Joint Species Distribution Modeling with species competition and non-stationary spatial random effects

Juho Kettunen
School of Computing, University of Eastern Finland

Lauri Mehtätalo
Bioeconomy and Environment Unit. Natural Resources Institute Finland (Luke)

Eeva-Stiina Tuittila
School of Forest Sciences, University of Eastern Finland

Aino Korrensalo
School of Forest Sciences, University of Eastern Finland

Jarno Vanhatalo
Department of Mathematics and Statistics, University of Helsinki

Abstract

Joint species distribution models (JSDM) are among the most important statistical tools in community ecology. However, existing JSDMs cannot model mutual exclusion between species. We tackle this deficiency by developing a novel hierarchical JSDM with Dirichlet-Multinomial observation process for mutually exclusive species groups. We apply non-stationary multivariate Gaussian processes to describe species’ niche preferences and conduct Bayesian inference using Markov chain Monte Carlo. We propose decision theoretic model comparison and validation methods to assess the goodness of the proposed model and its alternatives in a case study on modeling vegetation cover in a boreal peatland in Finland. Our results show that ignoring the interspecific interactions and competition significantly reduces models’ predictive performance and through that leads to biased estimates for total cover of individual species and over all species combined. Models’ relative predictive performance also depends on the predictive task highlighting that model comparison and assessment method should resemble the true predictive task. Our results also demonstrate that the proposed joint species distribution model can be used to simultaneously infer interspecific correlations in niche preference as well as mutual competition for space and through that provide novel insight into ecological research.

Keywords: non-stationary Gaussian process, predictive model comparison, Dirichlet-Multinomial, carbon cycling
1 Introduction

Community ecology has seen a fast and significant development of statistical methods for joint species distribution modeling in recent years (Warton et al., 2015; Ovaskainen and Abrego, 2020; Nordberg et al., 2019; Vanhatalo et al., 2020). Species distribution models are statistical models that describe the variation in the occurrence and abundance of a species in space and time as well as the abiotic driving factors that affect that variability (Gelfand et al., 2006). Joint species distribution models additionally describe species-to-species interactions through interspecific correlations between species responses to environment and random factors (Ovaskainen and Abrego, 2020; Vanhatalo et al., 2020). However, the state-of-the-art joint species distribution models lack realistic description for mutual exclusion – a process that can arise, for example, from interspecific competition for space or priority effects where a species arriving first excludes the later arriving species.

We propose a novel Bayesian hierarchical joint species distribution model to estimate abundances of mutually exclusive species groups. Our model is based on multivariate Gaussian processes (Vanhatalo et al., 2020) which we use to describe species specific niche preference (and relative competitive performance) over the study area. However, instead of modeling each species as conditionally independent given the Gaussian latent variable, we divide the species to mutually exclusive species groups and add group-wise Dirichlet-Multinomial hierarchical layer on top of the Gaussian process latent variable model. Dirichlet distributions model interspecific competition between species and Multinomial distributions model the uncertainty in observation process. We also account for the differing spatial patterns in species’ occurrences by extending the multivariate Gaussian processes of Vanhatalo et al. (2020) to non-stationary multivariate Gaussian processes. We present our model in the context of vegetation percentage covers where each plant species can occupy a proportion of a given area. Out of plant species, mosses are assumed to grow in one layer and therefore compete for space against others alike whereas vascular plants are assumed to grow in multiple layers together with others. Our model and methods are, however, equally well applicable to any compositional case where species can be assigned to exclusive groups. We apply our model to data collected from an extensively studied long-term monitoring site located in a boreal peatland in Finland (Figure 2).

Our aims are to produce species percentage cover maps and to estimate the total vegetation cover of moss and vascular plant species over the study site around an eddy covariance tower which is used for automatic monitoring of carbon gas fluxes within the surrounding area. These estimates could subsequently be used to calibrate the carbon gas flux estimates of the site. We compare our model to traditional stacked species distribution models, which model each species independently, as well as to joint species distribution models, which neglect species competition. We conduct the model comparison and validation using rigorous decision theoretic approach by calculating cross-validation predictive measures that mimic the actual prediction task at hand.
Figure 1: The directed acyclic graph (DAG) representation of the joint species distribution model. The columns represent the species groups and rows the model layers.

2 Statistical models

2.1 Hierarchical joint model for species-wise percentage covers

We model the species specific cover jointly using hierarchical Bayesian modeling where each hierarchical layer represents a part of the modeled process [Vanhatalo et al. (2020)], see Figure[1]. We denote by \( D \subset \mathbb{R}^2 \) the bounded study region of interest and by a set \( S = \{s_1, s_2, ..., s_n\} \) the locations (centroids) of the inventory plots such that \( s_i \in D, i \in \{1, ..., n\} \), where \( n \) is the total number of plots (see Figure[2]). The set of modeled species is denoted by \( E = \{e_1, ..., e_J\} \), where \( e_j \) is the identifier of the \( j \)-th species and \( J \) is the total number of species. We then allocate species into groups such that each group is formed by mutually exclusive species; that is, by species that cannot grow over each other. These groups form nonempty disjoint sets \( E_g \) which are defined such that \( E = \cup_{g=1}^{p} E_g \) where \( p \) is the total number of groups (Figure[1]).

We collect to a vector \( \phi_i = (\phi_{i1}, \phi_{i2}, ..., \phi_{ij})^\top \) the true species specific percentage covers in an inventory plot at spatial location \( s_i \); that is, \( \phi_{ij} \) is the percentage cover of species \( j \) at plot \( i \). Since the percentage covers are assumed to be independent between groups we can decompose the percentage cover model as

\[
\pi(\phi_i|f_i, \gamma) = \prod_{g=1}^{p} \pi(\phi_{ig}|f_{ig}, \gamma_g) \tag{1}
\]

where \( f_i \) is a vector of species specific latent variables, \( \gamma = (\gamma_1, ..., \gamma_p) \) is a vector of process parameters and \( \phi_{ig} = (\phi_{ij})_{j \in E_g} \) and \( f_{ig} = (f_{ij})_{j \in E_g} \) are respectively a vector of percentage covers and a vector of latent variables of species in group \( E_g \). The latent variable \( f_{ij} \) describes a species’ preference to site \( s_i \) (e.g., how suitable the environment is at that site) and its competition strength relative to other species in the same group. The model \( \pi(\phi_{ig}|f_{ig}, \gamma_g) \) then accounts for the randomness in the outcome of the species specific processes and interspecific competition that lead to the actual species specific percentage covers. We account for these random processes with Dirichlet distribution so that

\[
\pi([\phi_{ig}, \phi_{ig0}]|f_{ig}, \gamma_g) = \text{Dir}([\phi_{ig}, \phi_{ig0}]|\alpha(f_{ig}) \times \gamma_g) \tag{2}
\]
where $\phi_{i,g_0} = 1 - \sum_{j \in E_g} \phi_{ij}$ corresponds to the case that none of the species from group $E_g$ is present and $\alpha(f_{i,g}) = \alpha_{i,g} = [\alpha_{ij}]_{j \in E_g} = 1 - \sum_{j \in E_g} \alpha_{ij}$ is a vector of expected percentage covers defined through the soft-max function

$$\alpha_{ij} = \frac{\exp(f_{ij})}{1 + \sum_{j \in E_g} \exp(f_{ij})}. \quad (3)$$

Hence, the process parameter $\gamma_g$ governs the level of randomness in the process leading to percentage covers so that the randomness decreases with increasing $\gamma_g$, and $\alpha_{ij}$ summarizes a species’ relative competitive skill within its group. Hence, a natural measure for competition between two species is the correlation between their percentage covers:

$$\text{Corr}(\phi_{ij}, \phi_{ij'}) = -\sqrt{\frac{\alpha_{ij}\alpha_{ij'}}{(1-\alpha_{ij})(1-\alpha_{ij'})}}. \quad (4)$$

The correlation is strongest when $\alpha_{ij} = \alpha_{ij'} = 1/2$ and decreases towards zero when either or both of the $\alpha$ terms decrease to zero. Moreover, as $\alpha_{ij}$ varies in space the correlation varies in space as well and, hence, the model can capture spatial changes in the strength of competition. Note though, that we use competition in a technical manner to denote all processes that lead to species exclusion (modeled by the Dirichlet process) without direct link to a specific biological processes of competition. In single species groups, the Dirichlet distribution reduces to Beta distribution and the softmax function reduces to the inverse logit-link function. Before defining the latent function we introduce the observation model.

As typical, the percentage cover observations are expert assessments based on visual inspection of the inventory plots. To aid the visual assessment researchers divide the inventory plot (either mentally or using a mesh) into a homogeneous lattice. The reported percentage cover then roughly corresponds to the fraction of mesh vertexes occupied by a species. Hence, we denote by $y_i = (y_{i1}, \ldots, y_{iJ})^\top$ the number of mesh cells occupied by each species so that $y_{ij} = \lfloor$observed percentage cover of species $j$ $\rfloor \times N_{i,g}$, where $N_{i,g}$ corresponds to the mesh size for group $E_g$. We can then model the joint probability mass function of the observed percentage covers at location $s_i$ as

$$\pi(y_i|N_i, \phi_i) = \prod_{g=1}^p \text{Multinomial}([y_{i,g}, y_{i,g0}]|[\phi_{i,g}, \phi_{i,g0}], N_{ig}) \quad (5)$$

where $N_i = [N_{i1}, \ldots, N_{ip}]$ and $y_{i,g0} = N_{i,g} - \sum_{j \in E_g} y_{ij}$ corresponds to the number of mesh vertexes where none of the species from group $E_g$ is present. Hence, the parameter $N_{ig}$ is formally the sample size of a Multinomial distribution but in our application it actually corresponds to the measurement accuracy. The smaller $N_{ig}$ is, the more uncertainty is assumed in the expert assessments. In practice, we could also give an informative prior for $N_{ig}$ to account for the fact that the estimate for experts accuracy is not exact. However, we do not consider this option here.

