Published in New Phytologist 195, issue 2, 408-418, 2012
which should be used for any reference to this work

Glasshouse vs field experiments: do they yield ecologically similar results for assessing N impacts on peat mosses?

J. Limpens[1]*, G. Granath[2]*, R. Aerts[3], M. M. P. D. Heijmans[1], L. J. Sheppard[4], L. Bragazza[5,6,7], B. L. Williams[8], H. Rydin[2], J. Bubier[9], T. Moore[10], L. Rochefort[11], E. A. D. Mitchell[12], A. Buttler[6,7,13], L. J. L. van den Berg[14], U. Gunnarsson[2], A. -J. Francez[15], R. Gerdol[5], M. Thormann[16], P. Grosvernier[17], M. M. Wiedermann[18], M. B. Nilsson[18], M. R. Hoosbeek[19], S. Bayley[20], J. -F. Nordbakken[21], M. P. C. P. Paulissen[22], S. Hotes[23], A. Breeuwer[1], M. Ilomets[24], H. B. M. Tomassen[25], I. Leith[4] and B. Xu[26]

[1]Nature Conservation and Plant Ecology Group, Wageningen University, Droevendaalsesteeg 3a, 6708 PB Wageningen, the Netherlands; [2]Department of Plant Ecology and Evolution, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-752 36 Uppsala, Sweden; [3]Faculty of Earth and Life Sciences, Systems Ecology, Free University of Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, the Netherlands; [4]Centre for Ecology and Hydrology (CEH), Edinburgh Bush Estate Penicuik, EH26 0QB, Scotland; [5]Department of Biology and Evolution, University of Ferrara, Corso Ercole I d'Este 32, I-44121 Ferrara, Italy; [6]WSL Swiss Federal Institute for Forest, Snow and Landscape Research, Wetlands Research Group, Site Lausanne, Station 2, CH-1015 Lausanne, Switzerland; [7]École Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Laboratory of Ecological Systems (ECOS), Batiment GR, Station 2, CH-1015 Lausanne, Switzerland; [8]Macaulay Land Use Research Institute, Aberdeen, UK; [9]Mount Holyoke College, Environmental Studies Department, Clapp Laboratory, 50 College Street, South Hadley, Massachusetts 01075, USA; [10]Department of Geography, McGill University, 805 Sherbrooke St. W. Montreal, Quebec, Canada H3A 2K6; [11]Department of Plant Sciences, Université Laval, 2425 rue de l'Agriculture, Quebec, QC, Canada G1V 0A6; [12]Laboratory of Soil Biology, University of Neuchâtel, Rue Emile-Argand 11, CH-2009 Neuchâtel, Switzerland; [13]Laboratory of Chrono-Environnement, UMR 6249 CNRS – INRA, Université de Franche-Comté, Besançon, France; [14]Aquatic Ecology and Environmental Biology, Radboud University Nijmegen, 6525 AJ Nijmegen, the Netherlands; [15]UMR 6553 ECOBIO & FR90 CAREN, Rennes University, CNRS, Campus de Beaulieu, 263 avenue du Général Leclerc, 35042 Rennes Cedex, France; [16]Aquilon Environmental Consulting Ltd., 3111 Spence Wynd SW, Edmonton, Alberta, Canada T6X 0H7; [17]LIN'eco, Case postale 80, 2732 Reconvilier, Switzerland; [18]Soil Science, Biogeochemistry, Department of Forest Ecology & Management, Swedish University of Agricultural Sciences, S-901 83 Umeå, Sweden; [19]Earth System Science – Climate Change, Department of Environmental Sciences, Wageningen University, P.O. Box 47, 6700AA Wageningen, the Netherlands; [20]Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2E9; [21]The Norwegian Forest and Landscape Institute, Postbox 115, 1431 Ås, Norway; [22]Alterra, Team Ecological Modelling and Monitoring, Wageningen University and Research Centre, PO Box 47, 6700 AA Wageningen, the Netherlands; [23]Department of Animal Ecology, Justus-Liebig-University, Heinrich-Buff-Ring 26-32 (IFZ), D-35392 Giessen, Germany; [24]Department of Landscape Ecology, Institute of Ecology, Tallinn University, Uus-Sadama 5, EE-10120 Tallinn, Estonia; [25]B-Ware Research Centre, Nijmegen, the Netherlands; [26]Southern Illinois University Carbondale 1125 Lincoln Drive, Carbondale, IL 62901, USA

Author for correspondence: *Juul Limpens. Tel: +31 317483173. Email: juul.limpens@wur.nl*

Summary

- Peat bogs have accumulated more atmospheric carbon (C) than any other terrestrial ecosystem today. Most of this C is associated with peat moss (*Sphagnum*) litter. Atmospheric nitrogen (N) deposition can decrease *Sphagnum* production, compromising the C sequestration capacity of peat bogs. The mechanisms underlying the reduced production are uncertain, necessitating multifactorial experiments.
- We investigated whether glasshouse experiments are reliable proxies for field experiments for assessing interactions between N deposition and environment as controls on *Sphagnum* N concentration and production. We performed a meta-analysis over 115 glasshouse experiments and 107 field experiments.
- We found that glasshouse and field experiments gave similar qualitative and quantitative estimates of changes in *Sphagnum* N concentration in response to N application. However, glasshouse-based estimates of changes in production – even qualitative assessments – diverged from field experiments owing to a stronger N effect on production response in absence of vascular plants in the glasshouse, and a weaker N effect on production response in presence of vascular plants compared to field experiments.
- Thus, although we need glasshouse experiments to study how interacting environmental factors affect the response of *Sphagnum* to increased N deposition, we need field experiments to properly quantify these effects.

Keywords

Carbon (C), climate, experiments, meta-analysis, mires and peatlands, nitrogen deposition, productivity, *Sphagnum*.

*These authors contributed equally to this work.

Introduction

Peat bogs significantly affect the earth's atmosphere (Frolking & Roulet, 2007) by sequestering CO_2 in the form of partly decomposed organic material (peat), mostly formed by peat mosses (*Sphagnum* species; Rydin & Jeglum, 2006). Peat bogs are nutrient-poor wetlands that commonly rely on atmospheric inputs as their sole source of external nutrients, making the vegetation composition sensitive to increases in nitrogen (N) deposition (Bobbink *et al.*, 2010). The vegetation comprises mainly ericaceous dwarf shrubs and cyperaceous graminoids of low productivity, rooting in a matrix of living and dead *Sphagnum*. Once competition from *Sphagnum* is reduced, or the nutrient limitation is mitigated, vascular plants may gain a competitive advantage and, being taller, outcompete the mosses due to their better competitive ability for light (Hautier *et al.*, 2009). Shifts from a moss- to a vascular-plant dominated state can depress carbon (C) sequestration rates (Juutinen *et al.*, 2010) and even mobilize the N and C stored in the underlying peat by stimulating microbial activity (Freeman *et al.*, 2004). The peatland store is equivalent to between 34 and 46% of atmospheric CO_2 (IPCC, 2007), and hence the ensuing C release into the atmosphere and the local environment may be substantial (Limpens *et al.*, 2008).

