Deliverable

D2.2 Description of models and estimators

PathFinder Project

Version: 1.0

28 February 2025

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

I. DOCUMENT CONTROL

Project	PathFinder (101056907)
Project Title	Towards an integrated consistent European LULUCF monitoring and policy pathway assessment framework
Date	28 February 2025
Author/s	Mari Myllymäki (LUKE), Jukka Miettinen (VTT), Johannes Breidenbach (NIBIO)
Reviewer/s	Cédric Vega and Mirco Migliavacca
Activity	WP2, Task 2.2 Modelling and mapping forest parameters including carbon stocks, carbon stock changes and biodiversity
Filename	PathFinder_Deliverable_D2.2

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

<u>PU</u>	<u>Public, fully open access</u>
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
0.1	17.1.2025	See above	Full draft for internal review
1.0	28.2.2025	See above	Internal reviewers' comments integrated

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

1. Contents

IV. Summary

In this deliverable, we describe models and estimators that have been considered and are used in PathFinder to link the field and remote sensing data for mapping and estimation. A prototype version of the mapping and estimation platform was configured already, and European wide maps of volume, biomass, and conifer proportion were produced with the *k*-Nearest Neighbour (*k*-NN) method (Miettinen et al., 2024). We reviewed prerequisites and challenges in producing forest attribute maps in European context (Myllymäki et al., unpublished). We compared the *k*-NN method with alternative approaches allowing simultaneous prediction of several forest attributes for high-resolution map production in the model-based approach. The approaches were compared particularly with respect to the training data availability in the spatial neighbourhood (Balazs et al., unpublished). Further, regression and machine learning techniques were tested in national studies (Koma and Breidenbach, 2025; Toivonen et al., 2024). PathFinder supports these national studies, because more advanced remote sensing data are available on the national or sub-national scale that may also become available on the Pan-European scale in the future. Koma and Breidenbach (2025) found that resolution effects due to differences in plot sizes may be of low relevance in operational settings. Other studies on alternative modelling techniques including deep learning are ongoing (Koma and Breidenbach, unpublished). The models and estimators were compared also for producing small-area estimates (Kangas et al., 2025, unpublished). Further, we have aimed to enhance the quantification of the uncertainty associated with the local map predictions. First, we reviewed map uncertainty quantification (Kangas et al., 2023). Secondly, we provided pixel-wise uncertainty metrics for the European wide maps produced (Miettinen et al., 2024). Thirdly, we proposed new valid uncertainty intervals particularly for the *k*-NN method, which were shown to contain the true value with the desired probability (e.g., 90%) and perform comparable to those produced by conformal quantile regression (Kuronen et al., 2024). The outcomes of the analyses are reported in scientific publications. In this document we highlight the main findings and their implications to the final demonstration of the PathFinder platform.

1. Introduction

This deliverable can be understood as a summary of work done in PathFinder on models and estimators that link the field and remote sensing data for mapping and estimation (deliverable type 'OTHER'). While several modelling approaches use estimators to determine model parameters, this deliverable focuses on the description of models that have been used or tested. Estimators in the more common sense of utilizing National Forest Inventory (NFI) data will be described in Deliverable 2.3.

Below, we detail the research done regarding the first European wide maps using the *k*-NN method (Section 2), experiments with alternative modelling approaches (Section 3), uncertainty quantification for maps (Section 4) and model-based small-area estimation (Section 5). We conclude with an outlook on the choice of the models for the PathFinder platform and its demonstration (Section 6). We refer to the scientific publications for more details.

2. The first European wide maps and the *k*-Nearest Neighbor (*k*-NN) method

To work towards Objective 2.i) "Development of a processing system for big remote sensing and field plot datasets" and the second part of Objective 2.iv) "…creation of maps and estimates as inputs to other tasks", a prototype version of the mapping and estimation platform was configured. The *k*-NN method was chosen as the forest attribute model for the prototype version. The method is used to link the remote sensing and field plot data. The choice was based on the following reasons: a) the *k*-NN method has been successfully applied to national level mapping of forest attributes with NFI field plots and satellite imagery in some of the member countries, b) implementation of it was partly available in the platform component used for the mapping in PathFinder (Forestry TEP). As the *k*-NN method has been widely used for forest monitoring in Europe and globally, it also provides a good benchmark for testing and comparing other potential methods.