### 2.2 Species specific latent processes

The species specific latent processes are modeled as

$$f_{ij} = \beta_j + \epsilon_{ij}, \quad (6)$$
where $\beta_j$ is an intercept for species $j$ and $\epsilon_{ij}$ is a spatial latent process for species $j$ at a location $s_i$. Typically JSDMs incorporate both environmental covariate effects and spatial random effects into $\epsilon_{ij}$ (Ovaskainen and Abrego, 2020; Vanhatalo et al., 2020). However, since we did not have detailed covariate information from our study region, we did not incorporate covariate effects into $\epsilon_{ij}$ but used stationary and non-stationary multivariate spatial Gaussian processes.

### 2.2.1 Stationary spatial Gaussian processes

We model the spatial latent processes $\epsilon_{ij}$, either as mutually independent or jointly dependent. In the former approach, a species specific latent function is given a univariate Gaussian process prior

$$\epsilon_{ij} \sim GP(0, k_j^{(e)}(s_i, s_{i'})),$$

where $k_j^{(e)}(s_i, s_{i'})$ is the species specific spatial covariance function. In the baseline models, we use the stationary exponential covariance function

$$k_j^{(e)}(s_i, s_{i'}) = \sigma_j^2 \exp\left(-\frac{||s_i - s_{i'}||}{l_j}\right)$$

with species specific length scale parameter, $l_j$, governing how fast the spatial correlation decays and variance parameter $\sigma_j^2$ governing the magnitude of the variation. Independent species specific latent processes contain an implicit assumption that the processes behind species’ niche preferences are mutually independent (Vanhatalo et al., 2020; Ovaskainen and Abrego, 2020). This is an unrealistic assumption in practice since species typically show positive and negative associations in their niche preferences. Modeling these associations can significantly improve the predictive accuracy of species distribution models (Nordberg et al., 2019; Vanhatalo et al., 2020).

Hence, we extend the baseline model by allowing for interspecific correlations between the spatial latent processes. We model dependencies through linear model of coregionalization (LMC) such that spatial latent processes are expressed as a linear combination of $J$ zero mean univariate Gaussian processes having the covariance functions $k_j^{(e)}(s_i, s_{i'})$. This induces a multivariate Gaussian prior for the latent variables (Vanhatalo et al., 2020). By collecting all latent variables into vector $\mathbf{\epsilon}(S) = (\epsilon_1(S)^\top, \ldots, \epsilon_J(S)^\top)^\top$ such that $\epsilon_j(S)$ denotes the vector having latent variables for species $j$ at spatial locations $S$,

$$\mathbf{\epsilon}(S) \sim N\left(0, \sum_{j=1}^J B_j \otimes K_j^{(e)}\right)$$

where the covariance matrices are constructed with $[K_j^{(e)}]_{i,i'} = k_j^{(e)}(s_i, s_{i'}; l_j, \sigma_j^2 = 1)$ and the matrix $B_j = L_j L_j^\top$ where $L_j$ is the $j$th column of the Cholesky decomposition of the coregionalization covariance matrix $\Sigma_\epsilon = \sum_{j=1}^J L_j L_j^\top$, which models the interspecific dependencies between species niche preferences. The variance parameter $\sigma_j^2$ of $k_j^{(e)}(s_i, s_{i'})$ is set to 1 to ensure identifiability.

The latent process (9) is the most flexible version of the LMC models where each species is assumed to have its own process characteristics encoded by $k_j$. However, we can reduce the flexibility by reducing the number of unique covariance functions in the model. For
example, if all species share the same covariance function \( k_j^{(e)}(s_i, s_{i'}) = k_j^{(e)}(s_i, s_{i'}), \forall j = 1, \ldots, J \), we obtain model LMC(1), which is also called *intrinsic model for coregionalization*. In this model the species specific latent processes are (correlated) random draws from the same underlying Gaussian process whereas in \( LMC(k) \) the latent processes are (correlated) random draws from \( J \) different Gaussian processes. We tested also models in between these two extremes and denote by LMC\((k)\) a model where we have \( k \) distinct covariance functions so that \( k_j^{(e)}(s_i, s_{i'}) = k_k^{(e)}(s_i, s_{i'}), \forall j \geq k \) when \( k \) is less than \( J \). Note that the LMC\((k)\) models always have \( J \) unique spatial latent processes and the coregionalization matrix \( \Sigma \) is positive definite in all these models. Our approach is reasonable with our current application that has only moderate number of species. If extended for larger number of species, it would be reasonable to replace the LMC model with a latent factor model where we would have only \( p < J \) spatial Gaussian processes so that the interspecific covariance matrix would be low rank, such as in the hierarchical modelling of species communities approach (Ovaskainen and Abrego 2020).

### 2.2.2 Non-stationary spatial Gaussian processes

Stationary Gaussian processes work typically well if the region of interest is sampled relatively sparsely so that data does not allow inference for non-stationarity in spatial correlation. However, since the properties affecting peatlands vegetation can vary considerably within our study region, and our data is sampled with high resolution (see Section 4.3), we extend the stationary model by setting non-stationary covariance function for Gaussian process. We include non-stationarity into the model by changing stationary covariance function in equation (8) to a non-stationary Matérn (\( \nu = 3/2 \)) covariance function with spatially varying length scale parameter (e.g. Paciorek and Schervish 2006)

\[
k_j^{(m)}(s_i, s_{i'}) = |\Sigma_{ij}|^{1/4} |\Sigma_{i',j}^{1/4}| \left| \frac{\Sigma_{ij} + \Sigma_{i',j}}{2} \right|^{-1/2} (1 + \sqrt{3} \sqrt{Q_{i,i',j}} \exp(-\sqrt{3} \sqrt{Q_{i,i',j}}))
\]

where \( Q_{i,i',j} = (s_i - s_{i'})^\top \left( \Sigma_{ij} + \Sigma_{i',j} \right)^{-1} (s_i - s_{i'}) \) is the Mahalanobis spatial distance between locations \( s_i \) and \( s_{i'} \) for species \( j \) and \( \Sigma_{ij} = \Sigma_{i,i,j}^2 \left[ \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right] \) so that \( l_{ij} \) varies spatially. We model the spatially varying length scale parameter by giving a Gaussian process prior for its logarithm

\[
\log(l_{ij}) \sim GP(\mu_{ij}, k_j^{(e)}(s_i, s_{i'})),
\]

where the mean function \( \mu_{ij} \) specifies the expected value of \( \log(l_{ij}) \) and \( k_j^{(e)}(s_i, s_{i'}) \) is the stationary exponential covariance function (8). Gaussian process prior gives smoothly varying length scale and modelling logarithm of the length scale \( l_{ij} \) ensures positivity. Since the matrix \( \Sigma_{ij} \) is diagonal and all its diagonal elements are the same, the spatial correlation is isotropic at each location but the strength of the correlation decay varies within the area. The Non-stationary covariance function in equation (10) reduces to stationary Matérn (\( \nu = 3/2 \)) covariance function if the length scale parameter, \( l_{ij} \), is set to be the same at each location. As in the stationary case the parameterizations of the priors for the non-stationary covariance function parameters should be chosen according to the size of the study area and the sampling resolution. We summarize the prior distributions used in this work in Section 2.3.
The non-stationary GP was applied to models with and without interspecific dependence between the spatial latent processes. In the former case, each latent process was given an independent non-stationary GP prior. In the latter case, the latent processes were modeled jointly with LMC (equation (9)) where $K_j^{(e)}$ is replaced by $[K_j^{(m)}]_{i,i'} = k_j^{(m)}(s_i, s_{i'})$.

### 2.3 Priors

The model specification is completed by assigning prior distributions to the model parameters. In the stationary GP models, we gave weakly informative half-inverse-Student-$t$ prior for the length scale parameters, $1/l_j \sim \text{Student} - t_+ (\mu, s^2, \nu)$, which gives a priori more weight for the larger length scales. The parameter values of the prior were chosen according to the size of the study area (see Section 4.3). The location and scale parameters of the half-inverse-Student-$t$ prior for the length scale parameters of the stationary GP models were selected such that the length scale is less than 400 meters with probability 0.99. This gives $1/l_j \sim \text{Student} - t_+ (0, 0.19^2, 5)$ which corresponds to preferring smooth over small scale variability in vegetation composition.

In the non-stationary GP models, the prior for the mean of the log length-scale was $\mu_{l_j} \sim N(4.5, \sqrt{2^2})$, which favours relatively large $l_{ij}$ values ($> 100$ m). For the length scale and variance parameters of the GP prior for log $l_{ij}$ (11) we gave weakly informative priors, $1/l_j \sim \text{Student} - t_+ (0, 2^2, 4)$ and $\sigma_j^2 \sim \text{Student} - t_+ (0, 1, 4)$. The location and scale parameters in these priors were selected such that $l_j$ is less than 38 meters with probability 0.99.

