General words to Referees.

We truly thank both referees for their positive comments and highly valuable suggestions to improve the manuscript. We greatly appreciate their expertise. In this version, we erased the old traces from the two previous version. Only modification traces since the last version are left.

Referee 1:

Referee: Mao et al., submitted a revised version of their manuscript about the evaluation of the Yasso model over French forest. The paper has been well improved and the reading is now quite good.

Authors: Thank you for this positive remark to our manuscript!

Referee: Nevertheless I still disagree with the approach used to calculate the annual carbon stock changes for the simulation (eq. 7). I thought first it was a typo mistake in the first version of the manuscript but in their answers the authors confirmed that it was the good equation.

By doing ACCsim=(CSsim,t2-CSobs,t1)/(t2-t1) the author don't calculate the annual carbon stock changes in the simulation. Indeed, CSobs,t1 might be different than CSsim,t1 as shown by fig. S4 therefore if the error in the simulation of the carbon stocks at t1 is large the calculation of the ACC is mainly impact by this differences and not by the trends. Fig. S4 shown that Yasso tend to overestimate the carbon stocks for broadleaves whereas it underestimates the stocks for conifers. The impact for some points is probably not negligible. If you want to look to the capabilities of the model to reproduce the observed annual carbon stock changes ACCsim should be calculated as ACCsim=(CSsim,t2-CSsim,t1)/(t2-t1).

In my opinion this modification is mandatory before publication.

Authors: Thank you for re-raising this point. This indicates that we poorly explained this point in the previous version.

The concern raised by reviewer 1 probably comes from a misunderstanding of the initialization of the model. To estimate the change in soil carbon, the model was initialised based on the soil carbon stock observed at the first soil survey. We did not use the simulated carbon stock as it was obtained by supposing that the initial carbon stock was at steady state at one time (e.g. Ortiz et al., 2013). In our case, we have evidences that our soil systems were not at the steady state (Fig. S5).

This confusion is also probably due to the ambiguity of the newly added Fig. S5 (added after the R1 revision as demanded by the other reviewer, with the old name of "Fig. 5S"). Plotting "CS_sim at steady-state" against one of the two observed "CS_obs" (under the suggestion of Referee 2) should solely be regarded as a way to assess the disparity between theoretically obtained steady-state stock and the observed stock. We hope these explanations are now clearer.

As a result, we do not modify the equation and the associated results in the manuscript. Nevertheless, to improve the clarity regarding the reviewer's concerns we have (i) better justified the use of “CS_obs, t1” when presenting the Equation, see P15, LN10-18.; and (ii) clarified the captions of Figure S5 to make readers not think it is “CS_sim, t1”, but “CS_sim at steady-state” , see Fig. S5.

Referee 2:

Referee: The revised paper acknowledges my main concern of improper initialization of carbon quality distribution by a short discussion. The authors argue that comparing another initialization method would be too much work and should go to a new paper.

However, I still think that the entire part on annual carbon accumulation (ACC) cannot be trusted with the currently applied assumption of initializing carbon quality distribution. This might be only my personal interpretation that I defend below, but the authors should better defend their assumption and the validity of their results. Or they should clearly acknowledge that only an upcoming paper will supply more reliably conclusions. For example, one of the conclusions is that its “Yasso07s failure of too much penalizing loss of labile carbon” (P21L4ff). I disagree. To my view it’s the author’s assumptions of scaling the more labile carbon stocks during initialization that determines the ACC pattern in the simulations and its missing correlation with observations.

Authors: Thank you for the understanding and also the paths you suggested.

Regarding the initialization method, on one side, we do not deny the existence of your concern, as the popular steady-state hypothesis (SSH) is indeed a strong one. On the other hand, we do not have intension to change the model’s parameters or configuration as this is not the objective of this manuscript. So we continue to search compromises as follows.

First, we **added a new simulation which is based on an alternative initialization method** (but without modifying Yasso’s configuration and parameter values). To initialize Yasso07, both the quantity and the quality of the soil carbon must be fixed. In the previous versions of the paper, the total quantity was fixed to the soil C stock measured at the first soil survey of the RENECOFOR. The quality was determined by estimating the C amounts at steady state for each chemical fraction (A, W, E, N, H) based on the spin-up/matrix conversion and by calculating their proportions in relation to their sum. These proportions were then applied to the observed C stock to split it in various pools. This approach does not consider the difference in carbon stability among these pools. This is however most likely that the fast-cycling pools such as A, W and E were at steady state at the first soil survey while the H pool could still be far from it (depending on the site history) (as suggested by the Referee). In the new version of the paper, we alternatively considered the C quantity obtained from the spin-up/matrix conversion for A, W, E and N and deduced the H amount by difference with the measured C stock. Although this initialization method is not perfect, it follows well the main idea of Wutzler (2007): due to possible disturbance, the fast-cycling AWEN pools and the slow-cycling H pool do not stabilize at the same time. The text representing this argument can be found in P12-13, Section 2.5.

Next, we **carried out a new sensitivity analysis** on the effect carbon quality (complete versus partial steady-states), crossed with initial carbon quantity (until 40 cm versus until 100 cm) and fine root:leaf ratios (from 0.1 to 4.0) on model fit. See the newly added Fig. S3 related to this analysis. We confirm that the alternative assumption (i.e., partial steady-states assumption, Fig. S3c and S3d) indeed gives better results than those given by complete steady-state assumption, i.e., the one we used in previous versions (Fig. S3a and S3b). In fact, when using CS until 0.4m to initialize Yasso07 (Fig. S3c), the model fit is the best. However, since we obtained the soil carbon stock data until 1.0 m from RENECOFOR from the last version of the manuscript, we decided to use CS until 1.0 for all the simulations related to the main text (Fig. S3d).

So in this new version, we have **redone all the stats and post-analyses** based on this new carbon quality assumption. Accordingly, since the results related to the complete steady-state assumption will not play a major role in the manuscript any more, we have deleted the sensitivity analysis 2.6.1 (effect of litter quality on the steady-state soil carbon quality), as well as its associated texts (e.g., old Section 3.3 and 4.4) and figures (including two supplementary figures), and also old Figure 6 (which was added in the last version). Instead, the Figures 3, 4 and 5 are updated. The new Fig.6 is served for illustrating the disparity of carbon quality between two assumptions. In Supple. Mat., Figure S5 is new (fit with complete steady-state assumption); Table 2, Figures S7 (boxplots on soil type), S8 (geographical distribution of residual signs) and S9 (PCA of residuals) are all updated. The section "sensitivity analyses" in Results has become 3.2, i.e., before the residual analyses (Sect. 3.3). Because we would like to start from general results and justify why we choose that particular case for residual analyses.

Then, as what you suggested, we've **continued to tune down the conclusion**: we also added a phrase in the perspective that expresses the meaning of "upcoming paper will supply more reliable conclusions" in P29, LN19-23. This prudence you suggested is totally ok and acceptable.

We decided to delete the speculative sentence "Yasso07s failure of too much penalizing loss of labile carbon…", as we didn't do a sensitivity analysis on the model's response as a function labile carbon quantity compared with other models (it is not the objective of the paper). See P22, LN16-21.

Referee: Thanks for providing Fig S4. From lines p14/22ff of the manuscript, I infer that „simulated“ corresponds to the equilibrium carbon stocks and „observed“ corresponds to measured 40cm stocks at initial time extrapolated down to 1m, correct? This should be clarified in the figure title. Moreover, mentioning ACC in the figure title is probably an error here.

Authors: We apologize for the unclearness of the text. In Fig. S5 (Fig. S4 in the old version), CS of both axes correspond to stock until 1 m. We modified the figure labels and the captions of Fig. S5 to make it clearer.

Yes, mentioning ACC is an error which is now corrected.

Referee: For the conifers it shows that observed stocks are larger than simulated equilibrium, hinting to some model inadequacy, or mismatches in input fluxes, or underestimate of stead state stocks, e.g. by a too high decomposition rate of the slow pool.

From comparing Fig S4 and Fig. 3 I see that where initial stocks are overestimated, also the stock change is overestimated, and similarly where initial stocks were underestimated, stock change is underestimated. To me, this hints to the suspected large effect of scaling faster pools when transferring steady state to observed stocks.

Authors: We don't deny for the case of conifers, possibly due to the inadaptability of model parameters to the dataset. Thank you for such a good reasoning.

The Fig. S5 (Fig. S4 in old version) also shows that for most broadleaved sites, observed stocks are lower than its steady-state equilibrium, indicating that equilibrium may not yet be reached at these sites.

Since the ACC fit with the alternative initialization method based on the "partial steady-state assumption" (by using simulated absolute values of AWE and the revised N and H, notably H) is improved (see the new Fig. 3) compared to that based on complete steady-state assumption (see the new Fig. S6 for a depth of 1.0 m). We suppose that, this method may, to some extent, mitigate the impact of such discrepancies: for broadleaves, the proportion of A+W+E (that gives rise more CO_2 away) are more enhanced than that at complete steady-state, reducing the model's overestimation of ACC at steady-state; for coniferous sites, the proportion of A+W+E will be pressed, reducing the model's underestimation of ACC at steady-state.

Such phenomena have been mentioned and discussed in Discussion, see P21, LN11-20, P21-22, LN30-2 and P23-24 Section 4.2.

Fig. 4 could probably also be explained by the same issue. When initial steady state stocks are higher than observed stocks, you downscale the faster pools, leading to stronger positive decadal ACC when pools develop towards steady state again. Similarly, when initial steady state stocks are lower, you upscale the faster pools leading to negative decadal ACC during simulation. The missing of this pattern in observations suggests to me to better not scale the faster pools.

Authors: See our responses above. With such an alternative assumption (while keeping model parameter/configuration untouched), such effect is mitigated. In this version, we've put the figure based on complete steady-state assumption model fit as Fig. S6 (down to 1.0 m this time). Now the Fig. 3 in the main text is based on the partial steady-state assumption (down to 1.0 m).

Referee: My conclusion from you results, therefore, is that probably the mismatch in initial distribution of qualities determines your pattern of carbon accumulation. To me most of the results come down to the assumption of keeping carbon quality distribution constant when adjusting stocks.

Authors: See our responses above too.

Referee: Specific comments:

I cannot get the message from Fig. S9. However, it should be crucial to my interpretation of importance of changes to steady state litter quality distribution for ACC. Can you simplify it?

Authors: We keep the Fig. S4 (Fig. S9 in old version) to give a general idea how and how much initial carbon quality can alter the model's output. In this version, the description on Fig. S4 in the main text has been simplified to 9 lines.

Referee: P22l15ff: Thanks for including discussion on the relaxed equilibrium assumption. Lack of information on modified rate of H pool should not be a problem at the time scale of interest. Just use a significantly slower decomposition rate, as suggested by Wutzler 2007. The difference will only be relevant at longer time scale.

Authors: See our response above.

Referee: P23L7: Note that the intended usage of depth-dependent decomposition rates renders the calibrated model only applicable at the site of calibration. Depth is often just a surrogate factor for other stabilization mechanisms. Trying to capture better indicators for these stabilization mechanisms is better than relying on their relation with depth that will change across sites.

Authors: Thank you for this comment and suggestion. We rephrased the sentence and added "capture better indicators for these stabilization mechanisms." See P25, LN7-8.

Modeling soil organic carbon dynamics in temperate forests using Yasso07

Zhun Mao[1,8*], Delphine Derrien[1], Markus Didion[2], Jari Liski[3,9], Thomas Eglin[4], Manuel Nicolas[5], Mathieu Jonard[6], Laurent Saint-André[1,7]

[1] INRA, UR BEF – Biogéochimie des Ecosystèmes Forestiers, 54280 Champenoux, France

[2] Swiss Federal Institute for Forest, Snow and Landscape Research WSL, 8903 Birmensdorf, Switzerland

[3] Finnish Environment Institute, Ecosystem Change Unit, Natural Environment Centre, Mechelininkatu 34a, P.O.Box 140, 00251 Helsinki, Finland

[4] ADEME – DPED – Service Agriculture et Forêts, 49004 Angers, France

[5] Office National des Forêts Direction Forêts et Risques Naturels, Département Recherche et Développement - Bâtiment B, Boulevard de Constance, 77300 Fontainebleau, France

[6] Université Catholique de Louvain, Earth and Life Institute, Croix du Sud 2, L7.05.09, 1348 Louvain-la-Neuve, Belgium

[7] CIRAD, UMR ECO&Sols, place Viala, 34398 Montpellier Cedex 5, France

[8] Amap, Inra, University Montpellier, Cnrs, Ird, Cirad, Montpellier, France

[9] Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland

*Corresponding author: Zhun Mao; email address: maozhun04@126.com.

Abstract. Facing global changes, modeling and predicting the dynamics of soil carbon stock in forest ecosystems is vital but challenging. Yasso07 is considered as one of the most promising models for such a purpose. We aim at examining the accuracy of its prediction of the soil carbon dynamics over the whole French metropolitan territory at a decennial time scale.

We used data from 101 sites of the RENECOFOR network, which encompasses most of the French temperate forests. These data include (i) yearly measured quantity of aboveground litterfall from 1994 to 2008, and soil carbon stocks measured twice at an interval of c.a. 15 years (early 1990s versus around 2010). Using Yasso07, we simulated the annual carbon stock changes (tC ha^{-1} yr^{-1}) per site and compared them with the measured ones. We carried out meta-analyses to reveal the variability in litter biochemistry between different tree organs for conifers and broadleaves. We also performed sensitivity analyses to explore Yasso07's sensitivity to inputs.~~,~~ ~~including litter carbon quality and initial carbon stocks~~.

At the national level, the simulated annual carbon stock changes (ACC, +0.00~~45~~ ± 0.07~~9~~ tC ha^{-1} year^{-1}, mean ± standard error) stayed in the same order of magnitude as the observed ones (+0.34 ± 0.06 tC ha^{-1} year^{-1}). The correlation between predicted and measured ACC remained weak (R^2 <0.1). There was significant overestimation for broadleaved stands and underestimation for conifers sites. Sensitivity analyses showed that the final carbon stock was weakly affected by settings in model initialization, including litter and soil carbon quantity and quality, and ~~litter carbon quality, but strongly affected~~ also by simulation length~~ and initial soil~~

~~carbon quality~~. Carbon quality set with the partial steady-state assumption gave a better model fit than that with the complete steady-state assumption.

Taking Yasso07 as model support, we revealed the current bottleneck of soil carbon modelling due to lacking knowledge or data on soil ~~and litter~~ carbon quality and fine root litter quantity, rendering high uncertainties for model inputs.

Nomenclature and abbreviations

Name	Meaning
carbon stock (CS)	Quantity of soil organic carbon stock (in tC ha^{-1})
carbon stock change	Increment (positive value) or decrement (negative value) of soil organic carbon stock from the year t1 to the year t2 (in tC ha^{-1})
annual carbon stock change (ACC)	carbon stock change standardized by duration (in tC ha^{-1} $year^{-1}$)
carbon pools	The Yasso07 model contains a series of organic compounds differing in solubility in solvents and mean residence time in decomposition processes: water soluble compounds (W), acid-hydrolysable compounds (A); non-polar solvent, ethanol or dichloromethane compounds (E), non-soluble and non-hydrolyzable compounds (N). For soil, there is an extra recalcitrant pool named “humus” (H). Note: in this paper, “N” only denotes non-soluble and non-hydrolyzable compounds; nitrogen is spelled in full letter when mentioned.
coarse woody litter	Litter yield from either coarse aboveground residues due to either harvests or storms (including coarse branches, defined as branched of >4 cm in diameter and miscellaneous) and coarse roots (defined as those of >5 mm in diameter)
fine non-woody litter	Litter yield from either natural above-ground litterfall (leaves, small branches) or fine roots activities
litter carbon quality	Composition of litter carbon belonging to A, W, E and N carbon pools (in %)
litter quantity	Annual litter input (in tC ha^{-1} $year^{-1}$)
soil carbon quality	Composition of soil carbon belonging to A, W, E, N and H carbon pools (in %)

1 Introduction

The carbon stock in global soils, including litter and peatlands is 1500 to 2400 GtC, greatly exceeding that in vegetation (350 à 550 GtC, mainly in forests) and in the atmosphere (829 GtC in 2011, IPCC, 2014). Soils share a common interface with all the other spheres and play a key role in driving the global carbon cycle. Soil carbon stock dynamics are directly related to the greenhouse gas emissions (notably carbon dioxide; ~~(~~CO_2~~)~~) that are leading to the global warming effect (IPCC, 2014). An accurate estimation of soil carbon stock dynamics allows us to better understand the turnover rate and fate of soil carbon flux at both local and global geographical scales. Facing global changes, this task is essential for the evaluation of the climate change mitigation potentials of forests and the support of environmental policy decisions.

Significant challenges exist for accurate estimation of soil carbon stock changes. Current soil monitoring networks are generally not able to detect changes on timescales of less than 10 years (Saby et al. 2008). To obtain soil C stock change estimates at shorter intervals such as for the annual reporting to the United Nations Framework Convention on Climate Change and the Kyoto Protocol, the use of models is encouraged (IPCC, 2011). Numerous models have been elaborated for evaluating soil carbon dynamics (Manzoni and Porporato, 2009). The vast majority of terrestrial soil carbon models developed at the global or at the plot scales, e.g., CENTURY (Parton et al**.,** 1987), RothC (Coleman and Jenkinson, 1996) and ORCHIDEE (Krinner et al., 2005), assume that decomposition is the first order decay process accounting for the size of soil carbon pools, despite the existence of criticism to this, arguing that priming effect and the associated induced carbon pool interactions should be considered in model algorithms (Wutzler and Reichstein, 2013). The dynamics of carbon pools depend on the quantity and quality of litter inputs and on temperature, soil moisture and other soil parameters, e.g. texture, structure, chemical richness, pH etc. (Todd-Brown et al., 2012). Incorporating explicit mechanisms such as microbial activities or carbon protection by the soil matrix into soil carbon models has repeatedly been suggested in the last years (Schmidt et al., 2011; Lehmann and Kleber, 2015). However, for forest ecosystems, such refined mechanistic input data remain often limited. Accordingly, the typical time-step for litter input demanded by most of soil carbon models for forests is year, not month (but see RothC, Coleman and Jenkinson, 1996) or day (but see Romul, Chertov et al., 2001) (Didion et al., 2016). At this yearly-timescale, it is common to consider microbial communities and processes as a relatively stable factor (Todd-brown et al, 2012), and the assumption of carbon dynamics governed by first order decay may therefore be reasonable.

