Deliverable

3.2 Forest state modelling

PathFinder Project

Version: 2.0

13 March 2026

The research leading to these results has received funding from the European Union Horizon Europe (HORIZON) Research & Innovation programme under the Grant Agreement no. 101056907

Executive summary

In this deliverable, models created in WP3.2 are described, focusing on the development of forest growth, management, damage, and biodiversity. The deliverable reports on three sub-tasks: 1) "Forest structure and living biomass carbon", 2) "Resilience against disturbance", and 3) "Forest biodiversity".

In the first subtask, "Forest structure and living biomass carbon", a major focus was to parameterise EFDM activity and transition probabilities required for forecasting based on national forest inventory (NFI) plot data obtained from WP1 and derive the activity probabilities for the 1 km^2 grid from the management options provided by WP3 Task 3.1. The outputs of the EFDM models were used as inputs for running different disturbance models to assess forest "Resilience against disturbance" and to predict the impact of forest management and different climate scenarios on "Forest biodiversity".

The deliverable also reflects on advantages and disadvantages of map-based forecasting systems as they were developed in PathFinder.

Subtask: 1) Forest structure and living biomass carbon:

Preprocessing of the NFI plot data included the formation of pair data from the consecutive NFI plot data measurements and defining thresholds for categorizing the NFI data into different forest management actions of final felling (ff), thinning (thin) and no-management (noman). Forest management intensity information was obtained from WP3 task 3.1. (CLUMondo). While calculation of EFDM action probabilities, and transition probabilities for ff and thin, was fairly straightforward, some additional modelling was required to calculate "noman" transition probabilities, as NFI data collected in this study did not contain any data for regenerated stands. Thus regeneration was modelled using FORMIT-M, which contains precalculated data of potential maxGPP in Europe for modelling forest future growth changes. The modelled transition probabilities were calculated using observations on growth and management from NFIs (Austria, Belgium, Czech Republic, Germany, Poland, Ireland, Norway, Sweden, Slovenia and Spain). For developing pan-European parameterisations, regional EFDM probabilities were calculated for North, Central and South Europe.

Climate change and forest management intensity variations across time and space were accounted for by updating respective variables at every time step. To allow soil carbon modelling matching the EFDM scenarios, an approach was developed to estimate both litterfall and harvest residues from EFDM data. These were used as inputs for Yasso, which was also adapted to use climate data matching the EFDM future climate scenarios. Yasso was modified to run with EFDM inputs to ensure a consistent representation of simulated forest management effects between above- and below-ground C pools.

Subtask: 2) Resilience against disturbance

The assessment of the risk to disturbance has been separated into four interconnected components that represent the major hazards to European forestry. Specifically, these are:

1. Bark beetle risk: A new model called IpsR, based on an existing bark beetle damage model, was created in R. The model takes information about stand characteristics, drought intensity

over the last 3 years, and effective degree days to calculate the level of damage to Norway spruce stands
2. Storm damage risk. The existing ForestGALES/fgr model was used for calculating wind damage risk. The model takes stand characteristics, soil type and wind climate information to calculate the probability of wind damage.
3. Drought risk: A new model of drought risk has been developed based on drought mortality in French NFI plots from 2002–2020 for model calibration and using inventory data from Germany to assess model performance.
4. Fire Risk: The existing Forest Fire Weather Index (FWI) and Forest Fire Danger Rating System (CFFDRS) available in R are being used to assess the risk of fire damage.

The predictions of the EFDM model are used as inputs to the different disturbance models together with various climate variables calculated across Europe for the RCP4.5 and RCP8.5 climate scenarios. All calculations are made in the R modelling framework to simplify the modelling process and to ensure all disturbance interactions are incorporated.

Subtask: 3) Forest biodiversity

This subtask was broken down into three tasks:

1. Assessing the impact of different NFI plot designs on accurately assessing indices providing a measure of forest structure
2. Modelling deadwood quantity as a function of various stand, site and climate variable using the harmonized NFI assessments created in the PathFinder project and stored in the nFiesta repository.
3. Developing composite indices to describe forest biodiversity based on outputs from the EFDM model simulations.

I. DOCUMENT CONTROL

Project	PathFinder (101056907)
Project Title	Towards an integrated consistent European LULUCF monitoring and policy pathway assessment framework
Date	13 March 2026
Author/s	Titta Majasalmi (UH), Tam Do (IEFC), Barry Gardiner (IEFC), Iciar Alberdi (INIA), Daniel Moreno (INIA), Johannes Breidenbach (NIBIO), Rasmus Astrup (NIBIO)
Reviewer/s	Thales Pupo West (VUA)
Activity	WP3, Task 3.2. Forest state modelling
Filename	PathFinder_Deliverable_task3.2

DISSEMINATION LEVEL [Indicate the dissemination level by underlining the applicable one]

PU	Public, fully open access
RE	Restricted to a group specified by the PathFinder Consortium (including the Commission Services)
CO	Confidential, only for members of the PathFinder Consortium (including the Commission Services)

II. DOCUMENT HISTORY

Version	Date	Author	Change
1.1	13.08.2025	Titta Majasalmi	
1.2	25.08.2025	Barry Gardiner	Editing, formatting, figure and table numbering
2.0	13.03.2026	Johannes Breidenbach	Revision according to reviewer comments on the advantages and disadvantages of maps as input in forecasting systems.

III. Abbreviations

NIBIO	Norwegian Institute of Bioeconomy Research
ALU	Albert-Ludwigs University Freiburg
IGN	National Institute of Geographic and Forest Information
VUA	Vrije Universiteit Amsterdam
TI	Thünen Institute of Forest Ecosystems
CFRI	Croatian Forest Research Institute
LUKE	Natural Resources Institute Finland
BFW	Federal Research and Training Center for Forests, Natural Hazards and Landscape
GIS	Slovenian Forestry Institute
UHUL	Czech Forest Management Institute
VTT	Technical Research Centre of Finland Ltd.
CSIC	Consejo Superior de Investigaciones Científicas
CICERO	Center for International Climate Research
UGOE	University of Göttingen
UH	University of Helsinki
TM	TreeMetrics
EVINBO	Eigen Vermogen van het Instituut voor Natuur- en Bosonderzoek
ELO	European Landowners Organisation
IEFC	Institut Européen de la Forêt Cultivée
FMI	Finnish Meteorological Institute
WSL	Swiss Federal Research Institute for Forests Snow and Landscape Research
UB	University of Bristol
JRC	Joint Research Center
EEA	Environmental Agency

IV. List of tables

Table 1: PathFinder tree-species groups defined by WP2. 2
Table 2: NFI plot data description. Abbrevations: country = country code, N km2= number of NFI-plot data pairs, plots N per 1 km2 =number of km2 grid cells covered by NFI data, interv. ff (%) = percentage of plots classified as final felled based on NFI plot intervention information, interv. thin (%) = percentage of plots classified as thinned based on NFI plot intervention information, interv. noman (%) = percentage of plots classified as no-managed based on NFI plot intervention information, ff (%), thin(%) and noman (%) show the respective percentage values for countries and inventory periods without NFI plot intervention information, and inventory period = average time in years between consecutive plot inventories. Asterisk (*) denotes older NFI inventory round 3
Table 3: Regional forest management probabilities for 5-yr period. Abbreaviations: country group = NORTH: Sweden and Norway; SOUTH: Spain and Slovakia; CENTRAL: all other countries), ff (% 5y) = 5-yr final felling probability, thin (% 5y) = 5-yr thinning probability, ffThin (% 5y) = sum of ff and thin 5-yr probabilities, noman (% 5y) = 5-yr noman probability. 4
Table 4: Number of NFI plot-pair observations within each region and forest management intensity class (FM). 5
Table 5: EFDM activity probabilities by country_group, FM class and vol_bin. Color bars reflect the relative intensity of the action and are comparably sized within an activity type (i.e., ff, thin and noman). 8
Table 6: Forest regeneration table showing assumed planting density (in FORMIT-M)..... 12
Table 7a and b: The number of NFI plot pairdata observations (‘count’) and statistics for potential maxGPP for the mean NFI inventory year by sp_group and region. Colorbars visualise variations in data between sp_groups and regions. (a) shows the data sorted by sp_group and (b) by region. 15
Table 8: Number of NFI pairdata observations available for each generalized sp_group and P0_class (maxGPP value class). The respective maxGPP quantile values are shown in Table 9 17
Table 9: Potential maxGPPmax value thresholds (class bottom values) for the generalised sp_groups. Possible range of variation was 250-4050 (g C m-2 year-1). 18
Table 10: Area-weighted vol-classes at 80-yr age using constant P0_class values for comparing model predictions between regions, sp_groups and P0_classes. Generalised sp_groups, from which models were assigned to other sp_groups, are shown in bold. Values are shown for all sp_groups to allow for comparing model outcomes by sp_groups..... 19
Table 11: Fitted regression models for above-ground deadwood biomass by region and sp_group. R2 is the coefficient of determination. 22
Table 12: Predicted polynomial regression for stem number..... 23
Table 13a, b and c: Allometric biomass equations by Forrester et al. (2017). a),b),c) show the parameter values and statistics for the selected metamodels. For convenience separate tables are provided. Abbreviations: sp_group is the Pathfinder sp_group number, ‘Species’ and ‘Component’ columns show the original species and components the model was developed, and ‘Equation’, ‘ln(β0)’, ‘β for ln(d)’ and CF, are model information and parameters for Eq.1. The rest of the columns are listed for readers convenience (i.e. quick-look-overview), and thus please see the original publication for the clarification of other column variable acronyms. 25

Table 14: Ratio of below-ground biomass to above-ground biomass from IPCC (2006), Table 4.4, used in this study. For the SOUTH region, modified R value was used instead of the default value. 27
Table 15: Forest area by countries. Abbreviations: Country = country acronym, Country name= name of the country, Area = WP2 map based estimate of forest total area at 1 km2, FAO = forest areal estimate reported by FAO, Dif = absolute difference between (Area – FAO), and Dif (%) = percentage difference of areal values compared to FAO area. FAO statistics from the source were not provided for all countries, and thus those columns are empty. 29
Table 16: Dominant volume classes and sp_groups by countries. Please see section 2.4. for the definition of vol classes (i.e. 0, 1, 10, 25, 48, 80, 122, 204, 313, and 517 m3/ha (the intervals closed on the left and open on the right), and section 2.1 for description of sp_groups. 31
Table 17: Median, mean and standard deviation of the deadwood stocks (m^3 ha^{-1}) and number of plots by country according to the National Forest Inventories. 50
Table 18: Mean and standard error of the deadwood and predictors for the National Forest Inventories plots. 51
Table 19: Minimum input requirements for the fgr wind risk model 71
Table 20: Outputs from the fgr model 71
Table 21: Inputs for IpsR bark beetle damage model. 72
Table 22: Outputs from the IpsR bark beetle damage model. 73
Table 23: Inputs for the fwi and fbp fore models 74
Table 24: Outputs for the cffdrs fire danger rating system 76
Table 25: Inputs for the SPEI drought index and HWMI 77
Table 26: Outputs from the SPEI drought index and HWMI 77

V. List of figures

Figure 1: Activity probabilities of noman, thin and ff by country_group, forest management intensity (FM) and vol_bin (x-axis, 1:10). 9
Figure 2: Potential maxGPP for RCP45 for the year a) 2020, and b) the differences between potential maxGPP values at 2100 and 2020. 14
Figure 3: Workflow graph of calculations and responses in the bark beetle outbreak model in the ORCHIDEE land surface exchange model (Marie et al., 2024) 37
Figure 4: Structure and components of the Canadian Forest Fire Danger Rating System 41
Figure 5: Schematic of the potential interaction between disturbance models and growth models within the PATHFINDER project. 44
Figure 6: Deadwood stocks in the National Forest Inventory plots 50

VI. Contents

I. DOCUMENT CONTROL iii
II. DOCUMENT HISTORY iv
III. Abbreviations v
IV. List of tables vi
V. List of figures vii
VI. Contents ix
1. Introduction 1
1.1 Background 1
1.2 Advantages and disadvantages of maps in forest scenario models 1
2. Forest structure and living biomass carbon 2
2.1 National Forest Inventory plot data pairs 2
2.2 Forest management activities from NFI data 2
2.3 Forest management intensity from CLUMondo 4
2.4 Considerations regarding the EFDM framework 5
2.5 EFDM activity probabilities 6
2.6 EFDM transitions probabilities 9
2.6.1 Forest harvest 9
2.6.2 Forest regeneration using FORMIT-M 10
2.6.3 Considerations of forest growth (noman) modelling 12
2.6.4 Noman 14
2.7 EFDM CoefTable creation 20
2.7.1 Allometric biomass models 24
2.7.2 Litter models 27
2.7.3 Harvest residues 28
2.8 Preprocessing of the forest map data for EFDM 28
2.9. EFDM scenario modelling framework 32
3. Soil carbon pool modelling using Yasso 34
3.1 Linking EFDM and Yasso 34
4. Resilience against disturbance 36
4.1 Modelling bark beetle disturbance 36
4.1.1 Background 36
4.1.2 Method 36
4.2 Wind-storm models 37
4.2.1 Background 37

4.2.2 Methods 38
4.3 Modelling tree mortality influenced by drought 39
4.3.1 Background 39
4.3.2 Methods 39
4.4 Forest Fire Weather Index (FWI) and Forest Fire Danger Rating System (CFFDRS) 40
4.4.1 Background 40
4.4.2 Methods 40
4.5 Climate indices 41
4.5.1 Regional Climate modelling 41
4.5.2 Standardised Precipitation Evaporation Index (SPEI) 42
4.5.3 Effective Degree Days (DDeff) 42
4.5.4 Heatwave Magnitude Index (HWMI) 43
4.6 Model Operation 44
5. Forest biodiversity 45
5.1 Composition and structure 45
5.1.1 Indices integration in EFDM 45
5.1.2. Performance of forest structure-related indices under nested plot designs 45
5.1.2 Results 47
5.1.3 Discussion 48
5.1.4 Conclusions 48
5.2 Deadwood 49
5.2.1 Deadwood National Forest Inventory data and modelling 49
5.2.2 Results 51
5.2.3 Discussion 52
5.2.4 Further steps 52
5.3 Composite indices 53
References 54
ANNEX A. 61
List of NFI data used in parameterising the EFDM: 61
Transition matrices for noman sp_groups (n=10) for the three regions. 62
Future climate impact of P0_class ands sp_group on noman growth under RCP45. 67
Forest disturbance models parameterisation 71
Wind Damage Model 71
Insect Attack Model 72
Fire Risk Model 73
Drought Probability and Impact Model 77

1. Introduction

1.1 Background

The European Forestry Dynamics Model (EFDM) uses a set of activities to describe possible pathways for forest development during each step of a scenario simulation. These activities encompass a wide range of forest dynamics, including traditional management interventions (such as thinning or final felling), natural processes (like unmanaged growth), and possible external disturbances (e.g., windthrow, drought, pests, or snow damage). Each activity is assigned a name, which serves as an identifier in simulation outputs and is used to determine the associated probability of occurrence. Dynamic variables represent the state attributes that are modified by an activity. Commonly affected variables include forest volume and age. However, activities can also influence broader characteristics such as land use, in which case variables related to land-use classification become relevant.

More precisely, EFDM employs a matrix model structure defined by a preset number of fixed states, between which the area (i.e. forest) is moved over the simulation span. Moving areas between the fixed states is determined by dynamic state-variables, which means that all model dynamics in EFDM will result from these changes. The term statespace is used to denote all state variables involved in modelling. In EFDM, at every simulation step, the distribution of area within the statespace is changed by the transition probabilities. Activity probabilities are first used to split each cell of the statespace, and then activity-specific transition probabilities (with activity-specific state variables) are used to move areas in the dynamic statespace. The definition of the statespace variable is a choice that the modeler must make based on available data and formulation of the research problem.

1.2 Advantages and disadvantages of maps in forest scenario models

Well-known forest scenario models for continental scales such as EFISCEN-Space use NFI plot data as input for future scenarios. On the plot-level, uncertainty in input values is present (measurement errors, concentric or angle count plot designs give estimates, not observations), but can be considered very small. Vice-versa, this is the relatively large uncertainty in maps is the biggest disadvantage of using maps as inputs to forest scenario models. The uncertainty in maps will propagate into the forecasts of forest structure. We believe that there is a need for more models working wall2wall but also recognize that this comes with challenges that may not be solved by PathFinder alone. Despite all problems with uncertainties in maps, there are a number of reasons, why methods for wall2wall forecasting should be further investigated:

1) Forest management happens on stand-level, not on plots, and the spatial component in the management is important to consider. Indeed, many forest management planning forecasting systems use maps as input. 2) All Integrated Assessment Models which link to the economy, land use, and climate are map based. The connection to IAMs was specifically requested in the call. 3) Climate models are based on maps and our climate emulator for forest feedback requires maps. 4) Wind risk and bark beetle models depend, directly or indirectly, on landscape-level information. 5) Connectivity as a major component of biodiversity requires spatial data which is hard to model with plot-level information or statistics alone. 6) The FML, as originally proposed, required maps and this project supported the development of the law. 7) When interested in the development of small domains such as protected areas, even NFI data may be sparse and the landscape structure around them can be of importance. 8) Map data are becoming more available than plot-level field data. 9) The maps are method-consistent, especially with respect to time, which is not the case for plots. That is a major issue in plot-based forecasting systems. 10) Not much beats maps in communicating results. PathFinder also gives examples of how to represent uncertainties adequately.

2. Forest structure and living biomass carbon

2.1 National Forest Inventory plot data pairs

National Forest Inventory (NFI) plot data were available for ten European countries (Austria, Belgium, Czech Republic, Germany, Poland, Ireland, Norway, Sweden, Slovenia and Spain).

For each country the preprocessing of the NFI plot data contained the same steps: 1) select plots for which data is available at two consecutive time-steps, 2) save plot-level information 'area_domain_category' label (i.e., "no intervention", "intervention conducted") when available, 3) save plot-level information regarding inventory reference years, 4) combine target variable "vol" data from the two inventories using columns ("country", "cluster", "plot", "cell", "sub_population_category"), 5) calculate plot-level time-gap between inventories, 6) classify tree species into PathFinder tree species groups, which were defined by WP2 to map species diversity and abundance over Europe for 2020 (Table 1), 6) then for each plot, total stem volume was calculated for each PathFinder tree species group, and the dominating tree-species group was selected based on the highest stem volume (total plot volume was assigned to dominating tree-species). The list of raw NFI datasets used in the analysis is listed in **ANNEX A.**

Table 1: PathFinder tree-species groups defined by WP2.

Tree species group number:	Tree species contained by the group:
1	Betula, Carpinus, Fraxinus
2	Populus, Salix
3	Fagus
4	Quercus, Quercus cerris, Quercus lepidobalanus, Quercus robur, Quercus suber
5	Eucalyptus
6	Other broadleaves, Acer, Alnus, Castanea
7	Picea
8	Abies
9	Pinus, Pinus halepensis, Pinus nigra, Pinus pinaster, Pinus sylvestris
10	Other conifers, Larix, Pseudotsuga

2.2 Forest management activities from NFI data

NFI plot data pairs were then split in three activities: 'noman' (i.e. no-management), 'thin' (i.e. various types) and 'ff' (final felling, meaning clearcut harvest) to allow explicit parameterization in EFDM. Forest management activities were calculated using the data column 'area_domain_category' (i.e., "no intervention", "intervention conducted"), and labelled when available. Two approaches were used to split each NFI-plot data pair into noman, thin and ff activities depending on the availability of the availability of the 'area_domain_category' information. If label "no intervention" was available, a plot was assigned to 'noman´, if label was "intervention conducted" a plot was assigned to 'ff' if initial plot volume was >10m3 and the volume reduction between inventories was equal of more than 25% of the initial volume, otherwise it was classified into 'thin'. Information

delivered by 'area_domain_category' label was thus useful to separate plots which had damages from those that had been managed. However, for NFI plot data without 'area_domain_category' information, an approach for separating damaged stands from managed stands was needed. For this, a similar approach was used: if the initial plot volume was >10m^3 and the volume reduction between inventories was equal to or more than 25% of the initial volume, the plot was assigned to 'ff'. If plot was not assigned to 'ff', two different thresholds were used to split data into 'noman' and 'thin' activities: Activity label 'thin' was assigned if initial stem volume was >10m^3, and the removed stem volume was larger or equal to the 25% of the initial stem volume (for Spain 8% threshold was used instead of 25% due to differences in forest structure). NFI plot pair was classified as 'noman' if the previous 'ff' and 'thin' conditions were not met, and thus the resulting stem volume reduction was assumed to be caused by damage contained by 'noman' activity.

The countries from which NFI plot data pairs were available, number of NFI-plot data pairs, number of km2 grid cells covered by NFI data, and relative density of NFI-plot data pairs per 1 km^2 grid cell are listed in Table 2. All possible NFI plot data that was available was used in analyses, and if NFI plot intervention information was available, it was used in all later stages (i.e. estimated intervention information was only used for ES, PL, PL* and SI datasets). Noteworthy is that the average time period between different inventories varied between the two inventories in different countries, and thus the numbers presented in Table 1 are not directly comparable between countries.

Table 2: NFI plot data description. Abbrevations: country = country code, N km2= number of NFI-plot data pairs, plots N per 1 km2 =number of km2 grid cells covered by NFI data, interv. ff (%) = percentage of plots classified as final felled based on NFI plot intervention information, interv. thin (%) = percentage of plots classified as thinned based on NFI plot intervention information, interv. noman (%) = percentage of plots classified as no-managed based on NFI plot intervention information, ff (%), thin(%) and noman (%) show the respective percentage values for countries and inventory periods without NFI plot intervention information, and inventory period = average time in years between consecutive plot inventories. Asterisk (*) denotes older NFI inventory round .

country	N km2	plots N per 1km2	interv. ff (%)	interv. thin (%)	interv. noman (%)	ff (%)	thin (%)	noman (%)	inventory period
AT	1109.0	2.5	1.6	14.4	84.0	1.7	14.8	83.5	6.0
BE	1471.0	1.0	0.3	23.9	75.8	0.4	17.8	81.8	7.2
CZ	6683.0	1.0	2.3	20.2	77.5	2.4	16.0	81.6	5.0
DE	2903.0	2.2	0.3	21.4	78.3	0.3	22.5	77.2	5.0
ES	36938.0	1.0	0.0	0.0	0.0	2.1	6.8	91.1	13.4
ES*	12642.0	1.0	0.3	4.2	95.6	1.0	4.4	94.6	11.2
IE	1372.0	1.0	0.4	4.4	95.2	0.5	6.3	93.2	7.0
NO	11187.0	1.0	1.1	2.5	96.4	1.3	8.1	90.6	5.0
NO*	2221.0	1.0	1.1	2.6	96.3	1.3	7.8	90.9	5.0
PL	8696.0	2.1				0.9	12.6	86.5	5.0
PL*	10143.0	2.1				0.7	20.9	78.4	5.0
SE	9823.0	1.7	1.2	10.2	88.7	1.2	10.8	88.0	5.0
SI	729.0	1.0				1.0	24.1	74.9	6.0

Comparison of percentages assigned into ff, thin and noman activities using intervention information and without it in Table 2, shows that percentage values obtained using simple thresholding of the data fall within possible ranges, and align with literature-based numbers (e.g., Suvanto et al., 2025).

For example, in Table 2, percentages assigned into ff, thin and noman activities using intervention information and without it**,** produced fairly similar percentages for thin, ff and noman for Germany, Austria and Sweden, while for Spain, slightly higher forest management intensities were obtained using the approach without the intervention information.

For the North region (covering Norway, Sweden and Finland), forest management is active and final felling is more frequent than in Central and in South EU, where thinning is more common (Suvanto et al., 2025). South Europe (e.g., Spain) is characterised by areas of low and high total harvest rate, and has clear regional variations on both the frequency of harvest events and intensity of harvest events due to diversity in tree species abundance and harvest practices. Overall, in Spain, thinning where less than 25% of the basal area are harvested, are the most common (Suvanto et a., 2025). In Poland, the total harvest rate is similar to Central Europe, but the frequency of harvest events is higher than in Central EU, yet intensity of harvest events is low, indicating more frequent thinning compared to other EU countries, which aligns with Table 1 percentage values for ff and thin activities.

Then, NFI plot data from different countries were assigned to three country groups (NORTH=Sweden and Norway, SOUTH = Spain and Slovakia, and CENTRAL = all other countries) and for each country, the ff, thin and noman percentages were scaled to correspond 5yr inventory period. Forest management probabilities were calculated using intervention information (i.e. columns 'interv. ff (%)', 'interv. thin (%)' and 'interv. noman (%)' in Table 2) when that data was available. For the NFI plot pair datasets without intervention information (i.e. ES*, PL, PL and SI, in Table 2), the respective values were obtained using the alternative thresholding approach (i.e. columns: 'ff (%)', 'thin (%)', and 'noman (%)' in Table 2).

Regional forest management probabilities (Table 3) show that the probability of final felling increases (doubles) from South to North, whereas Central Europe has the highest probability for thinning. These numbers correspond with spatial variations and frequency of harvest events presented in the paper by Suvanto et al. (2025, Figure 2).

Table 3: Regional forest management probabilities for 5-yr period. Abbreaviations: country group = NORTH: Sweden and Norway; SOUTH: Spain and Slovakia; CENTRAL: all other countries), ff (% 5y) = 5-yr final felling probability, thin (% 5y) = 5-yr thinning probability, ffThin (% 5y) = sum of ff and thin 5-yr probabilities, noman (% 5y) = 5-yr noman probability.

country_group	ff (%, 5y)	thin (%, 5y)	ffThin (%, 5y)	noman (%, 5y)
SOUTH	0.6	2.5	3.1	96.9
CENTRAL	1.0	17.0	18.0	82.0
NORTH	1.2	5.7	6.9	93.1

Given the task to develop regional pan-European harvest models, some simplifications were necessary to harmonise the description/definition of forest management actions across different countries. In practice, this means that all various types of harvests were classified into ff, thin and noman activities. Also, as NFI plot data pairs were not available from all countries, and the number of NFI plot data pair observations varies between countries, the resulting regional averages were the most impacted by countries from which more NFI plot data pairs were available.

2.3 Forest management intensity from CLUMondo

Forest management intensity classes for the year 2020 were obtained from WP3 task 3.1, which classifies European forest areas into five distinct classes of 1) primary forest, 2) close-to-nature forestry, 3) Combined-objective forestry, 4) Intensive forestry, and 5) very Intensive forestry. The

spatial resolution of the map product is 1 km^2. Later, forest management intensity classes will be referred to as 'FM' classes.

Forest management intensity effects on forest management activities were obtained by extracting FM classes for each NFI plot data pair observation (Table 4). Results show that few forest plots were located in FM class 1, which corresponds only noman. The largest number of NFI plots were classified into FM classes of 2-4. In Central and South Europe, most NFI plots fell into FM class 3, whereas in North region forest areas were classified equally to FM classes 3 and 4; indicating more intensive forestry practices in North region. In South region, the number of NFI plot data pairs classified as Very intensive forestry practices (i.e., FM class 5) was notably higher than in North and Central Europe, which also reflects regional differentiation in forest management intensity (e.g., eucalyptus plantation locations).

Table 4: Number of NFI plot-pair observations within each region and forest management intensity class (FM).

FM	NORTH	CENTR	SOUTH
1	71	29	20
2	5427	8086	13962
3	8108	17331	26509
4	8108	5834	6304
5	1517	1097	3514

2.4 Considerations regarding the EFDM framework

After preprocessing of the NFI data and obtaining forest management intensity information from CLUMondo, we could move on to defining the EFDM simulation framework.

Based on available data, the stand total stem volume (vol) was selected as the dynamic statespace variable ('vol0' is the from-class and 'vol1' is the to-class). Other statespace variables were: tree species group (sp_group), geographic region (country_group), climate change indicator (P0_class) and forest management intensity (FM_class). The size of the EFDM spacespace is defined by the number of classes of each variable (i.e. vol=10, sp_group=10, country_group=3, P0_class=6 and FM_class=5; 10 × 10 × 3 × 6 × 5 = 9,000).

Other statespace variables were selected based on their crucial role for quantifying forest growth and forest management practices: For example, we decided to include as many sp_groups as possible in EFDM, as we know forest growth and management vary between different species. Separating region differences in forest management practices required including statespace variable country_group. Quantification of climate change impact on growth demanded inclusion of a statespace variable, P0_class, which is able to characterise the strength and direction of the climate impact on growth. Accounting for future changes in forest management intensity levels across the EU region demanded the incorporation of statespace variable FM_class.

The binning of the dynamic vol variable was determined based on initial NFI plot data pair total stem volumes using quantiles of 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.65, 0.8, and 0.95. The respective stem volume class lower limits were set as: 0, 1, 10, 25, 48, 80, 122, 204, 313, and 517 m^3/ha (the intervals closed on the left and open on the right). While the determination of the number and thresholding of

the vol_bins for the matrix model is not an easy task, and we could have used a larger number of vol_bins to better reflect regional growth patterns of a forest, we decided to take a pan-European approach and use the same vol_bins across EU regions. The number of vol_bins is not particularly large, but may be assumed to be sufficient for characterising forest growth across the EU, as it is based on quantiles from existing NFI plot data.

