Final report to Department of the Environment, Food and Rural Affairs, Joint Nature Conservation Committee and English Nature

Atmospheric nitrogen pollution impacts on biodiversity: Phase 1 – Model development and testing (CR0289)

Simon Smart[1], Chris Evans[2], Edwin Rowe[2], Wieger Wamelink[3], Simon Wright[1], Andy Scott[1], David Roy[4], Chris Preston[4], Mark Hill[4], Peter Rothery[4], James Bullock[5], Ibby Moy[5], Bridget Emmett[2], Lindsay Maskell[1]

August 2005

1 Centre for Ecology and Hydrology Lancaster, Library Avenue, Bailrigg, Lancaster LA4 4PD
2 Centre for Ecology and Hydrology Bangor, Deiniol Road, Bangor, Gwynedd LL57 2UP
3 Alterra, Green World Research, PO Box 47, 6700AA, Wageningen, The Netherlands
4 Centre for Ecology and Hydrology Monkswood, Abbots Ripton, Cambs PE28 2LS
5 Centre for Ecology and Hydrology Dorset, Winfrith Technology Centre, Dorset DT2 8ZD

CONTENTS:

APPENDICES

1 Executive Summary

OBJECTIVES

1. Review the current knowledge base for atmospheric nitrogen pollution impacts on biodiversity.
2. Further develop and test modelling techniques to help quantify the impacts of atmospheric nitrogen deposition on biodiversity nationally.
3. Apply the modelling techniques to a sample of habitats and sites to examine current and projected levels of the nitrogen threat (from atmospheric and other sources) to habitats and sites of high nature conservation importance.
4. Provide a preliminary interpretation of the results with respect to achievement of: i) the Public Service Agreement target for achieving favourable conditio n on SSSIs; and ii) Biodiversity Action Plan targets for priority habitats and species and related indicators of biodiversity (eg. Country Biodiversity Strategy indicators).
5. Develop proposals for Phase 2 of his work which should allow for a wider geographical application of the models.

METHODS

1. A program of research was undertaken to develop and test models for predicting the impact of nitrogen deposition on plant species composition and biomass in UK Priority Habitats. The ultimate goal was to produce a modelling capability that could test scenarios of the impact of multiple drivers on Priority Habitat patches in terms of policy relevant indicators especially change in Common Standards Monitoring attributes. The core activity involved developing and testing predictions of habitat suitability for higher and lower plants based on a dynamic soil model (MAGIC) whose outputs were used as inputs to empirical species niche models (GBMOVE).
2. Empirical niche models were constructed for 971 higher plants, 233 bryophytes and 74 lichens. These took the form of a multiple logistic regression equation for each species. Two sets of models were produced: A first generation had mean Ellenberg R, W, N and cover-weighted canopy height as explanatory variables. A second generation additionally included three climate variables (min Jan and max July temperatures and precipitation) however these abiotic+climate models were not tested in this project.
3. The linkage of the species models to outputs from the dynamic soil model MAGIC was based on MAGIC simulations of annual change in soil chemistry on eleven test sites predicted for annual time steps between 1850 and 2050. At each of these test sites, predictions of historical and future atmospheric S and N deposition were used as input to MAGIC. In addition MAGIC was paramaterised using present-day soil chemistry measured in each Priority Habitat.
4. Additional testing focussed on the uncertainty contributed by a series of calibration equations between mean Ellenberg R and N values and soil C/N and soil pH. These equations allow soil model outputs to be input into the GBMOVE species models.
5. The emphasis in this project was primarily on the impacts of N deposition but in order to develop a capacity for testing realistic scenarios of past and future ecological change we sought to develop linked models and filters sensitive to other key drivers that could constrain or exacerbate the effects of atmospheric nitrogen.. Hence, modelling biomass accumulation was based on testing and modifying the Dutch SUMO model. Two separate trait-based filters were also developed. The first was designed to allow ranking of

Common Standards Monitoring (CSM) species by their likely ease of dispersal into a modelled Priority Habitat patch from the local species pool. The second was designed to allow ranking of species by their expected ability to withstand or suffer from grazing by large herbivores. In addition, two further statistical models were developed to predict plant species-richness and probability of occurrence of rare species.
6. A parallel activity sought to complement the detailed modelling work by assessing the extent to which a database of risk factors, each of which could constrain or exacerbate the effects of increased atmospheric nitrogen deposition, could be assembled for all UK ASSI/SSSI. The objective being to scope the potential for a UK-wide risk assessment based on an integration of the information provided by risk factors coupled with empirical CL exceedance for nitrogen.

RESULTS
1. Three sites – Moorhouse Hard Hills, Porton Down and Rothamsted Park Grass - provided time-series long enough for testing predicted change in abundance of plant species against observations.
2. On two sites, weak and very weak yet significant positive correlations were detected between observed and predicted direction and rates of change. These results provide a degree of support for the modelling approach yet particular model components were associated with high levels of uncertainty and sensitivity that weaken the predictive power of the models. For example, the robustness of predicted *rates* of change rather than simply predicted direction of change was not well tested. At Rothamsted, predictions of change in species suitability were the reverse of those observed. This was because observed historical changes in soil C/N, which MAGIC finally reproduced, were actually inconsistent with the known reduction in vegetation productivity and associated shifts in species biomass contributions to the hay crop.
3. GBMOVE species models were partially validated by comparing predicted environmental optima with published Ellenberg numbers for higher plants and a new set of Ellenberg-style pH and fertility indices for bryophytes, constructed during this project. Correlations varied but were all positive. R-squared values ranged from 0.7 (wetness for higher plants) and 0.61 (soil pH for bryophytes) to lows of 0.33 (fertility for bryophytes) and 0.43 (soil pH for higher plants).
4. The low explanatory power of the calibration equations between soil properties and mean Ellenberg scores appeared to be the major influence on poor prediction of species' probabilities. The greatest uncertainty centred on application of the soil C/N versus mean Ellenberg N relationship in fertile (ie. low soil C/N) vegetation types. The problem can be addressed by further work creating within-vegetation type calibrations as well as testing explanatory variables in addition to soil C/N ratio. However this will lengthen the time required to ultimately develop a reliable link between species' response to soil fertility and changes in N availability output from MAGIC.
5. GBMOVE species models require further individual screening and validation. A workable strategy would be to target CSM indicator species for habitat types in which the models performed best; at present heath and bog.
6. In order to model management impacts on the vegetation, the Dutch soil and succession models SMART/SUMO were tested on a number of UK sites and various modifications made to adapt the SUMO model to British conditions. Testing and

validation produced promising results, for example the time series of biomass production in the Rothamsted Park Grass control plots was very well reproduced. Further testing is highly desirable yet limited by available data.

7. The empirical models developed for predicting the probability of occurrence of rare species showed promising results and more work is required to establish how many rare species could be effectively modelled. Testing against observed data is also desirable.

8. The empirical model for predicting above-ground plant species richness was associated with very high uncertainty around predicted values. So much so that the model is unlikely to be a reliable tool for predicting patch richness on designated sites.

9. The trait-based filters also varied in their reliability. The dispersal filter was considered reliable however a number of CSM indicators did not have sufficient trait data to have an index attributed. These need revisiting given the very recent availability of the Europe-wide LEDA database of traits. The grazing index was unsatisfactory and different analytical approaches are required.

10. Application of the MAGIC/GBMOVE and SMART/SUMO models to the prediction of current and future impacts on test sites showed how predictions could be used to estimate the likely impact of N deposition changes on CSM indicator species relative to the impact of management change. Predictions appeared to have greater reliability at peaty, acid and infertile sites ie. with soil C/N ratios greater than about 14. In these habitats MAGIC appeared to calibrate better to measured soil chemistry and GBMOVE estimates of habitat suitability were consistently high for species recorded at each site. These case-study tests should be considered a crude first pass. A more sophisticated capability for scenario testing will follow as models improve and are better integrated.

11. The relevance of the initial model tests to achievement of Habitat Action Plan targets and expected change in UK and country-level indicators of biodiversity was limited by the small number of test sites and the low reliability of model predictions in raised bog (Dead Island Bog), lowland meadow (Dromore Motte and Rothamsted Park Grass) and lowland calcareous grassland (Porton Down) Priority Habitats. Model tests suggested greater reliability in high soil C/N heath and bog habitats. Predictions of the impact of nitrogen deposition to 2010 in upland heath and blanket bog suggested that characteristic dominants would not see marked change in habitat suitability. Although predicted reduction in soil C/N was expected to result in conditions more favourable for negative CSM indicators such as Bracken, *Agrostis stolonifera* and *Deschampsia flexuosa*, the size of these changes was always small. The implication of minor nitrogen deposition impacts on site condition was supported by model test results at Moorhouse. Here, a reasonable degree of correspondence between observed and predicted plant species changes suggested that change in N and S deposition had indeed impacted the plant community between 1973 and 2001. However observed changes were clearly not large enough to have altered the currently favourable condition of the blanket bog unit whose maintenance seemed more obviously linked to appropriately low grazing pressure and long-rotation burning.

12. The final phase of this project assessed the feasibility of assembling a comprehensive database of external risk factors for UK designated sites. The aim was to characterise ASSI/SSSI by values of risk factors that could potentially exacerbate or constrain the responses of species and habitats to N deposition but that would be hard to quantify or measure by local conservation staff in a consistent manner across all sites.

13. The risk factors selected were site area and perimeter to area ratio, agricultural intensification history, empirical N Critical Load exceedance by Priority Habitat, flood risk, current growing season length (GSL) and recent change in GSL, and lastly the extent of intensive versus less intensive semi-natural habitat types around each site.
14. It is currently feasible to assemble a risk database for all factors for all British sites but time-series of agricultural census data and flood risk assessments do not appear to be readily available for Northern Ireland. Generating a standard index of change in agricultural productivity requires further research and input from agronomists.
15. Rather, than create a single integrated index of risk, variation in the values of risk factors would be better used to generate a site classification. This could help stratify and select sites for allocation of limited resources for monitoring and for testing scenarios of potential change using the developing models.
16. Further work is required to produce linked soil and vegetation models that are fit for the purpose of reliably testing scenarios of change on terrestrial Priority Habitats. The results reported here provide a foundation for a long-term campaign of model improvement and building credibility for their practical application.

2 Development and testing of models to predict change in plant species composition and biomass in UK terrestrial Priority Habitats

2.3 Introduction

2.3.1 Atmospheric N as a driver of ecological change

The negative impact of surplus N and P on ecosystems is a relatively recent problem, only emerging as a significant concern over the last 50 years (Dalton & Brand-Hardy 2003). The main drivers of increased N cycling in Europe over this period have been the use of artificial fertilizers, manufactured by fixing unreactive N from the atmosphere, and the burning of fossil fuels (van Egmond et al 2002). At the present time there is evidence that oxidised forms of N from fossil fuel combustion have recently declined, however agricultural emissions remain high (NEGTAP 2001). Inefficient utilisation of artificially fixed N, results in a surplus such that 42% of total artificial N inputs end up emitted to the atmosphere and deposited as reactive NO_y and NH_x or dispersed directly into terrestrial and aquatic ecosystems (van Egmond et al 2002). Since significantly more nutrients are applied than are removed in produce, surplus nutrients enrich soils and waters.

The impact of elevated nutrient loads on semi-natural ecosystems varies. Surplus N and P act as drivers of biological changes that can affect organisms at all trophic levels and can also alter the rates of key processes such as N leaching, carbon storage and litter decomposition (Achermann & Bobbink 2003). Unutilised nutrient surpluses can cause undesirable changes to semi-natural ecosystems previously unaffected by direct agricultural conversion but susceptible to hydrological or atmospheric inputs (Matson et al 1997). Many of the typical species in semi-natural ecosystems are adapted to inherently low nutrient supply and cannot respond rapidly to excess nutrients by greater biomass production. This leaves such species vulnerable to the effects of superior competitors that may have persisted at low abundance in the habitat or could invade from nearby source populations. Shifts in plant community composition and biomass can also have knock-on effects; increasing the susceptibility of characteristic species to frost and drought or attack by pests and pathogens. Resultant changes in characteristic species diversity and composition are often interpreted as reductions in conservation value, yet a major challenge is to attribute past change and predict the amount of future ecological change driven by atmospheric N deposition as opposed to other potential drivers.

Atmospheric N deposition may yield subtle but chronic effects that are difficult to reverse (eg. Strengbom et al 2003; Dupouey et al 2002). Moreover, other drivers such as climate change and reduced S deposition, may interact with N deposition and site management to produce a net pattern of change that is a complex function of their separate effects. Because N deposition effects may be subtle, cumulative yet hard to reverse, attribution of change is challenging but sorely needed. Hence techniques and tools are required that can estimate the contribution and consequences of future atmospheric pollutant deposition alongside other drivers as well as allowing tests of options for remedial land management.

2.3.2 Developing tools for modelling the impact of N deposition on UK Priority Habitats

We present the results of the initial phase of research into the development and testing of models for quantifying the likely impact of changing N deposition on uncommon plants and plant communities present in designated sites across the UK. The analytical approach has been influenced by four considerations.

1. If realistic predictions are to be made then we must consider the effects of other causes of vegetation change in parallel with changes in N deposition. These include site management, the ecological potential for change given local soil conditions and the availability of new colonists in the surrounding area.
2. If predictions are to be applicable to different, possibly new scenarios of pollutant deposition, climate and management, models should be able to express the consequences of synergistic and antagonistic interactions, and hence possible trade-offs, between these drivers. As far as possible this means modelling processes and not just repeating those ecological patterns that reflect current environments.
3. If models are to be applied across many species and habitats, the requirement for measured input data must be kept to a minimum. Data hungry models will be more costly to apply and less likely to be widely used.
4. New environments could lead to new plant communities. This coupled with the requirements of Species Action Plans means that modelling is needed at the individual species level and not at the coarser level of existing plant community units or land-cover types.

We report the results of model development and model testing against independent observations on 11 sites across the UK. Models were developed to predict change in species composition and biomass of key plant functional types. The principal drivers of change on each site were modelled estimates of N and S deposition (Box 1) and either known management regimes for model testing against historical site data, or scenarios of changing future management, such as stopping hay removal or dramatically increasing sheep grazing pressure. The impact of such scenarios on future changes were explored through a series of case-study applications on test sites to generate predicted annual change from the present day to either 2050 or 2100.

In order to be relevant to the demands of current conservation policy, predictions of change on test sites were conveyed in terms of Common Standards Monitoring (CSM) attributes relevant to each Priority Habitat and NVC unit where applicable. Hence, the objective was to predict change in specific CSM indicator species and in plant growth forms also consistent with CSM attributes and targets, for example cover of dwarf shrubs, trees and grass. The ultimate aim is to develop models that estimate the probable consequences of a driver such as N deposition and then allow identification of management strategies, such as changes in grazing regime, that could reduce the impact of the driver on the Priority Habitat under consideration. While the results of the project suggest we are some way off achieving this goal, significant and promising advances have been made. The next section introduces the models developed in this project as a first step toward this objective.

BOX 1 - Modelling atmospheric N deposition – model summary

FRAME Model

The FRAME (Fine Resolution Atmospheric Multi-pollutant Exchange) model is a Lagrangian atmospheric transport model operated by the Centre for Ecology and Hydrology, Edinburgh, to assess the long-term annual mean deposition of acidifying pollutants. The model simulates emission, transport, chemical transformation, and wet and dry deposition, of reduced and oxidised nitrogen and sulphur on a 5 km grid across the United Kingdom. Present day deposition estimates are obtained by calibrating FRAME to measurements of annual wet and dry deposition from the UK Acid Deposition Network, with separate estimates provided in each grid square for deposition to moorland and forest vegetation, reflecting the role of forests in 'filtering' pollutants from the atmosphere. The calibrated model is then re-run using current emissions predictions for 2010, to provide the deposition forecasts used in this study. In MAGIC applications, deposition is assumed to decrease linearly from the present day to this 2010 value, and to remain constant thereafter.

Further details of the FRAME model are available at http://www.frame.ceh.ac.uk/

Historic deposition Sequences

Since MAGIC simulates changes in soil chemistry from pre-industrial conditions, historic sequences of S and N deposition are required to drive the model. The shape of these sequences is standard for all sites, with spatial variations in deposition amount incorporated by scaling the sequence to fit FRAME-estimated present-day deposition for that site. The sulphur deposition sequence was derived from published reconstructions of historic emissions from 1850 through to 1980 (Bettelheim and Littler, 1979; Warren Spring Laboratory, 1987; Simpson et al., 1997). For the period 1980 to present day, deposition data from the longest available UK record at Eskdalemuir, SW Scotland (Hayman et al., 2001) were used to refine deposition sequences. For oxidised and reduced N, sequences were obtained from a recent study undertaken during the NERC GANE thematic programme (Fowler et al., 2004).

Specification of the core models

In order to predict the details of vegetation change over a specified time interval, a series of linked models are required (Fig 1). The main components are a soil model (MAGIC) that mimics the cycling of nutrients in the soil (Box 2) and a succession model (SUMO) that takes nutrients (N and P) out of the soil as plants grow but returns nutrients as plants die (Box 3). The Dutch SUMO model incorporates the effects of management on vegetation growth and successional stage because impacts such as grazing or mowing remove nutrients and also allow more light to be available for the growth of shorter types of plants such as herbs or dwarf shrubs. Both models are called dynamic because they mimic processes that operate over time such as biomass accumulation, decomposition and N mineralisation. Hence, predictions can be made explicitly over a 10, 20 or 100 year interval. A number of modifications to SUMO were implemented and tested to make the model more applicable to British ecosystems.

Prediction of change in individual species relies on a series of multiple regression equations that define the realized niche of each plant species. Regression coefficients were derived for a range of environmental gradients, particular parts of which will favour different species. For example, Cross-leaved heath (*Erica tetralix*) occupies the moderately wet part of the soil moisture gradient, the more acidic end of the soil pH gradient, the low fertility part of the soil fertility gradient and that part of the successional gradient associated with medium canopy height. Collectively these models are called GBMOVE. Changes in species composition in a particular place are modelled by firstly simulating the effect of N and S deposition on the soil. This is done by the MAGIC soil model. This model produces estimates of soil pH and C/N ratio for each yearly time step. Soil pH and soil C/N ratio are then translated into mean Ellenberg R and N values respectively using calibration equations (Smart et al 2003). Since the latter are terms in the GBMOVE regression equations, each equation for each relevant species can be solved at each time step resulting in a changing predicted probability of species occurrence as time passes. The initial set of GBMOVE equations were generated using extensive vegetation survey data representing the range of plant communities found in the UK (Box 4).

A critical component in the model chain are the three calibration equations used to convert soil C/N, pH and % soil moisture into mean unweighted Ellenberg fertility, pH and wetness values (Box 5).

Additional model components

The development of a series of additional empirical models and filters was necessary either to take account of further influences on vegetation species composition or to increase the relevance of the modelling approach to established conservation policy targets for rare species and indicator variables. These additional model components are described briefly below. The reader is referred to Boxes and Appendices for a more detailed account of their development and testing.

Species-richness prediction

A statistical model was developed to predict plant species richness given values of soil C/N, soil pH, % soil moisture and cover-weighted canopy height. Such a model was needed firstly because species richness is a simple indicator variable used to describe

BOX 2- Modelling change in soil properties – model summary (MAGIC)

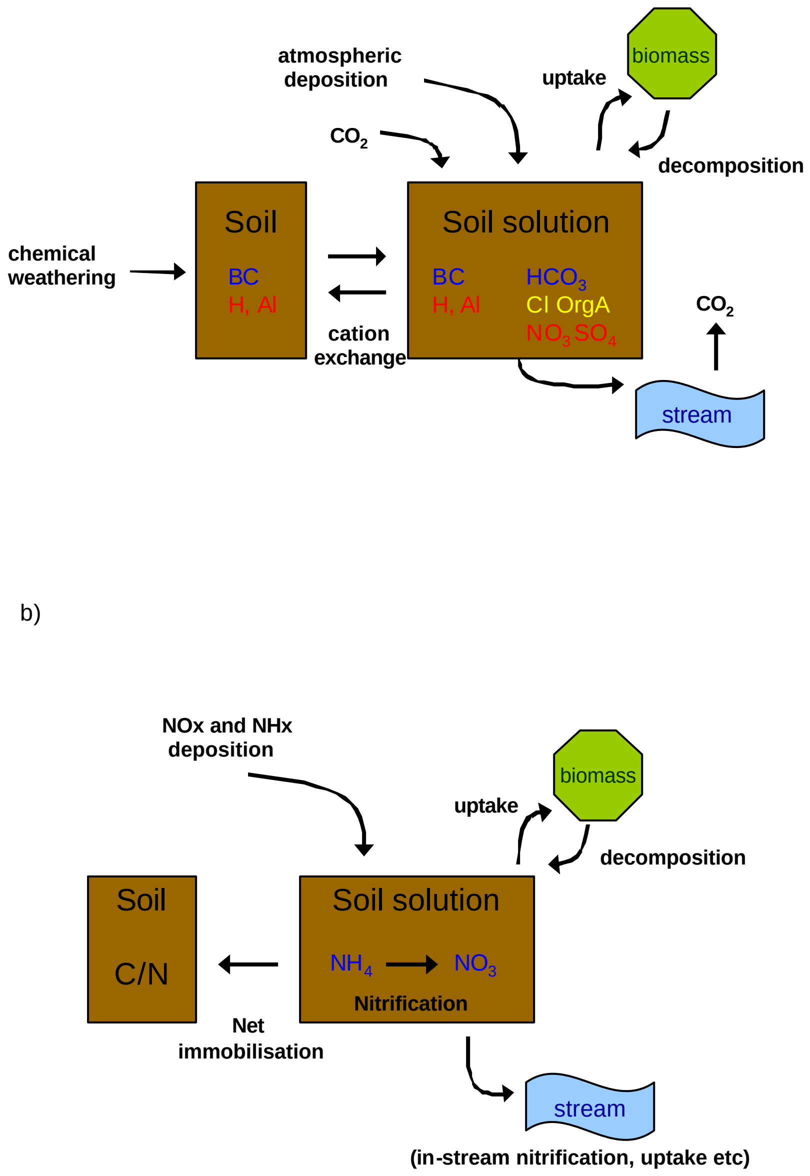
MAGIC (Model of Acidification of Groundwater In Catchments; Cosby et al., 1985; Cosby et al., 2001) simulates changes in soil, soil solution and groundwater chemistry resulting from acid and N deposition and land use. The model can simulate transfers between several different soil, wetland and stream compartments, but in the current project has been used in its simplest form, with one soil and one soil solution compartment. Soil properties are averaged over the soil column, and so the model can be applied to a soil, catchment or region with a comparatively small set of input data. The model has been widely applied and tested, and is one of those recommended by UN-ECE for mapping critical loads.

MAGIC consists of a set of equations describing equilibrium soil processes, a set of mass balance equations describing input-output relationships for base cations and strong acid anions in precipitation and streamwater, and a set of definitions relating the variables in the equilibrium equations to the variables in the mass-balance equations. Key parameters include the input and output fluxes of base cations and strong acid anions, the soil cation exchange capacity, and the fraction of this capacity that is occupied by Ca, Mg, Na and K ions. Nitrogen dynamics are simulated in a simple way, by calculating net retention at each time step according to the current C/N ratio in the soil. Plant uptake, and other sinks and sources, can be included where necessary.

Calibration of the model to a site involves fitting unknown terms, such as soil cation weathering rates and base cation selectivity coefficients, so that they are consistent with the measured soil and soil solution chemistry. The stream concentrations of SO_4 and Cl ions are calibrated first, assuming that transport through the soil is conservative and so output fluxes are equal to input fluxes. Next, N uptake functions are calibrated to match observed stream NO_3 and NH_4 concentrations. Finally, the base cation concentrations are calibrated using an optimisation procedure. The calibrations are performed on simulations run for 140 years to present day, based on historical deposition sequences. After each historical simulation, the model variables are compared to present-day observed data, the adjustable parameters are modified as necessary to improve the fit, and the historical simulation is re-run. The procedure is repeated until no further improvement in the fit is achieved. The resulting calibrated model can be used to forecast future changes in soil pH and N status as a result of changes to deposition and / or land management.

Figure 1. Modelling the acidification process a) and N dynamics b) using MAGIC.

a)



BOX 3 - Modelling biomass growth – model summary (SUMO)

The SUMO (SUccession MOdel) has been developed by Wieger Wamelink at Alterra, Netherlands. It is a process based model that simulates biomass growth under given soil, climate and management conditions. The basis of the model is an exponential growth equation consisting of a series of reduction factors that constrains maximum growth. These factors convey the effect of changes in the availability of light, nitrogen, phosphorous and water. Biomass growth is also a function of temperature and management.

The process that is modelled is competition for light and nutrients by five functional types of plant (climax trees, pioneer trees, shrubs, dwarf shrubs and herbs). The competitive balance between functional types is governed by canopy height and biomass of roots and leaves, which in turn reflect management and initial abiotic conditions. However these conditions change during yearly time steps as a result of the growth and death of functional types or by interventions in the form of changing pollutant deposition, climate or management. Soil dynamics are modelled by a linked soil model called SMART2. This is similar in many ways to the MAGIC model.

SUMO is not formally coupled to plant species niche models either in Britain or in the Netherlands. It's testing and modification as part of this project reflected the fact that SUMO could potentially predict biomass and hence cover of functional types that correspond to structural attributes used in Common Standards Monitoring guidance, for example cover of Dwarf Shrubs and trees.

BOX 4 - Modelling change in habitat suitability for plant species – model summary (GBMOVE).

Multiple logistic regression was used to construct empirical equations that could predict habitat suitability for as many higher and lower plants as possible based on their abundance along key environmental gradients as recorded by extensive botanical quadrat data representative of British plant communities (eg. Roy et al 2000). Each equation consists of regression coefficients that apply to either four explanatory variables or seven if climate variables were included (Table 1). The data used to derive each equation was assembled from a variety of sources so as to maximise the number of plant species covered (Table 2). Each logistic regression was then based on presence/absence data for each plant species in each plot paired with values of each of the explanatory variables. Variable selection was carried out by first testing the explanatory power of each variable separately and then entering those that were significant into a stepwise procedure. The result is an equation that produces a probability of the plant species being present under different sets of conditions specified by the values of the explanatory variables. Modelled changes in environmental conditions are then driven by output from the MAGIC soil model (Box 3), which in turn predicts changes in soil chemistry that follow from changes in pollutant deposition or fertiliser application. Canopy height can be changed arbitrarily using pre-existing knowledge of the pace of succession in a particular location, or on a more process-linked basis, by the SUMO succession model (Box 2). Climate variables could be changed to mimic expectations under different climate change scenarios. Likewise, soil moisture could also be change to mimic drainage or drought.

Table 1. Explanatory variables used in multiple logistic regression equations to define each species realised niche.

Drivers of change to which explanatory variables are responsive	Explanatory variable	Linked by calibration equation to measured..
Atmospheric N deposition, NPK originating directly or indirectly from agriculture	Mean unweighted Ellenberg fertility	Soil C/N ratio
SOx deposition, liming	Mean unweighted Ellenberg pH	Soil pH
Drainage, drought, flooding	Mean unweighted Ellenberg wetness	% soil moisture
Succession and disturbance	Cover-weighted mean canopy height	n/a
Climate change	Minimum January temperature	n/a
"	Maximum July temperature	n/a
"	Precipitation	n/a

Table 2. Datasets and sample numbers used to build GBMOVE models for British higher and lower plants. 1217 NVC quadrats had no grid reference and so were omitted from GBMOVE models that included climate variables.

Dataset	Number of quadrats
Key Habitats 1992	548
Countryside Survey 1998	7221
Broadleaved woods 1971	1648
National Vegetation Classification (various years)	31266

The resulting GBMOVE models constitute an empirical, statistical description of the realized niche of each species. The final number of higher and lower plants having models is shown below (Table 3).

Table 3. Number of species having GBMOVE regression models. The count is based on models with no climate variables. Figures in brackets indicate the number of species that have models but for which no optima and hence no maximum occurrences probability could be calculated (see Appendix 5 and Box 3). Coastal species were defined by Hill et al (2003).

	Bryophytes	Higher plants	Lichens
Coastal		75 (13)	
Non-coastal	233 (72)	971 (182)	74 (28)

BOX 5 – Construction and application of calibration equations between soil measurements and mean Ellenberg values

Calibration equations (Fig 1-3) were constructed to enable soil C/N ratio, soil pH and % soil moisture to be estimated for quadrats in which no soil measurements were recorded. The reason for this is purely pragmatic. Only a subset of quadrats in the database used to build GBMOVE models were associated with soil measurements. However, in order to develop as many individual regression models as possible with maximum information on the environmental preferences of each species, all the quadrat data available ought to have been used to contribute species presence /absence data and estimates of species' abundance along each environmental gradient. To solve this problem, equations were constructed that used mean Ellenberg scores to explain soil measurements. For this step, only the quadrats with soil measurements could be used. GBMOVE regression models were then constructed using all available quadrat data but with mean unweighted Ellenberg values as explanatory variables plus climate variables and mean cover-weighted canopy height. The calibration equations were then used to translate soil C/N and soil pH estimates from MAGIC into values of explanatory variables to solve each GBMOVE equation.

While these calibration equations solve an important problem, they contribute uncertainty related to the fact that soil pH, soil C/N and soil moisture do not explain total variation in mean Ellenberg scores. The greater the scatter about each regression line the more likely it is that predictions of mean Ellenberg values from soil measurements will vary from actual observations. Hence, testing of the calibration equations is a key part of the project. Calibration equations were all constructed using paired soil measurements and mean Ellenberg values from the Countryside Survey 1998 database (Smart et al 2003).

Figure 1. R-sqrd=62%, ln(C/N ratio) = 3.61 – 0.63 ln(mean Ellenberg fertility).

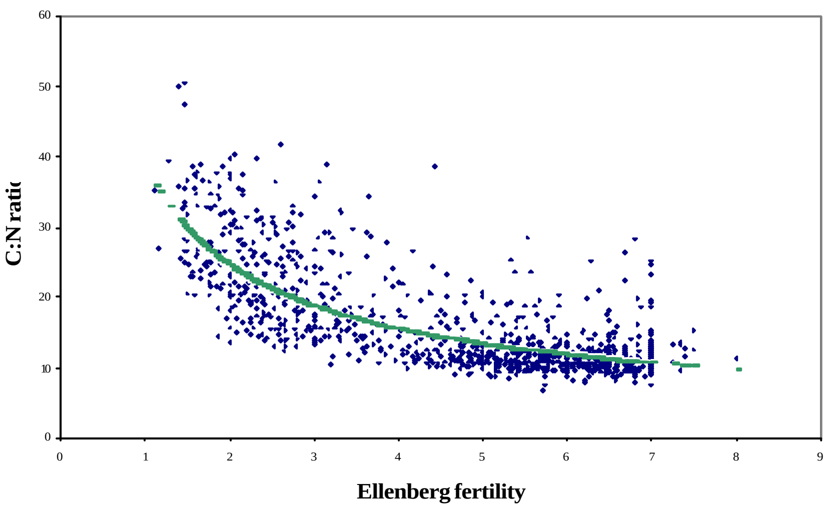


Figure 2. R-sqrd = 72%, ln(M%/100-M%) = 0.55 (mean Ellenberg wetness) – 3.27

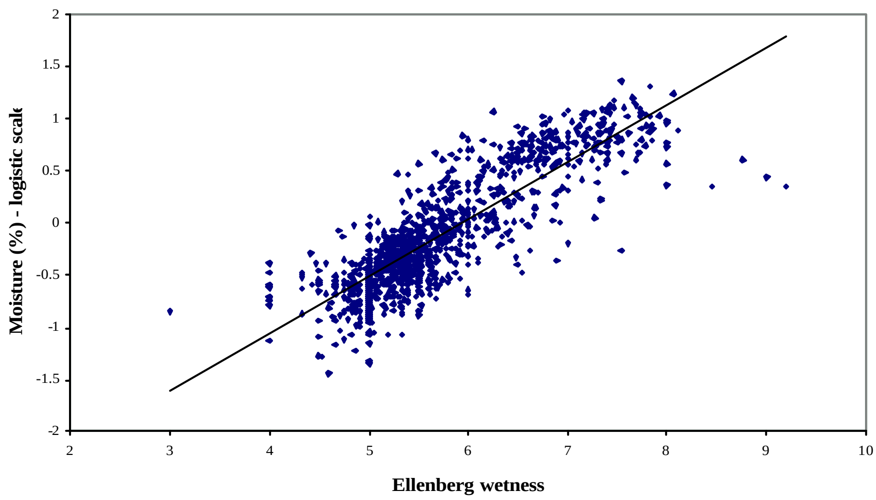
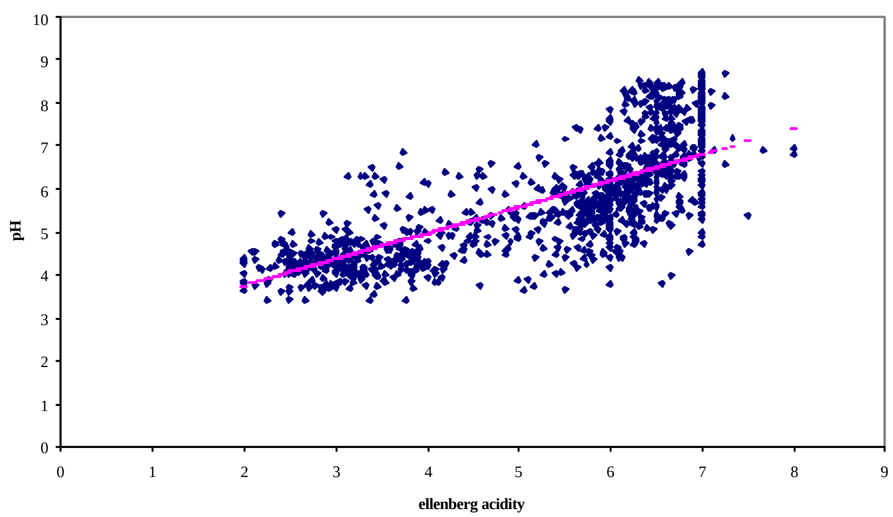


Figure 3. R-sqrd = 61%, soil pH = 2.5 + 0.61 (mean Ellenberg pH)



sectoral changes in habitat types for the UK[1] and England (DEFRA 2003). Also, if species richness could be reliably predicted, this would provide a useful means of populating simulated quadrat data with predicted plant species without having to interpret probabilities output from GBMOVE as expectations of percentage frequency in plots across a modelled site (see Box 6 for a full explanation of this issue). Details of species-richness model development are given in Appendix 1.

Grazing and dispersal filters

Trait-based indices of grazing and dispersal were developed so that predicted species lists from MAGIC+GBMOVE could be filtered to highlight those species most likely to benefit or suffer from changing grazing pressure, as well as those more or less likely to disperse into a target patch of Priority Habitat from the surrounding area. The development of these filters is described in Appendices 2 and 3. The dispersal filter was combined with an estimate of species abundance in the local species pool to derive an empirical guide to the probability of different CSM indicators appearing in a target Priority Habitat patch (Box 7).

Rare and subordinate species prediction

By definition, rare species are recorded in few extant quadrat datasets hence niche models of their environmental preferences could not be devised using GBMOVE methods. A novel technique was therefore developed to estimate their probability of occurrence based on the presence of more common species associated with the rarities in those few quadrats in which they did occur. Development of the method is described in Appendix 4 and a summary and example application for two rare species at Moorhouse NNR is given in Box 8.

2.3.3 Testing the vegetation (GBMOVE) and soil (MAGIC) models

Site selection for model testing

The core activity of the project was to test predictions from the linked soil and vegetation models against observations on a range of test sites. Three main criteria governed the choice of sites. Firstly, site-specific soil data were required for at least one time point to enable the MAGIC soil model to be paramaterised. The following soil variables need to have been measured for the Priority Habitat patch:

- Soil solution; Ca, Mg, Na, K, Al, pH, Dissolved Organic Carbon
- Soil exchangeable; Ca, Mg, Na, K, Cation Exchange Capacity
- Soil pH and soil C/N over as long a time series as possible.

Second, each site should support Priority Habitats associated with a high risk of threat from N deposition as determined by habitat experts. Third, each Priority Habitat on each site should be associated with long-term vegetation monitoring data and if possible but even more uncommon, time-series of soil C/N and pH. Testing of dynamic model

[1] www.sustainable-development.gov.uk/sustainable/quality99/sitemap.htm

BOX 6 - Why do we need to determine species' optima? The effect of commonness and rarity on interpretation of GBMOVE probabilities.

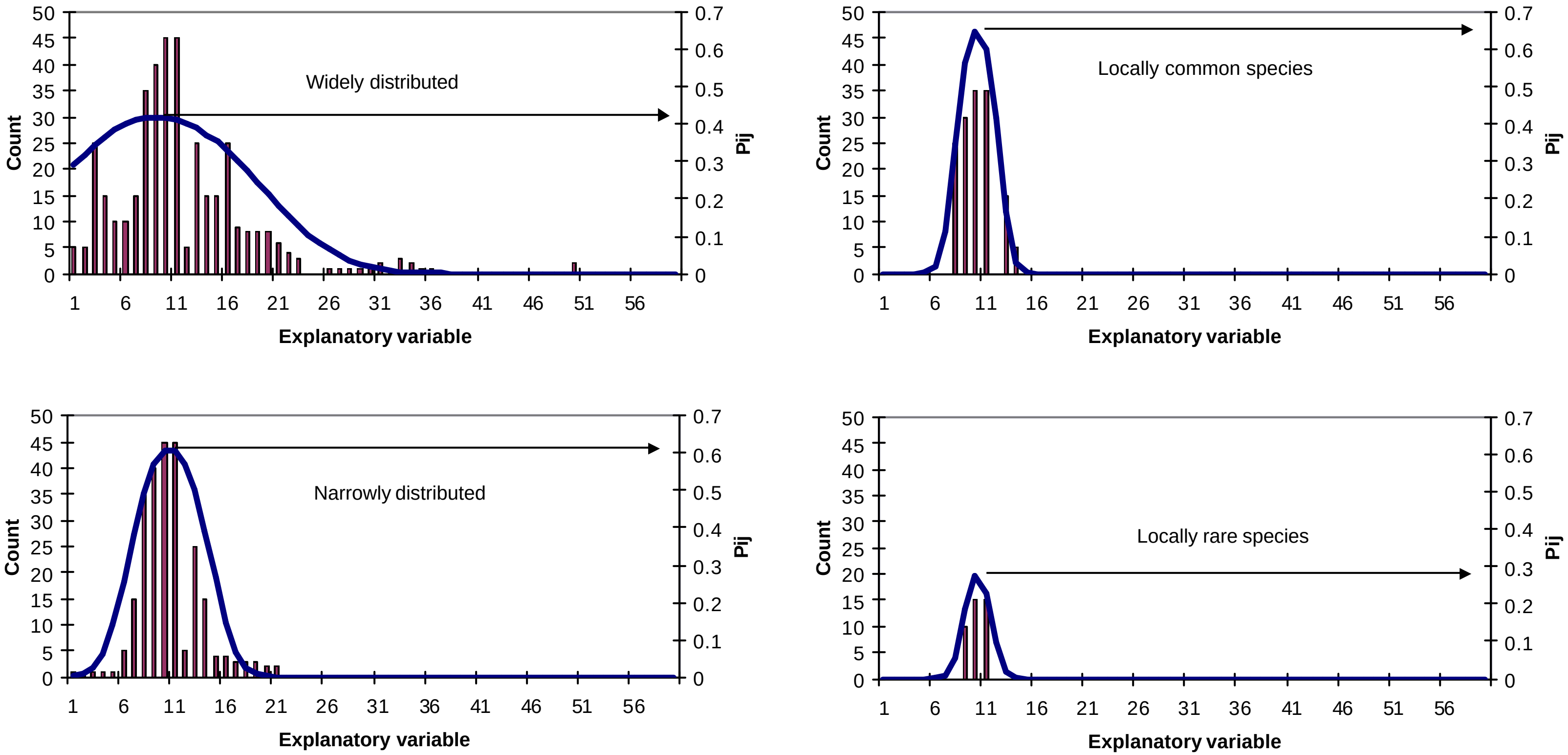
Fig 1 below shows the abundance of four different plant species along a single environmental gradient (the x-axis). The vertical bars indicate how many times a species was recorded in different positions along the entire length of the gradient. The curved line is a logistic regression curve fitted to each set of presence/absence data. The equation for this curve is the basis for each GBMOVE model for each species. However in the actual models, there is not one x-axis but four (or seven if we consider the models with three climate variables added in addition to soil pH, soil moisture, soil C/N ratio and canopy height). The probability of occurrence of a species given values of these explanatory variables is then the solution to each equation. This gives a value between 1 and 0, depicted in the figures below on the right hand y-axis.

The interpretation of these probabilities is critical. Although they represent a probability of occurrence, they do not translate directly into an indication of presence or absence in a specific place. In order to produce generally applicable models, and because of intractability of measurement, many factors which govern occurrence in a patch of habitat are not included as explanatory variables. These factors include accidents of dispersal in the past ie. a species which maybe a dominant may occupy that position because it colonized early on in the history of the vegetation stand, or is very abundant in the local area and so is always 'rescued' from local competitive exclusion by dispersal from nearby populations. On the other hand a species maybe absent from a particular place because of some past biogegraphical process or climatic constraint that makes it rare or absent even though it would be expected to find local abiotic conditions favourable. These subtle factors would be extremely hard to quantify in every patch. If we tried to do so we would end up with a model that was too demanding to parameterise and would only apply in each specific place where these additional factors were measured, assuming we knew what they were in the first place. Hence the cost of having general models that can be applied in most places is that they may not fully explain the species composition in any specific place. Therefore the maximum occurrence of a species is likely to be an average that conceals a much higher peak if only we had another explanatory variable that would solve the riddle of why an apparently rare species – such as that in the lower right graph – actually occurs on some sites and not others. The practical issue is that for any specific patch, the estimated maximum probability at the species abiotic optimum (the point where the horizontal arrows meet the right hand y-axis) will often turn out to be an over or under-estimate. The probability becomes more reliable the more patches we sample, but can be very unreliable if we use it as a prediction of occurrence in a single stand. The implication for development of a modelling capability for specific Priority Habitats on specific sites is that our general models may not accurately predict the species composition in a single patch. What we desire is a way of expressing the predicted probabilities in a way that factors out these unexplained differences in commonness and rarity while retaining the response curve and the information it gives us about the abiotic conditions that should favour each species. A solution to this problem is to rescale the predicted probabilities by the maximum probability. Thus in each of the cases below if abiotic conditions coincide with the peak of each curve the rescaled probabilities will all be 1 rather than all having different values.

This approach requires an important difference in the way modelled probabilities are interpreted. In essence, the rescaled values convey an estimate of habitat suitability *if the species was present in the patch or could disperse into it and establish*, rather than a prediction that the species is or is not present. Although this would seem to restrict the usefulness of the modelled probabilities, it all depends on the question being asked. In the case where we wish to predict change among species we know are present, we can still use predicted trends in the rescaled probabilities as a guide to expected changes in abundance over time. Moreover, given a set of other indicators that are known to be present nearby and could disperse into the monitored stand, we could, also usefully interpret their predicted changes in habitat suitability even though they are not present. Given their likelihood of reaching the stand, modelled changes in habitat suitability measure changes in the risk of a negative indicator being able to persist if it managed to colonize from a nearby source.

The rationale described above is exemplified in the case studies we have presented where MAGIC+GBMOVE has been used to explicitly target the fate of named CSM indicators for each Priority Habitat. Thus the key question is not which species from the entire flora or the entire local species pool could grow in this place. Rather the key questions are; what is likely to happen to the indicator species known to already be present given a particular deposition and management scenario, and secondly, how will expected changes in habitat suitability affect the chances of other indicators persisting if they could reach the target patch.