Northern hemisphere peat bogs are experiencing a hitherto unprecedented combination of stresses such as increases in N deposition (Galloway *et al.*, 2008), temperature, and drought frequency. Interactions between these factors may initiate a regime shift from a moss- to a vascular-plant dominated plant community, turning peat bogs from C sinks to C sources (Dise, 2009). Interactions between N deposition and other stresses are likely, as shown in a recent statistical meta-analysis of field experiments (Limpens *et al.*, 2011). The analysis suggested that several variables affect the response of *Sphagnum* species to N deposition including N influx, presence of vascular plant, phosphorus availability, temperature, water table depth and species. The benefit of a meta-analysis of many experiments is that it reveals the influence of factors that are uncontrolled in a field situation. Experimental underpinning is needed, however, to understand the mechanisms underlying these statistical correlations (Dise & Phoenix, 2011).

The design of glasshouse experiments, although artificial, can remove some of the variables allowing a closer examination of the influence of factors, alone or in combination with others, on the response to N application. One of the difficulties of this approach, however, is that the small spatial scale in glasshouse experiments may cause a different response to N (Irvine *et al.*, 2004). Such scale effects make translating the results from glasshouse to field difficult. This is particularly the case if processes operating at a larger scale (e.g. competitive interactions) start to dominate the processes at a small scale (e.g. physiological changes; Englund & Cooper, 2003). Comparing the outcomes of N-application experiments between glasshouse and field experiments, across different experimental spatial scales (from shoot to plot) and organizational complexity (with and without vascular plants), allows us to explore if the responses of *Sphagnum* to N application depend on the scale of experimentation and whether they are driven mainly by direct or indirect effects. We hypothesized that:

- the direction of the response (qualitative response) of *Sphagnum* production and tissue N concentration to N and its sensitivity to environmental modifiers (N influx, presence of vascular plants, phosphorus availability, temperature, water table depth, species) is similar for both glasshouse and field experiments;
- the strength of the response (quantitative response) of *Sphagnum* production is lower in glasshouse experiments because of superior growing conditions;
- the quantitative response of *Sphagnum* production is determined by the organizational complexity (such as the presence of vascular plants), rather than by the scale (area) of the experimental unit.

Description

Data acquisition

Nitrogen application studies conducted on *Sphagnum*-dominated vegetation were located by searching the Web of Science and Google Scholar using the key words *Sphagnum,* nitrogen, peatlands, bogs, mires, and fertilis(z)ation in 2009. In addition, we used our contacts within the peatland researcher community to acquire additional data and information about the experiments. Authors were asked to share raw data, enabling a uniform calculation of treatment effects. For those cases where raw data were irretrievable (nine studies; Supporting Information Table S1), we extracted the data from published manuscripts. In addition, we included unpublished N concentration, production or stem growth data related to published experiments of co-authors of this paper. For multiple-year experiments we only used data from the last year. We selected all studies where the control was subject to the same temperature regime as the fertilization treatments. These selection criteria left us with 44 studies covering 14 countries across North America and Eurasia. The resulting dataset (Table S1) contains information of 222 experiments (115 glasshouse; 107 field) reporting on *Sphagnum* N concentration (119 glasshouse; 87 field) and 206 reporting on *Sphagnum* production.

From the 44 studies, we included three response variables and 14 explanatory variables. Response variables were: *Sphagnum* N concentration in the upper 3 cm of the shoot, *Sphagnum* production and stem growth. Explanatory variables were: experiment type; number of seasons that N was applied; presence of vascular plants; background N deposition; N application; N application rate; P application; precipitation; temperature; position above the water table (microhabitat); scale of the experimental unit; *Sphagnum* species; N dose concentration; and form and frequency in which N was applied. Mean and standard deviation (SD) of the response variables were calculated or extracted for all N-treatments per study, treating different species subject to the same treatment, or the same species subject to different treatments, as separate experiments (Gurevitch & Hedges, 1999). In 14 studies, stem growth had been measured instead of

production. To maximize the number of experiments included in our analysis, we converted stem growth to production by using the relationship between stem growth and production derived from a subset of glasshouse and field experiments where both variables had been reported. Excluding the stem growth studies from our meta-analysis did not affect the conclusions but did slightly widen the 95% credible intervals.

Sphagnum responses to N application were standardized before the analyses, expressing the effect relative to the control. For each experiment, the effect size was calculated as the natural logarithm ($\log_e$) of the response ratio (*rr*) of *Sphagnum* production (PROD) or N concentration (N). The *rr* is defined as the mean of the experimental group (E) divided by the mean of the control group (C). The $\log_e$ of the response ratio was used to linearize the metric and achieve a more normal distribution (Hedges *et al.*, 1999). A negative $\text{Nlog}_e rr$ (hereafter written as Nlnrr) indicates that applying N reduced the N concentration, whereas a positive Nlnrr indicates that applying N increased the N concentration relative to the control. Assuming treatment and control are independent, the variance (Var) of $\log_e rr$ is Var ($\log_e E - \log_e C$) and is calculated as $(SD_E^2/n_E E^2) + (SD_C^2/n_C C^2)$ (Hedges *et al.*, 1999), where n is the sample size. To better evaluate the relative importance of the explanatory variables they were standardized by subtracting the mean and dividing by two times SD before the analyses (Gelman, 2008). Regression coefficients are then standardized in our models and are directly comparable with each other, including untransformed binary variables (Gelman, 2008).

Statistical model building

We tested our hypotheses using a meta-regression approach (Gurevitch & Mengersen, 2010) using the standardized response of *Sphagnum* production (PRODlnrr) and *Sphagnum* N concentration (Nlnrr) to N application as the response variables. Explanatory variables (Table 1) were assessed for collinearity to ensure that modelled variables could be estimated independently. Additionally, we investigated the distribution of the variables to ensure relatively even distribution of data within the N application range, and within categories. When building submodels for testing variables that could not be tested in the main models due to overfitting (Thompson & Higgins, 2002; Lajeunesse, 2010), or too uneven distribution of data within categories (Harrell, 2001), we included N application rate and, if possible, other variables that were significant in the main models (Tables 2, 3; model b). The influence of covariates associated with the N treatments that could potentially bias our results (Table 1) was tested in submodels. For more information on explanatory variables and model choices see Supporting Information Notes S1.

In order to test if glasshouse and field experiments differ in their overall response to N application, we fitted a model with experimental type as the only variable (Tables 2, 3; model a). To explore if glasshouse experiments are reliable proxies for field experiments for the assessment of the interactions between N deposition and environment, we used the models for the response of production (PRODlnrr) and *Sphagnum* N concentration (Nlnrr) from Limpens *et al.* (2011), adding interactions between the explanatory variables and experimental type (Tables 2, 3; model b). To test interactions with species we ran separate models for those species for which we had a substantial amount of data covering a broad range of our explanatory variables for both glasshouse and field experiments (Table S2). To investigate if there is a linear trend between *Sphagnum* response and scale we ran a separate model (Tables 2, 3; model c).