The *k*-NN method was tested in a pilot study by producing pan-European forest attribute maps, within the PathFinder platform system (Miettinen et al., 2024). Outputs of the pilot study include pan-European maps of volume, above-ground biomass (AGB), and conifer proportion in 10 m spatial resolution. The *k*-NN prediction for a forest attribute at a 10 m grid cell is a weighted average of the attribute of *k* observations from training data that are deemed to be the *k* nearest neighbours of the target cell. The neighbours are determined by a distance in the space of EO data. The choice of the *k*, as well as auxiliary Earth Observation (EO) based datasets, affect the predictions. Increasing the number of *k* would typically decrease the root mean squared error (RMSE) and smoothen the visual impression of the map, but it would also limit prediction of very high values and in some case induce bias. Based on the experiments done in selected test sites in different parts of Europe, the number of *k*=7 nearest neighbours turned out to be a good compromise. Regarding the auxiliary datasets, it was decided that the Copernicus High Resolution Layer (HRL), Forest Type (FTY) and Tree Cover Density (TCD) products as well as the University of Maryland Global Forest Change (GFC) product were used in the mapping in addition to the Sentinel-2 spectral information of seven spectral bands. The FTY and TCD products were noticed to improve the accuracy of the predictions slightly, while the GFC product was used to mask out plots with potential temporal mismatch with the remotely sensed data and field measurements.

The maps were validated computing RMSEs in a test data in different countries. Initial comparisons of accuracies were done with national feature banks, i.e. feature banks that include data from only one country. Feature bank refers to the collection of field measurements and spectral values (+FTY and TCD) which is used in the k-NN prediction. For each country, 1/3 of the plots were taken aside (with systematic sampling of plots ordered with respect to total volume) and used as independent testing dataset, while the remaining 2/3 of the plots were used in the k-NN prediction. The country-wise percent RMSEs computed from the test data varied between 45-72% for volume, between 45-66% for biomass, and 27-100% for the conifer proportion depending on the region. The final maps were done by dividing Europe into 14 areas based on ecological similarities and the availability of field reference data. Accuracy metrics were calculated separately for all these 14 areas. The area-specific percent RMSEs varied between 56-83% for volume, between for 53-72% biomass, and 30-110% for the conifer proportion depending on the region.

3. Alternative modelling approaches

The *k*-NN maps of the pilot study were considered satisfactory based on the validation metrics and visual inspection (Miettinen et al., 2024). However, when forest attributes are mapped in areas where no field plot data are available in the vicinity, the bias and RMSE can be large. Even though the field data collected under the PathFinder project and utilized for the mapping is extensive, there are some gaps in the data spatially, i.e., regions without NFI plots. To illustrate how large errors are possible in such situations, we computed validation metrics for the *k*-NN map product of the pilot study leaving out the data from different countries and trying to predict the forest attributes with other countries' data. Among six tested countries, the bias was 16% on average. Therefore, it is important to carefully choose the field plots to be used in the forest attribute model for countries without available field measurement data. The used plots should represent as closely as possible the ecological and geographic characteristics of the target area. The plots which are geographically closest to the empty area may often be a reasonable choice.

To better understand the quality of maps and differences between different modelling approaches, we compared the performance of the *k*-NN to alternative models, namely the Random Forests (RF) and Multi-Layer Perceptron (MLP) methods, for creating large-scale forest resource maps (Balazs et al., unpublished). The study was designed to better understand the bias due to the model choice and model extrapolation, which we considered an important, yet not so well understood, aspect for the usefulness of the map (Myllymäki et al., unpublished). We tested models that were fitted locally using training data from different neighborhoods. The experiment was done with data from Norway, Sweden, and Finland, where the training data available in PathFinder from the respective NFIs was spatially comprehensive and allowed us to validate the maps in case of (artificially created) spatial gaps of varying sizes. The same forest attributes as in the initial maps were used (volume and AGB of total growing stock, conifers, and broad-leaved species). We considered including other models, but finally settled down to these three alternatives because they allowed straightforwardly simultaneous modelling of all the attributes. We found that the RF and MLP led to slightly smaller RMSEs than *k*-NN at the country level, but MLP had clearly higher bias than the other two methods, particularly if the number of field plots used for training the local model was not sufficiently large. The effect of the spatial gaps was surprisingly minor on the country-level results. However, when evaluating model performances at smaller scales, we observed smaller RMSE and larger biases regardless of the model used as well as larger effects of the number of training data used for training local models and the spatial gaps.

