

Payne R, Caporn SJM, Stevens CJ, Carroll JA, Edmondson JL, Gowing DJ & Dise NB (2013) Inferring nitrogen deposition from plant community composition, *Ecological Indicators*, 26, pp. 1-4.

**This is the peer reviewed version of this article**

*NOTICE: this is the author's version of a work that was accepted for publication in Ecological Indicators. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Ecological Indicators, [VOL 26 (2013)] DOI: <http://dx.doi.org/10.1016/j.ecolind.2012.10.013>*

1 *Short Note*

2 **Inferring nitrogen deposition from plant community composition**

3 Richard J. Payne<sup>1\*</sup>, Simon J.M. Caporn<sup>1</sup>, Carly J. Stevens<sup>2,3</sup>, Jacky A. Carroll<sup>1,4</sup>, Jill L. Edmondson<sup>1,5</sup>, David J.  
4 Gowing<sup>3</sup>, Nancy B. Dise<sup>1</sup>

5 <sup>1</sup>School of Science and the Environment, Manchester Metropolitan University, Chester St., Manchester  
6 M1 5GD, United Kingdom.

7 <sup>2</sup>Lancaster Environment Centre, Lancaster University, Bailrigg, Lancaster LA1 4YQ, United Kingdom.

8 <sup>3</sup>Department of Earth, Environment and Ecosystems, Open University, Walton Hall, Milton Keynes MK7  
9 6AA, United Kingdom.

10 <sup>4</sup>Penny Anderson Associates Ltd, Park Lea, 60 Park Road, Buxton, Derbyshire, SK17 6SN, UK.

11 <sup>5</sup>Department of Animal and Plant Sciences, University of Sheffield, Sheffield, S10 2TN, UK.

12 \*Corresponding author. E-mail: [r.payne@mmu.ac.uk](mailto:r.payne@mmu.ac.uk), Tel: 0161 2476198.

13

14

15 ABSTRACT

16 Chronically elevated reactive nitrogen deposition has a severe impact on many ecosystems, and there is  
17 widespread interest in the possibility of using plant community composition to estimate the level of  
18 nitrogen deposition and consequent impacts. Existing approaches use a variety of simple measures  
19 including functional type ratios, Ellenberg numbers, and diversity indices. We propose an alternative  
20 approach in which species-environment models are constructed using national datasets designed to  
21 capture broad-scale deposition patterns. We construct models using partial least squares, weighted  
22 average, and maximum likelihood Gaussian logit regression for two British semi-natural habitats, and  
23 test how well they predict N deposition by cross-validation. We find that performance is good with  $R^2$   
24 values up to 0.7, and suggest that such models may be a useful addition to the bioindication toolbox.

25 KEYWORDS: Nitrogen; Bioindication; Pollution; Ellenberg indices; Transfer functions; Biodiversity

26

27 1. Introduction

28 Deposition of reactive nitrogen (N) largely derived from intensive agricultural and industrial  
29 activity is an increasingly urgent conservation concern. A wealth of evidence links N deposition to loss of  
30 biodiversity, plant community change and degradation of ecosystem services (Bobbink et al. 1998;  
31 2010). In developing countries N deposition is increasing rapidly and constitutes a clear threat to  
32 biodiversity hot-spots and protected areas (Bleeker et al. 2011; Phoenix et al. 2006). While N deposition  
33 is stabilising, or even falling in much of the industrialised world, ecosystems carry a legacy of past  
34 deposition which will not be quickly reversed. National- and international-scale models of N deposition  
35 (Jonson et al. 1998; Smith et al. 2000; Fagerli & Aas 2008) represent the large-scale distribution of  
36 pollution reasonably accurately, but cannot show the local-scale impacts of point sources such as  
37 individual industrial or agricultural units. It is these local-scale impacts which are usually the concern of  
38 practical conservation management, where the interest is often in the impacts of a specific polluter on  
39 an individual designated site. Protected areas are preferentially located in topographically complex  
40 regions (Joppa & Pfaff 2009) where large-scale deposition models perform less effectively (Sutton et al.  
41 2004). Considerable research attention has therefore focussed on the identification of bioindicator  
42 approaches. Bioindicators can be used to identify both the level of pollution and the impacts of pollution  
43 exposure, although this distinction is rarely made explicit. Among the many approaches to bioindication  
44 of nitrogen pollution (Sutton et al. 2004), studies have investigated the potential of plant community-  
45 based bioindicators using the occurrence or abundance of indicator species, or derived indices such as  
46 Ellenberg values, plant functional type ratios and diversity measures (Pitcairn et al. 2002, 2003; Stevens  
47 et al. 2009). Although results have often shown significant relationships with N deposition, the strength  
48 of this relationship is variable.

