Deliverable D1.3

Data curation for the adoption of YASSO and CBM models

PathFinder Project

Version: 1.1

June 21st, 2024

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	21st of June, 2024
Author/s	Heleen Deroo (EVINBO), Nathalie Cools (EVINBO), Claudia Guidi (WSL), Markus Didion (WSL)
Reviewer/s	Jari Liski (FMI) (TBC), Johannes Breidenbach (NIBIO)
Activity	WP1, Task 1.3 [Assessing carbon pools and fluxes]
Filename	PathFinder_Deliverable1.3_version_1.1

DISSEMINATION LEVEL

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
0.1	2024-05-21	Deroo, Cools	First draft
0.2	2024-06-07	Guidi, Didion	Data preparation for litterfall input functions has been included
0.3	2024-06-12	Liski, Breidenbach	Internal review
1.0	2024-06-21	Cools, Deroo	Incorporation of reviewer feedback
1.1	2024-06-21	Breidenbach	Final check and minor adjustments

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
SOC	Soil Organic Carbon
OC	Organic Carbon
CI	Confidence Interval
NFI	National Forest Inventory

Content

IV. Summary

The objective of Task 1.3 was to analyse the NFI, ICP Forests, and LUCAS data to harmonize and improve the assessment of carbon fluxes among living biomass, deadwood, litter and soil pools. After exploration of the available data sets, the ICP Forests Level II data set was found the most suitable to set the initial modelling conditions as it is the only data set that comprises above- and below-ground compartments and contains repeated assessments on a subset of about 200 plots across Europe. This deliverable describes the pre-processing of the soil data in order to calculate the soil carbon stocks and changes over time. The pre-processing of the growth and litterfall data is included in the Deliverable D.3.3. Based on the improved data sets where the field observations on growth, litterfall and soil were available, the Yasso model will be further adopted, validated, and improved within Task 3.3.

1. Introduction

In this deliverable, we describe how NFI, ICP Forests, and LUCAS data have been analysed to harmonize and improve the assessment of carbon fluxes among living biomass, deadwood, litter and soil pools. Data on these carbon fluxes are used as input, calibration and validation data of the Yasso model in Work Package 3, Task 3.3. Within the available timeframe (project month 1 -22) no concrete data request for application within the CBM model was made to Task 1.3 and so is not further considered in this report.

In particular, we focused on pre-processing data from the ICP Forests monitoring programme (the Level II intensive monitoring network in particular), which was found to be the most comprehensive and harmonized set of field data across European forests, encompassing both the above-ground and soil compartments (until a depth of 80 cm) on the same plots. Despite its potential, the data still required significant validation and harmonisation efforts, which constituted a large part of the time dedicated to Task 1.3 (soil data) and Task 3.3 (growth and litterfall data).

2. Analysis of the available data sources

2.1 Selection criteria

As input data for the modelling in Yasso, it was essential to start for a EU wide dataset that included carbon data on both the above-ground and below-ground carbon storing compartments.

2.2 NFI data

For our purpose of assessing litter and deadwood fluxes from living biomass to the soil compartment, the availability of consistent relevant data collected by NFIs is concentrated on the assessment of above-ground carbon stocks in the living and deadwood biomass. However, several NFIs include a soil data module within their monitoring programme. However, the soil information is very heterogeneous across different countries. So within the timeframe of this task and the available resources it was not possible to set up a centralized soil module to collect harmonised soil data collected by the NFIs (see Task 1.1).

Though, for a number of countries the plots where the harmonised ICP Forests soil monitoring programme was performed, form a subset of the NFI plots (see Table 1).

Table 1: Relation between National Forestry Inventories and ICP Forests Level I soil surveys.

Country	same sample	co-located	no relation	no information	Total
Norway	403		19	11	433
Romania	235			7	242
Germany	95		212	162	469
Sweden	40	167		807	1014
Austria		134		5	139
Hungary		77	1	1	79
Switzerland		47		1	48
Slovenia		44		2	46
Cyprus		15			15
Spain			604	8	612
France			564	7	571
Italy			236	13	249
Czech Republic			146	36	182
Slovak Republic			116		116
Russia			100		100
Croatia			98		98
Estonia			95	6	101
Latvia			88	9	97
Lithuania			61	13	74
Bulgaria			31	135	166
Belgium			31		31
Denmark			26		26
Ireland			22	13	35
Greece			13	2	15
Finland				585	585
Portugal				153	153
Serbia				131	131
United Kingdom				76	76
Poland				26	26
Canaries (Spain)				15	15
Netherlands				11	11
Azores (Portugal)				8	8
Luxembourg				4	4
Total	773	484	2463	2247	5967

Table 1 shows that the number of plots where the above and below ground soil assessments is rather limited to a few countries (Norway, Romania, Germany and Sweden). In a number of countries the location is the same, though assessments might not have been done on the same sample at the same time. Though for most of the ICP Forests plots there is no relation with the NFI inventory or it has not been documented in the ICP Forests database.

Likewise, some NFIs do measure litterfall and mortality rates of the above-ground living biomass, but the rather poorly harmonised nature of these efforts overall complicates the application of NFI plots for Yasso.

