Reassessing the observational evidence for nitrogen deposition impacts in acid grassland: Spatial Bayesian linear models indicate small and ambiguous effects on species richness

O.L. Pescott[1*] & M. Jitlal[2]

1. Centre for Ecology and Hydrology, Crowmarsh Gifford, Wallingford, Oxfordshire, United Kingdom
2. Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, United Kingdom
* Corresponding author, oliver.pescott@ceh.ac.uk

Abstract

Nitrogen deposition (Ndep) is a threat to plant diversity in grassland ecosystems around the world. The evidence supporting this conclusion comes from both observational and experimental research, with "space-for-time" substitution surveys of pollutant gradients a significant portion of the former. However, estimates of regression coefficients for Ndep impacts derived with a focus on causal inference are hard to locate in the observational literature. Some influential observational studies have presented estimates from univariate models, overlooking the effects of omitted variable bias, and/or have used *P*-value-based stepwise variable selection (PSVS) to infer impacts, a strategy now known to be poorly suited to the accurate estimation of regression coefficients. Broad-scale spatial autocorrelation has also generally been unaccounted for. We re-examine two UK observational datasets that have previously been used to investigate the relationship between Ndep and plant species richness in acid grasslands, a much-researched habitat in this context. One of these studies (Stevens et al. 2004) estimated a large negative impact of Ndep on richness through the use of PSVS; the other reported smaller impacts (Maskell et al. 2010), but did not explicitly report regression coefficients or partial effects, making the actual size of the estimated Ndep impact difficult to assess. We reanalyse both datasets using a spatial Bayesian linear model estimated using integrated nested Laplace approximation (INLA). Contrary to previous results, we found similar-sized estimates of the Ndep impact on plant richness between studies, both with and without bryophytes, albeit with some disagreement over the most likely direction of this effect. Our analyses suggest that some previous estimates of Ndep impacts from space-for-time substitution studies are likely to have been over-estimated, and that the evidence from observational studies could be fragile when confronted with alternative model specifications, although further work is required to investigate potentially nonlinear responses. Given the growing literature on the use of observational data to estimate the impacts of pollutants on biodiversity, we suggest that a greater focus on clearly reporting important outcomes with associated uncertainty, the use of techniques to account for spatial autocorrelation, and a clearer focus on the aims of a study, whether explanatory or predictive, are required.

Keywords

Nitrogen deposition, species richness, grasslands, observational studies, causal inference, spatial autocorrelation, INLA

Introduction

Nitrogen deposition (Ndep) is a significant threat to the plant diversity of various habitat types, both in north-western Europe (UK National Ecosystem Assessment, 2011) and around the world (Phoenix et al., 2006). The evidence for this position comes from a variety of sources, including correlative analyses of observational data (e.g. Maskell, Smart, Bullock, Thompson, & Stevens, 2010), typically conducted across relatively large areas as "space-for-time" gradient studies, and small scale experiments (e.g. Van der Eerden, Dueck, Berdowski, Greven, & Van Dobben, 1991), although the latter have also frequently been pooled across larger areas through meta-analytical approaches (Chris M. Clark et al., 2007; Phoenix et al., 2012; Soons et al., 2017). Reviews of Ndep impacts on plant biodiversity have typically drawn on all of this evidence (Bobbink et al., 2010; RoTAP, 2012; Rowe et al., 2017; Stevens, Smart, et al., 2011; UK National Ecosystem Assessment, 2011), in addition to other types of studies, such as Before-After surveys of historic plots (e.g. Britton, Beale, Towers, & Hewison, 2009). Observational and experimental studies are therefore both generally considered useful ways of understanding pollutant-driven biodiversity change in terrestrial ecosystems.

Different inferential approaches are often considered complementary, with large-scale, observational methods potentially allowing access to "treatment" effects across pre-existing gradients, with levels of replication that would likely be challenging to resource via an experimental route (but see Fraser et al., 2013). One cost of this approach is that the effect of interest is likely to be crossed in various complex ways with numerous other variables, including historic drivers for which data are likely to be inaccessible, leaving one with a large choice of covariates that could potentially be included in a model, including some which will be unknown, or suspected to be of importance but impossible to access. Furthermore, spatially autocorrelated variables that are not captured by the covariates included, or other processes causing spatial structure such as dispersal, may also need to be accounted for to ensure accurate estimation of regression coefficients (Beale, Lennon, Yearsley, Brewer, & Elston, 2010; Crase, Liedloff, Vesk, Fukuda, & Wintle, 2014). The ultimate purpose of a statistical model must also be taken into account in making analytical decisions: does one primarily wish to make predictions, or is the focus on unbiased effect estimation to develop causal understanding (Mac Nally, 2000; Shmueli, 2010; Stephens, Buskirk, & del Rio, 2007)? Although causal inference may imply predictive success, models that are constructed using methods that solely seek to maximise predictive accuracy will not necessarily capture causal processes accurately. Even when explicitly aiming for causal explanations through regression modelling in non-experimental situations, estimated coefficients may still only have a weak claim to be viewed as causal effects (Gelman & Hill, 2007; Young, 2018). Statistical issues then, in addition to domain-specific understanding, must also be at the forefront when attempting to make statements about cause and effect from observational data (Rubin, 2004; Young, 2018).

The work on Ndep presented here arose from a desire to use information from existing studies to inform the analysis of new data through the use of informative priors in a Bayesian framework (Lemoine, 2019). However, causality-focused estimates of regression coefficients from observational studies of Ndep impacts on plant richness proved hard to find in the current literature. For example, several studies have presented "headline" estimates from univariate models after multiple regression modelling (Field et al., 2014; Maskell et al., 2010; Stevens, Dise, Mountford, & Gowing, 2004), implying regression coefficient values for Ndep that do not necessarily have any causal meaning, and/or have used *P*-value-based stepwise variable selection (PSVS) to derive final models, a strategy now known to be poorly suited to the accurate estimation of regression coefficients for inferential purposes (Faraway, 2014; Harrell, 2015; Heinze, Wallisch, & Dunkler, 2018; Mac Nally,

2000; Whittingham, Stephens, Bradbury, & Freckleton, 2006). Re-analysis of existing studies with the primary aim of developing models that focus on causal inference may therefore be valuable in exploring the dependence of earlier conclusions on modelling choices. Fortunately, data focusing on UK acid grasslands from two such studies (Maskell et al., 2010; Stevens et al., 2004) were available for re-analysis; note also that similar work has been done across other habitats (e.g. Field et al., 2014; Maskell et al., 2010), and across larger areas (e.g. Dupré et al., 2010).

The main focus of the current work is on deriving estimates of the effect sizes of nitrogen deposition on plant species richness, rather than on their statistical significance (Amrhein, Greenland, & McShane, 2019), whilst accounting for previously unmodelled broad-scale spatial autocorrelation. The model forms investigated were specified *a priori*; in general these were delimited by the full sets of covariates previously investigated by the original studies, given that these all have good ecological reasons for inclusion. We do not make the (unprovable) claim that our "full" models are fully correct with respect to reality, merely that the inclusion of as many plausible covariates as possible is likely to help avoid omitted variable bias, given that the ignorability assumption (i.e. all confounders are measured) is more likely to be satisfied (Gelman & Hill, 2007; Young, 2018), and that even our worst case ratio of response data points to predictors meets rules of thumb put forward by statisticians to help ensure that coefficients can be estimated accurately (Harrell, 2015, p. 72; Heinze et al., 2018). In addition, we also include a model for the Stevens et al. (2004) data that focuses on a reduced set of covariates chosen for their expected similarity to the ecological impacts of the covariates used by Maskell et al. (2010) for comparative purposes. Two additional models for the Maskell et al. (2010) data were specified *post hoc*, due to a desire to investigate differences between our results and those of the original paper.

Materials and Methods

Datasets

These are discussed chronologically by date of the field survey that created the original dataset, rather than the date of publication of the analysis. The data analysed by Maskell et al. (2010), "MEA10" hereafter, are described in that paper, and were originally collected as a part of the 1998 UK Countryside Survey (UKCS; http://www.countrysidesurvey.org.uk). Briefly, the data analysed by MEA10 were selected from the 1998 UKCS on the basis of matches between plant communities in 2 x 2 m plots and particular National Vegetation Classification syntaxa; for acid grasslands, the chosen plots had a best fit to the acid grassland types U1-9 (Rodwell, 1992). For more details on the UKCS sampling strategy see Maskell et al. (2010) and Smart et al. (2003); we only note here that the UKCS is a stratified sample of so-called UK land classes (Firbank et al., 2003), with systematic random sampling within strata. We take the opportunity here to clarify some points relating to the datasets used in MEA10: the number of acid grassland plots used in analyses was 883, not 895 as reported; total nitrogen deposition from the modelled dataset of Smith et al. (2000) was specifically the estimated deposition over moorland, where moorland was defined according to the Land Cover Map 1990 (Fuller, Groom, Jones, & Thomson, 1993), and so is a better match to acid grassland than the grassland category of Smith et al. (2000); and, finally, the sulphur dioxide deposition covariate was not absolute deposition for a particular year, but the difference between the modelled values over moorland for 1998 and the modelled peak in 1970 at a 5 x 5 km resolution (all L.C. Maskell & S.M. Smart, pers. comms). This approach ensured that the covariate effectively measured the recent and substantial reduction in acidifying sulphur deposition across Britain.

Descriptions of the data collected by Stevens et al. (2004), "SEA04" hereafter, can be found in that paper, and several others (e.g. Stevens, Dupré, Gaudnik, et al., 2011). The dataset reanalysed here was archived by Stevens, Dupré, Dorland et al. (2011; Ecological Archives deposit E092-128). Briefly,

SEA04 surveyed a random sample of 68 sites (after applying size and accessibility filters) from a larger database of acid grasslands collated by British national conservation agencies (the grasslands in this database may not be representative of the total national habitat resource). The random sample was stratified across a nitrogen deposition gradient, the gradient being again the deposition model of Smith et al. (2000), as for MEA10. Within sites, five 2 x 2 m plots were recorded (i.e. n = 68 × 5 = 340) within a larger 100 x 100 m area chosen to contain at least 50% of NVC type U4 (*Festuca ovina - Agrostis capillaris - Galium saxatile*; Rodwell, 1992) acid grassland (Stevens et al. 2004). Note, however, that Stevens et al. (2011) archived data for 320 of the sites from Stevens et al. (2004), and it is this dataset that we reanalyse here.

Data preparation

Within each dataset, some covariates were re-scaled to allow for more direct comparisons between regression coefficients, and to allow for their more intuitive ecological interpretation (Gelman & Hill, 2007). For example, altitude, scaled in metres in the original datasets, were divided by 100 to produce regression coefficients that estimated the change in richness per 100 m. This produces a range of 9.75 (0.00–9.75), comparable to the other covariates, and a more ecologically-interpretable regression coefficient (see Table 1 for all covariates used and their ranges). Although standardisation to unit variance is also often recommended for improving the comparability of coefficients in regression modelling (Gelman & Hill, 2007), it can also make direct comparisons between studies more difficult (Baguley, 2009), and so we focus here on models estimated using our rescaled covariates (where deemed necessary; see Table 1). MEA10 reported little difference in conclusions with respect to the analysis of vascular plant richness only or vascular plant plus bryophyte richness. SEA04 included bryophytes in their analyses. We focus primarily on vascular plant richness responses only (given that they are more likely to be accurately estimated; for example, MEA10 note that UKCS surveys only include “a selected range of the more easily identifiable bryophytes”), but we also report results from analyses including bryophytes in Supplementary Information 2.

SEA04 considered a larger number of covariates than MEA10, therefore we present two reanalyses of the SEA04 dataset here: model 1, using a smaller set of covariates chosen to match those of MEA10 in terms of their likely ecological effects; and model 2, using a larger set of covariates, matching those considered by SEA04 as closely as possible. Note that the full set of covariates considered by SEA04 contains some that are linear combinations of each other (e.g. total acid deposition is given by SEA04 as total N plus total S; likewise, total N is normally calculated as reduced N plus oxidised N), for this reason, we only consider total N deposition and total S deposition as pollutant covariates in our models (Table 1). For SEA04 model 2, four of the available climate variables chosen to match the analysis of Stevens et al. (2004) had very high pairwise linear correlations (all with $r > 0.75$); these variables (mean annual potential evapotranspiration (PET), mean annual daily maximum and minimum temperatures, 1996-2006, all from the MARS dataset (see Stevens, Dupré, Gaudnik, et al., 2011); and mean annual potential evapotranspiration from Tanguy et al. (2018)) were combined using PCA and the first two principle components of this ordination used in their place (Table 1). PET data were unavailable for 3 sites (15 plots) in SEA04 (all on Lundy Island, England, 51° 10’ 57.82” N, 4° 40’ 11.46” W), and these values were imputed as the mean PET value across the remainder of the SEA04 dataset prior to PCA.