Table 1 Description and ranges of predictor variables and covariates used in our models (Tables 2, 3, S2)

Variables In models (Tables 2, 3)	Description	Range	
		Glasshouse	Field
Experiment type	Glasshouse (G) or field experiment (F)	–	–
Presence of vascular plants	Present vs removed by clipping aboveground shoots (G & F)	–	–
Number of seasons	6 months = 1 growing season (G), number of growing seasons (F)	0.3–3	1–6
Background N deposition	Wet N deposition rate (gN m^{-2} yr^{-1}) at collection site (G) or experimental site (F)	0.12–2.8	0.09-2
N application rate	N applied (gN m^{-2} yr^{-1}) to experimental unit (G & F)	0.25–11.26	0.46–15
P application	P applied vs no P applied (G & F)	Yes–no	Yes–no
Temperature	Mean glasshouse temperature (G) or mean July temperature (F) in °C	15–25	10.6–20
Microhabitat	Species characteristic of (G) wet lawn vs dry hummock microhabitat (F)	Lawn-hummock	Lawn-hummock
Scale	Experimental unit to which treatments where applied (G & F)	Shoot, pot	Pot, plot, field
Variables & Covariates			
In submodels (Table S2)			
Sphagnum species	Dominant species in the experimental unit (G & F)	–	–
N dose concentration	N concentration (g l^{-1}) of the fertilizer solution (G & F)	0.0002–2	0.002–2
N form	Form (NH_4^+, NO_3^-, NH_4NO_3) in which N was applied (G & F)	All three	All three
N frequency	Frequency: low < 6 times yr^{-1} ≤ medium ≤ 0.3 times wk^{-1} (high) of N application (G & F)	High	Low, medium, high

For a more extensive description of the variables see Table S1 and/or Notes S1. Note that some variables (Temperature, Microhabitat, Background N deposition and Number of seasons) were not fully comparable between experiment types.

Table 2 Results of three (a–c) Hierarchical Bayes Linear Models (HBLM) on the response of *Sphagnum* N concentration (NlnrrNlnrr, $\log_e$ (treatment/control)) with standardized model coefficients (St. coef.)

	N	St. coef.	Upper	Lower	*P*	τ^2	Rob.
Model a: overall differences							
Exp = 206							
Intercept (field)	87	0.31	0.41	0.21	<0.01	0.07	
Experiment type (glasshouse)	119	−0.01	0.15	−0.17	0.88		
Model b: interactions with environmental variables (Fig. 2a)							
Exp = 206							
Intercept (field, lawn species, no P)		0.41	0.50	0.33	<0.01	0.02	++
Experiment type (glasshouse)		−0.08	0.03	−0.18	0.13		
No. of seasons		0.18	0.26	0.10	<0.01		++
Background N deposition		−0.22	−0.09	−0.35	<0.01		++
Nitrogen (N) application rate		0.35	0.47	0.24	<0.01		++
Nitrogen (N) application rate^2		−0.37	−0.23	−0.52	<0.01		++
Phosphorus (P) application		−0.05	0.04	−0.14	0.27		
Temperature		0.11	0.28	−0.05	0.18		
Microhabitat (hummock)		−0.18	−0.08	−0.28	<0.01		++
Experiment type (glasshouse) × no. of seasons		0.02	0.22	−0.17	0.82		
Experiment type (glasshouse) × Background N deposition		0.28	0.42	0.13	<0.01		++
Experiment type (glasshouse) × Nitrogen (N) application rate		0.01	0.14	−0.13	0.92		
Experiment type (glasshouse) × Nitrogen (N) application rate^2		0.26	0.43	0.09	<0.01		++
Experiment type (glasshouse) × Phosphorus (P) application		−0.17	−0.04	−0.30	<0.01		+
Experiment type (glasshouse) × Temperature		−0.08	0.11	−0.28	0.40		
Experiment type (glasshouse) × Microhabitat (hummock)		0.36	0.50	0.22	<0.01		++
Model c: effect of scale (Fig. 2c)							
Exp = 206							
Intercept (Scale - shoot)	63	0.28	0.33	0.23	<0.01	0.02	++
Scale - pot	71	0.03	0.10	−0.04	0.37		
Scale - plot	72	0.08	0.15	0.01	0.02		+
Nitrogen (N) application rate		0.36	0.43	0.30	<0.01		++
Nitrogen (N) application rate^2		−0.19	−0.11	−0.27	<0.01		++

Note that Number of seasons replaces Presence of vascular plants, which affected PRODlnrr (Table 2) but not Nlnrr. Exp, total number of experiments; *N*, number of experiments within category. Negative coefficients indicate that an increase in the predictor variable depresses the response of *Sphagnum* to N addition. Categorical levels are compared to the intercept, which for model b is set to field experiments, in the lawn microhabitat, without phosphorus application. Interaction terms show the differences in intercept, or slope for continuous variables, between groups within a variable. For example, the slope for N application rate in the field is the term Nitrogen (N) application rate, while the slope for glasshouse experiments is acquired by adding the terms Nitrogen (N) application rate and Experiment type (glasshouse) × Nitrogen (N) application rate. Upper and lower = 95% credible intervals; *P,* two-sided *P*-value derived from the posterior probability that the regression coefficient is zero; τ^2, hierarchical variance (i.e. heterogeneity); Rob., robustness results assessed by a leave-one-out jack-knifing approach. 0, sensitive; +, robust; ++, very robust.

Sampling dependence and hierarchical Bayes linear model (HBLM)

The linear mixed-model of a meta-analysis can be expressed as: $\boldsymbol{y} = X\beta + \delta + \varepsilon$ ($\boldsymbol{y}$, vector of effect size estimates ($\log_e$ *rr*); *X*, design matrix with the explanatory variables; β, vector of parameters (including an intercept term and the effects of the explanatory variables); δ, identity matrix with τ^2 (i.e. the residual heterogeneity) along the diagonal). Residual heterogeneity is the variability among experimental outcomes that is not explained by the explanatory variables in the model. ε is the sampling variance–covariance matrix that is assumed to be known and has experiment-specific variances on the diagonal. As indicated earlier, we calculated the response ratio for different N application rates by dividing the mean of the experimental group by the mean of the control group. Because many studies used multiple N application rates and only one control treatment, the same samples were used as control for more than one experimental group, creating sampling dependence in the responses that had to be dealt with (Gurevitch & Hedges, 1999). We controlled for the sampling dependence by including co-variances between related experiments in ε as off-diagonal blocks (Hedges *et al.*, 2010) in a hierarchical Bayes linear model (HBLM; Kulmatiski *et al.*, 2008; Stevens & Taylor, 2009). For further details on calculations of co-variances between experiments and comparison of HBLM with the method of moments estimation see Limpens *et al.* (2011). Controlling for sampling dependence is crucial in a study such as ours as our dataset had many multiple-treatment studies (Table S1). We performed the analyses in R (R Development Core Team, 2011), employing the *metahdep* package (Stevens & Nicholas, 2009). In a HBLM set-up *metahdep* uses a noninformative normal prior on β|(τ), and a log-logistic prior on τ. See Stevens & Taylor (2009) for computational details. Regression coefficients are given with 95% credible intervals, meaning that β lies within the interval with a posterior probability of 0.95. Credible intervals were calculated as two

Table 3 Results of three (a–c) Hierarchical Bayes Linear Models (HBLM) on the response of *Sphagnum* production (PRODlnrr, $\log_e$ (treatment/control)) with standardized model coefficients (St. coef.)