2.3 LUCAS data

Although the LUCAS data (Land Use and Coverage Area frame statistical Survey) offer a very important data source on land use changes at the European scale, its potential to verify changes in carbon stocks across forests is rather limited because:

The LUCAS survey, in its current form, it has been run every 3 years from 2009 to 2018, except in 2022 which is 4 years after the last survey. Each survey makes a selection of 41,000 point out of 250,000 sampling points across all land uses across Europe. So comparing of SOC stocks over time at the same plot is for the moment not yet possible.

The LUCAS soil sampling is limited to one sample over a depth of 30 cm (2022 survey) or 20 cm (previous surveys), including the forest floor and mineral soil in one sample. As Yasso considers soil to a depth of 100 cm, LUCAS does not cover information on the subsoil.

In addition, bulk densities were only assessed from the 2018 survey onwards, and this on a subset of 4000 points. This implies that soil organic carbon stock estimates from previous surveys are rather uncertain.

Lastly, LUCAS focuses on the soil compartment only, and thus no corresponding data on litter or deadwood input from the living forest biomass compartment are available on the same plots.

2.4 The ICP Forests Level I database

The ICP Forests Level I database comprises soil data on 7960 plots across Europe. See Table 2.

Being a systematic but less intensive monitoring network, above-ground surveys within Level I do not include litterfall and growth assessments. Meanwhile, in a number of countries, these Level I plots form a subset of the NFI plots, so some above-ground biomass measurement data exist. However, querying the ICP Forests Level I database, revealed that although co-location is valid for the Fenno-Scandinavian countries, this is only rarely the case in the rest of Europe. Furthermore it was impossible within the timeframe of the project, to retrieve and harmonise above-ground flux data from the NFIs on the co-located plots.

Table 2: Total N° of Level I plots contained within the ICP Forests Level I soil database, given by country and survey year

Country	Total	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1998	2004	2006	2007	2008	2009	2014	2017	2020	2021	2022
Austria	142			37	61	32								1		38	43	54						
Belgium	31									10	21				10									
Bulgaria	176						3	7	33	88	45													
Canaries	15											12						3	10					
Croatia	96									78		3									55			
Cyprus	15																15					15		
Czechia	182											99				111	35							
Denmark	26										25							25						
Estonia	101						20	18	14	18	21					26	70					20	29	25
Finland	662		77	85	117	59						104					630							
France	572								1	22	494					175	370							
Germany	664			107		37	81	52	131	7						54	203	171	1					
Greece	15										15													
Hungary	79										67						78							
Ireland	36											22				1	32	2						
Italy	269											21	49			47	192							
Latvia	97							76										95						
Lithuania	74								74								62							
Luxembourg	4										4													
Netherlands	11											11												
Norway	440				95	22	154	106	63															
Poland	646											122				10	335	179						
Portugal	153											149					47	56						
Romania	242									171	70	1												
Russia	39																		39					
Serbia	131														27					130				
Slovakia	116									111						112								
Slovenia	51										7	27				45								
Spain	612									153	100	199						268	330					
Sweden	1983	407	440	324	3							75						793						
Switzerland	47									47														
UK	233									7	59	1						166						
Total	**7960**	**407**	**517**	**553**	**276**	**150**	**258**	**259**	**316**	**712**	**928**	**846**	**49**	**1**	**37**	**619**	**2112**	**1812**	**380**	**130**	**55**	**35**	**29**	**25**

2.5 The ICP Forests Level II database

The plots of the ICP Forests level II monitoring network comprise surveys of litterfall (so-called "Sampling and Analysis of Litterfall"), above-ground biomass accrual and mortality rates ("Growth and Yield") and the soil compartment ("Sampling and Analysis of Soil") up to a depth of 80 cm. For more information, consult the ICP Forests manuals at http://icp-forests.net/page/icp-forests-manual. This database forms the most harmonised and comprehensive European-wide forest monitoring database to date, allowing a comprehensive assessment of the carbon stocks and fluxes in and between the different forest compartments (except for deadwood and living below-ground biomass, which are not routinely assessed in ICP Forest Level II plots).

2.6 Selected dataset

We identified the ICP Forests Level II database as the most suitable for the purpose of Task 3.3 and thus proceeded with the preparation of ICP Forests data for its further analysis in Task 3.3.

In January 2023, the PathFinder project made an official data request to the programme coordinating centre of ICP Forests to obtain access to the ICP Forests database. Once granted, the surveys relevant to the project were downloaded. The data was forwarded to the other tasks within the project that required this information.

Although ICP Forests made many efforts to come to a harmonised pan-European forest monitoring database, still a lot of pre-processing of the data (especially for the soil and growth surveys) was required. Given the expertise on soil, EV INBO focused on the derivation of a quality-controlled soil dataset while WSL worked on deriving / pre-processing of growth and litterfall data on those ICP plots with comprehensive soil information.

Jointly with Task 3.3, a selection of plots with complete information (soil, growth, litterfall) is made for the Yasso application in Deliverable D3.3. Further details about pre-processing of growth data will be provided in Deliverable of Task 3.3. The approach for the calculation of the forest soil organic carbon stocks is described in this deliverable (D.1.3, see below).

