The following supplement accompanies the article

Impacts of pollution and climate change on ombrotrophic *Sphagnum* species in the UK: analysis of uncertainties in two empirical niche models

S. M. Smart[1,*], P. A. Henrys[1], W. A. Scott[1], J. R. Hall[2], C. D. Evans[2], A. Crowe[1], E. C. Rowe[2], U. Dragosits[3], T. Page[4], J. D. Whyatt[4], A. Sowerby[2], J. M. Clark[5,6,7]

[1]NERC Centre for Ecology and Hydrology, Library Avenue, Bailrigg, Lancaster LA1 4AP, UK
[2]NERC Centre for Ecology and Hydrology, Environment Centre Wales, Deiniol Road, Bangor, Gwynedd, LL57 2UW, UK
[3]NERC Centre for Ecology and Hydrology, Bush estate, Penicuik, Midlothian EH26 0QB, UK
[4]Lancaster Environment Centre, Lancaster University, Bailrigg, Lancaster LA1 4AP, UK
[5]Wolfson Carbon Capture Laboratory, School of Biological Sciences, Bangor University, Deiniol Road, Bangor, Gwynedd LL57 2UW, UK
[6]Grantham Institute for Climate Change, Department of Civil & Environmental Engineering, Imperial College London, South Kensington Campus, London, SW7 2AZ, UK
[7]*Present address:* Walker Institute for Climate Systems Research and Soils Research Centre, Geography and Environmental Science, School of Human and Environmental Sciences, University of Reading, Whiteknights Reading RG6 6DW, UK

*Email: ssma@ceh.ac.uk

Climate Research 45: 163–177 (2010)

1. Model testing

Dataset. Independent test data were selected from the 2007 Countryside Survey. Vegetation plots were selected from locations not sampled in the 1998 survey and therefore not spatially confounded with the survey data used to build the statistical models. In all other respects, selection criteria were the same as those applied to the training data. Hence, plots were only excluded where they sampled urban habitats, boundaries and linear features, improved grassland, inland rock, maritime habitats and arable land. We selected 4 m^2 plots (n = 97), of which 15 had observed occurrences of ombrotrophic *Sphagnum* species.

Model tests. Both the generalised additive mixed model (GAMM) and the generalised linear mixed model (GLMM) were solved using values of model covariates applying to each plot location and using the best parameter estimates from each model plus at the upper and lower 95% confidence intervals on the parameters.

RESULTS

GLMM. Predictions were highly uncertain given the large 95% confidence intervals on model parameters (see Table 2 in main text). In addition, several outliers were observed (Fig. S1). All outliers coincided with conditions that were predicted to be highly favourable for ombrotrophic *Sphagnum* cover with very high mean monthly rainfall and high soil C:N ratio. Three of these outlying plots were occupied. *Sphagnum* was absent from the rest despite high substrate C:N and high rainfall.

The outlying values partly reflect the non-linear scaling resulting from back-transformation of the logit cover values. However, with or without the outliers, the plots where *Sphagnum* was observed had higher predicted cover than where *Sphagnum* was absent (Fig. S2)

GAMM. Results were very similar to the GLMM test with greater predicted cover for plots in which ombrotrophic *Sphagnum* cover was present in the 2007 survey. Predictions were more variable, consistent with the greater sensitivity of the model to variation in the explanatory variables included (see Table 5 in main text).

Fig. S1. Testing the GLMM. Predicted percentage cover^0.5, is shown for plots in which *Sphagnum* was present or absent in an independent dataset of random, stratified vegetation plots recorded in the British Countryside Survey in 2007. The horizontal mid-line indicates the median prediction. Upper and lower box edges indicate the interquartile range. Whiskers indicate the non-outlier range

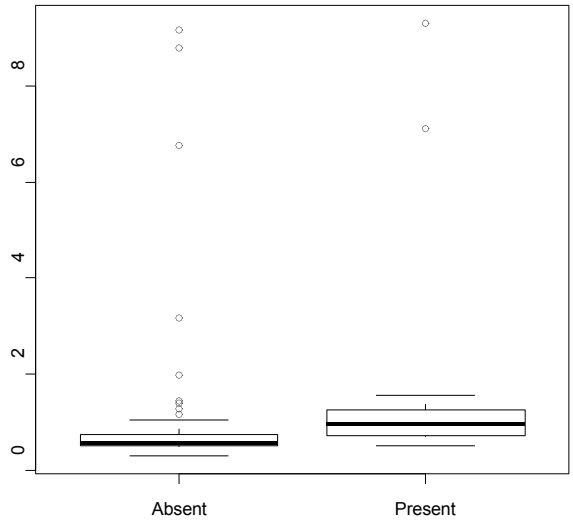


Fig. S2. Results of testing the GLMM. Predicted percentage cover^0.5 is shown for plots in which *Sphagnum* was present or absent in an independent dataset of random, stratified vegetation plots recorded in the British Countryside Survey in 2007. Outlying predicted values are excluded. The horizontal mid-line indicates the median prediction. Upper and lower box edges indicate the interquartile range. Whiskers indicate the non-outlier range

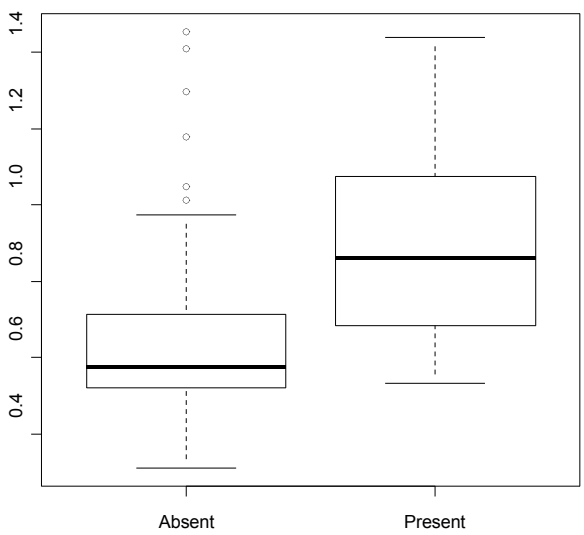


Fig. S3. Results of testing the GAMM. Predicted percentage cover^0.5 is shown for plots in which *Sphagnum* was present or absent in an independent dataset of random, stratified vegetation plots recorded in the British Countryside Survey in 2007. The horizontal mid-line indicates the median prediction. Upper and lower box edges indicate the interquartile range. Whiskers indicate the non-outlier range