	N	St. coef.	Upper	Lower	*P*	τ^2	Rob.
Model a: overall differences							
Exp = 222							
Intercept (field)	107	−0.07	0.01	−0.14	0.08	0.04	
Experiment type (glasshouse)	115	0.02	0.11	−0.08	0.74		
Model b: interactions with environmental variables (Fig. 3a)							
Exp = 222							
Intercept (field, no vascular plants, lawn species, no P)		0.25	0.40	0.09	<0.01	0.02	++
Experiment type (glasshouse)		−0.32	−0.15	−0.49	<0.01		++
Presence vascular plants		−0.52	−0.35	−0.68	<0.01		++
Background N deposition		−0.43	−0.23	−0.62	<0.01		++
Nitrogen (N) application rate		−0.15	−0.05	−0.25	<0.01		++
Phosphorus (P) application		0.27	0.43	0.12	<0.01		+
Temperature		−0.01	0.18	−0.20	0.92		
Microhabitat (hummock)		−0.16	0.00	−0.32	0.04		0
Experiment type (glasshouse) × Presence vascular plants		0.55	0.84	0.27	<0.01		++
Experiment type (glasshouse) × Background N deposition		0.41	0.63	0.20	<0.01		++
Experiment type (glasshouse) × Nitrogen (N) application rate		0.03	0.16	−0.10	0.64		
Experiment type (glasshouse) × Phosphorus (P) application		−0.24	−0.05	−0.43	0.01		0
Experiment type (glasshouse) × Temperature		0.10	0.32	−0.12	0.37		
Experiment type (glasshouse) × Microhabitat (hummock)		0.19	0.47	−0.08	0.15		
Microhabitat (hummock) × Temperature		−0.26	−0.02	−0.50	0.03		0
Model c: effect of scale (Fig. 3c)							
Intercept (Scale - plot - no vascular plants)		0.32	0.49	0.15	<0.01	0.03	+
Scale - field	5	−0.16	0.10	−0.43	0.21		
Scale - pot	17	−0.30	−0.11	−0.49	<0.01		0
Scale - shoot	62	−0.38	−0.20	−0.56	<0.01		+
Presence vascular plants (plot)	62	−0.59	−0.39	−0.78	<0.01		++
Scale - pot × Presence vascular plants (pot)	67	0.45	0.73	0.17	<0.01		+
Nitrogen (N) application rate	9	−0.12	−0.06	−0.19	<0.01		++

Exp, total number of experiments; *N*, number of experiments within category. Negative coefficients indicate that an increase in the predictor variable depresses the response of *Sphagnum* to adding N. Categorical levels are compared to the intercept, which for model b is set to field experiments, without vascular plants, in the lawn microhabitat without phosphorus application. For interpretation interaction terms see legend Table 2. Upper and lower = 95% credible intervals, *P*, two-sided *P*-value derived from the posterior probability that the regression coefficient is zero; τ^2, hierarchical variance (i.e. heterogeneity); Rob., robustness results assessed by a leave-one-out jack-knifing approach. 0, sensitive; +, robust; ++, very robust.

times the posterior SD of the coefficients. Two-sided *P*-values for the coefficients are also provided for a more familiar interpretation of significant effects.

Model checking

In order to examine sample size bias, we plotted effect size vs variance and number of replicates (Supporting Information Fig. S1a–d). In addition, we checked potential bias of our results owing to differences in within-study variance of the two experiment types. Barring one outlier (Fig. S1c), mean within-study variance and SD were almost identical between field and glasshouse experiments. For general model checking we used residual analyses. To check for robustness in our significant coefficient estimates (i.e. $P < 0.05$), a leave-one-out jack-knifing approach was used where we re-ran models after removing each observation, or study, sequentially from the initial dataset (Efron & Tibshirani, 1993). The result was considered weakly robust (+; Tables 2, 3) when the coefficient remained significant for all runs where one experiment was left out (i.e. $N_{experiment\ -\ 1}$ runs), and robust (++; Tables 2, 3) when the coefficient remained significant after a whole study was left out (i.e. $N_{study\ -\ 1}$ runs). The effect of extreme data points was investigated by comparing the model results before and after removing these extremes. For the jack-knifing approach we used a statistical definition of study. We collected data of 44, mainly published, sources. Some of these sources reported experiments that had been performed at multiple sites. Others reported multiple experiments that had been run under different conditions in the glasshouse. At this substudy level we had not 44 but 56 studies that we considered statistically independent and that were used for the jack-knifing approach.

Exploring N treatment artefacts

A major criticism of N-application experiments is that the level of the experimental treatment (in this case, N-application rate) is often substantially higher than the background N deposition received by even highly impacted ecosystems, leading to treatment artefacts (Pearce & van der Wal, 2008). Limpens *et al.* (2011) showed that the increase in *Sphagnum* tissue N

concentration with N influx across field experiments was indistinguishable from that measured across a natural gradient of N deposition (Bragazza *et al.*, 2005), suggesting that N application side effects, if present at all, were negligible. To test if the same applied to glasshouse experiments, we compared the relationships between nonstandardized *Sphagnum* N concentration and N influx between glasshouse and field experiments. For field experiments N influx (gN m^2 yr^{-1}) was defined as the sum of background N deposition and N application rate; for glasshouse experiments N influx equalled N application rate. The analysis was performed by fitting a generalized least square regression (GLS) to account for the within-study correlation, using the R package *nlme* (Pinheiro *et al.*, 2011). For more information see Notes S1.

Results

Sphagnum N concentration response

The relationship between *Sphagnum* N concentration and N influx in glasshouse experiments was almost identical (both in trend and scatter) to that in field experiments for a wide range in N influx (Fig. 1). This result was sustained when experiments at high N influx (10 g N m^{-2} yr^{-1}) were omitted. The choice of a nonlinear model (N concentration = intercept + $\beta \times \log_e$(N influx) was based on previous work (Bragazza *et al.*, 2005; Limpens *et al.*, 2011) to better model N saturation. Fitting a linear model (i.e. not transforming N influx with $\log_e$) did not yield any difference between glasshouse and field experiments either (P = 0.37), was less successful in predicting N concentration at low N influx, and generally displayed a worse fit (Akaike Information Criterion $(AIC)_{linear}$ = 648, $AIC_{non\text{-}linear}$ = 645). Our results suggest that artefacts associated with glasshouse experiments are negligible compared to the variability inherent to field experiments.

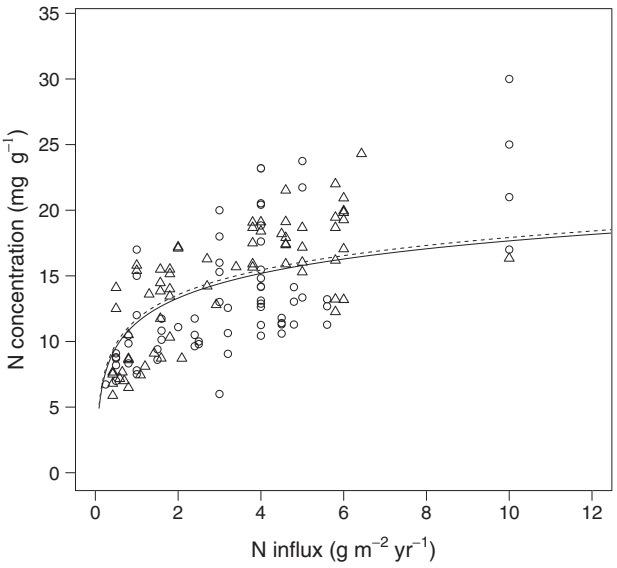

Fig. 1 Relationships between *Sphagnum* nitrogen (N) concentration in the top 0–3 cm of the shoot and N influx (the sum of background wet deposition and applied N) for glasshouse (circles) and field (triangles) experiments. The fitted model is: N concentration = μ + 2.7 ± 0.3 × $\log_e$ (N influx) (P < 0.01), and is illustrated by lines for glasshouse experiments (dashed line, μ = 11.7, n = 67) and field experiments (solid line, μ = 11.4, n = 69). The nonsignificant experiment type × N influx interaction (P = 0.42) was removed for simplicity but the main effect of experiment type was kept for illustrative purpose, that is, the difference between the lines, even though their effects were tiny and not significant.