Figure 1. Species maximum probabilities influenced by differences in the abundance of a plant species in the training data used to construct GBMOVE regression models.



BOX 7 - Estimating immigration potential based on dispersal traits, species' broad habitat preferences and local broad habitat composition

STEPS:

1. Determine 10km^2 species pool using BRC data or use site species lists where available.
2. Determine broad habitat extent in site (s), in 1500m buffer (b1) around site and in further 1500m buffer (b2). This step uses LCM 2000 and therefore assumes it is accurate at least in the identity and rank abundance of Broad Habitats.
3. Derive an approximate abundance weighting for each species in each zone (s, b1 and b2) using its preference index for each broad habitat as published by Hill et al (2003) in combination with the abundance of each broad habitat ie. high indices will reflect a large extent of the broad habitat with which each species is most associated. The index is worked out for species *j* as the sum of the products of multiplying each broad habitat proportion (a value between 0 and 1) by the preference index for the species and the broad habitat. This gives a maximum value of 4. So the index is rescaled by dividing this sum by 4.
4. Multiply the pool abundance index from 3 by the species' dispersal index.
5. Sort the CSM indicator table in descending order of the index in 4.
6. Interpret the table on the assumption that each component of the above index is reliable and realistic at the particular site, and that the dispersal ranking is an accurate reflection of real dispersal potential if appropriate vectors are in place.
7. Hence, species with high indices are expected to be most likely to disperse into the monitored patch. Establishment is then hypothesised to be favoured by increased habitat suitability, measured by change in Pi/Pmax (described in Box 6).

<u>Diagram of index construction for Purple Moor-grass at Budworth common</u>

Broad Habitat composition in site with preference indices in brackets

Bog 3% (4)
Dwarf Shrub Heath 80% (2)
Acid grassland 5% (1)

Sum of proportions * preference

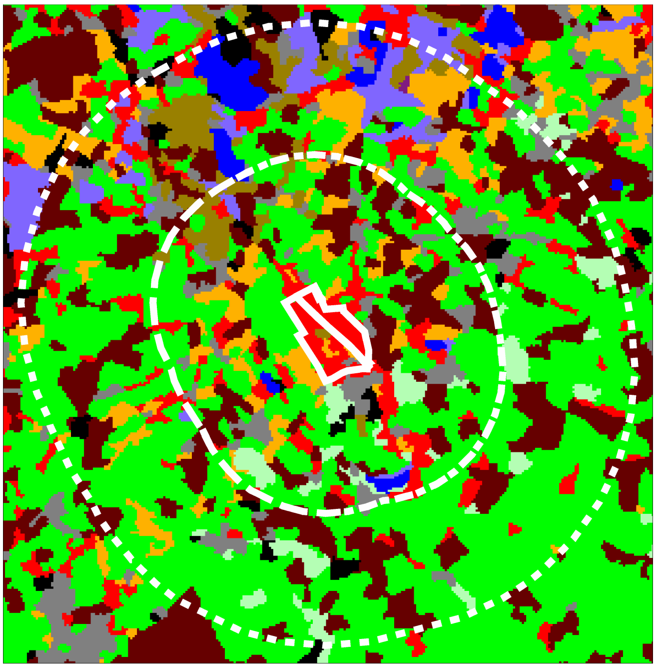
= (0.03 x 4) + (0.8 x 2) + (0.05 x 1)

= 1.77

Index = 1.77/ 4 = 0.44

Then multiply by dispersal index (high = more easily dispersed)

0.44 x 0.43 = **0.19**



Repeat for all CSM indicators NOT recorded in the monitored vegetation, then order table by index in descending order.

BOX 8 – Development and application of a method for predicting changes in probability of occurrence of rare and subordinate species

The logistic regression techniques used to develop the GBMOVE models cannot be readily applied to the problem of predicting rare species because there are usually too few samples available in which species have been recorded. In addition the reasons for species rarity may have more to do with factors other than those included as GBMOVE explanatory variables. This must be so because many rare species are absent from apparently favourable patches (eg. Piggot & Walters 1954). This observation suggests that application of rare species models is only likely to be useful in places where a species population is present or where there is reason to believe it could establish following re-introduction or invasion if the rare species is a non-native casual for example.

A method for predicting probability of occurrence of rare species was developed for this project and is fully described in Appendix 4. Rare and subordinate species were selected based on their scarcity in the training datasets used to build the GBMOVE models (Figs 1 and 2). 13 of these species were selected for model development. The method uses the pattern of association between rare species and other more common neighbouring plants to define a mix of species with which the rarity tends mostly to be found. The method assumes that the other common associates are to some extent indicators of appropriate ecological conditions but this is not explicit in the models; in essence the only explanatory variables are neighbouring species. The accuracy with which a rare species' optimum floristic context is modelled depends greatly on the availability of quadrat data in which each rare species has been recorded. Thus *Trifolium ochroleucon* and *Gentiana pneumonanthe*, only occurred in one plot each in the training data, so that the applicability of the estimated suitable mixture of associates maybe severely limited if its patterns of association are more varied. Although models were constructed for these species they may convey a narrower pattern of optimum species associations than actually occurs.

The method uses a species list for a quadrat as its only input. Therefore applying the method to predictions of change in habitat suitability generated by MAGIC+GBMOVE, will require habitat suitability index changes to be applied as weightings to initial species lists thus generating simulated changes in the composition of species lists over time. Further development is required to test options for applying habitat suitability indices in this way. The rare species model could still be readily applied to monitoring data without this further development work.

Figure 1. Histogram of all higher plant species abundance in the training dataset (n=43000) used to build GBMOVE models (see Box 1).

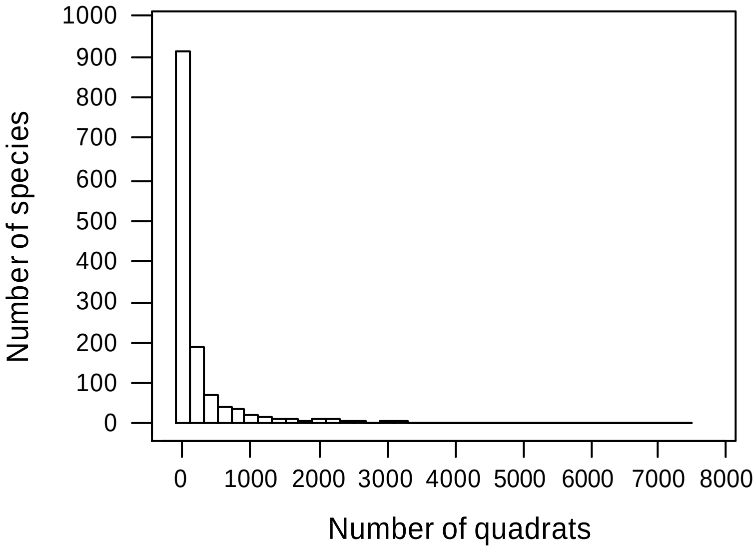
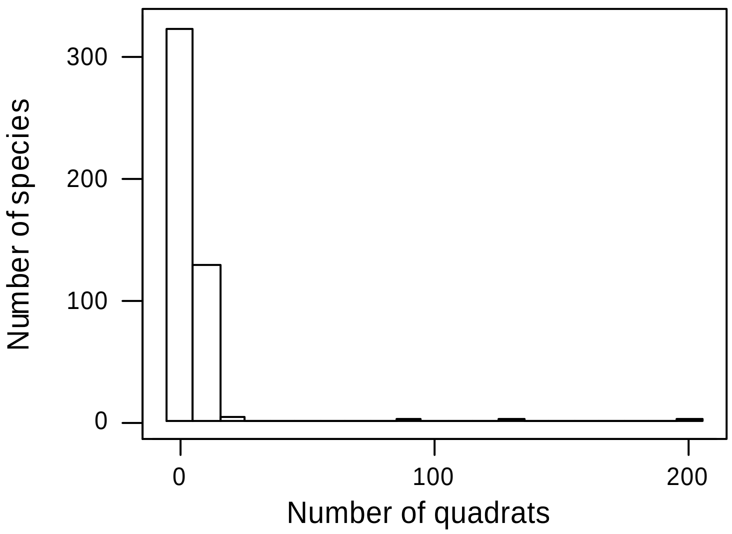


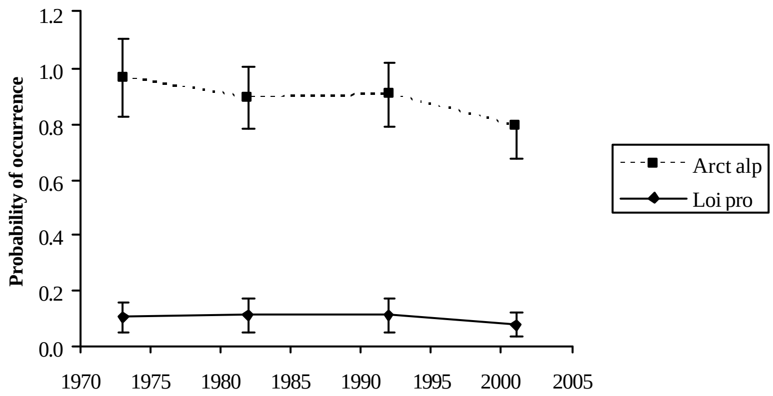
Figure 2. Histogram of rare higher plant species abundance in the training dataset used to build GBMOVE models.



Example application

To illustrate model application, changes in probability of occurrence for two rare species was calculated from the species composition of the Hard Hills control plots at Moorhouse. Although both species are associated with upland heath and blanket bog, neither have been recorded at Moorhouse.

Figure 3. Change in the mean probability of occurrence (+/- se, n=9) of two rare species, *Arctostaphylos alpinus* and *Loisleuria procumbens*, in the nine control plots of the Moorhouse Hard Hills burning and grazing experiment between 1973 and 2001.



The high probability generated for *A.alpinus* reflects the fact that the species present in the control plots form an assemblage typically found in places where the species occurs. For both species, their rarity means they will generally be absent even if an apparently favourable assemblage is present. This again emphasises that model application is only likely to be of interest as an aid to monitoring on sites where a rare species occurs or on potentially favourable recipient sites in the region in which the species occurs.

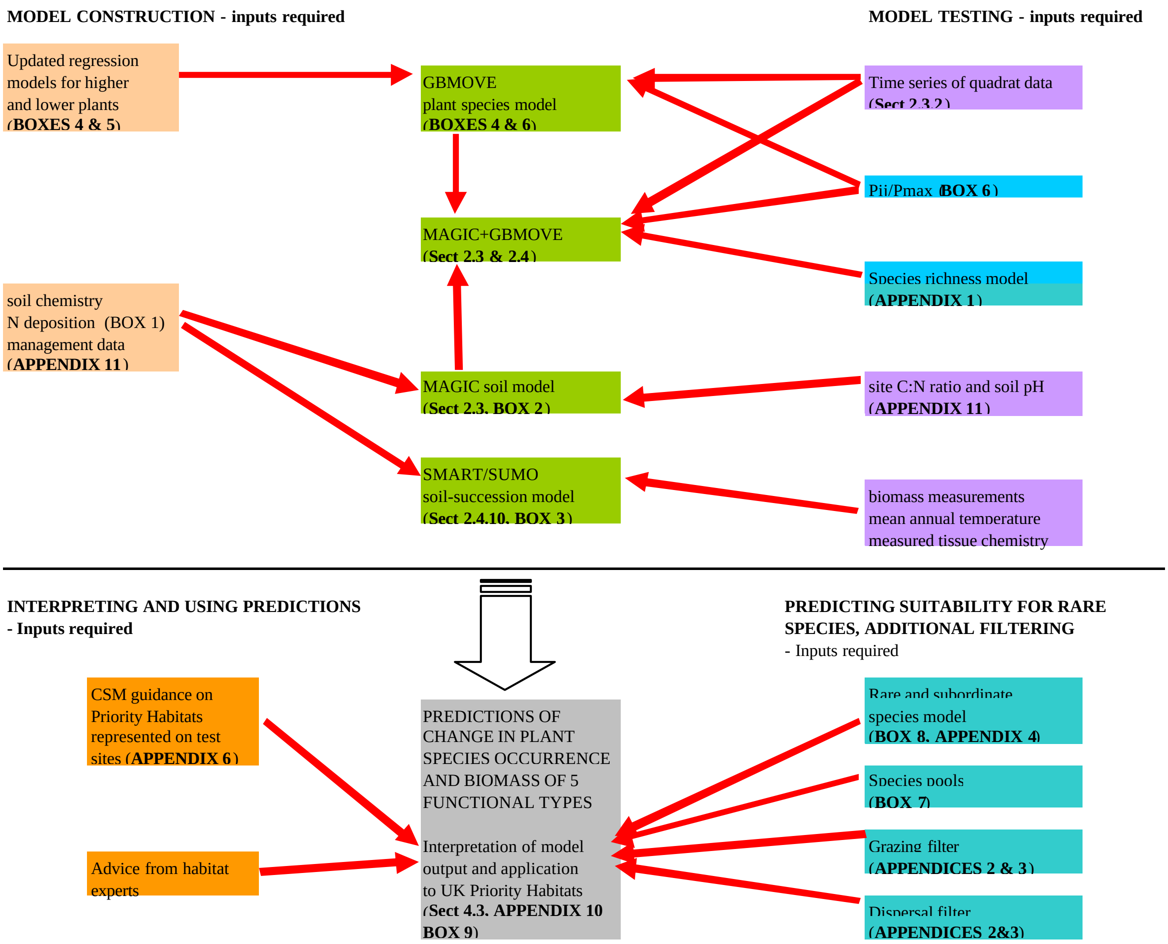
Further developments

The rare species modelling approach appears promising. The next step ought to be publication to secure scientific credibility. Generating models for the remaining 447 species would be a relatively straightforward task but there is no guarantee of model convergence for each of these species. In addition a further search is recommended for presence data for those species with very few records in the current training dataset.

Further technical modifications could see climate variables or distributional data included so that predictions are sensitive to the localised distributions of rare species. However, in reality most users are only likely to use rare species models in a general cross-site application when the focus is on assessing indirect changes in habitat suitability for native or non-native species likely to be undergoing range expansion and infilling.

components was carried out on the eleven sites identified in table 1. All but two are non-woodland Priority Habitats in which atmospheric N deposition was highlighted as a significant obstacle to progress in the 2002 Biodiversity Action Plan (BAP) progress report[2]. The importance of including management impacts within each site and habitat patch, in addition to N deposition, is borne out by the fact that other drivers of habitat loss and degradation were rated as the most important obstacles to Habitat and Species Action Plan[2] delivery in the 2002 progress report.

Fig 1. Road map to the project - model components and dependencies on input data for model testing and application. The reader is referred to relevant Boxes, Appendices and sections covering treatment of each component.



[2] See www.ukbap.org.uk/asp/2002_main.asp

Table 1. Experimental and ECN sites supporting those UK Priority Habitats for which atmospheric N deposition was highlighted as a factor causing loss or decline in the 2002 BAP progress report.

Priority Habitat	**Site name and type**	**ASSI/ SSSI**	**Rank for N deposition threat (1=highest risk)**	**Highest threat to PH (italics indicate that impacts can be modelled)**
Blanket bog	Moorhouse ECN site	Y	5	*Overgrazing*
Upland heath	Climoor experimental site	Y	2	*Overgrazing* / fire
Upland heath	Ruabon experimental site	N	2	“
Upland heath	Cairngorm ECN site	Y	2	“
Not PH (Upland acid grassland)	Plynlimon experimental site	N	n/a	n/a
Not PH (Upland acid grassland)	Pwllpeiran experimental site	N	n/a	n/a
Lowland heath	Budworth Common experimental site	Y	3	Invasive aliens / *scrub encroachment*
Lowland meadows	Dromore Motte (NI)	Y	5	*Agricultural intensification*
Lowland meadows	Rothamsted Park Grass	N	5	“
Lowland raised bog	Dead Island Bog (NI)	Y	4	Peat extraction / *drainage*
Lowland calcareous grassland	Porton Down ECN site	Y	-	*Neglect*

The top-ranking threats to each Priority Habitat are also listed in table 1 with an indication of whether their impacts can be modelled within the proposed work programme. Because, sites with both vegetation data and the soil chemistry data required to run MAGIC and SMART were so scarce, we added two non-Priority Habitat upland acid grassland sites. They provide additional opportunities for model testing, although in conservation terms they do not represent target community types.

Table 2. Time series of vegetation monitoring data available for testing model predictions of change in time.

Site	Years with vegetation data available
Moorhouse Hard Hills control plots	1973, 1982, 1992, 2001
Cairngorm ECN site	1998, 2002
Porton Down	ECN (fine-grained annual) 1997-2000 T.C.E. Wells baseline 1975 ECN baseline 1991 ECN coarse 1994
Ruabon experiment	Pre-burn 1995-1999 Post-burn 2001-2004
Rothamsted Park Grass control plots	Plot 3; 1914, 1919, 1921-26, 1936-40, 1947-48, 1975-76 Plot 2; 1862, 1867, 1872, 1877, 1914, 1919, 1949 Plot 12; 1862, 1867, 1872, 1877, 1914, 1919, 1949

Time series differences

The signal of N deposition is expected to be of low magnitude in terrestrial vegetation (Smart et al 2004). Hence, long time series of vegetation data stand the best chance of capturing the chronic, low level effect of often small resultant changes in soil C/N ratio and soil pH. Of the test sites selected, Rothamsted, Moorhouse and Porton Down provided the greatest scope for long-term signal detection (Table 2). Other test sites with short or single observations were still utilised but the test was restricted to comparisons of observed versus predicted species composition, species richness and soil variables at specific time points rather than testing for consistent trends over time.

The effect of differences in plot size

Differences in the size and numbers of plots from which species composition was recorded, are also liable to affect the match between observed and predicted variables. Large plots will tend to be more species rich while large plots are also likely to include species at low abundance that are less typical of the vegetation type and hence have atypical Ellenberg numbers that could be influential on unweighted means. The training datasets used to build the GBMOVE models, generally employed large quadrats that ranged from 4 to 200m^2 while most plots on test sites were relatively small (Table 3). Where possible, change in percentage frequency between replicates of the largest plot size were used as the response variable for comparison with predictions. Numbers of test plots also varied considerably from site to site

(Table 3). Because sample size affects standard errors but not standard deviation, we plot the latter as a measure of variation in observed data where possible.

Observed vegetation time-series for model testing

At Rothamsted, predicted trends were compared with 100 years of observations. However, the response variable on this site was biomass recorded in the annual hay crop. A large body of published evidence has demonstrated the operation of other constraints on biomass trends at this site including rainfall (Silvertown et al 1994), trait-based predisposition to external factors (Dodd et al 1995), species richness (Dodd et al 1994a) and other community-level phenomena grouped under the term assembly rules (Wilson et al 1996, Silvertown 1987). Since we have no way of co-varying out these additional effects, they are likely to contribute to unexplained variation. Conversely, Hill & Carey (1997) found that mean abundance weighted Ellenberg fertility values explained over 80% of the annual variation in biomass. However, changes in more productive species' biomass between years could still covary with rainfall since wet and warm Summers tend to favour more competitive species (Dunnett et al 1998). Hence, the eutrophication signal carried by observed mean Ellenberg fertility values may actually be driven by a complex of factors. A consequence is that the explanatory power of N and S deposition effects may be low even though attribution of a fraction of the ecological signal is possible.

At Porton Down, the first series of observations were taken from the published transect data of Wells et al (1976). We used the published location maps and also more detailed copies, kindly provided by Kevin Walker (CEH Monkswood), to locate a transect closest to the ECN baseline and coarse sampling plots with an additional criteria that all plots should have sampled the same vegetation type. Despite these efforts to ensure locational and floristic comparability, the original Wells plots may have been up to 200m from the ECN samples with which they were paired in the tests presented here. Hence, we cannot rule out differences in species abundance being confounded with plot location error.

At Moorhouse, observations were taken from the unburnt and ungrazed control plots of the Hard Hills experiment established in 1954. See Marrs et al (1989) for further details.

Table 3. Numbers of plots and plot sizes for test data sets.

Site & year	N	Size
Rothamsted Park Grass	1-3	c.100m^2
Porton Down '75 to '94	12	4 m^2
Cairngorm ECN baseline and coarse sampling	17	4 m^2
Budworth	4	0.25m^2
Pwllpeiran	24	4 x 1.2 m^2
Climoor	9	0.5 m^2
Moorhouse, Hard Hills	3	100m^2
Plynlimon	15	3 x 0.5 m^2
Dromore Motte	1	4m^2
Dead Island Bog	1	4m^2
Ruabon	15	1m^2

Testing strategy – key questions and their relevance to users

Predictions are based on, at most, five linked models, FRAME/GANE predictions of N and S deposition + soil models (MAGIC) + calibration equations + species richness model + species probability models (GBMOVE). Each model contributes uncertainty to the final prediction. The testing strategy aimed to estimate the uncertainty attributable to each model component by testing them separately and then when linked together. Such comparative tests could then suggest components responsible for contributing the greatest lack of fit to observed data. For example, the uncertainty contributed by the calibration equations (Box 5) was assessed by comparing observed mean Ellenberg values of N and R for test plots against the Ellenberg values predicted from measured soil data. At Rothamsted it was also possible to test the dynamic soil models MAGIC and SMART (the soil model that drives SUMO) by comparing observed soil C/N and pH at three time points with model predictions.

Two additional tests were carried out using other datasets. A test of the calibration equation for soil pH versus mean Ellenberg R was carried out by comparing observed soil pH with predictions based on the species composition of 244 quadrats recorded from English Environmentally Sensitive Areas (Critchley et al 2001). Also, the accuracy of the GBMOVE regression models was tested by comparing published Ellenberg numbers for higher plants and a newly generated series of numbers for lower plants against the mean Ellenberg numbers calculated at the environmental location where each species' probability of occurrence was maximum.

The initial matrix of model testing questions by model components is shown below (Table 4). Model test results are also considered in terms of how varying performance is likely to affect their application to assessment of species and habitat targets and indicators; in particular in contributing to prediction and scenario testing of changes in CSM attributes and the achievement or maintenance of changing condition on SSSI.

Table 4. Testing strategy based on comparing outputs of combinations of model components with observations on test sites. Pij/Pmax refers to the rescaling of a predicted probability for a species from GBMOVE by the maximum probability estimated at its niche optimum (see Box 6 and Appendix 5).

		MODEL MODIFICATIONS AND FILTERS				
MODELS	**QUESTION**	**Pij/Pmax**	**Species richness model**	**Grazing filter**	**Dispersal filter**	**Selection from species pools**
1a. GB_MOVE	Are species present in observed plots predicted to be present at times 1..n?	Y	Use observed richness	N	N	Y
2a. MAGIC + GB_MOVE	Are species present in observed plots predicted to be present at times 1..n?	Y	Y	N	N	Y
2b. MAGIC + GB_MOVE	Are predicted directions of change in intial species consistent with observed data?	Y	Not applicable as focus is on species present at time 1			
2c. MAGIC + GB_MOVE	Are predicted patterns of species compositional change consistent with observed data in terms of NVC match? and mean Ellenberg scores?	Y	Y	N	N	Y
3. SMART+ SUMO	Are predictions of biomass accumulation of five plant functional types accurately predicted.		Not applicable since all these filters operate on species composition not biomass.			
4. MAGIC	Are predictions of soil C:N and pH consistent with obesrvations and prediction from observed species composition?		Not applicable as focus is on soil variables.			
5. SPECIES RICHNESS	Is predicted species richness from soil measurements, species composition and MAGIC consistent with observations?		Not applicable as test is against observed species richness at time 1..n.			

2.4 Model testing results

2.4.1 Uncertainty and reliability of N and S deposition estimates

Sources of uncertainty surrounding N deposition estimates are well known. Large amounts of sub-grid square variation are attributable to topography and the location of high emitting agricultural sources such as pig and poultry units (Sutton et al 1993). In addition, difficulties in the measurement of dry deposition relate to uncertainty in the estimation of deposition velocities to different vegetation surfaces (Smith et al. 2000). While the importance of dry versus wet deposition varies across Britain (NEGTAP 2001), difficulties in measuring dry deposition mean that time-series of accurate observations usually do not exist. This makes assessment of the contribution of uncertainty in the deposition model estimates impossible to test rigorously for our sites. An important consequence is that modelled predictions can indicate empirical Critical Load exceedance for N whereas the real deposition may indicate otherwise or presumably *vice versa*. Again though, the difficulty of measuring dry deposition makes validation difficult. At Rothamsted, Goulding et al (1998) gave the measured deposition history in precipitation for Rothamsted (Figs 2,3), which indicates a maximum total wet N deposition of 16.5 kg N ha yr in 1980 falling to 8 kg N ha yr in 1995. However, the FRAME/GANE prediction is for total N deposition of 24.7 in 1980 and 23.4 in 2002. Therefore the model predicts empirical critical load exceedance (20-30 kg N ha yr for low and medium altitude hay meadows; Achermann & Bobbink 2003) for at least 20 years whereas wet deposition measurements indicate no exceedance in the same period. However, a problem with this comparison arises because the Rothamsted observations do not take account of dry deposited N hence this missing component could bring real deposition much closer to the model prediction. CEH Edinburgh provided model estimates of dry deposition at Rothamsted, which were added to measured data for the site to produce the best possible deposition history. Thus in model testing at Rothamsted, observed data were used (red lines in Figs 2 and 3) plus estimated dry deposition. This requirement was particularly important at this site because initial modelled deposition coupled with hay N content and biomass resulted in more N being removed from the system than was deposited, hence the soil was predicted to be much more fertile at the start of the experiment than was observed. Using site measurements of deposition substantially improved MAGIC performance (see section 2.11).

Figure 2. Oxidised N deposition at Rothamsted based on FRAME/GANE modelled estimates and actual site measurements of wet, oxidised N deposition.

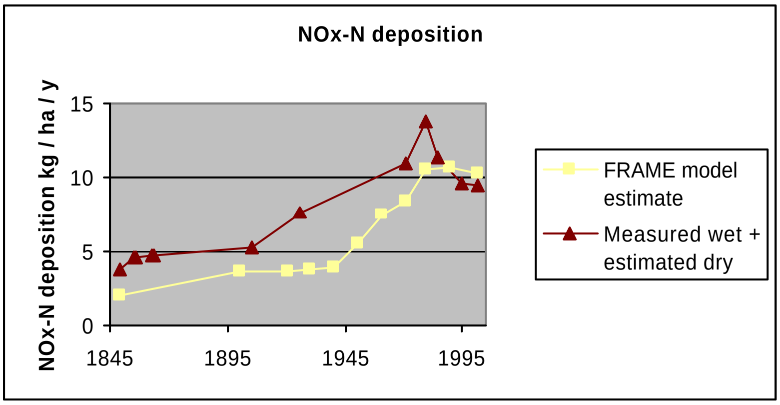
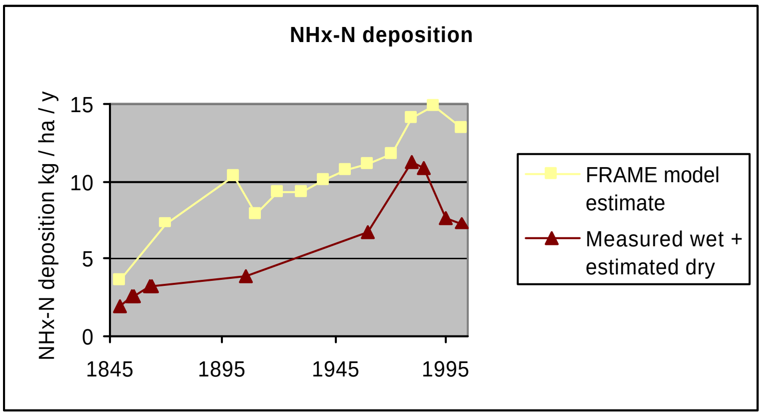


Figure 3. Reduced N deposition at Rothamsted based on FRAME/GANE modelled estimates and site measurements of wet, reduced N deposition.



In the absence of measured time series for test sites other than Rothamsted, we have used model predictions of wet and dry deposition to test linked soil and vegetation models and to construct predictions of species compositional change. The FRAME model and sources of historical deposition sequences are described in Box 1. Crucially, our tests hinge on the assumption that real deposition history corresponds sufficiently with modelled change that differences can be ruled out as an underlying reason for mismatches between observed and model driven changes in soil properties and species composition. Thus, where predicted species changes correspond well with those observed, there is a good chance that we will have detected a signal of N and S deposition impact as well as having tested the developing models.

Conclusion: While much faith is placed in modelled predictions of N deposition in other policy applications, the predictions are prone to error in both magnitude and direction of change. However validating hindcasts at grid or sub-grid scales is usually impossible because of an absence of corresponding observations.

2.4.2 Validation and uncertainty of MAGIC simulations

<u>Modelling acidity</u>

The MAGIC model was first developed during the 1980s (Cosby et al., 1985), and has been revised and updated on several occasions since, most recently to include a more detailed representation of N dynamics (Cosby et al., 2001). The model has been widely applied to simulate soil and surface water chemistry in forests and semi-natural ecosystems across much of Europe and North America. The model is normally calibrated to present-day measurements of soil, surface water and/or soil water chemistry, and the accuracy of the model simulation can be tested against long-term monitoring records, where available. In general, surface water (rather than soil or soil-water) records are most widely obtainable, and have the longest duration. A range of previous studies have demonstrated a generally good fit between MAGIC-simulated and observed water chemistry for acidity-related variables including pH and Acid Neutralising Capacity (ANC). These include model applications to sites which have undergone major chemical changes in central Europe (Hruska et al., 2002; Majer et al., 2003; Kopacek et al., 2003), lakes in Norway (Wright and Cosby, 2003), and upland lakes and streams in the UK Acid Waters Monitoring Network (Jenkins and Cullen, 2001). It has also proved successful in reproducing observed recovery from acidification at small-catchment 'clean rain' experiments in Norway and Sweden (Beier et al., 1995). Fewer long-term data are available to validate soil chemistry simulations, but given the close linkage between soil solution and surface water chemistry, these observations suggest that the model should generally give an effective simulation of soil solution pH.

A detailed uncertainty analysis of MAGIC was undertaken by Page et al. (2003), based on a model application to the Afon Gwy, a moorland stream in mid-Wales. A Monte Carlo approach was used to generate a large set of simulations for which a set of simulated soil and stream chemical variables were within a specified range of observed values. They concluded that the greatest sources of uncertainty were the initial conditions of the simulation (weathering rates and initial base saturation), and the shape of the deposition sequences. However, since they observed that pH simulations were 'quasi-parallel' once the initial conditions were set, and since initial conditions are normally calibrated when applying MAGIC, their findings suggest that if the model can be accurately calibrated to reproduce observed pH, then the range of possible pH forecasts should be fairly narrow. Larssen et al. (2004) obtained similar results from an uncertainty assessment for a Norwegian stream with a long monitoring record, and showed also that prediction uncertainty could be further reduced by constraining the calibration to fit observations from more than one period in the monitoring record.

Overall, assessments of model performance in simulating acidity suggest that the greatest uncertainties are associated with the input data, rather than the model structure itself. One important source of input data uncertainty appears to be the historic deposition sequences. Model calibration to observations is essential, and the quality of the model simulation is thus highly dependent on the quality of the input and calibration data used.

Modelling nitrogen

The treatment of N immobilisation and leaching within MAGIC is based on simple principles. Available N is calculated from additions of ammonium and nitrate in fertiliser or deposition, after subtraction of denitrification and plant uptake (Cosby et al, 2001). As long as the C/N ratio in soil organic matter is above an upper threshold, available N is subject to immobilisation and thus decreases the C/N ratio of the soil organic matter. If C/N is below a lower threshold, there is no immobilisation. Between the upper and lower thresholds the percentage of available N that is immobilised declines linearly from 100 % to 0 % (Fig 4). The N that is not immobilised is assumed to be leached.

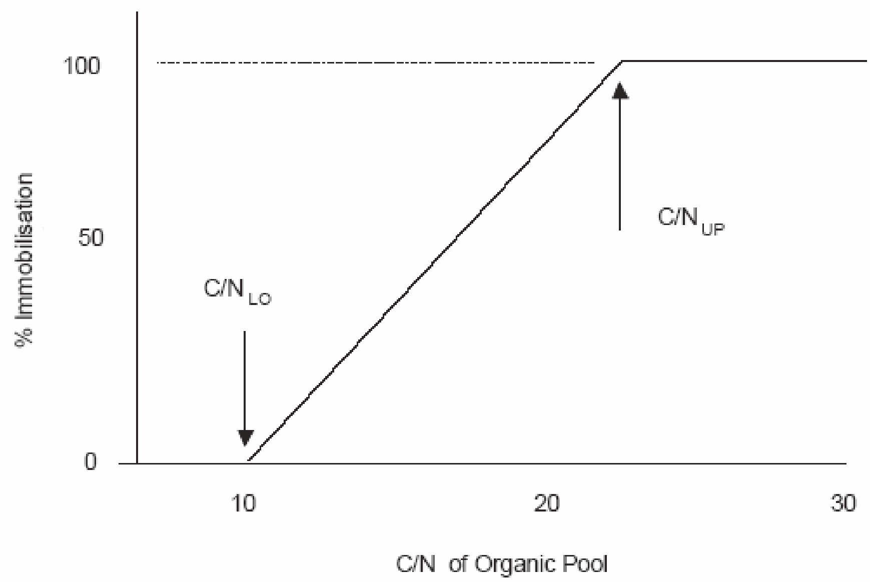
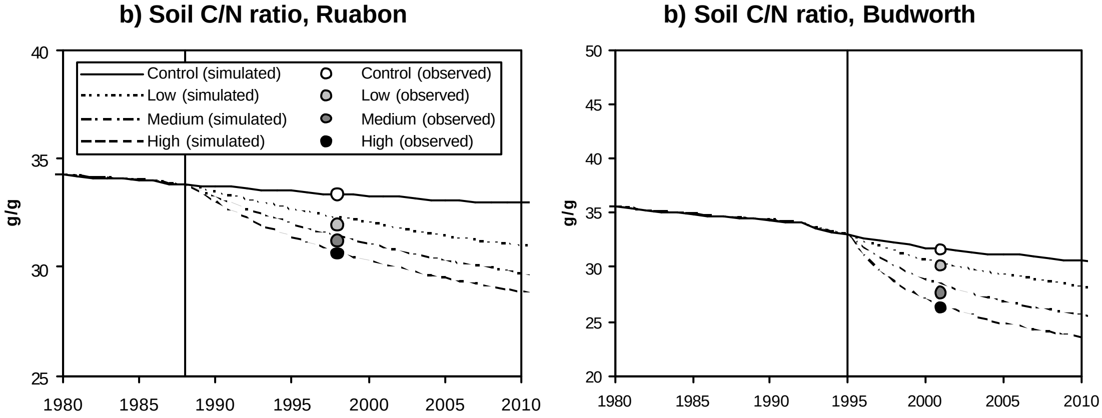


Figure 4. MAGIC simulation of N immobilisation and leaching. After subtraction of plant uptake and denitrification, immobilisation of the remaining inorganic N is a function of the C/N ratio of the soil organic pool. Upper and lower thresholds vary according to vegetation and/or soil type.

The N component of MAGIC and other similar models has been added relatively recently, and validation studies are limited by a number of factors. These include the large short-term climatic variability of surface water NO_3 concentrations, which make the underlying trends (as modelled by MAGIC) difficult to detect; the general paucity of long-term data on changes in C/N ratio at individual sites; and the difficulty of accurately quantifying changes in soil C and N pools, even where these are measured. However, rising trends in NO_3 concentrations since the 1980s in Northern Italian rivers, believed to be a consequence of terrestrial N saturation, were successfully reproduced based on an organic soil C/N ratio control (Rogora et al., 2003). Kopacek et al. (2004) have also shown that dramatic changes in lake NO_3 concentrations in the Tatra Mountains (Slovakia-Poland) since the 1950s can be satisfactorily explained in terms of changes in N deposition and declining terrestrial N retention as a function of declining C/N ratio. In the UK, few monitored surface waters show clear NO_3 trends, although Jenkins et al. (2001) used MAGIC to simulate NO_3 increases at Lochnagar, NE Scotland. Apart from the Rothamsted example (discussed below) there are few UK sites at which C/N has been repeatedly sampled over time, and therefore testing model predictions of C/N is problematic. However, data from plot-scale N manipulation studies have been used to test the ability of MAGIC to predict changes in observed N leaching and soil C/N under different addition levels (Evans et al., in prep.). For two sites with high quality soil C and N data (Fig 5), the model successfully reproduced observed decreases in C/N under three treatment levels. It should be noted that these simulations incorporated an (observed) increase in C storage as a consequence of N deposition, which slowed down the rate of C/N change. This was hypothesised to reflect additional liiter incorporation. This effect is being incorporated in MAGIC as part of the DEFRA 'Critical Loads and Dynamic Modelling' contract (www.ukreate.ceh.ac.uk).

Figure 5. Simulated and observed organic soil C/N ratio under ambient N deposition and three levels of long-term NH_4NO_3 addition at two heathland experimental sites. Vertical line indicates start of experiment.



Although a formal uncertainty analysis has yet to be undertaken for the N component of MAGIC, the key sources of uncertainty have been identified during extensive model development and testing. In a typical UK semi-natural ecosystem, N leaching forms a relatively small component of the overall N budget, and denitrification is negligible. Apart from intensively managed systems such as Rothamsted, N offtake (due to mowing, grazing etc) is also likely to be minor, and most of the N added must therefore be stored within the system. In general, vegetation provides a relatively small sink, so most of the N storage must occur within the soil. The calculation of change in soil C/N is therefore a relatively simple budgeting exercise, in which the major uncertainties are the amount of N deposition, the magnitude of current C and N pools (Evans et al., in press), and the extent of C accumulation as noted above. Soil C and N pools are difficult to quantify at large-landscape (e.g. catchment) scales, but may be measured relatively accurately at smaller scales. Consequently, the greatest uncertainty in the prediction of future C/N change in small-scale studies is considered to be the magnitude of current and future N deposition.

Because of the mismatch between MAGIC simulations of past C/N ratio on the Rothamsted Park Grass control plots and early measurements, the sensitivity of MAGIC predictions of C/N to uncertainty about the historic N offtake sequence was examined. Park Grass N offtake was calculated by multiplying hay offtake, for which accurate measurements based on decadal averages are available (Dodd et al 1994), by the proportion of N in hay biomass. For N proportion, a weighted average from three measurement periods (1920-23,1940-43 and 1956-59) was used (Warren & Johnston 1964). This calculation method is subject to uncertainty, since N availability during these periods may have differed from that in the earlier and later years of the experiment, leading to differences in N concentration in hay. This uncertainty has a large effect on the net addition (deposition minus offtake) and thus on the accurate simulation of the historic C/N trajectory (Figure 6). Note that raising N deposition would have the same effect as reducing N offtake on the simulation.

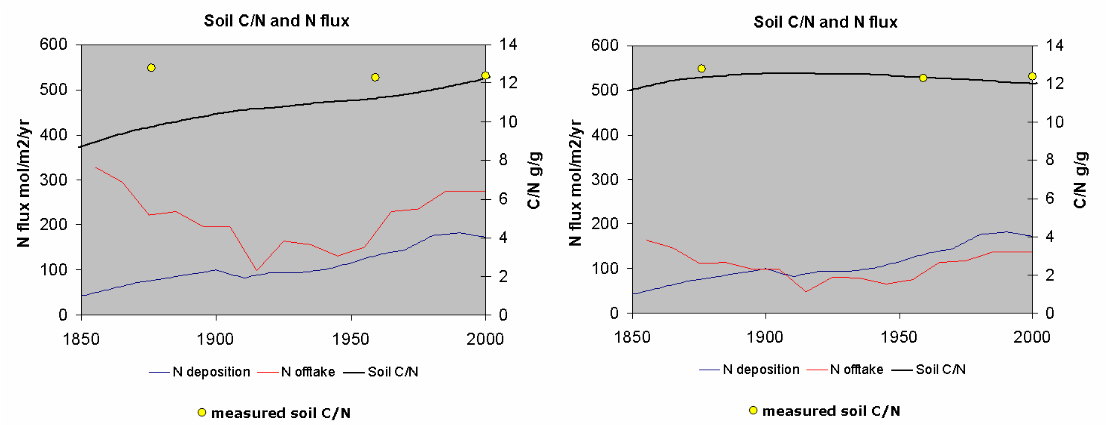
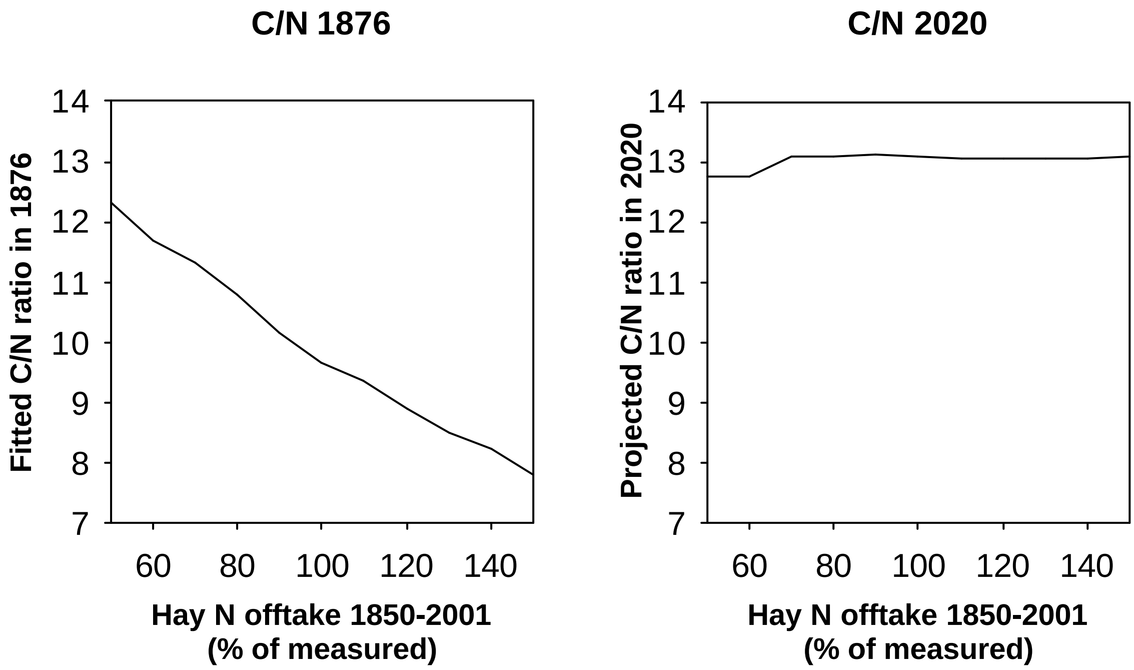


Figure 6. a) measured N offtake and deposition history, and comparison of MAGIC-simulated and measured C/N; b) as a) but with N offtake halved.

Simulated C/N in the early years of the model run was strongly affected by uncertainty in net N addition (Fig 7a). However, the projected C/N for 2020 was much less sensitive to uncertainty in this parameter (Fig 7b). This shows that while uncertain historic parameters can cause mismatches to historic measurements, projections are more robust than this mismatch would suggest. Accurate projections depend more strongly on accurate present-day measurements.

Fig 7. a) Past and b) future projection of soil C/N ratio at Rothamsted Park Grass.



Sensitivity analyses can be used to assess the uncertainty that should be attached to model outputs. A full uncertainty analysis of MAGIC's nitrogen module would be useful to determine which drivers are most important to measure in order to accurately predict future C/N, and to identify areas where the module could be improved to make better use of available data.

Conclusion: MAGIC has been widely used and validated although more testing of the N cycle component would be desirable. Uncertainty analyses can be used to locate possible errors in the input data when historical C/N trajectories are not well reproduced. The reliability of future projections depend largely on the accuracy of present-day soil measurements.

