Ecological ranges for the pH and NO3 of syntaxa: a new basis for the estimation of critical loads for acid and nitrogen deposition

Journal of Vegetation Science

Wamelink, G.W.W.; Goedhart, P.W.; Malinowska, A.H.; Frissel, J.Y.; Wegman, R.M.A. et al

https://doi.org/10.1111/j.1654-1103.2011.01286.x

Journal of Vegetation Science **22** (2011) 741–749

SPECIAL FEATURE: ECOINFORMATICS

Ecological ranges for the pH and NO_3 of syntaxa: a new basis for the estimation of critical loads for acid and nitrogen deposition

G.W.W. Wamelink, P.W. Goedhart, A.H. Malinowska, J.Y. Frissel, R.J.M. Wegman, P.A. Slim & H.F. van Dobben

Keywords
Distribution model; Eco informatics; Ecological range; Nitrate; pH; Syntaxon; Vegetation database

Received 27 May 2010
Accepted 4 March 2011
Co-ordinating Editor: Joerg Ewald

Wamelink, G.W.W. (corresponding author, wieger.wamelink@wur.nl); **Malinowska, A.H.** (agnieszka.malinowska@wur.nl); **Frissel, J.Y.** (joep.frissel@wur.nl); **Wegman, R.J.M.** (ruut.wegman@wur.nl); **Slim, P.A.** (pieter.slim@wur.nl); **van Dobben, H.F.** (han.vandobben@wur.nl): Alterra, Wageningen UR, Droevendaalsesteeg 3a, PO Box 47, NL–6700 AA, Wageningen, The Netherlands
Goedhart, P.W. (paul.goedhart@wur.nl): Biometris, Wageningen UR, PO Box 100, NL–6700 AC, Wageningen, The Netherlands.

Abstract

Question: Can the abiotic ranges of syntaxonomic units (associations) in terms of pH and nitrate concentration be estimated and then in principle be used to estimate critical loads for acid and nitrogen deposition?

Location: Europe.

Methods: Using splines, abiotic ranges of syntaxonomic units were estimated on the basis of measured soil pH and nitrate concentration and vegetation relevés. Owing to limited calibration data, this yielded responses for only a few syntaxa. In a second attempt, we used a 160 000-relevé training set; the syntaxon of each relevé was known but not their soil pH values and soil nitrate concentrations, so for each relevé, the soil pH and nitrate concentration were estimated by inference from species composition. We again estimated abiotic ranges using the spline method, including the 5th and 95th percentiles as a proxy for the ecological range.

Results: The second (indirect) method yielded ranges for soil pH and for nitrate concentration for many of the associations. The means and percentiles were corrected for regression to the mean.

Conclusions: It is not yet possible to directly estimate ranges for syntaxa for pH and nitrate on a large scale from the data available. However, indirectly estimated soil pH and nitrate concentrations are sufficiently available to derive ranges for many associations. The lower (5th) percentile of the indirectly estimated pH ranges for associations may be used as the starting point for the estimation of critical loads for acid deposition. The 95th percentile should still be regarded as a rather uncertain estimate of the maximum nitrate concentration.

Introduction

The response functions of plant species and vegetation types to abiotic variables such as soil pH or nitrogen availability are widely used in ecology (Diekmann 2003). This paper concentrates on the response functions of syntaxonomic units. It is a sequel to our earlier work on the response functions of plant species (Wamelink et al. 2005). We looked at syntaxa because they are the basis of habitat types under the European Habitats Directive (92/43/EEC, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31992L0043:EN:NOT). To be able to review the environmental status of the habitat types at the Natura 2000 sites information about the abiotic preferences of the habitat types is required.

Abiotic ranges of syntaxonomic unit have two possible applications. The first is to estimate the probability of occurrence of syntaxonomic units, as is done, for example, in the MOVE model (Bakkenes et al. 2002), NTM (Wamelink et al. 2003a) or by Witte et al. (2007). The response functions of these models are based on expert knowledge, such as Ellenberg indicator values (Ellenberg et al. 1991), without providing a direct link to actual soil properties. Responses based on physical measurements, and consequently, estimated in physical units, have the advantage that they can be linked directly to dynamic soil models (Ertsen et al. 1998; Schaffers & Sýkora 2000; Wamelink et al. 2002).

The second application is the estimation of critical loads for acidity and nitrogen. The calculation of critical loads by Van Dobben et al. (2006) was based on abiotic ranges (in terms of pH and nitrogen availability) of vegetation types defined by Ellenberg indicator values. At some stage during the calculation of the critical loads, the indicator values have to be converted to physical units (e.g. $mol\,ha\,yr^{-1}$ deposition). This tricky procedure would be unnecessary if ranges for vegetation types were available in physical units.

Main goal of this research was to develop a method to estimate ranges for pH and nitrate concentration in the soil for syntaxonomic units based on field data. Furthermore, the reliability of the ranges and their applicability for the estimation of critical loads are discussed.

The abiotic responses of syntaxonomic units estimated in this study include optima and ecological ranges. They may therefore have a wider range of applications than the two examples above, for example in nature management and nature development (Bakker 1989), in response modeling (as suggested by Witte et al. 2007), or to replace (at least partly) the expert-based indicator values in models such as MOVE (Bakkenes et al. 2002), PROBE (Witte et al. 2007), NTM (Wamelink et al. 2003a) and ForSAFE (Wallman et al. 2005).