We gave a Gamma prior, $\gamma_g \sim \text{Gamma}(3/2, 2/3)$, for the process parameters governing the level of randomness. It is a priori likely that randomness in the vegetation composition is high for which reason we assigned scale and rate parameters of the Gamma distributions so that they give more weight for the smaller values of the process parameters. For each intercept term we assigned weakly informative Student-$t$ priors $\beta_i \sim \text{Student} - t(0, 2.5^2, 4)$. In order to model coregionalization matrix $\Sigma_\epsilon$ efficiently we use a separation strategy (e.g. Barnard et al., 2000) where coregionalization matrix is decomposed into correlation matrix $\Omega$ and vector of standard deviations $\omega$ such that $\Sigma_\epsilon = \text{diag}(\omega)\Omega\text{diag}(\omega)$. Correlation matrix $\Omega$ describes the interspecific dependencies and we gave it LKJ-prior with unit shape (Lewandowski et al., 2009) that defines LKJ-prior distribution which is marginally uniform over all correlation parameters (Vanhatalo et al., 2020). We gave half-Student-$t$ prior for the standard deviations, $\omega \sim \text{Student} - t_+(0, 4^2, 4)$.

### 2.4 Posterior inference

All the alternative species distribution models were implemented using Stan via Rstan (Stan Development Team, 2018) with which we conducted posterior sampling for all the model parameters and latent variables. We ran four parallel Markov chains of 2000 iterations such that first 1000 iterations of each chain were discarded as warmup. Posterior sampling was done using dynamic Hamiltonian Monte Carlo Sampler as coded in Stan version 2.18.1. The convergence and effective sample sizes were checked using trace- and autocorrelation plots and through potential scale reduction factor and Geyer’s initial monotone sequence criterion. After conducting the posterior sampling, we drew posterior predictive samples of the species specific percentage covers at prediction locations covering the study area (see,
Table 1: Summary of the model comparison and validation methods that are described in detail in Section 3.

| Predictive task | Model comparison | Model validation |
|-----------------|------------------|------------------|
| Species specific percentage cover maps | \( CV_1 \): An average of species- and location-wise log predictive densities. Equation (12). | \( PIT_1 \): Species specific PIT histograms of location-wise predictive distributions. Equations (16) and (17). |
| Total percentage cover maps | \( CV_2 \): An average of location-wise log joint over species predictive densities. Equation (13). | \( PIT_2 \): PIT histogram of location-wise total over species predictive distributions. Equations (16) and (19). |
| Species specific total vegetation cover over the study area | \( CV_3 \): An average of species- and CV-fold-wise log joint over locations (within a CV-fold) predictive densities. Equation (14). | \( PIT_3 \): Species specific PIT histograms of total over locations (within a CV-fold) predictive distributions. Equation (16) and CDF is estimated analogously to (19). |
| Total vegetation cover over the study area | \( CV_4 \): An average of CV-fold-wise log joint over species and locations (within a CV-fold) predictive densities. Equation (15). | \( PIT_4 \): PIT histogram of CV-fold-wise total over species and locations (within a CV-fold) predictive distributions. Equation (16) and CDF is estimated analogously to (19). |

e.g., Vanhatalo et al. (2020).

3 Model comparison and validation

The central aims of our study are to provide posterior predictive maps for the vegetation cover within the study area, and to provide posterior predictive distributions for the total vegetation cover over the study area. Hence, it is natural to compare alternative models with the goodness of their out-of-sample posterior predictive distributions. This can be done with cross-validation (CV) using the log posterior predictive density diagnostics (Vehtari and Ojanen, 2012). Model comparison indicates the best model among the alternatives but does not tell whether any of the models has actually good fit for the purpose. Hence, after choosing the best model we assess the goodness of its predictive distributions using probability integral transform (PIT) statistics (Gneiting et al., 2007). All the model comparison and model validation methods are summarized in Table 1 and we explain them in detail next.

3.1 Predictive model comparison with cross-validation

Species specific percentage cover maps (so called, species distribution maps) are produced by forming a lattice mesh over the study area and calculating point-wise posterior predictive distributions for all the mesh cells. Later the summary statistics (mean, variance, quantiles, etc.) of these predictive distributions can be drawn as maps. On the other hand, when predicting the total vegetation cover, we need to calculate the joint predictive distribution over all the mesh cells. Moreover, we want to compare models predictive performance both
by species and for the total cover by all species. Hence, we constructed our own cross-validation splitting strategy for each of these tasks.

In order to compare models in terms of producing species-specific percentage cover maps, we divided the observed dataset randomly into $K$ distinct subsets, indexed by sets of locations $S_1, ..., S_K$ such that the full data set is given by $\bigcup_{k=1}^{K} S_k = S$. The single species, point-wise, predictive performance was then

$$CV_1 = \frac{1}{Jn} \sum_{i=1}^{n} \sum_{j=1}^{J} \log \left( \pi(y_{ij}|Y(S_{k(i)}), S_{k(i)}, s_i) \right),$$

(12)

where $\pi(y_{ij}|Y(S_{k(i)}), S_{k(i)}, s_i)$ is the cross-validation posterior predictive density for $y_{ij}$. The set $k(i)$ is the cross-validation set that contains location $s_i$, and set $Y(S_{k(i)})$ includes all other species observations except the observations in the cross-validation set $k(i)$. The comparison of models in the task of producing total vegetation cover maps was done analogously so that we calculated the average log (spatially) point-wise joint over $k$ species posterior predictive density. This leads to $K$-fold cross-validation criterion

$$CV_2 = \frac{1}{n} \sum_{i=1}^{n} \log \left( \pi(y_i|Y(S_{k(i)}), S_{k(i)}, s_i) \right)$$

(13)

where $\pi(y_i|Y(S_{k(i)}), S_{k(i)}, s_i)$ is the joint posterior predictive probability mass function of all the species at location $i$. To evaluate models’ performance in predicting per species total percentage cover over an area we calculated first the log joint over locations in a CV-fold posterior predictive density for all CV-folds and then took an average of them; that is

$$CV_3 = \frac{1}{KJ} \sum_{k=1}^{K} \sum_{j=1}^{J} \log \left( \pi(Y_j(S_k)|Y(S_{k(i)}), S_{k(i)}), s_i) \right),$$

(14)

where $Y_j(S_k)$ collects all the observations of the $j$th species at locations $S_k$. Similarly, to evaluate models’ performance in predicting the total percentage cover over all species and an area we calculated the average of CV-fold-wise log joint predictive densities with

$$CV_4 = \frac{1}{K} \sum_{k=1}^{K} \log \left( \pi(Y(S_k)|Y(S_{k(i)}), S_{k(i)}), s_i) \right),$$

(15)

where $Y(S_k)$ collects observations of all species at locations $S_k$. Note that in all above cross-validation criterion we have calculated the average over individual log predictive densities. Hence, the statistics are in comparable scale.

We used $k = 10$ in all cross-validation metrics. Moreover, we estimated the posterior predictive densities using Monte Carlo over Markov chain samples from the posterior distribution. For example, in $CV_1$ the posterior predictive density is approximated with

$$\pi(y_{ij}|Y(S_{k(i)}), S_{k(i)}, s_i) \approx \frac{1}{M} \sum_{m=1}^{M} \pi(y_{ij}|f^{(m)}(ij), \gamma^{(m)})$$

where $f^{(m)}(ij)$ and $\gamma^{(m)}$ denote the $m$th posterior sample from the joint posterior (predictive) distribution of the latent variable and Dirichlet model parameters respectively; that is $f^{(m)}(ij), \gamma^{(m)} \sim p(f_{ij}, \gamma|Y(S_{k(i)}), S_{k(i)}, s_i)$. In $CV_2$ the posterior predictive density is approximated with

$$\pi(y_i|Y(S_{k(i)}), S_{k(i)}, s_i) \approx \frac{1}{M} \sum_{m=1}^{M} \pi(y_i|f(s_i)^{(m)}, \gamma^{(m)})$$

and analogously in $CV_3$ and $CV_4$. After obtaining the Markov
chain samples for the model parameters and latent variables at data locations, the posterior predictive samples for the latent variables at prediction locations can be constructed in Gibbs style using the full Gaussian conditional distribution of the latent variables (see Vanhatalo et al. 2020, Section 4.1).

If the samples of posterior predictive densities, such as $\pi(y_{ij}|f_{ij}^{(m)}, \gamma^{(m)})$, have large variance, the Monte Carlo approximation for log predictive densities will have high variance as well and through that the sample based cross-validation criterion becomes unreliable. Large variance in the Monte Carlo approximation for the posterior predictive densities is likely, especially in case of $CV_2$ - $CV_4$ where we need to calculate the joint density over multiple observations. We used bootstrapping with 1000 replicates to estimate the uncertainty in the CV criterion induced by the Monte Carlo approximation for the log predictive densities. Each bootstrap replicate of the CV criterion was based on a random sample of size $M$ with replacement from the posterior sample of model parameters and latent variables.