Nitrogen application increased *Sphagnum* N concentration relative to the control (Nlnrr), irrespective of experiment type (Table 2, model a). Moreover, the variation in effect size of Nlnrr did not significantly differ between experiment types (SD glasshouse = 0.28, SD field = 0.25; P = 0.14). Examining the interaction effects between experiment type and environmental variables, such as microhabitat, background N deposition, N application rate and P application, did reveal a number of differences between glasshouse and field experiments, however (Table 2, model b). The largest interaction effect was with microhabitat (Fig. 2a). In glasshouse experiments, *Sphagnum* species characteristic of dry microsites (hummocks) increased more in N concentration than *Sphagnum* species characteristic of wet microsites (lawns), even at similar water table depth (not shown). In the field experiments, hummock species showed an opposite response, increasing less in N concentration than lawn species per unit of applied N. The response of Nlnrr to background N deposition also differed between experiment types. In glasshouse experiments, Nlnrr was not affected by the background N deposition the mosses had been subject to before the start of the experiment. For field experiments, however, *Sphagnum* subject to a high background N deposition increased less in N concentration per unit of applied N, than *Sphagnum* subject to low background N deposition. The interaction with P application was strongly driven by a few experiments and was, therefore, not robust.

Analyses based on individual species (Table S2) showed that *S. fuscum* was likely the driver behind the microhabitat × experiment type interaction (Table 2, model b). This species had the lowest response value in the field but the highest in the glasshouse, while *S. fallax* and *S. magellanicum* had very similar values for both experiment types (Fig. 2b, Table S2).

The scale of the experimental unit hardly affected the N concentration response. Although the data suggested a smaller increase in N concentration in shoot than in plot experiments (Fig. 2c; Table 2, model c), the effect was small and depended on a few experiments and/or studies. Note that none of the experiments using field as experimental unit reported N concentration.

The effect of N dose on Nlnrr did not differ between experiment types (Table S2). Analysis of N form was unreliable as a consequence of unequal distribution of sample sizes (Table S2).

Sphagnum production response

Without controlling for differences in environment, species or design (Table 3, model a), the only difference between the outcome of glasshouse and field experiments was the smaller variation of the production response (i.e. the variation of effect sizes for PRODlnrr) in glasshouse experiments (SD glasshouse = 0.28, SD field = 0.60; P < 0.01). After including environmental variables in the model (Table 3, model b), heterogeneity in the outcomes among the experiments was better accounted for and

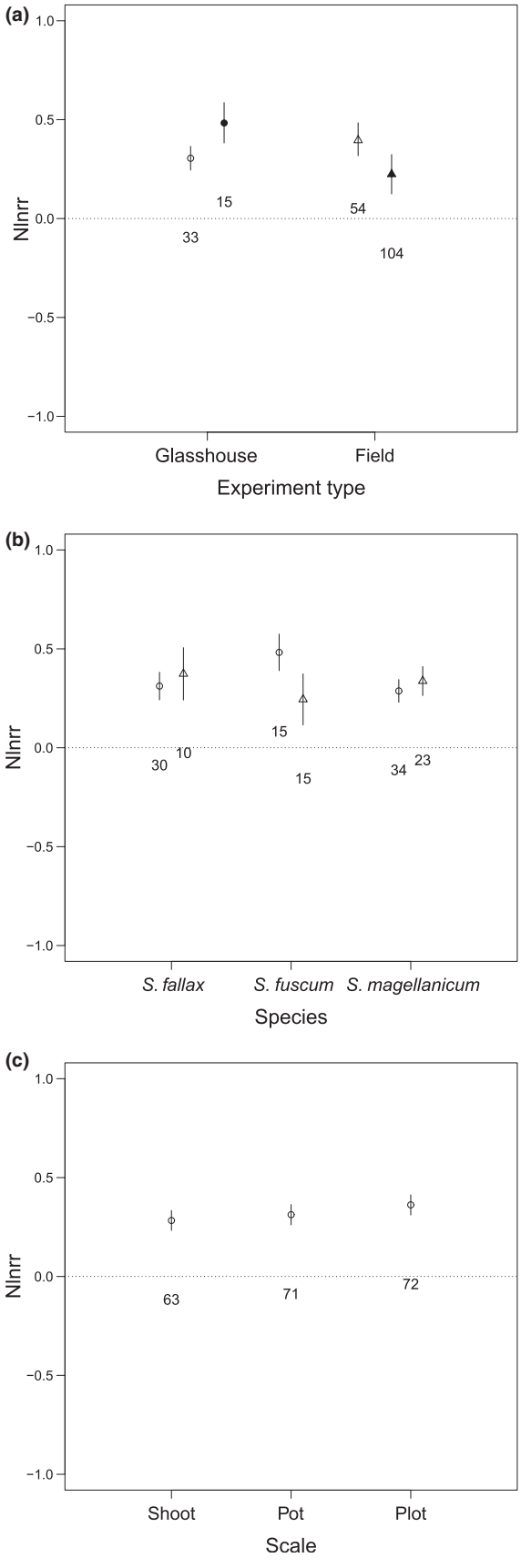

Fig. 2 Response of *Sphagnum* nitrogen (N) concentration (Nlnrr, $\log_e$ (treatment/control)) for: (a) glasshouse and field experiments with species characteristic of dry (hummocks, open symbols) and wet (lawns, filled symbols) microhabitats, (b) glasshouse (circles) and field (triangles) experiments for individual species (*Sphagnum fallax, S. fuscum, S. magellanicum*) and (c) for different scales of the experimental unit (shoot, pot and plot). Positive values for the response ratio, that is, above the thin horizontal line, indicate an increased production relative to the control treatment. Symbols represent fitted means conditioned on the variables in models b and c (Table 2). For tests at species level we added species interaction terms to model b (Table S2). Variables not shown in the figure are kept fixed at their mean value. Confidence in estimates is illustrated by 95% credible intervals. The numbers below the intervals indicate the number of experiments in each group.

differences in experiment type became evident. While N application in glasshouse experiments had a neutral to negative effect on PRODlnrr, adding similar rates of N in field experiments either stimulated or depressed *Sphagnum* production relative to the control, depending on the presence of vascular plants (Fig. 3a; Table 3, model b). Although the sample size for glasshouse experiments with vascular plants is low and the interaction experiment type × presence of vascular plants should be interpreted with some caution, the effect is consistent between experiments and studies. For the other interactions with experiment type, only the effects of background N deposition and P application differed between the field and glasshouse experiments (Table 3, model b). The negative effect of background N deposition was less pronounced in glasshouse experiments, and it was not driven by a few experiments or studies. By contrast, the interaction between experiment type and P application was less robust and should be treated with caution.

In order to explore whether the differences between glasshouse and field experiments could be attributed to the response of certain species, we fitted a model containing all explanatory variables significant in model b for a data subset including three species (*S. fallax*, *S. fuscum, S. magellanicum*) for which we had enough data (Table S2). The difference between glasshouse and field experiments was most pronounced for *S. magellanicum*, yielding a significant species × experiment type interaction (Fig. 3b, Table S2). While adding N to *S. magellanicum* in glasshouse experiments depressed production relative to the control, adding N at the same rate in field experiments stimulated production, resulting in a positive PRODlnrr.