Statistical models

As discussed by Blangiardo & Cameletti (2015), models for point-referenced data (i.e. those with measurements of some outcome across a set of specific locations, where the locations are indexed by a two (or three) dimensional vector) are one type of spatial data that can be modelled within a Bayesian framework (such point-referenced data are also known as geostatistical data). Hierarchical

approaches to regression modelling, where unstructured random effects are incorporated within the model, are often implemented using a Bayesian approach (Blangiardo & Cameletti, 2015; Gelman & Hill, 2007). Such models can be extended to include structured random effects that allow analysts to account for similarities based on temporal or spatial neighbourhoods. The integrated nested Laplace approximation (INLA) method of approximate Bayesian inference is particularly well-suited to this area of modelling given its speed and relative ease of implementation (Blangiardo & Cameletti, 2015).

We modelled both datasets using hierarchical Poisson regressions in R-INLA (Rue, Martino, & Chopin, 2009; www.r-inla.org). All models included a set of covariates and an additional random spatial field, to account for broad-scale spatial autocorrelation, as independent variables. The spatial field was evaluated using the stochastic partial differential equation (SPDE) approach developed by Lindgren et al. (2011), and was specified as a mesh constructed using a triangulation based on the 1 × 1 kilometre squares of the British national grid (EPSG identifier 27700). Our models also included a random effect for the 1 kilometre squares containing the plots (as per Maskell et al., 2010), because in both datasets there were instances of multiple plots being recorded within a single square. Therefore the random spatial field models broader scale spatial autocorrelation between squares, whilst the 1 km square random effect captures smaller scale autocorrelation between plots within squares (Maskell et al., 2010). For the MEA10 dataset, such models had considerably lower values of the Deviance Information Criterion (DIC) than models omitting a random effect of 1 km square (ΔDIC = 54.6 or 31.8, without and with bryophyte data respectively); for SEA04 models 1 and 2 the inclusion of the 1 km square random effect was less important, but still generally improved the model (model 1: ΔDIC = 1.3 or -3.7; model 2: ΔDIC = 2.6 or 2.0; both without and with bryophyte data respectively). The 1 km square random effect was therefore included for all models. The same check was made on the inclusion of the random spatial field based on excluding this term from a model including the 1 km square random effect. The results indicated that the spatial random field was required in five out of six models (MEA10: ΔDIC = 7.4 or 10.8, without and with bryophyte data respectively); for SEA04: model 1: ΔDIC = 7.4 or -1.1; model 2: ΔDIC = 11.1 or 10.6; both without and with bryophyte data respectively). The form of our final model can therefore be described as follows:

$$y_{ij} \sim Poisson(\lambda_{ij})$$

$$\lambda_{ij} = \exp(\beta_0 + \sum_{m=1}^{M} \beta_m X_{mij} + \beta_j Q_{j[i]} + \omega_{ij})$$

Where y_{ij} is an observation of plant species richness in any plot/square combination, *i* is a unique plot identifier, *j* is a unique 1 km square identifier, β_0 is a global intercept, β_m is the regression coefficient for covariate X_m (*m* = 1,…,M) and β_j is the hierarchical coefficient for each 1 km square, with $\beta_j \sim N(0, \theta^2)$ and $Q_{j[i]}$ indicates the association of the *i*th observation with the *j*th square (Gelman & Hill, 2007). The ω_{ij} term represents an additional spatial effect assumed to be a zero mean Matern Gaussian Markov random field (GMRF) evaluated by the SPDE solution estimated over the specified triangulated mesh (Figure A1.7). A detailed overview of geostatistical models estimated using SPDEs and GMRFs can be found in Blangiardo & Cameletti (2015). Priors for all parameters were left at the INLA defaults (see www.r-inla.org).

Results

The inclusion of both the 1 km square random effect and the spatial random field improved almost all the models of the effects of N deposition on plant species richness in acid grassland explored here (as judged by DIC, see “Statistical models” above). The one model (SEA04 model 1 with bryophyte

data) where ΔDIC was negative when either the 1 km square random effect or the spatial random field was dropped (indicating that the simplified model was preferred) still required a spatial random effect at some scale: dropping both random effects gave a DIC value of 1732.7 compared to a value of 1712.7 for the full model; the preferred model in this case was that with the spatial random field only (DIC = 1709.0). Estimates of the overall random effects variance explained by the structured spatial field varied between 87.4% (MEA10, vascular plants only) to 99.7% (SEA04 Model 1, vascular plant and bryophyte richness combined) across the models including both random effects.

The addition of an extra kilogram of total N deposition on plant species richness per hectare per year was consistently amongst the smallest effect estimated by our models, with low uncertainty in comparison to other covariates (Figs 1, 2, 3, A2.1, A2.2, A2.3). Exponentiating the fixed effects indicated that the impact of increasing total N deposition is equivocal across these datasets and models (Table 2): six models estimated a 1% drop in species richness as the most likely impact (Figs 2, 3, A2.2, A2.3), although two of these models (Supplementary Information 3; SI3) were formulated *post hoc* and were not favoured as the best models for the dataset by their DIC estimates. Two models, estimated from MEA10 using the DIC-favoured model, suggested that an increase in species richness had more support (Figs 1, A2.1). The reduced covariate set used for the SEA04 data ("model 2") for closer comparison to the covariate set used in the MEA10 analysis made no difference to the estimated effect of N deposition (Table 2; Figs 3, A2.3). Figure 4 illustrates the estimated marginal effect of N deposition (with other covariates set to zero) on vascular plant species richness for two of our models (MEA10 and SEA04 model 1), these demonstrate both the relatively small predicted mean impact on species richness, the disagreement in impact direction between the datasets, and the uncertainty attached to these predictions.

The results presented in SI3 were included as exploratory, *post hoc*, analyses after it was found that our DIC-favoured models of the MEA10 dataset indicated that the most likely overall influence of Ndep on species richness was more likely to be positive, contrary to the conclusions of Maskell et al. (2010). Our additional analyses (SI3), each with only one of the spatial random effects used in the full model, did show a reversal of average sign, with the analysis matching that of Maskell et al. (2010), i.e. using a 1 km square random effect, indicating the strongest negative effect of N deposition on vascular plant richness (a mean of a 2% loss of species richness per kg ha^{-1} yr^{-1} of total N deposition; Table 2, Fig. A3.1). However, this model had a DIC value 7.4 units higher than the favoured model.

Discussion

Previous work on the estimation of the effects of nitrogen deposition on plant species richness from space-for-time substitution studies is likely to have had a significant impact on both science and policy (assuming that the number of citations received can be used as an index to this). Therefore being clear about the reliability and accuracy of estimates from observational work is important, not only for fundamental reasons of improving ecological understanding and analyses, but also because policy decisions must be made concerning the utility of funding work in this area relative to other ways of assessing environmental change (although field experimental work may have issues of its own; Peters, 1991). Clear and accurate estimates of the likely sizes of effects are invaluable for informing future studies, for example, through power analyses, or the use of informative priors for Bayesian approaches. Our results suggest that only small effects of nitrogen deposition on richness may be detectable in these observational datasets.

The headline result of Stevens et al. (2004) reported "a reduction of one species per 4-m^2 quadrat for every 2.5 kg N ha^{-1} yr^{-1} of chronic nitrogen deposition". An estimated reduction in species

richness of 23% (based on a change from deposition of 5 kg N ha^{-1} yr^{-1} to 17 kg N ha^{-1} yr^{-1}) was also highlighted as a key message using a regression coefficient estimated from a model including only Ndep as a covariate. The estimate from our SEA04 model 1 implied a loss of around 1% of richness per kg N ha^{-1} yr^{-1}; however, this ignores the uncertainty associated with this particular coefficient estimate, which included an estimate of no effect at the 97.5% percentile, and the larger uncertainty associated with alternative model specifications (including ones that we have not explored here). Indeed, our estimates from the larger, random-stratified national survey of Maskell et al. (2010) suggested that a small average increase in richness along the total nitrogen deposition gradient was more plausible under the favoured model for that dataset. Furthermore, the marginal effects calculated from our full models (Fig. 4) indicated that the average effect on richness across Britain may only approach a difference of around one species in either direction. This is also notable because previous commentaries (e.g. Tipping et al. 2013) have suggested that differences in the apparent effect sizes of Ndep between the two observational studies reanalysed here were likely due to study design, whereas our work suggests that the differences are actually minimal, and that those found previously may simply be artefacts of the different statistical modelling procedures adopted by the original studies. The different median signs found in our analyses could also be due to differences in the statistical populations targeted, namely the differential sampling of grassland types, with the greater plant community coverage of the dataset of Maskell et al. (2010) potentially covering more sites on the left of a net primary productivity/diversity unimodal curve, resulting in an average increase in richness under fertilisation (see below for more discussion of the potential impacts of non-linear responses and other issues on the current work).

Our results should cause others to re-evaluate their approaches. For example, Field et al. (2014) also used PSVS, and other *P*-value based selection techniques, to formulate models in their space-for-time analyses of Ndep impacts for a number of semi-natural habitats across Britain, as well as highlighting simple univariate relationships through scatter plots. These authors, however, did note in their methods section that their results should be “interpreted with caution”, although this statement of uncertainty was not clearly carried through to other parts of their paper, nor to other research. For example, Payne et al. (2017) used data from both SEA04 and Field et al. (2014) to forecast the impacts of Ndep on plant species richness under future nitrogen deposition scenarios. Ndep and its polynomial transformations were the only covariates in these models. Payne et al. (2017) noted in their online supplementary material (their Web Panel S1) that such forecasting makes the strong assumption that covariances between variables remain the same at the future time point (i.e. the distribution of nitrogen deposition amounts will vary with, e.g., mean maximum July temperature in the same way in 2030 as they did when originally modelled), but they did not clearly acknowledge that the Ndep-only models in their forecasting exercise had received no validation in support of their status as the best predictive (or explanatory for that matter) models by the original studies or elsewhere. Ndep may be a useful covariate for producing good predictive models of species richness in certain habitats, but this should be demonstrated using measures of out-of-sample predictive ability, and the resulting partial effects of Ndep in such models has no intrinsic claim to be a reliable estimate of its causal relationship with richness (cf. Payne et al. 2017).

These issues from the literature should not detract from the fact that there are also several ways in which the work presented here may also be deficient. Our efforts are linear models (as for SEA04, MEA10, Field et al. 2014, among others), but there is increasing evidence that the response of species richness to Ndep may be better modelled as unimodal (Christopher M. Clark et al., 2019; Simkin et al., 2016; Tipping, Henrys, Maskell, & Smart, 2013). Simkin et al. (2016) still, however, reported declines in the plant species richness of open habitats above 8.7 kg N ha^{-1} yr^{-1}, very similar to the 7.9 kg N ha^{-1} yr^{-1} richness decline threshold identified for acid grassland by Tipping et al.

(2013), suggesting that a large part of the gradient studied by SEA04 and MEA10 may still be well-approximated by a linear relationship (although we note that the approach of Tipping et al. 2013 also makes the assumption that there are no omitted linear or non-linear variables correlated with Ndep that might change their estimated univariate breakpoint relationship). Certainly, different approaches to accounting for this likely non-linearity in multivariable models, such as the inclusion of interactions or smoothers, should be further explored, particularly given that novel methods focused on causal inference that can account for these issues continue to be developed (Dorie, Hill, Shalit, Scott, & Cervone, 2019). Other uncertainties relating to our conclusions pertain to the fact that the broad-scale spatial field used here may be accounting for information that, if known, would change the size of the Ndep impact regression coefficient. This could be in the form of additional covariates, or more highly resolved estimates of the Ndep load that a location has actually received. The Ndep estimates used here (and by the original studies) are resolved to a grain size of 5 x 5 km, and this additional uncertainty could have attenuated our estimate of the regression coefficient (the measurement error in explanatory variables problem; e.g. see Fox, 2016), even if there is no systematic bias relating to the association of 5 x 5 km-estimated Ndep levels with particular types of vegetation. This argument is likely to apply to many of the variables used here and in other studies, particularly given that many are estimated from other modelling exercises at relatively large spatial grain sizes (Table 1). Conversely, measurement error coupled with conditioning on statistical significance, as happens through PSVS (e.g. Field et al., 2014; Stevens et al., 2004), is likely to result in the overestimation of effects (Gelman & Carlin, 2014; Loken & Gelman, 2017). Whether or not the potential for these biases is more serious for inference than the absence of covariates that are completely unavailable, such as historic land management events (e.g. Rackham, 1986), is very difficult to say in the absence of focused simulation exercises.

