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

S. M. Smart1,*, P. A. Henrys1, W. A. Scott1, J. R. Hall2, C. D. Evans2, A. Crowe1, E. C. Rowe2, U. Dragosits3, T. Page4, J. D. Whyatt4, A. Sowerby2, J. M. Clark5,6,7

1NERC Centre for Ecology and Hydrology, Library Avenue, Bailrigg, Lancaster LA1 4AP, UK
2NERC Centre for Ecology and Hydrology, Environment Centre Wales, Deiniol Road, Bangor, Gwynedd LL57 2UW, UK
3NERC Centre for Ecology and Hydrology, Environment Centre Wales, Deiniol Road, Bangor, Gwynedd LL57 2UW, UK
4Lancaster Environment Centre, Lancaster University, Bailrigg, Lancaster LA1 4AP, UK
5Wolfson Carbon Capture Laboratory, School of Biological Sciences, Bangor University, Deiniol Road, Bangor, Gwynedd LL57 2UW, UK
6Grantham Institute for Climate Change Fellow, and Environmental Engineering, Imperial College London, South Kensington, London SW7 2AZ, UK
7Present 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

ABSTRACT: A significant challenge in the prediction of climate change impacts on ecosystems and biodiversity is quantifying the sources of uncertainty that emerge within and between different models. Statistical species niche models have grown in popularity, yet no single best technique has been identified reflecting differing performance in different situations. Our aim was to quantify uncertainties associated with the application of 2 complimentary modelling techniques. Generalised linear mixed models (GLMM) and generalised additive mixed models (GAMM) were used to model the realised niche of ombrotrophic Sphagnum species in British peatlands. These models were then used to predict changes in Sphagnum cover between 2020 and 2050 based on projections of climate change and atmospheric deposition of nitrogen and sulphur. Over 90% of the variation in the GLMM predictions was due to niche model parameter uncertainty, dropping to 14% for the GAMM. After having covaried out other factors, average variation in predicted values of Sphagnum cover across UK peatlands was the next largest source of variation (8% for the GLMM and 86% for the GAMM). The better performance of the GAMM needs to be weighed against its tendency to overfit the training data. While our niche models are only a first approximation, we used them to undertake a preliminary evaluation of the relative importance of climate change and nitrogen and sulphur deposition and the geographic locations of the largest expected changes in Sphagnum cover. Predicted changes in cover were all small (generally <1% in an average 4 m² unit area) but also highly uncertain. Peatlands expected to be most affected by climate change in combination with atmospheric pollution were Dartmoor, Brecon Beacons and the western Lake District.

KEY WORDS: Nitrogen · Sulphur · Generalised linear model · Generalised additive model · Uncertainty · Large scale · Peatlands · UKCP09 · UKCIP02

1. INTRODUCTION

Modelling approaches offer important insights into the impact of climate change and other interacting drivers on global ecosystems, their biodiversity and the services they provide. However, in the search for answers, it is too easy to ignore or fail to properly weigh the uncertainties that accompany predictions of future states. Here we focused on the impacts of climate change and pollutant deposition on UK peatlands. Using 2 simple niche models, we explored scenarios of possible change in the cover of Sphagnum

*Email: ssma@ceh.ac.uk
moss up to 2050. We quantified the uncertainty associated with the niche model predictions and then undertook a preliminary assessment of the location and magnitude of expected changes in Sphagnum cover associated with these 2 drivers.

Sphagnum mosses comprise a unique group of ecosystem engineers responsible for the development of peatland ecosystems. Where rainfall sufficiently exceeds evapotranspiration, extensive peat deposition and the formation of ombrogenous mires occurs. In the UK, these conditions are broadly associated with >1000 mm annual rainfall, >160 rain days yr⁻¹ and a mean July temperature below 15°C (Lindsay 1995). Peat masses form because Sphagnum moss litter is relatively recalcitrant (Turetsky et al. 2008, Lang et al. 2009): Sphagnum species reduce decomposition directly through the formation of phenolic compounds, and indirectly through water retention and acidification of their immediate environment (Clymo & Hayward 1982).

Ombrogenous mires provide important ecosystem services at small and large scales. They are a net sink for carbon dioxide (CO₂) and are thought to store approximately 455 gigatonnes of carbon globally, which amounts to 20 to 30% of all the soil carbon on earth (Gorham 1991, Worrall & Evans 2009). Upland peatlands cover about 8% of UK land area. Combined with Ireland, this amounts to about 15% of the total world resource of blanket peat bog (Tallis 1997). In the UK, peatlands are also recognised for their cultural value and characteristic biodiversity (Lindsay 1995). For example, ombrogenous mire ecosystems are classified as Priority Habitats (Blanket Bog and Raised Bog) under the UK Biodiversity Action Plan and as Annex I habitats under the European Habitats Directive. Whilst small-scale species diversity among most taxonomic groups is typically low in these ecosystems, characteristic species have a narrow ecological tolerance and are geographically restricted.

