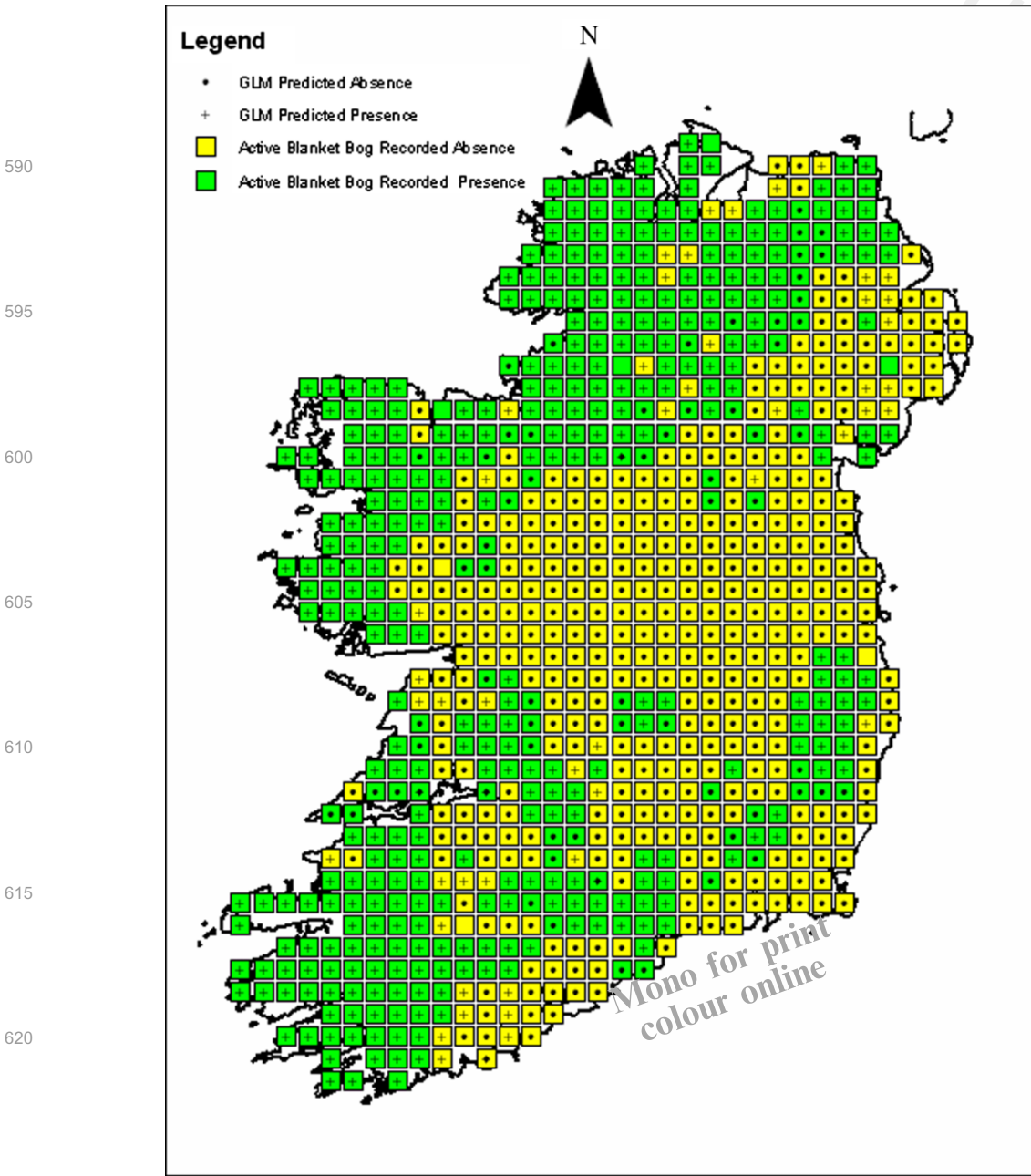
Figure 5. 10 x 10 km grid resolution recorded presence and absence for the Active Blanket Bog (code 7130) habitats (n = 841).
Note: Overlain black circles denote GLM predicted absences; overlain black crosses denote GLM predicted presences. Presence-absence coded grids with no symbols denote missing climate data (see text for explanation)

among candidate predictors and hence eliminating multicollinearity in results emphasises the importance of rigorous test routines at the model construction stage. Similarly, the parallel exploration of the spatial data in MV and UV environments in tandem with GIS-based routines emphasises the benefits of such an approach and contributes to an improved understanding when constructing models. This process also tends to emphasise there is no 'one size fits all' method or software when attempting to extract the complex spatial information contained in climatic and environmental data, and when trying to identify appropriate predictors for future climate driven changes.

Both climate and topography appear in this study as important determinants for the distribution of the Active Blanket Bog (code 7130) habitat communities. Again, this is not surprising, since the current paradigm is that climate governs species distribution on a broad scale (Currie 1991, Wright *et al*. 1993, Huntley *et al*. 1995, H-Acevedo and Currie 2003, Thuiller *et al*. 2004), whereas the topography, geology, land cover and spatial configuration of suitable habitats affect species occupancy patterns at finer spatial resolutions (Pearson *et al*. 2004, Thuiller *et al*. 2004). Given the importance of these controls in the present, changes to the key controls of temperature and precipitation identified here are likely drivers of future changes for the habitats and their associated communities and species, certainly at least at landscape and regional scales.

In this context the predictive models developed here only emphasise earlier findings in relation to the key controls on the Active Blanket Bog (code 7130) habitat **AQ2** distribution and the likely future drivers of change (Moore 1962, Hammond 1981, Schouten 1984, Lindsay *et al*. 1988, Crushell 2000, Byrne *et al*. 2003, Jones *et al*. 2006, Donnelly *et al*. 2008, Sottocornola *et al*. 2009). However, this study is the first so far applied to Irish bog habitats using climate and terrain variables as predictors in a GLM-based approach where such close attention has been paid to the wider issue of addressing multicollinearity in bioclimatic modelling generally. This study has also been the first in Ireland to bring together such a spatially coherent mix of geo-referenced climatic, elevation and habitat distribution data on an island-wide basis for use in the development of a predictive modelling capacity.

However, the limitations and assumptions involved in modelling at a 10×10 km grid resolution are recognised. Thus there is awareness that important controlling variables such as topographic and environmental heterogeneity will be lost at this resolution, together with important local micro-climatic controls. This is particularly important for the Active Blanket Bog (code 7130) habitats where high rates of change to surface patterning will be a site scale response, and hence between and within habitat diversity should respond quickly even along spatially short gradients in response to local topography. For similar scale-dependant reasons no account can be taken of the relative coherence or patchiness of the Active Blanket Bog (code 7130) habitat within individual 10×10 km grids where the community presence is recorded. The interested reader is therefore directed to the NPWS (2008) and JNCC (2007) survey records for further information on the survey methods. Nevertheless, the scale of analysis here represents a refinement on the use of 50×50 km grids elsewhere which are a blunt tool in regions such as the Alps and the Pyrenees where average climate on this scale is a poor representation of the conditions experienced within most of the square (Beale *et al.*, 2008). For similar scale-dependant reasons, the use of 50×50 km regional climate model (RCM)

generated climate change data also has limited utility in topographically diverse maritime regions (Coll *et al*. 2005, Coll *et al*. 2010).