In a scientific study, tree boosting with random effects (GPBoost) was compared to linear mixed effects (LME) modelling in predicting forest age (Toivonen et al., 2024). The GPBoost approach was found to improve forest age prediction with RMSE values down to 36.3% in comparison to 38.2% of the LME approach and to produce better predictions for old forests (over 150 years old). Thus, the GPBoost approach showed potential for better identification of old-growth forests. Other studies on the use of deep learning methods are ongoing and include Koma and Breidenbach (unpublished). Initial analysis indicate that deep learning can be slightly more precise than the state-of-the-art benchmark method (a linear mixed model).

Plot sizes differ among NFIs of different countries. When creating maps based on NFI data from different countries, this could increase the uncertainty of models. In PathFinder, we mitigate issues related to this by extracting satellite data always for a circle of 100 m^2 around the plot center. In this way, the extracted resolution corresponds to the predicted resolution (10 m pixel side length). For the sake of harmonization, we accept that the resolution of the field measurement differs from the resolution of the prediction maps. Related to this, Koma and Breidenbach (2025) compared forest attribute predictions across different resolutions using lidar and NFI data in Norway. A model was fit at the field-plot resolution (16 m) and applied on resolutions ranging from 1-30 m pixels. The prediction maps were then aggregated for independent validation stands. The effect of the resolution was moderate for additive variables such as volume and biomass, with the highest accuracy and precision achieved for 10 m pixels. For non-additive variables such as mean height, scaling effects are in fact severe.

4. Uncertainty quantification for forest attribute maps

Forest attribute maps without uncertainty assessment may lead to false impressions of precision (Kangas et al., 2023). Uncertainties provided with predictions, e.g., in terms of prediction intervals, can improve understanding and help to avoid wrong conclusions. Such prediction intervals can be provided with maps as additional layers or in an interactive manner (e.g. Norwegian SR16 forest resource maps).

For the *k*-NN predictions, the leave-one-out RMSE has been used conventionally as a grid cell level uncertainty measure (e.g. Finnish multi-source NFI maps). It can be provided for areas with sufficiently many field plots, but several other approaches have aimed to provide grid cell-specific uncertainties, and thus uncertainty layers. It is however not clear how to best characterize the uncertainty related to the *k*-NN predictions and several approaches have been suggested (see e.g. Kuronen et al., 2024). The initial *k*-NN maps were provided with pixel-level uncertainties based on the standard deviation of the *k* nearest neighbours (Miettinen et al., 2024). In small-area estimation context, we tested an alternative, the root mean squared error of the *k* nearest neighbours (Kangas et al., 2025). These prediction intervals provide information on the variability of the uncertainty across the space, but they do not generally achieve the desired (e.g., 90%) coverage, which means that the prediction intervals can cover the ground truth value less or more often than desired. In Kuronen et al. (2024), we proposed prediction intervals for forest attribute maps based on conformal prediction methodology, particularly for the *k*-NN method. Conformal prediction is a method to make proper uncertainty estimates (desired coverage) from heuristic uncertainty estimates (not desired coverage). We showed that the proposed conformal *k*-NN procedures produce valid prediction intervals in the sense that they contain the true value with the desired probability, for example 90%. We also showed that there are multiple methods for *k*-NN to produce prediction intervals competitive with those produced by conformal quantile regression, considered as one of the best ways in the literature for obtaining (continuous) prediction intervals. The validity of the conformal *k*-NN methods was shown for the so-called jackknife conformal prediction, which allows to utilize all the available training data without the need to leave part of the training data to a separate calibration set. We see conformal prediction as a recommendable unit-level uncertainty quantification method for forest attribute maps (please refer to Kuronen et al., 2024, for discussion of future considerations).

