CARBON SEQUESTRATION AND BIODIVERSITY IMPACTS IN FORESTED ECOSYSTEMS

Role of data uncertainty when identifying important areas for biodiversity and carbon in boreal forests

Heini Kujala , Francesco Minunno, Virpi Junttila,
Ninni Mikkonen, Annikki Mäkelä, Raimo Virkkala, Anu Akujärvi,
Niko Leikola, Risto K. Heikkinen

Abstract Forest conservation plays a central role in meeting national and international biodiversity and climate targets. Biodiversity and carbon values within forests are often estimated with models, introducing uncertainty to decision making on which forest stands to protect. Here, we explore how uncertainties in forest variable estimates affect modelled biodiversity and carbon patterns, and how this in turn introduces variability in the selection of new protected areas. We find that both biodiversity and carbon patterns were sensitive to alterations in forest attributes. Uncertainty in features that were rare and/or had dissimilar distributions with other features introduced most variation to conservation plans. The most critical data uncertainty also depended on what fraction of the landscape was being protected. Forests of highest conservation value were more robust to data uncertainties than forests of lesser conservation value. Identifying critical sources of model uncertainty helps to effectively reduce errors in conservation decisions.

Keywords Biodiversity conservation ·
Carbon sequestration · Forest growth model ·
Model uncertainty · Spatial conservation planning ·
Species distribution modelling

INTRODUCTION

Forests play a central role in the global efforts of halting biodiversity loss and mitigating climate change. In the boreal region, forests account for > 30% of global forest carbon stock and 20% of the annual global forest carbon sink (Pan et al. 2011; Gauthier et al. 2015). Boreal forests comprise one-third of global forest cover, providing habitat for a large number of species (IPBES 2019), yet nearly two-thirds of these forests are under economic use (Gauthier et al. 2015). Intensive forestry and clear-cut logging reduces habitat available for several forest-dwelling species (IPBES 2019; Mönkkönen et al. 2022), and although some early succession species may benefit from harvesting, regenerated stands hold less biological and structural diversity than those originating from natural disturbances (Gauthier et al. 2015). Forest management can increase their carbon sequestration but the benefits of this depend on whether the harvested wood is used for products that prevent the carbon being re-released back to atmosphere in long (e.g. hardwood) or short (e.g. pulp, paper) term (Soimakallio et al. 2016). Clear cutting also decreases the carbon stocks of forests and increases emissions from decomposing harvest residues (Kolari et al. 2004; Goulden et al. 2011).

In Finland, situated in the boreal region, high and low productive forests combined cover about 22.7 million hectares (75%) of the land area (Vaahtera et al. 2021). Around 90% of these are under even-aged rotation forestry, predominantly coniferous managed forests, while the rest are protected. The average size of clear-cut, where no or only a small number of trees are left as seed or retention trees, is 1.76 ha, and afterwards stands are re-planted or allowed to regenerate naturally, and the rotation between final fellings is on average 80–100 years (Kniivilä et al. 2020). The long history of logging and active management has significantly altered the age structure and functional heterogeneity of the Finnish forests (Gauthier et al. 2015; Korhonen 2021), decreasing the area of old-growth forests, number of large trees and the volume of dead wood

(Mönkkönen et al. 2022). These large-scale alterations have led to declines in forest biodiversity: 11% of forest species and 76% of the forest habitat types are threatened (Hyvärinen et al. 2019; Kontula and Raunio 2019). Currently, forest heterogeneity is predominantly driven by forestry treatments, forest site type (describing site fertility and aridity) and climate.

The synergistic benefits of forest conservation for both biodiversity and climate change mitigation are increasingly recognised in international and national policies, such as the EU 2030 Biodiversity Strategy, although how much carbon sinks should be emphasised over carbon storages in these policies is still debated (Soimakallio et al. 2016; Pukkala 2018). Nevertheless, through the identification of forests important for both biodiversity and carbon, forest conservation can be strategically targeted to areas that support meeting both policy targets cost-efficiently. In recent years, studies have demonstrated the use of spatial prioritisation (optimisation) tools (Moilanen et al. 2009) for identifying co-benefits and trade-offs between biodiversity and carbon services at various spatial scales (e.g. Forsius et al. 2021; Jung et al. 2021).

Such optimisation exercises require large-scale spatial data. Both biodiversity and carbon data are commonly created using models that estimate the habitat suitability of forest stands for different species and the amount of carbon stored in or sequestered by them (Forsius et al. 2021; Miettinen et al. 2021). As all model estimates contain uncertainty, the goodness of any spatial prioritisation result is inherently dependent on the accuracy of the input data that underpins it. The effects of model uncertainty on the estimated values have been extensively studied (e.g. Barry and Elith 2006; Mäkelä et al. 2020), but how these uncertainties impact spatial optimisation results is less well understood. Recent research has quantified how the impact of data uncertainties on prioritisation results depends on the alteration made, characteristics of the input data and the total data pool that is used to produce the priority result (Kujala et al. 2018a, b). The first-line improvements of modelled input data should focus on uncertainties that, when resolved, change the prioritisation result the most.

Here, we explore the impact of different data uncertainties on conservation prioritisation within a real-world forest conservation context. We look at how the uncertainty in forest variable estimates affects modelled distributions of biodiversity and carbon patterns, and how this in turn introduces variability in a hypothetical conservation plan. We use several features of carbon (size of carbon sinks, amount of carbon stored in trees, ground vegetation and soil) and biodiversity to explore their co-occurrences and role in multi-objective conservation planning, and three time periods from present to 2050 to understand their dynamics through time. We purposefully investigate the

data uncertainties in the absence of other major drivers, such as forest harvesting and climate change, so as to exclude their impact. Our purpose is not to identify what areas to protect, but to understand how this decision is impacted by data uncertainties. In particular, we explore whether forest conservation strategies are more sensitive to uncertainties associated with biodiversity vs carbon estimates, and which of the individual sources of uncertainty are most important to solve so to effectively reduce potential errors in spatial conservation decisions.

MATERIALS AND METHODS

Overview of the analysis design

We used the following approach in this study. First, using measurements from forest stands and a mechanistic forest growth model, we created spatial data for several forest attributes (structural variables, site type and mean age, hereafter called baseline forest variables, see "Forest and carbon data" section). These baseline estimates were then used to model the presence of carbon and biodiversity values across the study area (baseline carbon values, "Forest and carbon data" section and baseline biodiversity values, "Biodiversity data" section), and to identify most important forest areas for conservation at each time period (baseline priorities, "Spatial prioritisation" section) (Fig. 1). Next, we iteratively re-simulated 50 realisations of the forest variables (called sample forest variables, "Sample iterations" section). From each sample, we re-created the maps of carbon and biodiversity data and the respective conservation plan, keeping all other environmental variables constant. This resulted in 50 conservation plans (sample priorities). We measured the variability in the sample model outputs and used a canonical correlation analysis to understand how uncertainty in each forest variable and biodiversity and carbon value affects the final conservation plan ("Quantification of uncertainty" section). In next sections, we describe each analysis step and the data and methods used in more detail.