Methods

Training set

Our data set was similar to that used to estimate the response per species (Wamelink et al. 2002, 2005). However, more data were added and some shortcomings in the data that were revealed by a thorough quality check were corrected. The resulting data set comprised 6724 relevés from The Netherlands: 6128 had a measured soil $pH_{(H2O)}$ value and 596 had a measured nitrate concentration. The pH values ranged from 3.0 to 9.5, but most were between 3.5 and 8.0. The nitrate concentration, measured in $CaCl_2$ extraction, ranged from 0.1 to 805.5 $mg\,kg^{-1}$, with most between 0.9 and 125 $mg\,kg^{-1}$. There are several ways of extracting and measuring nitrate concentration in the soil. We chose the $CaCl_2$ extraction as it was one of the standard extraction methods for which we had most of the data available. The soil nitrate contents were expressed for the mass weight of NO_3 (so 62).

Assignment to syntaxonomic units

Each relevé was assigned to a syntaxonomic unit (syntaxon) using ASSOCIA software (Van Tongeren et al. 2008, see also Wamelink et al. 2003a). ASSOCIA identifies vegetation relevés by comparing them with a training set that consists of previously classified relevés. The training set we used was assembled to redefine the Dutch phytosociological system in the Dutch National Vegetation Classification Project (Schaminée et al. 1995a) and is assumed to be representative of the vegetation in The Netherlands. The classification of the training set was done by hand by the authors of the classification project (Schaminée et al. 1995a, 1995b, 1996, 1998; Stortelder et al. 1999). Each unclassified relevé was assigned to the syntaxon that was considered most likely by the ASSOCIA algorithm. Abiotic ranges were estimated for associations and for higher levels (alliance, order and class). The relevés identified above the level of association were used only for the estimations at the higher syntaxonomic levels. We did not review the quality parameters of ASSOCIA given for each relevé. This may have compromised the accuracy (relevés being assigned to an association that fitted less well with ASSOCIA's training set); however, we felt that this was justifiable given the paucity of data. We assumed that more data being used for the estimation of the ranges was more important than leaving out the less accurate assignments by ASSOCIA.

Estimation of abiotic ranges: direct method

The technique used to estimate abiotic ranges per syntaxon is analogous to the one used for individual species (see Wamelink et al. 2005). Here we give only a brief summary, with soil pH as an example. An identical technique was used to estimate the range for nitrate (and other abiotic parameters). The principle of the statistical technique is that the presence/absence of a syntaxon in each pH class can be used to relate the probability of occurrence per syntaxon to the pH. The curve of a syntaxon describes the probability P of that syntaxon occurring as a function of an environmental variable x. The range of x is divided into segments and the probability of occurrence is calculated for each segment. A penalized cubic B-spline (Eilers & Marx 1996) is used to estimate the curve. The number of degrees of freedom used for fitting the spline is important: the more degrees of freedom there are, the more closely the spline follows the data. We applied an automated procedure that determines the number of degrees of freedom used for the spline that controls the number of nodes. The number of equidistant nodes and the choice of the penalty λ have to be determined. The number of nodes is not crucial, provided that there are not too few. We chose 39 equidistant nodes (resulting in 40 intervals) between the minimum and maximum of the observed x values. Curves were fitted for degrees of freedom equal to 1, 2, . . ., 10. The 'best' curve was determined by backward

Doi: 10.1111/j.1654-1103.2011.01286.x

deviance testing: the degrees of freedom were decreased stepwise, from five to four, three and two, and the procedure was halted when the resulting decrease in fit was significant at the 1% level, as judged on the basis of a deviance test (Hastie & Tibshirani 1990). A curve was estimated when the database contained at least 25 measurements for the syntaxon.

A bootstrap sample with replacement, the same size as the number of relevés, was drawn from the relevés and the penalized spline was fitted to the bootstrapped data. Taking 1000 independent bootstrap samples yields 1000 curves. A 95% bootstrap interval is then given by the 2.5% and 97.5% percentiles of the fitted ranges for every abiotic value.

For each syntaxon, the ecological range for pH was estimated as the range between two complementary percentiles (e.g. the 80% interval represents the interval between the 10th and 90th percentiles of the abiotic values of all the relevés where that syntaxon is present). The lower percentile (e.g. the 5th percentile) was used as a basis for the critical load for acid deposition and the higher percentile (e.g. the 95 percentile) was used as a basis for the critical load for nitrogen deposition. The percentiles were also calculated on the basis of the curves, where, fro example, the range between the 10th and 90th percentiles represents 80% of the surface of the curve.

Estimation of curves and ranges: indirect method

The number of syntaxa for which curves could be estimated directly was severely restricted by the data available. In order to increase this number we also used an indirect method to estimate the curve per syntaxon for soil pH. It was based on a 160 000-relevé training data set (Hennekens & Schaminée 2001) that is assumed to be representative for the vegetation of The Netherlands (Hennekens & Schaminée 2001). For each relevé, the average pH value was inferred as the mean of the pH and nitrate optima of its constituent species (following the method described earlier by Wamelink et al. 2005). A value was calculated when at least five species with a known pH response were present. This criterion was not met by about 30 000 relevés for soil pH and by 57 735 for nitrate, but the remaining data were sufficient to estimate a curve for each syntaxonomic level, using the spline method described above. The minimum number of findings per syntaxon was set at 25, as had been done for the direct method.