This is the choice made by the group who built the Yasso model (Liski et al., 2005) and Yasso07 model (Tuomi et al., 2009; 2011a and 2011b), i.e. an improved version of Yasso with more refined carbon pooling and abundant data for calibration. The intention of the models' developers is to let their models be suitable for general forestry applications by taking into account the low availability of forest soil and litter data (Liski et al., 2005). Yasso07 explicitly defines several chemical pools of chemical compounds in litter carbon (Tuomi et al., 2011b) and possesses well-defined, biological meaningful and measurable parameters. Due to these qualities, Yasso and Yasso07 were applied in more than 70 case studies (URL: http://www.syke.fi/en-US/Research__Development/Research_and_development_projects/Projects/Soil_carbon_model_Yasso/) in forest ecosystems in the northern hemisphere with generally high satisfaction levels in comparison with measured carbon values (e.g. Karhu et al., 2011 ; Rantakari et al., 2012; Ortiz et al., 2013 ; Didion et al., 2014; Lu et al., 2015; Wu et al., 2015). Yet, so far most of these applications have been limited to local case studies, especially those on cold forests with limited tree species diversity (e.g. boreal or montane forests). Rarely have previous studies validated Yasso07 based on data (i) of long-term observations (here defined as data of >10 years), (ii) from temperate forests with a much higher diversity of tree species or (iii) on carbon stock changes (in tC ha^{-1} $year^{-1}$). This is partially due to the lack of extensive long term soil carbon monitoring in forest ecosystems which differ in climatic and soil conditions and species, stretch over a large territorial scale. Nevertheless, Yasso07 has been considered as one of potential models appropriate for evaluating national and continental inventories of forest carbon balance in Europe (Hernández et al. 2017). It is therefore of high interest to assess the ability of Yasso07 to reflect the carbon balance in different European forest ecosystems at large spatial-temporal scales. Moreover, as a carbon pool based model, Yasso07 shares certain similar principles to other prevailing soil carbon models in the same genre (e.g., RothC, CENTURY etc.). Via Yasso07 as an example, we may also learn from this application case for future carbon modelling for temperate forests

The measured data of carbon stock and litter quantity dynamics from the RENECOFOR network (URL: http://www.onf.fr/renecofor/@@index.html), National Forest Management Agency (ONF), France, offered us a valuable opportunity for model validation. The 101 forest sites considered from this network are located all over the French metropolitan territory and cover the most common forest types and tree species. For each site, annual measurements of litterfall were available in addition to two inventories of soil organic carbon stock with an average interval of 15 years (minimum 12 years and maximum 20 years). These data allowed us to use site-specific observed soil carbon stock and above-ground litterfall dynamics as model input estimates, thus reducing the uncertainties of the model input, which were identified as a major source of uncertainties for model estimates of soil carbon stock changes (Ortiz et al. 2013). By minimizing this source of uncertainty, we were able to focus on the inherent model structure.

Consistent with our objective to contribute to the further development of soil carbon modeling, we aim at (i) testing and characterizing the ability of Yasso07 to model soil carbon stock dynamics for temperate forests (ii) identifying limitations and providing suggestions for a better adaptation of the model for C dynamics in both deciduous and evergreen temperate forests and (iii) discussing the perspectives based on the current state-of-the-art of soil carbon modelling. Associated with the above aims, our null hypotheses are as follows: (i) Yasso07 predicts accurate and unbiased carbon stock changes at the national scale and (ii) the model's fit residuals (predicted data minus observed data) have null relationships with site characteristics (e.g. location, climate, forest type, soil type and initial carbon stock).

2 Materials and methods

2.1 The model Yasso07

The dynamic soil carbon model Yasso07 is based on the general assumption that the soil carbon stock is driven by decomposition of different litter types, which may differ in quantity and quality, and by climatic conditions. Litter carbon quality is represented by four chemical compound groups which have different decomposition rates (Tuomi et al., 2009). Soil organic carbon is divided into these four relatively labile carbon pools and one recalcitrant pool named “humus” (H) (Fig. S1). The five pools differ in specific mass loss rates and mass flows among them. As in many other pool-based models, the H pool is considered the oldest and most stable carbon pool, although recent studies doubted its physical existence and stability (see Lehmann and Kleber, 2015). Some mass flows correspond to CO_2 release (microbial respiration). The mean residence time of carbon in these pools varies from several months (i.e., water soluble compounds, W), a few years (i.e., acid-hydrolysable compounds, A; non-polar solvent, ethanol or dichloromethane compounds, E), several decades (i.e., non-soluble and non-hydrolyzable compounds, N), or even several centuries (i.e., H).

Mathematically, the kernel equation of Yasso07 can be written as follows:

$$\dot{\boldsymbol{X}}(t) = \boldsymbol{A_p K}(c)\boldsymbol{X}(t) + \boldsymbol{I}(t) \qquad \text{(Eq. 1a)}$$

where, symbols in capital letters in bold denote either vectors or matrices whilst those in small letters in parentheses denote scalars; $\boldsymbol{X}(t)$ and $\boldsymbol{X}(t)$ are vectors describing the masses of the five carbon pools (A, W, E, N, H) and carbon mass changes in soil at time (t), respectively; $\boldsymbol{A_p}$ is mass flow matrix describing carbon allocation among pools; $\boldsymbol{K}(c)$ is decomposition matrix describing the decomposition rates as a function of climatic conditions (c); $\boldsymbol{I}(t)$ is litter input to the soil, with the last element equal to 0, as “H” does not exist in litters. (Eq. 1a) can be expressed in a more detailed form:

$$\begin{pmatrix} \partial x_A/\partial t \\ \partial x_W/\partial t \\ \partial x_E/\partial t \\ \partial x_N/\partial t \\ \partial x_H/\partial t \end{pmatrix} = \begin{pmatrix} -1 & p_{W\to A} & p_{E\to A} & p_{N\to A} & 0 \\ p_{A\to W} & -1 & p_{E\to W} & p_{N\to W} & 0 \\ p_{A\to E} & p_{W\to E} & -1 & p_{N\to E} & 0 \\ p_{A\to N} & p_{W\to N} & p_{E\to N} & -1 & 0 \\ p_{A\to H} & p_{W\to H} & p_{E\to H} & p_{N\to H} & -1 \end{pmatrix} \begin{pmatrix} k_A & 0 & 0 & 0 & 0 \\ 0 & k_W & 0 & 0 & 0 \\ 0 & 0 & k_E & 0 & 0 \\ 0 & 0 & 0 & k_N & 0 \\ 0 & 0 & 0 & 0 & k_H \end{pmatrix} \begin{pmatrix} x_A \\ x_W \\ x_E \\ x_N \\ x_H \end{pmatrix} + \begin{pmatrix} I_A \\ I_W \\ I_E \\ I_N \\ 0 \end{pmatrix} \qquad \text{(Eq. 1b)}$$

where, $p_{F\to T}$ is the relative mass flow parameters between two pools (from F to T; F and T can be any two pools in A, W, E, N and H) in the soil (dimensionless, $p_{F\to T} \in [0, 1]$).

Temperature and precipitation are supposed not to affect the mass flows p, but influence the mass loss rates k_i (i = A, W, E, N or H) according to:

$$k_i\ c\ = \alpha_i \exp\ \beta_1 T + \beta_2 T^2\ [1 - \exp(\gamma P_a)] \qquad \text{(Eq. 2)}$$

where, α_i is the mass loss rate parameter of the chemical pool i; β_1, β_2 and γ are parameters related to temperature (T, in °C) and precipitation (P_a, in mm).

To consider the effect of litter size on the decomposition rate of litters, k_i was multiplied by a litter size factor (h_s), which allows making the distinction between different types of litters, e.g. foliage, coarse woody, stem etc., which differ in diameter (d, in mm):

$$h_s\ d\ = \min\ (1 + \varphi_1 d + \varphi_2 d^2)^r, 1 \qquad \text{(Eq. 3)}$$

where, φ_1, φ_1 and r are parameters related to litter size.

Yasso07 has 44 parameters calibrated using the Markov chain Monte Carlo (MCMC) method with the Metropolis-Hastings algorithm (Tuomi et al., 2011a). Currently, several calibrated parameter sets for Yasso07 are available, including the two most recent sets published by Tuomi et al. (2011) and Rantakari et al. (2012). In this present study, the Tuomi 2011 set was chosen to fit the RENECOFOR dataset containing various forest species, as it had been calibrated using a wider range of observed foliage and root decomposition data. The Tuomi 2011 set was calibrated using a combination of three sources of dataset: (i) a global dataset (n >9000) of litterbags for mass loss of non-woody litters from approximately 100 sites in Europe, Northern and Central America. These sites covered a wide range of climate and soil conditions, forest types and tree species; (ii) a dataset (n > 2000) of mass loss of decomposing woody litter measured in Northern Europe; (iii) measured accumulation rate of soil carbon pools of forest sites along a 5300 year soil chronosequence in southern Finland, for determining the residence time of the H carbon pool. The Tuomi 2011 parameter set contains 10000 parameter vectors (each vector contains the values of all the 44 Yasso07 parameters), which are randomly generated to take into account stochastic effect.

2.2 RENECOFOR network

The RENECOFOR network is part of the Level II network of the International Cooperative Program on Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forest). The 101 sites (Fig. 1) considered in this study cover the most common types of forest ecosystems in France, including even-aged forests in plain area, pine plantations and uneven-aged mountain forests. They also cover the majority of tree species in France and central Europe, including *Quercus robur*. *Quercus petraea*, *Pseudotsuga Menziesii*, *Picea abies*, *Fagus sylvatica*, *Pinus pinaster*, *Pinus sylvestris* and *Abies alba*. At each site, annual forest woody and non-woody litter quantities have been either directly measured or estimated based on the existing dendrometric data.

2.2.1 Soil carbon and physical and chemical properties

At each site, soil carbon stocks (CS) were measured twice with an interval of approximately 15 years (1993 – 95 for the first assessment and 2007 – 12 for the second one). ~~The temporal evolution of soil carbon stocks was analyzed by Jonard et al. (2017).~~ At each site and for each assessment, soils to a depth of 0.4 m were sampled from five points selected in each of the five subplots and divided into different layers (0 – 0.1 m, 0.1 – 0.2 m and 0.2 – 0.4 m), including both organic and mineral soil layers. The temporal evolution of the soil ~~carbon stocks~~CS until 0.4 m was analyzed by Jonard et al. (2017). Composite samples were produced for each layer and subplot, and analyzed for mass, bulk density, soil organic carbon and physical and chemical properties, including texture (percentages of clay, silt and sand, in %), pH value, total nitrogen stock (in t ha^{-1}), carbon:nitrogen ratio (dimensionless), total phosphor stock (in t ha^{-1}), stocks of exchangeable aluminum (Al), calcium (Ca), potassium (K) and magnesium (Mg, in kmol ha^{-1}). Soil physical and chemical properties data were used for residual analyses (see Sect. 2.7) and only those measured in the 1st inventories were used for this purpose.

Regarding the CS of depth 0.4 – 1.0 m, only the data of the first assessment (1993 – 95) are available. Soil samples were obtained from only one soil profile per site at two mineral layers (0.4 – 0.8 m and 0.8 – 1.0 m). Bulk density and carbon concentration measured at these layers were used to estimate soil carbon stock until a depth of 1.0 m. Table 2 provides a synthesis of the data source for each of the 101 sites of the RENECOFOR network (URL: http://www.onf.fr/renecofor/sommaire/renecofor/reseau/20090119-130815-828957/@@index.html). More detailed information about each site and soil sampling procedure is available in Supplementary Material I (Table S1) and Jonard et al. (2017).

2.2.2 Climate data

Necessary climate data required by Yasso07 includes annual mean precipitation (mm) and annual maximum, mean and minimum temperature (°C). These measured data were obtained from the nearest national meteorological stations of Météo-France (http://www.meteofrance.com) for each RENECOFOR site.

2.3 Litter quantity

Litter input (in tC ha^{-1} yr^{-1}) comes from several sources (Table 2) as follows. The conversion factor between biomass (dry matter) and carbon was assumed to be 0.5 (Thomas and Martin, 2012).

Aboveground litter input from living trees includes leaves for broadleaves and needles for conifers, small branches, fruits and miscellaneous (e.g., flower, bud etc.). Aboveground litterfall mass was annually measured between 1994 and 2008. For sites where litter quantity data from 1992 – 1993 and 2009 – 2012 were lacking, we used mean litter quantity of all the other years of the same site. The observed branch size in this category is below 2 cm (fine branches). Branches and stems bigger than 2 cm due to natural mortality should be rare (as some of them can be salvaged) and thus were not included.

Woody residues due to harvest or storms were estimated on the basis of repeated stand inventory data and species specific height-girth and biomass. Coarse woody litter inputs from harvesting residues or storms were estimated from full inventories performed by ONF since 1991. Missing years of litter input of this category are gap-filled using the average over the period. On average 3 years are missing per site but there are high differences amongst sites. The mode is one year, and 6 sites have 10-11 missing years. These residuals are assumed to be coarse branches (> 4 cm in diameter, confirmed with ONF) as a function of aboveground tree characteristics. Litter input from stems was set to 0, since in most cases stemwood was removed from the site after storm damage. Litter input from coarse woody roots is considered to be equal to total root biomass, which could be estimated using meta-analysis based allometric equations proposed by Cairns et al. (1997). More detailed information about forest inventories and storm events occurring at each site is available in Supplementary Material I (Table S1). Litter input from fine roots (here defined as roots of ≤ 5 mm in diameter), especially those finest ones with diameter ≤2 mm, can significantly contribute to carbon sequestration in soils (Brunner et al., 2013; Kögel-Knabner et al., 2002; Berg and McClaugherty, 2008). Fine root litter was supposed to be proportional to that of foliage, which was measured on the RENECOFOR sites. Jonard et al. (2017) suggested using the generic equation published by Raich and Nadelhoffer (1989) and, simultaneously, adopting the hypothesis that fine root litter production represents about one third of the carbon allocated to roots (Raich and Nadelhoffer~~ and Raich~~, 1989):

$$I_{fine\ root} = 0.333 \times 1.92 \times (100 \times I_{foliage}) + 130 \times 0.01 \quad \text{(Eq. 4)}$$

Where, $I_{fine\ root}$ and $I_{foliage}$ are litter input of fine root and foliage, respectively (in tC ha^{-1} year^{-1}).

~~However, t~~The relationship between fine root and foliage litter inputs can be highly variable as a function of tree species, stand characteristics and climate and such variability may not be represented in the generic equation. For this, we carried out a sensitivity analysis to investigate the response of model fit to the choice of fine root:foliage ratio varying from 0.1 to 4.0 (see Sect. 2.6 and 3.2).~~ Therefore, here we estimated litter input for Yasso07 simulations using fine root:foliage ratios ranging from 0.1 to 4.0. Based on a sensitivity an analysis on the effect of fine root:foliage ratio, we found that ratios of 0.1 for broadleaves and 1.9 for conifers achieved the best fit between simulated and observed soil C stock changes (Fig. S).~~ Yet, when applying Raich and Nadelhoffer (1989)'s equation (Eq. (4) over all the RENECOFOR sites, we found that fine root:foliage ratios had a median of 1.0 and a mean of 1.0 – 1.1 for both coniferous and broadleaved sites (Fig. S2). Hence, we chose to present the outcomes of model fit and residual analyses from the simulations using the ratio of 1.0 over all the RENECOFOR sites (see Sect. 3.3). Such a choice facilitates our evaluation of site factors (e.g. dominant tree functional type, climatic and soil features) without the additional source of variability introduced by litter quantity.

2.4 Litter carbon quality

There are no measured data of litter carbon quality, defined as composition of litter carbon belonging to different carbon pools (A, W, E and N) in the RENECOFOR network. Therefore, we carried out a meta-analysis on the data collected in literature where authors measured litter carbon quality via chemical fractioning procedures or near-infrared spectroscopy (NIRS) techniques. This data collection was restricted to non-tropical areas. Chemical data on litters of tree coarse organs (e.g. stems, coarse branches) are relatively scanty, so we used tree stemwood data compiled in Pettersen (1984), Rowell et al., (2005) and Rowell (2012). Assembly of these works covers a wide range of temperate tree species from North America, Japan and Russia, but no data are available for Europe. Data on foliage and root litter carbon quality were manually searched from either networks, e.g. CIDET (Trofymow et al., 1998) and LIDET (http://andrewsforest.oregonstate.edu/research/intersite/lidet.htm) or independent studies in northern hemisphere, including Europe. The database for the meta-analysis is available in Supplementary Material II. Root diameter or branching order can play a significant role in modifying the composition of the chemical compounds (Fahay et al., 1988; Tingey et al., 2003; Guo et al., 2004). All the measurements included in the meta-analysis on roots refer to fine roots (diameter < 5.0 mm), although in several studies, e.g. Aber et al. (1990), Aulen et al. (2011) and Stump and Binkley (1993), root size was not clearly indicated. Yet, we still included the data from these above studies, as available root data are less abundant than foliage. The collected coarse roots data in literature were too few for a meaningful meta-analysis and thus values for stemwood were used instead.

