This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.

Author(s): Parvez Rana & Anne Tolvanen

Title: Transferability of 34 red-listed peatland plant species models across boreal vegetation zone

Year: 2021

Version: Publisher's version
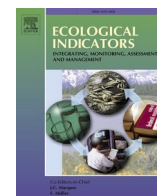
Copyright: The author(s) 2021

Rights: CC BY

Rights url: https://creativecommons.org/licenses/by/4.0/

Please cite the original version:

Parvez Rana & Anne Tolvanen. Transferability of 34 red-listed peatland plant species models across boreal vegetation zone. Ecological indicators 2021, 129, 107950. https://doi.org/10.1016/j.ecolind.2021.107950

Transferability of 34 red-listed peatland plant species models across boreal vegetation zone

Parvez Rana [*], Anne Tolvanen

Natural Resources Institute Finland (Luke), Paavo Havaksen tie 3, 90570 Oulu, Finland

ABSTRACT

Red-listed species are negatively affected by habitat degradation and fragmentation. They usually have small populations and may be affected by local extinctions which makes species distribution modeling (SDM) challenging. Although SDM has been used extensively for biodiversity protection and regional land-use planning, the transferability of SDM between regions is still at the developmental level. We show a first attempt to demonstrate model transferability for red-listed plant species over boreal regions. We modeled the distribution of 34 red-listed boreal peatland plant species at national and regional levels, using multiple streams of environmental data in Finland. The objectives were: (1) to evaluate how environmental characteristics explaining species distribution differ between three regions covering five vegetation zones (subarctic, northern, middle, southern boreal, and hemi-boreal vegetation zones); (2) to assess the performance of one national and three regional species distribution models (SDM: northern, middle, and southern regions); and (3) to test whether the regional models can be transferred to other regions and discuss alternative methods to improve transferability. The maximum entropy (maxent) algorithm was employed to predict suitable habitats for the assessed species. An SDM performance was measured with the area under the receiver operative characteristics (AUC), true skill statistics (TSS), and the continuous Boyce index (CBI). Three conclusions are relevant. First, the environmental variables explaining species distribution differed significantly ($p < 0.05$) between the three regions. Second, the internal measure of accuracy measured as cross-validation of AUC, TSS, and the CBI was quite similar in both the national and regional models, which indicates that realistic species distribution maps could be generated from all models. Last, the external measure of accuracy (i.e. transferability) in the regional models was lower than the internal measure of accuracy, which indicates that a good regional model could not automatically ensure good performance when transferred to another region. To improve the transferability of the regional models, we suggest the normalizing of environmental variable values. The data-driven evaluation of red-listed plant species provides an approach that can be used in biodiversity and nature conservation.

1. Introduction

The biodiversity conservation goal has gained increasing consideration globally due to the significant loss of floral and faunal diversity (IPBES, 2018). The Convention on Biological Diversity (CBD) estimates that 30% of the world's land needs to be protected in the next decade to halt the decline and extinction of global biodiversity (CBD, 2020). Degradation and fragmentation can negatively affect the distribution of red-listed species (Hanski, 2015). For example, in boreal Finland, a growing trend of the number of red-listed species (all taxa included) was evidenced, with an increase of 1,020 species between 2010 and 2019 (Hyvärinen et al., 2019). Every ninth species was considered red-listed,

the major risk being the degradation of suitable habitats caused by disturbance, construction, forest management activities, the reduction of old-growth forests, and climate change (Hyvärinen et al., 2019). To halt the biodiversity loss, international targets are being updated (CBD, 2020) or have just been set (i.e. European Union (EU) Biodiversity Strategy for 2030, European Commission, 2020) to protect significant amounts of the land and the sea area, or to restore biodiverse areas with high ecosystem services potential (e.g. biomass). The identification of potential nature protection sites and prediction of suitable habitats for red-listed species are core elements of future biodiversity protection and ecosystem restoration.

Peatlands (wetlands with peat soil) harbor a rich and unique range of

species and habitats worldwide (e.g. Yu, 2012). They are a highly space-effective reservoir of carbon and provide a wide set of other ecosystem services such as recreation, water management, timber, energy, and horticultural peat (e.g. Zedler and Kercher, 2005; Byg et al., 2017; Heikkinen et al., 2018). Globally, the area of peatland cover has been reduced to 10–20% (Cris et al., 2014), and 15% of peatlands have been drained, which is the main cause of change in peatland biodiversity (Harpenslager et al., 2015). Peatlands cover about 30% (approx. 10 million hectares) of the land area (Turunen, 2008) of Finland. A large percentage of forestry drainage (more than 50%, Finnish Forest Research Institute, 2014), conversion to agricultural uses, and extraction of peat and minerals are the main causes of peatland habitat degradation (Hyvärinen et al., 2019). Using predictive models, it was estimated that the 15% ecosystem restoration target set in the previous CBD and EU Biodiversity Strategy to 2020 targets would increase the predicted distribution of 71% of red-listed peatland plant species, whereas the potentially occupied area of species would increase only by 9% (Tolvanen et al., 2020). The models were developed for the aapa mire region that covers part of the northern and middle boreal regions (Saarimaa et al., 2019; Tolvanen et al., 2020). A new methodology is therefore needed that is transferable to the vegetation zones (i.e. subarctic, northern, middle, southern, and hemi-boreal zones) where boreal red-listed peatland plant species are found. Because of the large longitudinal dimension from the 60°N hemi-boreal vegetation zone to the 70°N subarctic zone, the diversity of peatland characteristics is large in Finland (Seppä, 2002). However, the environmental characteristics (e.g. temperature, rainfall) and the diversity of peatland site types reflect the richness of peatland vegetation in all vegetation zones (Aapala et al., 1998).

Species distribution modeling (SDM, also known as habitat suitability modeling) is widely used in biodiversity protection and regional land-use planning. The basic principle of SDM relies on correlating species occurrences with environmental variables, and employing geostatistical models or hypothetically inferred response curves (e.g. Araújo et al., 2019). Environmental data and recent progress in methods for SDM (e.g. Guisan and Thuiller, 2005; Zhang et al., 2018; Arenas-Castro et al., 2020; Li et al., 2020) provide opportunities to estimate species occurrences on a large spatial scale. Despite the wide use of SDMs, their usefulness as conservation tools has been doubted due to their tendency to have a larger internal measure of accuracy (also known as interpolation) and a lower external measure of accuracy (also known as transferability or extrapolation) (Datta et al., 2020; Heikkinen et al., 2012; Huang et al., 2018; Low et al., 2021; Manzoor et al., 2018; Qiao et al., 2019). These characteristics can constrain the reliable application of SDMs beyond surveyed regions. In addition, existing biodiversity models still struggle with data limitation, model transferability limitation (i.e. the extrapolation or generalization of the models), and complex structure of biological systems (Araújo et al., 2019). Earlier studies concentrated only on a small area within one vegetation zone (Parviainen et al., 2009, 2008, 2013) or on Aapa mires that cover parts of northern and middle boreal zones in Finnish boreal peatlands (Saarimaa et al., 2019; Tolvanen et al., 2020). In addition, the generalization of the models over other areas or areas covering multiple vegetation zones was missing the previous studies. Model transferability into new areas or other time frames is one of the challenges of SDM (Datta et al., 2020; Low et al., 2021; Qiao et al., 2019). Using SDM in an extrapolative manner (transferability or external measure of accuracy) is a complex task compared to the simple filling of gaps in SDM within partly sampled areas (i.e. interpolative estimation or internal measure of estimation) (Heikkinen et al., 2012; Huang et al., 2018; Manzoor et al., 2018). The trained SDM may fail to represent distinct environmental characteristics in the unsurveyed areas. Some recent SDMs use deep learning algorithms (e.g., convolutional neural networks) to better address the complexity of ecological niches (Deneu et al., 2021).

SDMs for predicting red-listed plant species distribution are becoming increasingly important to assess their extinction risk and the degree of protection (Marcer et al., 2013; McCune, 2016; Parviainen et al., 2008, 2013; Saarimaa et al., 2019; Thapa et al., 2020). In addition, SDM-derived hotspot maps can be used in tandem with spatial conservation prioritization to protected area network mapping (Marcer et al., 2013; McCune, 2016; Parviainen et al., 2013). However, the scarce occurrence data continues to challenge the use of SDM for red-listed plant species (Marcer et al., 2013; McCune, 2016). In previous studies, researchers have focused on optimizing SDM and interpolation performance of SDM (e.g. Bellamy et al., 2013; McCune, 2016; Arenas-Castro et al., 2020), but not on the extrapolation performance (Datta et al., 2020; Heikkinen et al., 2012; Randin et al., 2006; Wenger and Olden, 2012). Transferability assessment of SDM in spatio-temporal scale provides crucial information for red-listed plant species (Heikkinen et al., 2012) as it helps to evaluate whether there is suitable habitat for the species outside their present distribution range. This information helps the future management and protection measures of the species. Transferability of an SDM may also be important in northern ecosystems where species distributions are predicted to change more rapidly than in regions further south.

In this study, we modeled the distribution of 34 red-listed boreal peatland plant species at national and regional levels. Our objectives were (1) to evaluate how environmental characteristics explaining species distribution differ between three regions covering five vegetation zones (subarctic, and northern, middle, southern boreal, and hemi-boreal vegetation zones); (2) to assess the performance of the national and three regional SDM (northern, middle, and southern regions); and (3) to test the transferability of regional models to other regions and discuss alternative methods to improve the transferability. We utilized nationwide species occurrence data observed for the last 20 years and included four categories of environmental variables (i.e. climate, topography, geology, and habitat characteristics) as explanatory variables in the models. We extended the earlier models (Saarimaa et al., 2019; Tolvanen et al., 2020) to cover the five vegetation zones in Finland in which the studied red-listed species have been encountered. We generalized the results over the regions using a maximum entropy (Maxent) algorithm. To our knowledge, this study was the first attempt to demonstrate model transferability for red-listed species distribution over boreal regions.
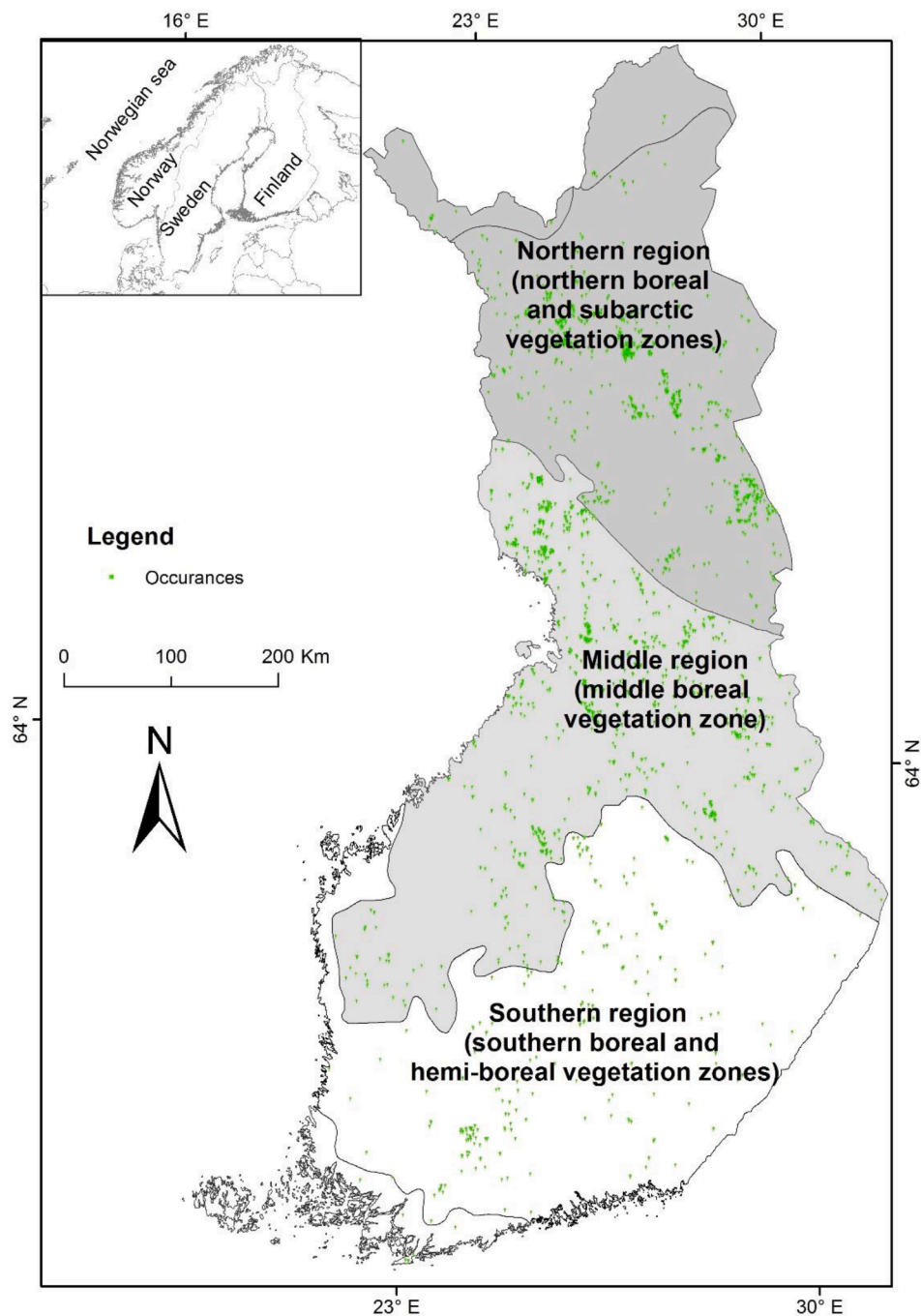
2. Materials and methods

2.1. Study area

The study area covers the whole of Finland (60°W, 70°N) in northern Europe (Fig. 1). The climate is mild, influenced by the Gulf Stream. The annual average temperature is 6 °C in the southwestern part of the country and −2 °C in the northeastern part of the country. Across the country, precipitation varies between 400 mm and 700 mm. The study area is mostly flat terrain, with an area of 75% forest cover (Lier et al., 2019).