Other aspect supporting using rather lower than higher number of vol_bins in EFDM was that as we need to extract the climate change impact from another data source which must be classified, using larger number of vol_bins would have resulted sparse transition probability matrices for P0_class (climate change impact), and using lower number of P0_classes would have resulted lesser sensitivity to climate change impacts on growth. Thus, the selection of the number of vol_bins and P0_classes was a bit of a trade-off between structural description and climate sensitivity of the model. Since our analysis did not contain any secondary dynamic variables (such as forest age) and vol_bin interval was not dense enough for this purpose, we could not define a prior-term to EFDM to modify dynamic variables in the case of missing data.

Scenario development for EFDM was decided to be implemented by running EFDM one time step (5-yr) at a time, and by updating the P0_class and FM_class data between every 5-yr. This approach also allows for extracting harvest residues from the harvests to be used as an input for Yasso. Thus, our approach for running EFDM differs from the previous EFDM set-ups, and demonstrates a new approach for using a matrix model in a context where inputs from other models are sequentially fed into EFDM (instead of the common approach for running EFDM using static parameters). We used the newest EFDM version that runs on R (Kuronen and Räty, 2023).

2.5 EFDM activity probabilities

To calculate EFDM activity probabilities for ff, thin and noman, the initial stem volumes were classified into the volume bins, as only the initial stem volume will be used to define action probabilities. The action probabilities were calculated as means of ff, thin and noman NFI plot pairdata properties, and averages were calculated by country_group, FM class and vol_bin. Also, mean NFI inventory period years were calculated, to scale the ff and thin probabilities to correspond with 5-yr EFDM period. This was achieved by dividing the mean ff and thin probabilities, calculated by region, FM and vol_bin, with the respective means of the inventory period, and multiplying the resulting annual activity probabilities by 5 to obtain 5-yr activity probabilities. Activity probabilities of ff and thin were summed, by region, FM and vol_bin, and respective noman probabilities were calculated as one minus the ff and thin probability sum. Action probabilities were assumed independent of sp_group and P0 class, thus statespace was defined by vol_bins (n=10), FM classes (n=5) and country_groups (n=3). Calculated action probabilities were merged with statespace to make sure probabilities are provided for all possible parameterisation options. For each row containing ff, thin and noman probabilities, the summation of these probabilities was confirmed to be 1. For rows that had NA values (i.e. did not NFI pair data based on activity probabilities), the noman was set as 1, and ff and thin as 0, respectively.

Some additional sensitivity checks were done to make sure activity probabilities are sensible (i.e., NFI data and FM data may contains errors): It was assumed that for the two volume bins (i.e., stem volume less than 10 m3/ha), only noman was possible (i.e. set to 1, and ff and thin to 0, respectively). Also, It was assumed that for FM class 1, only noman activity was possible (i.e., set to 1, and ff and

thin to 0, respectively). In addition, to avoid problems caused on lack of representative NFI-plot data observations to derive activity probabilities and misclassified FM class observations, some addition rules were set to make sure resulting parameterizations show increasing forest management intensity systematically across country_groups and vol_bins: in the case of missing data for FM 5 class (which had few NFI pair data observations available), the NA activity probabilities were gapfilled using the respective country_groups and respective vol_bins maximum ff and thin activity probabilities (i.e. across FM classes). Also, to make sure FM classes will predict increasing forest management activity probabilities, it was assumed that for each vol_bin, the noman activity must stay at the same level or increase with the increasing FM class. The country_group and vol_bin specific activity probability values, which failed to demonstrate increasing forest management intensity were gapfilled using lower (or, if it made a smaller impact, the higher) FM level ff and thin activity probabilities. For example, for North region only one such correction was done, and that was for FM class 3 vol_bin 10, which was replaced by FM class 2 vol_bin 10 activity probabilities for ff and thin. For Central Europe, a total of 7 activity probability replacements were done. Two activity value replacements for FM class 3 and 5 for FM class 5. For the South region, two activity replacements were done for FM class 4 vol bins 9 and 10.

Finally, the noman activity probability value was recalculated for the plots where activity probabilities for ff and thin were manually changed. Overall, differences in activity probabilities for ff and thin were small due to manual changes. Need for these changes arises from the fact that no data is perfect, and averages are influenced by both lack of data and variations in underlying ff and thin values. Making sure forest management activity probabilities of ff and thin will increase systematically with increasing FM class is important for the modelling perspective, which relies on the EFDM model's ability to quantify the direction and strength of the forest management activity levels with increasing FM class.

Table 5: EFDM activity probabilities by country_group, FM class and vol_bin. Color bars reflect the relative intensity of the action and are comparably sized within an activity type (i.e., ff, thin and noman).

		NORTH			CENTR			SOUTH		
Vol_bin	FM	noman	thin	ff	noman	thin	ff	noman	thin	ff
1	1	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
1	2	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
1	3	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
1	4	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
1	5	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
2	1	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
2	2	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
2	3	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
2	4	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
2	5	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
3	1	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
3	2	0.99	0.01	0.00	0.96	0.04	0.00	0.99	0.01	0.00
3	3	0.96	0.03	0.01	0.95	0.04	0.01	0.98	0.02	0.00
3	4	0.95	0.04	0.01	0.94	0.05	0.01	0.97	0.03	0.00
3	5	0.95	0.04	0.01	0.92	0.05	0.03	0.95	0.04	0.01
4	1	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
4	2	0.99	0.01	0.00	0.95	0.05	0.00	0.99	0.01	0.00
4	3	0.97	0.02	0.01	0.92	0.06	0.02	0.98	0.02	0.00
4	4	0.95	0.04	0.01	0.92	0.06	0.02	0.96	0.03	0.01
4	5	0.92	0.07	0.01	0.89	0.06	0.05	0.94	0.04	0.02
5	1	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
5	2	0.98	0.02	0.00	0.91	0.08	0.01	0.98	0.02	0.00
5	3	0.94	0.05	0.01	0.91	0.08	0.01	0.96	0.03	0.01
5	4	0.94	0.05	0.01	0.91	0.08	0.01	0.95	0.04	0.01
5	5	0.91	0.07	0.02	0.90	0.09	0.01	0.94	0.04	0.02
6	1	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
6	2	0.98	0.02	0.00	0.91	0.09	0.00	0.98	0.02	0.00
6	3	0.94	0.05	0.01	0.90	0.09	0.01	0.96	0.03	0.01
6	4	0.91	0.08	0.01	0.90	0.09	0.01	0.95	0.04	0.01
6	5	0.88	0.10	0.02	0.85	0.13	0.02	0.92	0.05	0.03
7	1	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
7	2	0.98	0.02	0.00	0.86	0.13	0.01	0.97	0.03	0.00
7	3	0.89	0.10	0.01	0.85	0.14	0.01	0.95	0.04	0.01
7	4	0.86	0.13	0.01	0.85	0.14	0.01	0.93	0.05	0.02
7	5	0.82	0.16	0.02	0.84	0.14	0.02	0.89	0.07	0.04
8	1	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
8	2	0.96	0.03	0.01	0.82	0.17	0.01	0.96	0.04	0.00
8	3	0.86	0.12	0.02	0.81	0.18	0.01	0.93	0.05	0.02
8	4	0.81	0.16	0.03	0.81	0.18	0.01	0.90	0.07	0.03
8	5	0.81	0.16	0.03	0.81	0.18	0.01	0.85	0.10	0.05
9	1	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
9	2	0.92	0.06	0.02	0.79	0.21	0.00	0.92	0.07	0.01
9	3	0.88	0.09	0.03	0.76	0.23	0.01	0.91	0.07	0.02
9	4	0.84	0.13	0.03	0.76	0.23	0.01	0.87	0.07	0.06
9	5	0.83	0.13	0.04	0.75	0.23	0.02	0.82	0.09	0.09
10	1	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00
10	2	0.88	0.12	0.00	0.73	0.26	0.01	0.86	0.12	0.02
10	3	0.83	0.12	0.05	0.70	0.28	0.02	0.82	0.14	0.04
10	4	0.77	0.18	0.05	0.68	0.30	0.02	0.76	0.14	0.10
10	5	0.77	0.18	0.05	0.68	0.30	0.02	0.66	0.14	0.20

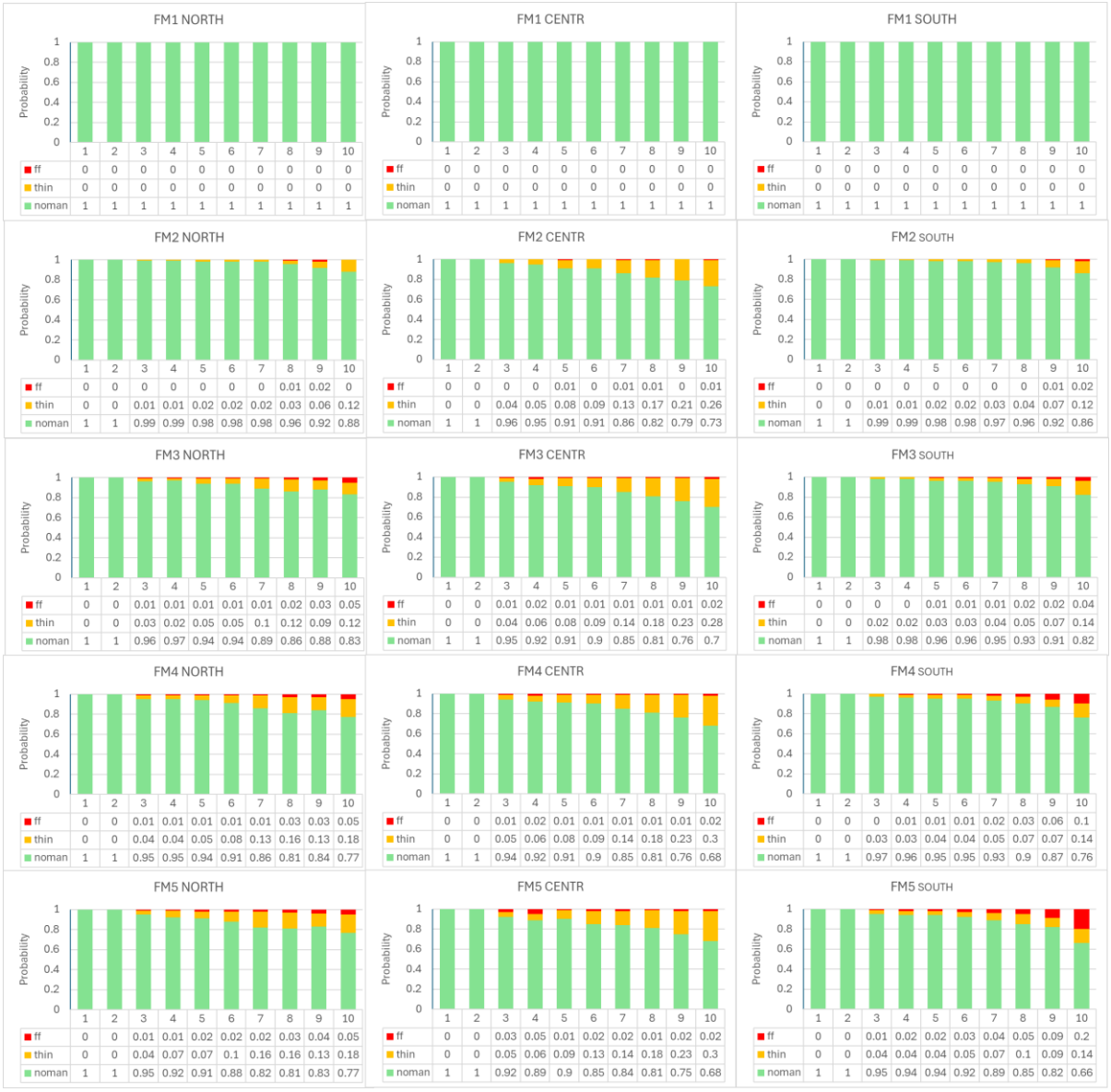


Figure 1: Activity probabilities of noman, thin and ff by country_group, forest management intensity (FM) and vol_bin (x-axis, 1:10).

2.6 EFDM transitions probabilities

2.6.1 Forest harvest

Transition probabilities for forest harvest (i.e., ff and thin) depend on initial stem volume classes. For ff, transition probability 1 was assigned to move the areas to the smallest volume bin. Final felling activity was assumed to depend only on initial volume bin number, and thus was modelled similarly across country_groups. This is a common assumption used in EFDM, as it allows forest regeneration to start. For thinning, transition probabilities were calculated from NFI pairdata that was thinned between the two consecutive inventories. Since there was no information regarding the exact timing of the thinning, stand total volumes were classified into volume bins, and thinning impact was evaluated based on number of observed bin-drops of vol classes. In EFDM, thinning impact on

volume was assumed to depend on initial volume bin number, country_group and tree species group (i.e., hereafter referred as 'sp_group'). After merging NFI plot pair data with statespace defined by vol_bins (n=10), sp_group (n=10) and country_group (n=3), in the case of NA data, the assumed volume class drop was set to 0. Values were manually inspected in the case of artefacts caused by a lack of NFI pair data. For example, NFI pair data did not have observation-based values for thinning the two largest volume bins (of vol_bin 9 and 10) of sp_group 1 in the South region, and thus it was manually set as 2 based on inspection of the closest other class drop values of that species. In the case of missing volume drop information, it was set to 1 (i.e., when thinning happens, the area to be thinned will move one volume bin down). However, for the two smallest vol_bin classes (i.e., 1 and 2 that contain forest areas less than 10 m^3/ha), no thinning impact was assumed to happen (this is in line with ´the definition of EFDM activity probabilities in the previous chapter, see Figure 1: Activity probabilities of noman, thin and ff by country_group, forest management intensity (FM) and vol_bin (x-axis, 1:10).).

2.6.2 Forest regeneration using FORMIT-M

Since NFI plot data do not contain any data containing small trees, a forest stand simulator FORMIT-M (Härkönen et al., 2019, Mäkela and Härkönen, 2019) was used to create some data to parametrise small trees.

FORMIT-M is an open-access, flexible forest management simulator designed to work with standard forest inventory data. It integrates a process-based carbon balance approach with empirical inventory data, making it both climate-sensitive and broadly applicable. Recently, FORMIT-M has been used by Härkönen et al. (2019) to simulate forest development across Europe through the 21st century to illustrate the model's structure and utility, comparing various management strategies under climate change. They noted that FORMIT-M showed strong agreement with national statistics on growing stock (R^2 = 0.938) and produced climate impact projections consistent with other studies. Their results indicated that forest management choices have a stronger influence on growing stock, harvest potential, and carbon balance than climate change—assuming disturbance rates remain stable. The benefit of the FORMIT-M is that it comes with precalculated GPPmax scenario data for several RCPs (RCP2.6, RCP45 and RCP8.5).

The precalculated maxGPP data for the period between 2010-2100 were obtained from the FORMIT-M model input data (Härkönen et al., 2019). The annual Gross Primary Production (GPP) values were calculated using PRELES (Mäkelä et al., 2008; Peltoniemi et al., 2015), a semi-empirical model originally developed for boreal and temperate coniferous forests and later parameterised for a Mediterranean *Pinus pinaster* stand (Mäkelä et al., 2008). PRELES applies the Light Use Efficiency (LUE) approach, where potential GPP is computed as a function of photosynthetically active radiation (PAR), which is subsequently adjusted by multiplicative modifiers that account for temperature, precipitation, and vapour pressure deficit. PRELES also includes a soil water balance sub-model and computes components of evapotranspiration, using daily rainfall as an additional input variable. The PRELES has been parameterised using flux tower data from boreal and temperate coniferous sites, and results suggest that parameter values are largely site-independent within a given vegetation type (Minunno et al., 2016). Based on prior studies, PRELES can be considered relatively robust for predicting GPP under stable atmospheric CO_2 conditions. More recently, the model has also been extended to include CO_2 effects (Kalliokoski et al., 2018). However, in FORMIT-M, they opted not to include an explicit CO_2 effect in our climate change projections. Although the direct impact of elevated CO_2 on potential GPP can be of a similar magnitude to the climate effect, growing evidence suggests this response is often down-regulated due to other limiting factors (Comins and McMurtrie,

1993; Hyvönen et al., 2007; Smith et al., 2016). For instance, in many FACE (Free Air Carbon Enrichment) experiments, no long-term growth enhancements have been observed, likely due to nutrient limitations (Norby and Zak, 2011; Reich and Hobbie, 2013). Nevertheless, the hypothesised interaction between elevated CO_2 and reduced stomatal conductance may still be relevant, particularly at drier sites where it may counteract drought effects (Ainsworth and Rogers, 2007; Kalliokoski et al., 2018). Consequently, the resulting maxGPP projections under future climate scenarios are intentionally conservative.

Forest regeneration was predicted using FORMIT-M using the following set-up: First, a shapefile was used to sample potential maxGPP (g C m^{-2} $year^{-1}$) values for the three regions (North, Central and South) from the precalculated maxGPP data for RCP45 corresponding present day climate. A total of 200 maxGPP values were selected using quantiles (selected between 0 and 0.995 using interval of 0.005), which resulted a sample of 200 maxGPP values, which represent regional variations in maxGPP values. These quantile values were topped up with minimum and maximum maxGPP values. Then the all EU maxGPP points were classified into bins formed by min, max and quantile values, and stratified sampling was used to draw a random sample plot from each bin, to represent variations across European maxGPP data in FORMIT-M regeneration simulations. This sampling resulted a draw of 58 plot locations from North, 86 plot locations from Centr, and 56 plot locations from South EU. This step was necessary to provide objectively selected plot locations to FORMIT-M, which uses NFI-like input data with plot coordinates (to extract maxGPP values from the data). This step was necessary to cover a wider range of variations in maxGPP values across EU land area, and to make sure that high and low maxGPP values are represented by the sample.

Then, FORMIT-M NFI-like input data was initiated by setting zeros to all forest variables (i.e. basal area, height, stand density, mean diameter), and replicating the forest structural input for every plot location and for each tree species group. FORMIT-M was run annually for years 2020-2024, and forest structural data from years 2020 and 2024 were used to create pair data needed to parametrise regeneration in EFDM.

Since tree species grouping varies between models, FORMIT-M tree species groups simulations were reclassified into PathFinder sp_groups so that, in the North region, FORMIT-M species groups were pine, spruce and birch, whereas Central and South regions also contained tree species group simulations for Quercus and Fagus. The regeneration data simulated using FORMIT-M was pre-processed to match with PathFinder sp_groups before it was added to noman NFI pair data. A new datatable was created for each pathfinder tree sp_group by duplicating the regional tree species simulations from FORMIT-M. For the North region, FORMIT-M simulations for birch were used to approximate growth of all deciduous sp_groups (except for Eucalyptus, which is not assumed to be able to grow in North), FORMIT-M spruce simulations were used to approximate Abies and Picea sp_groups, and FORMIT-M simulations for pine were used to approximate Pinus and other_conifers sp_groups. For the Central and South regions, the FORMIT-M species data were reclassified similarly, except that regeneration of PathFinder sp_group Fagus was approximated using respective FORMIT-M Fagus data, and PathFinder quercus and other_broadleaf sp_groups regeneration using FORMIT-M Quercus simulations. Eucalyptus regeneration was not added to the Central region, but in the South region, Eucalyptus regeneration was approximated using FORMIT-M birch tree species groups simulations data (in FORMIT-M, this class is suggested to be used also for Eucalyptus). For each regeneration plot, a mean maxGPP value was calculated. The results from FORMIT-M regeneration simulation are shown in Table 6**.** The regeneration pairdata was later combined with noman pairdata to include the growth of small trees into noman transition probabilities in EFDM. Regeneration was assumed to be done with the same species as was occurring before the final felling.

In FORMIT-M forest regeneration intensity is assumed to vary between regions and species. In the North and South regions, all species groups are assumed to be planted, and the stand density is 2000 saplings per hectare at the age of 5. In the Central region, forest planting intensity is assumed to be higher than in other regions, and the number of saplings at the age of 5 varies between 2500-7000 saplings per ha, depending on the species. Overall, political decisions regarding which tree species to plant and which planting densities may be subject to change fast and given the assumed direction towards more intensive forestry practices, that maximize forest growth potential, assuming intensive planting seems applicable solution.

Table 6: Forest regeneration table showing assumed planting density (in FORMIT-M).

	0-yr dbh, cm	5-yr dbh, cm	0-yr n/ha	5-yr n/ha	0-yr m3/ha	5-yr m3/ha
Pine:						
CENTR	0	2	0	4500	0	6.6
NORTH	0	2	0	2000	0	3.2
SOUTH	0	2.5	0	2000	0	3.7
Spruce:						
CENTR	0	2	0	2500	0	3.8
NORTH	0	2	0	2000	0	3.1
SOUTH	0	2.5	0	2000	0	3.8
Deciduous (fast growing, 4)						
CENTR	0	2	0	3000	0	5.1
NORTH	0	2	0	2000	0	3.8
SOUTH	0	4	0	2000	0	6.8
Deciduous (slow growing, 5)						
CENTR	0	1.3	0	7000	0	9.2
SOUTH	0	2.5	0	2000	0	4.9
Deciduous (slow growing, 6)						
CENTR	0	1.3	0	6000	0	6.8
SOUTH	0	2.5	0	2000	0	4.2

2.6.3 Considerations of forest growth (noman) modelling

In order to approximate forest growth as a function of climate change, an empirical approach was developed. An assumption was made that present-day forest cover reflects habitat suitability ranges of different sp_groups. The underlying problem in the development of growth trajectories for different sp_groups was that each sp_group contains several different tree species, which differ in their physiological properties. Since sp_groups were formed based on WP2 possibilities to separate different species in different geographical regions, rather than tree species physiological properties or forest management related properties, a generalised approach was needed for growth modelling.

Our approach to model forest growth is to create an empirical link between potential maxGPP values and observed stem volume growth (i.e. we overlook GPP allocation to respiration and NPP, and NPP allocation to stem volume growth, and assume increasing GPP to be linked with increasing growth).

This simplification was necessary, as our input data did not contain information of e.g., site quality (which has a direct effect on the NPP allocation to stem growth, i.e. in poor sites stem growth decreases while the below-ground allocation increases (Litton et al., 2007; Valentine and Mäkelä, 2012).

The use of an empirical approach may also be supported by the fact that our PathFinder NFI data contained ten fairly broad groupings of different tree species, which were created based on spatial separability of different species in Sentinel-2 satellite data, rather than physiological or forest management-related considerations. The use of generalised grouping of various tree species also made it challenging to evaluate future changes in tree species abundance, or future geographical shifts, and thus tree species abundance was assumed to change only within each sp_group, but not between sp_groups. Given the generalised sp_groups, this assumption may be considered acceptable, for example, for pine, while some species are more prone to suffer from climate change impacts, some pine species are more drought resistant and thus will face less climate stress. Overall, our approach assumes that present-day habitat suitability of different sp_groups is described by the NFI pair data, and the variation in present-day climate characterises both the potential high and low growth areas, and contains forest damages that are represented by the NFI pair data.

All these assumptions thus rely on NFI data to accurately characterise forest states. Yet, since NFI pairdata was available only from 10 European countries (i.e. Austria, Belgium, Czech Republic, Germany, Poland, Ireland, Norway, Sweden, Slovenia and Spain), and resulting transition matrix models are sensitive to number of available observations, the models are expected to be perform the best in countries from which more data was available.

Overall, predicting future changes in tree species cover and habitat suitability to different tree species is further complicated by the fact that political decisions/legislation, e.g., regarding which tree species to plant/prefer, can have a significant effect on the spatial distribution of different tree species and their abundance across the EU. In addition, there is a need for generally accepted trajectories which describe which tree species should be preferred in forest regeneration in different geographical areas. Although potential habitat suitability trajectories are available for many single species (e.g., de Rigoet al., 2016; Achille et al., 2022), no consensus on how these data should be used in scenario work exists.

A particular challenge in modelling forest and development is also caused by a very broad geographical extent, combined with the fact that sp_group NFI plot locations were not spatially uniformly distributed over the EU region. For example, quantiles of the potential maxGPP difference maps between 2100 and 2020 for RCP45 are: -879.7 (0%), 23.6 (25%), 107.9 (50%), 148.5 (75%) and 865.9 (100%) gC $m^{-2}yr^{-1}$, which indicates that for most forest areas very moderate climate change impact is to be assumed (Figure 2: Potential maxGPP for RCP45 for the year a) 2020, and b) the differences between potential maxGPP values at 2100 and 2020.).

The potential maxGPP data for RCP2.6, RCP4.5 and RCP8.5 are provided as FORMIT-M input files. The file structure is similar for all files. In each file, the maxGPP data are provided for three time periods (mid-years of 2005, 2045 and 2095) by latitude and longitude coordinates. Thus, for each grid location (n= 35657), the maxGPP value can be obtained by interpolation. The approach does not allow accounting for abrupt changes, although it allows evaluating climate effects before conducting actual model runs, which aligns well with the matrix model approach used in this study.

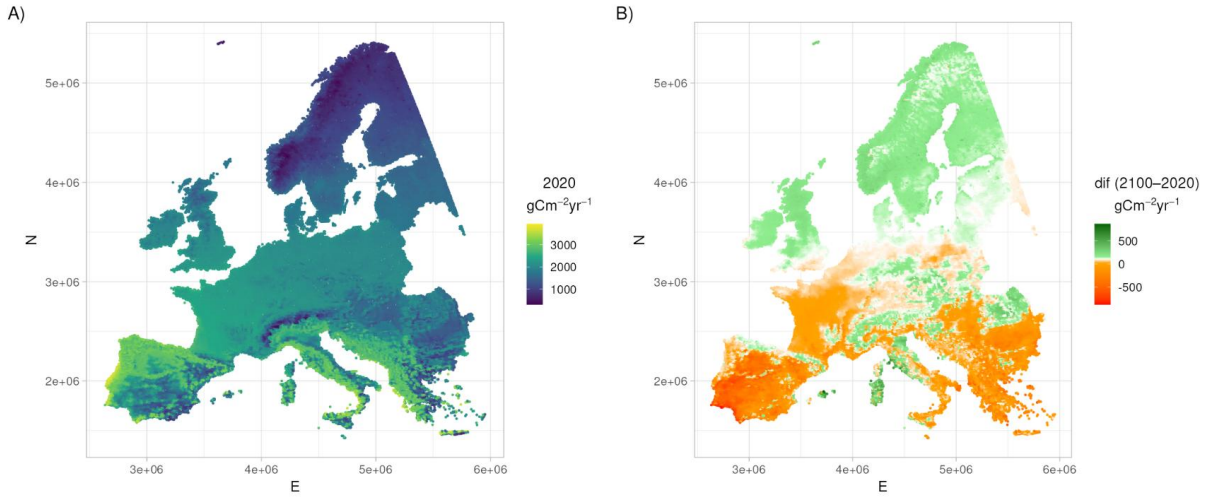


Figure 2: Potential maxGPP for RCP45 for the year a) 2020, and b) the differences between potential maxGPP values at 2100 and 2020.

2.6.4 Noman

Transition probabilities for noman were calculated using NFI plot data pairs, and regeneration pairdata from FORMIT-M simulations, since NFI data did not contain forest regeneration plots. Forest growth (noman) was modelled empirically by overlaying noman data pairs with the potential maxGPP data, and calculating potential maxGPP values corresponding to the mean inventory year of each of the NFI plot pairdata observation (Table 7).

As the NFI data interval varied between countries, only the initial NFI pair data stem volumes were directly classified into the vol bins. The latter NFI inventory stem volumes were scaled to correspond to a 5-yr growth period before they were classified into EFDM vol bins. This was done by calculating the average 5-yr growth increment and adding it to the initial plot stem volume. Then, the latter stem volume was classified into vol-bins. For regeneration plots, the stem volumes were directly classified into stem vol_bins as forest plots were grown for 5-yrs.

Calculation of transition probabilities in EFDM, requires that the maxGPP values are classified. Yet, classification of the maxGPP values was found challenging due to large variations in maxGPP values between country_groups and also within and between sp_groups. Although modelling growth of different sp_groups accurately is important from a forest management perspective, it is equally important to be able to capture the climate impact on growth. Thus, a trade-off solution was needed.

Since potential maxGPP varied greatly among regions (exceeding the climate change impact), developing generalised noman models for each sp_group over Europe was not possible. The means of country_group mean maxGPP values were: NORTH= 1152, CENTR= 1991 and SOUTH= 2526 gC m^{-2} yr^{-1}) (Table 7), and thus it was decided to fit a separate noman model for each region. This decision was supported by the fact that that the number of NFI plot pairdata observations from NORTH was notably lower than that from CENTR and SOUTH regions (NORTH n= 27192, CENTR n = 48021 and SOUTH n = 46409), and thus are not regionally biased by the number of available observations. Also, the development of regional models can be supported by the fact that NFI plot data contains influences of forest damages, the intensity of which may be assumed to vary between regions and sp_groups.

A simulation was conducted to test if the same maxGPP classification thresholds across all sp_groups and regions would perform sensibly, but this resulted in poor model performance (i.e., forest volume growth would increase and decrease along with increasing potential maxGPP for many sp_groups). This may be partly assumed to be caused by varying levels of forest damage in different regions and species, and different forest structural distributions.