Study area

The study area is located in Central Finland, Europe (Fig. 2). The area covers 41 503 km², of which $\sim$ 6 500 km² (15.7%) are freshwater bodies (lakes, ponds, rivers). Based on the CORINE Land Cover 2018 data, approximately 60% and 10% of the land area are forests and mires, respectively, 4% is urban and other built-up environments and 9% is in agricultural use (Fig. 2). The majority (71%) of the forests are coniferous (> 75% of trees conifers), while 4% and 24% are deciduous (> 75% deciduous trees)
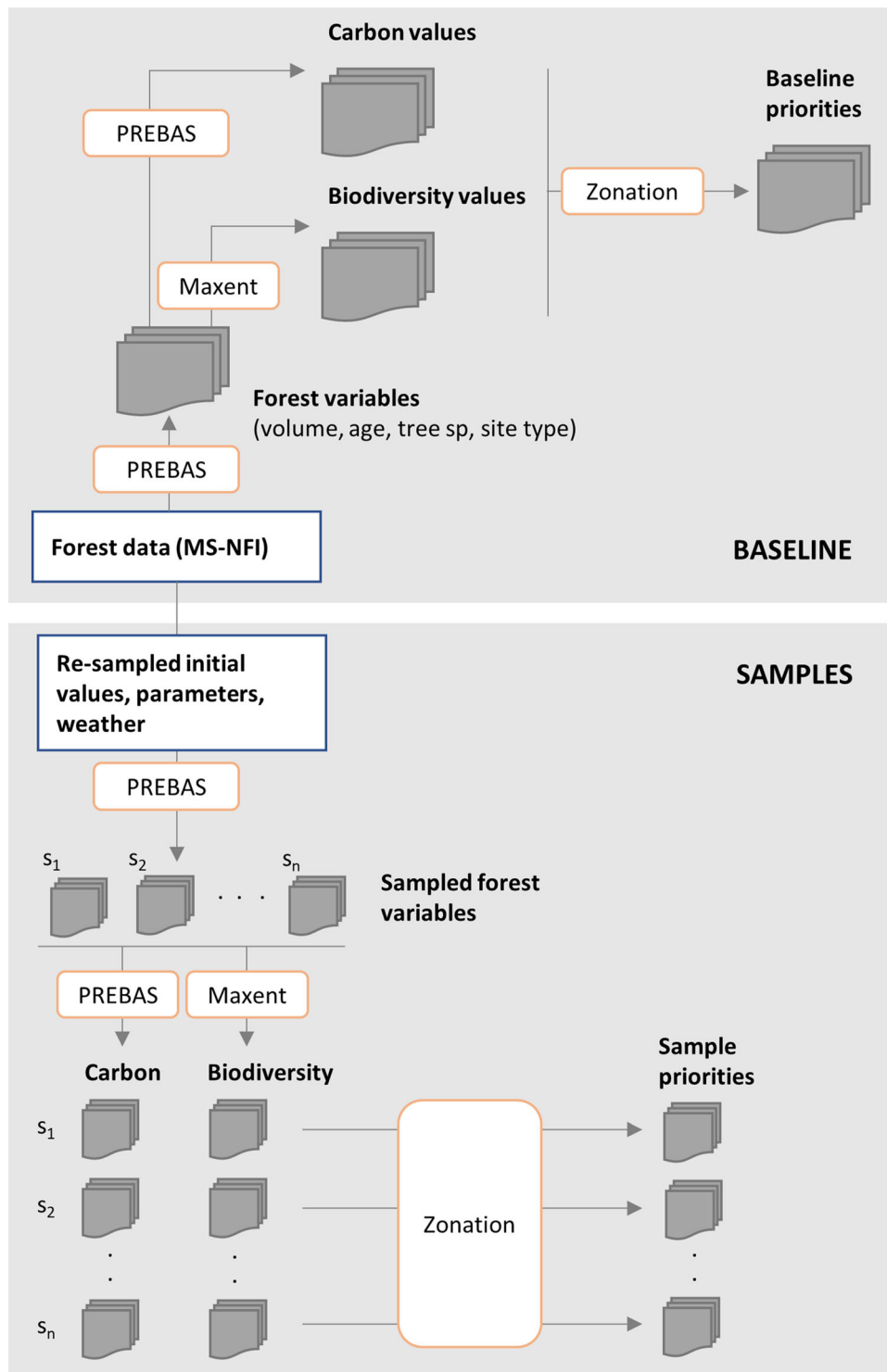
Fig. 1 Analysis design showing the different modelling steps and main methodological tools used. Data on the original state of forests came from the multi-source national forest inventories (MS-NFI). The carbon-balance-based stand growth and gas exchange model PREBAS were used to first simulate forest growth and stand characteristics (forest variables) and then to estimate the size of carbon storage and sinks in forest stands (carbon values). The machine learning tool Maxent was used to model habitat suitability of forests for six bird species (biodiversity values), while also accounting for other non-forest variables (climate, land-use types). Finally, Zonation was used to identify potential priority areas for forest conservation. The baseline forest variables were simulated using existing forest data (MS-NFI). For the samples (s_1, s_2, …, s_n), the PREBAS simulation was iteratively repeated while introducing different sources of uncertainty
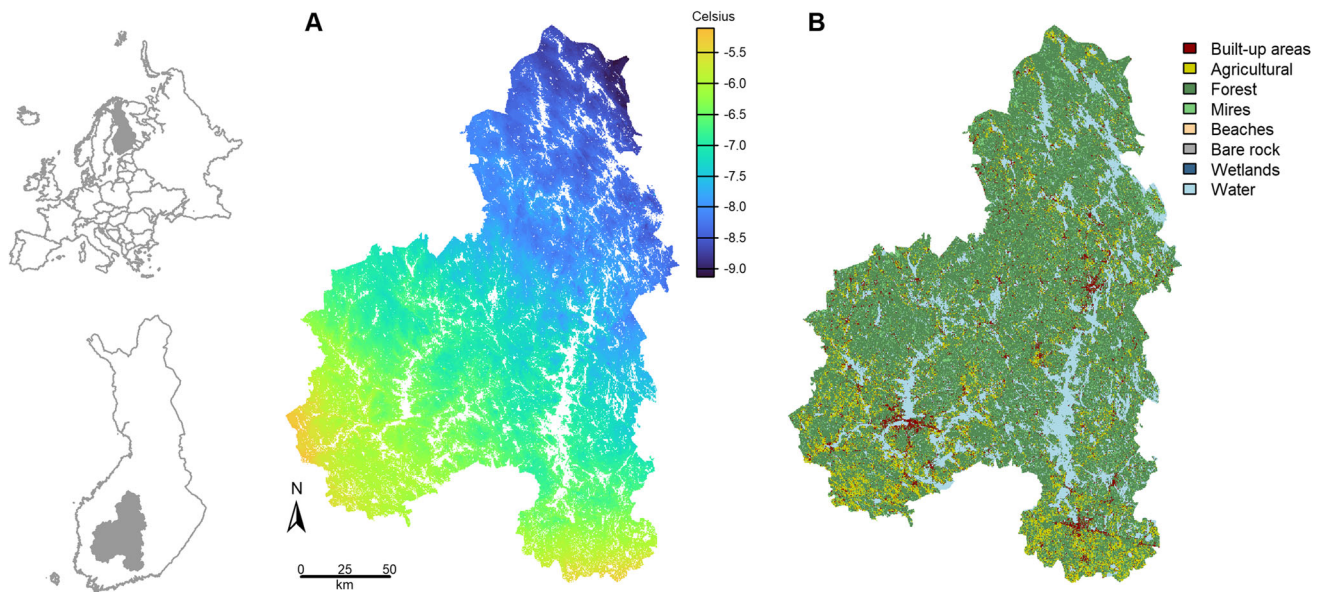
Fig. 2 Study area and its location in Finland, Europe. Maps **A** and **B** show the mean January temperature and major land-use categories, respectively

and mixed, respectively. The dominant tree species are Norway spruce (*Picea abies*), Scots pine (*Pinus sylvestris*), silver birch (*Betula pendula*) and downy birch (*Betula pubescens*). Of the forests, 12% are grown on drained peatlands. Approximately 25 130 ha of the forests are clear-cut, and another 93 000 ha managed (e.g. thinned) every year in this region (Vaahtera et al. 2021). All data used in this study were scaled to a uniform 96 × 96 m resolution grid covering the entire study area but excluding water bodies.

Forest and carbon data

We used PREBAS to simulate forest variables and carbon balance. PREBAS is a C-balance-based stand growth and gas exchange model, which combines a process-based forest growth model (called CROBAS, Valentine and Mäkelä 2005) and a daily canopy gas exchange model (PRELES, Peltoniemi et al. 2015). Photosynthesis (GPP) and evapotranspiration are calculated using a light-use efficiency approach linked to soil moisture and driven by daily climate information and ambient CO_2 concentration. GPP is allocated to mean-tree growth and respiration at an annual time step. To calculate net ecosystem production (NEP), PREBAS has been linked through annual litter inputs with the soil C model YASSO15, which has been calibrated to one meter depth (Viskari et al. 2020). PRE-BAS has been calibrated with Nordic eddy covariance sites and Finnish growth experiments (Minunno et al. 2019).

PREBAS produces an array of outputs. Of these, we used the stand volume, age, dominant tree species, mean tree height and mean diameter at breast height (DBH) (Table 1) together with information on forest site type to represent the forest variables in our analysis. To represent carbon values, we used the PREBAS estimated amount of carbon stored in trees, ground vegetation and soil and the annual size of carbon sinks of the stand, calculated from NEP.

To initiate PREBAS simulations, we used information on the initial state and type of the forest from the Multi Source National Forest Inventory 12 (MS-NFI 12, 2014–2018) maps, which describe the forest parameters across Finland at 16 × 16 m resolution (Tomppo et al. 2008). To reduce computational effort, the simulation was done on homogeneous segments consisting of multiple 16 m pixels (Haakana et al. 2022). Within segments, the initial value and growth in forest were assumed the same for all pixels. Information on forest site type was extracted directly from the MS-NFI data. Due to the sparsity of measurement-based estimates, setting initial carbon content of mineral soils is challenging and was therefore assumed to be in a steady state with the historical mean levels of harvest (round wood, pulpwood and energy wood). After each simulation, the results were restored back to the 16 m resolution and aggregated to the 96 × 96 m analysis resolution (Forsius et al. 2021). For the conservation prioritisation, we only considered carbon sinks by setting NEP values to zero in pixels that on average acted as a source across the simulated time period. This is in line with current conservation policies where areas of large sinks may be favoured but areas of large sources are not penalised if the site is otherwise important for preserving biodiversity.