Future prospects

The approach developed here and the spatial data used represents a significant advance in developing a predictive modelling capacity for Irish habitats and species, and one which will usefully inform future work and extend the available ecological modelling capacity. Thus by extending the ecological and environmental information contained in the database and by extending the available environmental and climate change data, additional candidate predictors and factors for other habitats and species can readily be incorporated. Similarly, by having the species presence data corresponding to these and other priority habitats, more ecologically coherent mixed effects models can be developed. Similarly, more refined Generalised Additive Models (GAMs) can be developed and used to incorporate habitat presence, for example, as a factor for predicting presence-absence of candidate species alongside other climatic and environmental data.

It has been suggested that implementing and applying a suite of modelling approaches in consort with ANNs to form a multi-model ensemble forecast would constitute a more rigorous approach (Araújo *et al*. 2005a). Such a modelling framework would be taken to include GLMs and GAMs using conventional statistics in tandem with ANNs. However, while some work has shown that ANNs project more accurately than GLMs and GAMs (Thuiller 2003, Araújo *et al*. 2005b), there is also a suggestion that ANNs are prone to overfitting. Despite this, ANNs are able to handle explanatory variables from different sources, such as categorical and boolean data (Heikkinen *et al*. 2006), and are better able to determine climatic envelopes that have non-linear responses to predictors (Pearson *et al*. 2002, 2004). Overall therefore, the advantages of automated model building for future applications are recognised, but it is considered that the approach developed here in manually formulating and checking models in an iterative and interactive way between applications will better inform the future application of these methods for Ireland. Thus if applied alongside an ANNs-based approach, the use of GLM and GAMs in parallel applications are likely to provide more insights into the contribution of the predictors in the prediction process, alongside the ability to examine species response curves to environmental gradients masked in ANNs.

Issues of statistical rigour associated with bioclimatic models aside, confidence in any predictive models derived using climate change data will be intrinsically linked to the quality of the RCM-generated data used to prime them. Fortunately, RCM data at 14×14 km grid resolution has recently been made available for Ireland (McGrath and Lynch 2008). These provide up to date climate change data for key climatic variables derived from an RCM double nested in two driving global climate models (GCMs) (McGrath and Lynch 2008), and are the data sources most likely to be used in subsequent modelling projecting the impacts of future climate change.

Conclusion

Predicting climate-change driven shifts in species and habitat distributions has been dominated by continental-scale studies, although regional scale applications are of

40   *J. Coll* et al.

most relevance in relation to conservation goals. Similarly, many of the 50×50 km grid-scale studies have taken no account of topographical heterogeneity or attempted to model habitat or species distributions at finer scales with an integrated topographic component. The omission of this sort of information is likely to introduce large and systemic biases in bioclimatic models, hence the inclusion of more localised environmental variables in relation to a reference baseline climate is a desirable goal if improved predictive modelling capacity for future climate-driven changes is to be achieved.

This paper has presented a finer scaled modelling approach for integrating climate and elevation variables at a regional scale, and has explored the important issues of careful variable selection and the elimination of collinearity at the model building stage. Extensions of the approach to other habitats and species, together with the incorporation of additional environmental variables should help further our understanding of climate-biota relationships in Ireland. By clarifying an understanding of these in the present, it is suggested that more accurate and locally-relevant climate change impact modelling assessments for candidate species and habitats can be undertaken.

References

Accadia, C., Mariani, S., Casaioli, M., Lavaqnini, A., and Speranza, A., 2005. Verification of precipitation forecasts from two limited-area models over Italy and comparison with ECMWF forecasts using a resampling technique. *Weather and Forecasting*, 20, 276–300.

Agresti, A., 2002. *Categorical data analysis: Second edition*. Hoboken: Wiley & Sons.

Allouche, O., Tsoar, A., and Kadmon, R., 2006. Assessing the accuracy of species distribution models: Prevalence, Kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43, 1223–1232.

Anderson, K. and Bows, A., 2008. Reframing the climate change challenge in light of post-2000 emission trends. *Philosophical Transactions of the Royal Society A*, 366, 3863–3882.

Araújo, M.B., Cabeza, M., Thuiller, W., Hannah, L., and Williams, P.H., 2004. Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. *Global Change Biology*, 10, 1618–26.

Araújo, M.B., Pearson, R.G., Thuiller, W., and Erhard, M., 2005a. Validation of species-climate impact models under climate change. *Global Change Biology*, 11, 1504–1513.

Araújo, M.B., Whittaker, R.J., Ladle, R.J., and Erhard, M., 2005b. Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*, 14, 529–538.

Austin, M.P., 1987. Models for the analysis of species response to environmental gradients. *Vegetatio*, 69, 35–45.

Baker, R.H.A., Sansford, C.E., Jarvis, C.H., Cannon, R.J.C., MacLeod, A., and Walters, K.F.A., 2000. The role of climatic mapping in predicting the potential geographical distribution of non-indigenous pests under current and future climates. *Agriculture Ecosystems and Environment*, 82, 57–71.

Bakkenes, M., Alkemade, J., Ihle, F., Leemans, R., and Latour, J., 2002. Assessing the effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology*, 8, 390–407.

Bambar, D., 1975. The area above the ordinal dominance graph and the area below the receiver operating graph. *Journal of Mathematical Psychology*, 12, 387–415.

Bässler, C., Muller, J., Hothorn, T., Kneib, T., Badeck, F., and Dziock, F., 2010. Estimation of the extinction risk for high-montane species as a consequence of global warming and assessment of their suitability as cross-taxon indicators. *Ecological indicators*, 10 (2), 341–352.

Beale, C.M., Lennon, J.J., and Gimona, A., 2008. Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences*, 105, 14908–14912.

Beaumont, L.J. and Hughes, L., 2002. Potential changes in the distributions of latitudinally restricted Australian butterfly species in response to climate change. *Global Change Biology*, 8, 954–971.

Beaumont, L.J., Hughes, L., and Poulsen, M., 2005. Predicting species distributions: Use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecological Modelling*, 186, 250–269.

Beerling, D.J., Huntley, B., and Bailey, J.P., 1995. Climate and the distribution of Fallopia japonica: Use of an introduced species to test the predictive capacity of response surfaces. *Journal of Vegetation Science*, 6, 269–282.

Berry, P.M., Dawson, T.P., Harrison, P.A., and Pearson, R.G., 2002. Modelling potential impacts of climate change on the bioclimatic envelope of species in Britain and Ireland. *Global Ecology and Biogeography*, 11, 453–62.

Box, E.O., Crumpacker, D.W., and Hardin, E.D., 1993. A climatic model for location of plant species in Florida, USA. *Journal of Biogeography*, 20, 629–644.

Box, E.O., Crumpacker, D.W., and Hardin, E.D., 1999. Predicted effects of climatic change on distribution of ecologically important native tree and shrub species in Florida. *Climatic Change*, 41, 213–248.

Boylan, N., Jennings, R., and Long, M., 2008. Peat slope failure in Ireland. *Quarterly Journal of Engineering Geology and Hydrogeology*, 41, 93–108.