Table 7a and b: The number of NFI plot pairdata observations ('count') and statistics for potential maxGPP for the mean NFI inventory year by sp_group and region. Colorbars visualise variations in data between sp_groups and regions. (a) shows the data sorted by sp_group and (b) by region.

a)

		maxGPP gC m^{-2} yr^{-1}				
sp_group	**country_group**	**count**	**min**	**mean**	**sd**	**max**
1	CENTR	4127	996	2001	101	2788
1	NORTH	5780	308	1037	314	1888
1	SOUTH	509	1108	2760	497	3834
2	CENTR	870	1386	2001	106	2323
2	NORTH	402	486	1279	297	1880
2	SOUTH	556	1064	2586	479	3676
3	CENTR	3326	996	2004	120	2300
3	NORTH	80	1423	1704	113	1868
3	SOUTH	2351	1311	2875	315	3676
4	CENTR	3789	1341	2018	107	2315
4	NORTH	230	1034	1628	146	1930
4	SOUTH	14489	568	2511	586	3977
5	SOUTH	1544	725	2851	574	3977
6	CENTR	3959	1140	1996	107	2347
6	NORTH	581	461	1282	332	1886
6	SOUTH	2264	887	2708	535	3977
7	CENTR	7662	625	1930	178	2574
7	NORTH	8949	450	1169	288	1867
7	SOUTH	153	1343	2038	259	3032
8	CENTR	1387	1213	1971	120	2324
8	NORTH	3	1181	1276	146	1444
8	SOUTH	165	1372	2512	489	3421
9	CENTR	21578	954	2006	66	2301
9	NORTH	11143	327	1170	288	1930
9	SOUTH	23178	568	2453	628	3676
10	CENTR	1323	754	1974	190	2416
10	NORTH	24	1336	1635	137	1861
10	SOUTH	1200	1001	2607	453	3379

b)

			maxGPP gC m^{-2} yr^{-1}			
sp_group	**country_group**	**count**	**min**	**mean**	**sd**	**max**
1	NORTH	5780	308	1037	314	1888
2	NORTH	402	486	1279	297	1880
3	NORTH	80	1423	1704	113	1868
4	NORTH	230	1034	1628	146	1930
6	NORTH	581	461	1282	332	1886
7	NORTH	8949	450	1169	288	1867
8	NORTH	3	1181	1276	146	1444
9	NORTH	11143	327	1170	288	1930
10	NORTH	24	1336	1635	137	1861
1	CENTR	4127	996	2001	101	2788
2	CENTR	870	1386	2001	106	2323
3	CENTR	3326	996	2004	120	2300
4	CENTR	3789	1341	2018	107	2315
6	CENTR	3959	1140	1996	107	2347
7	CENTR	7662	625	1930	178	2574
8	CENTR	1387	1213	1971	120	2324
9	CENTR	21578	954	2006	66	2301
10	CENTR	1323	754	1974	190	2416
1	SOUTH	509	1108	2760	497	3834
2	SOUTH	556	1064	2586	479	3676
3	SOUTH	2351	1311	2875	315	3676
4	SOUTH	14489	568	2511	586	3977
5	SOUTH	1544	725	2851	574	3977
6	SOUTH	2264	887	2708	535	3977
7	SOUTH	153	1343	2038	259	3032
8	SOUTH	165	1372	2512	489	3421
9	SOUTH	23178	568	2453	628	3676
10	SOUTH	1200	1001	2607	453	3379

While some tree sp_groups had a large number of NFI plot datapairs, such as Quercus (sp_group of 4) and pine (sp_group of 9), for most sp_groups few observations were available (or observations had low variation in maxGPP values). Thus, an approach was required to model the growth of all ten sp_groups at the regional-level. For each region, all available data were used to create 'generalised' noman models. The generalised noman models, were then, in turn, used to predict the growth of all sp_groups that did not have enough pairdata to develop separate models.

For NORTH region, three main sp_groups were created for noman modelling (Table 7**b**). These re-groupings were: **1)** deciduous species (sp_groups of 1-6 in Table 7**b**), **2)** picea + abies (7, 8), and **3)** pines + other conifers (9, 10) (Table 8**,** Table 9). NFI plot data was better available for CENTR region, and thus five separate models were developed for: **1)** deciduous species (sp_groups of 1, 2, 5 and 6), **2)** Fagus (3), **3)** Quercus (4), **4)** Picea (7), and **5)** conifers (8, 9, 10). For the SOUTH region, also five generalised sp_groups were created based on the NFI data availability: **1)** deciduous (1, 2, 3, 6), **2)** Fagus (3), **3)** Quercus, **4)** Eucalyptus (5) and **5)** conifers (7, 8, 9, 10). Although the number of

eucalyptus observations was low, it was considered enough as eucalyptus is not as widely spread compared to other sp_groups. These generalised models were used to parameterise all sp_groups, which had contributed to the model development. Although the generalised models were compiled primarily to represent local variations in species abundance and growth, other considered factors were: 1) to obtain representative regional transition probability matrices for each sp_group, and 2) to be able to capture regional climate change effects.

Creation of representative transition probability matrices was done by inspecting pairdata availability by decreasing the number of maxGPP quantiles (these classes required to capture growth changes) by main sp_groups and regions until sensible results were obtained. This was investigated by visually assessing resulting noman matrices (see the EFDM noman matrices with simulated forest regeneration pairdata in **ANNEX A**), and by evaluating predicted outcomes in regard to maxGPP values (i.e., later referred to as 'P0_classes').

Transition probability matrix performance was investigated by predicting noman growth for each region, sp_group and P0_class, and calculating area-weighted-volume class value at the age of 80 years. First, quantiles were used to split NFI plot data pairs into similar-sized bins by sp_group and region. Then, noman models were fit, and used to predict forest development starting from volume class 1 regeneration state, for 80 years, keeping the P0_class class value constant. The 80-yr simulation was repeated for all generalised sp_groups and P0_classes and regions. Finally, the number of P0_classes was reduced until predicted area-weighted-volume class values showed sensible growth trends (i.e., growth was not allowed to siz-saz between consecutive P0_classes, as the aim was to capture either increasing or decreasing trend (although decrease was allowed for the final P0_class to reflect the fact that sp_group is growing in the habitat suitability extreme indicating decreasing growth). In short, 80-yr simulations were necessary to determine the number of P0_classes by region and sp_group, that can be modelled based on NFI pairdata observations, and to compare different matrix model performances over sp_groups and regions (Table 10).

Number of NFI pairdata observations available for each generalized sp_group and P0_ are shown in Table 8**,** and respective maxGPP quantile values in Table 9**.** P0_class bottom value was always set as 250 and maximum was 4050. While the quantiles approach ensured that each P0_class contained enough NFI plot pairdata observations, it also resulted that P0_class maxGPP value thresholds varying by sp_group, region and number of P0_classes.

Table 8: Number of NFI pairdata observations available for each generalized sp_group and P0_class (maxGPP value class). The respective maxGPP quantile values are shown in Table 9

			P0_class					
	grouping	**sp_group**	**1**	**2**	**3**	**4**	**5**	**6**
NORTH	**decid**	1,2,3,4,5,6	1229	1245	1234	1237	1236	892
	spruce	7,8	1550	1574	1574	1565	1567	1122
	pine	9,10	1939	1952	1968	1949	1951	1408
CENTR	**fagus**	3	1085	1088	1153			
	quercus	4	942	939	947	961		
	decid	1,2,5,6	2933	2947	3076			
	spruce	7	1328	1350	1307	1369	1347	961
	conif	8,9,10	7956	7868	8464			
SOUTH	**decid**	1,2,3,6	967	1014	982	1002	1004	711

	quercus	4	2529	2531	2525	2533	2518	1853
	eucal	5	503	510	531			
	pine	9	4042	4062	4041	4055	4073	2905
	conif	7,8,10	499	500	519			

Table 9: Potential maxGPPmax value thresholds (class bottom values) for the generalised sp_groups. Possible range of variation was 250-4050 (g C m-2 year-1).

			P0_class					
	grouping	**sp_group**	**1**	**2**	**3**	**4**	**5**	**6**
NORTH	**decid**	1,2,3,4,5,6	250	757	913	1058	1281	1573
	spruce	7,8	250	874	1005	1153	1351	1558
	pine	9,10	250	888	1016	1162	1346	1552
CENTR	**fagus**	3	250	1979	2052			
	quercus	4	250	1979	2029	2071		
	decid	1,2,5,6	250	1979	2037			
	spruce	7	250	1805	1916	1976	2030	2080
	conif	8,9,10	250	1984	2038			
SOUTH	**decid**	1,2,3,6	250	2230	2824	2951	3034	3143
	quercus	4	250	1827	2316	2725	2937	3104
	eucal	5	250	2982	3124			
	pine	9	250	1774	2291	2697	2903	3053
	conif	7,8,10	250	2300	2859			

Based on simulations (Table 10) the number of P0_classes by region was determined to be three for the NORTH region and five for the CENTR and SOUTH regions. For the NORTH region clear increasing growth trend with an increasing number of P0_classes was observed, and there were no signs of growth decline.

However, for CENTR Europe decreasing growth trend was observed for the deciduous species and conifer groups (i.e., sp_group), a slightly increasing and decreasing trend for Fagus, increasing trend for Quercus and Spruce (Table 10). Due to the low variation in maxGPP values in the CENTR region (Table 7) for all other than Spruce, it was only possible to define three P0_classes. For the SOUTH region, large variation in maxGPP values resulted in that for three species groups, six P0_classes could be defined. For Quercus, Eucalyptus, conifer and pine groups increasing growth trend was observed (Table 10). Also, for the deciduous group, an increasing trend was observed, except for the last P0_class, which showed slight decrease in growth.

Growth changes may be assumed reasonable, as future climate change (Figure 2: Potential maxGPP for RCP45 for the year a) 2020, and b) the differences between potential maxGPP values at 2100 and 2020.) will likely result in positive advances in P0_class values for the NORTH region, while the SOUTH region will mainly face decreases in present-day P0_class values. CENTR region is expected to face both increases and decreases in P0_class values. Overall, most dynamic increasing changes in potential growth are to be expected in the NORTH region. Highest potential growth is to be expected

in CENTR Europe, whereas in SOUTH Europe, for all sp_groups, except Eucalyptus, the potential maximum growth is expected to be the lowest.

Table 10: Area-weighted vol-classes at 80-yr age using constant P0_class values for comparing model predictions between regions, sp_groups and P0_classes. Generalised sp_groups, from which models were assigned to other sp_groups, are shown in bold. Values are shown for all sp_groups to allow for comparing model outcomes by sp_groups.

		sp_group									
	P0_class	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
NORTH	**1**	**3.52**	3.52	3.52	3.52	3.52	3.52	**6.04**	6.04	**5.92**	5.92
	2	**3.99**	3.99	3.99	3.99	3.99	3.99	**6.94**	6.94	**6.96**	6.96
	3	**5.26**	5.26	5.26	5.26	5.26	5.26	**7.94**	7.94	**6.93**	6.93
	4	**6.60**	6.60	6.60	6.60	6.60	6.60	**8.78**	8.78	**7.43**	7.43
	5	**7.92**	7.92	7.92	7.92	7.92	7.92	**9.09**	9.09	**7.78**	7.78
	6	**8.57**	8.57	8.57	8.57	8.57	8.57	**9.34**	9.34	**8.16**	8.16
CENTR	**1**	9.31	9.31	**9.55**	**9.28**	9.31	**9.31**	**9.60**	9.55	**9.55**	9.55
	2	9.20	9.20	**9.59**	**9.32**	9.20	**9.20**	**9.73**	9.49	**9.49**	9.49
	3	9.13	9.13	**9.35**	**9.47**	9.13	**9.13**	**9.79**	9.52	**9.52**	9.52
	4							**9.81**			
	5							**9.86**			
	6							**9.82**			
SOUTH	**1**	6.51	6.51	6.51	**3.33**	**8.99**	**6.51**	4.45	4.45	**5.87**	**4.45**
	2	6.77	6.77	6.77	**3.65**	**9.66**	**6.77**	4.10	4.10	**6.30**	**4.10**
	3	7.36	7.36	7.36	**4.38**	**9.78**	**7.36**	5.78	5.78	**6.84**	**5.78**
	4	7.37	7.37	7.37	**5.70**		**7.37**			**7.45**	
	5	7.84	7.84	7.84	**6.39**		**7.84**			**8.32**	
	6	7.32	7.32	7.32	**6.43**		**7.32**			**8.71**	

Models were fit using the R library *efdm* (Kuronen M, Räty M (2023). After the number of P0_classes and the class thresholds were determined, the forest regeneration plot data were added to the NFI plot pair data. EFDM statespace for noman transition probabilities calculation was set up with variables P0_class (n=6 classes), sp_group (n=10) and vol_bins (n=10). Transition probabilities were calculated by assigning P0_class and sp_group into 'by' term, which indicates that a separate model is fit to each P0_class and sp_group. Prior was set to 'nochange' as volume is our only dynamic variable (i.e., assuming any increase or decrease in vol_bin value would not be sensible as the vol_bins are not uniformly distributed, and the gap between class mean m^3/ha values may be larger than that the forest can realistically grow within the 5-yr period. The final transition matrices after including simulated forest regeneration pairdata are shown in **ANNEX A**.

Theoretical test of matrix model capability to capture climate change effects was conducted using the potential maxGPP data for RCP45, which was selected as the baseline scenario. First, each potential maxGPP gridcell membership into regions (i.e., NORTH, CENTR, and SOUTH) was determined using a shapefile (Eurostat, 2025) showing administrative areas of different countries. Then, the RCP45 potential maxGPP values were calculated for all NFI-mid periods between 2020-2100 (i.e. 2022, 2027 … 2092, 2097). Then, each gridcell maxGPP value was classified into predefined

P0_classes by sp_groups (Table 9**).** An assumption was made that each grid cell would contain all sp_groups. This is thus a theoretical test of the matrix model's capability to capture climate change effects. After each grid cell was classified into P0_classes by sp-groups and regions. The 'total' climate impact on P0_class transitions was investigated by counting the total number of times the grid cell P0_class changes between 2020-2100 (total of 16 time-steps) within each sp_groups, and dividing the total counts by the total number of sp_groups P0_classes minus 1.

This resulted in images that show areas where the largest changes in P0_classes will be expected. The largest differences are to be expected in CENTR Europe, while the smallest in NORTH region (see **ANNEX A**). To provide better insight regarding whether future climate change results in positive or negative changes, the calculation was modified by counting the total number of times the gridcell P0_class changes are positive (or negative) between 2020-2100 (total of 16 time-steps), and total counts are divided by the total number of sp_groups P0_classes minus 1.

Dividing the total counts of changes by the number of sp_group P0_classes -1, is an attempt to rescale the total numbers of P0_class changes, which vary between 3-6 by sp_group and region. Noteworthy, is that since the number of P0_classes varied by sp_group and region, the illustrated positive and negative percentage change values are not directly comparable. Yet, they serve to illustrate the potential climate change effects by sp_group and region following RCP45.

2.7 EFDM CoefTable creation

In EFDM all model output variables are produced using a look-up table (LUT, hereafter referred to as coeftable), by multiplying areas belonging to different vol_bins by respective vol_bin (by country_group and sp_group) specific LUT values. Thus, all model dynamics must be derived from the area changes between different forest classes. While the approach is well suited for quantifying the temporal development of forest structural properties, it is less suited for characterising forest properties that are indirectly linked with forest structure (such as above-ground deadwood). Also, accounting situations with huge variation in forest variable properties (such as stem density in final felled stands, which may have been planted or left to be naturally forested), estimation of LUT values remains challenging.

Growing stock (m2/ha) coeftable values were calculated as averages of the volume bin values, and thus these were the same across sp_groups and regions. Estimate of harvested volume in thinning varied by sp_group and region, and thus it was calculated by vol_bin, sp_group and region by merging with EFDM thinning activity matrix (in practice this was done by merging thinning activity-matrix with coeftable first by vol0 (before) and then by vol1 (after), and obtaining an estimate of harvested volume in thinning by subtracting the initial volumes from the latter volumes). Harvested volume in final felling was calculated by subtracting the growing stock volume of the smallest volume bin from the vol_bin mean volume. Thus, the coeftable contained three volume estimates by vol_bin, sp_group and region: growing stock volume, volume obtained in thinning (drain) and volume harvested in final felling (drain).

For other variables, some additional modelling and external datasets were needed, since NFI-plot datapairs did not contain all required input variables, or the dynamics of the variables were showing abrupt zigzag patterns (increasing and decreasing values) or missing values along the consecutive vol_bins.

A particularly challenging variable was above ground deadwood biomass (dwb). Although this variable was contained from the NFI pairdata, the data had a lot of missing values, outliers and abrupt zigzag pattern (increasing and decreasing values) along the consecutive vol-bins. Since there is no obvious reason why dwb would act (randomly) like that as the forest grows (i.e. past forest management history remains unknown), some preprocessing was necessary: First, clear outliers were excluded using regional vol_bin 80 percentile values (i.e., if vol_bin dwb value exceeded that value, it was excluded). Then, the missing values were replaced using regional mean dwb 80 percentile values of each vol_bin. Then, to smooth out zigzag patterns of vol_bin dwb values by region and sp_group, polynomial regression models were fit to data.A second-order polynomial regression model form was used, modeling dwb as a function of vol_bin and vol_bin squared. Original dwb values were replaced by modelled ones (Table 11). For the two vol_bins for which negative dwb values were obtained, the dwb value was set to zero.

Although the coefficient of determination values were low for some models, this step was necessary to smooth the unexplained observation values for scenario simulations. Simplified smoothing approach can be justified in this matrix model context, as above-ground volume, by itself, is a poor predictor of dwb (i.e., dwb dynamics do not follow forest growth dynamics, and is a result of past regional forestry practices over several decades). Thus, it is clear that future studies should focus on developing an approach for including a variable to quantify the impacts of past forest management practices on forest aboveground deadwood biomass.

Table 11: Fitted regression models for above-ground deadwood biomass by region and sp_group. R2 is the coefficient of determination.

region	sp_group	(Intercept)	vol_bin	I(vol_bin^2)	R2
NORTH	1	3.96	-0.36	0.11	0.91
NORTH	2	5.66	-2.38	0.43	0.79
NORTH	3	10.08	-3.52	0.43	0.74
NORTH	4	4.56	-1.04	0.22	0.90
NORTH	5	11.49	-4.26	0.58	0.91
NORTH	6	0.79	1.54	-0.05	0.82
NORTH	7	9.31	-2.98	0.40	0.90
NORTH	8	11.06	-3.73	0.48	0.55
NORTH	9	3.88	-1.17	0.19	0.88
NORTH	10	10.07	-4.44	0.58	0.67
CENTR	1	7.02	-1.30	0.24	0.95
CENTR	2	3.00	-0.11	0.16	0.90
CENTR	3	13.14	-1.12	0.12	0.34
CENTR	4	4.60	0.01	0.10	0.70
CENTR	5	13.92	-1.63	0.20	0.80
CENTR	6	12.89	-1.85	0.21	0.64
CENTR	7	15.11	-2.74	0.25	0.89
CENTR	8	17.38	-4.37	0.44	0.68
CENTR	9	3.64	0.04	0.06	0.81
CENTR	10	7.73	-0.60	0.10	0.36
SOUTH	1	-1.85	1.44	-0.03	0.48
SOUTH	2	3.04	-1.70	0.32	0.98
SOUTH	3	2.08	-0.99	0.23	0.96
SOUTH	4	2.52	-1.49	0.29	0.98
SOUTH	5	0.18	0.56	-0.02	0.89
SOUTH	6	-2.83	2.07	-0.09	0.55
SOUTH	7	1.16	-0.40	0.17	0.96
SOUTH	8	2.51	-1.59	0.34	0.99
SOUTH	9	1.65	-0.95	0.20	0.99
SOUTH	10	1.43	-0.93	0.20	0.83

Since the NFI plot pairdata did not contain all forest structural variables for all countries, FORMIT-M simulations were used to obtain estimates of diameter at breast-height, tree height, and plot basal area (i.e., hereafter referred to as D, H and BA, respectively). From FORMIT-M simulated data, vol_bin specific values for D, H and BA were calculated taking averages by vol_bin, sp_group and region.

Since FORMIT-M sp_groups differed from those used in Pathfinder analysis, FORMIT-M sp_groups were reclassified into Pathfinder sp_groups before the averages were calculated: For the NORTH region, FORMIT-M simulations for birch were used to approximate growth of all deciduous

sp_groups, FORMIT-M spruce simulations were used to approximate Abies and Picea sp_groups, and FORMIT-M simulations for pine were used to approximate Pinus and other_conifers sp_groups. For CENTR and SOUTH regions, the FORMIT-M species data were reclassified similarly, except that PathFinder sp_group Fagus was approximated using respective FORMIT-M fagus data, and PathFinder 'Quecus and other_broadlesves' sp_groups regeneration using FORMIT-M Quercus simulations.

Another challenging variable to predict was stem number (N). Since our data did not contain regeneration plots, these had to be included. As a result, vol_bin specific N values by region and sp_group showed clear, unnatural, zigzag patterns, which hardly represent real forest structural trajectories. Since allometric biomass and litter models, which are needed to parametrise the coeftable, are based on LUT-table forest structural variables, the relations between consecutive vol_bin specific forest variables must be sensible.

Table 12: Predicted polynomial regression for stem number.

region	sp_group	vol_bin	I(vol_bin^2)	R2
NORTH	1	1436.99	-148.49	0.79
NORTH	2	1436.99	-148.49	0.79
NORTH	3	1436.99	-148.49	0.79
NORTH	4	1436.99	-148.49	0.79
NORTH	5	1436.99	-148.49	0.79
NORTH	6	1436.99	-148.49	0.79
NORTH	7	1299.15	-133.79	0.82
NORTH	8	1299.15	-133.79	0.82
NORTH	9	1318.99	-136.08	0.82
NORTH	10	1318.99	-136.08	0.82
CENTR	1	1278.81	-115.58	0.80
CENTR	2	1278.81	-115.58	0.80
CENTR	3	1407.26	-130.60	0.79
CENTR	4	1491.05	-141.70	0.85
CENTR	5	1278.81	-115.58	0.80
CENTR	6	1491.05	-141.70	0.85
CENTR	7	1278.35	-114.96	0.80
CENTR	8	1278.35	-114.96	0.80
CENTR	9	1269.71	-116.51	0.80
CENTR	10	1269.71	-116.51	0.80
SOUTH	1	1331.92	-126.85	0.83
SOUTH	2	1331.92	-126.85	0.83
SOUTH	3	1345.86	-128.48	0.84
SOUTH	4	1320.88	-126.72	0.84
SOUTH	5	1331.92	-126.85	0.83
SOUTH	6	1320.88	-126.72	0.84
SOUTH	7	1397.16	-133.42	0.84
SOUTH	8	1397.16	-133.42	0.84
SOUTH	9	1288.01	-123.15	0.84
SOUTH	10	1288.01	-123.15	0.84

Thus, an approach was developed to first smooth the variations in variable N using polynomial regression, and then to predict D based on N and BA. A linear models were fitted to predict N, including indicator variables for each level of vol_bins (excluding the intercept), along with a quadratic term for vol_bin. The model was fit without intercept as the stem number of the smallest bin should go through the origin due to forest final felling and regeneration. For vol_bins where the predicted N was negative, the original stem number was used (these were all the largest vol_bin observations). Predicted N values were too high for the smallest volume bin, and thus, for the smallest volume bin, the N was set to 2 (an arbitrary value selected to represent stands for which single large trees were left to aid natural regeneration or to preserve biodiversity).

Then, vol_bin D-values were predicted based on the new N-values (sqrt((4*BA) / (pi*N))) *100). Predicted D values were too high for the smallest volume bin, and thus, for the smallest volume bin, the D was set to 0.8 (an arbitrary value selected to represent planted saplings on regeneration stands).

Although the resulting vol_bin mean values were not perfectly aligned with each other, they demonstrated sensible dynamics (i.e., smooth transitions) between consecutive vol_bins. Also, as EFDM coeftable often uses class means values, which are sensitive to extreme values, approximating variable dynamics and ranges dynamically seems both like a practical solution and sufficient for the modelling task.

2.7.1 Allometric biomass models

Biomass models were needed to convert forest structural variables into component biomasses. Since most of the available biomass component models are site-specific and thus restricted to the local climatic and structural conditions, we decided to use regional species-specific metamodel equations from the database by Forrester et al. (2017) that can be used in a wide range of stand conditions. We decided to use a model form that employs dbh as a predictor (i.e. that being the most common predictor to be measured in allometric biomass studies), to make sure metamodels are compiled using as large a number of individual studies as possible.

Selection of individual metamodels was also challenging, the fact that the models had to represent the sp_groups, which were by themselves mixtures of different tree species. Overall, a set of equations was selected from the Forrester database considering the following factors: species, availability of different biomass component models, sample size, bdh-range, BA-range, and latitude range. No specific decision rules were determined for model selection, as it was not possible to decide exact or relative weights for the aforementioned factors. Biomass components for which separate biomass models were not available were calculated based on modelled estimates. Total agb was calculated as the sum of individual biomass components. Fineroot biomass was assumed equal to the foliage biomass, in the absence of better information.

