METHOD

Global Ecology and Biogeography | A Journal of Macroecology

WILEY

Bayesian joint species distribution model selection for community-level prediction

Malcolm S. Itter[1,2] | **Elina Kaarlejärvi**[2] | **Anna-Liisa Laine**[2] | **Leena Hamberg**[3] | **Tiina Tonteri**[3] | **Jarno Vanhatalo**[2,4]

[1]Department of Environmental Conservation, University of Massachusetts Amherst, Amherst, Massachusetts, USA

[2]Research Centre for Ecological Change, Organismal and Evolutionary Biology Research Program, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland

[3]Natural Resources Institute Finland, Helsinki, Finland

[4]Department of Mathematics and Statistics, Faculty of Science, University of Helsinki, Helsinki, Finland

Correspondence

Malcolm S. Itter, Department of Environmental Conservation, University of Massachusetts Amherst, 160 Holdsworth Way, Amherst 01003, MA, USA.
Email: mitter@umass.edu

Funding information

Suomen Kulttuurirahasto; Jane ja Aatos Erkon Säätiö; Research Council of Finland; Grant/Award Number: 317255

Handling Editor: Antoine Guisan

Abstract

Aim: Joint species distribution models (JSDMs) are an important tool for predicting ecosystem diversity and function under global change. The growing complexity of modern JSDMs necessitates careful model selection tailored to the challenges of community prediction under novel conditions (i.e., transferable models). Common approaches to evaluate the performance of JSDMs for community-level prediction are based on individual species predictions that do not account for the species correlation structures inherent in JSDMs. Here, we formalize a Bayesian model selection approach that accounts for species correlation structures and apply it to compare the community-level predictive performance of alternative JSDMs across broad environmental gradients emulating transferable applications.

Innovation: We connect the evaluation of JSDM predictions to Bayesian model selection theory under which the log score is the preferred performance measure for probabilistic prediction. We define the joint log score for community-level prediction and distinguish it from more commonly applied JSDM evaluation metrics. We then apply the joint community log score to evaluate predictions of 1918 out-of-sample boreal forest understory communities spanning 39 species generated using a novel multinomial JSDM framework that supports alternative species correlation structures: independent, compositional dependence and residual dependence.

Main conclusions: The best performing JSDM included all observed environmental variables and compositional dependence modelled using a multinomial likelihood. The addition of flexible residual species correlations improved model predictions only within JSDMs applying a reduced set of environmental variables highlighting potential confounding between unobserved environmental conditions and residual species dependence. The best performing JSDM was consistent across successional and bioclimatic gradients regardless of whether interest was in species- or community-level prediction. Our study demonstrates the utility of the joint community log score to compare the predictive performance of JSDMs and highlights the importance of accounting for species dependence when interest is in community composition under novel conditions.

KEYWORDS

Bayesian model selection, boreal forest, global change, log score, model transferability, multinomial likelihood, probabilistic prediction, species dependence

1 | INTRODUCTION

Joint species distribution models (JSDMs) are an important tool to predict community composition in an era of global change (Blowes et al., 2019). Joint species distribution model predictions may be used to advance the understanding of biodiversity change (Antão et al., 2022), identify at risk populations or natural communities (Araújo et al., 2011) and inform conservation management decisions (Guisan et al., 2013; Pollock et al., 2020). Joint species distribution models differ from traditional species distribution models (SDMs) in that they model the joint occurrence or abundance of species as a function of fixed environmental effects and residual species correlations that account for statistical dependence among species (Clark et al., 2014; Pollock et al., 2014; Warton et al., 2015). In addition to quantifying residual species correlations, contemporary JSDMs apply a hierarchical structure to estimate species-specific responses to environmental variables allowing for sharing of information among species (Clark et al., 2017; Ovaskainen et al., 2017; Pollock et al., 2014; Warton et al., 2015). Further, they quantify spatio-temporal residual autocorrelation among sample units to the extent that it is present in model data (Ovaskainen et al., 2016; Taylor-Rodríguez et al., 2017; Tikhonov, Duan, et al., 2020). While the increased complexity of contemporary JSDMs may contribute to more accurate predictions of species occurrence/abundance, it also raises the risk of model overfitting and increased uncertainty due to poorly identified model parameters. These risks are particularly high when JSDMs are applied to make predictions under novel conditions (i.e., transferable applications) such as predicting changes in species distributions under global change (Yates et al., 2018). The growing complexity of contemporary JSDMs combined with the challenges of making predictions under novel conditions necessitates careful model selection to ensure meaningful predictions of communities under global change (Roberts et al., 2017).

A key consideration when conducting model selection for JSDMs is the prediction type of interest. As highlighted in recent work (Wilkinson et al., 2021), JSDMs allow for several types of prediction including individual species (marginal), all species simultaneously (joint), and simultaneous prediction of a subset of species conditional on a second, disjoint subset (conditional). JSDMs are expected to provide the greatest improvement in prediction (relative to SDMs) under the joint and conditional cases given that they account for the statistical dependence among species (Poggiato et al., 2021). This improvement is illustrated by an early JSDM study wherein summing the predicted abundance of tree species among independent SDMs (i.e., stacked SDMs) led to unrealistically high total abundance values relative to a JSDM, which accounted for dependence among species induced by fixed growing space (Clark et al., 2014). In contrast, SDMs and JSDMs are expected to yield similar or, under certain model specifications, identical marginal predictions of species occurrence or abundance (Poggiato et al., 2021; Wilkinson et al., 2019, 2021). Despite this expectation, the predictive performance of JSDMs is often evaluated based on the prediction of individual species even in cases where interest is in community composition (Caradima et al., 2019; Zurell et al., 2020).

Recent work describes the diversity of JSDM prediction types along with statistical methods to evaluate their accuracy (Broms et al., 2016; Wilkinson et al., 2021). In general, measures of predictive accuracy should be tailored to the type of predictions a model will be applied to make (Gelman et al., 2014; Hooten & Hobbs, 2015). In the context of JSDMs, this means that if interest is in community composition, models should be evaluated based on joint predictions that account for statistical dependence among co-occurring species (Poggiato et al., 2021; Wilkinson et al., 2021). In this study, we formalize and demonstrate an approach for evaluating community-level predictions applying Bayesian JSDMs (i.e., joint prediction type as defined in Wilkinson et al., 2021). We focus on Bayesian JSDMs given both their prevalence (Clark et al., 2017; Hui, 2016; Ovaskainen et al., 2017) and the capacity of Bayesian models to accommodate and quantify varying sources of uncertainty in ecological data and models (Cressie et al., 2009).

Statistical measures to evaluate community-level predictions (i.e., simultaneous predictions of all species in a community) include community dissimilarity indices (e.g., Bray–Curtis dissimilarity and Jaccard distance) and likelihood-based methods including the joint log score (Wilkinson et al., 2021). In some cases, measures commonly used to assess predictions of individual species including the root mean square error (RMSE) or the coefficient of determination (R-squared) are applied to univariate summaries of community composition such as species richness or diversity indices (Norberg et al., 2019). Community dissimilarity indices are well known in ecology and have been widely applied to assess the community-level predictive performance of JSDMs or stacked SDMs (Broms et al., 2016; D'Amen et al., 2015; Maguire et al., 2016; Norberg et al., 2019; Wilkinson et al., 2021). The joint log score is less well known in ecology and has been applied to assess the predictive performance of JSDMs in only a small number of cases (Harris, 2015; Ingram et al., 2020; Vanhatalo et al., 2020; Wilkinson et al., 2021). Despite its limited application to JSDMs, the log score is widely applied in Bayesian model selection when the objective is prediction of future data given important theoretical properties accounting for uncertainty in model parameters and the unknown, true data generating mechanism (Hooten & Hobbs, 2015).