Conclusions

Whilst we do not dispute the negative impacts of heavy nitrogen deposition on plant species richness, the evidence put forward through analyses of observational data appears to have been overstated. We estimate a possible decline in richness of around 1% per kg ha^{-1} yr^{-1} of total N deposition from two spatially and temporally separated British space-for-time gradient studies, considerably less than the estimates implied previously by described and/or plotted relationships from primary studies (Maskell et al., 2010; Stevens et al., 2004). Moreover, even this estimate appears uncertain, and our favoured model for the acid grassland data of Maskell et al. (2010) suggests that an average increase in richness of a similar magnitude may be more likely. The previous lack of presented regression coefficients derived with causal inference as their main focus, and of models that account for broad-scale spatial autocorrelation, is important to note, because scientists wishing to use the estimated effects of Ndep for the design of future studies, or for the construction of informative priors in new analyses, may be misled as to the size of effect that is expected to be detectable in datasets of this type. The various models presented here could be thought of as a small section of the “multiverse” of potential approaches to these data, an approach that has been put forward as an additional route to transparency and reproducibility in science, and which can provide insights into the fragility or robustness of particular conclusions (Steegen, Tuerlinckx, Gelman, & Vanpaemel, 2016). As such, this work is unlikely to be the last word on these datasets, or in the general area of observational studies of nitrogen deposition impacts. We hope that our re-analyses inspire further efforts to accurately extract the maximum available knowledge from these valuable datasets, whether for explanatory or predictive purposes.

References

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, *567*(7748), 305. doi: 10.1038/d41586-019-00857-9

Baguley, T. (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology*, *100*(3), 603–617.

Beale, C. M., Lennon, J. J., Yearsley, J. M., Brewer, M. J., & Elston, D. A. (2010). Regression analysis of spatial data. *Ecology Letters*, *13*(2), 246–264. doi: 10.1111/j.1461-0248.2009.01422.x

Blangiardo, M., & Cameletti, M. (2015). *Spatial and Spatio-temporal Bayesian Models with R - INLA*. Chichester, West Sussex: John Wiley & Sons.

Bobbink, R., Hicks, K., Galloway, J., Spranger, T., Alkemade, R., Ashmore, M., … De Vries, W. (2010). Global assessment of nitrogen deposition effects on terrestrial plant diversity: a synthesis. *Ecological Applications*, *20*(1), 30–59. doi: 10.1890/08-1140.1

Britton, A. J., Beale, C. M., Towers, W., & Hewison, R. L. (2009). Biodiversity gains and losses: Evidence for homogenisation of Scottish alpine vegetation. *Biological Conservation*, *142*(8), 1728–1739. doi: 10.1016/j.biocon.2009.03.010

Clark, Chris M., Cleland, E. E., Collins, S. L., Fargione, J. E., Gough, L., Gross, K. L., … Grace, J. B. (2007). Environmental and plant community determinants of species loss following nitrogen enrichment. *Ecology Letters*, *10*(7), 596–607. doi: 10.1111/j.1461-0248.2007.01053.x

Clark, Christopher M., Simkin, S. M., Allen, E. B., Bowman, W. D., Belnap, J., Brooks, M. L., … Waller, D. M. (2019). Potential vulnerability of 348 herbaceous species to atmospheric deposition of nitrogen and sulfur in the United States. *Nature Plants*, *5*(7), 697–705. doi: 10.1038/s41477-019-0442-8

Crase, B., Liedloff, A., Vesk, P. A., Fukuda, Y., & Wintle, B. A. (2014). Incorporating spatial autocorrelation into species distribution models alters forecasts of climate-mediated range shifts. *Global Change Biology*, *20*(8), 2566–2579. doi: 10.1111/gcb.12598

Dorie, V., Hill, J., Shalit, U., Scott, M., & Cervone, D. (2019). Automated versus Do-It-Yourself Methods for Causal Inference: Lessons Learned from a Data Analysis Competition. *Statistical Science*, *34*(1), 43–68. doi: 10.1214/18-STS667

Dupré, C., Stevens, C. J., Ranke, T., Bleeker, A., Peppler-Lisbach, C., Gowing, D. J., … Diekmann, M. (2010). Changes in species richness and composition in European acidic grasslands over the past 70 years: the contribution of cumulative atmospheric nitrogen deposition. *Global Change Biology*, *16*(1), 344–357.

Faraway, J. J. (2014). *Linear models with R* (2nd ed.). Boca Raton, FL: CRC Press.

Field, C. D., Dise, N. B., Payne, R. J., Britton, A. J., Emmett, B. A., Helliwell, R. C., … Caporn, S. J. M. (2014). The role of nitrogen deposition in widespread plant community change across semi-natural habitats. *Ecosystems*, *17*(5), 864–877. doi: 10.1007/s10021-014-9765-5

Firbank, L. G., Barr, C. J., Bunce, R. G. H., Furse, M. T., Haines-Young, R., Hornung, M., … Smart, S. M. (2003). Assessing stock and change in land cover and biodiversity in GB: an introduction to Countryside Survey 2000. *Journal of Environmental Management*, *67*(3), 207–218.

Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models* (3rd ed.). California, USA: SAGE Publications.

Fraser, L. H., Henry, H. A., Carlyle, C. N., White, S. R., Beierkuhnlein, C., Cahill, J. F., … Turkington, R. (2013). Coordinated distributed experiments: an emerging tool for testing global hypotheses in ecology and environmental science. *Frontiers in Ecology and the Environment*, *11*(3), 147–155. doi: 10.1890/110279

Fuller, R. M., Groom, G. B., Jones, A. R., & Thomson, A. G. (1993). *Countryside Survey 1990. Mapping the land cover of Great Britain using Landsat imagery: a demonstrator project in remote sensing. Final report.* (No. ITE Project No: T02052m5; p. 71). Retrieved from NERC/Institute of Terrestrial Ecology website: http://nora.nerc.ac.uk/id/eprint/4361/

Gelman, A., & Carlin, J. (2014). Beyond power calculations: assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science*, *9*(6), 641–651.

Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge University Press.

Harrell, F. E. (2015). *Regression Modelling Strategies* (2nd ed.). Switzerland: Springer International Publishing.

Heinze, G., Wallisch, C., & Dunkler, D. (2018). Variable selection – A review and recommendations for the practicing statistician. *Biometrical Journal*, *60*(3), 431–449. doi: 10.1002/bimj.201700067

Hough, M. N., & Jones, R. J. A. (1997). The United Kingdom Meteorological Office rainfall and evaporation calculation system: MORECS version 2.0-an overview. *Hydrology and Earth System Sciences Discussions*, *1*(2), 227–239.

Lemoine, N. P. (2019). Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. *Oikos*, *128*(7), 912–928. doi: 10.1111/oik.05985

Lindgren, F., Rue, H., & Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(4), 423–498.

Loken, E., & Gelman, A. (2017). Measurement error and the replication crisis. *Science*, *355*(6325), 584–585.

Mac Nally, R. (2000). Regression and model-building in conservation biology, biogeography and ecology: the distinction between–and reconciliation of–'predictive'and 'explanatory'models. *Biodiversity & Conservation*, *9*(5), 655–671.

Maskell, L. C., Smart, S. M., Bullock, J. M., Thompson, K., & Stevens, C. J. (2010). Nitrogen deposition causes widespread loss of species richness in British habitats. *Global Change Biology*, *16*(2), 671–679. doi: 10.1111/j.1365-2486.2009.02022.x

Payne, R. J., Dise, N. B., Field, C. D., Dore, A. J., Caporn, S. J., & Stevens, C. J. (2017). Nitrogen deposition and plant biodiversity: past, present, and future. *Frontiers in Ecology and the Environment*, *15*(8), 431–436. doi: 10.1002/fee.1528

Peters, R. H. (1991). *A critique for ecology*. Cambridge, UK: Cambridge University Press.

Phoenix, G. K., Emmett, B. A., Britton, A. J., Caporn, S. J. M., Dise, N. B., Helliwell, R., … Power, S. A. (2012). Impacts of atmospheric nitrogen deposition: responses of multiple plant and soil parameters across contrasting ecosystems in long-term field experiments. *Global Change Biology*, *18*(4), 1197–1215. doi: 10.1111/j.1365-2486.2011.02590.x

Phoenix, G. K., Hicks, W. K., Cinderby, S., Kuylenstierna, J. C. I., Stock, W. D., Dentener, F. J., … Ineson, P. (2006). Atmospheric nitrogen deposition in world biodiversity hotspots: the need for a greater global perspective in assessing N deposition impacts. *Global Change Biology*, *12*(3), 470–476. doi: 10.1111/j.1365-2486.2006.01104.x

Rackham, O. (1986). *The History of the Countryside*. London: Dent.

Rodwell, J. S. (Ed.). (1992). *British Plant Communities Volume 3. Grasslands and montane communities*. Cambridge, UK: Cambridge University Press.

RoTAP. (2012). *Review of Transboundary Air Pollution: Acidification, Eutrophication, Ground Level Ozone and Heavy Metals in the UK* [Contract Report to the Department for Environment, Food and Rural Affairs.]. Centre for Ecology & Hydrology.

Rowe, E. C., Jones, L., Dise, N. B., Evans, C. D., Mills, G., Hall, J., … Emmett, B. A. (2017). Metrics for evaluating the ecological benefits of decreased nitrogen deposition. *Biological Conservation*, *212*(Part B), 454–463. doi: 10.1016/j.biocon.2016.11.022

Rubin, D. B. (2004). Teaching statistical inference for causal effects in experiments and observational studies. *Journal of Educational and Behavioral Statistics*, *29*(3), 343–367.

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *71*(2), 319–392.

Shmueli, G. (2010). To explain or to predict? *Statistical Science*, 289–310.

Simkin, S. M., Allen, E. B., Bowman, W. D., Clark, C. M., Belnap, J., Brooks, M. L., … Waller, D. M. (2016). Conditional vulnerability of plant diversity to atmospheric nitrogen deposition across the United States. *Proceedings of the National Academy of Sciences*, *113*(15), 4086–4091. doi: 10.1073/pnas.1515241113

Smart, S. M., Clarke, R. T., van de Poll, H. M., Robertson, E. J., Shield, E. R., Bunce, R. G. H., & Maskell, L. C. (2003). National-scale vegetation change across Britain; an analysis of sample-based surveillance data from the Countryside Surveys of 1990 and 1998. *Journal of Environmental Management*, *67*(3), 239–254. doi: 10.1016/S0301-4797(02)00177-9

Smith, R. I., Fowler, D., Sutton, M. A., Flechard, C., & Coyle, M. (2000). Regional estimation of pollutant gas dry deposition in the UK: model description, sensitivity analyses and outputs. *Atmospheric Environment*, *34*(22), 3757–3777.

Soons, M. B., Hefting, M. M., Dorland, E., Lamers, L. P. M., Versteeg, C., & Bobbink, R. (2017). Nitrogen effects on plant species richness in herbaceous communities are more widespread and stronger than those of phosphorus. *Biological Conservation*, *212*, 390–397. doi: 10.1016/j.biocon.2016.12.006

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing Transparency Through a Multiverse Analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. doi: 10.1177/1745691616658637

Stephens, P. A., Buskirk, S. W., & del Rio, C. M. (2007). Inference in ecology and evolution. *Trends in Ecology & Evolution*, *22*(4), 192–197. doi: 10.1016/j.tree.2006.12.003

Stevens, C. J., Dise, N. B., Mountford, J. O., & Gowing, D. J. (2004). Impact of nitrogen deposition on the species richness of grasslands. *Science*, *303*(5665), 1876–1879.