There are two types of peatland in Finland: Aapa mires, which are located mostly in the middle and northern boreal vegetation regions; and raised bogs, located mainly in the southern boreal region. Aapa mires are fens characterized by sedges and moss peat, whereas raised bogs are composed of *Sphagnum* peat and *Eriophorum* peat.

The study area covered five vegetation zones (i.e. subarctic, and northern, middle, southern boreal, and hemi-boreal vegetation zones) (Ahti et al., 1968), which were merged into the northern region (subarctic and northern boreal vegetation zones), middle region (middle boreal vegetation zone), and southern region (southern boreal and hemi-boreal vegetation zones) (Fig. 1) to increase the species occurrences in each region. The study area was segmented into a spatial grid of 500 m by 500 m (25 ha) squares, and squares with more than 10% peatland cover were retained in the analysis. As a result, we had a total of 2,941,262 spatial grid squares (735,316 km^2) for analysis.

Fig. 1. Map of the study area, with red-listed plant species occurrence records. The northern region includes northern boreal and subarctic vegetation regions, the middle region includes the middle boreal vegetation region, and the southern region includes southern boreal and hemi-boreal vegetation regions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2. Red-listed plant species data

Peatlands are the primary habitat of 280 red-listed species, which is 4.5% of the total red-listed species in Finland (Hyvärinen et al., 2019). The national database was utilized to extract the presence-only occurrence records of red-listed peatland plant species (Hyvärinen et al., 2019). The term red-listed includes the IUCN categories critically endangered (CR), endangered (EN), vulnerable (VU), and near threatened (NT). We utilized the following standards for data extraction from the database: (*i*) the occurrences of those vascular plants and bryophytes which occur primarily on undrained mires; (*ii*) the newest occurrence records (i.e. from 2000 or later); (*iii*) only observations with accuracy better than 100 m; (*iv*) species with at least ten or more records of occurrence. Species absences are not included in the database. Finally, we had a total of 8,846 occurrence records of 34 red-listed plant species, which were used for model calibration and validation (Appendix 1).

The red-listed plant species (n = 34) were subdivided into four species groups according to their principal habitats (see also Saarimaa et al., 2019): spruce swamp forest species (SSF); mesotrophic fen species; rich fen species; and calcareous fen species. Spruce swamp forests are wooded minerotrophic nutrient-rich peatlands, characterized by Norway spruce (*Picea abies*), sparse deciduous trees, and abundant deadwood (Similä et al., 2014). Mesotrophic fens are characterized by thick peat layers formed from *Sphagnum* moss and other bryophytes, and

by sedges and herbaceous plants in the field layer. The main attributes of rich fens are large species diversity, neutral or mildly acidic soil, and an open or sparsely wooded landscape (Saarimaa et al., 2019). A large proportion of calcareous rock in the soil and high mineral concentration in groundwater are the main requirements for a calcareous fens habitat (Similä et al., 2014). The total number of species in each species group was 8, 5, 9, and 12 respectively for SSF, mesotrophic fen, rich fen, and calcareous fen species. The northern region had a total of 14 species, the middle region 21 species, and the southern region 13 species.

2.3. Environmental data

A total of 16 environmental variables within four categories (i.e. climate (n = 2), topography (n = 1), geology (n = 1), and habitat (n = 12)) were employed in the model building (Appendix 2). We ensured using Spearman's rank correlation (pair-wise absolute correlation cutoff = 0.70) that multicollinearity among the environmental variables was absent (e.g. Jiménez-Valverde and Lobo, 2007; McCune, 2016). In the climate category, growing degree days (GDD) and mean water balance were calculated for 1990–2013 from the climate data of the Finnish Meteorological Institute (Pirinen et al., 2012). Despite that the most recent data (2014–2019) were not available at the same resolution, the 23-year time period (1990–2013) covered a considerable amount of the years when species observations had been made. Hence the data could capture the GDD and mean water balance conditions of the species quite well. The GDD reflecting the extent of the growing season was calculated according to Equation (1).

$$GDD = \sum \max\{0, (T_m - T_t)\} \tag{1}$$

where T_m is the mean daily temperature, and T_t is a threshold temperature (5° C). Precipitation alone is insufficient to indicate the moisture available for plant growth; water balance was therefore calculated as the monthly difference between precipitation and potential evapotranspiration. The monthly water balance was summed per annum. The potential evapotranspiration (PET) was calculated according to Equation (2) (Skov and Svenning, 2004).

$$PET = \frac{58.93 \times T_{(above\,0°\,C)}}{12} \tag{2}$$

In the topography category, the topographic wetness index (TWI) (Beven and Kirkby, 1979) indicates the local relative differences in moisture conditions and was calculated using Equation (3).

$$TWI = \ln\left(\frac{\alpha}{tan\beta}\right), \tag{3}$$

where α is the local upslope area draining through a certain point per unit contour length, and $tan\beta$ is the local slope in radians. A high TWI value indicates a wet area.

In the geology category, the proportion of calcareous rock in each grid square was assessed by utilizing a digital map of the Quaternary deposit and pre-Quaternary rock database.

In the habitat category, we employed 12 environmental variables, which provide information on tree species' mean volume, site type, drainage status, and biodiversity index (BI) (Appendix 2). We derived eight variables from the multi-source national forest inventory (MS-NFI) data: the mean volume (m³/ha) of pine (*Pinus sylvestries*), spruce (*Picea abies*), birch (*Betula pendula* and *B. pubescens*), and other broadleaved trees, and the proportion of herb-rich *Vaccinium myrtillus*, *Vaccinium vitis-idaea*, and *Cladina* forest site types in each grid square (%). The proportion of undrained, drained, and open peatlands in each grid square was calculated using a drainage map by the Finnish Environment Institute (SYKE, 2009). The drainage map is based on the topographic database of the Finnish National Land Survey. Last but not least, we assessed a BI value developed for forests (Mikkonen et al., 2018) for each grid square. The BI value is normalized from 0 to 1, with high values

indicating high biodiversity. The index is formed from three main elements: 1: local decaying wood potential, number of large trees, and tree species diversity; 2: penalties for forestry operations (e.g. logging, clear-cutting); and 3: connectivity, based on ecological similarity, distance, and quality between forest stands. For a detailed description, see Mikkonen et al. (2018).

2.4. Statistical testing

Since our data were not normally distributed (confirmed by the Kolmogorov-Smirnov test, Marsaglia et al., 2003), we used non-parametric Wilcoxon signed rank to assess whether the environmental variable values differed between the three regions (Hollander and Wolfe, 1973). Testing was done in a pairwise manner: northern vs. middle region; northern vs. southern region; middle vs. southern region. We used Wilcoxon signed rank, which uses signed ranks and is an appropriate test in our study.

The Bhattacharyya distance was used as a measure of divergence (Davis, 1996; Fukunaga, 2013) to realize the necessity of data normalization for model transferability (Huang et al., 2018; Qiao et al., 2019). The advantage of the Bhattacharyya distance separability criterion is the fact that it is a simple criterion that can be easily implemented. The Bhattacharyya distance can be defined as follows: Let *i* and *j* be the two variable distributions that are usually modeled as a Gaussian distribution with mean μ and covariance $\sum$. The Bhattacharyya distance between two probability distributions *Pi* and *Pj* denoted by B_{ij} is defined by

$$B_{ij} = \frac{1}{8}(\mu_i - \mu_j)^T \left(\frac{\sum_i + \sum_j}{2}\right)^{-1}(\mu_i - \mu_j) + \frac{1}{2}\ln\left(\frac{|\Sigma_i + \Sigma_j|}{\sqrt{|\Sigma_i||\Sigma_j|}}\right), \tag{4}$$

where μ_i and μ_j refer to the mean of the distributions, $\sum_i$ and $\sum_j$ refers to the covariance matrix of the distributions. The Bhattacharyya distance of environmental variable distributions between two regions (e.g. northern vs. middle region) will be zero if there is a complete overlap (Davis, 1996; Fukunaga, 2013).

2.5. Model calibration and validation

The overall workflow, from the initial environmental input files and species occurrences to the delineated hotspot maps, is presented in Fig. 2.

We employed a maxent algorithm to predict the habitat suitability for the red-listed plant species. Maxent generates a continuous prediction of occurrences, ranging from 0 (least suitable) to 1 (best suitable) (e. g. Elith et al., 2011; Elith and Leathwick, 2009; Low et al., 2021; Merow et al., 2013; Phillips et al., 2006; Phillips and Dubik, 2008; Saarimaa et al., 2019). It has been shown effective in predicting the distributions of peatland plant species (McCune, 2016; Oke and Hager, 2017; Parviainen et al., 2013; Saarimaa et al., 2019; Tolvanen et al., 2020). We applied the default parameters, i.e. regularization multiplier = 1, auto-features, maximum iterations = 500, and convergence threshold = 10^{-5}, to generate reasonably simple models that did not over-fit the calibration data (Kramer-Schadt et al., 2013; Phillips and Dubik, 2008; Virzi et al., 2017). The logistic output was used to compare the prediction maps between species (Li et al., 2020; Saarimaa et al., 2019). We employed the same maxent setting and the same environmental variables for all species so that we could aggregate models for several species.

Since we had presence-only data, we used pseudo-absences in the models to help control the sampling bias, and avoid erroneous or false absences (Bellamy et al., 2013; Brown et al., 2017). Following the previous literature (e.g. Mateo et al., 2010; Barbet-Massin et al., 2012; Connor et al., 2018), we generated 10,000 pseudo-absence points, and we gave equal weight to the presences and pseudo-absence records (prevalence of 0.5) (Elith et al., 2011; Jiménez-Valverde and Lobo,
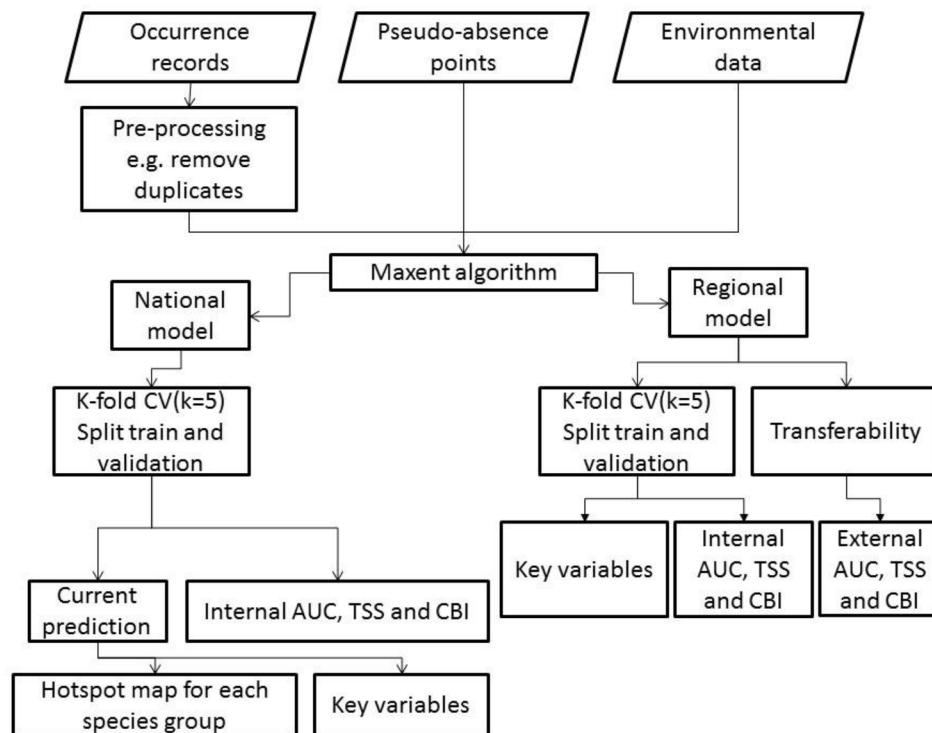
Fig. 2. Overall flowchart of the model building and validation.