Brauner, N. and Shacham, M., 1998. Role of range and precision of the independent variable in regression of data. *American Institute of Chemical Engineering Journal*, 44, 603–611.

Brereton, R., Bennett, S., and Mansergh, I., 1995. Enhanced greenhouse climate change and its potential effect on selected fauna of south-eastern Australia: A trend analysis. *Biological Conservation*, 72, 339–354.

Brown, D.G., 1994. Predicting vegetation types at treeline using topography and biophysical disturbance variables. *Journal of Vegetation Science*, 5, 641–656.

Bruun, H.H., Moen, J., Virtanen, R., Grytnes, J.A., Oksanen, L., and Angerbjorn, A., 2006. Effects of altitude and topography on species richness of vascular plants, bryophytes and lichens in alpine communities. *Journal of Vegetation Science*, 17, 37–46.

Burns, C.E., Johnston, K.M., and Schmitz, O.J., 2003. Global climate change and mammalian species diversity in USA national parks. *Proceedings of the National Academy of Sciences*, 100, 11474–11477.

Byrne, C., Jones, M., Donnelly, A., and Wilson, J., 2003. Assessment of the impacts of climate change on biodiversity in Ireland. *In*: J. Sweeney, ed. *Climate change scenarios and impacts for Ireland*. Dublin: Environmental Protection Agency, 121–140.

Canadell, J.G., Le Quere, C., Raupach, M.R., Field, C.B., Buitenhaus, E.T., Ciais, P., Conway, T.J., Gillet, N.P., Houghton, R.A., and Marland, G., 2007. Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity and efficiency of natural sinks. *Proceedings of the National Academy of Sciences*, 104 (47), 18866–18870.

Carling, P.A., 1986. Peat slides in Teesdale and Weardale, Northern Pennines, July 1983: Description and failure mechanisms. *Earth Surface Processes and Landforms*, 11, 193–206.

42   *J. Coll* et al.

Carpenter, G., Gillison, A.N., and Winter, J., 1993. DOMAIN: A flexible modelling procedure for mapping potential distributions of plants and animals. *Biodiversity and Conservation*, 2, 667–680.

Carter, T.M., Hulme, M., and Lal, M., 1999. *Guidelines on the use of scenario data for climate impact and adaptation assessment: Version 1*. IPCC, Task Group on Scenarios for Climate Impact Assessment.

Cawsey, E.M., Austin, M.P., and Baker, B.L., 2002. Regional vegetation mapping in Australia: A case study in the practical use of statistical modelling. *Biodiversity and Conservation*, 11, 2239–2274.

Charman, D., 2002. *Peatlands and environmental change*. London: Wiley & Sons.

Convention on Biological Diversity (CBD), 2007. Emerging Issues for Biodiversity Conservation in a Changing Climate. *Abstracts of Poster Presentations at the 12th Meeting of the Subsidiary Body on Scientific*, Technical and Technological Advice of the Convention on Biological Diversity, 2–6 July 2007.Paris, Montreal: Secretariat of the Convention on Biological Diversity.

Coll, J., Gibb, S.W., and Harrison, J., 2005. Modelling future climates in the Scottish Highlands – An approach integrating local climatic variables and regional climate model outputs. *In*: D.B.A. Thompson, M.F. Price and C.A. Galbraith, eds. *Mountains of Northern Europe: Conservation, management, people and nature*. Edinburgh: The Stationery Office, 103–119.

Coll, J., Gibb, S.W., Price, M.F., McClatchey, J., and Harrison, J., 2010. Developing site scale projections of climate change in the Scottish Highlands. *Climate Research*, 45, 71–85.

Cowley, M.J.R., Wilson, R.J., Leon-Cortes, J.L., Gutiérrez, D., Bulman, C.R., and Thomas, C.D., 2000. Habitat-based statistical models for predicting the spatial distribution of butterflies and the day-flying moths in a fragmented landscape. *Journal of Applied Ecology*, 37 (1), 60–72.

Crawley, M.J., 1993. *GLIM for ecologists*. Oxford: Blackwell.

Crisp, D.T., Rawes, M., and Welch, D., 1964. A Pennine peat slide. *Geographical Journal*, 130, 519–524.

Crushell, P., 2000. *Irish fen inventory-A review of the status of fens in Ireland*. Dublin: IPCC.

Crumpacker, D.W., Box, E.O., and Hardin, E.D., 2001. Implications of climatic warming for conservation of native trees and shrubs in Florida. *Conservation Biology*, 15, 1008–1020.

Currie, D.J., 1991. Energy and large-scale patterns of animal and plant-species richness. *The American Naturalist*, 137, 27–49.

Czado, C., Erhardt, V., Min, A., and Wagner, S., 2007. Zero-inflated generalized poisson models with regression effects on the mean, dispersion and zero-inflation level applied to patent outsourcing rates. *Statistical Modelling*, 7 (2), 125–153.

De Veaux, R.D. and Unger, L.H., 1994. Multicollinearity: A tale of two nonparametric regressions. *In*: P. Cheeseman and R.W. Oldford, eds. *Selecting models from data: AI and Statistics IV*. New York: Springer-Verlag.

DOEHLG, 2007. *Ireland National Climate Change Strategy 2007-2012*. Dublin: Department of Environment, Heritage and Local Government.

Donald, P.F., Sanderson, F.J., Burfield, I.J., Bierman, S.M., Gregory, R.E., and Waliczky, Z., 2008. International conservation policy delivers benefits for birds in Europe. *Science*, 317, 810–813.

Donnelly, A., Caffarra, A., Albanito, F., and Jones, M., 2008. The impact of climate change on semi-natural ecosytems in Ireland (Chapter 5). *In*: J. Sweeney, *et al*., eds. *Climate change in Ireland: Refining the impacts*. Dublin: Environmental Protection Agency, 132–159.

Douglas, C., 1998. Blanket bog conservation *In*: G.F. O'Leary, ed., *Towards a conservation strategy for the bogs of Ireland*. Dublin: Irish Peatland Conservation Council, 205–222.

Dykes, A.P. and Kirk, K.J., 2001. Initiation of a multiple peat slide on Cuilcagh Mountain, Northern Ireland. *Earth Surface Processes and Landforms*, 26 (4), 395–408.

Dykes, A.P., Gunn, J., and Convery, K.J., 2008. Landslides in blanket peat on Cuilcagh Mountain, northwest Ireland. *Geomorphology*, 102 (3-4), 325–340.

Elith, J. and Burgman, M., 2002. Predictions and their validation: Rare plants in the Central Highlands, Victoria, Australia. *In*: J.M. Scott, *et al*., eds. *Predicting species occurrences. Issues of accuracy and scale*. Washington DC: Island Press, 303–313.

Elith, J., Graham, C.H., Anderson, R.P., Dudik, 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., Peterson, A.T., 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.

Elmore, K.L., Weiss, S.J., and Banacos, P.C., 2003. Operational ensemble cloud model forecasts. Some preliminary results. *Weather and Forecasting*, 18, 953–964.

European Commission, 2009. *Adapting to climate change: Towards a European framework for action*. Brussels: European Commission.

Evans, D., 2006. The habitats of the European Union Habitats Directive. *Biology and Environment*, 106B (3), 167–173.