Regression to the mean

The indirect method used to estimate abiotic ranges for syntaxa has the disadvantage that the x-axis contracts as a result of averaging, as was shown by Roy et al. (2000) when they were estimating Ellenberg indicator values for the UK. The contraction can be severe and may influence the predictive power (Table 1). It is clear that every time an average of an environmental variable is estimated the respective axis is shortened; thus, the lowest pH optimum rises from 3.0 in the raw (field) data to 4.6 in inferred data, and for the highest pH optimum the decrease is even larger (two pH units). The contraction is an artifact inherent in estimating optima by reciprocal averaging and requires rescaling.

To overcome the problem we used a regression equation to correct for regression to the mean for the abiotic ranges for soil pH and nitrate concentration. The regression is based on the relationship between measured and calculated pH and nitrate concentration in the original data set with measurements. We applied indirect estimation of pH and nitrate concentration to those relevés with measured values available, and regressed measured values (y) on estimated values (x). For pH we obtained the regression equation $y = -7.19 + 2.24^{*}x$ ($n = 5{,}680$, $R^2 = 0.51$). For nitrate, the correction for regression to the mean did not improve the estimated response curves and percentiles at this stage (i.e. the percentage explained variance did not improve and the regression coefficient was not closer to the ideal one). Therefore, the nitrate concentration was not corrected for regression to the mean.

Results

Direct estimation of curves and ranges

ASSOCIA assigned relevés mostly to hierarchical levels above the association. Only a few syntaxa were assigned to

Table 1. Minimum (min) and maximum (max) values for pH and nitrate in the field data, the minimum and maximum averages of the species responses used for the indirectly estimated association range, the minimum and maximum estimates for the relevés used to estimate the indirectly estimated associations range, and the minimum and maximum estimated averages of the indirectly estimated associations range.

	pH		NO_3 (mg kg^{-1})	
	Min	Max	Min	Max
Field data	3.0	9.5	0.1	805.5
Average species response	3.8	7.9	4.6	128.9
Estimated averages for the relevés	4.1	7.6	7.0	82.7
Average indirectly estimated associations range	4.6	7.4	11.0	59.9

associations or sub-associations. As a result, response curves to soil pH could be estimated directly for only 23 associations and 13 sub-associations (Table 2). The estimated curves are not very reliable (see example in Fig. 1; all results can be found at http://www.abiotic.wur.nl). The higher hierarchical syntaxon levels are better represented, but here, too, many syntaxa are missing and the abiotic ranges are very uncertain. That is why we also used the indirect approach to estimate abiotic range of associations.

Because the results were so poor for soil pH and even fewer measurements were available for nitrate than for soil pH, we did not use the direct method to estimate the curves and abiotic ranges for nitrate.

Indirect estimation

The indirect method of estimating syntaxon ranges for pH yielded far more syntaxa for which a curve and abiotic range could be estimated (Table 2). There are 372 associations (including fragmentary basal and derivative communities) defined for The Netherlands. Curves for soil pH could be estimated for 224 and curves for nitrate for 156 of them. For the 270 sub-associations defined for The Netherlands, 235 response curves could be estimated for soil pH and 199 responses for nitrate.

The uncertainty surrounding the estimated curves is small (for examples see Fig. 2 and http://www.abiotic.wur.nl). Table 3 gives an example of the most important results for the curves at the association level for soil pH after correction for regression to the mean. Ecological ranges (based on the 5th and 95th percentiles, Fig. 3) vary greatly, from narrow (< 0.5 pH units) to a large part of the pH range (> 2.5 pH units). Most of the ecological ranges comprise between 0.5 and 1.5 pH units (Fig. 3a). Most of the ranges for nitrate are between 10 and 30 mg kg^{-1} NO_3 (Fig. 3b).

Discussion

Reliable abiotic ranges of syntaxa are essential for the estimation of critical loads for habitat types (Van Dobben et al. 2006). However, as exact estimation of the lower end of the environmental range is crucial in defining critical loads (Wamelink & Van Dobben 2003b), we believe that our direct method is not sufficiently accurate.

Table 2. Number of directly and indirectly estimated abiotic ranges for syntaxa.

Syntaxa	Number of Syntaxa	Directly Estimated Range (pH)	Indirectly Estimated Range (pH)	Indirectly Estimated Range (NO_3)
Association*	372	23	224	156
Sub-association	270	13	235	199

*Including frame and derivate communities.

Moreover, even though our dataset was large, it did not allow to estimated abiotic ranges for soil pH and nitrate concentration for a sufficient number of associations. Although it is expected that the dataset will grow, it is unlikely to become large enough within the next 5 yr. The indirect method seemed to be more powerful in number and accuracy of estimated ranges for soil pH and nitrate content. However, it does have some shortcomings: (1) nitrate concentration as a proxy for nitrogen availability, (2) the effect of regression to the mean and the correction for it, (3) testing/validating the responses of the associations. These are discussed below.

Nitrate as a proxy for nutrient availability

We chose nitrate concentration in the soil extracted in a $CaCl_2$ solution for the estimation of the association responses. There are several methods to measure the nitrate concentration of the soil, and extraction in water is one of the more important. The nitrate concentration in the soil is the result of numerous processes in the soil and changes during the year and during the day (Stevens et al. 2011). However, especially in areas with long ongoing deposition we believe that the nitrate concentration is a good indicator of the nutrient status of the soil. There are other proxies available for the nitrogen availability for plant species, such as ammonium concentration, carbon/nitrogen (C/N) ratio, total nitrogen content of the soil and mineralization. All could be used and we have sufficient data to estimate the responses for all, except for mineralization. Unfortunately, none of them is known to be the ultimate predictor of species and vegetation type occurrence. From the work of Rowe et al. (2006, 2011) it can be concluded that C/N ratio (or rather N/C) and nitrate concentration may give the best proxy. For The Netherlands however, C/N does not give much information any more because of the ongoing nitrogen deposition and the resulting drop in C/N. That leaves us with nitrate concentration as an indicator of nitrogen availability, which can also be modeled by SMART2. Ammonium is not considered at this point, because the results were less reliable. More research on other determinants of nitrogen-availability such as ammonium and C/N ratio is needed to improve the prediction of species composition.