3. ICP Forests soil data

Soil surveys have been conducted on both the ICP Forests Level I and Level II monitoring network. Locations for Level II plots were selected by individual ICP Forests partners around the end of the 1980s based on their suitability for intensive forest monitoring, but do not necessarily adhere to a spatially representative statistical sampling grid across Europe. Each network includes soil survey data based on fixed-depth layers and soil characterisation information based on pedogenic horizons. In the database, the survey forms for fixed-depth layers and pedogenic horizons describe the different soil layers that make up a soil profile, encompassing both the forest floor (negative depths) and the mineral or peat soil (positive depths). For each plot, data are reported for one or multiple survey years, with one or multiple replicate profiles/repetitions per survey year. Soil sampling at fixed depths happens in at least 24 sampling points per plot (covering approximately 5000 m^2), to cover the typically large spatial variability of forest soils at a small scale. Per layer, samples from the different sampling points are compiled into one or multiple composite samples (i.e. repetitions). On the other hand, since the depths of the transitions between different pedogenic horizons vary spatially and digging profile pits is laborious, data for pedogenic horizons originate from soil samples taken in a single soil profile pit. As such, each repetition (profile) represents one sampling point in the plot and is not necessarily representative for the whole plot.

Both survey forms contain data on a range of physical (i.e. soil texture, bulk density and the content of coarse fragments) and chemical (amongst others carbon, nitrogen, plant-available and semi-total elements, pH) soil properties. These properties were analysed by national forest soil laboratories according to standardised protocols, for which the lab quality was validated through ring tests. Analytical results are submitted by the partners to the central database managed by the Programme Coordinating Centre (PCC) of ICP Forests.

The objective is to repeat the soil surveys (composite samples at fixed-depth layers only) every ten years. This frequency should suffice to evaluate changes in soil organic carbon stocks, which are typically slow. The combination of (i) the detailed sampling design with a large amount of sampling points per Level II plot compiled into composite samples; and (ii) the temporal consistency of the soil sampling and analysis protocol (with quality control of the lab analyses); is supposed to provide a strong foundation for tracking changes in forest soil organic carbon stocks at a local plot scale. Even though the selected Level II plots may not be spatially representative across Europe, outcomes of the Yasso model (Task 3.3) can be effectively extrapolated to achieve sound conclusions at a regional or at the European scale, by combining them with the systematic Level I network (covering a 16 x 16-km grid across Europe) (De Vos et al., 2015).

4. Pre-processing of soil data

This section describes the process of transforming raw layer- or horizon-specific soil data - as submitted by the partners to the PCC database - into a cleaned, harmonised and gap-filled version which is necessary for further data analyses. The first step of this transformation involved identifying inconsistencies in the data, which required feedback from the data-submitting partners. Such validation checks were conducted using R scripts, and assessed the compliance, conformity and uniformity of the data. For example, the checks evaluated the availability of primary information and mandatory data in the correct format, the complete coverage of each soil profile by different layers without gaps or overlaps, and the plausibility of reported values. All R-scripts are available on GitHub.

In March 2023, EVINBO, that is leading Task 1.3, communicated these inconsistencies to the data-submitting partners. Partners were asked to evaluate each inconsistency, with the objective to either submit the corrected data to the PCC database, to confirm extreme values as correct or to indicate that no (correct) data could be reported for a given inconsistency. While 37 out of 41 Level II partners responded, not all of them provided complete feedback on the inconsistencies. Only a few countries resubmitted corrected data to the PCC database, resulting in a decrease in raw data inconsistencies by 15% for Level II. In fact, since data (re)submissions are relatively time-consuming, some partners opted to directly provide corrected data to EVINBO along with their inconsistency reports. Since raw data can only be modified by the data-submitting partners, certain corrections were implemented solely in the improved version of the data instead of in the raw data.

After this phase with direct partner communication, EVINBO implemented further data corrections into an improved version of the data using their expert knowledge. These data corrections were automated using R scripts, which consistently applied safety checks to ensure that the conditions under which the data needed to be corrected remained valid. Overall, assumptions in the data transformation process align with the data requirements as outlined in the ICP Forests soil manual.

Briefly, this workflow consisted of the following steps:

4.1 Initial data conversions

Columns in the data are converted to the best data format, and some additional columns with layer information (“layer_type”) and with unique identifiers (e.g. “plot_id”) are added. Plot codes are harmonised so that each unique location corresponds to one plot (meanwhile maintaining the links with other ICP Forests surveys). Harmonised coordinates per plot are generated.

4.2 Solve record inconsistencies

Redundant records are removed or merged with other records, so that each sample corresponds to one record. This action was necessary for: (i) old surveys, since fixed-depth layer data were previously split into a mandatory as well as an optional dataset in old versions of the database; (ii) some profiles for which data have been resubmitted under a different survey year in the PCC database (because of which the original data were not overwritten); (iii) some forest floor layers with no information reported; and (iv) records removed at the request of the responsible soil experts.

4.3 Apply gap-filling and corrections reported in partner inconsistency reports

Data for which corrected values are reported by partners along with their feedback on the inconsistencies in their data, and which were not updated by the partners in the PCC database since March 2023 (when inconsistency reports were shared with all the partners), are corrected as indicated by the partners.

4.4 Gap-fill based on other external data sources

Not all soil data of the ICP Forests Level II plots ended up in the current PCC database. In this step, existing records are gap-filled based on these historical versions of the database, and missing records are added accordingly.

4.5 Solve inconsistencies in primary keys

To ensure that each layer within a soil profile has a unique layer code, some codes for soil layers or master soil horizons are updated or generated (if missing). Some plots contain records that clearly represent the depth range of one single profile, yet different adjacent survey years or repetition codes are reported across the different layers of the profile. In those cases, survey years or codes for the repetition or profile pit are corrected accordingly.

