

 Open access • Journal Article • DOI:10.1111/ECOG.01129

## **Biotic interactions boost spatial models of species richness** — [Source link](#)

[Heidi K. Mod](#), [Peter Christiaan le Roux](#), [Peter Christiaan le Roux](#), [Antoine Guisan](#) ...+1 more authors

**Institutions:** [University of Helsinki](#), [University of Pretoria](#), [University of Lausanne](#)

**Published on:** 01 Sep 2015 - [Ecography](#) (Wiley)

**Topics:** [Species richness](#), [Vegetation](#) and [Plant community](#)

Related papers:

- [The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling](#)
- [The importance of biotic interactions for modelling species distributions under climate change](#)
- [Species Distribution Models: Ecological Explanation and Prediction Across Space and Time](#)
- [Predicting species distribution: offering more than simple habitat models.](#)
- [Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model \(JSDM\)](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/biotic-interactions-boost-spatial-models-of-species-richness-2ohymrw3w0>

# Biotic interactions boost spatial models of species richness

Heidi K. Mod<sup>1\*</sup>, Peter C. le Roux<sup>1,2</sup>, Antoine Guisan<sup>3,4</sup> and Miska Luoto<sup>1</sup>

<sup>1</sup> Dept of Geosciences and Geography, Univ. of Helsinki, Univ. of Helsinki, Finland

<sup>2</sup> Dept of Plant Science, Univ. of Pretoria, Private Bag X20, Pretoria, South Africa

<sup>3</sup> Dept of Ecology and Evolution, Univ. of Lausanne, Biophore, Lausanne, Switzerland

<sup>4</sup> Inst. of Earth Sciences, Univ. of Lausanne, Géopolis, Lausanne, Switzerland

\*Email: Heidi K. Mod (heidi.mod@helsinki.fi)

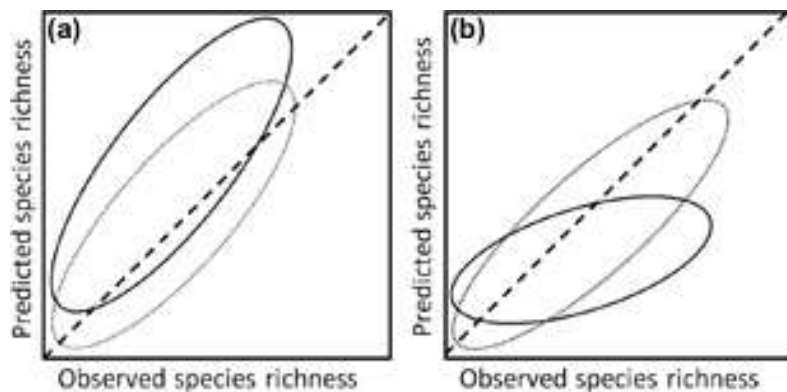
## Abstract

Biotic interactions are known to affect the composition of species assemblages via several mechanisms, such as competition and facilitation. However, most spatial models of species richness do not explicitly consider inter-specific interactions. Here, we test whether incorporating biotic interactions into high-resolution models alters predictions of species richness as hypothesised. We included key biotic variables (cover of three dominant arctic-alpine plant species) into two methodologically divergent species richness modelling frameworks – stacked species distribution models (SSDM) and macroecological models (MEM) – for three ecologically and evolutionary distinct taxonomic groups (vascular plants, bryophytes and lichens). Predictions from models including biotic interactions were compared to the predictions of models based on climatic and abiotic data only. Including plant–plant interactions consistently and significantly lowered bias in species richness predictions and increased predictive power for independent evaluation data when compared to the conventional climatic and abiotic data based models. Improvements in predictions were constant irrespective of the modelling framework or taxonomic group used. The global biodiversity crisis necessitates accurate predictions of how changes in biotic and abiotic conditions will potentially affect species richness patterns. Here, we demonstrate that models of the spatial distribution of species richness can be improved by incorporating biotic interactions, and thus that these key predictor factors must be accounted for in biodiversity forecasts.

Global environmental change requires enhanced predictions of how biotic and abiotic conditions impact biodiversity. As currently implemented, the two frameworks most commonly used for modelling species richness – stacked species distribution (SSDM) and macroecological models (MEM; Ferrier and Guisan 2006) – are fundamentally limited. Their main weakness is that neither routinely considers biotic interactions as a key community assembly process (Pearson and Dawson 2003, Soberón and Peterson 2005, Kearney and Porter 2009, Wisz et al. 2013). These correlative models are thus a priori critically flawed, both in theory and in practice, for addressing one of ecology's core questions (Brown 1995, Gaston 2000): how does biodiversity vary through space and time? Answering this question

is a key to addressing the impact of global environmental change on biodiversity and ecosystems. For an accurate answer, both frameworks, SSDM and MEM, need modifications as in their current forms they neglect these important drivers that impact on biodiversity.

The SSDM framework creates individual species distribution predictions and subsequently models total species richness by blindly stacking the predicted species occurrences (Ferrier and Guisan 2006). Stacking predicted species occurrences (based on threshold values for transforming probabilities of occurrence to presence–absence values), however, consistently overpredicts species richness (Newbold et al. 2009, Dubuis et al. 2011, Pottier et al. 2013; Fig. 1a, although see Calabrese et al. 2014 for stacking probabilities of species occurrence to restrict the overprediction of SSDMs). A proposed explanation for this systematic overprediction in stacked species richness predictions is that the models do not account for biotic interactions, which act as filters determining which subset of species co-occur in an area under certain environmental conditions (Guisan and Rahbek 2011). By omitting biotic interactions, stacked species distribution models are built under the expectation that no intrinsic environmental carrying capacity exists which limits the number of species that can coexist (Dubuis et al. 2011). The inclusion of biotic interactions is thus expected to reduce the distribution of some species (although facilitation may also enlarge a species distributional range), and thus limit the number of co-occurring species (i.e. lowering species richness) when stacking the predicted occurrences of individual species. In contrast, the MEM models species richness values directly through a single model that implicitly considers the environment to be limiting the number of species (Ferrier and Guisan 2006), sacrificing all information about community composition and species identities (Dubuis et al. 2011). Therefore, MEM cannot distinguish between sites that have similar abiotic conditions but very different biotic composition. In contrast to SSDM, MEM tends to overpredict richness in species-poor sites and underpredict richness in species-rich sites (Newbold et al. 2009, Dubuis et al. 2011; Fig. 1b). SSDM and MEM frameworks thus both have limitations as currently implemented. Indeed, these limitations appear to result from a common cause in both cases: biotic interactions remain unquantified and treated as a uniform neutral background in the models.



**Figure 1.** Hypothesized effect of inclusion of biotic interactions into species richness models for (a) stacked species distribution modelling (= SSDM) and (b) macroecological modelling (= MEM). The solid ellipse represents the relationship between predicted and observed species richness, generally reported from conventional models. The dotted ellipse represents the same relationship as expected after the explicit inclusion of key biotic interactions: for SSDM the inclusion of biotic interactions is expected to filter the number of co-occurring species under certain environmental conditions (i.e. reflecting an inherent environmental carrying capacity) and thus, in general, reduce species richness. In MEM, the inclusion of biotic interactions assists to identify the areas with similar abiotic, but differing biotic conditions, and thus correct the bias in both end of the gradient. A 1:1 relationship represents a perfect prediction, i.e. predicted values match the observed values.

Biotic interactions, both positive and negative, are known to shape species realized environmental niches via several mechanisms, including competition, mutualistic interactions and facilitation, especially at fine scales (Callaway 1995, Aerts 1999, Anderson et al. 2002, Case et al. 2005, Araújo and Guisan 2006). As observed species occurrence patterns are determined by both abiotic and biotic drivers, these two components must both be accounted for in species richness models. Indeed, recent species distribution model (SDM) studies have started incorporating biotic interactions into models for individual species, showing how biotic interactions can reduce bias in occurrence predictions (Davis et al. 1998, Araújo and Luoto 2007, Heikkinen et al. 2007, Meier et al. 2010, Pellissier et al. 2010, le Roux et al. 2012, 2013c, Giannini et al. 2013, González-Salazar et al. 2013). Consequently, it is equally likely that, through the accumulated impacts on individual species, biotic interactions can also strongly shape species assemblages and community richness patterns (Lortie et al. 2004, Guisan and Thuiller 2005, Gotelli et al. 2010, Kissling et al. 2012, Thuiller et al. 2013, Wisz et al. 2013).

We therefore hypothesize that explicitly incorporating the cover of dominant species as proxies for biotic interactions should improve predictions of species richness, both for SSDM and MEM (following the approach of le Roux et al. 2014). These frameworks differ considerably in their underlying theoretical basis, methodology and working assumptions (Ferrier and Guisan 2006, Dubuis et al. 2011, Guisan and Rahbek 2011). As a consequence, the effects of incorporating biotic interactions in species richness predictions may also differ between the two modelling frameworks. We expect that for SSDM, considering biotic interactions will reduce overprediction at all richness levels (Guisan and Rahbek 2011; Fig. 1a), while in MEM the inclusion of biotic interactions will correct bias at both ends of the richness gradient (Fig. 1b). Thus, for both frameworks, the inclusion of biotic interactions is expected to provide predictions closer to observed species richness (i.e. 1:1 line in Fig. 1), which, if verified, would demonstrate the importance of biotic interactions on predictions of species richness.

We test these hypotheses by incorporating the cover of three dominant arctic-alpine plant species (*Empetrum nigrum* ssp. *hermaphroditum*, *Betula nana*, *Betula pubescens* ssp. *czerepanovii*; nomenclature following Hämet-Ahti et al. 1998), which are previously identified as important biotic predictors in the study area (le Roux et al. 2013c, 2014), into high-resolution SSDM and MEM species richness models based on climatic (mean temperature of coldest quarter, growing degree days, water balance) and local abiotic (soil quality, solar radiation and soil wetness) predictor variables. To assess the impact of including biotic interactions on the model's prediction bias, we averaged the cross-validated results of three statistical techniques (generalized linear models (GLM), generalized additive models (GAM) and generalized boosting methods (GBM)) for both species richness modelling frameworks (Araújo and New 2007, Elith and Graham 2009, Guisan and Rahbek 2011), and ran the analysis for three ecologically and evolutionary distinct taxa: vascular plants, bryophytes and lichens. We used a large dataset ( $n = 1080$ ) from arctic-alpine northern Fennoscandia, a system characterized by broad environmental gradients with relatively few interacting species (le Roux et al. 2013c). It therefore represents a powerful study system to test the importance of biotic interactions on species richness (Wisz et al. 2013).

## Material and methods

### Study area and data collection

The study area is situated in north-western Finland and Norway (ca 69°N, 20°50'E) with an annual mean temperature of  $-1.9^{\circ}\text{C}$  (January  $-12.9^{\circ}\text{C}$ , July  $11.2^{\circ}\text{C}$ ) at the nearby Kilpisjärvi meteorological station (Finnish Meteorological Inst.). In the study area vegetation consists of arctic- alpine tundra with some boreal features at lower elevations (Virtanen et al. 2010). Species data were collected from 270 sites, along 18 elevational transects at 20 m intervals during 2008–2011. Four  $1\text{ m}^2$  plots, located 5 m from the centre of each site, were surveyed, resulting in a total of 1080 plots for analysis. Maximum distance between sites was 80 km, with altitude ranging between 460 and 1360 m a.s.l. In each plot, the cover and identity of all vascular plants, bryophytes and lichens were recorded (see Virtanen et al. 2010, le Roux et al. 2012 for more details). Individuals were identified to species level, with the exception of a few groups that could not be distinguished reliably or have unresolved taxonomy ( $\sim 3.5\%$  of records; including *Taraxacum* spp. some *Bryum* and *Stereocaulon* spp.). Cover data was transformed to presence–absence data (species noted as present if cover  $> 0\%$ ) and presences were summed to determine species richness per plot for each taxon separately.

Environmental predictor variables for modelling were selected from three different groups of variables: climatic, local abiotic and local biotic, based on previous studies in this area (Sormunen et al. 2011, le Roux et al. 2012, 2013b; see also Guisan and Zimmermann 2000). The three climatic variables used were mean air temperature of the coldest quarter (TCQ: December–February), growing degree days (GDD3: sum of the days when air temperature  $> 3^{\circ}\text{C}$ ) and water balance (WAB: ratio of precipitation to evaporation). Temperature- and moisture-related climatic variables represent broad-scale environmental drivers that have a strong impact on vegetation, although their variability, and therefore potential explanatory power, at finer scales is more moderate (Pearson et al. 2002, Virtanen et al. 2006). The variables were downscaled for each  $1\text{ m}^2$  plot following Aalto et al. (2014) using a digital elevation model (DEM; National Land Survey of Finland) with a resolution of  $5 \times 5\text{ m}$ . The method provides adequate climatic data by accounting for, in addition to geographic location, the effects of topography and land cover when modelling temperature. Climatic data were derived from 61 weather stations, which cover the northern parts of Fennoscandia (Finnish Meteorological Inst., Norwegian Meteorological Inst., Swedish Meteorological and Hydrological Inst. and Abisko Scientific Research Station) for the normal period 1971–2000 (most recent normal period from which climatic data from all 61 weather stations were available; see comparison for normal period 1981–2010 for eight weather stations in Supplementary material Appendix 1).