The scale of the experimental unit interacted with N application, generally giving a more negative response at smaller scales (Fig. 3c; Table 3, model c). The effects in this model were not very robust, however, and should be treated as tentative. Note that the fitted model was adapted to only fit the interaction with vascular plants at pot and plot scale – the only scales for which we had experiments with and without removal of vascular plants.

The effect of N dose was assessed for data subsets containing all experiments for which information on the N concentration of the fertilization solution was available (*c.* 90% of all experiments). The analyses showed no effect of experiment type (interaction coefficient estimate: − 0.08, $P = 0.85$; Table S2).

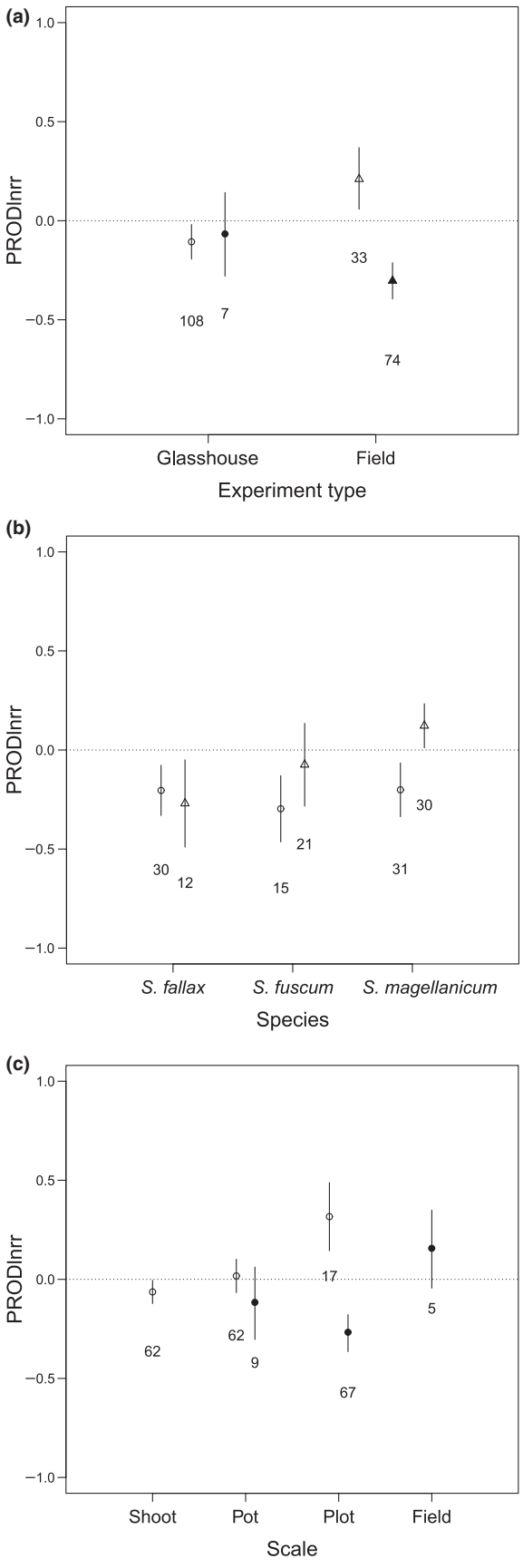

Fig. 3 Response of *Sphagnum* production (PRODlnrr, $\log_e$ (treatment/control)) for: (a) glasshouse and field experiments without (open symbols) and with (filled symbols) vascular plants, (b) glasshouse (circles) and field (triangles) experiments for individual species (*Sphagnum fallax*, *S. fuscum*, *S. magellanicum*) and (c) for different scales of the experimental unit (shoot, pot, plot and field) without (open symbol) and with (filled symbol) vascular plants. Note: these categories could be distinguished for the pot and plot scale only. Positive values for the response ratio, that is, over the thin horizontal line, indicate an increased production relative to the control treatment. Symbols represent fitted means conditioned on the variables in models b and c (Table 3). For tests at species level we added species interaction terms to model b (Table S2). Variables not shown in the figure are kept fixed at their mean value. Confidence in estimates is illustrated by 95% credible intervals. The numbers below the intervals indicate the number of experiments in each group.

Likewise, no clear effect of N form could be detected; However, as most experiments used NH_4NO_3, these results should be treated with caution (Table S2). Finally, we tested if the difference between glasshouse and field experiments could be related to differences in absolute production because one could argue that the variability of low production rates may be less than that of high production rates. However, when comparing production in control treatments between glasshouse and field experiments, there was no indication of such bias (mean ± SD; glasshouse: 215 ± 155 g m^{-2} yr^{-1}, field: 229 ± 207 g m^{-2} yr^{-1}).

Discussion

How similar are the outcomes of glasshouse and field experiments?

The response of *Sphagnum* N concentration to N application was similar between glasshouse and field experiments, suggesting that glasshouse studies are suitable for detailed physiological investigations on *Sphagnum* N uptake and assimilation. The response of *Sphagnum* production, however, diverged between the experiment types in several important respects. Glasshouse experiments overestimated the negative N-effect on production compared to field experiments in the absence of vascular plant shoots, but underestimated the effect of adding N with vascular plant shoots present. The discrepancy between glasshouse and field experiments was particularly evident at species level where, on average, adding N stimulated production of *S. magellanicum* in the field, but depressed production in the glasshouse. The above implies that an important component of the N-effect mediated through vascular plants cannot be studied in glasshouse experiments until we understand the mechanisms underlying the interaction with vascular plants better. Similar discrepancies between glasshouse and field experiments when assessing direct and indirect effects of perturbations can be found in other studies, although such effects have hitherto received little attention in plant ecology. For example, in a recent meta-analysis on the effects of water and light on plant performance, Holmgren *et al.* (2011) showed that physiological (direct) effects were very similar between glasshouse and field experiments, whereas seedling survival (indirect effect) was not.

We hypothesized that the responses of *Sphagnum* production (PRODlnrr) and tissue N concentration (Nlnrr) in the glasshouse would be modified by the same variables that operate in field experiments. For some variables, the effect was consistent between approaches: in both glasshouse and field experiments production decreased more at high than at low rates of N application. Most other variables, such as presence of vascular plants (PRODlnrr, not significant in glasshouse), temperature (PRODlnrr, not significant in glasshouse), P application (PRODlnrr, weaker effect in glasshouse), background N deposition (PRODlnrr, weaker effect in glasshouse) and microhabitat (Nlnrr, positive in glasshouse, negative in field) were of less significance in the glasshouse than in the field. The reasons behind the lack of sensitivity in the glasshouse remain speculative. Either these variables are mainly controlled for in glasshouse experiments (background N deposition), take effect at time scales beyond the duration of most glasshouse experiments, are mediated through other factors controlled for in glasshouse experiments (presence of vascular plants, temperature, microhabitat), are not fully comparable between the experiment types (temperature, microhabitat), or results reflect a limited amount of data. Sensitivity analyses suggested that the absence of a P effect in the glasshouse is driven by a few influential experiments only. The other effects were fairly robust, however, suggesting alternative explanations. The most interesting difference with field experiments, the vascular plant effect aside, is the absence of an interaction between N application and temperature in glasshouse experiments. Assuming we can equate the effect of average glasshouse temperature with that of mean July temperature, the absence of a temperature effect in glasshouse experiments suggests that the temperature sensitivity reported for field fertilization experiments (Limpens *et al.*, 2011), is likely an indirect effect mediated through other (environmental) factors and not a direct effect of temperature-induced changes in respiration or photosynthetic activity (Limpens *et al.*, 2011). The above clearly highlights the current lack of experiments targeted at elucidating interactions between N application and other environmental variables. In view of the potential consequences of such interactions for the C sequestration potential of peat bogs (Limpens *et al.*, 2011), this lack of mechanistic knowledge is worrying.