### 3.2 Model validation with PIT histograms

A key property of a predictive distribution is its calibration (Gneiting et al., 2007) which we evaluated using randomized probability integral transform (PIT) histograms proposed by Denuit and Lambert (2005). Randomized PIT histogram is a graphical method to evaluate whether data, $y_{ij}$, can be considered as a random sample from the discrete predictive distribution that is given by the fitted model or not. To construct randomized PIT histogram for species- and location-wise predictive distributions we define

$$u_{ij} = F(y_{ij} - r_{\text{min}}) + v_{ij} (F(y_{ij}) - F(y_{ij} - r_{\text{min}})),$$

where $v_{ij}$ is a draw from standard uniform distribution, $F(y_{ij})$ is the posterior predictive cumulative distribution function (CDF) for observation $y_{ij}$, and $r_{\text{min}}$ is the minimum gap between any possible adjacent values of $y$. We set $F(y_{ij} - r_{\text{min}}) = 0$ if $y_{ij} - r_{\text{min}} < 0$ since percentage cover cannot be negative. Calibration is evaluated graphically by plotting histogram of the values $u_{ij}$ since they follow standard uniform distribution if $F(y_{ij})$ corresponds to the true data generating process (Gneiting et al., 2007).

The PIT histograms were drawn for cross-validation predictive distributions using the same splitting strategy as in the cross-validation tests. Hence, $PIT_1$ corresponds to species- and location-wise predictions for $y_{ij}$, $PIT_2$ to location-wise predictions for the sum over species, $\bar{y}_i = \sum_j y_{ij}$, $PIT_3$ to species-wise predictions for the sum over locations, $\bar{y}_j = \sum_i y_{ij}$, and $PIT_4$ to predictions for the sum over species and locations, $\bar{y} = \sum_{ij} y_{ij}$ (Table 1). We used Monte Carlo approximation for the posterior predictive CDFs. For $PIT_1$ we approximated the posterior predictive CDF directly as

$$F_1(y_{ij}|Y(S\setminus k(i)), S\setminus k(i), s_i) = \sum_{z=0}^{y_{ij}} \frac{1}{M} \sum_{m=1}^{M} \pi(z|f_{ij}^{(m)}, \gamma^{(m)}).$$

In $PIT_2$–$PIT_4$ we first constructed an estimators for the posterior predictive densities for the sums over species and/or locations conditional on latent variables and parameters.
These were calculated for each posterior sample using Monte Carlo so that, for example,

\[
\pi(\bar{y}_i | f^{(m)}_1, \gamma^{(m)}) = \frac{1}{B} \sum_{b=1}^{B} \bar{y}_i \left( \sum_{j=1}^{J} \bar{y}_{ij}^{(b,m)} \right)
\]  

where \( [\bar{y}_{i,1}^{h,m}, \ldots, \bar{y}_{i,J}^{h,m}] \sim \pi(y_i | f^{(m)}_1, \gamma^{(m)}) \). After this we calculated

\[
F_2(\bar{y}_i | Y(S_{(k(i))}, S_{(k(i))}, s_i) = \sum_{z=0}^{M} \frac{1}{M} \sum_{m=1}^{M} \pi(z | f^{(m)}_1, \gamma^{(m)})
\]

The CDFs for \( PIT_3 \) and \( PIT_4 \) were estimated analogously.

### 4 Application to greenhouse gas flux modeling

#### 4.1 The role of leaf area in greenhouse gas flux

The carbon gas exchange between an ecosystem and atmosphere is primarily determined by the amount of photosynthetically active biomass and the area of green leaves – even though abiotic factors, such as the light and water availability and temperature, play a role as well (e.g. Peichl et al., 2018). Therefore, knowledge about the abundance and structure of the vegetation is an essential input for modeling the ecosystem climatic impact (e.g. Korrensalo et al., 2019) and for ecosystem models in general. In this context, peatland ecosystems are of special interest because they store approximately one third of global terrestrial carbon in their peat layer (Gorham, 1991) and are the largest natural source of atmospheric methane (e.g. Heilig, 1994). Different peatland species have distinct photosynthetic capacities and differ by the substrate quality they provide for decomposition processes. Further, the abundance of different vascular plant species is an important control for ecosystem-scale methane efflux, as certain species act as conduits of methane from the peat to the atmosphere (Bhullar et al., 2013), bypassing the microbial methane oxidation in the oxidation peat layers (Larmola et al., 2010). Plant species groups also differ by their emissions of biogenic organic compounds, that have a net cooling effect on climate (Tiiva et al., 2009; Faubert et al., 2011).

#### 4.2 Study site and data

The study site is located at a boreal poor fen, which is part of Siikaneva peatland complex in Southern Finland (Figure 2). The study site has an eddy covariance tower which is used for monitoring the CO\(_2\) and CH\(_4\) fluxes within a 200-meter radius footprint around the tower that extends from the margin of the peatland towards the center. Typical to aapa mires, the margin is poorer in nutrients and slightly drier than the wetter center. To quantify the spatial variation of the vegetation within the eddy covariance tower footprint, a vegetation inventory was done in the summer of 2017. The grid sampling design with 328 inventory plots was applied within 200 m distance from the eddy covariance tower (Figure 2). Locations that hit to ditch and mineral soil were excluded from the grid and thus the formed grid is not symmetric. We selected randomly 200 inventory plots to fit the models.
The bottom layer of the site is formed by mosses (mainly Sphagnum L. species) and the field layer consists of vascular plants adapted to the conditions where water table prevails close to the surface. Both of these vegetation layers have a significant role in the greenhouse gas cycling of the site. For this work, we selected the six most common Sphagnum (S.) mosses: Sphagnum papillosum, S. balticum, S. fallax, S. magellanicum, S. majus and S. angustifolium and the following eight ecologically interesting vascular plants: Carex lasiocarpa, Carex limosa, Carex rostrata, Empetrum nigrum, Eriophorum vaginatum, Pinus sylvestris, Rubus chamaemorus and Scheuchzeria palustris. The Sphagnum species are mutually exclusive (Gong et al., 2020) but the vascular plants can grow in layers that overlap with each other and the Sphagnum layer. To minimize human-related errors, such as misclassified or omitted species (Kennedy and Addison, 1987), each plot was inventoried by two persons. For every inventory plot, the percentage cover of each species within a circular frame of 0.071 m² (radius 15 cm) was estimated visually. For each species the percentage cover was measured as a total vertical projection on the ground and reported in 0.25 % percentage unit accuracy at cover below 1 % and in 1 % percentage accuracy at cover above it. For species names we followed The Plant List (2013).

4.3 Case study analyses

There is one species group formed by the Sphagnum moss species that we consider to be mutually exclusive. Each vascular plant forms its own species group since they can grow in overlapping layers and on top of Sphagnum mosses. Since Sphagnum mosses had larger than 1% cover throughout the study region we set $N_{i,a} = 100$ for them corresponding to the used 1 percentage unit measurement accuracy. Vascular plant percentage cover measurements ranged from below to above 1% so we assumed in average 0.5 percentage unit accuracy for them leading to $N_{i,a} = 200$ for vascular plants. To select the most suitable model structure and to evaluate the effect of alternative model structures to leaf area predictions we compared 12 models with different combinations of alternative descriptions for species.
niche preferences and competition. All the compared models are summarized in Table 2 and their DAGs are summarized in figures 1 and S1–S3.

Table 2: Summary of the alternative models compared in the case study. The abbreviations in the model names are: C for constant latent function ($\epsilon_j(s_i) = 0$) modeling constant niche preference; IGP$_S$ and IGP$_{NS}$ denote respectively independent stationary and non-stationary spatial GPs modeling independent heterogeneous niche preferences; LMC($k$)$_S$ and LMC($k$)$_{NS}$ denote respectively stationary and non-stationary linear model of coregionalization with $k$ distinct covariance functions modeling dependent heterogeneous niche preferences; BB and DM denote respectively Beta-Binomial and Dirichlet-Multinomial processes modeling mutually independent or competing species.

| Model          | Species niche preference                          | Competition |
|----------------|--------------------------------------------------|-------------|
|                | Constant  | Independent heterogeneous | Dependent heterogeneous | No | Yes |
| C+BB           |          |                         |                         |    |     |
| C+DM           |          |                         |                         |    |     |
| IGP$_S$+BB     |          |                         |                         |    |     |
| IGP$_{NS}$+BB  |          |                         |                         |    |     |
| IGP$_S$+DM     |          |                         |                         |    |     |
| IGP$_{NS}$+DM  |          |                         |                         |    |     |
| LMC(1)$_S$+BB  |          |                         |                         |    |     |
| LMC(1)$_{NS}$+BB |        |                         |                         |    |     |
| LMC(1)$_S$+DM  |          |                         |                         |    |     |
| LMC(1)$_{NS}$+DM |       |                         |                         |    |     |
| LMC(2)$_S$+DM  |          |                         |                         |    |     |
| LMC(2)$_{NS}$+DM |       |                         |                         |    |     |

We compared the alternative models and conducted the model validation by using the CV criteria and the randomized PIT histograms summarized in Table 1. We chose the model with the best overall predictive performance for producing the final results concerning species specific and total vegetation cover. We calculated the vegetation cover for each species separately, across all species and across both species groups (mosses and vascular plants) by first constructing a regular mesh, $D_M$, over the study area, $D$, and then sampling from the posterior predictive distribution for vegetation cover at each mesh cell. The grid cell size for the mesh ($A = 4$ m$^2$) was chosen so that change in the estimate for posterior predictive variance of the total vegetation cover $Var(\tilde{\phi})$ was negligible when grid cell size was further reduced.

5 Results

5.1 Model comparison

Table 3 summarizes the model comparison results for the alternative models. In all predictive tasks, the models with constant latent functions (C+BB and C+DM) had the lowest cross-validation predictive performance. This gives strong support for spatial heterogeneity in species specific vegetation covers. However, in the species-wise total over the region (CV$_3$) prediction, the difference to the best model was not significant if measured relative to standard error of the cross-validation estimates. In general, models with Dirichlet-Multinomial observation model outperform the otherwise same models with Beta-Binomial
in joint species predictions (CV₂ and CV₄) and perform practically equally well in single species predictions (CV₁ and CV₃). This indicates that accounting for species competition is more important for community predictions than for single species predictions.