Forrester et al. (2017) database contains biomass equations for 27 European tree species that are currently considered important by European foresters. For most species, the equations included a wide range of sites across the current species distributions within Europe and are thus assumed to be representative of the given species within Europe. Regional models of Forrester et al. (2017) were developed based on equations that predict the biomass components of foliage, branch wood (total, live or dead, all including the bark), stem (including wood and bark), stem and branch (including stem + branch wood and bark), aboveground (including stem + branch wood and bark + foliage) and

roots. Where the underlying studies split the components even further (e.g., stem wood and stem bark), they used both equations to calculate the stem or combined the aboveground components to estimate aboveground mass. The compiled equations were used to generate pseudo-observations to fit the regional allometric models. An equal number of pseudo-observations were used for each equation to give equal relative weight in fitting regional models. Regional equations (pseudo models) for each species and component were fitted using linear mixed models. The fixed effects included the independent variables, such as tree diameter (D, cm) measured at 1.3 m. To correct for the bias introduced by back-transforming predictions of Y, a correction factor (CF) was applied. This CF was calculated as the ratio of the mean of the pseudo-observations to the mean of the values predicted by the regional equation (Snowden, 1991). The list of selected biomass metamodels are model fitting statistics are listed in Table 13. Equation 3 of Forrester et al. (2017), which is referred to in Table 13**,** is shown below as Eq. 1:

$$ln(y) = ln(\beta 0) + \beta 1\, ln(D) + \varepsilon \qquad \text{Eq. 1}$$

Table 13a, b and c: Allometric biomass equations by Forrester et al. (2017). a),b),c) show the parameter values and statistics for the selected metamodels. For convenience separate tables are provided. Abbreviations: sp_group is the Pathfinder sp_group number, 'Species' and 'Component' columns show the original species and components the model was developed, and 'Equation', 'ln(β0)', 'β for ln(d)' and CF, are model information and parameters for Eq.1. The rest of the columns are listed for readers convenience (i.e. quick-look-overview), and thus please see the original publication for the clarification of other column variable acronyms.

a)

sp_group	Species	Component	Equation	ln(β0)	β for ln(d)	CF	N	AIC	BIC	RMSE	RMSEreal	MAPE	Bias
1	Betula pendula Or B. pubescens	Stem mass	3	-2.32	2.34	0.97	337	-271.0	-255.7	0.14	0.26	9.99	-2.13
1	Betula pendula Or B. pubescens	Total branch	3	-3.82	2.33	1.15	317	106.9	121.9	0.25	0.54	23.18	18.14
1	Betula pendula Or B. pubescens	Foliage mass	3	-4.14	1.89	1.16	231	253.6	267.3	0.38	0.65	40.58	30.14
3	Fagus sylvatica	Stem mass	3	-1.45	2.17	1.00	429	57.7	73.9	0.24	0.32	15.24	2.67
3	Fagus sylvatica	Aboveground	3	-1.66	2.36	1.00	330	-220.0	-204.8	0.16	0.21	9.35	0.92
3	Fagus sylvatica	Foliage mass	3	-4.48	1.91	1.09	330	-52.2	-37.0	0.2	0.39	15.16	10.59
4	Quercus robur	Stem + branch mass	3	-2.91	2.74	1.03	134	-135.5	-124.0	0.12	0.26	7.7	3.43
4	Quercus robur	Stem mass	3	-2.71	2.53	1.01	100	-115.7	-105.4	0.11	0.2	6.27	1.34
4	Quercus robur	Foliage mass	3	-4.47	2.14	0.97	99	-56.0	-45.7	0.16	0.42	12.14	-2.16
2,5,6	Broad-leaved species	Stem mass	3	-2.45	2.41	0.94	2663	68.6	92.2	0.22	0.32	15.24	-3.44
2,5,6	Broad-leaved species	Total branch	3	-3.72	2.41	1.39	2496	2779.5	2802.8	0.38	0.65	55.43	49.55
2,5,6	Broad-leaved species	Foliage mass	3	-4.23	1.86	1.06	1824	1822.0	1844.0	0.36	0.54	27.92	14.32
7,8	Picea abies	Foliage mass	3	-2.80	1.87	1.02	1007	270.2	289.9	0.26	0.4	17.63	5.5
7,8	Picea abies	Stem + branch mass	3	-2.05	2.30	1.03	668	-533.4	-515.4	0.15	0.23	10.67	4.41
7,8	Picea abies	Stem mass	3	-2.50	2.34	1.06	444	-221.6	-205.3	0.17	0.44	12.42	7.66
9,10	Conifers	Foliage mass	3	-3.40	1.91	1.00	3910	2227.9	2253.0	0.3	0.45	20.96	4.96
9,10	Conifers	Stem mass	3	-2.62	2.42	1.06	2569	-1410.2	-1386.8	0.17	0.44	12.4	7.2
9,10	Conifers	Total branch	3	-3.63	2.22	0.99	2443	848.3	871.5	0.26	0.38	17.38	2.56

b)

sp_group	Species	Component	Confidence interval for intercept of Observed vs Predicted	Confidence interval for slope of Observed vs Predicted	R2	R2c	d range (cm)	BA range (m2 per ha)
1	Betula pendula Or B. pubescens	Stem mass	1.1 / 3.5	0.93 / 0.95	0.976624	0.993642	0.6 / 8.7 / 38	0.5 / 33.2 / 75
1	Betula pendula Or B. pubescens	Total branch	-0.9 / -0.3	1.07 / 1.09	0.920371	0.980829	0.6 / 8.3 / 38	0.5 / 33.3 / 75
1	Betula pendula Or B. pubescens	Foliage mass	-0.2 / -0.1	1.11 / 1.15	0.902837	0.939841	0.6 / 7.7 / 38	0.5 / 33.4 / 75
3	Fagus sylvatica	Stem mass	4.4 / 21	0.96 / 0.98	0.970869	0.988287	0.5 / 24.1 / 79.4	12.2 / 22.3 / 32.5
3	Fagus sylvatica	Aboveground	-26.7 / 0.6	1.01 / 1.02	0.981266	0.993925	1 / 27.3 / 84	0.5 / 29.4 / 75
3	Fagus sylvatica	Foliage mass	-1.8 / -1.2	1.12 / 1.17	0.883213	0.981926	0.8 / 29.8 / 73	0.5 / 28.2 / 70.1
4	Quercus robur	Stem + branch mass	-46.9 / -12.5	1.02 / 1.05	0.984866	0.996242	4 / 28.9 / 67.5	10.6 / 27.7 / 45.6
4	Quercus robur	Stem mass	-0.2 / 4.9	0.99 / 1	0.98936	0.996485	5 / 25.2 / 67.5	10.6 / 27.7 / 45.6
4	Quercus robur	Foliage mass	2.2 / 4.1	0.87 / 0.92	0.91076	0.984323	5.9 / 34.6 / 67.5	10.6 / 28.3 / 45.6
2,5,6	Broad-leaved species	Stem mass	13.5 / 19.1	0.92 / 0.93	0.970458	0.99135	0.1 / 16 / 85.9	0.5 / 29.6 / 70.1
2,5,6	Broad-leaved species	Total branch	-78.3 / -51.1	1.55 / 1.63	0.86948	0.977409	0.1 / 15.4 / 79.4	0.5 / 29.6 / 70.1
2,5,6	Broad-leaved species	Foliage mass	-0.1 / 0.1	0.99 / 1.01	0.846904	0.96808	0.1 / 16.8 / 88.2	0.5 / 33.3 / 75
7,8	Picea abies	Foliage mass	0.9 / 2.3	0.92 / 0.96	0.906286	0.969362	1 / 23.8 / 76.4	10.3 / 23.3 / 33.4
7,8	Picea abies	Stem + branch mass	-16.3 / -4.2	1.02 / 1.04	0.969919	0.991386	1 / 22.6 / 76.4	12.2 / 26.6 / 33.4
7,8	Picea abies	Stem mass	-21 / -12.5	1.08 / 1.11	0.965298	0.990763	1 / 19.8 / 70	12.2 / 26.6 / 33.4
9,10	Conifers	Foliage mass	1.6 / 2.3	0.89 / 0.9	0.830559	0.963918	1 / 23.3 / 163	2.4 / 28.5 / 73.8
9,10	Conifers	Stem mass	-18.8 / -9.2	1.03 / 1.03	0.962209	0.991642	1 / 24 / 163	0.5 / 29.2 / 75
9,10	Conifers	Total branch	3.1 / 5.5	0.92 / 0.93	0.930917	0.976858	1 / 23.2 / 163	0.5 / 28.2 / 70.1

c)

sp_group	Species	Component	TPH range (trees per ha)	Latitude range (degrees)	Mean annual precipitation (mm)	Mean annual temperature (degrees Celcius)
1	Betula pendula Or B. pubescens	Stem mass	223 / 5294 / 69527.4	32.9 / 49.9 / 69.3	525 / 680.5 / 809	2.2 / 5.4 / 9.5
1	Betula pendula Or B. pubescens	Total branch	223 / 5294 / 69527.4	32.9 / 49.9 / 69.3	381 / 875.6 / 2284	0.8 / 7.2 / 14
1	Betula pendula Or B. pubescens	Foliage mass	223 / 4639.9 / 69527.4	32.9 / 49.9 / 69.3	381 / 863.1 / 2284	0.8 / 7.7 / 14
3	Fagus sylvatica	Stem mass	1114 / 8756.3 / 16398.5	42.4 / 54.7 / 59.5	569 / 653.2 / 753	2.9 / 4.7 / 6.7
3	Fagus sylvatica	Aboveground	168 / 7401.6 / 69527.4	32.9 / 50.1 / 69.7	394 / 765.3 / 2315	1.2 / 7.8 / 15.9
3	Fagus sylvatica	Foliage mass	208 / 7002.3 / 69527.4	32.9 / 50.8 / 69.3	394 / 765.3 / 2315	1.2 / 7.8 / 15.9
4	Quercus robur	Stem + branch mass	375 / 9209.7 / 69527.4	32.9 / 50.9 / 65	418 / 717.5 / 1064	0.8 / 7.9 / 15.8
4	Quercus robur	Stem mass	375 / 9209.7 / 69527.4	32.9 / 50.9 / 65	630 / 726.8 / 801	5.1 / 7.3 / 8.2
4	Quercus robur	Foliage mass	167 / 9285.7 / 69527.4	36.4 / 50.9 / 64.2	630 / 726.8 / 801	5.1 / 7.3 / 8.2
2,5,6	Broad-leaved species	Stem mass	223 / 5421.4 / 17387.7	42.7 / 54 / 65	414 / 503 / 570	-1.1 / 0.6 / 2.2
2,5,6	Broad-leaved species	Total branch	223 / 5557 / 17387.7	42.7 / 55.7 / 65	381 / 870.5 / 2284	1.2 / 7.7 / 14
2,5,6	Broad-leaved species	Foliage mass	223 / 5294 / 69527.4	32.9 / 49.9 / 69.3	570 / 570 / 570	0.8 / 0.8 / 0.8
7,8	Picea abies	Foliage mass	1114 / 10857.2 / 27719.4	57.3 / 59.4 / 61.5	569 / 644.5 / 753	5.8 / 8.6 / 15.6
7,8	Picea abies	Stem + branch mass	1114 / 11771.5 / 27719.4	42.4 / 57.3 / 61.5	569 / 651.7 / 753	5.8 / 6.3 / 6.7
7,8	Picea abies	Stem mass	1114 / 11771.5 / 27719.4	58 / 59.8 / 61.5	394 / 765.3 / 2315	1.2 / 7.8 / 15.9
9,10	Conifers	Foliage mass	208 / 6990.6 / 69527.4	24.8 / 50.2 / 69.3	569 / 681.8 / 753	5.9 / 6.4 / 6.7
9,10	Conifers	Stem mass	208 / 6836.3 / 69527.4	32.9 / 50.6 / 69.3	640 / 654.7 / 669	2.9 / 3.2 / 3.4
9,10	Conifers	Total branch	208 / 7286.5 / 69527.4	32.9 / 50.5 / 69.3	640 / 654.7 / 669	2.9 / 3.2 / 3.4

Since below-ground-biomass (bgb) models were not available from Forrester et al., (2017) database, root-to-shoot ratios (of bgb to agb) from IPCC (2006, Table 4.4) were used (Table 14). Total bgb was calculated by multiplying agb with root to shoot ratio. Values from the IPCC report were used as is, except for dry forests in SOUTH region where the ratio value (i.e. R) was adjusted down to allow smoother transition between bgb values of between consecutive vol_bins. Coarseroot biomass component was calculated by subtracting fineroot biomass from the total bgb.

Table 14: Ratio of below-ground biomass to above-ground biomass from IPCC (2006), Table 4.4, used in this study. For the SOUTH region, modified R value was used instead of the default value.

Domain	Species / ecological zone	Above-ground-biomass (t/ha)	R [tonne root d.m. (tonne shoot d.m.)$^{-1}$]	R (modifed value)
Boreal (NORTH)	all	< 75	0.39 (0.23-0.96)	
	all	>= 75	0.24 (0.15-0.37)	
Temperate (CENTR)	conifers	< 50	0.40 (0.21-1.06)	
	conifers	50-150	0.29 (0.24-0.50)	
	conifers	> 150	0.20 (0.12-0.49)	
	quercus	> 70	0.30 (0.20-1.16)	
	eucalyptus	< 50	0.44 (0.29-0.81)	
	eucalyptus	50-150	0.28 (0.15-0.81)	
	eucalyptus	>150	0.20 (0.10-0.33)	
	broadleaf	< 75	0.46 (0.12-0.93)	
	broadleaf	>=75-150	0.23 (0.13-0.37)	
	broadleaf	>150	0.24 (0.17-0.44)	
Subtropical (SOUTH)	dry forest	< 20	0.56 (0.28-0.68)	0.40
	dry forest	>= 20	0.28 (0.27-0.28)	

2.7.2 Litter models

Annual foliar and non-foliar litterfall (less than 2 cm diameter) was calculated using models based on ICP Forest data. Yet for other litter components literature-based values were used. Equations to derive foliar and nonfoliage (i.e., bark, seeds, cones, etc.) litterfall were obtained from Task 3.3:

$$Foliage_litter_broadleaf = (\,175.803 + 3.873 * BA\,) * 0.01 \qquad \textit{Eq. 2.}$$

$$Foliage_litter_conifer = (\,114.423 + 3.116 * BA) * 0.01 \qquad \textit{Eq. 3.}$$

$$Nonfoliage_litter_broadleaf = (119.782 + 2.920 * BA\,) * 0.01 \qquad \textit{Eq. 4.}$$

$$Nonfoliage_litter_conifer = (\,98.586 + 1.178 * BA) * 0.01 \qquad \textit{Eq. 5.}$$

Fineroot litter input was assumed equal to that of the foliage litter calculated using Eqs. 2. and 3. If fineroot litter estimate exceeded fineroot biomass, it was set equal to fineroot biomass. In NORTH region, foliage litter of deciduous sp_groups was assumed equal to foliage mass, as all deciduous trees shed their leaves annually. For CENTR and SOUTH regions, it was assumed that if foliage litter exceeds foliage mass, it is equal to foliage mass. Since the range of BA values used to fit Eqs. 2-5 was 5.6 - 56.5 m^2/ha for broadleafs and 5 – 83.3 m^2/ha for conifers; some edits were necessary to obtain to sensible estimates for the smallest vol_bins. Thus, the foliage litter values of all conifer sp_groups second vol_bin were halved. For nonfoliage litter estimates, corrections were done for the two smallest vol_bins: The nonfoliage litter values of the smallest vol_bin were replaced with the respective vol_bin stem litter estimates (in the absence of better estimates), and the second smallest vol_bin nonfoliage litter estimates were halved.

Annual litter turnover rates of other biomass components were taken from the paper by Table 3 in Kutz et al. (2009), to match the numbers with CBM. For coarseRoot litter input, annual 2% turnover

was assumed. Branch litter input was calculated assuming an annual turnover of 3% for conifers and 4% for deciduous species. Stem litter input was calculated assuming an annual turnover of 0.45% for conifers and 0.67% for deciduous species.

2.7.3 Harvest residues

Harvest residues were calculated assuming all biomass of foliage, branches, coarseroot, fineroots and 10% of the merchantable stem volume (i.e. stump) will remain in the forest. Since EFDM employs 5-yr time-steps, harvests are assumed to be distributed evenly over the simulation period (5-years) (i.e. both harvest residues and litter input for Yasso are defined as annual).

To estimate harvest residues of thinned areas, a variable 'thinning_harvest_residue_percentage' (i.e., thin_hres_perc) was added for the coeftable. It quantifies the percentage of stem volume that is removed in thinning, and is calculated using coeftable variables of harvested stem volume (i.e. "thin_V_m3ha") and growing stock volume (i.e. "V_m3ha"). For final felling areas, all biomass of foliage, branches, coarseroot, fineroots and 10% of the merchantable stem volume (i.e. stump) was assumed to stay in the forest (ie., to go to the dead pool).

2.8 Preprocessing of the forest map data for EFDM

To realistically model the effects of forest management, forward predictions must be initialised using the measured current state of the forests. The initial state0 data for running EFDM was obtained from WP2, which mapped the volumes and probability coverage of each of the 10 sp_groups at 10m spatial resolution across EU region. Forest data were delivered as raster maps (48 tiles for each forest variable), which had been masked to cover only forest areas. A total of eleven raster map sets were used (i.e. separate raster sets were available for stem volume, and each of the ten sp_groups contained the percentage of that sp_group in each 10m resolution pixel).

Preprocessing of the stem volumes raster data contained the following steps: 1) Calculate the most abundant stem volume value for each vol-tile excluding NoData and Noforest pixels (32766, 32767), 2) Split each of the 48 tiles into 25 smaller tiles, 3) gapfill all NoData cells (32767) using larger tile mean volume , 4) classify volumes into vol_bins, 5) use EU country shapefile (CNTR_RG_20M_2024_3035.shp) to mask oceans which had been classified into NoData. This raster is later referred to as vRast.

Preprocessing of the raster datasets containing sp_group percentage values consisted the following steps: 1) Split each of the 48 tiles into 25 smaller tiles, 2) Create stacks the ten sp_group smaller tiles, 3) Set NoData and NoForest cells into NA, 4) Rename raster layers following sp_group numbers (i.e. 1:10), 5) Calculate dominant sp_group for each 10m pixel (random tie break), 6) Calculate dominant sp_group pixel-counts for all small tiles, 7) For gap filling purposes, aggregate dominant sp_group pixel-counts of the smaller tiles corresponding each of the 48 larger tiles. The resulting dominant sp_group raster will be referred to as spRast. The gap filling of missing sp_groups data is explained next.

Then, raster data preprocessing continued using both raster datasets: 1) Stack spRast (not-gapfilled) and vRast (gapfilled and country-masked), 2) Mask out parts of spRast where vRast lacks data, 3)

Combine masked spRast and Vrast, 3) Gapfill spRast pixels where vRast has data, using the larger tile dominant sp_group number.

Next, high spatial resolution pixels had to be assigned into low spatial resolution pixels (km^2), and aggregated to obtain distributions of sp_groups and vol_bin areas of low spatial resolution pixels: This was done as: 1) extract coordinates of the vRast grid centre points and assign them into low spatial resolution grid cells (1km*1km, CLUMondo FM map pixels), 2) convent raster data into datatable, 3) add spRast values into datatable, 4) change 1km cell coordinates to refer lower left corners, 5) define cell 'key' by combining X and Y coordinates of the grid lower left corners, 6) assign each high spatial resolution raster cell an area of 0.01 (ha), 7) aggregate areas by vol_bin, sp_group and grid cell key, 8) save output datatable.

Since EFDM data is needed at country-level, it was combined with the $1km^2$ mapping of countries defined based on CLUMondo FM raster data, which already had grid-cell classification into countries. For forest areas, which were not classified into countries, the closest country grid cells were used to assign country information. Then, all data were split into country-wise folders.

Noteworthy is that the WP2 based estimates of forest total area differ from official statistics (e.g., FAO 2018-2021 (FAO, 2025)) (Table 15), since the Copernicus masks (Copernicus, 2025), used by the WP2 map creation process, are essentially tree-cover masks, rather than forest masks (which is a land use issue, and more of a political/administrative decision). Also, since the country mapping was done using $1km^2$ cells, the differences in forest areal estimates are very large for countries that have little forest area such as Cyprus, Luxemburg and Malta (Table 15)

Table 15: Forest area by countries. Abbreviations: Country = country acronym, Country name= name of the country, Area = WP2 map based estimate of forest total area at 1 km2, FAO = forest areal estimate reported by FAO, Dif = absolute difference between (Area – FAO), and Dif (%) = percentage difference of areal values compared to FAO area. FAO statistics from the source were not provided for all countries, and thus those columns are empty.

Country	**Country name**	**Area (1000 km^2)**	**FAO (1000 km^2)**	**Dif**	**Dif (%)**
SE	Sweden	262.88	279.8	-16.9	-6.05
FI	Finland	211.88	224.1	-12.2	-5.45
FR	France	178.28	177.0	1.3	0.73
ES	Spain	158.17	185.8	-27.6	-14.85
NO	Norway	129.72	121.3	8.4	6.96
DE	Germany	120.62	114.8	5.8	5.06
IT	Italy	114.72	96.2	18.5	19.25
PL	Poland	109.81	94.8	15.0	15.87
RO	Romania	84.61	69.9	14.7	21.06
EL	Greece	53.22	39.0	14.2	36.39
BG	Bulgaria	48.38	39.0	9.4	23.99

AT	Austria	38.75	38.9	-0.2	-0.44
LV	Latvia	34.09	34.1	-0.1	-0.16
CZ	Czech Republic	30.83	26.8	4.0	15.09
PT	Portugal	29.62	33.4	-3.8	-11.29
BA	Bosnia&Herzegovina	29.54	15.9	13.7	85.99
HR	Croatia	26.08	19.4	6.6	34.16
EE	Estonia	23.74	24.4	-0.6	-2.63
HU	Hungary	23.22	20.7	2.5	12.03
SK	Slovakia	23.20	19.5	3.7	18.67
LT	Lithuania	23.17	22.1	1.1	4.87
SI	Slovenia	12.64	11.8	0.8	7.12
MK	North Macedonia	12.09	10.1	2.0	19.74
IE	Ireland	9.13	8.1	1.0	12.68
DK	Denmark	8.53	6.9	1.6	23.79
BE	Belgium	8.38	6.9	1.5	21.61
NL	Netherlands	4.41	3.6	0.8	21.27
CY	Cyprus	2.65	1.7	0.9	54.52
LU	Luxembourg	0.99	0.9	0.1	5.92
MT	Malta	0.01	0.0	0.0	120.86
VA	Vatican City	0.00			
MC	Monaco	0.00			
RS	Serbia	37.60			
MD	Moldova	33.20			
UK	United Kingdom	31.98			
CH	Switzerland	13.80			
AL	Albania	10.74			
ME	Montenegro	7.51			
AD	Andorra	0.18			
LI	Liechtenstein	0.07			

SM	San Marino	0.02			

At the European level, volume class distributions seem sensible (Table 16). Differences in actual forest volumes and those obtained from WP2 data may originate from the fact that some regions did not have their own NFI plots, so forest properties were then predicted based on data from other countries (For example, the Balkan countries of Serbia and Montenegro). Also, mapping very mature forest areas such as those in e.g., Romania and Central Europe, is by itself a difficult task as satellite pixel reflectance values saturate in forest that have a lot of volume. These have probably been mostly classified based on Austrian experimental plots. Same applies to Cyprus, which seems to have relatively high stem volumes. Evaluation of tree species distributions at European level is perhaps even more challenging as each sp_group is a combination of many different tree species. Pine (sp_group=9) and birch (sp_group=1) appears dominating in North, while Quercus (sp_group=4) seems dominating large parts of CENTR and SOUTH Europe. Some errors in sp_groups dominance may have resulted from the preprocessing of the WP2 data, as e.g., gap filling of NoData with most dominant sp_groups (i.e. the most abundant become even more abundant).

Table 16: Dominant volume classes and sp_groups by countries. Please see section 2.4. for the definition of vol classes (i.e. 0, 1, 10, 25, 48, 80, 122, 204, 313, and 517 m3/ha (the intervals closed on the left and open on the right), and section 2.1 for description of sp_groups.

Country	**Country_name**	**Most common vol class**	**Sp_dom1**
AD	Andorra	7	9
AL	Albania	5	4
AT	Austria	9	7
BA	Bosnia&Herzegovina	8	3
BE	Belgium	8	4
BG	Bulgaria	5	4
CH	Switzerland	9	3
CY	Cyprus	5	4
CZ	Czech Republic	8	7
DE	Germany	8	3
DK	Denmark	7	1
EE	Estonia	7	1
EL	Greece	5	4
ES	Spain	3	4
FI	Finland	7	9
FR	France	7	4

HR	Croatia	8	3
HU	Hungary	8	6
IE	Ireland	7	7
IT	Italy	5	4
LI	Liechtenstein	9	7
LT	Lithuania	8	6
LU	Luxembourg	8	3
LV	Latvia	8	1
MC	Monaco	5	9
MD	Moldova	6	3
ME	Montenegro	9	3
MK	North Macedonia	5	4
MT	Malta	4	5
NL	Netherlands	8	4
NO	Norway	4	1
PL	Poland	8	9
PT	Portugal	3	4
RO	Romania	9	3
RS	Serbia	8	4
SE	Sweden	7	9
SI	Slovenia	8	3
SK	Slovakia	8	3
SM	San Marino	5	4
UK	United Kingdom	7	7
VA	Vatican City	5	4

2.9. EFDM scenario modelling framework

Set-up the EFDM was the following: EFDM scenarios were run using a loop which updates P0-values and FM classes of all gridcells at the beginning of each 5-yr period following the given input data. For each gridcell the closest P0 values were searched from the input file containing data, which contains precalculated P0-values for the time-step mean-years between 2020-2100. The number of

preselected locations was 35657 (i.e. the same as in FORMIT-M). These P0_values were then classified into predefined P0_classes, of which class thresholds and number of classes vary between sp_groups and regions (see section 'Noman' for further details). Grid cell FM_classes were read from a separate file, which contains FM_classes of all grid cells for each period (after initial FM_classes were assigned to all pixels, only FM_classes which had changes in their FM_class values were updated for EFDM). Thus, the EFDM was run using a single time-steps, and an output files from each simulation round were saved to separate files to derive Yasso inputs.

3. Soil carbon pool modelling using Yasso

3.1 Linking EFDM and Yasso

Linking EFDM with Yasso ensures a consistent representation of simulated forest management effects between above- and below-ground C pools. Yasso (2025), also called as Ryassofortran provides convenient R-functions for calling the Fortran90-release of the soil carbon model Yasso20 (there is an error in the R-package version description). The Yasso model (Tuomi et al. 2009, Tuomi et al. 2011) was developed for even faster prediction of C cycling in mineral soil, litter and harvest residues using EFDM inputs. In Yasso, the decomposition of the different components of C inputs is modelled based on climate, chemical composition, and size of woody parts. Yasso is by default parameterized for conifer and deciduous species, and thus a variable called sp_group_rev (deciduous coded as 2, conifers as 1) was introduced to allow aggregation of the data of the ten sp_groups into these two groups. Yasso contains by default six litter inputs (i.e. stem, branch, root, nonfoliage, fineroot an foliage), and litter element size parameter varies between different input types. While this parameter is set as zero for fineroots and foliage, it is notably larger for other woody litter types (i.e. stem = 20, roots = 7, branch = 4, and non-foliage = 2).

Preprocessing of the EFDM outputs to Yasso inputs was done after running a full EFDM simulation containing 7 timesteps (2020-2050). Coeftable was used to convert the areal changes of different activities into forest structural properties. For each time-step EFDM output table was merged with coeftable by "country_group", "sp_group", "vol_bin". From the output table, it was possible to calculate all result values as e.g., as tonnes per hectare (tha) or hectares per gridcell (1 km^2). Yasso litter and harvest residue inputs were calculated from EFDM outputs by aggregating total gridcell sums of stem, branch, root, nonfoliage, foliage, and fineroot Yasso inputs.

For noman areas, results were obtained directly by multiplying areas with coeftable values, and aggregating results by sp_group_rev (i.e. two groups of conifer and deciduous species), key (i.e., gridcell), time, and forest management class (i.e FM_class). For the calculation of harvest residues for final felling and thinning areas, an additional step was needed. This step contained additional multiplication of coeftable mass values (e.g. branch litter tha) with respective ‘thin_hres_perc’ value, which denotes vol_bin, sp_group and region - specific percentage of volume that is removed in thinning. The resulting tons per hectare coefficients (by vol_bin, sp_group and region) were multiplied with respective vol_bin areas, to obtain estimates in tonnes, and aggregating the values by conifer and deciduous species (i.e. sp_group_rev), key, time, and FM_class. Noteworthy is that as litter inputs for Yasso are annual, thus also EFDM-based harvest residues estimates were divided with 5 (i.e. number of simulations years). Harvest residues were summed with the litter inputs prior Yasso simulation. Since EFDM activities happen between time-steps of t and t+1, the calculated harvest residue estimates from time-step t were added to the Yasso biomass input for the t+1 period. As EFDM first year 2020 does not have any harvest residues, required for Yasso stable state calculation, the initial year for Yasso simulations is 2025.

Yasso temperature and precipitation data selection was done in collaboration with task 3.5. Based on investigations done by task 3.5, the model combination MPI-M-MPI-ESM-LR/SMHI-RCA4_v1 was selected. This set of simulations has climate evolution close to the ensemble mean (see e.g. IPCC, 2025), and has a fairly good general performance (Vautard, et al., 2021). Climate was obtained from

EURO-CORDEX for rcp45 (search_url='https://esg-dn1.nsc.liu.se/esg-search/wget/?distrib=false&dataset_id=cordex.output.EUR-11.SMHI.MPI-M-MPI-ESM-LR.rcp45.r1i1p1.RCA4.v1a.mon.tas.v20160803|esg-dn1.nsc.liu.se'). The units in the climate data ncfiles from CORDEX are Kelvin and “kg m-2 -s”. To get temperature in Celsius the annual mean values were subtracted 273.15, and for annual precipitation sum the precipitation in mm the values were multiplied with the number of seconds in a year (i.e. 31540000). Noteworthy is that, the potential maximum GPP used in EFDM forest growth modelling was parameterized using different model combination (Härkönen et al. (2019): “The weather data was generated by the MPI-M-MPI-ESM-LR model version CCLM4-8-17”). To match Yasso climate data with EFDM maxGPP data, Yasso climate data was obtained for the same period mean years and locations as the potential maxGPP values (i.e. P0_values). Similar to EFDM model application, for each gridcell location the closest climate data (i.e. 5-year means of annual temperature means [C] and precipitation sums [mm]) were searched from the predefined climate data table. Yasso climate data was input followingly: For a particular location and time step, one would supply temperature as a 1-dimensional matrix with 12 constant values and precipitation as a single integer. Thus, running X-timesteps, the temperature matrix contains X rows representing gridcell temperature values at one particular time-step, and precipitation is a vector length of X.

4. Resilience against disturbance

4.1 Modelling bark beetle disturbance

4.1.1 Background

Outbreaks caused by the European spruce bark beetle (*Ips typographus*) can result in extensive tree mortality, significantly impacting forest ecosystems and leading to substantial economic losses. The frequency and severity of these outbreaks are magnified under climate change, particularly with rising temperatures and more frequent extreme drought events. Numerous predictive models have been developed to investigate the relationship between host tree characteristics and bark beetle-induced mortality at the plot level (Francis et al., 2025). These models typically incorporate host tree attributes, stand structure, and climatic variables to estimate the probability of tree attack and subsequent mortality. Recent research efforts have focused on simulating bark beetle outbreaks, leading to the development of models implemented in an R package (*IpsR*: available as an installable library) to estimate the biomass affected and killed by bark beetle infestations.

Past studies used a variety of approaches to model the impacts of bark beetles on forests. Some models treated bark beetle outbreaks as background mortality, while others dynamically modelled these outbreaks within the ecosystems. Dynamic modelling of beetle outbreaks provides a more comprehensive view by incorporating the lifecycle of bark beetles, tree defence mechanisms and ensuing alterations in forest composition and functionality (Marie et al., 2023).

4.1.2 Method

The bark beetle model developed in Pathfinder originated from the LandClim landscape model (Temperli et al., 2013) developed further by Marie et al. (2024) and implemented in the ORCHIDEE land surface exchange model. It is used for calculating the annual biomass of spruce trees attacked and killed by bark beetles. The model requires as inputs the volume or mass of living wood, the volume or mass of wood killed within the last year by wind, drought or bark beetles (provides food for the bark beetles), the proportion of Norway spruce (*Picea abies*) in the stand or forest, the mean diameter of the Norway spruce trees, the stand density, the number of effective degree days when the beetles are active and the maximum of the drought stress in the last three years (Figure 3: Workflow graph of calculations and responses in the bark beetle outbreak model in the ORCHIDEE land surface exchange model (Marie et al., 2024)).

The model incorporates key forest attributes and structural characteristics, including average tree diameter, plot or stand density, and forest type (primarily distinguishing between boreal and temperate forests). Figure 3: Workflow graph of calculations and responses in the bark beetle outbreak model in the ORCHIDEE land surface exchange model (Marie et al., 2024) shows that host attractiveness is influenced by water stress, a species-specific spruce factor, and a proxy for reduced nitrogen and carbohydrate reserves resulting from intense resource competition. These variables collectively determine the susceptibility of a stand to colonisation by a new bark beetle population.