Problems of indirect estimation

For the calculation of the pH and nitrate concentration of the relevés with unknown abiotic variables we used an imputation method. Tichý et al. (2010) tested imputation for pH and conductivity and found that it was a powerful and reliable method to estimate unknown abiotic conditions. However, the method here applied involves regression to the mean, which brings additional uncertainty in

Journal of Vegetation Science
Doi: 10.1111/j.1654-1103.2011.01286.x

Table 3. An example of the available characteristics in the data files accompanying the syntaxon abiotic ranges for soil pH. Here the data for associations based on the indirectly estimated curves are shown after correction for regression to the mean (http://www.abiotic.wur.nl). A similar file is available for nitrate. In the table C stands for curve, indicating that the results are based on the response curve. C_025 gives the 2.5 percentile of the data and so on. [1]According to Schaminée et al. (1995a, b).

Latin	SpecNr	Count	Cmean	Cmedian	Csd	C_025	C_050	C_100	C_250	C_500	C_750	C_900	C_950	C_975
Cirsio dissecti-Molinietum	156	1120	5.15	5.18	0.15	4.40	4.56	4.72	4.96	5.18	5.38	5.55	5.65	5.74
Crepido-Juncetum acutiflori	157	595	5.51	5.53	0.15	4.79	4.94	5.09	5.32	5.53	5.73	5.89	5.99	6.08
Rhinantho-Orchietum morionis	158	113	6.09	6.10	0.16	5.39	5.52	5.67	5.88	6.10	6.31	6.50	6.62	6.74
Lychnido-Hypericetum tetrapteri	159	500	5.48	5.49	0.18	4.66	4.81	4.98	5.23	5.49	5.74	5.96	6.11	6.24
Ranunculo-Senecionetum aquatici	160	950	5.88	5.86	0.14	5.32	5.42	5.52	5.68	5.86	6.05	6.25	6.39	6.53
Scirpetum sylvatici	161	239	5.91	5.92	0.18	5.10	5.26	5.42	5.67	5.92	6.15	6.38	6.53	6.67
Angelico-Cirsietum oleracei	162	444	5.85	5.91	0.14	5.09	5.25	5.43	5.70	5.91	6.06	6.18	6.25	6.31
Fritillario-Alopecuretum pratensis	163	149	6.20	6.21	0.08	5.88	5.94	6.01	6.11	6.21	6.30	6.40	6.46	6.51
Sanguisorbo-Silaetum	164	135	6.45	6.47	0.11	5.96	6.06	6.18	6.33	6.47	6.60	6.71	6.79	6.86
Arrhenatheretum elatioris	165	3049	6.94	6.94	0.12	6.43	6.52	6.62	6.77	6.94	7.11	7.27	7.37	7.47
Lolio-Cynosuretum	166	3046	6.41	6.37	0.13	5.94	6.01	6.09	6.21	6.37	6.57	6.84	6.98	7.09
Galio-Trifolietum	167	119	7.33	7.33	0.16	6.63	6.76	6.89	7.10	7.33	7.55	7.78	7.93	8.07
RG Holcus lanatus-Lolium perenne-[Molinio-Arrhenatheretea]	168	3287	6.29	6.28	0.11	5.79	5.89	5.99	6.14	6.28	6.44	6.60	6.71	6.82
RG *Holcus lanatus-Lychnis flos-cuculi*-[*Molinietalia*]	169	877	6.01	6.02	0.10	5.54	5.64	5.75	5.89	6.02	6.15	6.26	6.34	6.41
RG *Festuca rubra-Lotus uliginosus*-[*Molinietalia*]	170	371	5.92	5.93	0.16	5.18	5.34	5.49	5.72	5.93	6.13	6.32	6.44	6.57
RG *Juncus effusus*-[*Molinietalia/ Lolio-Potentillion*]	171	361	5.67	5.72	0.23	4.51	4.77	5.03	5.40	5.72	5.99	6.23	6.40	6.56
RG *Carex panicea-Succisa pratensis*-[*Junco-Molinion*]	172	309	5.36	5.41	0.15	4.53	4.72	4.93	5.20	5.41	5.59	5.73	5.81	5.89
RG *Carex disticha*-[*Calthion palustris*]	173	465	6.10	6.10	0.14	5.48	5.61	5.73	5.92	6.10	6.28	6.46	6.58	6.72
RG *Alopecurus pratensis-Lychnis flos-cuculi*-[*Alopecurion/ Molinietalia*]	175	493	6.34	6.33	0.08	6.00	6.06	6.13	6.23	6.33	6.44	6.55	6.62	6.68
RG *Alopecurus pratensis-Elymus repens*-[*Arrhenatheretalia*]	176	1300	6.60	6.60	0.10	6.16	6.24	6.33	6.46	6.60	6.74	6.87	6.95	7.03
RG *Alopecurus pratensis-Hordeum secalinum*-[*Alopecurion/Cynosurion*]	177	79	6.78	6.63	0.28	6.26	6.33	6.40	6.51	6.63	6.77	7.01	8.54	8.99
RG *Anthriscus sylvestris*-[*Arrhenatheretalia*]	178	266	6.67	6.68	0.14	6.05	6.18	6.30	6.49	6.68	6.86	7.05	7.17	7.28
RG *Hieracium lactucella*-[*Cynosurion/Plantagini-Festucion*]	179	25	6.16	6.16	0.18	5.43	5.58	5.73	5.95	6.16	6.38	6.58	6.72	6.85