Fealy, R. and Sweeney, J., 2007. Statistical downscaling of precipitation for a selection of sites in Ireland employing a generalised linear modelling approach. *International Journal of Climatology*, 27, 2083–2094.

Fielding, A.H. and Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24, 38–49.

Finley, J.P., 1884. Tornado predictions. *American Meteorological Journal*, 1, 85–88.

Fitzmaurice, G.M., Laird, N.M., and Ware, J.H., 2004. *Applied longitudinal analysis*. Hoboken, NJ: Wiley and Sons.

Foden, W., Midgley, G.F., Hughes, G.O., Bond, W.J., Thuiller, W., Hoffman, N.T., Kaleme, P., Underhill, L.G., Rebelo, A.G., and Hannah, L., 2007. A changing climate is eroding the geographical range of the Namib Desert tree aloe through population declines and dispersal lags. *Diversity and Distribution*, 13, 645–653.

Foss, P.J., O'Connell, C.A., and Crushell, P.H., 2001. Bogs and fens of Ireland. *Conservation Plan 2005*. Dublin: Irish Peatland Conservation Council.

Gaston, K., Jackson, S.F., Nagy, A., Cantu-Salazar, L., and Johnson, M., 2008. Protected areas in Europe: Principle and practice. *Annals of the New York Academy of Science*, 1134, 97–119.

Gates, S. and Donald, P.F., 2000. Local extinction of British farmland birds and the prediction of further loss. *Journal of Applied Ecology*, 37, 806–820.

Graham, M.H., 2003. Confronting multicollinearity in ecological multiple regression. *Ecology*, 84, 2809–2815.

Guisan, A., Edwards, T.C., and Hastie, T., 2002. Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecological Modelling*, 157, 89–100.

Guisan, A. and Theurillat, J.P., 2000. Assessing alpine plant vulnerability to climate change: a modeling perspective. *Integrated Assessment*, 1 (4), 307–20.

Guisan, A. and Thuiller, W., 2005. Predicting species distributions: Offering more than simple habitat models. *Ecology Letters*, 9, 993–1009.

Guisan, A. and Zimmerman, N.E., 2000. Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147–186.

H-Acevedo, D. and Currie, D.J., 2003. Does climate determine broad-scale patterns of species richness? A test of the causal link by natural experiment. *Global Ecology and Biogeography*, 12, 461–473.

Hammond, R.F., 1979. The peatlands of Ireland *In*:*Soil and Survey Bulletin*, 35. Dublin: An Foras Talúntais.

Hammond, R.F., 1984. The Classification of Irish peats as surveyed by the National Soil Survey of Ireland. *Proceedings of the 7th International Peat Congress*, Dublin.

Hanley, J.A. and McNeil, B.J., 1983. A method of comparing the areas under Receiver Operating Characteristic curves derived from the same cases. *Radiology*, 148, 839–843.

Harrell Jr, F.E., 2001. *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis*. New York: Springer.

Hauck, W.W. and Donner, A., 1977. Wald's Test as applied to hypotheses in logit analysis. *Journal of the American Statistical Association*, 72, 851–853.

44 *J. Coll* et al.

Heikkinen, R.K., Luoto, M., Kuussaari, M., and Poyry, J., 2004. Effects of habitat cover, landscape structure and spatial variables on the abundance of birds in an agricultural-forest mosaic. *Journal of Applied Ecology*, 41, 824–835.

Heikkinen, R.K., Luoto, M., Araújo, M.B., Virkalla, M., Thuiler, W., and Sykes, M.T., 2006. Methods and uncertainties in bioclimatic envelope modelling under climate change. *Progress in Physical Geography*, 30 (6), 751–777.

Hensher, D. and Johnson, L.W., 1981. *Applied discrete choice modelling*. London: Croom Helm.

Hill, J.K., Thomas, C.D., and Huntley, B., 1999. Climate and habitat availability determine 20th century changes in a butterfly's range margin. *Proceedings of the Royal Society of London Series B Biological Sciences*, 266, 1197–206.

Hill, J.K., Thomas, C.D., Fox, R., Telfer, M.G., Willis, S.G., Asher, J., and Huntley, B., 2002. Responses of butterflies to twentieth century climate warming: Implications for future ranges. *Proceedings of the Royal Society of London Series B Biological Sciences*, 269, 2163–71.

Hopkins, J.J., Allison, H.M., Walmsley, C.A., Gaywood, M., and Thurgate, G., 2007. *Conserving biodiversity in a changing climate: Guidance on building capacity to adapt*. London: Department for the Environment, Food and Rural Affairs.

Hosmer, D.W. and Lemeshaw, S., 1989. *Applied Logistic Regression*. New York: Wiley & Sons.

Huntley, B., Berry, P.M., Cramer, W., and McDonald, A.P., 1995. Modelling present and potential future ranges of some European higher plants using climate response surfaces. *Journal of Biogeography*, 22, 967–1001.

Huntley, B., Green, R.E., Collingham, Y.C., Hill, J.K., Willis, S.G., Bartlein, P.J., Cramer, W., Hagemeijer, W.J.M., and Thomas, C.D., 2004. The performance of models relating species geographical distributions to climate is independent of trophic level. *Ecology Letters*, 7, 417–426.

Iverson, L.R. and Prasad, A.M., 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecological Monographs*, 68, 465–485.

Iverson, L.R. and Prasad, A.M., 2001. Potential changes in tree species richness and forest community types following climate change. *Ecosystems*, 4, 186–199.

Iverson, L.R. and Prasad, A.M., 2002. Potential redistribution of tree species habitat under five climate change scenarios in the eastern US. *Forest Ecology and Management*, 155, 205–222.

JNCC, 2007. *Second report by the UK under Article 17 on the implementation of the Habitats Directive from January 2001 to December 2006*. Peterborough: Joint Nature Conservation Committee.

Jones, M.B., Donnelly, A., and Albanito, F., 2006. Responses of Irish vegetation to future climate change. *Biology and Environment: Proceedings of the Royal Irish Academy*, 106B (3), 323–334.

Kadmon, R., Farber, O., and Danin, A., 2003. A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, 13, 853–867.

Kleyer, M., Kratz, R., Lutze, G., and Schroder, B., 1999. Habitatmodelle fur Tierarten: Entwicklung, Methoden und Perspektiven fur die Anwendung. *Zeitschrift fur Okologie und Naturschutz*, 8, 177–182.

Kurbatova, J., Li, C., Tatarinov, F., Varlagin, A., Shalukhina, N., and Olchev, A., 2009. Modeling of the carbon dioxide fluxes in European Russia peat bogs. *Environmental Research Letters*, 4 (4), 1–5.

Lavergne, S., Thuiller, W., Molina, J., and Debussche, M., 2005. Environmental and human factors influencing rare plant local occurrence, extinctions and persistence: A 115 year study in the Mediterranean region. *Journal of Biogeography*, 32 (5), 799–811.

Lawler, J.J., White, D., Neilson, R.P., and Blaustein, A.R., 2006. Predicting climate-induced range shifts: Model differences and model reliability. *Global Change Biology*, 12, 1–17.