We then used the litter carbon quality database to assign the quality of litter input of each site of our study. Partitioning of litter inputs in biochemical classes respects the following order of priority: (i) values for the target species, when available in the database (ii) mean values of the species from the same genus, if data for the target species are absent, and (iii) mean values of the species from the same tree functional type (conifers versus broadleaves), if data are available at neither species nor genus level for a target species (see Table 1).

2.5 Initialization of soil carbon quantity and quality

To initialize Yasso07, both the quantity and the quality of the soil carbon are required. Here, the initial carbon stock quantity was fixed to the soil carbon stock measured at the first soil carbon assessment of the RENECOFOR (i.e. a model input). Measurement uncertainties of soil carbon stock were not considered as a source of stochastic effect when Yasso07 was fed, as we were more interested in the output uncertainties related to the model per se (i.e., the choice of model parameter set) and carbon quality settings in model initialization (see below). The carbon quality, defined as the proportions of soil carbon pools (A, W, E, N and H) in relation to their sum, can be initialized following two approaches. The classical approach is based on the assumption that carbon quality at initial state is identical to that at the complete steady-state, which can be calculated using the analytical matrix inversion approach based on Eq. 1a. ~~To calculate steady-state carbon stock, we used an analytical approach on the basis of (Eq. 1a)~~. At steady-state carbon stock ($t = t_s$), carbon gain is equal to carbon loss. Setting $\dot{\boldsymbol{X}}\ t_s\ = 0$, (Eq. 1a) becomes:

$$\boldsymbol{A_p K}\ c\ \boldsymbol{X}\ t_s\ + \boldsymbol{I}\ t_s\ = 0 \quad \text{(Eq. 5)}$$

Solving (Eq. 5), we obtained steady-state carbon stock at time t_s: $\boldsymbol{X}\ t_s$:

$$\boldsymbol{X}\ t_s\ = -(\boldsymbol{A_p K}(c))^{-1}\boldsymbol{I}\ t_s \quad \text{(Eq. 6)}$$

Where $\boldsymbol{I}\ t_s$ is a constant vector.

The estimated carbon quality in steady-state carbon stock $\boldsymbol{X}\ t_s$ to the depth of 1.0 m (also noted as $C_{steady\text{-}state}$, in tC ha^{-1}) was then applied to the observed carbon stock to split it in various carbon pools~~This steady-state carbon stock to the depth of 1.0 m ($C_{steady\text{-}state}$, in tC ha^{-1})~~

~~was only used to calculate the soil carbon quality distribution, here defined as the composition of soil carbon pools (A, W, E, N and H)~~. ~~Such calculation was performed for each site and for each randomly chosen Yasso07 parameter vector (see Sect. 2.7). Regarding the initial soil carbon quantity, we used the measured one during the first period of assessment of the RENECOFOR network.~~

The complete steady-state assumption is commonly used in literature despite high controversy~~,~~ as such assumption does not consider the difference in stabilization among these pools (Elliot et al., 1996; Foereid et al., 2012). Soil carbon pools (especially those at sites that underwent disturbances in recent centuries) may not be in a complete steady-state, but in a transient or partial steady-state. In such states, the slow-cycling pools can be still accumulating carbon, while the relatively rapid-cycling pools are able to recover until a dynamic equilibrium (Wutzler and Reichstein, 2007). In this study, we adopted the partial steady-state assumption to mimic such a circumstance. More precisely, we assumed that the rapid-cycling pools such as A, W and E were at steady-state at the first soil survey, while the slow-cycling N and H pools might not have reached the steady-state yet. Accordingly, while directly considering the steady-state CS obtained from matrix inversion as A, W and E, we revised N and H amounts by calculating the difference with the observed CS until 1.0 m. In most cases, the sum of steady-state A, W, E and N was lower than the observed CS; the revised H was then equal to the difference between the latter and the former. Very occasionally, the sum of steady-state A, W, E and N could be greater than the observed CS; the revised N was then calculated by the difference between observed carbon stock and pool H was forced to zero. The new carbon quality, which corresponds to the proportions among the steady-state A, W and E and the revised N and H, will be used to split the observed CS in real simulations.

2.6 Sensitivity analyses on the impact of initial soil and litter settings on model output

It is important to gain a general idea of the magnitude of impact of our choices of initial soil and litter settings in the process of model initialization on model output and fit. To this end, we carried out a sensitivity analysis to assess how assumptions on carbon quality (complete steady-state versus partial steady-state) and carbon quantity as a function of soil depth (observed CS until 1.0 m versus observed CS until 0.4 m) and of fine root:shoot ratios (from 0.1 to 4.0) affected model predictions. Model fit is expressed via the comparison between simulated and observed annual carbon stock changes in soil (ACC).

~~To assess the effects of initial litter and soil carbon quality on model outputs, we conducted two modules of sensitivity analyses differing (see below).~~

~~2.6.1 Module I - Effect of litter carbon quality on steady-state carbon stock~~

~~First, we investigated the effect of all the theoretical possibilities of litter carbon quality on steady-state carbon quality. For this, we permuted the carbon percentage in each pool with the following constraint: the minimal and maximum percentages are 5 and 85%, respectively (In permutations, the unitary increment or decrement of each pool is ± 5 %).~~

~~Second, we investigated the impact of tree functional type on the steady state of soil carbon quality. For this, we used the mean and standard deviation of broadleaved and coniferous litter carbon quality calculated from the meta-analysis in Sect. 2.4. To only focus on the effect of litter carbon quality, the litter quantity was the same for broadleaves and conifers. Outcomes were calculated using the matrix method stated in Sect. 2.5 and the Tuomi-2011 parameter set. Possible correlations between A, W, E and N were not considered in simulations.~~

~~2.6.2 Module II - Effect of initial soil carbon quality and simulation length on final soil carbon stock~~

Besides, to fully explore the effects of all the theoretical initial soil carbon quality and that of simulation length on model outputs, we conducted another sensitivity analysis~~With a fixed initial soil carbon stock, we investigated the response of simulated final soil carbon quantity and quality to the setting of initial soil carbon quality and that of simulation length~~. For this, we created a virtual site where the climatic condition and litter input were constant and equal to the average values of the RENECOFOR sites. By~~we~~ fixing its initial soil carbon stock to 100 tC ha^{-1}, we permuted the initial percentage of soil carbon pools with the following constraint: the minimal and maximum percentages are 5% and 80%, respectively. We used four levels of simulation length (1, 10, 100, 1 000 and 10 000 years) for each combination of soil carbon quality distribution. ~~Initial carbon stock was fixed to 100 tC ha^{-1}.~~ Based on averaged soil and litter carbon data of RENECOFOR sites, the simulations were carried out for both broadleaved and coniferous forest stand cases. Here, only the results of broadleaved stand case were presented, as results between conifers and broadleaves did not change much, especially in long term.

2.7 Running Yasso07 and statistical analyses

We used the same FORTRAN code of the Yasso07 version 1.0.1 used in Didion et al. (2014) for all the model simulations. For each analysis (both RENECOFOR site specific and sensitivity analyses), we conducted 10 simulations. In each simulation, one parameter vector was randomly chosen from the 10 000 parameter vectors.

For each site, we calculated annual carbon stock changes (ACC, in tC ha^{-1} year^{-1}), i.e., the difference of carbon stock between the two national inventories standardized by the temporal interval (t_2 - t_1) as follows:

$$ACC_{obs} = (CS_{obs,t2} - CS_{obs,t1})/(t_2 - t_1)$$
$$ACC_{sim} = (CS_{sim,t2} - CS_{obs,t1})/(t_2 - t_1) \quad \text{(Eq. 7a and 7b)}$$

Where, $CS_{sim,t2}$, $CS_{obs,t2}$ and $CS_{obs,t1}$ are the simulated carbon stock until 1.0 m at the year t_2, observed carbon stock at the year t_2 and t_1, which are around the year of 1994 and 2010 depending on each site, respectively.

To compute ACC_{sim} (Eq. 7b), some studies used a simulated CS at the starting year instead of an observed one (e.g. Ortiz et al., 2013). In such a case, it is of primary importance to judge a "steady-state year" prior to the starting year from which observed data are available. From the estimated steady-state year, a spin-up or real model simulation is then followed to obtain a simulated CS at the starting year. In our simulations, the observed soil carbon stock at t_1 was served as a model input to set initial soil quantity and to calculate ACC (Eq. 7b). This allows avoiding such a judgement on steady-state year, which can be sometimes subjective. This also allows better focusing on the effect of initialized soil carbon quality, for which we attempted both complete or partial steady-state assumptions (see Sect. 2.5).

Two reasons support our general preference of comparing ACC_{sim} with ACC_{obs} over comparing $CS_{sim,t2}$ with $CS_{obs,t2}$. First, the parameter sets of Yasso07 were calibrated for a soil depth of 1.0 m, while carbon stock data~~s~~ from two assessments at the RENECOFOR sites were only available until ~~down to~~ 0.4 m (because the data of 0.4 - 1.0 m depth from the 2nd assessment are unavailable). It is thus reasonable to speculate that the observed carbon stock data are not comparable with Yasso07 estimates. However, focusing on carbon changes instead of carbon stocks may largely erase this bias, because previous studies have evidenced that carbon dynamics are much less active at deep soil layers than at superficial layers (Jandl et al., 2014; Balesdent et al., 2018). Second, ACC indicates if a site is gaining or losing soil carbon and this information is sometimes more important than the site's carbon stock value. Using a standardized metric (by year) such as ACC can also facilitate result comparison for future studies. The only exception came to the sensitivity analysis on the effect of initial soil carbon quality (Sect. 2.6), in which we showed $CS_{sim,t2}$ instead of ACC_{sim}, as the initial soil carbon stock was fixed at 100 tC ha^{-1}. Despite the primary focus on ACC, we additionally compared the simulated steady-state carbon stock ($CS_{steady\text{-}state}$, in tC ha^{-1}), which was obtained from the initialization procedure (see Sect. 2.5), with the $CS_{obs,t1}$ down to 1 m soil depth in order to check if Yasso07's predicted stocks to 1.0 m depth reach the level of observed stocks (see Fig. S4). ~~Then, we calculated the steady-state carbon quality for all the 101 sites, using site-dependent climatic data, litter input quality (broadleaves versus conifers) and quantity.~~

In order to test the performance of Yasso07 in estimating soil carbon changes at the RENECOFOR sites, we analyzed the residuals of carbon changes, here defined as the~~i.e.~~ difference between the simulated and observed values, using analysis of variance (ANOVA). The following environmental and biological factors were tested: site geographical location (latitude, longitude, and altitude), climatic conditions (temperature and precipitation), soil types, tree functional type and tree species. Before each ANOVA, we tested the normality of data using a Shapiro – Wilk test. For the sensitivity analyses, we performed loess regressions (Fox and Weisberg, 2011) to characterize the variation of soil carbon stock as a function of initial soil carbon stock settings and simulation length (1 – 10000 years). Statistical analyses were performed using R 2.13.0 (R Core Team, 2013).

3 Results

3.1 Litter carbon quality of northern temperate tree species

Our meta-analysis (Fig. 2) showed that the litter carbon quality, i.e., carbon composition, of northern temperate tree species significantly differed between tree organs. For woody litters (only using stem data) the percentage of A carbon pool attained up to 80% of the total carbon pool; the sum of A and N carbon pools corresponded to at least > 75% and, in most cases, >90%, with consequently only small percentages of W and E (Fig. 2a). Nevertheless, this dominance of A and N over W and E was much less pronounced in foliage and root litters (Figs. 2b and 2c). Generally, the different tree organs can be ranked according to the sum of the proportions of A and N as follows: wood (>90%) > roots (70 – 80%) > foliage (60 – 70%, Fig. 2d).

The effect of tree functional type on litter carbon quality strongly interacted with that of tree organs. For wood, broadleaves and conifers had clearly shifted point clouds for the relationship between A and N carbon pools: greater proportion of A, but lower proportion of N in broadleaves compared to those in conifers. In foliage and root litter, the effect of tree functional type on proportions of A and W was less pronounced than in wood. The main difference between broadleaves and conifers occurred in N rather than in A (Fig. 2d). Broadleaved litter had lower proportion of N than coniferous litter regardless of tree organ (Fig. 2d). The proportions of A and N relative to those of E and W were quite stable between broadleaves and conifers regardless of tree organs (Fig. 2d).

3.2 Sensitivity analyses on the impact of initial soil and litter settings on model output

Fig. S3 showed the impact of different settings of litter and carbon quantity and quality on model fit over the RENECOFOR sites. For soil carbon quality, the partial steady-state assumption (Fig. S3c and S3d) achieved significant better model fits (with lower model root-mean-square-error) than the complete steady-state assumption (Fig. S3a and S3b). Then, we found that model fits were better when using observed CS until 0.4 m as initial carbon quantity than that with CS until 1.0 m (Fig. S3a and S3c). Nevertheless, the choice of the observed CS until 1.0 m at the first assessment as model input is more advantageous, because Yasso07 predicts CS down to 1.0 m depth due to its used datasets for model calibration (Rantaraki et al., 2012).

Different choices of fine root:foliage ratio for fine root litter input also significantly influenced Yasso07's performance in predicting soil C changes (Fig. S3). Ratios of 0.1 – 0.8

for broadleaves and 1.8 – 3.0 for conifers achieved the best fits between simulated and observed soil CS changes according to different scenarios (Fig. S3). Using a constant value of 1.0 for both broadleaved and coniferous sites seems to be an acceptable compromise between both tree functional types, although such a choice is not optimal for each single functional type.

Based on the above diagnoses, only fit and residual analysis results based on the simulations with partial steady-state assumption, the observed CS until 1.0 m and fine root: shoot ratio of 1.0 (Fig. S3d) were shown in the Sect.3.3.

Fig. S4 visualized all the theoretically possible final carbon stocks by varying initial carbon stocks and simulation length (from 1 to 10 000 years). The initial soil carbon quality had a pronounced impact on the final soil organic carbon stocks at annual and decennial scales. For example, when the initial proportion of A pool increased from 0 to 80%, the final proportion of A could increase by +30 to +40 tC ha^{-1} (Fig. S4a) and the final total carbon stock could decrease by c.a. -20 to -30 tC ha^{-1}(Fig. S4u) at annual and decennial scales. When simulations were performed over millennium timescale, the initial soil carbon quality did not impact the final soil carbon quality anymore. In other words, the same final soil carbon quality was obtained regardless what the initial soil quality was (Fig. S4).

3.3 Simulated versus observed carbon data

~~The choice of fine root:foliage ratio significantly influenced Yasso07's performance in predicting soil C changes (Fig.). Based on the criteria of minimum root mean square error (RMSE), the ideal ratio for conifers appeared between 1.8 and 2.2, while the ideal ratio for broadleaves was the smallest ratio tested (0.1).~~

Using only mean litter input, the theoretical carbon stock ($CS_{steady\text{-}state}$) simulated from the initialization method and the observed $CS_{obs,t1}$ to 1 m depth shared the same order of magnitude and were even comparable (Fig. S5). However, the carbon stock were overestimated for most coniferous stands, and underestimated for broadleaved stands (Fig. S5).

When simulated annual carbon stock changes (ACC) were plotted against observed ones, the point clouds were distributed around the 1:1 diagonal line despite fairly high dispersion (Fig. 3~~4~~). The correlation between predicted and measured ACC remained weak ($R^2 < 0.1$). The mean observed and simulated annual carbon stock changes (ACC) of all sites are +0.34 ± 0.06 tC ha^{-1} year^{-1} (+0.20 ± 0.06 tC ha^{-1} year^{-1} for broadleaved stands and +0.48 ± 0.10 tC ha^{-1} year^{-1} for coniferous stands) and +0.~~45~~00± 0.~~09~~ 07 tC ha^{-1} year^{-1} (+0.~~96~~ 28 ± 0.~~10~~ 09 tC ha^{-1}

$year^{-1}$ for broadleaved stands and -0.~~05~~ 28 ± 0.11~~0~~ tC ha^{-1} $year^{-1}$ for coniferous stands), respectively. ~~48~~32% of ~~coniferous~~ broadleaved stands and 39% of coniferous stands showed significant differences between observed and simulated ACC (Fig. 3a). In only c.a. ~~25~~17% of the sites, ACC were significantly different from 0 for both simulated and observed results (i.e. the case 3 in Fig. 3b). There is a significant effect of the tree functional type on the observed and simulated values. The model tended to overestimate ACC in broadleaved stands but to underestimate ACC in coniferous stands. The quantity of sites in which estimates and observed carbon stock changes share the same tendency (i.e. data points in the zone I, IV, III and VI, Fig. 3~~4~~) was approximately two thirds of the total sites. c.a. one third of sites are in the remaining zones (II, and V) where the predicted tendency was contrary to the observed tendency. From the residual distribution, we could also find that model fit with carbon quality set by partial steady-state assumption (Fig. 3) was better than that set by complete steady-state assumption (Fig. S6).