the absolute values of the 5th and 95th percentiles. The effect of regression to the mean was corrected for pH by applying a regression equation based on individual plant species rather than syntaxa. Too few relevés were available to enable abiotic ranges for syntaxa to be estimated directly. Moreover, too few relevés were available for the estimation of a regression based on syntaxa to correct for regression to the mean. The correction would have been based on only a few syntaxa, and would therefore not be suitable for widespread application. However, a regression between the directly estimated optima of the associations on the optima of the indirectly estimated associations is possible. The regression illustrates that contraction of the x-axis occurs. When there is no regression to the mean, the regression coefficient is expected to be 1.00. There was a highly significant relationship between the estimated optima ($R^2 = 0.85$; see also Fig. 4). However, the regression coefficient was significantly different from 1.00 (0.54, $P < 0.0001$). The relation also indicates that there is a strong relation between the direct and indirect method. It cannot be regarded as a proper validation, but at the moment it is the only practical way to test the method.

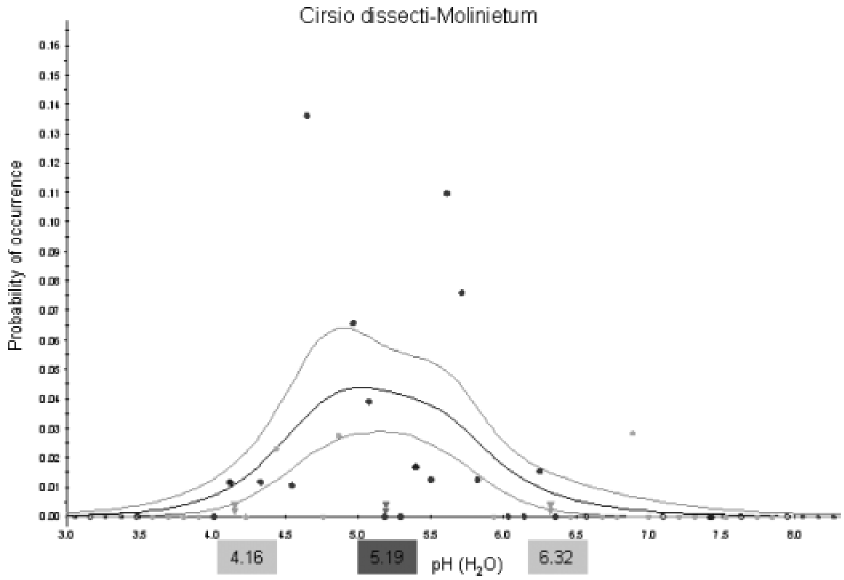

Fig. 1. Response curve of the association *Cirsio dissecti-Molinietum* for soil pH, directly estimated from field measurements ($n = 45$). The difference in color of the dots indicates the number of observations in each pH class. To this end, the pH axes is split up into 40 parts. For each part the probability of occurrence is calculated for the association, based on the findings for pH for the association and the total number of findings for the class in the database (so including all other vegetation types). The curve (using the spline method) is given including the 95% uncertainty interval of the curve. On the *x*-axis the optimum (i.e. the 50th percentile of the curve; highlighted in dark grey) and the 5th and 95th percentiles (highlighted in light grey) of the curve are shown.

Correction by using the species responses

A problem is the low reliability of the directly estimated ranges. To address this we used the regression for soil pH for species to correct for regression to the mean (see Wamelink et al. 2005). This indirect method had proved to be a good method for improving the estimation of the soil pH for the international forest dataset used by Wamelink et al. (2005). In this study, applying the indirect estimated responses for species and correcting for regression to the mean reduced the average deviation between measured and predicted soil pH from 0.53 to 0.34 pH units. The effect of correction depends on the pH value to be corrected, but can be as high as one pH unit, especially for the extreme pH values (i.e. below pH 4.5). For nitrate, the correction for regression to the mean did not improve the predicted responses. Although this might be because there is no effect of contraction of the *x*-axis, we think that the absence of contraction of the *x*-axis is highly unlikely. Better ways to correct for regression to the mean remain to be investigated, not only for nitrate but for all abiotic variables.

Testing and validating the association responses

The responses of plant species have been extensively tested (Wamelink et al. 2005) for soil pH. It is difficult, though not impossible, to test the ecological responses for syntaxa. The method we applied could be repeated on a second, large data set for the indirect method; the estimated ecological ranges should be similar to those we obtained. However, no such database is currently available and it remains a challenge to build such a database on a European scale. From the tests performed and the validation for species responses it is clear that the ranges for soil pH are less uncertain than for nitrate concentration. We expect that the ecological ranges given here for syntaxa will be similarly reliable as for species. The results for soil pH are expected to be applicable but those for nitrate concentration will probably still be rather poor. Further measurements of nitrate are being made in the field, which will help to improve the outcome for nitrate.