2007; Liu et al., 2016; Tehrani et al., 2020). We sampled the pseudo-absence points with a buffered minimum-convex polygon (buffer radius = 100 km, labeled as bias grid) based on known occurrences for each species (Brown et al., 2017). Bias grids regulate the pseudo-absence point's allocation and sampling density.

We assessed the SDMs at the national and regional levels, i.e in the northern, middle and southern regions (Fig. 2). In the national model, we postulated that we could build species-level SDMs that covered all occurrences of the respective species in all three regions. We calculated k-fold cross-validation (k-fold CV) accuracy metrics (see Section 2.6 for details) and employed k = 5 with a random partition of 80% training and 20% validation (Saarimaa et al., 2019; Tehrani et al., 2020; Araújo et al., 2005).

In the regional models, we postulated that we could build species-level SDMs for each region and transfer them to other regions that were not covered by the training data. For example, we built a model using all occurrence records of a species (e.g. *Carex appropinquata*) found in the northern region and reported the internal measure of accuracy metrics. We then validated (i.e. transferability) it on the other region (i. e. middle and southern regions) and reported the external measure of accuracy metrics. Similarly, we repeated the above calibration and validation for the middle and southern regions.

2.6. Assessing model performance and key environmental variable

The area under the receiver operative characteristics (AUC) offers widely used statistics for assessing maxent model performance (Marcer et al., 2013; Parviainen et al., 2009, 2008, 2013; Saarimaa et al., 2019; Syfert et al., 2014; Zhang et al., 2018), although it has recently been criticized (Fernandes et al., 2019; Manzoor et al., 2018). True skill statistics (TSS, Allouche et al., 2006; Fernandes et al., 2019; Li et al., 2020; Tehrani et al., 2020), and the continuous Boyce index (CBI, Hirzel et al., 2006; Progin, 2018) were therefore considered the most reliable statistics for evaluating presence-only data (Zurell et al., 2020). CBI is also an appropriate metric to validate model transferability (Cianfrani et al., 2010; Manzoor et al., 2018). Zurell et al. (2020), however, reported a

standard protocol for reporting SDM performance where AUC, TSS and CBI were advised for evaluating presence-only data and this is why we decided to use these statistical metrics.

The AUC, TSS, and CBI values were calculated for each species and species group (see Section 2.2 for species group description). AUC value ranges between 0 and 1 values < 0.7 were considered *fair*, 0.7–0.8 *good*, 0.8–0.9 *very good*, and >0.9 *excellent* (Manzoor et al., 2018; Swets, 1988). The range of TSS is −1 to + 1, where values close to + 1 indicate that the model can distinguish between false negative and false positive occurrences. The CBI avoids using a confusion matrix and is threshold-independent (Hirzel et al., 2006; Low et al., 2021). The CBI ranges between −1 and + 1, where values close to + 1 indicate that the predicted model distribution is consistent with the distribution of presences in the evaluation dataset (Boyce et al., 2002; Hirzel et al., 2006).

We employed maxent permutation importance analysis to identify key environmental variables for each species, and we then summarized the permutation percentage values for each species group (Table 3). In assessing the contribution of each variable, maxent permutation importance analysis performs better than percentage contribution analysis, because permutation importance is path (algorithm) -independent (e.g. Saarimaa et al., 2019), whereas the percentage contribution is based on the order of the variable in a model. The contribution of each variable relies on randomly permuting the values of that variable among the training points (both presence and pseudo-absence), and the final values are normalized to percentages (Connor et al., 2018; Li et al., 2020; Saarimaa et al., 2019).

2.7. Hotspot maps

To produce hotspot maps for the species groups and for all species, we utilized the predicted distribution map of the national model, because the national model covered the entire study area (Figs. 6 and 7). In addition, we produced a hotspots map using finer grain data of 25 ha (0.25 km^2), because a criticism has been that coarse-grain data (e.g. 10 km^2) may overestimate the hotspots value (Di Marco et al., 2017; McKerrow et al., 2018). Maxent produced a continuous probability map,

which was reclassified into a binary map for simplicity, presenting unsuitable (0) and suitable (1) areas. Threshold values for maximum sensitivity and specificity (Max SS) were used for the reclassification. The Max SS threshold sets the limit at which the sum of sensitivity (true positive rate) and specificity (true negative rate) is highest (Liu et al., 2013). The Max SS also reduces the mean error rate and minimizes the risk of selecting inappropriate areas by identifying only those areas with the maximum suitability (Saarimaa et al., 2019).

We stacked the individual species-level suitability values for the species groups and created a hotspot map for each species group. The top 5% of grid squares ranked by stacked suitability value were presented in the map (e.g. Williams et al., 1996; Orme et al., 2005). To produce a hotspot map for all red-listed species (n = 34), the summing of suitability values was done across all species, and 5% of the grid squares ranked by stacked suitability value were presented.

2.8. Data processing and analysis environment

Data processing, mapping and statistical analysis were done using R statistical programming (libraries, i.e. raster, dismo, SDMtune, ecospat, sf, rJava) (R Core Team, 2020), Python programming (libraries, i.e. arcpy, pandas, numpy), and ArcGIS version 10.6.1 (ESRI, 2018).

3. Results

3.1. Environmental variables in the three regions

The distribution of 13 out of 16 environmental variables was statistically different ($p < 0.05$) between all three regions (Fig. 3, Fig. 4, Table 1, Appendix 3). For example, temperatures (GDD) were coldest in the northern region and warmest in the southern region (Appendix 3), which was confirmed by the large Bhattacharyya distance (3.34) between the northern and southern regions (Table 1) and by the probability density function where northern and southern regions were not overlapping (Fig. 3). TWI was less varied in the three modeling regions which was visible in the Bhattacharyya distance (Table 1) and probability density function (Fig. 3). The proportion of calcareous rock was largest in the northern region, whereas the middle vs. the southern region showed no difference (Bhattacharyya distance 0.0). Average pine, spruce, and birch tree volumes (m^3/ha) were largest in the southern region (Appendix 3), which was confirmed by the large Bhattacharyya distance between the northern and southern regions (Table 1). Tree volumes distribution differences were also visible in the probability density function, for instance, the northern region had a spike at zero percent of their pine tree volume distribution (Fig. 3). The proportion of the *Vaccinium myrtillus* site type was largest in the southern region (Appendix 3), and the Bhattacharyya distance was largest (0.37) between the middle and southern regions (Table 1). The percentage of undrained peatlands was highest in the northern region and lowest in
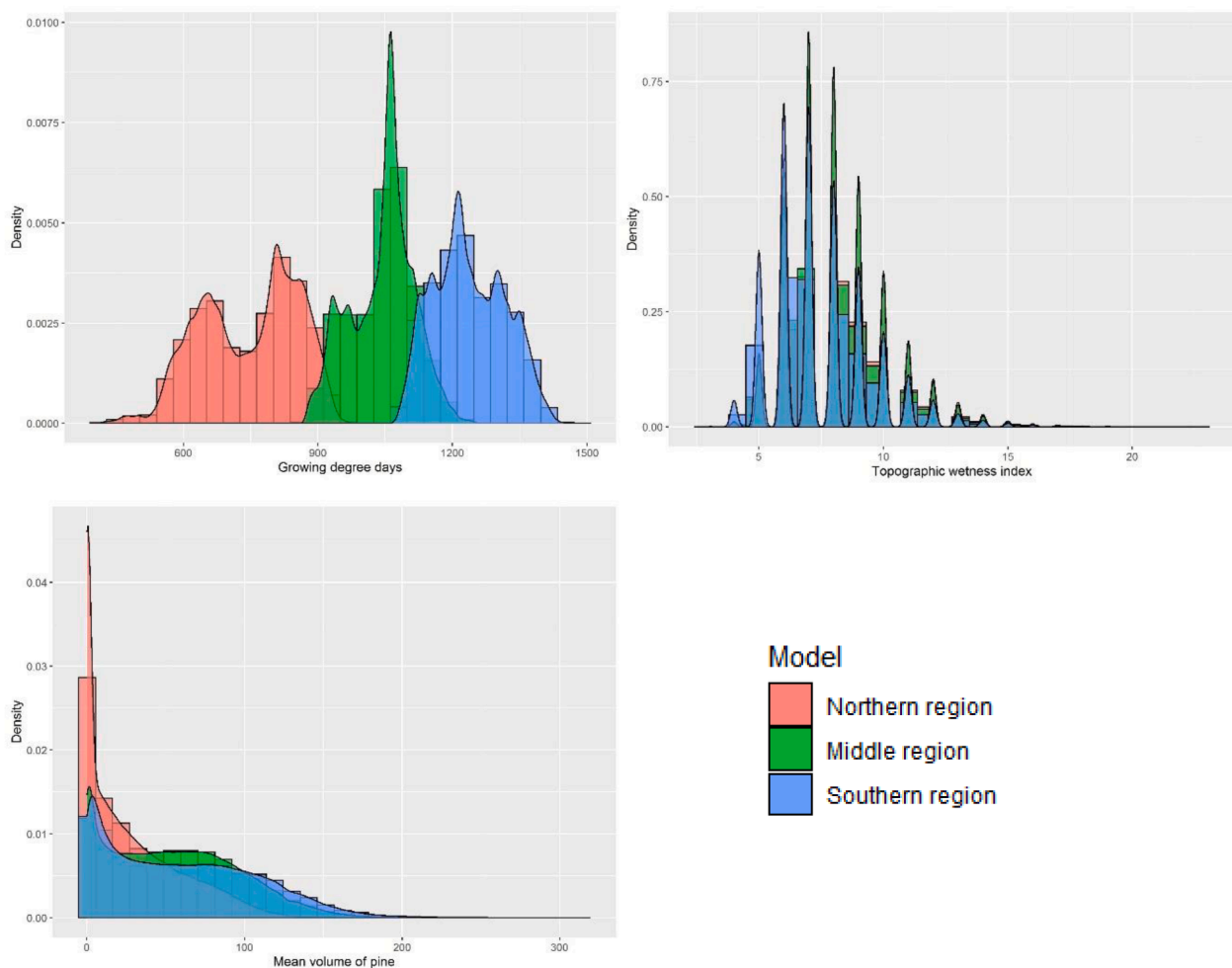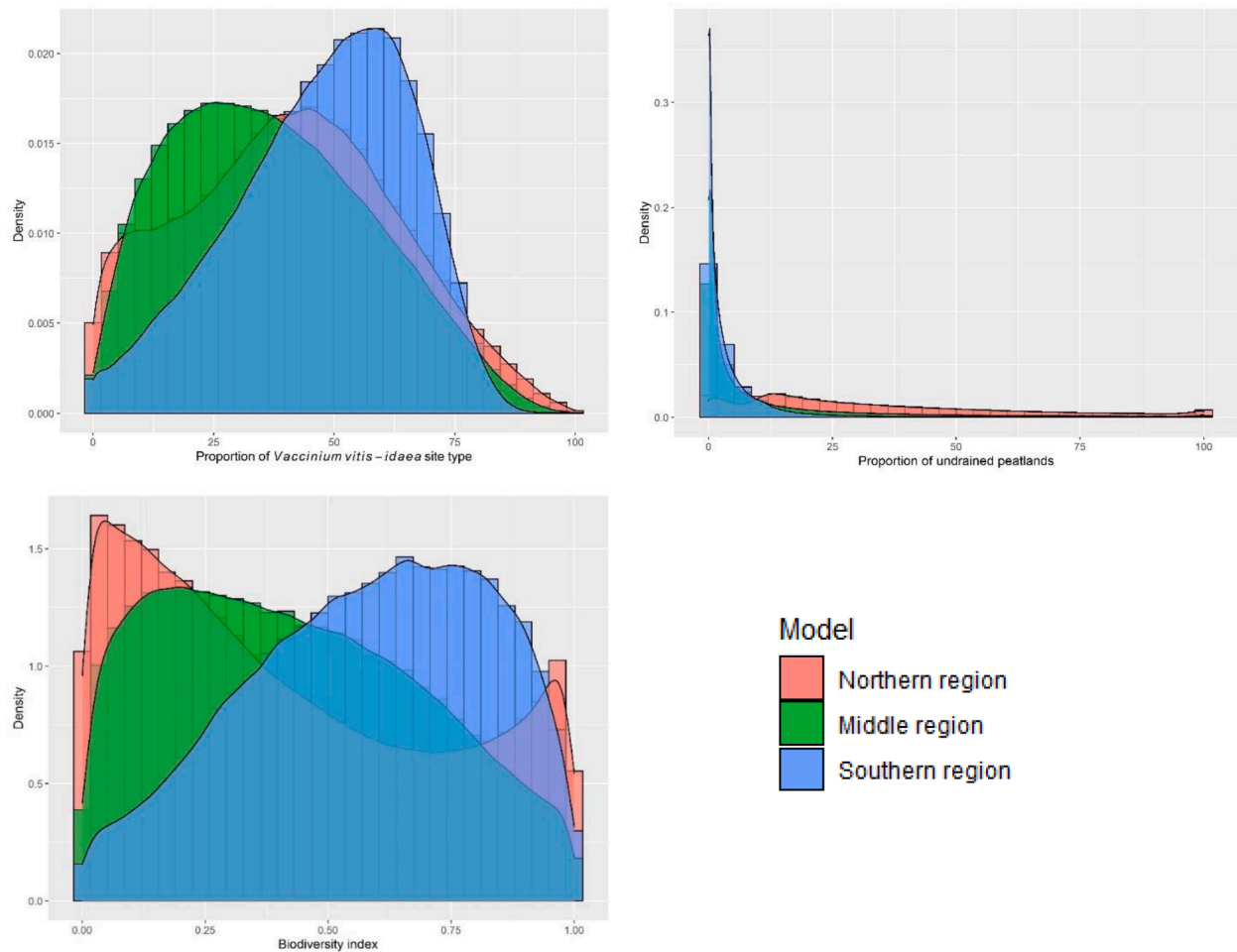

Fig. 3. Growing degree days, topographic wetness index, and mean volume of pine in the three regions. Wilcoxon signed-rank test confirmed that the distributions were statistically significantly different ($p < 0.05$) in the three regions for model building.