49           We propose an alternative concept in which relationships between plant species composition  
50 and N deposition are modelled using national vegetation datasets and N deposition models. These  
51 species-environment models can be calibrated using sites better suited to national-scale models (away  
52 from point sources and complex topography), and then applied to predict deposition in situations where  
53 national-scale deposition models are less appropriate. The approach essentially uses vegetation-N  
54 relationships to down-scale national deposition models. In this paper we test the concept that the  
55 vegetation composition of a set of plant communities that fulfil certain criteria can be used to develop  
56 models to predict the cumulative N deposition at other sites comprising a similar vegetation type.

## 57           2. Material and Methods

58           We model the relationship between species abundance and N deposition for two semi-natural  
59 UK vegetation types using three alternative regression approaches. We use UK vegetation datasets of  
60 average species cover for acid grasslands (%), and frequency (occurrence per quadrat) for heather  
61 moorlands. The acid grassland dataset encompasses 64 sites of UK National Vegetation Classification  
62 (NVC: Rodwell 1992) type U4, (*Festuca ovina*-*Agrostis capillaris*-*Galium saxatile* grassland) sampled in  
63 2002-3 (Stevens et al. 2004, 2006). The heathlands dataset (NVC type H12, *Calluna vulgaris*-*Vaccinium*  
64 *myrtillus* heath) combines the data of Edmondson et al. (2010) and Caporn et al. (2009) giving 36 sites  
65 sampled in 2005 and 2006. While the grasslands data includes all plant species the heathlands data  
66 includes bryophytes alone. All studies used five quadrats per site.

67           In both datasets, N deposition is the strongest environmental correlate with community  
68 composition, and appears to be a key agent of vegetation change (Payne et al. 2011; unpublished). We  
69 use modelled cumulative N deposition between 1900 and the year of sampling rather than modelled  
70 current N deposition, reflecting an increasing weight of opinion that cumulative deposition better  
71 represents how ecosystems respond to N pollution (Duprè et al. 2010, De Schrijver et al. 2011). We

72 apply the scaling factors of Fowler et al. (2004) to output from the Centre for Ecology and Hydrology  
73 CBED model (Smith et al. 2000) to calculate cumulative nitrogen deposition since 1900 on a 5x5 km grid  
74 basis. The cumulative N deposition range was 430-2856 kg N ha<sup>-1</sup> (mean: 1742, sd: 720) for the grassland  
75 sites and 459-3067 kg N ha<sup>-1</sup> (mean: 1886, sd: 671) for the heathlands.

76 We test three regression techniques based on two contrasting concepts of how species  
77 abundance may respond to N deposition. The simplest concept assumes that species respond linearly to  
78 N deposition: an increase in N deposition produces an increase or decline in each species. Multiple  
79 linear regression performs poorly for ecological data with large number of species whose abundances  
80 are strongly correlated (e.g. ter Braak & van Dam 1989). We test an alternative approach: partial least  
81 squares (PLS) regression. PLS attempts to extract a minimal number of latent factors or components  
82 from a training set which explain the variability in the environmental data (Geladi & Kowalski 1986). PLS  
83 has been applied in several previous ecological studies (e.g. Charman 1997) and has been used for the  
84 bioindication of nitrogen deposition with metabolic finger-print data (Gidman et al. 2006).

85 The assumption of a linear relationship between species abundance and cumulative N  
86 deposition may be valid if impacts are due to direct toxicity, species are at the edge of their  
87 environmental tolerances or where there is a limited range of deposition values. However, N is an  
88 essential nutrient for plants so an alternative hypothesis is that, for many species, small inputs may be  
89 beneficial but larger additions deleterious, producing a unimodal response. We therefore also test two  
90 regression techniques which assume a unimodal response of species abundance to N deposition.