Leathwick, J.R., Whitehead, D., and McLeod, M., 1996. Predicting changes in the composition of New Zealand's indigenous forests in response to global warming: A modelling approach. *Environmental Software*, 11, 81–90.

Leps, J. and Smilauer, P., 2003. *Multivariate analysis of ecological data using CANOCO*. Cambridge: Cambridge University Press.

Lindsay, R.A., Charman, D.J., Everingham, F., O'Reilly, R.M., Palmer, M.A., Rowell, T.A., and Stroud, D.A., 1988. *The flow country: The peatlands of Caithness and Sutherland*. Peterborough: Nature Conservancy Council.

Liu, C., Berry, P.M., Dawson, T.P., and Pearson, R.G., 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28, 385–393.

Liu, C., White, M., and Newell, G., 2009. Measuring the accuracy of species distribution models: A review. *18th World IMACS/MODSIM Congress*, Cairns, Australia, 13–19 July, 4241–4247.

Lobo, J.M., Jimenez-Valverde, A., and Real, R., 2008. AUC: A misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, 17, 145–151.

Luoto, M., Toivonen, T., and Heikkinen, R.K., 2002. Prediction of total and rare plant species richness in agricultural landscapes from satellite images and topographic data. *Landscape Ecology*, 17, 195–217.

Luoto, M., Pöyry, J., Heikkinen, R.K., and Saarinen, K., 2005. Uncertainty of bioclimate envelope models based on geographical distribution of species. *Global Ecology and Biogeography*, 14, 575–84.

Luoto, M., Heikkinen, R.K., Poyry, J., and Saarinen, K., 2006. Determinants of biogeographical distribution of butterflies in boreal regions. *Journal of Biogeography*, 33 (10), 1764–1768.

MacNally, R., 2000. Regression and model-building in conservation biology, biogeography and ecology: The distinction between – and reconciliation of –'predictive' and explanatory models. *Biodiversity and Conservation*, 9, 655–671.

Manel, S., Dias, J-M., and Ormerod, S.J., 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: A case study with a Himalayan river bird. *Ecological Modelling*, 120, 337–347.

Marmion, M., Luoto, M., Heikkinen, R.K., and Thuiller, W., 2009. The performance of state-of-the-art modelling techniques depends on geographical distribution of species. *Ecological Modelling*, 220, 3512–3520.

~~McBride, J.L. and Ebert, E.E., 2000. Verification of quantitative precipitation forecasts from operational numerical weather prediction models over Australia. *Weather and Forecasting*, 15, 103–121.~~

McCullagh, P. and Nelder, J., 1989. *Generalized linear models*. London: Chapman and Hall.

McGrath, R. and Lynch, P., 2008. *Ireland in a warmer world: Scientific predictions of the Irish climate in the twenty-first century*. Dublin: Community Climate Change Consortium for Ireland (), C4I.

Menzel, A. and Fabian, P., 1999. Growing season extended in Europe. *Nature*, 397, 659.

Metz, C.E., 1986. ROC methodology in radiologic imaging. *Investigative Radiology*, 21, 720–733.

Meynecke, J.O., 2004. Effects of global climate change on geographic distributions of vertebrates in North Queensland. *Ecological Modelling*, 174, 347–57.

Mitchell, R.J., Morecroft, M.D., Acreman, M., Crick, H.Q.P., Frost, M., Harley, M., Maclean, I.M.D., Mountford, O., Piper, J., Pontier, H., Rehfisch, M.M., Ross, L.C., Smithers, R.J., Stott, A., Walmsley, C.A., Watts, O., and Wilson, E., 2007. *England Biodiversity Strategy-Towards adaptation to climate change*. London: Department for the Environment, Food and Rural Affairs.

Monserud, R.A. and Leemans, R., 1992. Comparing global vegetation maps with Kappa statistic. *Ecological Modelling*, 62, 275–293.

Moore, J.J., 1962. The Braun-Blanquet system: A reassessment. *Journal of Ecology*, 50, 761–769.

Munoz, J. and Felicismo, A.M., 2004. Comparison of statistical methods commonly used in predictive modelling. *Journal of Vegetation Science*, 15, 285–292.

Nagelkerke, N.J.D., 1991. A note on a general definition of the coefficient of determination. *Biometrica*, 78, 691–692.

Nichols, A.O., 1989. How to make biological surveys go further with generalised linear models. *Biological Conservation*, 50, 51–75.

46   *J. Coll* et al.

NPWS, 2008. *The status of EU protected habitats and species in Ireland*. Dublin: National Parks and Wildlife Service, Department of the Environment, Heritage and Local Government.

Ostermann, O.P., 1998. The need for management of Natura conservation designated under Natura 2000. *Journal of Applied Ecology*, 35, 968–973.

Parmesan, C., 2006. Influences of species, latitudes and methodologies on estimates of phenological response to global warming. *Global Change Biology*, 13, 1860–1872.

Parmesan, C. and Yohe, G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421, 37–42.

Parry, M., Lowe, J., and Hanson, C., 2009. Overshoot, adapt and recover. *Nature*, 458 (30), 1102–1103.

Pauli, H., Gottfried, M., Reiter, K., Klettner, C., and Grabherr, G., 2007. Signals of range expansions and contractions of vascular plants in the high Alps: Observations (1994-2004) at the GLORIA master site Schrankogel, Tyro, Austria. *Global Change Biology*, 13, 147–156.

Pausas, J.G., Carreras, J., Ferre, A., and Font, X., 2003. Coarse-scale plant species richness in relation to environmental heterogeneity. *Journal of Vegetation Science*, 14, 661–668.

Pearce, J. and Ferrier, S., 2000a. An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecological Modelling*, 128, 127–147.

Pearce, J. and Ferrier, S., 2000b. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133, 225–245.

Pearson, R.G., Dawson, T.P., Berry, P.M., and Harrison, P.A., 2002. SPECIES: A spatial evaluation of climate impact on the envelope of species. *Ecological Modelling*, 154, 289–300.

Pearson, R.G., Dawson, T.P., and Liu, C., 2004. Modelling species distributions in Britain: A hierarchical integration of climate and land-cover data. *Ecography*, 27, 285–298.

Peterson, A.T., 2001. Predicting species' geographic distributions based on ecological niche modelling. *The Condor*, 103, 599–605.

Peterson, A.T., Ball, L.G., and Cohoon, K.P., 2002. Predicting distributions of Mexican birds using ecological niche modelling methods. *Ibis*, 144, 27–32.

Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sánchez-Cordero, V., Soberón, J., Buddemeier, R.H., and Stockwell, D.R.B., 2002. Future projections for Mexican faunas under global climate change scenarios. *Nature*, 416, 626–29.

Peterson, A.T., Martínez-Meyer, E., González-Salazar, C., and Hall, P.W., 2004. Modelled climate change effects on distributions of Canadian butterfly species. *Canadian Journal of Zoology*, 82, 851–58.

Pounds, A.J., Bustamante, M.R., Coloma, L.A., Consuegra, J.A., Fogden, M.P.L., Foster, P.N., La Marca, E., Masters, K.L., Merino-Viteri, A., Puschendorf, R., Ron, S.R., Sanchez-Azofeifa, G.A., Still, C.J., and Young, B.E., 2006. Widespread amphibian extinctions from epidemic disease driven by global warming. *Nature*, 439, 161–167.