Application to critical loads

We expect that the 5th percentile for pH could be used as a starting point to estimate critical loads, but only after correction for regression to the mean. We assume that this percentile gives the minimum value for which a syntaxon may occur in a well-developed state. The deposition of acid compounds is related to soil pH. Lower pH values (e.g. caused by acid deposition) would in time lead to species loss, invasion by tolerant species and, ultimately, loss of the habitat type. Van Dobben et al. (2006) used this knowledge to estimate critical loads for acid deposition by applying the dynamic soil model SMART2 (Mol Dijkstra et al. 2009). This model simulates

Journal of Vegetation Science
Doi: 10.1111/j.1654-1103.2011.01286.x

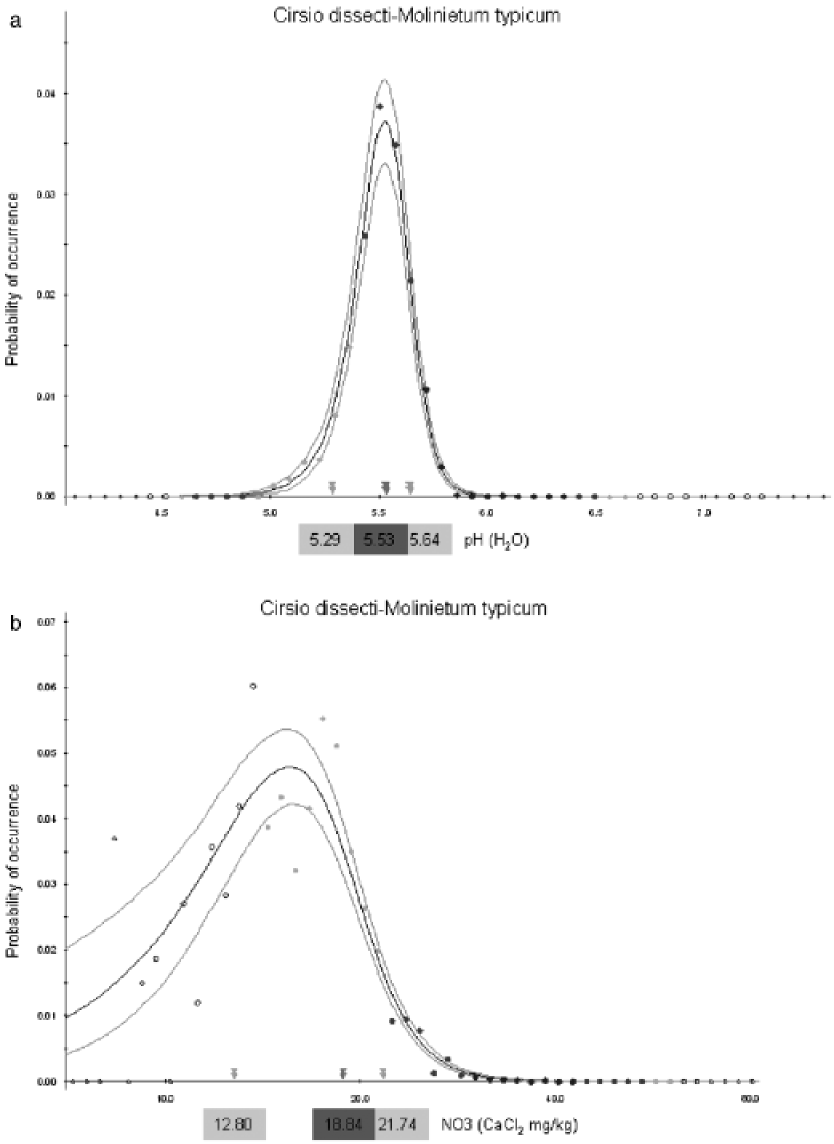

Fig. 2. Response curves of the *Cirsio dissecti-Molinietum* associations for soil pH (**a**) and nitrate concentration (**b**), estimated using an indirect method (i.e. based on a calibration data set of relevés without soil measurements). The size of the dots indicates the number of observations in each class. The curve (using the spline method) is given, including the 95% uncertainty interval of the curve. On the *x*-axis the mean of the curve (highlighted in dark grey) and the 5th and 95th percentiles (highlighted in light grey) of the curve are shown. The response curve is not corrected for regression to the mean.

soil processes including a complete nitrogen cycle and calculates soil pH. Cycles of base cations (Ca, Mg and K) are also included, thus accounting for their effects on soil pH. Normally, this model is used to evaluate the effects of nitrogen and acid deposition on, for example, nitrogen availability and soil pH. By running this model backwards based on a minimum Ellenberg indicator value for acidity for a syntaxon the maximum acid deposition (i.e. the critical load) was calculated. Defining the critical load as the 5th percentile of the syntaxon range avoids the awkward step of translating Ellenberg indicators for acidity into soil pH (Wamelink et al. 2002; Wamelink & Van Dobben 2003b).