Local abiotic variables included in models were soil quality (SQ), soil wetness (WET) and solar radiation (SOL), which were derived using GIS techniques in ArcGIS (ESRI 2012). Soil quality reflects patterns in bedrock composition, which exert considerable influence over the distribution of calciphilous and acidophilous plant species (Baack et al. 2006) and affects nutrient availability for generalist species. Bedrock composition was estimated from the proportion of calcareous (nutrient-rich) and siliceous (nutrient-poor) bedrock with data extracted from a bedrock map (Korsman 1997) at a resolution of 100 m and interpolated using the ArcGIS interpolation tool (ESRI 2012). Soil moisture is a key determinant of species richness in arctic and alpine areas, and can show high heterogeneity over short distances due to finescale variation in topography and physical soil properties (le Roux et al. 2013a). Relative soil moisture at each plot was estimated by calculating the topographic

wetness index (TWI; Beven and Kirkby 1979) from the same DEM using the ArcView Spatial Analyst extension (ESRI 2012). TWI was calculated as:

$$\text{TWI} = \ln [\text{upslope catchment area per unit width orthogonal to the flow direction} \times \tan(\text{slope angle})]$$

with slope angle measured in radians. Variation in the fine-scale intensity of incident solar radiation is strongly determined by topographic variables, and therefore often varies considerably in rugged high latitude and altitude environments, often creating pronounced fine-scale soil temperature gradients (Billings and Mooney 1968, Dymond and Johnson 2002). Potential annual solar radiation ( $\text{Mj cm}^{-2} \text{ yr}^{-1}$ ) for each plot was calculated from the DEM using Solar Analyst extension in ArcView (McCune and Keon 2002, ESRI 2012). This value reflects the maximum potential radiation, assuming clear skies, based on slope aspect and angle.

Biotic variables added to the species richness models comprised the cover of *Empetrum nigrum* ssp. *hermaphroditum*, *Betula nana* and *Betula pubescens* ssp. *czerepanovii* which are dominant species in the area with demonstrated impacts on the distribution of other species (Grytnes et al. 2006, Pellissier et al. 2010, le Roux et al. 2014), also from other taxa, i.e. bryophytes and lichens (Cornelissen et al. 2001, Pajunen et al. 2011). *Empetrum nigrum* is an evergreen dwarf shrub that is a strong competitor in low-nutrient, acidic arctic-alpine ecosystems due to its allelopathy and dense growth form (Tybirk et al. 2000, Aerts 2010). This species has positive effects on some co-occurring species by providing shelter from the wind, trapping protective snow cover and creating a beneficial microclimate (Williams 1988, Shevtsova et al. 1995, Mod et al. 2014). *Betula nana* is a deciduous dwarf shrub, and like *E. nigrum*, one of the dominant species in the study area (Eskelinen et al. 2009). It is hypothesized to be a potential competitor, but as a tall and dense dwarf-shrub it also provides sheltered and shaded conditions for short stature species (Pellissier et al. 2010). *Betula pubescens* forms the tree line in the region at ca 600–700 m a.s.l. (Eskelinen et al. 2009), and strongly buffers abiotic conditions, favouring many boreal species and therefore affecting the distribution of some species at lower elevations.

## Statistical analysis

To test our hypothesis that the inclusion of biotic interaction improves predictions of species richness, we applied the two prevailing species richness modelling frameworks, stacked species distribution models (SSDM) and macroecological modelling (MEM) using three sets of different explanatory variables. As a result three different models were run (Eq. 1, 2, and 3, following Zimmermann et al. 2009 and le Roux et al. 2013c):

$$\text{Climate model: TCQ} + \text{GDD} + \text{WAB}(1)$$

$$\text{Abiotic model: Climate model} + \text{SQ} + \text{WET} + \text{SOL}(2)$$

$$\text{Biotic model: Abiotic model} + \text{E. nigrum} + \text{B. nana} + \text{B. pubescens}(3)$$

To account for different modelling algorithms giving different predictions, three alternative state-of-the-art techniques were used: generalized linear models (GLMs), generalized additive models (GAMs) and generalized boosting method (GBM; Elith et al. 2008, Franklin and Miller 2009, Dubuis et al. 2011, Guisan and Rahbek 2011). All three modelling

algorithms were used both for SSDM and MEM and repeated for all models and species groups. All analyses were conducted in R (R Development Core Team). GLMs were fitted using the stats package (R Development Core Team), and GAMs using the mgcv package (Wood 2011). For GAMs, the initial degrees of smoothness for each univariate term were set to four. GBMs were fitted using functions from the gbm and dismo packages (Ridgeway 2005, Elith et al. 2008), setting interaction depth to six, learning rate to 0.001 (reduced for species where inadequate trees were calculated) and bagging fraction to 0.75. In SSDM, for single SDMs all three modelling algorithms were used, applying a logit link function and assuming a binomial distribution of errors for GLM and GAM and a Bernoulli distribution for GBM. Predicted probabilities of occurrence were converted to presence/absence predictions using the threshold value maximizing sensitivity and specificity (Liu et al. 2005, Levinsky et al. 2013). For the MEM, a logarithmic link function was applied assuming a Poisson distribution of errors for all modelling algorithms.

To account for different algorithms giving different predictions, we averaged their results by using ensemble averaging methods for both frameworks (Ferrier and Guisan 2006, Heikkinen et al. 2006, Araújo and New 2007, Algar et al. 2009). For SSDM, committee averaging was used, which is an ensemble forecasting method that assigns a presence prediction only when the majority of models (i.e.  $\geq 2$ ) predicts a presence (Gallien et al. 2012). For MEM, to average the results based on different algorithms, their arithmetic mean was calculated (consensus method) (Araújo and New 2007, Dubuis et al. 2011).

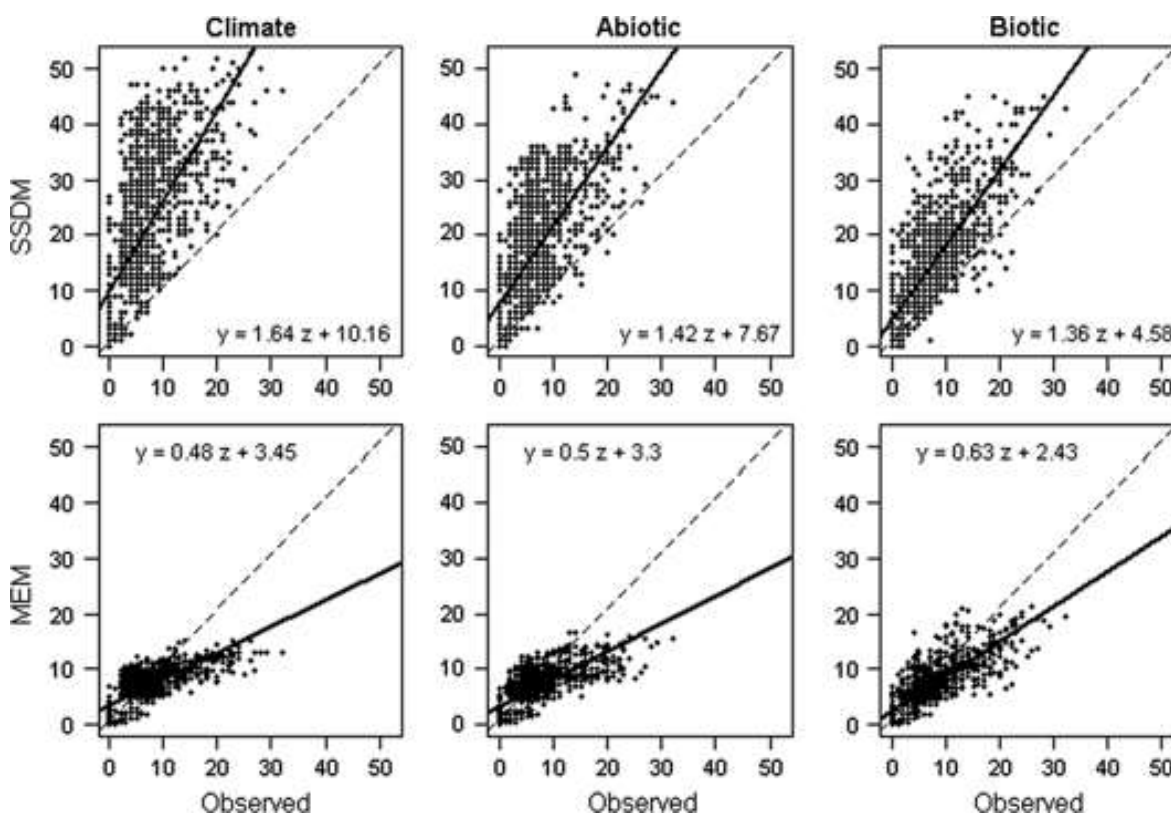
A linear regression was fitted to the plot of predicted vs observed species richness to compare the fit of different models (assessed by examining the slope, intercept and adjusted  $R^2$  values of the regression, with slope = 1 and intercept = 0 indicating perfect prediction; following Guisan and Rahbek 2011). The statistical significance of changes in slope, intercept and  $R^2$  of the regression between different models were assessed by using bootstrapping procedure for the linear models with 199 permutations and post-hoc-Tukey test using boot and multcomp packages in R (R Development Core Team). To test model improvement we examined the Spearman correlation (for MEM) and area under curve (AUC) and true skill statistics (TSS; for SSDM) for all three levels of model complexity.

All models were run using four-fold cross-validation with random assignment to quantify the predictive power of the models (Fielding and Bell 1997). Using cross-validation also robustly accounts for possible non-independence (i.e. spatial autocorrelation) of the data (Hijmans 2012). To examine spatial autocorrelation, correlograms (Moran's  $I$ ) with a lag of 50 m were calculated for observed species richness and for the residuals of the models. Only species with at least 12 presence records in each calibration dataset and 8 presence records in each validation dataset were included in analyses (excluding *E. nigrum*, *B. nana* and *B. pubescens*), resulting in 91 vascular plant species, 67 bryophyte species and 54 lichen species being modelled.

Based on the SSDM and MEM, species richness was projected for a  $3.1 \times 5.2$  km subarea in the centre of the study region (see also le Roux et al. 2013b for details). The area was divided in  $25 \times 25$  m cells, with a  $1 \text{ m}^2$  plot in the centre of each cell used for all calculations and predictions.

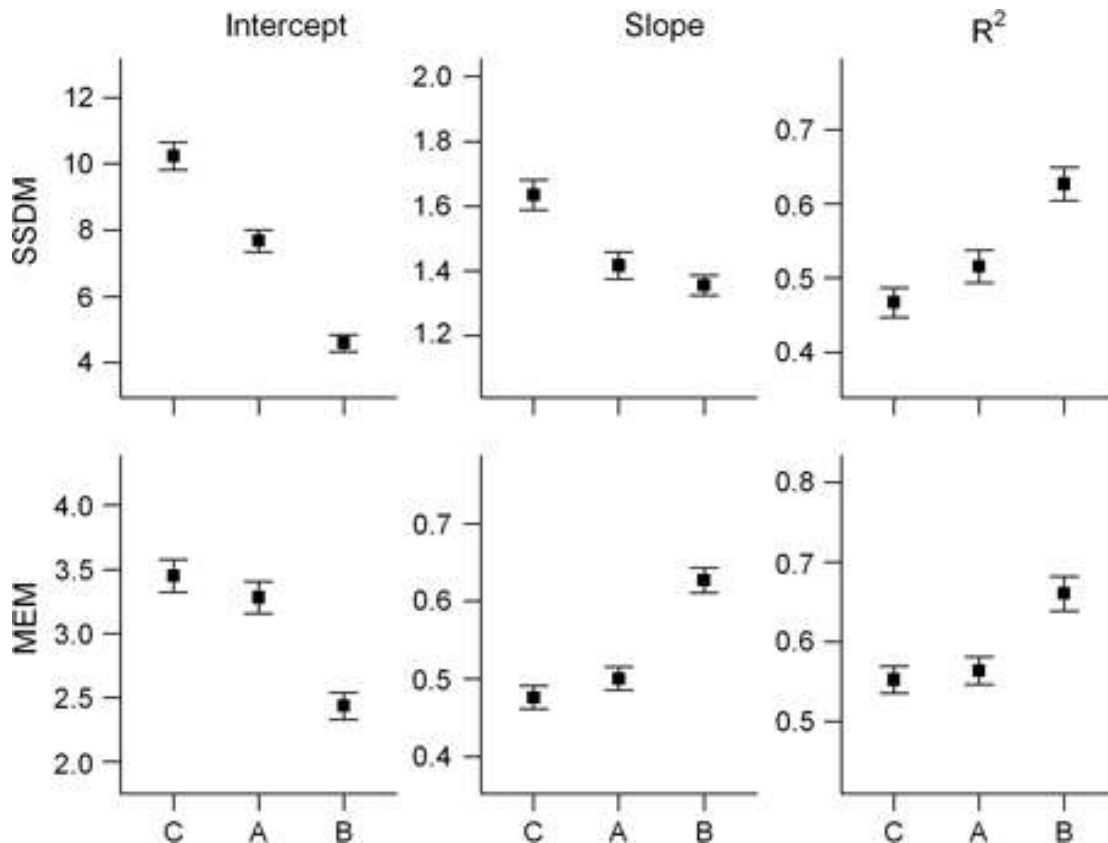
## Results

The inclusion of biotic variables in both SSDM and MEM improved species richness predictions for all three taxonomic groups, as robustly assessed by four-fold cross-validation (Fig. 2, 3 and Supplementary material Appendix 2, Fig. A1–A4). Results thus support our hypothesis about the importance of biotic interactions in species richness modelling. Species richness predictions from SSDM without biotic interaction constantly overpredicted richness, resulting in an overall positive bias. In contrast, MEM mainly overpredicted species richness in low-diversity cells, but constantly underpredicted in high-richness cells. Nonetheless, for both frameworks the inclusion of biotic variables reduced these biases, with the linear fit between observed and predicted species richness converging towards perfect prediction (i.e. the regression slope became significantly closer to unity [ $= 1$ ] and the intercept significantly smaller [ideally 0]) (Fig. 2, 3). In SSDM, the mean predicted richness declined when adding biotic variables to the models for all three taxa (Supplementary material Appendix 3, Table A1).