2.4.3 Testing calibration equations between soil and mean Ellenberg values

In an analysis of error propagation in Dutch soil and vegetation model chains, the residual variation about the calibration equations contributed the greatest uncertainty to predictions of change in species composition (Schouwenberg et al 2001, Van Dobben et al 2004). Calibration equations developed for British soils and plant communities have higher r-sqrd values than their Dutch equivalents. For example Ertsen et al (1998) found an r-sqrd of 54% for Ellenberg fertility and standing crop (62% for British calibration between soil C/N and mean Ellenberg fertility), 51% for soil moisture (72% for British calibration) and 54% for soil pH (61% for British calibration). Even so much unexplained variation is still evident in both nation's calibration results (see Box 5).

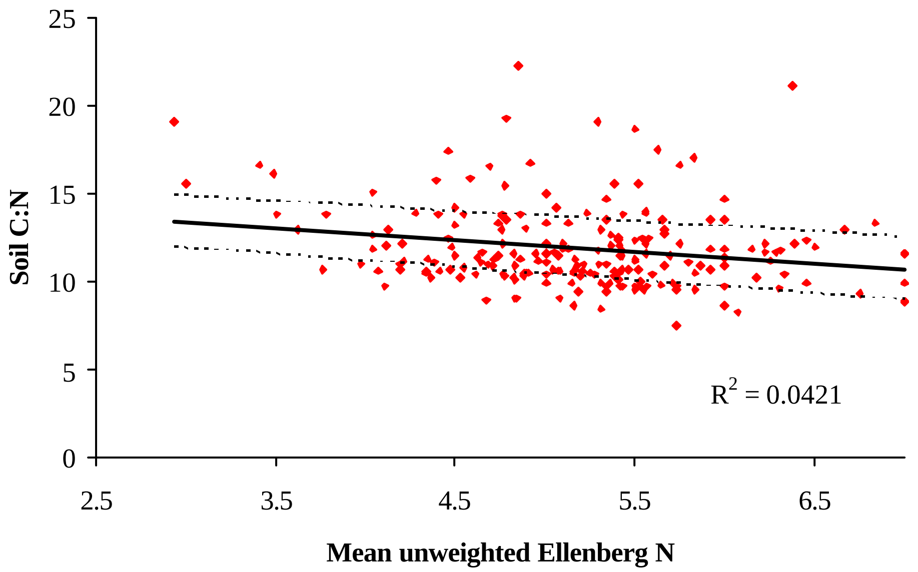
Global or vegetation-type specific calibrations?

Wamelink et al. (2002) showed that the variation explained in calibration equations could be improved dramatically if equations were developed separately for each vegetation type. The reasons why 'global' cross-community calibrations should perform less well relates to the fact that additional explanatory variables that are correlated with vegetation type are not included in each calibration (Smart and Scott 2004; Wamelink et al 2002). Differences can be a function of shade and moisture (Schaffers 2000) and can also reflect the subtle correlations between Ellenberg numbers. For example in Scandinavian forest plots, mean Ellenberg R values are negatively correlated with mean Ellenberg light values because in these species pools increased shade tends to filter out the more acidophilous species that favour more open conditions (Diekmann 2003). The reverse pattern occurs in British broadleaved woodlands since species of more open woodland conditions tend to be more circum-neutral in their affinities and hence have higher Ellenberg R and N numbers (Kirby et al 2005). The fact that woodland tends to be shaded will therefore affect the overall pool of Ellenberg numbers from which means are drawn. Thus shade will ultimately impact the calibration between soil C/N or pH and mean Ellenberg N and R values. Such vegetation-type specific phenomena support the use of vegetation-type specific calibration equations (Wamelink et al. 2002). The problem they create is that change between vegetation types is no longer conveyed by one equation.

The desirability of within-vegetation type calibrations in our model tests arose when translating initial MAGIC predictions for Rothamsted into predicted species composition. The exponential form of the British calibration between mean Ellenberg N and soil C/N resulted in very large increases in mean Ellenberg value with only very small reductions in soil C/N . This must partly reflect the reality of the relationship but also reflected shortcomings in the early MAGIC runs at the site. A new calibration was developed for just infertile neutral and calcareous grasslands. It performed better and was adopted in further model tests yet residual variation was still very high (Fig 8) such that many of the plots in the dataset used to build the calibration apparently have soil C/N and pH values that deviate markedly from expectation given above-ground plant species composition. An obvious conclusion is that soil C/N on its own is a poor predictor of mean Ellenberg N and this lack of explanatory power is worst toward the lower end of the soil C/N range, which includes neutral grasslands. Hill & Carey (1997) concluded that Ellenberg N values are better treated as overall indices of fertility rather than explicitly N availability.

In certain situations, the apparent disequilibrium between soil C/N and above-ground species composition can be understood. A good example being drained, peaty, high-grade soils under cultivation or supporting sown improved grassland, yet associated with rapid mineralization (eg. Grootjans et al 1985).

Fig 8. Calibration relationship between mean Ellenberg N score and soil C/N based on Countryside Survey plots in the infertile grassland aggregate class only. Upper and lower 95% prediction intervals are also shown.



<u>Calibration equations and the impact of their uncertainty on predicted soil and vegetation measurements</u>

The objective here was to assess the impact of uncertainty about the calibrated relationship between soil variables and mean Ellenberg values. The first step was to estimate soil C/N and pH values from the species composition of the vegetation on each test site using the calibration equations described above and in Box 5. Uncertainty about the predicted soil C/N and pH values was represented by 95% confidence intervals for each point. These intervals were derived from the parametric uncertainty in the coefficients of the calibration equation, ie. based on the standard errors of each equation term. When soil variables estimated from each calibration equation were plotted against observed measurements for the same plot locations, all observed values except those in the unimproved grasslands at Porton Down (soil C/N), Rothamsted Park Grass (soil C/N and pH) and Moorhouse (pH), fell outside the confidence interval of each prediction (Fig 9). This was despite the fact that observed and predicted values were positively correlated. One possible hypothesis is that soil C/N and pH in the other largely acidic and peaty habitats are to some extent out of equilibrium with above-ground species composition because species compositional changes lag behind soil changes induced by pollutant deposition or climate change.

The Rothamsted Park Grass control vegetation, on the other hand, has been considered a stable equilibrium community (Silvertown 1987).

The apparent disequilibrium between soil and vegetation goes in both directions; some plots had soil values less than would be expected while others had higher C/N and pH (Fig 9, Table 5). These mismatches could however simply be a matter of sampling scale effects. For example a single C/N measurement based on the top 10-15cm of soil minus litter inevitably paints a homogenous picture that will deviate to some extent from the C/N variation experienced by plant species with different below ground foraging strategies and hence rooting depth (eg. Fitter 1994; Ettema & Wardle 2002). Soil developments at Rothamsted provide another example. Measured C/N ratios in later years in experimental addition plots have carefully avoided a thin mat of persistent litter that has developed in the O horizon over the course of the experiment (Paul Poulton pers.comm. and Warren & Johnston 1964). This would result in C/N measurements indicating a higher fertility rooting zone than that encountered by at least some of the more shallow rooting species present. Small scale heterogeneity will be most marked in vegetation types well known for their micro-topographic variation in turn associated with variation in pH and nutrient availability. Examples include blanket bog, topogenous and soligenous base-rich fens (eg. Giller & Wheeler 1988, Clapham 1940, Bellamy & Rieley 1967). Such variation is unlikely to be properly conveyed by one or two soil samples targeted on a standard horizon depth. Another scale-dependent source of discrepancy is that the plot area over which species are censused is always larger than the point location of soil samples. While the benefit of mean Ellenberg values is that they integrate the abiotic environment over a certain quadrat size, this averaging effect is likely to result in differences between mean Ellenberg values predicted from soil samples versus above-ground vegetation. The problem is that such unexplained system variation will be very demanding, if not impossible to measure from place to place (van Dobben et al 2004).

The comparison of observed versus predicted pH values revealed an apparent systematic under-prediction (Fig 9b). This is almost certainly explained by the fact that the calibration equation was built on the relationship between mean Ellenberg R and soil slurry pH while observations and MAGIC predictions relate to soil water pH, which tends to be higher. Further research should focus on deriving a correction factor to address this problem.

Fig 9. Comparison of observed soil C/N (a) and soil pH (b) versus soil C/N and pH predicted by back-transforming from mean observed Ellenberg N and R scores. Empty squares indicate locations where the observed value fell inside the 95% CI for the predicted value.

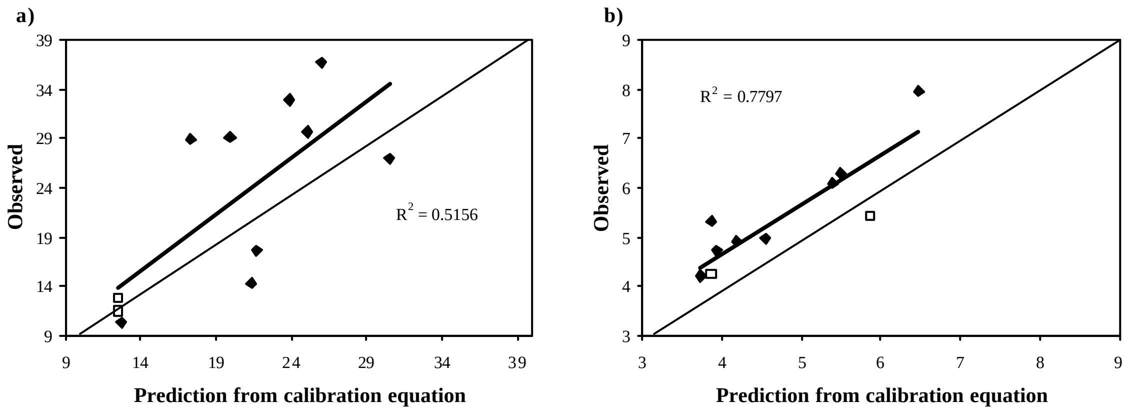


Table 5. Data table for figure 9.

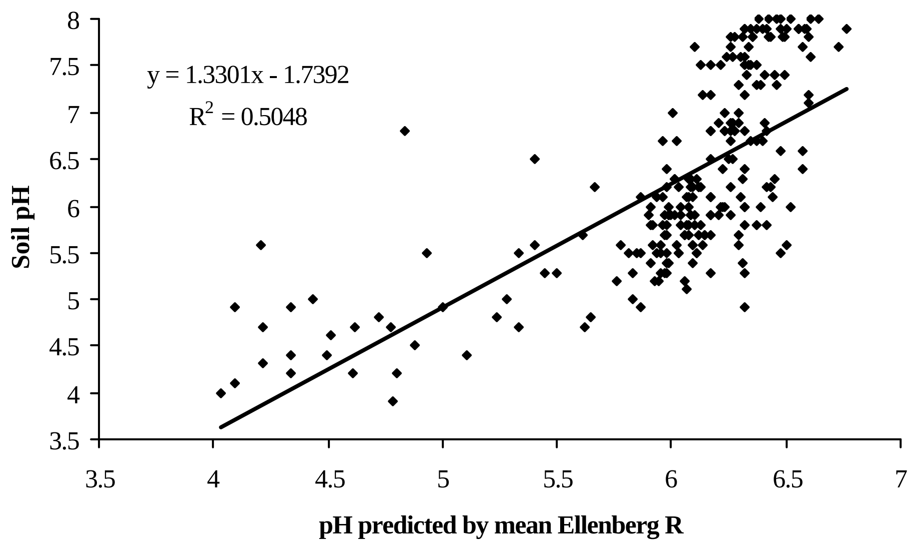
Site & year	Predicted C/N	Observed C/N	Predicted pH	Observed pH
Roth PG 1877/76	12.54	12.77	5.87	5.40
Porton 2000	12.55	11.39	6.48	7.96
Cairngorm 2000	25.03	29.69	3.87	5.32
Pwllpeiran 2003	21.33	14.30	4.18	4.92
Climoor 1998	25.98	36.70	3.73	4.20
Moorhouse 2000	30.49	26.97	3.87	4.24
Budworth 1998	19.91	29.10	3.93	4.72
Plynlimon 2002	21.70	17.60	4.55	4.98
Ruabon 2005	23.89	32.90	3.73	4.21
Dead Island Bog 2005	17.31	28.89	5.49	6.30
Dromore Motte 2005	12.71	10.34	5.40	6.10

Testing the pH calibration equation against independent data

Soil pH has been measured in a sample of monitoring plots located in English Environmentally Sensitive Areas (Critchley et al 2001). Tests of both the global calibration equation and the grassland-only calibration equation were carried out by generating a predicted soil pH from the mean Ellenberg R score calculated from the species composition, and comparing this with measured soil pH (Fig 10). The r-squared values differed by a negligible amount. Results showed that only half the variation in pH predicted from the species composition was explained by measured pH. The scatter of points suggests however, that the plots fall into two groups defined by two different relationships. Quadrats with higher pH values had a lower range of predicted pH values. This is an inevitable consequence of the fact that average Ellenberg values are likely to fall short of actual maximum and minimum Ellenberg numbers because of averaging effects and the fact that Ellenberg numbers cannot be more or less extreme than their upper and lower limits. For example a plot associated with high pH might contain 15 species but among these 15 there will be a range of Ellenberg R values present because species differ in the width of their tolerance about their Ellenberg optimum. However none of these species will have values higher than the highest possible value. Thus the average Ellenberg value in the plot will reflect this variation in Ellenberg values and fall short of the Ellenberg maximum even if soil pH is high.

It seems likely that further subdivision by vegetation type of the training datasets used to build the calibration equations could generate better fitting within-vegetation type relationships. This is an area of much needed further research because of the unacceptably high uncertainties in the current equations.

Figure 10. Soil pH in monitoring plots from English Environmentally Sensitive Areas predicted by calibration equation relating observed mean Ellenberg R to soil pH (see Box 5)



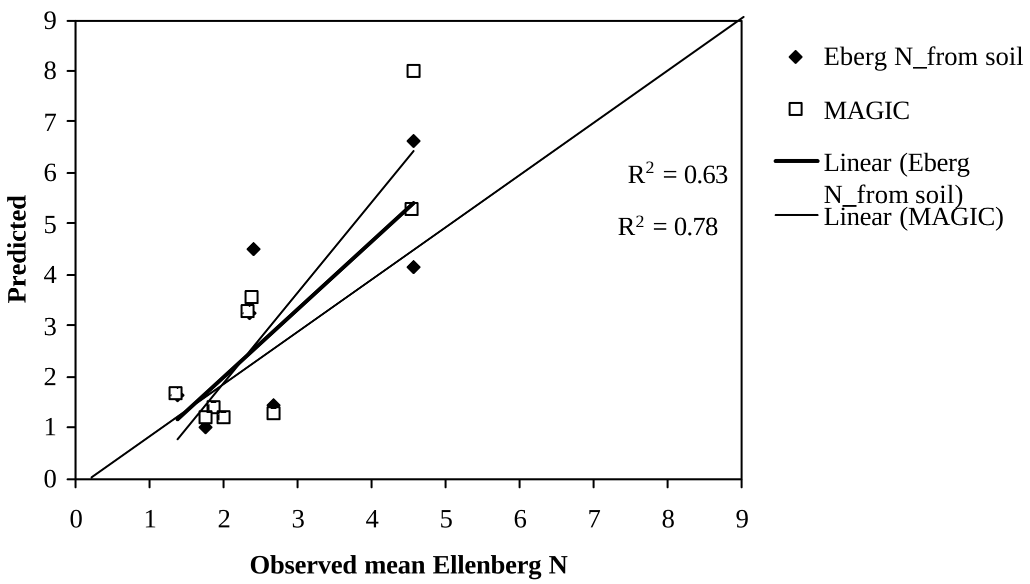
Conclusion: After taking into account uncertainty in calibration equations, results imply that some plant communities have above-ground vegetation whose composition is at odds with predictions derived from calibration equations. This is to be expected for a variety of reasons. Quantifying such sources of residual error at the plot level will be too resource-demanding if the general applicability of models across many species and vegetation types is to be retained. A systematic bias due to differences between soil water versus soil slurry pH should be easier to address by constructing an empirical correction factor. Better calibration curves are likely to be derived from finer subdivisions of quadrat data into different vegetation types. This is also an area of promising further research.

2.4.4 Comparison of observed versus predicted mean Ellenberg values

In this section observed mean Ellenberg values calculated from the plant species composition of test quadrats are compared with mean values predicted from observed soil data and MAGIC generated soil C/N and pH. On all sites, even Rothamsted, paired soil measurements and species composition were only available for a single occasion, usually a recent measurement to which MAGIC calibrates its hindcast. All sites are therefore plotted together (Figs 11a and b). Again, correlations were positive and r-squared values high. However, the high values for MAGIC-derived versus observed Ellenberg scores are partly a consequence of the fact that MAGIC predictions calibrate to current soil C/N and pH. Over-prediction of Ellenberg N was a particular feature of plots with higher observed Ellenberg N values – the neutral and calcareous grasslands at Rothamsted and Porton respectively – and is indicated by the fitted slopes being much steeper than the y=x line (Fig 11a). This reflects the form of the calibration curve explored in the previous section. Whether 'global' or within-vegetation type, at low soil C/N values, a small reduction results in a large increase in mean Ellenberg N. Mean Ellenberg R values were better predicted by both soil measurements and MAGIC estimates. The largest discrepancy was at Porton Down (Fig 11b and Table 6) where the observed mean Ellenberg R was substantially lower than both predictions; in fact MAGIC did not calibrate well to the present-day soil pH measurement at the site. Overall the positive correlations and high r-sqrd values between observations and predictions give a degree of confidence in the calibration equations. Yet the critical sensitivity of concern, is the over-prediction of mean Ellenberg N in more fertile semi-natural habitats.

Figure 11. Comparison between mean unweighted Ellenberg N (a) and R (b) based on observed species composition scores versus scores predicted by using calibration equations to transform soil C/N and soil pH values measured on each site and values predicted by MAGIC.

a)



b)

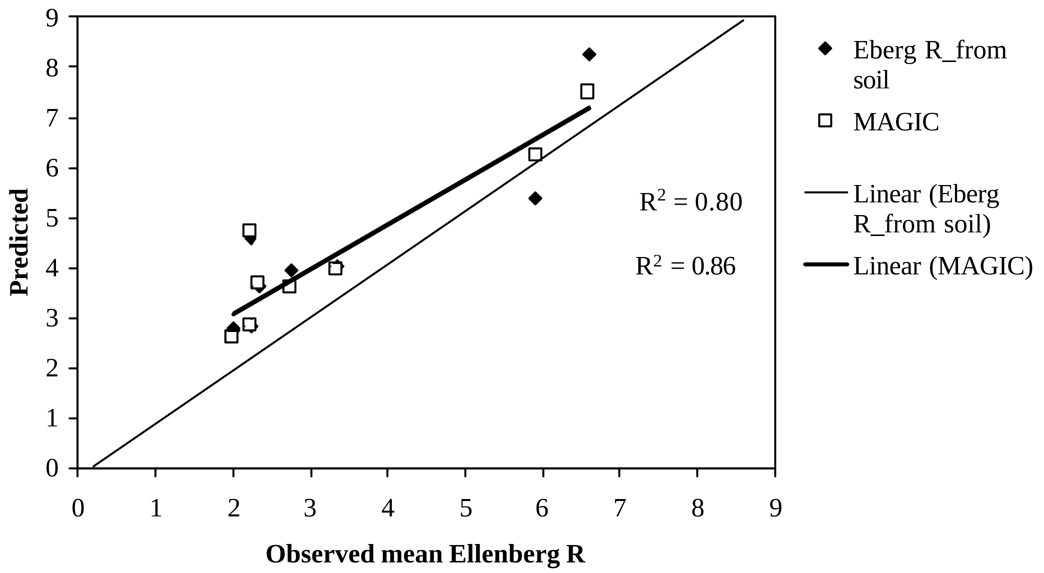


Table 6. Data table for figure 11 above. Note that predicted mean Ellenberg N and R values were derived by back-transformation from the calibration equations between Ellenberg scores and soil variables (see Box 5).

Site & year	Observed mean Ellenberg N	Mean Ellenberg N predicted from measured soil C/N	Mean Ellenberg N derived from MAGIC C/N ratio prediction
Rothamsted Park Grass 1877/76	4.57	4.15	7.96
Porton Down2000	4.55	6.64	5.29
Cairngorm 2000	1.86	1.42	1.38
Pwllpeiran 2001	2.39	4.52	3.55
Climoor 1998	1.75	1.01	1.19
Moorhouse 2001	1.36	1.65	1.65
Budworth 1998	2.67	1.46	1.27
Plynlimon 2002	2.33	3.25	3.26
Ruabon 2005	2.00	1.20	1.19

Site & year	Observed mean Ellenberg R	Mean Ellenberg R predicted from measured soil pH	Mean Ellenberg R derived from MAGIC pH
Rothamsted Park Grass 1877/76	5.91	5.38	6.24
Porton Down 2000	6.60	8.26	7.50
Cairngorm 2000	2.23	4.61	4.72
Pwllpeiran 2001	2.74	3.95	3.62
Climoor 1998	2.00	2.77	2.61
Moorhouse 2001	2.23	2.84	2.86
Budworth 1998	2.33	3.62	3.70
Plynlimon 2002	3.34	4.05	3.99
Ruabon 2005	2.00	2.79	2.61

Conclusion: Positive correlations between mean Ellenberg scores derived from soil measurements versus means based on observed species composition, offer a general validation of the calibration equations between soil and mean Ellenberg values. Discrepancies are more likely in neutral and calcareous grassland Priority Habitats because a wider range of mean Ellenberg N values are possible over a very narrow range of soil C/N. However, calibration equations are also typified by large unexplained variation that is ignored when soil C/N, pH and % soil moisture are transformed into mean Ellenberg values.

2.4.5 Modelled versus observed species richness

A test of observed versus predicted plant species richness was carried out in the same way as for mean Ellenberg values in the previous section. The test aimed to assess the performance of the statistical model of species richness developed for terrestrial plant communities in Britain, where species richness is predicted by the explanatory variables cover-weighted canopy height, soil C/N, soil pH and soil moisture content.

Model uncertainty was high (Appendix 1) hence predictions would have large prediction intervals, reducing the precision of the model for forecasting or testing scenarios of change. Results confirmed this (Fig 12) but also showed that observed species richness on test sites was much more variable than predictions based on MAGIC estimates of soil C/N and pH. Thus, while a positive correlation was seen between observed and predicted values, the slope of the relationship deviated considerably from the y=x line (Fig 12). In addition, considerable variability attached to both observed and predicted values. Since species richness is typically a logarithmic function of area sampled, the natural log of test plot size was graphed against observed richness to identify any area-related bias. None was apparent (Fig 13).

Figure 12. Comparison of observed species richness versus species richness predicted from MAGIC soil C/N and soil pH predictions (+/- standard deviation).

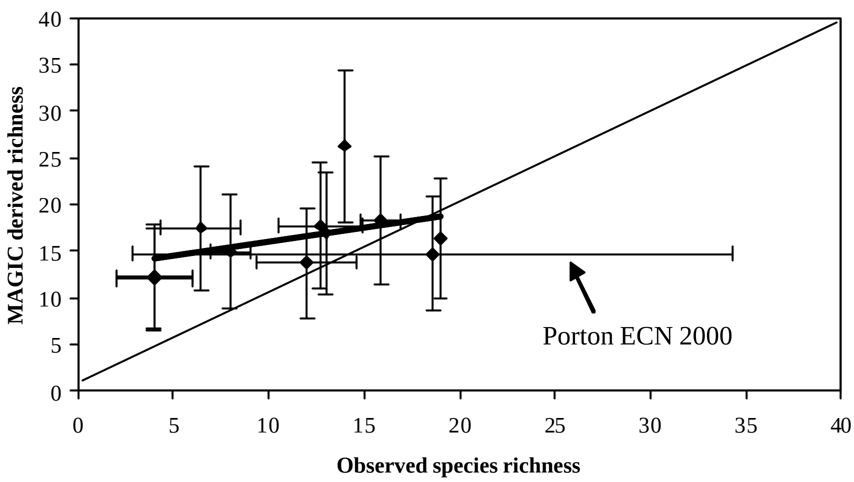
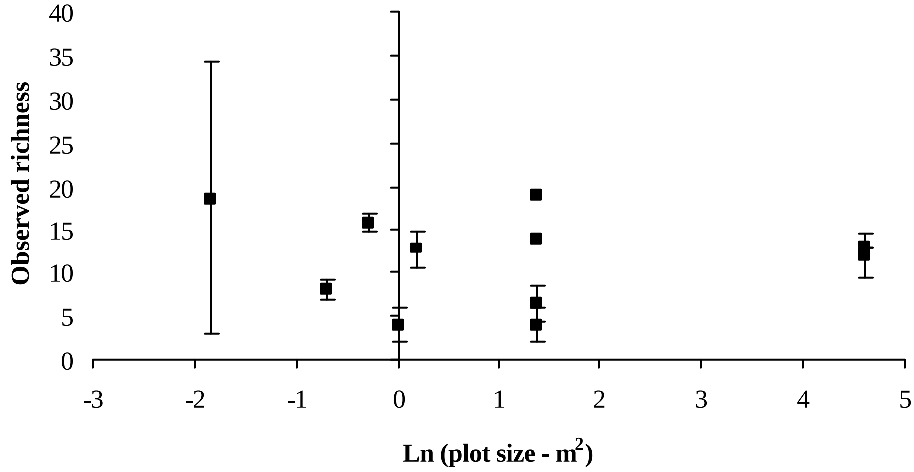


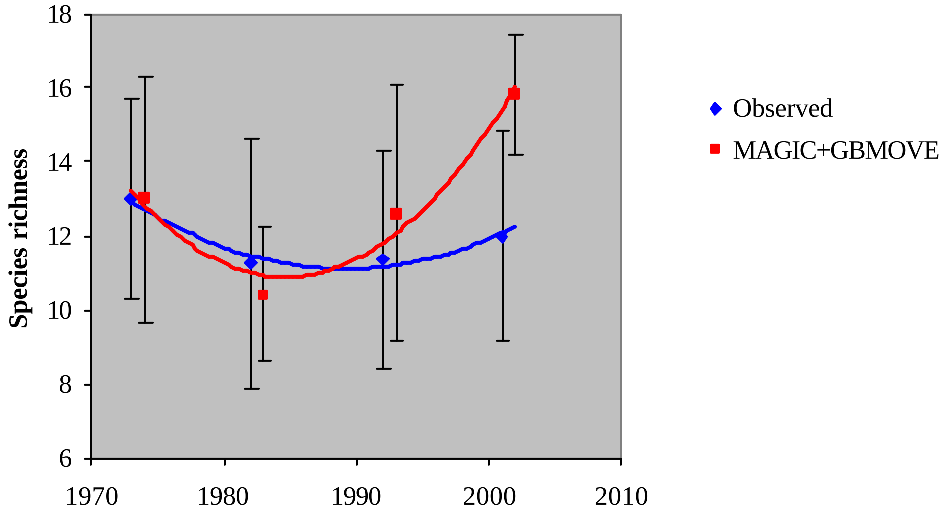
Figure 13. Observed species richness (+/-sdev) versus plot size on model testing sites.



2.4.6 Species richness through time

A comparison of observed versus predicted species richness over time was carried out using the Moorhouse Hard Hills control plots; this being the site with the greatest number of time intervals and where MAGIC performed well (Fig 14). As expected, variation about observed and predicted values was large yet similar trajectories of change were seen. Species richness was predicted to increase between 1981 and 2001 to a greater extent than observed.

Figure 14. Predicted and observed changes in species richness (+/-sdev) in the Moorhouse Hard Hills control plots with fitted polynomial regression lines. Predictions were based on MAGIC simulations of C/N ratio and soil pH with %soil moisture and canopy height held constant.



Conclusion: Observed values of species richness will generally be poorly predicted from the best statistical model. However there is weak evidence that the direction of change in richness given change in soil conditions and

successional stage can be correctly predicted albeit with high levels of uncertainty surrounding each estimate. Because of this variability, modelled species richness values will over-populate predicted pseudo-quadrats. Further testing is required to confirm the ability of the model to correctly predict direction and rates of change in species richness.

2.4.7 Validating GBMOVE regression models

While each individual species model could, in theory, be validated against independent test data, such an exercise would be extremely time consuming although probably desirable if credibility is to be properly won for the majority of species models. This activity is best seen as a long-term campaign involving the gradual accumulation of other datasets for testing against GBMOVE. CSM indicators could be prioritised in this way, which would give a manageable subset of species to focus on.

An overall assessment of the accuracy of the GBMOVE niche models can however be gained by correlating species' Ellenberg numbers for fertility (N), soil pH (R) and soil wetness (W) with the estimated optimum position of each species along each gradient derived algebraically from each GBMOVE equation (see Appendix 5 and Box 6). While this approach does not validate the predictive accuracy of the entire response surface, it does verify that a critical attribute of the GBMOVE models is consistent with independent data.

Initial correlations were disappointing but r-squared values improved substantially when certain groups of models were excluded (Fig 15). Species were excluded for which no algebraic optimum existed or where a saddle point, ie. a minima, was estimated rather than a maximum (Appendix 5). Next, only species were included that had significant quadratic terms for every gradient. Finally, species were excluded for which Ellenberg did not give numbers in his original publication because they ranged widely across environmental gradients (Schaffers & Sykora 2000; ter Braak & Barendregt 1986).

Figure 15. Change in r-squared values between optima predicted by GBMOVE models and published Ellenberg numbers for higher plants. R-squared values increased as certain groups of species were progressively excluded

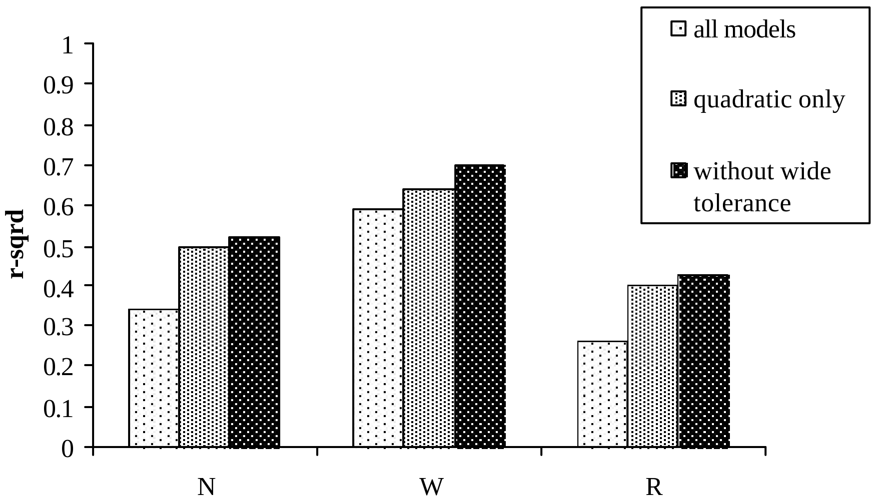
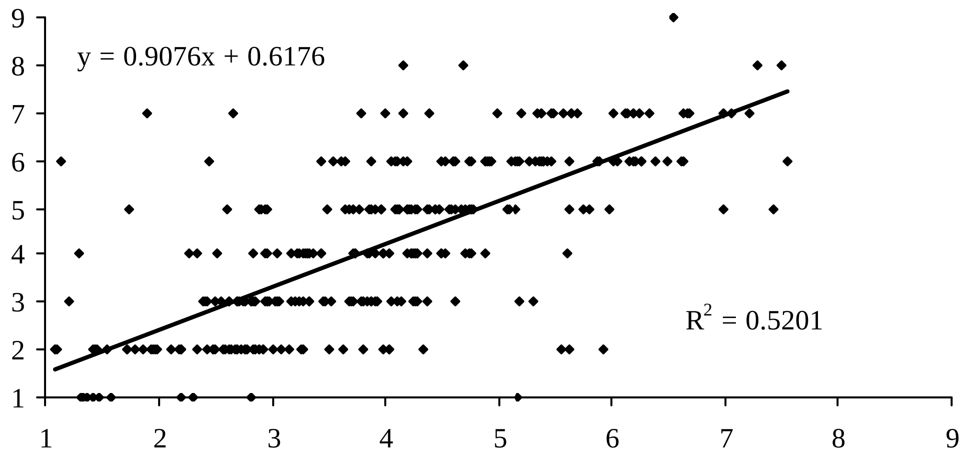
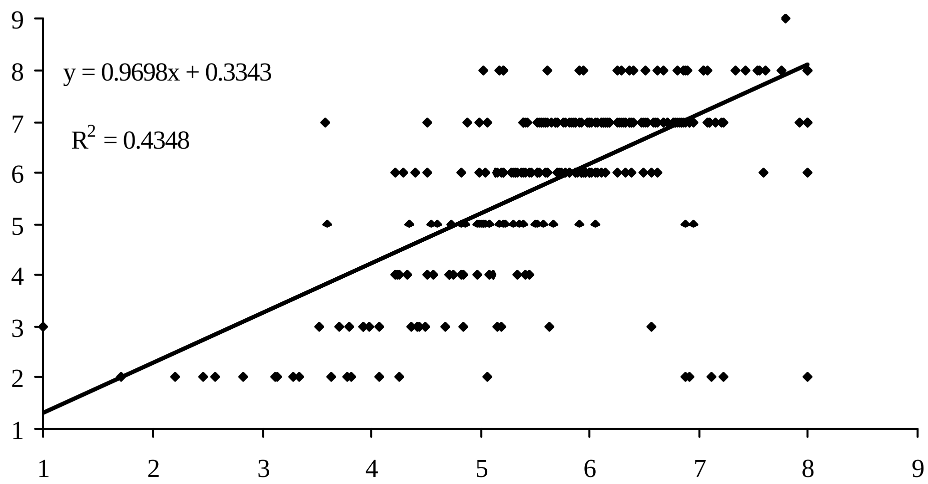


Figure 16. Scatter plots of predicted optimum Ellenberg values from GBMOVE models (x-axis) versus published Ellenberg numbers (y-axis) for higher plants. a) Fertility, b) soil pH, c) Wetness.

a)



b)



c)

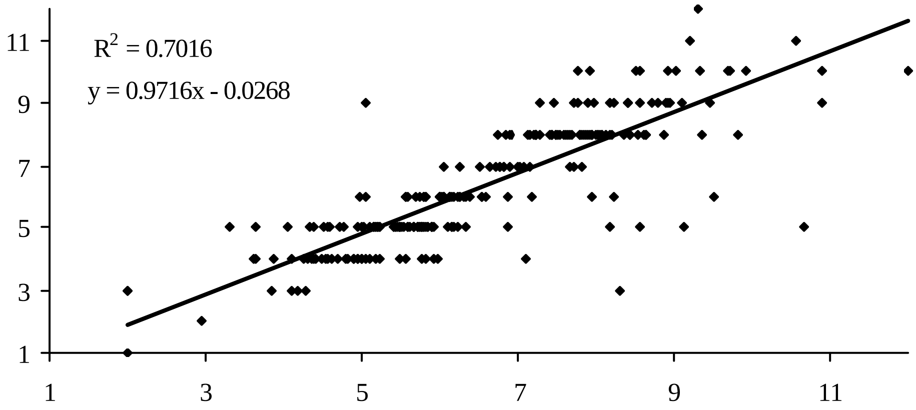
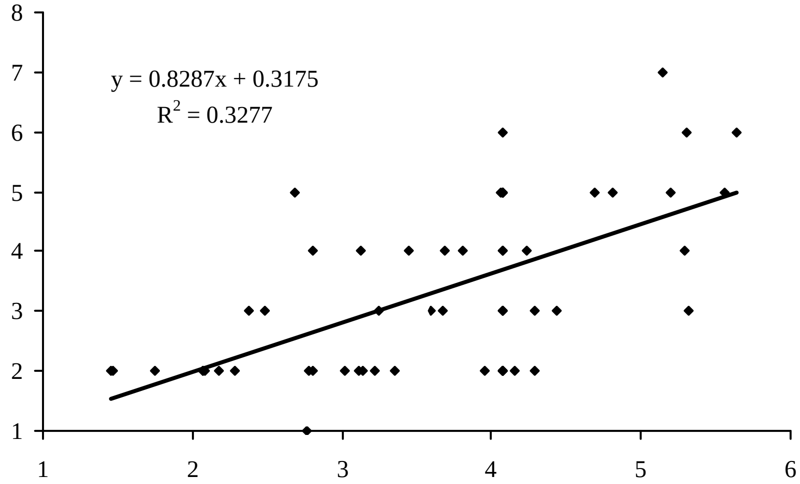
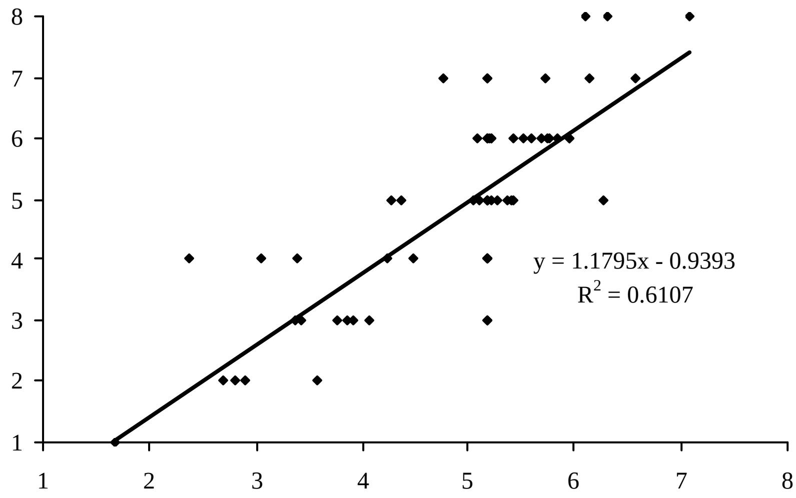


Figure 17. Scatter plots of predicted optimum Ellenberg values from GBMOVE models (x-axis) versus newly generated Ellenberg numbers (y-axis) for bryophytes. a) Fertility, b) soil pH.

a)



b)



Results showed that higher plants were on average better correlated with published Ellenberg numbers than bryophytes. Also no substantial bias was evident in that slopes were all near to one and intercepts between zero and one (Fig 16a-c). Bryophyte optima for fertility were poorly correlated with the newly produced Ellenberg N values with a bias toward lower GBMOVE optima at the more fertile end of the gradient. Oddly, soil pH optima were reasonably well correlated (Fig 17b) and in fact faired much better than soil pH optima for higher plants (Fig 16b). The results for bryophytes are encouraging. Models were constructed for bryophytes and lichens in exactly the same way as for vascular plants on the basis that different species exhibit preference for different community types and hence for different ranges of abiotic variables including soil pH and C/N. Although their interaction with their abiotic environment will be different from rooting plants, the empirical basis of the models allows for the fact that the match between species and environment remains correlative rather than mechanistic. The fact that Ellenberg R, N and L values have been newly computed for bryophytes as part of this project indicates that soil pH, substrate fertility and shade are indeed significant constraints on species occurrence despite the clear importance of substrate type and other factors.

Conclusions: GBMOVE optima showed variable correlations with Ellenberg numbers. The results are encouraging and suggest that a core subset of reliable GBMOVE models have been developed but some models warrant closer examination as their optima deviate strongly from their Ellenberg values. Examination of individual species models would be the only way of further selecting the best species. Focussing on CSM indicators would be a user-oriented and convenient criteria for selecting models for closer inspection.

2.4.8 Modelled versus observed species composition

The objective in this section was to determine how successfully MAGIC+GBMOVE could reproduce the observed species composition in test plots. Results were compared with predictions based on mean Ellenberg values and species richness taken from observed plot data. Hence, the comparison is against predictions generated without calibration equations, species richness prediction or MAGIC predictions of soil C/N and pH and therefore with the least uncertainty. The goal must be near complete prediction of all species present, arbitrarily over 90%, if the models are to be used to reliably predict community assembly.

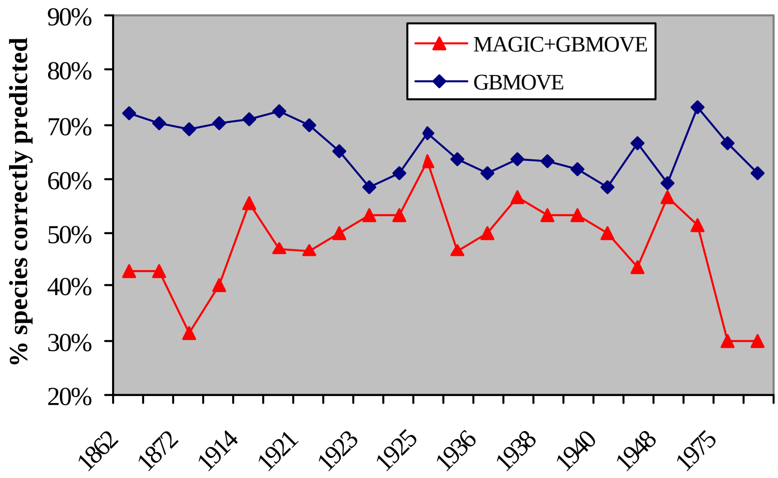
The prediction of species composition involved the most complicated manipulation of predicted outputs of all the model tests. This is because the goal was to assemble a set of simulated quadrats each populated with an expected species composition. Because predicted probabilities from each GBMOVE model cannot be directly interpreted as expectations of occurrence in a stand (see Box 6) we used species richness along with GBMOVE output to build pseudo-quadrats in the following stages.

1. MAGIC predictions of soil C/N and pH were used in combination with fixed soil moisture and canopy height to generate GBMOVE probabilities of occurrence for all species present in the 10km square species pool.
2. These input data were also used to generate a mean species richness and an associated variance of richness values which were assumed to be normally distributed about the mean.
3. A set of simulated richness values was generated conditioned on this distribution.
4. Richness values were drawn at random from this distribution up to the number of observed quadrats available on each site. Hence each pseudo-quadrat had an assigned predicted species richness.
5. Pseudo-quadrats were then populated by drawing a number of species up to the predicted richness for that pseudo-quadrat from the list of predicted GBMOVE probabilities sorted in descending order.
6. Having generated a set of simulated quadrats, the predicted frequency table was matched against the National Vegetation Classification so that community level affinities could be assessed in addition to the key comparison of how many species present in the observed data were actually present in pseudo-quadrats.

Rothamsted Park Grass

At Rothamsted, predictions were possible for the entire time-series. When observed mean Ellenberg scores were used to predict species composition, an average of 67% of all species observed were actually predicted (Fig 18). When predictions were based on MAGIC estimates of soil C/N and pH as input to GBMOVE, the percentage correctly predicted decreased substantially (Fig 18).

Figure 18. Percentage of species correctly predicted in the three Park Grass control plots based on pseudo-quadrats whose richness and composition were predicted by MAGIC+GBMOVE based on FRAME N and S deposition versus predictions based on observed mean Ellenberg scores only as input to GBMOVE.



Over the 150 years of the experiment at Rothamsted, a modest but significant reduction in yield and a shift toward less fertile assemblages has been seen in control plots reflecting continuous hay offtake with no N addition except from the atmosphere. In the last 30 years there has been a tendency for increasing yield (Silvertown et al 1994) that may reflect N deposition, although this was not reflected in the botanical data transcribed from Williams (1978) since the time series stopped at 1976. Consistent with the moderate but long-term reduction in yield in control plots through most of the time period, Dodd et al (1994b) reported a best fit to MG5a *Centaurea nigra – Cynosurus cristatus* grassland, *Lathyrus pratensis* sub-community throughout the time period but with an increase in similarity to the somewhat less fertile and more calcicolous MG5b *Galium verum* sub-community from the early 1900s onward.