Prasad, A.M., and Iverson, L.R., 2000. Predictive vegetation mapping using a custom built model-chooser: comparison of regression tree analysis and multivariate adaptive regression splines. CD-ROM *In*: *4th International Conference on Integrating GIS and Environmental Modeling (GIS/EM4): Problems, Prospects and Research Needs*, Banff, Alberta, Canada, September 2–8, 2000.

Price, J., 2000. Modeling the potential impacts of climate change on the summer distributions of Massachusetts passerines. *Bird Observer*, 28, 224–30.

Rangel, T.F.L.V.B., Diniz-Filho, J.A.F., and Bini, L.M., 2006. Towards an integrated computational tool for spatial analysis in macroecology and biogeography. *Global Ecology and Biogeography*, 15, 321–327.

Raupach, M.R., Marland, G., Ciais, P., Le Quere, C., Canadell, J.G., Klepper, G., and Field, C.B., 2007. Global and regional drivers of accelerating CO2 emissions. *Proceedings of the National Academy of Sciences*, 104 (24), 10288–10293.

Real, R., Barbosa, A.M., and Vargas, J.M., 2006. Obtaining environmental favourability functions from logistic regression. *Environmental and Ecological Statistics*, 13, 237–245.

Reineking, B. and Schroder, B., 2006. Constrain to perform: Regularization of habitat models. *Ecological Modelling*, 193, 675–690.

Root, T.L., MacMynowski, D.P., Mastrandrea, M.D., and Schneider, S.H., 2005. Human-modified temperatures induce species changes: Joint attribution. *Proceedings of the National Academy of Science*, 102, 7465–7469.

Saseendran, S.A., Singh, S.V., Rathore, L.S., and Das, S., 2002. Characterization of weekly cumulative rainfall forecasts over meteorological subdivisions of India using a GCM. *Weather and Forecasting*, 17, 832–844.

Schouten, M.G.C., 1984. Some aspects of the ecogeographical gradient in the Irish ombrotrophic bogs *In*: *Proceedings of the 7th International Peat Congress*, The Irish National Peat Committee, 18–23 June, 1984, Dublin, Ireland.

Sheehy Skeffington, M.J. and O'Connell, C., 1998. Peatlands of Ireland. *In*: P. Giller, ed. *Studies in Irish Limnology*. Dublin: Societas Internationalis Limnologiae, 39–66.

Skov, F. and Svenning, J.C., 2004. Potential impact of climatic change on the distribution of forest herbs in Europe. *Ecography*, 27, 366–380.

Sokal, R.R. and Rohlf, F.J., 1981. *Biometry*. 2nd ed. New York: W.H. Freeman.

Sottocornola, M., Laine, A., Kiely, G., Byrne, K.A., and Tuitilla, E.S., 2009. Vegetation and environmental variation in an Atlantic blanket bog in South-western Ireland. *Plant Ecology*, 203 (1), 69–81.

Steyerberg, E.W., Harrell, F.E.J., Borsboom, G.J.J.M., Eijkemans, M.J.C., Vergouwe, Y., and Habbema, J.D.F., 2001. Internal Validation of predictive models-efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology*, 54, 774–781.

Suarez-Seoane, S., Osborne, P.E., and Alonso, J.C., 2002. Large-scale habitat selection by agricultural steppe birds in Spain: Identifying species-habitat responses using generalized additive models. *Journal of Applied Ecology*, 39, 755–771.

Svenning, J.C. and Skov, F., 2004. Limited filling of the potential range in European tree species. *Ecology Letters*, 7, 565–73.

Swart, R., Biesbroek, R., Binnerup, S., Carter, T.R., Cowan, C., Henrichs, T., Loquen, S., Mela, H., Morecroft, M., Reese, M., and Rey, D., 2009. *Europe adapts to climate change: Comparing national adaptation strategies*. Helsinki: Partnership for European Environmental Research.

Sweeney, J. and Fealy, R., 2002. A preliminary investigation of future climate scenarios for Ireland. *Biology and Environment: Proceedings of the Royal Irish Academy*, 102B (3), 121–128.

Sweeney, J. and Fealy, R., 2003. Establishing reference climate scenarios. *In*: J. Sweeney, ed. *Climate change scenarios and impacts for Ireland*. Dublin: Environmental Protection Agency.

Sweeney, K., Fealy, R., McElwain, L., Siggins, L., Sweeney, J., and Trinies, T., 2008. *Changing shades of green: The environmental and cultural impacts of climate change in Ireland*. California: The Irish American Climate Project & Rockefeller Family Fund.

Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. *Science*, 240, 1285–1293.

Sykes, M.T., Prentice, I.C., and Cramer, W., 1996. A bioclimatic model for the potential distributions of north European tree species under present and future climates. *Journal of Biogeography*, 23, 203–233.

Thuiller, W., 2003. BIOMOD: Optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology*, 9, 1353–1362.

Thuiller, W., Araújo, M.B., and Lavorel, S., 2004. Do we need land-cover data to predict species distributions in Europe?. *Journal of Biogeography*, 31, 353–361.

Thuiller, W., Albert, C., Araújo, M.B., Berry, P.M., Cabeza, M., Guisan, A., Hickler, T., Midgley, G.F., Paterson, J., Schurr, F.M., Sykes, M.T., and Zimmerman, N.E., 2008. Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology*. *Evolution and Systematics*, 9, 137–152.

Tomlinson, R.W., 2005. Soil carbon stocks and changes in the Republic of Ireland. *Journal of Environmental Management*, 76, 87–106.

Visser, M.E. and Holleman, L.J.M., 2001. Warmer springs disrupt the synchrony of oak and winter moth phenology. *Proceedings of the Royal Society London Series B*, 268, 289–294.

Walther, G.R., Beissner, S., and Burga, C.A., 2005. Trends in the upward shift of alpine plants. *Journal of Vegetation Science*, 16 (5), 541–548.

48   *J. Coll* et al.

AQ1,9 Walther, G.R., Berger, S., and Sykes, M.T., 2005. An ecological 'footprint' of climate change. *Proceedings of the Royal Society of London Series B*, 272, 1427–1432.

Warburton, J., Holden, J., and Mills, A. J., 2004. Hydrological controls of surficial mass movements in peat. *Earth-Science Reviews*, 67 (1-2), 139–156.

White, M.A., Brunsell, N., and Schwartz, M.D., 2003. Vegetation phenology in global change studies. *In*: M.D.E. Schwartz, ed. *Phenology: An integrative environmental science.* Dordrecht: Kluwer Academic Publishers, 453–466.

Wright, D.H., Currie, D.J., and Maurer, B.A., 1993. Energy supply and patterns of species richness on local and regional scales. *In*: R.E. Ricklefs and D. Schluter, eds. *Species diversity in ecological communities: Historical and geographical perspectives.* Chicago: University of Chicago Press.

Zavaleta, E.S., Shaw, M.R., Chiariello, N.R., Thomas, B.D., Cleland, E.E., Field, C.B., and Mooney, H.A., 2003. Grassland responses to three years of elevated temperature, CO2, **AQ9** precipitation and N deposition. *Ecological Monographs*, 73, 585–604.