The simulated carbon stock changes exhibited a negative linear relationship with the initial soil carbon stock (Fig. 4b), whereas this tendency was not observed for the observed carbon stock changes (Fig. 4a). Storm damage and soil type could not provide clear tendencies in explaining the residuals. Only for coniferous stands, residuals showed significantly differences among the three major types of soil (*n* of sites >5): cambisol > luvisol > podzol (Fig. S7). Tree ages in coniferous stands tend to be smaller than those in broadleaved stands. When considering both tree functional types and tree ages, neither the latter nor their interaction had a significant effect on residuals. With all sites together, residuals become higher with increasing latitude, indicating that simulated ACC was more overestimated in northern zones (ANCOVA, F = ~~14.9~~11.2, P<0.001). This pattern was particularly strong for broadleaved stands~~, with the exception of several ones in Pyrenees Mountains~~ (Fig. S8a). Yet, this tendency was not clear for coniferous stands (Fig. S8e). ~~Identical~~ Both residual signs ~~is~~ were generally present ~~in clusters~~for ~~in~~ all of the main species (Fig. S8b, S8c, S8d, S8f, S8g and S8h). Broadleaved and coniferous stands differed in their responses to environmental factors: for coniferous stands, both temperature and precipitation had ~~no~~ little effect on residuals~~ (Fig. Sa)~~, whilst for broadleaves, precipitation was negatively correlated with residuals (ANCOVA, F = ~~7~~10.~~17~~8, P<0.001~~, Fig. Sb~~).

Regarding soil physical and chemical properties, total nitrogen stock soil were significantly correlated with residuals for both broadleaved and coniferous stands (Fig. 5). Then, soil texture (proportions of clay and sand) and exchangeable magnesium and potassium were significantly correlated with residuals only for broadleaved stands (Figs. 5 and S9~~;~~ Table S2).

The remaining tested variables, such as exchangeable aluminum and calcium, ~~proportion of silt,~~ pH, total phosphorus and carbon:nitrogen ratio, had no relationship with the residuals~~, except for exchangeable aluminum, which showed a weak correlation with ACC residuals (*P*<0.05*) only for coniferous stands~~ (Table S2).

~~3.3 Effect of litter carbon quality on model prediction (Sensitivity analyses 2.6.1)~~

~~Variation of litter carbon quality (without distinction of original organ) altered the carbon quality at steady-state distribution of soil carbon pools (Fig.). The carbon belonging toproportion of soil A, W and E carbon pools remained below 15% regardless , whatever the biochemistry of litter inputs. The percentages of soil N and H pools were more susceptible to the variation of litter carbon quality than the more labile ones (e.g., A, W and E) (Fig. 6). The size of soil N and H always varied between 25% and 65% of, whenever the pools in litter varied from 5% to 80% (Fig. 6).~~

~~The strong sensitivity of the carbon steady state distribution to litter carbon quality was *de facto* greatly discounted in reality, because the variation in chemical composition of tree species was very limited (Fig.). Using average compositions of broadleaves and conifers species, we found that, at the steady- state, the H pool contains 30 – 40% of soil carbon, the N pool 45 to 55 %, the A pool <5% and W and E pools <2% (Fig. 7). Broadleaves dominated sites differed from conifers dominated sites with a slightly lower percentage N-carbon in the steady-state soil carbon stock, but a higher percentage of H-carbon (Fig. 7).~~

~~3.4 Impact of initial condition of soil carbon stock on model prediction (Sensitivity analyses 2.6.2)~~

~~Fig. visualized all the theoretically possible final carbon stocks by varying initial carbon stocks and simulation length (from 1 to 10 000 years). The initial soil carbon quality had a pronounced impact on the final soil organic carbon stocks at annual and decennial scales. For example, when the initial proportion of A pool increased from 0 to 80%, the final proportion of A could increase by +30 to +40 tC ha^{-1} (Fig. a) and the final total carbon stock could decrease by c.a. -20 to -30 tC ha^{-1}(Fig. u) at annual and decennial scales. When simulations were performed over millennium timescale, the initial soil carbon quality did not impact the final soil carbon quality anymore. In other words, the same final soil carbon quality was obtained regardless what the initial soil quality was (Fig.).~~

4 Discussion

4.1 Agreement between simulated and observed annual soil carbon stock changes

Testing widely popularized soil carbon models using large dataset is highly meaningful work that enables not only assessing the model's ability over various climatic and ecosystem types, but also providing lessons and implications for future modelling work. Here, based on the observed carbon stock data to 1.0 m soil depth from the RENECOFOR network, ~~,~~ we found the simulated and observed carbon stocks ($CS_{steady\text{-}state}$ versus $CS_{obs,\ t1}$) to 1.0 m showed the same order of magnitude, validating Yasso07's good capability to predict carbon stock in average at the scale of the French territory. Such good performance at the national scale is consistent with Yasso's aim for generality and supported by previous studies (see Ortiz et al. 2013; Lehtonen et al. 2016; Hernández et al. 2017). Nevertheless, the observed CS until 1.0 m at t1 exceeded already $CS_{steady\text{-}state}$ for most coniferous stands (Fig. 5S), suggesting, to some extent, some inadaptability of the model parameters to the RENECOFOR dataset. Such inadaptability may simply be due to the setting of an over-high decomposition rate of the slow carbon pools in the model. Or, as the coniferous stands are on average younger and were afforested more recently than the broadleaved stands (Jonard et al., 2017), the model does not account for such landuse change history to calculate the SOC stock at steady state. Fig. S5 also showed that for most broadleaved stands, observed stocks are lower than their $CS_{steady\text{-}state}$, forming the evidence that that steady-state equilibrium may have not yet been reached at these sites.

Then, based on the observed annual soil carbon stock changes (ACC) with average 15-year interval between the two inventories, we found the simulated ACC ~~using Yasso07~~ were significantly biased for more than one third of the French RENECOFOR sites. Particularly, Yasso07 generally overestimated the ACC at the broadleaved stands located in the north of France (Fig. S8~~6~~a-d) and the overestimation can be exacerbated with lower precipitation. Yasso07 tended to underestimate the ACC in our coniferous stands. Nevertheless, we would expect slightly better performance of Yasso07 in coniferous stands than in broadleaved ones, since the model's estimates have shown good correspondence to measurements (of stocks and/or changes) in coniferous forests, especially the Nordic boreal ones (e.g., Karhu et al., 2011; Ortiz et al., 2013). Probably due to the younger age of the coniferous stands, observed ACC of the coniferous stands were greater than those of the broadleaved stands (Fig. 3, Jonard et al;, 2017). Again, Yasso07 was unable to reproduce this observed effect of tree functional type on ACC, as it lacks consideration of landuse change history, i.e., the same reason with the case of steady-state carbon stock mentioned above.

Except for tree functional type and geographical location (e.g. latitude, which is correlated with climatic variables), qualitative ecological variables that are assumed as key factors influencing carbon sequestration processes, e.g. soil type (except for coniferous stands), storm damage and stand age range, showed limited tendencies in explaining residuals. Note that those factors were not fully crossed in the 101 sites, rendering testing each signer factor difficult.

The simulated ACC showed strongly negative correlation with the observed initial soil carbon stock ($CS_{obs,t1}$) , with an overestimation of ACC at sites of lower CS $_{obs,t1}$ and an underestimation at sites of higher CS $_{obs,t1}$ (Figs. 4 and ~~S7~~S9). Such phenomenon can be logically explained by the model's mechanism. With increasing initial carbon stock, ~~due to the fairly stable steady-state carbon quality (Fig. 6),~~ there is an increase in the quantity of those easily decomposable compounds, i.e. A, W and E, in soil, which triggers a more substantial mass loss at a decennial scale. However, the observed data on carbon stock changes did not support this trend. ~~, suggesting that Yasso07's configuration tends to penalize too much the loss of labile carbon at decennial scale. Compared to broadleaved stands, the slightly steeper slope for coniferous stands in Fig. 4b might be attributed to their higher steady-state proportion of the extremely labile pools (A, W and E) in soil at a given soil carbon stock (Fig. 6a) due to the higher proportion of A, W and E pools in the litter quality of broadleaves (Fig.2).~~

Several quantitative soil physical and chemical properties showed clear correlations (especially for broadleaved stands) with ACC residuals (Fig. 5). Also, in the principle component analyses (Fig. ~~S7~~S9), the arrows standing for soil variables are ~~generally~~ slightly closer to the pivoting axis of "initial carbon stock – ACC residuals" than those standing for climatic and geographic variables, notably for broadleaved stands. ~~The correlations (Table S2 and Fig. S7) may indicate that texture and nitrogen content contribute to lower ACC for broadleaved stands compared to model predictions and that aluminum and perhaps also pH (Fig.S7) could be involved in the mechanisms that allow increasing microbial activities and carbon mineralization in soils of coniferous stands compared to model predictions. All t~~These results suggest a potential interest of incorporating soil properties into new versions of Yasso model family, in which soil parameters are lacking or only implicitly incorporated. Indeed, there are numerous evidences that soil physical and chemical properties can greatly govern soil carbon dynamics and stock capacity (Beare et al., 2014; Dignac et al., 2017; Rasmussen et al., 2018).

The limitations of the model at the site-scale are not surprising as the model was developed for primarily large-scale application integrating processes that dominate at the site scale. Despite Yasso07's significant prediction bias at a number of sites, it is unreasonable to simply attribute the bias to the model *per se*, as multiple uncertainties affecting the quality of the model's input data can be identified (see Sects. 4.2 – 4.3~~4~~). These uncertainties can occur not only with Yasso07, but also with other prevailing models one may choose, highlighting large knowledge gaps in ecology and soil carbon modelling.

4.2 Setting soil carbon quality for model initialization: a recurrent challenge in soil carbon modelling

A great uncertainty is associated with the model initialization of soil carbon quality, as it was not measured, but usually estimated, for example, ~~obtained~~ by matrix inversion with the assumption that the litter input has been the same for decades. Compared to total soil carbon stock, measuring soil carbon quality is much labour-intensive and time-consuming. Moreover, data of soil carbon quality from different sources are partly or totally incompatible due to the use of different chemical pools or protocols of fractionation (Blair et al., 1995). Therefore, measured data of soil carbon quality are generally lacking at worldwide scale. Such lack of information is a recurrent issue for soil carbon dynamics modeling (see Elliot et al. (1996), who has discussed the issue of "Measuring the modelable"). Many prevailing soil carbon models require setting carbon quality besides carbon quantity, e.g., Romul (Chertov et al., 2001), RothC (Coleman and Jenkinson, 1996), CENTURY versions Parton et al., 1987; Metherell et al., 1993, CBM-CFS3 (Kurz et al., 2009). Inappropriate setting of carbon quality in models may greatly change carbon stock predicts (Wutzler and Reichstein, 2007; Carvalhais et al., 2008; 2010).

In the present study, soil carbon quality data were unavailable at the French RENECOFOR sites. ~~As a result, w~~We ~~used~~ tested both complete and partial steady-state assumptions to set the initial ~~the simulate~~ carbon quality~~ at steady-state to feed Yasso07~~. Compared to the complete steady-state assumption, the partial steady-state assumption allows that slow cycling pools can be still accumulating carbon while ~~more labile~~fast cycling pools are in equilibrium (Wutzler and Reichstein, 2007). ~~This is a strong, but widely adopted hypothesis in soil carbon modelling work (Foereid et al., 2012). Following the general idea of the partial steady-state assumption, we~~ In this study, we did not use the exact method to ~~relaxed equilibrium~~

~~assumption~~ estimate initial carbon quality ~~for simulations~~as proposed in Wutzler and Reichstein (2007) due to the lack of information for setting the modified the decomposition-accumulation dynamics of H pool~~ required by the assumption~~. ~~However~~Nevertheless, following the same idea of partial steady-state assumption, we revised the proportions of N and H pools by assuming that A, W and E pools are in equilibrium and equal to the simulated steady-state ones and that the sum of all pools at t1 is constant to observed stock. We found that our partial steady-state assumption gave rise to generally ~~significant~~ better model fits than the complete one (Fig. S3; see also Figs. 3 and S6), hinting its good suitability to the RENECOFOR sites. When plotting $CS_{stead\text{-}state}$ against CS_{obs} (Fig. S5), we visualized the discrepancy that, while CS_{obs} of most of broadleaved stands were smaller than $CS_{stead\text{-}state}$, CS_{obs} of most of coniferous stands were greater than $CS_{stead\text{-}state}$. Such a discrepancy was then brought into ACC fit when the complete steady-state assumption was adopted (Fig. S6). Nevertheless, the partial steady-steate assumption can, to some extent, mitigate such discrepancies: for broadleaved stands, the revised proportions of A+W+E pools became higher than those at complete steady-state (Fig. 6; with 70% of stands above the the steady-state strip), thus reducing the model's overestimation of ACC; for coniferous sites, the proportions of A+W+E pools are often compressed (Fig. 6; with >50% of stands below the steady-state strip), reducing the model's underestimation of ACC at steady-state. ~~, f~~For future work, it would be definitely worthwhile to have both assumptions compared using prevailing carbon models (e.g., Yasso07, RothC, Century etc.), as studies comparing initialization assumptions still remain scanty compared to those on model comparisons.

In order to gain a global overview on Yasso07's sensitivity to initial soil carbon quality, here we equally conducted a sensitivity analysis that computed the final soil carbon stocks using all the possible combinations of the composition of chemical pools. This sensitivity analysis confirmed the high influence of initial soil carbon quality on soil carbon stock estimates (Fig. S4), notably at short temporal scales (i.e., yearly and decennial). This result is in line with the previous carbon stock modelling studies (Parton et al., 1993; Kelly et al., 1997; Smith et al., 2009; Foereid et al., 2012), confirming that it is a crucial step ~~a general problem~~ for all of the chemical pool based carbon models. Besides this consensus, our sensitivity analysis further showed that such effect of initial composition carbon stocks will gradually vanish with increasing length of simulation and especially when the length is up to several centuries or millenniums. Our analysis provides new insights on the sensitivity of model estimated carbon stocks to the method and assumptions used in model initialization. Such analysis can be transplanted to the other carbon models to test their theoretical performance and robustness of each model at different temporal scales and also, to compare models.

Finally, solely testing different initialization assumptions or performing sensitivity analysis does not allow radically solving the prediction issue related to uncertainties of soil carbon quality. Based on ground truth data, Balesdent et al. (2018) showed that carbon age shows strong patterns as a function of soil depth and ecosystem type. It appears highly necessary for future modelling work to capture better indicators for carbon stabilization mechanisms~~consider such specific or generic patterns, as shown in Balesdent et al. (2018)~~, into the procedure of model initialization. For this, it is to be noted that Yasso07's particular model configuration, i.e. the use of measurable chemical pools, may open the possibility of using measured data of soil carbon quality for model initialization instead of ~~simulated~~ steady-state ~~ones~~assumptions. Future measurements on soil carbon radiocarbon age of the RENECOFOR sites may offer an ideal opportunity to compare the impact of the two sources of soil carbon quality on Yasso07's predictions.

4.3 A precise estimation of root litter quantity helps improve Yasso07 prediction

An important source of uncertainty in the estimates of litter quantity at the RENECOFOR sites was the fine root litter input. Many studies have revealed that fine roots act as a major source contributing to total litter quantity due to their fast turnover rates (Brunner et al., 2013; Kögel-Knabner et al., 2002; Berg and McClaugherty, 2008). In some forest ecosystems, the proportion of fine root litter is even comparable to that of foliage (Freschet et al., 2013; Xia et al. 2015). However, estimating fine root litter inputs is, again, a time-consuming and challenging task. Due to this reason, so far rarely have national wide forest inventory projects ever incorporated direct measurement of the dynamics of fine root litter input (i.e. the case of RENECOFOR network). Fine root turn-overs of forest species are variable depending on climate, tree species and management scenarios (Kögel-Knabner et al., 2002; Litton et al., 2003; Mokany et al., 2006), rending the choice of model input values highly subjective and difficult. By testing variable fine root:foliage ratios of litter input, we observed a significant shift in the predicted carbon stock changes by Yasso07 (Fig. S2~~1~~). This finding not only highlights the importance of precisely quantification of fine root litter input, but also suggests that broadleaves and conifers may have separated quantification of fine root litter input with regard to that of foliage, although here we chose the same ratio for both broadleaved and coniferous stands. We also noted that using one ratio per tree functional type (conifers versus broadleaves) could only change the overall prediction baseline, but cannot reduce the data dispersion. Consequently, it is of great interest to estimate root litter input quantity at species level on the basis of direct measurement and then couple specific data with Yasso07.

Another potentially important litter inputs may come from the understory shrubby and herbaceous species, which were not taken into account in this study due to data unavailability. Herb and shrub layer are typically not estimated in forest inventories but they can contribute significantly to the annual litter production in forests (eg. de Wit et al. 2006, Gilliam 2007, Lehtonen et al. 2016). Muukkonen and Mäkipää (2006) estimated that the carbon inputs from herb and shrub vegetation in Finnish forests were in the range of 0.50 to 0.66 tC ha^{-1} $year^{-1}$. Such value is apparently high, as it attains 12% - 23% of the mean total tree litter inputs of all the RENECOFOR sites (Table 1). This is in line with the preliminary data from Etzold et al. (2014), who suggested that understory vegetation contributed c.a. 12% (0.1 to 36.8%) to the total observed annual C turnover at six sites of the Long-term Forest Ecosystem Research Programme LWF (ICP-Level II plots).

Also, Yasso07's parameter set was calibrated using one of the richest litterbag datasets in the world in terms of number of observation. The state-of-the-art of soil carbon modeling is based on the litter input and decomposition processes as the driving forces in soil carbon accumulation where measured mass loss of litter is used to fit model parameters. Our knowledge on the importance of other sources of biological carbon input, e.g. soil fauna and rhizodepostion, as well as how to take them into account in modelling processes still remains poor. Accordingly, whether and to which extent the bias of Yasso07 is related to these alternative sources of biological carbon input is unknown.

