HoliSoils

Working together for forest soils

D2.1 Model ensemble script

Holistic management practices, modelling and monitoring for European forest soils, HoliSoils

Project duration: 1.5.2021–31.10.2025

H2020 Grant Agreement No 101000289

Coordinator: Natural Resources Institute Finland (Luke)

Deliverable D2.1: Model ensemble script		
D2.1 describes a R script that launches a multi-model ensemble to simulate the time evolution of soil organic carbon (SOC) stock and greenhouse gas (GHG) fluxes. A test site is used with: preliminary flux data from the Spanish Gámiz experiment of WP5, soil data from the LUCAS database, ISIMIP weather forcing, and other data derived from the literature.		
Due date	30.9.2022	
Authors	Elisa Bruni, Bertrand Guenet et al.	
Date of publication	6.10.2022	
Dissemination level		
PU	Public, fully open e.g. web	X
CO	Confidential, only for members of the consortium (including the Commission Services)	
CI	Classified, information as referred to in Commission Decision 2001/844/EC	
Nature of the Deliverable		
R	Document, report	
DEM	Demonstration, pilot, prototype, plan design	
DEC	Websites, patents filing, market studies, press & media actions, videos etc.	
OTHER	Software, technical diagram etc.	X
Ethics	Ethics deliverables	

This project has received funding from the European Union's H2020 research and innovation programme under grant agreement No 101000289.

Contents

1. Introduction

The HoliSoils H2020 project aims to advance modelling tools for monitoring soil properties and GHG fluxes. One of the objectives of the WP2 is to use several state-of-the-art models to simulate forest soil responses to management practices, natural disturbances and overall impacts of climate change. In particular, within the WP2 a multi-model ensemble is developed to simulate the evolution in time of the SOC stocks and GHG fluxes, and to take into account the uncertainty of the simulations linked to the different model structures and calibration data sets. This document describes the models included in the multi-model ensemble, the R script that launches the first version (v0) of the ensemble, and the data used for the test run. This v0 of the ensemble is designed as a test tool that will be used by SOC modelers in order to calibrate model parameters with experimental data, and that will be further developed as a web app for end users that wish to study the effect of forest management and climate change scenarios on SOC stocks and GHG fluxes at their sites.

2. Models in the ensemble

The models that were included in the first version of the multi-model ensemble are:

- AMG (Andriulo et al., 1999)
- Century (Parton et al., 1988)
- ICBM (Andrén and Kätterer, 1997)
- Roth-C (Coleman and Jenkinson, 1996)
- SG (Hashimoto et al., 2011)
- Yasso07 (Tuomi et al., 2009)

2.1 Brief description of the models

Apart from SG, all models represent SOC with a conventional multi-compartmental structure (see Figure 1) that can be summarized with the following equation:

$$\frac{d\boldsymbol{C}}{dt} = \boldsymbol{I}(t) - \mathbf{A} \cdot \mathbf{K} \cdot \xi(t) \cdot \boldsymbol{C}(t)$$

where $\boldsymbol{C}(t)$ is a vector describing the masses of C of the *n* compartments as a function of time (t); $\boldsymbol{I}(t)$ is the vector of the C input to the soil; $\mathbf{A}$ is a matrix describing the mass flow within each pool; $\mathbf{K}$ is a diagonal matrix containing the decomposition coefficients of the compartments, and $\xi(t)$ is the scalar effect of the pedo-climatic conditions on the decomposition of C.

The models differ from each other by: the number of pools in which the soil is conceptually divided; the transfers within these pools; and the functions that describe the environmental and/or edaphic effects on the decomposition rates of the soil compartments [$\xi(t)$] (Figure 1). Each model is briefly described below.