Table 1 Selected PREBAS outputs used in this study. The interlinked submodels of PREBAS simulate process-based forest growth (CROBAS), daily canopy gas exchange (PRELES) and soil carbon through annual litter inputs (YASSO). For references and further details of the submodels, see the main text

Variable name	Meaning	Unit	Submodel
DBH	Stand mean diameter at breast height	cm	CROBAS
Height	Stand mean height	m	CROBAS
Vol	Stand stem volume	$m^3\ ha^{-1}$	CROBAS
Dec. vol	Deciduous stem volume	$m^3\ ha^{-1}$	CROBAS
Age	Dominant tree age	years	CROBAS
Species	Dominant tree species	Pine, spruce, deciduous	CROBAS
Soil C	Total soil carbon	$kg\ C\ ha^{-1}$	YASSO
Ground veg C	Total carbon in ground vegetation	$kg\ C\ ha^{-1}$	Ground vegetation submodel (see Junttila et al. 2023)
Tree C	Total carbon in tree biomass	$kg\ C\ ha^{-1}$	CROBAS
NEP/C sink	Net ecosystem production/size of carbon sinks	$g\ C\ m^{-2}\ year^{-1}$	PRELES + CROBAS + YASSO

For calculating regional carbon fluxes, both sinks and sources were considered.

Forest growth was simulated from 2015 to 2050 with a two-year initiation period (2015–2016). For transparency, during the simulations we assumed no climate change or further forest harvesting to take place, as our goal was to understand how input data uncertainty affects the spatial distribution of conservation priorities. The output variables from the simulations were produced as averages for three time periods: years 2017–2025 (T1, 9 years), 2026–2033 (T2, 8 years) and 2034–2050 (T3, 17 years).

Biodiversity data

For biodiversity, we used the nesting suitability maps of six forest-dwelling bird species: European honey buzzard (*Pernis apivorus,* PERAPI), northern goshawk (*Accipiter gentilis,* ACCGEN), common buzzard (*Buteo buteo,* BUTBUT), white-backed woodpecker (*Dendrocopos leucotos*, DENLEU), lesser-spotted woodpecker (*Dryobates minor,* DRYMIN) and Eurasian three-toed woodpecker (*Picoides tridactylus*, PICTRI). The nesting sites of these species have been shown to be good indicators of forest biodiversity, as they prefer mature or undisturbed forest stands, mosaics of forest types and deciduous forests that have high dead wood volume and richness of wood-decaying polypore species (Roberge and Angelstam 2006; Burgas et al. 2014). All three hawk species and the white-backed woodpecker are red-listed (threatened or near-threatened, Hyvärinen et al. 2019) and the lesser-spotted woodpecker and three-toed woodpecker have shown considerable long-term declines in Finland (Väisänen et al. 1998).

The nesting suitability maps were produced using the species distribution modelling (SDM) tool Maxent (Phillips et al. 2006). As SDMs can account for a broader suite of

variables (climate, non-forest habitats) and spatially relevant ecological drivers, such as the size and configuration of suitable forests and the quality of matrix (area between forests), they provide ecologically more realistic estimate of the goodness of forests for biodiversity than simple stand characteristics. We used bird ringing data independent from the MS-NFI to mark the locations of nesting sites of each species, and sets of environmental variables that describe the climate, forest and land cover characteristics at nesting site and landscape level (Table 2). Virkkala et al. (2022) modelled the nesting suitability of these species at the national scale and here we replicated these models using the same nesting records and species-specific sets of variables identified by Virkkala et al. (2022) (Table 2) but cut to the study area. The modelling was done using R v.4.1.0, Rstudio v.2022.07.01 and the package *dismo* v.1.3-5. We used the default settings of Maxent. Likely spatial biases in the ringing data were accounted for through the use of a bias grid and target group sampling (Phillips et al. 2009; Virkkala et al. 2022). The models were validated using a fivefold cross-validation to calculate the mean and standard deviation of the Area Under the receiver operating characteristic Curve (AUC), which is a commonly used metric to describe how well the model discriminates between known presences and absences. AUC values > 0.7 are thought to indicate an informative model (Fielding and Bell 1997) (Table 2). For the uncertainty analysis, we used models that were fitted using all available nesting records. We did not threshold the predicted values.

Sample iterations

The baseline forest variables were simulated using existing forest data (MS-NFI). To create sample forest variables, we iteratively repeated the PREBAS simulation while introducing three sources of uncertainty: (1) variable input

Table 2 Number of nesting records and the environmental variables used in the nesting suitability models for each bird species, and the average test AUC value and standard deviation of the models across fivefold cross-validation. For each variable, the values give the percent contribution and permutation importance (in brackets) of the full model using all records. The first value describes how much the variable contributes to the full model. The second value shows how much the model AUC is reduced when the variable is not included in the model. The three most important variables for each species in terms of contribution and permutation are in bold. LS refers to variables that were calculated at landscape level, either as averages (forest and climate variables) or proportion (land-use category), within a radius of either 500 m or 1 km around the nesting site, depending on the species (woodpeckers and hawks, respectively). Fixed variables are those that were kept constant in all model samples, whereas the others were replaced with a newly sampled data layer, see main text for details. *PERAPI* European honey buzzard, *ACCGEN* northern goshawk, *BUTBUT* common buzzard, *DENLEU* white-backed woodpecker, *DRYMIN* lesser-spotted woodpecker, *PICTRI* Eurasian three-toed woodpecker, *DBH* Diameter at breast height, *LS* Variable calculated at landscape level around the nesting site

		Hawks			Woodpeckers		
		ACCGEN	BUTBUT	PERAPI	DENLEU	DRYMIN	PICTRI
Number of records		928	356	75	64	27	56
Test AUC		0.795 ± 0.014	0.780 ± 0.024	0.726 ± 0.052	0.976 ± 0.003	0.835 ± 0.082	0.880 ± 0.063
Environmental variable	**Fixed**						
Stand volume (m³)		8.8 **(43.1)**	**39** (18.7)	**58.3 (48.3)**		0.0 (0.0)	0.0 (0.0)
Stand DBH (cm)		**22** (1.3)			2.7 (5.0)		
Stand height (m)			7.2 (13.2)	**10.8** (0.0)	4.7 (0.0)		**25.9** (4.6)
Dominant tree age (years)		**38.6 (18.9)**	7.7 (10.2)			1.9 (2.1)	1.4 (2.1)
Deciduous tree volume (m³)			4.1 (8.2)	3.3 (13.4)	**28.3 (37.6)**	5.0 **(24.1)**	
Main tree species				**16.5 (14.5)**	1.6 (0.8)	1.5 (0.0)	6.7 (2.7)
Forest site type		3.2 (2.2)		1.4 (3.4)	4.6 (0.1)		3.0 (1.4)
Shoreline forest (%)	x					11.7 (5.3)	
Agricultural areas (%)	x	1.4 (6.4)					2.3 (6.9)
Urban areas (%)	x	1.2 (0.1)	4.8 (1.0)	4.5 (5.0)			
Water areas (%)	x	3.1 (6.7)	1.9 (3.5)				
LS volume at 500 m radius (m³)					1.4 (2.5)		**21** (1.0)
LS volume at 1 km radius (m³)		0.2 (0.3)	2.6 (4.0)				
LS site type at 500 m radius (%)					5.3 (0.1)	**16.2** (4.6)	11.8 **(17)**
LS site type at 1 km radius (%)			1.9 (2.7)	0.6 (0.4)			
LS shoreline forest at 500 m radius (%)	x				**17.1** (4.7)	20.9 **(36.6)**	
LS forest on peatland at 500 m radius (%)	x				7.1 **(30.9)**	0.4 (1.8)	
LS marshlands at 500 m radius (%)	x				2.5 (0.0)	3.6 (2.0)	
LS agricultural areas at 1 km radius (%)	x		2.2 (2.0)	4.5 **(15.0)**			
LS urban areas at 500 m radius (%)	x					**34.2 (17.9)**	8.1 (8.4)
LS urban areas at 1 km radius (%)	x	**17 (18)**	**7.8 (18.6)**				
LS water areas at 500 m radius (%)	x				2.4 (0.2)	0.9 (5.7)	3.9 **(19)**
LS water areas at 1 km radius (%)	x		**18.9 (13.8)**				
January mean temp (°C)	x	4.4 (2.9)			**22.4 (18.1)**	3.7 (0.0)	**15.9 (36.8)**
GDD5	x		1.8 (4.2)	0.0 (0.0)			

uncertainty; (2) model parameter uncertainty and (3) uncertainty in weather conditions (Table S1).

To introduce variable input uncertainty, the initial segment-level variables (structural, site type and age) were resampled from their respective estimate distributions. For structural variable values (basal area split to proportions of pine, spruce and birch, stand mean height and mean DBH), estimate distributions were produced using data from three remotely sensed 100 × 100 km tiles across Central Finland from Miettinen et al. (2021) and assuming multivariate normal distribution. Variables were resampled simultaneously for each segment using a covariance matrix of errors. To avoid negative or unrealistically large values, and to maintain the covariance structure between variables, the segment samples were post-processed using the quantile matching procedure (Junttila and Kauranne 2018). For site

type, data on the satellite-based site type and forest structural variables from Miettinen et al. (2021) were used together with a probit model to calculate the probability of a segment belonging to each of the site type classes. Those probabilities were then used to resample the forest site type at the start of each iteration to reflect potential uncertainty in the initial classification of the MS-NFI. Uncertainty in the segment mean age was estimated using random samples from a normal distribution around MS-NFI-based mean age (across all tree species) and 10% standard deviation.