~~4.4 Limited but potentially strong effect of litter carbon quality on Yasso07 prediction~~

~~Litter carbon quality, especially the content of litter carbon in the N carbon pool, controls the bulk litter decomposition rate and this has been well-known (De Deyn et al., 2008). Indeed, the meta-analysis (Fig. 3) confirmed the significant disparity of carbon allocation between litters of broadleaves and conifers in all the investigated organs. However, little has been known about how this disparity of litter carbon quality between broadleaved and coniferous stands will be projected into the long-term prediction of soil carbon stock. Our sensitivity analysis Module I (Sect. 2.6.1) with Yasso07 showed a generally limited impact of such disparity on the soil carbon quality of steady-state (Fig. 6). Litter carbon quality seems to be a less important factor determining the model predictions via affecting soil stock initialization. This is especially true for the three more labile carbon pools (i.e. A, W and E) and their mean residence time has quite low disparity between themselves (Fig. 1). This seems to more or~~

~~less weaken the meaningfulness of splitting litter and soil labile carbon compounds into the three carbon pools (A, W and E) in Yasso07.~~

4.4~~5~~ Suggestions for model improvement in the future

First of all, we found the model structure and algorithm good, clear and simple to operate and this goes along well with the positive remarks toward Yasso and Yasso07 in literature (Rantakari et al., 2012; Didion et al., 2014; Lu et al., 2015; Wu et al., 2015). Fig. S1 only showed the mass flows that are statistically significant for the case of using the Tuomi 2011 parameter set. Yasso07 keeps all the theoretical mass flow possibilities in the $\boldsymbol{A_p}$ matrix in (Eq. 1b). However, a mass flow parameter with a statistical significance does not signify that it is biologically meaningful. For this we can quote the flow N → A of the model (Fig. S1), for which the modeler had assigned an astonishingly high percentage: $p_{N \to A}$ = 83%. This quantity is disputable in the angle of soil biochemistry, because as lignin, i.e. the major component constituting the N carbon pool, likely does not turn into the A pool, but would condense with other nearby phenol, peptides or saccharides (Burns et al., 2013).

As a model aiming at predicting soil carbon dynamics, Yasso07 is still highly simple in the description of soil variables that are known to strongly impact decomposition processes in non organic-soil.~~,~~ For example, the effect of soil mineralogy or aggregation ~~have~~ has not been considered in Yasso07 yet. Indeed, the model was often applied on soils fairly rich in organic matter (e.g., Karhu et al., 2011), where the consideration of soil mineral properties was not particularly relevant, and where the authors' assumption that litter quantity is a good proxy for soil properties was reasonable. In addition, when Yasso, i.e., Yasso07's prototype, came up in 2005 (Liski et al., 2005), information on mineral soil properties in the various forest soil horizons was not commonly available, but nowadays it is easier to obtain it, although there is still a lack of such detailed data for consistent application across large regions or at the national scale (Didion et al., 2016).

5 Conclusions

We tested the performance of the soil carbon model Yasso07 using the decennial scale French national wide forest data thank to the RENECOFOR network, as well as a meta-analysis database for litter carbon quality and sensitivity analyses to characterize the effect of inputs of initial litter and soil carbon quality on the model's predicts. We showed that while the model's predicts of the carbon stock until 1.0 m soil depth and annual soil organic carbon changes (ACC) stay within the same order of magnitude with the observed ones, ~~correlation~~ accordance between the observed and simulated ACC at the site scale remained weak. There was a bias of model prediction for the carbon change tendency at more than one third of the French sites. The performance of Yasso07, as well as the other soil carbon models, should be examined before their application for management guidelines and policy-making for forest ecosystems at any study scales.

Such bias can be attributed to multiple reasons concerning model input, such as (i) large uncertainty in the measured soil carbon stock and changes; (ii) lack of information on initial soil carbon quality at the site level and (iii) lack of information on below ground litter production. ~~For the latter two aspects, their importance was explicitly confirmed by our sensitivity analyses.~~ These reasons are valid for the whole state-of-the-art of soil carbon modelling, regardless of the model that one uses. For the latter two aspects, their importance was explicitly confirmed by our sensitivity analyses. Setting soil carbon quality should be oene of the most crucial step influencing the model's fit. To set soil carbon quality, we found that partial soil steady-state assumption gives rise to significant better model fit than the complete steady-state assumption. Some of the model's parameters governing the transfer among soil pools are statistically derived but not directly measured, and thus may poorly represent the real biochemical processes of decomposition. Residual analysis also suggests a potentially important role of soil physical and chemical properties in explaining the model's prediction.

These findings allow us to provide a series of suggestions to modelers, users and policy makers:

- To Yasso07 modelers, we suggest keeping the current model structure, algorithm and parameter natures, but incorporating more refined some biochemical processes, including (i) revising certain mass flows to achieve both statistically and biologically meaningful process (especially the N→ A flow) (ii) refining decomposition process (i.e., the residence times between the A, W and E soil carbon pools) and possibly, (iii) explicitly incorporating easy-measured soil parameters to better represent biophysical and biochemical interactions in soil carbon cycling.

- To Yasso07 users, we suggest working in conjunction with modelers in order to better reduce the uncertainties in ~~both~~ model initialization of soil carbon stock. We also suggest ~~using~~ measuring~~ement~~ ~~based~~ forest ~~litter input~~carbon quality and quantity, and also ~~especially the~~ belowground fine root litter data to better feed the model.
- To policy makers, we suggest keeping prudent toward diagnosis from based on a single carbon model, especially when long term trend is predicted. Predictions from multiple models served as a cross-validation procedure are preconized for both global and local scales areas.

Our decennial observation sites spreading at a large spatial scale that covers different ecosystems can facilitate and provide good opportunities for future calibration, improvement, and re-assessment of the model. Finally, taking Yasso07 as an example, this work highlighted the bottleneck of soil carbon modelling due to lacking knowledge or data on soil and litter carbon quality and fine root litter quantity, rendering high uncertainties for model inputs, and also demonstrated. Simultaneously, this study demonstrated methodologies of testing the other soil carbon models via sensitivity analyses, which enable us to better understand the limits of the model and of data input for future improvements in soil organic carbon modelling. In this study, we used the published model structure and parameters from Tuomi et al. (2011a) ~~developed for Scandinavia~~without any modifications. Upcoming work of sensitivity analyses incorporating modifications of both the settings of carbon quality and litter inputs and Yasso07's configuration and parameters should be performed to ultimately confirm the reliability of the current diagnoses.

Acknowledgement

This study was funded by the French Agence de l'Environnement et de la Maîtrise de l'Energie (ADEME, Contract ref. : 14-60-C0082). The UR1138 BEF and this study was supported by a grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (ANR-11-LABX-0002-01, Lab of Excellence ARBRE) – QLSPIMS project. This study is an outcome of a project under task "Input to improving the comparability in MRV across EU MS" within the LULUCF MRV project: "Analysis of and proposals for enhancing Monitoring, Reporting and Verification (MRV) of land use, land use change and forestry (LULUCF) in the EU" funded by the European Commission and funding for M. Didion by the Swiss Federal Office for the Environment. We thank several French colleagues Dr. I.~~sabelle~~ Feix (ADEME), Dr. A.~~rnaud~~ Legout (INRA) and Dr. B.~~ertrand~~ Guenet (CNRS) for their valuable comments ~~toward~~ to this work. We are also grateful to ~~two finish colleagues~~ Dr. A.~~nna~~ Repo (FEI - SYKE) and Dr. E.~~mmi~~ Hilasvuori (FEI - SYKE) for their ~~patient~~ explanations ~~concerning the~~of Yasso07~~ model parameters and carbon fraction protocols~~. ~~Finally, we thank Dr. Hélène Vogt-Schilb (WOAINI) for her help in R coding for the spatial visualization of data points.~~

References

Aber, J. D., Melillo, J. M., and McClaugherty, C. A.: Predicting long-term patterns of mass loss, nitrogen dynamics, and soil organic matter formation from initial fine litter chemistry in temperate forest ecosystems, Canadian Journal of Botany, 68(10), 2201–2208, doi.org/10.1139/b90-287, 1990.

Aulen, M., Shipley, B., and Bradley, R.: Prediction of in situ root decomposition rates in an interspecific context from chemical and morphological traits, Annals of botany, mcr259, doi.org/10.1093/aob/mcr259, 2011.

Balesdent, J., Basile-Doelsch, I., Chadoeuf, J., Cornu, S., Derrien, D., Fekiacova, Z. and Hatté, C.: Atmosphere–soil carbon transfer as a function of soil depth. 559, 599–602, ~~Carbon transfer from atmosphere to deep soil layers over the last 50 years~~, doi: 10.1038/s41586-018-0328-3. Nature, ~~revision 2,~~ 2018.

Beare, M., McNeill, S., Curtin, D., Parfitt, R., Jones, H., Dodd, M., and Sharp. J: Estimating the organic carbon stabilisation capacity and saturation deficit of soils: a New Zealand case study. Biogeochemistry, Springer Science + Business Media, doi: 10.1007/s10533-014-9982-1. 2014

Berg, B. and McClaugherty, C. Plant litter: decomposition, humus formation, carbon sequestration, Second edition. Springer-Verlag Heidelberg Berlin. doi.org/10.5860/choice.51-6172, 2008.

Brunner, I., Bakker, M. R., Björk, R. G., Hirano, Y., Lukac, M., Aranda, X., Børja, I., Eldhuset, T. D., Helmisaari, H. S., Jourdan, C., Konôpka, B., López, B. C., Miguel Pérez, C., Persson, H. and Ostonen, I.: Fine-root turnover rates of European forests revisited: an analysis of data from sequential coring and ingrowth cores, Plant and Soil, 362(1–2), 357–372. doi.org/10.1007/s11104-012-1313-5, 2013.

Burns, R. G., DeForest, J. L., Marxsen, J., Sinsabaugh, R. L., Stromberger, M. E., Wallenstein, M. D., Weintraub, M. N. and Zoppini, A.: Soil enzymes in a changing environment: current knowledge and future directions, Soil Biology & Biochemistry 58, 216–234. doi.org/10.1016/j.soilbio.2012.11.009, 2013.

Carvalhais, N., Reichstein, M., Seixas, J., Collatz, G. J., Pereira, J. S., Berbigier, P., Carrara, A., Granier, A., Montagnani, L., Papale, D. and Rambal, S.: Implications of the carbon cycle steady state assumption for biogeochemical modeling performance and inverse parameter retrieval, Global Biogeochemical Cycles, 22, GB2007. doi.org/10.1029/2007gb003033, 2008.

Carvalhais, N., Reichstein, M., Ciais, P., Collatz, G.J., Mahecha, M.D., Montagnani, L., Papale, D., Rambal, S. and Seixas, J.: Identification of vegetation and soil carbon pools out of equilibrium in a process model via eddy covariance and biometric constraints, Global Change Biology, 16, 2813–2829. doi.org/10.1111/j.1365-2486.2010.02173.x, 2010.

Chertov, O. G., Komarov, A. S., Nadporozhskaya, M., Bykhovets, S. S. and Zudin, S. L., ROMUL – a model of forest soil or- ganic matter dynamics as a substantial tool for forest ecosystem modeling, Ecol. Model, 138, 289–308, doi.org/10.1016/s0304-3800(00)00409-9, 2001.

Coleman, K., Jenkinson, D.S., RothC-26.3 – A Model for the turnover of carbon in soil. In: Powlson, D.S., Smith, P., Smith, J.U. (Eds.), Evaluation of Soil organic matter models, Using Existing Long-Term Datasets. Springer-Verlag, Heidelberg, pp. 237–246, doi.org/10.1007/978-3-642-61094-3_17, 1996.

De Deyn, G. B., Cornelissen, J. H., & Bardgett, R. D.: Plant functional traits and soil carbon sequestration in contrasting biomes. Ecology letters, 11(5), 516–531. doi.org/10.1111/j.1461-0248.2008.01164.x, 2008

Didion, M., B. Frey, N. Rogiers, and E. Thürig. : Validating tree litter decomposition in the Yasso07 carbon model. Ecological Modelling, 291, 58–68, doi.org/10.1016/j.ecolmodel.2014.07.028, 2014.

Didion, M., Blujdea, V., Grassi, G., Hernández, L., Jandl, R., Kriiska, K., Lehtonen, A. and, Saint-André, L.: Models for reporting forest litter and soil C pools in national greenhouse gas inventories: methodological considerations and requirements, Carbon Management, 1–14, doi.org/10.1080/17583004.2016.1166457, 2016.

Dignac, M. F., Derrien, D., Barré, P., Barot, S., Cécillon, L., Chenu, C., Chevallier, T., Freschet, G.T., Garnier, P., Guenet, B. and Hedde, M: Increasing soil carbon storage: mechanisms, effects of agricultural practices and proxies. A review. Agronomy for sustainable development, 37(2), 14. doi.org/10.1007/s13593-017-0421-2, 2017.

Etzold, S., Helfenstein, J., Thimonier, A., Schmitt, M. and Waldner, P.: Final Report: The role of the forest understory within the forest nutrient and carbon cycle of LWF sites, Birmensdorf: Swiss Federal Research Institute for Forest, Snow and Landscape Research, 2014.

Fox, J., Weisberg, S.: An R companion to applied regression, Sage Publications, 2011.

Freschet, G. T., Cornwell, W. K., Wardle, D. A., Elumeeva, T. G., Liu, W., Jackson, B. G., ... & Cornelissen, J. H.: Linking litter decomposition of above-and below-ground organs to plant–soil feedbacks worldwide, Journal of Ecology, 101(4), 943–952, doi.org/10.1111/1365-2745.12092, 2013.

Guo, D. L., Mitchell, R. J., and Hendricks, J. J.: Fine root branch orders respond differentially to carbon source-sink manipulations in a longleaf pine forest, Oecologia, 140(3), 450–457, doi.org/10.1007/s00442-004-1596-1, 2004.

Hernández, L., R. Jandl, V. N. B. Blujdea, A. Lehtonen, K. Kriiska, I. Alberdi, V. Adermann, I. Cañellas, G. Marin, D. Moreno-Fernández, I. Ostonen, M. Varik, and M. Didion.: Towards complete and harmonized assessment of soil carbon stocks and balance in forests: The ability of the Yasso07 model across a wide gradient of climatic and forest conditions in Europe. Science of The Total Environment 599–600:1171-1180. doi.org/10.1016/j.scitotenv.2017.03.298, 2017.

IPCC: Use of Models and Facility-Level Data in Greenhouse Gas Inventories (Report of IPCC Expert Meeting on Use of Models and Measurements in Greenhouse Gas Inventories 9-11 August 2010, Sydney, Australia). Institute for Global Environmental Strategies (IGES), Hayama, Japan, 2011.

IPCC: Climate Change 2014: Mitigation of Climate Change. *Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

Jandl, R., Rodeghiero, M., Martinez, C., Cotrufo, M.F., Bampa, F., van Wesemael, B., Harrison, R.B., Guerrini, I.A., Richter, D.D., Rustad, L., Lorenz, K., Chabbi, A., Miglietta, F., 2014. Current status, uncertainty and future needs in soil organic carbon monitoring. Sci. Total Environ. 468, 376–383.

Jonard, M., Nicolas, M., Coomes, D. A., Caignet, I., Saenger, A., and Ponette, Q.: Forest soils in France are sequestering substantial amounts of carbon. Science of The Total Environment, 574, 616–628. doi.org/10.1016/j.scitotenv.2016.09.028, 2017.

Karhu, K., Wall, A., Vanhala, P., Liski, J., Esala, M., & Regina, K.: Effects of afforestation and deforestation on boreal soil carbon stocks—comparison of measured C stocks with Yasso07 model results. Geoderma, 164(1–2), 33–45. doi.org/10.1016/j.geoderma.2011.05.008, 2011.

Kelly, R.H., Parton, W.J., Crocker, G.J., Grace, P.R., Klír, J., Körschens, M., Poulton, P.R. and Richter, D.D.: Simulating trends in soil organic carbon in long-term experiments using the Century model, Geoderma, 81, 75–90, doi.org/10.1016/s0016-7061(97)00082-7, 1997.

Kögel-Knabner, I.: The macromolecular organic composition of plant and microbial residues as inputs to soil organic matter, Soil Biology and Biochemistry, 34(2), 139–162. doi.org/10.1016/s0038-0717(01)00158-4, 2002.

Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S. and Prentice, I.C.: A dynamic global vegetation model for studies of the couple atmosphere-biosphere system, Global Biogeochemical Cycles, 19(GB1015), doi.org/10.1029/2003gb002199, 2005.

Kurz, W. A., Dymond, C. C., White, T. M., Stinson, G., Shaw, C. H., Rampley, G. J., Smyth, C., Simpson, B. N., Neilson, E. T., Trofymow, J. A. and Metsaranta, J.: CBM-CFS3: a model of carbon-dynamics in forestry and land-use change implementing IPCC standards, Ecological modelling, 220(4), 480–504. doi.org/10.1016/j.ecolmodel.2008.10.018, 2009.

Lehmann, J. and Kleber, M.: The contentious nature of soil organic matter. Nature, 528, 60–68, doi.org/10.1038/nature16069, 2015.

Lehtonen, A., Linkosalo, T., Peltoniemi, M., Sievänen, R., Mäkipää, R., Tamminen, P., Salemaa, M., Nieminen, T., Tupek, B., Heikkinen, J. and Komarov, A.: Forest soil carbon stock estimates in a nationwide inventory: evaluating performance of the ROMULv and Yasso07 models in Finland. Geosci. Model Dev., 9, 4169–4183, doi:10.5194/gmd-9-4169-2016, 2016.

Liski, J., Palosuo, T., Peltoniemi, M., Sievänen, R.: Carbon and decomposition model Yasso for forest soils. Ecol. Modell., 189, 168–182, doi.org/10.1016/j.ecolmodel.2005.03.005, 2005.

Litton, C. M., Ryan, M. G., Tinker, D. B., and Knight, D. H.: Belowground and aboveground biomass in young postfire lodgepole pine forests of contrasting tree density, Canadian Journal of Forest Research, 33(2), 351–363, doi.org/10.1139/x02-181, 2003.