- AMG has one fresh organic matter pool and two SOC pools: an active and a stable pool. The stable pool is calculated as a fraction of the total observed initial SOC, and is considered constant throughout the simulation length. The fraction of stable C depends on the historic land-use of the site (Clivot et al., 2019). The active pool receives the C input, which is first partly mineralized, with a rate that depends on the type of fresh organic material.
- Century is a seven-pool model, with four litter C pools (aboveground and belowground structural, and aboveground and belowground metabolic C pools), and three SOC pools (active, slow and passive). In both models, in absence of data on the aboveground:belowground ratio, the C that enters the soil is separated into its aboveground and belowground part according to the mean annual precipitation, with higher precipitations driving a higher aboveground biomass proportion, and with the proportion of both the aboveground and belowground biomass approximately ranging between 0.48 and 0.52 (see Figure S 1). The estimated proportions were kept constant throughout the simulation length. The metabolic and structural parts are calculated as a function of the lignin to nitrogen ratio of the litter input. ICBM has two pools: a young C pool that directly receives the C input, and an old C pool that receives part of the C output from the young pool.

- Roth-C has four active pools (decomposable and resistant plant material, microbial biomass and humified organic matter), and one stable pool (inert organic C). The fraction of inert C is calculated from the level of initial SOC using the Falloon equation (Falloon et al., 1998).
- Yasso07 is a five-pool model, where C is divided into different compound groups (acid hydrolysable, water soluble, ethanol soluble, neither soluble nor hydrolysable compounds, and humus). Except for the humus pool, all compartments receive part of the C input.
- SG is a group of empirical models that predict the CO_2, CH_4 and N_2O effluxes in forest soils, using soil physiochemical properties, water filled pore space and soil temperature data.

For details on the environmental functions and on the model equations, refer to the model references.

This v0 of the multi-model ensemble is designed to be used by SOC modelers for calibration of model parameters with site-scale experimental data, and for estimation of uncertainties linked to the parametrization.

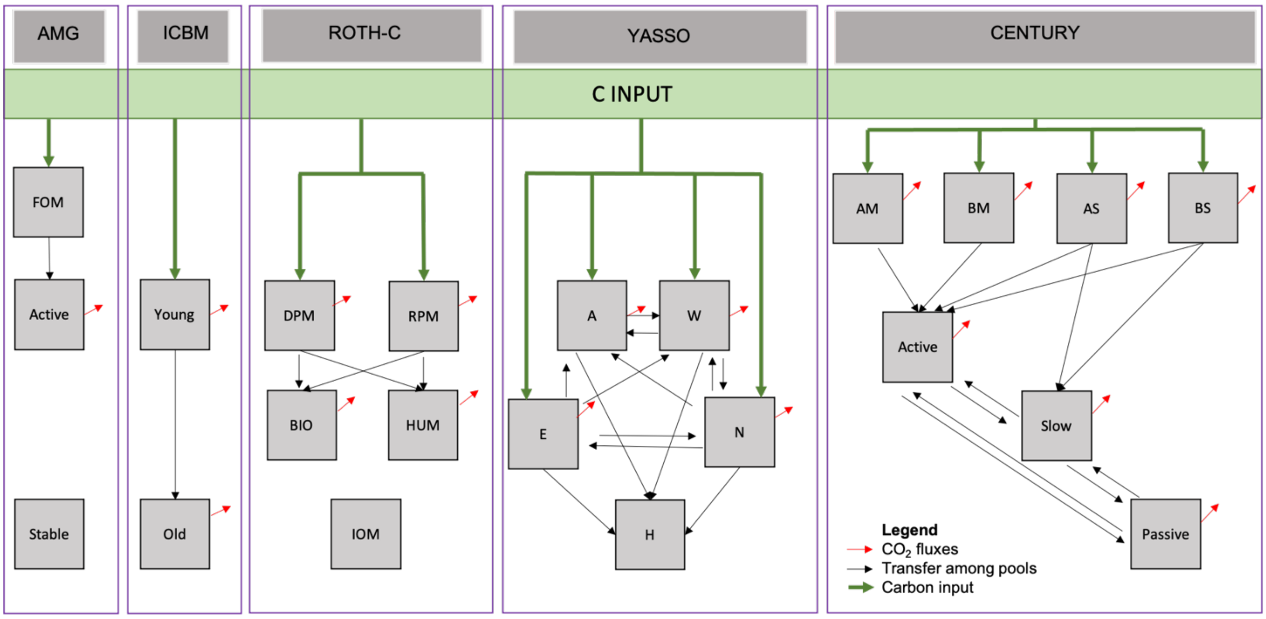

Figure 1 Schematization of the SOC models. Pool names: FOM = fresh organic matter, DPM = decomposable plant material, RPM = resistant plant material, BIO = microbial biomass, HUM = humified organic matter, IOM = inert organic matter, A = acid hydrolysable, W= water soluble, E= ethanol soluble, N = neither soluble nor hydrolysable compounds, H = humus, AM = aboveground metabolic litter, BM = belowground metabolic litter, AS = aboveground structural litter, BS = belowground structural litter.