Fig. 4. Proportion of *Vaccinium vitis-idaea* site type, proportion of undrained peatlands, and biodiversity index in the three regions. Wilcoxon signed-rank test confirmed that the distributions were statistically significantly different (p < 0.05) in the three regions for model building.

Table 1

Bhattacharyya distances of environmental variables between the three regions. The distance between the two regions is zero if there is complete overlap.

Environmental variable	Northern vs. middle region	Northern vs. southern region	Middle vs. southern region
GDD	1.36	3.34	0.84
WAB	0.04	0.07	0.09
TWI	0.00	0.02	0.01
Rock	0.04	0.02	0.00
Pine	0.09	0.12	0.01
Spruce	0.20	0.53	0.13
Birch	0.12	0.25	0.03
Broadleaf other	0.27	0.65	0.16
Site type 1	0.01	0.02	0.06
Site type 2	0.06	0.17	0.37
Site type 3	0.00	0.02	0.04
Site type 4	0.08	0.02	0.04
Undrained	0.14	0.43	0.10
Drained	0.16	0.06	0.04
Open	0.54	0.43	0.02
BI	0.01	0.05	0.05

Abbreviations are explained in Appendix 3.

the southern region (Appendix 3), which was confirmed by the large Bhattacharyya distance (0.43) between the northern and southern regions (Table 1) and by the probability density function where northern and southern regions were less overlapping (Fig. 4).
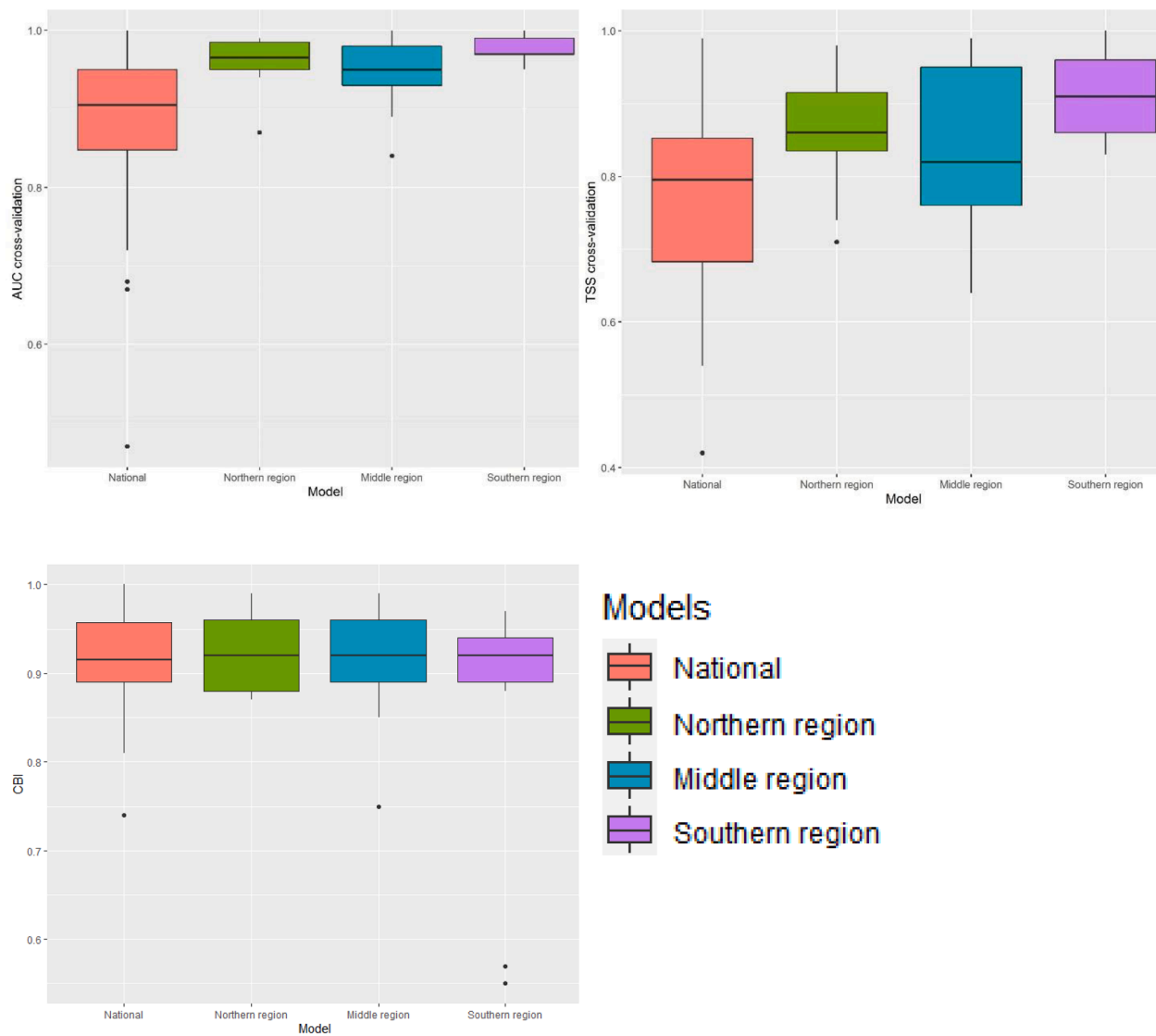
3.2. National models

The national model performed well for all 34 plant species, which was shown by all three accuracy metrics. The mean AUCs were *excellent* (average 0.97, range 0.89–1) and *very good* (average 0.88, range 0.47–1) for training and cross-validation respectively (Fig. 5, Appendix 4). Only for two species (*Plagiothecium latebricola, Sphagnum subnitens*) were the mean cross-validation AUC values <0.70 (*fair*). The mean TSSs values were 0.87 (range 0.62–1) and 0.76 (range 0.42–0.99) for training and cross-validation respectively (Appendix 4, Fig. 5). The mean CBI was 0.91 (range 0.74–1).

Concerning the species groups, the highest mean (±SD) AUC value of 0.91 (±0.09) at validation was recorded for the calcareous species, whereas the lowest value of 0.81 (±0.15) was recorded for SSF species (Table 2). The TSS value showed a similar trend with AUC values. The CBI was highest for the mesotrophic fen species and lowest for the SSF species.

GDD, the proportion of undrained peatlands, and the BI were the three most important variables in the national models (Table 3). The GDD had the most relative importance (%) for the mesotrophic fen species (52.3%), for SSF (30%), and for calcareous species (21.1%), whereas the proportion of undrained peatland areas had the most relative importance (36.2%) for the rich fen species. The BI had the most relative importance (19.4%) for the SSF species.

Fig. 5. Internal measure of accuracy (i.e. cross-validation AUC, cross-validation TSS, and CBI) for validation data in national and regional models (i.e. northern, middle and southern regions). Boxplots visualize five summary statistics (the median, two hinges (the 25th and 75th percentiles) and two whiskers: the upper whisker extends from the hinge to the largest value; the lower whisker extends from the hinge to the smallest value).
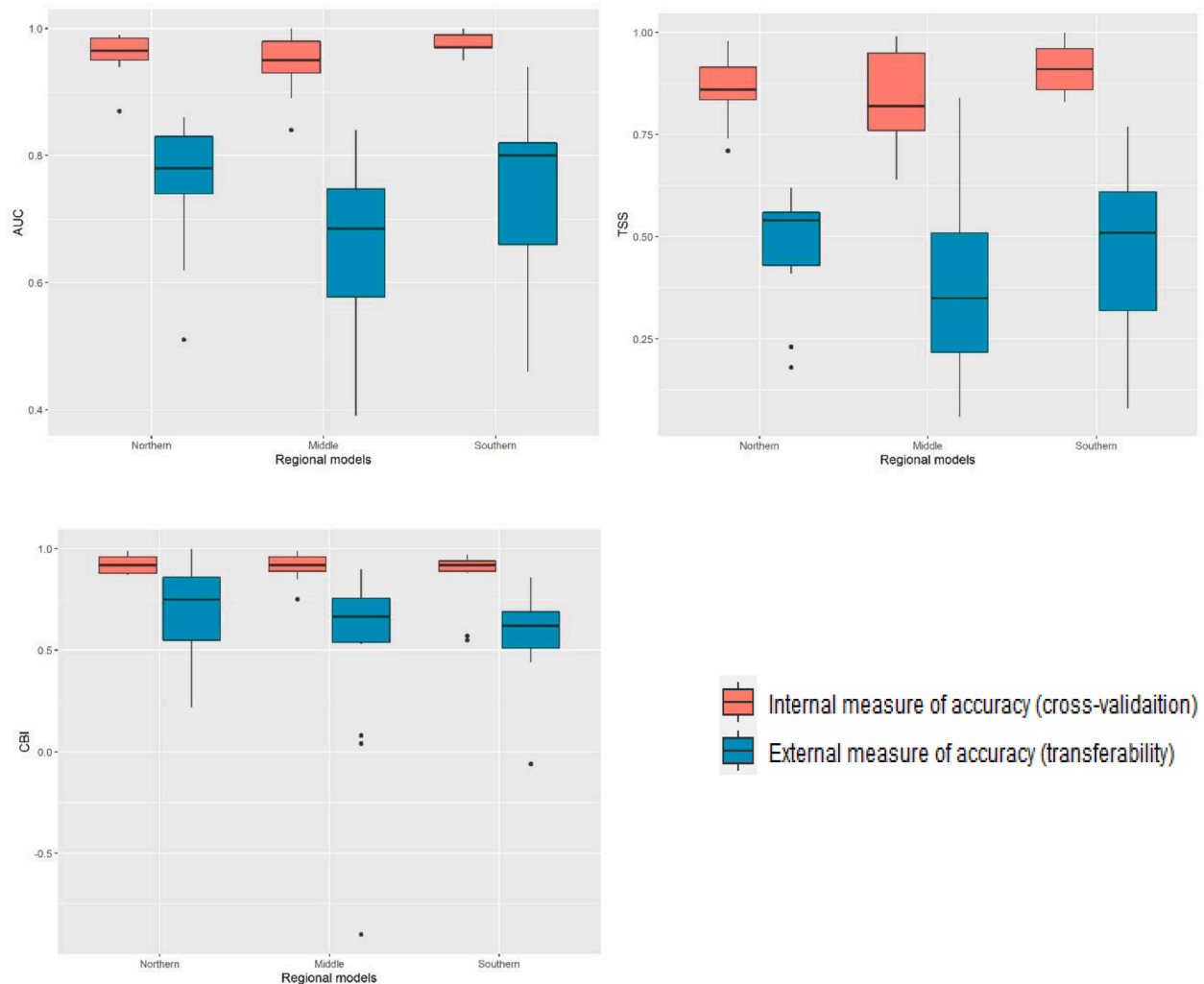
3.3. Regional models

The regional models performed well, which was shown by all three-accuracy metrics (Fig. 5, Appendix 5, 7, 9). For example, all the SDMs had a cross-validation AUC value that was higher than 0.84, which indicates the consistent excellent performance of regional models. Similar results were observed with the CBI values. The median value of cross-validated AUC and CBI values were close to each other in three regional models (Fig. 5). However, the hinges (the 25th and 75th percentiles) and whiskers (upper and lower whisker extends) of cross-validated TSS value were higher in the middle region compared to the northern and southern regions (Fig. 5).