Historically, the functioning of upland mires in Britain has been disrupted by a range of human activities. These include drainage (Holden et al. 2004), overgrazing (Fuller & Gough 1999), burning (Yallop et al. 2009), afforestation (Cannell et al. 1993) and the atmospheric deposition of sulphur and nitrogen (Caporn & Emmett 2009). In the UK, the largest areas of these habitats are now afforded considerable legal protection against many of these pressures, yet climate change poses a new, unregulated threat. Globally, ombrotrophic peatlands (i.e. those in which water and nutrient inputs are largely derived from direct precipitation) are a feature of high latitude ecosystems, all of which are considered at particular risk of change due to warming and changes in the seasonality of rainfall (Pachauri & Reisinger 2007). The expected impact of climate warming is to shift peatlands from net carbon sinks to carbon sources (Gorham 1991, Worrall & Evans 2009). However, considerable uncertainties remain over the timescales and dynamics of peatland responses to climate warming (Limpens et al. 2008, Wania et al. 2009). This reflects uncertainties in climate projections (Zaehle et al. 2007) as well as imprecise knowledge regarding the resilience of mire ecosystems and their ability to withstand or abruptly change in response to a warming climate which may include extreme weather events (Heijmans et al. 2008, Dise 2009). Prediction of climate change impacts are further complicated because future changes are likely to involve interactions between multiple drivers, such as nitrogen deposition, the legacy of previous sulphur deposition and other factors affecting current habitat conditions (Caporn & Emmett 2009). For example, in northwest Europe, atmospheric pollutant deposition of both sulphur and nitrogen have been associated with historical declines in Sphagnum cover (Ferguson et al. 1978, Tallis 1987, Lee 1998), and recent studies have shown that nitrogen deposition has accelerated decomposition of Sphagnum (Braganza et al. 2006). The effects of further changes in pollutant deposition will vary, reflecting the interplay between the persistent effects of historical and ongoing nitrogen deposition and recovery from reductions in sulphur deposition since the early 1970s (Fowler et al. 2004). Like climate change, prediction of pollutant deposition impacts is problematic because models of pollutant emission, deposition and biogeochemistry also carry many separate sources of uncertainty (Schouwenberg et al. 2001, Page et al. 2004).

In this paper, we focus specifically on uncertainties associated with modelling changes in the suitability of ecological conditions for Sphagnum growth in response to climate change and pollutant deposition. As Sphagnum is a key peat-building genus, changes in its distribution are likely to affect long-term net carbon accumulation by altering the supply of resistant organic matter to the peat mass and also by changing the risk of surface erosion. Hence, understanding whether future environmental conditions are within the ecological niche for Sphagnum will help to assess ecosystem vulnerability to change. We first modelled the realised niche of aggregated ombrotrophic Sphagnum species using paired species abundance and environmental data from Britain (Fig. 1). Two popular techniques, generalised linear mixed models (GLMM) and generalised additive mixed models (GAMM), were used to generate empirical niche models. These were used to provide a first approximation of the potential for change with respect to climate and deposition drivers. Scenarios of climate change and pollutant deposition for the years 2020, 2030 and 2050 were then applied singly and in combination to drive changes in the predicted suitability of the niche for ombrotrophic Sphag-
num species in upland mires of the UK (Fig. 1). Our objective was to model impacts at the 3 time steps, and then to evaluate the relative contribution of the uncertainty on the parameters of the empirical niche models by comparison with other sources of variability in the dataset of predictions. To quantify this relative contribution, we further decomposed variation in the predicted values into unique components attributable to variation between the 3 years considered, to climate change uncertainty and to variation between climate change and pollutant deposition scenarios over time, having averaged out variability due to niche model parameter and climate change uncertainty. Having averaged over all these factors, we then quantified the remaining spatial variation in the dataset of predictions between the 1 km peatland squares in Great Britain. This enabled us to express the proportions of variation in the total dataset of predictions that were attributable to these various factors. Our objective was not to develop the final, best possible Sphagnum niche model but to produce a sufficiently useful first approximation for the purposes of illustrating uncertainties between the techniques relative to other sources of variation in the predictions generated. Our predictions are discussed in terms of the relative importance of each driver in different parts of Britain, but caution should be exercised in interpreting these results. Specifically, we asked the following questions: (1) What are the best predictors of ombrotrophic Sphagnum cover across Britain? (2) What are the relative contributions of the following to the total range of variation in the dataset of model predictions across both techniques: (a) niche model parameter uncertainty, (b) variation in climate predictions, (c) variation in predictions of ombrotrophic Sphagnum cover between scenarios of climate change and pollution and (d) spatial variation in predicted Sphagnum cover across the UK having accounted for the effects of (a) to (c)? (3) Does predicted change in ombrotrophic Sphagnum cover due to climate and pollution impacts vary spatially across the UK?

2. METHODS

Analysis was carried out in 2 stages (Fig. 1); niche models were first produced for grouped ombrotrophic Sphagnum species. These models were then used to forecast species cover in 2020, 2030 and 2050 across bogs in the UK using inputs from climate change scenarios and values of soil variables derived from simulating the impact of sulphur and nitrogen deposition on the soil biogeochemistry of upland peats. This simulation was carried out by using modelled estimates of atmospheric pollutant deposition from the Fine Resolution Atmospheric Multi-pollutant Exchange Model (FRAME; Fournier et al. 2002) as input to the soil biogeochemical very simple dynamic (VSD) model (Posch & Reinds 2009). Modelling of deposition was carried out at the 5 x 5 km scale, and soil responses were modelled at the 1 km² scale. Sphagnum cover was modelled at a scale of 4 m² plots but averaged over each 1 km².