As the MS-NFI data does not include information about height of the crown base, we used species-specific empirical models fitted to data from permanent sample plots (Minunno et al. 2019) and varied the estimates with 10% standard deviation (estimate multiplied with $h \sim N(1, 0.1^2)$).Otherwise, model parameter uncertainty was generated by bootstrapping the joint posterior distributions of PRELES, CROBAS and YASSO parameters from Minunno et al. (2016), Minunno et al. (2019) and Viskari et al. (2020). The sampling strategy respected parameter correlations and interactions. Weather condition uncertainty was simulated by sampling yearly weather conditions from the last 46 years (1971–2016) to represent the local weather during the simulation period.

In total, we ran 50 iterations of the PREBAS simulation. In each iteration, the values for site type and PREBAS parameters were sampled at the start and then kept constant across the simulation time period. For more details on the uncertainty sampling, see Junttila et al. (2023).

Spatial prioritisation

Priority areas for protection were identified using the spatial prioritisation software Zonation v.5.0 (Moilanen et al. 2009; Moilanen et al. 2022). Zonation produces a priority ranking of each spatial unit (here 96 × 96 m grid cells) based on input feature data (6 bird species and 4 carbon layers), ordering the units from the least to the most important for conservation. This is done in an iterative optimisation algorithm where the software seeks to find a rank order that maximises the representation of all features in the top ranked grid cells. Consequently, a set of top ranked priority areas together typically capture high value areas for all input features. For each feature j, we used a benefit function $v_j (r_j) = r_j^z$ to describe the value of added protection v_j as a function of the accumulating fraction r_j of feature j's full distribution that is protected. The parameter z defines the shape of the function. We set $z = 0.25$ for all biodiversity features (bird species), following the well-established species-area relationship (Rosenzweig 1995). For carbon features, we used $z = 1$ as, unlike species

persistence, the persistence of carbon in the landscape does not have a non-linear relationship with the remaining area.

During the prioritisation, feature-specific benefits v_j of protecting a grid cell i were treated as additive. We produced a priority ranking for the baseline and each sampled iteration of the biodiversity and carbon features, and selected the top ranked 10% of the grid cells to represent the potential conservation solution for that iteration.

Quantification of uncertainty

We quantified uncertainty by measuring across iteration variability in (i) the PREBAS simulated forest variable and carbon values per hectare, (ii) amount of suitable bird habitat in the region and (iii) within grid-cell priority ranking and the frequency at which any one grid cell was included in the top 10% ranking. Although some of the forest variables correlate (volume, height, DBH), we examined them all as they have varying importance for dependent biodiversity (Table 2).

For each sample iteration and time period, we extracted forest variable, biodiversity, carbon and conservation priority values from 10 000 grid cells across the study region. We conducted a canonical correlation analysis (CCA), a multivariate extension of correlation analysis that allows identifying linear relationships between two datasets consisting of one or multiple variables (Stewart and Love 1968). We summarise the CCA results with the use of the redundancy index that expresses the amount of variance in the within-pixel priority rankings explained by each variable (van den Wollenberg 1977). The index receives values 0–1, where higher value indicates that the variable explains more of the variation, but there is no agreed interpretation of the exact value. Therefore, we examine the derived redundancy indices of each variable in relation to one another to reveal relative uncertainties. To better understand the explanatory power of the variables, we included three prioritisations in the CCA which were based on (i) both biodiversity and carbon features, (ii) biodiversity only and (iii) carbon only. We summarise the result for only (i) in the main text, see the Supplementary for all results.

RESULTS

Changes and uncertainty input data estimates

In the absence of harvesting, the tree height, DBH, volume and age of forests increased (Fig. 3). There was little variation in the across study area mean of forest variables between sample iterations (for segment-level variation, see Table S2). The volume and proportion of deciduous trees

also increased (Fig. 3, from 8.6 at T1 to 9.2% at T3), mainly due to intraspecific competition and the domination of deciduous tree species during the re-growth of recently clear-cut and early succession forests that were initially present in the region.

The amount of carbon stored in trees and soil increased from 473 to 571 million t C (+ 20%) across the region (Fig. 3), as no carbon was removed from the system through logging. Uncertainty around the tree carbon estimates decreased towards the last time period, whereas the estimates of soil carbon remained more variable. In contrast, ground vegetation carbon stayed relatively stable and became less variable through time. The total size of carbon sinks decreased across the region, from 3.6 to 2.3 million t C year^{-1} (− 36%), as tree growth slowed down in the maturing forests. Uncertainty around carbon sinks remained stable across time. There was more variability around the carbon estimates in comparison to forest variables (Fig. 3).

The undisturbed maturation of forest stands led into large increases in the amount of suitable nesting habitat for most bird species, in particular for the northern goshawk, common buzzard, the three-toed woodpecker but also the European honey buzzard (Fig. 3). Wood volume, tree height, DBH and age of forest stands were some of the strongest predictive variables for these species (Tables 2, S1) and all these increased through time. Although deciduous tree volume was a major predictor for the lesser-spotted and white-backed woodpeckers (Table 2), its increase did not translate into notable improvements in their nesting habitat (Fig. 3). This might be because: (i) the additional deciduous trees were mixed with conifers and these species prefer pure deciduous forests, (ii) the habitat increase of the lesser-spotted woodpecker is constrained to nearby water bodies and (iii) in the absence of disturbances, such as fires, old-growth deciduous forests preferred by the white-backed woodpecker tend to become overgrown by conifers (Fig. S2, Tables 2, S1). For all species, the estimated amount of available habitat became more uncertain through time, being largest for the common buzzard (Figs. 3, S1). There was no clear relationship between model fit (test AUC values, Table 2) and the variation in the amount of available habitat (Figs. 3, S9).

Changes and uncertainty in conservation priority rankings

Across the time periods, the priority rankings of forests for biodiversity and carbon increased in the southeast of the region, while forests in the northern parts of the region became less important (Fig. 4). By the last time period, there were only few forests ranked in the top 10% in the northern parts, while their locations shifted and became more evenly spread in the south-west (Figs. 4A, S4).

The ranking of individual grid cells varied notably across the sample iterations in each time period (Fig. 4C, vertical spread of points). This variation was reduced in the second time period, but increased again in the third time period (Figs. 4D, S5). There were no strong spatial patterns in the variability of cell rankings (Fig. S5). In all cases, grid cells that were on average ranked as the highest or the lowest priority across the 50 iterations showed least variation (Fig. 4C, D).

Still, even for the on average top 10% ranked grid cells (points right from the vertical line in Fig. 4C), this variation was large, and in some iterations, these cells were ranked as moderate or low priority. Furthermore, by the third time period, variation in the cell rankings increased more in the high ranked than low ranked cells (Fig. 4C, D). From this followed that the probability of any one grid cell to be included in the top 10% in all iterations decreased (Fig. 4B) and a larger number of grid cells could be ranked as top priority in at least one of the iterations (Fig. 4C, area of coloured points).

CCA

Across all grid cells, the variance in priority ranking was most strongly and consistently explained by the amount of carbon stored in trees and the distribution patterns of three-toed woodpecker, honey buzzard and common buzzard (Fig. 5A, upper row). Of the initial forest variables, tree volume followed by tree height and DBH explained most of the variation seen in priority rankings of all grid cells, particularly in the first two time periods (Fig. 5B). Tree volume and tree carbon storage correlate closely, in addition to which tree volume is one of the strongest predictors of the nesting suitability for several of the bird species (Table 1). Hence, minor changes in tree volume, and thereby tree carbon storage, may trickle into the prioritisation results via multiple input features, even if carbon features are not included in the prioritisation (Fig. S7).

By the third time period, most of the forest stands had become similar in terms of biomass and structural variables (Figs. 3, S3), and the forest variables alone had little effect on how grid cells were ranked. Concurrently, the importance of the tree carbon storage, common buzzard and honey buzzard on the ranking were reduced, while the importance of the three-toed woodpecker remained high (Fig. 5A, upper row). In the last time period, changes in the carbon sink estimates greatly affected the priority rankings, although this effect in itself was highly variable across the 50 iterations.