Zuur, A.F., Ieno, E.N., and Smith, G.M., 2007. *Analysing ecological data.* New York: Springer.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., and Smith, G.M., 2009. *Mixed effects models and extensions in ecology with R.* New York: Springer.

APPENDIX
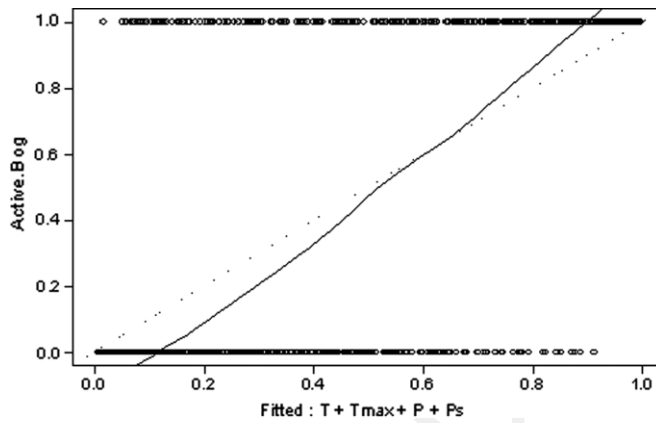
Figure A1. Observed and fitted probability values for the Active Blanket Bog (code 7130) habitat. Presence-absence data (0s and 1s denoted as circles). Hypothetical fitted probability (dotted line) and fitted initial logistic regression model (solid line).
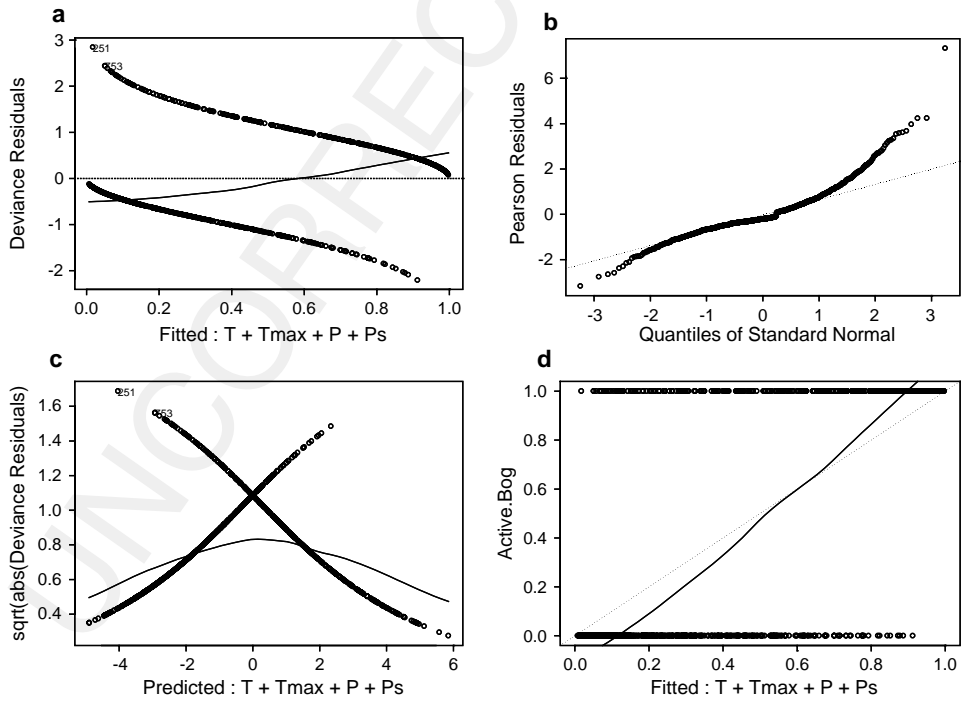
Figure A2. S+ validation graphs for blanket bog presence-absence data obtained by the initial logistic regression model: a) residuals versus fitted; b) normal Q-Q plot; c) scale-location plot; d) observed and fitted probability.
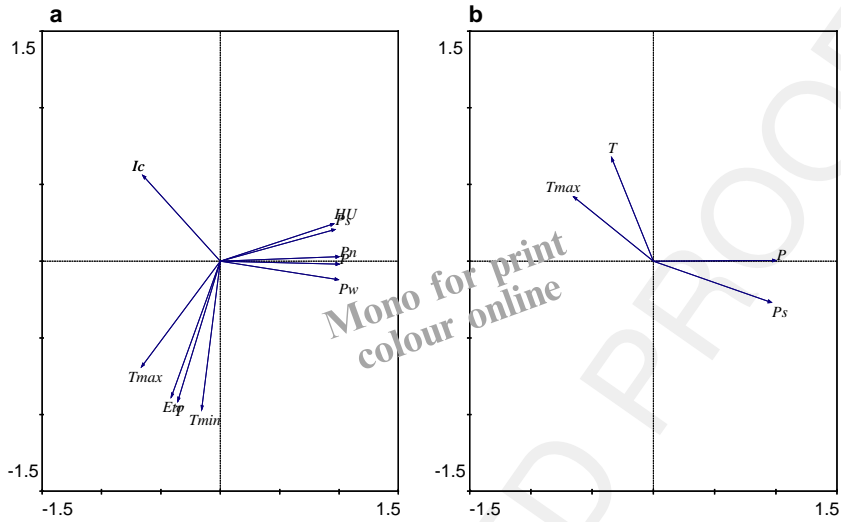
Figure A3. Principal components analysis (PCA) of candidate climate predictor variables (n = 841).
Note: a) Full set of available candidate predictors for 10km grid cells (left). b) candidate predictor variables identified in the initial GLM. See Table 1 for explanation of abbreviations
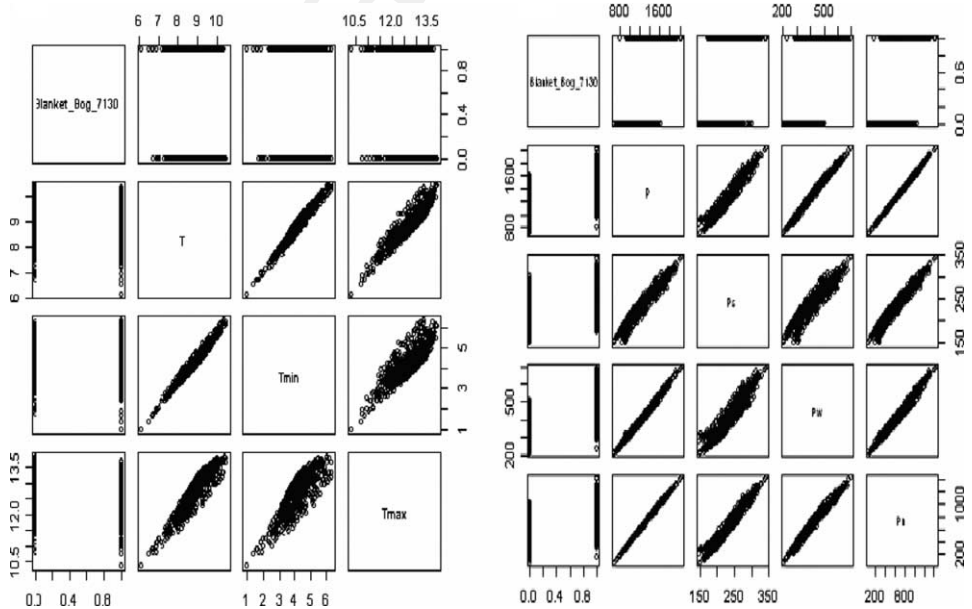
Figure A4. Pairplots: a) temperature and; b) precipitation variables.
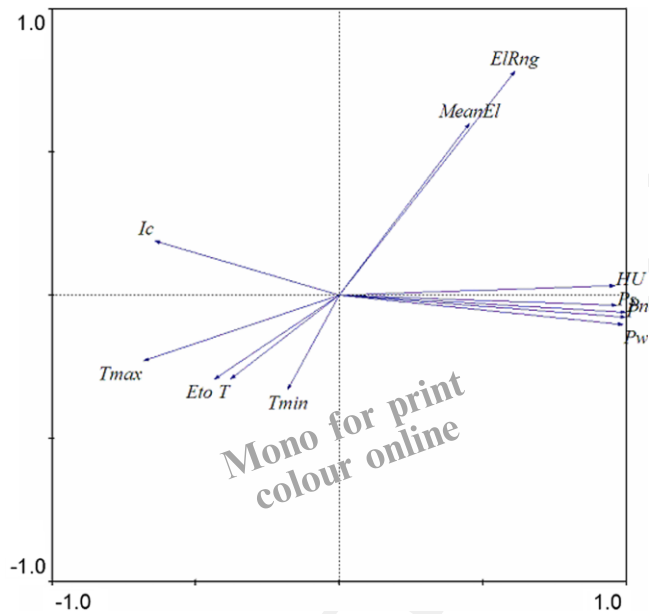
Figure A5. PCA of combined climate and elevation variables indicating collinearity of the elevation variables. See Table 1 for explanation of abbreviations.