The best performing model in terms of CV₁ and CV₃ was LMC(1)ₛ+BB with practically equal performance by LMC(1)ₛ+DM and LMC(2)ₛ+DM. However, the CV₁ and CV₃ estimates of all spatially heterogeneous models except IGPₛ+DM in CV₁ were within one standard error of the best cross-validation estimate indicating that these differences were practically negligible. When looking at CV₁ and CV₃ for each species separately, the relative differences between alternative models were in average small but the best model for all species was always a model that had either stationary or non-stationary LMC latent function prior. In terms of CV₂ and CV₄ the best performing models were LMC(1)ₛ+DM, LMC(2)ₛ+DM and LMC(1)ₙₛ+DM. The former two models had practically equal CV₂ and CV₄ estimates and those of the latter were within one standard error estimate from the former two. The relatively larger difference between LMC(1)ₙₛ+DM and the best model might be, at least partly, explained by worse mixing of its Markov chain compared to the Markov chains of LMC(1)ₛ+DM and LMC(2)ₛ+DM which induces more weight to occasional low log predictive density values. The CV₂ and CV₄ estimates of all other models were significantly worse than those of the best three models. The Markov chains for IGPₙₛ+DM and LMC(2)ₙₛ+DM converged slowly and their mixing was poor resulting into considerably smaller effective sample size for the log predictive densities compared to other models so we excluded them from the cross-validation comparison.

Table 3: Model comparison results with 10-fold cross-validation using log predictive density utilities (CV₁, ..., CV₄; see Table [II]) together with the standard error estimates (se) and the Monte Carlo error estimate (me) for the CV estimates (see Section [3]). The best CV estimate is indicated by bold font and the CV estimates that are within one standard error from the best model are indicated by italic font. Rows with – indicate models whose effective sample size for log predictive density was considerably smaller than for the rest of the model and, hence, were excluded from the comparison.

| Model    | CV₁ (se/me)       | CV₂ (se/me)       | CV₃ (se/me)       | CV₄ (se/me)       |
|----------|------------------|------------------|------------------|------------------|
| C+BB     | -1.80 (4e-2/3e-5)| -25.2 (3e-1/5e-4)| -35.9 (2/6e-4)  | -503.0 (8/2e-2)  |
| C+DM     | -1.80 (4e-2/3e-5)| -23.8 (4e-1/5e-4)| -36.1 (2/7e-4)  | -474.9 (9/2e-2)  |
| IGPₛ+BB  | -1.69 (4e-2/1e-5)| -23.6 (4e-1/3e-3)| -33.5 (2/6e-3)  | -474.5 (10/4e-1) |
| IGPₙₛ+BB | -1.71 (4e-2/3e-4)| -23.9 (4e-1/7e-3)| -34.1 (2/2e-2)  | -487.1 (10/6e-1) |
| IGPₛ+DM  | -1.75 (4e-2/5e-4)| -23.1 (4e-1/1e-2)| -34.1 (2/2e-2)  | -460.3 (10/9e-1) |
| IGPₙₛ+DM |                |                  |                  |                  |
| LMC(1)ₛ+BB | -1.67 (4e-2/1e-4)| -23.3 (4e-1/4e-3)| -33.4 (2/6e-3)  | -470.3 (10/4e-1) |
| LMC(1)ₙₛ+BB | -1.70 (4e-2/7e-4)| -23.6 (5e-1/1e-2)| -34.5 (2/3e-2)  | -495.8 (14/9e-1) |
| LMC(1)ₛ+DM | -1.68 (4e-2/1e-4)| -21.9 (4e-1/3e-3)| -33.5 (2/6e-3)  | -440.1 (10/4e-1) |
| LMC(1)ₙₛ+DM | -1.71 (4e-2/4e-4)| -22.1 (4e-1/9e-3)| -34.2 (2/2e-2)  | -450.4 (12/8e-1) |
| LMC(2)ₛ+DM | -1.68 (4e-2/1e-4)| -21.9 (4e-1/3e-3)| -33.5 (2/5e-3)  | -439.2 (10/4e-1) |
| LMC(2)ₙₛ+DM |                |                  |                  |                  |
5.2 Model validation

Model comparison with cross-validation did not indicate clear differences between the best models: LMC(1)\textsubscript{S}+DM, LMC(2)\textsubscript{S}+DM and LMC(1)\textsubscript{NS}+DM. Moreover, the former two were practically the same model since LMC(2) effectively reduced to LMC(1). Hence, we conducted model validation for the latter two. The randomized PIT\textsubscript{2} and PIT\textsubscript{4} histograms for the LMC(1)\textsubscript{NS}+DM and LMC(2)\textsubscript{S}+DM models (Figures 3 and S4) show moderate departure from uniformity and they indicate that the total percentage cover is to some extent underestimated in average over the cross-validation folds. However, it should be noted that the PIT\textsubscript{2} and PIT\textsubscript{4} histograms are formed from only 200 and 32 posterior predictive distributions so the deviation from uniformity is likely also due randomness. The species wise PIT\textsubscript{1} histograms for LMC(2)\textsubscript{S}+DM and LMC(1)\textsubscript{NS}+DM models (Figures S5 and S6) are close to uniform for all other species except S. papillosum for which the PIT\textsubscript{1} histogram is slightly ∪-shaped indicating underestimation of predictive uncertainty. This deviation from uniformity was more distinct for LMC(2)\textsubscript{S}+DM than LMC(1)\textsubscript{NS}+DM. Out of all species, S. papillosum was the one showing the most clear non-stationarity in the posterior of the latent function of LMC(1)\textsubscript{NS}+DM model giving some evidence that non-stationary Gaussian process prior might have been beneficial in this case. The species wise PIT\textsubscript{3} histograms for LMC(2)\textsubscript{S}+DM and LMC(1)\textsubscript{NS}+DM models (Figures S7 and S8) do not show clear deviations from uniformity even though the histograms are more noisy than the PIT\textsubscript{1} histograms.

5.3 Percentage cover predictions and species interactions

Based on the model comparison and validation results we selected LMC(1)\textsubscript{NS}+DM for the final inference. The percentage covers of the species have clearly different spatial patterns (Figure 4). Spatial distribution of the vegetation is typical for minerotrophic fen which center is lower than its edges and therefore also water table is on average higher (wetter) at the center. As spatial composition of sphagna and vascular plants follows the variation in the water table (WT), plants adapted to grow in drier areas are concentrated to edges of the study area and flarks are inhabited by species that can withstand waterlogging.
Hummock species like *S. angustifolium* and *E. nigrum* favoring drier habitat are mostly growing at the edges of the study area. Lawn species *S. papillosum* dominates the composition of sphagna grows fairly evenly throughout the area. Lawn community type can be further divided into three subgroups specified by vascular plants (from drier to wetter) *E. vaginatum* lawn, *C. rostrata* lawn and *C. lasiocarpa* lawn. *C. rostrata* and *C. lasiocarpa* grow mainly on the southern edges of the study area while percentage cover of *E. vaginatum* tends to be higher on the northern edges. Species (*S. majus*, *C. limosa* and *S. palustris*) associated to wetter growing conditions (hollows) are concentrated in smaller hotspots at the center of the study region.

Combined sphagnum cover (sum of sphagnum species covers) is relatively stable (85-95%) over the study area although species-wise distributions have distinct spatial patterns. Contrarily, spatial distribution of combined vascular plants cover has high variation such that hotspots of higher combined covers are located on both wetter centre areas and drier parts of the study area.

The estimated interspecific correlations in niche preferences are summarized in Figure 5 and the interspecific competition between *spaghnum* species is summarized in Figure 6 in terms of the (negative) correlation in their percentage cover. Species can be differentiated into co-occurring groups according to the interspecific dependencies between species niche preferences. Sphagnums *S. majus* and *S. fallax* with vascular plants *S. palustris* and *C. limosa* adapted to grow in wetter conditions form one cluster. Lawn species *C. lasiocarpa* and *C. rostrata* can be classified into the second co-occurring cluster. High lawn-low hummock species *S. angustifolium*, *S. magellanicum* and *P. sylvestris* form the third cluster. Dwarf shrubs *E. nigrum* and *R. chamaemorus* form the fourth cluster (both species were also positively correlated with *P. sylvestris* but 80% credibility intervals for the correlation between *R. chamaemorus* and *P. sylvestris* overlapped zero). Lawn species *S. papillosum*, *S. balticum* and *E. vaginatum* were also positively correlated but 80% credibility intervals for correlations overlapped zero. The interspecific correlations in niche preference followed qualitatively similar pattern in all LMC models. However, models with Beta-Binomial observation model found more negative interspecific correlations than models with Dirichlet-Multinomial observation model as illustrated by comparison of figures 5 and S10.