Climatic factors are recognised as critical drivers of bark beetle development and outbreak dynamics. One such variable is *DDeff*, which represents the cumulative effective degree days required for beetle reproduction. The calculation of *DDeff* is based on the formulation in Wermelinger & Seifert (1998). The index is calculated using various temperature metrics, including air temperature (*Tair*), maximum and minimum developmental thresholds (*Tmax* and *Tmin*), bark temperature, and the

optimal bark temperature for beetle development. Drought is a major trigger of bark beetle outbreaks. Recent studies have demonstrated that water-limited conditions can significantly impair Norway spruce resistance to bark beetle attacks. Accordingly, the model accounts for the maximum monthly drought intensity over a three-year period. The drought intensity and the subsequent level of plant water stress (*PWSmax*) were calculated using the Standardised Precipitation Evaporation Index (Vicente-Serrano et al., 2010) by converting *SPEI* from a range of -2 (strong drought) to 2 (no drought).

Mixed-species forests have been shown to be more resilient to bark beetle infestations than pure stands of Norway spruce (Nardi et al., 2023). Therefore, the proportion of Norway spruce (*FN* = 0-1) within a stand is included as an indicator of vulnerability. Additional model inputs require inventory data from previous years, including the total dead biomass that has remained in the stand for less than one year, and the total living biomass of Norway spruce. The model estimates the probability of a stand being attacked and the corresponding biomass affected. Finally, by incorporating the probability of successful attack and tree mortality, the model provides an estimation of the total biomass killed by bark beetles.

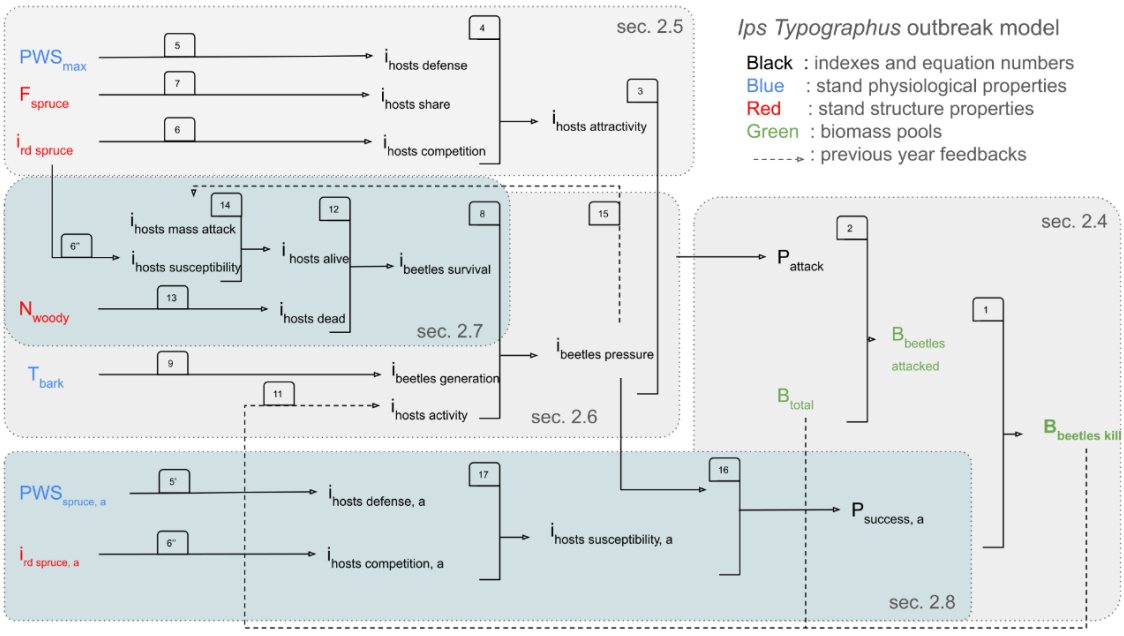


Figure 3: Workflow graph of calculations and responses in the bark beetle outbreak model in the ORCHIDEE land surface exchange model (Marie et al., 2024)

4.2 Wind-storm models

4.2.1 Background

The increased understanding of the mechanisms of wind damage to trees and forests and the availability of good quality data following windstorms have enabled the construction of an accurate predictability of the level of the risk to forests. Empirical models and hybrid-mechanistic models have

been developed and widely used (Gardiner et al., 2008). ForestGALES is a computer-based decision support system that has been used to evaluate the risk of wind damage to forests in the United Kingdom and has been adopted in several other countries, facilitating the improvements and comparison of impacts on different silvicultural practices (Hale et al., 2015). Such models are useful for developing hypotheses and investigating component processes, organising experimental results and making predictions for new management scenarios (Gardiner et al., 2008). In Pathfinder, we use the R version of ForestGALES (*fgr*) that is available for download from Forest Research in the UK (https://www.forestresearch.gov.uk/tools-and-resources/fthr/forestgales/).

4.2.2 Methods

ForestGALES uses a standard risk framework, where risk is determined by both the vulnerability of a system and the probability of the hazard occurring at a level sufficient to cause damage. The model calculates risk in two steps:

- Vulnerability Assessment: The model estimates the critical wind speed (*CWS*) at which trees are likely to be damaged—either by uprooting or stem breakage, by using biomechanical models parameterised with tree pulling experiments (Nicoll et al., 2006) and green wood strength measurements (e.g. Lavers, 1969). The information needed for the calculation includes tree species, tree height, stem diameter, between-tree spacing, soil type and rooting depth, and the presence of forest edges and the width of any upwind gap (Locatelli et al., 2022). The model incorporates tree-specific parameters for calculating tree volume, stem taper, root anchorage strength, and crown characteristics to estimate the wind force required to cause structural failure. The calculations account for aerodynamic drag and the mechanical stress exerted on stems and roots, derived from fundamental principles of fluid dynamics and material resistance. By simulating these physical interactions, ForestGALES/*fgr* predicts the wind speed threshold at which damage has a 50% probability of occurring for either individual trees within a stand or the entire stand treated together.
- Hazard Assessment: To assess the wind hazard, ForestGALES/*fgr* analyses the local wind climate, typically characterised by a Weibull distribution of hourly mean wind speeds at the stand's location. In British forestry, this is often simplified using the DAMS (Detailed Aspect Method of Scoring) system (Quine and White, 1993), which classifies sites by wind exposure based on topography and regional wind data. However, wind damage is primarily caused by extreme gusts, not mean winds. The relationship between mean winds and gusts—and the factors influencing this interaction—has been widely studied in tree biomechanics and wind engineering. ForestGALES/*fgr* integrates these processes to provide a conversion from the calculated mean wind load to the estimated gust load as a function of tree height, between tree spacing, distance from any forest edge and the size of any upwind gap (Quine, Gardiner and Moore, 2021).

The risk of wind damage is computed as the probability that the local wind climate (including allowance for gusts) will exceed the stand's critical wind speed. This probability is expressed as a return period—the estimated frequency of damaging storms—based on historical wind data and future climate projections. By coupling these physical and statistical approaches, the model provides actionable insights for forest management, such as identifying high-risk stands or optimising thinning regimes to mitigate windthrow.

4.3 Modelling tree mortality influenced by drought

4.3.1 Background

Although there are a number of existing physiological studies which have examined drought-induced tree mortality at local scales, quantitative understanding across large spatial and temporal scales in Europe remains limited. Recent research has primarily focused on single ecosystems or a few species, often neglecting long-term dynamics, creating gaps in our ability to generalise drought effects in heterogeneous forests (Senf et al., 2020). However, studies such as Senf et al. (2020) and Lawal et al. (2024) have begun addressing this by analysing 30-year climatic trends, linking drought indices (e.g., SPEI, HWMI, and climatic water balance) to tree mortality across diverse species. These approaches integrate forest attributes (e.g., tree size, genus) with climatic variables to quantify mortality risk, providing a framework for large-scale, probabilistic assessments of drought impacts.

4.3.2 Methods

The model incorporates two drought-related climate indices: the Standardised Precipitation Evapotranspiration Index (SPEI) (Vicente-Serrano et al., 2010) and the Heatwave Magnitude Index (HWMI) (Russo et al., 2014)

The study utilises French NFI data (2002–2020; ~19,000 plots) for model calibration, supplemented by inventory data from Germany and to assess model predictability. Preliminary analysis identified outliers (e.g., trees >100 cm DBH, excluded via 0.025 percentile trimming) and data imbalances. Key attributes include diameter, height, species, and tree status (measured at tree level, aggregated to plot level: quantitative metrics as means, qualitative metrics as dominant categories). Species were consolidated into eight dominant genera (Oak, Pine, Beech, Ash, Fir, Spruce, Birch, European larch). Missing data (e.g., tree density, age) were noted as potential limitations.

SPEI and HWMI values were matched to plot coordinates using nearest-neighbour interpolation (11 km grid resolution, converted from NetCDF to CSV). While computationally efficient, this method introduces spatial aggregation bias, as multiple plots often share identical climate values due to the low spatial resolution of the climate data.

We employed empirical modelling to assess the contributions of drought (SPEI) and extreme heat (HWMI) to tree mortality using two statistical approaches: Random Forest (RF), Generalised Additive Models (GAM).

Parametric Alternatives (GAM): GAMs accommodated non-linearity (e.g mortality vs. SPEI). Though less robust to imbalanced data, the model helped quantify predictor effects and explain variance.

Ensemble Methods (RF): Ensemble tree-based algorithms excel at handling correlated predictors (e.g., multiscale SPEI/HWMI at 0–36 month lags) and require minimal assumptions. Data were pre-processed via standard scaling to ensure equal feature weighting and oversampling to address class imbalance. Hyperparameters (e.g., tree depth, number of parameters) were optimised via grid search with cross-validation, and performance was evaluated using classification metrics (ROC-AUC).

4.4 Forest Fire Weather Index (FWI) and Forest Fire Danger Rating System (CFFDRS)

4.4.1 Background

With respect to forest fires, the impacts of warmer and drier weather observed in the past are expected to become stronger in the future under projected climate change (Khabarov et al., 2016). Fires are one of the main disturbances that affect terrestrial ecosystems and have profound consequences on vegetation structures and functioning. The intensity and duration of drought, bark beetles and windthrow will probably intensify the susceptibility of forest fires, as they provide additional fuel for fire ignition. Meteorologically based fire danger indices can evaluate and summarise the fire danger considering past and current weather and provide seasonal summaries to compare the overall wildfire potential of a given year due to meteorological conditions (Camia et al., 2017)

4.4.2 Methods

The CFFDRS is a comprehensive, science-based system used in Canada and adopted or adapted in many other countries to assess wildland fire danger and predict fire behaviour. Available package *cffdrs* in R provides fire managers with critical information for preparedness, resource allocation, and operational decision-making (Wang et al., 2017). The CFFDRS allows the quantification of the likelihood of fire ignition, fire spread, difficulty of control, and potential impact. It also provides estimates of how a fire might behave under specific conditions (e.g., rate of spread, intensity).

The CFFDRS is a series of four linked subsystems that integrate various factors influencing fire danger, presented in **Error! Reference source not found.**:

Fire Weather Index (FWI) System: This is the core meteorological component, providing a daily numerical rating of fire danger based solely on weather conditions.

Fire Behaviour Prediction (FBP) System: This subsystem provides quantitative estimates of fire behaviour (rate of spread, fuel consumption, frontal fire intensity) for specific fuel types and topography. It uses outputs from the FWI System as inputs.

Accessory Fuel Moisture (AFM) System: This component addresses moisture in specific fuel types not directly covered by the FWI's general fuel layers (e.g., specific grass types, logging slash).

Fire Occurrence Prediction (FOP) Systems: These are regionally developed systems that predict the likelihood of lightning- and human-caused fires based on historical data and current conditions.

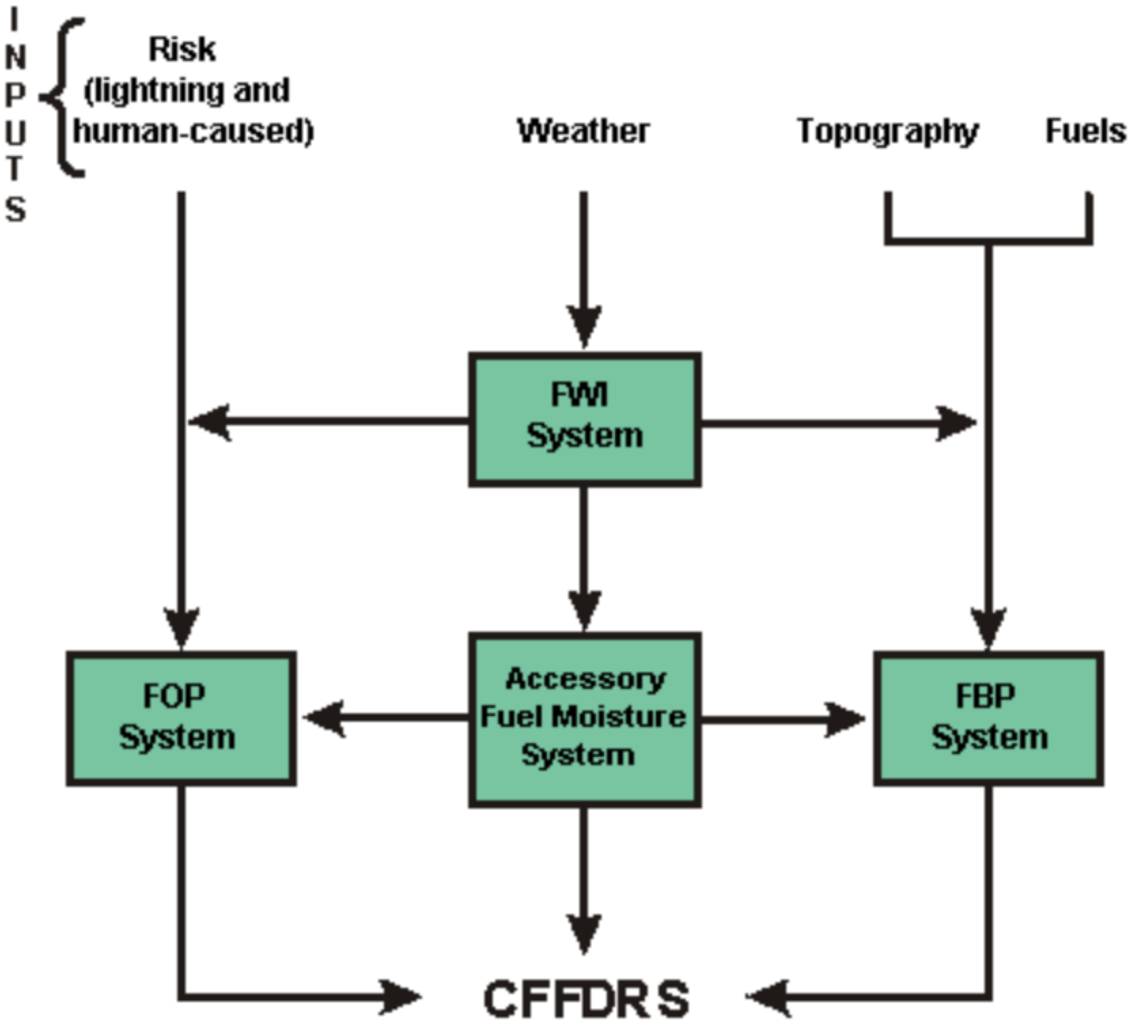


Figure 4: Structure and components of the Canadian Forest Fire Danger Rating System

The CFFDRS will be implemented once outputs from other models, specifically those providing detailed fuel information, are produced. Since fuels and fuel types are critical inputs for the FBP System, they are essential for assessing key outputs such as fuel consumption. This necessary fuel input can be derived from computations of forest damage, including volume affected by windthrow, bark beetle infestations, and drought. While the full CFFDRS is yet to be implemented in this project, it is recognised as an essential model for incorporating and understanding the effects of various forest disturbances on fire danger and behaviour and will be run once outputs of the other three models are available.

4.5 Climate indices

4.5.1 Regional Climate modelling

The EURO-CORDEX ensemble of regional climate model simulations at 0.11 degree horizontal resolution covering Europe. Each regional climate simulation is forced by a global climate simulation. The simulations of CMIP5 are based on Representative Concentration Pathways (RCPs) (Vautard et

al., 2021). A set of 3 pathways is used in this project, which span the range of radiative forcing values from 2.6 to 8.5W/m2 in 2100. RCP8.5 describes a continuous increase in greenhouse gases, RCP2.6 represents a mitigation scenario that leads to a radiative forcing around 2040 of 3W/m2 and reduces to the end of the century to 2.6 W/m2, and RCP 4.5 represents a stabilisation (without overshoot) in radiative forcing at 4.5W/m2 post 2100. Simulations were run for climate scenarios RCP45 and RCP85 calculated with the MPI-M-MPI-ESM-LR German global climate model and the Swedish SMHI-RCA4 regional climate model, as discussed in Vautard et al. (2021)

4.5.2 Standardised Precipitation Evaporation Index (SPEI)

It is a meteorological drought index that includes precipitation and temperature data developed by (Vicente-Serrano et al., 2010). This represents a simple climatic water balance that is calculated at different time scales to measure both wet and dry moisture extremes. We calculated SPEI by using gridded monthly climate data from EURO-CORDEX at spatial resolution of 0.11 degree spanning from 1986–2020 and includes monthly mean temperature (T_{mean}), maximum temperature (T_{max}), minimum temperature (T_{min}), relative humidity, potential evaporation (*PET*), and precipitation (*P*), provided by dynamically downscaled GCM-RCM combinations selected for their performance in European climate simulations. The computation of monthly SPEI is by subtracting *PET* from *P* (Eq. 6), the aggregate measurement of SPEI at different timescales is obtained through a mean of individual monthly timescales. For example, a 12-month timescale is obtained from the mean of the previous 12 months, including the current month. As SPEI is an essential factor for both bark beetle model and drought mortality model, we calculated SPEI mean and SPEI min at 12-month, 24-month, and 36-month scales to provide alternative depictions of the climatological phenomenon (e.g., hot drought-triggered tree mortality in certain time periods) (Lawal et al., 2024).

$$SPEI = P - PET \quad \text{Eq. 6}$$

4.5.3 Effective Degree Days (DDeff)

DD_{eff} represents the sum of effective temperature for bark beetle reproduction in °C days. DD_{eff} is calculated from 1 January until the diapause of the first generation. A diapause is triggered when the day length exceeds 14.5 h. Each day before the diapause with a daily average bark temperature above 8.3 °C (T_{min}) and below 38.4 °C (T_{max}) is included in the calculation of DD_{eff} (Eq. 7). This simulates the phenology of bark beetles, which breed earlier when winter and spring are warmer, allowing for multiple generations in the same year (Hlásny et al., 2021). The sum of degree-days for one bark beetle generation is equal to 547 °C days.

$$DD_{eff} = \sum_{n_{diapause}}^{i=1} \left(T_{opt} - T_{\min}\right) \cdot \left(e^{\left(0.0288 \cdot T_{bark,i}\right)} - e^{\left(0.0288 \cdot b_{eff} - \left(40.99 - T_{bark,i}\right)/3.59\right)} - 1.25\right) \quad \text{Eq. 7}$$

where *i* is each day, and $n_{diapause}$ is the number of days between the 1 January and the day of the diapause. T_{opt} (30.3 °C) is the optimal bark temperature for beetle development, and T_{min} (8.3 °C) is the temperature below which beetle development stops. $T_{bark,\,i}$ is the average daily bark temperature calculated as the daily average air temperature plus 2 °C. All parameter values are from Temperli et

al. (2013) and Baier et al. (2007). Note that there are differences between this formulation and the one in (Marie et al., 2023),which contains a few errors.

4.5.4 Heatwave Magnitude Index (HWMI)

HWMI quantifies heatwave intensity and duration into a single numerical value, which allows comparison of heatwaves across different regions and periods; thus, *HWMI* is also a multiscale index (Russo et al., 2014). A heatwave is typically defined as a period of at least three consecutive days where the daily maximum temperature (T_{max}) exceeds a specific threshold. This threshold is usually the 90th percentile of daily maximum temperatures for a 31-day running window centred on that day, calculated over a historical reference period (e.g., 1981-2010). This percentile-based approach accounts for local climatological conditions. Once the daily thresholds are established, the actual daily T_{max} time series is compared against these thresholds. Any period of three or more consecutive days where the T_{max} consistently remains above its corresponding daily threshold is identified as a heatwave event. The *HWMI* daily value is then calculated as the sum of the magnitude of consecutive days comprising a heatwave (Eq. 8), and in addition, considered 25th and 75th percentile thresholds derived from the distribution of annual maximum temperatures over the reference period, adopted from a study of Lawal et al. (2024) (Eq. 9). The index is scaled from 0 (normal) to 100 (ultra-extreme) and was calculated using the *extRemes* package in R. *HWMI* at 12-month, 24-month and 36-month scales are calculated to evaluate the lasting effect on tree mortality.

$$A_d = \bigcup_{y=1981}^{2010} \bigcup_{i=d-15}^{d+15} T_{y,i} \qquad \text{Eq. 8}$$

$$M_d\left(T_d\right) = \begin{cases} \dfrac{T_d - T_{30y25p}}{T_{30y75p} - T_{30y25p}} & \text{if } T_d > T_{30y25p} \\ 0 & \text{if } T_d \leq T_{30y25p} \end{cases} \qquad \text{Eq. 9}$$

4.6 Model Operation

All models were created in the R environment to allow for ease of operation, ease of data exchange, and to allow for sharing of the models during and after the Pathfinder project. All analyses were conducted in R 4.4.3. Figure 5: Schematic of the potential interaction between disturbance models and growth models within the PATHFINDER project. gives a graphical diagram of the potential links between climate and soil data and the EFDM growth model and the disturbance models.

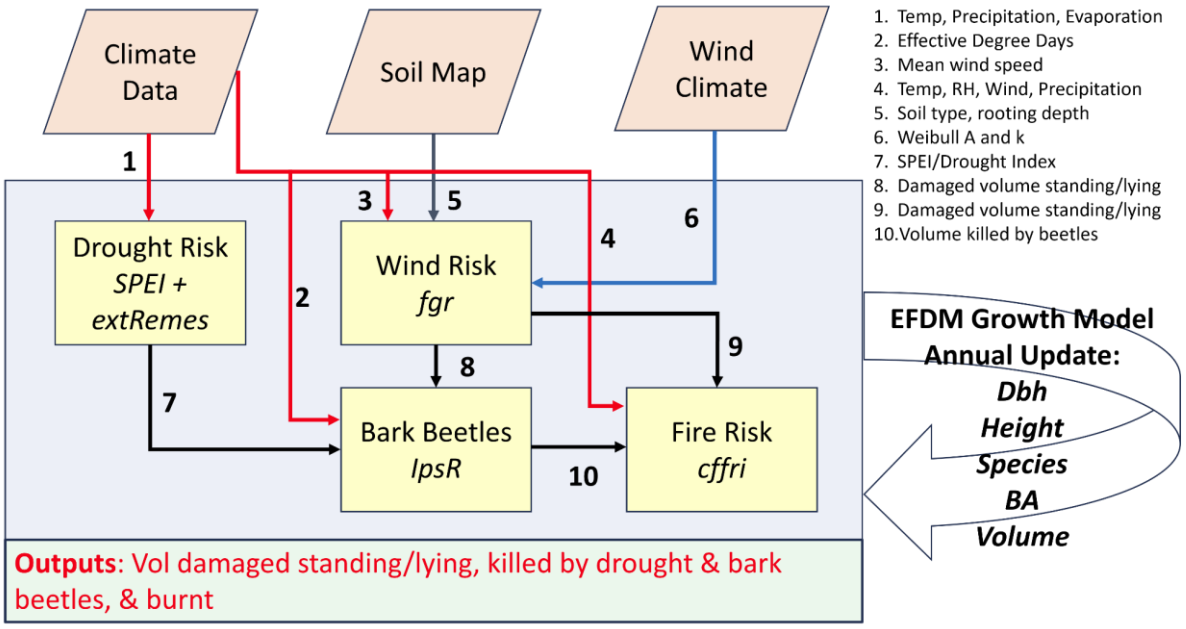


Figure 5: Schematic of the potential interaction between disturbance models and growth models within the PATHFINDER project.

Currently, the following model packages are available in R to use in modelling forest disturbance:

- *fgr*: Wind damage risk model
- *IpsR*: Bark beetle damage risk model
- *Cffri*: Canadian forest fire risk index model
- *SPEI*: Package for calculating the Standardised Precipitation Evaporation Index (*SPEI*)
- *extRemes*: Package for calculating a number of climatic extremes, including the Heat Wave Magnitude Index (*HWMI*).

The drought model will be made available as an R package after model completion and validation.

5. Forest biodiversity

5.1 Composition and structure

5.1.1 Indices integration in EFDM

The absence of NFI tree-level data from most countries severely limits the development of studies to analyse the shifts in main native European tree species and to forecast main horizontal and vertical forest structural indices, as initially planned.

However, we have integrated a functional composition index and a structural diversity index, considering EFDM outputs, to obtain congruent results (see Section 5.4, Composite Index).

Additionally, we have conducted a study to evaluate the performance of structural indices under different nested designs to identify the best operational structural indices that could be applied based on the NFIs data sources.

5.1.2. Performance of forest structure-related indices under nested plot designs

5.1.1.1 Introduction

Forest structure is a complex term described by various features like stand variables (e.g., basal area, tree density) and, more specifically, by spatial tree distribution, species composition, and tree size diversity (del Río et al. 2016). Unlike simple stand variables, forest structure is not directly measurable and requires surrogate indices (Pommerening 2002; Hui et al. 2019; Lin et al. 2020).

Accurately describing forest structure requires extensive sampling. National Forest Inventories (NFIs) are key for providing unbiased estimations over wide areas (Moreno-Fernández et al. 2020a). However, NFI plot designs vary, with fixed-area sampling (all trees above a minimum size) and nested designs (concentric subplots with increasing minimum diameter thresholds) being common (Tomppo et al. 2010).

A challenge arises with nested designs, as traditional structural indices are not valid anymore. Methodological approaches are being developed to correct this bias, often by accounting for tree sampling probability (Moreno-Fernández et al. 2021). However, information on the performance of distance-independent indices in nested designs is scarce but highly valuable. This is crucial given international agreements and regulations that increasingly require accurate structural indicators applicable across various monitoring networks.

Therefore, we aim to compare species composition and tree size diversity indices in nested designs with those in fixed-area designs using a simulation approach to understand their performance.

5.1.1.2 Simulation scheme

We simulated 1-hectare forest plots to study different forest structures. We then adjusted three key forest characteristics:

- Species richness (S): From single species plots up to 10 tree species.
- Basal area (BA): Ranging from 10 to 50 m^2 ha^{-1} in 5 m^2 ha^{-1} increments.
- Tree density (N): From 100 to 2,000 trees ha^{-1} in 100 tree ha^{-1} increments.

To keep the simulations simple, all species within a plot had an equal number of trees. We also considered three distinct spatial patterns:

- Regular: Trees were placed on a grid.
- Random: Tree coordinates were generated randomly within the plot.
- Clustered: Trees were grouped into 3 to 5 randomly located clusters, each at least 10 meters from the plot edge.

The different combinations of forest characteristics and spatial patterns resulted in 1,980 plot configurations. Furthermore, we created 20 repetitions for each combination (1,980 plot configurations x 20 repetitions = 39,600 plots). Diameter distributions were simulated using Gamma and Normal distributions.

To evaluate different plot designs, we placed plots from the Spanish and Finnish National Forest Inventories (NFIs) at the centre of their 1-ha simulated forests. These two NFIs were chosen because they represent the greatest differences in plot configuration, specifically in their diameter at breast height (dbh) thresholds and plot sizes. Then, we ensured we covered a wide range of intermediate designs.

Below, we show a brief description of the Spanish and Finnish NFI:

- Spanish NFI: Uses a nested plot design with four concentric subplots. As the radius increases (5m, 10m, 15m, 25m), so does the minimum dbh for trees measured (7.5cm, 12.5cm, 22.5cm, 42.5cm, respectively). Basic tree data, like dbh and species, are collected (Alberdi et al. 2016).
- Finnish NFI (13th iteration): Also uses a nested plot design, but with two concentric subplots. Trees within a 4m radius must have a minimum dbh of 4.5 cm, while those within a 9m radius must have a minimum dbh of 9.5 cm (Korhonen et al. 2024).

5.1.1.3 Structural indices

We focused on distance-independent structural indices, considering NFIs data design limitations. We categorised the chosen structural indices into three main groups:

1. Species composition-related indices: These measure the diversity of tree species.
2. Tree size diversity indices: These quantify the variation in tree sizes (Pommerening 2002).
3. Combined indices: These integrate both species composition and tree size diversity (Staudhammer and LeMay 2001).

For species composition, we included:

- Species richness (S) (Chao and Chiu 2016).
- Shannon's Diversity Index (SDI_sp) (Shannon 1948).
- Simpson's Diversity Index (D) (Simpson 1949).
- Margalef's Richness Index (MRI) (Margalef 1958).

To quantify forest tree size diversity, we used:

- Sum of square roots of dbh differences (SQRI) (Barbeito et al. 2009).
- Shannon's Diversity Index applied to 10 cm diameter classes (SDI_dbh) (Staudhammer and LeMay 2001).
- Gini Index (Gini).