The external measure of accuracy (i.e. transferability) was lower than the internal measure (Fig. 6), which indicates that regional models provided more accurate prediction in internal estimation (filing the gaps within the region) than external estimation (predictions outside the region). This suggests that the transferability of regional models into new regions is challenging. The highest mean $\pm$ SD transferability of AUC, TSS, and CBI was found in the northern region (0.76 $\pm$ 0.01, 0.47 $\pm$ 0.02, 0.69 $\pm$ 0.16), whereas the lowest transferability AUC was found in the middle region (0.67 $\pm$ 0.03, 0.38 $\pm$ 0.03, 0.50 $\pm$ 0.18) (Appendix

6, 8, 10). Furthermore, 2 (out of 14, northern region), 10 (out of 21, middle region), and 6 (out of 13, southern region) species SDMs had an external validation AUC value of <0.7 (*fair*). In the northern region, the most important variables were GDD, the proportion of undrained peatland, and drained peatlands respectively (Appendix 11). GDD, the proportion of undrained peatland, and BI were the three most important variables respectively in the middle and southern regions (Appendix 12, 13).

3.4. Hotspot maps

Suitable habitats (i.e. based on the stacked suitability value) for the calcareous species group were mostly predicted in the northern (65% of all suitable grid squares of calcareous species) and middle (28%) regions (Fig. 7), whereas suitable habitats for mesotrophic fen species were predicted especially in the middle region (82% of suitable grid squares) (Fig. 7). The rich fen species had very suitable habitats in both the northern (54%) and middle regions (42%) (Fig. 7), whereas it was predicted that the majority (84%) of suitable habitats for SSF species were in the southern region (Fig. 7). The all-species hotspots were scattered throughout the country, although they were mostly located in

Fig. 6. Internal (cross-validation) and external (transferability) measures of accuracy (i.e. AUC, TSS, and CBI) of regional models (i.e. northern, middle and southern regions). Boxplots visualizes five summary statistics (the median, two hinges (the 25th and 75th percentiles) and two whiskers: the upper whisker extends from the hinge to the largest value; the lower whisker extends from the hinge to the smallest value).

the middle (48%) and northern regions (36%) (Fig. 8).

4. Discussion

This study shows that although the national and regional SDM performed very well, the transferability of the regional models to other regions was not as good. The reason appears to be the very different environmental characteristics between the regions, which limits the distributions of individual species. In addition, the statistical challenges distressed the transferability of the regional model as it forces the algorithm to make predictions for outside of training environmental conditions, which may often be erroneous. To our knowledge, this was the first study to test whether the SDM of red-listed species could be transferred across boreal regions. We discuss the benefits, limitations, and implications of our method below.

4.1. Benefits and limitations of SDM

Our model was capable of producing a promising accuracy, and it was quite stable at both training and cross-validation stages, indicating that realistic species distribution maps can be generated from both national and regional modeling. Saarimaa et al. (2019) studied threatened plant species in aapa mires in Finland, recording a mean cross-validation AUC of 86%, which is close to our study (88%). Besides, Parviainen et al.

(2013) utilized Landsat images to estimate the SDM of 28 red-listed plant species in northeastern Finland, mentioning an accuracy of 87%. However, it must be noted that apart from a larger study area covering the whole country, we used a different database from earlier studies (Parviainen et al., 2009, 2008, 2013; Saarimaa et al., 2019; Tolvanen et al., 2020) due to the newest red-list assessment available for this study (Hyvärinen et al., 2019).

SDM can provide a useful means to predict species occurrences in unsurveyed areas (e.g. Datta et al., 2020; Heikkinen et al., 2012; Huang et al., 2018; Regos et al., 2019). However, the transferability of regional models is the key to a successful application of SDM, required in ecological and biodiversity conservation (Datta et al., 2020; Duque-Lazo et al., 2016; Huang et al., 2018; Qiao et al., 2019; Regos et al., 2019; Torres et al., 2015). However, the trained SDM from an area may not represent the distinct characteristics well (e.g. habitat structure and climate condition) in the unsurveyed areas (Datta et al., 2020; Huang et al., 2018; Low et al., 2021; Manzoor et al., 2018; Torres et al., 2015). The transferability of a regional model in our study performed similarly to each other. However, the external measure of accuracy (i.e. transferability) was lower than the internal measure of accuracy in all three regions. A good regional model thus does not automatically ensure equally good performance when they are transferred to a new case study area.

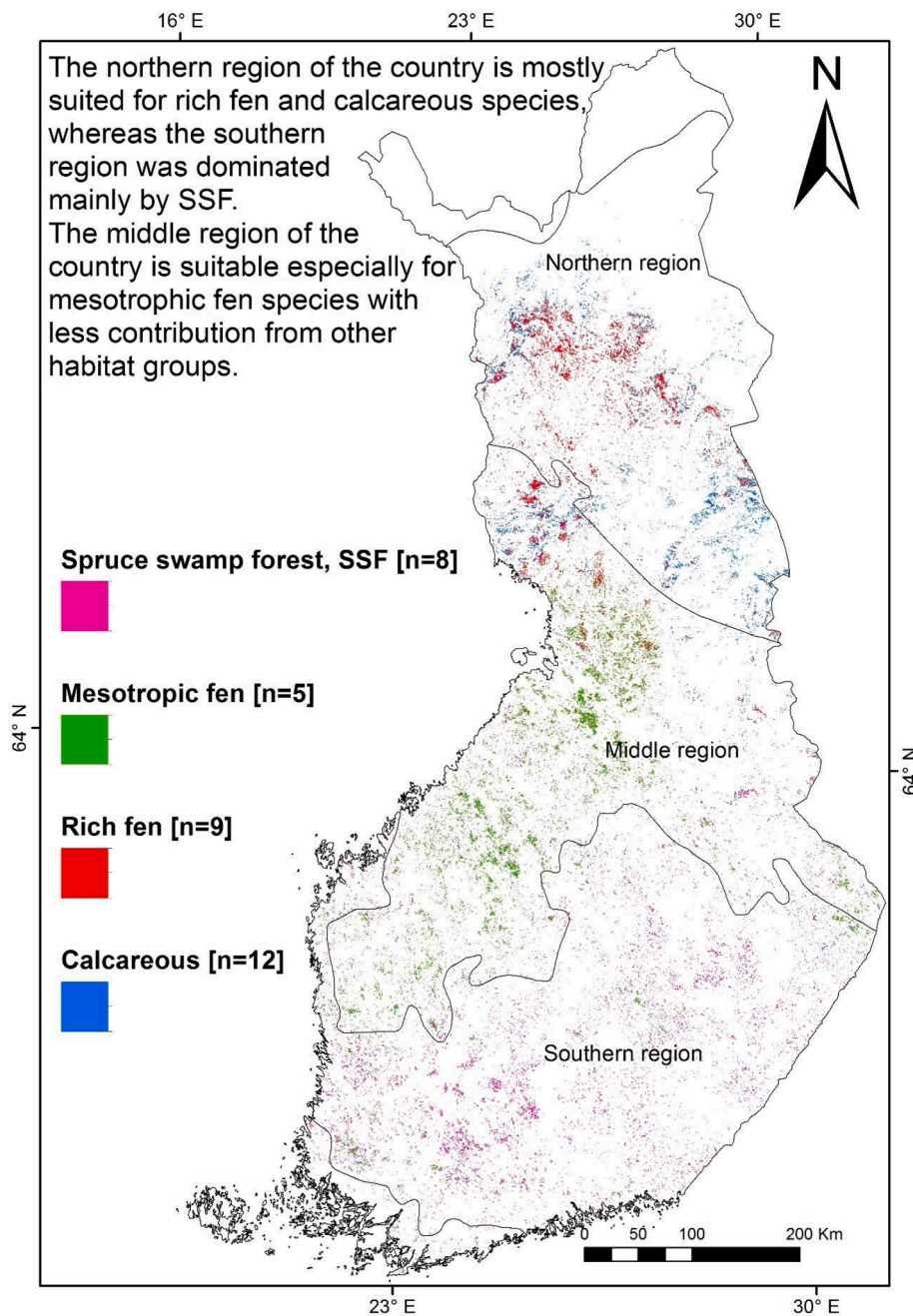Heikkinen et al. (2012) studied the model transferability of vascular

Fig. 7. Hotspots for each species group were based on the national model. The top 5% of grid squares ranked by stacked suitability value were presented in the map.

Table 2
Measure of accuracy for training and validation data for each species group summarized from Appendix 4, national model.

Species group	AUC training mean ± SD	AUC validation mean ± SD	TSS training mean ± SD	TSS validation mean ± SD	Internal CBI mean ± SD
Calcareous	0.97 ± 0.03	0.91 ± 0.09	0.88 ± 0.12	0.81 ± 0.12	0.91 ± 0.05
Mesotrophic fen	0.96 ± 0.03	0.90 ± 0.06	0.80 ± 0.10	0.73 ± 0.13	0.95 ± 0.03
Rich fen	0.96 ± 0.03	0.88 ± 0.11	0.83 ± 0.10	0.73 ± 0.12	0.94 ± 0.04
SSF	0.98 ± 0.01	0.82 ± 0.15	0.92 ± 0.08	0.75 ± 0.15	0.87 ± 0.07

plants in Finland, recording a mean transferability AUC value of 0.70, which resembled our estimation accuracy (0.73). However, our study modeling design is distinct from Heikkinen et al. (2012) in that we assessed the transferability of each regional model based on vegetation zone, whereas Heikkinen et al. (2012) combined occurrence records from all the vegetation zones. By mingling our study findings and the results of Heikkinen et al. (2012), we can see that the model transferability of plant species is a challenge, irrespective of modeling design.

The transferability of regional models provides several benefits, which have also been observed in other studies (e.g. Randin et al., 2006; Heikkinen et al., 2012; Wenger and Olden, 2012; Datta et al., 2020). The temporal scale can be extended by generating new species distribution maps and hotspot maps when new environmental data is available over the same area, and the spatial scale can be expanded by applying the models to new areas. Assessing at a temporal scale would be a

Table 3
Variable contribution (%) summarized for each species group in the national model.

Geospatial variable	SSF %	Mesotrophic fen %	Rich fen %	Calcareous %	Mean %
GDD	**33.01**	**52.26**	**21.24**	**21.1**	**31.90**
WAB	0.90	**4.83**	3.60	2.43	2.94
TWI	1.65	0.39	0.80	4.30	1.79
Rock	0.81	0.33	2.93	**12.48**	4.14
Pine	7.43	0.57	1.55	4.44	3.50
Spruce	1.42	1.35	2.25	1.07	1.52
Birch	1.49	1.19	1.24	0.88	1.20
Broadleaf other	4.51	0.93	1.07	2.06	2.14
Site type 1	1.18	0.96	2.29	6.54	2.74
Site type 2	1.60	2.65	**11.41**	4.92	5.15
Site type 3	**8.03**	1.42	2.86	3.15	3.87
Site type 4	4.87	4.31	4.74	**12.52**	6.61
Undrained	5.84	**25.36**	**36.21**	11.27	**19.67**
Drained	7.61	2.08	2.30	1.79	3.45
Open	0.25	0.06	0.07	0.17	0.14
BI	**19.38**	1.31	5.44	10.88	**9.25**
Total	100	100	100	100	100

*Three most important variables for each species group and on average are in bold. Abbreviations are explained in Appendix 3.
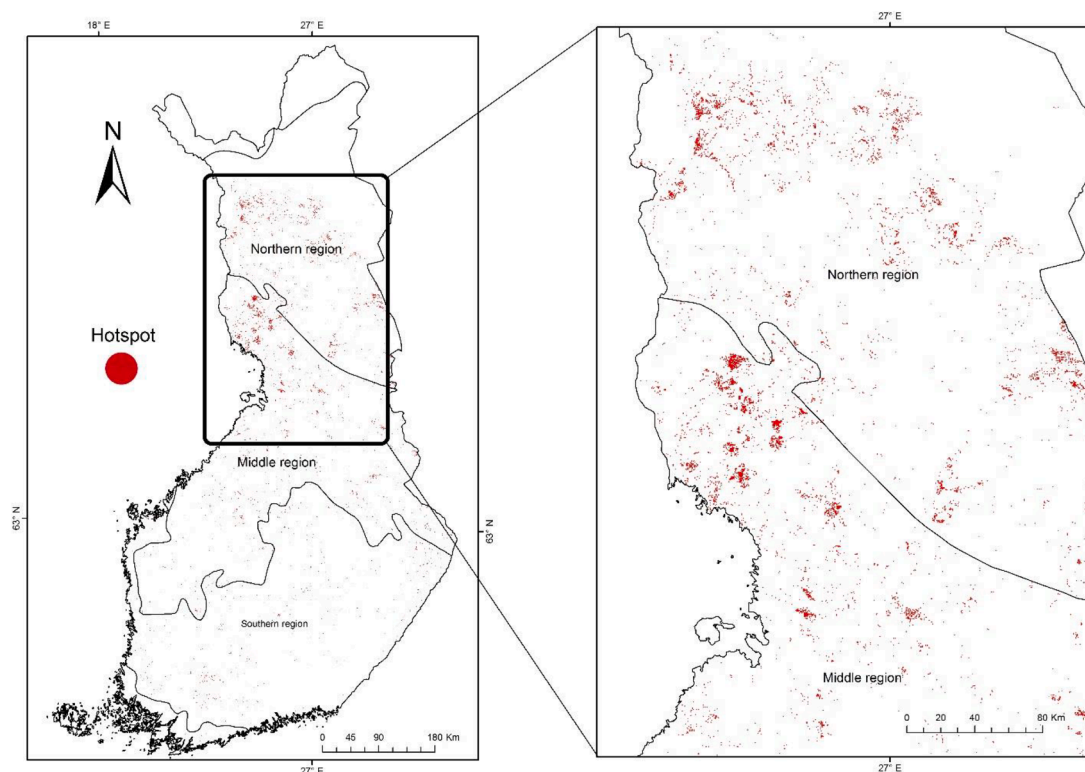
motivating methodological link for future investigation.