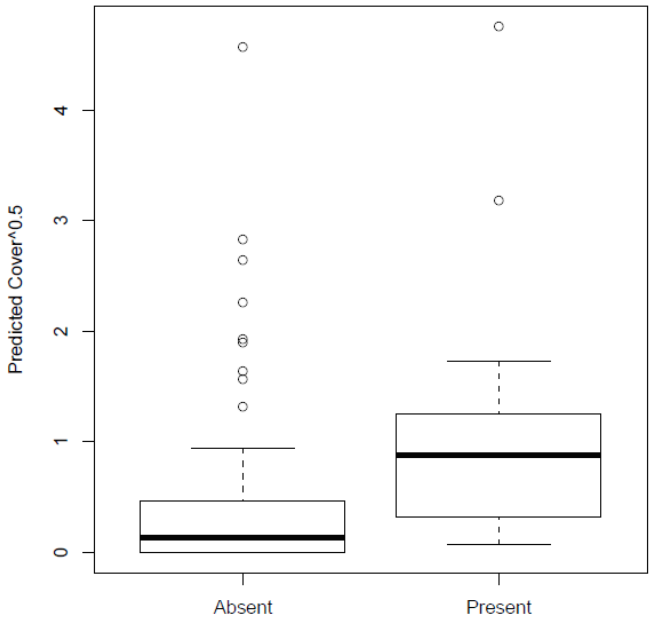
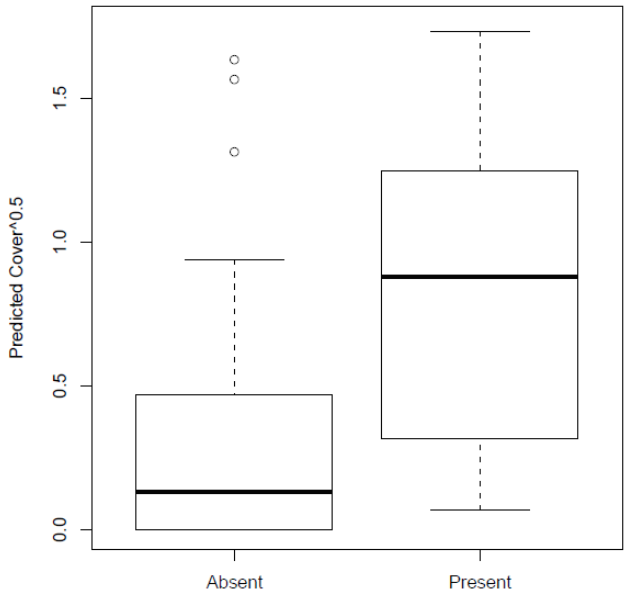
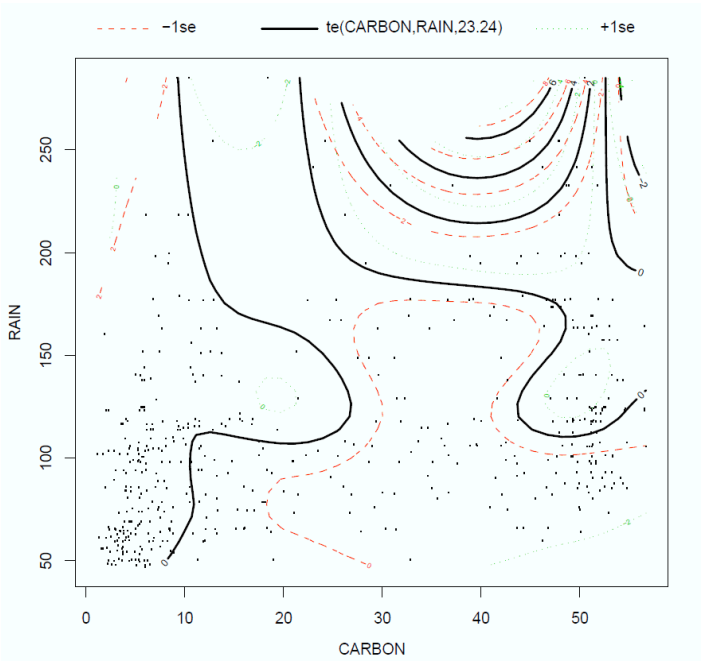


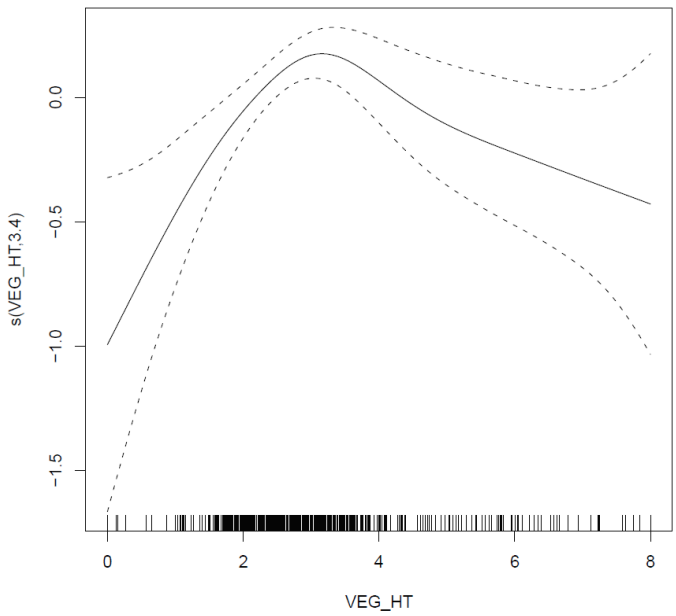
Fig. S4. Results of testing the GAMM. Predicted percentage cover^0.5 is shown for plots in which *Sphagnum* was present or absent in an independent dataset of random, stratified vegetation plots recorded in the British Countryside Survey in 2007. Outlying predicted values are excluded. The horizontal mid-line indicates the median prediction. Upper and lower box edges indicate the interquartile range. Whiskers indicate the non-outlier range

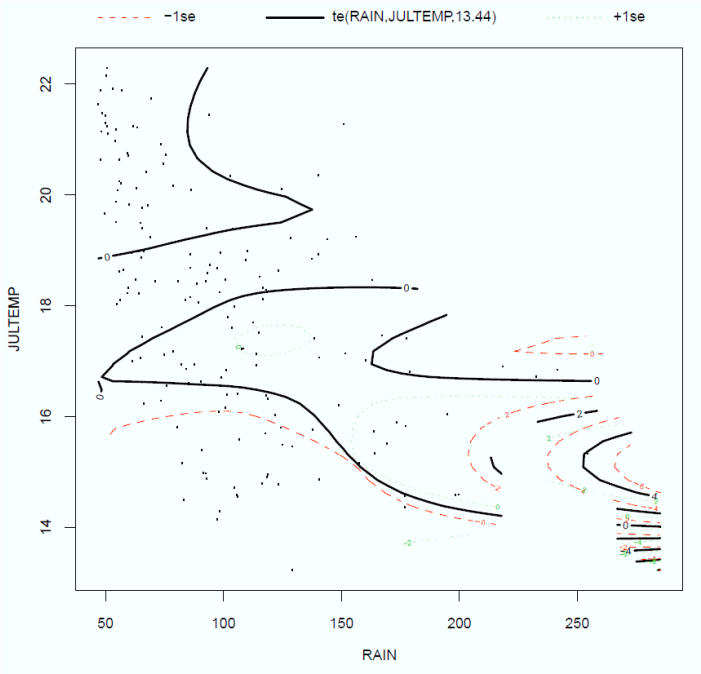


2. Additional diagnostic plots of modelled covariate relationships

Fig. S5. Marginal effect plots for the GAMM. The 2 interaction effects are shown on the left where contours demark the 2-dimensional smooth surface fitted by the model and the black dots are the observed data points. Marginal effects plots for % substrate carbon and cover-weighted vegetation height are shown on the right. Density of the observed data is indicated by vertical tick marks along the *x*-axes. The relative contribution of each term to the predicted cover of ombrotrophic *Sphagnum* is shown on the *y*-axes. RAIN = mean monthly rainfall (mm) (long term average), JULTEMP = Long term annual average maximum July temperature (°C), carbon = % organic carbon (top 15 cm). See Section 2 in main paper for further details.