2.1. Niche modelling datasets

The models were trained on observed percentage cover data for aggregated ombrotrophic Sphagnum species matched with potential explanatory variables (Table 1). Cover data were recorded as part of the Countryside Survey (CS) of Great Britain carried out in 1998 (see Smart et al. 2003 for details) from a stratified random sample of 623 small quadrats (4 m²) nested within a stratified random sample of 172 squares (each 1 km²) from across the whole of England, Scotland and Wales. Explanatory variables were assembled either at the plot scale (canopy height and substrate measurements) or within the wider 5 x 5 km (interpolated long-term annual average climate data, 1961–1990, from
the UK Met Office; see Table 1). Selected explanatory variables were judged sufficient to represent the principal factors inhibiting or favouring Sphagnum growth. These comprised explanatory variables able to track climatic gradients, historical sulphur deposition, successional stage of the vegetation and substrate conditions. Long-term annual averages (1961–1990) for the British climate were assembled for variables that could adequately represent the small temperature range, cool conditions and high rainfall that broadly define the optimum oceanic envelope for ombrotrophic Sphagnum species in Britain (Lindsay et al. 1988, Lindsay 1995, Clark et al. 2010, this Special). Therefore, we used climate variables likely to best describe gradients of temperature range, cloudiness and precipitation (Table 1). We used long-term annual averages from 1961–1990 so as to cover a long enough window to capture the pre-1980 rise in the Central England Temperature Record; however, we also wanted the climate data to be as close as possible to the period during which the training data were recorded. This does mean that some unknown proportion of the spatial pattern in observed Sphagnum cover could reflect recent climate driven change between 1980 and 1990.

In addition, measurements made at the 4 m² plot scale were available for a range of substrate factors as well as cover-weighted canopy height (see Smart et al. 2010 for details) for the vascular plant layer. Soil pH, % organic nitrogen, % organic carbon and the C:N ratio were included as predictors. These variables track gradients in litter quality, decomposition rate and the rapidity of nutrient cycling and hence correlate with changes in peatland ecosystem functioning from conditions favourable to peat growth through to conditions more favourable for vascular plants with lower C:N ratio litter inputs. Such directional shifts can be driven by pollutant deposition and climate change and so are likely to translate into changes in the favourability of conditions for ombrotrophic Sphagnum growth. Cover-weighted canopy height was also used to capture the favourability of successional stage of the vegetation within the niche models. The contribution of ombrotrophic Sphagnum to the height of the vegetation was excluded from the calculation of cover-weighted vegetation height to avoid circularity; however, the same cannot be said for the substrate measurements. The issue of including explanatory variables in niche models when the focal species can strongly affect the variable is discussed below. Our premise is that inclusion of these variables is informative if external drivers can drive change along these gradients in competition with the species’ ability to maintain favourable conditions. These more finely resolved data were included to help explain smaller-scale spatial variation in Sphagnum abundance within each 1 km square. Explanatory variables also included those whose values could be changed in accordance with scenarios of climate change and pollutant deposition (Table 1).

In the CS, cover of Sphagnum species was estimated using the coarse categories of red/thin, red/fat, green/thin and green/fat. These categories reflect pragmatism in the face of a national deficit of readily deployable field botanists whose knowledge includes the ability to assign Sphagnum to species level. Most individual species can be exclusively assigned to each category, although some species will have been recorded in different categories given their variation in colour and size, e.g. S. fallax, S. capillifolium and S. palustre. After consulting with the British expert on the genus (M. O. Hill pers. comm.), we amalgamated records for red/thin and red/fat and proceeded on the basis that these data would largely encapsulate the ombrotrophic species S. capillifolium and S. magellanicum with occasional additional records for S. papillosum, S. russowii, S. fallax and other rarer taxa.

Cover data were recorded to 5% intervals with <5% cover coded as 1% and absences coded as 0.0005% cover. Data were square-root transformed prior to coding absences. A logit transform was then applied to ensure that predictions were restricted to the range of the root-transformed cover data. Plot level observations from all habitat types were included in the model building process, with the exception of those that fell into urban, boundaries and linear features, improved grassland, inland rock, maritime habitats and arable land (Jackson 2000).

Table 1. Explanatory variables used to build minimum adequate models for ombrotrophic Sphagnum species cover in Britain. All climate variables were calculated as long-term averages for the years 1961–1990 inclusive. Responsive?: responsive to scenario (P: pollution, C: climate change). Climate data were downloaded from www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09.download/access_gd/index.html#ita. For determination of substrate variables in Countryside Survey (CS) plots, see Emmett et al. (2010). For calculation of cover-weighted canopy height see Smart et al. (2010)

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Source</th>
<th>Scale</th>
<th>Range in training data</th>
<th>Responsive?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substrate % organic C</td>
<td>CS</td>
<td>4 m² plots</td>
<td>1.18–56.70</td>
<td>P</td>
</tr>
<tr>
<td>Substrate C:N ratio</td>
<td>CS</td>
<td>4 m² plots</td>
<td>8.70–50.29</td>
<td>P</td>
</tr>
<tr>
<td>Cover-weighted canopy height of associated species</td>
<td>CS</td>
<td>4 m² plots</td>
<td>0–8</td>
<td></td>
</tr>
<tr>
<td>Mean monthly rainfall per year (mm)</td>
<td>UK Met Office</td>
<td>25 km²</td>
<td>46.4–275.5</td>
<td>C</td>
</tr>
<tr>
<td>Mean maximum July temperature (°C)</td>
<td>UK Met Office</td>
<td>25 km²</td>
<td>11.21–22.64</td>
<td>C</td>
</tr>
</tbody>
</table>
2.2. GLMMs