5. Use of models in small-area estimation

Forest attribute maps are often desired to be utilized in estimating results for small areas (e.g., municipalities). Direct design-based estimators utilizing data only from the inspected domain would be desirable, as they provide unbiased estimates. They are available if there are field sample plots within the domain. If the number of sample plots is small, they can be used, but they may produce unacceptably large standard errors for the estimates. On the other hand, indirect estimators that use data also from outside the domain (e.g. the mean of the *k*-NN predictions within the domain) are the only alternative, when no samples are available from the domain of interest, but they can exhibit bias. Thus, in the case of a few sample plots available from the small area, a composite estimator, i.e. a weighted sum of a direct and an indirect estimator, may be an attractive option balancing the pros and cons of both estimators. In Kangas et al. (2025), we studied direct, indirect, and composite estimators for small-area estimation in a simulation experiment with varying size small areas and varying quality auxiliary data. Based on the study, composite estimators were indeed the most recommendable of the tested estimators whenever there was at least one sample plot available from the target domain, but further studies are still required on many aspects including variance estimators for the model-based k-NN estimator as their uncertainty is likely to be underestimated (see Kangas et al., 2025). In a submitted work, we also investigate the effect of sampling design on model-based small-area estimates (Kangas et al., unpublished). The differences between the sampling designs were small. With all designs and estimators, poor results concentrated on certain areas, usually those with most valuable forests. In ongoing work, we explore means to improve predictions in these areas.

6. Outlook to the final demonstration of the platform system

Due to the positive results obtained by the *k*-NN method in producing the initial European wide maps, it was decided that the final demonstration of the PathFinder platform system will also utilize the *k*-NN method. The studies with alternative models have shown that it is possible to improve precision for the forest attribute predictions using other machine learning methods. However, the *k*-NN method had rather comparable behaviour in many cases, and it has several key advantages that make it suitable for the PathFinder platform. Firstly, it is fast to run compared to many other models, which makes it attractive for operational continental level mapping. Secondly, the derivation of required error metrics and layers is more feasible with *k*-NN method than other alternatives. This is particularly important for the PathFinder platform that utilizes model-assisted estimation. Furthermore, there is little knowledge on the optimal technical and operational approaches to apply other methods for continental level mapping utilizing hundreds of thousands of NFI plots from around Europe.

Although the final demonstration maps will be made with the *k*-NN method, the tests with other forest attribute mapping models will continue because *k*-NN also has some disadvantages. One of the disadvantages is that the data are an integral part of the model. The data cannot be shared, because the extracted satellite data may be used to reverse-engineer the exact plot coordinates. Therefore, the *k*-NN models cannot be shared either. In case an enhanced model is found for continental level operational (annual) mapping, it can be taken in use in the future. The PathFinder platform's mapping module is not limited to using any particular method but can accommodate the use of any mapping algorithms. Already during the PathFinder project there may be an opportunity to test and compare other mapping methods to the *k*-NN approach at large scale, but subcontinental level. Feasibility of these tests will be evaluated after the final demonstration maps that are required by other WPs have been produced.

Based on the findings of the pilot and other studies made, several improvements were made to the way the *k*-NN method is applied in the final demonstration. Firstly, a leave-one-out cross validation (CV) approach was developed and implemented for deriving error metrics for model-assisted estimation. Using this CV approach allows to provide valid uncertainty metrics in terms of leave-one-out RMSE utilizing all field sample plots in the map-making, compared to only 2/3 of plots used in the pilot study. Another major improvement relates to the way that the NFI plots are used in *k*-NN. In the pilot study, several feature banks were used and the continental map was processed in 14 different processing areas (Miettinen et al., 2024). In the final demonstration, one continental feature bank is created and *n* closest NFI plots for each Sentinel-2 tile are used in the mapping. This approach ensures the use of the closest NFI plots and the smooth transition of the feature banks, reducing visible borders between processing areas.

In addition to the two major improvements described above, several minor modifications have been implemented. To prepare for future annual mapping, only the most recent EO datasets will be used in the mapping to ensure temporal consistency. No possibly outdated auxiliary products (such as the 2018 Copernicus tree cover density or forest type products) will be used. The plan is to run the final demonstration with 2024 satellite data and include environmental variables (namely elevation as well as mean annual precipitation and temperature over 30 years) into the explanatory features. Furthermore, an automated system for outlier screening was implemented to operationalize the feature bank preprocessing.

The final demonstration maps will also include more output variables than the pilot study maps. In addition to the volume, above ground biomass and conifer proportion (and their uncertainties), several other variables will be mapped. These include gross and net annual increment, diameter, basal area, height, species count and the proportions of ten species groups. The accuracy of the new output layers will be evaluated before deciding on which final demonstration maps to publish.