In contrast to all grid cells, variation in the top 10% ranked cells was more evenly explained by both carbon and
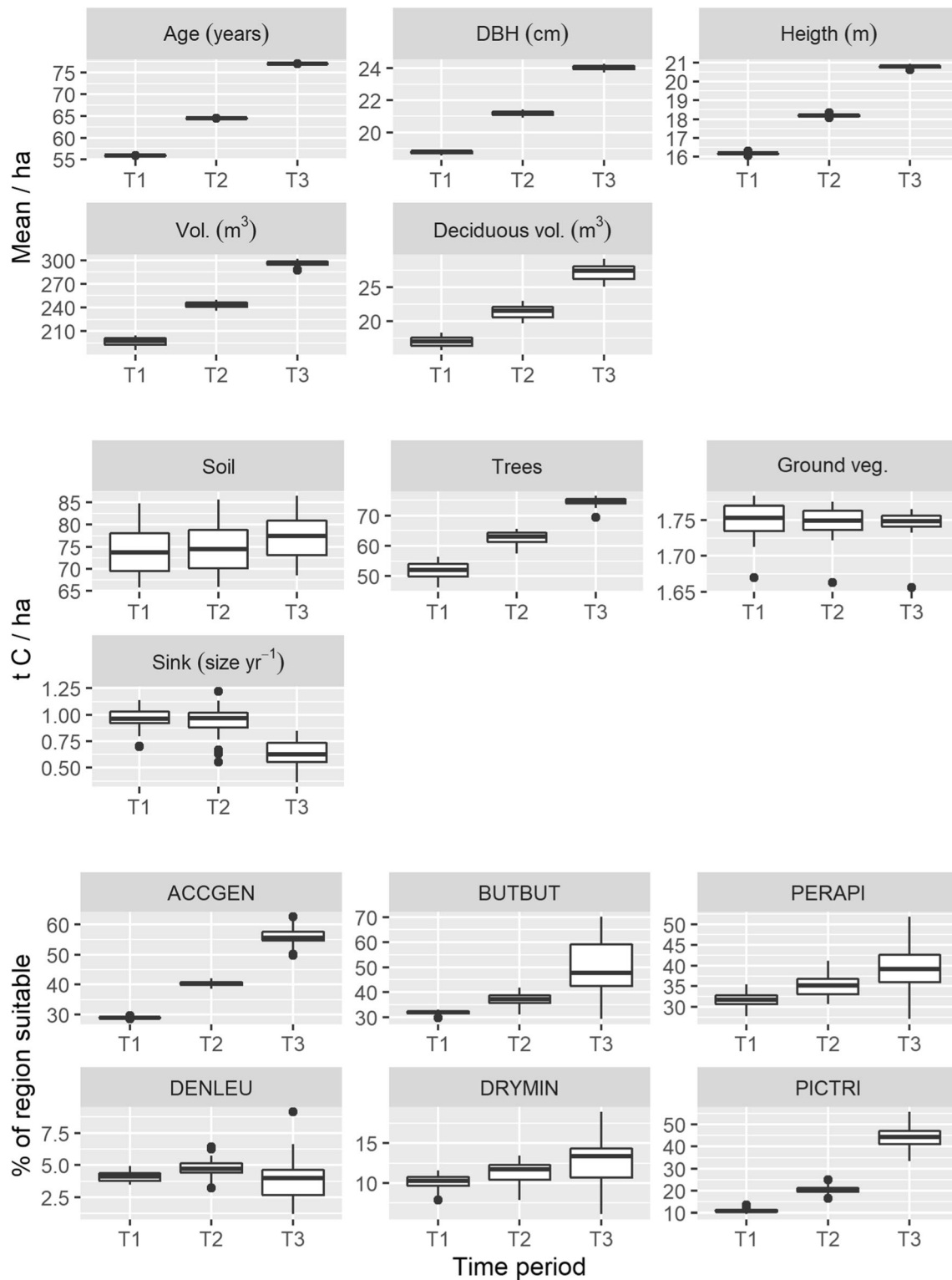
Fig. 3 Variation in the sampled forest variables and the respective carbon and biodiversity values. For the forest variables and carbon features, the values are given as means across study region. For segment-level means and variation, see Table S2. For the biodiversity features, the proportion of study area that is suitable is given, calculated as the sum of predicted values divided by the study area. *DBH* tree diameter at breast height, *ACCGEN* northern goshawk, *BUTBUT* common buzzard, *PERAPI* European honey buzzard, *DENLEU* white-backed woodpecker, *DRYMIN* lesser-spotted woodpecker, *PICTRI* Eurasian three-toed woodpecker. Note that the y-axes are not fixed but are on different scales
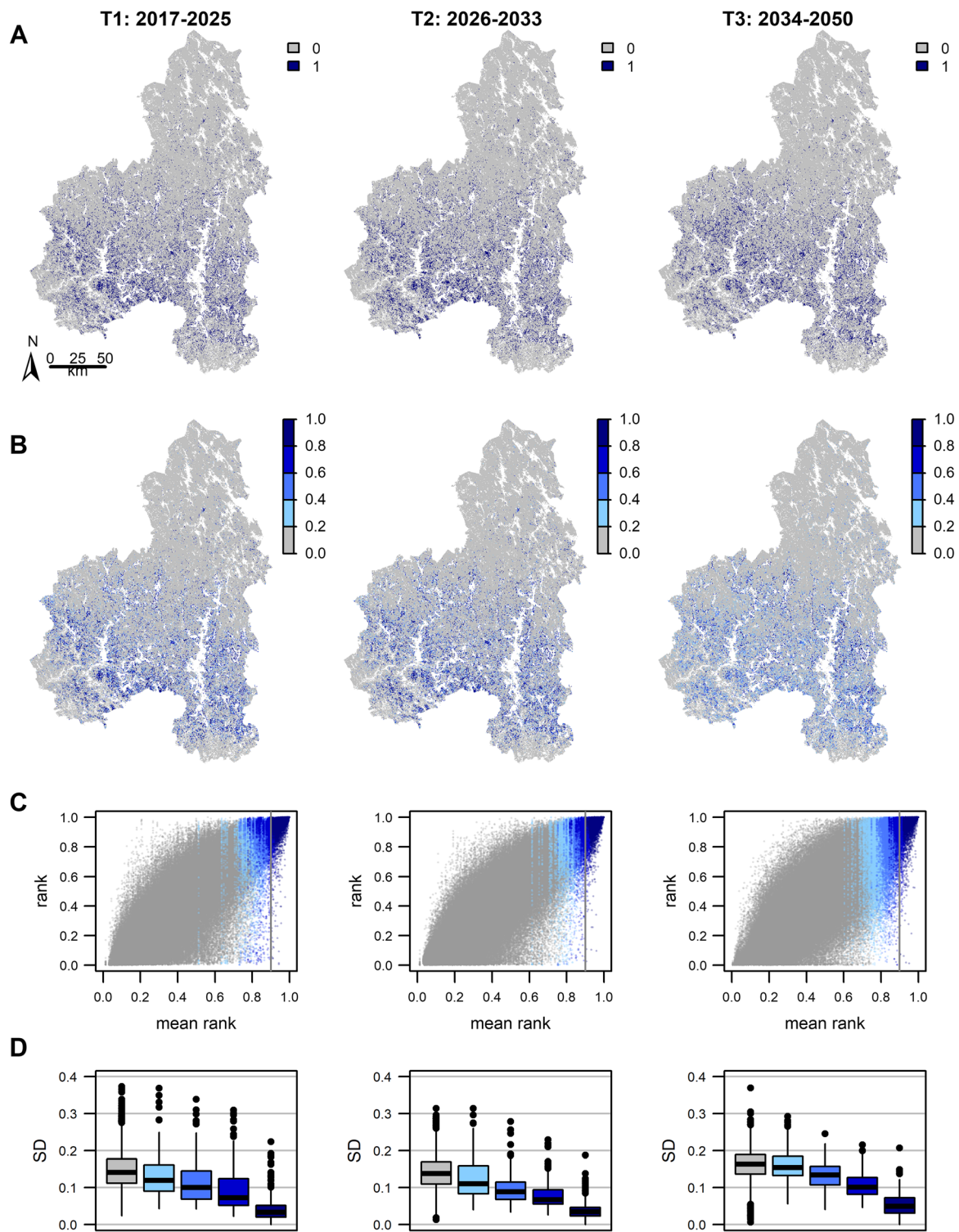
Fig. 4 Priority areas for conservation and their associated uncertainty in each time period. For each time period, a baseline and 50 sample prioritisations were produced using the resampled forest variables and their corresponding carbon and bird distribution maps. Panel **A** shows the top ranked 10% of grid cells (dark blue) for baseline data at each time period. Panel **B** gives the probability of a grid cell to be in the top 10% across the 50 sample prioritisations. The scatterplots (**C**) show the mean (*x*-axis) and range (*y*-axis) in rank values for each grid cell across the sample solutions. Here grid cells have been ordered based on their mean rank values from lowest (left) to highest (right) priority and for each cell the points give the rank value for one iteration. For visibility, these are shown for a subset of 4000 grid cells. The point colours correspond to the probability of the cell to be within the top 10% across samples, as in **B**. The boxplots in panel **D** show the distribution of standard deviations (SD) in the rank values for grid cells. The colours correspond to each of the probability groups shown in **B** and **C**

biodiversity features (Fig. 5A, lower row). This is logical as the prioritisation seeks to rank highest those areas that capture the high-value locations of all input features. Variation was predominantly explained by the three-toed woodpecker, northern goshawk and tree carbon storage on the first time period and became balanced between tree carbon storage and sinks and the three woodpecker species in the last time period. Unlike in the full ranking, the variation in the top ranked 10% grid cells was consistently explained most by the forest structural variables describing the mean height, volume and DBH together with dominant tree age of the forest stands, although the explanatory power was low.