GLMMs were fitted using the SAS procedure MIXED (Little et al. 2000). The best minimum adequate model was selected by a series of manual steps. Firstly, matrix plots of explanatory variables were examined. Those that were significantly correlated (Spearman rank correlation, p < 0.05) were fitted individually in the absence of all other variables, and, where statistically significant, the variable that most minimised the Akaike Information Criterion (AIC) value (Akaike 1974) was retained for further fitting. A final set of potential explanatory variables was then tested in all possible combinations to determine the model with the lowest AIC value. Quadratic and interaction terms were only tested where ecologically justified. Given the nesting of plot-level observations within the 1 km squares of the CS, we adopted a mixed modelling approach. Hence, the 1 km square was treated as a random fixed effect, and degrees of freedom were down-weighted using the method of Satterthwaite (1946) to account for the relative non-independence of plots within grid squares.

2.3. GAMMs

In the same way that GLMMs extend the standard GLM theory to include extra random components in the model framework, GAMMs (Lin & Zhang 1999) extend the framework of the standard GAM. The random components included in the mixed model account for unobserved effects that could influence the outcome of the response variable. Hence, the 1 km square was again included as a random class variable. The GAMM can be written as:

\[
g(\mathbb{E}[y \mid x, b]) = \alpha + \sum_{j=1}^{k} f_j(x_j) + Z_b \tag{1}
\]

where \( y \) is the response variable, \( f_j \) represents smooth functions (a piecewise continuous set of cubic polynomials, Poirier 1973), \( g \) is the link function and \( \alpha \) is the intercept term. \( Z \) represents different grouping levels and \( b \sim N(\mathbf{0}, \sigma^2) \) represents the differing variation assigned to each of the groups defining the random factor \( Z \). GAMM modelling was carried out using the R statistical language and the mgcv package (Wood 2008). Regression parameters were estimated using a penalised quasi-likelihood based approach.

As in the case of the GLMM, parameter selection in the model was dictated firstly by ecological plausibility. All combinations of variables were considered together with a number of specific 2-dimensional joint-effect terms, which were included as 2-dimensional tensor product smooths, due to the differing scales of each variable. We also allowed for any residual spatial variation in the data resulting from possible spatial autocorrelation by specifying a 2-dimensional smooth term in the model for easting and northing corresponding to the location of the survey square. Having fitted each model, plots of residuals were checked together with quantile-quantile plots of the observed and predicted values. The final model selected was that which minimised the AIC and the cross-validation score. Model training was based on a 'leave one out' cross-validation, such that the developing model was used to predict an observation omitted from the model training set. Each observation was then put back in and the next omitted in turn and the model optimised until every observation was left out once.

2.4. Model testing

Final models were validated against an independent dataset. Vegetation plot data were drawn from the Great Britain CS carried out in 2007 with the proviso that none of the locations were recorded in 1998, avoiding the possibility that the test data included observations highly spatially correlated with the training data from the 1998 survey. We selected 97 plots, of which 15 had observed cover of ombrotrophic Sphagnum (see Section 1 in the Supplement at www.int-res.com/articles/suppl/c045_p163_supp.pdf). The distribution of the predicted values from each model was graphically compared between occupied versus unoccupied plots. No formal statistics measuring accuracy or classification rate were calculated because our models, though useful first approximations, were primarily developed to investigate the uncertainty in their parameters. We considered our models sufficiently validated for our purposes if the average predicted cover of ombrotrophic Sphagnum was higher in occupied plots than in unoccupied plots.

2.5. Climate change projections

Future climate data were obtained from the UK climate projections website http://ukclimateprojections.defra.gov.uk/ © UK Climate Projections 2009; UKCP09). Climate predictions are presented as averages, be they annual, seasonal or monthly values, taken over a 30 yr window from 2010–2039 up to and including the time period of 2070–2099. Projections were available on a 25 × 25 km grid covering the whole of the UK under various emissions scenarios graded according to the Special Report on Emissions Scenarios (Nakicenovic & Swart 2000).

The user interface was used to access the climate projections database, and predictions were obtained for 2020, 2030 and 2050, using the 30 yr windows of
2010–2039, 2020–2049 and 2040–2069, respectively, under the high emissions scenario. The UKCP forecasts are based on model averaging across an ensemble of regional climate models (RCMs). Variation in the average predictions for each time interval and across the 25 × 25 km grid is available from the website, and we ran predictions based on climate values at the 33rd, 50th (the central estimate) and the 67th percentiles. The full uncertainty within each RCM ensemble member is not included in these outputs, while the effect of averaging at the 25 × 25 km scale will also have the effect of reducing the total range of the predicted values and making the predictions insensitive to variation in climate change within each grid cell (Trivedi et al. 2008).