Stevens, C. J., Dupré, C., Dorland, E., Gaudnik, C., Gowing, D. J. G., Diekmann, M., … Dise, N. B. (2011). Grassland species composition and biogeochemistry in 153 sites along environmental gradients in Europe. *Ecology*, *92*(7), 1544–1544. doi: 10.1890/11-0115.1

Stevens, C. J., Dupré, C., Gaudnik, C., Dorland, E., Dise, N., Gowing, D., … Diekmann, M. (2011). Changes in species composition of European acid grasslands observed along a gradient of nitrogen deposition. *Journal of Vegetation Science*, *22*(2), 207–215. doi: 10.1111/j.1654-1103.2010.01254.x

Stevens, C. J., Smart, S. M., Henrys, P., Maskell, L. C., Walker, K. J., Preston, C. D., … Emmett, B. A. (2011). *Collation of evidence of nitrogen impacts on vegetation in relation to UK biodiversity objectives* (JNCC Report No. 447; p. 187). Retrieved from JNCC website: http://archive.jncc.gov.uk/pdf/jncc447_web.pdf

Tanguy, M., Prudhomme, C., Smith, K., & Hannaford, J. (2018). Historical gridded reconstruction of potential evapotranspiration for the UK. *Earth System Science Data*, *10*(2), 951–968. doi: https://doi.org/10.5194/essd-10-951-2018

Tipping, E., Henrys, P. A., Maskell, L. C., & Smart, S. M. (2013). Nitrogen deposition effects on plant species diversity; threshold loads from field data. *Environmental Pollution*, *179*(Supplement C), 218–223. doi: 10.1016/j.envpol.2013.04.008

UK National Ecosystem Assessment. (2011). *The UK National Ecosystem Assessment Technical Report*. Cambridge: UNEP-WCMC.

Van der Eerden, L. J., Dueck, T. A., Berdowski, J. J. M., Greven, H., & Van Dobben, H. F. (1991). Influence of NH_3 and $(NH_4)_2SO_4$ on heathland vegetation. *Acta Botanica Neerlandica*, *40*(4), 281–296.

Whittingham, M. J., Stephens, P. A., Bradbury, R. B., & Freckleton, R. P. (2006). Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology*, *75*(5), 1182–1189. doi: 10.1111/j.1365-2656.2006.01141.x

Young, C. (2018). The Difference Between Causal Analysis and Predictive Models: Response to "Comment on Young and Holsteen (2017)." *Sociological Methods & Research*, 0049124118782542. doi: 10.1177/0049124118782542

Acknowledgements

We thank the authors of the original studies reanalysed for publishing their data or otherwise making it available to us. We also thank P. Henrys, S.M. Smart and L.C. Maskell for useful comments on the manuscript. The initial work that led to the current manuscript was funded by the Joint Nature Conservation Committee to CEH under project NEC06730 (National Plant Monitoring Scheme). This work was also supported by the Natural Environment Research Council award number NE/R016429/1 as part of the UK-SCAPE programme delivering National Capability.

Data accessibility

The data of Stevens et al. (2004) are available in Stevens, Dupré, Dorland et al. (2011) (DOI: 10.1890/11-0115.1). UK Countryside Survey data access policies can be read at https://countrysidesurvey.org.uk/content/data-access

Tables

Table 1. All covariates with their original spatial grain sizes, original ranges, and re-scalings as used in the reanalyses of Maskell et al. (2010) (MEA10) and Stevens et al. (2004) (SEA04) presented here.

Covariate	Supporting refs or sources	Original data grain size	Original range	Original units	Rescaled range where relevant	Rescaled units where relevant	Relevant reanalysis
Total Ndep estimated over moorland	Maskell et al. (2010); Smith et al. (2000)	5 × 5 km	4.9–40.0	kg ha^{-1} yr^{-1}	-	-	MEA10
Change (1970 to 1998) from peak Sdep estimated over moorland	Smith et al. (2000); L. Maskell (pers. comm.)	5 × 5 km	-5.36–0.00	Δ kg S ha^{-1} yr^{-1}	-	-	
Max. altitude	Maskell et al. (2010)	1 × 1 km	0–975	m	0.00–9.75	100 m	
Mean min. Jan. temp (1961-1999)	Maskell et al. (2010)	5 × 5 km	-8.16–0.08	°C	-	-	
Mean max. Jul. temp (1961-1999)	Maskell et al. (2010)	5 × 5 km	14.11–26.67	°C	-	-	
Mean annual precipitation (1961-1999)	Maskell et al. (2010)	5 × 5 km	554.33–3305.80	mm	2.22–13.22	250 mm	
Change in sheep numbers	Maskell et al. (2010)	2 × 2 km	-11.19–88.47	Δ sheep per year (1969 to 2000)	-1.12–8.85	Δ 10 sheep per year (1969 to 2000)	
Mean annual precipitation (1996-2006)	Stevens, Dupré, Dorland et al. (2011)	25 × 25 km	604.9–1773.3	mm	0.42–7.09	250 mm	SEA04: Model 1
Mean annual daily max. temp. (1996-2006)	Stevens, Dupré, Dorland et al. (2011)	25 × 25 km	11.5–14.6	°C	-	-	
Mean annual daily min. temp. (1996-2006)	Stevens, Dupré, Dorland et al. (2011)	25 × 25 km	4.2–8.1	°C	-	-	
Topsoil aluminium	Stevens, Duprè, Gaudnik et al. (2011)	Empirical plot data	11.60–1318.75	mg kg^{-1} dry soil	0.06–6.59	200 mg kg^{-1} dry soil	
Total Sdep estimated over grassland	Stevens et al. (2004); Smith et al. (2000)	5 × 5 km	3.20–13.44	kg ha^{-1} yr^{-1}	-	-	SEA04: Models 1 & 2
Total Ndep estimated over grassland	Stevens et al. (2004); Smith et al. (2000)	5 × 5 km	7.70–40.86	kg ha^{-1} yr^{-1}	-	-	
Topsoil pH	Stevens et al. (2004); Stevens, Duprè,	Empirical plot data	3.69–5.37	pH unit	7.38–10.74	0.5 pH unit	

	Gaudnik et al. (2011)						
Max. altitude	Stevens et al. (2004); Stevens, Duprè, Gaudnik et al. (2011)	Empirical plot data	15–500	m	0.15–5.00	100 m	
Grazing intensity	Stevens et al. (2004); Stevens, Duprè, Gaudnik et al. (2011)	Empirical plot data	Coded as low/ medium/ high	-	-	-	
Climate PC1	See methods	See methods	-4.90–3.43	-	-	-	SEA04: Model 2 only
Climate PC2	See methods	See methods	-1.47–1.20	-	-	-	
C:N	Stevens et al. (2004); Stevens, Duprè, Gaudnik et al. (2011)	Empirical plot data	13.34–30.58	topsoil mass ratio	-	-	
Slope	Stevens et al. (2004); Stevens, Duprè, Gaudnik et al. (2011)	Empirical plot data	0–60	°	0–6	10°	
Soil %N	Stevens et al. (2004); Stevens, Duprè, Gaudnik et al. (2011)	Empirical plot data	0.12–1.57	topsoil %N	-	-	
Soil moisture deficit (SMD)	Stevens et al. (2004); Hough & Jones (1997)	40 × 40 km	1.66–48.94	mm	0.17–4.89	10 mm	

Table 2. Exponentiated total nitrogen deposition regression coefficient medians, 2.5 and 97.5% quantiles for all models (all given to 2 decimal places). For example, for model MEA10, vascular plants only, a median value of 1.01 implies that a 1% gain of species richness per kg ha^{-1} yr^{-1} of total N deposition is highly compatible with the data; however, the 2.5% quantile value of 0.99 for this model also suggests that losses up to and around 1% are plausible (all values being conditional on model assumptions and data accuracy).

Model, dependent richness variable	**Spatial error structure**	**2.5% quantile**	**Median**	**97.5% quantile**	**Model location**
MEA10, vascular plants only	Mesh + square	0.99	1.01	1.03	Main text
MEA10, vascular plants and bryophytes	Mesh + square	0.99	1.00	1.02	SI2
SEA04 model 1, vascular plants only	Mesh + square	0.98	0.99	1.00	Main text
SEA04 model 2, vascular plants only	Mesh + square	0.98	0.99	1.01	Main text
SEA04 model 1, vascular plants and bryophytes	Mesh + square	0.98	0.99	1.01	SI2
SEA04 model 2, vascular plants and bryophytes	Mesh + square	0.98	0.99	1.01	SI2
MEA10, vascular plants only	Square only	0.97	0.98	1.00	SI3
MEA10, vascular plants only	Mesh only	0.98	0.99	1.01	SI3

Figures

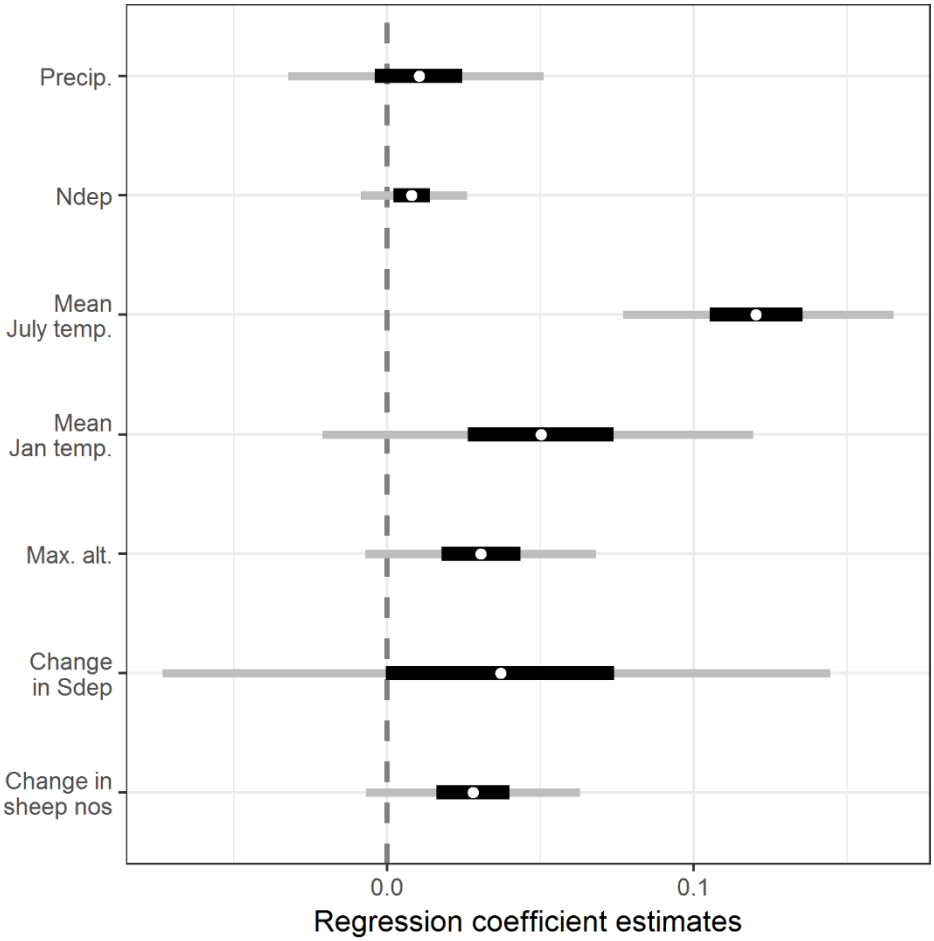

Figure 1. Estimated regression coefficients for the reanalysis of Maskell et al. (2010). The dependent variable was vascular plant species richness. White circles represent the posterior median estimate, black bars the posterior 50% credible interval, grey bars the posterior 95% credible interval. All covariates are described in Table 1.