91 In maximum likelihood (ML) Gaussian logit regression, the relationship between an  
92 environmental variable and abundance of each species is modelled as a Gaussian curve. Maximum  
93 likelihood estimation is used to determine the value of the environmental variable with the highest  
94 probability of being associated with a particular community composition; this estimate is the model

95 prediction (for details of computation see Birks 1995). The method has been shown to perform well with  
96 simulated data and real ecological datasets (ter Braak & Looman 1986), but it is relatively complex and  
97 computationally intensive.

98           An alternative unimodal technique is weighted average (WA) regression, in which it is assumed  
99 that a species will be most abundant in a site with environmental conditions close to the species  
100 optimum; a reasonable approximation of the species optimum is therefore made by calculating the  
101 average environmental values of all the sites in which the species occurs, weighted by the abundance of  
102 the species in those sites. An estimate of the environmental variable for an unknown site is provided by  
103 a weighted average of the optima of all species present. As this procedure serves to compress the  
104 environmental gradient a de-shrinking regression is applied to remove this compression. WA is less  
105 statistically rigorous than ML but is computationally simpler and often has superior performance in  
106 practise (Birks et al. 1990; ter Braak & van Dam 1989).

107           Assumptions of all these models include the independence of samples, lack of confounding  
108 secondary gradients and the presence of a direct (or indirect but linear) relationship between the  
109 species and the environmental variable of interest. These assumptions, and the consequences of their  
110 violation, are discussed in greater depth by Birks (1990, 1995, 1998), Belyea (2007) and ter Braak &  
111 Prentice (1988).

112           We applied all three techniques (PLS, ML & WA) to both of the vegetation datasets. We assessed  
113 model performance statistically by applying the model to the same dataset used to construct the model.  
114 To avoid overly optimistic estimates of performance if the same data are used to both build and test  
115 models we used jack-knife ('leave-one-out') cross-validation in which models are successively  
116 constructed using  $n-1$  samples with the remaining sample serving as a test. Performance statistics used  
117 are the  $R^2$  between observed and predicted values, the root mean squared error of prediction (RMSEP)

118 and the Maximum Bias (cross-validated values are denoted  $R^2_{\text{jack}}$ ,  $\text{RMSEP}_{\text{jack}}$  and  $\text{Max Bias}_{\text{jack}}$ ). These  
119 three measures provide distinct but complementary information about the performance of models:  $R^2$   
120 gives a measure of the overall strength of relationship between observed and predicted values,  $\text{RMSEP}$   
121 gives a measure of average errors along the gradient and Maximum Bias gives a measure of maximum  
122 mean error for any one tenth of the gradient. Models were developed using  $C^2$  (Juggins 2003).

### 123 3. Results

124 All models produced  $\text{RMSEP}$  values below the standard deviation of the nitrogen deposition  
125 data, so all can be considered to have predictive power despite the limited size of the training sets.  
126  $\text{RMSEP}$  values suggest that these models may be able to predict cumulative N deposition with a mean  
127 error as low as  $367 \text{ kg N ha}^{-1}$ , 15% of the range captured by the vegetation dataset.

128 For the acid grasslands data the best-performing model in terms of  $R^2$  and  $\text{RMSEP}$  is ML and in  
129 terms of maximum bias is a 3-component PLS model. For the heathlands data the best-performing  
130 model is a single-component PLS model for  $R^2$  and  $\text{RMSEP}$  and ML for maximum bias. Model  
131 performance with the heathland data is superior to that of the grasslands, despite the smaller dataset  
132 size and use of frequency rather than cover data. This may represent greater importance of other  
133 environmental variables in the grasslands or greater sensitivity of bryophytes in comparison to vascular  
134 plants. The strength of the relationships between observed data (i.e. the output of deposition models)  
135 and predicted values, using the ML model for acid grasslands and the PLS model for heathlands, is  
136 shown in Fig. 1. Regression lines are close to the 1:1 line, and scatter is relatively limited. Both models  
137 tend to over-predict N deposition at the low end of the gradient and under-predict N deposition at the  
138 high end of the gradient.