Lu, N., Akujärvi, A., Wu, X., Liski, J., Wen, Z., Holmberg, M., Feng, X., Zeng, Y. and Fu, B.: Changes in soil carbon stock predicted by a process-based soil carbon model (Yasso07) in the Yanhe watershed of the Loess Plateau, Landscape Ecology, 30(3), 399–413, doi.org/10.1007/s10980-014-0132-x, 2015.

Manzoni, S., and Porporato, A.: Soil carbon and nitrogen mineralization: theory and models across scales. Soil Biology and Biochemistry, 41(7), 1355–1379. doi.org/10.1016/j.soilbio.2009.02.031, 2009.

Metherell, A., Harding, L.A., Cole, C.V. and Parton, W.J.: Technical Documentation Agroecosystem Version 4.0. Great Plains, System Research Unit, USDA-ARS, Fort Collins, CO. 1993.

Mokany, K., Raison, R., & Prokushkin, A. S.: Critical analysis of root: shoot ratios in terrestrial biomes. Global Change Biology, 12(1), 84–96, doi.org/10.1111/j.1365-2486.2005.001043.x, 2006.

Muukkonen, P., and Mäkipää, R.: Empirical biomass models of understorey vegetation in boreal forests according to stand and site attributes, Boreal Environment Research, 11, 355–369, 2006.

~~Nadelhoffer, K. J., Raich, J. W.: Fine root production estimates and below-ground carbon allocation in forest ecosystems, Ecology 73, 1139–1147, , 1992.~~

Ortiz, C. A., Liski, J., Gärdenäs, A. I., Lehtonen, A., Lundblad, M., Stendahl, J., Ågren, G. I. and Karltun, E.: Soil organic carbon stock changes in Swedish forest soils—a comparison of uncertainties and their sources through a national inventory and two simulation models, Ecological Modelling, 251, 221–231. doi.org/10.1016/j.ecolmodel.2012.12.017, 2013.

Parton, W. J., Schimel, D. S., Cole, C. V., Ojima, D. S.: Analysis of factors controlling soil organic-matter levels in Great-Plains grasslands, Soil Science Society of America Journal, 51, 1173–1179, doi.org/10.2136/sssaj1987.03615995005100050015x, 1987.

Parton, W. J., Scurlock, J. M. O., Ojima, D. S.: Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide, Global Biogeochemical Cycles, 7, 785–809, doi.org/10.1029/93gb02042, 1993.

Pettersen, R. C.: The chemical composition of wood. The chemistry of solid wood, 207, 57–126, 1984.

Raich, J.W. and Nadelhoffer, K.J.: Below-ground carbon allocation in forest ecosystems: global trends. Ecology 70, 1346–1354. doi.org/10.2307/1938194, 1989.

Rantakari, M., Lehtonen, A., Linkosalo, T., Tuomi, M., Tamminen, P., Heikkinen, Liski J. Mäkipää R., Ilvesniemi H. & Sievänen, R.: The Yasso07 soil carbon model–Testing against repeated soil carbon inventory, Forest Ecology and Management, 286, 137–147, doi.org/10.1016/j.foreco.2012.08.041, 2012.

Rasmussen C, Heckman K, Wieder W, Keiluweit M, Lawrence C, Berhe A, Blankinship J, Crow S, Druhan J, Pries C, Marin-Spiotta E, Plante A, Schädel C, Schimel J, SierraC, Thompson A & Wagai R (2018) Beyond clay: towards an improved set of variables for predicting soil organic matter content. Biogeochemistry, Springer Nature, 137, 297-306 10.1007/s10533-018-0424-3

Rowell, R. M., Pettersen, R., Han, J. S., Rowell, J. S., & Tshabalala, M. A.: Cell wall chemistry. Handbook of wood chemistry and wood composites, 35–74, doi.org/10.1201/b12487-5, 2005.

Rowell, R. M. (Ed.): Handbook of wood chemistry and wood composites. CRC press. doi.org/10.1201/b12487, 2012.

Saby, N. P. A., P. H. Bellamy, X. Morvan, D. Arrouays, R. J. A. Jones, F. G. A. Verheijen, M. G. Kibblewhite, A. N. N. Verdoodt, J. B. ÜVeges, A. Freudenschuß, and C. Simota.: Will European soil-monitoring networks be able to detect changes in topsoil organic carbon content? Global Change Biology, 14, 2432–2442, doi.org/10.1111/j.1365-2486.2008.01658.x, 2008.

Schmidt, M. W. I., Torn, M. S., Abiven, S., Dittmar, T., Guggenberg, G., Janssens, I. A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, M., Nannipieri, P., Rasse, D. P., Weiner, S. and Trumbore, S. E.: Persistence of soil organic matter as an ecosystem property, Nature, 478, 49–56, doi.org/10.1038/nature10386, 2011.

Smith, W. N., Grant, B. B., Desjardins, R. L., Qian, B., Hutchinson, J. and Gameda, S.: Potential impact of climate change on carbon in agricultural soils in Canada 2000–2099, Climatic Change, 93, 319–333. doi.org/10.1007/s10584-008-9493-y, 2009.

Stump, L. M., and Binkley, D.: Relationships between litter quality and nitrogen availability in Rocky Mountain forests, Canadian Journal of Forest Research, 23(3), 492–502, doi.org/10.1139/x93-067, 1993.

Thomas, S. C., Martin, A. R.: Carbon content of tree tissues: a synthesis. Forests, 3(2), 332–352, doi.org/10.3390/f3020332, 2012.

Tingey, D. T., Mckane, R. B., Olszyk, D. M., Johnson, M. G., Rygiewicz, P. T., and Henry Lee, E.: Elevated CO_2 and temperature alter nitrogen allocation in Douglas-fir, Global Change Biology, 9(7), 1038–1050. doi.org/10.1046/j.1365-2486.2003.00646.x, 2003.

Todd-Brown, K. E. O., Hopkins, F. M. H., Kivlin, S. N., Talbot, J. M. and Allison, S. D., A framework for representing microbial decomposition in coupled climate models. Biogeochemistry, 109, 19–33, doi.org/10.1007/s10533-011-9635-6, 2012.

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

Tuomi, M., Laiho, R., Repo, A., Liski, J.: Wood decomposition model for boreal forests, Ecol. Modell., 222, 709–718. doi.org/10.1016/j.ecolmodel.2010.10.025, 2011.

Tuomi, M., Rasinmaki, J., Repo, A., Vanhala, P., Liski, J.: Soil carbon model Yasso07 graphical user interface. Environ. Modell. Softw., 26, 1358–1362, doi.org/10.1016/j.envsoft.2011.05.009, 2011b.

Wu, X., Akujärvi, A., Lu, N., Liski, J., Liu, G., Wang, Y., Holmberg, M., Li, F., Zeng, Y. and Fu, B.: Dynamics of soil organic carbon stock in a typical catchment of the Loess Plateau: comparison of model simulations with measurements, Landscape Ecology, 30(3), 381–397, doi.org/10.1007/s10980-014-0110-3, 2015.

Wutzler, T. and Reichstein, M.: Soils apart from equilibrium? consequences for soil carbon balance modelling, Biogeosciences~~ Discussions~~, 3(5), 1679–1714, doi.org/10.5194/bg-4-125-2007, 2007~~6~~.

Wutzler, T. and Reichstein, M.: Priming and substrate quality interactions in soil organic matter models. Biogeosciences, 10, 2089–2103, doi: 2089-2103 10.5194/bg-10-2089-2013. 2013.

Xia, M., Talhelm, A. F. and Pregitzer, K. S.: Fine roots are the dominant source of recalcitrant plant litter in sugar maple-dominated northern hardwood forests, New Phytologist, 208(3), 715–726, doi.org/10.1111/nph.13494, 2015.

Tables

Functional type	Species	Organ	Case	No. of obs.				Mean (%)				SD (%)			
				A	W	E	N	A	W	E	N	A	W	E	N
Broadleaves	*Fagus sylvatica* L.	wood	4	4	4	4	4	74.5	2.8	1.2	21.5	1.4	1	0.5	1.4
		leaf	2	2	1	1	2	39.6	22.1	12.5	25.8	3.5	NA	NA	1.7
		root	3	1	9	9	1	31.5	8.8	18.6	41.1	NA	1.2	1.2	NA
	Quercus petraea (Matt.) Liebl.	wood	4	19	19	19	19	67.5	6.1	3.5	22.9	4.9	2.3	1.7	2.6
		leaf	4	12	12	12	12	40.8	16.3	14.2	28.7	3.5	4.7	9.3	7.1
		root	5	15	9	9	15	34.9	7.6	16.2	41.3	8.0	1.1	1.1	10.4
	Quercus robur L.	wood	4	19	19	19	19	67.5	6.1	3.5	22.9	4.9	2.3	1.7	2.6
		leaf	2	1	12	12	1	37.7	21.6	17.3	23.4	NA	7.3	7.3	NA
		root	3	1	9	9	1	28.6	11.1	23.4	36.9	NA	1.5	1.5	NA
Conifers	*Abies alba* Mill.	wood	4	14	14	14	14	66.7	2.7	2.4	28.2	1.9	1.3	0.8	1.3
		leaf	2	1	6	6	1	32.4	26.4	10.7	30.5	NA	1.4	1.4	NA
		root	3	1	13	13	1	25.3	19.1	21.5	34.1	NA	6.2	6.2	NA
	Larix deciduas Mill.	wood	4	6	6	6	6	65.3	5.9	1.9	26.9	3.2	2.4	0.9	1.5
		leaf	2	2	4	4	2	33.3	30.2	10.1	26.4	2.5	1.6	1.6	7.7
		root	3	1	13	13	1	32.5	16.2	18.2	33.1	NA	5.2	5.2	NA
	Picea abies (L.) H. Karst	wood	1	1	1	1	1	69.5	1.9	1.0	27.6	NA	NA	NA	NA
		leaf	2	1	6	6	1	37.0	29.5	12.0	21.5	NA	2.2	2.2	NA
		root	3	3	13	13	3	36.6	14.8	16.6	32.0	7.8	4.8	4.8	2
	Pseudotsuga menziesii (Mirb.) Franco	wood	1	1	1	1	1	65.3	4.0	4.0	26.7	NA	NA	NA	NA
		leaf	1	6	6	6	6	36.4	25.1	10.9	27.6	6.8	13.1	1.2	6.3
		root	1	2	2	2	2	41.7	16.9	8.4	33.0	2.4	5.5	0.3	3.3
	Pinus nigra var. corsicana (J.W. Loudon) Hyl.	wood	4	22	22	22	22	66.6	3.3	4.0	26.1	2.9	1.5	2.4	1.3
		leaf	2	1	27	27	1	47.1	15.2	13.8	23.9	NA	6.3	6.3	NA
		root	4	10	10	10	10	36.0	9.2	11.9	42.9	4.9	4.4	3.1	7.3
	Pinus pinaster Aiton	wood	4	22	22	22	22	66.6	3.3	4.0	26.1	2.9	1.5	2.4	1.3
		leaf	2	1	27	27	1	43.2	18.2	16.5	22.1	NA	7.5	7.5	NA
		root	4	10	10	10	10	36.0	9.2	11.9	42.9	4.9	4.4	3.1	7.3
	Pinus sylvestris L.	wood	1	1	1	1	1	71.7	0.9	1.0	26.4	NA	NA	NA	NA
		leaf	1	3	3	3	3	40.7	17.0	16.0	26.3	3.8	7.5	6.5	2.4
		root	2	4	10	10	4	51.2	4.4	6.0	38.4	3.7	1.4	1.4	4.5

Table 1 Litter carbon quality of the species present in the French RENCOFOR network estimated based on literature. In the column "Case," each number corresponds to one case of data availability in literature: 1- at least one dataset of complete chemical composition (i.e. for AWEN) exists at species level; 2 - at least one dataset of incomplete chemical composition (only for A, N and the sum of W and E) exists at species level; in this case, the mean proportion of W and E at genus level is used; 3 – no data are available at species level, but at least one complete dataset of chemical composition exists at genus level; 4 - no data are available at species level, but at least one dataset of chemical composition exists at genus level; in this case, the mean proportion of W and E at tree functional type level is used; 5 – no data are available at neither species nor genus level, in this case, the mean AWEN composition at tree functional type level is used. From Case 1 to 5 is in descending order of priority.

Data	Observed litter input quantity (mean ± SD, in tC ha^{-1} yr^{-1})		Year																
	Conifers (51 sites)	Broadleaves (50 sites)	1961 - 1990	1991	1992	1993	1994	1995	1996	1997 - 2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Climate	-	-	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M
Organic matter inputs via forests																			
Fruits and miscellaneous	0.36 ± 0.28	0.64 ± 0.41					M	M	M	M	M	M	M						
Leaves	1.12 ± 0.35	1.28 ± 0.31					M	M	M	M	M	M	M						
Fine branches	0.29 ± 0.14	0.45 ± 0.14					M	M	M	M	M	M	M						
Coarse woody branches*	0.32 ± 0.14	0.72 ± 0.29					M	M	M	M	M	M	M	M	M	M	M	M	M
Stems*	0	0					0	0	0	0	0	0	0	0	0	0	0	0	0
Coarse woody roots*	0.83 ± 0.36	1.03 ± 0.38					E	E	E	E	E	E	E	M	M	M	M	M	M
Fine roots	-	-					E	E	E	E	E	E	E						
Soil carbon stock	-	-						M									M		

Table 2 A summary of the data used for Yasso07 simulations in the present study. In the "Year" columns: M - measured data; E - estimated data according to the measured ones; 0 – noted, but the contribution to litter is negligible~~ignorable~~. For soil carbon stock measurement, dashed line zones denote the inventory duration. For each year, each symbol (M and E) only account for the general case and hence it is possible that measurement was occasionally omitted at some sites. * - litter input caused by harvest or storms were included (once they occurred); SD - standard deviation; litter inputs are dry matters. Diameters used for defining each litter type: ≤2 cm for fine branches, >4 cm for coarse woody branches, > 5 mm for coarse woody roots and ≤ 5 mm for fine roots.

Figures

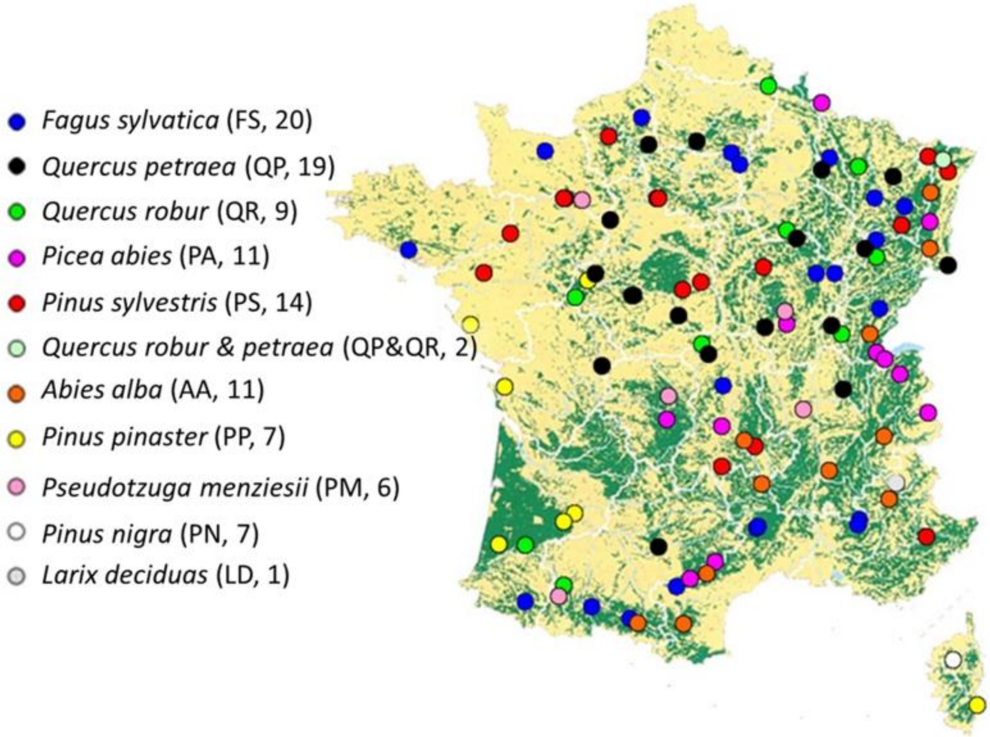


Figure 1 Geographical distribution of the sites of RENECOFOR network used for testing the performance of Yasso07 (see also Jonard et al., 2017). Forested areas are represented in green. Each circle represents one site; the color represents the dominant tree species of the plot. In each pair of parentheses, the species abbreviation and number of sites by species are indicated.