Data were downloaded for the mean daily maximum July temperature, the mean daily minimum January temperature, both of which are readily available from the website, and mean rainfall. These variables were selected reflecting their established use in defining range limits for British species (see for example Hill et al. 2004, 2007). UKCP09 presents rainfall as mean mm d⁻¹, whereas rainfall data used in the model training were mean monthly values. Therefore, to match the scaling of the 2 sources, the UKCP09 rainfall data (in mm d⁻¹) were multiplied by a factor of 365/12.

Scenarios were also run for the same time window and the same variables but using the earlier UK Climate Impacts Programme 2002 (UKCIP02) climate projections (www.ukcip.org.uk). This was done as a validation step. While the UKCP09 projections are based on probability distributions of climate data for each 25 × 25 km grid cell, these distributions do not form a spatially coherent series. Each grid cell probability distribution is independent of neighbouring cells. This changes the way climate impact maps must be interpreted. Because it was not clear what effect this change in approach would have, we produced separate predictions of Sphagnum cover across British peatlands using both projections (see Figs. S11 & S12 in Section 3 of the Supplement at www.int-res.com/articles/suppl/c045_p163_supp.pdf).

2.6. Pollutant deposition data

Atmospheric sulphur (S) and nitrogen (N) impacts were simulated using the VSD model. This biogeochemical model was developed for large-scale application under the UN Convention on Long-range Transboundary Air Pollution, as an extension of the widely used critical loads concept, and is described in detail by Posch & Reinds (2009). Briefly, the model simulates soil acid–base chemistry on an annual time step as a function of deposition and weathering inputs, element sinks (such as net vegetation uptake), and soil cation exchange equilibria. N enrichment is modelled based on an annual mass budget, whereby deposited N is lost to vegetation uptake, denitrification and long-term soil formation at prespecified rates. Residual N is either accumulated in the organic matter pool or leached, as a function of the pool C:N ratio; above a threshold value, all N is assumed to be retained, while below this threshold the proportion of N leached increases linearly as a function of declining C:N ratio. The VSD model is designed to run with the limited data available at national scales, and for the UK has been applied on a 1 km grid to 6 broad habitats (including bog) with input parameters derived from UK soil and vegetation databases and critical load default values (for further details on model parameterisation and application, see Hall et al. 2008). The model was run with present-day S and N deposition inputs derived from the Centre for Ecology and Hydrology (CEH) Edinburgh Concentration Based Estimated Deposition (CBED) 5 × 5 km gridded dataset (Smith et al. 2000), with historic and future deposition sequences (relative to present-day values) derived from the UK deposition FRAME model (Fournier et al. 2004). Deposition forecasts assumed that currently legislated S and N emissions reductions occurred linearly from 2005 to 2020, with constant deposition thereafter. VSD model outputs utilised in the Sphagnum niche models were soil pH, and substrate C (%), N (%) and C:N ratio.

2.7. Niche model application

Model predictions were analysed for 2020, 2030 and 2050 because of their coincidence with the moving climate prediction windows and because the 30 yr interval between 2020 and 2050 is a reasonably realistic horizon over which to evaluate expected ecological changes linked to the 2 drivers. For each of the 3 years, values for the final explanatory variables in each niche model were changed reflecting the 2 scenarios of change in climate and pollutant deposition (Fig. 1). Uncertainty was introduced into the model predictions in 2 ways. Sets of predictions for each of the 18 261 squares (1 × 1 km) in Britain containing peat bog were generated for both the GLMM and GAMM. Predictions were made at (1) the best estimate of all model parameters and then at (2) the upper and lower 95% confidence intervals of all model parameters. These predictions were made at the median, 33rd and 67th percentile values of predicted rainfall and temperature estimates from UKCP09 for each 25 × 25 km square.

2.8. Uncertainty analysis

Uncertainties associated with estimates of deposition and soil biogeochemical processes were not available.
Analysis therefore focussed only on the contribution of the niche model uncertainty relative to the spatial and temporal variation in the dataset of predictions, part of which could also be attributed to the percentile variation in the climatic inputs. The proportions of the total variance attributable to each type of uncertainty in each dataset of predictions from either the GLMM or GAMM were calculated using the SAS procedure VARCOMP (SAS Institute 1989). Maps were also generated to show the uncertainty around the best estimates of Sphagnum cover at each time step.

3. RESULTS

3.1. Best predictors of ombrotrophic Sphagnum cover across Britain

The best fitting GLMM and GAMM differed in the explanatory variables selected. The GLMM included a quadratic relationship with vegetation canopy height, indicating an optimum for ombrotrophic cover in short vegetation, with Sphagnum less likely to be found in the shortest vegetation and not favoured under taller canopies (See Fig. S5 in Section 2 of the Supplement and Fig. S13 in Section 4 of the Supplement at www.int-res.com/articles/suppl/c045_p163_supp.pdf). Other terms reflected expected positive responses to substrate C:N ratio and mean monthly rainfall per year based on a significant interaction between the 2 variables. This indicated that an increase in the favourability of conditions for Sphagnum at higher C:N and rainfall was greater than just the sum of the 2 variables (Table 2, see Fig. S7 in Section 2 of the Supplement and Figs. S15 & S16 in Section 4 of the Supplement). Trying all possible combinations of parameters together with hypothesised 2-dimensional joint effects resulted in a final best fitting GAMM consisting of terms for soil carbon content (%); vegetation height; 2-dimensional tensor product smooth terms for the joint effect of rainfall and carbon content and the joint effect of rainfall and the mean maximum temperature in July; a 2-dimensional smooth surface of the easting and northing of the location of each 1 km square; and a random effect of survey square on between-plot variation.