4.6 Solve inconsistencies in data ranges

Parameter data which are obviously reported in incorrect units (e.g. bulk density in g cm^{-3} instead of kg m^{-3}, or total organic carbon in % instead of g kg^{-1}) are identified and corrected. Parameter values outside the possible range for a given parameter (e.g. negative values, bulk density values above the mineral density of 2650 kg m^{-3}) are removed. Unknown codes in parameters with categorical classes (e.g. texture classes) are also removed. Obvious errors in total organic carbon data (e.g. values swapped between layers, based on comparison with other profiles of the same plot, or forest floor organic carbon data reported to be below the limit of quantification, which is highly implausible) are manually detected and resolved or removed. Although some total organic carbon data, particularly from surveys in the ‘80s and ‘90s, appeared less reliable in absolute terms or by comparison with other profiles or survey years from the same plot, there was usually insufficient evidence to correct them. As a result, only a limited number of total organic carbon values is corrected. Lastly, limits of quantification are assigned to each of the analytical parameters for each record, and values falling below these limits are replaced by 0.5 times the limit of quantification.

4.7 Harmonise layers within a profile

All actions in this step have been manually validated. Firstly, layer depth limits are corrected or completed (based on the layer code) as much as possible. The layer type (forest floor, peat or mineral) is updated according to the total organic carbon content and WRB soil classification, provided that there is sufficient evidence. According to the manual, organic layers on top of the mineral soil are supposed to be designated as forest floor if their thickness is less than 40 cm, and as peat if their thickness exceeds 40 cm, regardless of their composition, and with the 0-cm depth at the transition from the forest floor to the mineral or peat soil. Consequently, layer limits for some profiles were adjusted upward or downward. This adjustment is only made if the location of the resulting 0-cm depth is consistent across the different profiles and survey years of a certain plot. Finally, a sequential layer number is assigned to each layer of a profile, starting with 1 for the uppermost layer and incrementing for each successive layer below it, based on their layer limits or layer code (e.g. for the forest floor). Redundant layers within the soil profile (i.e. layers which fully overlap with other more complete or detailed layers within the same profile, or which represent parent material at the bottom of the profile) are excluded from the layer numbering (i.e. “NA” is assigned).

4.8 Add derived attributes

Additional attributes (such as layer thickness, C-to-N ratio, the sum of the base cations and acid cations) are calculated if possible. Coarse fragments reported as a mass percentage are converted to a volumetric percentage.

4.9 Apply additional manual corrections

Any additional individual corrections suggested by members of the Expert Panel are implemented at this stage. The number of corrections is limited in this step, because most of the suggestions raised by the soil experts concerned structural issues that were implemented in other steps of the data transformation.

4.10 Gap-fill internally (based on directly reported data)

Data gaps for physical parameters (i.e. bulk density, including the density of organic layers; the content of coarse fragments; and clay, silt and sand content) are filled under the assumption that these parameters remain constant at a certain depth over time. To achieve this, the best data for each plot are selected at each depth interval, based on fixed-depth layers, pedogenic horizons (including bulk densities and coarse fragments estimated by data-submitting partners), as well as bulk density data reported in the soil water content survey (“sw_swc”) part of the ICP Forests survey on Meteorology (see ICP Forests Manual Part IX). Hereby, bulk density data which are considered implausible based on pedotransfer predictions using the total organic carbon content, supplemented with a plausibility range of ± 160 kg m^{-3}, are excluded. Data are matched at a depth basis (i.e. not through their layer code), and data from other than the supposedly best survey years are also updated. Overall, the selection of the most plausible physical data is guided by expert knowledge (e.g. prioritising data originating from the same survey form, and measured data over data estimated by partners). The uncertainty associated with this selection is propagated in the uncertainty intervals of the resulting data forms. Similarly, data gaps in the total organic carbon content of the forest floor are filled using data from other repetitions, survey years or survey forms, as long as the survey year of the data source is within three years of the target survey year. Finally, uncertainty ranges for total organic carbon are generated, using the results of the study by Hiederer et al. (2011) in which roughly one tenth of the forest floor and topsoil samples from the ‘90s and ‘00s was reanalysed by a central soil analytical laboratory. Even though the reported uncertainty ranges are limited to lab analytical uncertainty (within and between labs), uncertainty in soil sampling at the plot scale is assumed to be minimized by including a large number sampling points per plot (horizontally) and by adhering to the correct fixed-depth limits (vertically). In addition, the former uncertainty component, along with sample preparation and within-lab analytical uncertainty, is supposed to be somehow reflected in the variation between different repetitions, if available. For values below the limit of quantification, uncertainty ranges minimally cover the range from 0 to the limit of quantification.

4.11 Gap-fill with pedotransfer predictions

Gaps in bulk densities for mineral layers are filled using predictions from a new robust pedotransfer function, in which total organic carbon serves as the sole predictor. This function was established using the Level I bulk density data and validated using Level II data, resulting in an uncertainty range of ± 160 kg m^{-3}. Gaps in densities of organic layers are filled using average densities for the respective layer code or type (“OF”, “OH”, “O” excluding “OL”, versus peat layers), reported across both Level I and Level II, along with their associated uncertainties.