**Figure 2.** Predicted vs observed vascular plant species richness in 1080 plots in arctic-alpine north-western Finland and Norway, based on three levels of model complexity: climate model (3 explanatory variables; left column), abiotic model (6 explanatory variables; middle column), and biotic model (9 explanatory variables in total; right column). Top row: vascular plant species richness as predicted by stacked species distribution modelling (SSDM; values are the mean of predictions from three modelling techniques). Bottom row: vascular plant species richness as predicted by macroecological modelling (MEM; values are the mean of predictions from three modelling techniques). The details of the best linear fit between observed and predicted species richness are provided in each panel, with the fit represented by a solid line. For perfect prediction regression slope would be 1, and intercept zero, i.e. the dashed line, which represents the 1:1 relationship between observed and predicted values. For bryophytes and lichens, see Supplementary material Appendix 2, Fig. A1, A2.

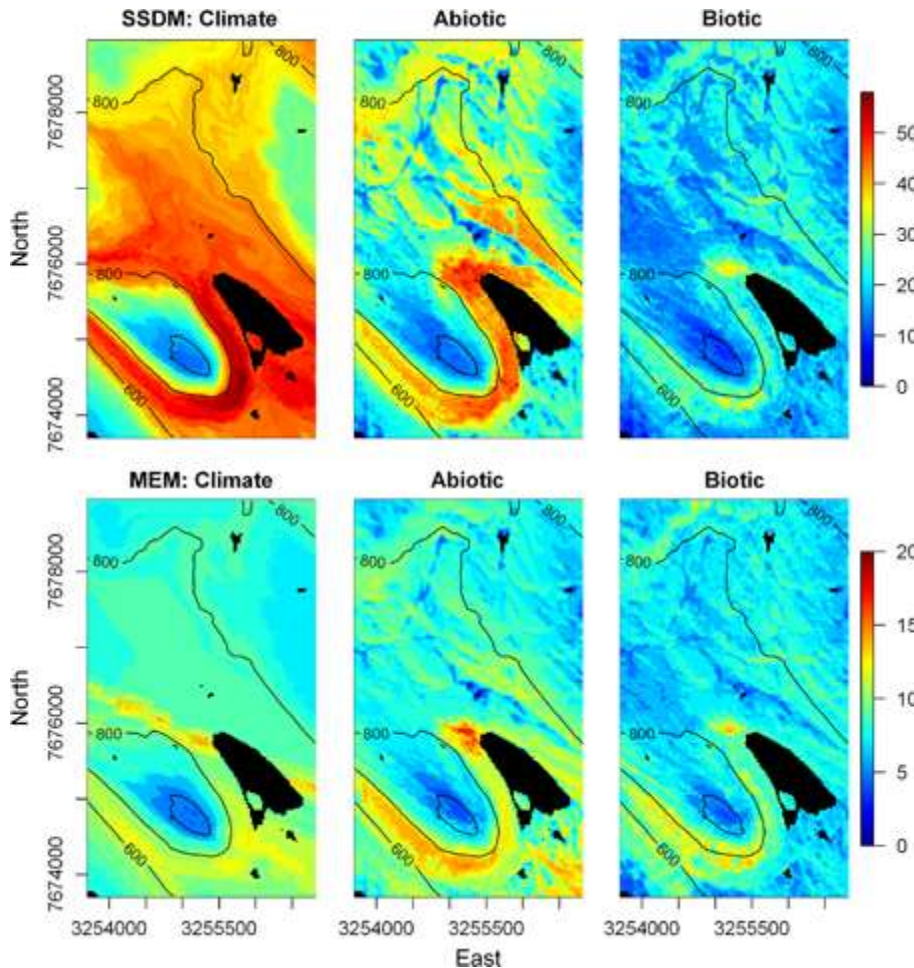




**Figure 3.** Bootstrapped linear model parameters (mean  $\pm$  SD) from regressions of predicted vs observed vascular plant species richness (top row: stacked species distribution modelling = SSDM; bottom row: macroecological modelling = MEM), using three different levels of model complexity (C = climate model; A = abiotic model; B = biotic model). Left column: regression intercept. Middle column: regression slope. Right column: coefficient of determination. Regression parameters from the three different models differ significantly ( $p < 0.05$ ) in all comparisons. Inclusion of biotic interactions significantly improved the species richness models, both in SSDM and MEM, with the intercept and slope of the regression converging towards zero and one respectively. For bryophytes and lichens, see Supplementary material Appendix 2, Fig. A3, A4.

Not only did overall species richness predictions improve, but the inclusion of biotic interactions also improved other aspects of the SSDM- and MEM-based models. Adjusted  $R^2$  (or predictive deviance for GBM) values were improved for all algorithms and taxa when comparing biotic models to climatic or abiotic models. For SSDM, when averaged across all modelled species, mean AUC values were higher for biotic models than climate or abiotic model in seven cases out of nine (six of these being statistically significant) and TSS was improved for four out of nine cases (two improvements being statistically significant) (e.g. for lichens the improvement in AUC is 0.024 and in TSS 0.021 when comparing climatic and biotic GLM improvement; for all taxa and algorithms see Supplementary material Appendix 3, Table A1, A2). For MEM, Spearman correlations were significantly higher for all biotic models compared to climate and abiotic models, indicating less unexplained variance when predicting species richness (e.g. mean improvement in correlation coefficients when comparing climatic and biotic models for vascular plant species is 0.1; for correlation and adjusted  $R^2$  values for all taxa and algorithms see Supplementary material Appendix 3, Table A3). Because spatial autocorrelation was relatively weak in model residuals (especially for MEM), we did not consider further evaluation of spatial autocorrelation necessary (Supplementary material Appendix 4, Fig. A1–A3).

Projections of species richness for a subarea of the study region were refined by including biotic variables (Fig. 4 and Supplementary material Appendix 2, Fig. A5, A6). In particular, as visually interpreted, overprediction is reduced in SSDM, and both over- and underprediction are reduced in MEM in these projections with the incorporation of biotic variables.



**Figure 4.** Incorporation of local abiotic (abiotic model) and biotic variables (biotic model) into the climate model generates more fine-scale variability in predictions for vascular plant species, for both the stacked species distribution models = SSDM (top row) and macroecological modelling = MEM (bottom row) (for bryophytes and lichens, see Supplementary material Appendix 2, Fig. A5, A6). From this visualization it is also evident that the predictions of the SSDM and MEM converge (and that the prediction biases are reduced) with the sequential addition of the local abiotic and biotic variables. Geographic coordinates are shown in the Finnish coordinate system.

## Discussion

Recent reviews call for the integration of biotic interactions into biodiversity modelling (Kissling et al. 2012), especially when using high-resolution data (Wisiz et al. 2013). Biotic interactions are known to impact individual species distributions, and therefore are expected to shape species richness patterns too (Araújo and Luoto 2007, Wisiz et al. 2013). Our study is the first to provide a formal test whether adding biotic interactions into the two prevailing modelling frameworks of species richness, stacked species distribution (SSDM) and macroecological modelling (MEM), improves biodiversity estimates. We show unambiguously that biotic interactions are important in determining high-resolution patterns

of species richness: inclusion of the cover of dominant species as predictor variables reduced previously reported biases in richness predictions in both SSDM and MEM of vascular plants, bryophytes and lichens (Newbold et al. 2009, Dubuis et al. 2011). Improvements in predictions, irrespective of the framework or taxa used, provide clear support for our hypothesis that the incorporation of biotic interactions improves species richness models by more explicitly considering community-level assembly processes (Pearson and Dawson 2003).

Due to fundamental differences between the SSDM and MEM frameworks, the incorporation of biotic interactions improves species richness predictions in markedly different ways. Conventional SSDM using only abiotic predictors treat each species in isolation, modelling its occurrence under the assumption that other species do not affect its presence in the environment (Dubuis et al. 2011, Guisan and Rahbek 2011). In other words, each species distribution is predicted without considering competition for resources or space (Pearson and Dawson 2003, Guisan and Thuiller 2005) or other interactions (Araújo and Luoto 2007, Pellissier et al. 2010). The inclusion of biotic interactions into individual species distribution models generally constrain species predicted ranges (although broader distributions may also be predicted; see also le Roux et al. (2012)), because the species probability of occurrence becomes contingent on both the abiotic and biotic variables in the models. Consequently, this may result in a reduction of overprediction of stacked richness predictions.

In contrast, MEM relates total species richness directly to explanatory variables (Ferrier and Guisan 2006, Thuiller et al. 2006, Rahbek et al. 2007, Algar et al. 2009). The resources or limitations expressed by the explanatory variables are assumed to influence the whole pool of species alike, adopting an approach that is fundamentally different from SSDM where explanatory variables are correlated with each individual species (Dubuis et al. 2011, Guisan and Rahbek 2011). The inclusion of biotic interactions into MEM had the effect of reducing the biases previously reported as affecting both ends of the richness gradient, with at least two distinct mechanisms involved. The biotic variables help to distinguish sites that have similar abiotic conditions but different species composition, and therefore different levels (or dominant types) of inter-specific interactions. For example, in our study system, the cover of the three dominant species are assumed to provide an estimate of the level of competition at each site (following le Roux et al. 2014). Thus, under abiotically-favourable conditions, higher total richness is associated with low dominant cover (i.e. lower competition), while fewer species co-occur in plots with a high cover of dominants.

The inclusion of biotic interactions improved predicted species richness of all three taxa, underlining the potential for biotic interactions to determine species richness in diverse taxa. However, there was some variability in the intensity of the effect of biotic interactions between taxa. Vascular plants and lichens responded more strongly to biotic interactions than bryophytes, suggesting bryophytes' greater independence from the impact of the three dominant species. Such differences between bryophytes and vascular plant species and lichens has been demonstrated in other studies too: Grytnes et al. (2006) found that altitude did not impact bryophyte species richness, suggesting that these species respond more strongly to fine-scale environmental drivers, including possibly soil moisture. For vascular plant species both positive and negative responses towards other species, including *E. nigrum* and *B. pubescens* in tundra environments, have been demonstrated in multiple studies (Grytnes et al. 2006, Pellissier et al. 2010, le Roux et al. 2013c, Mod et al. 2014). In these studies a range of different potential mechanisms were suggested as driving the interactions, including allelopathy and providing shelter from disturbance and extreme microclimatic

conditions. For lichens, the impact of biotic interactions presumably results from shading, which decreases above treeline (Cornelissen et al. 2001, Grytnes et al. 2006). Varying outcomes between taxa also suggested that other classifications (e.g. based on species traits) may more accurately group species by their sensitivity to biotic interactions (see also Pellissier et al. 2010, le Roux et al. 2012, Thuiller et al. 2013) and thus further improve species richness predictions for these subgroups.