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>Intercept</th>
<th>Cover-weighted canopy height</th>
<th>(Cover-weighted canopy height)$^2$</th>
<th>Substrate C:N ratio</th>
<th>Mean monthly precipitation per year (mm)</th>
<th>Interaction (C:N × Precipitation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>−3.2238</td>
<td>0.3599</td>
<td>−0.04437</td>
<td>−0.04363</td>
<td>−0.00532</td>
<td>0.000667</td>
</tr>
<tr>
<td>Lower 95</td>
<td>−3.7713</td>
<td>0.1971</td>
<td>−0.06299</td>
<td>−0.06803</td>
<td>−0.00937</td>
<td>0.000461</td>
</tr>
<tr>
<td>Upper 95</td>
<td>−2.6764</td>
<td>0.5227</td>
<td>−0.02395</td>
<td>−0.01923</td>
<td>−0.00127</td>
<td>0.000873</td>
</tr>
</tbody>
</table>

Table 2. Generalised linear mixed model (GLMM) for modelling cover of ‘red’ ombrotrophic Sphagnum species across British peat bogs. Model parameters and confidence intervals for best minimum adequate model

Table 3. Decomposition of total variance in ombrotrophic Sphagnum cover explained by each modelling technique based on the same training data from the British Countryside Survey carried out in 1998 (n = 623 vegetation plots in 172 squares of 1 × 1 km). GAMM: generalised additive mixed model; GLMM: generalised linear mixed model

<table>
<thead>
<tr>
<th></th>
<th>Between squares</th>
<th>Within squares</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAMM</td>
<td>77.1</td>
<td>48.1</td>
<td>51.2</td>
</tr>
<tr>
<td>GLMM</td>
<td>69.1</td>
<td>16.9</td>
<td>37.1</td>
</tr>
</tbody>
</table>

3.2. Explanatory power of the ‘best’ models

As expected, the GAMM explained substantially more variation in the training data than the GLMM (Table 3). In particular, the ability of the GAMM smoothing functions to account for local variation in the covariate space was evident by the much higher amount of between-plot, within-square variation accounted for. The proportions of between-square variation explained by each model were more similar. Despite the better performance of the GAMM, it still only managed to explain 51% of the variation in the cover data used to build the model (Table 3).

3.3. Model testing

Comparisons of the range of predictions for unoccupied versus occupied test plots showed that predicted values were higher for those plots in which ombrotrophic Sphagnum had been observed (see Figs. S1–S4 in Section 1 of the Supplement). However, the high uncertainty in the GLMM parameters meant that predictions covered a very large range and would be of little use as expectations of cover in any specific location. The GAMM test produced similar results to the GLMM but with a larger range of variability in predicted values and a higher median prediction for plots with ombrotrophic Sphagnum species present. Eight outliers coincided with conditions predicted to be highly favourable for ombrotrophic Sphagnum. Sphagnum species were observed to be present in 3 of these plots but absent from the other 5.
3.4. Projected change in driving variables

The VSD biogeochemical model suggested a widespread reduction in C:N ratio across UK peatlands. Reductions were smallest in the less historically polluted far north, and largest towards the south, through Wales, northern Ireland and the Pennines (see Fig. S8 in Section 3 at the Supplement). UKCP09 projections for maximum July temperatures followed a similar pattern (see Fig. S9 in Section 3 of the Supplement), but average monthly rainfall did not change appreciably by 2050 (see Fig. S10 in Section 3 of the Supplement). Predictions of Sphagnum cover in 2050 were made using both UKCP09 and UKCIP02 datasets (see Figs. S11 & S12 in Section 3 of the Supplement). Although spatial patterns in the predictions were very similar, the most obvious difference relates to the averaging of predicted values within the large 25 × 25 km grid squares used by UKCP09 (Fig. S11) as opposed to the 5 × 5 km UKCIP02 grid (Fig. S12).

3.5. Uncertainty analysis

Predictions from the GAMM and GLMM differed in several respects (Table 4). Average variation in the GLMM output across the UK for the 3 time steps was overwhelmingly dominated by the influence of uncertainty in the niche model parameters. The variation in predictions made at the best estimates, and upper and lower 95% confidence intervals absorbed over 90% of the variance in the dataset (Table 4). The next largest source of uncertainty was the average variation in predicted Sphagnum cover across the UK (7.6%), whilst variation between years, between scenarios and between scenarios conditional on year together explained a very small amount (0.006%) of the total variability (Table 4). Variation in predictions made with the GAMM had a much lower contribution from uncertainty on the model parameters at 14%, with 86% of the variability in the dataset attributable to residual spatial variation in predicted Sphagnum cover across the UK (Table 4). As with the GLMM, a miniscule amount of variability was attributable to average change between time steps. For both models, the contributions of variation due to mean differences in predictions made at the 33rd, 50th and 67th percentiles of the UKCP09 climate data were also very small (Table 4).