4.12 Tidy up data

Finally, the data forms are cleaned up by removing redundant columns and rearranging the rows.

Overall, these efforts resulted in a clearly improved data quality, in spite of the fact that several gaps and inconsistencies could not be resolved. This is typically a problem for data for which the original data-providing partners could no longer be contacted, or when the analyses were conducted so long ago that their accuracy could no longer be verified.

5. Calculation of forest soil organic carbon stocks and fluxes

Building upon this enhanced version of the data, soil organic carbon (SOC) stocks were calculated for each profile. This calculation was performed separately for the forest floor versus below-ground soil section (mineral and peat soil). The OC stock of the forest floor was computed as the sum of the OC stocks of its organic sublayers [t OC ha^{-1} $layer^{-1}$], assessed as follows:

$$c\_layer\_stock = organic\_layer\_weight\ [kg\ soil\ m^{-2}\ layer^{-1}] \times organic\_carbon\_total\ [g\ kg^{-1}] \times 1E\text{-}2 \quad [1]$$

For the mineral soil profile, the objective was to calculate OC stocks until the effective soil depth, set to a maximum of 100 cm. To achieve this, the C density [t C ha^{-1} cm $depth^{-1}$] was calculated for each layer:

$$c\_density = bulk\_density\ [kg\ m^{-3}] \times (1 - coarse\_fragment\_vol\ [v/v]) \times organic\_carbon\_total\ [g\ kg^{-1}] \times 1E\text{-}4 \quad [2]$$

As described in De Vos et al. (2015), we used the concept of C density, a volume-based measure that enables (i) easy comparison among layers and (ii) easy computing of OC stocks by integrating over any depth interval of interest. For peat layers, a combination of both approaches (eq. [1] or [2]) was used, depending on availability of the data. Since peat layer thicknesses are typically known, C densities were also calculated using the forest floor layer stock approach (eq. [1]) was used, by further dividing the layer OC stock by its thickness. For both organic and mineral layers, we additionally estimated uncertainty intervals by propagating uncertainties from the source data.

From this stage onward, we proceeded with calculating the OC stock for a profile only if the C density is known for at least two layers (including the uppermost below-ground layer), or at least the uppermost layer if a soil profile is known to be very shallow.

We further filled any remaining gaps in C densities in the depth range up to the effective soil depth, using the following assumptions:

Below 40 cm, the C density of a plot remains constant.

Below 80 cm, the C density is usually low and similar across soil types, and estimated to be 0.31 t C ha^{-1} cm $depth^{-1}$ based on available data in the soil database.

These assumptions are consistent with the ICP Forests manual on Sampling and Analysis of soil (Cools and De Vos, 2020, http://icp-forests.net/page/icp-forests-manual), and the associated uncertainty was also considered. However, since these gap-filling steps sometimes did not suffice to cover the entire target depth range, we filled remaining carbon density gaps by systematically applying: (i) a mass-preserving quadratic smoothing spline [using the R package *mpspline2* by O'Brien (2022), with lambda set to 0.1] to the below-ground C density data of each soil profile; and subsequently, (ii) a natural spline (using the R package *stats*) to the output of the latter, in order to extrapolate the C density data beyond their original depth range.

Next, the total OC stock [t C ha^{-1}] of a profile was calculated by integrating the final below-ground C densities and the forest floor layer OC stocks. In addition, topsoil OC stocks [t C ha^{-1}] were calculated in the same manner, but with below-ground C densities limited to a maximum depth of 30 cm. Stocks for which the original observations were all shallower than 30 cm, while the effective soil depth exceeded that, were flagged with the recommendation to use topsoil OC stocks rather than the regular OC stocks. Finally, average (topsoil) OC stocks and their uncertainties were calculated per plot and survey year. Absolute (t C ha^{-1} $year^{-1}$) and relative (% ha^{-1} $year^{-1}$) annual increases in OC stock were derived and considered as a proxy for the net annual SOC sequestration rate.

6. Results and discussion

Forest soil organic carbon stocks of the soil and forest floor could be calculated for at least one soil survey for 612 ICP Forests Level II plots (Fig. 1 and Table 3). Annual (net) sequestration rates of forest soil organic carbon stocks in the soil and forest floor could be calculated for 244 Level II plots (Table 4), of which 241 plots do not contain peat. The total SOC stock includes forest floor and the mineral and peat soil up to a depth of 100 cm (or the effective soil depth if shallower).