Why do glasshouse experiments give different results?

Glasshouse experiments underestimated the (indirect) effect of adding N compared to field experiments when vascular plant shoots are present, indicating that the vascular plant effect found in field experiments at sparse cover is not so much a consequence of competitive interaction, but, rather, it is more likely to be mediated through other factors controlled for in glasshouse experiments. The presence of vascular plant shoots in the field could potentially aggravate the effect of N by increasing interception of snow (Dorrepaal *et al.*, 2003) dry N deposition (Limpens *et al.*, 2004) or, perhaps, enable more intensive competition with algae (Gilbert *et al.*, 1998) by altering the microclimate (Chong *et al.*, 2012). Why the direct effect of adding N in the absence of vascular plant shoots was overestimated in the glasshouse compared to the field remains speculative. The contrasting response cannot be attributed to differences in overall production, at least not under our assumptions that 6 months in the glasshouse equalled one growing season in the field.

One explanation could be that the frequent application of low doses of N characteristic of glasshouse experiments has a more negative effect than infrequent application of high doses common in field experiments. As we found no indication of interactions with N dose and N frequency, this explanation seems unlikely. Another explanation could be that N uptake, or N mineralization, was higher in the glasshouse than in the field, resulting in N saturation and associated negative effects occurring more rapidly (Limpens *et al.*, 2011). Although the relationship between N influx and the response of *Sphagnum* N concentration seemed to be more linear in the glasshouse, suggestive of more efficient N uptake, there was no consistent difference in unstandardized *Sphagnum* N concentrations between glasshouse and field experiments (Fig. 1).

Alternatively, the deviating microclimate (low light intensity, low relative humidity, absence of seasonality) in glasshouse experiments may have interacted with the N effect, particularly because *Sphagnum* performance is sensitive to microclimate (Chong *et al.*, 2012). Although most *Sphagnum* species saturate photosynthesis at low to intermediate photosynthetically active radiation (PAR) up to 550 μmol m^{-2} s^{-1} (Marchall & Proctor, 2004; Hájek *et al.*, 2009), light intensities in glasshouse experiments are often lower. The latter particularly applies to glasshouse experiments conducted in winter and climate chamber experiments where the only source of light is artificial. As such we cannot exclude that at low light different N allocation patterns (Harley *et al.*, 1986) or, perhaps, changes in leaf morphology (Manninen *et al.*, 2011) may have interacted with N addition, even though recent evidence for *S. capillifolium* suggests otherwise (Bonnett *et al.*, 2010) and photosynthetic capacity is not negatively affected by N application in glasshouse or field experiments (Granath *et al.*, 2009, 2012). Interactions between N influx and microclimate may also explain the contrasting response of production between glasshouse and field experiments observed for *S. magellanicum* (Fig. 2b). For this species, adding the same amount of N depressed production in the glasshouse, but stimulated production in the field, despite controlling for differences in environmental variables. Changes in the sign of responses clearly indicate that the production response to N in the glasshouse was constrained compared with responses in the field. Glasshouse experiments focusing on the effect of adding N on *Sphagnum* at contrasting irradiance levels (Bonnett *et al.*, 2010) may help to explain some of the uncertainty regarding the underlying mechanisms.

Do scale or organizational complexity matter?

Organizational complexity affected the response of *Sphagnum* production to N additions, but neither scale nor organizational complexity modified the response of *Sphagnum* N concentration. The response of *Sphagnum* N concentration to adding N was similar among shoot, pot or plot experiments, corroborating the

direct cause–effect relationship between N deposition and tissue N concentration of bryophytes (Markert *et al.*, 2003). The main process determining the tissue N concentration, N uptake by *Sphagnum*, is the same for a single shoot, hummock or carpet and, thus, relatively insensitive to the type of experiment, its scale or the organizational complexity (Levin, 1992; Englund & Cooper, 2003). By contrast, the response of production to N application turned out to be far more sensitive to differences in experimental approach, reflecting its more indirect relationship with N deposition (Manning *et al.*, 2006). While removing vascular plant shoots, that is, reducing organizational complexity, affected the direction of the response of production to N application (qualitative response; Irvine *et al.*, 2004), the scale of the experimental unit only affected the strength of the response (quantitative response). The negligible effect of scale in assessing qualitative responses to N application corroborates earlier work by Wiedermann *et al.* (2008) who found high correspondence between N effects in a medium-scale field N application experiment and a large-scale field survey along a N deposition gradient in Sweden in the presence of vascular plants.

Conclusion

Glasshouse and field experiments gave similar qualitative and quantitative estimates of changes in *Sphagnum* N concentration in response to adding N. However, glasshouse-based estimates of changes in production, even qualitative assessments, diverged from field experiments owing to a stronger N effect on the production response in the absence of vascular plants in the glasshouse, presumably caused by the artificial microclimate; and a weaker N effect in the presence of vascular plants compared to field experiments. Thus, although we need glasshouse experiments to study the effects of interacting environmental factors on the response of *Sphagnum* to increased N deposition, we need field experiments to properly quantify these effects.

Acknowledgements

We are indebted to Dr John Stevens for his statistical advice on meta-regressions. We thank Uppsala University and Wageningen University for facilitating contact between the lead authors. Finally, a warm thanks to all researchers and sponsors associated with the N application experiments included in this meta-analysis: without their work and support, ours would not have been possible.

References

Bobbink R, Hicks K, Galloway J, Spranger T, Alkemade R, Ashmore M, Bustamante M, Cinderby S, Davidson E, Dentener F *et al.* 2010. Global assessment of nitrogen deposition effects on terrestrial plant diversity: a synthesis. *Ecological Applications* **20**: 30–59.

Bonnett SAF, Ostle N, Freeman C. 2010. Short term effect of deep shade and enhanced nitrogen supply on *Sphagnum capillifolium* morphophysiology. *Plant Ecology* **207**: 347–358.

Bragazza L, Limpens J, Gerdol R, Grosvernier P, Hájek M, Hajkova P, Iacumin P, Kutnar L, Rydin H, Tahvanainen T. 2005. Nitrogen content and $\delta^{15}N$ signature of ombrotrophic *Sphagnum* plants in Europe: to what extent is the increasing atmospheric N deposition altering the N-status of nutrient-poor mires? *Global Change Biology* **11**: 106–114.

Chong M, Humphreys ER, Moore TR. 2012. Microclimatic responses to increasing shrub cover and its effect on *Sphagnum* CO_2 exchange in an ombrotrophic bog. *Ecoscience*, in press.

Dise NB. 2009. Peatland response to global change. *Science* **326**: 810–811.

Dise NB, Phoenix GK. 2011. Peatlands in a changing world. *New Phytologist* **191**: 309–311.

Dorrepaal E, Aerts R, Cornelissen JHC, Callaghan TV, Van Logtestijn RSP. 2003. Summer warming and increased winter snow cover affect *Sphagnum fuscum* growth, structure and production in a sub-arctic bog. *Global Change Biology* **10**: 93–104.