### 139 4. Discussion



140 Our study shows that species-environment models may be effective at predicting cumulative  
141 nitrogen deposition from vegetation composition. Plant community-based bioindication has several  
142 appealing features. In contrast to other techniques, vegetation surveys require no more equipment than  
143 a quadrat and no more expense than the time of an experienced observer. The skills for plant  
144 identification are fairly widespread and much vegetation data is collected routinely. We show that  
145 relatively simple statistical models perform well, and appear to have potential for bioindication. By  
146 making better use of community data, such models may allow more accurate estimates of deposition.  
147 For instance in the acid grasslands data, the  $R^2$  between N deposition and Ellenberg R was only 0.06  
148 (there was no significant relationship with Ellenberg N: Stevens et al. 2010), compared to an  $R^2_{\text{jack}}$  of 0.64  
149 here. Once such models are developed, the application of the model to new data is straightforward and  
150 software is freely available.

151 A strength of our approach is that the relative performance of models imparts some information  
152 about the underlying relationship between N deposition and community composition. Consequently, it  
153 is best to consider several different approaches in evaluating data from such studies. For instance, if a  
154 maximum likelihood regression fits the data better than a partial least squares regression, it follows that  
155 the relationship is better modelled as a unimodal curve than a straight line. For species abundance  
156 along an environmental gradient this in turn suggests that the full range of the gradient has been  
157 sampled, since species respond by increasing, and then decreasing, in abundance along the gradient.

158 Our models are developed on the basis of national-scale models rather than local deposition  
159 monitoring data for the individual sites. This is partly for pragmatic reasons, as it is clearly impractical to  
160 monitor deposition at a hundred, widely dispersed and often remote sites for a period of many years.  
161 The use of modelled data is justified as the deposition model performs well at capturing the broad-scale  
162 patterns of deposition and the sampling sites have been carefully selected to capture this broad-scale

163 pattern, avoiding local factors which may modify deposition. The regression models developed using  
164 these data-sets should therefore be able to predict N deposition in situations where the national models  
165 cannot (e.g. in the rain shadow of a woodland, an exposed slope, or downwind of a long-term point  
166 source). The scatter in our results probably partly relates to the limitations of the deposition models.  
167 Indeed, the comparison of values predicted by our regression models and the national deposition model  
168 could potentially indicate the contribution of local factors.

169           Our study is a proof-of-concept but further validation and testing of these models is necessary.  
170 It is possible that model performance in practice may be poorer than the cross-validated results we  
171 present here due to the influence of other environmental factors, taxonomic biases, spatial  
172 autocorrelation (Telford & Birks 2005) and surveyor variability. However there are also alternative  
173 modelling approaches which might offer superior performance (including neural nets: Malmgren &  
174 Nordlund 1997, analogue techniques: Simpson 2007, and Bayesian approaches: Toivonen et al. 2001).  
175 Further studies including a greater variety of techniques and testing with independently collected field  
176 data will be required to reveal the full potential of this approach. Studies of local-scale gradients near to  
177 point-sources would be a particularly interesting test of model performance. Although we use  
178 cumulative deposition data, models could equally be constructed using current deposition values for  
179 easier comparison with measured data: correlations between cumulative and modern deposition are  
180 strong so we believe such models will have broadly similar efficacy. Our results indicate that models  
181 relating nitrogen deposition to vegetation in large-scale gradient studies could be a useful addition to  
182 the suite of techniques used for the bioindication of pollution, and deserve further study.

183

184

185 Acknowledgements

186 RJP, NBD and SJMC were supported by the UK Natural Environment Research Council (NERC) through  
187 the European Union FP6 BiodivERsA (ERA-NET) project PEATBOG. Heathland data collection was  
188 supported by DEFRA through the UKREATE programme and HEFCE. Grassland data collection was  
189 supported by the The Open University, Ferguson Trust and NERC. Thanks to two anonymous reviewers  
190 for comments on a previous version of the manuscript.