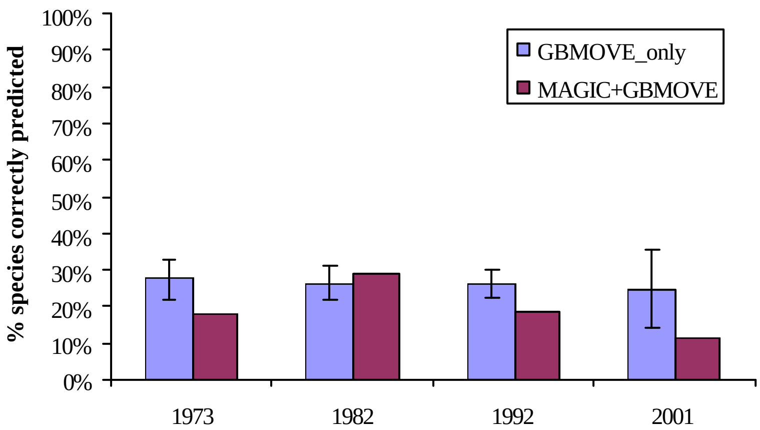
When predicted species composition derived from observed soil pH, C/N ratio and canopy height were matched against the NVC, plots recorded in 1876 were predicted to be either the more fertile MG6 *Lolium perenne – Cynosurus cristatus* grassland or MG5 (plots 2 and 12) or CG2 *Festuca ovina –Avenula pratensis* grassland (plot 3 in 1876) with no obvious indication of any change in trophic status throughout the period, consistent with negligible change in C/N ratio. In plot 3 a consistent change was predicted with a shift from CG2 to MG6. Only in this plot did measured soil C/N decrease appreciably from 13.2 in 1876 to 12.6 in 2000. When pseudo-quadrats were generated from the best current MAGIC run, the best fit at the start of the experiment was with MG5 changing to the more fertile, semi-improved MG6 by 1976.

The results for Rothamsted indicated that even when mean Ellenberg scores were used as input to GBMOVE based on observed species composition, substantially less than 90% of species present were predicted to be present. Although the NVC community type was correctly predicted by MAGIC+GBMOVE at the start, the observed and modelled decline in soil C/N resulted in a shift to a more fertile community type, which is the reverse of the exens tifying trend observed in the control plots (Williams 1978; Dodd et al 1995). Therefore the main reason for poor performance of MAGIC+GBMOVE appears to be that observed soil changes were actually inconsistent with observed vegetation change.

Moorhouse

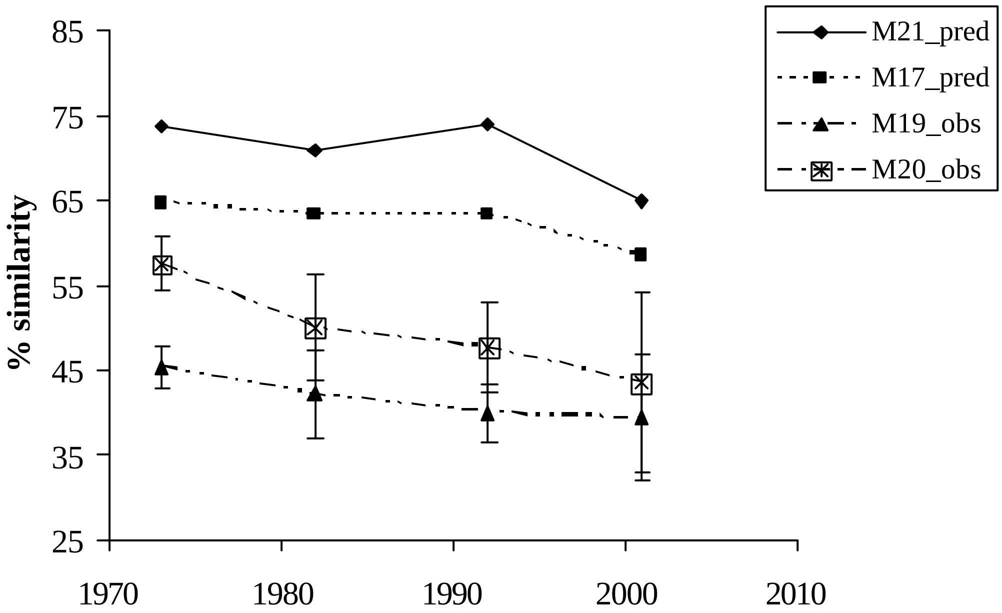
The same testing strategy was adopted as for Rothamsted. A comparison was made between the proportion of species present in each year that were correctly predicted using observed mean Ellenberg values and observed species richness, versus predictions from MAGIC linked to GBMOVE and a statistical model of above-ground species richness (Fig 19).

Fig 19. Percent of species in control plots that were correctly predicted to be present based on observed Ellenberg values (+/-sdev) ie GBMOVE only, versus the percentage correctly predicted in pseudo-quadrats generated by MAGIC+GBMOVE.



When predicted species lists for both GBMOVE and MAGIC+GBMOVE were examined, key absences included a range of bryophytes. However, the key dominants in the vegetation were predicted to be present in both model runs. These included *Calluna vulgaris*, *Eriophorum vaginatum*, *E.angustifolium*,the liverwort *Cephalozia connivens* and *Sphagnum capillifolium*. Both predictions lacked the important diagnostic species *Rubus chamaemorus* but this was because this species has no GBMOVE model. Predictions generated by MAGIC+GBMOVE also included diagnostic species for the lowland valley mire community M21 such as *Narthecium ossifragum* with *Erica tetralix*, *Molinia caerulea* and occasional *Vaccinium oxycoccus*. The result was that predicted and observed quadrats were assigned to different NVC communities (Fig 20). MAGIC+GBMOVE predictions tended to be most similar to M21 *Narthecium ossifragum – Sphagnum papillosum* lowland valley mire and the Western oceanic lowland M17 *Scirpus cespitosus – Eriophorum vaginatum* blanket mire. Observed control plot data tended, as expected, to be most similar to M19 *Calluna vulgaris - Eriophorum vaginatum* and M20 *Eriophorum vaginatum* blanket mires (Fig 20).

Figure 20. Change in matching coefficients with the top two best-fitting NVC communities for observed control plots at Hard Hills and predicted pseudo-quadrats generated using MAGIC+GBMOVE for the same time period (+/-sdev).

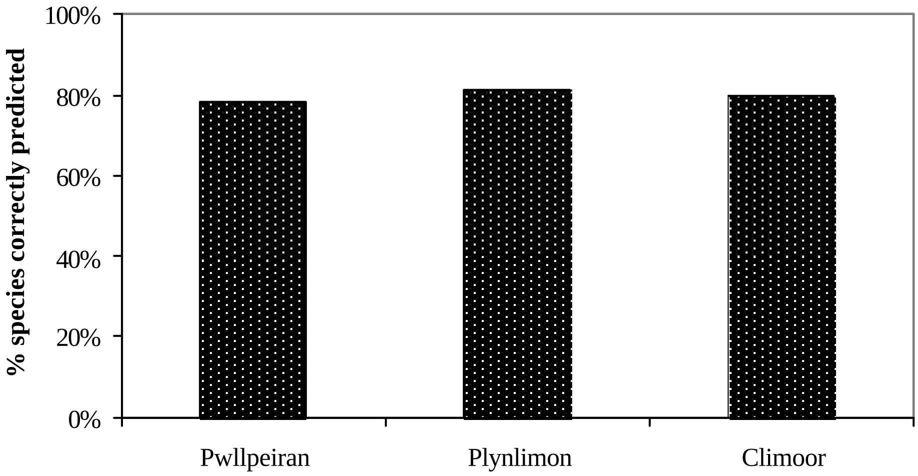


As seen at Rothamsted, the linked models did not perform particularly well in predicting species actually observed in control plots. Yet neither did predictions based solely on observed mean Ellenberg values.

Sites with only one time point

Contrary to the poor predictions at Moorhouse and Rothamsted, relatively high proportions of observed species were actually predicted at three experimental sites (two on upland acid grassland and one on upland heath – see Table 1, Fig 21). However, these results are influenced by over-prediction of species richness, hence many more species populate pseudo-quadrats than were actually observed yet this larger pool then gives a greater likelihood of including those actually present. This is reinforced by the low Jaccard similarity coefficients between observed and predicted quadrats for the three sites. These coefficients take into account species predicted to be present but that were in fact absent. Values of 1 indicate all species in common while 0.5 would indicate that half the species present in observed and predicted plots were not shared. Results for the three sites were as follows; Pwllpeiran – 0.44, Plynlimon – 0.59, Climoor – 0.22.

Figure 21. Proportions of species present at each site that were predicted to occur in pseudo-quadrats assembled using predicted species-richness values and occurrence probabilities generated by MAGIC+GBMOVE.



Overall, the results suggest generally poor performance if the goal is to accurately predict at least 90% of the species present in a sample plot. However, when observed mean Ellenberg values were used, percentage concordance also fell short of the arbitrary 90% level. It is unsurprising that a very high level of accuracy could not be achieved by a modelling capability designed to apply generally across many species and vegetation types. As shown previously, a number of sources of error are liable to reduce model accuracy including uncertainty in the calibration equations and species-richness predictions as well as the scaling and sampling error that can leave relatively small soil samples unrepresentative of the heterogeneity in the rooting zones of plant species across a quadrat. It seems therefore unlikely that a modelling capability will ever be developed that can always predict over 90% of the species present in any Priority Habitat patch. More detailed predictive models could inevitably be built based on additional measurements of explanatory variables on specific sites but such models would then lack general applicability across many sites.

This is not a counsel of despair however, since there are other questions of probably greater relevance to policy-users and site managers, that can be addressed without having to predict the species composition of an entire community and its species richness. In particular, the CSM approach conveniently reduces the number of species of interest into a restricted range of indicators. Some will already be present in a targeted Priority Habitat patch at time 1 so that model application could be focussed on predicting changes in habitat suitability *for species already present*. Key CSM indicators that are not present can also be analysed to determine predicted changes in habitat suitability. Such an assessment would highlight prospects for establishment and persistence of each species should it appear in a monitored patch. Answering these questions does not require the species composition of a patch to be predicted from scratch nor does it require application of the species richness model. The approach is to predict change in habitat suitability of species already present and to separately tabulate predicted change in species absent from the patch at time 1 but that are predicted to find habitat conditions highly suitable. In order to test the applicability of models to this approach, the final set of tests focussed on the consistent prediction of *change* over time and thus avoid the need to create a simulated community based on pseudo-quadrats and predicted species richness.

Conclusion: Performance varied greatly between sites for reasons that can be listed overall but would be costly to research across sites. While some sites had relatively high success rates, it is unlikely that both general yet highly accurate models can be developed even with much more work. This is because of the dependence of current species composition on site-specific aspects of patch and wider landscape history.

2.4.9 Predicting temporal change among species already present at time 1

Three sites are examined in turn, each of which was thought to have had a long enough time series to enable detection and attribution of vegetation and soil change driven by changes in N and S deposition.

Approach to model testing

Changes in species abundance over time comprise cyclic, random and directional components. The expectation is that model predictions based on N and S deposition history driving MAGIC and then GBMOVE ought to explain a significant fraction of directional change, assuming a pollutant deposition signal is not eclipsed by other effects such as succession, sampling error and the weather. A combined test of observed versus expected change across all species in the observed datasets was carried out for each site based on a comparison of the slope coefficients for each species when a linear regression line was fitted to observed abundance across each year and predicted habitat suitability across the same time period. Contingency tables were used to simply determine whether most species that were observed to increase over time were predicted to do so and *vice versa*.

At Rothamsted, the known decrease and then recent increase in productivity in Park Grass control plots suggested that change in some species might not be linear. This was certainly demonstrated by Dodd et al (1995) for visual species cover assessments up to 1991. Although the observations used in the tests reported here stopped at 1976 and therefore prior to the recent trend for increasing productivity (Williams 1978), quadratic models were tested over linear models so that species trends best fitting the former could be separately compared with observations.

So as to ensure that all species and observations could be compared graphically, both predicted habitat suitability values and observed abundance changes were standardised to between 0 and 1 by dividing each yearly value by the maximum value across the years for which values occurred. This means that directions of change in observations and predictions can be compared on a standard basis. Such standardisation does not however obscure the possibility of poor fits between observations and predictions due to scale effects or the multitude of other unexplained influences, such as the weather, herbivory and recorder error.

Significant results for chi-square tests of contingency tables and correlations between slope coefficients would suggest that a) changes among taxa appeared consistent with a pollutant deposition signal, and b) that the model chain can generate predictions consistent with this signal against which observations can be usefully compared.

Moorhouse

Observed test data from the Hard Hills control plots coincided with a period during which MAGIC predicted an increase in soil pH and a steady decline in soil C/N ratio (Fig 22). Recent intensive ECN monitoring has indeed shown a consistent increase in pH whose rate of change actually exceeded that predicted (Fig 23). While this trend is partly consistent with recovery from previous S deposition and with the eutrophying effect of N, it is also thought to be consistent with a marked trend toward warmer winter and early Spring temperatures (J.Adamson pers.comm., Holden & Adamson 2002) as well as the deposition of sea salt. In passing, it is also worth noting the periodic downward spikes in pH at Moorhouse (Fig 23). These have been attributed to sulphate release during very dry summers.

Fig 22. MAGIC prediction of change at Moorhouse in response to FRAME/GANE modelled deposition history. The yellow and blue triangles are observed soil measurements to which MAGIC calibrates. Vertical lines indicate the interval covered by test data from the control plots of the Hard Hills experiment.

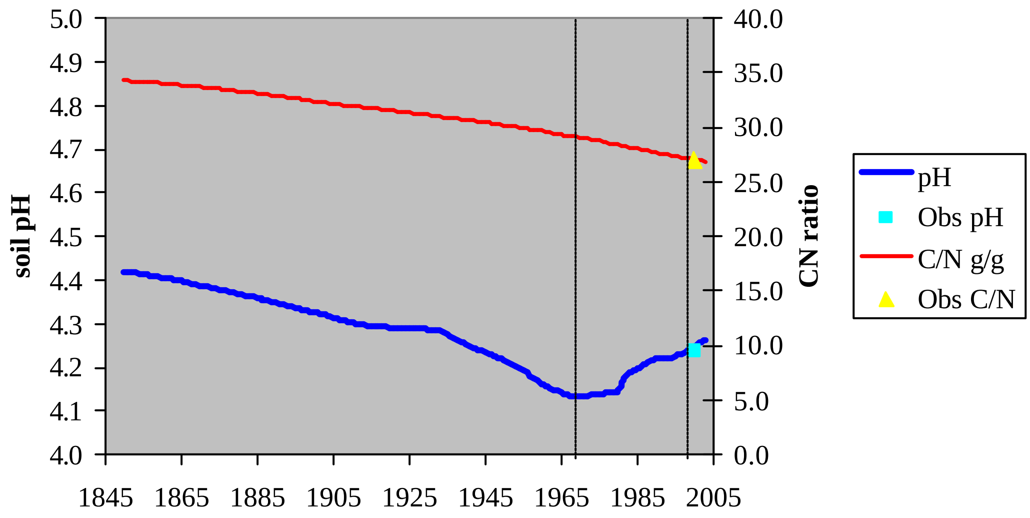
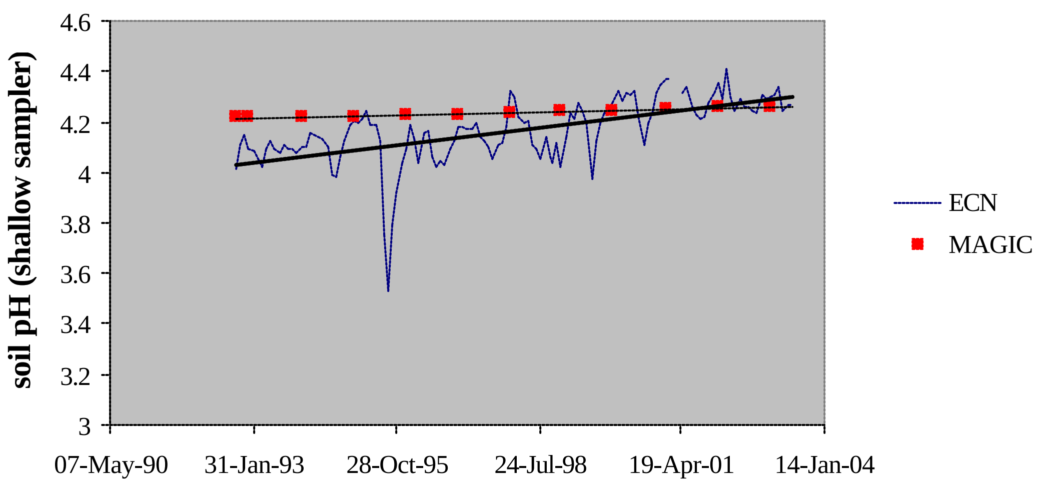


Fig 23. Measured and modelled change in soil water pH at Moorhouse from 1992 to 2003. ECN sampling plus MAGIC predictions.



Despite considerable scatter there was a positive correlation between observed change in species frequency and predicted change in habitat suitability at Moorhouse (Fig 24). The reasons for the residual variation are illustrated by some example plots for individual species in Fig 25. A chi-square test of observed versus predicted directions of change was significant (p=0.016). While the correlation between observed and predicted slopes was also significant, predicted rates of change covered a narrower range than observed species changes.

Fig 24. Predicted versus observed change in individual species in the Moorhouse Hard Hills control plots. Predicted change is the slope coefficient of a linear regression on occurrence probabilities predicted by MAGIC+GBMOVE for each year between 1973 and 2001. Observed change is the slope coefficient of a linear regression on % frequency in sample plots in each survey year. Pearson correlation coefficient = 0.568, p=0.002.

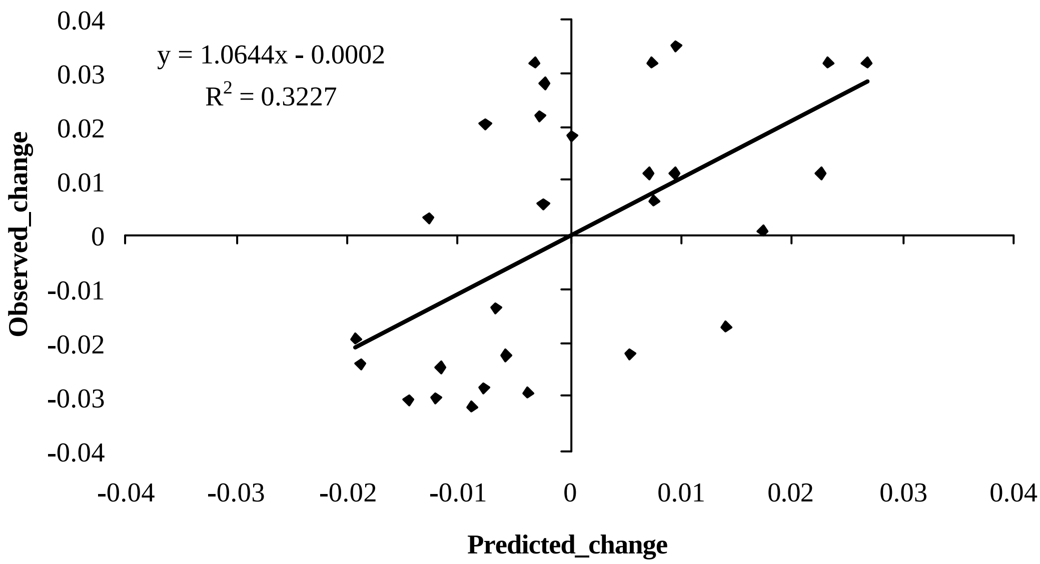
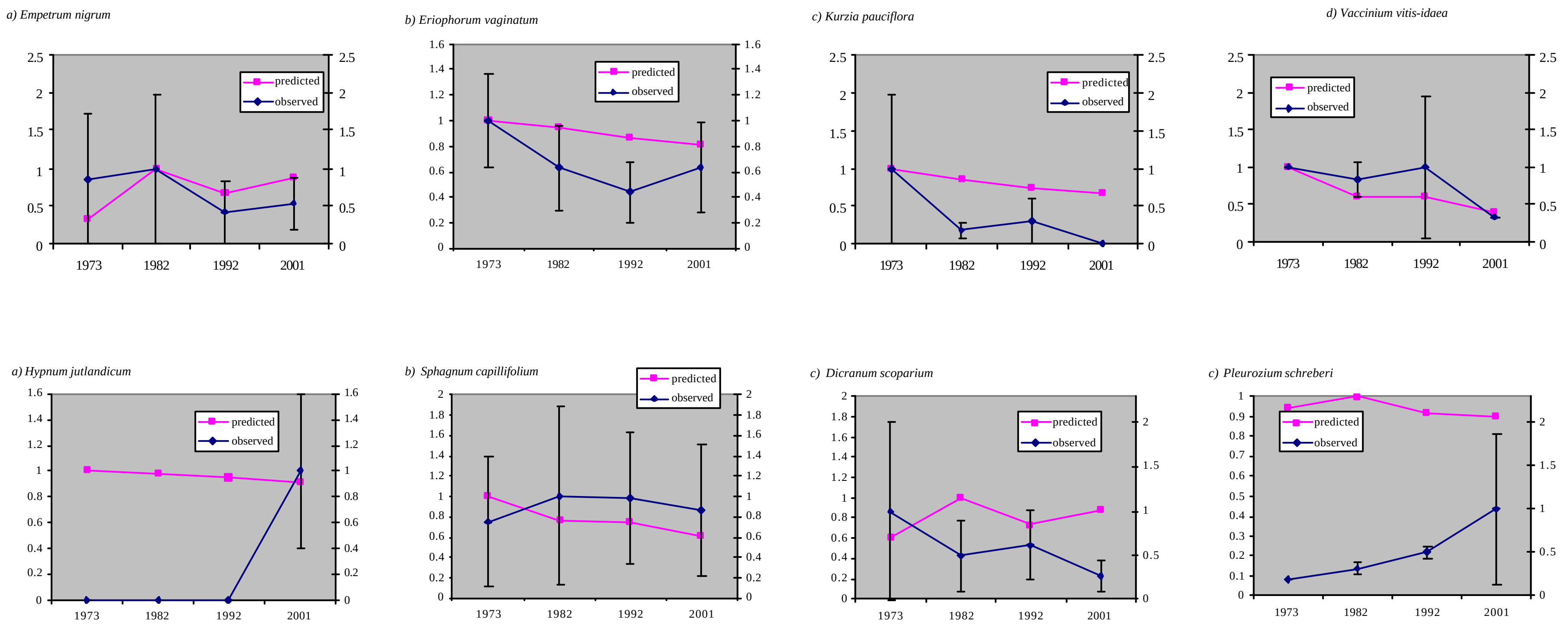


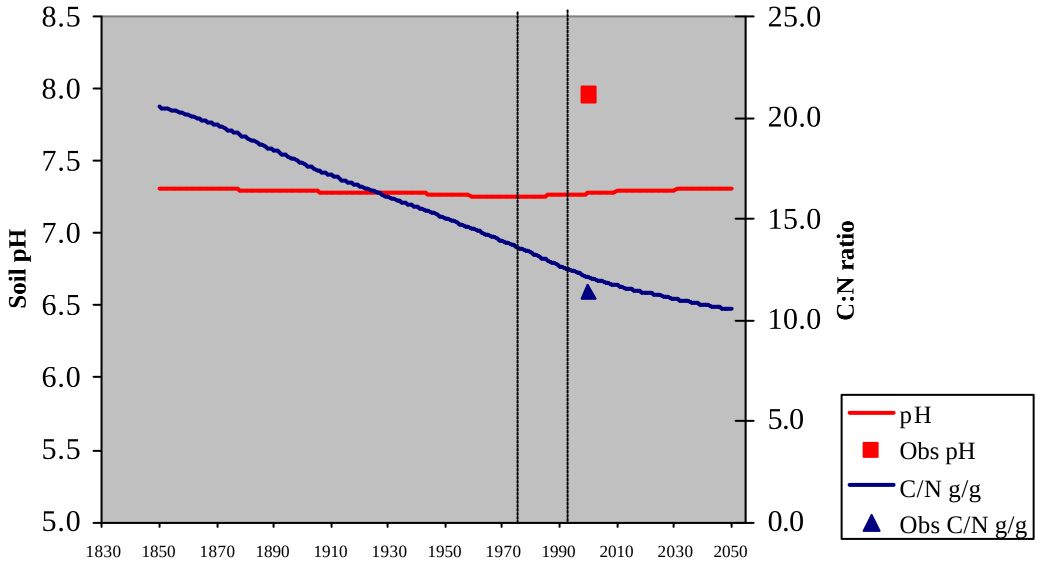
Fig 25. Examples of individual species changes at Moorhouse. Frequency and predicted probabilities were standardised to range between zero and 1 across the time series to enable comparability of the direction of change. a-d are good model fits, e-h poor model fits. The standard deviation of observed counts are shown.



Porton Down

The MAGIC prediction for the sampling period showed a continuation of a long term decline in soil C/N ratio implying increased fertility. Soil pH was predicted to increase to a minor extent however MAGIC did not calibrate to the current pH measurement and so could under or overestimate change in habitat suitability (Fig 26).

Figure 26. Predicted change in soil C/N and pH in response to modelled N and S deposition at Porton Down. Vertical lines indicate the interval covered by test data from monitoring plots recorded by T.C.E. Wells in 1974 and by ECN sampling in 1991 and 1994.



ECN recording is fragmentary for soil pH at this site and does not provide enough data to adequately validate the MAGIC prediction of change (Fig 27). Measurements indicate a good deal of variability and the time series is probably too short to reach any firm conclusions.

Fig 27. ECN pH sampling at Porton Down versus MAGIC predictions for the same period.

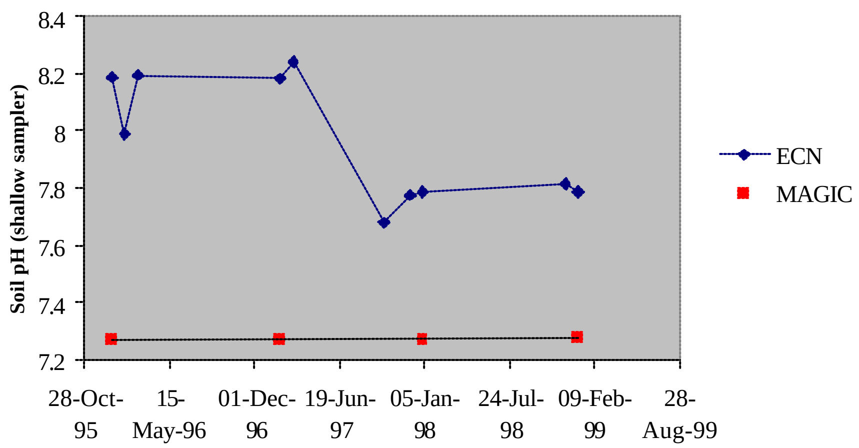
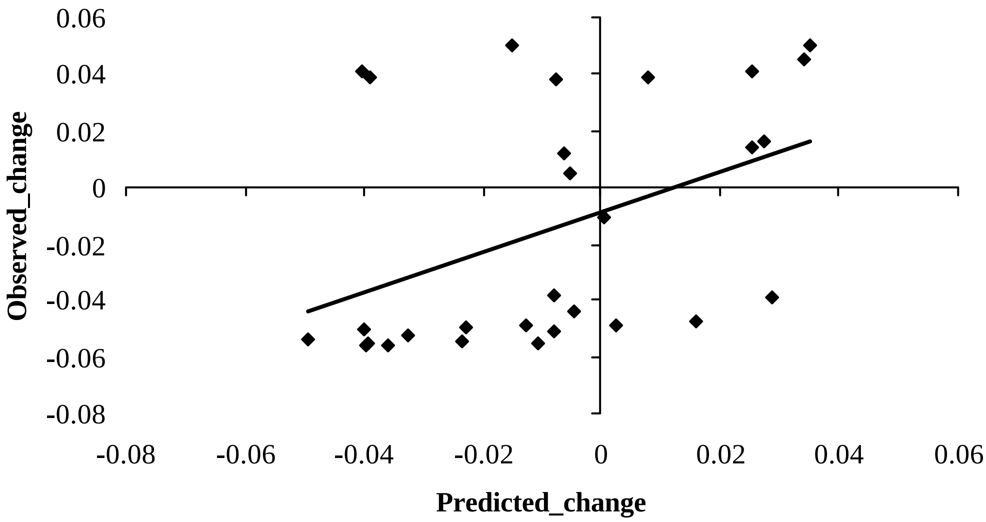
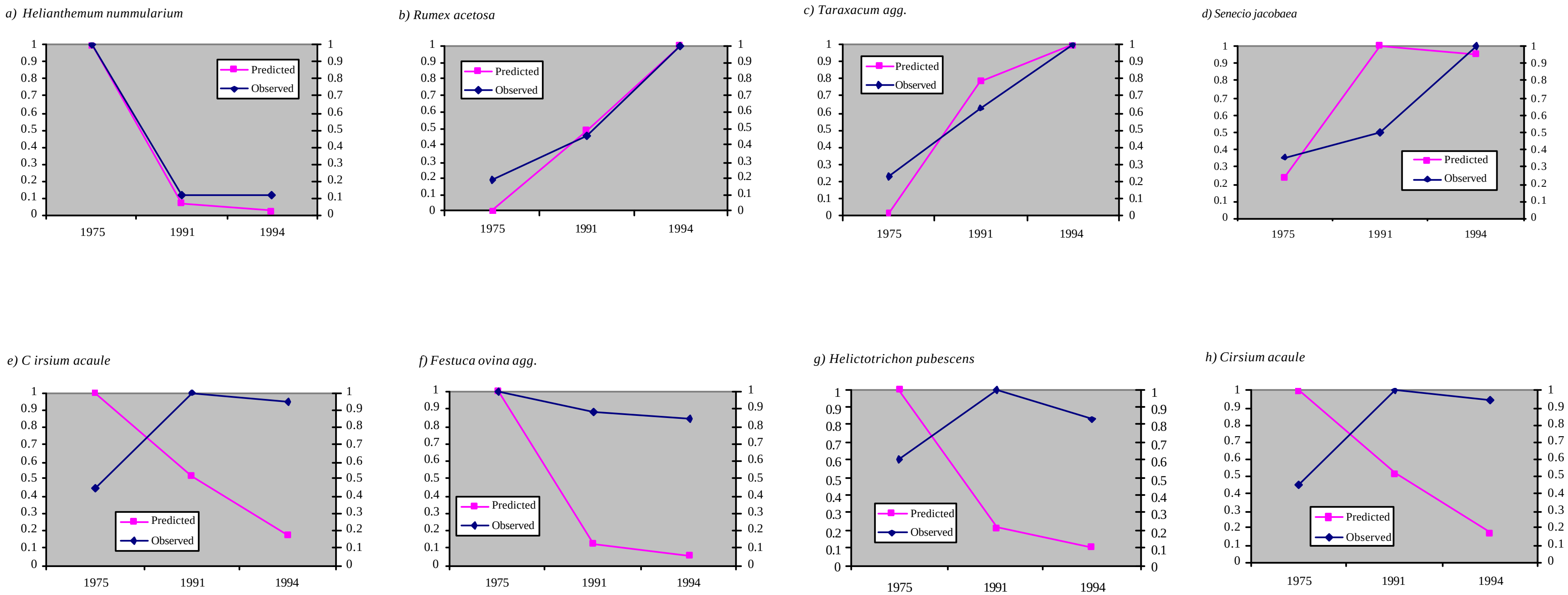


Figure 28. Predicted versus observed change in individual species at Porton Down National Nature Reserve. Predicted change is the slope coefficient of a linear regression on occurrence probabilities predicted by MAGIC+GBMOVE for each year between 1975 and 1994. Observed change is the slope coefficient of a linear regression on % frequency in sample plots in each survey year; 1974, 1991 and 1994. Pearson correlation coefficient = 0.431, p=0.020.



A significant positive correlation was seen between observed and predicted change at Porton Down, despite considerable scatter about the $y=x$ line (Fig 28). However, a chi-square test of directions of change was not significant (p=0.14). Examples of modelled species are shown in Fig 29.

Fig 29. Examples of individual species changes at Porton Down. Frequency and predicted probabilities were standardised to range between zero and 1 across the time series to enable comparability of the direction and magnitude of change. a-d are good model fits, e-h poor model fits.



Rothamsted Park Grass control plots (annual hay offtake with no fertilizer addition)

Considerable time was spent attempting to generate a satisfactory MAGIC simulation that not only calibrated to current soil measurements but that also fitted a unique time series of historical soil data (Williams 1978; Paul Poulton unpublished data; Warren & Johnston 1964). Problems centred on the fact that in order to balance atmospheric N inputs with known yield figures and N content in the hay, a much more productive system was predicted at the start of the time period in 1850. Analysis of model uncertainty indicated that estimates of N output or input must have been in error. When estimates of dry deposition of N were obtained from CEH Edinburgh (D.Fowler pers.comm.), a better simulation of historical C/N measurements was obtained (Fig 30). Soil pH is still overestimated by MAGIC partly because historical data refer to soil slurry pH whereas MAGIC predicts soil water pH. The best final MAGIC simulation is shown below.

Fig 30. MAGIC simulation of historical change in soil C: and pH at Rothamsted Park Grass.

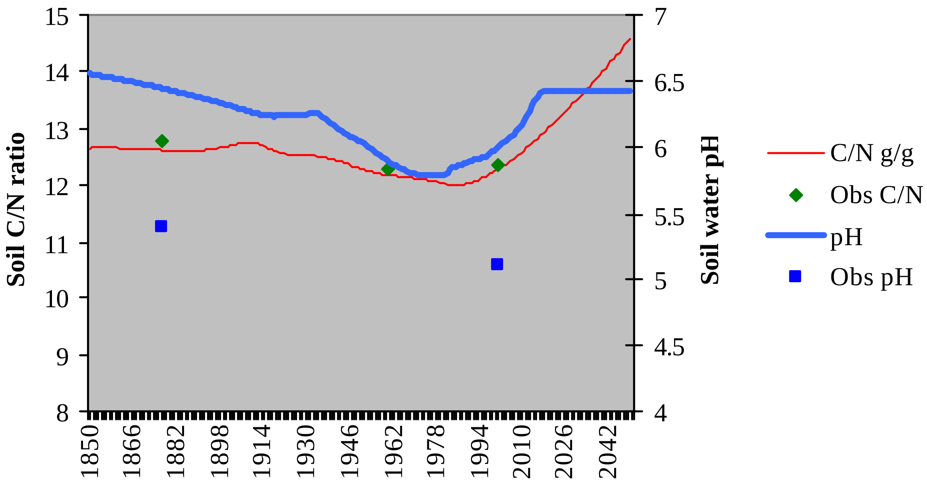
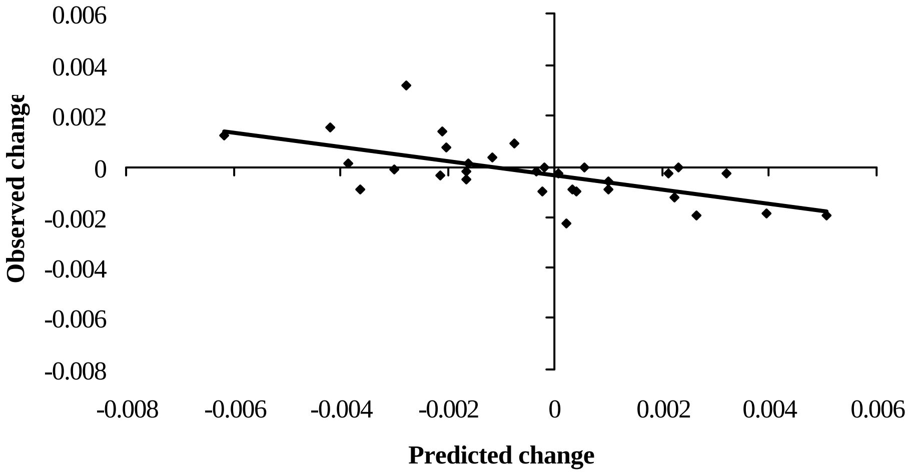


Fig 31. Predicted versus observed species trends at Rothamsted Park Grass based on the MAGIC run above that showed the best fit to observed soil measurements data.



Predicted directions of change in species suitability tended to be the reverse of changes actually observed resulting in a negative correlation between observed and predicted change (Fig 31). This is because historical soil C/N measurements actually saw a small decline over the period of observation (1862-1976) suggesting increasing fertility contrary to the observed decline in yields and consistent switches in the contribution of species to harvested biomass (Williams 1978). Because only three historical soil C/N measurements were available for comparison, the test of observed versus predicted soil data is weak.

Conclusion: Two out of three sites with long time-series showed significant positive correlations between observed and predicted rates of change among species present in monitoring plots. Only at Moorhouse did this relationship have a slope close to 1. Because the signal of pollutant deposition impacts is probably small, high r-sqrd values were not to be expected. However, the evidence supporting model performance remains weak at Moorhouse, very weak at Porton Down and completely contradictory at Rothamsted because at this site historical soil C/N ratio decreased by a small amount even though the grassland system saw a reduction in fertility as N offtake exceeded inputs. These results are equivocal. They offer only weak support for the modelling approach and indicate that more testing is required against observed spatial or temporal vegetation change.

2.4.10 Testing and validation of SMART/SUMO in UK habitats

This section reports the testing and validation of SMART-SUMO after implementation of extra features required to optimise application across British habitats (Table 7). All validations were carried out on sites in the UK, using the regional version of SMART-SUMO. This version needs only minor site specific input. Site specific data were used for pollutant deposition, soil type, soil chemistry, management and annual average temperature.

Table 7. Modifications in SUMO and the performed tests.

SUMO modification	Tested	Validated	Site(s)
More complex sequence of management	Yes	Yes	Rothamsted, Moorhouse
Split up of the functional type herbs/grasses into grasses, herbs and legumes	Yes	Yes	Rothamsted
Differences in vegetation growth due to climate; effect of temperature	Yes	No	But tested at Moorhouse-
Effect of fire	Yes	Yes	Moorhouse, Ruabon
Parameterisation of SUMO for UK, including P modelling	Yes	Yes	All sites

Table 8. Available and missing data for test sites.

Site	Missing key data
Wardlow hay cop	Groundwater table
ECN site (Moorhouse)	Groundwater table
Moorhouse – burning/grazing experiment	Groundwater table
Pwllpeiran	Groundwater table
Rothamstead	Groundwater table
Ruabon	Biomass after the burn

Validation

The validation was carried out for several sites, where different measurements were available. For each site the results are shown and discussed.

Wardlow hay cop.

Wardlow hay cop is a grazed meadow, where a fertilisation experiment was carried out. We used the data from the control of the calcareous grassland experiment. The data used for the validation were retrieved from Caroll et al. (Environmental pollution 121, 363-376). The results for aboveground biomass and N content of the aboveground biomass are shown in figs 32 and 33. The aboveground biomass is slightly underestimated, but well within the standard error. The N content is estimated very well.

Moorhouse - ECN site & Hard Hills burning and grazing experimenmt

This is a nature reserve, with very extensive grazing that stopped in 2000. SUMO was use to simulate chnages in the unburnt control plots and in the burnt plots. The site specific temperature was used for the simulations. In control plots, the effects of grazing (0.1 sheep/ha) are negligible. More surprising is the fact that the herbs/grasses become totally overgrown by the dwafshrubs (heath, Fig 34). In the treatment plots experimental site burning is applied in two intervals (every 20 and 10 years, with the last burn in 1994 and 1995 respectively). The effect of burning on the biomass is clearly visible (Fig 35). The biomass becomes totally dominated by dwarf shrubs for the unburned site. This is also the case for the burned sites, except for the first years after the burn.

Compared to the measured data there is a mixed result. The total biomasses for the ungrazed situation compared to predicted total biomass are a reasonable match. Yet, for the grazed plots the model overestimates the present biomass. The model is initiated with a very low number of sheep/ha (0.1). It is doubtful that such a low density can have an immense impact on the biomass. In fact much higher grazing pressure is applied to the bog in the season post-burn. This is because the stimulation of new *Eriophorum vaginatum* shoots provide an unusually nutritious bite – called the 'moss crop'.The approximate grazing intensity in this period is however unkown (J.Adamson pers.comm.) except that it is substantially higher than the 0.1 sheep per hectare that usually prevails.

Pwllpeiran

The Pwllpeiran site is an experimental site on peat where grazing density and N addition are combined. In table 9 the measured and simulated biomass averaged for the N load is given for 1999. The total biomass for the non-grazed and with 4 sheep/ha grazed sites are simulated quite well. The simulated total biomass for the highest sheep density is however, overestimated. This may be caused by a too high density of sheep in the simulation. The biomass simulation per functional type is not very accurate. Production of dwarf shrub is overestimated for every grazing density and the amount of grass is underestimated. The deviations from the measured values tend to become smaller the longer the model run, i.e. the biomass of the grasses/herbs tends to become higher at the expense of the dwarfshrubs, especially at higher grazing densities.

Table 9. Measured and simulated aboveground biomass in 1999 for three grazing regimes for Pwllpeiran. The shown values are the averages for the four N treatments (0, 10 and 20 kg NH4 and 20 kg NO3).

functional type	**non-grazed**		**4 sheep/ha**		**8 sheep/ha**	
	measured	simulated	measured	simulated	measured	simulated
grasses/herbs	3.25	0.06	2.17	0.52	2.12	0.51
dwarf shrubs	1.78	5.30	1.94	4.05	0.54	0.87
sum	5.03	5.36	4.11	4.59	2.66	1.44

Rothamsted

This is a long term grassland experiment that started in 1850. Data were retrieved from Jenkinson et al (1994) and additional information on yield, deposition and hay N content, provided by CEH. The site was extensively fertilised till around 1890, which was included in the model run. The fertilising stopped in 1890. Fig 36 shows the simulated and observed biomass harvest (the latter as an average of ten years). The biomass is slightly underestimated, but reacts pretty well to the higher N deposition occurring after 1940 and is a good validation of the model performance.

Modelled estimates of the proportions of grasses, herbs and legumes in the hay crop were consistent with observed data, although modelled changes were not tracked by equivalent changes in observed data (Fig 36). This maybe due to the inherent stability of the Rothamsted grassland community (eg. Wilson et al 1996; Dodd et al 1994a).

Ruabon

This site was burned in 2000. Unfortunately, biomass data are only available from before the burn in 2000 (Table 10). As for the Pwllpeiran site the total biomass was simulated quite well, but the amount of woody parts was underestimated and the amount of leaves overestimated. The N and P contents are simulated quite well, although the N-content of the woody parts is underestimated (as is the P-content) and of the leaves overestimated. The effect of the burn is visualised in Fig 38. The biomass amount decreases tremendously, but regrowth starts directly afterwards and is quite large. As found for other sites the simulation shows more or a less a monoculture of dwarf shrubs; grasses and herbs tend to be outcompeted. As for all the vegetation types SUMO simulates biomass for shrubs and two tree species as well. The amount is minor, but not negligible, though during the run they become suppressed by the dwarfshrubs. In the Dutch situation succession to forest most likely would occur. This is not happening here because the model is initialised with less biomass for seedlings than in the Dutch situation. Moreover the seed input is much lower. This all assumes that there are no seed sources in the neighbourhood. If that is not the case initialisation of the model should be different, which may give different results. The run for this site was repeated assuming that a ‘normal’ amount of seed (300 g/ha/y) of trees reaches the site. The biomass for the trees is higher than in the shown plot, but still insignificant, trees were not able to establish themselves in the plot even after the burn until 2020 (the end of the model run).