3. Scripts

The scripts provided to run the first version of the multi-model ensemble are written in R (R Core Team, 2022) and can be found at the following link: https://github.com/elisabruni/Holisoils-multimodel/tree/main/R. The tool relies on the **SoilR package (version 1.2.105)** developed by Sierra et al. (2012) and requires a **R software version >= 3.5.0**. Note that without the required versions of SoilR and R, the scripts will not run properly.

For information on how to install or update R consult the following link: https://rstudio-education.github.io/hopr/starting.html. The SoilR documentation can be found here: https://cran.r-project.org/web/packages/SoilR/SoilR.pdf.

The following scripts are provided and will be detailed in the next subchapter:

- Run_Holisoils_multimodel_v0.R
- Holisoils_multimodel_v0.R
- Holisoils_AMG_v0.R
- AMG_environmental_functions.R
- Holisoils_Century_v0.R
- Holisoils_ICBM_v0.R
- Holisoils_RothC_v0.R
- Holisoils_SG_v0.R
- Holisoils_Yasso07_v0.R
- Forcing_data_test_run_v0.R

To solve equations of the SOC models (i.e., AMG, Century, ICBM, Roth-C, and Yasso07) the initial SOC stock level in each compartment needs to be provided. In AMG, the partitioning of total initial SOC into the active and stable pools is defined by default (i.e., 65% of C is stable for conditions of land use with long-term arable history and 40% for conditions of long-term grassland history). For long-term forest soils, we hypothesised similar conditions as for historical grassland soils. For the other models, the initial partitioning of C in the different pools is not defined by default. Hence, it needs to be estimated. To do that, we run the models with average climate conditions for 1000 years, until all the SOC pools reach a steady-state. The steady-state is considered attained if the annual variation of SOC in all pools is lower than 0.1% for at least 10 years. Then, we rescale the total simulated SOC using the value of measured SOC provided by the user, and we keep the same estimated proportion of C in each pool as at steady-state (Dimassi et al., 2018). Finally, we solve the matrix differential equation for the defined simulation length (i.e., 10 years for our test run). The user can specify a different simulation length.

3.1 Description of the scripts

The following table describes the scripts that are needed to run the simulations and the actions that are required from the user. Note that these scripts require several R packages that need to be installed prior to utilization. To install packages in R, the following command can be used:

install.packages("PACKAGE_REQUIRED"),

where PACKAGE_REQUIRED is the name of the required package. An example is provided in the Run_Holisoils_multimodel_v0.R script.

Script name	Function	Required actions from the user
Run_Holisoils_multimodel_v0.R	Loads the required packages, calls all the other scripts and	<u>Define:</u> **PATH_functions** Path to the folder where the R scripts are stored

	launches a test simulation	
Holisoils_multimodel_v0.R	Calls all the models and plots the evolution of the SOC stocks and CO_2 fluxes	None
Holisoils_AMG_v0.R	Calls the AMG model and plots the SOC stock evolution for each pool	None
AMG_environmental_functions.R	Defines the environmental functions of the AMG model	None
Holisoils_Century_v0.R	Calls the Century model and plots the SOC stock evolution for each pool, for the spin-up and forward runs	None
Holisoils_ICBM_v0.R	Calls the ICBM model and plots the SOC stock evolution for each pool, for the spin-up and forward runs	None
Holisoils_RothC_v0.R	Calls the Roth-C model and plots the SOC stock evolution for each pool, for the spin-up and forward runs	None
Holisoils_SG_v0.R	Calls the SG models and plots the CH_4 and N_2O fluxes	None
Holisoils_Yasso07_v0.R	Calls the Yasso07 model and plots the SOC stock evolution for each pool, for the spin-up and forward runs	None
Yasso07Model_fixed.R	Matrix equations for the Yasso07 model	None
Forcing_data_test_run_v0.R	Defines all the data that should be provided by the end user	Define: #Define input directory **loc_forc** Path to the folder where the climate data is stored **plot_figures** (TRUE/FALSE) Decide whether to plot figures or not **computation_time_step_fwd** (scalar) Set the time step at which the solution of the forward run is sought [in years] (e.g., 1/12 if the solution is sought once per month)