Despite the fundamental differences between the two frameworks and the three taxa, incorporating the cover of three dominant species as a proxy for biotic interactions into the models consistently significantly improved species richness predictions. This approach is cost-efficient and statistically simple where suitable data is available (although, see Ovaskainen et al. 2010 and Kissling et al. 2012) for other methods to incorporate biotic interactions to the models); no additional data or analyses are required and results are easy to interpret. Therefore our approach could be successfully implemented in other frameworks; for example to supplement ecological assembly rules (Götzenberger et al. 2012) in spatially explicit species assemblage modelling (Guisan and Rahbek 2011). This methodology can account for both negative and positive interactions (Araújo and Luoto 2007, Pellissier et al. 2010) and for plant–animal (le Roux et al. 2013c) and predator–prey (Sandom et al. 2013) interactions, even though in our study negative interactions appear to dominate (presumably due to the higher frequency of the species having negative responses to the three dominant plant species) and the hypothesis was tested for plant–plant interactions. Further, the data used in this study emphasise the importance of detailed and field-quantified data, as well as the need for identifying and incorporating relevant biotic interactions (i.e. potentially competitive and/or facilitative species) when implementing high-resolution studies (see also Wisz et al. 2013). As our research demonstrates the increased accuracy of high-resolution models when accounting for biotic interactions, we believe this approach should be utilized where possible; while it may require slightly more time-consuming data collection, this cost is offset by the benefit of more accurate models of species richness.

While the improvements in species richness predictions were highly significant, unexplained variation remains (i.e. slope and intercept estimates for the regression between observed and predicted species richness did not reach one and zero, respectively, after including biotic interactions into the models). Ecological theory suggests that species richness patterns are affected by a diversity of factors in addition to environmental conditions and biotic interactions (e.g. other remaining biotic processes and dispersal; Graves and Rahbek 2005, Field et al. 2009, Gotelli et al. 2009, Thuiller et al. 2013). Further improvements to MEM could include adding important environmental constraints on the number of species that can coexist through, e.g. accounting more explicitly for differences in local and regional species pools (Tofts and Silvertown 2002). For SSDM, further improvements would be the addition of dispersal limitations and population dynamics for individual species (Lortie et al. 2004, Guisan and Thuiller 2005, Guisan and Rahbek 2011, Thuiller et al. 2013). In this study, however, dispersal processes likely have little impact, since neither the geographical context nor the spatial extent of our study area are expected to pose serious barriers to the dispersal of these species over long time periods.

To fully understand ecosystem dynamics and, thus, the potential impacts of environmental change thereon, the factors determining biodiversity patterns need to be accurately characterized. Species richness modelling studies are still mainly based solely on climatic data, although the importance of other variables and processes for the distribution of species have been demonstrated in myriad of studies (Araújo and Luoto 2007, Meier et al. 2010, le

Roux et al. 2013b, Lenoir et al. 2013). This study provides strong evidence that interspecific interactions, in combination with abiotic factors, are important drivers of species richness (although, it does not rule out the impact of other additional drivers). Furthermore, the methodology used here provides a practical solution to account for biotic interactions in species richness models, offering the promising potential for more robust species richness predictions. With the extensive data and suitable methods currently available, modelling approaches disregarding biotic predictors are short-sighted and may lead to serious misinterpretations in predictions of biodiversity.

## Acknowledgements

We acknowledge funding from the Academy of Finland (Project Number 1140873). HKM received also funding from Research Foundation of the Univ. of Helsinki. We also thank R. Virtanen for distribution data, and T. Rämä for assistance with the field work.

## References

- Aalto, J. et al. 2014. The meso-scale drivers of temperature extremes in high-latitude Fennoscandia. – *Clim. Dyn.* 42: 237–252.
- Aerts, R. 1999. Interspecific competition in natural plant communities: mechanisms, trade-offs and plant–soil feedbacks. – *J. Exp. Bot.* 50: 29–37.
- Aerts, R. 2010. Nitrogen-dependent recovery of subarctic tundra vegetation after simulation of extreme winter warming damage to *Empetrum hermaphroditum*. – *Global Change Biol.* 16: 1071–1081.
- Algar, A. C. et al. 2009. Predicting the future of species diversity: macroecological theory, climate change, and direct tests of alternative forecasting methods. – *Ecography* 32: 22–33.
- Anderson, R. P. et al. 2002. Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. – *Oikos* 98: 3–16.
- Araújo, M. B. and Guisan, A. 2006. Five (or so) challenges for species distribution modelling. – *J. Biogeogr.* 33: 1677–1688.
- Araújo, M. B. and Luoto, M. 2007. The importance of biotic interactions for modelling species distributions under climate change. – *Global Ecol. Biogeogr.* 16: 743–753.
- Araújo, M. B. and New, M. 2007. Ensemble forecasting of species distributions. – *Trends Ecol. Evol.* 22: 42–47.
- Baack, E. J. et al. 2006. Ecological factors limiting the distribution of *Gilia tricolor* in a California grassland mosaic. – *Ecology* 87: 2736–2745.
- Beven, K. and Kirkby, M. 1979. A physically based, variable contributing area model of basin hydrology. – *Hydrol. Sci. J.* 24: 43–69.

- Billings, W. D. and Mooney, H. A. 1968. The ecology of arctic and alpine plants. – *Biol. Rev.* 43: 481–529.
- Brown, J. H. 1995. *Macroecology*. – Univ. of Chicago Press.
- Calabrese, J. M. et al. 2014. Stacking species distribution models and adjusting bias by linking them to macroecological models. – *Global Ecol. Biogeogr.* 23: 99–112.
- Callaway, R. 1995. Positive interactions among plants. – *Bot. Rev.* 61: 306–349.
- Case, T. J. et al. 2005. The community context of species' borders: ecological and evolutionary perspectives. – *Oikos* 108: 28–46.
- Cornelissen, J. H. C. et al. 2001. Global change and arctic ecosystems: is lichen decline a function of increases in vascular plant biomass? – *J. Ecol.* 89: 984–994.
- Davis, A. J. et al. 1998. Individualistic species responses invalidate simple physiological models of community dynamics under global environmental change. – *J. Anim. Ecol.* 67: 600–612.
- Dubuis, A. et al. 2011. Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. – *Divers. Distrib.* 17: 1122–1131.
- Dymond, C. C. and Johnson, E. A. 2002. Mapping vegetation spatial patterns from modeled water, temperature and solar radiation gradients. – *ISPRS J. Photogramm.* 57: 69–85.
- Elith, J. and Graham, C. H. 2009. Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. – *Ecography* 32: 66–77.
- Elith, J. et al. 2008. A working guide to boosted regression trees. – *J. Anim. Ecol.* 77: 802–813.
- Eskelinen, A. et al. 2009. Links between plant community composition, soil organic matter quality and microbial communities in contrasting tundra habitats. – *Oecologia* 161: 113–123.
- ESRI 2012. *ArcGIS 10.1*. – Environment Systems Research Inst.
- Ferrier, S. and Guisan, A. 2006. Spatial modelling of biodiversity at the community level. – *J. Appl. Ecol.* 43: 393–404.
- Field, R. et al. 2009. Spatial species-richness gradients across scales: a meta-analysis. – *J. Biogeogr.* 36: 132–147.
- Fielding, A. H. and Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. – *Environ. Conserv.* 24: 38–49.
- Franklin, J. and Miller, J. A. 2009. *Mapping species distributions: spatial inference and prediction*. – Cambridge Univ. Press.

- Gallien, L. et al. 2012. Invasive species distribution models – how violating the equilibrium assumption can create new insights. – *Global Ecol. Biogeogr.* 21: 1126–1136.
- Gaston, K. J. 2000. Global patterns in biodiversity. – *Nature* 405: 220–227.
- Giannini, T. C. et al. 2013. Improving species distribution models using biotic interactions: a case study of parasites, pollinators and plants. – *Ecography* 36: 649–656.
- González-Salazar, C. et al. 2013. Comparing the relative contributions of biotic and abiotic factors as mediators of species' distributions. – *Ecol. Model.* 248: 57–70.
- Gotelli, N. J. et al. 2009. Patterns and causes of species richness: a general simulation model for macroecology. – *Ecol. Lett.* 12: 873–886.
- Gotelli, N. J. et al. 2010. Macroecological signals of species interactions in the Danish avifauna. – *Proc. Natl Acad. Sci. USA* 107: 5030–5035.
- Götzenberger, L. et al. 2012. Ecological assembly rules in plant communities – approaches, patterns and prospects. – *Biol. Rev.* 87: 111–127.
- Graves, G. R. and Rahbek, C. 2005. Source pool geometry and the assembly of continental avifaunas. – *Proc. Natl Acad. USA* 102: 7871–7876.
- Grytnes, J. A. et al. 2006. Species richness of vascular plants, bryophytes, and lichens along an altitudinal gradient in western Norway. – *Acta Oecol.* 29: 241–246.
- Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. – *Ecol. Model.* 135: 147–186.
- Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – *Ecol. Lett.* 8: 993–1009.
- Guisan, A. and Rahbek, C. 2011. SESAM – a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. – *J. Biogeogr.* 38: 1433–1444.
- Hämet-Ahti, L. et al. 1998. *Retkeilykasvio (Field flora of Finland)*. – Finnish Museum of Natural History, Botanical Museum.
- Heikkinen, R. K. et al. 2006. Methods and uncertainties in bioclimatic envelope modelling under climate change. – *Prog. Phys. Geogr.* 30: 751–777.
- Heikkinen, R. K. et al. 2007. Biotic interactions improve prediction of boreal bird distributions at macro-scales. – *Global Ecol. Biogeogr.* 16: 754–763.
- Hijmans, R. J. 2012. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. – *Ecology* 93: 679–688.
- Kearney, M. R. and Porter, W. P. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. – *Ecol. Lett.* 12: 334–350.

- Kissling, W. D. et al. 2012. Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. – *J. Biogeogr.* 39: 2163–2178.
- Korsman, K. et al. (eds) 1997. Suomen kallioperäkarta (Bedrock map of Finland) 1:1 000 000. – Geological Survey of Finland.
- le Roux, P. C. et al. 2012. Biotic interactions affect the elevational ranges of high-latitude plant species. – *Ecography* 35: 1048–1056.
- le Roux, P. C. et al. 2013a. Soil moisture's underestimated role in climate change impact modelling in low-energy systems. – *Global Change Biol.* 19: 2965–2975.
- le Roux, P. C. et al. 2013b. Geomorphological disturbance is necessary for predicting fine-scale species distributions. – *Ecography* 36: 800–808.
- le Roux, P. C. et al. 2013c. Horizontal, but not vertical, biotic interactions affect fine-scale plant distribution patterns in a low energy system. – *Ecology* 94: 671–682.
- le Roux, P. C. et al. 2014. Incorporating dominant species as proxies for biotic interactions strengthens plant community models. – *J. Ecol.* 102: 767–775.
- Lenoir, J. et al. 2013. Local temperatures inferred from plant communities suggest strong spatial buffering of climate warming across northern Europe. – *Global Change Biol.* 19: 1470–1481.
- Levinsky, I. et al. 2013. Climate envelope models suggest spatio-temporal co-occurrence of refugia of African birds and mammals. – *Global Ecol. Biogeogr.* 22: 351–363.
- Liu, C. et al. 2005. Selecting thresholds of occurrence in the prediction of species distributions. – *Ecography* 28: 385–393.
- Lortie, C. J. et al. 2004. Rethinking plant community theory. – *Oikos* 107: 433–438.
- McCune, B. and Keon, D. 2002. Equations for potential annual direct incident radiation and heat load. – *J. Veg. Sci.* 13: 603–606.
- Meier, E. S. et al. 2010. Biotic and abiotic variables show little redundancy in explaining tree species distributions. – *Ecography* 33: 1038–1048.
- Mod, H. K. et al. 2014. Outcomes of biotic interactions are dependent on multiple environmental variables. – *J. Veg. Sci.* 25: 1024–1032.
- Newbold, T. et al. 2009. Climate-based models of spatial patterns of species richness in Egypt's butterfly and mammal fauna. – *J. Biogeogr.* 36: 2085–2095.
- Ovaskainen, O. et al. 2010. Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. – *Ecology* 91: 2514–2521.
- Pajunen, A. M. et al. 2011. Impact of shrub canopies on understorey vegetation in western Eurasian tundra. – *J. Veg. Sci.* 22: 837–846.



Pearson, R. G. and Dawson, T. P. 2003. Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? – *Global Ecol. Biogeogr.* 12: 361–371.

Pearson, R. G. et al. 2002. SPECIES: a spatial evaluation of climate impact on the envelope of species. – *Ecol. Model.* 154: 289–300.

Pellissier, L. et al. 2010. Species distribution models reveal apparent competitive and facilitative effects of a dominant species on the distribution of tundra plants. – *Ecography* 33: 1004–1014.

Pottier, J. et al. 2013. The accuracy of plant assemblage prediction from species distribution models varies along environmental gradients. – *Global Ecol. Biogeogr.* 22: 52–63.

Rahbek, C. et al. 2007. Predicting continental-scale patterns of bird species richness with spatially explicit models. – *Proc. R. Soc. B* 274: 165–174.