- Coefficient of variation of the mean plot dbh (CV) (Zeller and Pretzsch 2019; Astigarraga et al. 2020).
- Interquartile range (IQR).
- Second L-moment or L-scale (L2): An analogue to standard deviation (Table 17).

Finally, we employed two formulations of Shannon's Diversity Index to describe species composition and structure together (Staudhammer and LeMay 2001):

- Shannon's Diversity Index post-hoc method (SDI_ph): Calculated as the arithmetic mean of SDI_sp and SDI_dbh.
- Shannon's Diversity Index combination method (SDI_cm): This applies Shannon's Diversity Index to combinations of species and 10 cm diameter classes.

All indices were calculated under three distinct plot configurations:

- Total: For the entire 1-hectare simulated plot.
- Nested: For the nested plot design specific to Spanish and Finnish NFI.
- Fixed: For a fixed-radius plot, using the largest radius defined for each NFI's nested design (Spain = 25 m radius and Finland = 8 m radius).

The nested plot design changes the probability of sampling trees based on their size and location. Because of this, trees within a nested design must be weighted to account for these shifts in sampling probability. Thus, we corrected all the indices with the exception of S and MRI.

Finally, we calculated the R^2, Bias and Root mean squared error (RMSE) to evaluate the performance of the indices considering values of Total as observed values and Nested as predicted. Furthermore, we calculated these three metrics with Total as observed and Fixed as predicted and Fixed as observed and Nested as predicted.

All the statistical analyses were performed in R 4.4.3. (R Core Team 2025).

5.1.2 Results

In the Spanish NFI, the highest R^2 was found when comparing Total vs Fixed plots, while the lowest occurred in Total vs Nested comparisons. Similarly, the comparison of Total against Fixed gave the values closest to zero for the error metrics (Bias and RMSE). IQR, S, and MRI showed the most variation across plot types, while SDI_sp and S were the most stable indicators of species composition. For forest complexity, L2, SDI_dbh, and SQRI performed best. SDI_cm and SDI_ph showed similar performance in describing complexity and diversity relationships.

In the Finnish NFI, the comparison between Fixed vs Nested yielded the best results (highest R^2 and lowest Bias and RMSE). S, D, and SDI_sp were most effective in characterising species composition, while L2 and SQRI were the best indices for describing forest complexity. SDI_ph outperformed SDI_cm in describing complexity with species diversity.

In Spain, R^2 values for the comparison between Total and Nested ranged from 0.42 to 0.88, while in Finland they ranged from 0.13 to 0.72 (Figure 6). When comparing Total with Fixed, R^2 values were

higher in Spain (0.83 to 0.98) than in Finland (0.17 to 0.75). For the comparison between Fixed and Nested, R^2 ranged from 0.47 to 0.94 in Spain and from 0.68 to 0.97 in Finland.

5.1.3 Discussion

The study found that Spanish Fixed and Nested designs better capture forest structure compared to the Finnish designs when benchmarked against Total, likely due to Spain's larger plot size (25 m vs. 9 m) and wider plot spacing. However, greater consistency between Nested and Fixed plots was observed in Finland, possibly due to its smaller dbh thresholds and smaller plot size (Alberdi et al. 2016; Korhonen et al. 2024). Given the significant differences in national forest structures and inventory designs, cross-country comparisons should be avoided. Instead, the performance of individual indices should be compared.

A key limitation is the use of simulated plots, which may not fully represent real forest conditions, particularly for boreal forests with typically higher stocks compared to more open forests and more dominant species in Mediterranean areas (Svenning and Skov 2005; Moreno-Fernández et al. 2020a). This could affect the accuracy of certain structural indices.

Some indices, such as MRI, are biased in nested designs due to underrepresentation of rare or small-diameter species (Lin et al. 2020; Moreno-Fernández et al. 2024), while SDI_sp and D perform better because they factor in abundance (basal area). However, results across studies regarding the effectiveness of SDI_sp are mixed (Franc 1998; Neumann and Starlinger 2001; Mulya et al. 2021; Kitikidou et al. 2024).

SDI_dbh was found suitable for nested designs but has limitations due to its dependence on classifying continuous data (Staudhammer and LeMay 2001; Valbuena et al. 2012). Despite this, it is effective in describing forest complexity and correlates with SQRI, which avoids the need for data binning (Staudhammer and LeMay 2001; Barbeito et al. 2009). Both SDI_dbh, Gini, and CV are size-insensitive, making them useful for comparing stands of different ages, unlike IQR and L2, which are size-sensitive.

L2 and other L-moments are robust to outliers and better suited for small samples, making them valuable for nested designs (Hosking 1990).

Finally, SDI_ph and SDI_cm showed similar effectiveness in jointly assessing structure and composition, combining the strengths of SDI_sp and SDI_dbh. Among them, SDI_ph may be more informative, as it allows independent evaluation of dbh and species diversity (Staudhammer and LeMay 2001) .

5.1.4 Conclusions

To accurately describe species composition in both Nested and Fixed plot designs, we recommend using either Shannon's Diversity Index (SDI_sp) or Simpson's Diversity Index (D). For quantifying forest complexity, the most effective and comparable indices across plot designs were the Sum of Square Roots of dbh Differences (SQRI), Shannon's Index applied to diameter classes (SDI_dbh), and the second L-moment (L2). Additionally, both post-hoc and combined approaches using Shannon's Diversity Index proved effective in jointly capturing forest complexity and species composition.

We suggest replicating this study under different simulation settings (e.g., varying tree density by species) and using alternative NFI plot designs. Extending the analysis to field data would further validate the findings.

5.2 Deadwood

Historically, deadwood was considered a threat to forest productivity and fire risk (Wijas et al. 2024). However, recent decades have highlighted its ecological importance, particularly in nutrient cycling and carbon storage, with deadwood storing ca 8% of forest carbon (Pan et al. 2011; Stockland et al. 2016; Martin et al. 2021). While tropical forests store the most deadwood carbon, temperate and subtropical forests hold the lowest (Anderson-Teixeira et al. 2015; Wijas et al. 2024). Deadwood is also a key biodiversity indicator, providing habitat for many species (Parisi et al. 2016; SoEF 2020; Parajuli and Markwith 2023). As a result, accurately quantifying its stocks and dynamics is essential for sustainable forest management.

Under this context, extensive surveys such as National Forest Inventories have incorporated deadwood measurements in the field protocols. Thus, this data has been compiled in different studies at the national level to evaluate the deadwood stocks (Woodall et al. 2008; Alberdi et al. 2020; Bujoczek et al. 2021). However, at the pan-European level, there is just one attempt to model deadwood using the ICP Level I Network (Augustynczik et al. 2024). The strength of this network is the number of countries covered and the harmonized protocols, but the main drawback is the lower number of plots in comparison with the NFIs.

Deadwood has been estimated in Pathfinder using EFDM- approach and disturbance outputs. However, considering the variability of deadwood and the data provided from NFIS, we are additionally working to identify the main variables explaining the amounts of deadwood in Europe, relying on National Forest Inventory data. We will continue working on a model to forecast deadwood.

5.2.1 Deadwood National Forest Inventory data and modelling

To achieve our goals, we leveraged the NFI data stored in the nFiesta repository within the framework of the current project, PathFinder HE. The data has been cross-validated, standardized and harmonized to common definitions for all the NFIs. Examples of variables included in the repository are living aboveground volume, basal area, deadwood stocks, living stock increment, or mean quadratic diameter.

In order to protect privacy rights, the exact plot coordinates were not provided but each plot is linked to a corresponding cell in the 1 × 1 km INSPIRE Grid (ETRS89-extended / LAEA Europe, EPSG: 3035), which spans the whole Europe. Thus, we randomly assigned coordinates to each plot within its respective grid cell.

Our target variable was the root square of the deadwood stock data (in m^3 ha^{-1}), which includes standing dead trees (snags) and lying deadwood with a diameter of 10 cm or more, while stumps were excluded. This size threshold was chosen as it aligns with the most commonly adopted standard in international frameworks (SoEF 2020; FRA 2023). The total dataset included 70,704 from nine countries (Austria, Belgium (Walloon region), the Czech Republic, Germany, Italy, Ireland, Norway, Spain and Sweden) (Table 17 and Figure 6).

Table 17: Median, mean and standard deviation of the deadwood stocks (m^3 ha^{-1}) and number of plots by country according to the National Forest Inventories.

Country	Median	Mean	Standard deviation	Number of plots
Austria	6.7	13.8	21.1	737
Belgium (Walloon region)	1.1	4.4	8.9	1,458
Czech Republic	5.0	10.3	18.2	6,657
Germany	1.9	6.7	19.2	6,501
Spain	0.7	3.3	14.3	18,506
Ireland	1.4	6.2	15.3	1,538
Italy	1.8	6.2	11.4	6,722
Norway	1.2	4.9	10.1	11,498
Sweden	0.7	2.7	6.9	17,087

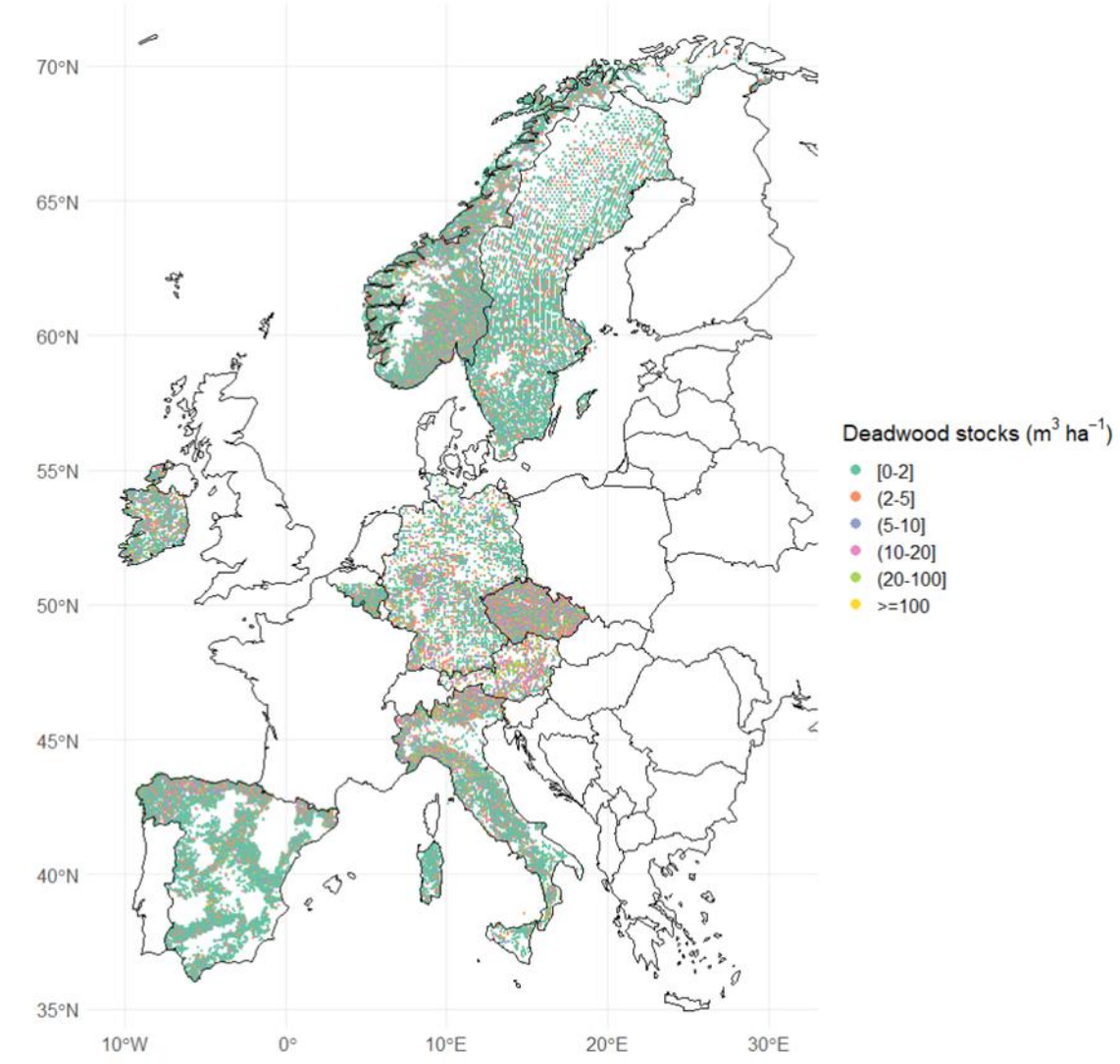


Figure 6: Deadwood stocks in the National Forest Inventory plots

As predictors, we considered four categories of variables: forest, climatic, physiographic, and human-related factors (Table 18).

- Forest variables (from the nFiesta repository) included quadratic mean diameter (in cm), living aboveground biomass (in Mg ha^{-1}) and tree mortality (calculated as the difference between gross and net increments (in m^3 ha^{-1} $year^{-1}$).
- Climatic data (from Chelsa 2.1) included mean annual temperature (°C) and total annual precipitation (mm) (Karger et al. 2017).
- Physiographic variables (from the Copernicus DEM) included altitude (m asl), slope (°), and aspect (cos-transformed) (European Commission 2012)
- Human-related variables included protection status (protected vs. unprotected, from the European Environmental Agency) (EEA 2025).

Table 18: Mean and standard error of the deadwood and predictors for the National Forest Inventories plots.

Variable	Mean	Standard deviation
Deadwood (m^3 ha^{-1})	4.8	13.3
Quadratic mean diameter (cm)	19.9	11.7
Living aboveground biomass (Mg ha^{-1})	117.8	106.7
Mortality (m^3 ha^{-1} $year^{-1}$)	0.2	1.7
Proportion of conifers (%)	56.9	43.2
Mean annual temperature (°C)	7.9	4.2
Cumulative annual precipitation (mm)	1006.4	473.4
Altitude (m asl)	507.3	392.6
Slope (°)	8.2	8.9
Cosine of the aspect	0.0	0.7
Percentage of plots located in protected areas (%)	27	-

We fitted a Random Forest algorithm to model deadwood stocks as a function of these predictors using the "ranger" package (Wright and Ziegler 2017), implemented in R (R Core Team 2025). Furthermore, we calculated the variance importance to understand the degree of association between the predictors and the response, as well as the out-of-the-bag prediction error (OBB MAE) and the out-of-the-bag R^2 (OBB R^2) as a measure of the goodness of the fit. Finally, we carried out a five-fold cross-validation to assess the performance of the model on external data and calculated the Root Mean Square Error (RMSE), R^2 and the mean absolute error (MAE).

5.2.2 Results

The OBB R^2 was 0.31 and the OBB MAE was 1.79. Regarding the cross-validation statistics, RMSE was 1.34, R^2 0.31 and MAE 0.92. According to the variance importance, the predictors were ranked as follows: mortality, living aboveground biomass, mean annual temperature, altitude, quadratic mean diameter, cumulative annual rainfall, slope, proportion of conifers, protection status and aspect. All

the variables exhibited a positive relationship with deadwood stocks except the proportion of conifers and the mean annual temperature. Concerning the protection status, protected forests tend to display larger stocks of deadwood than unprotected forests.

5.2.3 Discussion

In agreement with our results, several authors have reported that deadwood accumulation is influenced by tree mortality and dieback processes (Alberdi et al. 2020; Bouget and Cours 2025), as well as the shedding of branches from living trees (Zuleta et al. 2023; Wijas et al. 2024). The aboveground biomass of living trees was identified as an important predictor of deadwood. A positive relationship between standing biomass and deadwood has been previously documented (Moreno-Fernández et al. 2020b; Augustynczik et al. 2024), but this association can vary depending on the intensity and type of forest management (Bujoczek et al. 2021). The positive relationship between the quadratic mean diameter and deadwood can be explained because mature forests store more deadwood than younger stands (Bujoczek et al. 2021; Augustynczik et al. 2024). Additionally, mature forests are often associated with protected areas, which, as supported by our results and previous studies, tend to contain higher amounts of deadwood (Paletto et al. 2014; Bujoczek et al. 2021; Oettel et al. 2023). Finally, our results show that conifer-dominated forests generally contain less deadwood. This finding is consistent with Augustynczik et al. (2024), who reported lower deadwood levels in conifer forests compared to mixed or broadleaved stands.

We also found that **mean annual temperature is negatively related to deadwood stocks**, likely due to faster decomposition driven by increased microbial and detritivore activity in warmer climates (Přívětivý et al. 2016; Harmon et al. 2020; Zanne et al. 2022; Wijas et al. 2024). Although higher temperatures can also increase tree mortality (Rebollo et al. 2024), leading to more deadwood input, the overall effect is a net reduction due to rapid decay. In contrast**, rainfall shows a positive relationship with deadwood**, as it supports greater forest productivity and biomass (Gustafson et al. 2017), enhances pathogen activity leading to dieback (Chen et al. 2021), and increases susceptibility to windthrow in wetter soils (Mitchell 2013).

The study confirms that **deadwood amounts are positively associated with slope and altitude.** We hypothesized that steeper and higher-elevation forests have limited access and reduce forest management activities (Crecente-Campo et al. 2016; Bujoczek et al. 2021; Oettel et al. 2023). This finding aligns with previous research showing tha**t less-managed, mountainous forests tend to accumulate more deadwood** (Alberdi et al. 2020; Bujoczek et al. 2021; Augustynczik et al. 2024)**.**

5.2.4 Further steps

We plan to use this model to create wall-to-wall maps of deadwood under the current conditions as well as upcoming scenarios.

5.3 Composite indices

We have proposed a composite index to describe forest biodiversity based on the EFDM output. The indezx aggregates three subindices, which are defined as follows:

(1) Amount of deadwood by pixel (*DW*)
(2) Number of groups of species as a proxy of species or functional diversity (*Ngroups*)
(3) Shannon Diversity Index applied to the volume classes (*Complexity*). The idea of this index is to account for forest complexity. The index is defined as follows:

$$Complexity = -\sum_{v=1}^{V} a_v \, ln \, a_v$$

where *v* are the number of volume classes (0, 1, 10, 25, 48, 80, 122, 204, 313, 517 and 3000, $m^3\ ha^{-1}$) and *a* is the area occupied by each class.

To eliminate the differences in units, the three indices were normalized as follows:

$$Index\ normalized_i = \frac{(observed\ value_i - \min(Index))}{(\max(Index) - \min(Index))}$$

where *i* is each pixel, *observed value* is the value of a given index in the *i* pixel, min(*Index*) and max(*Index*) are the minimum and maximum value of the index

After the normalization, the three indices were summed up to obtain the aggregated index:

$$Biodiversity\ Index = Normalized.DW + Normalized.Ngroups + Normalized.Complexity$$

References

Achille, M., Girardello, M., Strona, G., Beck, B., Forzieri, G., Caudullo, G., Manca, F., Cescatti, A. "EU-Trees4F, a dataset on the future distribution of European tree species." Scientific data 9, no. 1 (2022): 37.

Ainsworth, E. A., & Rogers, A. (2007). The response of photosynthesis and stomatal conductance to rising [CO2]: mechanisms and environmental interactions. Plant, cell & environment, 30(3), 258-270.

Alberdi I, Hernández L, Condés S, Cañellas I (2016) Spain. In: Vidal C, I Alberdi, Hernández L, Redmond J (eds) National Forest Inventories. Assessment of wood availability and use. Springer, pp 749–769

Alberdi I, Moreno-Fernández D, Cañellas I, et al (2020) Deadwood stocks in south-western European forests: Ecological patterns and large scale assessments. Science of The Total Environment 747:141237. https://doi.org/10.1016/j.scitotenv.2020.141237

Anderson-Teixeira KJ, Davies SJ, Bennett AC, et al (2015) CTFS-ForestGEO: a worldwide network monitoring forests in an era of global change. Glob Chang Biol 21:528–549. https://doi.org/10.1111/GCB.12712

Astigarraga J, Andivia E, Zavala MA, et al (2020) Evidence of non-stationary relationships between climate and forest responses: Increased sensitivity to climate change in Iberian forests. Glob Chang Biol 26:5063–5076. https://doi.org/10.1111/gcb.15198

Augustynczik ALD, Gusti M, di Fulvio F, et al (2024) Modelling the effects of climate and management on the distribution of deadwood in European forests. J Environ Manage 354:120382. https://doi.org/https://doi.org/10.1016/j.jenvman.2024.120382

Baier, P., Pennerstorfer, J., & Schopf, A. (2007). PHENIPS—A comprehensive phenology model of Ips typographus (L.) (Col., Scolytinae) as a tool for hazard rating of bark beetle infestation. *Forest Ecology and Management*, *249*(3), 171–186. https://doi.org/https://doi.org/10.1016/j.foreco.2007.05.020

Barbeito I, Cañellas I, Montes F (2009) Evaluating the behaviour of vertical structure indices in Scots pine forests. Ann For Sci 66:710–710. https://doi.org/10.1051/forest/2009056

Bouget C, Cours J (2025) Effects of forest dieback on deadwood patterns: Large scale trends from a cross-analysis of European databases. J Environ Manage 375:124315. https://doi.org/10.1016/j.jenvman.2025.124315

Bujoczek L, Bujoczek M, Zięba S (2021) How much, why and where? Deadwood in forest ecosystems: The case of Poland. Ecol Indic 121:107027. https://doi.org/10.1016/j.ecolind.2020.107027

Camia, A., Libertá, G., & San-Miguel-Ayanz, J. (2017). *Modeling the impacts of climate change on forest fire danger in Europe Sectorial results of the PESETA II Project*. https://doi.org/10.2760/768481

Chao A, Chiu C (2016) Species Richness: Estimation and Comparison. In: Wiley StatsRef: Statistics Reference Online. Wiley, pp 1–26

Chen Q, Hu H, Yan Z, et al (2021) Precipitation increases the abundance of fungal plant pathogens in *Eucalyptus* phyllosphere. Environ Microbiol 23:7688–7700. https://doi.org/10.1111/1462-2920.15728

Comins, H. N., & McMurtrie, R. E. (1993). Long-Term Response of Nutrient-Limited Forests to CO" 2 Enrichment; Equilibrium Behavior of Plant-Soil Models. Ecological Applications, 3(4), 666-681.

Crecente-Campo F, Pasalodos-Tato M, Alberdi I, et al (2016) Assessing and modelling the status and dynamics of deadwood through national forest inventory data in Spain. For Ecol Manage 360:297–310. https://doi.org/10.1016/j.foreco.2015.10.029

de Rigo, D., Caudullo, G., Houston Durrant, T., San-Miguel-Ayanz, J., 2016. The European Atlas of Forest Tree Species: modelling, data and information on forest tree species. In: San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A. (Eds.), European Atlas of Forest Tree Species. Publications Office of the European Union, Luxembourg, pp. e01aa69+. https://w3id.org/mtv/FISE-Comm/v01/e01aa69

Copernicus, 2025.https://land.copernicus.eu/en/technical-library/product-user-manual-tree-cover-and-forests-2018-2021/@@download/file [accessed 16th of June 2025]

del Río M, Pretzsch H, Alberdi I, et al (2016) Characterization of the structure, dynamics, and productivity of mixed-species stands: review and perspectives. Eur J For Res 135:23–49. https://doi.org/10.1007/s10342-015-0927-6

EEA (2025) Nationally designated areas. In: European Environment Agency. https://www.eea.europa.eu/en/datahub/datahubitem-view/f60cec02-6494-4d08-b12d-17a37012cb28?activeAccordion=. Accessed 10 May 2025

European Comission (2012) EU-DEM Version 1, available from http:// http://epp.eurostat.ec.europa.eu/portal/page/portal/gisco_Geographical_information_maps/introduction and http://www.eea.europa.eu/data-and-maps/data/eu-dem

Eurostat, 2025. https://ec.europa.eu/eurostat/web/gisco/geodata/administrative-units/countries [Loaded: 27.02.2025]

FAO, 2025. Forests, forestry and logging - Statistics Explained - Eurostat https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Forests,_forestry_and_logging#:~:text=In%20absolute%20terms%2C%20the%20countries,)%2C%20as%20estimated%20in%202022.&text=The%20stocks%20of%20timber%20in,(over%20bark)%20in%202022 [accessed 16th of June 2025]

Forrester, D. I., Tachauer, I. H. H., Annighoefer, P., Barbeito, I., Pretzsch, H., Ruiz-Peinado, R., ... & Sileshi, G. W. (2017). Generalized biomass and leaf area allometric equations for European tree species incorporating stand structure, tree age and climate. *Forest ecology and management*, *396*, 160-175.

FRA (2023) Global Forest Resources Assessment - FRA 2025 - Terms and Definitions (fao.org). https://www.fao.org/3/cc4691en/cc4691en.pdf

Franc A (1998) Some mathematical remarks on forest biodiversity. In: Bachmann P, Köhl M, Päivinen R (eds) Assessment of Biodiversity for Improved Forest Planning. Kluwer Academic Publishers, Dordrecht, pp 159–169

Francis, E. J., Jung, C. G., Hicke, J. A., & Hurteau, M. D. (2025). Modeling the probability of bark beetle-caused tree mortality as a function of watershed-scale host species presence and basal area. *Forest Ecology and Management*, *580*, 122549. https://doi.org/https://doi.org/10.1016/j.foreco.2025.122549

Gardiner, B., Byrne, K., Hale, S., Kamimura, K., Mitchell, S. J., Peltola, H., & Ruel, J.-C. (2008). A review of mechanistic modelling of wind damage risk to forests. *Forestry: An International Journal of Forest Research*, *81*(3), 447–463. https://doi.org/10.1093/forestry/cpn022

Gustafson EJ, Miranda BR, De Bruijn AMG, et al (2017) Do rising temperatures always increase forest productivity? Interacting effects of temperature, precipitation, cloudiness and soil texture on tree species growth and competition. Environmental Modelling & Software 97:171–183. https://doi.org/10.1016/J.ENVSOFT.2017.08.001

Hale, S. E., Gardiner, B., Peace, A., Nicoll, B., Taylor, P., & Pizzirani, S. (2015). Comparison and validation of three versions of a forest wind risk model. *Environmental Modelling & Software*, *68*, 27–41. https://doi.org/https://doi.org/10.1016/j.envsoft.2015.01.016

Harmon ME, Fasth BG, Yatskov M, et al (2020) Release of coarse woody detritus-related carbon: a synthesis across forest biomes. Carbon Balance Manag 15:1. https://doi.org/10.1186/s13021-019-0136-6

Hosking JRM (1990) L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. J R Stat Soc Series B Stat Methodol 52:105–124. https://doi.org/10.1111/j.2517-6161.1990.tb01775.x

Hui G, Zhang G, Zhao Z, Yang A (2019) Methods of Forest Structure Research: a Review. Current Forestry Reports 5:142–154. https://doi.org/10.1007/s40725-019-00090-7

Hyvönen, R., Ågren, G. I., Linder, S., Persson, T., Cotrufo, M. F., Ekblad, A., ... & Wallin, G. (2007). The likely impact of elevated [CO2], nitrogen deposition, increased temperature and management on carbon sequestration in temperate and boreal forest ecosystems: a literature review. New Phytologist. 173 (3): 463-480., 173(3), 463-480.

IPCC 2006, IPCC Guidelines for National Greenhouse Gas Inventories. (2006). Prepared by the National Greenhouse Gas Inventories Programme Eggleston H.S., Buendia L., Miwa K., Ngara T. and Tanabe K. (eds). Published: IGES, Japan.

IPCC, 2025. IPCC WGI Interactive Atlas [accessed 27th of June 2025]

Kalliokoski, T., Mäkelä, A., Fronzek, S., Minunno, F., & Peltoniemi, M. (2018). Decomposing sources of uncertainty in climate change projections of boreal forest primary production. Agricultural and Forest Meteorology, 262, 192-205.

Karger DN, Conrad O, Böhner J, et al (2017) Climatologies at high resolution for the earth's land surface areas. Sci Data 4:170122. https://doi.org/10.1038/sdata.2017.122

Khabarov, N., Krasovskii, A., Obersteiner, M., Swart, R., Dosio, A., San-Miguel-Ayanz, J., Durrant, T., Camia, A., & Migliavacca, M. (2016). Forest fires and adaptation options in Europe. *Regional Environmental Change*, *16*(1), 21–30. https://doi.org/10.1007/s10113-014-0621-0

Kitikidou K, Milios E, Stampoulidis A, et al (2024) Using Biodiversity Indices Effectively: Considerations for Forest Management. Ecologies 5:42–51. https://doi.org/10.3390/ecologies5010003

Kolby Smith, W., Reed, S. C., Cleveland, C. C., Ballantyne, A. P., Anderegg, W. R., Wieder, W. R., ... & Running, S. W. (2016). Large divergence of satellite and Earth system model estimates of global terrestrial CO2 fertilization. Nature climate change, 6(3), 306-310.

Korhonen KT, Räty M, Haakana H, et al (2024) Forests of Finland 2019–2023 and their development 1921–2023. Silva Fennica 58:1–47. https://doi.org/10.14214/sf.24045Kuronen M, Räty M (2023). efdm: Simulate Forest Resources with the European Forestry Dynamics Model. R package version 0.2.1. https://CRAN.R-project.org/package=efdm.