To improve the transferability of the regional model, one possibility is to normalize the environmental variable values, which may differ considerably between regions. Non-linear response functions between species and environmental data may affect the transferability (Heikkinen et al., 2012). Possible normalizing methods could be histogram matching, median-based adjustment, and regression-based adjustment. Histogram matching is a widely used technique for the normalization of variable values employing the cumulative probability function (Gonzalez and Woods, 2017). For median-based adjustment, the median value per environmental variable is used to normalize the value between the reference area variable value (training dataset) and the dependent area variable value (validation dataset) (Kappal, 2019; Välikangas et al., 2018). For regression-based adjustment, the variable normalization is based on simple linear regression, using the median value of variable per plant species. We suggest that future studies could test possible variable normalization methods that might improve the transferability of the regional model.

The appropriate selection of pseudo-absence points is a crucial parameter in SDM, because it greatly influences the model's accuracy and uncertainty (Alvarez-martinez et al., 2015; Benito et al., 2017; Hengl et al., 2009; Iturbide et al., 2015; Li et al., 2011; Norris et al., 2011; Ranc et al., 2017; Senay et al., 2013; Wisz and Guisan, 2009). In this study, we randomly sampled the pseudo-absence points with a buffered minimum-convex polygon (bias grid), based on known occurrences for each species separately with a prevalence of 0.5. The bias grid helps to avert the sampling of habitat located outside a species' known occurrences (Bellamy et al., 2013; Brown et al., 2017; Lentini et al., 2018; Martin et al., 2020). However, to improve the model performance, alternative pseudo-absence point selection methods, e.g. random sampling with environmental constraint (Iturbide et al., 2015; Lentini et al., 2018), random sampling with a combination of geographical and environmental constraints (Iturbide et al., 2015; Senay et al., 2013), target group sampling (Andrew and Fox, 2020; Ranc et al., 2017), random sampling with k-means clustering (Iturbide et al., 2015; Senay et al., 2013), and random sampling with a self-organizing map (Kohonen, 2001), could be tested in future studies.

SDM was also affected by other parameters, including threshold selection during model development. An incorrect threshold can lead to an inferior SDM, with a false identification of suitable areas and hotspot maps (McKerrow et al., 2018; Trotta-Moreu and Lobo, 2010). However, threshold selection for each species during model building is neither an easy nor a straightforward task. In our study, we employed the maximum sensitivity and specificity threshold, which is quite practical, because it minimizes the risk of over prediction (Liu et al., 2013, 2016).

Fig. 8. A hotspot map of all red-listed plant species (n = 34). The top 5% of grid squares ranked by stacked suitability value were presented in the map. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Threshold selection can also be affected by the use of presence-only data, and even a small margin of the threshold can lead to poorer distribution maps (Jiménez-Valverde and Lobo, 2007; Liu et al., 2013, 2016).

We used maxent algorithm for SDM which was extensively used in the literature (e.g. Elith et al., 2011; Elith and Leathwick, 2009; Low et al., 2021; Merow et al., 2013; Phillips et al., 2006; Phillips and Dubik, 2008; Saarimaa et al., 2019). However, other algorithms for assessing the transferability of SDM can also be tested in the future, including deep learning algorithm, i.e., convolutional neural networks (CNN), radial basis function networks (RBFN), multilayer perceptrons (MP), long-short term memory (LSTM), Kohonen network or self-organizing map (SOM) (Botella et al., 2018; Deneu et al., 2021; Reisenbüchler et al., 2021; Shamshirband et al., 2019; Zhou, 2020). The possible benefit of using a deep learning algorithm is to tackle the statistical challenges which concerned the transferability of the regional model.

SDM was adopted to evaluate how environmental characteristics explaining species distribution differ between three regions in our study. However, the use of process-based model or biophysical model (Kearney and Porter, 2009), and geographically weighted regression (Páez et al., 2011) might be a good alternative for future studies. Higgins et al. (2020) compared that process-based model may be more appropriate for making projection outside of the training data whereas, SDM has superior ability to describe the data in the training area.

4.2. Environmental variables influencing the distribution of the species groups

Our models show that most hotspots are located in the middle and northern regions of the country. The situation is partly connected with the greater land-use pressure and low protection status of southern Finland, where only 5% of forest lands (including peatlands) are protected, compared to 80% in the north (LUKE, 2019). Furthermore, the unprotected old-growth forest area is <2% in the total forest area in the southern region (Tikkanen et al., 2006). Nevertheless, the greater coverage of hotspot areas in the middle and northern regions also reflects the different habitat requirements between the species groups and the availability of the preferred habitats. The northern region was most suitable for rich fen and calcareous species, whereas the southern region was dominated by SSF species. For example, calcareous habitats occur naturally in the northern part of the country. The middle region of the country was especially suitable for mesotrophic fen species.

Our results support previous observations on the importance of GDD and undrained peatland variables for red-listed species in the aapa mire region (Parviainen et al., 2013; Saarimaa et al., 2019), confirming that these variables also explain their distribution in a larger geographical area. Although the choice of variables in the national and regional models was similar, the contribution of each variable varied between the models. GDD seems the most useful variable, because it provides crucial information about the plants' growth conditions. The mean GDD differed between habitat groups, which indicates the specific GDD preferences in these groups.

Undrained peatlands are characterized by high TWI, diverse soil fertility, diverse micro-topography, and high peat layers (Saarimaa et al., 2019; Sallinen et al., 2019). These characteristics favor rich fen and mesotrophic fen species. Undrained mires, and hence rich fen and mesotrophic fen species, are mostly located in the northern and middle regions of Finland, whereas drained peatlands are located in the southern region due to the denser human population, and more intensive forest management and agriculture. In the southern region, peatlands are also smaller (e.g. < 2 ha). They are confined to separate basins, which makes them more exposed to disturbance than undrained mires (Sallinen et al., 2019). Since rich fen and mesotrophic fen species require undrained conditions, their habitat suitability decreases after drainage due to the drawdown of the water table. The choice of BI in the top list was unsurprising, because the BI includes forest-related variables, e.g. decaying wood potential, forest connectivity, and connectivity with the key woodland habitat, which are either directly or indirectly linked with the distribution of red-listed species, especially SSF.

We used four categories (climate, topography, geology, and habitat) of environmental variables in this study, as our studied 34 red-listed plant species are well correlated with the 16 environmental variables in these four categories (e.g. Parviainen et al., 2008, 2009, 2013; Saarimaa et al., 2019; Tolvanen et al., 2020). New studies shown that snow persistence, freezing degree days, thawing degree days, and permafrost occurrence may have an impact on the distribution of arctic plant species (Niittynen et al., 2020; Rissanen et al., 2021). However, the impact of the above variables on red-listed plant species need to be assessed in the future studies for better clarification.

The correlations between species occurrence points and environmental variable values are not always apparent (e.g. Araújo et al., 2019). SDM tends to overestimate the number of suitable habitats (i.e. false-positive occurrences error) (e.g. Svenning et al., 2008; Syfert et al., 2014; Araújo et al., 2019). To overcome this potential error, historical phylogeographic factors, anthropogenic effects (e.g. deforestation), ecological competitions, historical disturbance, geographical pattern, and non-equilibrium species dynamics should also be included in the models (Guisan and Thuiller, 2005; Svenning et al., 2008; Syfert et al., 2014). However, it is worth to note that few of the variables were not available in our analysis such as historical disturbance, ecological competition.

4.3. Implications for biodiversity conservation

Our model results can be utilized to locate potential new areas for red-listed species and to reassess their red-listed status (Syfert et al., 2014; McKerrow et al., 2018), which serves the data-driven evaluation of conservation status endorsed by the IUCN, for example. The models also provide information that helps to interpret the ecology of red-listed plant species (Di Marco et al., 2017). The maps provide objective evidence for biodiversity conservation without the probably subjective bias in maps where the habitat suitabilities are developed using expert opinion (Fourcade et al., 2013; Syfert et al., 2014).

The hotspot maps were produced by summing the habitat suitability maps of individual species, which may overestimate hotspot values, because the maps do not take into account constraints such as competition and the dispersal limits of species (Bellamy et al., 2013; Tehrani et al., 2020). Furthermore, red-listed plant species usually have small populations and may be affected by local extinctions. They may therefore be missing in some areas, despite the suitable environmental variables they contain (Syfert et al., 2014).

The hotspot maps of red-listed plant species offer a potential tool for measuring the overall red-listed plant diversity for the entire country. The approach has been generally proven efficient in conservation biology (McKerrow et al., 2018; Orme et al., 2005; Tehrani et al., 2020). Our models can also be used to extract cold spot maps. These are the reverse of hotspots, in that they reveal areas unsuitable for red-listed species due to the undesirable condition of crucial characteristics such as drainage or large human settlements. Cold spot maps generally show a large number of habitats, and like hotspots, cold spots need to be monitored to evaluate whether conservation or restoration efforts would improve their suitability for red-listed species.

One way to use the prediction models is to include the historical land use and management impacts to help evaluate whether the area of hotspots in the southern region would be larger than the present if the land use were not as heavy as it is now. A related approach was used in Tolvanen et al. (2020), in which the drainage level was theoretically altered between 0% and the present level to observe the impact on the distribution of red-listed species. Nevertheless, habitat variables other than drainage remained unaltered in their assessment. Hence, a question remains concerning the total impact of land use on the historic distribution and hotspots of red-listed species until the present. Finding answers to this difficult question would help evaluate the real need for the

conservation and restoration of habitats for red-listed species.

5. Conclusions

This study demonstrated the SDM that incorporates national and regional modeling approaches. Three significant conclusions are relevant. First, the distribution of environmental variable values differed between the three regions. Second, both national and regional models had similar performances, which indicates that realistic species distribution maps can be generated. Last, the transferability (an external measure of accuracy) of the regional models (northern, middle, and southern regions) was relatively low, which indicates that a good regional model does not automatically ensure equally good performance when they are transferred to a new case study area. Our study corroborates that transferring SDMs for red-listed plants into unsurveyed areas is difficult compared to the internal measure of estimation. To improve the transferability of the regional models, we suggest the normalizing of environmental variable values. Although we used the maxent algorithm for statistical model building, other algorithms might be interesting, including deep learning algorithms. Our future research work will lie in this direction.

CRediT authorship contribution statement

Parvez Rana: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **Anne Tolvanen:** Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2021.107950.

References

Aapala, K., Heikkilä, R., Lindholm, T., 1998. Protecting the diversity of Finnish mires. In: Vasander, H. (Ed.), Mires of Finland. Finnish Peatland Society, Helsinki, pp. 45–57.

Ahti, T., Hämet-Ahti, L., Jalas, J., 1968. Vegetation zones and their sections in northwestern Europe. Ann. Bot. Fenn. 5, 169–211.

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43, 1223–1232. https://doi.org/10.1111/j.1365-2664.2006.01214.x.

Alvarez-martinez, J.M., Suarez-Seoane, S., Palac, C., Sanz, J., Alonso, J.C., 2015. Can Eltonian processes explain species distributions at large scale? A case study. Divers. Distrib. 21, 123–138. https://doi.org/10.1111/ddi.12256.

Andrew, M.E., Fox, E., 2020. Modelling species distributions in dynamic landscapes : the importance of the temporal dimension. J. Biogeogr. 47, 1510–1529. https://doi.org/10.1111/jbi.13832.

Araújo, M.B., Pearson, R.G., Thuiller, W., Erhard, M., 2005. Validation of species–climate impact models under climate change. Global Change Biol. 11(9), 1504–1513. https://doi.org/10.1111/j.1365-2486.2005.01000.x.