−1se
te(RAIN,JULTEMP,13.44)
+1se
JULTEMP
RAIN

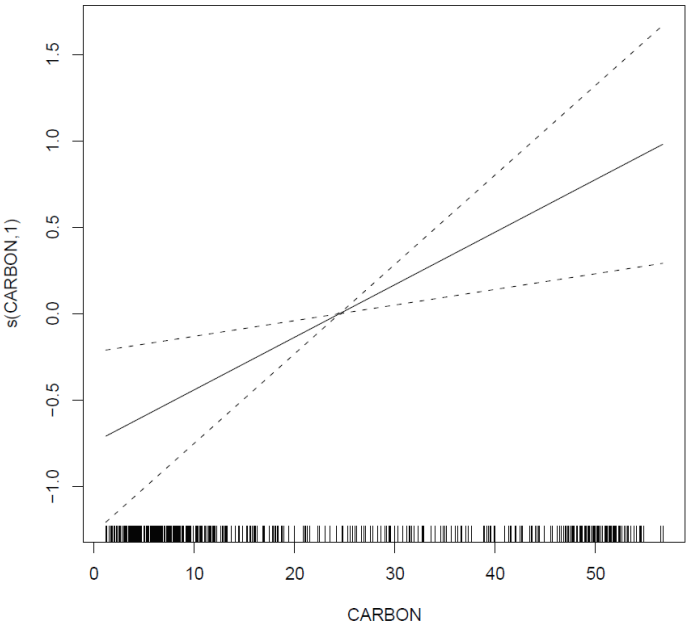

s(CARBON,1)
CARBON

Fig. S6. Marginal effect plots of model explanatory variables and observed data for the GLMM. The *y*-axis shows the residuals between observed cover and cover predicted by the fitted model with the term corresponding to the variable on the *x*-axis removed. The solid blue line is the fitted linear term in the model, corresponding to the parameters given in Table 2 in the main text

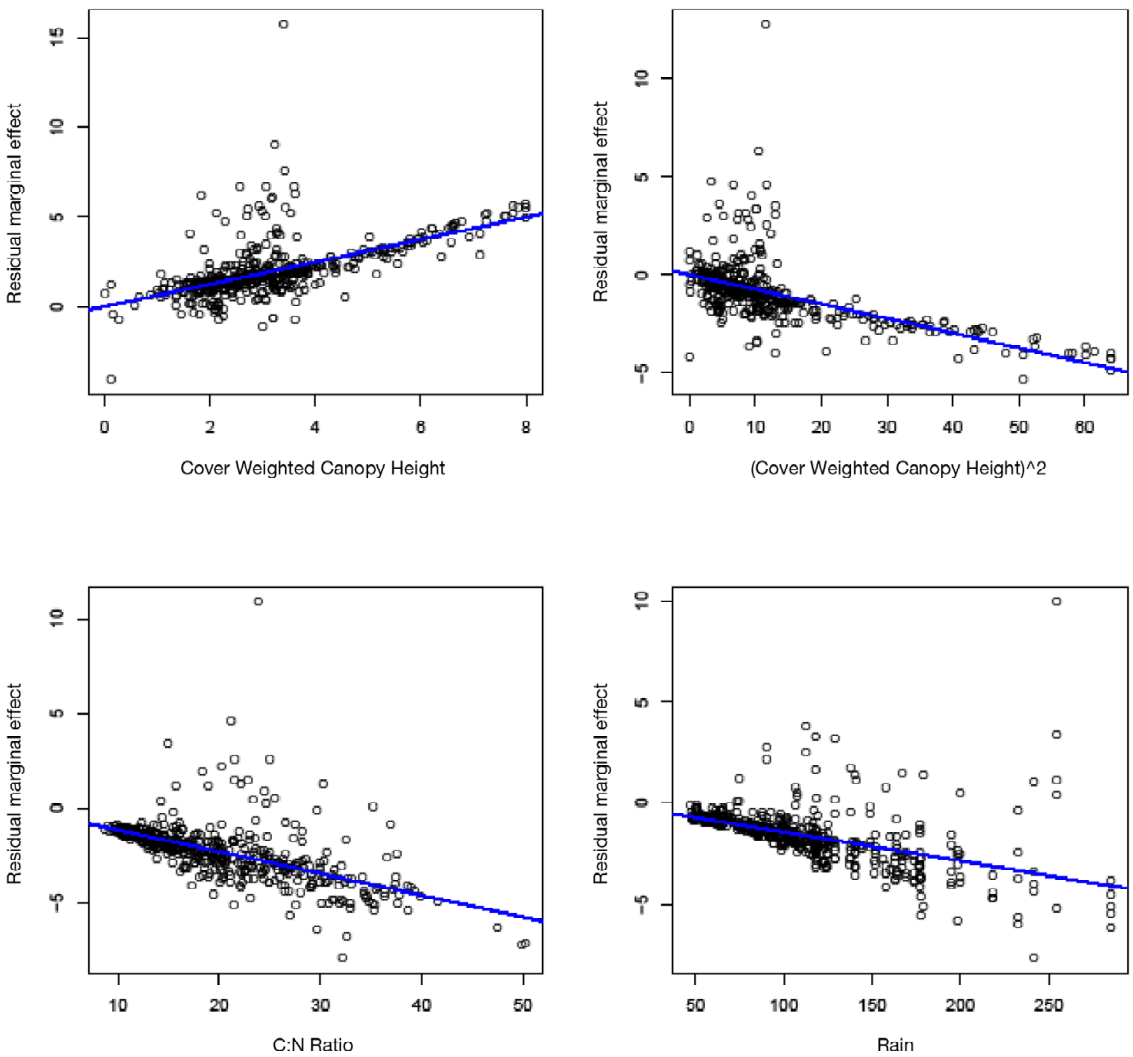
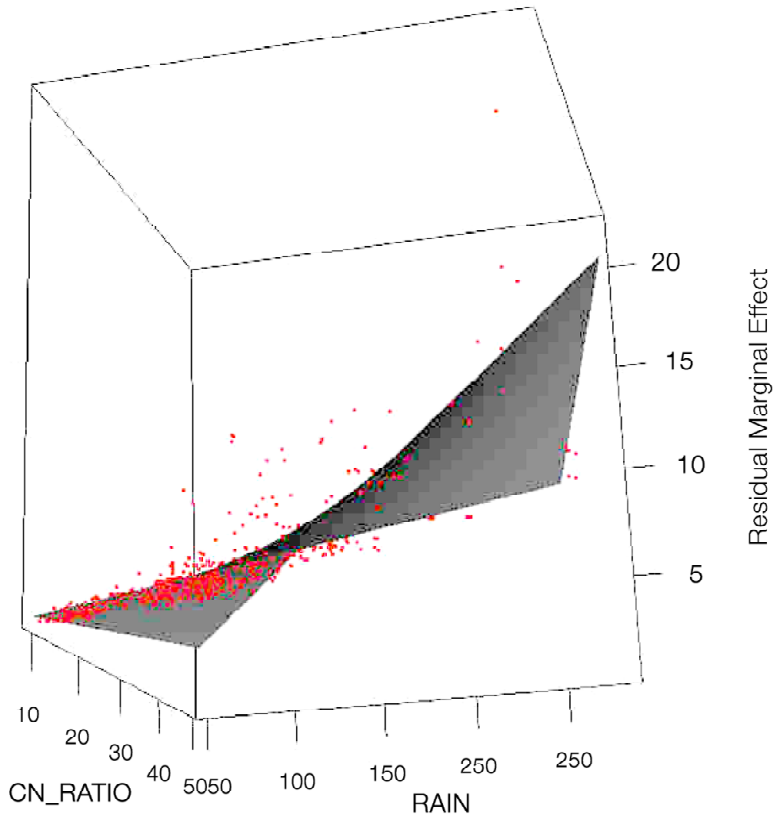