Kurz, W. A., Dymond, C. C., White, T. M., Stinson, G., Shaw, C. H., Rampley, G. J., ... & Apps, M. J. (2009). CBM-CFS3: a model of carbon-dynamics in forestry and land-use change implementing IPCC standards. Ecological modelling, 220(4), 480-504.

Lawal, S., Costanza, J., Koch, F., & Scheller, R. (2024). Modeling the impacts of hot drought on forests in Texas. *Frontiers in Forests and Global Change*, *7*. https://doi.org/10.3389/ffgc.2024.1280254

Lin HT, Lam TY, von Gadow K, Kershaw JA (2020) Effects of nested plot designs on assessing stand attributes, species diversity, and spatial forest structures. For Ecol Manage 457:117658. https://doi.org/10.1016/j.foreco.2019.117658

Litton, C. M., Raich, J. W., & Ryan, M. G. (2007). Carbon allocation in forest ecosystems. Global Change Biology, 13(10), 2089-2109.

Locatelli, T., Hale, S., Nicoll, B., & Gardiner, B. (2022). The ForestGALES wind risk model and the fgr R package. *Edinburgh (UK): UK Forestry Commision*.

Mäkela, A., Härkönen, S. (2019), "FORMIT-M simulator", Mendeley Data, V1, doi: 10.17632/344n6ts3tg.1 [accessed: 5th of June 2025]

Mäkelä, A., Pulkkinen, M., Kolari, P., Lagergren, F., Berbigier, P., Lindroth, A., ... & Hari, P. (2008). Developing an empirical model of stand GPP with the LUE approach: analysis of eddy covariance data at five contrasting conifer sites in Europe. Global change biology, 14(1), 92-108.

Margalef R (1958) Information Theory in Ecology. General Systems 3:36–71

Marie, G., Jeong, J., Jactel, H., Petter, G., Cailleret, M., McGrath, M., Bastrikov, V., Ghattas, J., Guenet, B., Lansø, A., Naudts, K., Valade, A., Yue, C., & Luyssaert, S. (2023). *Simulating Bark Beetle Outbreak Dynamics and their Influence on Carbon Balance Estimates with ORCHIDEE r7791*. https://doi.org/10.5194/egusphere-2023-1216

Martin AR, Domke GM, Doraisami M, Thomas SC (2021) Carbon fractions in the world's dead wood. Nat Commun 12:889. https://doi.org/10.1038/s41467-021-21149-9

Minunno, F., Peltoniemi, M., Launiainen, S., Aurela, M., Lindroth, A., Lohila, A., ... & Mäkelä, A. (2016). Calibration and validation of a semi-empirical flux ecosystem model for coniferous forests in the Boreal region. Ecological Modelling, 341, 37-52.

Mitchell SJ (2013) Wind as a natural disturbance agent in forests: a synthesis. Forestry 86:147–157. https://doi.org/10.1093/forestry/cps058

Moreno-Fernández D, Cañellas I, Alberdi I, Montes F (2021) Improved stand structure characterization from nested plot designs in the Spanish National Forest Inventory. Forestry: An International Journal Of Forest Research 94:244–257. https://doi.org/10.1093/forestry/cpaa031

Moreno-Fernández D, Cañellas I, Hernández L, et al (2024) Nested plot designs used in forest inventory do not accurately capture tree species richness in Southwestern European forests. Ann For Sci 81:20. https://doi.org/10.1186/s13595-024-01237-8

Moreno-Fernández D, Cañellas I, Rubio-Cuadrado Á, Alberdi I (2020a) National scale variability in forest stand variables among regions of provenances in Spain. Ann For Sci 77:44. https://doi.org/10.1007/s13595-020-00943-3

Moreno-Fernández D, Hernández L, Cañellas I, et al (2020b) Analyzing the dynamics of the deadwood carbon pool in Spain through the European Level I Monitoring Programme. For Ecol Manage 463:118020. https://doi.org/10.1016/j.foreco.2020.118020

Mulya H, Santosa Y, Hilwan I (2021) Comparison of four species diversity indices in mangrove community. Biodiversitas 22:. https://doi.org/10.13057/biodiv/d220906

Nardi, D., Jactel, H., Pagot, E., Samalens, J.-C., & Marini, L. (2023). Drought and stand susceptibility to attacks by the European spruce bark beetle: A remote sensing approach. *Agricultural and Forest Entomology*, *25*(1), 119–129. https://doi.org/https://doi.org/10.1111/afe.12536

Neumann M, Starlinger F (2001) The significance of different indices for stand structure and diversity in forests. For Ecol Manage 145:91–106. https://doi.org/10.1016/S0378-1127(00)00577-6

Nicoll, B. C., Gardiner, B. A., Rayner, B., & Peace, A. J. (2006). Anchorage of coniferous trees in relation to species, soil type, and rooting depth. *Canadian Journal of Forest Research*, *36*(7), 1871–1883. https://doi.org/10.1139/x06-072

Norby, R. J., & Zak, D. R. (2011). Ecological lessons from free-air CO2 enrichment (FACE) experiments. Annual review of ecology, evolution, and systematics, 42(1), 181-203.

Oettel J, Zolles A, Gschwantner T, et al (2023) Dynamics of standing deadwood in Austrian forests under varying forest management and climatic conditions. Journal of Applied Ecology 60:696–713. https://doi.org/10.1111/1365-2664.14359

Paletto A, De Meo I, Cantiani P, Ferretti F (2014) Effects of forest management on the amount of deadwood in Mediterranean oak ecosystems. Ann For Sci 71:791–800. https://doi.org/10.1007/s13595-014-0377-1

Pan Y, Birdsey RA, Fang J, et al (2011) A large and persistent carbon sink in the world's forests. Science 333:988–93. https://doi.org/10.1126/science.1201609

Parajuli R, Markwith SH (2023) Quantity is foremost but quality matters: A global meta-analysis of correlations of dead wood volume and biodiversity in forest ecosystems. Biol Conserv 283:110100. https://doi.org/10.1016/j.biocon.2023.110100

Parisi F, Lombardi F, Sciarretta A, et al (2016) Spatial patterns of saproxylic beetles in a relic silver fir forest (Central Italy), relationships with forest structure and biodiversity indicators. For Ecol Manage 381:217–234. https://doi.org/10.1016/j.foreco.2016.09.041

Peltoniemi, M., Pulkkinen, M., Aurela, M., Pumpanen, J., Kolari, P., & Mäkelä, A. (2015). A semi-empirical model of boreal-forest gross primary production, evapotranspiration, and soil water-calibration and sensitivity analysis.

Pommerening A (2002) Approaches to quantifying forest structures. Forestry 75:305–324. https://doi.org/10.1093/forestry/75.3.305

Přívětivý T, Janík D, Unar P, et al (2016) How do environmental conditions affect the deadwood decomposition of European beech (*Fagus sylvatica* L.)? For Ecol Manage 381:177–187. https://doi.org/10.1016/J.FORECO.2016.09.033

Quine, C.P. and White, I.M.S. (1993). Revised windiness scores for the Windthrow Hazard Classification. Forestry Commission Research Information Note 230, FC, Edinburgh.

R Core Team (2025) R: A language and environment for statistical computing. Version 4.4.3.

Rebollo P, Moreno-Fernández D, Cruz-Alonso V, et al (2024) Recent increase in tree damage and mortality and their spatial dependence on drought intensity in Mediterranean forests. Landsc Ecol 39:38. https://doi.org/10.1007/s10980-024-01837-9

Reich, P. B., & Hobbie, S. E. (2013). Decade-long soil nitrogen constraint on the CO2 fertilization of plant biomass. Nature Climate Change, 3(3), 278-282.

Russo, S., Dosio, A., Graversen, R. G., Sillmann, J., Carrao, H., Dunbar, M. B., Singleton, A., Montagna, P., Barbola, P., & Vogt, J. V. (2014). Magnitude of extreme heat waves in present climate and their projection in a warming world. *Journal of Geophysical Research: Atmospheres*, *119*(22), 12, 500–512, 512. https://doi.org/https://doi.org/10.1002/2014JD022098

Senf, C., Buras, A., Zang, C. S., Rammig, A., & Seidl, R. (2020). Excess forest mortality is consistently linked to drought across Europe. *Nature Communications*, *11*(1), 6200. https://doi.org/10.1038/s41467-020-19924-1

Shannon CE (1948) A Mathematical Theory of Communication. Bell System Technical Journal 27:379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Simpson EH (1949) Measurement of Diversity. Nature 163:688–688. https://doi.org/10.1038/163688a0

Snowdon, P. (1991). A ratio estimator for bias correction in logarithmic regressions. Canadian Journal of Forest Research, 21(5), 720-724.

SoEF (2020) State of Europe's Forests

Staudhammer CL, LeMay VM (2001) Introduction and evaluation of possible indices of stand structural diversity. Canadian Journal of Forest Research 31:1105–1115. https://doi.org/10.1139/cjfr-31-7-1105

Stockland JN, Woodall C, Fridman J, Ståhl G (2016) Burial of downed deadwood is strongly affected by log attributes, forest ground vegetation, edaphic conditions, and climate zones. Canadian Journal of Forest Research2 46:1451–1457. https://doi.org/10.1139/cjfr-2015-0461

Suvanto, S., Esquivel-Muelbert, A., Schelhaas, M. J., Astigarraga, J., Astrup, R., Cienciala, E., ... & Pugh, T. A. (2025). Understanding Europe's forest harvesting regimes. Earth's Future, 13(2), e2024EF005225.

Svenning J, Skov F (2005) The relative roles of environment and history as controls of tree species composition and richness in Europe. J Biogeogr 32:1019–1033. https://doi.org/10.1111/j.1365-2699.2005.01219.x

Temperli, C., Bugmann, H., & Elkin, C. (2013). Cross-scale interactions among bark beetles, climate change, and wind disturbances: a landscape modeling approach. *Ecological Monographs*, *83*(3), 383–402. https://doi.org/https://doi.org/10.1890/12-1503.1

Tomppo E, Gschwantner T, Lawrence M, McRoberts R (2010) National Forest Inventories. Pathways for common reporting. Springer

Tuomi, M., Laiho, R., Repo, A., & Liski. J. 2011. Wood decomposition model for boreal forests. Ecological Modelling 222 (3): 709-718.

Tuomi, M., Thum, T., Järvinen, H., Fronzek, S., Berg, B., Harmon, M., Trofymow, J.A., Sevanto, S. & Liski, J. 2009. Leaf litter decomposition - Estimates of global variability based on Yasso07 model. Ecological Modelling 220 (23):3362-3371.

Valbuena R, Packalén P, Martı´n-Fernández S, Maltamo M (2012) Diversity and equitability ordering profiles applied to study forest structure. For Ecol Manage 276:185–195. https://doi.org/10.1016/j.foreco.2012.03.036

Valentine, H. T., & Mäkelä, A. (2012). Modeling forest stand dynamics from optimal balances of carbon and nitrogen. New Phytologist, 194(4), 961-971.

Vautard, R., Kadygrov, N., Iles, C., Boberg, F., Buonomo, E., Bülow, K., Coppola, E., Corre, L., van Meijgaard, E., Nogherotto, R., Sandstad, M., Schwingshackl, C., Somot, S., Aalbers, E., Christensen, O. B., Ciarlo, J. M., Demory, M.-E., Giorgi, F., Jacob, D., … Wulfmeyer, V. (2021). Evaluation of the Large EURO-CORDEX Regional Climate Model Ensemble. *Journal of Geophysical Research: Atmospheres*, *126*(17), e2019JD032344. https://doi.org/https://doi.org/10.1029/2019JD032344

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate*, *23*(7), 1696–1718. https://doi.org/https://doi.org/10.1175/2009JCLI2909.1

Wang, X., Wotton, B. M., Cantin, A. S., Parisien, M.-A., Anderson, K., Moore, B., & Flannigan, M. D. (2017). cffdrs: an R package for the Canadian Forest Fire Danger Rating System. *Ecological Processes*, *6*(1), 5. https://doi.org/10.1186/s13717-017-0070-z

Wermelinger, B., & Seifert, M. (1998). Analysis of the temperature dependent development of the spruce bark beetle Ips typographus (L) (Col., Scolytidae). *Journal of Applied Entomology*, *122*(1–5), 185–191. https://doi.org/https://doi.org/10.1111/j.1439-0418.1998.tb01482.x

Wijas BJ, Allison SD, Austin AT, et al (2024) The Role of Deadwood in the Carbon Cycle: Implications for Models, Forest Management, and Future Climates. Annu Rev Ecol Evol Syst 55:133–155. https://doi.org/10.1146/annurev-ecolsys-110421-102327

Woodall CW, Heath LS, Smith JE (2008) National inventories of down and dead woody material forest carbon stocks in the United States: Challenges and opportunities. For Ecol Manage 256:221–228. https://doi.org/10.1016/j.foreco.2008.04.003

Wright MN, Ziegler A (2017) ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. J Stat Softw 77:1–17. https://doi.org/10.18637/jss.v077.i01

Yasso, 2025. https://github.com/YASSOmodel/Ryassofortran [accessed: 5th of June 2025]

Zanne AE, Flores-Moreno H, Powell JR, et al (2022) Termite sensitivity to temperature affects global wood decay rates. Science (1979) 377:1440–1444. https://doi.org/10.1126/science.abo3856

Zeller L, Pretzsch H (2019) Effect of forest structure on stand productivity in Central European forests depends on developmental stage and tree species diversity. For Ecol Manage 434:193–204. https://doi.org/10.1016/j.foreco.2018.12.024

Zuleta D, Arellano G, McMahon SM, et al (2023) Damage to living trees contributes to almost half of the biomass losses in tropical forests. Glob Chang Biol 29:3409–3420. https://doi.org/10.1111/gcb.16687

Yasso, 2025. https://github.com/YASSOmodel/Ryassofortran [accessed: 5th of June 2025]

ANNEX A.

List of NFI data used in parameterising the EFDM:

Germany_plot_data_2025-Feb-17-08-57-09.csv

Germany_plots_2025-Feb-17-08-57-09.csv

plot_data_Austria_2024-Dec-02-15-51-04.csv

plot_data_Belgium_2024-Aug-07-14-59-05.csv

plot_data_Czechia_2025-Jan-07-09-17-16.csv

plot_data_France_2024-Sep-10-00-00-00.csv

plot_data_Ireland_2024-Nov-21-14-23-07.csv

plot_data_Italy_2024-Nov-27-08-50-36.csv

plot_data_Norway_2025-Jan-14-15-17-11.csv

plot_data_Poland_2024-Jan-16-07-07-00.csv

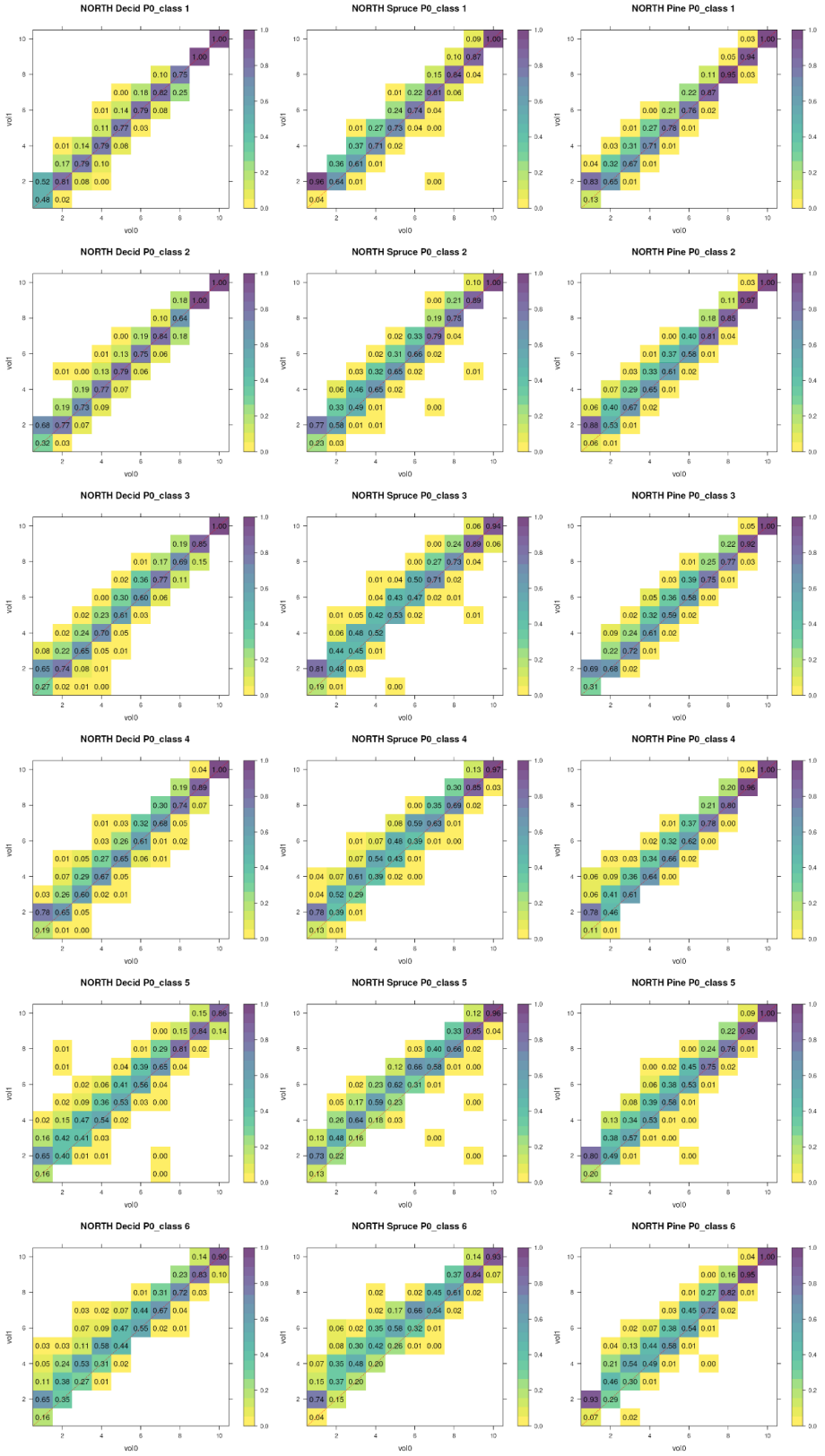
plot_data_Slovenia_2024-Jul-31-13-24-08.csv"

plot_data_Spain_2024-Nov-06-18-00-00.csv

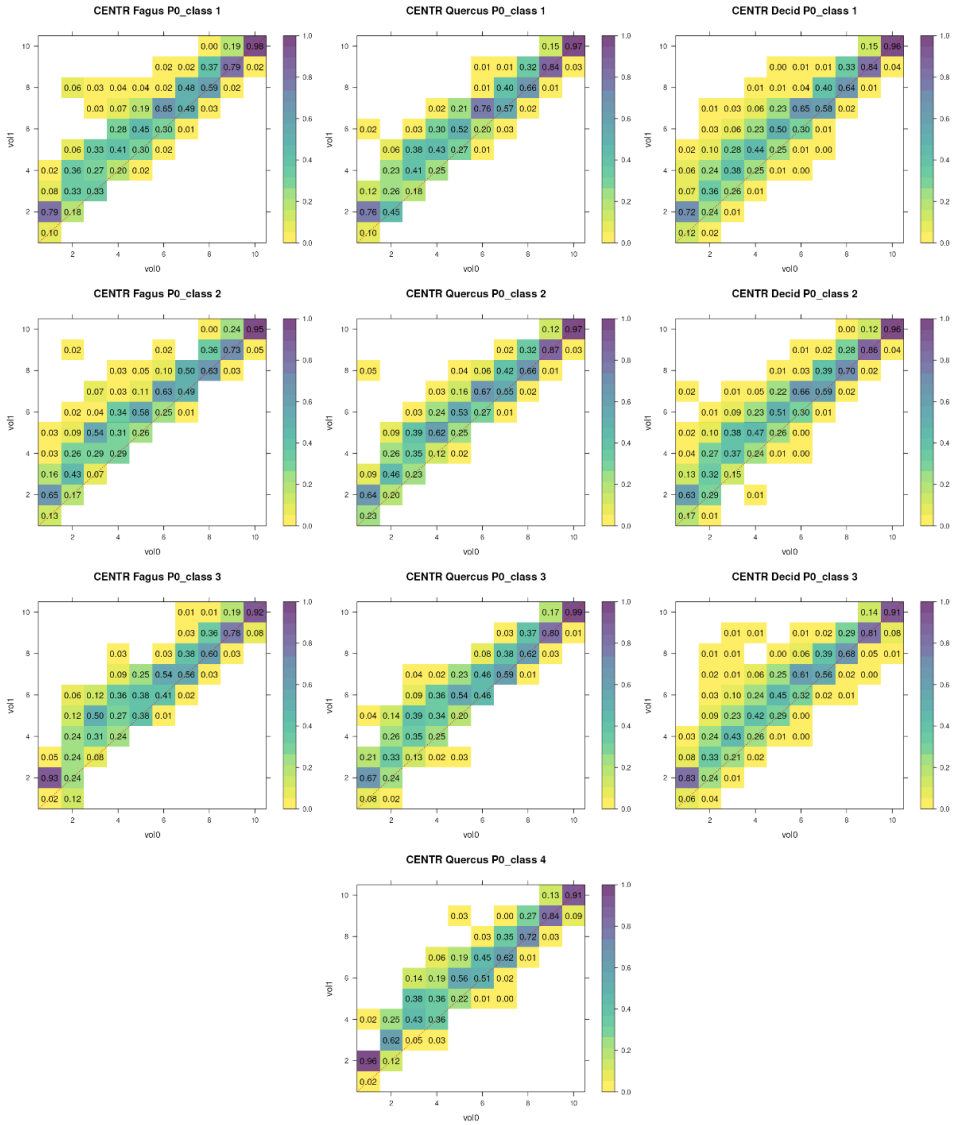
plot_data_Sweden_2025-Jan-21-15-19-13.csv

Transition matrices for noman sp_groups (n=10) for the three regions.

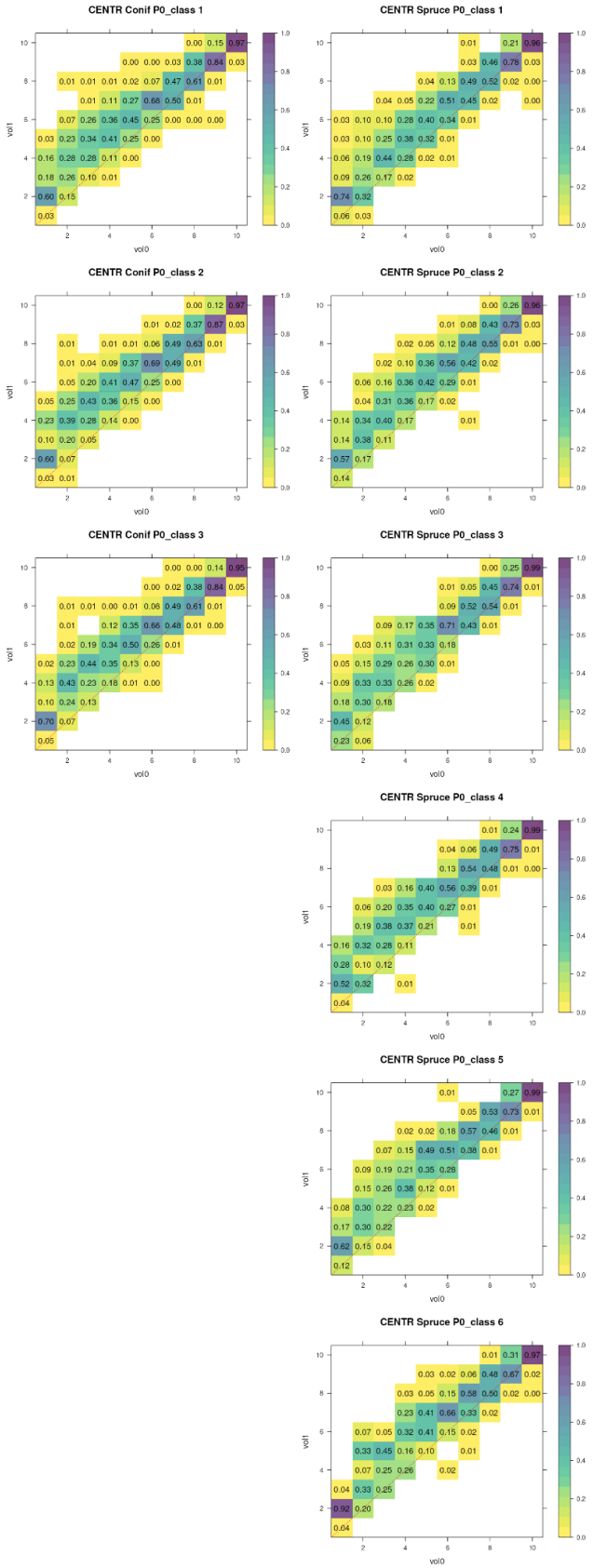
Transition matrices for the NORTH region for other deciduous, spruce and pine sp-groups. Transition matrices for each sp-group are in columns.



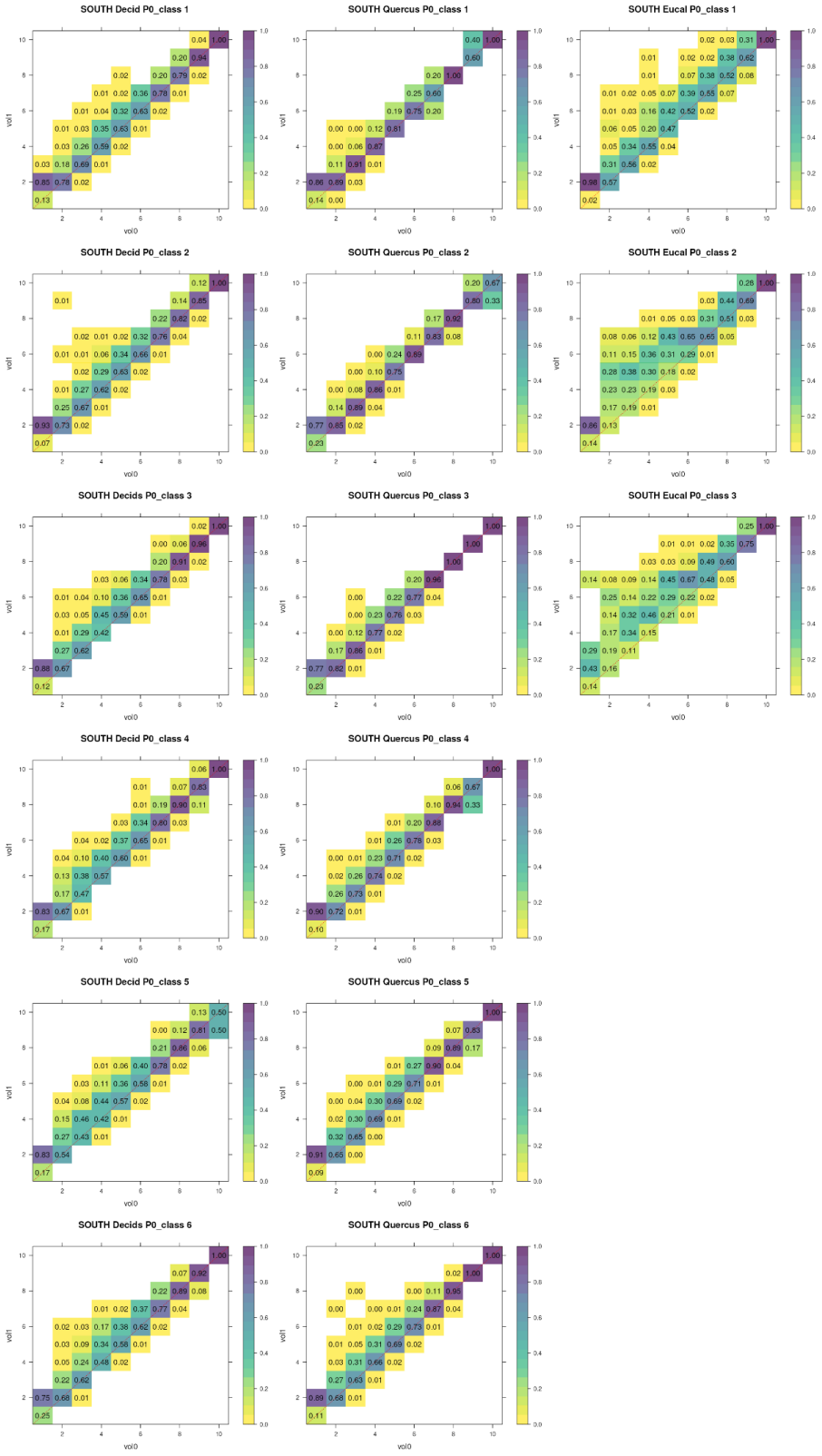
Transition matrices for the CENTR region for Fagus, Quercus and other deciduous sp-groups. Transition matrices for each sp-group are in columns.



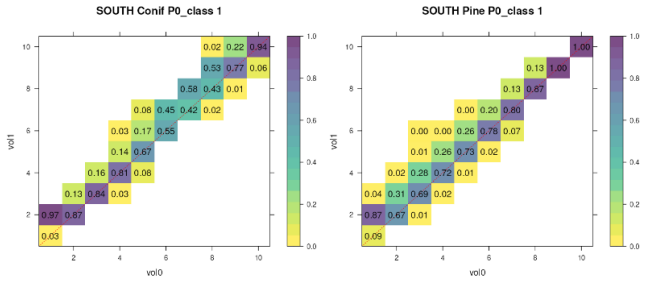
Transition matrices for the CENTR region for other conifers and spruce sp-groups. Transition matrices for each sp-group are in columns.