Araújo, M.B., Anderson, R.P., Barbosa, A.M., Beale, C.M., Dormann, C.F., Early, R., Garcia, R.A., Guisan, A., Maiorano, L., Naimi, B., O'Hara, R.B., Zimmermann, N.E., Rahbek, C., 2019. Standards for distribution models in biodiversity assessments. Sci. Adv. 5, 1–12. https://doi.org/10.1126/sciadv.aat4858.

Arenas-Castro, S., Gonçalves, J.F., Moreno, M., Villar, R., 2020. Projected climate changes are expected to decrease the suitability and production of olive varieties in southern Spain. Sci. Total Environ. 709, 136161. https://doi.org/10.1016/j.scitotenv.2019.136161.

Barbet-Massin, M., Jiguet, F., Albert, C.H., Thuiller, W., 2012. Selecting pseudo-absences for species distribution models: how, where and how many? Methods Ecol. Evol. 3, 327–338. https://doi.org/10.1111/j.2041-210X.2011.00172.x.

Bellamy, C., Scott, C., Altringham, J., 2013. Multiscale, presence-only habitat suitability models: fine-resolution maps for eight bat species. J. Appl. Ecol. 50 (4), 892–901. https://doi.org/10.1111/1365-2664.12117.

Benito, B.M., Svenning, J., Kellberg-nielsen, T., Riede, F., Gil-romera, G., Mailund, T., Kjaergaard, P.C., Sandel, B.S., 2017. The ecological niche and distribution of Neanderthals during the Last Interglacial. J. Biogeogr. 51–61 https://doi.org/10.1111/jbi.12845.

Beven, K.J., Kirkby, M.J., 1979. A physically based, variable contributing area model of basin hydrology. Hydrol. Sci. Bull. 24, 43–69. https://doi.org/10.1080/02626667909491834.

Botella, C., Joly, A., Bonnet, P., Monestiez, P., Botella, C., Joly, A., Bonnet, P., Monestiez, P., Munoz, F., Distribution, S., et al., 2018. A deep learning approach to species distribution modelling. Multimed. Tools Appl. Environ. 169–199.

Boyce, M.S., Vernier, P.R., Nielsen, S.E., Schmiegelow, F.K.A., 2002. Evaluating resource selection functions. Ecol. Modell. 157, 281–300.

Brown, J.L., Bennett, J.R., French, C.M., 2017. SDMtoolbox 2.0: The next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. PeerJ 5, e4095.

Byg, A., Martin-Ortega, J., Glenk, K., Novo, P., 2017. Conservation in the face of ambivalent public perceptions – The case of peatlands as 'the good, the bad and the ugly'. Biol. Conserv. 206, 181–189. https://doi.org/10.1016/j.biocon.2016.12.022.

CBD, 2020. Zero Draft of post-2020 biodiversity framework, Secretariat of the Convention on Biological Diversity, CBD/WG2020/2/3. Kunming, China.

Cianfrani, C., Lay, L., Hirzel, A.H., Loy, A., 2010. Do habitat suitability models reliably predict the recovery areas of threatened species? J. Appl. Ecol. 47, 421–430.

Connor, T., Hull, V., Viña, A., Shortridge, A., Tang, Y., Zhang, J., Wang, F., 2018. Effects of grain size and niche breadth on species distribution modeling. Ecography 41, 1270–1282. https://doi.org/10.1111/ecog.03416.

Cris, R., Buckmaster, S., Bain, C., Reed, M. (Eds.), 2014. Global Peatland Restoration demonstrating SUCCESS. IUCN UK National Committee Peatland Programme, Edinburgh.

Datta, A., Schweiger, O., Kühn, I., 2020. Origin of climatic data can determine the transferability of species distribution models. NeoBiota 59, 61–76. https://doi.org/10.3897/neobiota.59.36299.

Davis, J.C., 1996. Introduction to statistical pattern recognition. Comput. Geosci. 7 (22), 833–834.

Deneu, B., Servajean, M., Bonnet, P., Botella, C., Munoz, F., Joly, A., 2021. Convolutional neural networks improve species distribution modelling by capturing the spatial structure of the environment. PLoS Comput. Biol. 17, 1–21. https://doi.org/10.1371/journal.pcbi.1008856.

Di Marco, M., Watson, J.E.M., Possingham, H.P., Venter, O., 2017. Limitations and trade-offs in the use of species distribution maps for protected area planning. J. Appl. Ecol. 54, 402–411. https://doi.org/10.1111/1365-2664.12771.

Duque-Lazo, J., Van Gils, H.A., Groen, T.A., Navarro-Cerrillo, R.M., 2016. Transferability of species distribution models: the case of *Phytophthora cinnamomi* in Southwest Spain and Southwest Australia. Ecol. Model. 24, 62–70.

Elith, J., Leathwick, J.R., 2009. Species distribution models: ecological explanation and prediction across space and time. Annu. Rev. Ecol. Evol. Syst. 40, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists. Divers. Distrib. 17, 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x.

ESRI, 2018. ArcGIS Desktop: Release 10. Environmental Systems Research Institute, Redlands, CA.

European commision, 2020. EU Biodiversity Strategy for 2030, Bringing nature back into our lives. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52020DC0380.

Fernandes, R.F., Scherrer, D., Guisan, A., 2019. Effects of simulated observation errors on the performance of species distribution models. Divers. Distrib. 25 (3), 400–413. https://doi.org/10.1111/ddi.12868.

Finnish Forest Research Institute, 2014. Finnish statistical yearbook of forestry 2014. Vantaa.

Fourcade, Y., Engler, J.O., Besnard, A.G., Rödder, D., Secondi, J., 2013. Confronting expert-based and modelled distributions for species with uncertain conservation status : a case study from the corncrake (*Crex crex*). Biol. Conserv. 167, 161–171. https://doi.org/10.1016/j.biocon.2013.08.009.

Fukunaga, K., 2013. Introduction to statistical pattern recognition. Elsevier.

Gonzalez and Woods, 2017. Digital Image Processing, 4th edition. Pearson.

Guisan, A., Thuiller, W., 2005. Predicting species distribution: Offering more than simple habitat models. Ecol. Lett. 8, 993–1009. https://doi.org/10.1111/j.1461-0248.2005.00792.x.

Hanski, I., 2015. Habitat fragmentation and species richness. J. Biogeogr. 42, 989–993. https://doi.org/10.1111/jbi.12478.

Harpenslager, S.F., van den Elzen, E., Kox, M.A.R., Smolders, A.J.P., Ettwig, K.F., Lamers, L.P.M., 2015. Rewetting former agricultural peatlands: Topsoil removal as a prerequisite to avoid strong nutrient and greenhouse gas emissions. Ecol. Eng. 84, 159–168. https://doi.org/10.1016/j.ecoleng.2015.08.002.

Heikkinen, R.K., Marmion, M., Luoto, M., 2012. Does the interpolation accuracy of species distribution models come at the expense of transferability? Ecography 35, 276–288.

Heikkinen, K., Karppinen, A., Karjalainen, S.M., Postila, H., Hadzic, M., Tolkkinen, M., Marttila, H., Ihme, R., Kløve, B., 2018. Long-term purification efficiency and factors affecting performance in peatland-based treatment wetlands: an analysis of 28 peat extraction sites in Finland. Ecol. Eng. 117, 153–164. https://doi.org/10.1016/j.ecoleng.2018.04.006.

Hengl, T., Sierdsema, H., Radović, A., Dilo, A., 2009. Spatial prediction of species' distributions from occurrence-only records: combining point pattern analysis, ENFA and regression-kriging. Ecol. Modell. 220, 3499–3511. https://doi.org/10.1016/j.ecolmodel.2009.06.038.

Higgins, S.I., Larcombe, M.J., Beeton, N.J., Conradi, T., Nottebrock, H., 2020. Predictive ability of a process-based versus a correlative species distribution model. Ecol. Evol. 10 (20), 11043–11054.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C., Guisan, A., 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecol. Modell. 199 (2), 142–152. https://doi.org/10.1016/j.ecolmodel.2006.05.017.

Hollander, M., Wolfe, D.A., 1973. Nonparametric Statistical Methods. John Wiley & Sons, New York.

Huang, Q., Fleming, C.H., Robb, B., Lothspeich, A., Songer, M., 2018. How different are species distribution model predictions?—Application of a new measure of dissimilarity and level of significance to giant panda *Ailuropoda melanoleuca*. Ecol. Inf. 1 (46), 114–124.

Hyvärinen, E., Juslén, A., Kemppainen, E., Uddström, A., Liukko, U.M. (Eds.), 2019. The 2019 Red List of Finnish Species. Ympäristöministeriö & Suomen ympäristökeskus, Helsinki, p. 704.

IPBES, 2018. In: The IPBES assessment report on land degradation and restoration. Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Bonn, Germany, p. 744. https://doi.org/10.4324/9781315640051-105.

Iturbide, M., Bedia, J., Herrera, S., del Hierro, O., Pinto, M., Gutiérrez, J.M., 2015. A framework for species distribution modelling with improved pseudo-absence generation. Ecol. Modell. 312, 166–174. https://doi.org/10.1016/j.ecolmodel.2015.05.018.

Jiménez-Valverde, A., Lobo, J.M., 2007. Threshold criteria for conversion of probability of species presence to either-or presence-absence. Acta Oecologica 31, 361–369. https://doi.org/10.1016/j.actao.2007.02.001.

Kappal, S., 2019. Data normalization using median median absolute deviation MMAD based Z-score for robust predictions vs. min–max normalization. London J. Res. Sci. Nat. Formal. 19, 39–44.

Kearney, M., Porter, W., 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. Ecol. Lett. 12 (4), 334–350. https://doi.org/10.1111/j.1461-0248.2008.01277.x.

Kohonen, T., 2001. Self-Organizing Maps, 3rd edition. Springer-Verlag, Berlin, Germany.

Kramer-Schadt, S., Niedballa, J., Pilgrim, J.D., Schröder, B., Lindenborn, J., Reinfelder, V., Stillfried, M., Heckmann, I., Scharf, A.K., Augeri, D.M., Cheyne, S.M., Hearn, A.J., Ross, J., Macdonald, D.W., Mathai, J., Eaton, J., Marshall, A.J., Semiadi, G., Rustam, R., Bernard, H., Alfred, R., Samejima, H., Duckworth, J.W., Breitenmoser-Wuersten, C., Belant, J.L., Hofer, H., Wilting, A., 2013. The importance of correcting for sampling bias in MaxEnt species distribution models. Divers. Distrib. 19, 1366–1379. https://doi.org/10.1111/ddi.12096.

Lentini, P.E., Stirnemann, I.A., Stojanovic, D., Worthy, T.H., Stein, J.A., 2018. Using fossil records to inform reintroduction of the kakapo as a refugee species. Biol. Conserv. 217, 157–165. https://doi.org/10.1016/j.biocon.2017.10.027.

Li, J., Fan, G., He, Y., 2020. Predicting the current and future distribution of three Coptis herbs in China under climate change conditions, using the MaxEnt model and chemical analysis. Sci. Total Environ. 698, 134141. https://doi.org/10.1016/j.scitotenv.2019.134141.

Li, W., Guo, Q., Elkan, C., 2011. Can we model the probability of presence of species without absence data? Ecography 34, 1096–1105. https://doi.org/10.1111/j.1600-0587.2011.06888.x.

Lier, M., Korhonen, K.T., Packalen, T., Sauvula-Seppälä, T., Tuomainen, T., Viitanen, J., Mutanen, A., Vaahtera, E., Hyvärinen, J., 2019. Finland's forests 2019. Helsinki, Natural Resources Institute Finland (Luke).

Liu, C., Newell, G., White, M., 2016. On the selection of thresholds for predicting species occurrence with presence-only data. Ecol. Evol. 6, 337–348. https://doi.org/10.1002/ece3.1878.

Liu, C., White, M., Newell, G., 2013. Selecting thresholds for the prediction of species occurrence with presence-only data. J. Biogeogr. J, 778–789. https://doi.org/10.1111/jbi.12058.

Low, B.W., Zeng, Y., Tan, H.H., Yeo, D.C., 2021. Predictor complexity and feature selection affect Maxent model transferability: evidence from global freshwater invasive species. Divers. Distrib. 27, 497–511.

LUKE, 2019. Finnish statistical yearbook of forestry 2019. Natural Resources Institute Finland, Helsinki.

Manzoor, S.A., Griffiths, G., Lukac, M., 2018. Species distribution model transferability and model grain size – finer may not always be better. Sci. Rep. 1–9 https://doi.org/10.1038/s41598-018-25437-1.

Marcer, A., Sáez, L., Molowny-horas, R., Pons, X., Pino, J., 2013. Using species distribution modelling to disentangle realised versus potential distributions for rare species conservation. Biol. Conserv. 166, 221–230. https://doi.org/10.1016/j.biocon.2013.07.001.