Fig. S7. Residual marginal effect of the interaction between effect of substrate C:N ratio (4 m^2 sample plots) and mean monthly rainfall (mm) in the GLMM model. The black surface shows the model fit of the interaction term, hence the contribution to the overall model that the interaction between C:N ratio and mean monthly rainfall has. The red dots are the observed data. The interaction manifests itself as a much steeper increase in predicted percentage ombrotrophic *Sphagnum* cover at high rainfall and high C:N than at lower values of these explanatory variables



3. Additional maps

Fig. S8. Very simple dynamic (VSD) model prediction of substrate C:N change at the 1 × 1 km scale across Great Britain between 2020 and 2050. The biggest reductions are in areas of highest current N deposition and highest historical reduction in S deposition

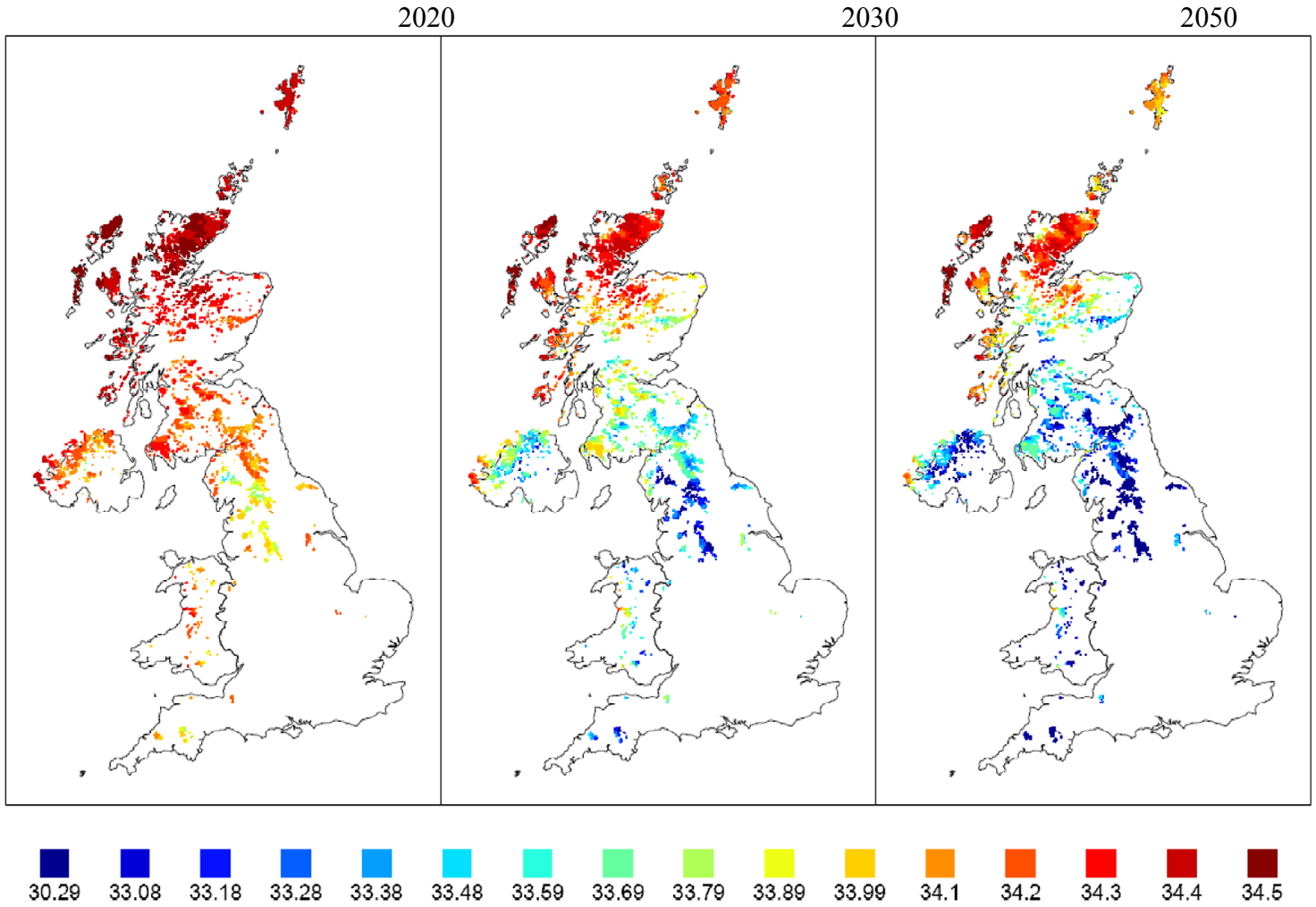


Fig. S9. Predicted July temperature in °C at the 50th percentile for 2020, 2030 and 2050 based on UKCP09 projections at the 25 × 25 km scale. Data obtained from © http://ukclimateprojections.defra.gov.uk/