The CCA patterns were consistently similar for prioritisation based only on biodiversity, whereas in the carbon-only prioritisation uncertainty in the ranks were predominantly only explained by carbon sinks (Figs. S7, S8).

DISCUSSION

Our results illustrate the complex nonlinearities between model input and output uncertainties across the different modelling levels (forest variables, carbon and biodiversity features, priority rankings). The seemingly small alterations in the simulated forest variables often resulted in much larger variation in carbon and biodiversity estimates (Fig. 3). To some degree, these are explained by the different ways uncertainty accumulates in these estimates. For example, the large variation around soil carbon is likely driven by the uncertainty in the initial estimate (not tested here), which assumes steady state and is in turn affected by uncertainty around PREBAS parameters and forest management (Junttila et al. 2023). Variation around biomass and carbon stored in trees and ground vegetation mostly stems from the input variable uncertainty, that is, variation in the initial structural forest variables and mean age, which becomes reduced as the simulated forest growth evens out differences between forest stands (Junttila et al. 2023). In contrast, uncertainty around the carbon sink estimate accumulates from multiple submodels of PRE-BAS (Table 1). Unlike input variable uncertainty, the effect of PREBAS parameter uncertainty increases with time as model outputs generated with different parameter vectors become more diverged (Mäkelä et al. 2020; Junttila et al. 2023), increasing uncertainty in the sink estimate. Similarly, despite the perceived certainty around forest variable estimates, the bird habitat estimates become more uncertain with time (Fig. 3) as the combined environmental conditions (forest and non-forest variables) of the region becoming increasingly different from those under which the model was calibrated (Barry and Elith 2006).

On the other hand, the priority rankings of grid cells showed both spatial stability (Figs. 4B, S5) and disproportionally large within-pixel variability (Fig. 4C) in comparison to uncertainty around carbon and biodiversity estimates (Fig. 3). The latter is explained by the way values within each biodiversity and carbon feature are distributed (Fig. S6). For all features, the highest values tend to be rare. Most of the grid cells, therefore, have low (biodiversity) or intermediate (carbon) values and their relative conservation importance becomes similar. Hence, even modest alterations in the raw carbon and biodiversity values can drastically change the priority of these middle-low ranked cells (Kujala et al. 2018b). In contrast, the few grid cells with the highest biodiversity and carbon values constrain the top rankings to these locations. As the sample iterations do not significantly alter the shape of the value distribution, the top ranked grid cells remain more stable. Finally, the study area has some 3.8 million grid cells: even very large variations in single pixel rankings (Fig. 4C) do not necessarily translate into changes in the overall spatial patterns (Fig. S5). These observations show that the connections between input data uncertainty (forest attributes, carbon and biodiversity patterns) and decision uncertainty (priority rankings) are complex, and that care must be taken in interpreting how one influences or is driven by the other.

The need to reduce uncertainty around input data depends on how it affects the decision being made (Howard 1966). Refining estimates of those input data that explain most of the variation in priority ranks is the most logical starting point, and in our study case, the uncertainty around placement of protected areas (top 10% ranked sites) would be most reduced by improving our understanding on the distribution of the three-toed woodpecker, the northern goshawk, tree carbon storage and carbon sinks (Fig. 5). A more detailed uncertainty analysis of the different modelling approaches, such as done by Junttila et al. (2023) for PREBAS, can reveal strategies for how to best achieve this. However, our results reveal also other important observations. First, which input data introduces most of the uncertainty in the priority solution depends on what fraction of the landscape we focus on (Fig. 5). Second, high variability in predicted values or differences in model fit (AUC, Table 2) indicated poorly which input feature introduced most uncertainty in the priority ranking (Fig. S9). Instead, the spatial rarity, nestedness and co-occurrence with other input features explains how much of the ranking is driven by an input variable (Kujala et al. 2018b). For example, as tree carbon storage, common buzzard and honey buzzard became spatially more common with time (Figs. 3, S1, S3), there were more equally good options to protect them and they explained less of the ranking variability (Fig. 5a). The habitat of the three-toed
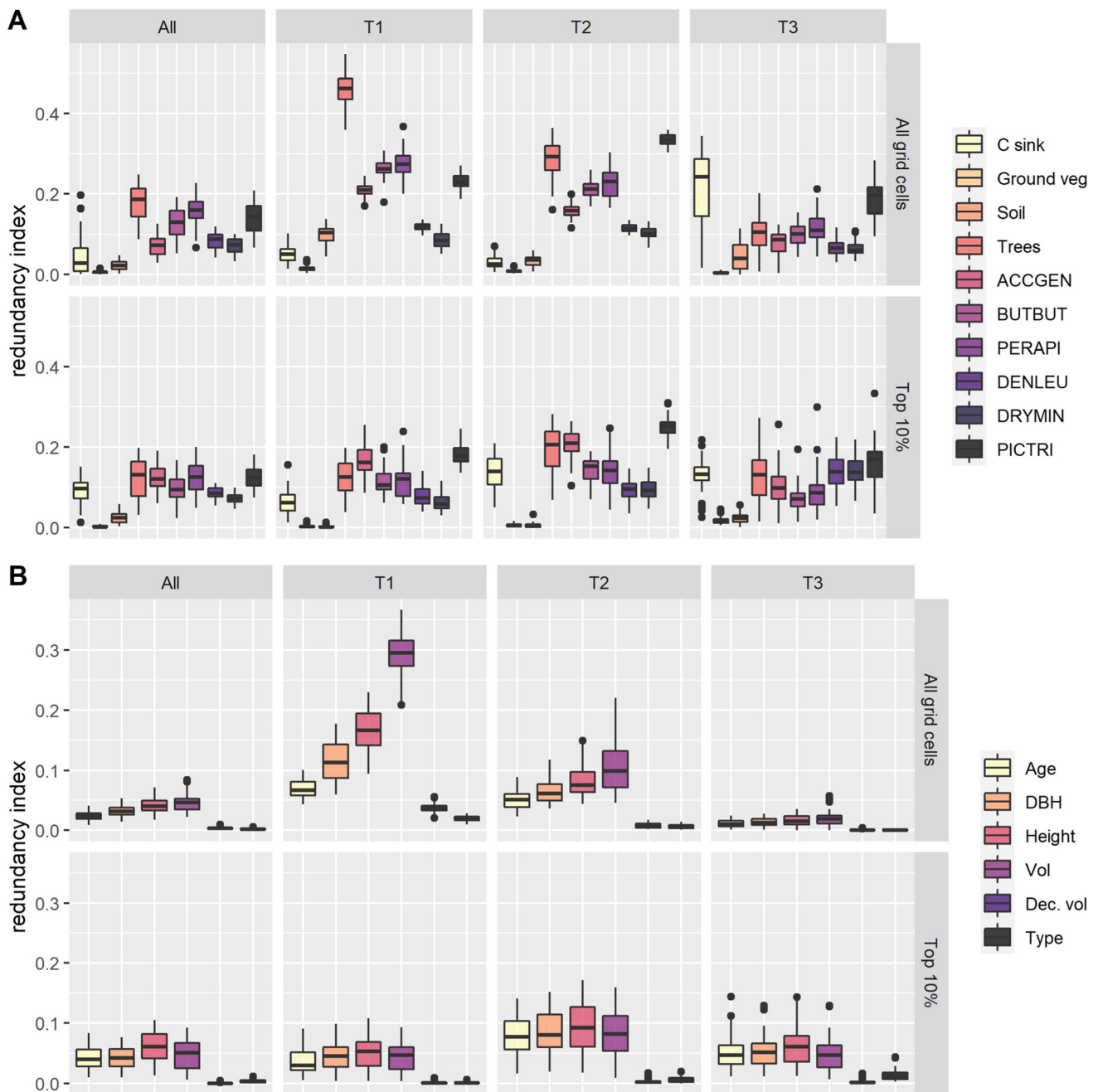
Fig. 5 Redundancy indices of different modelled input variables from the canonical correlation analysis. Higher values of the redundancy index mean the feature explains more of the variation in the priority rankings. The boxplots show how different input features (A. carbon and biodiversity, B. forest variables) explain the variance in the priority rankings of all grid cells (upper row) and in the top 10% ranked grid cells (lower row) and how much this varies across the 50 sample iterations (boxplot). The results are broken down for each time period (T1, T2, T3). The leftmost panel ("All") shows the canonical correlation across all time periods. The redundancy index values are comparable between columns, but not between rows, as these measure correlations between different sets of input features and priority solutions (see text). *Ground veg* ground vegetation carbon, *ACCGEN* northern goshawk, *BUTBUT* common buzzard, *PERAPI* European honey buzzard, *DENLEU* white-backed woodpecker, *DRYMIN* lesser-spotted woodpecker, *PICTRI* Eurasian three-toed woodpecker, *DBH* tree diameter at breast height, *Dec. vol* volume of deciduous trees, *Type* forest site type (fertility). Note that the scales of the y-axis differ across panels

woodpecker also increased (Fig. 3) but remained spatially confined (Fig. S2) and retained its high influence on the ranks. In contrast, the decrease in carbon sinks (Figs. 3, S3) increased its influence on the priority ranks. Carbon sinks

also tend to have different spatial patterns in comparison to stored carbon and biodiversity (Fig. S6), as NEP is highest in relatively young, fast growing forest stands (Minunno et al. 2019). Since the spatial prioritisation aims to capture

high-value areas of all input features, one feature with a strongly dissimilar distribution from others can influence the ranking more (Kujala et al. 2018b). This effect varied greatly between sample iterations (Fig. 5a), even though the increase in uncertainty around carbon sinks was only moderate.