Here, we describe the log score in the context of JSDMs. We define different ways it can be quantified to assess community-level predictions specifying what is meant by an independent versus joint log score (Wilkinson et al., 2021). We then apply the log score to compare alternative JSDMs and identify the best performing model for community-level prediction of 1918 out-of-sample boreal forest understory plant communities comprising 39 species. The applied model framework extends existing JSDM approaches (e.g., the Hierarchical Model of Species Communities; Ovaskainen et al., 2017) to account for compositional dependence commonly encountered in natural community data in addition to residual correlation among species. Candidate models include nested sets of environmental variables and alternative species dependence structures. We further assess model transferability by evaluating the performance of alternative models in predicting individual species and

communities across out-of-sample data partitions spanning broad successional and bioclimatic gradients. Finally, we compare the log score to other commonly used measures to assess community-level predictions of JSDMs. In doing so, we demonstrate the flexibility and utility of the log score for Bayesian JSDM selection and highlight the role of species correlations in community-level predictions.

2 | METHODS

2.1 | Log score calculation

In the context of Bayesian model prediction, the log score refers to the log predictive density. It is the log of the probability distribution of new data integrating over uncertainty in model parameters (Gelman et al., 2014; Hooten & Hobbs, 2015). The log score is used to evaluate the out-of-sample predictive performance of models. That is, how well the model predicts data not used in model fitting. The log score evaluates the posterior predictive density at out-of-sample data points reflecting uncertainty and variability in model predictions. It is a preferred metric for Bayesian model selection when the goal is probabilistic prediction (Hooten & Hobbs, 2015). Probabilistic prediction refers to the case where interest is in the probability distribution of new data as opposed to point prediction where interest is in a point estimate of new data, most commonly the mean (Gelman et al., 2014).

We first define the general form of the log score before describing the different ways it can be quantified in the context of JSDMs. We adapt definitions and notation from Hooten and Hobbs (2015) to facilitate connection to a comprehensive review of Bayesian model selection methods in ecology. Let y^{obs} and y^{oos} denote observed and out-of-sample data, respectively. Further, let θ denote unknown model parameters. Finally, let $[x]$ denote a generic probability distribution of a random variable x. The log score as defined in Hooten and Hobbs (2015) is given by,

$$\log[y^{oos}|y^{obs}] = \log \int [y^{oos}|\theta][\theta|y^{obs}]d\theta. \quad (1)$$

Here, $[\theta|y^{obs}]$ is the posterior distribution of the model parameters, while $[y^{oos}|\theta]$ is the likelihood of out-of-sample data conditional on the model parameters (the likelihood function follows from the model applied). The integral provides the posterior predictive density of out-of-sample data: $[y^{oos}|y^{obs}]$. In practice, the integral is approximated numerically applying Monte Carlo integration based on posterior samples of model parameters (θ).

In the context of JSDMs, the log score can be quantified for individual species when interest is in species-level prediction or for all species simultaneously when interest is in community-level prediction. We refer to the former as the 'species log score' and the latter as the 'community log score'. We focus on the community log score here given our objective is to evaluate the performance of JSDMs for community-level prediction. Definitions of the species log score can be found in a number of existing studies (Ingram et al., 2020;

Wilkinson et al., 2021). The community log score refers to the log of the combined likelihood for all species in a natural community. One method to quantify the community log score is to sum over the marginal log predictive density of each species in a community. This method has been referred to as the independent log score given its connection to stacked SDMs under which the predictions of each species are evaluated independently of one another (Wilkinson et al., 2021). An alternative method to quantify the community log score is to utilize the joint log predictive density of all species in a community taking advantage of the joint likelihood provided by the JSDM. This method has been referred to as the joint log score given its use of the joint log-likelihood (Harris, 2015; Ingram et al., 2020; Wilkinson et al., 2021).

To formalize the independent and joint community log scores (and demonstrate the difference between the two), we first define a general JSDM fit using observations of m species across n sample sites denoted as y_{ij}. We use j to index species ($j = 1, \ldots, m$) and i to index sample sites ($i = 1, \ldots, n$). Here, y_{ij} may be observations of presence/absence ($y_{ij} = 0$; $y_{ij} = 1$) or abundance ($0 \leqslant y_{ij} < \infty$). Adopting a generalized linear mixed model framework consistent with current JSDM approaches (Ovaskainen et al., 2017; Poggiato et al., 2021; Pollock et al., 2014; Warton et al., 2015; Wilkinson et al., 2019), observations are modelled conditionally on a normally distributed, latent variable denoted as z_{ij} applying an appropriate link function. In the case of count data, for example, we apply a Poisson likelihood and a log link function such that the likelihood is defined as,

$$y_{ij} | z_{ij} \sim \text{Poisson}(\exp(z_{ij})) \quad (2)$$

with $\exp(z_{ij})$ representing the inverse link function mapping the continuous latent variables modelled on the log scale back to the scale of the data. In the case of normally distributed data, an identity link is applied, while in the case of presence/absence data, a logit or probit link is applied (see Ovaskainen et al., 2017; Poggiato et al., 2021; Wilkinson et al., 2019, for detailed examples).

Regardless of the observation type and link function applied, the statistical dependence among species is modelled through the latent variable. Specifically,

$$z_{ij} = \mathbf{x}_i^T \boldsymbol{\beta}_j + w_{ij} \quad (3)$$

where $\mathbf{x}_i$ is a vector of environmental variables observed at site i (note bold font is used to denote a vector and the 'T' superscript denotes the transpose of a vector or matrix), $\boldsymbol{\beta}_j$ is an equally-sized vector of environmental responses for species j, and w_{ij} is a correlated error term accounting for variation not attributable to observed environmental conditions. By correlated, we mean that w_{ij} arises from a multivariate normal distribution that models statistical dependence among species. Grouping the error term for each species into an m-dimensional vector, $\mathbf{w}_i = (w_{i1}, w_{i2}, \ldots, w_{im})^T$, we have,

$$\mathbf{w}_i \sim \text{MVNorm}(\mathbf{0}, \Sigma) \quad (4)$$

where $\mathbf{0}$ is an m-dimensional vector of zeros and Σ is an m-dimensional covariance matrix expressing the statistical dependence among

species. Following from Equations (3) and (4), the joint distribution of latent variables for all species at a site i is given by,

$$\mathbf{z}_i \sim \text{MVNorm}(\mathbf{Bx}_i, \Sigma) \qquad (5)$$

where $\mathbf{z}_i = (z_{i1}, \ldots, z_{im})^{\mathsf{T}}$ is the collection of latent variables for all m species and $\mathbf{B}$ is a matrix containing the responses of each species to the observed environmental variables in $\mathbf{x}_i$. Following from properties of the multivariate normal distribution, the marginal distribution of the latent variable for a single species j at site i is given by,

$$z_{ij} \sim \text{Norm}\left(\mathbf{x}_i^{\mathsf{T}} \beta_j, \sigma_j^2\right) \qquad (6)$$

where σ_j^2 is the variance for species j (the diagonal elements of Σ). Note that the covariances among species are not reflected in the marginal distribution of the latent variable for a single species—the distribution depends only on each species' response to the observed environment and its species-specific variance.