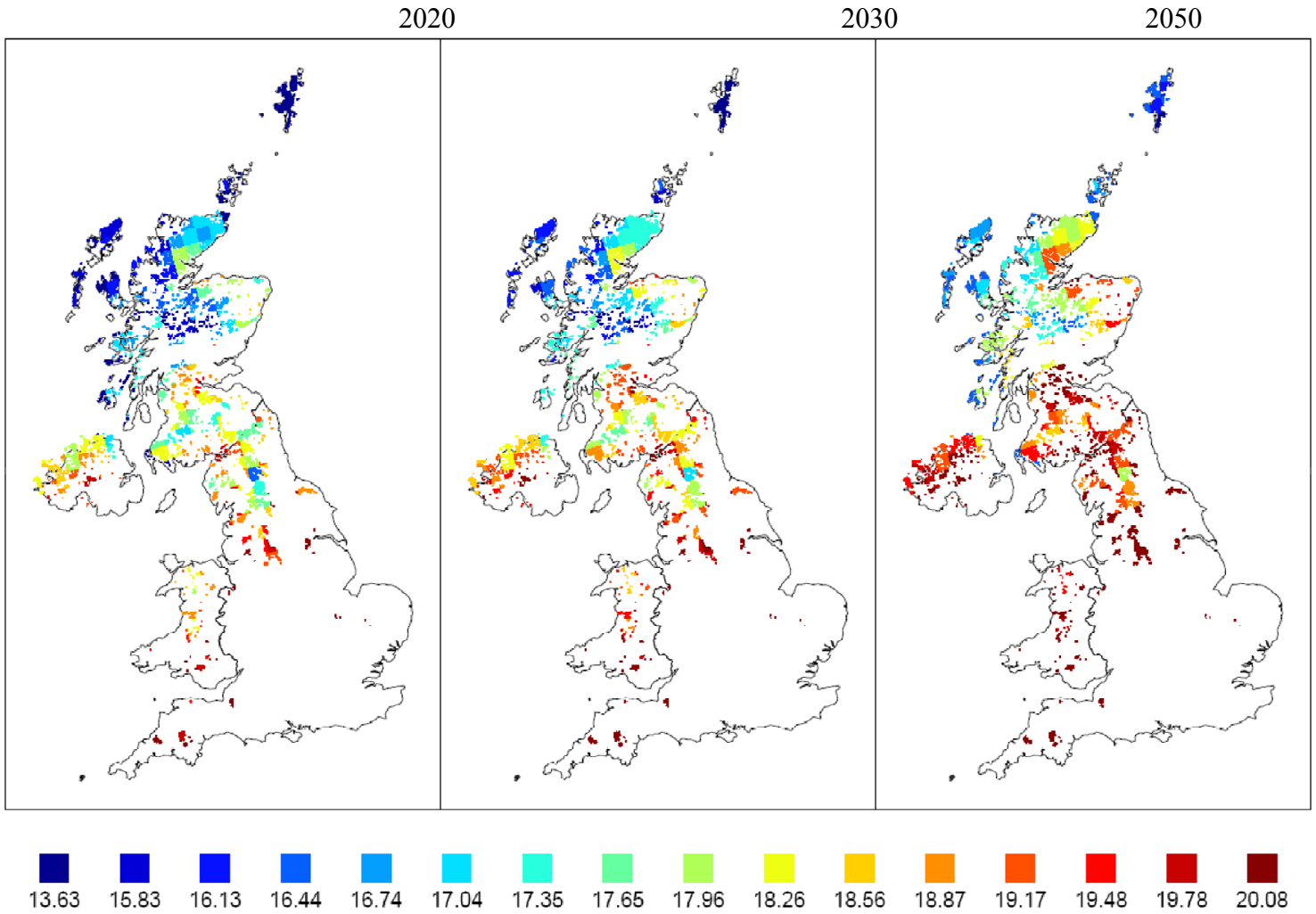


Fig. S10. Predicted mean monthly rainfall (mm) at the 50th percentile for 2020, 2030 and 2050 based on UKCP09 projections at the 25 × 25 km scale. Data obtained from © http://ukclimateprojections.defra.gov.uk/

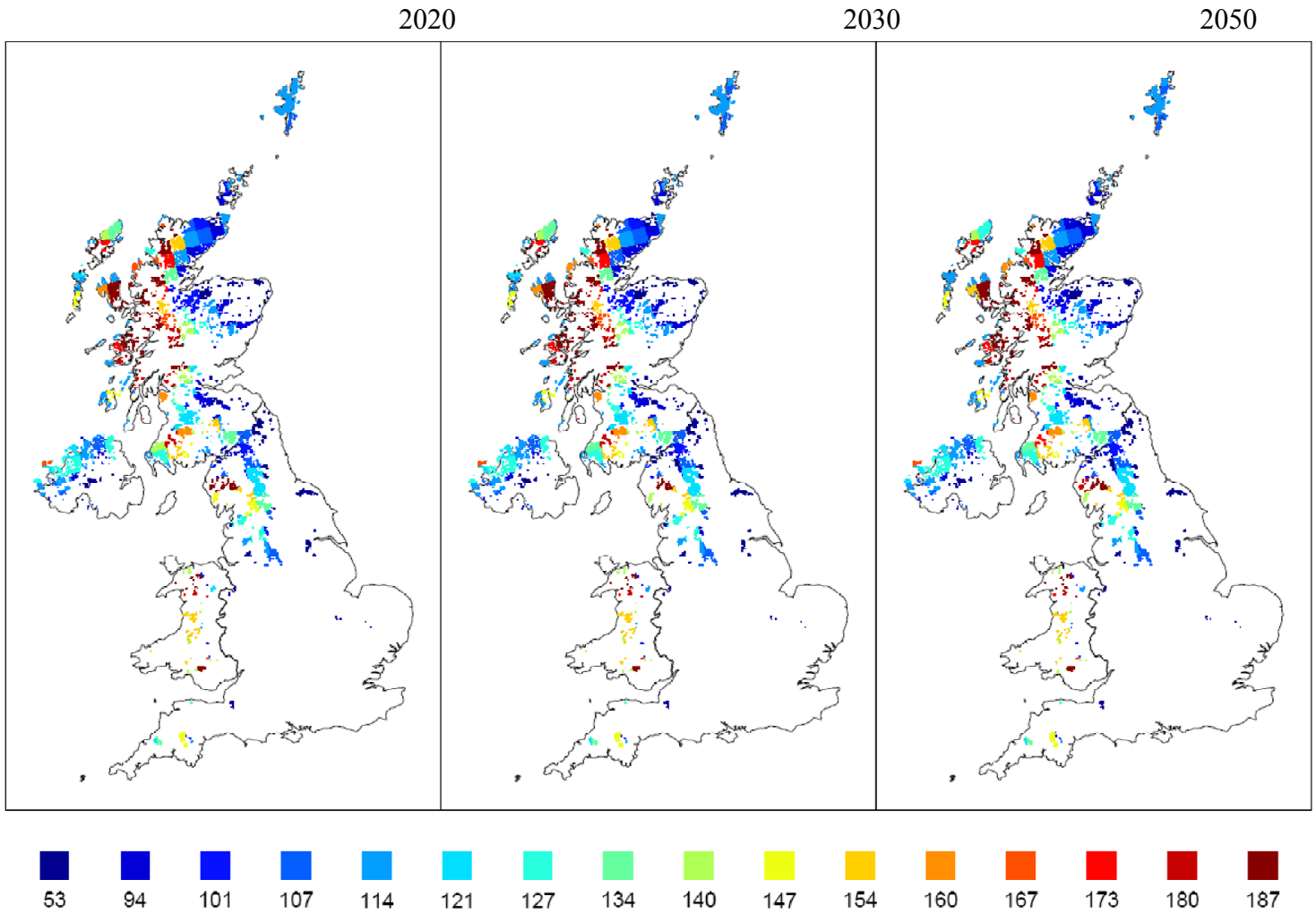


Fig. S11. GAMM predictions of potential *Sphagnum* cover (%) at the 4 m^2 scale for 2050 based on median estimates of UKCP09 climate variables (25 × 25 km). Maps show (a) lower, (b) central and (c) upper estimates based on the niche model parameter estimates and their 95% confidence intervals