Ridgeway, G. 2005. Generalized boosted regression models. – [www.i-pensieri.com/gregr/gbm.shtml](http://www.i-pensieri.com/gregr/gbm.shtml).

Sandom, C. et al. 2013. Mammal predator and prey species richness are strongly linked at macroscales. – *Ecology* 94: 1112–1122.

Shevtsova, A. et al. 1995. Growth and reproduction of dwarf shrubs in a subarctic plant community: annual variation and above-ground interactions with neighbours. – *J. Ecol.* 263–275.

Soberón, J. and Peterson, A. T. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. – *Biodivers. Inform.* 2: 1–10.

Sormunen, H. et al. 2011. Inclusion of local environmental conditions alters high-latitude vegetation change predictions based on bioclimatic models. – *Polar Biol.* 34: 883–897.

Thuiller, W. et al. 2006. Predicting patterns of plant species richness in megadiverse South Africa. – *Ecography* 29: 733–744.

Thuiller, W. et al. 2013. A road map for integrating eco-evolutionary processes into biodiversity models. – *Ecol. Lett.* 16: 94–105.

Tofts, R. and Silvertown, J. 2002. Community assembly from the local species pool: an experimental study using congeneric species pairs. – *J. Ecol.* 90: 385–393. Tybirk, K. et al. 2000. Nordic Empetrum dominated ecosystems: function and susceptibility to environmental changes. – *Ambio* 29: 90–97.

Virtanen, R. et al. 2006. Broad-scale vegetation-environment relationships in Eurasian high-latitude areas. – *J. Veg. Sci.* 17: 519–528.

Virtanen, R. et al. 2010. Recent vegetation changes at the high-latitude tree line ecotone are controlled by geomorphological disturbance, productivity and diversity. – *Global Ecol. Biogeogr.* 19: 810–821.

Williams, R. B. G. 1988. The biogeomorphology of periglacial environments. – In: Viles, H. (ed.), *Biogeomorphology*. Blackwell, pp. 222–252.

Wisz, M. S. et al. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. – *Biol. Rev.* 88: 15–30.

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. – *J. R. Stat. Soc. B* 73: 3–36.

Zimmermann, N. E. et al. 2009. Climatic extremes improve predictions of spatial patterns of tree species. – *Proc. Natl Acad. USA* 106: 19723–19728.

## Supplementary material

### Appendix 1: Comparison of observed climatic data between the normal periods of 1971-2000 and 1981-2010.

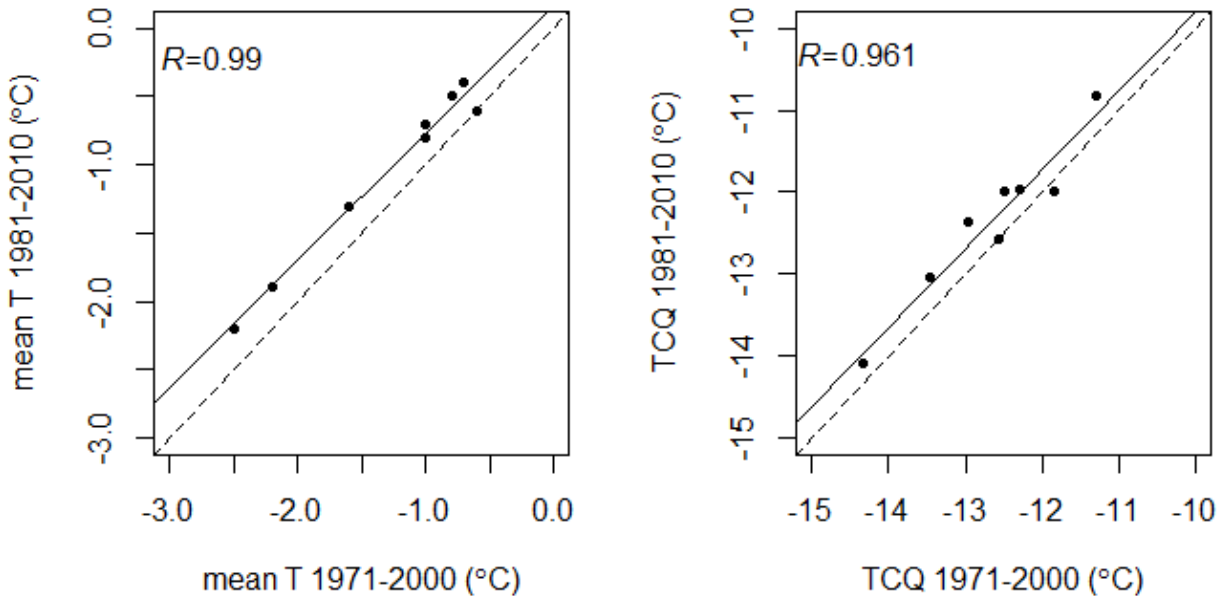


Figure A1. Comparison of observed mean annual temperature (mean T; left panel) and temperature of coldest quarter (TCQ; right panel) between the normal periods of 1971-2000 and 1981-2010 from eight weather stations close to the study area. Solid line represents linear fit (simple regression) between observed temperatures from the two normal periods and dashed line represent 1:1 relationship. Correlation between the two datasets is reported in top-left corner of each panel ( $p<0.01$ ). Based on the 8 weather stations average increase in annual mean temperature between normal periods is  $0.3\text{ }^{\circ}\text{C}$  and average increase in TCQ is  $0.45\text{ }^{\circ}\text{C}$ . As no climatic data from the most recent normal period (1981-2010) was available from all used 61 stations, the normal period of 1971-2000 was chosen for the climatic data. However, the temporal difference between species data collection (2008-2011) and climatic data should not reduce the reliability of the results, as climatic data from the different time periods shows broadly similar geographic patterns (i.e. despite the increase in temperatures, the conditions follow mainly the same spatial patterns in both normal periods).

**Appendix 2: Corresponding figures for bryophytes and lichens as represented for vascular plant species in the article.**

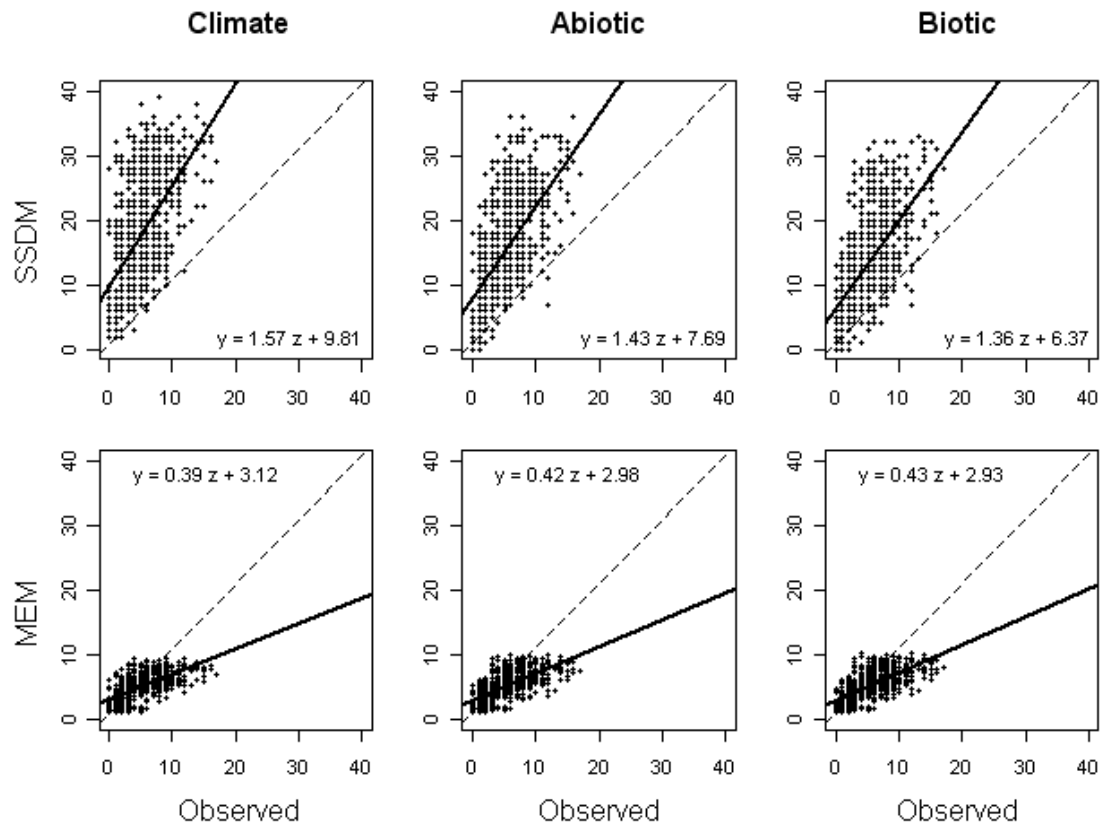


Fig. A1. Predicted vs. observed bryophyte species richness in 1080 plots in arctic-alpine north-western Finland and Norway, based on three levels of model complexity: climate model (3 explanatory variables; left column), abiotic model (6 explanatory variables; middle column), and biotic model (9 explanatory variables in total; right column). *Top row:* Bryophyte species richness as predicted by stacked species distribution modelling (SSDM; committee averaging method). *Bottom row:* Bryophyte species richness as predicted by macroecological modelling (MEM; values are the mean of predictions from three modelling techniques). The details of the best linear fit between observed and predicted species richness are provided in each panel, with the fit represented by a solid line. For perfect prediction regression slope would be 1, and intercept zero, i.e. the dashed line, which represents the 1:1 relationship between observed and predicted values.

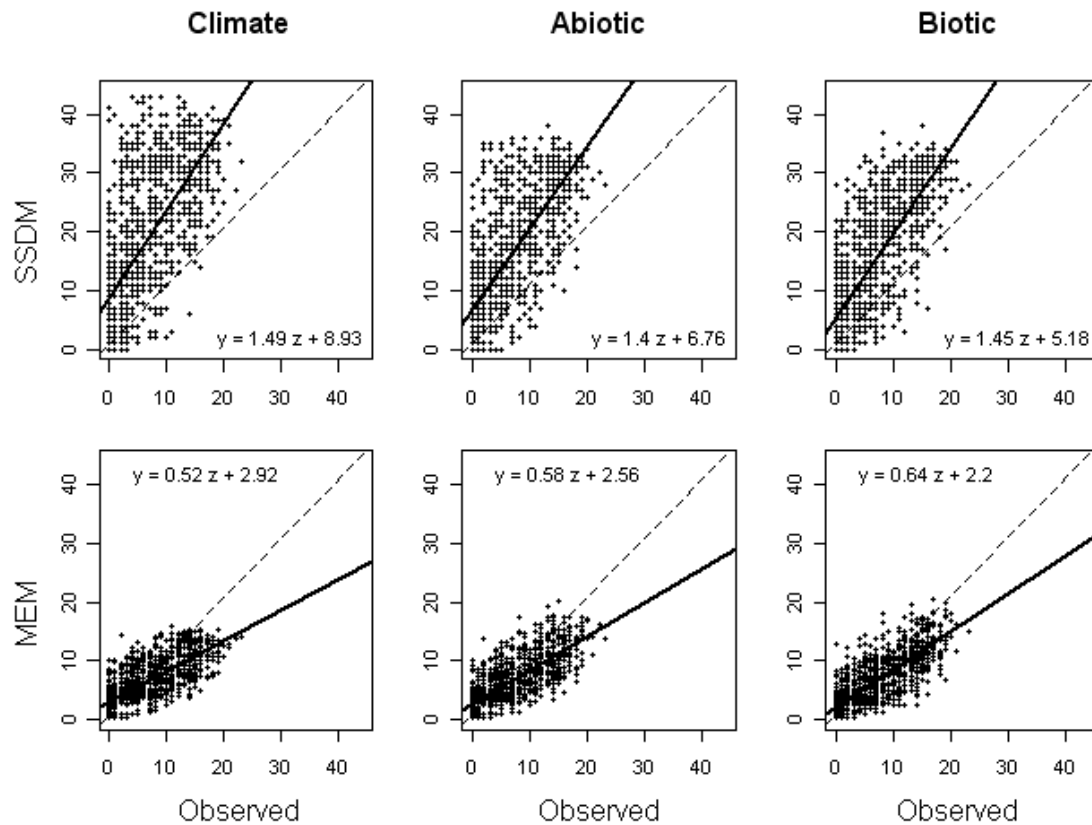


Fig. A2. Predicted vs. observed lichen species richness in 1080 plots in arctic-alpine north-western Finland and Norway, based on three levels of model complexity: climate model (3 explanatory variables; left column), abiotic model (6 explanatory variables; middle column), and biotic model (9 explanatory variables in total; right column). *Top row*: Lichen species richness as predicted by stacked species distribution modelling (SSDM; committee averaging method). *Bottom row*: Lichen species richness as predicted by macroecological modelling (MEM; values are the mean of predictions from three modelling techniques). The details of the best linear fit between observed and predicted species richness are provided in each panel, with the fit represented by a solid line. For perfect prediction regression slope would be 1, and intercept zero, i.e. the dashed line, which represents the 1:1 relationship between observed and predicted values.