191

192 References

- 193 Belyea, L.R., 2007. Revealing the emperor's new clothes: niche based palaeoenvironmental  
194 reconstruction in the light of recent ecological theory. *Holocene*. 17, 683-688.
- 195 Birks, H.J.B., 1995. Quantitative palaeoenvironmental reconstructions, in: Maddy, D., Brew, J.S. (Eds.),  
196 Statistical modelling of quaternary science data, Technical Guide 5. Quaternary Research Association,  
197 Cambridge, pp. 116-254.
- 198 Birks, H.J.B., 1998. Numerical tools in palaeolimnology: progress, potentialities and problems. *J.*  
199 *Paleolimnol.* 20, 307-332.
- 200 Birks, H.J.B., Line, J.M., Juggins, S., Stevenson, A.C., Ter Braak, C.J.F., 1990. Diatoms and pH  
201 reconstruction. *Phil. Trans. Roy. Soc. Lond. B.* 327, 263-278.
- 202 Bleeker, A., Hicks, W.K., Dentener, F., Galloway, J., Erisman, J.W., 2011. N deposition as a threat to the  
203 World's protected areas under the Convention on Biological Diversity. *Environ. Pollut.* 159, 2280-2288.
- 204 Bobbink, R., Hornung, M., Roelofs, J.G.M., 1998. The effects of air-borne nitrogen pollutants on species  
205 diversity in natural and semi-natural European vegetation. *J. Ecol.* 86, 717-738.
- 206 Bobbink, R., Hicks, K., Galloway, J., Spranger, T., Alkemade, R., Ashmore, M., Bustamente, M., Cinderby,  
207 S., Davidson, E., Dentener, F., Emmett, B., Erisman, J-W., Fenn, M., Gilliam, F., Nordin, A., Pardo, L., De  
208 Vries, W., 2010. Global assessment of nitrogen deposition effects on terrestrial plant diversity: a  
209 synthesis. *Ecol. Appl.* 20, 30-59.
- 210 Caporn, S.J.M., Edmondson, J., Carroll, J.A., Price, E.A.C., 2009. Long-term impacts of enhanced and  
211 reduced nitrogen deposition on semi-natural vegetation. In: UKREATE (2009) Terrestrial Umbrella:  
212 Effects of eutrophication and acidification on terrestrial exosystems. CEH contract report C03425.
- 213 Charman, D.J., 1997. Modelling hydrological relationships of testate amoebae (Protozoa: Rhizopoda) on  
214 New Zealand peatlands. *J. Roy. Soc. N. Z.* 27, 465-483.

215 De Schrijver, A., De Frenne, P., Ampoorter, E., Van Nevel, L., Demey, A., Wuyts, K., Verheyen, K., 2011.  
216 Cumulative nitrogen input drives species loss in terrestrial ecosystems. *Glob. Ecol. Biogeogr.* 20, 803-  
217 816.

218 Duprè, C., Stevens, C.J., Ranke, T., Bleeker, A., Pepler-Lisbach, C., Gowing, D.J., Dise, N.B., Dorland, E.,  
219 Bobbink, R., Diekmann, M., 2010. Changes in species richness and composition in European acidic  
220 grasslands over the past 70 years: the contribution of cumulative atmospheric nitrogen deposition.  
221 *Glob. Change Biol.* 16, 344-357.

222 Edmondson, J.L., Carroll, J.A., Price, E.A.C., Caporn, S.J.M., 2010. Bio-indicators of nitrogen pollution in  
223 heather moorland. *Sci. Total. Environ.* 6202-6209.

224 Fagerli, H., Aas, W., 2008. Trends of nitrogen in air and precipitation: Model results and observations at  
225 EMEP sites in Europe, 1980–2003. *Environ. Pollut.* 154, 448-461.

226 Fowler, D., O'Donoghue, M., Muller, J.B.A., Smith, R.I., Dragosits, U., Skiba, U., Sutton, M.A.,  
227 Brimblecombe, P., 2004. A chronology of nitrogen deposition in the UK between 1900 and 2000. *Water*  
228 *Air Soil Pollut.* 4, 9-23.

229 Geladi, P., Kowalski, B.R., 1986. Partial least-squares regression: A tutorial. *Anal. Chim. Acta.* 185, 1-17.

230 Gidman, E.A., Stevens, C.J., Goodacre, R., Broadhurst, D., Emmett, B., Gwynn-Jones, D., 2006. Using  
231 metabolic fingerprinting of plants for evaluating nitrogen deposition impacts on the landscape level.  
232 *Glob. Change Biol.* 12, 1460-1465.

233 Jonson, J.E., Bartnicki, J., Olendrzynski, K., Jakobsen, H.A., Berge, E., 1998. EMEP Eulerian model for  
234 atmospheric transport and deposition of nitrogen species over Europe. *Environ. Pollut.* 102, 289-298.

235 Joppa, L.N., Pfaff, A., 2009. High and far: biases in the location of protected areas. *PLoS ONE.* 4, e8273.

236 Juggins, S., 2003. *C<sup>2</sup> user guide*. Software for ecological and palaeoecological data analysis and  
237 visualisation. University of Newcastle, Newcastle Upon Tyne, UK.