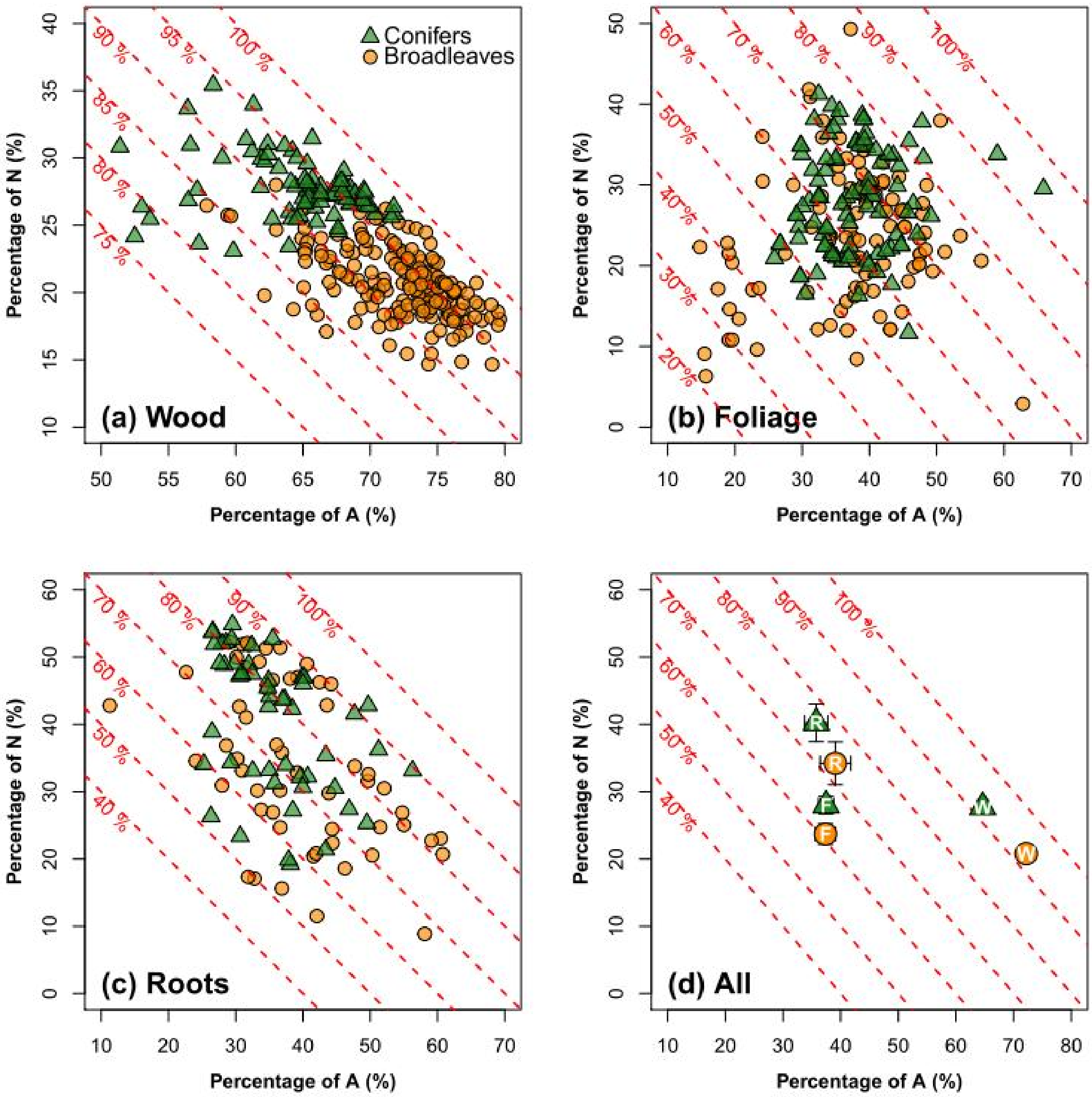


Figure 2 A meta-analysis of the carbon composition for northern temperate tree species: *x*-axis represents the percentage of acid-hydrolysable compounds (e.g. cellulose, noted by A, in %) and *y*-axis represent the percentage of non-soluble and non-hydrolyzable compround (e.g. lignin, noted by N, in %). The oblique dashed red lines notify the sum of A and N, the values of which are shown here. The remaining percentage, i.e. 100 - A - N, refers to the portion of compounds like non-polar extractives, ethanol ordichloromethane (E), or in water (W). (a) Analysis conducted for wood (106 data points for broadleaves; 79 for conifers), (b) for foliage litter (b, 106 data points for broadleaves; 83 for conifers) and (c) for root litter (58 data points for broadleaves; 49 for conifers); (d) is a statistical synthesis (symbols – means and error bars – 1.96 * standard error) of wood (W), foliage (F) and roots (R) in a common coordinates system. Attention to the use of different axis graduations in each plot. See Supplementary Material II for the data sources. Note the different y-axis scales.

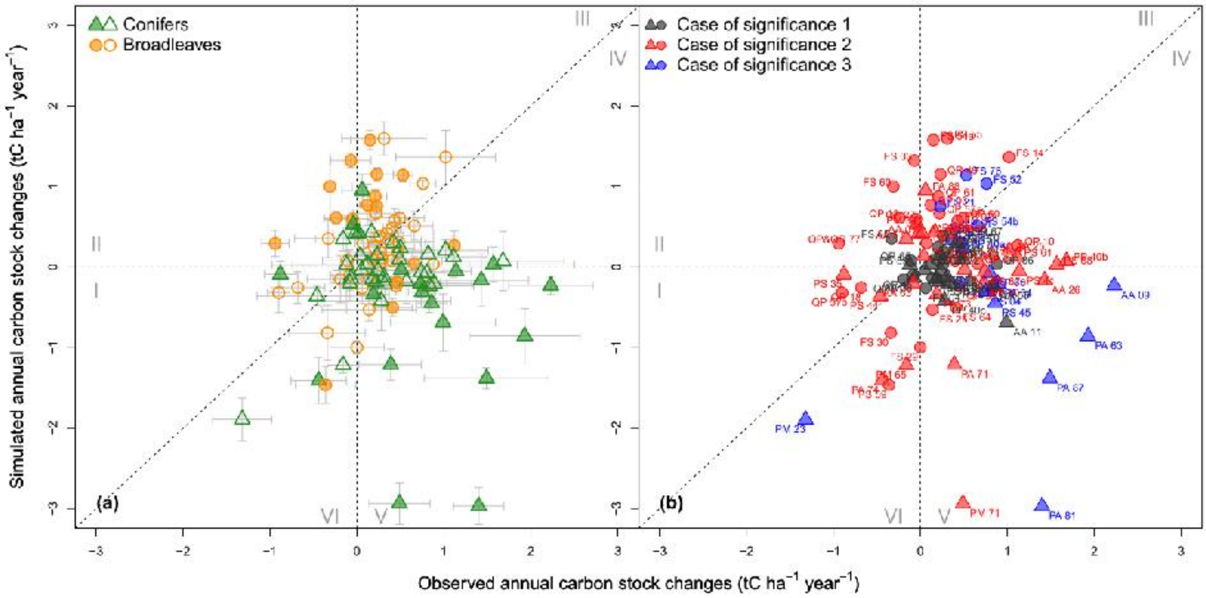


Figure 3 Comparison between simulated and observed annual carbon stock changes (ACC, in tC ha^{-1} year^{-1}). Round and triangle symbols represent sites dominated by broadleaves and conifers, respectively. Partial steady-state assumption was used for initializing carbon quality of the stock until 1.0 m. The chosen fine root:foliage ratio for broadleaves and conifers is 1.0. To facilitate discussions, we set Roman numbers (I-VI) denoting the six zones in which data points are distributed. In (a), error bars represent standard errors; hollow and filled points represent non-significant and significant differences between simulated and observed ACC according to t-test (at 95% confidence level). In (b), case of significance: 1 – no significant difference from 0 for neither observed nor simulated ACC; 2 - a significant difference from 0 for either observed or simulated ACC and 3: - a significant difference from 0 for both observed and simulated ACC.

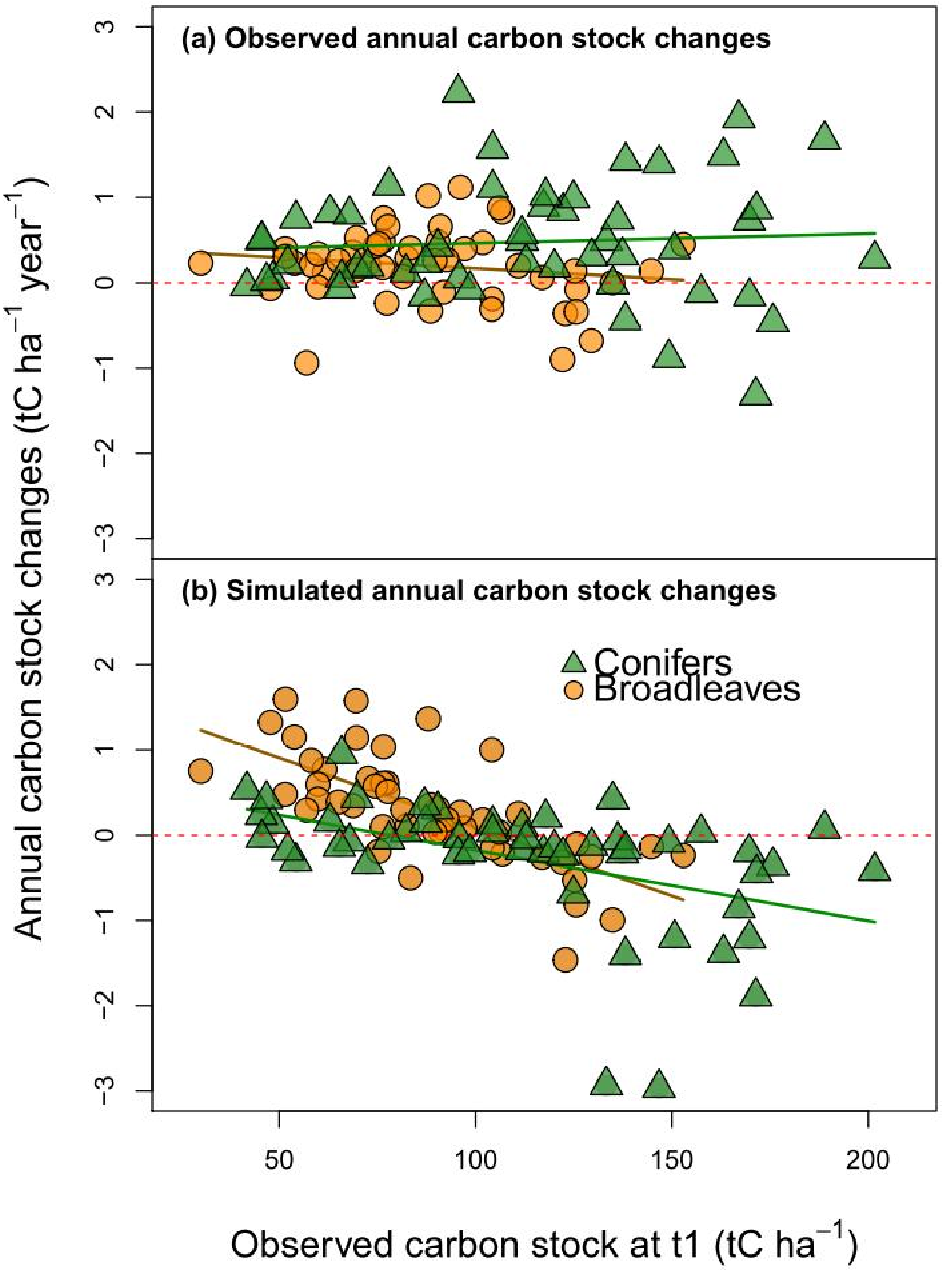


Figure 4 Observed (y-axis, a) and simulated annual change changes (y-axis, b) plotted against the observed carbon stock until 1.0 m (x-axis) during the first soil carbon stock inventory. Regressions: $y = -0.003x + 0.422$ ($R^2 = 0.03$) for observed values in the sites dominated by broadleaves; $y = 0.001x + 0.353$ ($R^2 = 0.01$) for the sites dominated by conifers; $y = -0.016x + 1.715$ ($R^2 = 0.62$) for simulated values of the sites dominated by broadleaves; $y = -0.008x + 0.648$($R^2 = 0.60$) for simulated values of the sites dominated conifers.

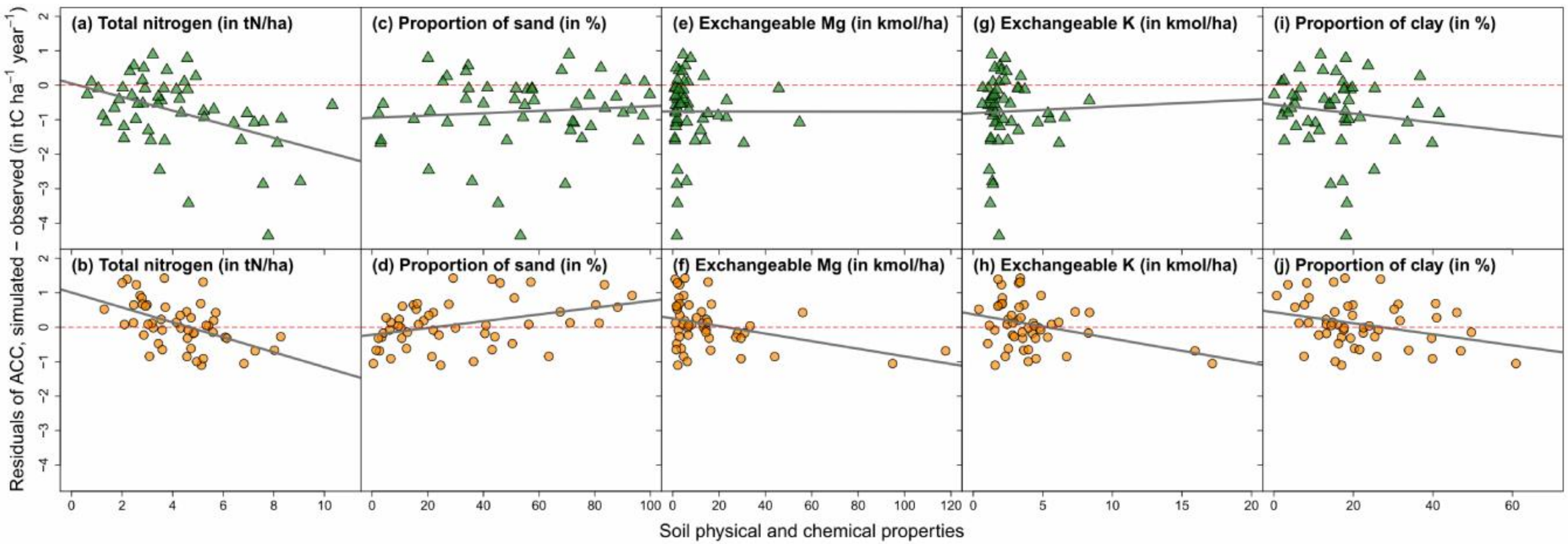
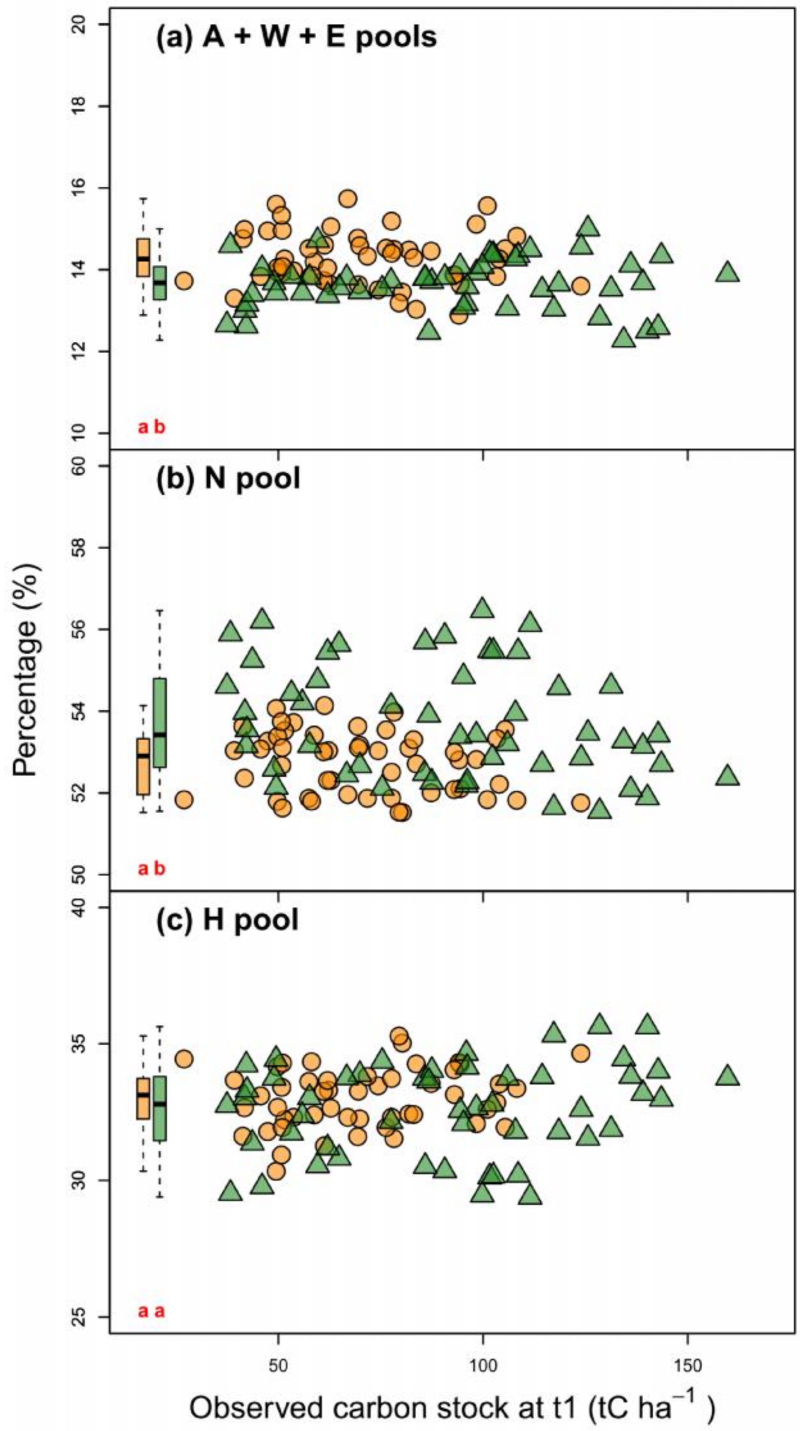


Figure 5 Residuals plotted against selected soil physical and chemical properties. Top plots with green triangles stand for the sites dominated by conifers and bottom plots with orange dots stand for the sites dominated by broadleaves. Regressions in all the five subplots for the broadleaved sites (b, d, f, h and i) and in one subplot for the stands dominated by conifers (a) are significant ($P<0.5^*$). See Table S2 for results of linear regressions of all the 11 soil variables. Red dashed line indicates the zero line.