The difference between the independent versus the joint community log score is how each utilizes estimates of species covariance within Σ when calculating the likelihood of out-of-sample data, $[y^{\text{oos}} | \theta]$ in Equation (1). Continuing with the same general JSDM as above, suppose we fit the model using n^{obs} sample sites and evaluate community-level predictions at n^{oos} out-of-sample sites. The joint community log score is calculated applying the joint likelihood of species at an out-of-sample site conditional on the joint distribution of the latent variables at that site,

$$\text{Joint community likelihood} = [\mathbf{y}_i^{\text{oos}} | \mathbf{z}_i] \qquad (7)$$

where we have grouped out-of-sample observations for each species, $\mathbf{y}_i^{\text{oos}} = (y_{i1}^{\text{oos}}, \ldots, y_{im}^{\text{oos}})^{\mathsf{T}}$. By conditioning on the joint distribution of the latent variables as defined in Equation (5), the joint community log score reflects the statistical dependence among species as estimated within Σ. The independent community log score also calculates the combined likelihood for all species in an out-of-sample site, but does so conditional on the marginal distribution of each latent variable,

$$\text{Independent community likelihood} = \prod_{j=1}^{m} \left[y_{ij}^{\text{oos}} | z_{ij} \right] \qquad (8)$$

such that the combined likelihood for all species is calculated as the product of their marginal likelihoods. By conditioning on the marginal distribution of each latent variable as defined in Equation (6), the independent community log score depends only on species-level variances and does not account for the covariance among species as estimated with Σ. In general, when interest is in community-level prediction applying a JSDM that includes statistical dependence among species (as in Equation 5), the joint community log score is preferred since it accounts for the modelled dependence structure when evaluating model predictions. In the remaining sections, we demonstrate the application of the joint community log score in Bayesian JSDM selection using a large forest understory plant community data set. Before doing so, we extend the general JSDM

framework defined above to allow for a multinomial likelihood for use with compositional data common in community ecology. Note that the use of the joint community log score and the broader model selection approach presented here will work for any JSDM and data type.

2.2 | Multinomial JSDM for compositional data

Compositional community data are frequently encountered in natural community surveys wherein a fixed area or volume is used to sample the count or relative abundance of each species present. Negative statistical dependence in species observations is common in this setting given some fixed total available growing space or resource pool. For example, the fixed area of a survey plot might be taken as the total available growing space containing a limited amount of resources. When all available growing space and/or resources are utilized, the relative abundance of a given species cannot increase without a corresponding decrease in the total abundance of all other species present. A multinomial likelihood is a natural choice for modelling compositional data given that it estimates the joint abundance of species conditional on a total abundance value reflecting the available growing space or fixed resources at a site.

We extend the JSDM framework defined above to allow for a multinomial likelihood. Similar versions of the multinomial model have been developed and applied to model microbiome data as well as plant compositional cover data, but have not been formally integrated within a JSDM framework (Damgaard, 2015, 2018; Xia et al., 2013). Under a multinomial likelihood, a species' relative abundance is estimated as a function of its expected success relative to the expected success of all species given the observed environment. We apply the same notation for a local community as above with $\mathbf{y}_i$ denoting an m-dimensional vector of species' relative abundances. The relative abundance of each species is modelled jointly applying a multinomial likelihood conditional on species-specific probabilities denoted by π_{ij}. Species-specific probabilities are estimated as a function of the normally distributed, latent variables $\mathbf{z}_i$. Specifically, the relative abundance of all species at site i is modelled as,

$$\mathbf{y}_i | \mathbf{z}_i \sim \text{Multinom}(y_{i\cdot}, \pi_i) \qquad (9)$$

where $\pi_i = (\pi_{i1}, \ldots, \pi_{im})^{\mathsf{T}}$ defines the probabilities for each species and $y_{i\cdot}$ is the total abundance at site i across all m species reflecting available growing space. Species-specific probabilities are modelled applying a softmax link function of the latent variables,

$$\pi_{ij} = \frac{\exp z_{ij}}{\sum_{j=1}^{m} \exp z_{ij}} \qquad (10)$$

which ensures each $\pi_{ij} \in (0, 1)$ and $\sum_{j=1}^{m} \pi_{ij} = 1$. The latent variables are modelled as before according to Equation (3).

As is common in existing JSDM frameworks (Poggiato et al., 2021; Wilkinson et al., 2019), species-specific responses to the observed environment (β_j's) are estimated applying a hierarchical prior allowing for partial pooling of data among species. Specifically,

we model the response of species j to the kth environmental variable ($k = 1, \ldots, p$) as,

$$\beta_{jk} \sim \text{Norm}\left(\mu_k, \tau_k^2\right) \qquad (11)$$

where μ_k is the mean response to the kth variable across all species, and τ_k^2 is the variance in species-level responses.

The multinomial likelihood accounts for negative statistical dependence among species given some fixed total available growing space. Specifically, the multinomial correlation between any two species j and j' ($j \neq j'$) is given by,

$$\text{Cor}\left(y_{ij}, y_{ij'}\right) = -\sqrt{\frac{\pi_{ij}\pi_{ij'}}{\left(1 - \pi_{ij}\right)\left(1 - \pi_{ij'}\right)}}. \qquad (12)$$

Note that multinomial correlations are constrained to be between −1.0 and 0.0 and account only for the fact that as the relative abundance of one species increases, the total relative abundance of all other species must decrease for a fixed community size ($\pi_{ij} = 1 - \sum_{j' \neq j} \pi_{ij'}$). We refer to the narrow form of species dependence captured by the multinomial likelihood as 'compositional dependence'. In contrast, the correlated residual error term ($\mathbf{w}_i$) defined in Equation (4) is flexible and supports both positive and negative pairwise correlations among species. We refer to the flexible species dependence arising through the correlated residual error term ($\mathbf{w}_i$) as 'residual dependence' as it captures residual correlation among species not accounted for by the effects of the observed environment.

2.3 | Natural community data

We applied the multinomial JSDM to observations of Finnish boreal forest understory communities collected as part of national vegetation inventories (Itter, 2023). Understory vegetation was surveyed in 1985–1986, 1995, and 2006 across 1700 unique plot locations established on mineral-soil in forested land. Plots were part of the systematic sampling network connected to the 8th Finnish National Forest Inventory (Reinikainen et al., 2000). This network consisted of clusters located 16 km apart in southern Finland, and 24 and 32 km apart in northern Finland along east–west and north–south axes, respectively. Each cluster consisted of four linearly located sampling sites 400 m apart in southern Finland and three sampling sites 600 m apart in northern Finland. Data include 1494 sites measured in 1985–1986, 1673 sites in 1995, and 435 sites in 2006 (3602 sites total). The survey in 2006 was part of the BioSoil project carried out under the Forest Focus scheme, a subset of the pan-European International Co-operative Programme on Assessment and Monitoring of Air Pollution Effects on Forests plot network (Level I, Lorenz & Fischer, 2013). This survey included only one site per cluster (hence the smaller sample size), but the spatial extent of the inventory was comparable to the 1985 and 1995 surveys and spanned the entirety of Finland (Figure S1). The surveys included observations of a total of 380 understory plant species. To ensure that meaningful values of species-specific model parameters could be estimated, we focussed only on those species which occurred in at least five percent of sites in each inventory year reducing the total number of species to 39 (a common practice in high-dimensional JSDM settings, see Clark et al., 2017).

In all three surveys, vascular plant species were identified and each species' per cent cover (0.1% to 100%) was visually estimated within four permanent 2 m^2 quadrats located 5 m apart within each site. There was no requirement that the per cent cover values of all species sum to 100% given that a quadrat may not be fully occupied or may have a total cover greater than 100% if plants over top one another (a common occurrence in understory communities). To facilitate modelling, per cent cover values were treated as numeric counts (percent cover values were rounded up to the nearest integer) representing the relative abundance of each species with the total abundance given by the sum across all species. When modelling understory communities, we calculated the average per cent cover across the four sampled quadrats before converting percentages to estimates of relative and total abundance per site.

A number of environmental variables characterizing local conditions were measured along with understory community composition. The basal area of overstory trees at each site was estimated using measurements of stem diameter at breast height (Tomppo et al., 2011) and provides information on past forest management given basal area is inversely related to past harvest intensity (Figure S2). Shrub cover at each site was quantified as the projected per cent cover of shrubs and 0.5–1.5 m tall trees within a 9.8 m radius circular plot centred on the permanent vegetation survey site. Soil fertility at each site was determined in the field using six ordinal classes based on vegetation (Cajander, 1949; Tomppo et al., 2011). For the purposes of this study, we re-classified these into two groups representing 'high' and 'low' fertility. Finally, growing degree days were estimated as the average annual sum of daily mean temperatures exceeding +5°C per site over the decade preceding the inventory year based on 10 km^2-resolution interpolated daily temperature values modelled by the Finnish Meteorological Institute (Venäläinen et al., 2005).