Estimated total vegetation cover distributions for each model presented in Figure S9 show that non-stationary JSDM models predicted combined covers that are close to the observed sample means. All stationary models tend to predict combined cover to be smaller than their non-stationary counterparts. In each case 95% prediction intervals for non-stationary JSDM models overlap observed sample mean but for stationary models 95% prediction intervals do not overlap observed sample mean. This suggest that there is non-stationarity in the observation process that should be taken into account or received vegetation cover estimates might be biassed. Stacked species distribution models tend to overpredict combined vascular plants cover but underpredict combined sphagnum and total vegetation covers.
Figure 4: The posterior predictive means of the percentage covers and probability distributions of the combined percentage covers over the study area as predicted by LMC(1)_{NS+DM} model. (A) species-wise percentage covers of the sphagnums, (B) total percentage cover of all sphagnum species combined and the total combined cover over the study area, (C) species-wise percentage covers of vascular plants, (D) total percentage cover of vascular plants combined and the total combined cover over the study area and (E) combined total cover of sphagnums and vascular plants. For combined totals over the area probability distributions are estimated using kernel density estimation on MCMC samples. Solid vertical lines show the posterior means for the combined covers, dashed lines show the sample means computed from the training data.
Figure 5: Posterior mean estimates of the interspecific correlations in the species niche preference in the LMC(1)_{NS+DM} model. White cells indicate that the 80% posterior credible interval of the correlation overlapped zero (i.e., weak support for interspecific correlation) and stars indicate that the 95% posterior credible interval did not overlap zero (i.e. strong support for interspecific correlation).

Figure 6: The interspecific competition between *spagnum* species. The maps on the lower left corner show the spatial distribution of the (negative) interspecific correlation in percentage covers for all pairs of *spagnum* species. The matrix on the upper right corner shows the average of these correlations over the whole study area.
In this work, we proposed a novel joint species distribution model which extends the existing JSDMs in two important ways. First, we incorporated species competition into JSDM through Dirichlet-multinomial distribution and, second, we introduced non-stationary multivariate Gaussian process to the spatial random effects of JSDMs. Moreover, we proposed a decision theoretic model comparison and validation approach that applies cross-validation to measure models’ performance in four different types of predictive tasks. We also applied our model to a topical application concerning estimation of vegetation cover in a peatland that is used in global greenhouse gas flux studies.

Dirichlet-Multinomial model has been used to model vegetation cover data, for example, in ordination settings (Damgaard et al., 2020) but our model is the first one combining it with formal joint species distribution modeling framework. Current state-of-the-art JSDMs, such as the hierarchical modeling of species communities model (Ovaskainen and Abrego, 2020), capture species dependence through an associated correlation matrix arising from a set of latent multivariate normal variables. However, these associations offer limited insight into realized dependence behavior between species at sites (Gelfand and Shirota, 2021) and cannot, in reality, separate environmental effects from biotic interactions (Poggiato et al., 2021). Our model overcomes this deficiency in a setting where species are known to be exclusive due to limited resource, such as growing space. The basis of the proposed model is the use of hierarchical modelling approach which allows estimation of species niche preferences in one latent layer of the model (the latent Gaussian layer) and species competition in the other (the Dirichlet process layer).

Modelling dependencies between species niche preferences has been shown to be useful in many studies since it enables information sharing between species which then improves the model estimates and predictions especially for rare species (Ovaskainen and Soininen, 2011; Vanhatalo et al., 2020; Nordberg et al., 2019). Also our case study results support these findings. Our case study results also show that explicit modelling of species competition improves models predictive performance especially in predictions concerning community structure and total percentage cover ($CV_2$ and $CV_4$ in table 3). Hence, we believe that combining the Dirichlet-Multinomial layer with the hierarchical latent Gaussian layer of, for example, the hierarchical modeling of species communities model, can improve its predictive and inference performance considerably in applications where interspecific competition is likely. In this work, we assigned the species into competing and non-competing groups manually but the model could be extended to allow for uncertain group labels by modeling the group associations with, for example, Dirichlet processes (Taylor-Rodríguez et al., 2017).

Our results showed also that the relative performances of alternative models are different in different predictive tasks which highlights the importance of tailoring the model comparison method to the specific task at hand. The best models (LMC(1)$_S$+DM, LMC(2)$_S$+DM and LMC(1)$_{NS}$+DM) were among the best in all predictive tasks whereas other models performed equally well as they only in a few of the predictive tasks. Namely, all spatially heterogeneous models were practically equally good in producing species distribution maps and per species total cover estimates in terms of their log predictive density estimates ($CV_1$ and $CV_3$ see table 3). However, in joint species predictions the best models stood out as the best clearly. Interestingly, the spatially homogeneous models C+BB and C+DM
did not differ significantly from the other models in per-species total over area predictions ($CV_3$) indicating that a sample mean over uniform random locations can give a reasonable estimate for species-wise total covers. This is supported also by Figure S9 where the total percentage cover estimates of C+BB and C+DM do not differ noticeably from the rest of the total percentage cover estimates. However, 

Foster et al. (2021) give an example on how sample mean over uniform random locations may fail if the underlying field is very heterogeneous.

Even though the differences between the best models were small, the proposed non-stationary JSDM model seemed to work somewhat better in general than its stationary counterpart. The main reason for this was that some species, most clearly S. papillosum, had non-stationary patterns in their latent Gaussian field (see Figure 4). Moreover, models did not differ only in their predictive properties but they lead to different inference results concerning interspecific correlations in niche preference. Models with Beta-Binomial observation model found more negative interspecific correlations in species niche preferences than models with Dirichlet-Multinomial observation model (see figures 5 and S10).

The case study results are highly congruent with the earlier ecological knowledge of the species in the study area. Results also suggest that proposed model performs well in interpolating vegetation cover over the area and also describing general patterns and composition of the vegetation. Natural extensions to presented model is to include additional predictors to explain distribution of the vegetation cover. For example, height of water table could improve the model fit further as this is a commonly known driver of vegetation patterns in peatlands (Andersen et al. 2011). Also covariates related to soil properties could be beneficial in describing small scale variation of vegetation (Andersen et al. 2011). However, information on such covariates was not available for our case study area.

Proposed model framework can be broadly used in applications where estimates of vegetation distributions are needed. It is expected that estimating areal vegetation composition more precisely can be beneficial for example in areal greenhouse gas emission models since vegetation is important explanatory factor in carbon dioxide and methane flux models. Case study results demonstrate that presented model offers valid method to estimate areal vegetation cover for those purposes.

Acknowledgements

This work was funded by the Doctoral Programme in Science, Technology and Computing of the University of Eastern Finland through support to JK, by University of Eastern Finland, Faculty of Science and Forestry, through support to AK, and by the Academy of Finland through a grant 317255 to JV, grants 337550 and 337549 to E-ST and grant 338980 to AK, and by Tiina and Antti Herlin Foundation through a grant to JK.
References

Andersen, R., M. Poulin, D. Borcard, R. Laiho, J. Laine, H. Vasander, and E.-S. Tuittila (2011). Environmental control and spatial structures in peatland vegetation. *Journal of Vegetation Science* 22, 1–13.

Barnard, J., R. McCulloch, and X. L. Meng (2000). Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. *Statistica Sinica* 10, 121–1311.

Bluhlar, G., M. Iravani, P. Edwards, and H. Venterink (2013). Methane transport and emissions from soil as affected by water table and vascular plants. *BMC Ecology* 13.

Damgaard, C., R. R. Hansen, and F. K. Hui (2020). Model-based ordination of pin-point cover data: Effect of management on dry heathland. *Ecological Informatics* 60, 101155.

Demui, M. and P. Lambert (2005). Constraints on concordance measures in bivariate discrete data. *Journal of Multivariate Analysis* 93, 40–57.

Faubert, P., P. Tiiva, T. Nakam, J. Holopainen, and R. Rinnan (2011). Non-methane biogenic volatile organic compound emissions from boreal peatland microcosms under warming and water table drawdown. *Biogeochemistry* 106, 503–516.

Foster, S. D., J. Vanhatalo, V. M. Trenkel, T. Schulz, E. Lawrence, R. Przeslawski, and G. R. Hosack (2021). Effects of ignoring survey design information for data reuse. *Ecological Applications* 31(6), e02360.

Gelfand, A. E., M. Holder, A. Latimer, P. O. Lewis, A. G. Rebelo, J. A. Silander Jr., and S. Wu (2006). Explaining species distribution patterns through hierarchical modeling. *Bayesian Analysis* 1, 41–92.

Gelfand, A. E. and S. Shirota (2021). The role of odds ratios in joint species distribution modeling. *Environmental and Ecological Statistics* 28, 287–302.

Gneiting, T., F. Balabdaoui, and A. E. Raftery (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society, Series B* 69, 243–268.

Gong, J., N. Roulet, S. Frolking, H. Peltola, M. Laine, N. Kokkonen, and E.-S. Tuittila (2020). Modelling the habitat preference of two key sphagnum species in a poor fen as controlled by capitulum water content. *Biogeosciences* 17, 5693–5719.

Gorham, E. (1991). Northern peatlands: Role in the carbon cycle and probable responses to climatic warming. *Ecological Applications* 1, 182–195.

Heilig, G. K. (1994). The greenhouse gas methane (ch4): Sources and sinks, the impact of population growth, possible interventions. *Population and Environment* 16, 109–137.

Kennedy, K. A. and P. A. Addison (1987). Some considerations for the use of visual estimates of plant cover in biomonitoring. *Journal of Ecology* 75, 151–157.
Korrensalo, A., L. Mehtätalo, P. Alekseychik, S. Uljas, I. Mammarella, T. Vesala, and E.-S. Tuittila (2019). Varying vegetation composition, respiration and photosynthesis decrease temporal variability of the CO2 sink in a boreal bog. *Ecosystems* 23, 842–858.