Our work focused solely on input variable uncertainty. Most notably, we did not account for uncertainty arising from forest management actions and future climate change, which can have an even larger impact on forest and biodiversity estimates (Buisson et al. 2010; Mäkelä et al. 2020). Although this was necessary to single out the impacts of target uncertainties, real-world conservation planning would rarely happen in isolation from these threats, and our results particularly for future time steps need to be interpreted with this in mind. However, observing the changes in forests in the absence of harvesting allowed us to better understand why certain forest, carbon and biodiversity values become such important drivers of the spatial prioritisation solution.

Our results also need to be interpreted in light of the made modelling choices and their potential limitations. Although the area on which the original bird species distribution models (SDMs, Virkkala et al. 2022) were built overlaps with ours, and thus, the risk of model extrapolating is small, the transferability of these types of models can be poor (Barry and Elith 2006). We also focused on just one region and used one type of forest model, SDM method and spatial optimisation. The benefit of the chosen spatial prioritisation algorithm is that it does not require pre-defined conservation targets (such as percentage of area protected) to be set for each feature, which would add an arbitrary constraint and reduce the transparency and interpretability of the results. Nevertheless, there are many other options for modelling tree growth, carbon fluxes and species distribution patterns, and the selection of modelling method is a known source of uncertainty in itself (Buisson et al. 2010; Mäkelä et al. 2020). Including all these factors would likely result in higher uncertainty in the priority areas and carbon and biodiversity estimates than shown here.

CONCLUSION

Input data uncertainty may lead to misguided spatial plans and decisions, resulting in excessive spending of resources or the neglect of targeted values in areas mistaken as unimportant. Although the conservation priority of individual grid cells was highly sensitive to input data uncertainty, the spatial patterns remained rather stable. In particular, the top priority areas were least sensitive, giving some confidence that spatial prioritisation-based conservation plans can be robust against data uncertainty.

We did not find evidence that forest conservation strategies would be more sensitive to uncertainties in either biodiversity or carbon data. The amount of uncertainty in feature estimates themselves, whether measured by variability in predicted values or model fit (AUC), was a poor indicator of which the input features introduced most variation in the priority rankings. Rather, their impact was explained by the combination of what fraction of the forests was being considered and the spatial rarity and co-occurrence of the input features. Thus, even small estimate variation in a very influential input feature (e.g. tree carbon storage or three-toed woodpecker at time period T1) could have large effect on the location of priority areas and vice versa (e.g. common buzzard at T3).

The presented approach allows identifying input features that most strongly define the conservation priority of forest stands, and which are less important to account for. This helps us to understand the relative importance of data uncertainties and how to most effectively reduce it when improving the accuracy of spatial prioritisation-based conservation plans.

Declarations

Conflict of interest The authors declare no conflict of interest.

REFERENCES

Barry, S., and J. Elith. 2006. Error and uncertainty in habitat models. *Journal of Applied Ecology* 43: 413–423. https://doi.org/10.1111/j.1365-2664.2006.01136.x.

Buisson, L., W. Thuiller, N. Casajus, S. Lek, and G. Grenouillet. 2010. Uncertainty in ensemble forecasting of species distribution. *Global Change Biology* 16: 1145–1157. https://doi.org/10.1111/j.1365-2486.2009.02000.x.

Burgas, D., P. Byholm, and T. Parkkima. 2014. Raptors as surrogates of biodiversity along a landscape gradient. *Journal of Applied Ecology* 51: 786–794. https://doi.org/10.1111/1365-2664.12229.

Fielding, A.H., and J.F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24: 38–49. https://doi.org/10.1017/S0376892997000088.

Forsius, M., H. Kujala, F. Minunno, M. Holmberg, N. Leikola, N. Mikkonen, I. Autio, V.-V. Paunu, et al. 2021. Developing a spatially explicit modelling and evaluation framework for integrated carbon sequestration and biodiversity conservation: Application in southern Finland. *Science of The Total Environment* 775: 145847. https://doi.org/10.1016/j.scitotenv.2021.145847.

Gauthier, S., P. Bernier, T. Kuuluvainen, A.Z. Shvidenko, and D.G. Schepaschenko. 2015. Boreal forest health and global change. *Science* 349: 819–822. https://doi.org/10.1126/science.aaa9092.

Goulden, M.L., A.M.S. Mcmillan, G.C. Winston, A.V. Rocha, K.L. Manies, J.W. Harden, and B.P. Bond-Lamberty. 2011. Patterns of NPP, GPP, respiration, and NEP during boreal forest succession. *Global Change Biology* 17: 855–871. https://doi.org/10.1111/j.1365-2486.2010.02274.x.

Haakana, M., S. Tuominen, J. Heikkinen, M. Peltoniemi, and A. Lehtonen. 2022. Spatial patterns of biomass change across Finland in 2009–2015. *BioRxiv*. https://doi.org/10.1101/2022.02.15.480479.

Howard, R.A. 1966. Information value theory. *IEEE Transactions on Systems Science and Cybernetics* 2: 22–26. https://doi.org/10.1109/TSSC.1966.300074.

Hyvärinen, E., A. Juslén, E. Kemppainen, A. Uddström, and U.-M. Liukko. 2019. *The Red List of Finnish Species 2019*. Ympäristöministeriö & Suomen ympäristökeskus.

IPBES. 2019. Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (Version 1). Zenodo. https://doi.org/10.5281/ZENODO.3831673.

Jung, M., A. Arnell, X. de Lamo, S. García-Rangel, M. Lewis, J. Mark, C. Merow, L. Miles, et al. 2021. Areas of global importance for conserving terrestrial biodiversity, carbon and water. *Nature Ecology & Evolution* 5: 1499–1509. https://doi.org/10.1038/s41559-021-01528-7.

Junttila, V., and T. Kauranne. 2018. Distribution statistics preserving post-processing method with plot level uncertainty analysis for remotely sensed data-based forest inventory predictions. *Remote Sensing* 10: 1677. https://doi.org/10.3390/rs10111677.

Junttila, V., F. Minunno, M. Peltoniemi, M. Forsius, A. Akujärvi, P. Ojanen, and A. Mäkelä. 2023. Quantification of forest carbon flux and stock uncertainties under climate change and their use in regionally explicit decision making: Case study in Finland. *Ambio*. https://doi.org/10.1007/s13280-023-01906-4.

Kniivilä, M., J. Hantula, J.-P. Hotanen, J. Hynynen, H. Hänninen, K. T. Korhonen, J. Leppänen, M. Melin, et al. 2020. *Metsälain ja metsätuholain muutosten arviointi*. 3/2020. Luonnonvara- ja biotalouden tutkimus. Helsinki: Luonnonvarakeskus.

Kolari, P., J. Pumpanen, Ü. Rannik, H. Ilvesniemi, P. Hari, and F. Berninger. 2004. Carbon balance of different aged Scots pine forests in Southern Finland. *Global Change Biology* 10: 1106–1119. https://doi.org/10.1111/j.1529-8817.2003.00797.x.

Kontula, T., and A. Raunio. 2019. *Threatened habitat types in Finland 2018—Red List of habitats results and basis for assessment*. 2/2019. The Finnish Environment. Helsinki.

Korhonen, K.T. 2021. Forests of Finland 2014–2018 and their development 1921–2018. *Silva Fennica* 55: 49. https://doi.org/10.14214/sf.10662.

Kujala, H., J.J. Lahoz-Monfort, J. Elith, and A. Moilanen. 2018a. Not all data are equal: Influence of data type and amount in spatial conservation prioritisation. Edited by Andres Lopez-Sepulcre. *Methods in Ecology and Evolution* 9: 2249–2261. https://doi.org/10.1111/2041-210X.13084.

Kujala, H., A. Moilanen, and A. Gordon. 2018b. Spatial characteristics of species distributions as drivers in conservation prioritization. Edited by Justin Travis. *Methods in Ecology and Evolution* 9: 1121–1132. https://doi.org/10.1111/2041-210X.12939.