2.4 | Model selection

2.4.1 | Candidate models

We defined seven candidate models representing alternative sets of environmental variables and alternative species dependence types. Three alternative sets of environmental variables were considered. A 'stochastic' set included an intercept term alone to estimate the mean abundance of each species across all sites. A 'management' set included an intercept term and basal area per hectare (BA) to account for time since forest overstory harvest. An 'environment' set

included an intercept term and BA along with growing degree days (GDD), soil fertility (SoilFert) and per cent shrub cover (ShrubCover) including two- and three-way interactions between BA, GDD and SoilFert to account for past management and local environmental conditions.

Three species dependence types were considered. An 'independent' type assumed no correlation among species and applied an independent log-normal Poisson likelihood rather than the multinomial likelihood defined in Equation (9). Specifically, the relative abundance of all species in a local community under the independent type was modelled as,

$$\mathbf{y}_i \mid \mathbf{z}_i \sim \prod_{j=1}^{m} \mathrm{Pois}\left(\exp z_{ij}\right) \tag{13}$$

with z_{ij} as defined in in Equation (3), but with an important difference in the way the residual random error terms (w_{ij}) were estimated. Specifically, under the independent species dependence type, the random error terms were modelled as independent, normally distributed random variables,

$$w_{ij} \sim \mathrm{Norm}\left(0, \sigma_j^2\right) \tag{14}$$

such that there was no statistical dependence among species estimated. A 'compositional' type modelled species dependence applying the multinomial likelihood (Equation 9), but did not estimate correlation in the residual random error terms, again estimating these as independent, normal random variables according to Equation (14). Note that this approach captures compositional dependence among species alone. Finally, a 'residual' species dependence type applied the multinomial likelihood as defined in Equation (9) and estimated residual correlation among species modelling the random error terms according to Equation (4). Note that the residual type accounts for both compositional and residual dependence among species.

The seven candidate models were named according to the environmental variable set and species dependence type applied within the model: (1) Stochastic-independent, (2) Management-independent, (3) Environment-independent, (4) Management-compositional, (5) Environment-compositional, (6) Management-residual and (7) Environment-residual (Table 1). We

considered the Stochastic-independent model as a null model for comparison with the more ecologically realistic models 2–7. We considered quadratic terms for BA and GDD during the exploratory phase of our analysis, but found limited evidence for such quadratic relationships. All candidate models were fit applying the Hierarchical Model of Species Communities (HMSC) R package (Ovaskainen et al., 2017; Tikhonov et al., 2021; Tikhonov, Opedal, et al., 2020). The HMSC package applies the same JSDM structure to the one presented here, but does not include a multinomial likelihood (Equation 9). To address this, we modified the HMSC package to apply the well-established Poisson approximation to the multinomial, which induces a multinomial likelihood for the latent variables ($\mathbf{z}_i$), which uniquely define species-specific probabilities (Baker, 1994; McCullagh & Nelder, 1989). Details on the Poisson approximation to the multinomial and its implementation within HMSC are provided in the Supplemental material.

2.4.2 | Model assessment

We compared candidate models based on their ability to predict out-of-sample understory communities. Model predictions were evaluated using the joint community log score calculated using the joint community likelihood as defined in Equation (7). We applied a sequential approach to construct in-sample and out-of-sample data using subsets of the 1985, 1995 and 2006 understory vegetation inventories. Specifically, we fit and evaluated each candidate model twice: once using the 2006 inventory data alone to predict 435 communities within the 1995 data (1995|2006), and once using the 1995 and 2006 inventory data combined to predict 1483 communities within the 1985 data (1985|1995, 2006). We then summed the calculated community log scores across all 1918 (435 + 1483) out-of-sample communities to assess predictive performance. The statistical basis for our sequential out-of-sample predictive approach is provided in the Supplemental material.

We chose to assess models based on the joint community log score given our interest is in probabilistic community-level prediction. Potential alternatives commonly applied to assess community-level JSDM predictions include the independent community log score, community dissimilarity metrics, and point estimates of

TABLE 1 Summary of candidate models.

Model	Environmental variables ($\mathbf{x}_i$)	Species dependence	Data model
Stochastic-independent	None	Independent	Log-normal Poisson
Management-independent	BA	Independent	Log-normal Poisson
Management-compositional	BA	Compositional	Multinomial
Management-residual	BA	Compositional + residual	Multinomial
Environment-independent	BA, SoilFert, GDD, ShrubCover	Independent	Log-normal Poisson
Environment-compositional	BA, SoilFert, GDD, ShrubCover	Compositional	Multinomial
Environment-residual	BA, SoilFert, GDD, ShrubCover	Compositional + residual	Multinomial

Variable abbreviations are defined as follows: BA, basal area per hectare; GDD, growing degree days; ShrubCover, percent shrub cover; SoilFert, soil fertility.

species richness or diversity (Wilkinson et al., 2021). In order to compare the model selection results applying the joint community log score to other measures of JSDM predictive performance, we calculated the independent community log score using the independent community likelihood (Equation 8), the Jaccard community dissimilarity index (Chao et al., 2005), and the RMSE of the Shannon true diversity index (Jost, 2006) for the 1918 out-of-sample communities. Details on the calculation of each additional evaluation metric are provided in the Supplemental material.

Lastly, we compared the performance of the seven candidate models to predict individual species and communities across out-of-sample data partitions representing different combinations of successional stage and climate in order to assess model transferability (i.e., predictive performance under novel conditions; Roberts et al., 2017). Specifically, we partitioned the 1985 and 1995 inventory data based on BA and regional bioclimatic conditions. The BA variable was divided into three ranges to represent management history: 0–10, 10–30 and >30 m^2/ha. The low BA range is indicative of understory communities early in ecological succession, while the high BA range is indicative of later successional communities given prolonged time since forest overstory harvest. We utilized three bioclimatic zones spanning the latitudinal gradient of Finland (south, mid and north boreal) to represent regional climatic conditions (Table S1, Figure S1; Ahti et al., 1968). This resulted in eight data partitions (the highest BA range was dropped for the north boreal zone given lack of sufficient out-of-sample data) with a minimum of 54 communities (midboreal, high basal area) and a maximum of 565 communities (southboreal, mid basal area; Table S2). Within each partition, we estimated the mean species log score for all 39 species (to assess species-level prediction) and the mean joint community log score (to assess community-level prediction).

3 | RESULTS

Joint community log scores for each candidate model are provided in Figure 1. Based on the log score values, the Environment compositional model including the full set of environmental predictors and compositional dependence among species had the highest community-level predictive performance across the 1918 out-of-sample communities. In terms of predictor variables, the top three performing models all included the complete set of environmental variables (environment models). The next best three models included forest overstory density (BA) as the only environmental variable (management models). The Stochastic-independent model without any environmental variables (included as a null model) had the poorest predictive performance. In terms of species dependence type, the top-performing model included compositional species dependence, while the second-best performing model assumed statistical independence among species. The Environment-residual model with both compositional and residual species dependence had the poorest performance of the three environment models although its log

score value was quite close to that of the Environment-independent model. In-sample and out-of-sample posterior predictive model checks under the best predicting Environment-compositional model revealed high correlation with observed data and species-level *R*-squared values exceeding 0.50 for the majority of species across all inventory years (Table S5, Figure S3; see Supplemental material for model checking details).

The relative ordering of candidate model predictive performance was largely maintained across the alternative JSDM evaluation metrics applied including the independent community log score, RMSE of the Shannon true diversity index, and the Jaccard community dissimilarity index (Table S3). The results of the independent community log score closely matched those of the joint community log score with the ordering of models remaining the same (Table S3). The results of the RMSE of community diversity indicated different predictive performance of the management models (Table S3). Specifically, the Management-compositional model had nearly identical predictive performance as the Environment-residual model, while the Management-independent and Management-residual models both had slightly poorer predictive performance than the Stochastic-independent model. All other RMSE results followed the joint community log score values with the Environment-compositional model having the best predictive performance in terms of community diversity. The three environment models were again preferred based on the Jaccard community dissimilarity index, followed by the three management models, and finally the Stochastic-independent model (Table S3). The relative ordering of predictive performance of alternative species dependence types within the environment and management models was residual > compositional > independent. The Environment-residual model had the best predictive performance based on the Jaccard community dissimilarity index values, however, the Environment-compositional model had near equal predictive performance.