Table 10. Measured and simulated biomass and N- and P-content for the Ruabon site in 2000.

	biomass (ton/ha)		**N-content (%)**		**P-content (%)**	
	measured	**simulated**	**measured**	**simulated**	**measured**	**simulated**
wood	19.7	15.7	0.36	0.25	0.03	0.014
leaves	5.2	7.2	1.35	1.48	0.09	0.088
total	24.9	22.9				

Forest Level 2 plots

The level 2 plots consist of ten sites spread over the UK. They are all forest plots, with either oak, spruce or Scots pine as planted dominant tree species. The measured and simulated N-contents of wood and leaves of the oak are given in table 12, the regression analyses for all the sites is given in Appendix 1. Although the simulated N-contents for wood is in the same range as the measured values the correlation is not significant (Appendix 1). Especially the simulations for oak are poor, where the N-content is underestimated. The relation between measured and simulated N-content in the leaves is much better and highly significant, though the N-content is overestimated (Appendix 1). The simulations of the biomass are in general not too good (Fig 39-43). The simulations for oak are reasonable, but the simulations for the fast growing sites with Scots pine and sitka spruce are very poor. The realised growth in the field is much higher than where SUMO originally was built for, and this may cause part of the large differences. There may also be differences in the strains of cultivated tree species such that fast-growing, high rainfall-tolerant genotypes selected for timber in upland Britain require different growth parameters than cultivars typically selected for Dutch plantation.

Table 11. Measured and simulated N content for wood and leaves of the planted tree species for ten sites in 2000.

site	**species**	**N-content wood trees (%)**		**N-content leaves Trees (%)**	
		measured	**simulated**	**measured**	**simulated**
Alice Holt	Oak	0.31	0.08	2.58	3.71
Savernake	Oak	0.35	0.09	2.55	3.70
Lakes	Oak	0.29	0.14	2.25	1.84
Thetford	Scots pine	0.08	0.14	1.92	1.83
Sherwood	Scots pine	0.09	0.13	1.82	1.96
Rannoch	Scots pine	0.08	0.13	1.57	1.96
Coalburn	Sitka spruce	0.09	0.13	1.53	1.63
Tummel	Sitka spruce	0.11	0.20	1.56	2.53
Loch Awe	Sitka spruce	0.06	0.20	1.55	2.53
Llyn Brianne	Sitka spruce	0.09	0.12	1.44	1.53

Discussion

The results of the validation for the different sites and vegetation types vary widely. For some of the sites the simulation give acceptable results, but less so for for others. In particular, the simulations for some of the forest sites need further investigation before SUMO can be successfully applied for these forests in the UK. This may lead to a different parameterisation or even model changes.

There are several reasons why the performance of SUMO is not as good as we would like.

1. The validation is performed with an absolute minimum of data available from the test sites. If more information about the sites could be gathered the results will most likely improve. However, the more data-demanding the model then the more costly it is to parameterise and hence likely to be less widely applied.
2. Sumo was developed for the Netherlands and many of the parameters and the initial input for the model was not changed. It may be necessary to use deviating parameters and initial values for the UK (as already was done for some sites on a minor scale). This may be maximum growth rate, seed input

(site specific), initial biomass (site specific), maximum and minimum N and P-content in the organs.

3. It is possible that an influential relationship or factor in UK plant community dynamics has not been accounted for in the modified model because it has no or only minor effects on the simulations for the Netherlands. Identification of such a factor, when present, may be difficult. Some obvious candidates include a) the difference in competitive relations between dominant grass species and dwarf shrubs in upland Britain as opposed to lowland Dutch habitats, b) the importance of bryophyte biomass in upland vegetation types, c) controls on tree recruitment due to climate, d) extra parameterisation for individual tree species that are significant in many common British broadleaved woodlands; in particular Sycamore, Field Maple, Lime and Elm.

We think that the model SUMO can be applied in the UK, at least for grassland and heathland and some forest types, but more validation would still be desirable. Further development ought to focus on some of the desirable yet missing features listed above and on the detailed interactions between grazing, burning and climate change in heathland.

Year to year variation in SUMO is also not very well captured, the model gives merely the long term average results. Year to year variation due to differences in precipitation, temperature and deposition can in principle be modelled, but it only makes sense when the information about the variation is available, in the form of yearly measurements or scenarios against which model predictions can be compared.

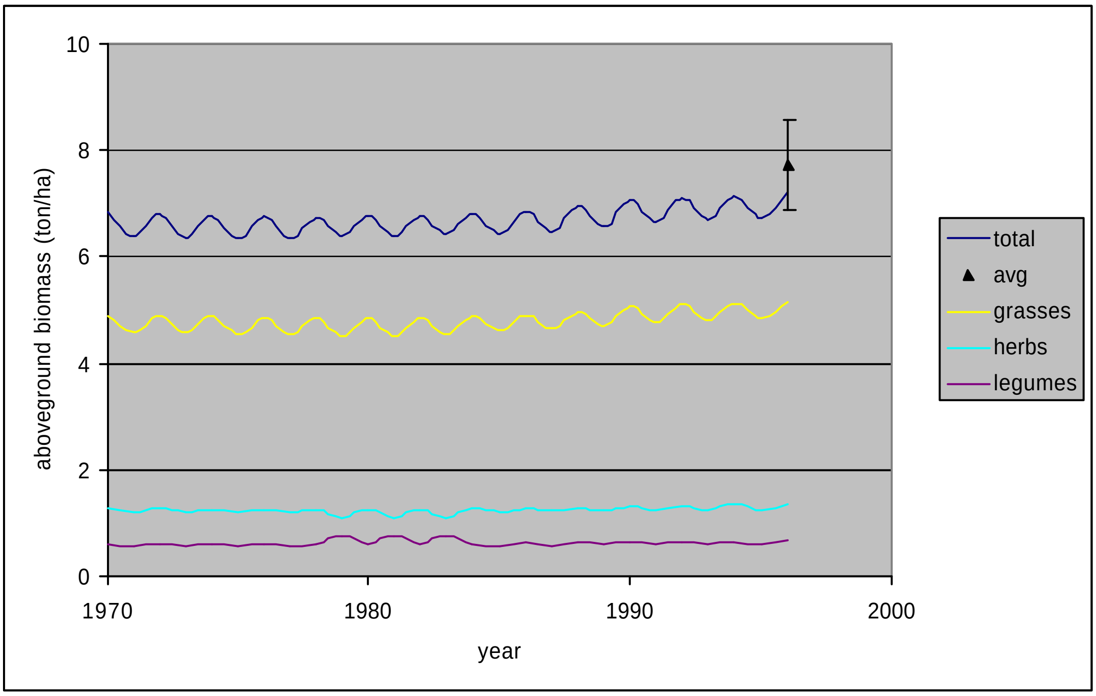


Fig 32. Aboveground biomass simulated by SMART-SUMO for Wardlow hay cop. The total aboveground biomass is broken down for grasses, herbs and legumes. The triangle indicates the measured aboveground biomass in 1996 with standard error.

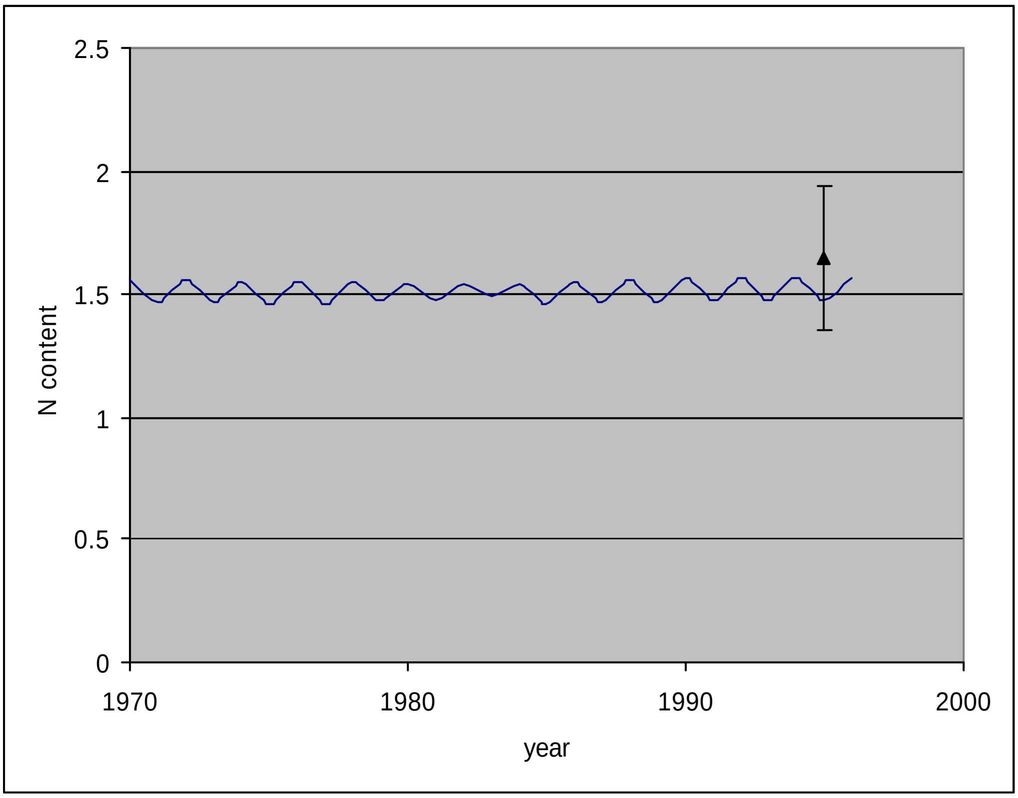


Fig 33. Nitrogen content of the aboveground biomass simulated by SMART-SUMO for Wardlow hay cop. The triangle indicates the averaged measured N content of the aboveground biomass in 1996 with standard error.

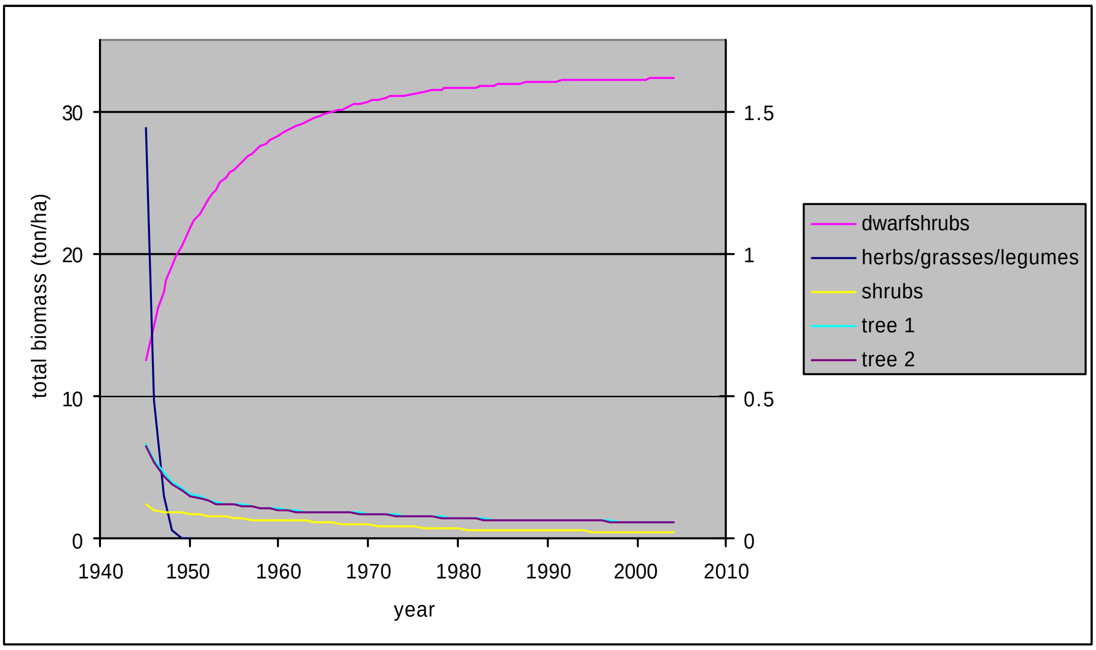


Fig 34. Total biomass per functional types simulated by SMART-SUMO for the ECN site at Moorhouse. The total biomass for dwarfshrubs is given on the left axis, for the other functional types on the right axis.

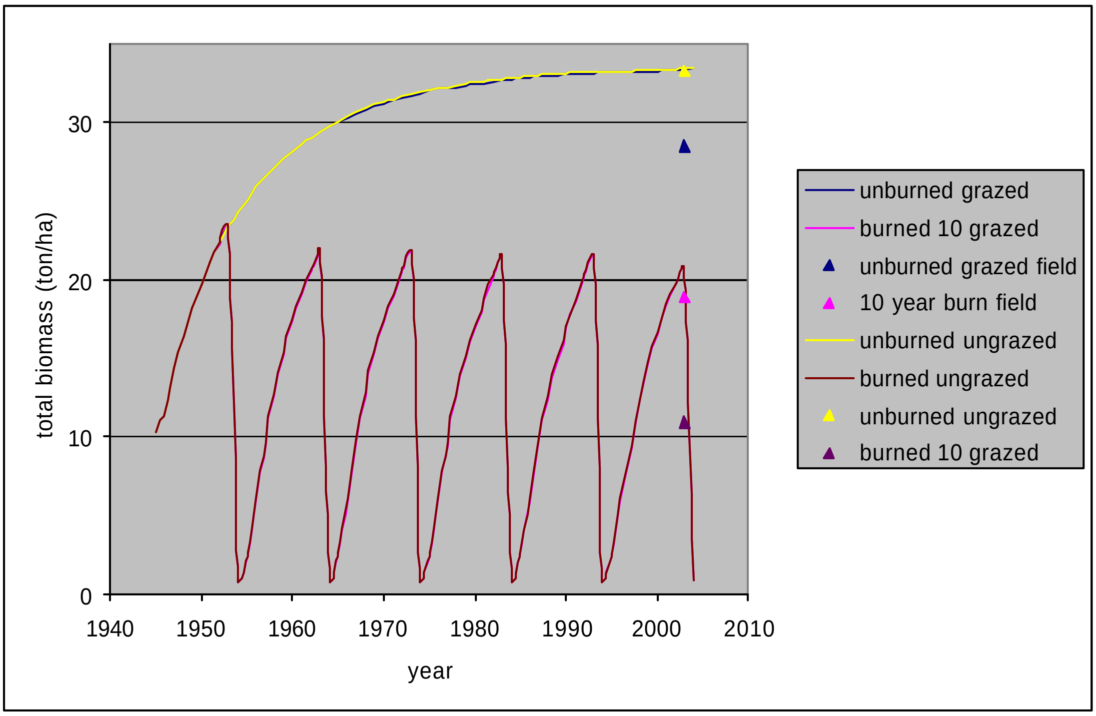


Fig 35. Total biomass at Moorhouse experimental site simulated by SMART-SUMO. The management consist of grazing combined with burning (unburned, burned 10; burned every ten years, burned 20; burned every 20 years). The triangles indicate field measurements for the total aboveground biomass.

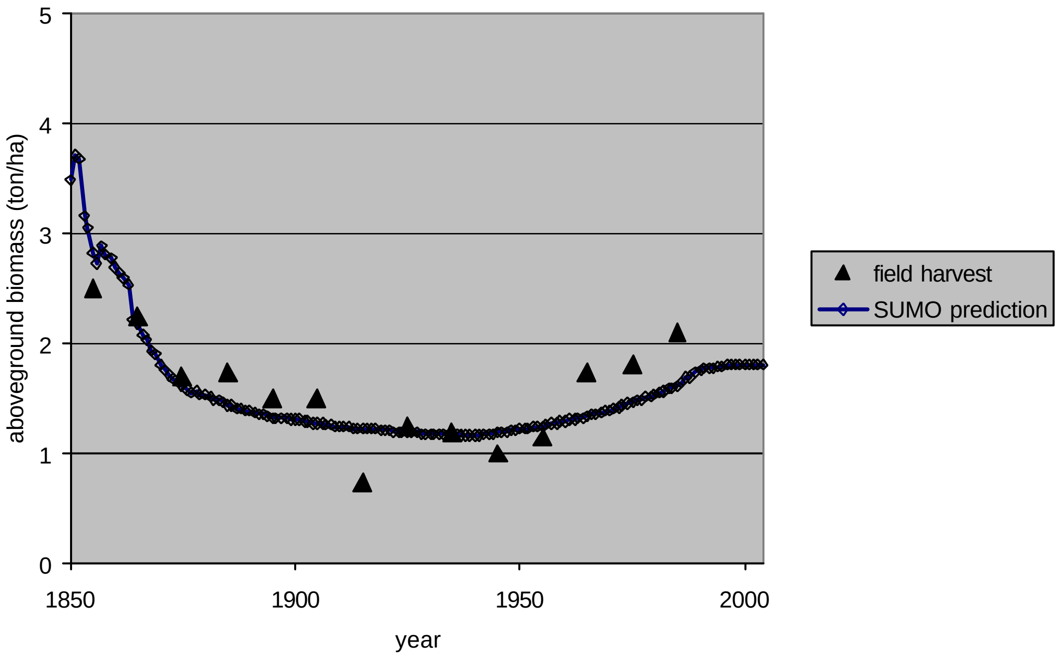


Fig 36. Harvested biomass for the Rothamstead site. The black triangles (with s.e.) indicate the average biomass harvest over the five previous and five following years. The simulated plot is fertilized extensively till 1890 (according to Jenkinson et al. 1994) and mown twice a year.

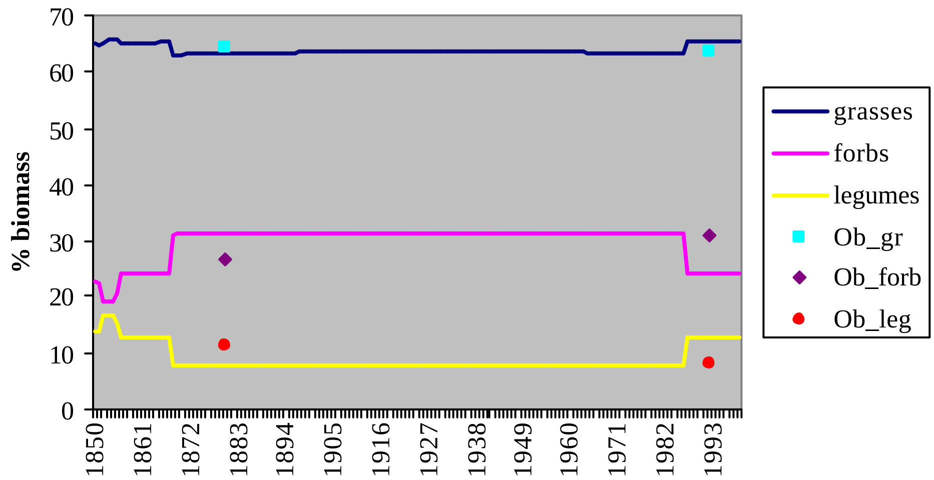


Fig 37. Predicted and obersved changes in the three functional sub-types that make up the herb group in Rothamsted PG control plots. Scenario based on FRAME N and S deposition and observed annual N offtake in hay crop with no added fertilizer after 1890.

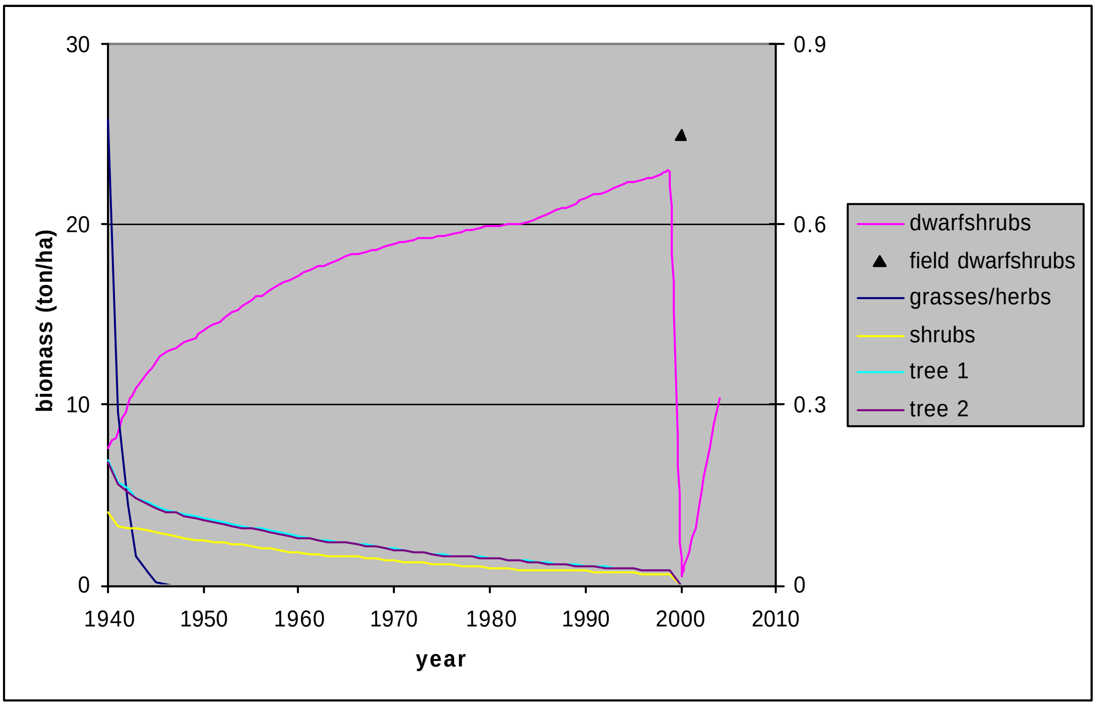


Fig 39. The simulated aboveground biomass for the Ruabon site. The biomass for the dwarfshrubs is given on the left axis, for the other functional types on the right axis. The site was burned in 2000.

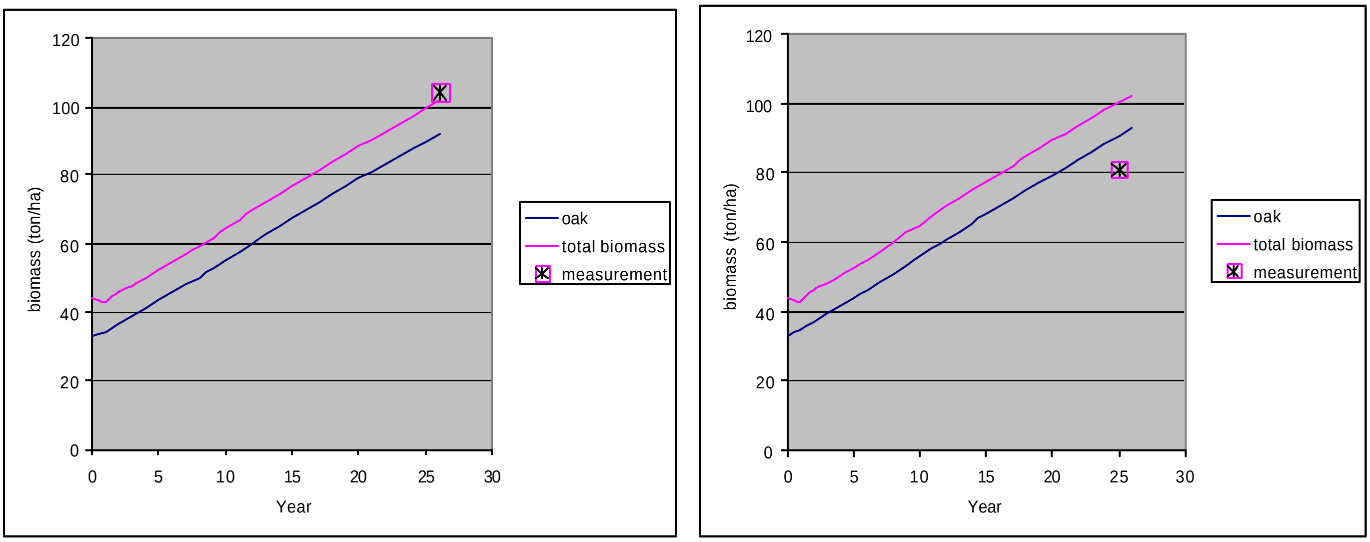


Fig 39. Simulated total biomass and tree biomass and measured biomass for Alice Holt (left; oak) and Savernake (right; oak).

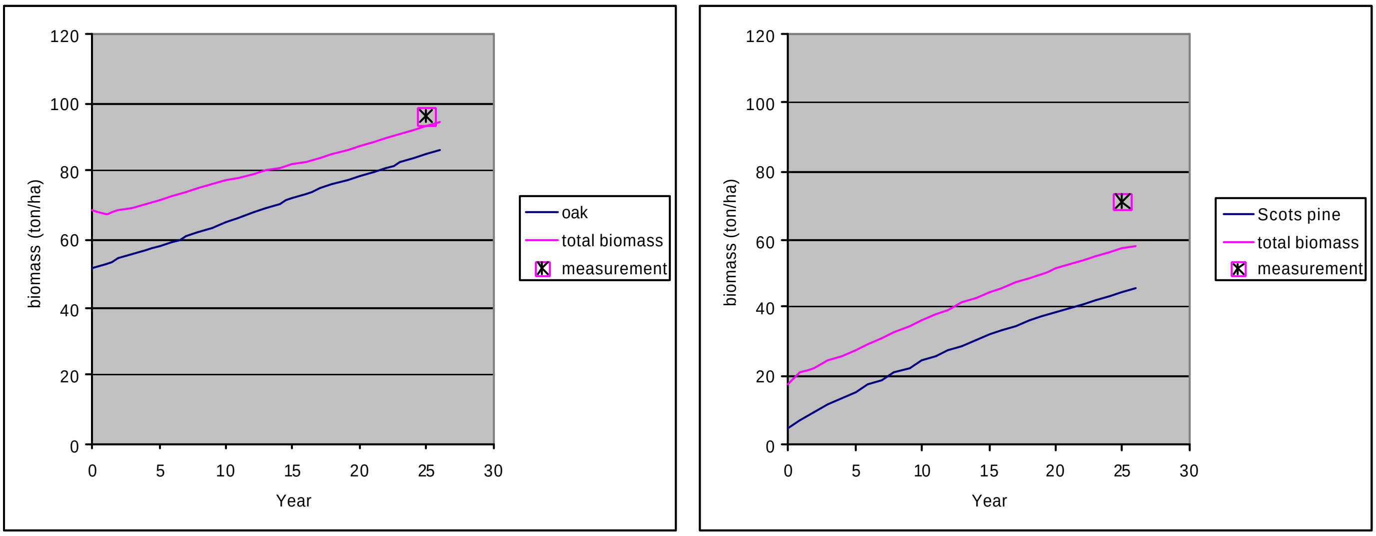


Fig 40. Simulated total biomass and tree biomass and measured biomass for Lakes (left; oak) and Thetford (right; scots pine).

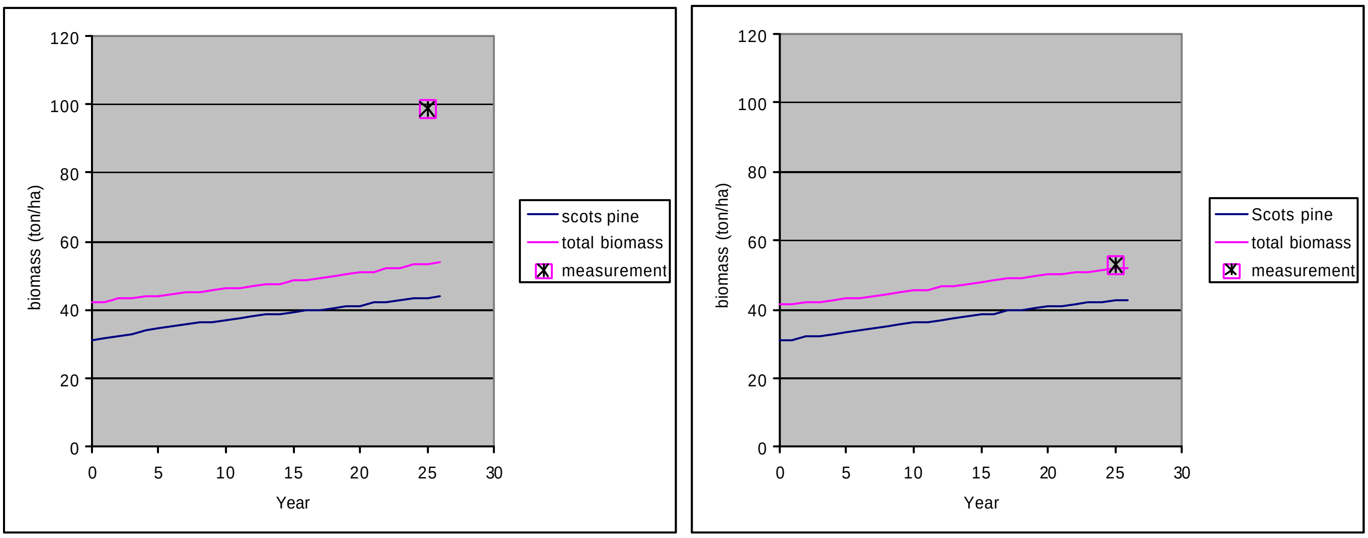
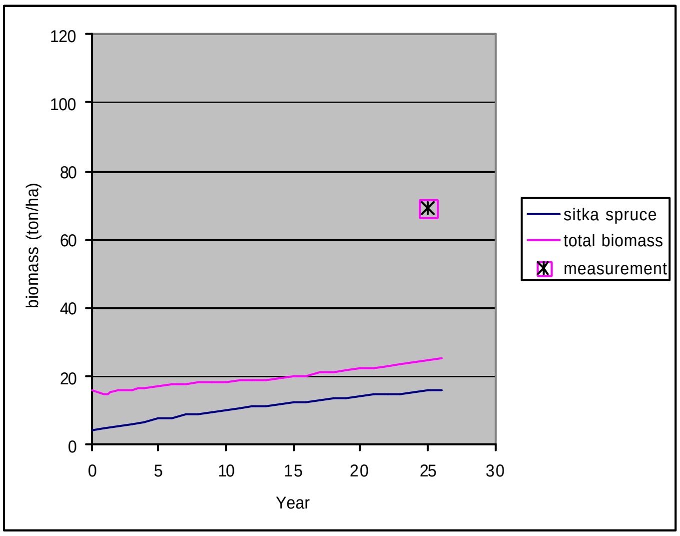


Fig 41. Simulated total biomass and tree biomass and measured biomass for Sherwood (left; scots pine) and Rannoch (right; scots pine).



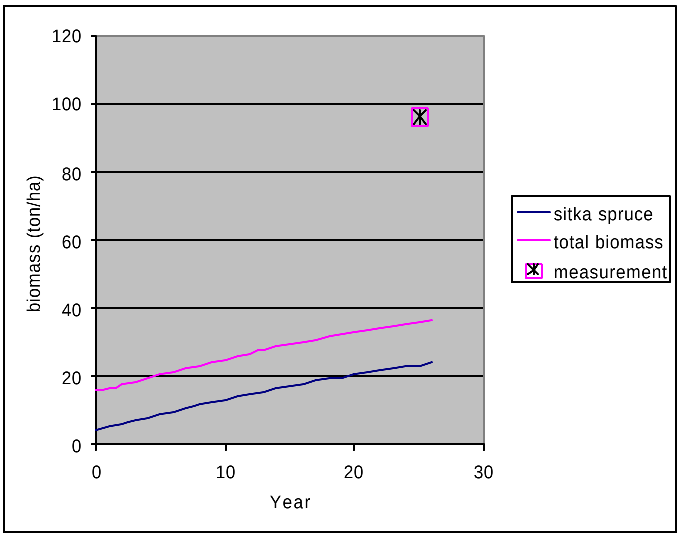
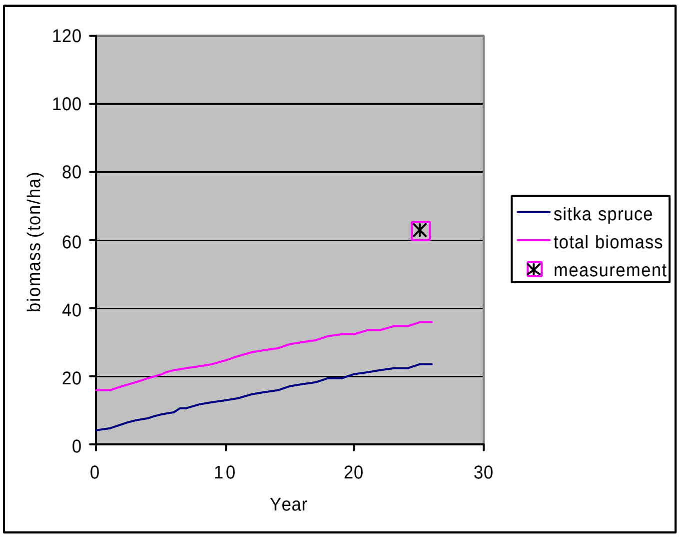


Fig 42. Simulated total biomass and tree biomass and measured biomass for Coalburn (top; sitka spruce) and Tummel (bottom; sitka spruce).



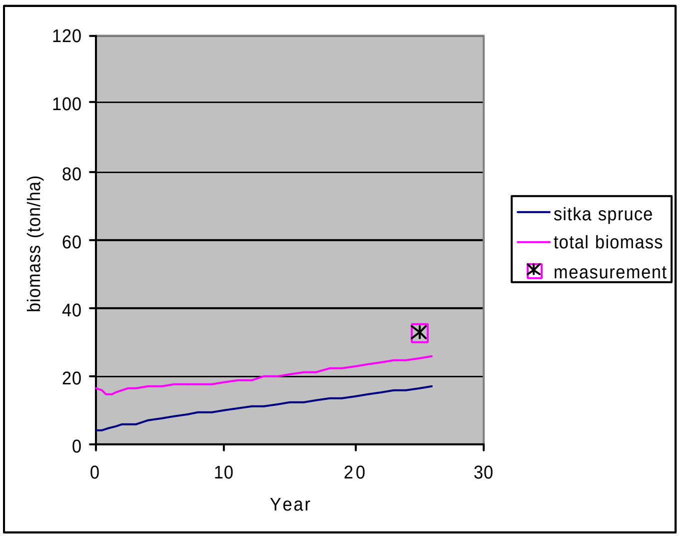


Fig 43. Simulated total biomass and tree biomass and measured biomass for Loch Awe (top; sitka spruce) and LlynBrianne (bottom; sitka spruce).

Appendix 1. Regression analysis of the relation between measured and simulated N-content of wood and leaves for the level2 plots. The regression analyses is carried out with Microsoft Excel.

wood

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.599731
R Square	0.359677
Adjusted R	0.279637
Standard E	0.033283
Observatio	10

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressior	1	0.004978	0.004978	4.493699	0.066842
Residual	8	0.008862	0.001108		
Total	9	0.01384			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Ipper 95.0%*
Intercept	0.167838	0.01834	9.151611	1.64E-05	0.125546	0.210129	0.125546	0.210129
X Variable	-0.205827	0.097096	-2.119835	0.066842	-0.429731	0.018076	-0.429731	0.018076

leaves

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.729983
R Square	0.532875
Adjusted R	0.474484
Standard E	0.579666
Observatio	10

ANOVA

	df	*SS*	*MS*	*F*	*ignificance F*
Regressior	1	3.066459	3.066459	9.126022	0.016536
Residual	8	2.688101	0.336013		
Total	9	5.75456			

	Coefficients	*andard Err*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*ower 95.0%*	*Ipper 95.0%*
Intercept	-0.195287	0.853206	-0.228886	0.8247	-2.162782	1.772209	-2.162782	1.772209
X Variable	1.342184	0.444295	3.020931	0.016536	0.317638	2.36673	0.317638	2.36673

3 Developing indices and datasets for risk assessment of N impacts on UK protected sites

3.1 Introduction

This section of the report explores constraints and opportunities for assembling data on external risk factors for UK ASSI/SSSI. Risk factors were identified as those relating to pressures that could potentially exacerbate the eutrophying effect of atmospheric N deposition on plant communities within designated sites.

Work focussed on identifying and assembling spatially explicit information on risk factors for each of the test sites used in part 1 of the project. This was done to scope the potential for building a database of risk factor information for all UK ASSI/SSSI. While local knowledge possessed by stakeholders in specific areas will serve to define the history and severity of certain types of risk, this information maybe hard to assemble in a consistent format for all designated sites. However, a UK-wide assessment of the state of key risk factors may well be desirable because it provides decision makers with a strategic overview of the designated site series. Such an overview may prove useful for the following reasons:

1. Provides contextual information that can help identify groups of sites where external factors could prevent achievement of SSSI condition targets despite positive site management.
2. Helps plan cross-site monitoring networks since sites can be readily stratified by levels of different risk factors to ensure that gradients of interest are long and well-replicated eg. N deposition, but other key factors are held as constant as possible or classified into groups eg. large increase in growing season length versus little change, small sites versus large, sites with a history of marked agricultural conversion around their periphery versus sites where adjacent management intensity has remained stable.
3. Identifying groups of sites associated with particular patterns of risk helps hypothesise the kinds and rates of ecological change that should be expected. This could aid choice of monitoring methods, target organisms and frequency of recording.

Objectives for this part of the project were as follows:

a) Assemble and present risk factors for all test sites used in part 1 of the project.
b) Determine the feasibility of assembling risk factor information for all ASSI/SSSI in the UK.
c) Assess options for summarising risk factor values for each ASSI/SSSI

3.2 Favourable Conservation Status and external factors

Defining Favourable Conservation Status (FCS) on designated sites requires knowledge about the status of factors affecting site interest features. These include factors internal and external to the site (Alexander 2003). Internal factors are probably easier to assess by agency staff. In particular, site management and its effects will be more apparent and easier to influence. Management prescriptions maybe agreed with land-owners or competent authorities such as Wildlife Trusts, while management impacts maybe recorded as part of the site monitoring process. A key difficulty however, is the attribution of past change and the prediction of future change to different potential drivers. While simple techniques such as the establishment of grazing exclosures within a site, can provide invaluable evidence of the positive effects of management in comparison with change that would be expected if a reserve was abandoned, it is much more difficult estimating how managed change will interact positively or negatively with external factors such as increases in growing season length (Hossell et al 2000), atmospheric pollutant deposition (Achermann & Bobbink 2003) and exposure to enriched agricultural run-off (ECUS 2003). Yet evaluation of the added risk posed by external factors is implied by the definition of FCS, which stipulates that both internal and external factors that are agents of change for species and habitats should be under control (Alexander 2003).

In this part of the project, approaches are developed for estimating external factors that could specifically modify and interact with the increased supply of N to terrestrial ecosystems from atmospheric deposition. The emphasis is on external factors because internal factors are assumed to be more readily apparent to site managers (Alexander 2003) while external factors lack quantification using standard approaches across the ASSI/SSSI series in the UK.

By analysing readily available databases we assess the feasibility of estimating the risk of vegetation change posed by different levels of a series of external factors. Since the focus of the project is on impacts of N deposition, emphasis is placed on exceedance of the empirical critical load for N as a threshold below which different levels of external factors may play out their own impacts but with no antagonistic or suppressive interplay with N deposition (eg. van der Wal et al 2003). This was the approach taken in the MIRABEL project that set out to estimate the risk of eutrophication impacts on Natura 2000 sites posed by combinations of risk factors (Petit et al 2003). A similar rationale was adopted here, however rather than condense risk factor information into a single index, the desirability and feasibility of such a step is examined more critically. Primary importance is given to N deposition because that is the focus of the project and not because it is assumed that this is the most important threat to the SSSI series.

3.3 Initial selection of risk assessment factors

In all cases the identity of the risk factors reflects published evidence of their importance from survey and monitoring or experiments. The risk factors can be considered as likely to amplify or constrain the eutrophying effect of increased N availability. The challenge in assembling datasets on each factor is to achieve consistency of measurements across all UK designated sites but at scales that are not so coarse that sensitivity is lost through averaging across large spatial units. Brief descriptions of the initial list of risk factors follow:

<u>Phosphorous limitation</u>
Phosphorous availability is well known as a constraint on annual primary production at low and high pH (Cunha et al 2002;Venterink et al 2003; Hogg et al 1994; Aerts et al 2001). Its role in constraining vegetation response to increased N supply has also led to its adoption as a modifier to empirical critical loads (Achermann & Bobbink 2003) However difficulties in deriving site and Priority Habitat specific measures of P limitation centre on the lack of a general relationship between P availability and other more easily measured surrogates that can be used at appropriate scales. A simple approach would be to use soil pH as a simple guide to expected limitation. For example nutrient availability is thought to be least limited between 5.5 and 6 while P becomes limiting above pH 7 and below 5.5 (Schaffers 2000). Yet UK soil maps and related soil series attribute data are too coarsely resolved given that soil characteristics, vegetation and P limitation can vary in parallel over scales of 1 – 10m (Boyer & Wheeler, 1989). Applying UK soil map data would therefore probably not substitute for an estimate of the possibility of P limitation simply based on type of vegetation present. After discussion with the project steering group, no further attempt was made to assemble data for this risk factor.

Growing season length (GSL)
Primary productivity is a function of temperature, closely correlated with precipitation, altitude and latitude (Polis 1999). The longer the growing season then the greater the potential response to increased N deposition because more biomass can accumulate each year. This biomass increment is then more likely to reflect enhanced growth of more nutrient-demanding species such as tall grasses and high Specific Leaf Area herbs (Achermann & Bobbink 2003). Thus the main risk posed by increased growing season length is assumed to be in favouring competitive dominants (eg. Dunnett et al 1998). Clearly a whole range of other N deposition and climate interactions are also possible including vegetation impacts such as increased drought and frost susceptibility of community dominants (Caporn et al 2004; Lee & Caporn 2001; Terry et al 2004), drought-induced gap creation and increased community susceptibility to invasion (Buckland et al 2001), and soil impacts such as changes in N mineralization following seasonal or long-term drought (van Vuuren et al 1992). Data was assembled on current growing season length and an estimate of recent change in growing season length.

Site size and shape
Landscape ecological theory indicates that larger areas of semi-natural habitat ought to support larger species populations, which are less susceptible to decline as a result of random dynamics and environmental effects (Woodland Trust 2000; Petit et al 2004). Hence, fragmented habitats will experience species loss as population sizes fall below species-specific minimum thresholds (eg. Lindborg & Errikson 2004; Piessens et al 2005). Other things being equal, sites with greater edge to area are also assumed to be more readily influenced by adjacent habitats both in terms of species immigration and exposure to nutrients and disturbance. The importance of such phenomena is likely to vary between species and ecosystems although few species have been studied well enough to confidently estimate the sensitivity of their dynamics to area and isolation effects. The general principle is well established however, hence site size and perimeter to area ratio are used as risk factors.

Area of donor habitats for nutrient-demanding species
This factor is a surrogate intended to reflect the availability of competitive plant species around the site or habitat patch by measuring the proportional cover of Broad Habitat types known to be associated with disturbance and high productivity. Thus, a polluted habitat isolated from source populations of responsive plant species is hypothesised to be at lower risk of change in plant species composition than a polluted yet small habitat patch surrounded by large source populations of nitrophiles. Area of donor habitats was measured within designated site boundaries and in buffer zones around each site.

Area of vulnerable semi-natural habitats
Conversely, the area of buffering semi-natural habitats around each designated site was also measured. This risk factor is thus consistent with the idea of semi-natural ‘core area’ thought desirable for buffering impacts of land-use and climate change on broadleaved woodland (Woodland Trust 2000) and could therefore be considered applicable to other semi-natural habitats.