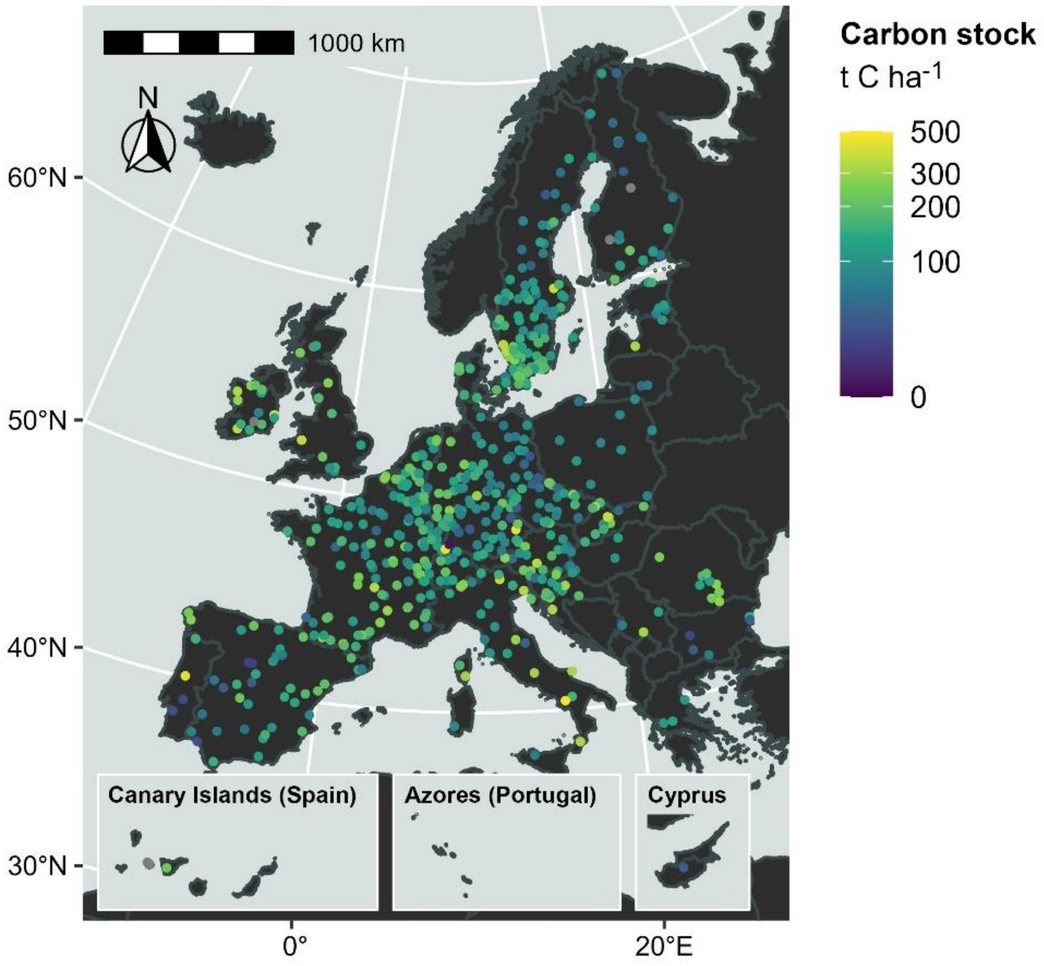

Figure 1: Map of the total soil organic carbon stock (forest floor and soil; mean over survey period 1990 - 2023) in the ICP Forests Level II network on 612 plots up to a depth of 100 cm or until the effective soil depth if shallower.

Table 3: Soil organic carbon stock in t C ha^{-1} (mean over survey period 1990 – 2023). Confidence intervals (CI) are bootstrapped, bias-corrected and accelerated (i.e. adjusted for both bias and skewness in the bootstrapped samples).

Compartment	n	Mean	95 % CI mean	Median	95 % CI median
Forest floor SOC stock	625	20.3	[18.0; 24.2]	11.5	[10.4; 12.8]
Below-ground SOC stock (mineral soil)	766	117.3	[112.5; 122.7]	100.6	[96.2; 106.3]
Below-ground SOC stock (peat soil)	17	334.7	[226.6; 466.1]	225.9	[106.1; 524.9]
Total SOC stock (forest floor + soil)	612	152.5	[145.5; 161.2]	131.4	[125.1; 136.2]

Overall, the OC stocks of the forest floors were estimated at a median value of 11.5 t C ha^{-1} and a mean of 20.3 t C ha^{-1} which are close to the estimates (mean of 22.1 t C ha^{-1}) previously obtained at the ICP Forests Level I network in the 2006-2008 survey (De Vos et al. 2015).

For the below-ground stocks, we separated between the mineral soil profiles and the peat soils. So the mineral soil profiles store on average 117.3 t C ha^{-1} with a median of 100.6 t C ha^{-1}. These values are consistent with the previous estimates at the Level I network by De Vos et al. (2015), who observed an average stock of 108 t C ha^{-1} in mineral soil at the same depth range. Based on 17 peatland plots, we obtained a mean SOC of 334.7 t C ha^{-1} and a median value of 225.9 t C ha^{-1}. This is rather a small number to draw any conclusions at the European level. On the other hand, as the Yasso model does not include peatland SOC fluxes, this estimate will not affect the further modelling work in Task 3.3.

We observed quite large differences between the median and mean values. This is due to the skewed distribution due to several outliers on the higher end of the range. The FSCC is still in exchange with the individual data providers to find any plausible explanations of the sometimes, very high SOC stocks values.

On a total of 267 plots, the forest floor was sampled twice in the considered time interval. In the forest floor we observe an annual statistically significant mean decrease of -0.51 t C ha^{-1} year^{-1} and a median decrease of -0.13 t C ha^{-1} year^{-1} (Table 4).

Table 4: Annual net sequestration rate of soil organic carbon stock in t C ha^{-1} year^{-1}. Confidence intervals (CI) are bootstrapped.