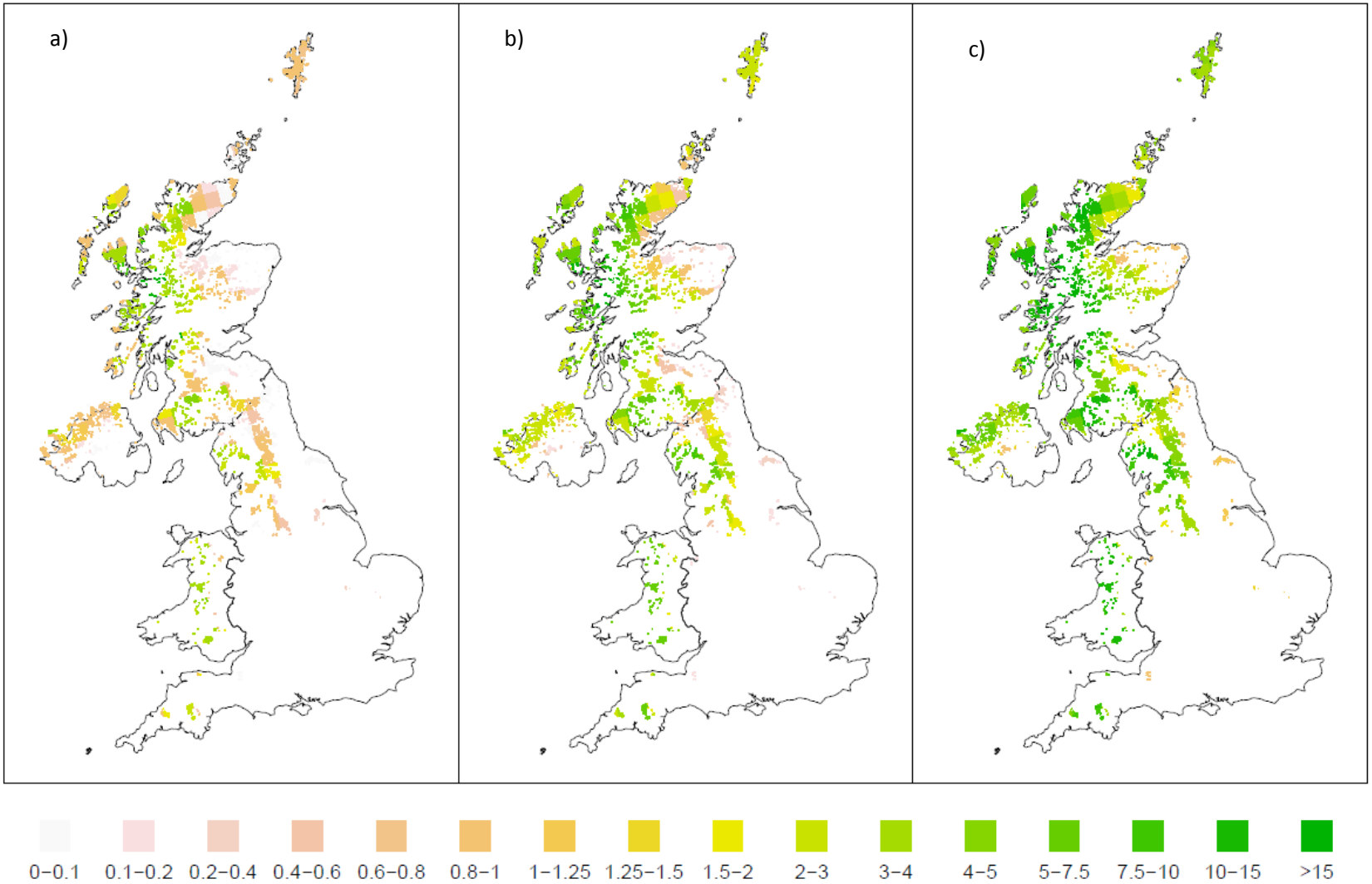
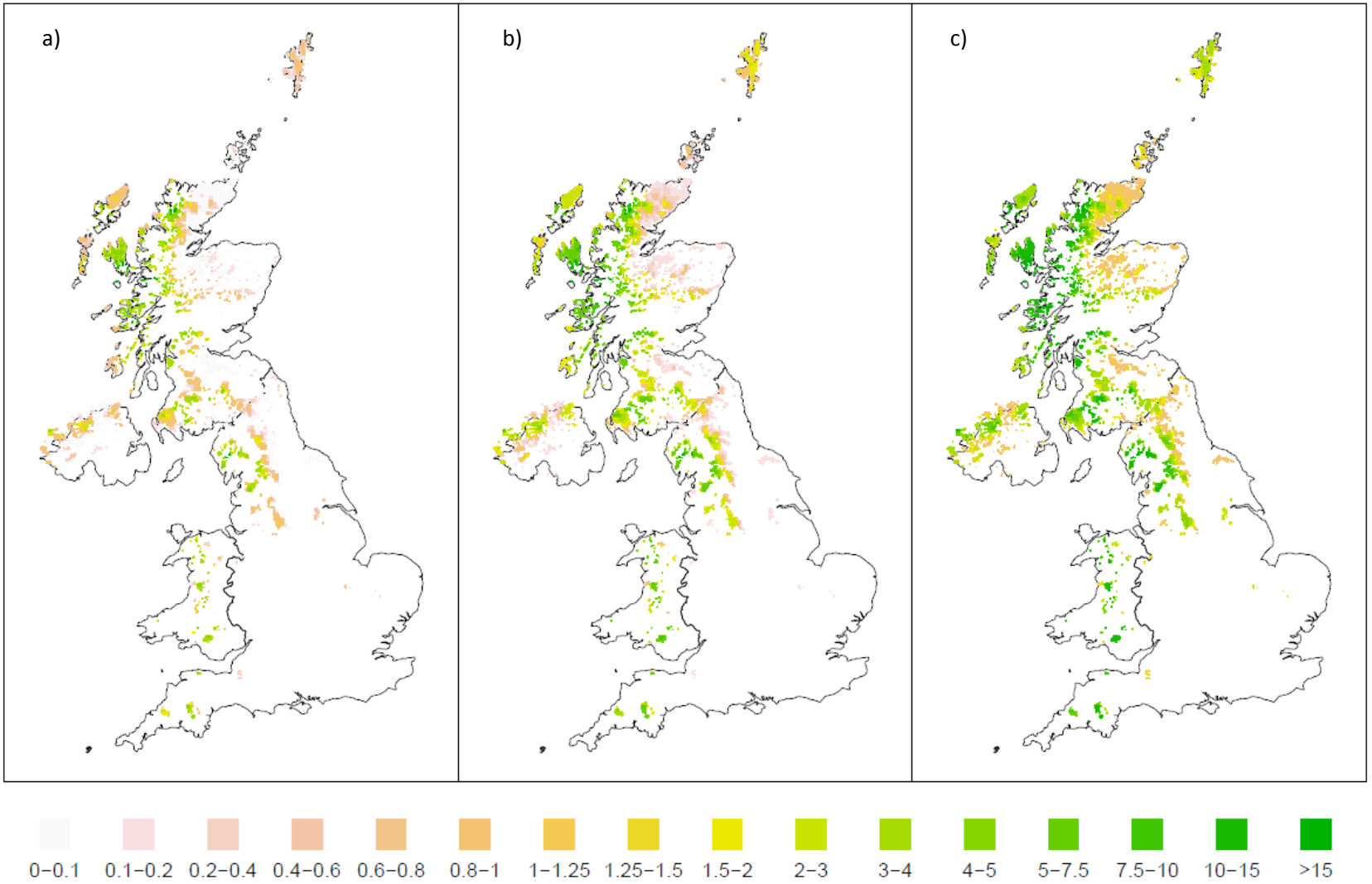


Fig. S12. GAMM predictions of potential *Sphagnum* cover (%) at the 4 m^2 scale for 2050 based on UKCIP02 climate variables (5 × 5 km). Maps show (a) lower, (b) central and (c) upper estimates based on the niche model parameter estimates and their 95% confidence intervals



4. Plots of 'red' ombrotrophic *Sphagnum* abundance versus explanatory variables

Maximum cover and the proportion of plots occupied are plotted below against intervals of each of the explanatory variables featured in either the final GLMM or GAMM

Fig. S13. Ombrotrophic *Sphagnum* cover and presence versus cover-weighted canopy height in each vegetation plot from the Countryside Survey of 1998. Cover-weighted canopy height was based on an average of an ordinal classification of the average canopy height of vascular plants into 8 height intervals (see Smart et al. 2010 for details). The bars in the top panel indicate the maximum cover of ombrotrophic *Sphagnum* spp. recorded in the model training data within the relevant interval of each explanatory variable (indicated by the lower X-axis label). The (] notation for the histogram intervals indicates that the right hand value is included in the interval but the left hand one is not