Grazing pressure
Its hypothesised importance stems partly from the results of analyses of CS2000 data that revealed a modest but significant incursion of semi-improved grassland plants into

unenclosed upland habitats from 1978 onwards (Smart et al 2005). While the processes behind this require further investigation, it seems likely that increased sheep numbers and improvement of in-bye land have increased the abundance of invading plants in their source areas as well as their chances of dispersal into atypical upland situations. Climate change and increased N deposition could also promote their persistence once arrived (van der Wal et al 2003). Some measure of grazing intensity should therefore feature in the risk assessment alongside information on the presence of donor fertile habitats. Options for testing the approach included using AgCENSUS data at 2km square resolution for GB. Sheep density estimates could also be modified so that only densities that exceeded maintenance levels recommended for the particular Priority Habitat[3] were entered into the calculation of overall risk. For SSSI in unenclosed upland situations we would expect that site managers would be in possession of accurate and current grazing density information. Such local knowledge is bound to be more detailed than AgCENSUS 2km square estimates yet local knowledge is not so easy to assemble in a consistent fashion across all designated sites.

Flood risk

This risk factor should identify flood-plain sites associated with a high flooding risk. Flooding is important because of the potency of flood waters as a vector for dispersal of plant propagules (Bischoff 2002; Geertsema et al 2002), as a force for dramatic vegetation change over short periods (Critchley et al 1996) and as a major source of N and P inputs (Mainstone et al 1994; ECUS 2003). Given the increasing likelihood of extreme weather events in the coming decades (www.heatisonline.org/weather.cfm) it is useful to include flood risk estimates along with the range of other risk factors for designated sites.

Empirical critical load exceedance for Nitrogen

Site estimates can be constructed from the recently revised critical loads for relevant Priority Habitats (Achermann & Bobbink 2003) in combination with modelled N deposition estimates for the 5x5 km UK square containing the site (NEGTAP 2001) or from site measurements.

Agricultural intensification history

Nutrient inputs to designated sites originate from gaseous emissions but also from enriched run-off. The latter source is particularly important for phosphorous transport as well as nitrogen (DEFRA 2002; Heathwaite et al 1996). Therefore information on recent change in agricultural productivity and the current level of productivity should provide useful contextual information on potential exposure to nutrient surpluses and P inputs from peripheral land-use.

3.4 Assembling spatial data sets

A constraint on this work was that necessary spatial data had to be readily available and, if spatial data were to be purchased the cost could not be too onerous. All data except for the agricultural census data were obtained at no cost. Most data were downloaded from online sources. Data have been used under licence terms specified by the data providers.

3.4.1 Designated sites

Spatial boundary data sets for selected designated sites – National Nature Reserves (NNR) and Sites of Special Scientific Interest (SSSI) in Great Britain and Areas of Special Scientific

[3] See Upland Management Handbook at www.english-nature.org.uk/pubs/handbooks/

Interest (ASSI) in Northern Ireland - were obtained from national statutory agencies (English Nature[4], Scottish Natural Heritage[5], Countryside Council for Wales[6] and Environmental and Heritage Service of Northern Ireland[7]) or digitised from paper maps[8] or Ordnance Survey MasterMap spatial data[9].

3.4.2 Donor and Vulnerable habitats

The CEH Land Cover Map of Great Britain (LCMGB) was used to identify Donor and Vulnerable habitats. The LCMGB is a classification of multi-temporal multi-spectral satellite remotely sensed imagery and identifies 22 Broad Habitats (Table 1) at a resolution of 25 × 25 m. Three and seven of the broad habitats were selected to represent Donor and Vulnerable habitats respectively (Table 12). The proportion of donor and vulnerable habitats within and surrounding the designated sites has been quantified using a custom macro run within a GIS system.

Table 12. Broad Habitats identified by the LCMGB and used to identify Donor or Vulnerable habitats.

Broad habitat	Donor habitat	Vulnerable habitat
Broad-leaved woodland		✓
Coniferous woodland		
Boundaries and linear features		
Arable & horticultural	✓	
Improved grassland	✓	
Neutral grassland		
Calcareous grassland		✓
Acid grassland		✓
Bracken		
Dwarf shrub heath		✓
Fen, marsh and swamp		✓
Bog		✓
Standing water/canals		
Rivers and streams		
Montane habitats		✓
Inland rock		
Built up areas, gardens	✓	
Supra-littoral rock		
Supra-littoral sediment		
Littoral rock		
Littoral sediment		
Inshore sublittoral		

3.4.3 Growing season length

Growing season length spatial data have been obtained at European and national scales. Rötzer and Chmielewski (2001) predicted the spatial variation in growing season length

[4] http://www.english-nature.org.uk/pubs/gis/GIS_register.asp
[5] Site boundaries for the Caingorms SSSI and NNR were obtained directly from Scottish Natural Heritage.
[6] http://www.ccw.gov.uk/ccwdigitaldownload/index.html
[7] Site boundaries for the Dead Island Bog ASSI were obtained directly from the Environment and Heritage Service of Northern Ireland.
[8] Dromore Motte.
[9] Climoor and Rothamsted.

across Europe using data derived from a network of 66 International Phenological Gardens. Since 1957, observations have been made of phenophases (the timing of plant growth phases e.g. flowering, first ripe fruits, leaf fall, etc.) of woody species within these gardens. The observations were used to develop a regression model that relates growing season start, end and length to altitude, longitude and latitude. Rötzer and Chmielewski (2001) applied the regression model using a European wide DTM to estimate the spatial variation in mean (1961-1998) growing season start, finish and length across Europe with a resolution of 30 arc seconds. For this work the average growing season length data were re-projected using GIS to derive average growing season length for the UK with a 500 × 500 m resolution. These data were then used to examine average growing season length within and around each designated site.

Growing season length data for the UK for 1961-2000 with a resolution of 5 × 5 km are available from the UK Climate Impacts Programme (UKCIP[10]) for 1961-2000. These data have been derived from regression and interpolation of weather data derived from approximately 500 weather stations. These data have been used to quantify the change in growing season length between 1961-2000 within and around each designated site.

3.4.4 Agricultural production

Each June, agricultural census information for individual farms is collected by questionnaire by UK government departments dealing with agriculture and rural affairs. These data are amalgamated for various geographies, the finest scale being at the parish level, although to prevent the disclosure of commercially sensitive information, some census attributes may not be available at finer geographies. A more significant obstacle to the use of data at the sub-parish scale is the difficulty in extraction and assembly. In a trial carried out by DEFRA staff, it took 5 days staff time to extract and summarise census data for two buffer zones around a single SSSI.

Agricultural census information for selected countries and years is also available as grid square estimates at 2 × 2 km, 5 × 5 km and 10 × 10 km resolutions from the Edinburgh University Data Library (EDINA[11]). These data have been transformed by EDINA using an algorithm to convert the data from recognised geographies to gridded estimates. Essentially, grid squares at a 1 × 1 km resolution are identified as belonging to one of seven land uses (agricultural land, upland, woodland, restricted agriculture – natural, restricted agriculture – artificial, urban or inland water) and the algorithm distributes agricultural census information over those land uses suitable for the census item in question. Agricultural census information available from EDINA is summarised by country and year in Table 13. Note that AgCENSUS does not hold data for Northern Ireland.

[10] http://www.met-office.gov.uk/research/hadleycentre/obsdata/ukcip/index.html
[11] http://www.edina.ac.uk/

Table 13. Gridded agricultural census data available from EDINA.

Country	Year						
	2002	2000	1994	1988	1981	1976	1969
England	-	✓ 73 census items	-	-	-	-	-
Scotland	-	✓ 165 census items	✓ 188 census items	✓ 186 census items	✓ 163 census items	✓ 131 census items	✓ 150 census items
Wales	✓ 35 census items	✓ 33 census items	-	-	-	-	-
England & Wales	-	-	✓ 48 census items	✓ 200 census items	✓ 198 census items	✓ 222 census items	✓ 208 census items
Great Britain	-	-	✓ 2 census items	✓ 2 census items	✓ 2 census items	✓ 2 census items	✓ 2 census items

Permission was granted by DEFRA for access to parish scale agricultural census information but because of difficulties experienced by DEFRA in extracting the data it was decided to use gridded 2 × 2 km resolution agricultural census information from EDINA. Therefore a licence was purchased by CEH for access to these data. These data were downloaded from EDINA and imported into the GIS to allow the temporal variation in the production of selected agricultural products around the designated sites to be quantified. However, the available census items vary for each country and year (Table 2) so it was not possible to summarise agricultural production for all years.

3.4.5 Site area and shape

The perimeter and area of polygons used to delimit the designated sites are standard items maintained by the GIS. Using a custom algorithm these were extracted to a data file and the perimeter : area ratio for each site calculated.

3.4.6 Environmental characteristics within and around each designated site

A custom algorithm was developed to summarise environmental characteristics both within designated sites and within selected buffers (0-1500 m, 1500-3000 m and 0-3000 m) surrounding the designated sites. As environmental characteristics were all maintained as gridded data sets the analysis identified individual pixels as falling within the designated site or surrounding buffer based upon the pixel centre. The algorithm could also be implemented based upon the area of a pixel falling within a site or buffer but this would likely bias the analysis towards those environmental characteristics with the greatest area.

3.5 Results

3.5.1 Site area and shape

Figure 44 shows site area and perimeter : area ratio plotted for selected designated sites. The smaller sites such as Rothamsted, Little Budworth Common and Climoor have larger perimeter:area ratios partly reflecting their small size. Clearly site area data are available for all SSSI while additional variables such as perimeter : area ratio can be easily generated by any standard GIS system.

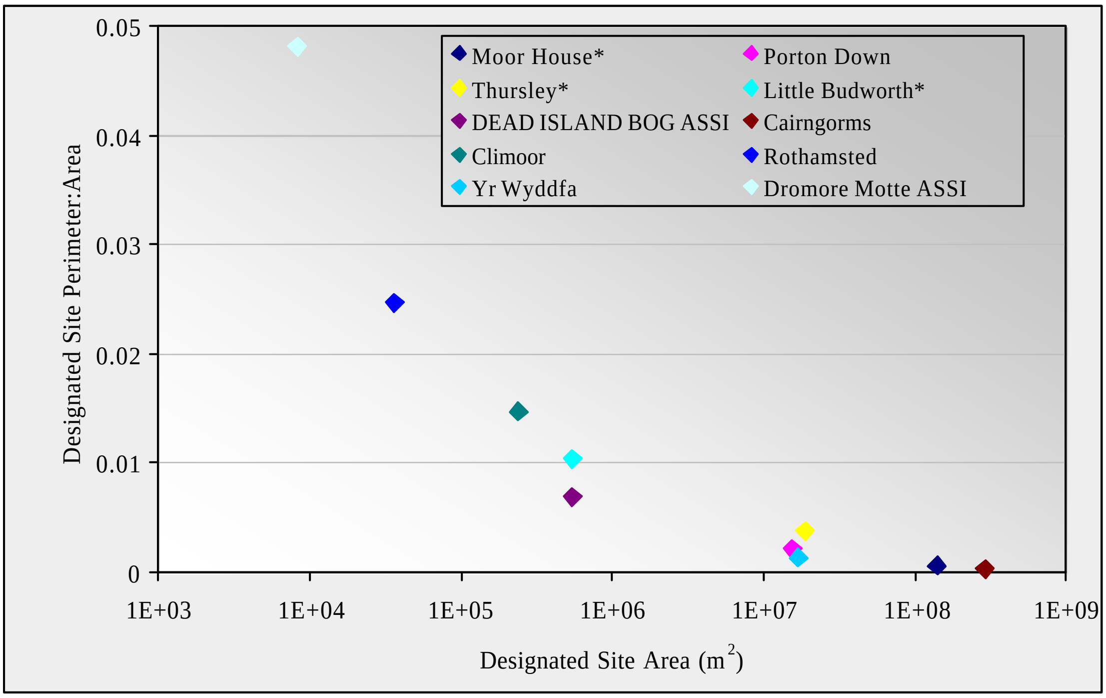


Figure 44. Area versus perimeter : area ratio for selected designated sites and the Rothamsted Park Grass plot. Yr Wyddfa is the Snowdon National Nature Reserve.

3.5.2 Donor and Vulnerable habitats

LCM2000 was used to quantify the area of Donor and Vulnerable habitats within and in 0-1500 and 1500-3000 m buffers surrounding each designated site. Examples are shown in Tables 14a and 14b.

Table 14a. Area of Donor and Vulnerable habitats within and in 0-1500 m and 1500-3000 m buffers around Moor House & Cross Fell SSSI.

Broad Habitat	Area (m^2)		
	Site	0-1500 m buffer	1500-3000m buffer
Total area	138162500	107631875	107514375
Donor habitats			
Arable & horticultural	13125	454375	4450000
Improved grassland	2070625	16022500	27184375
Built up areas, gardens	0	67500	265000
Vulnerable habitats			
Broad-leaved woodland	803750	2316875	2175625
Calcareous grassland	2093125	7486250	5105625
Acid grassland	49960000	31596875	18771250
Dwarf shrub heath	30239375	10936250	8100000
Fen, marsh and swamp	0	0	0
Bog	35291875	15255000	21628125
Montane habitats	0	0	0

Table 14b. Area of Donor and Vulnerable habitats within and in 0-1500 m and 1500-3000 m buffers around Porton Down SSSI.

Broad Habitat	Area (m^2)		
	Site	0-1500 m buffer	1500-3000m buffer
Total area	15614375	39628750	50339375
Donor habitats			
Arable & horticultural	1023125	20296250	23367500
Improved grassland	2381250	11142500	17273125
Built up areas, gardens	140000	1208750	2158750
Vulnerable habitats			
Broad-leaved woodland	1420625	1947500	4110000
Calcareous grassland	9434375	2037500	2138125
Acid grassland	0	0	0
Dwarf shrub heath	0	33750	0
Fen, marsh and swamp	0	0	0
Bog	0	0	0
Montane habitats	0	0	0

A similar approach was used to identify the proportion of Donor and Vulnerable habitats in a 0-3000 m buffer around **all** 4101 SSSI's in England and **all** 1021 SSSI's and 67 NNR's in Wales. Figures 45a and 45b show the distribution of Donor and Vulnerable habitats around SSSI's in England whilst Figures 46a and 46b show the same for SSSI's in Wales. On these figures, the position within these distributions of individual SSSI's – Moor House & Cross Fell, Little Budworth Common and Porton Down in England and Aber Afon Conwy in Wales – is also shown, enabling the proportion of Donor and Vulnerable habitats around individual designated sites to be compared to national distributions. This analysis suggests that for England most SSSI's are surrounded by higher proportions of Donor habitats and lower proportions of Vulnerable habitats ie. the site periphery is composed of habitat likely to act as a source of nutrients and propagules of widespread species. For Wales, the histograms for each habitat category were less skewed suggesting a network of sites more buffered by semi-natural habitat and with less dominance of the wider hinterland by intensive or built broad habitat.

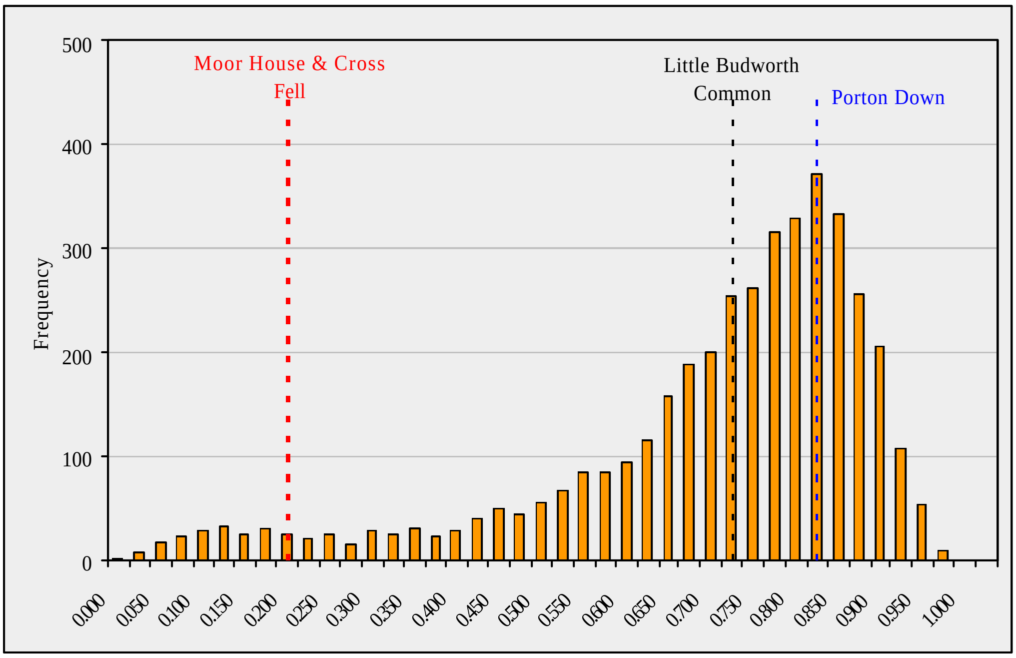


Figure 45a. Distribution of proportion of Donor habitats within 0-3000 m buffer surrounding SSSI's in England.

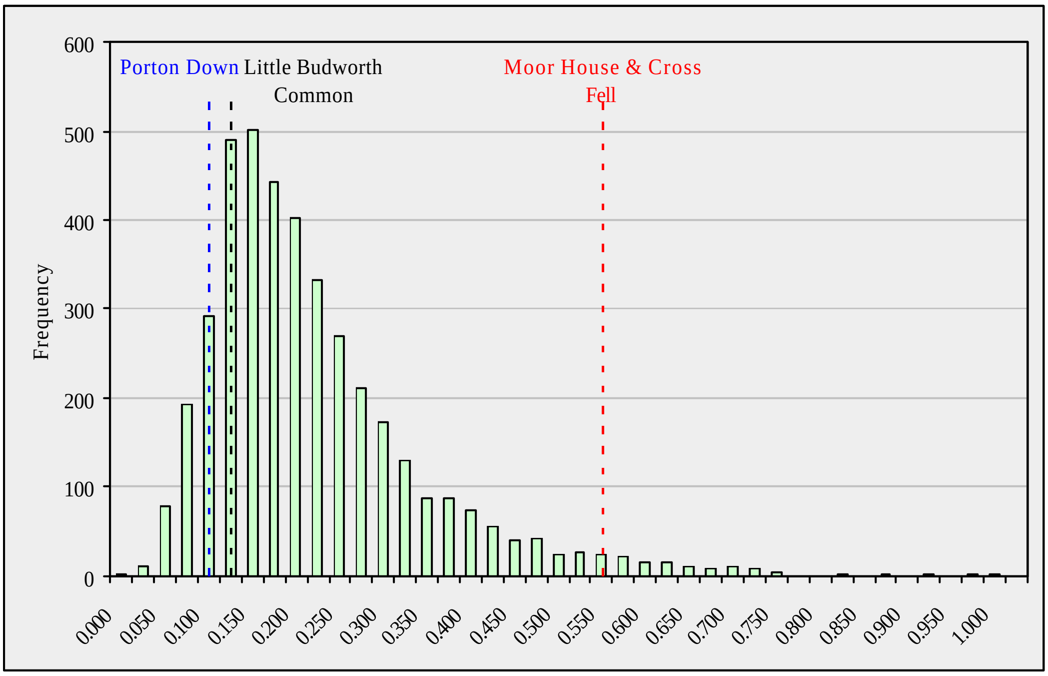


Figure 45b. Distribution of proportion of Vulnerable habitats within 0-3000 m buffer surrounding SSSI's in England. The X axis gives intervals of proportional cover ie. 0.5 = 50% cover of Vulnerable habitats in the the surrounding buffer zone.

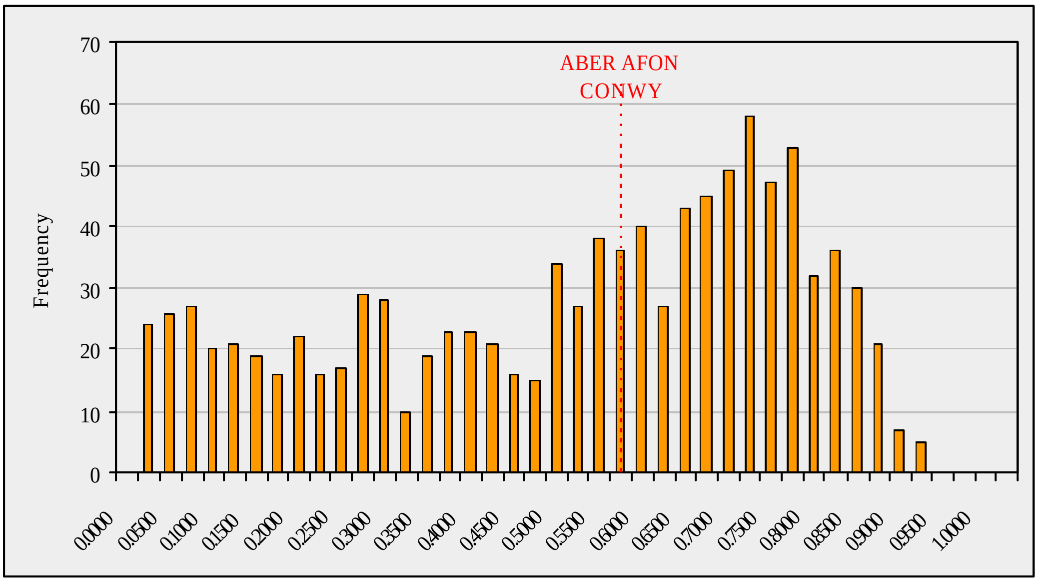


Figure 46a. Distribution of proportion of Donor habitats within 0-3000 m buffer surrounding SSSI's in Wales.

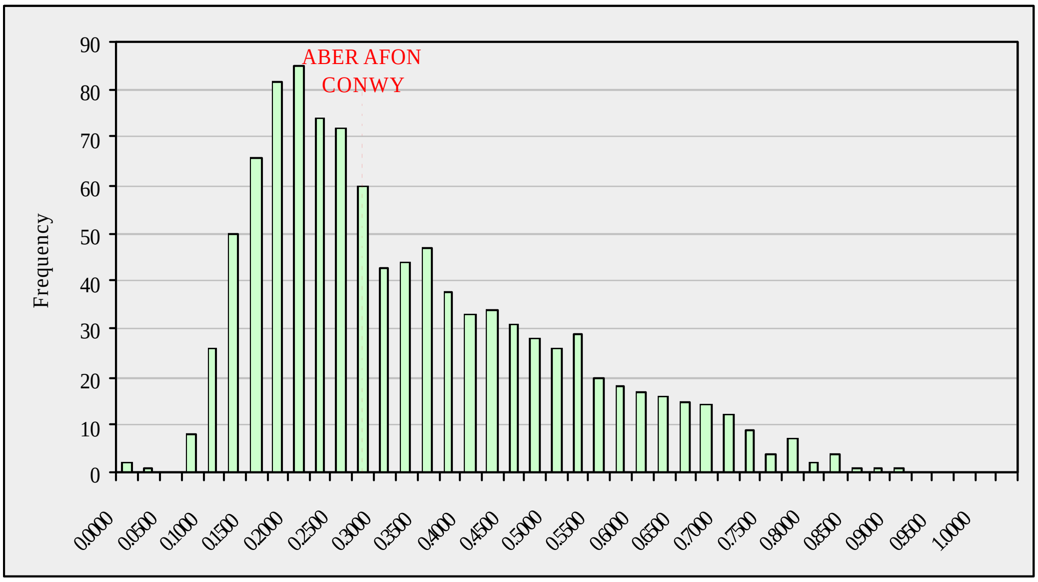


Figure 46b. Distribution of proportion of Vulnerable habitats within 0-3000 m buffer surrounding SSSI's in Wales.

3.5.3 Change in growing season length

The re-projected European average (1961-1998) growing season length data was used to summarise growing season length for individual nations, and in buffers 0-1500 m and 1500-3000 m surrounding each selected designated site (Figures 47a and 47b). For the UK, average growing season length varies between 140-220 days with the highest frequency occurring around 205 days. For individual designated sites, latitude is obviously correlated with growing season length with more northerly sites have lower growing season length than southerly sites. Furthermore, and as would be expected, larger sites, which extend over greater altitudinal ranges (e.g. Cairngorms SSSI; Figure 47a), have a greater range in growing season length then smaller sites (e.g. Porton Down; Figure 47b).

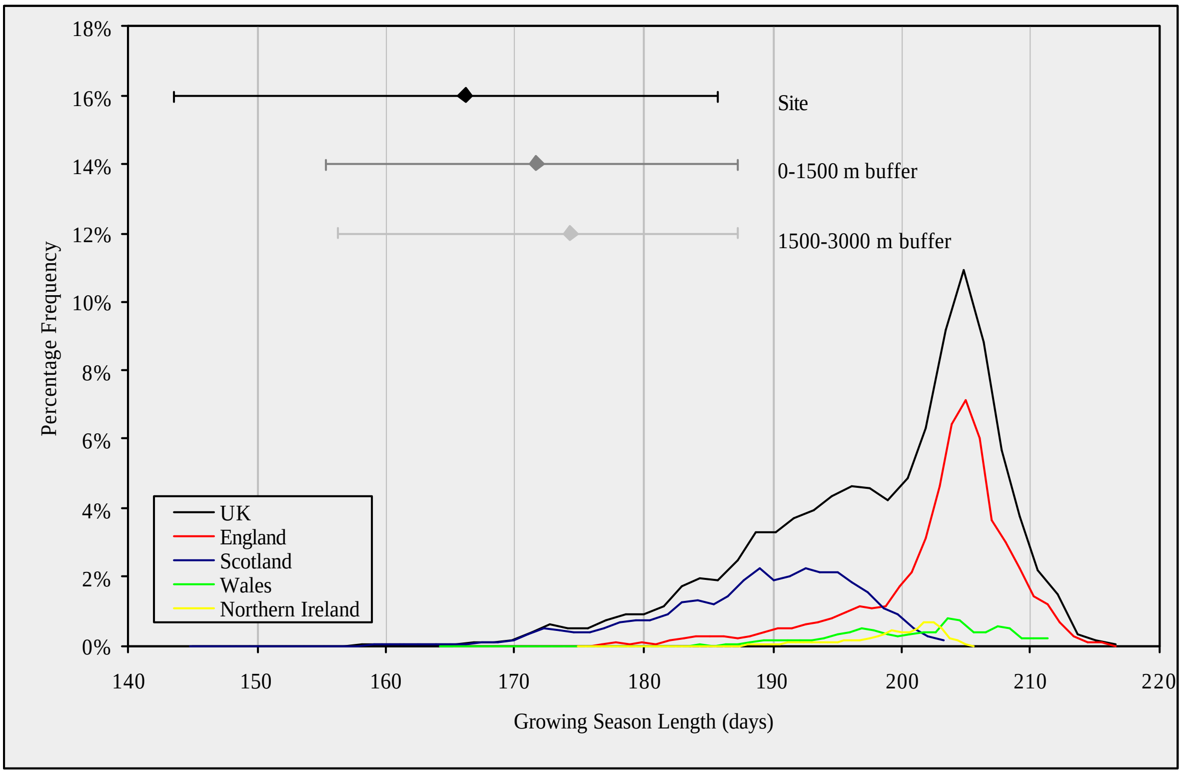


Figure 47a. Distribution of average (1961-1998) growing season length for England, Scotland, Wales and the UK. Also shown is the mean (◆) and range for Cairngorms SSSI and surrounding 0-1500 m and 1500-3000 m buffers.

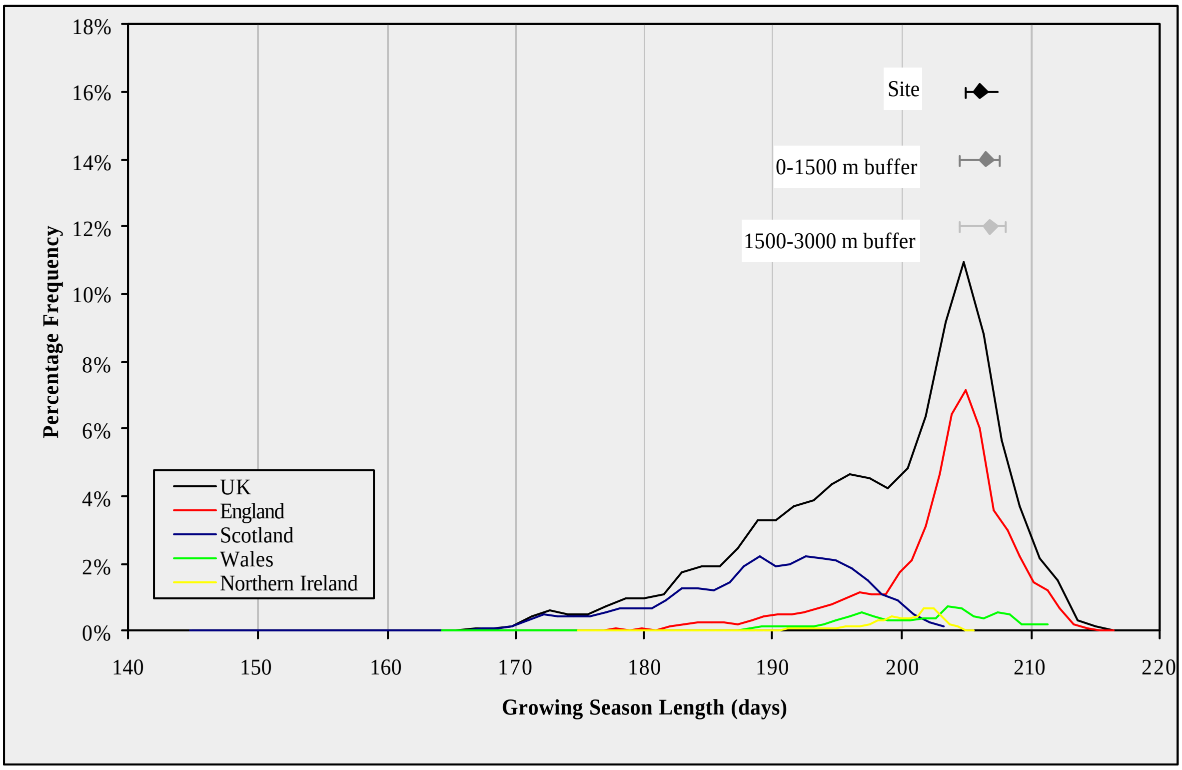


Figure 47b. Distribution of average (1961-1998) growing season length for England, Scotland, Wales and the UK. Also shown is the mean (◆) and range for Porton Down SSSI and surrounding 0-1500 m and 1500-3000 m buffers.

National growing season length data were used to identify the temporal change in growing season length between 1961-2000. For each designated site the annual growing season length was identified using a custom algorithm and a fitted linear regression used to identify the change with time (the slope of the fitted linear regression). Figure 48 shows change in growing season 1961-2000 plotted against growing season length in 2000 for the selected designated sites. Figure 48 suggests that the greatest changes in growing season length are affecting those sites with the shortest growing season length (the more northerly and more elevated sites – Cairngorms SSSI, Moor House & Cross Fell, Climoor and Yr Wyddfa (Snowdon) NNR).

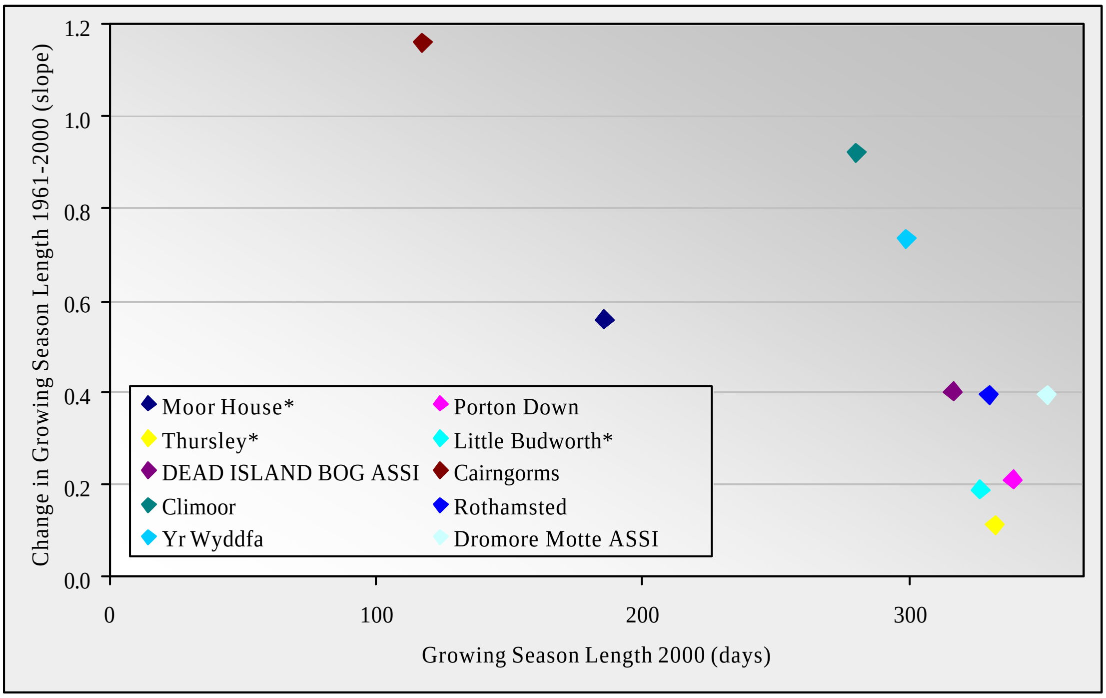


Figure 48. Growing season length in 2000 versus change in growing season length between 1961-2000 expressed as the slope of a fitted linear regression for selected designated sites.

3.5.4 Change in agricultural production

Gridded agricultural census data was used to compare the temporal change in agricultural production in and around selected designated sites to the change in agricultural production nationally. As mentioned previously, the census items maintained by EDINA are inconsistent between the different countries and years. Therefore, for some years it is not possible to derive sufficient agricultural production. For example, it was not possible to generate total sheep production for England and Wales for 1994 as only three census items (ewes kept for breeding, two-tooth ewes, and lambs < 1 year old) were available from the EDINA gridded data whereas the agricultural census for this year required information on a further three items (rams for service, draft and cast ewes, and wethers and other sheep) for a total sheep and lamb production to be calculated.

Figures 49a and b show examples of the comparison of agricultural production for individual sites – total sheep production for Moor House & Cross Fell SSSI in Figure 49a and total arable production (defined as the sum of barley, wheat, maize, oil seed rape (not for stock feeding) and beet (not for stock feeding)) for Porton Down SSSI in Figure 49b – to national production. These figures show the national production and the production in and around each designated site (assuming an approximate 3000 m buffer) for each year. The graph (at the lower right in Figures 49a and b) summarise the mean and range in national and site production for each year. Tables 15a and b compare the change in total sheep production (Table 15a) and total arable production (Table 15b) for the selected designated sites with national trends. National-scale change in agricultural production has been determined by fitting a linear regression to the annual production data. Total sheep production has increased both at the national and individual designated site levels whereas for total arable production national production has increased whereas in and around the selected designated sites there has tended to be a decline in total arable production. Note that the slope coefficients for each site regression can be compared to the national slope to rapidly judge whether local change in agricultural production (within site or in site buffer zones) implies intensification or extensification that deviates substantially from the British or national average. However, if trends over time are curvi-linear then linear regression, and hence a single slope coefficient, will not be appropriate and a better summary will be in graphic form. The capacity to generate summary graphic output for major crop groups for each SSSI is readily achievable in a GIS system manipulating the AgCENSUS data products.

While such one page summaries of major crop types are informative, further data reduction would be required for a strategic overview of agricultural production history around each British SSSI (note Northern Ireland data are not available in AgCENSUS). The goal here would be to summarise productivity using a standard higher level unit that subsumed the range of products. Hence, a common index would apply to all livestock and plant crop types. Possibilities include energy input per unit area per year or, even better, estimated nutrient surplus per unit area per year. Options for constructing an initial index focussed on quantifying the N input required to support a certain level of production over each farmed product. Published information on recommended fertiliser application levels showed promise as a basis for an index (ADAS 2000). Recommended levels depend on rainfall, soil type and previous crop. The first two gradients can be addressed using available GIS datasets, however previous crop type is problematic since this would require farm census information at the parcel level. Our conclusion at the end of this initial phase of exploration is that a standard index of agricultural productivity change would be a useful addition to the database of risk factors as well as yielding a valuable explanatory variable for other signal attribution studies such as Countryside Survey and the BSBI Local Change project (D.Roy pers.comm.). Development of a common index appears feasible but requires further input from livestock and crop production experts.

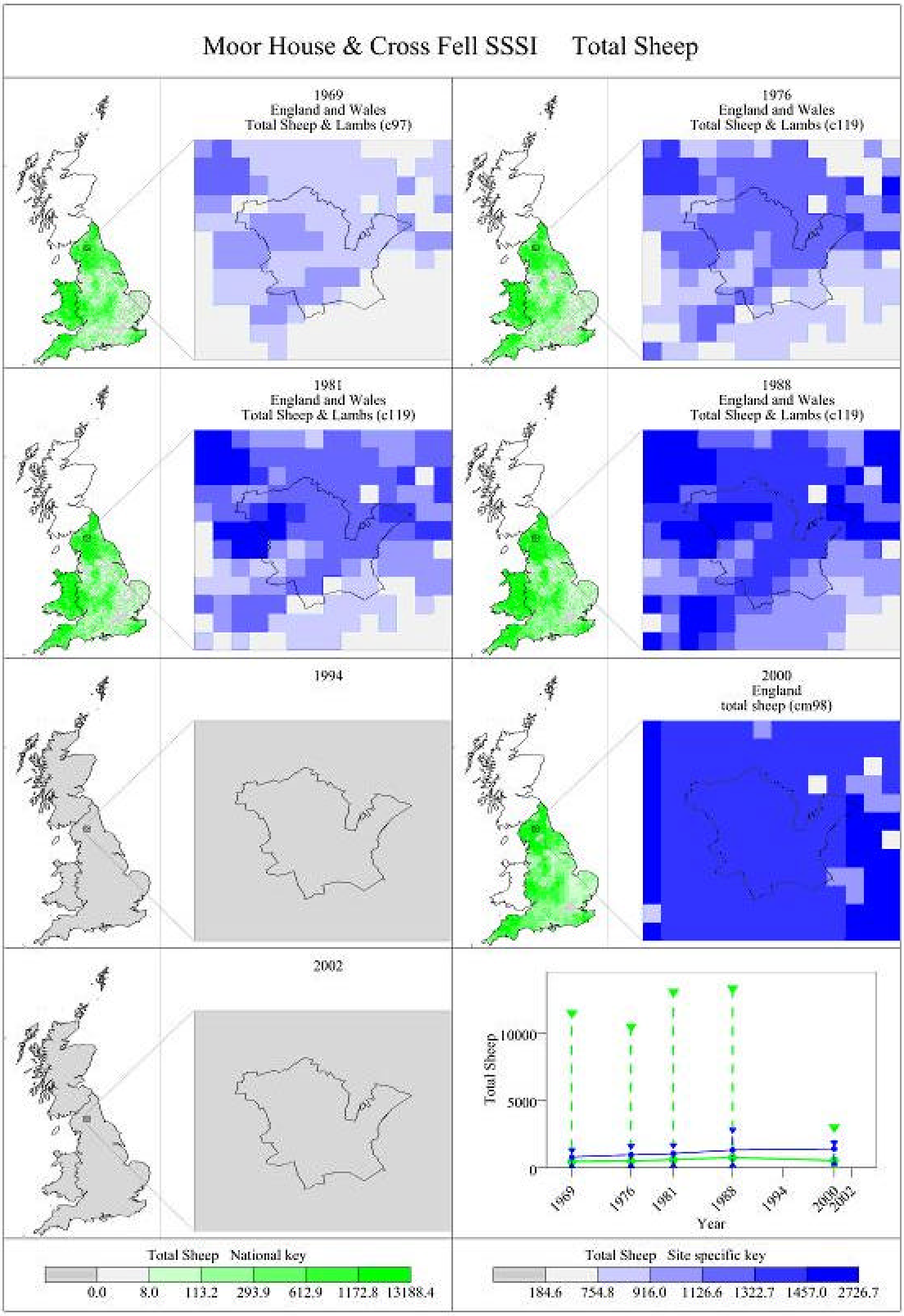


Figure 49a. Temporal variation in total sheep for Moor House & Cross Fell SSSI.

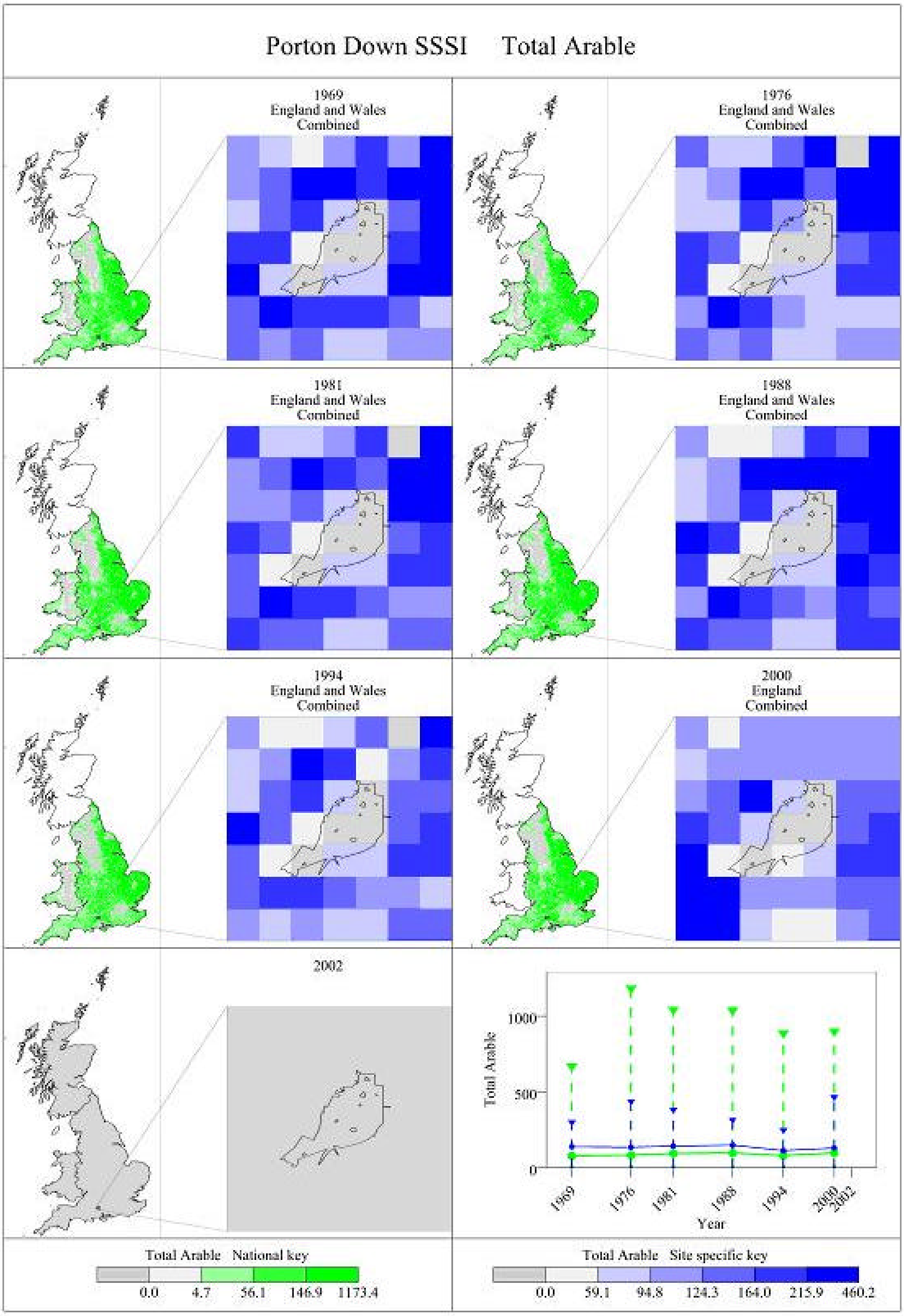


Figure 49b. Temporal variation in total arable for Porton Down SSSI.