		#----Soil data------ **clay_site** [%] = percentage of clay in the soil **silt_site** [%] = percentage of silt in the soil **sand_site** [%] = percentage of sand in the soil **ph_site** [unitless] = pH in the soil **CaCO3_site** [g kg^{-1}] = calcium carbonate in the soil **SOC_initial** [MgC ha^{-1}] = initial SOC stock **soil_OM_thick** [cm] = thickness of the soil organic matter layer #----C input data------ **Cinput_spinup** [MgC ha^{-1} yr^{-1}] = average C input to the soil during spin-up **Cinput_fwd** [MgC ha^{-1} yr^{-1}] = C input to the soil during forward run **lignin_to_nitrogen_ratio** [unitless] = lignin to nitrogen ratio of the litter input **structural_in_lignin_ratio** [unitless] = proportion of lignin that is structural **woodylittersize_scalar** [cm] = size of the woody litter input #----Historical land-use of the site------ **historical_LU** ['grassland'or 'arable' or 'forest'] = long-term land-use history #----Other data that should be @5cm depth------ **CN_site** [unitless] = C to nitrogen ratio **BD_site** [Mg m^{-3}] = bulk density #----Optional------ **water_filled_pore_space** [unitless] = water filled pore space (*optional*) **volumetric_soil_water_cont** [%] = volumetric soil water content (*needed if* water_filled_pore_space *not provided*) **CH4_conc_site** [ppb] = atmospheric CH_4 concentration (*optional*) #----Starting date of simulations------ **start_date_simulations_site** (date object YYYY-MM-DD) = starting date of the simulations #----Decomposition rates and model parameters------ **ksRothC** [yr^{-1}] = decomposition rate parameters of the Roth-C model **param_ICBM** = decomposition rate parameters [yr^{-1}] and humification coefficient of the ICBM model

		ksCent [month^{-1}] = decomposition rate parameters of the Century model **paramYasso07** = decomposition rates [yr^{-1}] and parameters of the Yasso07 model **ksAMG** = decomposition rate [yr^{-1}] and humification coefficient parameters of the AMG model #----Climate forcing------ To change the climate forcing, the user should provide ".txt" files with the climate data: • Soil temperature [°C] • precipitation [mm] • potential evapotranspiration [mm/month) The minimum resolution is months. The script should be adjusted depending on how and which data are provided. Ultimately, the following data needs to be provided to the multi-model function: **Temp_month** (monthly average soil temperature, °C) a dataframe with two columns: first column contains the dates (named "Date") and second column the values of temperature (named "Temp") **Precip_month** (monthly average precipitation, mm) a dataframe with two columns: first column contains the dates (named "Date") and second column the values of precipitation (named "Precip") **Potevap_month** (monthly average potential evapotranspiration, mm/month) a dataframe with two columns: first column contains the dates (named "Date") and second column the values of potential evapotranspiration (named "Potevap")

4. Data

To launch the test simulation, we used data provided by the colleagues from WP5 for a HoliSoils test site established in Northern Spain (Gámiz) in 2021 (see holisoils.eu), as well as data from global and regional databases, and data directly derived from the literature.

4.1 Soil data

Soil data (e.g., soil texture, pH, $CaCO_3$ and SOC stocks) were derived from the Land Use and Coverage Area frame Survey (LUCAS) database (European Commission Joint Research Centre, 2020). Data for the test run are directly provided in the "Forcing_data_test_run_v0.R" for the Spanish site of Gámiz (coordinates: 42°49'2.32"N; 2°37'10.67"W).