The same procedure may be applied for nitrate concentration in the soil, which is also modeled by the SMART2 model. The nitrate concentration in the model is a result of a full nitrogen cycle in the soil, including mineralization, nitrification and denitrification and nitrogen deposition. As for pH, a threshold value of nitrate concentration for occurrence of a syntaxon can be established, as described in this paper. If the nitrate concentration rises above the threshold value we expect the loss of diagnostic

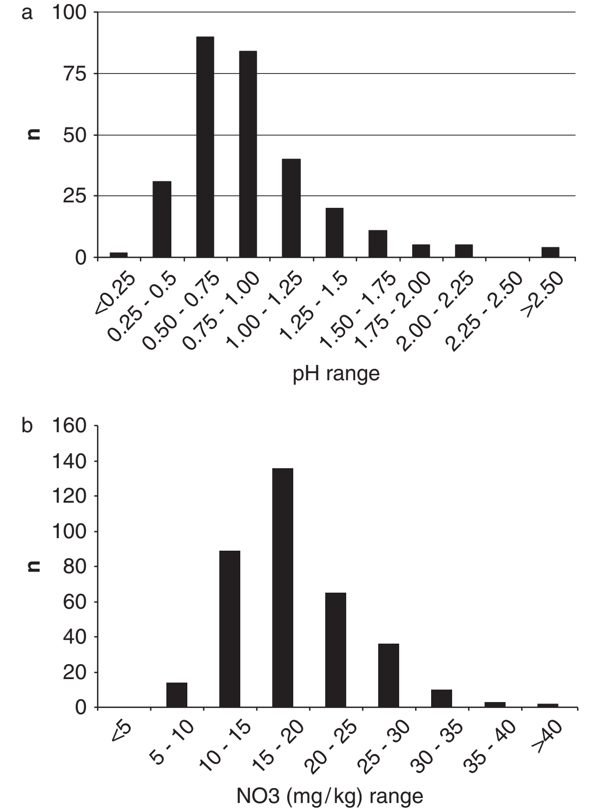

Fig. 3. Histograms of the ecological ranges for pH (**a**) and nitrate ($CaCl_2$ extraction, (**b**) of the associations (including sub-associations and frame and derivate communities). The ecological range is defined as the difference between the 5th and 95th percentiles.

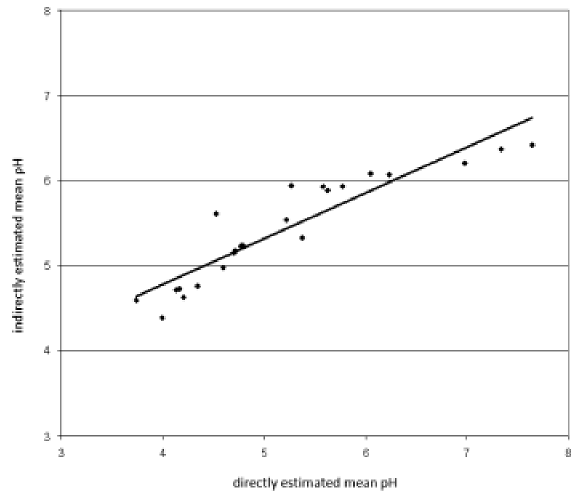

Fig. 4. Relation between the direct and indirect estimated mean pH for associations ($n = 23$).

species (e.g. as a result of nitrogen deposition). By using the threshold value as a starting point for the backwards model run with SMART2 it becomes possible to calculate the maximum nitrogen deposition, the critical load, for which no effects on the vegetation may be expected. This can be done in a site-specific manner and scaled up to regional or the national level, resulting in national critical load values (see also De Vries et al. 2007).

Our estimated ranges for syntaxa are based on the species composition of the syntaxa, but this composition may change in future in response to climate change. The changes will probably come about mainly because of differences in the dispersal capacity of the different species and because barriers in the landscape limit the dispersal of certain species. These expected changes will probably also change the ecological ranges of the syntaxa given here, eventually making them less or even no longer useful. This underlines the need for continued monitoring of species composition and abiotic conditions across all vegetation types and the long-term storage and management of these data in databases.

Acknowledgements

We are grateful for the comments of Wim Ozinga on the ecological range discussion, and the comments of two anonymous reviewers. The response curves are based on the data obtained from numerous scientists: see the literature list on http://www.abiotic.wur.nl. This research was funded by the Department of Science and Knowledge Transfer (DWK) of the Ministry of Agriculture Nature and Food Quality, project number BO-02-004-004. The authors' editor was Joy Burrough.

References

Bakkenes, M.D., de Zwart, D. & Alkemade, J.R.M. 2002. *MOVE nationaal Model voor de Vegetatie versie 3.2 Achtergronden en analyse van modelvarianten*. Report 408657006. RIVM, Bilthoven.

Bakker, J.P. 1989. *Nature management by grazing and cutting, on the ecological significance of grazing and cutting regimes applied to restore former species-rich grassland communities in the Netherlands*. Kluwer Academic Publishers, Dordrecht, NL.

De Vries, W., Kros, H., Reinds, G.J., Wamelink, G.W.W., Mol, J., Van Dobben, H., Bobbink, R., Emmett, B., Smart, S., Evans, C., Schlutow, A., Kraft, P., Belyazid, S., Sverdrup, H., Van Hinsberg, A., Posch, M. & Hettelingh, J.-P. 2007. *Developments in deriving critical limits and modeling critical loads of nitrogen for terrestrial ecosystems in Europe*. Alterra Report 1382. Alterra, Wageningen, NL.

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

Ellenberg, H., Weber, H.E., Düll, R., Wirth, V., Werner, W. & Pauliβen, D. 1991. Zeigerwerte von Pflanzen in Mitteleuropa. *Scripta Geobotanica* 18: 9–166.

Journal of Vegetation Science
Doi: 10.1111/j.1654-1103.2011.01286.x

Eilers, H.C. & Marx, B.D. 1996. Flexible smoothing with B-splines and penalties. *Statistical Science* 11: 89–121.