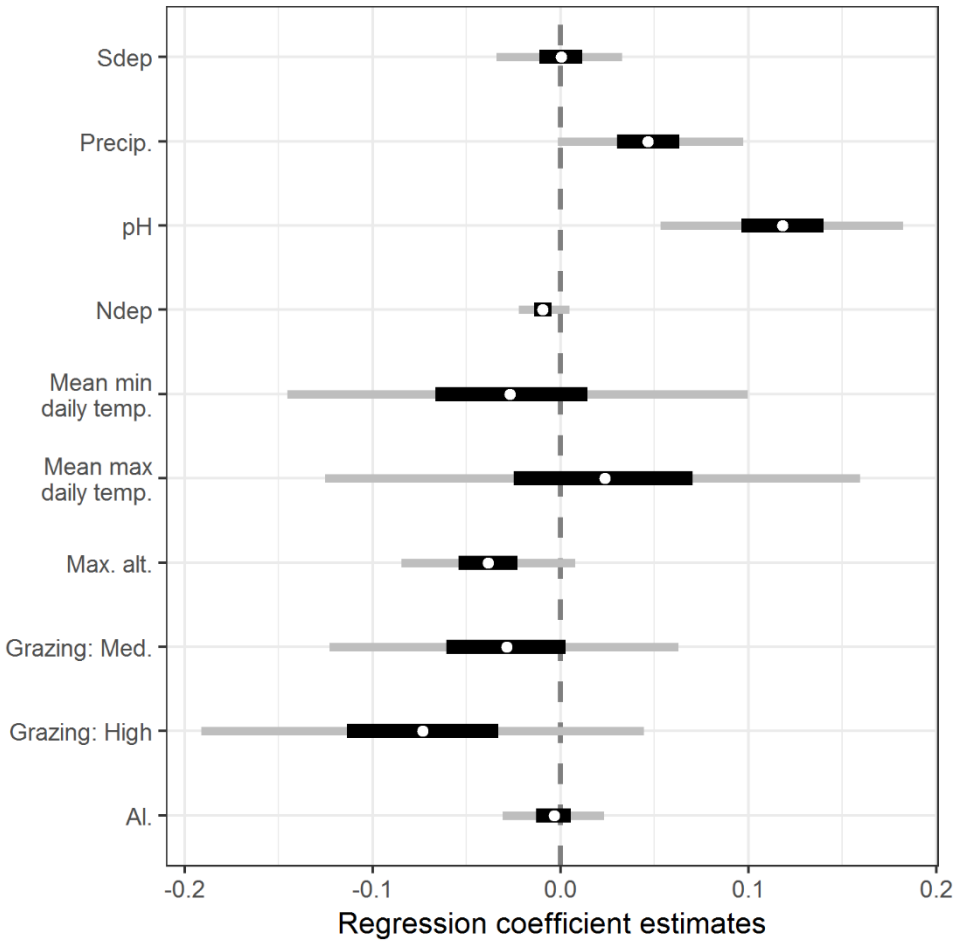

Figure 2. Estimated regression coefficients for the reanalysis of Stevens et al. (2004) using a reduced set of covariates chosen for their similar ecological status to the covariates used by Maskell et al. (2010), referred to in this paper as SEA04 model 1. The dependent variable was vascular plant species richness. White circles represent the posterior median estimate, black bars the posterior 50% credible interval, grey bars the posterior 95% credible interval. All covariates are described in Table 1.

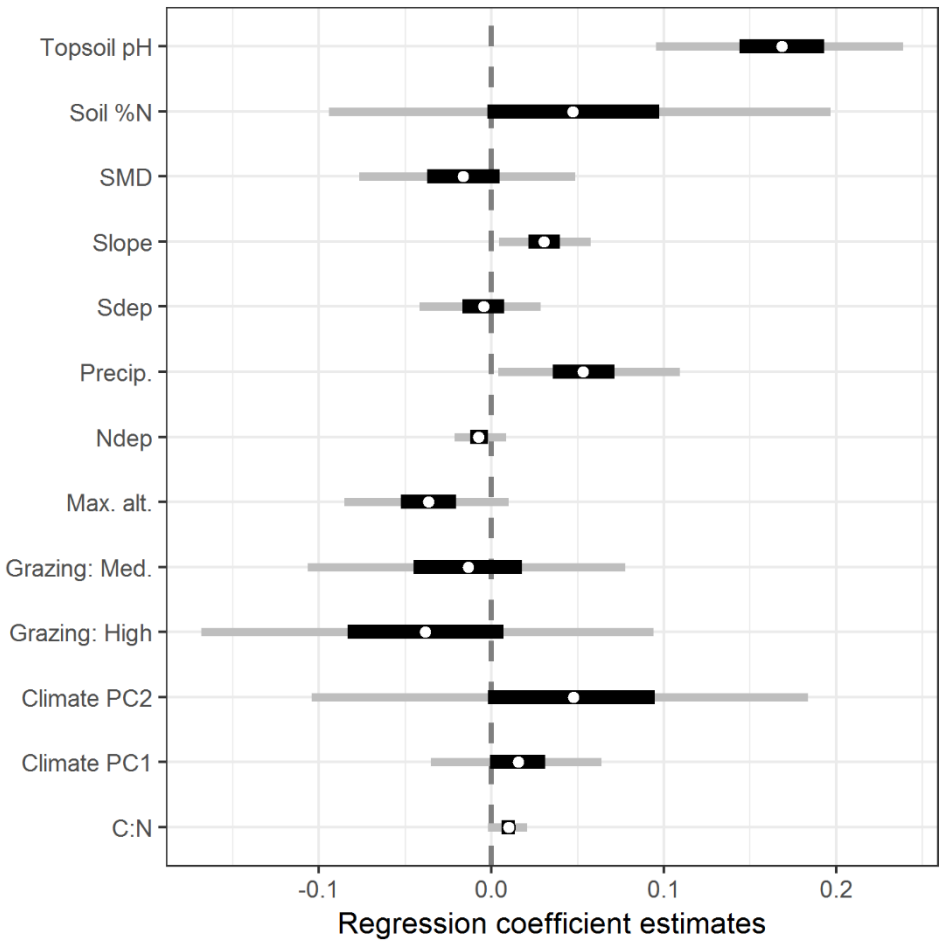

Figure 3. Estimated regression coefficients for the reanalysis of Stevens et al. (2004) using a set of covariates designed to match the original analysis of that paper as closely as possible, referred to in this paper as SEA04 model 2. The dependent variable was vascular plant species richness. White circles represent the posterior median estimate, black bars the posterior 50% credible interval, grey bars the posterior 95% credible interval. All covariates are described in Table 1.

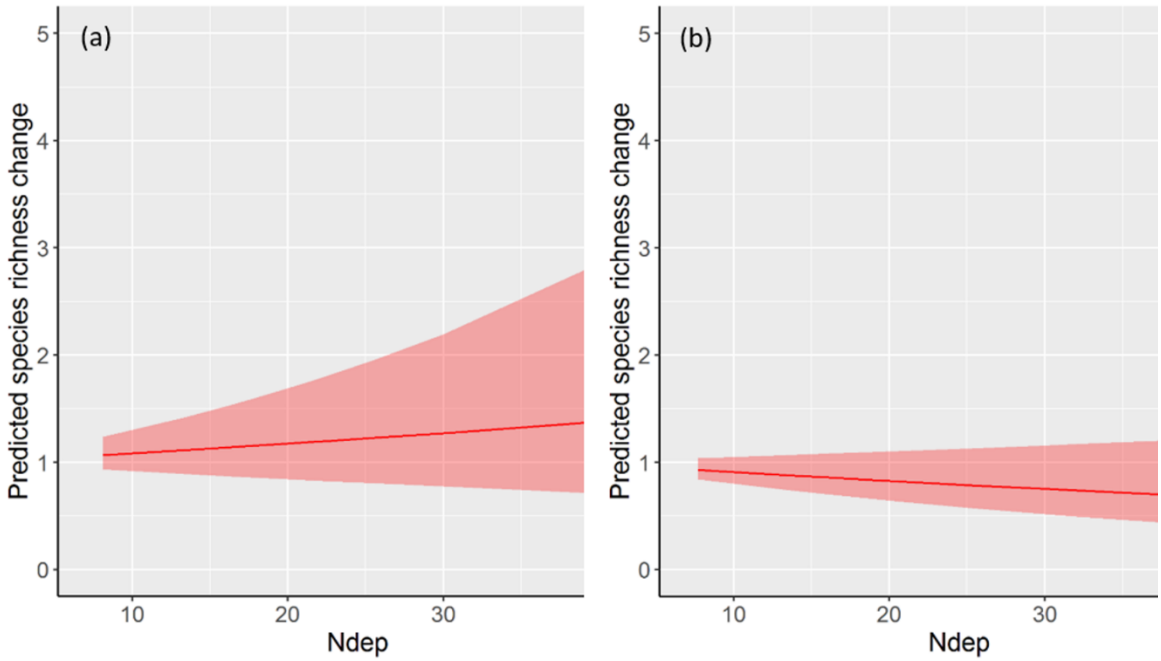

Figure 4. INLA model predictions for the marginal effect of total nitrogen deposition (Ndep; kg ha^{-1} yr^{-1}) on acid grassland estimated for (a) MEA10 and (b) SEA04 model 1; both the models used for these predictions used vascular plant species richness only as the dependent variable. Predictions were estimated for each 10th-percentile of the respective Ndep ranges covered by each study using the linear combinations option of INLA. Linear combinations were estimated for each point in the Ndep range with other covariates set to zero.

Supplementary Information

SI1: Spatial random fields

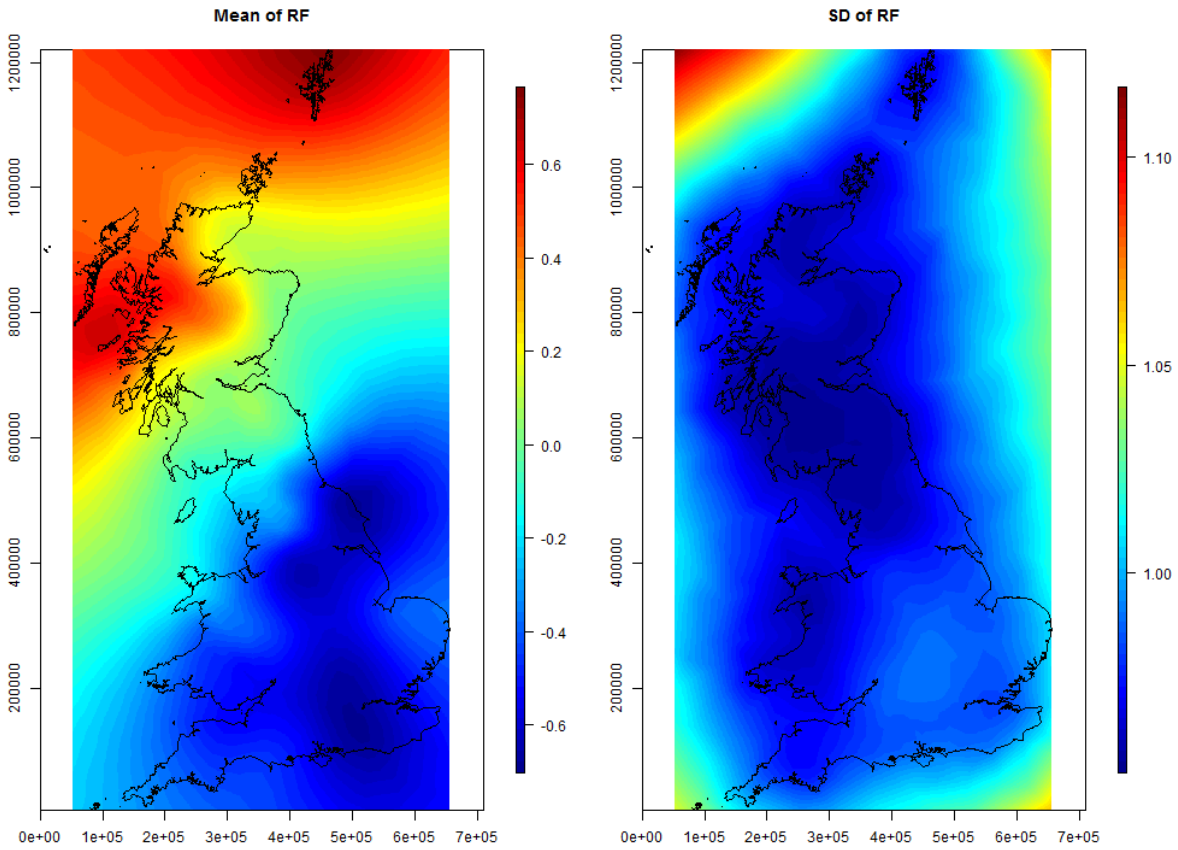

Figure A1.1. MEA10 model (vascular plant richness only). Mean and standard deviation of spatial random fields.