The much wider distribution of predicted values for the GLMM (Fig. 2a) compared to the GAMM (Fig. 2b) indicates the relatively larger impact of niche model parameter uncertainty on the GLMM. When predictions were compared between GLMM and GAMM based on the best parameter estimates, but including uncertainty in the climate predictions, the difference in distributions was much lower overall. This highlights the lesser impact of the uncertainty in the climatic means of the 25 × 25 km squares (Fig. 3). These differences are also vividly expressed when projections are mapped across the UK. The uncertainty around the GLMM predictions for Sphagnum cover in 2050 is so large that predictions at the lower and upper confidence intervals indicate an unvaryingly low or high cover irrespective of spatial location (Fig. 4). This contrasts with mapped values for the GAMM, where niche model parameter uncertainty is not high enough to obscure spatial variation in predicted Sphagnum cover at the upper and lower confidence intervals of the model parameters (Fig. 5).

3.6. Projected change in Sphagnum cover varies across the UK

Both GLMM and GAMM predicted either stability or a decrease in ombrotrophic Sphagnum cover across the UK between 2020 and 2050 (Fig. 6). However, all projected changes were generally small and uncertain, especially for the GLMM. The GLMM predictions did take into account pollutant deposition in addition to climate change. Therefore, it is useful to compare the 2 model outputs on the basis that where they agree spatially, the less uncertain GAMM predictions should provide a cross-validated impression of where climate-induced change may be more important than pollutant-induced change.
Peatlands thought to be most negatively impacted by climate change in both models were in northern Scotland (scattered through Assynt, Wester Ross and down to Lochaber). Bogs on Mull and neighbouring Morvern were also expected to be impacted in both niche models, as were areas in the south of Scotland (Galloway and south of Peebles). In England and Wales, both models predicted that the largest reductions in ombrotrophic *Sphagnum* cover would be in the western Lake District and the Brecon Beacons in south Wales (Fig. 6). Areas that were only predicted to be impacted by climate change by 2050 in the GAMM were Dartmoor, peatlands on the western border of northern Ireland and north Lewis in the Outer Hebrides. Comparison with the GLMM predictions of changing *Sphagnum* cover in response to climate and pollution highlighted areas where pollution seemed to be the more important driver (Forest of Bowland and peatlands in Wales) or where pollution exacerbates a predicted impact due to climate change (Dartmoor, Brecon Beacons and the western Lake District; Fig. 6).

4. DISCUSSION

4.1. Niche modelling—the benefits of using more than one modelling technique

We applied 2 robust and routinely applied techniques that offered useful and complementary perspectives on the uncertainties involved in modelling *Sphagnum* cover as well as the possible locations of maximum vulnerability to climate change and pollutant deposition up to 2050. The GAMM, as expected, outperformed the GLMM in terms of explanatory power (Table 3). This is because the weighted smoothing functions were very effective in capturing fine-scale changes in the relationship between response and explanatory variables in the data space. Also, the addition of the easting and northing 2-dimensional smooth in the GAMM, not present in the GLMM, can account for spatial correlation between observations unrelated to the explanatory variables, whereas the fixed terms and associated errors must absorb any spatial correlation in the GLMM. Inspection of model fit
along the rainfall gradient in the training data showed how the GAMM sensitively reflected responses in the ecological data with low uncertainty, while the GLMM provided a much less sensitive response curve with greater uncertainty because the systematically specified component did not capture the detail in the response data (Fig. 7). However, this example also illustrates the greater potential for GAMMs to overfit training data and to thereby produce a model with high explanatory power but potentially low predictive power. Although clearly an under-sampled section of the covariate space, at the highest levels of precipitation the GAMM is sensitive to the apparent reduction in Sphagnum cover and generates a turning point, whereas the GLMM is restricted to a monotonic trend. The tendency of the GAMM to overfit training datasets results in a typically better fit to observed data than the GLMM but with possibly poorer transferability than the GLMM when applied to the same species outside the modelled covariate space (Randin & Dirnbock 2006).

The GAMM and GLMM differed markedly in their parameter uncertainty, yet each has advantages that can be exploited if predictions are used together. Hence, while our predictions are first approximations intended principally to explore uncertainties in the niche models, application to UK ombrotrophic bogs showed that both models highlighted peatlands in the western Lake District and the Brecon Beacons as particularly sensitive to climate change, which is likely to be exacerbated by pollutant deposition. However, the magnitudes of predicted decreases in habitat suitability between 2020 and 2050 were very small.

4.2. Prospects for improved modelling of the niche of ombrotrophic Sphagnum species and its response to multiple drivers

Applying traditional niche theory to Sphagnum is somewhat problematic. Being an ecosystem engineer, Sphagnum shapes its own environment within the constraints of a favourable regional climate and local hydrology (Belyea & Baird 2006, Rydin et al. 2006). In this context, using substrate measurements to indicate habitat suitability for Sphagnum is somewhat circular, since, with greater cover, substrate measurements will increasingly reflect Sphagnum peat properties and
hence are not reasonably interpreted as precursor conditions that are independent of the species’ presence. However, building the niche model only on coarsely resolved climate variables misses an opportunity to add sensitivity to finer-scale variation in substrate factors. The issue is not clear-cut, since many plants adjust their environments to a greater or lesser extent (e.g. Kulmatiski et al. 2008, Orwin et al. 2010), although *Sphagnum* is an extreme case (Turetsky et al. 2008). We included substrate properties that are correlated with *Sphagnum* abundance because drivers such as climate change as well as restoration management can cause change along these abiotic gradients. For example, climate warming, N deposition and increased CO₂ concentration all appear able to favour increased growth of vascular plants relative to *Sphagnum* (Berendse et al. 2001, Freeman et al. 2004, Breeuwer et al. 2008, Gerdol et al. 2008, Heijmans et al. 2008). This can in turn change litter quality, increase litter decomposition rate and drive a reduction in C:N ratio and increase substrate pH. Niche models that capture small-scale spatial changes in the training data along these axes are therefore likely to be better at predicting temporal responses to important drivers of change.