Table 4a. Fitted linear regression of change in total sheep with time for selected designated sites.

Nation	Site	National basis			Site basis		
		Slope	Constant	R^2	Slope	Constant	R^2
England	Moor House SSSI	4.04	-7429	0.17	20.60	-39729	0.94
"	Porton Down SSSI	"	"	"	4.20	-8143	0.80
"	Thursley Common SSSI	"	"	"	4.95	-9731	0.82
"	Little Budworth SSSI	"	"	"	5.20	-10196	0.61
"	Moor House NNR	"	"	"	21.57	-41649	0.98
"	Thursley Common NNR	"	"	"	3.97	-7775	0.85
"	Rothamsted	"	"	"	3.10	-6086	0.52
Scotland	Cairngorms SSSI	3.29	-6140	0.82	-0.05	217	0.00
Wales	Yr Wyddfa NNR	50.28	-98808	0.86	34.47	-66760	0.65
"	Climoor	"	"	"	5.21	-9140	0.02

Table 4b. Fitted linear regression of change in total arable with time for selected designated sites.

Nation	Site	National basis			Site basis		
		Slope	Constant	R^2	Slope	Constant	R^2
England	Moor House SSSI	0.41	-733	0.31	-0.04	98	0.09
"	Porton Down SSSI	"	"	0.31	-0.47	1072	0.19
"	Thursley Common SSSI	"	"	0.31	-0.58	1181	0.91
"	Little Budworth SSSI	"	"	0.31	0.54	-1025	0.80
"	Moor House NNR	"	"	0.31	-0.05	111	0.28
"	Thursley Common NNR	"	"	0.31	-0.60	1211	0.78
"	Rothamsted	"	"	0.31	-0.09	327	0.00
Scotland	Cairngorms SSSI	-	-	-	-	-	-
Wales	Yr Wyddfa NNR	0.28	-468	0.11	0.00	3	0.08
"	Climoor	"	"	0.11	-0.05	106	0.59

3.5.5 Flood risk

Flood risk information for ASSI/SSSI is available from a number of sources but varies in coverage between UK territories. No Northern Ireland estimates are currently available although a national flood risk assessment is due to start soon. The best currently available datasets are held by Norwich Union Insurance and the Environment Agency. Norwich Union commissioned a detailed flood risk mapping project which estimates the probability of inundation over a range of return periods. This refers to the likelihood of inundation over different time intervals (see Bradbrook in press for further details). The advantage of the Norwich Union product is that it specifies return periods of between 5 and 100 years and covers Britain. The Environment Agency National Flood Risk Assessment (NaFRA) only covers England and Wales but also includes a range of return periods although more restricted in number than Norwich Union. Both assessments are scaled to a 5x5m Digital Terrain Model. Risk of inundation for a small number of sites can also be readily determined from the EA flood risk assessment web-site (www.environment-agency.gov.uk/subjects/flood). This specifies categories of risk based on post code searching. For example Rothamsted Park Grass has a low risk of sea or river flooding at 1 in 1000 per year. Lack of readily available databases prevented further interrogation for test sites and the wider population of ASSI/SSSI.

3.5.6 Critical Load exceedance

Modelled estimates of total N deposition are readily available for the UK at 5x5km square resolution (see NEGTAP 2001; Fowler et al 2004). Uncertainties relate to the deposition velocity of dry deposited N (Smith et al 2001; Leith et al 2004) and sub-grid variability in topography and in the distribution of agricultural point sources (Sutton et al 2003). Modelled or measured N deposition can then be compared with the Critical Load for the vegetation type of concern to produce an estimated exceedance of either the upper or lower limit or a mid-point. Empirical Critical Loads for N are available for all the Priority Habitats on project test sites (Table 16) but are not available for all Priority Habitats identified in the UK Biodiversity Action Plan (C.Whitfield pers.comm.).

Table 16. Total N deposition measured or modelled (FRAME/GANE) for project test sites between 2000 and 2004. Simon Caporn and Mike Pilkington (Manchester Metropolitan University) provided measured N deposition estimates for Ruabon and Budworth.

	FRAME/ NEGTAP			Critical Load exceedance (empirical N)			
Test site	**Kg N ha^{-1} yr^{-1}**	**Measured**	**Priority Habitat**	**Kg N ha^{-1} yr^{-1} > lower limit**	**Kg N ha^{-1} yr^{-1}> upper limit**	**CL lower**	**CL upper**
Plynlimon	13.3		n/a (upland acid grassland)	3.3	0	10	20
Climoor	17.5		Upland Heath	7.5	0	10	20
Ruabon	18.2	10.6	Upland Heath	0.6	0	10	20
Pwllpeiran	11.1		n/a (upland acid grassland)	1.1	0	10	20
Moorhouse	14.8		Blanket Bog	9.8	4.8	5	10
Cairngorm	4.3		Upland Heath	0	0	10	20
Rothamsted	23.8		Lowland Meadow	3.8	0	20	30
Porton	22.7		Lowland Calcareous Grassland	7.7	0	15	25
Budworth	20.7	25	Lowland Heath	15	5	10	20
Dromore Motte	30.1		Lowland Meadow	10.1	0.1	20	30
Dead Island Bog	20.1		Raised bog (lagg fen)	5.1	0	15	35

3.6 Opportunities for UK SSSI coverage

Selected risk factors vary in the extent to which data at meaningful scales can be readily obtained for UK ASSI/SSSI (Table 17). Some factors, such as exposure to phosphate surplus, are hard to estimate using simple extant data sources, hence case-study approaches coupled with direct consultation with local staff have been used to provide the most accurate yet geographically comprehensive risk assessment for UK designated wildlife sites (ECUS 2003).

Table 17. Status of selected risk factors and supporting datasets for UK ASSI/SSSI.

Risk factor	Resolution	UK coverage	Ease of computation	Further development required
P limitation	n/a	Not estimable at useful scales	n/a	Estimate based on local pH measurements and vegetation type
Growing Season Length	500 m x 500 m	UK	Simple	No
GSL change	5x5 km		Simple	No
Site geometry	Site scale	UK	Simple	No
Buffer zone habitats	25x25 m	UK	Simple	Possible integration with Immigration Potential approach developed for part 1
Flood risk	5 x 5 m on Digital Terrain Model	GB	Probably straightforward	Integration with EA water quality data for the catchment
N deposition	5x5 km	UK	Already available	Strategy of model validation against site measurements desirable?
Agricultural production history	2 x 2 km	GB	Medium complexity envisaged	Further research required with involvement of agronomists

P limitation is also not estimable across UK sites at a useful resolution (within Priority Habitat patch) using current knowledge and data sources and hence must rely on local assessments. Other datasets are either readily available for assembly at the UK scale or are potentially applicable given that the data exists but pending cost and accessibility. This applies to the Norwich Union flood risk assessment.

Two risk factors could be further developed to improve their ability to convey risk of eutrophication impacts. The use of Broad Habitat cover to estimate source areas for nutrient-demanding immigrant plants is simple and merely rests on the assumption that arable, improved grassland and built land support significant populations of plant species suited to higher nutrient levels than semi-natural Priority Habitats. A more advanced estimate of the abundance of such immigrants in adjacent buffer zones was developed in part 1 of this project where the 10km square species list was combined with Broad Habitat cover, species preference indices and dispersal indices to rank the immigration potential of named indicator species. Extension of this approach to all UK ASSI/SSSI is limited by the absence of dispersal indices for a residue of CSM indicators and by a lack of habitat preference indices for bryophytes. The latter are however near completion by Chris Preston at CEH Monkswood.

The risk factor that remains undeveloped is a common index of change in agricultural productivity that could summarise nutrient inputs or surpluses across the range of farmed products. This appears feasible but would require further research.

3.7 Options for synthesis of risk factor information across UK sites

The MIRABEL project experimented with the construction of a single index that combined values of a range of risk factors (Petit et al 2003). Although simple, this approach concealed the relative influence of the different factors in generating particular index values on particular sites. Users applauded the reduction of multiple sources of risk to one score but then inevitably sought to unpack the index so as to determine the highest source of risk for Natura 2000 sites in a particular biogeographic zone in Europe. This experience suggests that a better option would be a data reduction step that still retained a clear view of how different risk factors contributed to heightened risk across groups of sites. A simple technique would be to generate a multivariate classification of ASSI/SSSI by risk factor levels. This would identify groups of sites by a profile of shared values of risk factors. This would effectively divide site series into risk-based strata across UK territories.

Defining site clusters by combinations of risk factor levels ought to help identify sites where heightened vigilance is required for eutrophication impacts. This knowledge should therefore help direct the stratification of sites for monitoring and help target limited resources for monitoring on sites at most risk of change resulting from the operation of multiple but identifiable drivers. In the absence of deeper, quantitative understanding about the way risk factor levels exacerbate or constrain the impact of N deposition or even operate to cause eutrophication impacts on their own, two approaches are possible. Sites can be ranked in terms of likely risk simply by reference to the distribution of factor values across the population of ASSI/SSSI, as shown above for Growing Season Length, cover of donor and vulnerable habitats, and agricultural outputs (and see Box 9). In addition, further interpretation and scoring could be based on expert judgement taking into account the weight of published evidence (Suter 1993). This is consistent with the evolution of the empirical critical loads for N (Achermann & Bobbink 2003).

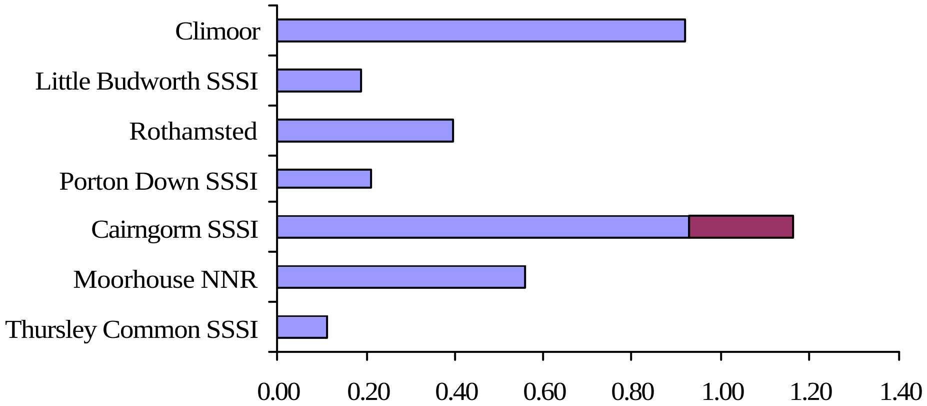
As a start in assembling the evidence base for such an assessment, a review of the evidence from interaction studies where N deposition had been manipulated alongside other impacts was carried out as part of this project. Results indicated that accurate yet general models applicable across habitat types and climate gradients are limited but accumulating evidence provides important clues as to the probable importance of such interactions (Appendix 8).

BOX 9 – Graphing risk factor values for test sites

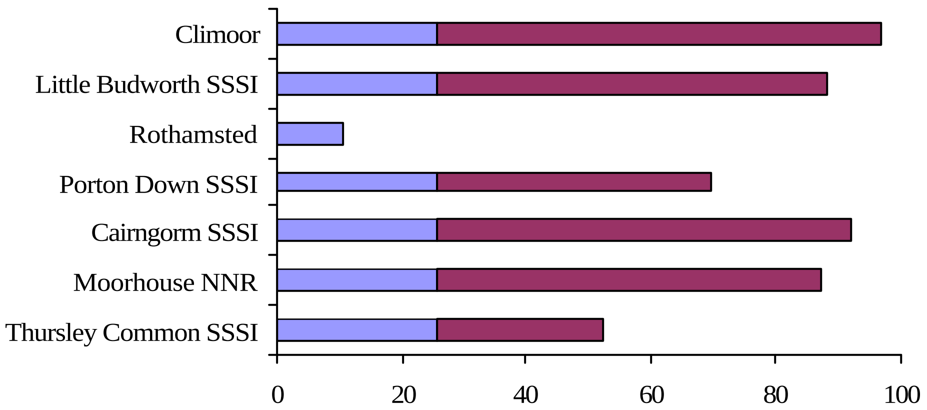
Risk factor values can be graphed separately for each site rather than construct a combined index, which conceals the separate contribution of each factor (Figs 1a-d). In the MIRABEL project (Petit et al 2003) an arbitrary threshold for defining sites at heightened risk was set at the 75%tile of index risk values. In the graphs shown below, the red part of each bar indicates the amount by which each factor exceeded the 75%tile value based on the total population of UK ASSI/ SSSI. In Fig 1d, the red part of each bar shows the amount of total N deposition by which the lower empirical Critical Load for the Priority Habitat associated with each site, was exceeded in 2000 (see page 17 for a listing of Priority Habitats by site). Such graphs provide a rapid overview of the magnitude of the range of risk factors but when assessing national site series as a whole, a Priority Habitat within site classification by risk factor values is likely to be more useful.

The results show that Cairngorm SSSI was the only test site to have exceeded the 75%tile value for GSL change yet zero cover of intensive broad habitat in the site periphery plus very low background N deposition clearly lessen the level of risk due to eutrophying N and interactions with other factors. Rothamsted Park Grass represents the other extreme since the lower CL has been recently exceeded while the site periphery is well dominated by intensive habitats that could act as sources of additional N and P as well as propagules of more nutrient-demanding colonists.

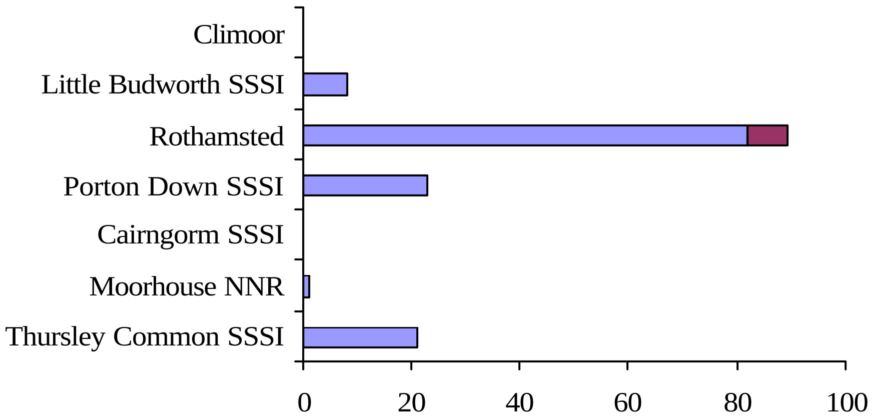
a) Change in Growing Season Length (1961-'99)



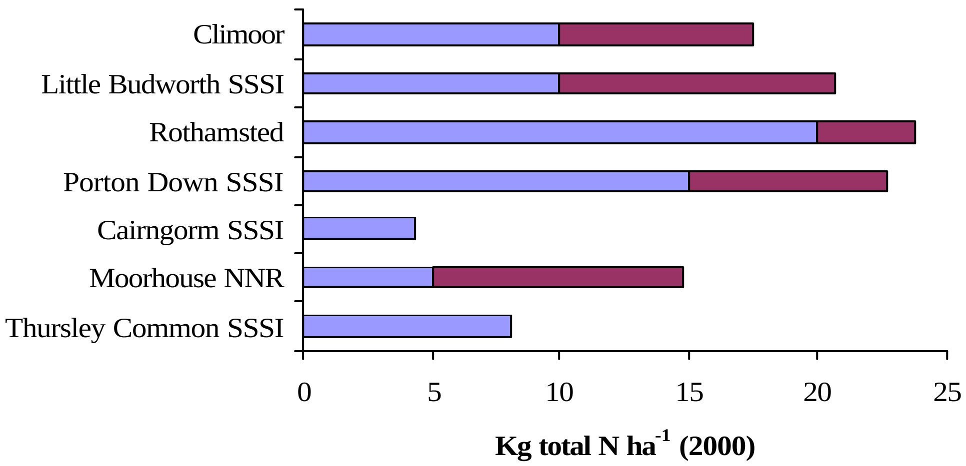
b) % semi-natural habitats in 1500m site buffer



c) % intensive habitats in 1500m buffer



d) Total N deposition and exceedance of lower empirical CL



4 Discussion and synthesis

4.1 Results from model testing – summary of the evidence

The MAGIC soil model

Although MAGIC has a long history of application and the acidification component has been well validated on many occasions, it initially found it difficult to replicate the Rothamsted historical time series. This ultimately reflected inaccurate information on dry N deposition. When updated input information was applied, a more satisfactory prediction was generated. However, although three historical values of soil C/N were replicated, the slight decline in soil C/N was inconsistent with observed vegetation change and reduced hay yield. This probably highlights the poor quality of the observed data in that a good test of MAGIC and GBMOVE ought to be based on a greater number of observed soil measurements.

While more validation and testing of the N dynamic component is essential the process basis of the model is considered robust. Hence, the main reasons for poor model performance are either inaccurate data on chemical inputs and outputs or poor soil chemistry data for calibrating the model. Uncertainty analysis can be used to back-calculate what the correct inputs/outputs should be to match historical soil data Unfortunately such data are scarce. With good present-day measurements of soil chemistry, future projections of soil change are considered reliable because the process basis of the model has been largely well studied and validated.

Given the need to model impacts of management as well as pollutant deposition, linkage to a sucession model such as SUMO would be highly desirable. This would allow rates of vegetation change and nutrient cycling under different regimes of biomass growth to be accurately modelled. For example, the simulation of woodland growth at Rothamsted (scenario R2 in Appendix 9) applied MAGIC/GBMOVE in one simulation and SMART/SUMO in a separate simulation. In the MAGIC/GBMOVE simulation, an appropriate woodland community type (currently present on the Broadbalk Wilderness) was very accurately predicted after simulating abandonment of the Park Grass control plots. However, succession was modelled as a simple empirical increase in canopy height to a final value typical of lowland broadleaved woodland in Britain. If SUMO could be linked to GBMOVE this would provide a process-based trajectory of change in canopy height based on modelling biomass growth. SUMO modifications also allow temperature change to affect biomass growth and decomposition. This is clearly important for simulations of the impact of future climate change.

SUMO

Validation and testing on UK sites resulted in a mixture of reasonable though somewhat inconclusive results. SUMO's successes outweighed its failures, while the successional processes that are the basis of the model appear sound from work carried out by the author of the model on Dutch test datasets. Indeed the accurate simulation of biomass change at Rothamsted was the most impressive of all the SUMO testing results.

Some aspects of poor performance could be understood. For example under-prediction of biomass growth of some planted conifer species on Forest Level II sites may reflect differences in growth rate of conifer cultivars used in high precipitation areas of upland Britain as opposed to the Netherlands. Other aspects need more development work and testing. The empirical method for dividing up herbaceous biomass into grass, legume and forb production is poorly tested. The method was not sensitive to the subtle changes in production seen at Rothamsted. A process-based model of grass, forb and legume production would be a complex research project in its own right. Since the herb functional type is already modelled in SUMO, an extension of the empirical approach may still offer promise if explanatory variables are added in addition to substrate fertility. In order to maintain the general applicability of SUMO to UK ecosystems, further work on the herb split would also need to contend with differences in the grazing-mediated competitive relations between grasses and dwarf shrub heaths in the Netherlands as opposed to upland Britain. Possible differences in the palatability of grass species between habitats, for example *Nardus stricta* versus *Deschampsia flexuosa*, could provide a basis for understanding such relationships. This is certainly an area for more model development but such research needs to be placed in the context of a wider strategy for integrating SMART/SUMO with GBMOVE or SUMO with MAGIC/GBMOVE.

If SUMO is to be more widely applied to UK ecosystems, further testing is required to progressively win more credibility for the approach. The critical requirement is therefore measured biomass data for each functional type across a chrono-sequence or time series of vegetation change plus the soil chemistry and N deposition data required to run the soil model component. Further work should therefore involve a concerted, thorough search for existing datasets prior to any decision to collect new data. The benefit of collecting new data is that the process can be completely controlled. This is an important issue given the sensitivity of soil and vegetation models to variation in input data. The disadvantage is the inevitable cost.

At present SUMO could be applied to test a range of simple management scenarios on British heathland and grasslands. Examples have been produced in this report.

<u>Calibration equations</u>

Calibration equations are a necessary consequence of the scarcity of paired species records and soil measurements. The uncertainty around each relationship is effectively ignored when MAGIC outputs are translated into mean Ellenberg values so that GBMOVE regressions can be solved. However, this variation is expressed when predictions are compared with observations. Two conclusions can be drawn from the results. First, the calibration between soil C/N and mean Ellenberg fertility has greatest uncertainty and highest sensitivity at the most fertile end of the vegetation gradient. In these situations, soil C/N appears to be a poor predictor of the correlation between species composition and vegetation productivity but a better predictor in less fertile, soils with higher C/N ratios. In an attempt to account for P availability, measurements of Olsen's P were used in addition to soil C/N in the initial construction of calibration equations using Countryside Survey data, however this did not significantly reduce unexplained variation (Smart et al 2003). Further testing of the calibration equations is feasible since all that is required is quadrat data with measured soil C/N, soil slurry pH and % soil moisture rather than the range of soil chemical measurements required for MAGIC calibration. Such data could also be used to update the calibration equations but since Countryside Survey data is a representative random sample of British soils and vegetation, it is unlikely that simply adding more data will solve the basic problem that vegetation productivity is better explained by factors in addition to soil C/N. The explanatory power of other soil variables is likely to be tested during the pilot soil-sampling work for the next

Countryside Survey (H.Black pers.comm.). These should include biomass and available N. Further work will also be carried out under the DEFRA Air Pollution Umbrella (www.ukreate.ceh.ac.uk), to refine existing calibration equations. The most promising option is to subdivide the CS training data into a larger number of vegetation types and then to derive separate calibration equations that take account of obvious differences in the form of mean Ellenberg versus soil relationships (see Fig 10 for example and Wamelink et al (2002)).

GBMOVE

Initial comparison of GBMOVE models against published Ellenberg values showed promising positive correlations for both higher and lower plants. Although published Ellenberg numbers are in no sense error-free, high scatter around relationships particularly for Ellenberg fertility and pH did suggest that species by species validation should be undertaken to either accept the GBMOVE optimum and hence the Ellenberg number becomes questionable or to modify or reject the GBMOVE model. Because there are many species models, validation will be a long-term campaign. However, identifying a core subset of the most reliable models could be achieved by focussing on CSM indicators as a species group that has the highest priority for the user community, and then prioritising species showing the largest differences between predicted optima and Ellenberg numbers. For example, CSM indicators for upland heath, lowland heath, raised bog and blanket bog Priority Habitats could be targeted since MAGIC+GBMOVE predictions appear to be more reliable in these systems than lower soil C/N grassland systems.

Model tests using MAGIC+GBMOVE linked by calibration equations showed that on two of the three test sites with long-time series of vegetation monitoring data, significant positive correlations were seen between observed and predicted directions of change. However these relationships were always characterised by large amounts of unexplained variation albeit that this was expected for several obvious reasons. The test outcomes, although not conclusive, suggest that the modelling approach has promise but requires further technical refinement and then further testing if real model applications are to be considered robust. The implication of the results is that pollutant deposition signals were detected at Porton and Moorhouse. At Rothamsted, an acceptable MAGIC run was finally generated, in the process illustrating the importance of reliable soil chemistry measurements and accurate information on N inputs and outputs. However, MAGIC output generated GBMOVE predictions that were the reverse of those observed even though MAGIC was consistent with observed soil C/N changes. It seems that soil C/N declined even though the grassland system became less productive.

Attempts to predict total community species composition using only abiotic input data largely failed and suggest that model applications should focus on predicting change in habitat suitability for key subsets of species comprising those present at a monitored site and CSM indicators at large in the local species pool. This approach is consistent with the assessment of impacts on CSM indicator species but not with prediction of policy-relevant indicator variables such as species richness and mean Ellenberg scores.

Rare and subordinate species modelling

A novel and promising method was developed during this project (see Appendix 4 and Box 8). Further work is needed to generate association models for the remaining

447 rare species not covered by GBMOVE. This work would be rapid yet necessary to determine how many SAP taxa were likely to be covered by the approach. Four out of the thirteen species initially modelled, did not converge. Taking this as a guide it may be that as many as a third of the remaining species would be left without models.

4.2 Application of dispersal and species pool filters to CSM indicator species not present in the starting vegetation

An empirical approach was adopted to the estimation of dispersal potential and abundance of indicators in local species pools. This approach was adopted because of variation in dispersal dynamics between species (Freckleton & Watkinson 2002) and spatial variation in the abundance of species in habitats in and around protected sites. Both aspects pose severe and probably insurmountable obstacles to the development of spatially accurate yet generalisable dispersal models based on parameterisation of each species' dynamics. Hence we adopted a simple, informatics based approach that links together databases of habitat preference and dispersal adaptive traits with spatially explicit information on land-cover extent and species pool composition around each test site. Because of the empirical nature of the approach we apply the 'dispersal filter' in a transparent and very simple manner. The approach depends upon the following chain of reasoning. First that the local population of a species will be larger where its most favoured broad habitat occupies a larger area in the site or site buffer zone. Second, a species can be assigned a dispersal index that reflects possession of attributes that promote dispersal by multiple vectors. Third, a poor disperser will find it just as hard to reach a target site even if its preferred broad habitat is extensive, as a good disperser whose preferred broad habitat is rare. The weighting of each species in the site or site buffer pool is based on these crude assumptions. The approach has been to apply this filter as a simple ranking of species pool members on the basis of their broad habitat abundance index multiplied by the dispersal index (Box 7). The results should be used as an indication of a need for heightened vigilance regarding good dispersers likely to be locally abundant, rather than a direct prediction of the order or timing of the appearance of propagules.

The reason for developing and applying this filter is also related to the fact that the predicted probabilities from GBMOVE are not validly interpreted as predictions of actual presence on a site but predictions of the suitability of abiotic conditions for each species. This important aspect is explained in Box 6. From this it follows that predictions of change in CSM indicator species can be given and interpreted as changes in the suitability of the monitored patch for these species even though they might be absent. The reason for doing this is that such changes amount to an estimate of the ease with which an indicator could establish and persist if it managed to disperse into the patch. To use a medical analogy; it is clearly reasonable and useful to measure changes in the susceptibility of a group of people to some disease even if they do not actually have the disease. It amounts to a risk assessment. The empirical approach to assessing potential immigration from the local species pool based on broad habitat composition and dispersal ability, compliments the modelled changes in habitat suitability (see Box 4). To extend the medical analogy again, it attempts to estimate the likelihood of exposure of a group of subjects (monitored Priority Habitat patches) to a pathogen (propagules of a negative indicator species) independently of an assessment of variation in susceptibility to infection among the subjects (ie. modelled habitat suitability for the absent indicator species).

A good example of a pattern that we would hope to capture by quantifying how regional change in habitat abundance and species dispersal combines with changes in habitat suitability, comes from the study of trends and outbreaks in species in the Rothamsted Park Grass time series (Dodd et al 1995). In the acidified plots, *Chamerion angustifolium* appeared in 1944, reached peak abundance in 1946 then rapidly declined. During this period source populations were widespread in the bomb sites of post-war London but disappeared under new buildings. Assuming that this species is a CSM indicator of interest, our approach to modelling change in habitat suitability would be aimed at estimating whether conditions in the experimental plots were becoming more or less favourable for persistence of *C.angustifolium* even though absent. Estimating immigration potential then rests on an empirical synthesis of information on the abundance of favoured habitat in the surrounding area combined with the fact that the species is readily dispersed by wind. The method is crude yet simple to apply to every protected site in the UK and every CSM indicator species. Validation would however be desirable at least by conducting a sample of site visits and rapid vegetation survey to confirm the local accuracy of LCM2000 and the relative abundance of preferential species between Broad Habitats.

4.3 Assessment of the case-study results in terms of predicted impacts on CSM attributes and condition change

In the following table, predictions of change in CSM indicator species, species present at time 1 and plant functional types on test sites, are summarised (see Appendix 9 for full details of results and methods). The predictions should be treated with caution given the likely variation in reliability of GBMOVE models and the uncertainties in the calibration equations, which are particularly evident in fertile systems.

<table>
<tr><th>Site</th><th>Scenario</th></tr>
<tr><td>R1. Rothamsted Park Grass – lowland meadow</td><td>FRAME/GANE prediction of N and S to 2050.
Continued annual hay crop with no fertilizer addition.</td></tr>
<tr><td colspan="2">• Hay offtake expected to completely offset atmospheric N inputs. Condition expected to remain stable or improve as positive indicators are favoured while conditions remain less favourable for negative indicators.</td></tr>
<tr><td>R2. “</td><td>As above but hay cropping ceased in 2005.</td></tr>
<tr><td colspan="2">• Harmer et al (2001) reported succession to woodland in 20-40 years at Broadbalk following abandonment in the mid-19th century. SUMO predicted grass and herb dominance would gradually give way to woodland after about 60 years. GBMOVE predicted that exactly the same community type would develop on the control plots as present at Broadbalk. Condition would expected to deteriorate within 10 years as positive indicators declined and negative indicators increased finally giving way to a eutrophic secondary broadleaved woodland assemblage.</td></tr>
<tr><td>C1. Climoor – upland heath</td><td>FRAME/GANE prediction of N and S to 2050.
Management stable.</td></tr>
</table>

• Negative indicators, particularly *Agrostis stolonifera*, expected to see a very slight increase in habitat suitability but overall conditions remain much more favourable to the dwarf shrubs and mosses that currently dominate. Condition would probably not be impacted because predicted changes were so small. The high habitat suitability indices for species present gave confidence in this prediction.	
M1. Moorhouse – **blanket bog**	FRAME/GANE prediction of N and S between 1973 and 2001. Match of observed versus predicted change in CSM indicator species in the Hard Hills control plots; 0.1 sheep per hectare and no burn since 1954.
• Moorhouse showed the best relationship between observed and predicted species changes. The results are summarised for CSM indicators showing that pleurocarpous mosses increased as expected while *Vaccinium vitis-idaea* and *Sphagnum* species decreased. Despite being consistent, observed and predicted changes were small. At the scale of the whole plot, condition is very obviously still favourable. Pollutant deposition impacts are expected to be low key in the next 10 to 50 years despite estimated exceedance of the upper CL for blanket bog (see Table 5 in Part 2).	
P1. Porton Down – **lowland calcareous grassland**	FRAME/GANE prediction of N and S to 2050. Site management unchanged.
• The lower CL for nitrogen is estimated to be currently exceeded at Porton. The effects of ongoing predicted decrease in soil C/N are expected to outweigh the very small predicted increase in soil pH. The consequence is a decline in habitat suitability for positive indicators that is expected to be very rapid for some species over the next five years. Hence, a negative impact is expected on condition status. The predictions are best treated cautiously as MAGIC did not calibrate to the measured soil pH, underestimating it by 1 unit. High pH could be critical in constraining predicted declines in calcicoles and increases in more nutrient-demanding species.	
Ca 1. Cairngorm – **upland heath**	FRAME/GANE prediction of N and S to 2050. Zero deer and sheep grazing pressure.
• Ongoing succession at the site (R.Brooker pers.comm., Thurlow et al (1999)) is expected to favour negative indicators such as Bracken and *Ranunculus repens* with *Calluna* declining as shade builds up. Very small N deposition impacts would presumably be expected since the lower CL is not exceeded. Hence the major driver is the incremental increase in canopy height applied to mimic succession. The predictions are best treated cautiously as species *in situ* had lower predicted habitat suitability indices than a range of absent negative indicators.	
B1. Budworth Common – **lowland heath**.	FRAME/GANE prediction of N and S to 2050. Site management unchanged.

<table>
<tr><td colspan="2">Although the upper CL is exceeded, N deposition impacts are expected to be very slight. Calluna is predicted to remain dominant but with a a slight increase in habitat suitability for Deschampsia flexuosa and the negative indicator Bracken. While condition would probably not be expected to change negatively, vigilance might be required to monitor the status of negative indicators and grasses. High habitat suitability indices for the species present lend confidence to the prediction.</td></tr>
<tr><td>NI 1. Dead Island Bog – raised bog</td><td>FRAME/GANE prediction of N and S to 2050.
Site management unchanged.</td></tr>
<tr><td colspan="2">MAGIC did not calibrate well at this site probably because soil sampling focussed on the sloppy peat below the lagg fen surface. Soil C/N, wetness and soil chemistry probably did not reflect the grassy more mesophytic vegetation developed over the quaking surface. Consequently, habitat suitability indices were very low for the species present hence the MAGIC and GBMOVE prediction must be considered very uncertain. Little change in habitat suitability indices was expected even though soil C/N is expected to continue declining. It is possible that high pH and very wet conditions constrain the responses of more nutrient-demanding species. Condition might therefore change rapidly if drying out or succession occurred.</td></tr>
<tr><td>NI 2. Dromore Motte – lowland meadow</td><td>FRAME/GANE prediction of N and S to 2050.
Site management unchanged.</td></tr>
<tr><td colspan="2">The low measured soil C/N appeared to be inconsistent with the unimproved meadow assemblage supported at the site. MAGIC again did not calibrate well to observed soil pH resulting in low habitat suitability indices for the species present. However, predicted changes in suitability were expected to be negligible despite exceedance of the upper CL for lowland meadow. The prediction at this must however be treated with a great deal of caution.</td></tr>
</table>

Overall, predictions appear to be more reliable for more peaty, less-fertile systems, such as Climoor and Budworth Common. MAGIC appears to calibrate better to measured soil chemistry at these sites while GBMOVE predictions at higher soil C/N values are less influenced by the uncertainty in the soil C/N versus mean Ellenberg calibration equation.

With these caveats in mind, expected impacts of modelled N and S deposition on condition status can be seen to vary between study site. The least impact on positive and negative indicator species was predicted for the upland heathland site at Climoor lowland heath at Budworth and blanket bog at Moorhouse. In all cases characteristic species composition is predicted to remain reasonably stable. Negative indicators, such as *Agrostis stolonifera* (Climoor), *Chamerion angustifolium* and *Pteridium aquilinum* (Budworth), are expected to find conditions more favourable in the next 10 to 50 years but predicted increases in habitat suitability were always very small. This was despite current exceedance of their lower empirical Critical Loads for N (upper limits also currently exceeded at Budworth and Moorhouse). At these three sites predictions maybe more reliable because MAGIC calibrated well to current soil measurements and species observed on site were matched by high predicted habitat suitability indices from GBMOVE.

At Porton Down, Dead Island Bog and Rothamsted, predicted changes are much less reliable because MAGIC did not calibrate satisfactorily to either soil pH or soil

C/N. The size of these discrepancies is likely to have a significant impact on the realism of GBMOVE predictions. For example, at Porton Down predicted soil pH was 1 unit lower than current values. This reduced the predicted suitability of conditions for a range of characteristic calcicoles present on site and is also likely to have resulted in overestimation of the vulnerability of the system to ongoing predicted reductions in soil C/N. For example, the suitability of conditions for positive indicators for CG2 grassland present in monitoring quadrats in 2000 were predicted to decrease dramatically to near zero by 2010. On the face of it, the results suggest that Porton Down is highly vulnerable to N deposition impacts and this will manifest itself by dramatic reductions in suitability for positive indicators before the PSA reporting year. However the predictions at Porton as well as at Dead Island Bog and Dromore Motte are probably not reliable because MAGIC did not calibrate satisfactorily to current conditions.

The two case studies at Rothamsted suggest that the predicted impact of the cumulative effect of historical and future changes in N and S deposition can be offset by management. Continued hay offtake with no added fertilizer is predicted to more than compensate for the exceeded Critical Load leading to an increasingly unproductive system unfavourable to negative indicators. Conversely, ceasing hay offtake triggers a successional sequence where conditions immediately start to become more favourable to nutrient-demanding forbs and grasses leading to plausible increases in negative indicators within 10 years.

Test site predictions of future change in species already present and CSM indicators in local species pools suggest a number of key features about the impact of atmospheric N deposition in the next 10 to 50 years:

1. Atmospheric deposition impacts are generally expected to be minor on heath and bog test sites even where the empirical nitrogen Critical Load is exceeded and even where dispersal of negative indicators is not limiting, for example at Budworth.
2. The heathlands at Budworth and Climoor, appear to show a degree of resistance to the impact of decreasing soil C/N. This is partly because consistently low soil pH keeps habitat suitability low for negative indicators. Similarly, the suspicion at Dead Island Bog was that continued high soil pH and soil saturation would limit habitat suitability for negative indicators even if soil C/N declined.
3. The further implication is that sudden shifts along other key gradients driven by succession, drought or drainage could trigger an accumulated ecosystem response to N deposition driven reduction in soil C/N.
4. Tantalizing glimpses of the potential impact of climate change were also apparent but could only be explored in a limited way by the developing models. For example at Moorhouse, the current annual average temperature places the Hard Hills site at the tree line. Under a UKCIP high emissions scenario of increase in annual average temperature for the site, SUMO predicts increased Dwarf Shrub Heath productivity in the next 100 years, with tree establishment effectively still limited by soil pH and ground wetness. However several ecosystem level changes could occur together to make future outcomes much more uncertain. For example, recent periodic Summer drought is already known to have been associated with acute drops in soil pH. Does this then portend a future threshold effect on the integrity of the surface peat and on soil water pH?. What will be the antagonistic effects of increasing rainfall or

rainfall intensity? For example will peat still form yet will erosion risk increase? How will these changes interact or be deflected by tree establishment? If climate change becomes more important then models need to increasingly incorporate their effects and interactions with chronic responses to changing pollutant deposition.

4.4 Interpretation of project results in terms of policy indicators and targets

Habitat Action Plan targets[12]

The relevance of model test results to the achievement of current policy targets is limited by their level of reliability and consequent limitation on progress in applying these models to serious scenario testing. Predictions of change in habitat suitability for CSM indicators are thought more reliable for heath and bog habitats but ought still to be treated as hypotheses to be tested against observed change. The implications of testing results are therefore summarised below for relevant Priority Habitat Action Plan targets:

<u>Upland heathland</u>

T2: Achieve favourable condition on all upland heathland SSSIs/ASSSIs by 2010.

T3: Achieve demonstrable improvements in the condition of at least 50% of semi-natural upland heath outside SSSI/ASSSs by 2010.

Case-study application of MAGIC and GBMOVE at Climoor indicated minor expected impact of N deposition on plant species composition despite lower CL exceedance in 2000. If other upland heaths exhibit a similar degree of expected response, it is possible that any minor shifts in favour of negative indicator species are very unlikely to outweigh continued expected dominance by dwarf shrubs. Such minor changes are also likely to be much smaller than those induced by managed reduction in grazing pressure carried out to address the highest current threat to the Priority Habitat.

<u>Blanket bog</u>

T1: Maintain the current extent and overall distribution of blanket mire currently in favourable condition.

T2: Introduce management regimes to improve to, and subsequently maintain in favourable condition a further 280,000ha of degraded blanket mire by 2010.

The relevance of test results is limited to the model application at Moorhouse. On this site recent changes were consistent with the predicted impact of changing N and S deposition, yet changes since 1973 were minor and the blanket bog unit remains in favourable condition largely because sheep grazing is very low with long rotation burning only. Pollutant deposition impacts in the next 10 years are expected to be

[12] www.ukbap.org.uk

similarly modest despite estimated exceedance of the upper empirical CL in 2000. The extent to which the results from Moorhouse can be generalized across similarly managed systems remains questionable. If they are more widely applicable, the results suggest that N deposition is not likely to represent a marked constraint to achievement of these HAP targets.

Lowland heathland

T2: Improve by management all existing lowland heathland currently in unfavourable condition.

The highest threats to lowland heath are considered to be invasive aliens and scrub encroachment followed by N deposition (see Table 1). Because of the limited extent of lowland heath in the UK, the target presumably requires that some of the heathland targeted for remedial management will have been impacted by chronic N deposition. Experimental N addition has shown that soil C/N does decrease in response, however much evidence exists to support the efficacy of interventions that vary in the amount of accumulated N that can be removed. This includes turf stripping, mowing, burning and grazing (Barker et al in press).

Results from this project were based on model testing at Budworth where only minor changes in habitat suitability were expected despite exceedance of the upper CL. The key constraint here as in other high soil C/N, low pH sites maybe P limitation (J.Carroll pers.comm.; Manning et al 2004) although some monocots such as *Molinia caerulea*, which is present at Budworth but restricted by ground wetness, can still apparently capitalise on elevated N despite low P supply (Hogg et al 1995). The implication is that Budworth is buffered to some extent against high ambient N deposition although the nature of this buffering and its relevance to other lowland heathland sites is not made any clearer from our model applications.

Lowland meadows

T2-T5: These targets all refer to achievement of favourable condition by 2005 to 2015.

Since the highest threat to the Priority Habitat in 2002 was considered to be agricultural improvement (see Table 1), site safeguard is presumably the highest priority. Once under positive management, case-study application of models at Rothamsted suggests that annual hay offtake can more than compensate for N deposition impacts resulting in greater N outputs than inputs and increased soil C/N. Because unimproved meadows are agricultural systems they are still typically maintained by N and P inputs from farm yard manure that often exceed ambient N deposition. This is inevitably part of the reason why N deposition was ranked fifth as a threat to the Priority Habitat in 2002 (Table 1). The implication is that lowland meadows are at much less risk from atmospheric N deposition than other threat factors.

Lowland raised bog

T1: Maintain the current distribution of primary near-natural lowland raised bog peat in the UK.

T2: Ensure that the condition of the current resource is maintained, where favourable, or enhanced where unfavourable, through appropriate management.

Model testing was only carried out at Dead Island Bog, Northern Ireland where MAGIC did not calibrate to current soil conditions. Hence the results provide a very uncertain basis for assessing the vulnerability of the resource to N deposition.

Regional and national indicators of biodiversity

A key constraint on using the developing models for larger scale scenario testing is that soil chemistry parameters used to initialise MAGIC must be based on representative values for prediction of change in a wider extent of a particular Priority Habitat or even Broad Habitat and its associated soil type. Such up-scaling issues are being addressed as one of the core objectives of linked soil-vegetation model development in the DEFRA Terrestrial Pollution Umbrella project. The objective is that regional assessments could test scenarios of habitat change based on differences in management between designated and undesignated sites as well as generating meaningful UK or GB-wide maps of interpolated change for particular soil types and associated plant species. Translation of such outputs into current large-scale biodiversity indicators is challenging but could be rapidly tested for species richness using the statistical model developed in this project although unacceptably high uncertainty about predictions would be expected given our test results.