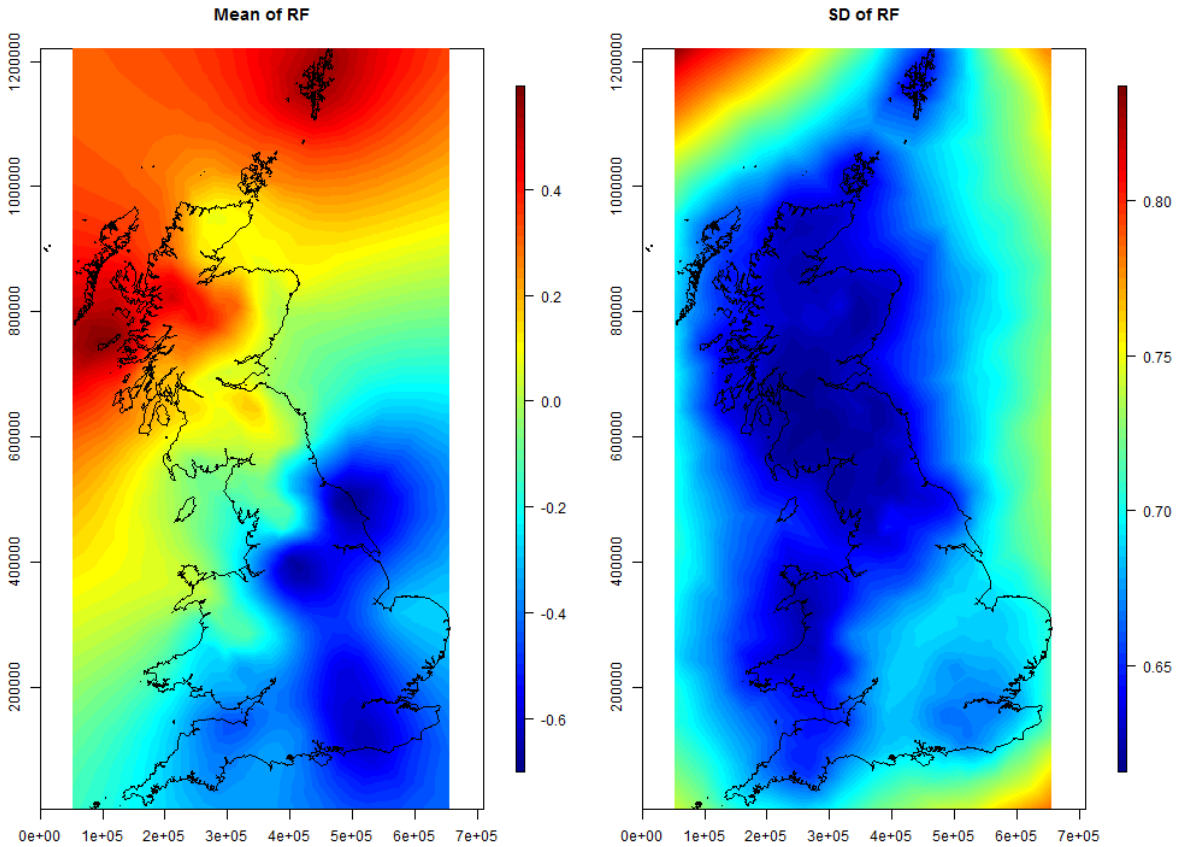

Figure A1.2. MEA10 model (vascular plant and bryophyte richness combined). Mean and standard deviation of spatial random fields.

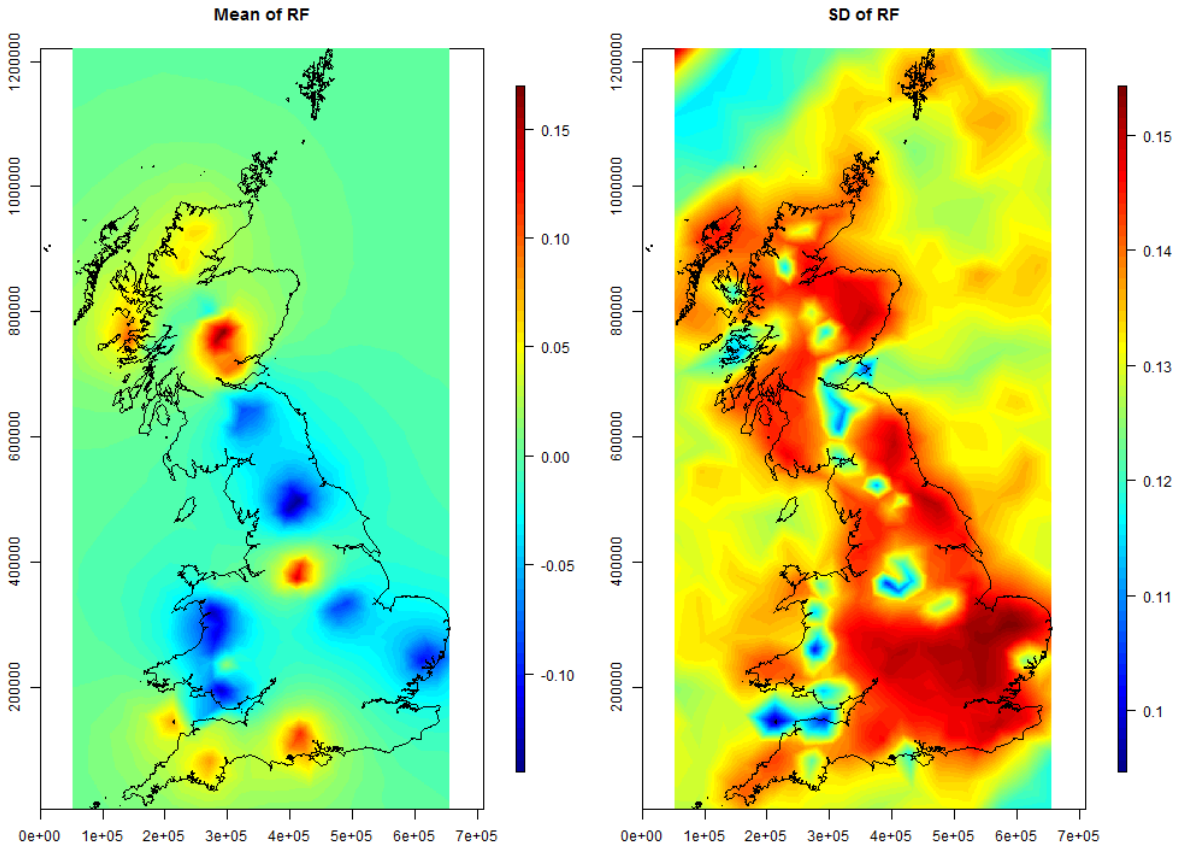

Figure A1.3. SEA04 model 1 (vascular plant richness only). Mean and standard deviation of spatial random fields.

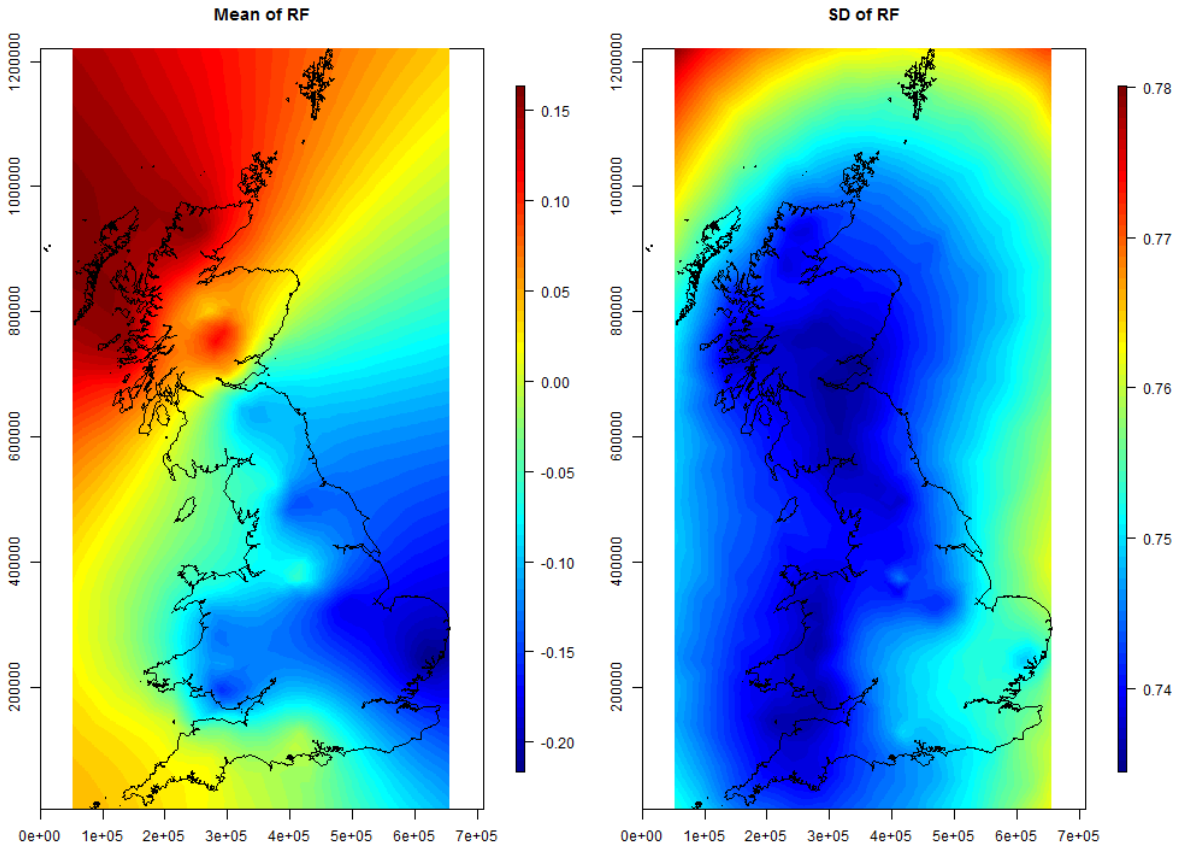

Figure A1.4. SEA04 model 1 (vascular plant and bryophyte richness combined). Mean and standard deviation of spatial random fields.

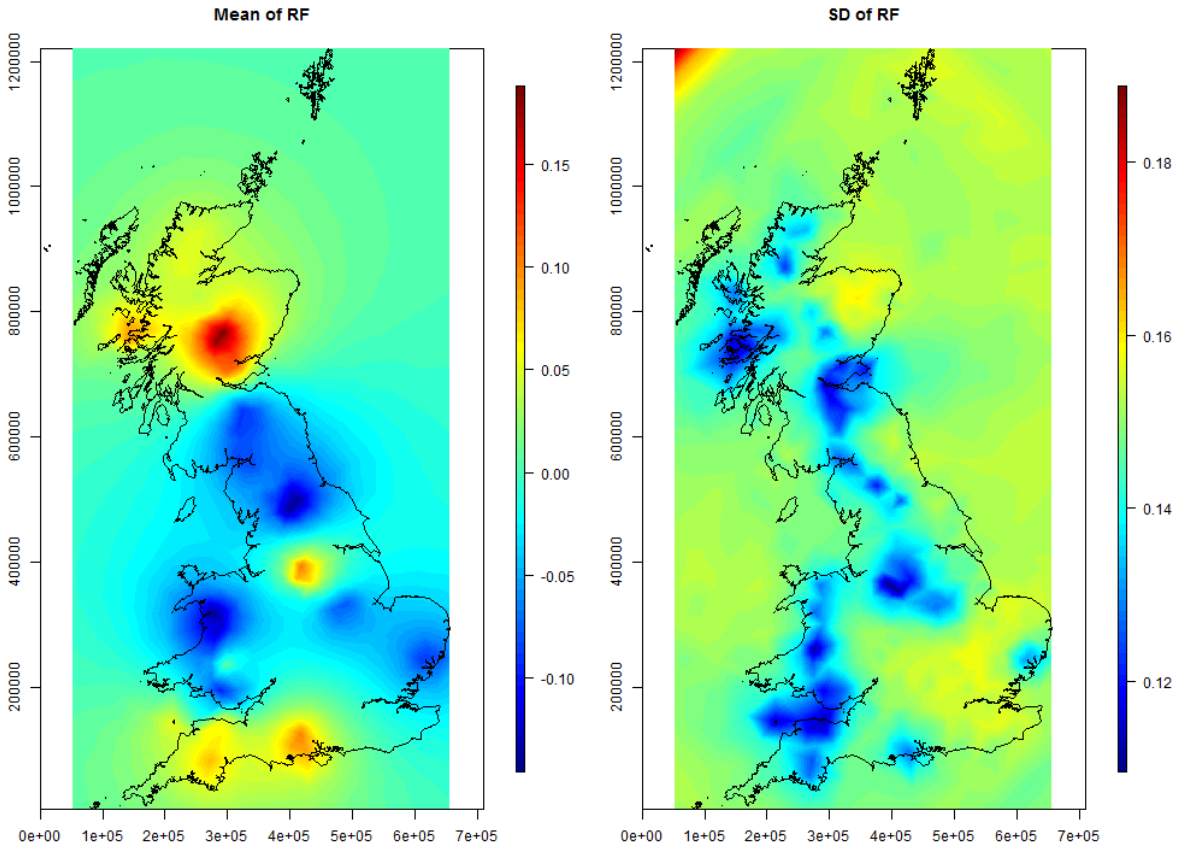

Figure A1.5. SEA04 model 2 (vascular plant richness only). Mean and standard deviation of spatial random fields.