238 Malmgren, B.A., Nordlund, U., 1997. Application of artificial neural networks to paleoceanographic data.  
239 *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 136, 359-373.

240 Payne, R.J., Stevens, C.J., Dise, N.B., Gowing, D.J., Pilkington, M., Phoenix, G., Emmet, B.A., Ashmore, M.,  
241 2011. Impacts of atmospheric pollution on the community composition of British acid grasslands.  
242 *Environ. Pollut.* 159, 2602-2608.

243 Phoenix, G.K., Hicks, W.K., Cinderby, S., Kuylenstierna, J.C.I., Stock, W.D., Dentener, F.J., Giller, K.E.,  
244 Austin, A.T., Lefroy, R.D.B., Gimeno, B.S., Ashmore, M.R., Ineson, P., 2006. Atmospheric nitrogen  
245 deposition in world biodiversity hotspots: the need for a greater global perspective in assessing N  
246 deposition impacts. *Glob. Change Biol.* 12, 470–476.

247 Pitcairn, C.E.R., Skiba, U.M., Sutton, M.A., Fowler, D., Munro R., Kennedy, V., 2002. Defining the spatial  
248 impacts of poultry farm ammonia emissions on species composition of adjacent woodland groundflora  
249 using Ellenberg Nitrogen Index, nitrous oxide and nitric oxide emissions and foliar nitrogen as marker  
250 variables. *Environ. Pollut.* 119, 9-21.

251 Pitcairn, C.E.R., Fowler, D., Leith, I.D., Sheppard, L.J., Sutton, M.A., Kennedy, V., Okello, E., 2003.  
252 Bioindicators of enhanced nitrogen deposition. *Environ. Pollut.* 126, 353–361.

253 Rodwell, J.S., 1992. *British Plant Communities : Volume 2 Mires and Heaths.* Cambridge University Press,  
254 Cambridge.

255 Simpson, G.L., 2007. Analogue methods in palaeoecology: Using the Analogue package. *J. Stat. Softw.*  
256 22, 2.

257 Smith, R.I., Fowler, D., Sutton, M.A., Flechard, C., Coyle, M., 2000. Regional estimation of pollutant gas  
258 dry deposition in the UK: model description, sensitivity analyses and outputs. *Atmos. Environ.* 34, 3757-  
259 3777.

260 Stevens, C.J., Dise, N.B., Mountford J.O., Gowing D.J., 2004. Impact of nitrogen deposition on the  
261 species richness of grasslands. *Science.* 303, 1876-1879.

262 Stevens, C.J., Dise, N.B., Gowing D.J., Mountford J.O., 2006. Loss of forb diversity in relation to nitrogen  
263 deposition in the UK: regional trends and potential controls. *Glob. Change Biol.* 12, 1823–1833.

264 Stevens, C.J., Maskell, L.C., Smart, S.M., Caporn, S.J.M., Dise, N.B., Gowing, D.J., 2009. Identifying  
265 indicators of atmospheric nitrogen deposition in acid grasslands. *Biol. Conserv.* 142, 2069-2075.

266 Stevens, C.J., Thompson, K., Grime, J.P., Long, C.J., Gowing, D.J., 2010. Contribution of acidification and  
267 eutrophication to declines in species richness of calcifuge grasslands along a gradient of atmospheric  
268 nitrogen deposition. *Funct. Ecol.* 24, 478–484.

269 Sutton, M.A., Pitcairn, C.E.R., Whitfield, C.P., 2004. Bioindicator and biomonitoring methods for  
270 assessing the effects of atmospheric nitrogen on statutory nature conservation sites, JNCC report 356.  
271 JNCC, Peterborough.

272 Telford, R.J., Birks, H.J.B., 2005. The secret assumption of transfer functions: problems with spatial  
273 autocorrelation in evaluating model performance. *Quaternary Sci. Rev.* 24, 2173-2179.

274 Ter Braak, C.J.F., Looman, C.W.N., 1986. Weighted averaging, logistic regression and the Gaussian  
275 response model. *Plant Ecol.* 65, 3-11.

276 Ter Braak, C.J.F., Prentice, I.C., 1988. A theory of gradient analysis. *Adv. Ecol. Res.* 18, 271-317.

277 Ter Braak, C.J.F., van Dam, H., 1989. Inferring pH from diatoms: a comparison of old and new calibration  
278 methods. *Hydrobiologia.* 178, 209-223.