Similar to the community log score results for the complete out-of-sample data (Figure 1), the Environment-compositional model had the highest community-level predictive performance across all out-of-sample data partitions (Figure 2). The differences in the predictive performance of the three environment models were small with 95 per cent confidence intervals for the three models overlapping in all partitions. There was clear separation of the environment models relative to the Stochastic-independent (null) model, but the 95 per cent confidence intervals for all seven candidate models overlapped in half of the out-of-sample data partitions (Figure 2). The Environment-compositional model also had the highest predictive performance for the majority of species within every out-of-sample data partition as assessed by the species log score (Table 2 and Figure 3). Depending on the data partition, the Environment-independent model or the Environment-residual model had the best predictive performance for the second-highest number of species (Table 2). Among these two, the Environment-independent model performed slightly better with the best species log scores for the second-highest number of species in six out-of eight partitions.
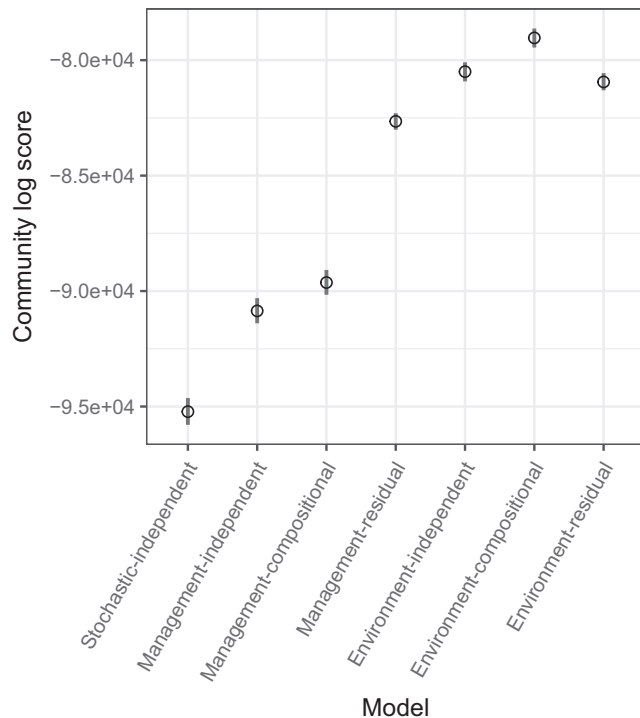
FIGURE 1 Community log scores for alternative models (as calculated by Equations S.4 or S.5). Points represent bootstrapped median values, while lines indicate bootstrapped 95 percent confidence intervals. Higher log score values indicate better out-of-sample prediction. Model names are defined in Table 1. Details on bootstrap approach are provided in Supplemental material.

4 | DISCUSSION

We found clear evidence for the inclusion of the complete set of environmental variables for the prediction of boreal forest understory communities based on the joint community log score results (Figure 1). However, the role of species dependence in community-level prediction was more nuanced. Species dependence was included in the candidate models in three ways: (1) no dependence (independent models), (2) compositional dependence alone (compositional models), (3) compositional and residual dependence allowing for both positive and negative species correlations (residual models; Table 1). We expected that the Environment-residual model would have the highest predictive performance given the flexibility of the correlated random error term allowing for both positive and negative associations among species as opposed to compositional dependence, which captures negative associations only. Contrary to this expectation, we found that the Environment-compositional model including the full set of environmental variables and compositional species dependence alone, had the best predictive performance as measured by its joint community log score (Figure 1). The Environment-residual model, which included both compositional and residual dependence, performed similarly to the Environment-independent model that assumed no species dependence.

Several factors may have contributed to the better predictive performance of the Environment-compositional model relative to the Environment-residual model. First, the correlated random error term (Equation 4) may have contributed to less accurate or more
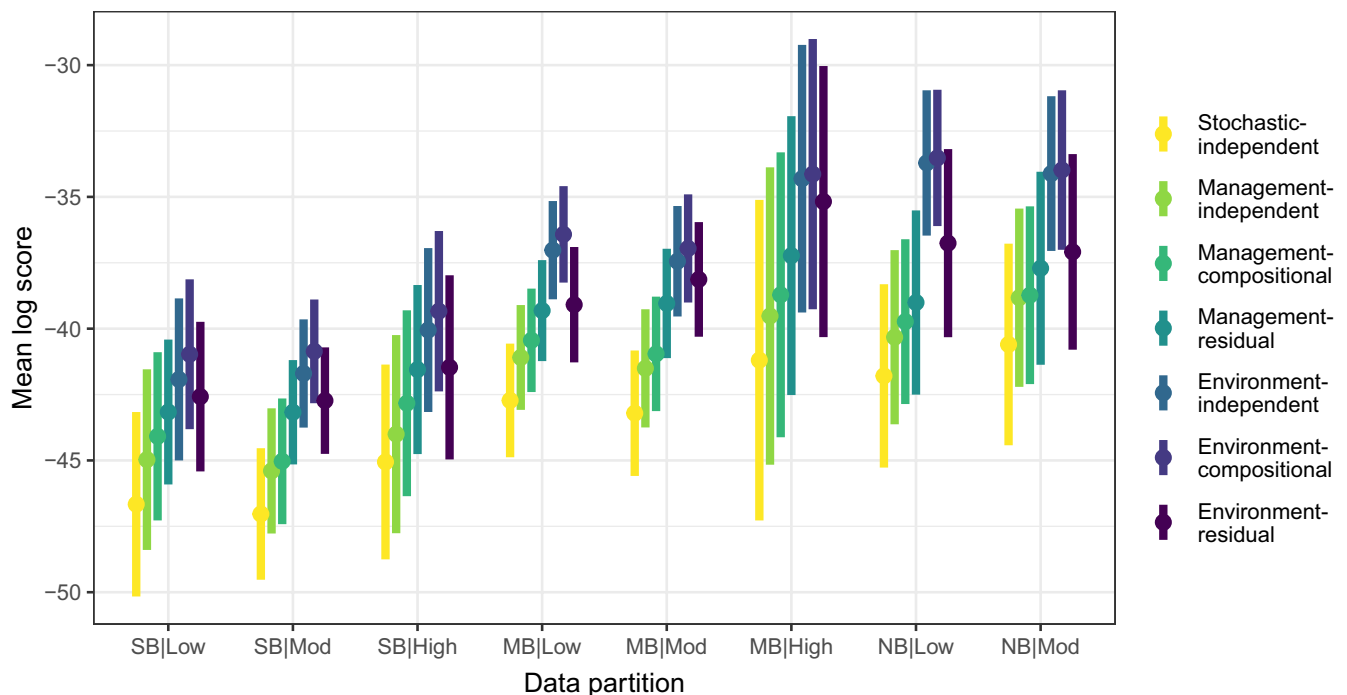
FIGURE 2 Mean, per-site community log scores for community-level prediction under candidate models applied within each out-of-sample data partition (as calculated by Equation S.9). Points represent the mean across all sites within a partition, while lines indicate 95% confidence intervals. Data partition labels are defined as: bioclimatic zone|basal area (BA). Bioclimatic zones: MB, mid boreal; NB, north boreal; SB, south boreal; BA levels (m^2·ha^{-1}): Low = [0–10], Mod = [10–30], High = ⩾30.

TABLE 2 Count of the number of species (out of 39) for which each candidate model had the best mean species log score (as calculated by Equation S.7) within each out-of-sample data partition.