Efron B, Tibshirani R. 1993. *An introduction to the bootstrap*. New York, NY, USA: Chapman and Hall.

Englund G, Cooper SD. 2003. Scale effects and extrapolations in ecological experiments. *Advances in Ecological Research* **33**: 161–202.

Freeman C, Fenner N, Ostle NJ, Kang H, Dowrick DJ, Reynolds B, Lock MA, Sleep D, Hughes S, Hudson J. 2004. Export of dissolved organic carbon from peatlands under elevated carbon dioxide levels. *Nature* **430**: 195–198.

Frolking S, Roulet NT. 2007. Holocene radiative forcing impact of northern peatland carbon accumulation and methane emissions. *Global Change Biology* **13**: 1079–1088.

Galloway JN, Townsend AR, Erisman JW, Becunda M, Cai Z, Freney JR, Martinelli LA, Seitzinger SP, Sutton MA. 2008. Transformation of the nitrogen cycle: recent trends, questions, and potential solutions. *Science* **320**: 889–892.

Gelman A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine* **27**: 2865–2873.

Gilbert D, Amblard C, Bourdier G, Francez A-J. 1998. Short-term effect of nitrogen enrichment on the microbial communities of a peatland. *Hydrobiologia* **373**: 111–119.

Granath G, Strengbom J, Rydin H. 2012. Direct physiological effects of nitrogen on *Sphagnum*: a greenhouse experiment. *Functional Ecology* **26**: 353–364.

Granath G, Wiedermann MM, Strengbom J. 2009. Physiological responses to nitrogen and sulphur addition and raised temperature in *Sphagnum balticum*. *Oecologia* **161**: 481–490.

Gurevitch J, Hedges V. 1999. Statistical issues in ecological meta-analyses. *Ecology* **80**: 1142–1149.

Gurevitch J, Mengersen K. 2010. A statistical view of synthesizing patterns of species richness along productivity gradients: devils, forests, and trees. *Ecology* **91**: 2553–2560.

Hájek T, Tuittila E-S, Ilomets M, Laiho R. 2009. Light responses of mire mosses – a key to survival after water-level drawdown? *Oikos* **118**: 240–250.

Harley PC, Tenhunen JD, Murray KJ, Beyers J. 1986. Irradiance and temperature effects on photosynthesis of tussock tundra *Sphagnum* mosses from the foothills of the Philip Smith mountains, Alaska. *Oecologia* **79**: 251–259.

Harrell FE. 2001. *Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis*. New York, NY, USA: Springer-Verlag.

Hautier Y, Niklaus PA, Hector A. 2009. Competition for light causes plant biodiversity loss after eutrophication. *Science* **324**: 636–638.

Hedges LV, Gurevitch J, Curtis PS. 1999. The meta-analysis of response ratios in experimental ecology. *Ecology* **80**: 1150–1156.

Hedges LV, Tipton E, Johnson MC. 2010. Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods* **1**: 39–65.

Holmgren M, Gomez-Aparicio L, Quero JL, Valladares F. 2011. Non-linear effects of drought under shade: reconciling physiological and ecological models in plant communities. *Oecologia*. doi: 10.1007/s00442-011-2196-5.

IPCC. 2007. Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL, eds. *Climate change 2007: the physical science basis. Contribution of Working Group I to the fourth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge, UK & New York, NY, USA: Cambridge University Press.

Irvine RL, Crone EE, Jackson LJ, MacIsaac EA. 2004. Does scale affect ecological model predictions? A test with lake responses to fertilization. *Ecological Applications* **14**: 1178–1188.

Juutinen S, Bubier JL, Moore TR. 2010. Responses of vegetation and ecosystem CO_2 exchange to 9 years of nutrient addition at Mer Bleue bog. *Ecosystems* **13**: 874–887.

Kulmatiski A, Beard KH, Stevens JR, Cobbold SM. 2008. Plant–soil feedbacks: a meta-analytical review. *Ecology Letters* **11**: 980–992.

Lajeunesse MJ. 2010. Achieving synthesis with meta-analysis by combining and comparing all available studies. *Ecology* **91**: 2561–2564.

Levin SA. 1992. The problem of pattern and scale in ecology. *Ecology* **73**: 1943–1967.

Limpens J, Berendse F, Blodau C, Canadell J, Freeman C, Holden J, Roulet N, Rydin H, Schaepman-Strub G. 2008. Peatlands and the carbon cycle: from local processes to global implications – a synthesis. *Biogeosciences* **5**: 1475–1491.

Limpens J, Berendse F, Klees H. 2004. How P affects the impact of N deposition on *Sphagnum* and vascular plants in bogs. *Ecosystems* **7**: 793–804.

Limpens J, Granath G, Gunnarsson U, Aerts R, Bayley S, Bragazza L, Bubier J, Buttler A, Van den Berg L, Francez A-J *et al.* 2011. Climatic modifiers of the response to N deposition in peat-forming *Sphagnum* mosses: a meta-analysis. *New Phytologist* **191**: 496–507.

Manninen S, Woods C, Leith ID, Sheppard LJ. 2011. Physiological and morphological effects of long-term ammonium or nitrate deposition on the green and red (shade and open grown) *Sphagnum capillifolium*. *Environmental and Experimental Botany* **72**: 140–148.

Manning P, Newington JE, Robson HR, Saunders M, Bradford MA, Eggers T, Ellis RJ, Gange AC, Marhan S, Kandeler E *et al.* 2006. Decoupling the direct and indirect effects of nitrogen deposition on ecosystem function. *Ecology Letters* **9**: 1015–1024.

Marchall M, Proctor MCF. 2004. Are bryophytes shade plants? Photosynthetic light responses and proportions of chlorophyll *a*, chlorophyll *b* and total carotenoids. *Annals of Botany* **94**: 593–603.

Markert B, Breure A, Zechmeister HG. 2003. *Bioindicators and biomonitors – principles, concepts and applications*. Amsterdam, the Netherlands: Elsevier.

Pearce ISK, van der Wal R. 2008. Interpreting nitrogen pollution thresholds for sensitive habitats: the importance of concentration versus dose. *Environmental Pollution* **152**: 253–256.

Pinheiro J, Bates D, DebRoy S, Sarkar D, the R Development Core Team. 2011. *nlme: linear and nonlinear mixed effects models*. R package version 3.1-102.

R Development Core Team. 2011. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: http://www.R-project.org.

Rydin H, Jeglum JK. 2006. *The biology of peatlands*. Oxford, UK: Oxford University Press.

Stevens JR, Nicholas B. 2009. Metahdep: meta-analysis of hierarchically dependent gene expression studies. *Bioinformatics* **25**: 2619–2620.

Stevens JR, Taylor AM. 2009. Hierarchical dependence in meta-analysis. *Journal of Educational and Behavioural Statistics* **34**: 46–73.

Thompson SG, Higgins JP. 2002. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* **21**: 1559–1573.

Wiedermann MM, Gunnarsson U, Nilsson MB, Nordin A, Ericson L. 2008. Can small-scale experiments predict ecosystem responses? An example from peatlands. *Oikos* **118**: 449–456.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Scatterplots of effect size vs sample size and sample variance (i.e. within-study variance) for PRODlnrr (a, c), and Nlnrr (b, d).

Table S1 Data file with overview of data sources and explanatory variables

Table S2 Supplementary models for PRODlnrr and Nlnrr testing effects of species, N dose concentration, N form and depth of the water table

Notes S1 Additional methods information: model choices and variable description.