Mäkelä, J., F. Minunno, T. Aalto, A. Mäkelä, T. Markkanen, and M. Peltoniemi. 2020. Sensitivity of 21st century simulated ecosystem indicators to model parameters, prescribed climate drivers, RCP scenarios and forest management actions for two Finnish boreal forest sites. *Biogeosciences* 17: 2681–2700. https://doi.org/10.5194/bg-17-2681-2020.

Miettinen, J., S. Carlier, L. Häme, A. Mäkelä, F. Minunno, J. Penttilä, J. Pisl, J. Rasinmäki, et al. 2021. Demonstration of large area forest volume and primary production estimation approach based on Sentinel-2 imagery and process based ecosystem modelling. *International Journal of Remote Sensing* 42: 9467–9489. https://doi.org/10.1080/01431161.2021.1998715.

Minunno, F., M. Peltoniemi, S. Härkönen, T. Kalliokoski, H. Makinen, and A. Mäkelä. 2019. Bayesian calibration of a carbon balance model PREBAS using data from permanent growth experiments and national forest inventory. *Forest Ecology and Management* 440: 208–257. https://doi.org/10.1016/j.foreco.2019.02.041.

Minunno, F., M. Peltoniemi, S. Launiainen, M. Aurela, A. Lindroth, A. Lohila, I. Mammarella, K. Minkkinen, et al. 2016. Calibration and validation of a semi-empirical flux ecosystem model for coniferous forests in the Boreal region. *Ecological Modelling* 341: 37–52. https://doi.org/10.1016/j.ecolmodel.2016.09.020.

Moilanen, A., P. Lehtinen, I. Kohonen, J. Jalkanen, E.A. Virtanen, and H. Kujala. 2022. Novel methods for spatial prioritization with applications in conservation, land use planning and ecological impact avoidance. *Methods in Ecology and Evolution* 13: 1062–1072. https://doi.org/10.1111/2041-210X.13819.

Moilanen, A., K.A. Wilson, and H. Possingham, eds. 2009. *Spatial conservation prioritization: Quantitative methods and computational tools*. Illustrated. Oxford: Oxford University Press.

Mönkkönen, M., T. Aakala, and C. Blattert. 2022. More wood but less biodiversity in forests in Finland: A historical evaluation: 12.

Pan, Y., R.A. Birdsey, J. Fang, R. Houghton, P.E. Kauppi, W.A. Kurz, O.L. Phillips, A. Shvidenko, et al. 2011. A large and persistent carbon sink in the world's forests. *Science* 333: 988–993. https://doi.org/10.1126/science.1201609.

Peltoniemi, M., M. Pulkkinen, M. Aurela, J. Pumpanen, P. Kolari, and A. Mäkelä. 2015. A semi-empirical model of boreal-forest gross primary production, evapotranspiration, and soil water—calibration and sensitivity analysis. *Boreal Environmental Research* 20: 51–171.

Phillips, S.J., R.P. Anderson, and R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026.

Phillips, S.J., M. Dudík, J. Elith, C.H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2009. Sample selection bias and

presence-only distribution models: Implications for background and pseudo-absence data. *Ecological Applications* 19: 181–197. https://doi.org/10.1890/07-2153.1.

Pukkala, T. 2018. Carbon forestry is surprising. *Forest Ecosystems* 5: 11. https://doi.org/10.1186/s40663-018-0131-5.

Roberge, J.-M., and P. Angelstam. 2006. Indicator species among resident forest birds—A cross-regional evaluation in northern Europe. *Biological Conservation* 130: 134–147. https://doi.org/10.1016/j.biocon.2005.12.008.

Rosenzweig, M.L. 1995. *Species diversity in space and time*, 1st ed. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511623387.

Soimakallio, S., L. Saikku, L. Valsta, and K. Pingoud. 2016. Climate change mitigation challenge for wood utilization—the case of Finland. *Environmental Science & Technology* 50: 5127–5134. https://doi.org/10.1021/acs.est.6b00122.

Stewart, D., and W. Love. 1968. A general canonical correlation index. *Psychological Bulletin* 70: 160–163. https://doi.org/10.1037/h0026143.

Tomppo, E., H. Olsson, G. Ståhl, M. Nilsson, O. Hagner, and M. Katila. 2008. Combining national forest inventory field plots and remote sensing data for forest databases. *Remote Sensing of Environment* 112: 1982–1999. https://doi.org/10.1016/j.rse.2007.03.032.

Vaahtera, E., T. Niinistö, A. Peltola, M. Räty, T. Sauvula-Seppälä, J. Torvelainen, and E. Uotila. 2021. *Metsätilastollinen vuosikirja—Finnish Statistical Yearbook of Forestry 2021*. Luonnonvarakeskus.

Väisänen, R.A., E. Lammi, and P. Koskimies. 1998. *Muuttuva pesimälinnusto—Distribution, numbers and population changes of finnish breeding birds*. Helsinki: Otava (in Finnish with an English summary).

Valentine, H.T., and A. Mäkelä. 2005. Bridging process-based and empirical approaches to modeling tree growth. *Tree Physiology* 25: 769–779. https://doi.org/10.1093/treephys/25.7.769.

Virkkala, R., N. Leikola, H. Kujala, S. Kivinen, P. Hurskainen, S. Kuusela, J. Valkama, and R.K. Heikkinen. 2022. Developing fine-grained nationwide predictions of valuable forests using biodiversity indicator bird species. *Ecological Applications* 32: e2505. https://doi.org/10.1002/eap.2505.

Viskari, T., M. Laine, L. Kulmala, J. Mäkelä, I. Fer, and J. Liski. 2020. Improving Yasso15 soil carbon model estimates with ensemble adjustment Kalman filter state data assimilation. *Geoscientific Model Development* 13: 5959–5971. https://doi.org/10.5194/gmd-13-5959-2020.

van den Wollenberg, A.L. 1977. Redundancy analysis an alternative for canonical correlation analysis. *Psychometrika* 42: 207–219. https://doi.org/10.1007/BF02294050.

AUTHOR BIOGRAPHIES

Heini Kujala (✉) is a University Researcher at the Finnish Natural History Museum, University of Helsinki. Her research interests include spatial conservation planning, biodiversity modelling and decision making under uncertainty.
Address: Finnish Natural History Museum, University of Helsinki, (Pohjoinen Rautatiekatu 13), P.O. Box 17, 00014 Helsinki, Finland.
e-mail: heini.kujala@helsinki.fi

Francesco Minunno is a post-doc at the Department of Forest Sciences, University of Helsinki. His research work focuses on modelling the growth and the carbon balance of forests, integrating data and models in order to quantify uncertainties.
Address: Department of Forest Science, University of Helsinki, P.O. Box 27, 00014 Helsinki, Finland.
e-mail: francesco.minunno@helsinki.fi

Virpi Junttila is a senior research scientist at the Finnish Environment Institute. Her research interests include modelling of ecosystems, especially forests and uncertainty analysis.
Address: Finnish Environment Institute Syke, Latokartanonkaari 11, 00790 Helsinki, Finland.
e-mail: virpi.junttila@syke.fi

Ninni Mikkonen is a coordinator and spatial prioritisation specialist at the Finnish Environment Institute, and a doctoral candidate at the University of Helsinki. Her research interests include data concerning conservation biology of threatened species and habitats, and ecological decision making especially in forest environments.
Address: Finnish Environment Institute Syke, Latokartanonkaari 11, 00790 Helsinki, Finland.
Address: Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland.
e-mail: ninni.mikkonen@syke.fi

Annikki Mäkelä is a Research Director at the University of Helsinki, Institute for Atmospheric and Earth System Research (INAR). Her work focuses on process-based modelling of forest growth and material balances, and model applications to forest management under climate change.
Address: Department of Forest Science, University of Helsinki, P.O. Box 27, 00014 Helsinki, Finland.
e-mail: annikki.makela@helsinki.fi

Raimo Virkkala is a Leading Researcher at the Finnish Environment Institute. His research work focuses on the effects of land-use and climate change on species and habitats and significance of protected area network for biodiversity.
Address: Finnish Environment Institute Syke, Latokartanonkaari 11, 00790 Helsinki, Finland.
e-mail: raimo.virkkala@syke.fi

Anu Akujärvi is a Researcher at the Finnish Environment Institute. Her work focuses on model applications to forest carbon balance under different forest management and climate change scenarios.
Address: Finnish Environment Institute Syke, Latokartanonkaari 11, 00790 Helsinki, Finland.
e-mail: anu.akujarvi@syke.fi

Niko Leikola is a senior research scientist at the Finnish Environment Institute. His research interests include conservation biology of endangered species and habitat types.
Address: Finnish Environment Institute Syke, Latokartanonkaari 11, 00790 Helsinki, Finland.
e-mail: niko.leikola@syke.fi

Risto K. Heikkinen is a Leading Researcher at the Finnish Environment Institute. His research interests include conservation biology of endangered species and habitat types, species distribution modelling and impacts of climate change on biodiversity.
Address: Finnish Environment Institute Syke, Latokartanonkaari 11, 00790 Helsinki, Finland.
e-mail: risto.heikkinen@syke.fi