Marsaglia, G., Tsang, W.W., Wang, J., 2003. Evaluating Kolmogorov's distribution. J. Stat. Softw. 8 (18), 1–4. https://doi.org/10.18637/jss.v008.i18.

Martin, G.D., Magengelele, N.L., Paterson, I.D., Sutton, G.F., 2020. South African Journal of Botany Climate modelling suggests a review of the legal status of Brazilian pepper *Schinus terebinthifolia* in South Africa is required. South African J. Bot. 132, 95–102. https://doi.org/10.1016/j.sajb.2020.04.019.

Mateo, R.G., Croat, T.B., Felicísimo, Á.M., Muñoz, J., 2010. Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. Divers. Distrib. 16, 84–94. https://doi.org/10.1111/j.1472-4642.2009.00617.x.

McCune, J.L., 2016. Species distribution models predict rare species occurrences despite significant effects of landscape context. J. Appl. Ecol. 53, 1871–1879. https://doi.org/10.1111/1365-2664.12702.

McKerrow, A.J., Tarr, N.M., Rubino, M.J., Williams, S.G., 2018. Patterns of species richness hotspots and estimates of their protection are sensitive to spatial resolution. Divers. Distrib. 24, 1464–1477. https://doi.org/10.1111/ddi.12779.

Merow, C., Smith, M.J., Silander, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions : what it does, and why inputs and settings matter. Ecography 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x.

Mikkonen, N., Leikola, N., Lahtinen, A., Lehtomäki, J., Halme, P., 2018. Forest areas important for biodiversity in Finland - Final report of the Zonation analyzes of the diversity values of wooded habitats, Reports of the Finnish Environment Institute 9/2018. Finnish Environment Institute.

Niittynen, P., Heikkinen, R.K., Aalto, J., Guisan, A., Kemppinen, J., Luoto, M., 2020. Fine-scale tundra vegetation patterns are strongly related to winter thermal conditions. Nat. Clim. Chang. 10, 1143–1148. https://doi.org/10.1038/s41558-020-00916-4.

Norris, D., Ferraz, S.F.D.B., Villani, J.P., Galetti, M., 2011. How to not inflate population estimates ? Spatial density distribution of white-lipped peccaries in a continuous Atlantic forest. Anim. Conserv. 14, 492–501. https://doi.org/10.1111/j.1469-1795.2011.00450.x.

Oke, T.A., Hager, H.A., 2017. Assessing environmental attributes and effects of climate change on Sphagnum peatland distributions in North America using single- and multi-species models. PLoS ONE 12, 1–16. https://doi.org/10.1371/journal.pone.0175978.

Orme, C.D.L., Davies, R.G., Burgess, M., Eigenbrod, F., Pickup, N., Olson, V.A., Webster, A.J., Ding, T.S., Rasmussen, P.C., Ridgely, R.S., Stattersfield, A.J., Bennett, P.M., Blackburn, T.M., Gaston, K.J., Owens, I.P.F., 2005. Global hotspots of species richness are not congruent with endemism or threat. Nature 436, 1016–1019. https://doi.org/10.1038/nature03850.

Páez, A., Farber, S., Wheeler, D., 2011. A simulation-based study of geographically weighted regression as a method for investigating spatially varying relationships. Environ. Plann. A 43 (12), 2992–3010.

Parviainen, M., Luoto, M., Ryttäri, T., Heikkinen, R.K., 2008. Modelling the occurrence of threatened plant species in taiga landscapes: methodological and ecological perspectives. J. Biogeogr. 35, 1888–1905. https://doi.org/10.1111/j.1365-2699.2008.01922.x.

Parviainen, M., Marmion, M., Luoto, M., Thuiller, W., Heikkinen, R.K., 2009. Using summed individual species models and state-of-the-art modelling techniques to identify threatened plant species hotspots. Biol. Conserv. 142, 2501–2509. https://doi.org/10.1016/j.biocon.2009.05.030.

Parviainen, M., Zimmermann, N.E., Heikkinen, R.K., Luoto, M., 2013. Using unclassified continuous remote sensing data to improve distribution models of red-listed plant species. Biodivers. Conserv. 22, 1731–1754. https://doi.org/10.1007/s10531-013-0509-1.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Modell. 190 (3–4), 231–259.

Phillips, S.J., Dubik, M., 2008. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31, 161–175. https://doi.org/10.1111/j.2007.0906-7590.05203.x.

Pirinen, P., Simola, H., Juha, A., Kaukoranta, J.-P., Karlsson, P., Ruuhela, R., 2012. Climatological statistics of Finland 1981–2010. Finnish Meteorological Institute, Reports 2012:1, Finnish Meteorological Institute p83.

Progin, D., 2018. Modelling habitat suitability of bats in the Prealps. Master thesis. Lausanne University, Lausanne, Switzerland.

Qiao, H., Feng, X., Escobar, L.E., Peterson, A.T., Soberón, J., Zhu, G., Papeş, M., 2019. An evaluation of transferability of ecological niche models. Ecography 42 (3), 521–534.

R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria http://www.R-project.org/.

Ranc, N., Santini, L., Rondinini, C., Boitani, L., Poitevin, F., Angerbjörn, A., Maiorano, L., 2017. Performance tradeoffs in target-group bias correction for species distribution models. Ecography 40, 1076–1087. https://doi.org/10.1111/ecog.02414.

Randin, C.F., Dirnböck, T., Dullinger, S., Zimmermann, N.E., Zappa, M., Guisan, A., 2006. Are niche-based species distribution models transferable in space? J. Biogeogr. 33, 1689–1703. https://doi.org/10.1111/j.1365-2699.2006.01466.x.

Regos, A., Gagne, L., Alcaraz-Segura, D., Honrado, J.P., Domínguez, J., 2019. Effects of species traits and environmental predictors on performance and transferability of ecological niche models. Scientific reports 9, 4221.

Reisenbüchler, M., Bui, M.D., Rutschmann, P., 2021. Reservoir sediment management using artificial neural networks: A case study of the lower section of the alpine saalach river. Water 13, 818. https://doi.org/10.3390/w13060818.

Rissanen, T., Niittynen, P., Soininen, J., Luoto, M., 2021. Snow information is required in subcontinental scale predictions of mountain plant distributions. Glob. Ecol. Biogeogr. 30, 1502–1513. https://doi.org/10.1111/geb.13315.

Saarimaa, M., Aapala, K., Tuominen, S., Karhu, J., Parkkari, M., Tolvanen, A., 2019. Predicting hotspots for threatened plant species in boreal peatlands. Biodivers. Conserv 28 (5), 1173–1204.

Sallinen, A., Tuominen, S., Kumpula, T., Tahvanainen, T., 2019. Undrained peatland areas disturbed by surrounding drainage: a large scale GIS analysis in Finland with a special focus on aapa mires. Mires Peat 24, 1–22.

Senay, S.D., Worner, S.P., Ikeda, T., 2013. Novel three-step pseudo-absence selection technique for improved species distribution modelling. PLoS ONE 8, e71218. https://doi.org/10.1371/journal.pone.0071218.

Seppä, H., 2002. Mires of Finland: regional and local controls of vegetation, landforms, and long-term dynamics. Fennia-Int. J. Geogr. 180 (1–2), 43–60.

Shamshirband, S., Jafari Nodoushan, E., Adolf, J.E., Abdul Manaf, A., Mosavi, A., Chau, K.W., 2019. Ensemble models with uncertainty analysis for multi-day ahead forecasting of chlorophyll a concentration in coastal waters. Eng. Appl. Comput. Fluid Mech. 13, 91–101. https://doi.org/10.1080/19942060.2018.1553742.

Similä, M., Aapala, K., Penttinen, J. (Eds.), 2014. Ecological restoration in drained peatlands – best practices from Finland. Metsähallitus, Natural Heritage Services, Vantaa 2014, p. 84.

Skov, F., Svenning, J.C., 2004. Potential impact of climatic change on the distribution of forest herbs in Europe. Ecography 27, 366–380. https://doi.org/10.1111/j.0906-7590.2004.03823.x.

Svenning, J., Normand, S., Kageyama, M., 2008. Glacial refugia of temperate trees in Europe : insights from species distribution modelling. J. Ecol. 96, 1117–1127. https://doi.org/10.1111/j.1365-2745.2008.01422.x.

Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. Science 240, 1285–1293.

Syfert, M.M., Joppa, L., Smith, M.J., Coomes, D.A., Bachman, S.P., Brummitt, N.A., 2014. Using species distribution models to inform IUCN Red List assessments. Biol. Conserv. 177, 174–184. https://doi.org/10.1016/j.biocon.2014.06.012.

SYKE, 2009. Finnish Environment Institute spatial drainage data on peatlands. Helsinki.

Tehrani, N.A., Naimi, B., Jaboyedoff, M., 2020. Toward community predictions: Multi-scale modelling of mountain breeding birds' habitat suitability, landscape preferences, and environmental drivers. Ecol. Evol. 10, 5544–5557.

Thapa, A., Hu, Y., Chandra Aryal, P., Singh, P.B., Shah, K.B., Wei, F., 2020. The endangered red panda in Himalayas: potential distribution and ecological habitat associates. Glob. Ecol. Conserv. 21, e00890. https://doi.org/10.1016/j.gecco.2019.e00890.

Tikkanen, O.P., Martikainen, P., Hyvärinen, E., Junninen, K., Kouki, J., 2006. Red-listed boreal forest species of Finland: associations with forest structure, tree species, and decaying wood. Ann. Zool. Fennici 43, 373–383.

Tolvanen, A., Saarimaa, M., Tuominen, S., Aapala, K., 2020. Is 15% restoration sufficient to safeguard the habitats of boreal red-listed mire plant species? Glob. Ecol. Conserv. 23, e01160. https://doi.org/10.1016/j.gecco.2020.e01160.

Torres, L.G., Sutton, P.J., Thompson, D.R., Delord, K., Weimerskirch, H., Sagar, P.M., Sommer, E., Dilley, B.J., Ryan, P.G., Phillips, R.A., 2015. Poor transferability of species distribution models for a pelagic predator, the grey petrel, indicates contrasting habitat preferences across ocean basins. PLoS ONE 10 (3), e0120014.

Trotta-Moreu, N., Lobo, J.M., 2010. Deriving the species richness distribution of geotrupinae (Coleoptera: Scarabaeoidea) in Mexico From the overlap of individual model predictions. Environ. Entomol. 39, 42–49. https://doi.org/10.1603/EN08179.

Turunen, J., 2008. Development of Finnish peatland area and carbon storage 1950–2000. Boreal Environ. Res. 13, 319–334.

Välikangas, T., Suomi, T., Elo, L.L., 2018. A systematic evaluation of normalization methods in quantitative label-free proteomics. Brief. Bioinformat. 19, 1–11.

Virzi, T., Lockwood, J.L., Lathrop, R.G., Grodsky, S.M., Drake, D., 2017. Predicting American Oystercatcher (*Haematopus palliatus*) Breeding Distribution in an Urbanized Coastal Ecosystem Using Maximum Entropy Modeling. Waterbirds 40, 104–122. https://doi.org/10.1675/063.040.sp112.

Wenger, S.J., Olden, J.D., 2012. Assessing transferability of ecological models: an underappreciated aspect of statistical validation. Methods Ecol. Evol. 3, 260–267. https://doi.org/10.1111/j.2041-210X.2011.00170.x.

Williams, P., Gibbons, D., Margules, C., Rebelo, A., Humphries, C., Pressey, R., 1996. A comparison of richness hotspots, rarity hotspots, and complementary areas for conserving diversity of British birds. Conserv. Biol. 10, 155–174. https://doi.org/10.1046/j.1523-1739.1996.10010155.x.

Wisz, M.S., Guisan, A., 2009. Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. BMC Ecol. 9, 1–13. https://doi.org/10.1186/1472-6785-9-8.

Yu, Z.C., 2012. Northern peatland carbon stocks and dynamics: a review. Biogeosciences 9, 4071–4085. https://doi.org/10.5194/bg-9-4071-2012.

Zedler, J.B., Kercher, S., 2005. Wetland resources: status, trends, ecosystem services, and restorability. Annu. Rev. Environ. Resour. 30, 39–74. https://doi.org/10.1146/annurev.energy.30.050504.144248.

Zhang, K., Yao, L., Meng, J., Tao, J., 2018. Maxent modeling for predicting the potential geographical distribution of two peony species under climate change. Sci. Total Environ. 634, 1326–1334. https://doi.org/10.1016/j.scitotenv.2018.04.112.

Zhou, Y., 2020. Real-time probabilistic forecasting of river water quality under data missing situation: deep learning plus post-processing techniques. J. Hydrol. 589, 125164. https://doi.org/10.1016/j.jhydrol.2020.125164.

Zurell, D., Franklin, J., König, C., Bouchet, P.J., Dormann, C.F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A., et al., 2020. A standard protocol for reporting species distribution models. Ecography 43, 1261–1277. https://doi.org/10.1111/ecog.04960.