Ertsen, A.C.D., Alkemade, J.R.M. & Wassen, M.J. 1998. Calibrating Ellenberg indicator values for moisture, acidity, nutrient availability and salinity in the Netherlands. *Plant Ecology* 135: 113–124.

Hastie, T. & Tibshirani, R. 1990. *Generalized additive models*. Chapman & Hall, London, UK.

Hennekens, S.M. & Schaminée, J.H.J. 2001. TURBOVEG, a comprehensive data base management system for vegetation data. *Journal of Vegetation Science* 12: 589–591.

Mol Dijkstra, J.P., Reinds, G.J., Kros, J., Berg, B. & De Vries, W. 2009. Modelling soil carbon sequestration of intensively monitored forest plots in Europe by three different approaches. *Forest Ecology and Management* 258: 1780–1793.

Rowe, E.C., Evans, C.D., Emmett, B.A, Reynolds, B., Helliwell, R.C., Coull, M.C. & Curtis, C.J. 2006. Vegetation type affects the relationship between soil carbon to nitrogen ratio and nitrogen leaching. *Water, Air and soil Pollution* 177: 335–347.

Rowe, E.C., Emmett, B.A., Smart, S.M. & Frogbrook, Z.L. 2011. A new net mineralizable nitrogen assay improves predictions of floristic composition. *Journal of Vegetation Science* 22: 251–261.

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, A.P. & Sýkora, K.V. 2000. Reliability of Ellenberg indicator values for moisture, nitrogen and soil reaction, comparison with field measurements. *Journal of Vegetation Science* 11: 225–244.

Schaminée, J.H.J., Stortelder, A.H.F. & Westhoff, V. 1995a. *De vegetatie van Nederland*. Vol. 2, Opulus Press, Upsala, SE.

Schaminée, J.H.J., Weeda, E.J. & Westhoff, V. 1995b. *De Vegetatie van Nederland. Deel 2. Plantengemeenschappen van wateren, moerassen en natte heiden*. Opulus Press, Uppsala, SE.

Schaminée, J.H.J., Stortelder, A.H.F. & Weeda, E.J. 1996. *De Vegetatie van Nederland. Deel 3. Plantengemeenschappen van graslanden, zomen en droge heiden*. Opulus Press, Uppsala, SE.

Schaminée, J.H.J., Weeda, E.J. & Westhoff, V. 1998. *De Vegetatie van Nederland. Deel 4. Plantengemeenschappen van de kust en van binnenlandse pioniermilieus*. Opulus Press, Uppsala, SE.

Stevens, C.J.S., Manning, P., van den Berg, L.J.L., de Graaf, M.C.C., Wamelink, G.W.W., Boxman, A.W., Bleeker, A., Vergeer, P., Arroniz-Crespo, M., Limpens, J., Lamers, L.P.M., Bobbink, R. & Dorland, E. 2011. Ecosystem responses to reduced and oxidised nitrogen inputs in European terrestrial habitats. *Environmental Pollution* 159: 665–676.

Stortelder, A.F.H., Schaminée, J.H.J. & Hommel, P.W.F.M. 1999. *De Vegetatie van Nederland. Deel 5. Plantengemeenschappen van ruigten, struwelen en bossen*. Opulus Press, Uppsala, SE.

Tichý, L., Hájek, M. & Zelený, D. 2010. Imputation of environmental variables for vegetation plots based on compositional similarity. *Journal of Vegetation Science* 21: 88–95.

Van Dobben, H.F., Van Hinsberg, A., Schouwenberg, E.P.A.G., Jansen, M., Mol-Dijkstra, J.P., Wieggers, H.J.J., Kros, J. & De Vries, W. 2006. Simulation of critical loads for nitrogen for terrestrial plant communities in the Netherlands. *Ecosystems* 9: 32–45.

Van Tongeren, O., Gremmen, N.J.M. & Hennekens, S.M. 2008. Assignment of relevés to pre-defined classes by supervised clustering of plant communities using a new composite index. *Journal of Vegetation Science* 19: 525–536.

Wallman, P., Svensson, M.G.E., Sverdrup, H. & Belyazid, S. 2005. ForSAFE – an integrated process-oriented forest model for long-term sustainability assessments. *Forest Ecology and Management* 207: 19–36.

Wamelink, G.W.W. & Van Dobben, H.F. 2003b. Uncertainty of critical loads based on the Ellenberg indicator value for acidity. *Basic and Applied Ecology* 4: 515–523.

Wamelink, G.W.W., Joosten, V., van Dobben, H.F. & Berendse, F. 2002. Validity of Ellenberg indicator values judged from physico-chemical field measurements. *Journal of Vegetation Science* 13: 269–278.

Wamelink, G.W.W., Ter Braak, C.J.F. & Van Dobben, H.F. 2003a. Changes in large-scale patterns of plant biodiversity predicted from environmental economic scenarios. *Landscape Ecology* 18: 513–527.

Wamelink, G.W.W, Goedhart, P.W., Van Dobben, H.F. & Berendse, F. 2005. Plant species as predictors of soil pH: replacing expert judgment by measurements. *Journal of Vegetation Science* 16: 461–470.

Witte, J.-P.M., Wójcik, R., Torfs, P.J.J.F., De Haan, M.W.H. & Hennekens, S. 2007. Bayesian classification of vegetation types with Gaussian mixture density fitting to indicator values. *Journal of Vegetation Science* 18: 605–612.