4.5 Combining model based analysis of change with a national database of risk values for Priority Habitats in designated sites

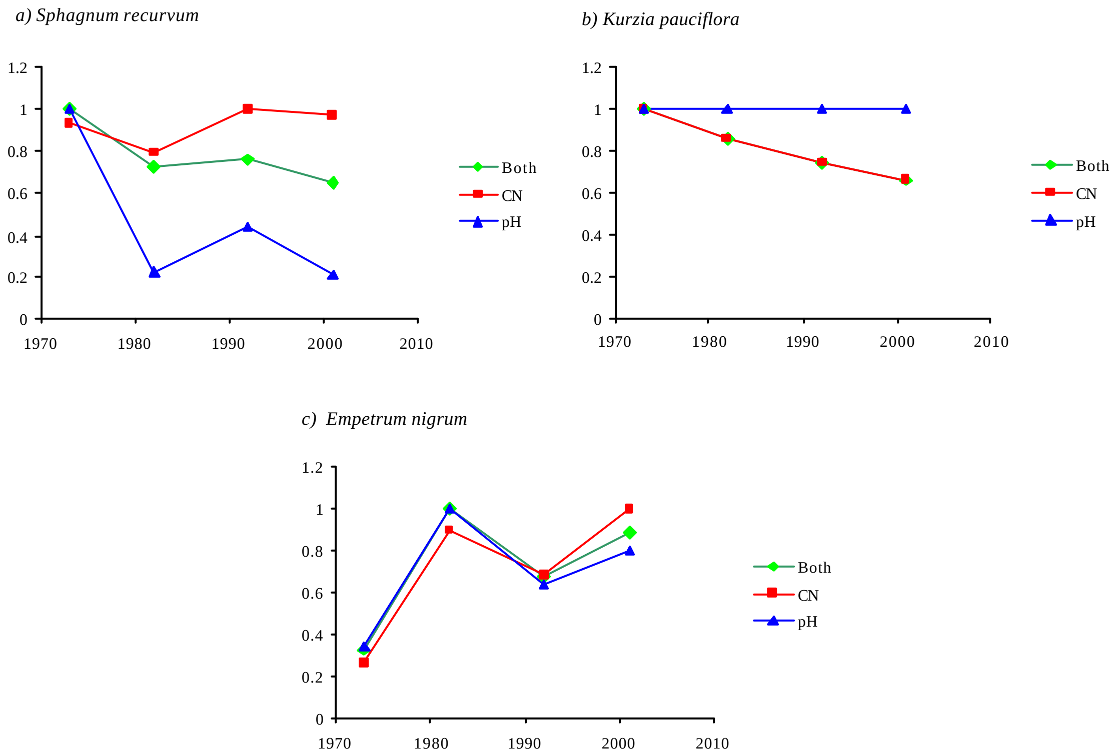
The two parts of this project have sought to develop complimentary approaches to analysing the eutrophying impacts of N deposition on Priority Habitats. Part 1 developed and tested linked soil-vegetation and succession models for testing scenarios of different drivers on condition change. Part 2 scoped the availability of UK-wide data for risk factors potentially impacting Priority Habitats on designated sites. Pending further model development it is possible to envisage how scenario testing tools could be combined with a comprehensive classification of sites and Priority Habitats by risk factor values to select sites where model-based analyses of different combinations of potential drivers are most relevant. For example, sites at high risk of flooding and surrounded by a high proportion of intensive land-use would be most relevant for modelling the relative magnitude of predicted atmospheric N and S deposition impacts on a Priority Habitat plus the additional impact of a simulated major increase in N and % soil moisture from a flood event. Hence, on such a site, attribution of observed change and risk assessment of future change could be informed by modelled scenarios of each effect separately and in combination (see Box 10). Likewise, small sites that have seen the largest increase in growing season length might be most relevant for testing scenarios combining atmospheric N and S deposition with and without expected changes in maximum July, minimum January temperatures and precipitation. Such sophisticated risk analyses rely on further model development to overcome some of the reasons for poor performance revealed in this project. The current status of the linked models is summarised in the next section.

BOX 10 - Application of MAGIC and GBMOVE species predictions for signal attribution and scenario testing

Pending ongoing validation and testing, predictions from MAGIC/GBMOVE and SUMO can be used in signal attribution analyses as well as in testing the sensitivity of species to different drivers of change. As a simple example, species changes at Moorhouse were modelled using MAGIC and GBMOVE. Model runs were based on allowing either soil C/N or pH to change while holding the other variable constant, versus a prediction based on both variables changing. Clearly, the two soil properties are in reality correlated so that when C/N decreases pH may increase independently of an S deposition effect. However, this example illustrates how the estimated importance of pH versus soil C/N changes are expected to differ between species with different response profiles at particular segments of key environmental gradients. For example the GBMOVE model for *Sphagnum recurvum* shows that under conditions at Moorhouse, it is expected to benefit from reduced soil C/N but this is outweighed by the greater negative impact of increasing pH (Fig 1a). The small liverwort *Kurzia pauciflora* is different and is not predicted to show any response to the predicted magnitude of increase in pH at Moorhouse but is negatively impacted by reduced soil C/N. Hence, changing soil C/N explains the entirety of this species' response (Fig 1b). *Empetrum nigrum* appears to be a species where the effects of pH and soil C/N change are equivalent (Fig 1c). Since species do not respond instantaneously to condition changes and are also impacted by other factors, the results of these kinds of model application are best interpreted as hypotheses of change in habitat suitability rather than predictions of change in abundance.

Two caveats apply to such sensitivity analyses. Firstly, the emphasis is on predicting change in habitat suitability so that even if a species is present it is not assumed that it responds instantaneously to changing abiotic conditions. Secondly, the predicted sensitivity is a reflection of whether linear and quadratic terms were significant in each GBMOVE model for a particular environmental gradient. A non-significant term for wetness would suggest that the species could occur equally under wet or dry conditions, whereas in reality this could reflect lack of effective sampling in the training data. This issue emphasises the importance of closer inspection of GBMOVE as part of a further process of identifying a core set of the most reliable models.

Fig 1. Partitioning species sensitivity to modelled soil C/N versus pH change. An analysis of the relative sensitivity of each modelled species at Moorhouse was carried out by running MAGIC+GBMOVE predictions with either soil C/N or soil pH held constant over the time period and comparing against predicted change when both parameters were allowed to vary. The y-axis is predicted habitat suitability standardised to between 0 and 1.



4.6 Is the linked series of core model components currently fit for the purpose of predicting scenarios of condition status on UK Priority Habitats?

At present, MAGIC and GBMOVE can be applied to Priority Habitat patches given reliable soil chemistry data and scenarios of future N inputs and N offtake. The uncertainties in the current calibration equations suggest that results should be very cautiously interpreted in low soil C/N systems. An initial assessment of the reliability of MAGIC+GBMOVE predictions on a specific site can be gained by examining firstly whether MAGIC calibrates to current soil measurements, and secondly whether species present on site match with high predicted habitat suitability scores (see section 4.3).

The case study applications show the range of outputs possible (Appendix 9). Coverage of CSM indicator species by GBMOVE is very comprehensive (Appendix 6) while SUMO can also predict structural attributes such as cover of trees and Dwarf Shrubs. More work is however required to accurately model grass versus forb cover particularly in response to grazing. Also, if scenarios of the simultaneous impact of changing vegetation management and pollutant deposition on species composition are to be explored, a succession model such as SUMO ought to be formally integrated with the soil and plant species models. Currently this lack of integration means that succession is implemented by changing canopy height according to published evidence or expert knowledge on rates of vegetation change in particular habitat locations. Although crude, this is a simple approach and actually worked well in an attempt to simulate assembly of an observed woodland community type at Rothamsted. In the absence of formal integration canopy height predictions from SUMO can be manually fed into MAGIC/GBMOVE. Hence, at present MAGIC and GBMOVE can be used to simply predict future change in habitat suitability of CSM indicators and species present. Predictions are likely to be more unreliable in vegetation types with soil C/N ratios below about 13.

Assuming MAGIC, GBMOVE and SUMO produce increasingly robust predictions in higher C/N systems as a result of ongoing testing and refinement, what could these models currently achieve in terms of scenario testing and risk analysis? The simplest strategy requiring the least model integration would involve modelling the impact of managed disturbance using SUMO canopy height output as input to GBMOVE and the impact of pollutant deposition using soil C/N and soil pH from MAGIC as input to GBMOVE or by directly changing mean Ellenberg R and N values by amounts that reflect published evidence of the effects of specified levels of N input on particular community types (eg. Jones 2004; Smart et al 2004). The latter strategy would also have the advantage of not relying on the soil C/N calibration equation hence the uncertainty contributed by this equation would not impact predictions of change. Changes in soil moisture, resulting for example, from drainage and chronic drought could also be empirically applied using rates of change in soil moisture content from the published literature and observational studies. Climate change impacts on species' habitat suitability ought to be feasibly modelled by GBMOVE pending further testing of GBMOVE+climate variable models.

4.7 Recommendations for further work

<u>Model testing</u>

The evidence presented is neither enough to justify discarding the approach presented nor to place complete trust in predictions based on the state of current model development. While this situation is unsatisfactory there is no further foolproof single test that will guarantee a clear pass or fail. Instead credibility will accumulate gradually as cycles of testing and development occur. As a result of this work, it is clear that further validation and testing of GBMOVE and calibration equations is desirable. Useful datasets would include soil measurements and botanical records from the other Rothamsted experimental plots. Regarding further tests of species temporal change consistent with N deposition effects, it should be possible to exploit other long term vegetation change data from fixed locations although this would require new soil chemistry analysis, biomass measurement for SUMO testing, and a strategy to investigate the quality and availability of these data. Possibilities include schemes recorded in Hill and Radford's (1984) inventory of fixed plots, the DOE calcareous grassland plots located across Britain some of which have been recently re-recorded as part of a bryophyte monitoring project (M.Ashmore pers.comm.) and the long-term Bibury road verge sampling scheme (Dunnett et al 1998). Given the possibility of methodological and sampling error and lack of representativeness of soil sampling data gathered by third parties it is probably desirable for further soil sampling to be under the strict control of the model testing team or at least carried out to the specification set out in Appendix 11.

Further testing of GBMOVE and calibrated links to soil C/N, % soil moisture and pH only require paired vegetation and soil measurements at any point in time, hence other ESA monitoring quadrat datasets and published or extant experimental datasets could provide useful additional test material.

Uncertainty and sensitivity analysis

The production of easily applicable routines for uncertainty and sensitivity analysis of GBMOVE models and calibration equations are sorely needed. This work will be carried out as part of the currently funded DEFRA air pollution umbrella program.

Further development of within-vegetation type calibration equations

This work will also be carried out as part of the air pollution umbrella program. Also included under this heading is the requirement for a transfer function to allow soil water pH to be estimated from soil slurry pH and *vice versa*.

Further testing of SUMO and integration with MAGIC and GBMOVE

Integration of SUMO with MAGIC and GBMOVE is ultimately a necessity if a capability is to be developed for modelling the simultaneous impact of succession, land-management, climate change and pollutant deposition. SUMO has shown sufficient promise in this project that we believe plans for integration plus further testing under UK conditions, are justified.

Analysis and targeting of the best GBMOVE models

A species-by-species assessment of the robustness of GBMOVE models would be desirable to identify a core set of reliable models. It would make sense to target CSM indicator species applicable to heath and bog Priority Habitats as an initial selection criteria.

Coupled with the need for sensitivity analyses, it would be relatively straightforward but useful to generate measures of the sharpness of species' optima as a guide to their predicted sensitivity to changing conditions. This may be better handled under the general heading of sensitivity analysis. It would certainly be of use to be able to rank CSM species according to their likely sensitivity to change along each abiotic gradient.

<u>Further development of rare/subordinate species models</u>

Initial work showed promise. Further model generation would not be time consuming. More challenging would be identifying test sites against which to test predictions of temporal change in rare species suitability against observed data.

<u>Completion of dispersal indices for CSM indicators and validation of immigration rankings around test sites</u>

A number of CSM indicators were not attributed with dispersal indices because attribute data were lacking. With the imminent completion of the European LEDA trait database, there are good prospects for filling in indices for the remaining indicator species. In addition, validation of the local abundance ranking of CSM indicators based on LCM2000 broad habitat coverage ought to be carried out for test sites to validate the approach.

These further avenues of model development are prioritised below considering which steps will lead to the most rapid progress in the shortest time.

<u>1. Development of within vegetation type calibration equations.</u> The soil C/N and soil pH calibrations both contribute very high uncertainty but are essential for linking MAGIC with GBMOVE. The soil C/N relationships is particularly problematic in more fertile, grassland systems while the soil water versus soil slurry pH issue must also be addressed. Further construction of calibration relationships based on smaller subsets of floristically similar quadrat data should rapidly highlight whether models with considerably greater explanatory power can be achieved based on soil C/N or whether soil C/N is ultimately inadequate as a robust correlate of vegetation productivity in some Priority Habitats.

<u>2. Determination and further testing of a core subset of of GBMOVE models for selected CSM indicators.</u> Given the better performance of MAGIC and GBMOVE in heath and bog systems, an obvious option would be to focus on these Priority Habitats and build credibility and understanding by a more focussed application to these ecosystems. This would require additional searching for test data and new soil chemistry analysis.

<u>3. Assembly of risk factor values for UK/GB and construction of a Priority Habitat by site classification.</u> Assembly of UK or GB wide information would be relatively straightforward and reasonably quick. It would not require development of a standardised agricultural production index if major products were grouped. A classification of Priority Habitats within sites by risk factor values would provide a rapid way of locating sites at risk of N deposition exacerbated by multiple external factors. It would also help decision makers in site selection for monitoring air pollution impacts.

Other developments are of lower priority since the feasibility of linking GBMOVE and MAGIC ultimately depends upon robust calibration equations. This problem therefore needs to be addressed urgently.

4.8 Matching project outputs to the original aims and objectives

1. Review the current knowledge base for atmospheric nitrogen pollution impacts on biodiversity.

Given the recent completion of major reviews of this nature, the requirement was amended to focus specifically on two topics. First, an assessment and summary of the evidence-base for the effects of interactions between N deposition and other key drivers of change (Appendix 7). Secondly, a literature review also focussed on analytical approaches to signal attribution where the challenge is to partition an observed response between different potential drivers (Appendix 8).

2. Further develop and test modelling techniques to help quantify the impacts of atmospheric nitrogen deposition on biodiversity nationally.

3. Apply the modelling techniques to a sample of habitats and sites to examine current and projected levels of he nitrogen threat (from atmospheric and other sources) to habitats and sites of high nature conservation importance.

The majority of the work carried out was devoted to these two tasks. Specifically, model development and testing is reported in section 2. Applications to future projection are reported in Appendix 9.

4. Provide a preliminary interpretation of theresults with respect to achievement of: i) the Public Service Agreement target for achieving favourable condition on SSSIs; and ii) Biodiversity Action Plan targets for priority habitats and species and related indicators of biodiversity (eg. Country Biodiversity Strategy indicators).

Interpretation of model applications to policy targets is discussed in section 4.4.

5. Develop proposals for Phase 2 of this work which should allow for a wider geographical application of the models.

Recommendations for further work are outlined in section 4.7. These recommendations do not include proposals for a wider geographical application since more fundamental developments are yet required to secure a modelling capability that could be rolled out to a larger number of designated sites across the UK.

Assessing the relative impact of sources of N and additional risk factors

The requirement to model differential impacts of wet *vs* dry deposition, reduced N *vs.* oxidised N deposition, and fertiliser *vs.* atmospheric inputs, has not been met. Effects of wet *vs* dry and reduced N *vs* oxidised N are not well-established and may require more experimental work. There is effectively no difference between fertiliser and atmospheric N inputs, although fertiliser N is often accompanied by P and K applications which have additional effects, while fertiliser N is typically applied for agricultural conversion and maintenance at loads in excess of atmospheric N deposition. Fertiliser effects are in any event less relevant to eutrophication impacts on designated sites since these are unlikely to be deliberately fertilised. There may be scope in future for using the DEFRA census data held on EDINA to separate effects of fertiliser and deposition. With this in mind, the role and importance of adjacent intensive land-use as a contributory risk factor to designated sites was covered by the scoping activity in section 3.

5 Acknowledgements

We thank the following for model testing data; Elena Vanguelova at the Forestry Commission, Paul Poulton and Paul Hargreaves at IACR Rothamsted, John Adamson, Mandy Lane, Lorna Sherrin, Rob Brooker, Clive Bealey and Mike Morecroft for ECN data and advice. We thank staff at DEFRA York who helped access fine-scale agricultural census data. Also we are very grateful for the guidance, encouragement and advice of Jeff Kirby who managed the contract along with the project Steering Group comprising Samantha Baker, Ian Strachan, Claire Whitfield, Jane Goodwin, Andy Stott, Richard Brand-Hardy, Mike Ashmore and Peter Coleman. We also thank the following for advice and information; Clare Whitfield, Jacky Carroll, Simon Caporn, Lawrence Jones, Han van Dobben, Dave Morris, Isabel Alonso, Eleanor Hewins and Vikki Morgan. Dan Osborne also arranged a contribution from the CEH science budget to purchase the AgCENSUS licence. Particular thanks are extended to Mike Ashmore for his advice as an expert reviewer of this work. The project was funded jointly by JNCC, EN and DEFRA and was also supported by work carried out under the Terrestrial Pollution Umbrella Project jointly funded by DEFRA and NERC.

6 Glossary of acronyms

ASSI: Area of Special Scientific Interest. The name of statutory designated wildlife sites in northern Ireland.

BAP: Biodiversity Action Plan. Published in its first part in 2000, the UK BAP includes action plans for species, priority habitats and statements for broad habitats, which are a coarser and inclusive classification of British habitats.

BRC: Biological Records Centre based at the NERC Centre for Ecology and Hydrology at Monkswood (www.brc.ac.uk).

C/N: Soil carbon to nitrogen ratio.

CCW: Countryside Council for Wales (www.ccw.gov.uk).

CSM: Common Standards Monitoring. This is the cross-agency approach to monitoring change in the condition of habitats on ASSI/SSSI across the UK. Guidance has been prepared for upland and lowland habitats (see guidance notes available online at www.jncc.gov.uk) and sets out habitat definitions, attributes and criteria for measuring condition. Attributes include structural features such as the extent of dwarf shrub heath and bare ground plus presence of key indicator species. Application of CSM ensures that condition assessments on all designated sites provide consistent data that can be used to measure national policy targets for SSSI condition (www.defra.gov.uk/corporate/busplan/psa2004.htm).

DEFRA: Department of the Environment, Food and Rural Affairs.

ECN: Environmental Change Network. The network was established in 1992 and currently consists of 12 terrestrial and 44 aquatic sites. At each site a core set of environmental and ecological measurements are made within particular habitats. Measurements range from daily to decadal. All are carried out according to standard protocols devised by the ECN co-ordination section at CEH Lancaster. See www.ecn.ac.uk.

EN: English Nature (www.english-nature.org.uk).

FCS: Favourable Conservation Status refers to the condition of designated wildlife sites and species across the UK as defined by the EEC Habitats Directive. See Alexander (2003) and www.jncc.gov.uk/pdf/comm02d07.pdf for discussion and definitions.

FRAME: Fine Resolution Atmospheric Multi-pollutant Exchange Model. A model developed and operated by CEH Edinburgh. It generates predictions of the deposition of wet and dry, reduced and oxidised nitrogen and sulphur deposition at the $5km^2$ scale across Britain. See Box 1 for a summary description and www.frame.ceh.ac.uk.

GANE: Global Atmospheric Nitrogen Emissions. The name of a thematic program of cross-institutional research established and funded by the UK Natural Environment Research Council. The program ran between 2001 and 2003 and explored the pathways, fate and ecological impacts of nitrogen in the natural environment. Key papers were published in Water, Air & Soil Pollution: Focus, volume 4, number 6 (2004).

GB: Great Britain. England, Wales and Scotland excluding northern Ireland.

GBMOVE: The original MOVE model was developed as a series of statistical niche descriptions for Dutch higher plants. As part of this project, models were developed for British plants using similar methods. Each species model consists of a multiple logistic regression equation where probability of occurrence is predicted by explanatory variables relating to abiotic and climatic gradients. See Box 4 for further details.

GSL: Growing Season Length. Number of days per year when the average daytime temperature exceeds 5^oC..

HAP: Habitat Action Plans. These plans set out action required for the conservation and restoration of specified areas of each of 45 marine and terrestrial priority habitats in the UK. Plans are monitored and implemented at national and local levels (www.ukbap.org.uk).

JNCC: Joint Nature Conservation Committee (www.jncc.gov.uk).

LCMGB: Land Cover Map of Great Britain. Satellite recorded, census map of GB. The first LCM was produced in 1990. A new product was generated in 2000 using updated technology for classifying pixels into new categories that equate with the UK broad habitat classification. LCMGB is produced by the Earth Observation section at CEH Monkswood (see http://science.ceh.ac.uk/data/lcm/LCM2000.shtm).

MAGIC: Model of Acidification of Groundwater In Catchments. This is one of a series of extant dynamic soil models. It predicts changes in soil chemistry based on equilibration between changing inputs and outputs of cations and anions. In this project MAGIC was used to generate annual predicted changes in soil C/N ratio and pH in response to atmospheric deposition of S and N and, in some situations, managed offtake of N. See Box 2 for a summary description.

MIRABEL: A project that developed methods for assessing risks to European landscapes of the impacts of atmospheric N deposition, farming intensification and land abandonment. See Petit et al (2003) for further details.

NEGTAP: National Expert Group on Transboundary Air Pollution. A wide-ranging review was carried out by national experts to assess the current measurements and models of pollutant emission, deposition and ecological impacts. The report is available on-line at www.nbu.ac.uk/negtap/.

NVC: National Vegetation Classification. A comprehensive phytosociological classification of British plant communities compiled by Professor John Rodwell at the Unit of Vegetation Science at Lancaster University.

SAP: Species Action Plans. These plans set out action required for the conservation and restoration of specified areas of each of 391 scarce species in the UK. Plans are monitored and implemented at national and local levels (www.ukbap.org.uk).

SNH: Scottish Natural Heritage (www.snh.org.uk).

SSSI: Site of Special Scientific Interest. The name of statutory designated wildlife sites in GB.

SUMO: SUccesion MOdel. A process model developed by Wieger Wamelink at Alterra, Netherlands (www.alterra.wur.nl/UK/). It models biomass growth, death and decay driven by succession and constrained by management, climate and other abiotic conditions. The principal output is biomass production at annual time steps.

UK: Great Britain plus northern Ireland.

UKCIP: UK Climate Impacts Prediction (See www.ukcip.org.uk)..

7 References

Achermann, B, Bobbink, R (2003). *Empirical Critical Loads for Nitrogen*. Proceedings of Expert Workshop held in Berne, Switzerland 11 - 13 November 2002. Environmental Documentation No. 164. Air. Swiss Agency for the Environment, Forests and Landscape SAEFL, Berne.

ADAS (2000) *Fertiliser Recommendations for Agricultural and Hotricultural Crops* (RB209). 7TH edition. MAFF. HM Stationery Office, London.

Aerts, R, Wallen, B, Malmer, N, De Caluwe, H (2001) Nutritional constraints on *Sphagnum* growth and potential decay in northern peatlands. *Journal of Ecology* 89, 292-299.

Alexander, M (2003) *Management Planning for SSSIs, Natura 2000 sites and RAMSAR sites*. Countryside Council for Wales, Plas Penrhos, Bangor.

Baker, CG, Power, SA, Bell, JNB, Orme, CDL (in press) Effects of habitat management on heathland response to atmospheric nitrogen deposition. *Biological Conservation.*

Beier, C., Hultberg, H., Moldan, F., and Wright, R.F. (1995). MAGIC applied to roof experiments (Risdalsheia, N; Gardsjon, S; Klosterhede, DK) to evaluate the rate of reversibility of acidification following experimentally reduced acid deposition. *Water, Air, Soil Pollut*., 85, 1745-1751.

Bellamy, DJ, Rieley, J (1967) Some ecological statistics of a miniature bog. *Oikos* 18, 33-40.

Bettelheim J, Littler A. (1979) Historical trends of sulphur oxide emissions in Europe since 1865. CEGB report PL-GS/E/1/79, CEGB, London

Bischoff, A (2002) Dispersal and establishment of floodplain grassland species as limiting factors in restoration. *Biol.Cons*. **104**, 25-33

Boyer, MLH, Wheeler, BD (1989) Vegetation patterns in spring-fed calcareous fens – calcite precipitation and constraints on fertility. *Journal of Ecology* 77, 597-609.

Bradbrook, K, Waller, S, Morris, D (in press) National Floodplain Mapping: Datasets and Methods – 160,000 km in 12 months. *Natural Hazards*.

Buckland SM, Thompson K, Hodgson JG, Grime JP 2001. Grassland invasions: effects of manipulations of climate and management. *Journal of Applied Ecology* 38, 301-309.

Bullock, JM, Clarke, RT (2002) Long distance seed dispersal by wind: measuring and modelling the tail of the curve. *Oecologia* 124: 506-521

Caporn, SJM, Carroll, JA, Cresswell, N, Edmondson, J, Pilkington, MG, Price, EAC, Ray, N, Wilson, DB (2004) Long-term impacts of enhanced and reduced nitrogen deposition on upland and lowland heath. In: Emmett, BA, McShane, G (2004) Terrestrial Umbrella Final Report. Work Package 2, Task 5. Long-term impacts of enhanced and reduced nitrogen deposition on semi-natural vegetation. Final report to DEFRA (on-line at www.bangor.ceh.ac.uk/terrestrial-umbrella/PDF/report%202004/Task2_2004.pdf).

Clapham, AR (1940) The role of bryophytes in calcareous fens of the Oxford District. *Journal of Ecology* 22, 71-80.

Cosby, BJ, Ferrier, RC. Jenkins, A, Wright, RF (2001) Modelling the effects of acid deposition: refinements, adjustments and inclusion of nitrogen dynamics in the MAGIC model, *Hydrol. Earth Syst. Sci.*, 5, 499-517.

Cosby, BJ, Hornberger, GM, Galloway, JN, Wright, RF (1985) Modelling the effects of acid deposition: assessment of a lumped parameter model of soil water and streamwater chemistry. *Water Res.*, 21, 51-63.

Critchley, CNR, Chambers, BJ, Fowbert, JA, Sanderson, RA, Bhogal, A, Rose, SC (2001) Association between lowland grassland plant communities and soil properties. *Biological Conservation* 67, 100-117.

Critchley, CNR, Smart, SM, Poulton, SMC, Myers, GM (1996) Monitoring the consequences of vegetation management in Environmentally Sensitive Areas. *Aspects of Applied Biology* 44, 193-201.

Cunha, A., Power, S.A., Ashmore, M.R., Green, P.R.S., Haworth, B.J. & Bobbink, R. (2002). Whole ecosystem nitrogen manipulation: an updated review. JNCC report no. 31. JNCC, Peterborough.

Dalton, H, Brand-Hardy, R (2003) Nitrogen: the essential public enemy. *Journal of Applied Ecology* 40, 771-781.

DEFRA (2002) *The government's strategic review of diffuse water pollution from agriculture in England: Agriculture and Water*. Department for Environment, Food and Rural Affairs, London. June 2002.

DEFRA (2003) *Working with the grain of nature: A biodiversity strategy for England*. Department for Environment, Food and Rural Affairs, London.

Diekmann, M (2003) Species indicator values as an important tool in applied plant ecology - a review. *Basic and Applied Ecology* 4, 493-506.

Dodd, ME, Silvertown, J, McConway, K, Potts, J, Crawley, MJ (1994a) Stability in the plant communities of the Park Grass experiment: the relationships between species richness, soil pH and biomass variability. *Phil. Trans. R. Soc. Lon. B* 346, 185-193.

Dodd, ME, Silvertown, J, McConway, K, Potts, J, Crawley, M (1994b) Application of the British National Vegetation Classification to the communities of the Park Grass experiment through time. *Folia Geobotanica Phytotax*. 29, 321-334.

Dodd, ME, Silvertown, J, McConway, K, Potts, J, Crawley, MJ (1995) Community stability: a 60 year record of trends and outbreaks in the occurrence of species in the Park Grass experiment. *Journal of Ecology* 83, 277-285.

Dunnett NP, Willis AJ, Hunt R, Grime JP (1998) A 38-year study of relations between weather and vegetation dynamics in road verges near Bibury, Gloucestershire. *J.Ecol.* 86: 610-623.

Dupouey, JL, Dambrine, E, Laffite, JD, Moares, C (2002) Irreversible impact of past land use on forest soils and biodiversity. *Ecology* 83, 2978-2984.

ECUS (2003) *Prioritising designated wildlife sites at risk from diffuse agricultural pollution.* English Nature Research reports number 551. English Nature, Peterborough.

Ertsen, ACD, Alkemade JRM, Wassen, MJ (1998) Calibrating Ellenberg indicator values for moisture, acidity, nutrient availability and salinity in the Netherlands. *Plant Ecol.* 135: 113-124.

Ettema,CH, Wardle, DA (2002) Spatial soil ecology. *Trends in Ecology and Evolution* 17, 177-182.

Evans, C.D., Pilkington, M.G., Caporn, S.J.M., Wilson, D.B., Carroll, J.A. (In prep) Modelling nitrogen saturation and carbon accumulation in moorland soils in response to nitrogen deposition.

Evans, C.D., Reynolds, B., Jenkins, A., Helliwell, R.C., Curtis, C.J., Goodale, C.L., Ferrier, R.C., Emmett, B.A., Pilkington, M.G., Caporn, S.J.M., Carroll, J.A., Norris, D., Davies, J., Coull, M.C. (In press) Evidence that soil carbon pool determines susceptibility of semi-natural ecosystems to elevated nitrogen leaching. *Ecosystems*.

Fitter, AA (1994) Architecture and biomass allocation as components of the plastic response of root systems to soil heterogeneity. In: *Exploitation of Environmental Heterogeneity by Plants* (eds. Caldwell, MM, Pearcy, RW). Pp305-323. Academic Press: New York.

Fowler, DA, O'Donoghue, M, Muller, JBA, Smith, RI, Dragosits, U, Skiba, U, Sutton, MA, Brimblecombe, P (2004) A chronology of nitrogen deposition in the UK between 1900 and 2000. *Water, Air & Soil Pollution: Focus* 4, 9-23.

Freckleton RP, Watkinson AR (2002) Large-scale spatial dynamics of plants: metapopulations, regional ensembles and patchy populations. *J. Ecol.* 90: 419-434

Geertsema, W, Opdam, P, Kropff, MJ (2002) Plant strategies and agricultural landscapes: survival in spatially and temporally fragmented habitat. *Landscape Ecology* 17, 263-279.

Giller, KE, Wheeler, BD (1988) Acidification and succession in a floodplain mire in the Norfolk Broadland, UK. *J.Ecol.* 76, 849-866.

Goulding, KWT, Bailey, NJ, Bradbury, NJ, Hargreaves, P, Howe, M, Murphy, DV, Poulton, PR, Willison, TW (1998) Nitrogen deposition and its contribution to nitrogen cycling and associated soil processes. *New Phytologist* 139, 49-58.

Grootjans, AP, Schipper, PC, Van der Windt, HJ (1985) Influence of drainage on N-mineralisation and vegetation response in wet meadows. I. *Calthion palustris* stands. Acta Oecologia 6, 403-417.

Harmer R, Peterken G, Kerr G, Poulton P (2001) Vegetation changes during 100 years of development of two secondary woodlands on abandoned arable land. *Biological Conservation* 101, 291-304.

Hayman G, Hasler S, Vincent K, Baker S, Donovan B, Smith M, Davies M, Sutton M, Tang Y S, Dragosits U, Love L, Fowler D, Sansom L, Page H (2001) Operation and Management of the UK Acid Deposition Monitoring Networks: Data Summary for 2000. Report AEAT/ENV/R/0740 Issue 1, AEA Technology, Abingdon, UK.

Heathwaite, AL, Johnes, PJ, Peters, NE (1996) Trends in nutrients and water quality. *Hydrological Processes* 10, 263-293.

Hill, MO, Carey, PD (1997) Prediction of yield in the Rothamsted Park Grass Experiment by Ellenberg indicator values. *J.Veg.Sci.* 8: 579-586.

Hill, MO, Preston, CD, Roy, DB (2003) *PLANTATTR. Attributes of British and Irish plants: Status in Britain, Size, Life history, Geography and Habitats.* CEH Monkswood, Abbotts Ripton.

Hill, MO, Radford, J (1984) Inventory of Permament Vegetation Monitoring Plots. ITE Monkswood, Abbots Ripton.

Hogg, EH, Malmer, N, Wallen, B (1994) Microsite and regional variation in the potential decay of *Sphagnum magellanicum* in south Swedish raised bogs. *Ecography* 17, 50-59.

Hogg, P, Squires, P, Fitter, A (1995) Acidification, nitrogen deposition and rapid vegetational change in a small valley mire in Yorkshire. *Biological Conservation* 71, 143-153.

Holden, J, Adamson, JK (2002) The Moorhouse long-term upland temperature record. *Weather* 57, 119-126.

Hossell, JE, Briggs, B, Hepburn, IR (2000) *Climate Change and UK Nature Conservation: a review of the impact of climate change on UK species and habitat conservation policy*. Department of the Environment, Transport and the Regions, London.

Hruška, J., Moldan, F., Kram, P. (2002). Recovery from acidification in Central Europe – observed and predicted changes of soil and streamwater chemistry in the Lysina catchment, Czech Republic. *Environ. Pollut.*, 120, 261-274.

Jenkins, A., and Cullen, J.M. (2001). An assessment of the potential impact of the Gothenburg Protocol on surface water chemistry using the dynamic MAGIC model at acid sensitive sites in the UK. *Hydrol. Earth Syst. Sci.*, 5, 529-541

Jenkins, A., Ferrier, R.F., Helliwell, R.C. (2001). Modelling nitrogen dynamics at Lochnagar, N.E. Scotland. *Hydrol. Earth Syst. Sci.*, 5, 519-528.

Jones, MLM (2004) The effects of reduced and enhanced nitrogen deposition in acid grassland and the interactive effects of management. In: Emmett, BA, McShane, G (eds.) *Terrestrial Umbrella Final Report 2004*. CEH Bangor.

Kirby, KJ, Smart, SM, Black, HIJ, Bunce, RGH, Corney, PM, Smithers, RJ (2005) *Long-term ecological change in British woodland (1971-2001)*. English Nature Research report 653. EN, Peterborough.

Kopácek, J, Cosby, BJ, Majer, V, Stuchlík, E, Veselý, J (2003) Modelling reversibility of Central European lakes from acidification: Part II – the Tatra Mountains. *Hydrol. Earth Syst. Sci.*, 7, 510-524.

Kopácek, ., Stuchlík, E, Wright, RF (2005). Long-term trends and spatial variability in nitrate leaching from alpine catchment-lake ecosystems in the Tatra Mountains (Slovakia-Poland). *Environ. Pollut.*

Larssen, T., Cosby, B.J., Høgåsen, T. (2004) Uncertainties in predictions of surface water acidity using the MAGIC model. *Water, Air, Soil Pollut: Focus*, 4, 125-137.

Lee, JA, Caporn, SJM (2001) Effects of enhanced atmospheric nitrogen deposition on semi-natural ecosystems. Terrestrial Umbrella progress report to DEFRA.

Leith, ID, Sheppard, LJ, Fowler, D, Cape, JN, Jones, M, Crossley, A, Hargreaves, KJ, Tang, SY, Theobald, M, Sutton, MA (2004) Quantifying dry NH3 deposition to an ombrotrophic bog from an automated NH3 field release system. *Water, Air & Soil Pollution: Focus* 4, 207-218.

Lindborg R, Eriksson, O (2004) Historical landscape connectivity affects present plant species diversity. *Ecology* 85, 1840-1845.

Mainstone, C, Gulson, J, Parr, W, (1994) *Phosphates in freshwater – standards for nature conservation. English Nature Research reports.* No. 73, English Nature, Peterborough.

Majer, V., Cosby, B.J., Kopácek, J., Stuchlík, E., and Veselý, J. (2003). Modelling reversibility of Central European lakes from acidification: Part I – the Bohemian forest. *Hydrol. Earth Syst. Sci.*, 7, 494-509.

Manning, P, Putwain, PD, Webb, NR (2004) Identifying and modelling the determinants of woody plant invasion of lowland heath. *Journal of Ecology* 92, 868-881.

Marrs, RH, Rizand, A, Harrison, AF (1989) The effects of removing sheep grazing on soil chemistry, above-ground nutrient distribution and selected aspects of soil fertiliy in long-term experiments at Moorhouse National Nature Reserve. *Journal of Applied Ecology* 26, 647-661.

Matson, PA, Parton, WJ, Power, AG (1997) Agricultural intensification and ecosystem properties. *Science* 77, 504-508.

NEGTAP (2001) *National Expert Group on Transboundary Air Pollution: Acidification, Eutrophication and Ground Level Ozone in the UK. 1st report*. On behalf of the UK Department of the Environment, Transport and the regions and the devolved administrations. On-line at: www.nbu.ac.uk/negtap/finalreport.htm

Page, T., Beven, K.J., Freer, J., and Jenkins, A. (2003). Investigating the uncertainty in predicting responses to atmospheric deposition using the model of acidification of groundwater in catchments (MAGIC) with a generalised uncertainty estimation (GLUE) framework. *Water, Air, Soil Pollut.*, 142, 71-94.

Petit, S, Maskell, LC, Smart, SM, Wright, SJ (2003) *MIRABEL II: Models for predicting the impact of terrestrial eutrophication, farming intensification and land abandonment on biodiversity in European landscapes.* Final report to the European Environment Agency. CEH Lancaster.

Petit, SP, Griffiths, L, Smart, SM, Stuart, RC, Wright, SM (2004) Effects of area and isolation of woodland patches on herbaceous plant species richness across Great Britain. *Landscape Ecology* **32**, 1-10.

Piessens, K, Honnay, Hermy, M (2005) The role of fragment area and isolation in the conservation of heathland species. *Biological Conservation* 122, 61-69.

Pigott, CD, Walters, SM (1954) On the interpretation of the discontinuous distributions shown by certain British species of open habitats. *Journal of Ecology* **42**, 95-116.

Polis, GA (1999) Why are parts of the world green? Multiple factors control productivity and the distribution of biomass. *Oikos* **86**, 3-15.

Rogora, M., Marchetto, A., and Mosello, R. (2003). Modelling the effects of atmospheric sulphur and nitrogen deposition on selected lakes and streams on the Central Alps (Italy). *Hydrol. Earth Syst. Sci.*, 7, 540-551.

Rötzer, T, Chmielewski, FM (2001) Phenological maps of Europe. *Climate Research*, 18, 249-257.

Roy, D.B., Hill, M.O., Rothery, P., & Bunce, R.G.H. (2000) Ecological indicator values of British species: an application of Gaussian logistic regression. *Annales Botanici Fennici*, 37, 219-226.

Schaffers, AP (2000) Soil, biomass and management of semi-natural vegetation. *Plant Ecology* 158, 229-246.

Schaffers, AP, Sykora, KV (2000) Reliability of Ellenberg indicator values for moisture, nitrogen and soil reaction: a comparison with field measurements. *J.Veg.Sci.* 11: 225-244.

Schouwenberg, E.P.A.G., Houweling, H., Jansen, M.J.W., Kros, J., Mol-Dijkstra, J.P 2001. Uncertainty propagation in model chains: a case study in nature conservancy. Alterra-rapport 001, ISSN 1566-7197. Wageningen, Netherlands.

Silvertown, J (1987) Ecological stability: a test case. *American Naturalist* 130, 807-810.

Silvertown, J, Dodd, M, McConway, K, Potts, J, Crawley, M (1994) Rainfall, biomass variation and community composition in the Park Grass experiment. *Journal of Applied Ecology* 17, 491-504.

Simpson D, Olendrzynski K, Semb A, Støren E, Unger S (1997) Photochemical oxidant modelling in Europe: multi-annual modelling and source-receptor relationships. EMEP/MSC-W Report 3/97, Norwegian Meteorological Institute, Oslo.

Smart, SM, Scott, WA, Whittaker, J, Evans, C, Emmett, B (2003) Development and testing of linked soil-plant model systems. Pollution Umbrella Progress Report. May 2003.

Smart, SM, Scott, WA (2004) Bias in Ellenberg indicator values – problems with detection of the effect of vegetation type. *Journal of Vegetation Science* 15, 843-846.

Smart, SM, Ashmore, MR, Hornung, M, Scott, WA, Fowler, DA, Dragosits, U, Howard, DC, Sutton, MA, Famulari, D (2004) Detecting the signal of atmospheric N deposition in recent national-scale vegetation change across Britain. *Water, Air, Soil Poll. Focus* 4: 269-278.

Smart, SM, Bunce, RGH, Marrs, RH, LeDuc, M, Firbank, LG, Maskell, LC, Scott, WA, Thompson, K, Walker, KJ (2005) Large-scale changes in the abundance of common higher plant species across Britain between 1978, 1990 and 1998 as a consequence of human activity: tests of hypothesised changes in trait representation. *Biological Conservation* 124, 355-371.

Smith, RI, Fowler, D, Sutton, MA, Flechard, C, Coyle, M.(2000) Regional estimation of pollutant gas dry deposition in the UK: model description, sensitivity analyses and outputs. *Atmos. Env.* 34, 3757-3777.

Strengbom, J, Nordin, A, Näsholm, T, Ericson, L (2001) Slow recovery of boreal forest ecosystem following decreased nitrogen input. *Functional Ecology* 15, 451-457.

Suter, GW (1993) Introduction to ecological risk assessment. In: (Suter, GW ed.) *Ecological Risk Assessment*. Pp 3-19. Lewis Publishers, Ann Arbor: Michigan.

Sutton, MA, Cape, JN, Rihm, B, Sheppard, LJ, Smith, RI, Spranger, T, Fowler, D (2003) The importance of accurate background atmospheric deposition estimates in setting critical loads for nitrogen. In B. Achermann, and R. Bobbink (eds.), *Empirical Critical Loads of Nitrogen*, Environmental Documentation 164, Swiss Agency for Environment Forests and Landscape, Berne, pp. 231-257.

Sverdrup, H, Warfvinge, P, Blake, L, Goulding, K (1995) Modeling recent and historic soil data from the Rothamsted-experimental-station, UK using SAFE. *Agric. Eco. Env*. 53, 161-177.

terBraak, CJF, Barendregt, LG (1986) Weighted averaging of species indicator values: its efficiency in environmental calibration. *Math.Biosci.* 78: 57-72.

Terry AC, Ashmore MR, Power SA, Allchin EA, Heil GW 2004. Modelling the impacts of atmospheric nitrogen deposition on Calluna-dominated ecosystems in the UK. J.Appl.Ecol. 41, 897-909.

Thurlow, M., French, D.D. & Bayfield, N.G. (1999) Colonisation by *Pinus sylvestris*. In: *Land- Use Changes in European Mountain Ecosystems.* Eds. Cernusca, A., Tappeiner, U. & Bayfield, N. Blackwell Wissensgafts, Berlin. pp 304-310.

Van der Wal, R., Pearce, I., Brooker, R., Scott, D., Welch D., Woodin, S. (2003) Interplay between nitrogen deposition and grazing causes habitat degradation. *Ecology Letters* **6,** 141-147.

Van Dobben, HF, Schouwenberg, EPAG, Mol, JP, Wieggers, HJJ, Jansen, MJW, Kros, J, de Vries, W (2004) Simulation of critical loads for nitrogen for terrestrial plant communities in the Netherlands. Alterra-rapport 953. Wageningen.

Van Egmond, K, Bresser, T, Bouwman, L, (2002) The European nitrogen case. *Ambio* 31, 72-78.

Van Vuuren, MMI, Aerts, R, Berendse, F, De Visser, W (1992) Nitrogen mineralization in heathland ecosystems dominated by different plant species. *Biogeochemistry* 16, 151-166.

Venterink, HO, Wassen, MJ, Verkroost, AWM (2003) Species richness-productivity patterns differ between N-,P- and K-limited wetlands. *Ecology* **84**, 2191-2199.

Wamelink, GWW, Joosten, V, van Dobben, HF, Berendse, F (2002) Validity of Ellenberg indicator values judged from physico-chemical measurements. *Journal of Vegetation Science* 13, 269-278.

Warren, RG, Johnston, AE (1964) The Park Grass Experiment. Rothamsted Experimental Station Research Report for 1963.

Warren Spring Laboratory (1987) Acid Deposition in the United Kingdom, 1981-85. Warren Spring Laboratory, Stevenage, UK

Wells, TCE et al (1976) Ecological studies on the Porton ranges. *Journal of Ecology* 12, 583-627.

Williams, ED (1978) Botanical composition of the Park Grass plots at Rothamsted. 1856-1976. Rothamsted Experimental Station, Harpenden.

Wilson, JB, Crawley, MJ, Dodd, ME, Silvertown, J (1996) Evidence for constraint on species coexistence in vegetation of the Park Grass experiment. *Vegetatio* 124, 183-190.

Woodland Trust (2000) *Woodland biodiversity: Expanding our horizons*. Woodland Trust, Grantham.

Wright, RF, Cosby, BJ (2003) Future recovery of acidified lakes in southern Norway predicted by the MAGIC model. *Hydrol. Earth Syst. Sci.*, 7, 467-483.