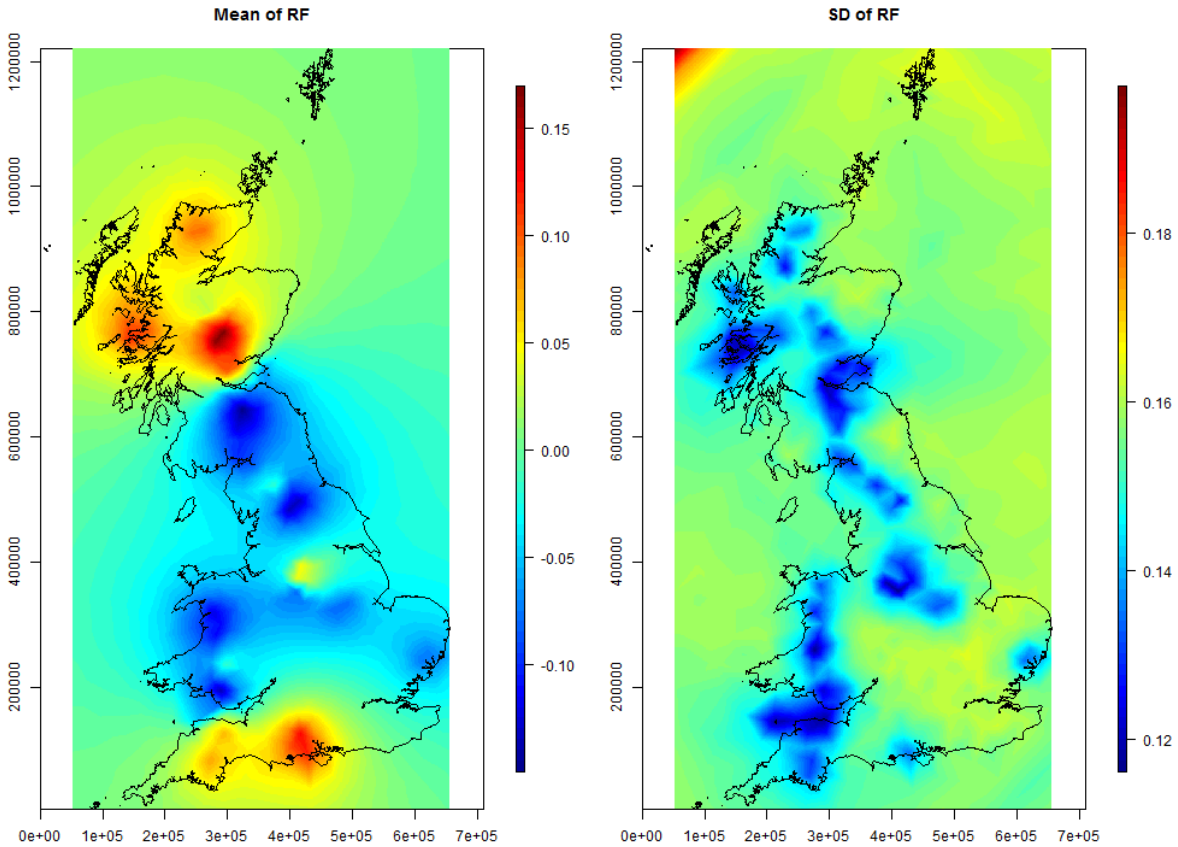

Figure A1.6. SEA04 model 2 (vascular plant and bryophyte richness combined). Mean and standard deviation of spatial random fields.

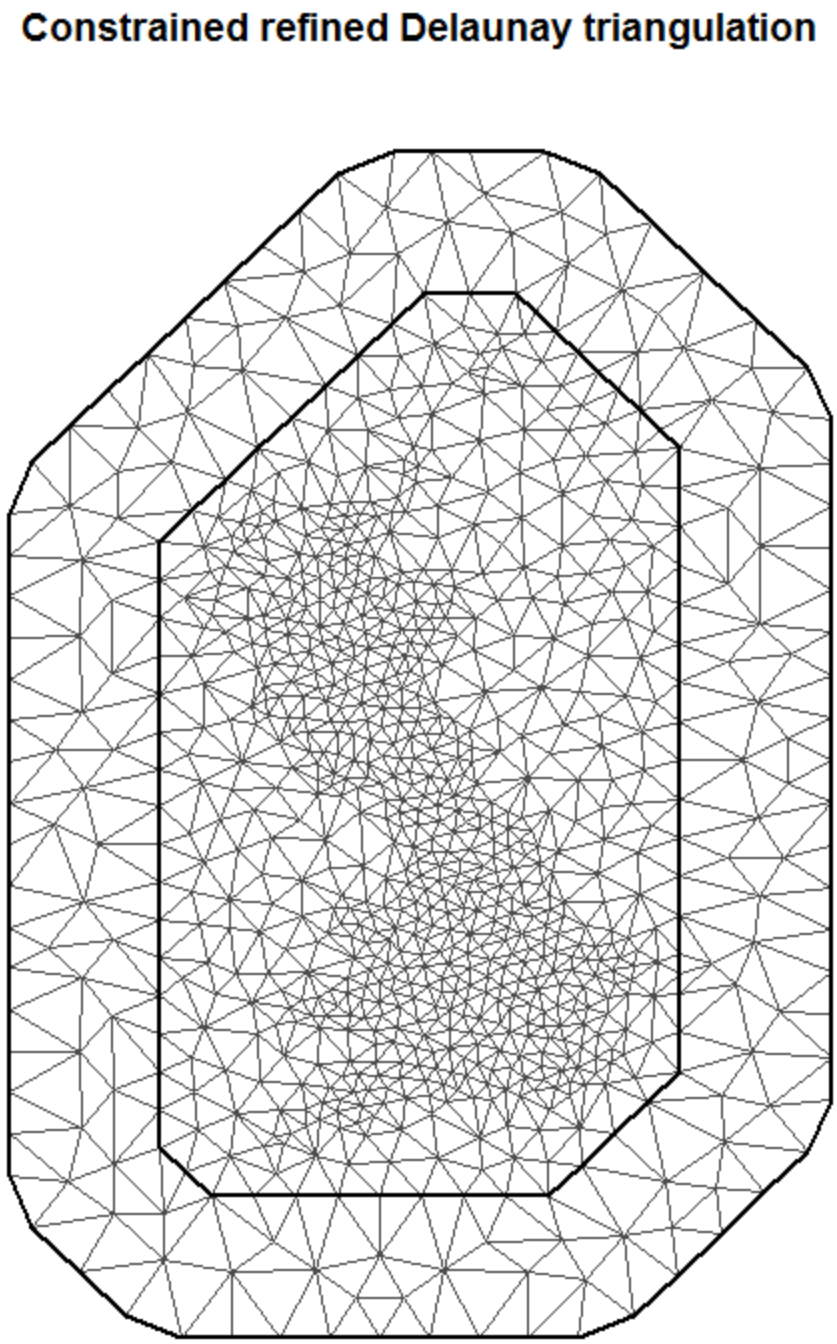

Figure A1.7. Constrained refined Delaunay triangulation of Britain over which the SPDE solution to the zero mean Matern Gaussian Markov random field (GMRF) is estimated.

SI2: Regression coefficients from analyses including bryophytes

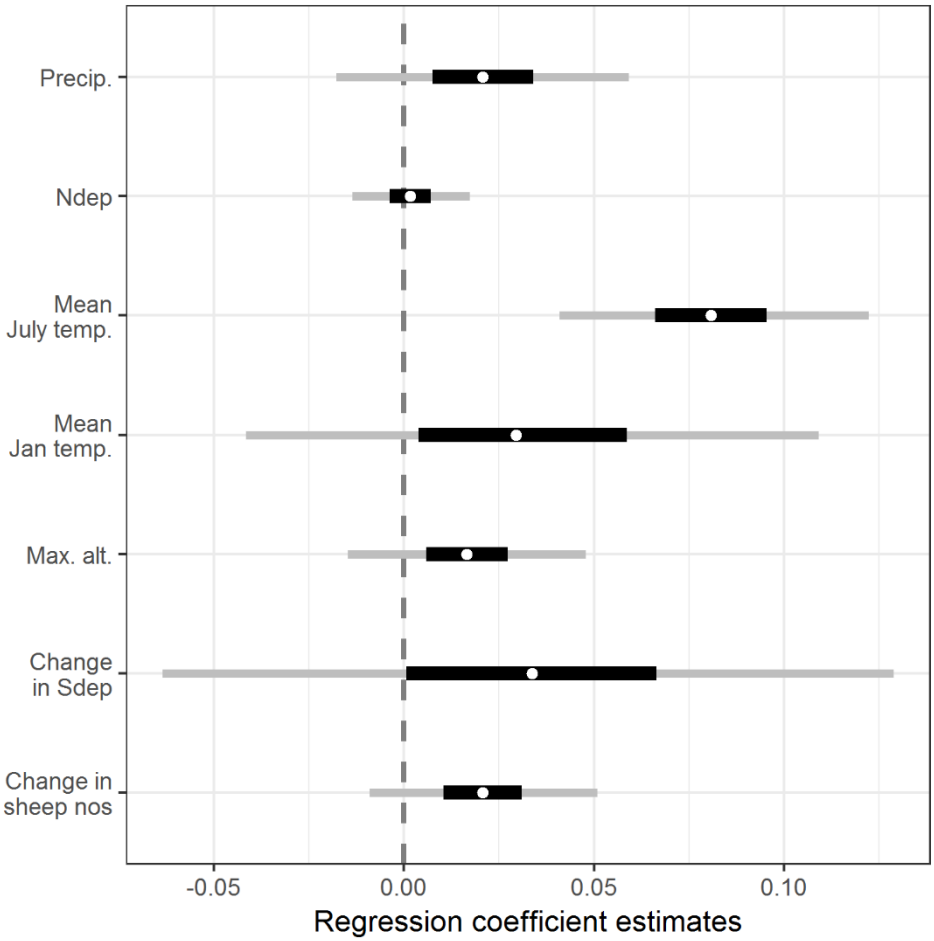

Figure A2.1. Estimated regression coefficients for the reanalysis of Maskell et al. (2010). The dependent variable was vascular plant plus bryophyte species richness. White circles represent the posterior median estimate, black bars the posterior 50% credible interval, grey bars the posterior 95% credible interval. All covariates are described in Table 1.