Partition	Stoc.-indep.	Man.-indep.	Man.-comp.	Man.-resid.	Env.-indep.	Env.-comp.	Env.-resid.
SB|Low	0	1	7	0	9	20	2
SB|Mod	0	0	6	0	6	25	2
SB|High	1	3	3	1	6	22	3
MB|Low	0	0	6	0	8	22	3
MB|Mod	0	1	1	0	9	22	6
MB|High	1	1	0	1	10	14	12
NB|Low	0	1	0	0	8	19	11
NB|Mod	0	0	1	3	10	18	7

Model abbreviations are defined as follows: comp., compositional; Env., environment; Indep., independent; Man., management; resid., residual; Stoc., stochastic. Data partition labels are defined as: bioclimatic zone|basal area (BA). Bioclimatic zones: MB, mid boreal; NB, north boreal; SB, south boreal; BA levels (m$^2 \cdot$ha^{-1}): Low = [0–10), Mod = [10–30), High = ⩾30.

FIGURE 3 Summary of mean, per-site species log scores for species-level prediction under candidate models applied within each out-of-sample data partition (as calculated by Equation S.10). Coloured cells indicate the model with the highest mean species log score for each species-by-partition combination. Data partition labels are defined as: bioclimatic zone|basal area (BA). Bioclimatic zones: MB, mid boreal, NB, north boreal; SB, south boreal; BA levels (m$^2 \cdot$ha^{-1}): Low = [0–10), Mod = [10–30), High = ⩾30. Species codes are defined in Table S4.
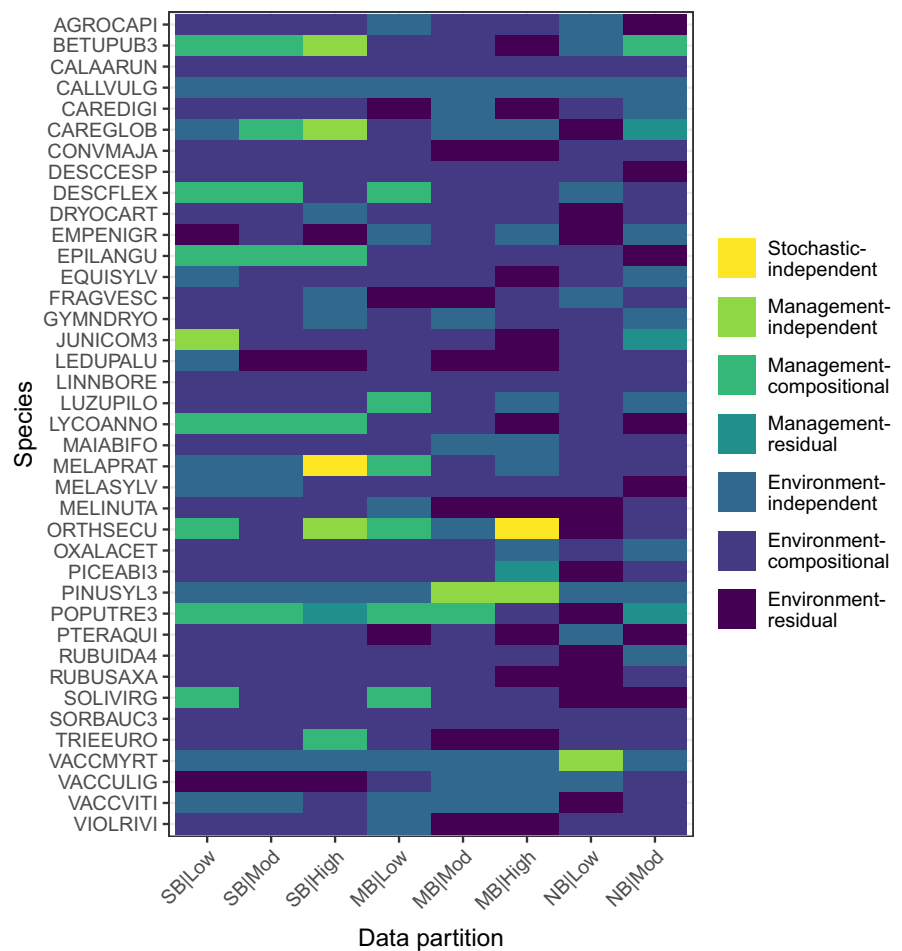
variable predictions after accounting for compositional dependence in the understory data. Specifically, there was evidence of a number of positive and negative residual species correlations under the Environment-residual model (Figure S4). These correlations constrain the random error terms for each species. If much of the statistical dependence among species is already accounted for through compositional dependence, these constrained random error terms may not yield improved model predictions. This does not suggest a lack of complex interactions among species in boreal forest understory communities; rather, it is an indication that model predictions are not improved by allowing for additional residual dependence among species after accounting for compositional dependence. Consistent with this interpretation, we found that the Environment-compositional and Environment-independent models had lower precision (higher posterior predictive standard deviation), but better accuracy (lower RMSE) than the Environment-residual model at the

individual species level, although this was not universally true across all species (Table S6).

There is good reason to suspect that compositional dependence plays a strong role in the modelled understory plant communities. Understory communities were measured using fixed area, $2\,m^2$ quadrats, with high relative abundance of a given species resulting in less available growing space and other resources for other species in the community. This type of resource limitation is observed regularly in forest ecosystems and is the basis for common conceptual models of forest development driven by density-dependent competition (e.g., Reineke's self-thinning law; Reineke, 1933). Although compositional dependence does not necessarily indicate resource competition among species, model predictions combined with previous experimental results suggest such a connection may exist in the modelled boreal forest understory communities (Gundale et al., 2012). Functionally, the compositional dependence modelled by the multinomial likelihood constrains predictions of individual species abundance (Equation 10) limiting the potential for overly high predictions of total abundance previously identified as a key benefit of JSDMs applied to plant abundance data (Clark et al., 2014).

In addition to the potential role of compositional dependence, the strong performance of the three environmental models, and the Environment-independent model in particular, suggests that much of the variability in out-of-sample communities is captured through the estimated effects of the observed environment. Note that the three environment models were the top three predicting models regardless of the metric used to evaluate community-level predictions (Figure 1; Table S3) and across all out-of-sample data partitions considered (Figures 2 and 3; Table 2). The inclusion of this informative set of environmental variables combined with compositional dependence that constrains the relative abundance of species as under the Environment-compositional model may leave little residual variation to be explained such that the correlated random error term included in the Environment-residual model led to less accurate predictions. While variance partitioning approaches (such as those described in Ovaskainen et al., 2017) would be useful to determine the proportion of variance explained by the environmental variables, such approaches are not immediately applicable to the Environment-compositional and Environment-residual models given the softmax link function (Equation 10) used under the multinomial model. Extension of variance partitioning methods for compositional data is beyond the scope of the current study. Despite the lack of variance partitioning, the strength of the Environment-compositional model is highlighted by the high Bayesian R-squared values estimated for each species (Table S5). Specifically, an out-of-sample Bayesian R-squared of greater than 0.50 was estimated for the majority of species under the Environment-compositional model with many species having a value greater than 0.90.

Interestingly, the relative importance of alternative species dependence structures changed when forest density alone was included in candidate models (management models). Among the management models, there was statistical support for including both compositional and residual species dependence to predict community composition as evidenced by the ordering of the joint

community log scores of the three models: Management-residual > Management-compositional > Management-independent (Figure 1). This ordering is what we would expect to find based on ecological theory. That is, community composition related to complex combinations of abiotic and biotic filters, here potentially reflected by the compositional and residual correlations among species (Vellend, 2016). While the predictive performance of the management models matches ecological expectations, when compared to the predictive performance of the more complex environment models it hints at a broader and known limitation of JSDMs: the confounding of unobserved environmental variables and species dependence (Poggiato et al., 2021). Specifically, there was strong evidence that including residual species dependence improved community-level prediction when important environmental variables were excluded as under the management models (Figure 1). When the full set of environmental variables was included, however, there was near equal predictive performance of independent and residual dependence models despite the former assuming no species dependence and the latter including both compositional and residual species dependence (Figure 1). This suggests that the variability accounted for by residual species dependence was captured by the additional fixed environmental effects in the three environment models. The difference in predictive performance among the three environment models may also be due to confounding of the flexible correlated random error term and observed environmental variation (Van Ee et al., 2022).