4.2 Climate data

Climate data (e.g., soil temperature, precipitation and potential evapotranspiration) were derived from the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) framework (Frieler et al., 2017). Climate data for the test run can be downloaded from the github platform (https://github.com/elisabruni/Holisoils-multimodel/tree/main/data).

4.3 Site-specific data

Carbon input data under anthropogenically disturbed (i.e. long-term exploitation) and non-disturbed conditions for oak forests (*Quercus suber*) in the Mediterranean region were derived from the literature (Mahmoudi et al., 2021) (see "Forcing_data_test_run_v0.R") .

CO_2 fluxes and experimental design were provided by colleagues from WP5.

5. Next steps

This v0 of the multi-model ensemble will be further developed to include other SOC models (e.g., Millennial and Yasso20), and then calibrated and evaluated to ensure the correct representation of measured SOC stocks and GHG fluxes. Finally, it will be transformed into a user-friendly tool that can be used by a larger public, in order to simulate the effect of forest management and climate change scenarios on SOC stocks and GHG fluxes at specific sites.

6. References

Andrén, O. and Kätterer, T.: ICBM: THE INTRODUCTORY CARBON BALANCE MODEL FOR EXPLORATION OF SOIL CARBON BALANCES, Ecological Applications, 7, 1226–1236, https://doi.org/10.1890/1051-0761(1997)007[1226:ITICBM]2.0.CO;2, 1997.

Andriulo, A., Mary, B., and Guerif, J.: Modelling soil carbon dynamics with various cropping sequences on the rolling pampas, Agronomie, 19, 365–377, https://doi.org/10.1051/agro:19990504, 1999.

Clivot, H., Mouny, J.-C., Duparque, A., Dinh, J.-L., Denoroy, P., Houot, S., Vertès, F., Trochard, R., Bouthier, A., Sagot, S., and Mary, B.: Modeling soil organic carbon evolution in long-term arable experiments with AMG model, Environmental Modelling & Software, 118, 99–113, https://doi.org/10.1016/j.envsoft.2019.04.004, 2019.

Coleman, K. and Jenkinson, D. S.: RothC-26.3 - A Model for the turnover of carbon in soil, in: Evaluation of Soil Organic Matter Models, edited by: Powlson, D. S., Smith, P., and Smith, J. U., Springer Berlin Heidelberg, Berlin, Heidelberg, 237–246, https://doi.org/10.1007/978-3-642-61094-3_17, 1996.

Dimassi, B., Guenet, B., Saby, N. P. A., Munoz, F., Bardy, M., Millet, F., and Martin, M. P.: The impacts of CENTURY model initialization scenarios on soil organic carbon dynamics simulation in French long-term experiments, Geoderma, 311, 25–36, https://doi.org/10.1016/j.geoderma.2017.09.038, 2018.

European Commission. Joint Research Centre.: LUCAS 2015 topsoil survey: presentation of dataset and results., Publications Office, LU, 2020.

Falloon, P., Smith, P., Coleman, K., and Marshall, S.: ESTIMATING THE SIZE OF THE INERT ORGANIC MATTER POOL FROM TOTAL SOIL ORGANIC CARBON CONTENT FOR USE IN THE ROTHAMSTED CARBON, 5, 1998.

Frieler, K., Lange, S., Piontek, F., Reyer, C. P. O., Schewe, J., Warszawski, L., Zhao, F., Chini, L., Denvil, S., Emanuel, K., Geiger, T., Halladay, K., Hurtt, G., Mengel, M., Murakami, D., Ostberg, S., Popp, A., Riva, R., Stevanovic, M., Suzuki, T., Volkholz, J., Burke, E., Ciais, P., Ebi, K., Eddy, T. D., Elliott, J., Galbraith, E., Gosling, S. N., Hattermann, F., Hickler, T., Hinkel, J., Hof, C., Huber, V., Jägermeyr, J., Krysanova, V., Marcé, R., Müller Schmied, H., Mouratiadou, I., Pierson, D., Tittensor, D. P., Vautard, R., van Vliet, M., Biber, M. F., Betts, R. A., Bodirsky, B. L., Deryng, D., Frolking, S., Jones, C. D., Lotze, H. K., Lotze-Campen, H., Sahajpal, R., Thonicke, K., Tian, H., and Yamagata, Y.: Assessing the impacts of 1.5 °C global warming – simulation protocol of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2b), Geosci. Model Dev., 10, 4321–4345, https://doi.org/10.5194/gmd-10-4321-2017, 2017.