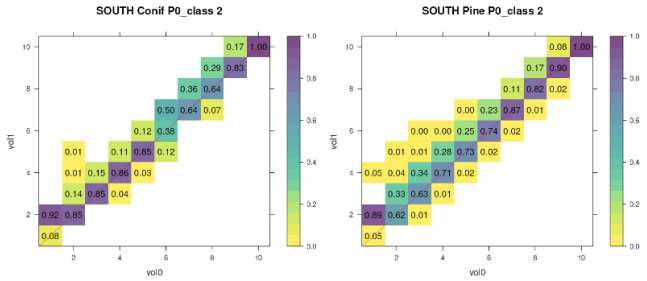


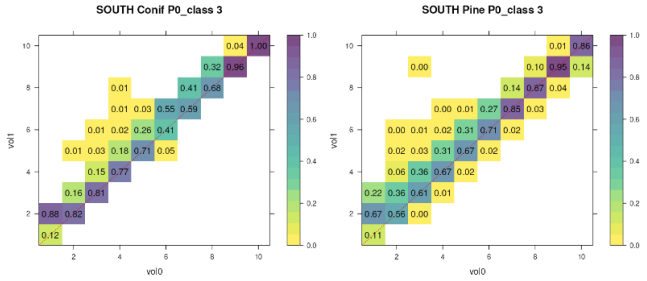
Transition matrices for the SOUTH region for other deciduous, Quercus and Eucalyptus sp-groups. Transition matrices for each sp-group are in columns.

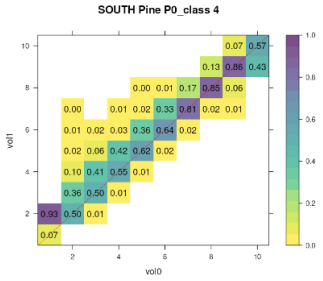


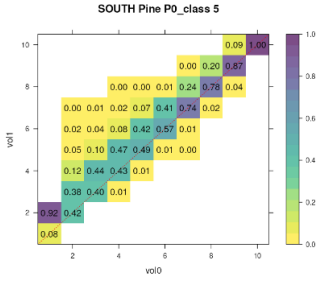
Transition matrices for the SOUTH region for other conifers and pine sp-groups. Transition matrices for each sp-group are in columns.

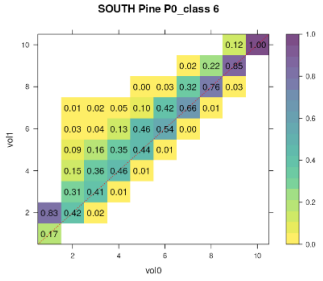






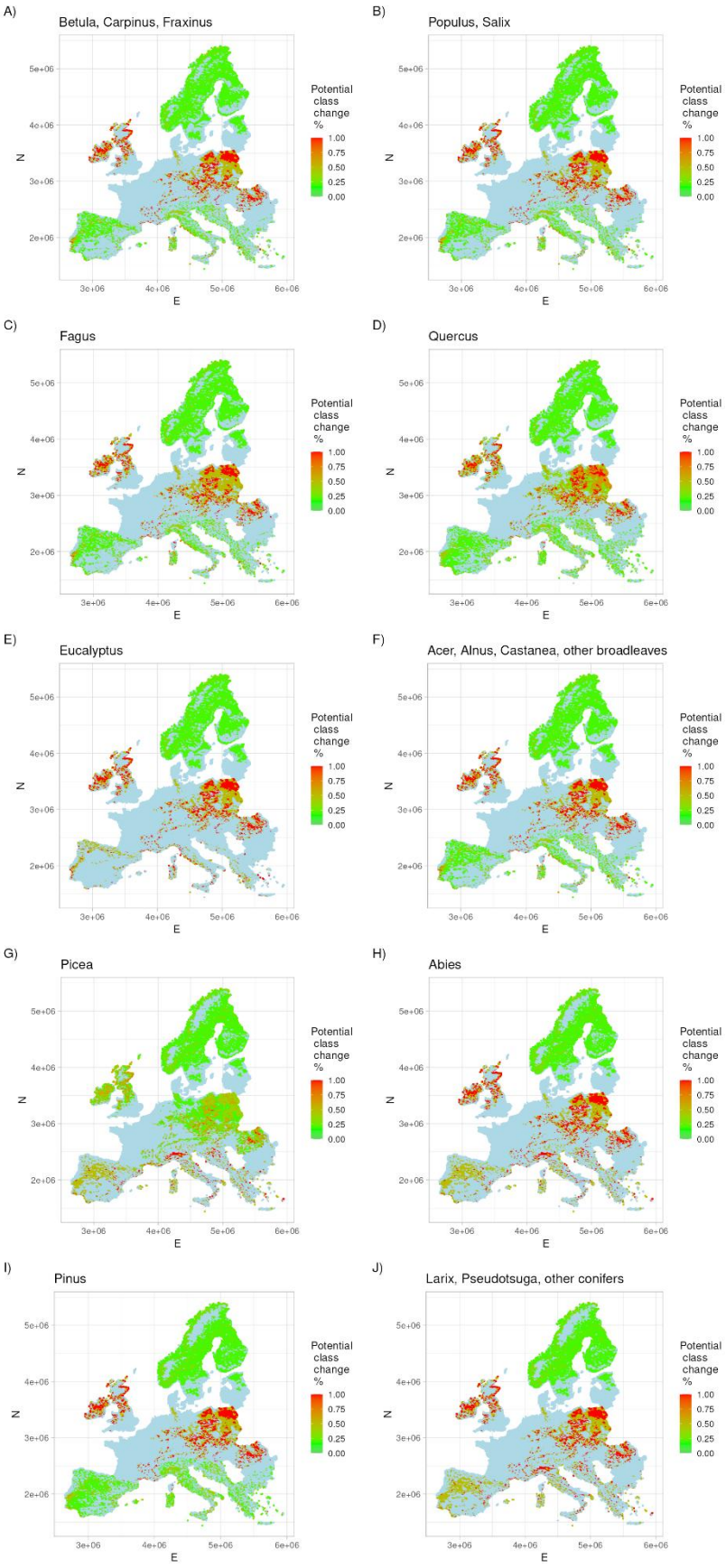




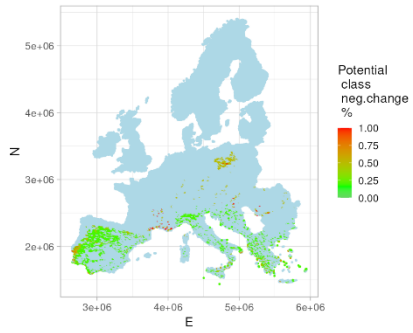


Future climate impact of P0_class ands sp_group on noman growth under RCP45.

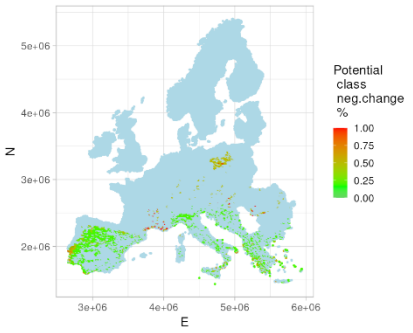
The below figures assume each grid cell contains each of ten sp_groups, and thus illustrate the potential climate change effects on P0_class changes. Figures are calculated by counting the total number of times the gridcell P0_class changes between 2020-2100 (total of 16 time-steps), and total counts are divided by the total number of sp_groups P0_classes minus 1. The two figures below are calculated similarly, but counting the total number of times the gridcell P0_class changes are positive (or negative) between 2020-2100 (total of 16 time-steps), and total counts are divided by the total number of sp_groups P0_classes minus 1. This is an attempt to rescale the total number of P0_class changes, which vary between 3-6 by sp_group and region.


A)
Betula, Carpinus, Fraxinus
B)
Populus, Salix
C)
Fagus
D)
Quercus
E)
Eucalyptus
F)
Acer, Alnus, Castanea, other broadleaves
G)
Picea
H)
Abies
I)
Pinus
J)
Larix, Pseudotsuga, other conifers
Potential class change %
1.00
0.75
0.50
0.25
0.00
N
E
5e+06
4e+06
3e+06
2e+06
3e+06
4e+06
5e+06
6e+06

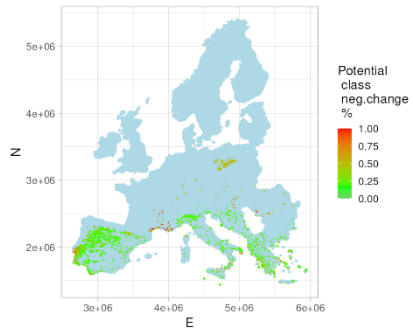
A) Betula, Carpinus, Fraxinus



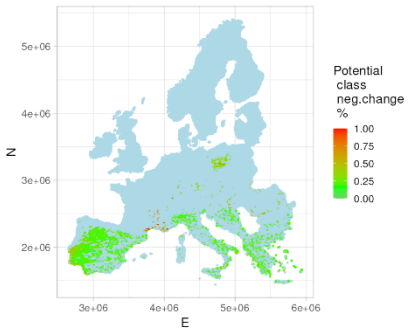
B) Populus, Salix



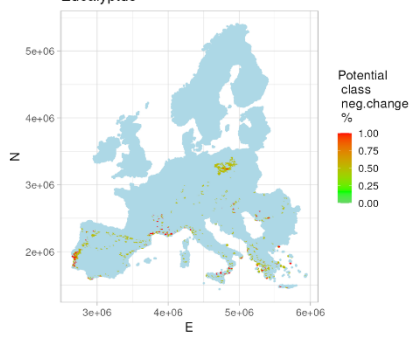
C) Fagus



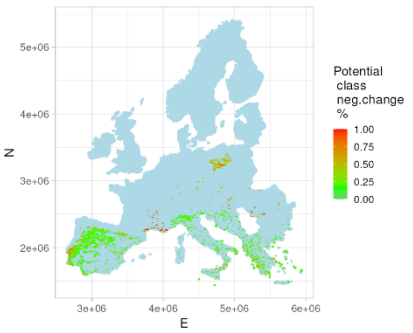
D) Quercus



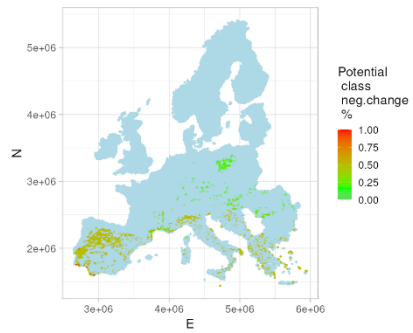
E) Eucalyptus



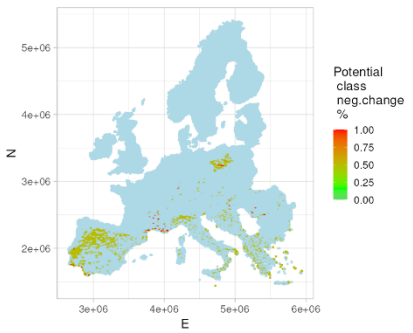
F) Acer, Alnus, Castanea, other broadleaves



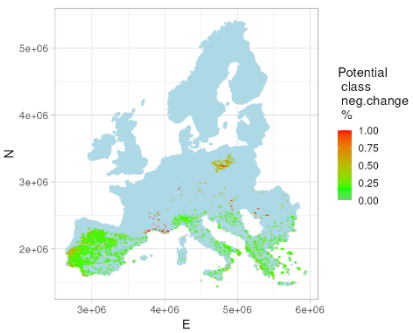
G) Picea



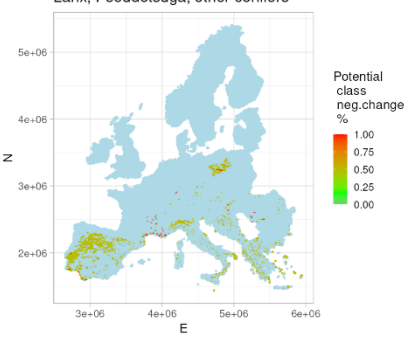
H) Abies

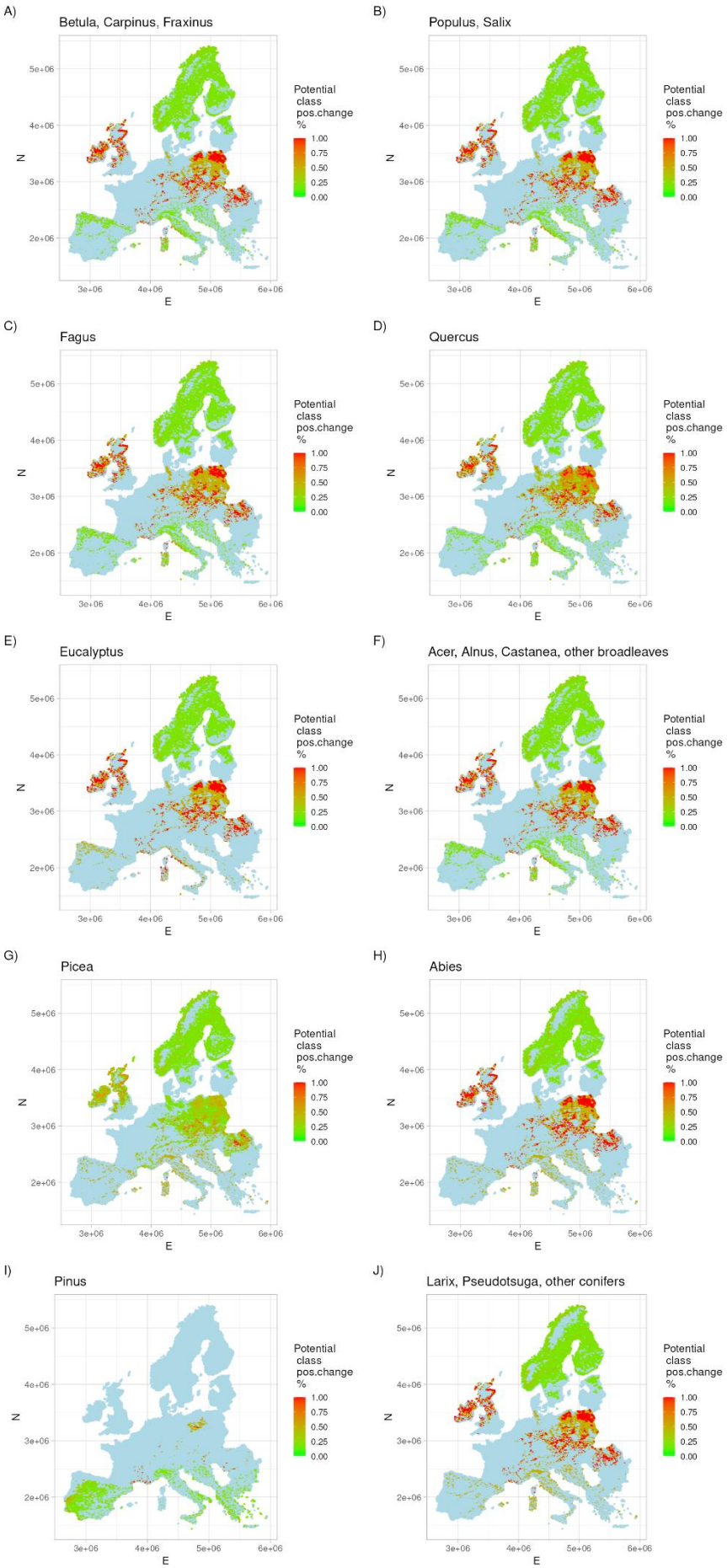


I) Pinus



J) Larix, Pseudotsuga, other conifers




A)
Betula, Carpinus, Fraxinus
B)
Populus, Salix
C)
Fagus
D)
Quercus
E)
Eucalyptus
F)
Acer, Alnus, Castanea, other broadleaves
G)
Picea
H)
Abies
I)
Pinus
J)
Larix, Pseudotsuga, other conifers
Potential class pos.change %
1.00
0.75
0.50
0.25
0.00
N
E
5e+06
4e+06
3e+06
2e+06
3e+06
4e+06
5e+06
6e+06

ANNEX B

Forest disturbance models parameterisation

Wind Damage Model

The ForestGALES wind risk model has been coded in R (*fgr*) (Hale et al., 2015; Locatelli et al., 2022; Quine et al., 2021).

The use of the *fgr* model in R is extensively described in Locatelli et al. (2022). The model takes basic forest stand data and calculates the wind speeds required to cause damage (breakage and overturning) and the probabilities of such damage if information on the wind climate is provided in the form of the Weibull A and k parameters.

Package is available as a zip (fgr_1.0.zip) or tar file (fgr_1.0.tar.gz).

The "Essential" and "Desirable" inputs for the *fgr* model are listed in Table 19. The full range of possible inputs to the model is provided in Table 1 of Locatelli et al. (2022).

Table 19: Minimum input requirements for the fgr wind risk model

Input	Essential (E)/ Optional (O)	Notes	Units
stand_id	E	Label for stand	Characters
tree_id	O	Label for tree	Characters
date	O		Characters
species	E	Uses 2 letter code for different species	Characters
mean_ht/top_ht	E	Stand mean height or stand top height is required	metres
mean_dbh	E	Stand mean dbh (diameter at 1.3 m)	centimetres
spacing	E	Average distance between trees	metres
weib_a	O	Weibull A parameter	Metres/sec
weib_k	O	Weibull k parameter	Unitless

The model outputs from *fgr* are given in Table 20. The probability of stem breakage (prob_b) and probability of overturning (prob_o) can then be multiplied by the standing live volume each time step in the simulation to give the average wind-damaged volume lying on the ground or standing.

Table 20: Outputs from the fgr model

Output	Notes	Units
stand_id	As for inputs	Characters
tree_id	As for inputs	Characters
date	As for inputs	Characters
species	As for inputs	Characters
u_elev_b	Wind speed for stem breakage at set distance above the zero-plane displacement (*d*). Usually set at 10 m above *d* but can be set to any height.	Metres/sec
prob_b	Probability of stem breakage	$Years^{-1}$
u_elev_o	Wind speed for tree overturning at set distance above the zero-plane displacement (d). Usually set at 10 m above d but can be set to any height.	Metres/sec
prob_o	Probability of tree overturning	$Years^{-1}$

Insect Attack Model

A bark beetle population model that originates from LandClim[1] (Schumacher et al., 2004) developed further by Temperli et al. (2013) and implemented in ORCHIDEE by Marie et al. (2024). The model has been written as a package (*IpsR*) that can be installed in R. The Whole model has been written as a new package as part of the Pathfinder project.

Package is available as a zip (*IpsR_0.2.2.zip*) or tar file (*IpsR_0.2.2.tar.gz*).

Table 21: Inputs for IpsR bark beetle damage model.

Input	Essential (E)/ Optional (O)	Notes	Units
id	O	Plot/Stand/Compartment/Cell/Location Number	Characters
Year	O	Year of simulation	Characters
Longitude	O	Longitude in designated coordinate system	Degrees/Number
Latitude	O	Latitude in designated coordinate system	Degrees/Number
Coord_System	O	Coordinate system using ESPG numbering system	Characters
Forest_Type	E	"Boreal" or "Temperate" only at the moment	Characters
av_trees_ha	E	Number of trees per ha on average in compartment/plot/stand/cell	Number
D_cm	E	Average tree diameter	centimetres
DDeff	E	Cumulative effective degree days (degrees x days). Reference value is 547 C days	Degrees x Days
PWSmax	E	Maximum monthly drought level in last 3 years (0 = no drought, 1= extreme drought)	Unitless
FNS_area	E	Fraction of Norway spruce by area in the compartment (0 = no Norway spruce, 1 = pure Norway spruce)	Unitless
Totaldead	E	Total dead biomass that is 1 year or less old. Includes recent mortality from drought and wind damage	Tonnes/Volume/ Carbon mass
bio_total	E	Total living biomass of Norway spruce in the plot/stand/compartment/Cell/Location	Tonnes/Volume/ Carbon mass
ibeetlesmassattack_lastyear	E	Beetle activity in the previous year (o = no activity, 1 = maximum activity)	Unitless
Bbeetleskill_lastyear	E	Biomass killed in the previous year	Tonnes/Volume/ Carbon mass
full_output	O	Designate whether you want a full output with all calculated values or short output with only essential outputs	0 or 1
units_biomass	E	Units that living and dead biomass are submitted to the calculation. Can be any units including tonnes biomass, m^3, kg of carbon, etc. Results will be in the same units.	Characters

[1] http://www.highendsolutions.eu/page/LandClim

Table 22: Outputs from the IpsR bark beetle damage model.

Output	Notes	Units
id	Plot/Stand/Compartment/Cell/Location Number	Characters
Year	Year of simulation	Characters
Longitude	Longitude in designated coordinate system	Degrees/Number
Latitude	Latitude in designated coordinate system	Degrees/Number
Coord_System	Coordinate system using ESPG numbering system	Characters
Psuccesss	Probability of successful attack	0 - 1
Pattacked,	Probability of attack	0 - 1
Total_Dead_Biomass	Total dead biomass that is 1 year or less old. Includes recent mortality from drought and wind damage. Same as Totaldead in Input	Tonnes/Volume/ Carbon mass
Total_Live_Biomass,	Total living biomass of Norway spruce in the plot/stand/compartment/cell/location. Same as bio_total in Input	Tonnes/Volume/ Carbon mass)
units_biomass,	Units that living and dead biomass are submitted to the calculation	Characters
Bbeetlesattacked,	Biomass attacked by beetles in the simulation year.	Tonnes/Volume/ Carbon mass
Bbeetleskilll	Biomass killed by beetles in the simulation year.	Tonnes/Volume/ Carbon mass

Fire Risk Model

The Canadian Forest Fire Danger Rating System[2] in R (*cffdrs::fwi* and *cffdrs::fwiraster*) (Van Wagner, 1987; Wang et al., 2015) is recognized for wildland fire risk assessment, main components are considered in the project include: Fire Weather Index System (FWI) and Fire Behavior Prediction System (FBP). The *cffdrs* package is available on R repository for instant installation and includes both *fwi* and *fbp* as well as other 9 functions to calculate FWI and FBP.

[2] https://osdp-psdo.canada.ca/dp/en/search/metadata/NRCAN-CFS-1-19973

Table 23: Inputs for the fwi and fbp fore models

Input	Essential (E)/ Optional (O)	Notes	Units
fwi			
id	O	Weather station or spatial point number	characters
lat	O	Latitude	decimal degree
long	O	Longitude	decimal degree
yr	O	Year of observation	numbers
mon	O	Month of the year	numbers
day	O	Day of the month	numbers
temp	E	Temperature	degree Celcius
rh	E	Relative humidity	percentage (%)
ws	E	10-m height wind speed	kilometers per hour
prec	E	24-hour rainfall	milimeter
ffmc	E	Previous day Fine Fuel Moisture Code	unitless
dmc	E	Previous day Duff Moisture Code	unitless
dc	E	Previous Day Drought Code	unitless
fbp			
FuelType	E	FBP System Fuel Type (16 types)	characters
Lat	E	Latitude	decimal degree
Long	E	Longitude	decimal degree
ffmc	E	Fine fuel moisture code (output of fwi system)	
bui	E	Buildup Index (output of fwi system)	
ws	E	Wind speed	Kilometer per hour
gs	E	Ground slope	percentage (%)
Dj	E	Julian day	numbers
Aspect	E	Aspect of the slope	decimal degrees
fbp (optional inputs if relevant fuel types do not appear in the input data)			
PC	O	Percent Conifer for M1/M2	percentage (%)
PDF	O	Percent Dead Fir for M3/M4	percentage (%)
cc	O	Percent Cured for O1a/O1b	percentage (%)
GFL	O	Grass Fuel Load	g/m^2
fbp (optional without causing impacts to primary outputs)			
CBH	O	Crown to Base Height	meters
WD	O	Wind direction	decimal degrees
Accel	O	Acceleration: 1 = point, 0 = line	1, 0
ELV	O	Elevation	meters above sea level
BUIEff	O	Buildup Index effect: 1=yes, 0=no	1, 0
D0	O	Julian day of minimum Foliar Moisture Content	numbers
hr	O	Hours since ignition	numbers
isi	O	Initial spread index (output of fwi system)	numbers
CFL	O	Crown Fuel Load	Grams per square meter

FMC	O	Foliar Moisture Content if known	percentage (%)
SH	O	C-6 Fuel Type Stand Height	meter
SD	O	C-6 Fuel Type Stand Density	stems per hectare
theta	O	Elliptical direction of calculation	degrees

Table 24: Outputs for the cffdrs fire danger rating system

Output	Notes	Units
fwi		
ffmc	Fine Fuel Moisture Code	unitless
dmc	Duff Moisture Code	unitless
dc	Drought Code	unitless
isi	Initial Spread Index	unitless
bui	Buildup Index	unitless
fwi	Fire Weather Index	unitless
dsr	Daily Severity Rating	unitless
fbp (primary outputs)		
CFB	Crown Fraction Burned by the head fire	numbers
CFC	Crown Fuel Consumption	Kg/m^2
FD	Fire description (1=Surface, 2=Intermittent, 3=Crown)	1, 2, 3
HFI	Head Fire Intensity	kW/m
RAZ	Spread direction azimuth	degrees
ROS	Equilibrium Head Fire Rate of Spread	m/min
SFC	Surface Fuel Consumption	Kg/m^2
TFC	Total Fuel Consumption	Kg/m^2
fbp (secondary outputs)		
BE	BUI effect on spread rate	numbers
SF	Slope Factor	numbers
ISI	Initial Spread Index	numbers
FFMC	Fine fuel moisture code (FWI component)	numbers
FMC	Foliar Moisture Content	percentage (%)
Do	Julian Date of minimum FMC	numbers
RSO	Critical spread rate for crowning	meters per minute
CSI	Critical Surface Intensity for crowning	kilowatts per meter
FROS	Equilibrium Flank Fire Rate of Spread	meters per minute
BROS	Equilibrium Back Fire Rate of Spread	meters per minute
HROSt	Head Fire Rate of Spread at time hr	meters per minute
FROSt	Flank Fire Rate of Spread at time hr	meters per minute
BROSt	Back Fire Rate of Spread at time hr	meters per minute
FCFB	Flank Fire Crown Fraction Burned	numbers
BCFB	Back Fire Crown Fraction Burned	numbers
FFI	Equilibrium Spread Flank Fire Intensity	kW/m
BFI	Equilibrium Spread Back Fire Intensity	kW/m
FTFC	Flank Fire Total Fuel Consumption	Kg/m^2
BTFC	Back Fire Total Fuel Consumption	Kg/m^2
DH	Head Fire Spread Distance after time hr	meters
DB	Back Fire Spread Distance after time hr	meters
DF	Flank Fire Spread Distance after time hr	meters
TI	Time to Crown Fire Initiation (hrs since ignition)	numbers
FTI	Time to Flank Fire Crown initiation (hrs since ignition)	numbers
BTI	Time to Back Fire Crown initiation (hrs since ignition)	numbers
LB	Length to Breadth ratio	numbers
LBt	Length to Breadth ratio after elapsed time hr	numbers
WSV	Net vectored wind speed	Km/hr
TROS	Equilibrium Rate of Spread at bearing theta	meters per minute
TROSt	Rate of Spread at bearing theta at time t	
TCFB	Crown Fraction Burned at bearing theta	numbers
TFI	Fire Intensity at bearing theta	kW/m
TTFC	Total Fuel Consumption at bearing theta	Kg/m^2
TTI	Time to Crown Fire initiation at bearing theta (hrs since ignition)	numbers

Currently the *fwi* and *fbp* models require hourly or daily inputs to run but the climate simulations are rovided monthly or annually. Need to work out a way to use the *cffdrs* fire danger rating system using monthly data.

Drought Probability and Impact Model

The SPEI index for the last 3 years is used to calculate PWSMax that is used in the *IpsR* model (see Table 21). SPEI and HWMI are both used for Drought tree mortality statistical model. *SPEI* and *extRemes* are available R packages to calculate these indices.

Table 25: Inputs for the SPEI drought index and HWMI

Input	Essential/ Optional	Notes	Units
T_{mean}	E	Mean 2 m air temperature in the last 24 hours	Degree celcius
T_{max}	E	Maximum 2m air temperature in the last 24 hours	Degree celcius
T_{min}	E	Minimum 2m air temperature in the last 24 hours	Degree celcius
EP	E	Evaporation	meter
P	E	Mean precipitation flux	meter
RH	O	Relative humidity	g/m^3

Table 26: Outputs from the SPEI drought index and HWMI

Output	Notes	Units
SPEI	Index of drought ranging typically from -2 (extreme drought) to +2 (very wet).	unitless
PWSMAx	Calculated as a conversion of the lowest annual SPEI value in the last 3 years ($1-(SPEI+2)/4$. Value of 0 = no water stress, 1 = maximum stress	unitless
HWMI	Index of extreme heatwave ranging typically from 0 (normal) to 100 (ultra extreme).	unitless