The relative predictive performance of the seven candidate models was maintained when predicting communities and individual species across the novel conditions reflected in out-of-sample data partitions spanning broad successional and bioclimatic gradients (Figures 2 and 3; Table 2). The applied out-of-sample data partitions emulate block cross-validation strategies advocated for in recent work to assess the transferability of JSDMs (Roberts et al., 2017). Based on recent work demonstrating spatio-temporal changes in the processes underlying community composition (Jabot et al., 2020), we expected candidate models to have varying predictive performance conditional on the successional stage and relative harshness of the regional climate. Contrary to our expectations, however, joint community log score results across all out-of-sample data partitions matched those found applying the complete set of out-of-sample data. Specifically, the Environment-compositional model had the highest predictive performance at the community-level across all partitions followed by the Environment-independent and Environment-residual models (Figure 2; note the relative differences among models varies by partition). Similar results were found when assessing species-level predictions based on the species log score (Table 2 and Figure 3). The strong predictive performance of the environment models across a range of successional stages and bioclimatic conditions at both the community and individual species scales highlights the robustness of model predictions when the full set of environmental variables is included and points to higher potential for model transferability.

Model selection results applying common metrics used to assess community-level JSDM predictions largely matched the joint community log score results (Table S3; Figure 1). There were,

however, some subtle, but important differences. In particular, the Environment-residual model had the best predictive model score based on the Jaccard community dissimilarity index, although the Environment-compositional model was a close second. The difference in model selection results is not surprising given that each alternative community-level metric assesses model predictions differently. The independent community log score is calculated based on the marginal prediction of species and does not take into account the species covariance included in JSDMs (see Log Score Calculation for details). RMSE is a measure of point prediction (based on the posterior predictive mean) rather than probabilistic prediction and hence does not account for the variability of future data (Gelman et al., 2014). The Jaccard community dissimilarity index is a measure of probabilistic prediction based on how closely predictions replicate observed community composition (Wilkinson et al., 2021).

In the context of predictive model selection, the metric used to assess out-of-sample predictions represents a statistical decision that affects how models are evaluated and ranked. As such, it is important that the metric used to assess out-of-sample predictions follow from the objectives of the model application. Here, we have applied the joint community log score given our interest in probabilistic community-level prediction. In Bayesian model selection, the log score is preferred when interest is in probabilistic prediction given its theoretical properties including the fact that its value is maximized under the true data generating mechanism (Hooten & Hobbs, 2015). A separate study with different objectives (e.g., point prediction of individual species), might apply a different scoring metric, and identify a different "best" predicting model among the same seven candidate models considered here.

The current study demonstrates the practical application of the joint community log score for community-level prediction and highlights its capacity to quantify differences in the predictive performance of complex JSDMs including alternative species dependence types. Further, the application of the joint community log score to assess predictive performance across out-of-sample data partitions spanning broad successional and bioclimatic gradients demonstrates its utility for transferable model selection where the goal is to identify the best performing model under novel conditions. Although not the direct objective of the current analysis, the multinomial JSDM provides a multivariate approach for modelling compositional data commonly encountered in natural community surveys. Given its implementation within the more general HMSC framework, the model can deal with a large number of species and locations and allows additional information on functional traits and phylogeny to be incorporated as predictors of environmental responses (Ovaskainen et al., 2017). As such, the model framework combined with the Bayesian model selection approach applied here, constitutes a step forward in our ability to predict natural communities under global change.

CONFLICT OF INTEREST STATEMENT

There are no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The R code and data that support the findings of this study are openly available on Dryad at https://doi.org/10.5061/dryad.tmpg4f53g.

ORCID

Malcolm S. Itter https://orcid.org/0000-0002-0184-2731
Elina Kaarlejärvi https://orcid.org/0000-0003-0014-0073

REFERENCES

Ahti, T., Hämet-Ahti, L., & Jalas, J. (1968). Vegetation zones and their sections in northwestern europe. *Annales Botanici Fennici*, 5(3), 169–211.

Antão, L. H., Weigel, B., Strona, G., Hällfors, M., Kaarlejärvi, E., Dallas, T., Opedal, Ø. H., Heliölä, J., Henttonen, H., Huitu, O., Korpimäki, E., Kuussaari, M., Lehikoinen, A., Leinonen, R., Lindén, A., Merilä, P., Pietiäinen, H., Pöyry, J., Salemaa, M., ... Laine, A. L. (2022). Climate change reshuffles northern species within their niches. *Nature Climate Change*, 12(6), 587–592.

Araújo, M. B., Alagador, D., Cabeza, M., Nogués-Bravo, D., & Thuiller, W. (2011). Climate change threatens european conservation areas. *Ecology Letters*, 14(5), 484–492.

Baker, S. G. (1994). The multinomial-poisson transformation. *Journal of the Royal Statistical Society: Series D (the Statistician)*, 43(4), 495–504.

Blowes, S. A., Supp, S. R., Antão, L. H., Bates, A., Bruelheide, H., Chase, J. M., Moyes, F., Magurran, A., McGill, B., Myers-Smith, I. H., Winter, M., Bjorkman, A. D., Bowler, D. E., Byrnes, J. E. K., Gonzalez, A., Hines, J., Isbell, F., Jones, H. P., Navarro, L. M., ... Dornelas, M. (2019). The geography of biodiversity change in marine and terrestrial assemblages. *Science*, 366(6463), 339–345.

Broms, K. M., Hooten, M. B., & Fitzpatrick, R. M. (2016). Model selection and assessment for multi-species occupancy models. *Ecology*, 97(7), 1759–1770.

Cajander, A. K. (1949). Forest types and their significance. *Acta For Fenn*, 56, 1–71.

Caradima, B., Schuwirth, N., & Reichert, P. (2019). From individual to joint species distribution models: A comparison of model complexity and predictive performance. *Journal of Biogeography*, 46(10), 2260–2274.

Chao, A., Chazdon, R. L., Colwell, R. K., & Shen, T.-J. (2005). A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters*, 8(2), 148–159.

Clark, J. S., Gelfand, A. E., Woodall, C. W., & Zhu, K. (2014). More than the sum of the parts: Forest climate response from joint species distribution models. *Ecological Applications*, 24(5), 990–999.

Clark, J. S., Nemergut, D., Seyednasrollah, B., Turner, P. J., & Zhang, S. (2017). Generalized joint attribute modeling for biodiversity analysis: Median-zero, multivariate, multifarious data. *Ecological Monographs*, 87(1), 34–56.

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., & Wikle, C. K. (2009). Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, 19(3), 553–570.

D'Amen, M., Dubuis, A., Fernandes, R. F., Pottier, J., Pellissier, L., & Guisan, A. (2015). Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models. *Journal of Biogeography*, 42(7), 1255–1266.

Damgaard, C. (2015). Modelling pin-point cover data of complementary vegetation classes. *Ecological Informatics*, 30, 179–184.

Damgaard, C. (2018). The joint distribution of pin-point plant cover data: A reparametrized dirichlet–multinomial distribution. arXiv preprint.

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24(6), 997–1016.

Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., … Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, 16(12), 1424–1435.

Gundale, M. J., Hyodo, F., Nilsson, M.-C., & Wardle, D. A. (2012). Nitrogen niches revealed through species and functional group removal in a boreal shrub community. *Ecology*, 93(7), 1695–1706.

Harris, D. J. (2015). Generating realistic assemblages with a joint species distribution model. *Methods in Ecology and Evolution*, 6(4), 465–473.

Hooten, M. B., & Hobbs, N. T. (2015). A guide to Bayesian model selection for ecologists. *Ecological Monographs*, 85(1), 3–28.