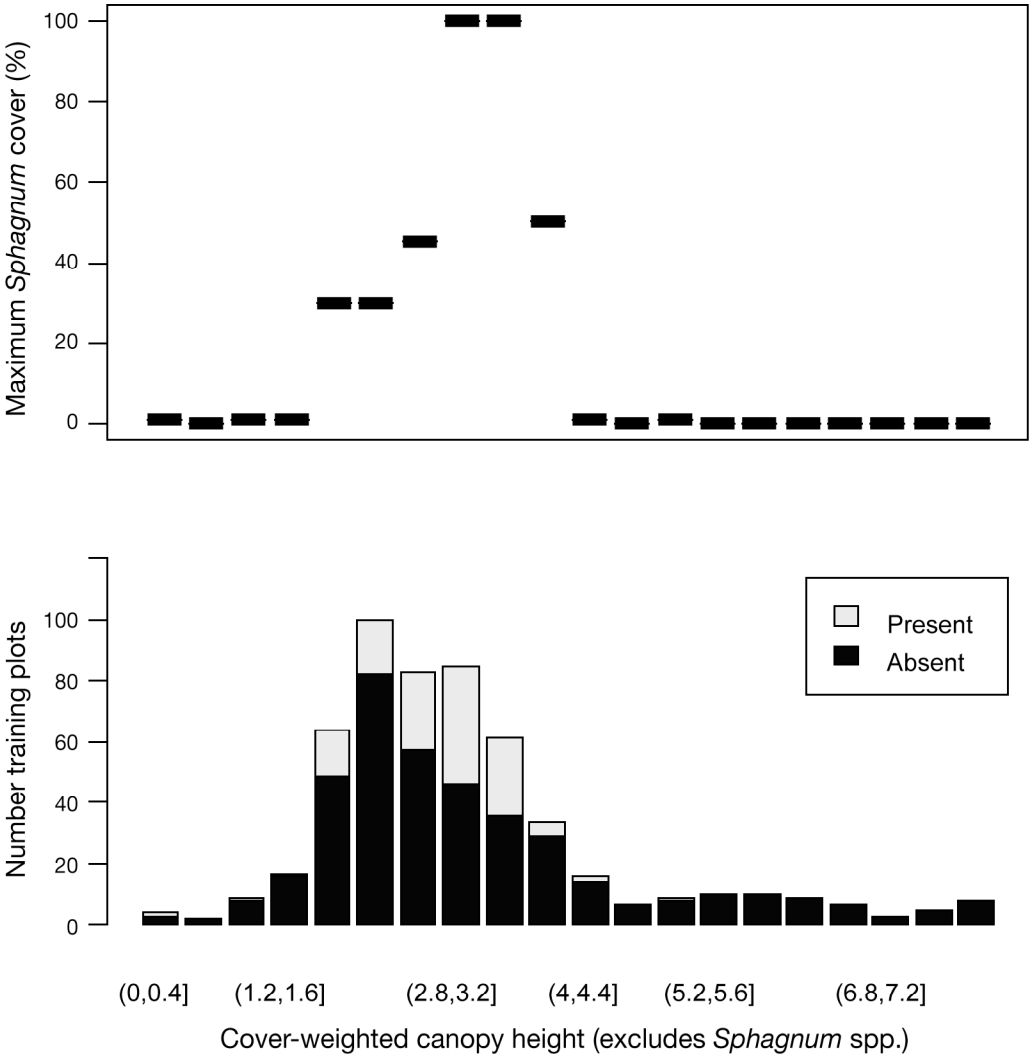


Fig. S14. Ombrotrophic *Sphagnum* cover and presence versus the long-term average (1961–1990) of maximum July temperature for the 5 × 5 km squares in which the Countryside Survey 1998 plots were located. Other details as in Fig. S13

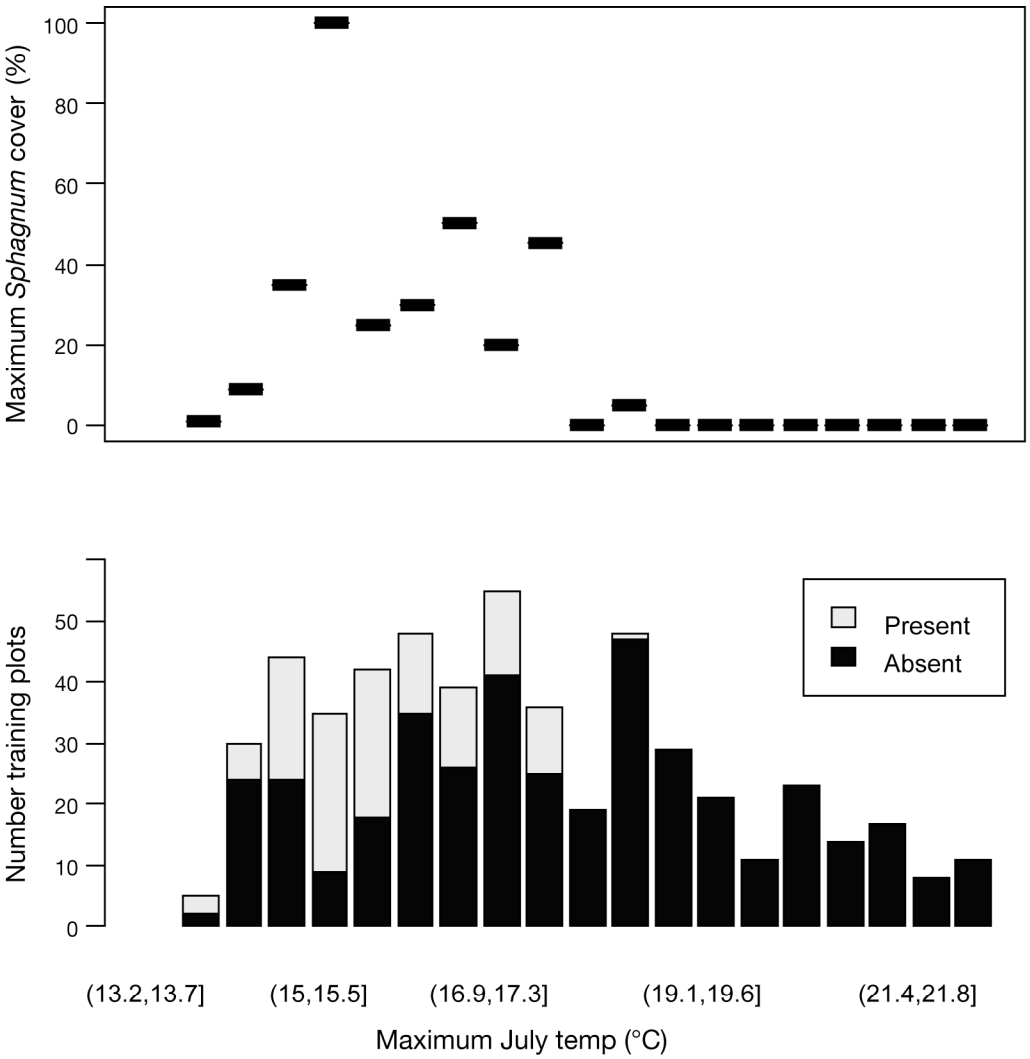


Fig. S15. Ombrotrophic *Sphagnum* cover and presence versus the long-term average (1961–1990) monthly precipitation for the 5 × 5 km squares in which the Countryside Survey 1998 plots were located. Other details as in Fig. S13