In addition to the annual map production, the final demonstration will allow evaluation of several change monitoring methods. In the optimal case, monitoring of changes in the PathFinder platform is conducted through the design-based and model-assisted estimation processes. However, this may not always be feasible e.g. due to field data limitations or temporal requirements. In these kinds of situations, change monitoring based on maps may be feasible to acquire coarse estimates of change trends (see Thoppil et al., 2024). The final demonstration of the PathFinder platform will allow evaluation of three main approaches:

1. Mapping of increment
2. Combining biomass basemap with change detection
3. Map-to-map comparison

As already highlighted above, the feasibility of the mapping gross and net annual increment will be evaluated in the final demonstration. This evaluation may also directly benefit the development of a standardized baseline for forestry, which will be inspected in the newly added PathFinder Task 2.5. If feasible, this approach would greatly improve the understanding of changes particularly in those forest areas that have not experienced any clearly detectable drastic changes. The increment maps could be used either alone or together with baseline maps and change detection results. In addition to overlaying forest clearance maps with baseline biomass maps to calculate loss of biomass, one could include also estimates on increment in the remaining forest areas. Finally, the plan to use 2024 satellite data in the final demonstration would also allow evaluation of map-to-map change comparison using the above ground biomass maps of the pilot study and the final demonstration (2020 vs. 2024). The feasibility of the evaluation of all the change monitoring approaches will depend on the decisions made during the final demonstration and the resulting output products.

References

Balazs, A., Miettinen, J., Nilsson, M., Breidenbach, J., Pitkänen, T.P., Myllymäki, Mari, unpublished. Large-scale forest resource mapping with local models and spatial gaps in the training data: Comparison of different modelling approaches.

Kangas, A., Myllymäki, M., Mehtätalo, L., 2023. Understanding uncertainty in forest resources maps. Silva Fenn. 57. https://doi.org/10.14214/sf.22026

Kangas, A., Myllymäki, M., Packalen, P., 2025. Small area estimators in a simulation test. Can. J. For. Res. 55, 1–17. https://doi.org/10.1139/cjfr-2024-0070

Kangas, A., Myllymäki, M., Packalen, P., unpublished. The effect of sampling design in model-based small area estimation.

Koma, Z., Breidenbach, J., 2025. Large-scale validation of forest attribute maps across different spatial resolutions. https://doi.org/10.5281/zenodo.14626363

Koma, Z., Breidenbach, J., unpublished. Forest attribute mapping with point cloud CNNs.

Kuronen, M., Räty, J., Packalen, P., Myllymäki, M., 2024. Uncertainty Quantification for Forest Attribute Maps with Conformal Prediction and K-Nearest Neighbor Method. https://doi.org/10.2139/ssrn.5035783

Miettinen, J., Adame, P., Adolt, R., Alberdi, I., Antropov, O., Arnarsson, Ó., Astrup, R., Berger, A., Bogason, J., Chirici, G., Corona, P., D'Amico, G., Fejfar, J., Fischer, C., Gohon, F., Gschwantner, T., Hertzler, J., Koma, Z., Korhonen, K.T., Krajnc, L., Latte, N., Lejeune, P., McCullagh, A., Mionskowski, M., Moreno, D., Myllymäki, M., Nilsson, M., Perin, J., Pitkänen, J., Redmond, J., Riedel, T., Schumacher, J., Seitsonen, L., Sirro, L., Skudnik, M., Snorrason, A., Sroga, R., Traub, B., Westerlund, B., Wurpillot, S., Breidenbach, J., 2024. High-Resolution Pan-European Forest Structure Maps: An Integration of Earth Observation and National Forest Inventory Data. https://doi.org/10.5281/zenodo.13143235

Myllymäki, M., Moreno-Fernández, D., Miettinen, J., Packalen, P., Kuronen, M., Räty, M., Berger, A., Vega, C., Morneau, F., Skudnik, M., Cescatti, A., Breidenbach, J., Kangas, A., unpublished. Prerequisites and challenges of producing forest resource maps with remote sensing in Europe.

Thoppil, M., House, J., Breidenbach, J., Astrup, R., 2024. A European reporting and accounting framework for the LULUCF Sector. Zenodo. https://doi.org/10.5281/zenodo.14616476

Toivonen, J., Kangas, A., Pitkänen, T.P., Myllymäki, M., Maltamo, M., Kukkonen, M., Packalen, P., 2024. Tree Boosting with Linear Forest-Type Effects Improves the Performance of Forest Age Predictions Using Airborne Laser Scanning and Satellite Data. https://doi.org/10.2139/ssrn.4886824