Larmola, T., E. Tuittila, M. Tiirola, H. Nykänen, P. J. Martikainen, K. Yrjälä, T. Tuomivirta, and H. Fritze (2010). The role of sphagnum mosses in the methane cycling of a boreal mire. *Ecology* 91, 2356–2365.

Lewandowski, D., D. Kurowicka, and H. Joe (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis* 100, 1989–2001.

Nordberg, A., N. Abrego, F. G. Blanchet, F. R. Adler, B. J. Anderson, J. Anttila, M. B. Araújo, T. Dallas, D. Dunson, J. Elith, S. D. Foster, R. Fox, J. Franklin, W. Godsoe, A. Guisan, B. O'Hara, N. A. Hill, R. D. Holt, F. K. C. Hui, M. Husby, J. A. Kålas, A. Lehikoinen, M. Luoto, H. K. Mod, G. Newell, I. Renner, T. Roslin, J. Soïninen, W. Thuiller, J. Vanhatalo, D. Warton, M. White, N. E. Zimmermann, D. Gravel, and O. Ovaskainen (2019). A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological Monographs* 89, e01370.

Ovaskainen, O. and N. Abrego (2020). *Joint Species Distribution Modelling With Applications in R*. Cambridge University Press.

Ovaskainen, O. and J. Soininen (2011). Making more out of sparse data: hierarchical modeling of species communities. *Ecology* 92, 289–295.

Paciorek, C. J. and M. J. Schervish (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics* 17, 483–506.

Peichl, M., M. Gažovič, I. Vermeij, E. de Goede, J. L. O. Sonnentag, and M. B. Nilsson (2018). Peatland vegetation composition and phenology drive the seasonal trajectory of maximum gross primary production. *Nature Scientific Reports* 8(1), 8012.

Poggiaito, G., T. Münkemüller, D. Bystrøva, J. Arbel, J. S. Clark, and W. Thuiller (2021). On the interpretations of joint modeling in community ecology. *Trends in Ecology & Evolution*.

Stan Development Team (2018). The stan core library, version 2.18.2. [http://mc-stan.org](http://mc-stan.org), visited on 2019-06-13.

Taylor-Rodríguez, D., K. Kaufeld, E. M. Schliep, J. S. Clark, and A. E. Gelfand (2017). Joint Species Distribution Modeling: Dimension Reduction Using Dirichlet Processes. *Bayesian Analysis* 12(4), 939 – 967.

The Plant List (2013). The plant list. [http://www.theplantlist.org](http://www.theplantlist.org), visited on 2021-10-18.

Tiiva, P., P. Faubert, S. Röy, J. Holopainen, and R. Rinnan (2009). Contribution of vegetation and water table on isoprene emission from boreal peatland microcosms. *Atmospheric Environment* 43, 5469–5475.
Vanhatalo, J., M. Hartmann, and L. Veneranta (2020). Additive multivariate gaussian processes for joint species distribution modeling with heterogeneous data. *Bayesian analysis* 15(2), 415–447.

Vehtari, A. and J. Ojanen (2012). A survey of Bayesian predictive methods for model assessment, selection and comparison. *Statistics Surveys* 6(none), 142 – 228.

Warton, D. I., F. G. Blanchet, R. B. O’hara, O. Ovaskainen, S. Taskinen, S. C. Walker, and F. K. Hui (2015). So many variables: Joint modelling in community ecology. *Trends in Ecology & Evolution* 30, 766–779.
**Supplementary material**

Figure S1: DAG of the stacked single species distribution models: C+BB (with \( f_{i,j} = \beta_j \)) and IGP+BB. The model does not contain species interaction in any form.

Figure S2: DAG of joint species distribution models which include interspecific competition but not include interspecific correlations between site preferences: C+DM and IGP+DM.
Figure S3: The directed acyclic graph (DAG) representation of a joint species distribution model which includes interspecific correlation between site preferences of species but does not include interspecific competition: LMC\( (k) + \text{BB} \). This is the DAG that corresponds to many state-of-the-art JSDMs such as the Hierarchical model of Species Communities (Ovaskainen and Abrego, 2020) and the Additive Multivariate GP model (Vanhatalo et al., 2020).

Figure S4: Randomized PIT\(_2\) and PIT\(_4\) histograms for the LMC(2)\(_S\)+DM model separately for sphagnums, vascular plants and combined vegetation.
Figure S5: Species specific randomized PIT$_1$ histograms for the LMC(1)$_{NS}$+DM model.
Figure S6: Species specific randomized PIT$_1$ histograms for the LMC(2)$_s$+DM model.
Figure S7: Species specific randomized PIT$_3$ histograms for the LMC(1)$_{NS+DM}$ model.
Figure S8: Species specific randomized PIT$_3$ histograms for the LMC(2)$_S$+DM model.
Figure S9: Posterior predictive distributions conditional on the full training data for the estimated combined sphagnum, vascular plants and total cover over the area for every model. Black dots show the posterior predictive means and blue squares represent the empirical average total percentage cover over the training data.

Figure S10: Posterior mean estimates of the interspecific correlations in the species niche preference in the LMC(1)_{NS}+BB model. White cells indicate that the 80% posterior credible interval of the correlation overlapped zero (i.e., weak support for interspecific correlation) and stars indicate that the 95% posterior credible interval did not overlap zero (i.e. strong support for interspecific correlation).
### Table S1: Species specific model comparison results with 10-fold cross-validation using log predictive density utility $CV_1$ together with the standard error estimates (se) and the Monte Carlo error estimate (me) for the CV estimates.

| Species                | 1) C+BB  | 2) C+DM  | 3a) IGP$_2$+BB | 3b) IGP$_{2d}$+BB |
|------------------------|----------|----------|----------------|------------------|
| $S. angustifolium$      | -1.46 (2e-1/5e-4) | -1.47 (1.6e-1/5e-4) | -1.47 (2e-1/5e-4) | -1.53 (2e-1/1e-3) |
| $S. balticum$          | -2.51 (2e-1/6e-4) | -2.51 (1.5e-1/5e-3) | -2.51 (1e-1/4e-4) | -2.52 (2e-1/8e-4) |
| $S. magellanicum$      | -3.03 (2e-1/7e-4) | -3.03 (1.5e-1/9e-4) | -3.03 (1e-1/3e-4) | -3.11 (2e-1/1e-3) |
| $S. majus$             | -0.91 (1e-1/3e-4) | -0.92 (1.4e-1/7e-4) | -0.92 (1e-1/5e-4) | -0.98 (2e-1/9e-4) |
| $S. papillosum$        | -4.10 (1e-1/8e-4) | -4.18 (1.7e-1/9e-3) | -4.18 (1e-1/9e-4) | -3.31 (2e-1/3e-3) |
| $S. palustris$         | -2.34 (2e-1/7e-4) | -2.34 (1.8e-1/7e-3) | -2.34 (2e-1/6e-4) | -2.37 (2e-1/2e-3) |
| $C. lasiocarpa$        | -0.87 (1e-1/3e-4) | -0.86 (1.1e-1/4e-4) | -0.87 (1e-1/3e-4) | -0.87 (1e-1/4e-4) |
| $C. limosa$            | -1.22 (1e-1/3e-4) | -1.22 (1.0e-1/4e-4) | -1.22 (1e-1/3e-4) | -1.22 (1e-1/3e-4) |
| $C. rostrata$          | -0.82 (1e-1/3e-4) | -0.81 (1.0e-1/3e-4) | -0.81 (1e-1/3e-4) | -0.81 (1e-1/3e-4) |
| $C. nigricans$         | -0.79 (1e-1/4e-4) | -0.79 (1.3e-1/6e-4) | -0.79 (1e-1/4e-4) | -0.80 (1e-1/4e-4) |
| $E. vaginatum$         | -2.86 (1e-1/4e-4) | -2.86 (1.2e-1/4e-4) | -2.86 (1e-1/5e-4) | -2.88 (1e-1/5e-4) |
| $P. alpestris$         | -0.54 (9e-2/3e-4) | -0.54 (8.8e-2/3e-4) | -0.54 (9e-2/3e-4) | -0.55 (9e-2/3e-4) |
| $R. chamaemorus$       | -0.59 (1e-1/4e-4) | -0.59 (1.1e-1/4e-4) | -0.59 (1e-1/4e-4) | -0.59 (1e-1/4e-4) |
| $S. palustris$         | -1.40 (1e-1/4e-4) | -1.40 (1.1e-1/4e-4) | -1.40 (1e-1/4e-4) | -1.42 (1e-1/5e-4) |