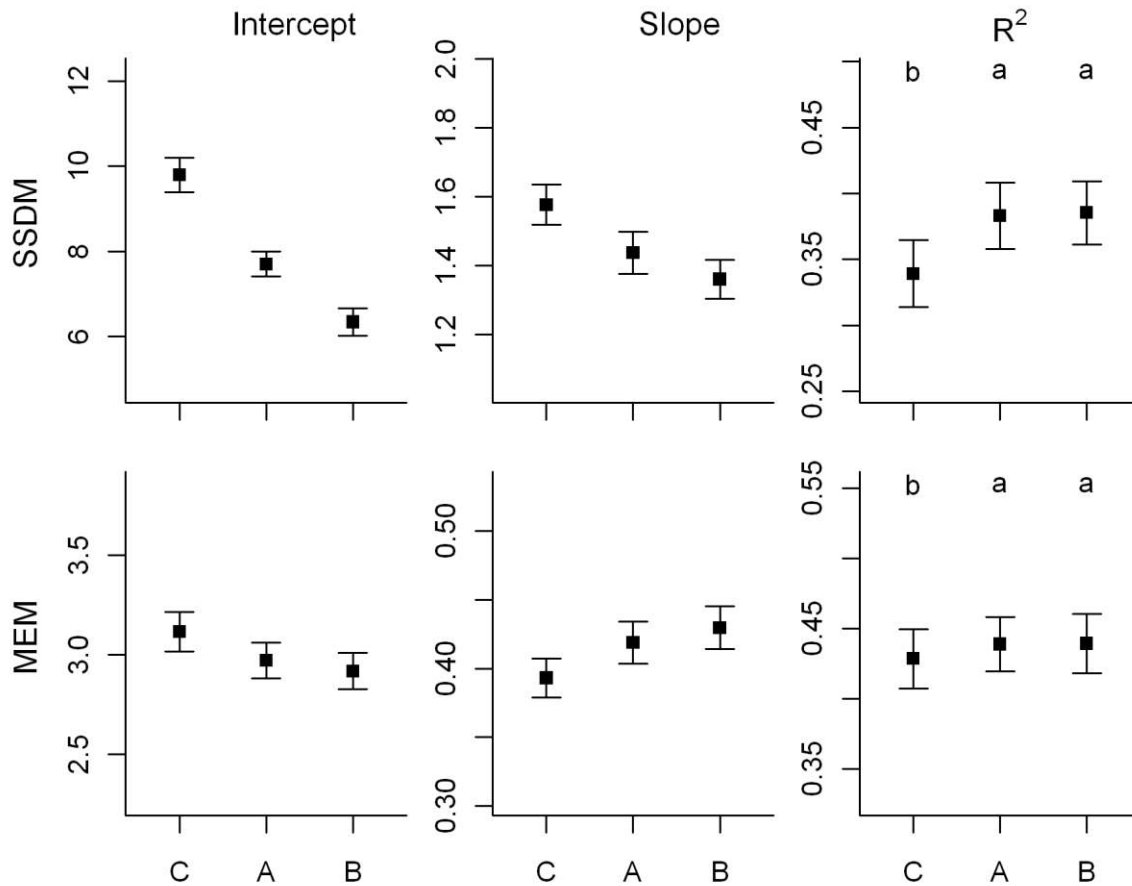


Fig. A3. Bootstrapped linear model parameters (mean  $\pm$  SD) from regressions of predicted vs. observed bryophyte species richness (top row: stacked species distribution modelling = SSDM; bottom row: macroecological modelling = MEM), using three different levels of model complexity (C = climate model; A = abiotic model; B = biotic model). *Left column:* regression intercept. *Middle column:* regression slope. *Right column:* coefficient of determination. Regression parameters from the three different models differ significantly ( $p < 0,05$ ) in all comparisons. Inclusion of biotic interactions significantly improved the species richness models, both in SSDM and MEM, with the intercept and slope of the regression converging towards zero and one respectively.

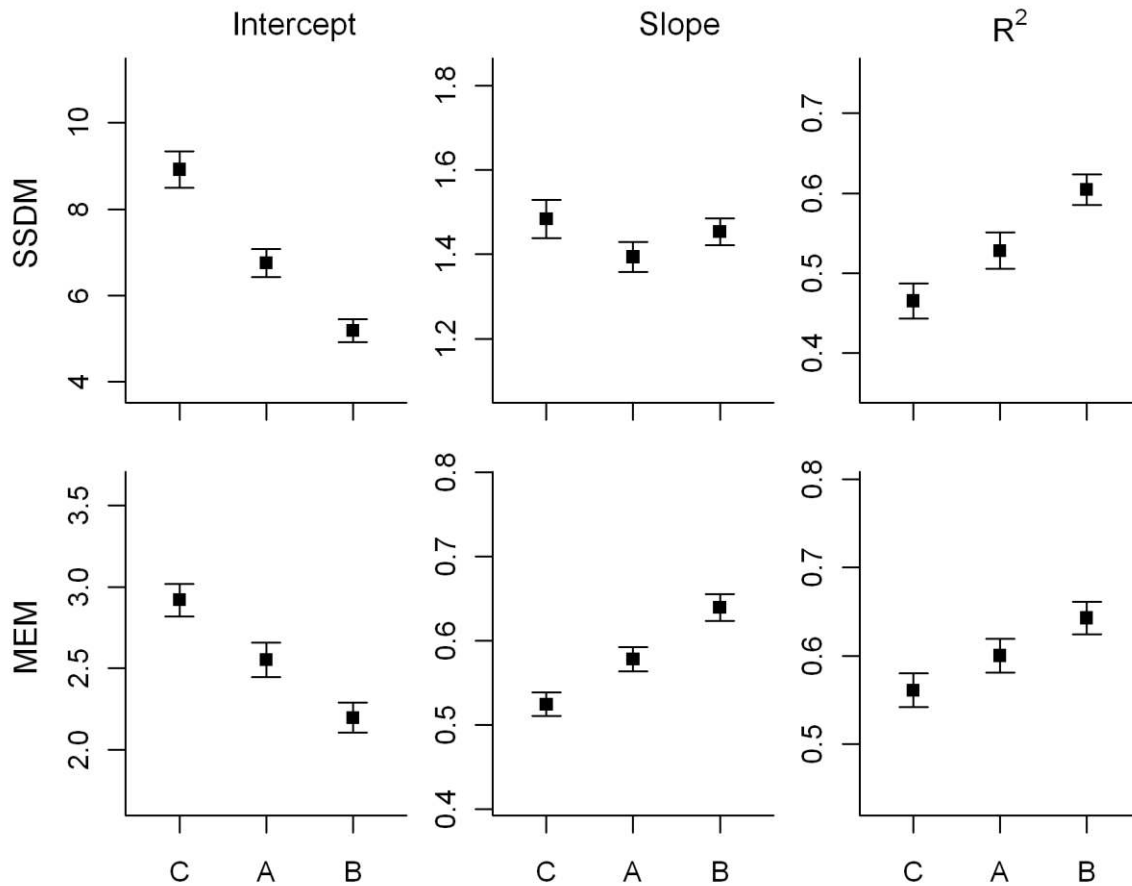


Fig. A4. Bootstrapped linear model parameters (mean  $\pm$  SD) from regressions of predicted vs. observed lichen species richness (top row: stacked species distribution modelling = SSDM; bottom row: macroecological modelling = MEM), using three different levels of model complexity (C = climate model; A = abiotic model; B = biotic model). *Left column:* regression intercept. *Middle column:* regression slope. *Right column:* coefficient of determination. Regression parameters from the three different models differ significantly ( $p < 0,05$ ) in all comparisons. Inclusion of biotic interactions significantly improved the species richness models, both in SSDM and MEM, with the intercept and slope of the regression converging towards zero and one respectively.

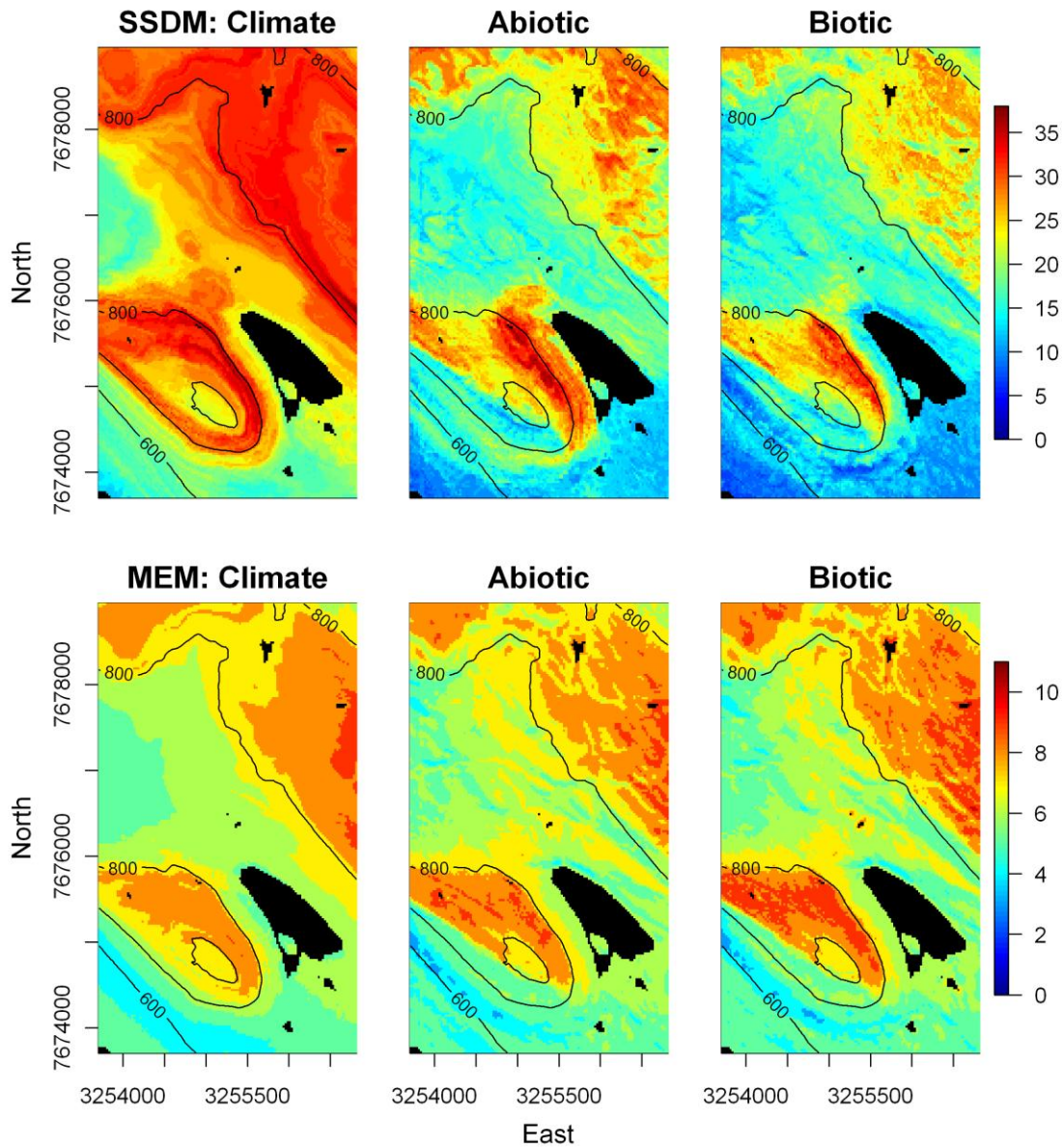


Fig. A5. Incorporation of local abiotic (abiotic model) and biotic variables (biotic model) into the climate model generates more fine-scale variability in predictions for bryophyte species, for both the stacked species distribution models = SSDM (top row) and macroecological modelling = MEM (bottom row). From this visualization it is also evident that the predictions of the SSDM and MEM converge (and that the prediction biases are reduced) with the sequential addition of the local abiotic and biotic variables. Geographic coordinates are shown in the Finnish coordinate system.