Compartment	n	Mean	95 % CI mean	Median	95 % CI median
Forest floor stock	267	-0.51	[-0.96; -0.25]	-0.13	[-0.16; -0.06]
Below-ground SOC stock (mineral soil)	264	0.39	[0.00; 0.76]	0.43	[0.25; 0.60]
Below-ground SOC stock (peat soil)	4	-3.20	[-4.74; -2.26]	-2.91	[-5.14: -2.06]
Total SOC stock (forest floor + soil)	244	-0.18	[-0.79; 0.32]	0.24	[0.08; 0.45]

When only considering the mineral soil part (up to 1 m or shallower) we obtain a annual statistically significant increase of the mean (0.39 t C ha^{-1} year^{-1}) and median (0.43 t C ha^{-1} year^{-1}) carbon sequestration rates, based on 264 plots. Concerning changes in the peat soils, our results remain indecisive as the sample size is rather small (4 plots only).

Based on the data coming from the ICP Forests Level II dataset, the overall mean carbon sequestration rates, including the full profile up to a maximum depth of 1, does not indicate any statistically significant changes of the carbon stocks over time. Though the median carbon sequestration rate does show an annual median increase of the SOC stock with 0.24 t C ha^{-1} year^{-1}.

These calculated carbon sequestration rates, based on field data, will be compared once the modelling outputs by the Yasso model have been generated. The results of this comparison will be reported in a later stage of the project, after the work within Task 3.3 is finalised. While the Level II network is not necessarily representative at a larger scale, they can be combined with results from the systematic Level I grid (De Vos et al., 2015) to give more robust insights about regional or pan-European patterns in carbon stock changes, which is however not planned within the frame of the PathFinder project.

V. Conclusions

The obtained SOC stocks, in combination with the derived and pre-processed data on litterfall and growth, will be further used by Task 3.3 in the adaption of the Yasso model. For the description of the pre-processing of the litterfall and growth data, we refer to Deliverable 3.3.

VI. References

Cools N, De Vos B, 2020: Part X: Sampling and Analysis of Soil. Version 2020-1. In: UNECE ICP Forests Programme Co-ordinating Centre (ed.): Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. Thünen Institute of Forest Ecosystems, Eberswalde, Germany, 29 p + Annex. http://www.icp-forests.org/manual.htm ISBN: 978-3-86576-162-0

De Vos, B. Cools, N. Ilvesniemi, H., Vesterdal, L., Vanguelova, E. , Carnicelli, S. 2015. Benchmark values for forest soil carbon stocks in Europe: Results from a large scale forest soil survey. Geoderma 251–252: 33-46. https://doi.org/10.1016/j.geoderma.2015.03.008

Hiederer, R. Micheli, E. and Durrant, T. 2011. Evaluation of BioSoil Demonstration Project. Soil Data Analysis. EUR 24729 EN. Doi:10.2788/56105. Publications Office of the European Union.

IUSS Working Group WRB. 2014. World Reference Base for Soil Resources 2014. Update 2015. International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106. FAO, Rome.

O'Brien, L., Brendan, B., Hengl, T., Bishop, T., Rossiter, D., Baudette, D. Brown, A. 2022. Package ‘mpspline2’. 13-oktober 2022.

Appendix: Description of attributes in the final plot-specific forest soil organic carbon stock dataset

- **altitude**:
  Altitude of the plot in meter above sea level

- **biogeographical_region**:
  Biogeographical region (EEA)

- **bs_class**:
  Qualitative class of base saturation (which is the percentage of the cation exchange capacity occupied by base cations)

- **c_stock**:
  (Soil) organic carbon stock (in t C ha^{-1}) of the complete forest soil (forest floor and below-ground; until "soil_depth"), obtained by summing up any "c_stock_below_ground" and "c_stock_forest_floor" values - either for a specific profile or the average across profiles within a plot survey

- **c_stock_below_ground**:
  Below-ground soil organic carbon stock (in t C ha-1), obtained through depth integration (i.e. summing up) of carbon densities (as returned from mass-preserving splines for a depth of i cm in t C ha^{-1} cm^{-1}) from 0 cm to "soil_depth" - either for a specific profile or the average across profiles within a plot survey

- **c_stock_below_ground_max**:
  Maximum of uncertainty range for "c_stock_below_ground", obtained by integrating all maxima of the uncertainty ranges for different (spline-estimated) carbon densities/layer stocks.

- **c_stock_below_ground_min**:
  Minimum of uncertainty range for "c_stock_below_ground", obtained by integrating all minima of the uncertainty ranges for different (spline-estimated) carbon densities/layer stocks.

- **c_stock_below_ground_topsoil**:
  Below-ground soil organic carbon stock until a depth of 30 cm (in t C ha^{-1}), obtained through depth integration (i.e. summing up) of carbon densities (as returned from mass-preserving splines for a depth of i cm in t C ha^{-1} cm^{-1}) from 0 to 30 cm or to "soil_depth" if the latter is shallower - either for a specific profile or the average across profiles within a plot survey

- **c_stock_below_ground_topsoil_max**:
  Maximum of uncertainty range for "c_stock_below_ground_topsoil", obtained by integrating all maxima of the uncertainty ranges for different (spline-estimated) carbon densities/layer stocks

- **c_stock_below_ground_topsoil_min**:
  Minimum of uncertainty range for "c_stock_below_ground_topsoil", obtained by integrating all minima of the uncertainty ranges for different (spline-estimated) carbon densities/layer stocks