Even though finely-resolved data on canopy height and substrate properties were included in model building, the explanatory power of our niche models was generally low. This was especially so for variation in cover between the small vegetation plots within each of the 1 km CS squares. Unexplained variation will be due to sampling error as well as the range of other variables that influenced *Sphagnum* cover at the time the data were recorded. Increasing the explanatory power of the niche models could probably be achieved with the inclusion of other variables. Examples include summer water table depth (Clymo 1984), slope and deviation in elevation from the mean bog surface, thus providing information on whether locations were hummocks, hollows or lawns (Moore et al. 2007). Topographic data could be derived from low-level remotely sensed imagery, but high-resolution data would be needed to discriminate the small spatial scales over which changes in slope and elevation are associated with changes in *Sphagnum* cover and species composition. Acquisition of further explanatory variables faces an inevitable trade-off between resolution and total area covered. Rather than assembling further explanatory variables, it is also possible that existing

---

**Fig. 5.** Generalised additive mixed model (GAMM) predictions of *Sphagnum* cover (%) per 4 m² sample unit averaged across peatland (1 km²) in Britain for 2050 based on the UKCP09 scenario with (a) lower, (b) central and (c) upper estimates shown.
datasets used to model the ombrotrophic Sphagnum niche could be combined to produce more sensitive explanatory variables. For example, Bragazza (2008) found that the precipitation:temperature ratio was an effective predictor of the impact of extreme drought on Sphagnum dieback. New response variables could also be created from existing data. For example, the ratio of vascular plant cover to Sphagnum cover could offer a more integrated expression of influential shifts in the balance of plant functional types across mire ecosystems (Bubier et al. 2007, Breeuwer et al. 2008).

4.3. Spatial niche as a predictor of temporal change

Since ombrotrophic peatlands depend upon the persistence and growth of Sphagnum mosses, a key question is how future scenarios of change in climate and pollutant deposition are likely to impact habitat suitability for...
future *Sphagnum* growth. In our example, this depends upon the plausibility of the spatial empirical niche as a generator of hypotheses about changes in time. Problems with this general approach are well known (e.g. Araújo & Rahbek 2005) and include an inability of empirical niche models to simulate rates and forms of temporal change that are impossible to generate from a static regression equation. This drawback is likely to apply to peatlands as much as other ecosystems, since evidence increasingly suggests that peatlands can behave as complex adaptive systems (Dise 2009). They exhibit resilience to gradual environmental change as a result of small and larger-scale stabilising interactions between dominant plant species, but undergo step changes in ecosystem properties when threshold levels of perturbation are exceeded (Belyea & Baird 2006). Stability seems to be achieved via feedback mechanisms between functional types that undergo compensatory shifts in cover in response to changes in abiotic conditions (Ellis 2008, Eppinga et al. 2009). The involvement of dynamic feedback mechanisms and non-linear change as stability thresholds are crossed suggests that static empirical niche models should be limited in their ability to forecast such dynamics. However, simple static niche models should still be able to predict directional changes in the suitability of habitat conditions for *Sphagnum* and therefore provide an early warning that change will occur even if it is suspected that the ecosystem may withstand a certain amount of external forcing prior to step changes in state (Araújo & Rahbek 2005). It is also possible that future climate regimes will be better tracked by climate variables that differ from the ones used here to train the niche models on contemporary patterns.

### 4.4. Need for further uncertainty analysis of model chains

Having explored the benefits and uncertainties of 2 complementary modelling techniques for the ombrotrophic *Sphagnum* niche, further steps are desirable to address uncertainties in the other components of the model chain. A unified uncertainty analysis where sufficient permutations of all parameter sets are jointly simulated across all chained models is a feasible but challenging task (e.g. Page et al. 2004, 2008). This is because the uncertainties in the parameter-rich atmospheric deposition and biogeochemical models are numerous. In our analysis of *Sphagnum* we showed that niche model parameter uncertainty depends on the type of modeling technique selected. However, a more complete evaluation of niche model uncertainty should be based on a comparison with the additional uncertainties associated with prediction of atmospheric deposition and soil biogeochemical processes (Page et al. 2008, 2004).

**Acknowledgements.** We acknowledge the support of DEFRA and NERC in funding parts of this work. The comments of 2 anonymous referees greatly improved a previous version of the paper.

**LITERATURE CITED**

- Fournier N, Pais VA, Sutton MA, Weston KJ, Dragosits U,


Submitted: February 23, 2010; Accepted: November 2, 2010

Proofs received from author(s): December 8, 2010