Hashimoto, S., Morishita, T., Sakata, T., Ishizuka, S., Kaneko, S., and Takahashi, M.: Simple models for soil CO2, CH4, and N2O fluxes calibrated using a Bayesian approach and multi-site data, Ecological Modelling, 11, 2011.

Mahmoudi, M. R., Bachtobji-Bouachir, B., Sebai, H., Ben-Attia, M., and Ghanem-Boughanmi, N.: Change of the litter fall, decomposition, and nutrient release in cork oak forest after anthropogenic disturbances in North West of Tunisia, Environ Sci Pollut Res, 28, 38584–38593, https://doi.org/10.1007/s11356-021-13294-x, 2021.

Parton, W. J., Stewart, J. W. B., and Cole, C. V.: Dynamics of C, N, P and S in grassland soils: a model, Biogeochemistry, 5, 109–131, https://doi.org/10.1007/BF02180320, 1988.

R Core Team: R: A language and environment for statistical computing, 2022.

Sierra, C. A., Müller, M., and Trumbore, S. E.: Models of soil organic matter decomposition: the SoilR package, version 1.0, Geosci. Model Dev., 5, 1045–1060, https://doi.org/10.5194/gmd-5-1045-2012, 2012.

Tuomi, M., Thum, T., Järvinen, H., Fronzek, S., Berg, B., Harmon, M., Trofymow, J. A., Sevanto, S., and Liski, J.: Leaf litter decomposition—Estimates of global variability based on Yasso07 model, Ecological Modelling, 220, 3362–3371, https://doi.org/10.1016/j.ecolmodel.2009.05.016, 2009.

7. Supplementary figures

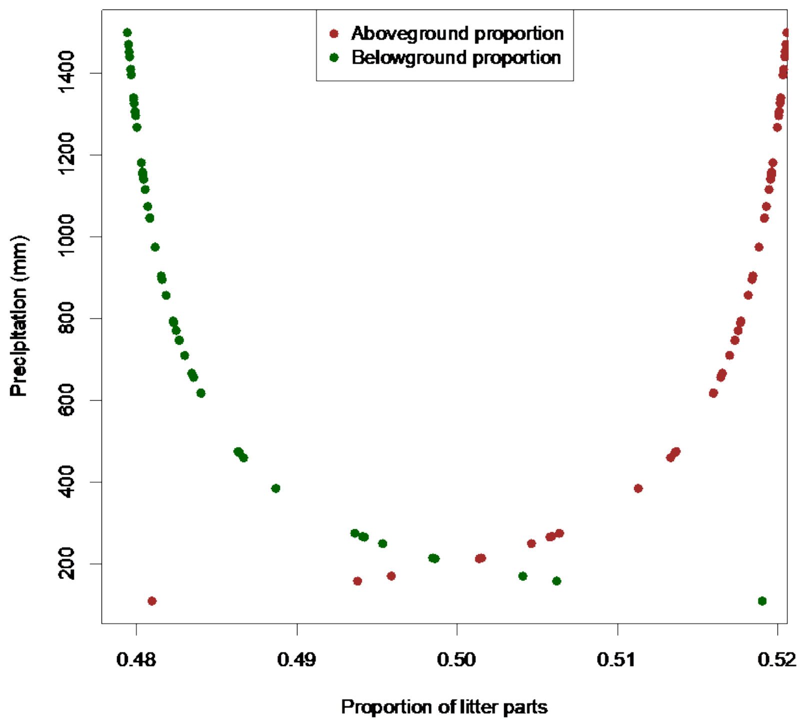

Figure S 1 Precipitation-dependent proportion of the aboveground and belowground litter input used for the AMG and Century models.