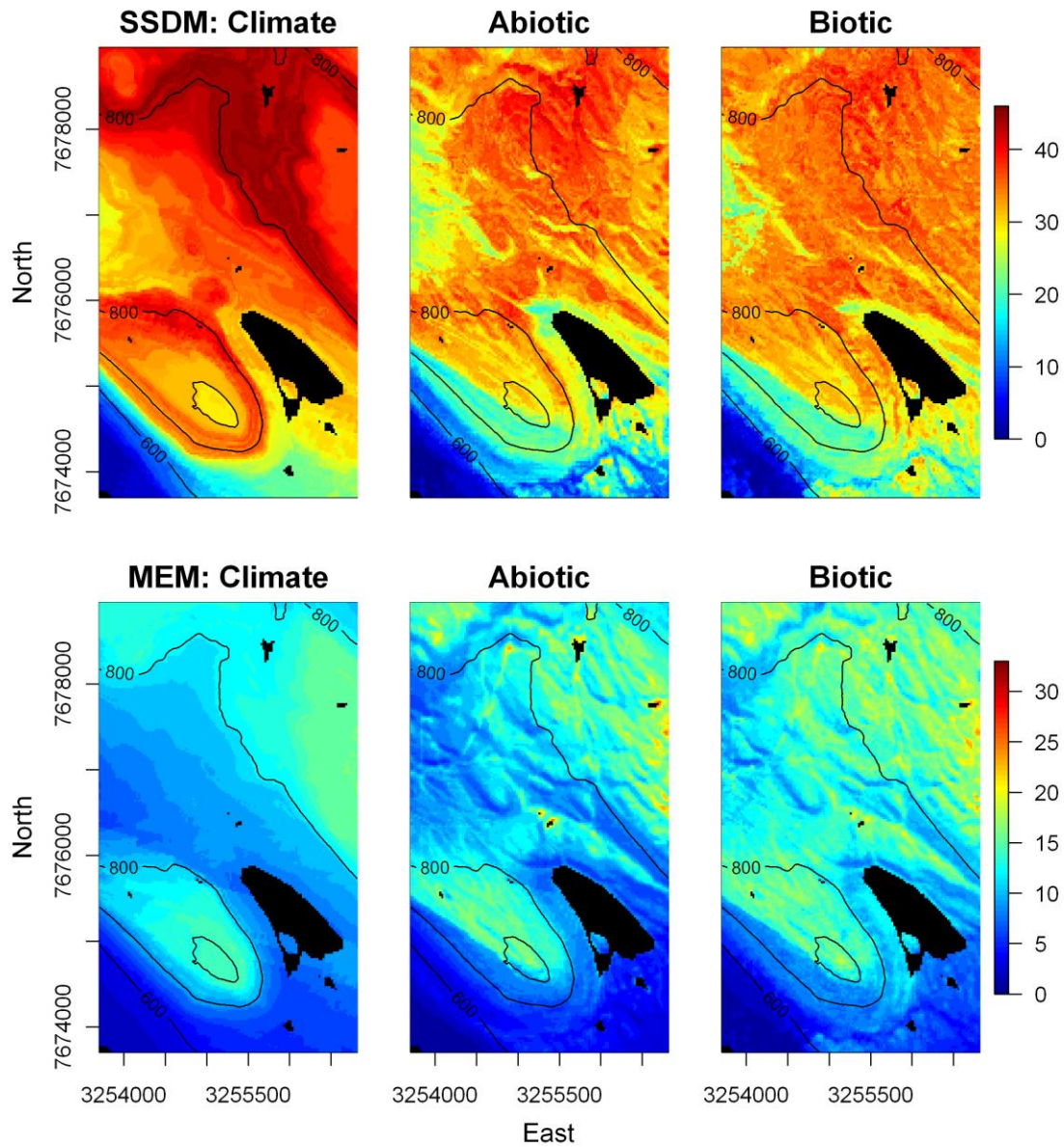


Fig. A6. Incorporation of local abiotic (abiotic model) and biotic variables (biotic model) into the climate model generates more fine-scale variability in predictions for lichen species, for both the stacked species distribution models = SSDM (top row) and macroecological modelling = MEM (bottom row). From this visualization it is also evident that the predictions of the SSDM and MEM converge (and that the prediction biases are reduced) with the sequential addition of the local abiotic and biotic variables. Geographic coordinates are shown in the Finnish coordinate system.

**Appendix 3: Model improvement by stacked species distribution models (SSDM) and macroecological modeling (MEM)**

Table A1. Mean observed and predicted richness by different models by SSDM. Predicted species richness decrease with sequential inclusion of variable sets.

	<b>Observed</b>	<b>Climate</b>	<b>Abiotic</b>	<b>Biotic</b>
<b>Vascular</b>	6.6	20.9	17.1	13.6
<b>Bryophyte</b>	5.2	17.9	15.1	13.4
<b>Lichen</b>	6.2	18.2	15.4	14.2

Table A2. Model improvement based on adjusted R<sup>2</sup>, area under curve (AUC) and true skills statistics (TSS) values and their statistical significance for SSDM. (C = Climate model, A = Abiotic model, B = Biotic model, Pred. dev. = Predictive deviance).

Species group	Algo-rithm	Metrics	C	A	B	(C→A)	(A→B)	(C→B)
			Mean (±SE)	Mean (±SE)	Mean (±SE)	<i>p</i>	<i>p</i>	<i>p</i>
Vascular	GLM	Adjusted R <sup>2</sup>	0.103 (±0.011)	0.125 (±0.011)	0.150 (±0.012)	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
		AUC	0.847 (±0.008)	0.860 (±0.008)	0.837 (±0.011)	<b>0.000</b>	0.561	<b>0.028</b>
		TSS	0.565 (±0.015)	0.583 (±0.015)	0.552 (±0.017)	<b>0.006</b>	0.790	0.072
	GAM	Adjusted R <sup>2</sup>	0.118 (±0.011)	0.143 (±0.012)	0.162 (±0.012)	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
		AUC	0.863 (±0.007)	0.877 (±0.007)	0.878 (±0.007)	<b>0.000</b>	0.069	<b>0.000</b>
		TSS	0.598 (±0.014)	0.615 (±0.014)	0.598 (±0.014)	<b>0.012</b>	0.827	0.071
	GBM	Pred. dev.	0.622 (±0.015)	0.624 (±0.014)	0.612 (±0.014)	0.526	<b>0.002</b>	<b>0.003</b>
		AUC	0.880 (±0.008)	0.876 (±0.008)	0.889 (±0.007)	0.834	<b>0.000</b>	<b>0.003</b>
		TSS	0.532 (±0.016)	0.534 (±0.016)	0.523 (±0.017)	0.601	0.692	0.700
Bryophyte	GLM	Adjusted R <sup>2</sup>	0.081 (±0.010)	0.087 (±0.011)	0.090 (±0.012)	<b>0.007</b>	0.073	<b>0.019</b>
		AUC	0.807 (±0.012)	0.795 (±0.011)	0.768 (±0.014)	0.057	0.851	0.548
		TSS	0.499 (±0.021)	0.451 (±0.021)	0.427 (±0.024)	0.804	0.317	0.807
	GAM	Adjusted R <sup>2</sup>	0.084 (±0.011)	0.098 (±0.011)	0.105 (±0.012)	<b>0.000</b>	<b>0.012</b>	<b>0.000</b>
		AUC	0.813 (±0.011)	0.825 (±0.010)	0.822 (±0.010)	<b>0.000</b>	0.325	<b>0.008</b>
		TSS	0.493 (±0.022)	0.507 (±0.020)	0.495 (±0.019)	0.158	0.398	0.183
	GBM	Pred. dev.	0.758 (±0.015)	0.753 (±0.016)	0.748 (±0.016)	0.061	<b>0.011</b>	<b>0.006</b>
		AUC	0.813 (±0.012)	0.815 (±0.012)	0.815 (±0.013)	0.141	0.202	0.350
		TSS	0.406 (±0.023)	0.398 (±0.023)	0.394 (±0.025)	0.851	0.673	0.980
Lichen	GLM	Adjusted R <sup>2</sup>	0.090 (±0.011)	0.100 (±0.014)	0.116 (±0.015)	<b>0.007</b>	<b>0.000</b>	<b>0.000</b>
		AUC	0.775 (±0.011)	0.787 (±0.011)	0.799 (±0.011)	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
		TSS	0.444 (±0.021)	0.447 (±0.021)	0.465 (±0.020)	0.108	<b>0.011</b>	<b>0.010</b>
	GAM	Adjusted R <sup>2</sup>	0.099 (±0.013)	0.104 (±0.013)	0.121 (±0.016)	0.067	<b>0.000</b>	<b>0.000</b>
		AUC	0.783 (±0.012)	0.798 (±0.010)	0.810 (±0.011)	<b>0.000</b>	<b>0.000</b>	<b>0.000</b>
		TSS	0.455 (±0.020)	0.456 (±0.020)	0.466 (±0.022)	0.155	0.116	<b>0.031</b>
	GBM	Pred. dev.	0.770 (±0.017)	0.767 (±0.018)	0.754 (±0.019)	0.103	<b>0.001</b>	<b>0.000</b>
		AUC	0.802 (±0.012)	0.801 (±0.012)	0.811 (±0.012)	0.352	<b>0.001</b>	<b>0.006</b>
		TSS	0.425 (±0.022)	0.422 (±0.023)	0.430 (±0.023)	0.625	0.551	0.050

Table A3. Model improvement based on Spearman correlation (cross-validated) for MEM approach. All correlations were statistically highly significant ( $p < 0,001$ ). (C = Climate model, A = Abiotic model, B = Biotic model).

<b>Species group</b>	<b>Algorithm</b>	<b>Metrics</b>	<b>C</b>	<b>A</b>	<b>B</b>
Vascular	GLM	Adjusted R <sup>2</sup>	0.346	0.382	0.552
		Cor	0.609	0.633	0.768
	GAM	Adjusted R <sup>2</sup>	0.424	0.438	0.585
		Cor	0.688	0.708	0.798
	GBM	Predictive deviance	0.265	0.260	0.239
		Cor	0.831	0.840	0.857
Bryophytes	GLM	Adjusted R <sup>2</sup>	0.329	0.350	0.367
		Cor	0.539	0.569	0.589
	GAM	Adjusted R <sup>2</sup>	0.378	0.405	0.405
		Cor	0.591	0.614	0.619
	GBM	Predictive deviance	0.497	0.501	0.493
		Cor	0.664	0.662	0.668
Lichen	GLM	Adjusted R <sup>2</sup>	0.439	0.505	0.577
		Cor	0.697	0.727	0.769
	GAM	Adjusted R <sup>2</sup>	0.479	0.538	0.612
		Cor	0.722	0.749	0.791
	GBM	Predictive deviance	0.372	0.359	0.338
		Cor	0.807	0.818	0.826

**Appendix 4: Correlograms for observed species richness and residuals of the models for vascular species, bryophytes and lichens.**

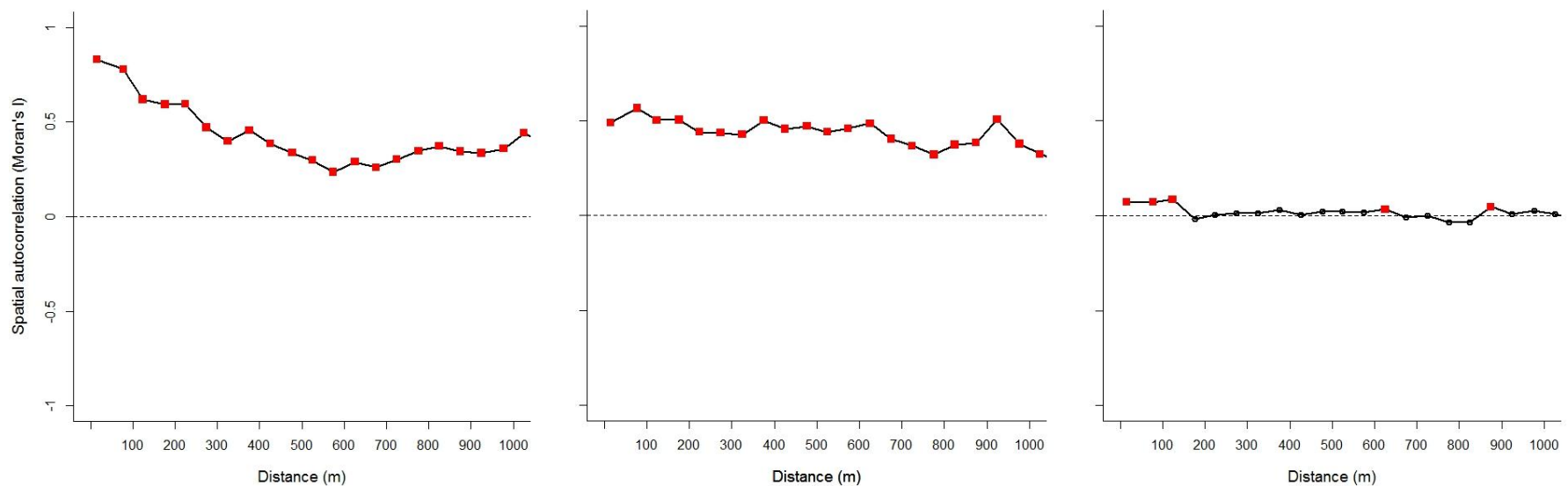


Fig A1. Spatial autocorrelation, expressed as Moran's *I* values for observed species richness of vascular plant species (left), bryophytes (middle) and lichens (right). Values are with 50 meter lags for the range of 1000 meters. Red dot indicates statistical significance ( $0,01 < p < 0,05$ ).