Table A1. Pearson correlation coefficients for the ten candidate predictor variables.

		Correlations									
		T	Tmin	Tmax	P	Ps	Pw	Ic	Eto	HU	Pn
T	Pearson Correlation Sig. (2-tailed) N	1	.978**	.914**	−.339**	−.511**	−.249**	−.266**	.996**	−.580**	−.385**
			.000	.000	.000	.000	.000	.000	.000	.000	.000
		848	848	848	848	848	848	848	848	848	848
Tmin	Pearson Correlation Sig. (2-tailed) N	.978**	1	.816**	−.138**	−.331**	−.043	−.439**	.961**	−.404**	−.186**
			.000	.000	.000	.000	.206	.000	.000	.000	.000
		848	848	848	848	848	848	848	848	848	848
Tmax	Pearson Correlation Sig. (2-tailed) N	.914**	.816**	1	−.648**	−.738**	−.577**	.134**	.942**	−.816**	−.683**
			.000	.000	.000	.000	.000	.000	.000	.000	.000
		848	848	848	848	848	848	848	848	848	848
P	Pearson Correlation Sig. (2-tailed) N	−.339**	−.138**	−.648**	1	.962**	.994**	−.660**	−.395**	.950**	.999**
			.000	.000	.000	.000	.000	.000	.000	.000	.000
		848	848	848	848	848	848	848	848	848	848
Ps	Pearson Correlation Sig. (2-tailed) N	−.511**	−.331**	−.738**	.962**	1	.932**	−.454**	−.553**	.963**	.970**
			.000	.000	.000	.000	.000	.000	.000	.000	.000
		848	848	848	848	848	848	848	848	848	848
Pw	Pearson Correlation Sig. (2-tailed) N	−.249**	−.043	−.577**	.994**	.932**	1	−.705**	−.306**	.921**	.989**
			.000	.206	.000	.000	.000	.000	.000	.000	.000
		848	848	848	848	848	848	848	848	848	848
Ic	Pearson Correlation Sig. (2-tailed) N	−.266**	−.439**	.134**	−.660**	−.454**	−.705**	1	−.194**	−.470**	−.635**
			.000	.000	.000	.000	.000	.000	.000	.000	.000

Table A1 (*Continued*)

		Correlations									
		T	Tmin	Tmax	P	Ps	Pw	Ic	Eto	HU	Pn
Eto	Pearson Correlation Sig. (2-tailed) N	848 .996**	848 .961**	848 .942**	848 −.395**	848 −.553**	848 −.306**	848 −.194**	848 1	848 −.629**	848 −.439**
			.000	.000	.000	.000	.000	.000	.000	.000	.000
HU	Pearson Correlation Sig. (2-tailed) N	848 −.580**	848 −.404**	848 −.816**	848 .950**	848 .963**	848 .921**	848 −.470**	848 −.629**	848 1	848 .962**
			.000	.000	.000	.000	.000	.000	.000	.000	.000
Pn	Pearson Correlation Sig. (2-tailed) N	848 −.385**	848 −.186**	848 −.683**	848 .999**	848 .970**	848 .989**	848 −.635**	848 −.439**	848 .962**	848 1
			.000	.000	.000	.000	.000	.000	.000	.000	.000
			848	848	848	848	848	848	848	848	848

Note: Abbreviations as in Table 1

54   *J. Coll* et al.

Table A2. S+ and Brodgar (R) model fit values, climate variables.

Variable	Regression coefficient	Std. Error	*t*-value	*p*-value
(Intercept)	− 100.73	12.67	− 7.95	<0.001
T	− 15.35	2.07	− 7.42	<0.001
Tmax	17.69	2.36	− 7.49	<0.001
P	0.05	0.01	− 8.38	<0.001
Ps	− 0.24	0.03	− 7.06	<0.001

Note: Null Deviance: 1143.66 on 847 degrees of freedom. Residual Deviance: 654.6251 on 843 degrees of freedom. AIC: 664.63

Table A3. SPSS and SAM model performance indicators, climate variables.

SPSS		SAM	
Measure	Score	Measure	Score
Chi-Square	489.04	Chi-Square	489.04
Nagelkerke R_{2N}	0.592	Mc Fadden's Rho^2	0.428
PA % predicted	86.8 (0s)	Kappa	0.534
	74.0 (1s)	TSS	0.497

Table A4. S+ and Brodgar (R) model fit values, climate and elevation variables.

Variable	Regression coefficient	Std. Error	*t*-value	*p*-value
(Intercept)	− 8.28	2.24	− 3.70	<0.001
T	− 0.44	0.22	− 2.00	0.045
P	0.01	0.00	10.82	<0.001
ElRng	0.01	0.00	10.62	<0.001

Note: Null Deviance: 1165.82 on 840 degrees of freedom. Residual Deviance: 469.67 on 837 degrees of freedom. AIC: 477.67

Table A5. SPSS and SAM model performance indicators, climate and elevation variables.

SPSS		SAM	
Measure	Score	Measure	Score
Chi-Square	696.15	Chi-Square	696.15
Nagelkerke R_{2N}	0.751	Mc Fadden's Rho^2	0.599
PA % predicted	89.6 (0s)	Kappa	0.755
	86.1 (1s)	TSS	0.755