| Species                | 5a) LMC(1)$_2$+BB  | 5b) LMC(1)$_{2d}$+BB  | 6a) LMC(1)$_2$+DM  | 6b) LMC(1)$_{2d}$+DM  |
|------------------------|----------------------|------------------------|----------------------|------------------------|
| $S. angustifolium$      | -1.47 (2e-1/4e-4) | -1.72 (1.9e-1/9e-3) | -1.50 (2e-1/1e-3) | -1.50 (2e-1/1e-3) |
| $S. balticum$          | -2.52 (1e-1/4e-4) | -3.05 (1.7e-1/9e-3) | -2.86 (2e-1/3e-3) | -2.86 (2e-1/3e-3) |
| $S. magellanicum$      | -3.03 (1e-1/4e-4) | -3.82 (1.9e-1/3e-3) | -3.36 (1e-1/3e-3) | -3.36 (1e-1/3e-3) |
| $S. majus$             | -0.91 (1e-1/3e-4) | -0.92 (1.4e-1/7e-4) | -0.98 (1e-1/4e-4) | -0.98 (1e-1/4e-4) |
| $S. papillosum$        | -4.16 (1e-1/7e-4) | -5.26 (1.8e-1/7e-3) | -4.30 (2e-1/3e-3) | -4.30 (2e-1/3e-3) |
| $S. palustris$         | -2.34 (2e-1/6e-4) | -3.36 (2.5e-1/3e-4) | -2.40 (7e-2/3e-4) | -2.40 (7e-2/4e-4) |
| $C. lasiocarpa$        | -0.87 (1e-1/3e-4) | -1.16 (1.4e-1/7e-4) | -0.87 (1e-1/3e-4) | -0.87 (1e-1/3e-4) |
| $C. limosa$            | -1.23 (1e-1/3e-4) | -1.56 (1.2e-1/5e-4) | -1.23 (1e-1/3e-4) | -1.23 (1e-1/3e-4) |
| $C. rostrata$          | -0.81 (1e-1/3e-4) | -0.96 (1.2e-1/4e-4) | -0.82 (1e-1/3e-4) | -0.82 (1e-1/3e-4) |
| $C. nigricans$         | -0.79 (1e-1/4e-4) | -0.95 (1.5e-1/6e-4) | -0.80 (1e-1/4e-4) | -0.80 (1e-1/4e-4) |
| $E. vaginatum$         | -2.86 (1e-1/4e-4) | -3.31 (1.3e-1/8e-4) | -2.86 (1e-1/4e-4) | -2.86 (1e-1/4e-4) |
| $P. alpestris$         | -0.54 (9e-2/3e-4) | -0.59 (9e-2/3e-4) | -0.55 (9e-2/3e-4) | -0.55 (9e-2/3e-4) |
| $R. chamaemorus$       | -0.59 (1e-1/4e-4) | -0.71 (1.3e-1/4e-4) | -0.60 (1e-1/3e-4) | -0.60 (1e-1/3e-4) |
| $S. palustris$         | -1.40 (1e-1/4e-4) | -1.96 (1.5e-1/9e-4) | -1.42 (1e-1/4e-4) | -1.42 (1e-1/4e-4) |
Table S2: Species specific model comparison results with 10-fold cross-validation using log predictive density utility \(CV_3\) together with the standard error estimates (se) and the Monte Carlo error estimate (me) for the CV estimates.

| Species          | 1) C+BB | 2) C+DM | 3a) IGP +BB | 3b) IGP NS +BB |
|------------------|---------|---------|-------------|----------------|
| S. angustifolium | -30.0 (4/3e-3) | -30.0 (3/2e-3) | -29.4 (4/2e-2) | -29.2 (3/2e-2) |
| S. balticum      | -54.0 (3/2e-3) | -54.0 (3/2e-3) | -50.1 (4/3e-2) | -49.9 (3/3e-2) |
| S. magellanicum  | -64.4 (4/3e-3) | -64.4 (4/3e-3) | -60.9 (4/3e-2) | -61.0 (4/2e-2) |
| S. majus         | -20.3 (2/2e-3) | -20.2 (2/1e-3) | -18.7 (3/2e-2) | -18.8 (2/1e-2) |
| S. papillosum    | -82.1 (3/2e-3) | -82.1 (3/2e-3) | -75.1 (4/3e-2) | -76.2 (4/2e-2) |
| S. fallax        | -49.3 (4/3e-3) | -49.7 (4/3e-3) | -46.4 (4/3e-2) | -46.2 (4/2e-2) |
| C. lasiocarpa    | -21.1 (2/2e-3) | -21.1 (2/2e-3) | -17.4 (2/8e-3) | -17.5 (2/6e-3) |
| C. limosa        | -26.7 (1/2e-3) | -26.7 (1/2e-3) | -24.5 (1/9e-3) | -24.7 (1/1e-3) |
| C. rostrata      | -18.0 (3/2e-3) | -18.0 (3/2e-3) | -16.3 (2/8e-3) | -16.4 (2/7e-3) |
| E. nigrum        | -16.6 (3/3e-3) | -16.6 (3/3e-3) | -15.9 (3/2e-2) | -16.0 (3/2e-2) |
| E. vaginatum     | -56.9 (2/2e-2) | -56.9 (2/2e-2) | -54.4 (2/1e-2) | -54.5 (2/1e-2) |
| P. sylvestris    | -10.8 (2/6e-3) | -10.8 (2/6e-3) | -10.3 (2/7e-3) | -10.4 (2/6e-3) |
| R. chamaemorus   | -11.6 (3/3e-2) | -11.8 (3/2e-2) | -11.3 (3/2e-2) | -11.4 (3/2e-2) |
| S. palustris     | -33.8 (2/2e-3) | -33.8 (2/2e-3) | -28.2 (2/1e-2) | -28.2 (2/1e-2) |

| Species          | 5a) LMC(1) + BB | 5b) LMC(1) + BB | 6a) LMC(1) + DM | 6b) LMC(1) + DM |
|------------------|-----------------|-----------------|-----------------|-----------------|
| S. angustifolium | -29.1 (3/1e-2)  | -28.8 (3/2e-2)  | -29.2 (3/2e-2)  | -30.5 (4/6e-2)  |
| S. balticum      | -50.1 (3/2e-2)  | -50.3 (3/1e-2)  | -50.0 (3/2e-2)  | -50.5 (4/2e-2)  |
| S. magellanicum  | -60.6 (4/2e-2)  | -61.6 (4/3e-2)  | -60.6 (4/2e-2)  | -62.3 (4/2e-2)  |
| S. majus         | -18.3 (2/1e-2)  | -18.8 (3/3e-2)  | -18.5 (3/2e-2)  | -19.7 (3/5e-2)  |
| S. papillosum    | -81.3 (3/5e-2)  | -94.1 (5/3e-2)  | -83.1 (3/5e-2)  | -86.3 (4/3e-1)  |
| S. fallax        | -46.5 (4/3e-2)  | -47.3 (4/3e-2)  | -46.7 (4/3e-2)  | -48.1 (5/3e-2)  |
| C. lasiocarpa    | -17.4 (2/9e-3)  | -17.2 (2/1e-2)  | -17.3 (2/9e-3)  | -17.2 (2/1e-2)  |
| C. limosa        | -24.3 (1/2e-2)  | -24.3 (1/2e-2)  | -24.4 (1/9e-3)  | -24.4 (1/2e-2)  |
| C. rostrata      | -16.3 (2/7e-3)  | -16.2 (2/8e-3)  | -16.3 (2/7e-3)  | -16.3 (2/9e-3)  |
| E. nigrum        | -15.7 (3/1e-2)  | -15.9 (3/4e-2)  | -15.7 (3/1e-2)  | -16.0 (3/2e-2)  |
| E. vaginatum     | -56.9 (2/2e-2)  | -57.5 (2/2e-2)  | -56.9 (2/2e-2)  | -57.3 (2/2e-2)  |
| P. sylvestris    | -10.8 (2/6e-3)  | -10.8 (2/6e-3)  | -10.8 (2/7e-3)  | -10.9 (2/7e-3)  |
| R. chamaemorus   | -11.6 (3/1e-2)  | -11.8 (3/1e-2)  | -11.6 (3/1e-2)  | -11.7 (3/1e-2)  |
| S. palustris     | -27.9 (2/1e-2)  | -28.1 (2/1e-2)  | -27.9 (2/1e-2)  | -28.3 (2/2e-2)  |

| Species          | 7a) LMC(2) + DM | 7b) LMC(2) + DM | 4a) IGP + DM |
|------------------|-----------------|-----------------|-------------|
| S. angustifolium | -29.3 (3/1e-2)  | -32.1 (4/6e-2)  | -30.5 (4/5e-2) |
| S. balticum      | -50.3 (3/1e-2)  | -58.0 (4/7e-2)  | -52.0 (3/6e-2) |
| S. magellanicum  | -60.6 (4/1e-2)  | -72.0 (4/1e-2)  | -68.2 (4/1e-2) |
| S. majus         | -18.5 (3/1e-2)  | -26.8 (3/2e-1)  | -18.7 (3/2e-2) |
| S. papillosum    | -82.7 (3/5e-2)  | -97.8 (3/5e-2)  | -93.2 (3/1e-1) |
| S. fallax        | -45.7 (4/3e-2)  | -61.5 (5/2e-2)  | -47.9 (4/1e-1) |
| C. lasiocarpa    | -17.2 (2/7e-3)  | -22.7 (3/4e-2)  | -17.4 (2/8e-3) |
| C. limosa        | -24.5 (1/8e-3)  | -29.6 (8/1-2e-2) | -24.5 (1/9e-3) |
| C. rostrata      | -16.2 (2/6e-3)  | -19.2 (2/1e-2)  | -16.3 (2/8e-3) |
| E. nigrum        | -15.6 (3/2e-2)  | -18.7 (3/3e-2)  | -15.9 (3/2e-2) |
| E. vaginatum     | -57.1 (2/1e-2)  | -65.2 (2/8e-2)  | -57.1 (2/1e-2) |
| P. sylvestris    | -10.7 (2/6e-3)  | -11.8 (2/9e-3)  | -10.9 (2/7e-3) |
| R. chamaemorus   | -11.7 (3/2e-2)  | -14.2 (3/2e-2)  | -11.8 (3/1e-2) |
| S. palustris     | -29.0 (2/1e-2)  | -38.5 (2/1e-1)  | -28.2 (2/1e-2) |