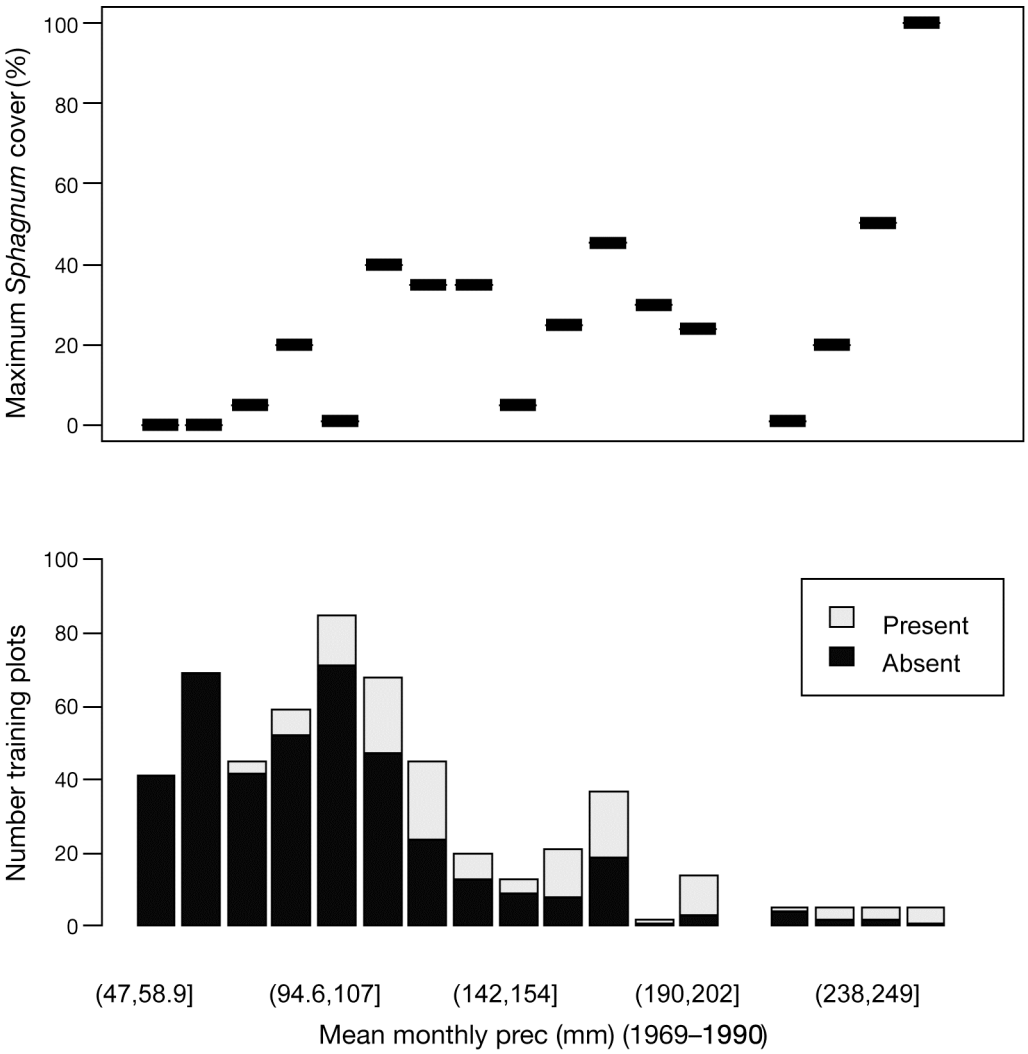


Fig. S16. Ombrotrophic *Sphagnum* cover and presence versus substrate C:N ratio for the top 15 cm of soil sampled in each plot in the Countryside Survey of 1998. Other details as in Fig. S13

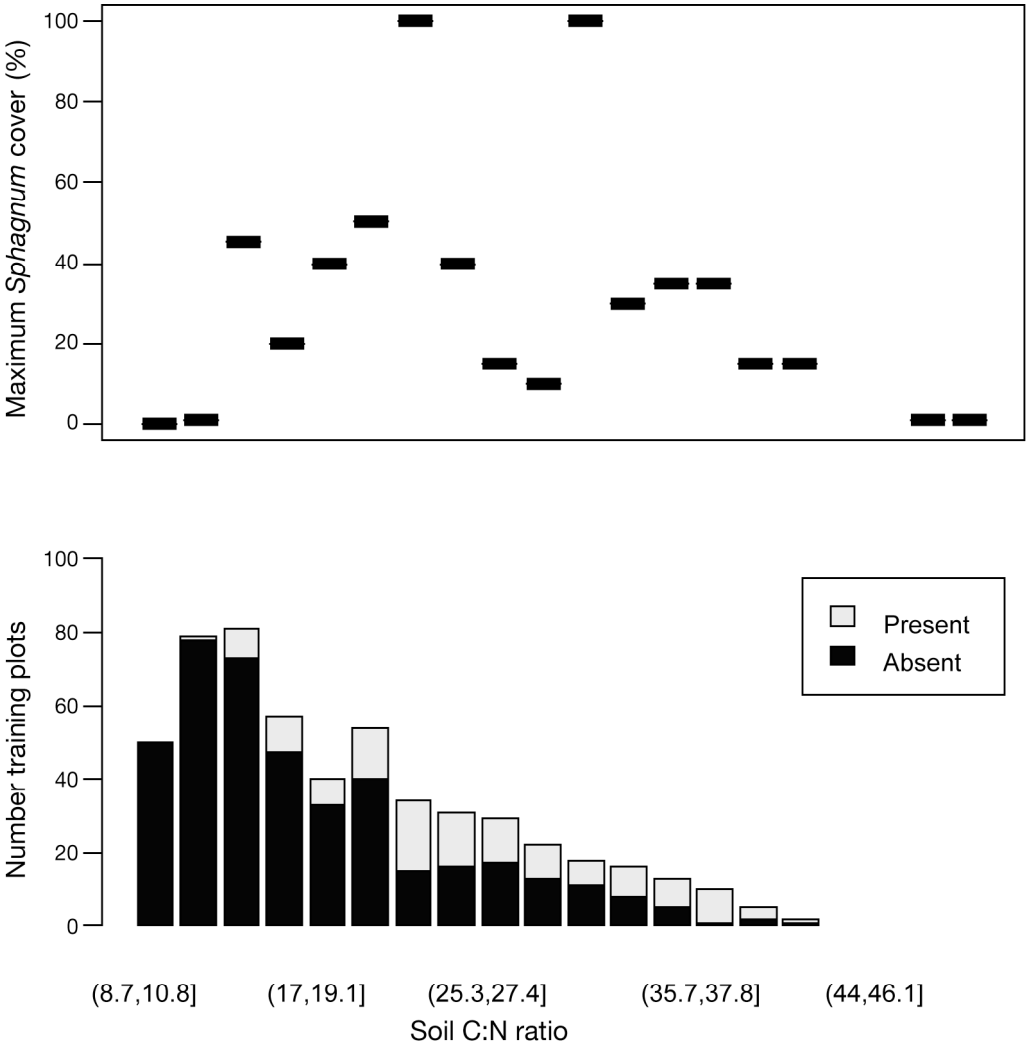


Fig. S17. Ombrotrophic *Sphagnum* cover and presence versus substrate organic carbon (%) in the top 15 cm of soil sampled in each plot in the Countryside Survey of 1998. Other details as in Fig. S13