279 Toivonen, H.T.T., Mannila, H., Korhola, A., Olander, H., 2001. Applying Bayesian statistics to organism-  
280 based environmental reconstruction. *Ecol. Appl.* 11, 618–630.

281  
282  
283  
284  
285  
286  
287  
288  
289  
290

291 Figure and Table

292

293 Table 1. Model performance for grassland and heathland data showing model structure and  $R^2$ , root  
294 mean squared error of prediction (RMSEP) and maximum bias (Max Bias) determined by jack-knife  
295 cross-validation. Model results illustrated in Fig. 1 marked '\*'.

	$R^2_{\text{jack}}$	$\text{RMSEP}_{\text{jack}}$	$\text{MaxBias}_{\text{jack}}$
<b>Grasslands</b>			
PLS (3 component)	0.43	548	623
ML*	0.64	464	1440
WA (inverse deshrinking)	0.46	526	677
<b>Heathlands</b>			
PLS (1 component)*	0.70	367	695
ML	0.62	456	586
WA (inverse deshrinking)	0.61	421	759

296

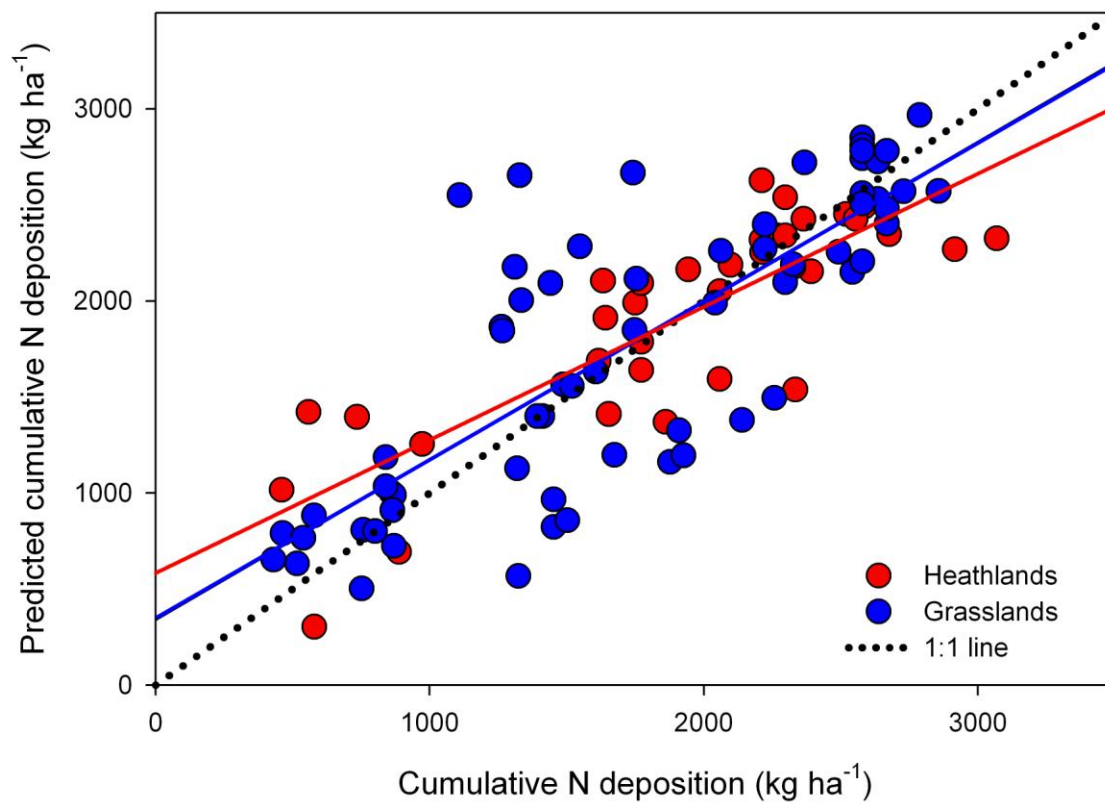
297

298

299



300 Figure 1. Cumulative N deposition (the output of a national deposition model with historical scaling  
301 factors) against vegetation-predicted cumulative N deposition for heathland and grassland data-sets  
302 under jack-knife cross validation.



303

304