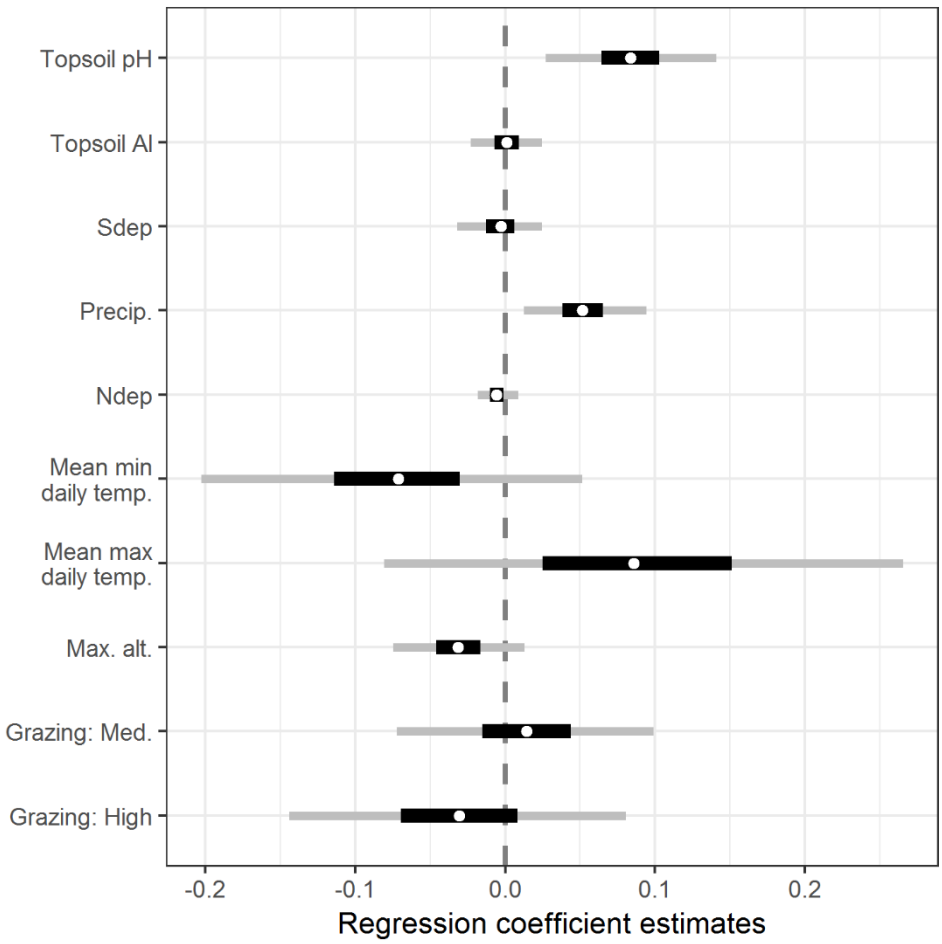

Figure A2.2. Estimated regression coefficients for the reanalysis of Stevens et al. (2004) using a reduced set of covariates chosen for their similar ecological status to the covariates used by Maskell et al. (2010), referred to in this paper as SEA04 model 1. The dependent variable was vascular plant plus bryophyte species richness. White circles represent the posterior median estimate, black bars the posterior 50% credible interval, grey bars the posterior 95% credible interval. All covariates are described in Table 1.

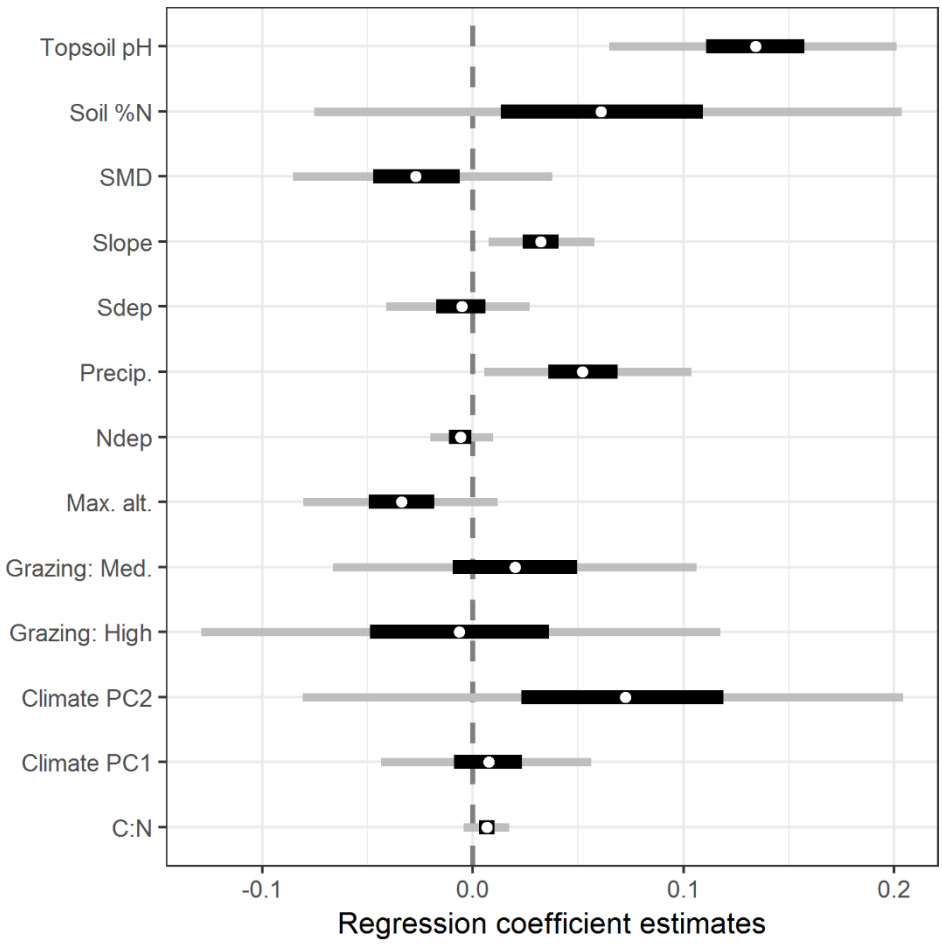

Figure A2.3. Estimated regression coefficients for the reanalysis of Stevens et al. (2004) using a set of covariates designed to match the original analysis of that paper as closely as possible, referred to in this paper as SEA04 model 2. The dependent variable was vascular plant species plus bryophyte richness. White circles represent the posterior median estimate, black bars the posterior 50% credible interval, grey bars the posterior 95% credible interval. All covariates are described in Table 1.

SI3: *Post hoc* models of MEA10 with partial accounting for spatial autocorrelation

Both models presented here use vascular plant richness only as the dependent variable, and were not the DIC-favoured models for these datasets.

1 km square-level random effect only

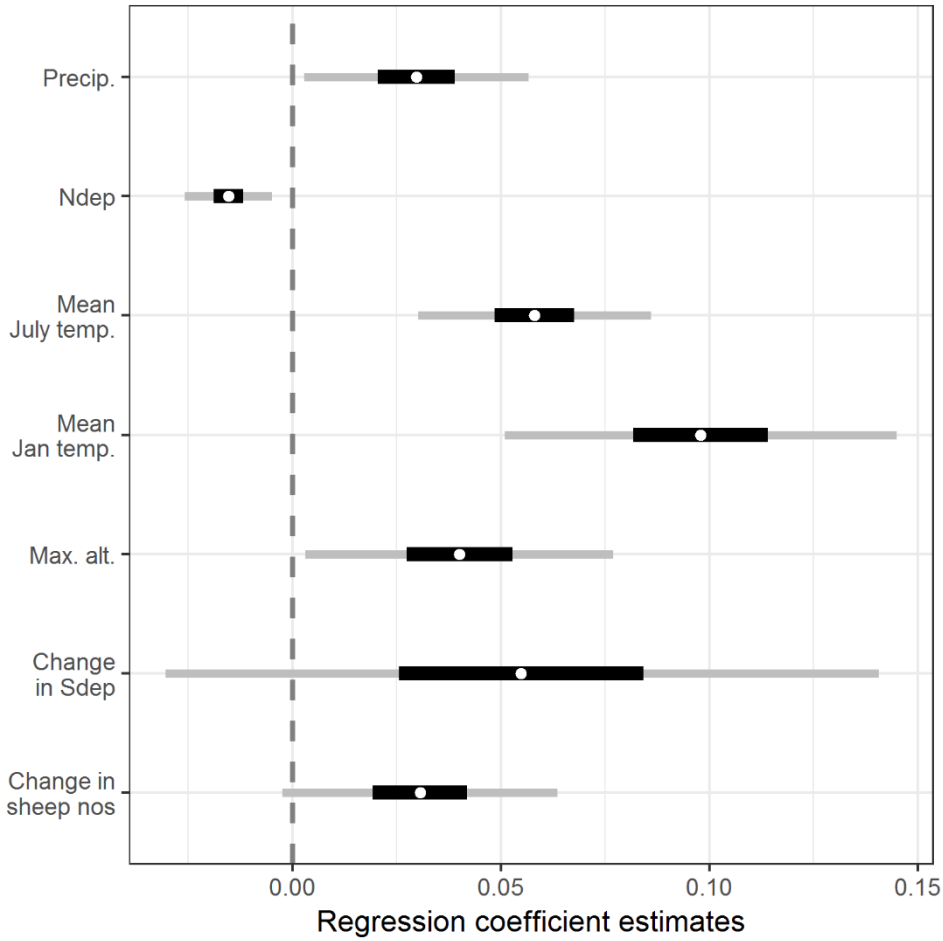

Figure A3.1. Estimated regression coefficients for the reanalysis of Maskell et al. (2010) using a 1 km square-level random effect only. The dependent variable was vascular plant species richness. White circles represent the posterior median estimate, black bars the posterior 50% credible interval, grey bars the posterior 95% credible interval. All covariates are described in Table 1.

Spatial mesh only

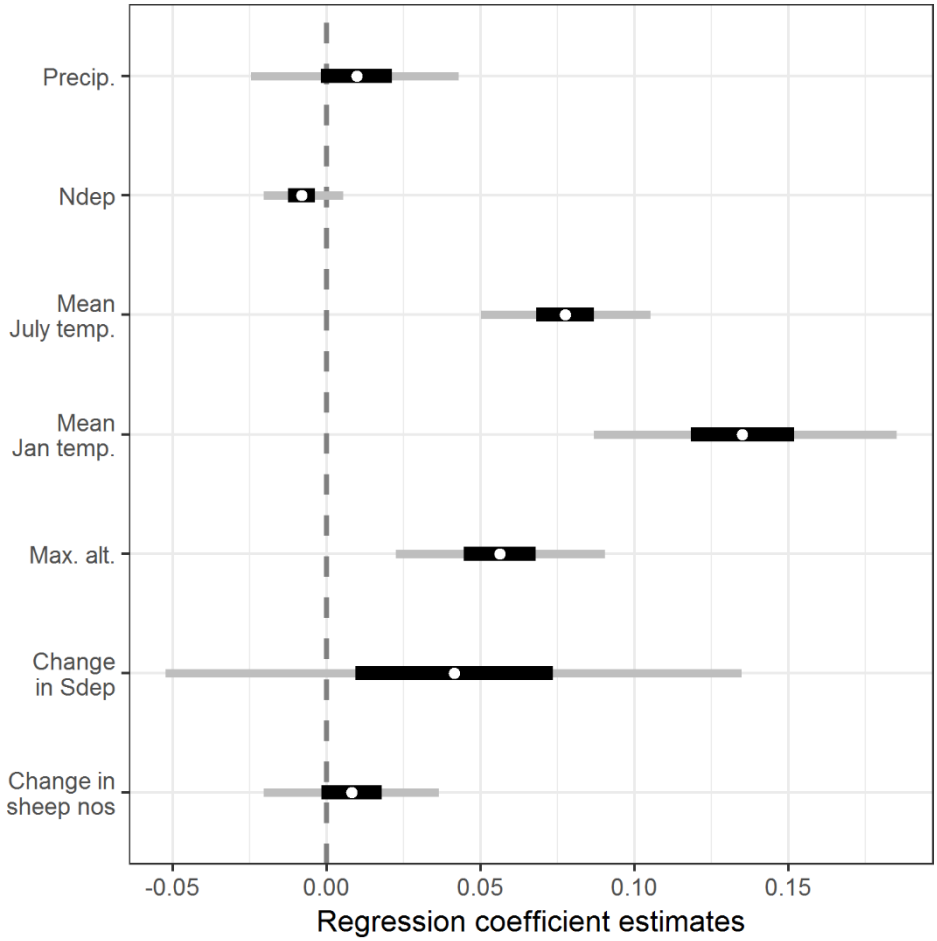

Figure A3.2. Estimated regression coefficients for the reanalysis of Maskell et al. (2010) using the spatial mesh only. The dependent variable was vascular plant species richness. White circles represent the posterior median estimate, black bars the posterior 50% credible interval, grey bars the posterior 95% credible interval. All covariates are described in Table 1.