- **c_stock_forest_floor**:
  Forest floor organic carbon stock (in t C ha^{-1}), obtained through depth integration (i.e. summing up) of carbon stocks in all individual above-ground forest floor layers - either for a specific profile or the average across profiles within a plot survey

- **c_stock_forest_floor_max**:
  Maximum of uncertainty range for "c_stock_forest_floor", obtained by integrating all maxima of the uncertainty ranges for different layer stocks

- **c_stock_forest_floor_min**:
  Minimum of uncertainty range for "c_stock_forest_floor", obtained by integrating all minima of the uncertainty ranges for different layer stocks

- **c_stock_max**:
  Maximum of uncertainty range for "c_stock", obtained by integrating all maxima of the uncertainty ranges for different (spline-estimated) carbon densities/layer stocks

- **c_stock_min**:
  Minimum of uncertainty range for "c_stock", obtained by integrating all minima of the uncertainty ranges for different (spline-estimated) carbon densities/layer stocks

- **c_stock_topsoil**:
  (Soil) organic carbon stock of the forest soil (forest floor and below-ground; until 30 cm or until "soil_depth" if the latter is shallower; in t C ha^{-1}), obtained by summing up any "c_stock_below_ground_topsoil" and "c_stock_forest_floor" values - either for a specific profile or the average across profiles within a plot survey

- **c_stock_topsoil_max**:
  Maximum of uncertainty range for "c_stock_topsoil", obtained by integrating all maxima of the uncertainty ranges for different (spline-estimated) carbon densities/layer stocks

- **c_stock_topsoil_min**:
  Minimum of uncertainty range for "c_stock_topsoil", obtained by integrating all minima of the uncertainty ranges for different (spline-estimated) carbon densities/layer stocks

- **code_country**:
  Unique ID of country in ICP Forests

- **code_plot**:
  Observation plot number

- **contains_peat**:
  Logical indicating whether the below-ground part of the profile contains any peat layers

- **country**:
  Name of country

- **eff_soil_depth**:
  Effective depth (in cm) of the soil profile until the continuous rock

- **eftc**:
  European Forest Type Category (EEA) of forest stand

- **forest_floor_layers_unique**:
  Unique value of "forest_floor_layers" if values of "forest_floor_layers" across different profiles within a plot survey are the same, else NA

- **forest_floor_thickness_avg**:
  Average of "forest_floor_thickness" values across profiles within a plot survey

- **humus_form**:
  Humus form

- **latitude_dec**:
  Latitude in decimal degrees (WGS84)

- **longitude_dec**:
  Longitude in decimal degrees (WGS84)

- **main_tree_species**:
  Main tree species in forest stand

- **map**:
  Mean annual precipitation (in mm) (source: WorldClim)

- **mat**:
  Mean annual temperature (in °C) (source: WorldClim)

- **nlay_below_ground_max**:
  Maximum of "nlay_below_ground" values across profiles within a plot survey

- **nlay_below_ground_min**:
  Minimum of "nlay_below_ground" values across profiles within a plot survey

- **nlay_forest_floor_max**:
  Maximum of "nlay_forest_floor" values across profiles within a plot survey

- **nlay_forest_floor_min**:
  Minimum of "nlay_forest_floor" values across profiles within a plot survey

- **nlay_max**:
  Maximum of "nlay" values across profiles within a plot survey

- **nlay_min**:
  Minimum of "nlay" values across profiles within a plot survey

- **obs_depth_avg**:
  Average of "obs_depth" values across profiles within a plot survey

- **parent_material**:
  Parent material

- **partner_code**:
  Unique ID of data-submitting partner in ICP Forests

- **partner_short**:
  Short name of data-submitting partner

- **plot_id**:
  Combination of "code_country" and "code_plot"

- **plot_id_form**:
  Combination of "survey_form" and "plot_id"

- **rmse_mpspline_max**:
  Maximum of "rmse_mpspline" values across profiles within a plot survey

- **rooting_depth**:
  Depth (in cm) until which roots were observed during the soil profile description. Note: this parameter is no longer mandatory.

- **slope**:
  Slope of the plot in decimal degrees

- **soil_depth_avg**:
  Average of "soil_depth" values across profiles within a plot survey

- **survey_form**:
  Name of the ICP Forests soil data form that serves as the source for the stock calculation data - either "s1_som", "s1_pfh", "so_som" or "so_pfh"

- **survey_year**:
  Year of the soil sampling

- **unknown_forest_floor**:
  Logical indicating whether stocks for the forest floor are known to be missing (i.e. there are any forest floor records in the database for which no forest floor layer stock could be calculated, e.g. due to missing organic_layer_weight and layer thickness information). If NA, (none of) the profile(s) contains any forest floor records.

- **use_stock_topsoil**:
  Logical indicating whether it is recommended to use the "c_stock_topsoil"/ "c_stock_below_ground_topsoil" data (up to 30 cm) rather than the "c_stock"/ "c_stock_below_ground" data (up to "soil_depth"). This recommendation arises from the observation depth being shallower than 30 cm, because of which extrapolation of the mass-preserving spline to the "soil_depth" introduces a considerable uncertainty for deeper layers.

- **wrb_qualifier_1**:
  First WRB qualifier (IUSS Working Group on WRB 2014/2015 version)

- **wrb_ref_soil_group**:
  WRB Reference Soil Group (IUSS Working Group on WRB 2014/2015 version)