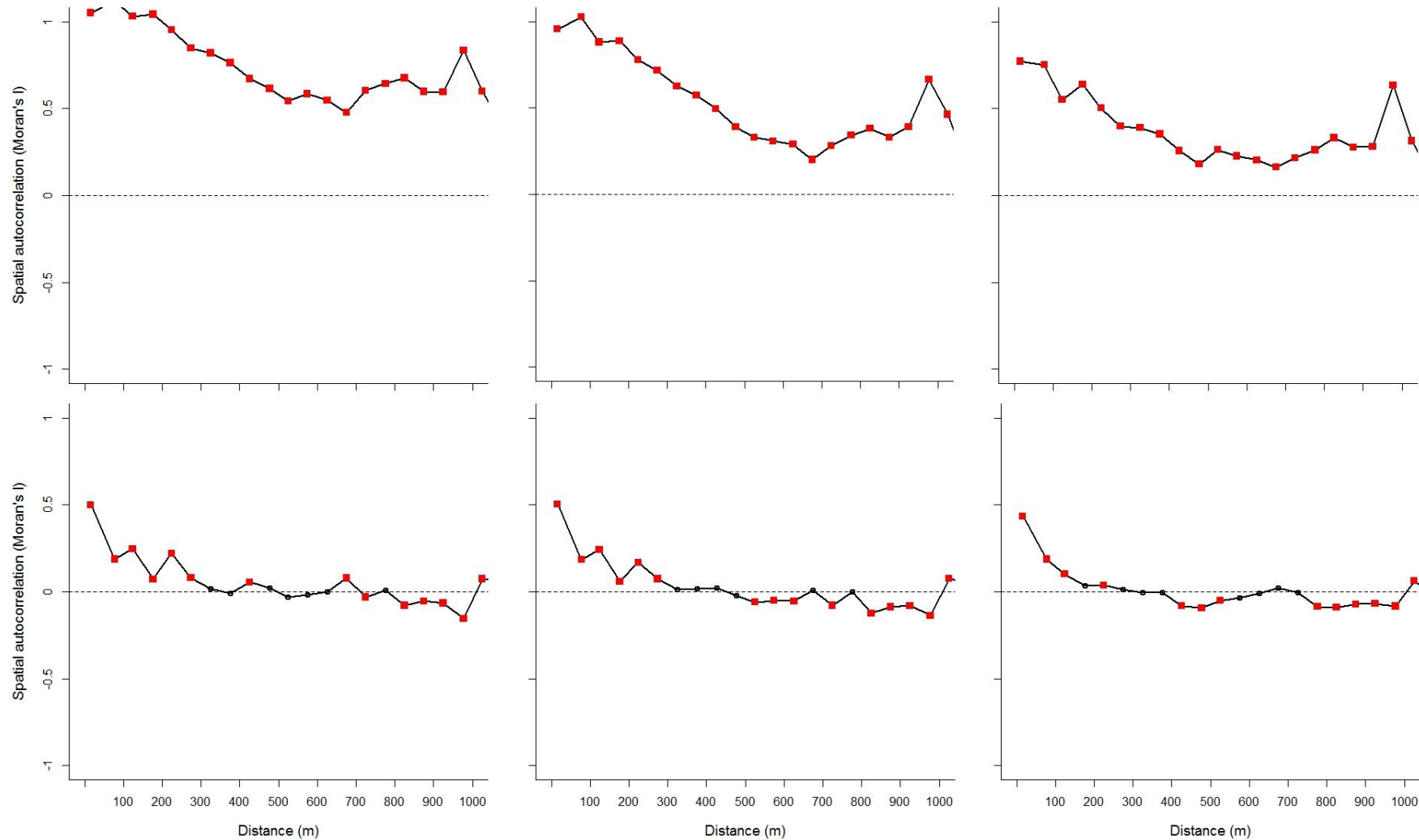


Fig A2. Spatial autocorrelation, expressed as Moran's  $I$  values for model residuals of stacked species distribution models (SSDM; top row) and macroecological models (MEM; bottom row) of vascular plant species based on climatic predictors (left column), climatic and abiotic predictors (center column) and climatic, abiotic and biotic predictors (right column). Values are with 50 metre lags for the range of 1000 metres. Red dot indicates statistical significance ( $0,01 < p < 0,05$ ).

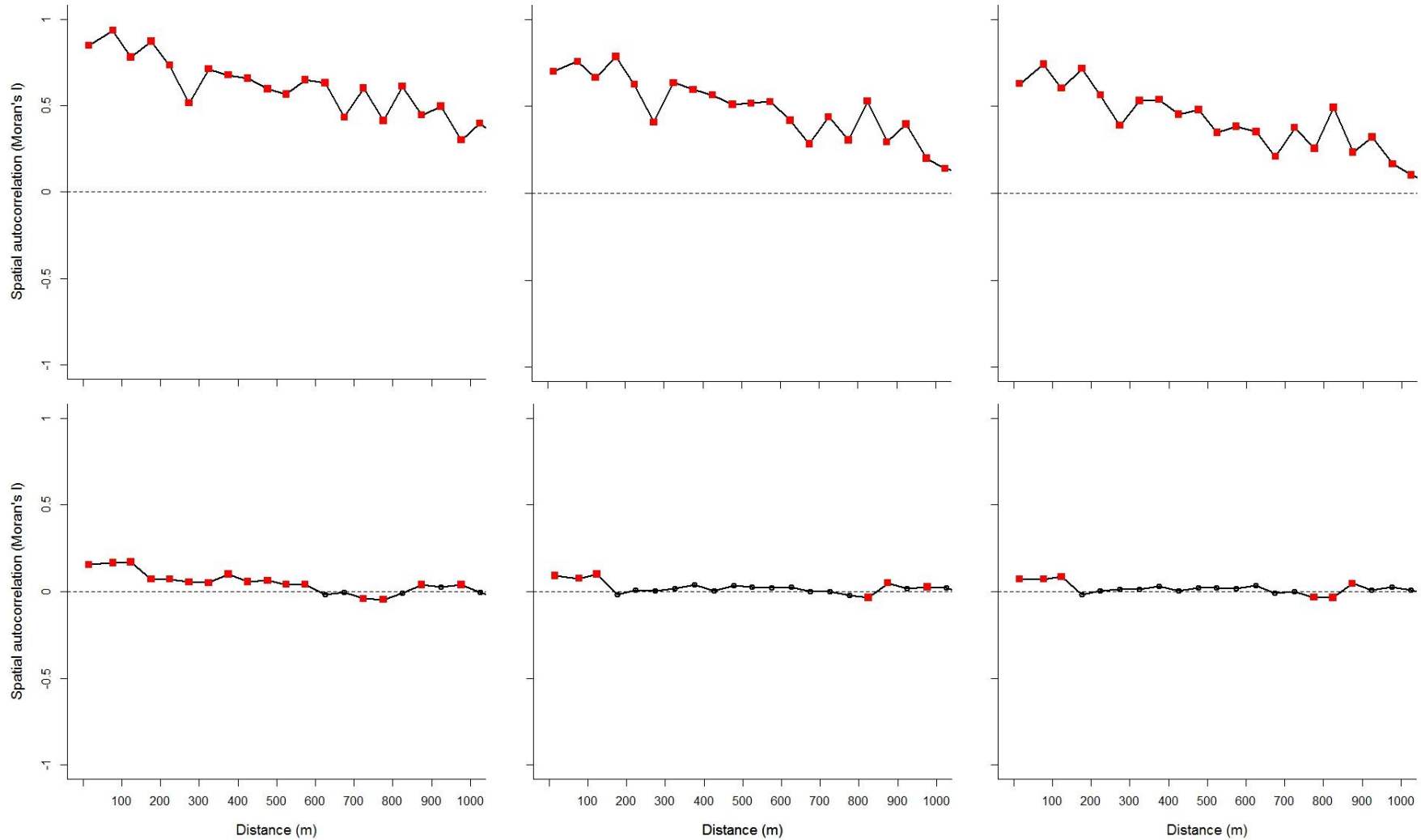


Fig A3. Spatial autocorrelation, expressed as Moran's  $I$  values for model residuals of stacked species distribution models (SSDM; top row) and macroecological models (MEM; bottom row) of bryophytes based on climatic predictors (left column), climatic and abiotic predictors (center column) and climatic, abiotic and biotic predictors (right column). Values are with 50 metre lags for the range of 1000 metres. Red dot indicates statistical significance ( $0.01 < p < 0.05$ ).

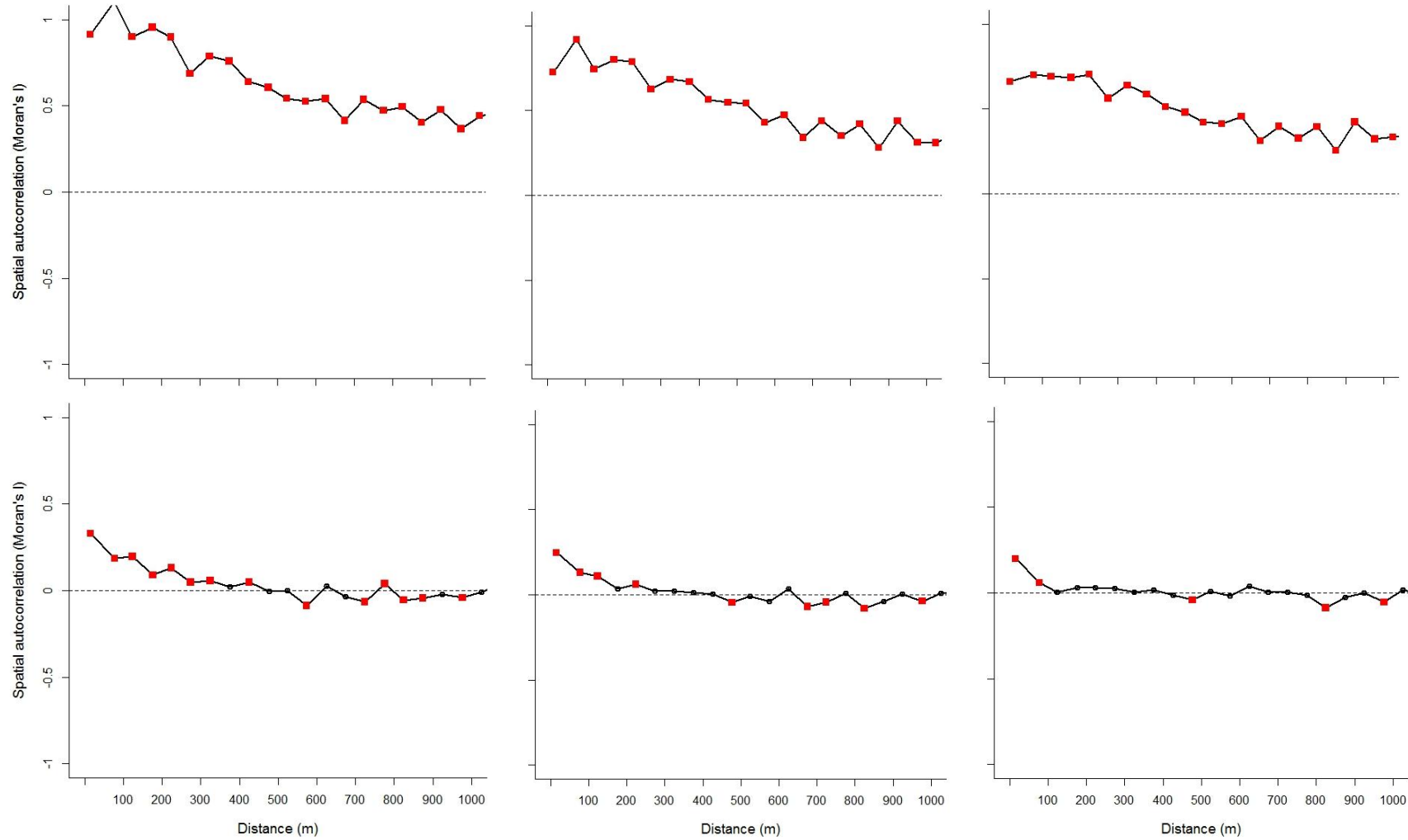


Fig A4. Spatial autocorrelation, expressed as Moran's  $I$  values for model residuals of stacked species distribution models (SSDM; top row) and macroecological models (MEM; bottom row) of lichens based on climatic predictors (left column), climatic and abiotic predictors (center column) and climatic, abiotic and biotic predictors (right column). Values are with 50 metre lags for the range of 1000 metres. Red dot indicates statistical significance ( $0,01 < p < 0,05$ ).