Hui, F. K. (2016). Boral–Bayesian ordination and regression analysis of multivariate abundance data in R. *Methods in Ecology and Evolution*, 7(6), 744–750.

Ingram, M., Vukcevic, D., & Golding, N. (2020). Multi-output Gaussian processes for species distribution modelling. *Methods in Ecology and Evolution*, 11(12), 1587–1598.

Itter, M., Kaarlejärvi, E., Laine, A. L., Hamberg, L., Tonteri, T., & Vanhatalo, J. (2023). *Code and data for Bayesian joint species distribution model selection for community-level prediction* [Data set]. Dryad. https://doi.org/10.5061/dryad.tmpg4f53g

Jabot, F., Laroche, F., Massol, F., Arthaud, F., Crabot, J., Dubart, M., Blanchet, S., Munoz, F., David, P., & Datry, T. (2020). Assessing metacommunity processes through signatures in spatiotemporal turnover of community composition. *Ecology Letters*, 23(9), 1330–1339.

Jost, L. (2006). Entropy and diversity. *Oikos*, 113(2), 363–375.

Lorenz, M., & Fischer, R. (2013). Pan-European forest monitoring: An overview. *Developments in Environmental Science*, 12, 19–32.

Maguire, K. C., Nieto-Lugilde, D., Blois, J. L., Fitzpatrick, M. C., Williams, J. W., Ferrier, S., & Lorenz, D. J. (2016). Controlled comparison of species-and community-level models across novel climates and communities. *Proceedings of the Royal Society B: Biological Sciences*, 283(1826), 20152817.

McCullagh, P., & Nelder, J. (1989). *Generalized linear models*. Chapman and Hall.

Norberg, A., Abrego, N., Blanchet, F. G., Adler, F. R., Anderson, B. J., Anttila, J., Araújo, M. B., Dallas, T., Dunson, D., Elith, J., Foster, S. D., Fox, R., Franklin, J., Godsoe, W., Guisan, A., O'Hara, B., Hill, N. A., Holt, R. D., Hui, F. K. C., … Ovaskainen, O. (2019). A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological Monographs*, 89(3), e01370.

Ovaskainen, O., Roy, D. B., Fox, R., & Anderson, B. J. (2016). Uncovering hidden spatial structure in species communities with spatially explicit joint species distribution models. *Methods in Ecology and Evolution*, 7(4), 428–436.

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T., & Abrego, N. (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters*, 20(5), 561–576.

Poggiato, G., Münkemüller, T., Bystrova, D., Arbel, J., Clark, J. S., & Thuiller, W. (2021). On the interpretations of joint modeling in community ecology. *Trends in Ecology & Evolution*, 36, 391–401.

Pollock, L. J., O'Connor, L. M., Mokany, K., Rosauer, D. F., Talluto, M. V., & Thuiller, W. (2020). Protecting biodiversity (in all its complexity): New models and methods. *Trends in Ecology & Evolution*, 35(12), 1119–1128.

Pollock, L. J., Tingley, R., Morris, W. K., Golding, N., O'Hara, R. B., Parris, K. M., Vesk, P. A., & McCarthy, M. A. (2014). Understanding co-occurrence by modelling species simultaneously with a joint species distribution model (JSDM). *Methods in Ecology and Evolution*, 5(5), 397–406.

Reineke, L. H. (1933). Perfecting a stand-density index for even-aged forest. *Journal of Agricultural Research*, 46, 627–638.

Reinikainen, A., Mäkipää, R., Vanha-Majamaa, I., & Hotanen, J.-P. (Eds.). (2000). *Kasvit muuttuvassa metsäluonnossa [summary in English: Changes in the frequency and abundance of forest and mire plants in Finland since 1950]*. Kustannusosakeyhtiö Tammi.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929.

Taylor-Rodríguez, D., Kaufeld, K., Schliep, E. M., Clark, J. S., & Gelfand, A. E. (2017). Joint species distribution modeling: Dimension reduction using Dirichlet processes. *Bayesian Analysis*, 12(4), 939–967.

Tikhonov, G., Duan, L., Abrego, N., Newell, G., White, M., Dunson, D., & Ovaskainen, O. (2020). Computationally efficient joint species distribution modeling of big spatial data. *Ecology*, 101(2), e02929.

Tikhonov, G., Opedal, Ø. H., Abrego, N., Lehikoinen, A., de Jonge, M. M., Oksanen, J., & Ovaskainen, O. (2020). Joint species distribution modelling with the R-package Hmsc. *Methods in Ecology and Evolution*, 11(3), 442–447.

Tikhonov, G., Ovaskainen, O., Oksanen, J., de Jonge, M., Opedal, O., & Dallas, T. (2021). Hmsc: Hierarchical model of species communities. R Package Version 3.0-11.

Tomppo, E., Heikkinen, J., Henttonen, H. M., Ihalainen, A., Katila, M., Mäkelä, H., et al. (2011). *Designing and conducting a forest inventory-case: 9th National Forest Inventory of Finland* (Vol. 22). Springer Science & Business Media.

Van Ee, J. J., Ivan, J. S., & Hooten, M. B. (2022). Community confounding in joint species distribution models. *Scientific Reports*, 12(1), 1–14.

Vanhatalo, J., Hartmann, M., & Veneranta, L. (2020). Additive multivariate Gaussian processes for joint species distribution modeling with heterogeneous data. *Bayesian Analysis*, 15(2), 415–447.

Vellend, M. (2016). *The theory of ecological communities (MPB-57)*. Princeton University Press.

Venäläinen, A., Tuomenvirta, H., Pirinen, P., & Drebs, A. (2005). A basic Finnish climate data set 1961–2000–description and illustrations. *Finnish Meteorological Institute, Reports*, 5, 1–27.

Warton, D. I., Blanchet, F. G., O'Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C., & Hui, F. K. C. (2015). So many variables: Joint modeling in community ecology. *Trends in Ecology & Evolution*, 30(12), 766–779.

Wilkinson, D. P., Golding, N., Guillera-Arroita, G., Tingley, R., & McCarthy, M. A. (2019). A comparison of joint species distribution models for presence–absence data. *Methods in Ecology and Evolution*, 10(2), 198–211.

Wilkinson, D. P., Golding, N., Guillera-Arroita, G., Tingley, R., & McCarthy, M. A. (2021). Defining and evaluating predictions of joint species distribution models. *Methods in Ecology and Evolution*, 12(3), 394–404.

Xia, F., Chen, J., Fung, W. K., & Li, H. (2013). A logistic normal multinomial regression model for microbiome compositional data analysis. *Biometrics*, 69(4), 1053–1063.

Yates, K. L., Bouchet, P. J., Caley, M. J., Mengersen, K., Randin, C. F., Parnell, S., Fielding, A. H., Bamford, A. J., Ban, S., Barbosa, A. M., Dormann, C. F., Elith, J., Embling, C. B., Ervin, G. N., Fisher, R., Gould, S., Graf, R. F., Gregr, E. J., Halpin, P. N., … Sequeira, A. M. M. (2018). Outstanding challenges in the transferability of ecological models. *Trends in Ecology & Evolution*, *33*(10), 790–802.

Zurell, D., Zimmermann, N. E., Gross, H., Baltensweiler, A., Sattler, T., & Wüest, R. O. (2020). Testing species assemblage predictions from stacked and joint species distribution models. *Journal of Biogeography*, *47*(1), 101–113.

BIOSKETCH

Malcolm S. Itter is a statistical and applied forest ecologist. His research interests are in the development of Bayesian spatial–temporal methodologies to improve understanding and prediction of community dynamics to inform conservation management. To this end, Dr. Itter works to synthesize observations of natural communities across space and time with process-based models of their underlying dynamics.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Itter, M. S., Kaarlejärvi, E., Laine, A.-L., Hamberg, L., Tonteri, T., & Vanhatalo, J. (2024). Bayesian joint species distribution model selection for community-level prediction. *Global Ecology and Biogeography*, *33*, e13827. https://doi.org/10.1111/geb.13827