~~Figure 6 Proportions of carbon pools (AWENH) at steady-state for all the RENECOFOR sites (*y*-axis) plotted against observed carbon stock at t1 until 0.4 m (*x*-axis). Each symbol represents one RENECOFOR site: green triangles stand for the sites dominated by conifers and orange dots stand for the sites dominated by broadleaves. For each boxplot, the lower and top edge of the box corresponds to the 25th and 75th percentile data points; lower and top bars the line within the box represents the median and the hollow points indicate outliers. Red letters below the boxplot denote the statistical diagnoses (t-test) with a significance level of *P*= 0.05*. No clear linear relationship was found between carbon quality and observed carbon stock at t1.~~

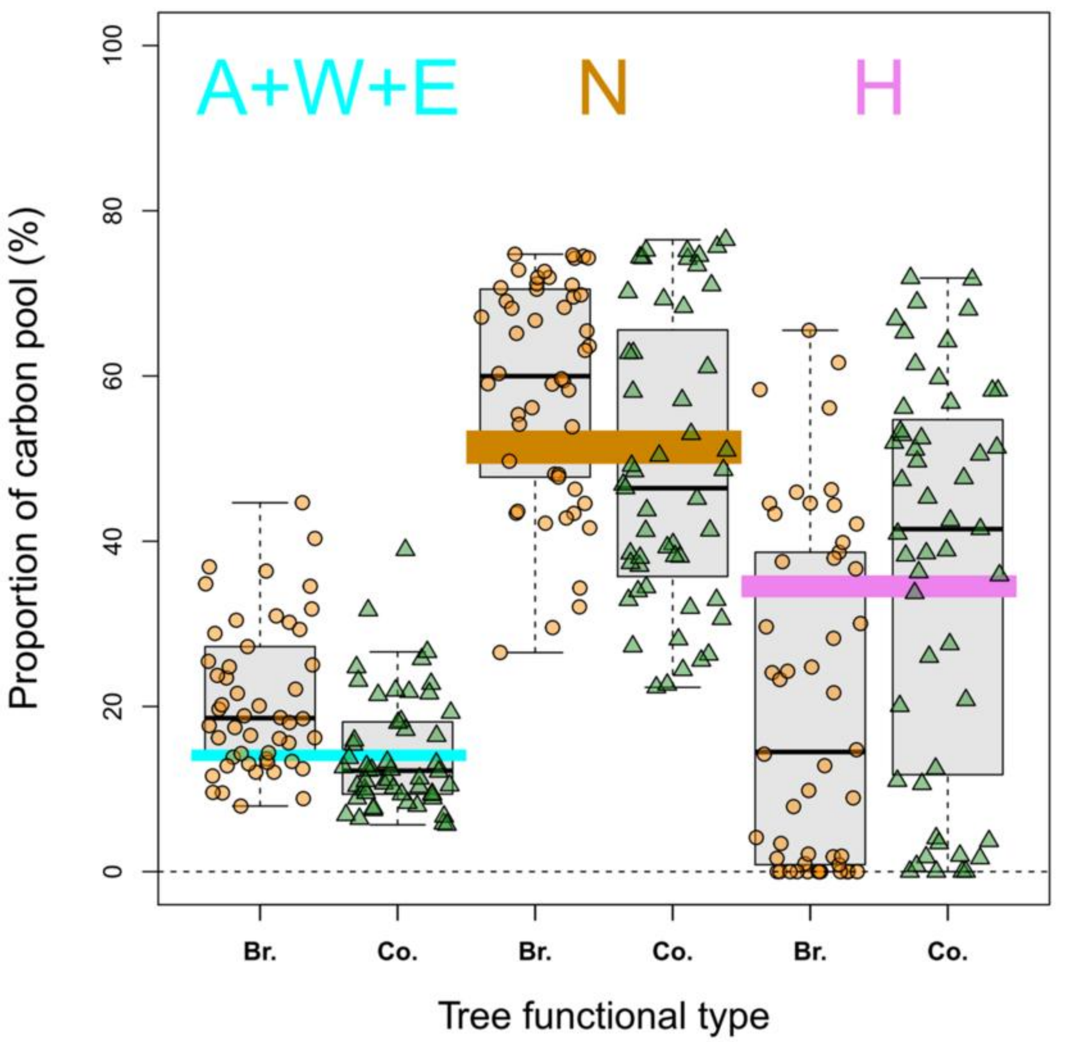


Figure 6 Distribution of estimated carbon qualities based on the partial steady-state assumption (boxplots) versus those based on the complete steady-state assumption (whose ranges are all very narrow and are expressed with strips in colour: 13 – 15 % for the sum of A, W and E (cyan); 49 – 53 % for N (brown); 33 – 36 % for H (purple)). For each boxplot, the lower and top edge of the box corresponds to the 25^{th} and 75^{th} percentile data points; lower and top bars the line within the box represents the median; no outlier points in this case. Br. – Broadleaves stands; Co. – Conifer stands.

Supplementary Materials

Supplementary Materials I: Supplementary tables and figures.

Supplementary Materials II: Database for the meta-analysis of wood and litter chemical composition.

Modeling soil organic carbon dynamics in temperate forests using Yasso07

Zhun Mao[1,8*], Delphine Derrien[1], Markus Didion[2], Jari Liski[3,9], Thomas Eglin[4], Manuel Nicolas[5], Mathieu Jonard[6], Laurent Saint-André[1,7]

[1] INRA, UR BEF – Biogéochimie des Ecosystèmes Forestiers, 54280 Champenoux, France

[2] Swiss Federal Institute for Forest, Snow and Landscape Research WSL, 8903 Birmensdorf, Switzerland

[3] Finnish Environment Institute, Ecosystem Change Unit, Natural Environment Centre, Mechelininkatu 34a, P.O.Box 140, 00251 Helsinki, Finland

[4] ADEME – DPED – Service Agriculture et Forêts, 49004 Angers, France

[5] Office National des Forêts Direction Forêts et Risques Naturels, Département Recherche et Développement - Bâtiment B, Boulevard de Constance, 77300 Fontainebleau, France

[6] Université Catholique de Louvain, Earth and Life Institute, Croix du Sud 2, L7.05.09, 1348 Louvain-la-Neuve, Belgium

[7] CIRAD, UMR ECO&Sols, place Viala, 34398 Montpellier Cedex 5, France

[8] Amap, Inra, University Montpellier, Cnrs, Ird, Cirad, Montpellier, France

[9] Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland

*Corresponding author: Zhun Mao; email address: maozhun04@126.com.

Supporting Material I: Supplementary tables and figures

Table S1 Information on forest inventories for stand biomass estimation

Site	Dominant Species	Soil	Forest inventory for stand biomass				Storm event (yr)	No. of thinnings
			Beginning(yr)	End (yr)	Span (yrs)	No. of inventories		
QR_10	*Quercus robur*	Calcisol	1991	2009	18	7		1
QR_18	*Quercus robur*	Planosol	1991	2009	18	7		3
QR_40	*Quercus robur*	Cambisol	1992	2011	19	9	2009	2
QR_49	*Quercus robur*	Planosol	1991	2010	19	6	2009	3
QR_55	*Quercus robur*	Calcisol	1992	2009	17	6		2
QR_59	*Quercus robur*	Luvisol	1991	2010	19	8		2
QR_65	*Quercus robur*	Cambisol	1992	2012	20	6		3
QR_70	*Quercus robur*	Luvisol	1992	2011	19	7		3
QR_71	*Quercus robur*	Luvisol	1991	2009	18	6		2
QP_1	*Quercus petraea*	Cambisol	1991	2011	20	7		2
QP_3	*Quercus petraea*	Cambisol	1991	2009	18	7		3
QP_10	*Quercus petraea*	Luvisol	1991	2010	19	9		3
QP_18	*Quercus petraea*	Luvisol	1991	2009	18	7		3
QP_21	*Quercus petraea*	Luvisol	1991	2012	21	7		2
QP_27	*Quercus petraea*	Luvisol	1992	2009	17	9	1999	2
QP_35	*Quercus petraea*	Luvisol	1991	2011	20	6		2

QP_41	*Quercus petraea*	Luvisol	1991	2010	19	7		1
QP_51	*Quercus petraea*	Cambisol	1992	2004	12	5	1999	0
QP_57a	*Quercus petraea*	Planosol	1992	2009	17	8	1999	1
QP_57b	*Quercus petraea*	Podzol	1992	2009	17	5		2
QP_58	*Quercus petraea*	Luvisol	1991	2009	18	6		3
QP_60	*Quercus petraea*	Planosol	1992	2009	17	7		2
QP_61	*Quercus petraea*	Luvisol	1991	2009	18	7	1999	2
QP_68	*Quercus petraea*	Calcisol	1992	2009	17	7	1999	2
QP_72	*Quercus petraea*	Luvisol	1991	2009	18	8		3
QP_81	*Quercus petraea*	Luvisol	1992	2009	17	6		1
QP_86	*Quercus petraea*	Luvisol	1991	2009	18	6	1999	4
QP_88	*Quercus petraea*	Cambisol	1992	2011	19	8		3
QP&QR_67	*Quercus petraea & Q. robur*	Cambisol	1992	2004	12	5	1999	3
QP&QR_77	*Quercus petraea & Q. robur*	Podzol	1991	2009	18	6	1999	2
PM_23	*Pseudotsuga menziesii*	Cambisol	1991	2008	17	7	1999	1
PM_34	*Pseudotsuga menziesii*	Cambisol	1991	2010	19	7		4
PM_61	*Pseudotsuga menziesii*	Luvisol	1991	2011	20	7		3
PM_65	*Pseudotsuga menziesii*	Cambisol	1992	2004	12	5		0

PM_69	*Pseudotsuga menziesii*	Cambisol	1991	2004	13	7	1999	1
PM_71	*Pseudotsuga menziesii*	Podzol	1993	2013	20	11		5
PA_8	*Picea abies*	Podzol	1992	2009	17	5		1
PA_34	*Picea abies*	Podzol	1991	2009	18	7	2009	3
PA_39a	*Picea abies*	Luvisol	1991	2004	13	5		1
PA_63	*Picea abies*	Andosol	1991	2009	18	7		3
PA_71	*Picea abies*	Podzol	1991	2004	13	5		1
PA_73	*Picea abies*	Cambisol	1992	2007	15	5		2
PA_74	*Picea abies*	Luvisol	1991	2009	18	8		3
PA_81	*Picea abies*	Podzol	1992	2004	12	5		1
PA_87	*Picea abies*	Podzol	1991	2009	18	7	1999	2
PA_88	*Picea abies*	Cambisol	1992	1999	7	3	1999	0
FS_2	*Fagus sylvatica*	Luvisol	1992	2009	17	6		2
FS_3	*Fagus sylvatica*	Cambisol	1991	2009	18	8		3
FS_4	*Fagus sylvatica*	Cambisol	1992	2009	17	5		0
FS_9	*Fagus sylvatica*	Podzol	1992	2009	17	6		1
FS_14	*Fagus sylvatica*	Cambisol	1991	2013	22	8		3
FS_21	*Fagus sylvatica*	Leptosol	1991	2009	18	6	1999	1

FS_25	*Fagus sylvatica*	Cambisol	1991	2009	18	7		3
FS_26	*Fagus sylvatica*	Leptosol	1991	2009	18	6		1
FS_29	*Fagus sylvatica*	Luvisol	1991	2009	18	6		3
FS_30	*Fagus sylvatica*	Podzol	1991	2012	21	7		2
FS_52	*Fagus sylvatica*	Leptosol	1991	2005	14	6	1999	2
FS_54a	*Fagus sylvatica*	Planosol	1992	1999	7	3	1999	1
FS_54b	*Fagus sylvatica*	Leptosol	1992	1999	7	4	1999	0
FS_55	*Fagus sylvatica*	Podzol	1992	2011	19	8	1999	2
FS_60	*Fagus sylvatica*	Luvisol	1992	2009	17	7	1999	1
FS_64	*Fagus sylvatica*	Cambisol	1992	2011	19	8		3
FS_65	*Fagus sylvatica*	Cambisol	1992	2009	17	7		2
FS_76	*Fagus sylvatica*	Luvisol	1991	2009	18	9		2
FS_81	*Fagus sylvatica*	Podzol	1992	2009	17	6		1
FS_88	*Fagus sylvatica*	Cambisol	1992	2009	17	8		2
LD_5	*Larix deciduas*	Regosol	1991	2014	23	6		2
PN_20	*Pinus nigra*	Cambisol	1991	2009	18	7		2
PN_41	*Pinus nigra*	Podzol	1991	2004	13	6	1999	2
PP_17	*Pinus pinaster*	Arenosol	1991	2009	18	7	1999	1

PP_20	*Pinus pinaster*	Cambisol	1991	2004	13	5		2
PP_40a	*Pinus pinaster*	Podzol	1992	2004	12	8		2
PP_40b	*Pinus pinaster*	Podzol	1992	2009	17	7	2009	2
PP_40c	*Pinus pinaster*	Podzol	1992	2009	17	8	2009	3
PP_72	*Pinus pinaster*	Podzol	1991	2010	19	7		4
PP_85	*Pinus pinaster*	Arenosol	1991	2011	20	6		3
PS_4	*Pinus sylvestris*	Leptosol	1991	2004	13	4		0
PS_15	*Pinus sylvestris*	Cambisol	1991	2011	20	7	1999	2
PS_35	*Pinus sylvestris*	Luvisol	1991	2013	22	7		3
PS_41	*Pinus sylvestris*	Podzol	1991	2004	13	6	1999	2
PS_44	*Pinus sylvestris*	Luvisol	1991	2010	19	7		3
PS_45	*Pinus sylvestris*	Planosol	1991	2005	14	7		2
PS_61	*Pinus sylvestris*	Luvisol	1991	1999	8	3	1999	0
PS_63	*Pinus sylvestris*	Cambisol	1991	2009	18	7	1999	0
PS_67a	*Pinus sylvestris*	Podzol	1992	2009	17	7	1999	1
PS_67b	*Pinus sylvestris*	Podzol	1992	2013	21	9	1999	3
PS_76	*Pinus sylvestris*	Podzol	1991	2009	18	7	1999	1
PS_78	*Pinus sylvestris*	Podzol	1992	2007	15	6	1999	1

PS_88	*Pinus sylvestris*	Podzol	1992	2007	15	6	1999	1
PS_89	*Pinus sylvestris*	Podzol	1991	1999	8	4	1999	1
AA_5	*Abies alba*	Cambisol	1991	2009	18	5		1
AA_7	*Abies alba*	Podzol	1991	2010	19	6		1
AA_9	*Abies alba*	Podzol	1992	2008	16	7	2009	2
AA_11	*Abies alba*	Luvisol	1992	2009	17	8		2
AA_25	*Abies alba*	Cambisol	1991	2012	21	8		3
AA_26	*Abies alba*	Cambisol	1991	2014	23	6		2
AA_38	*Abies alba*	Cambisol	1992	2009	17	6		1
AA_39	*Abies alba*	Cambisol	1991	2009	18	6		2
AA_57	*Abies alba*	Cambisol	1992	2009	17	9	1999	2
AA_63	*Abies alba*	Cambisol	1992	2004	12	6		1
AA_68	*Abies alba*	Cambisol	1992	2012	20	7		3

	Beginning (yr)	End (yr)	Mean span (yrs)	Mean no. of inventories	Frequency (storms/100 yrs)	Frequency (thinnings/10 yrs)
All sites:	1991	2014	17.0	6.6	2.1	1.1

Table S2 Linear regressions for explaining the variability of annual carbon change residuals using soil physical and chemical properties

	Broadleaves				Conifers			
Variable	**R^2**	**Slope**	**Intercept**	***P*-value**	**R^2**	**Slope**	**Intercept**	***P*-value**
Total nitrogen (in tN/ha)	0.257	-0.217	1.006	<0.001	0.191	-0.198	0.056	<0.01**
Proportion of sand (in %)	0.152	0.010	-0.221	<0.01**	0.008	0.003	-0.944	>0.05
Exchangeable Mg (in kmol/ha)	0.138	-0.011	0.255	<0.01**	0.000	0.000	-0.761	>0.05
Exchangeable K (in kmol/ha)	0.109	-0.071	0.374	<0.05*	0.001	0.020	-0.807	>0.05
Proportion of clay (in %)	0.099	-0.016	0.435	<0.05*	0.016	-0.013	-0.561	>0.05
Proportion of silt (in %)	0.094	-0.010	0.566	<0.05*	0.004	-0.003	-0.660	>0.05
Exchangeable Al (in kmol/ha)	0.070	-0.004	0.360	>0.05	0.002	-0.001	-0.704	>0.05
Total phosphorus (in tN/ha)	0.045	-0.011	0.304	>0.05	0.000	0.000	-0.770	>0.05
Exchangeable Ca (in kmol/ha)	0.016	0.000	0.135	>0.05	0.004	0.000	-0.729	>0.05
pH	0.005	0.042	-0.099	>0.05	0.000	0.018	-0.839	>0.05
Carbon:nitrogen ratio	0.000	0.001	0.069	>0.05	0.019	0.009	-1.063	>0.05

Note: the grey zone indicates the variables chosen for plotting the Figure 5 in the manuscript. R^2 = coefficient of determination;

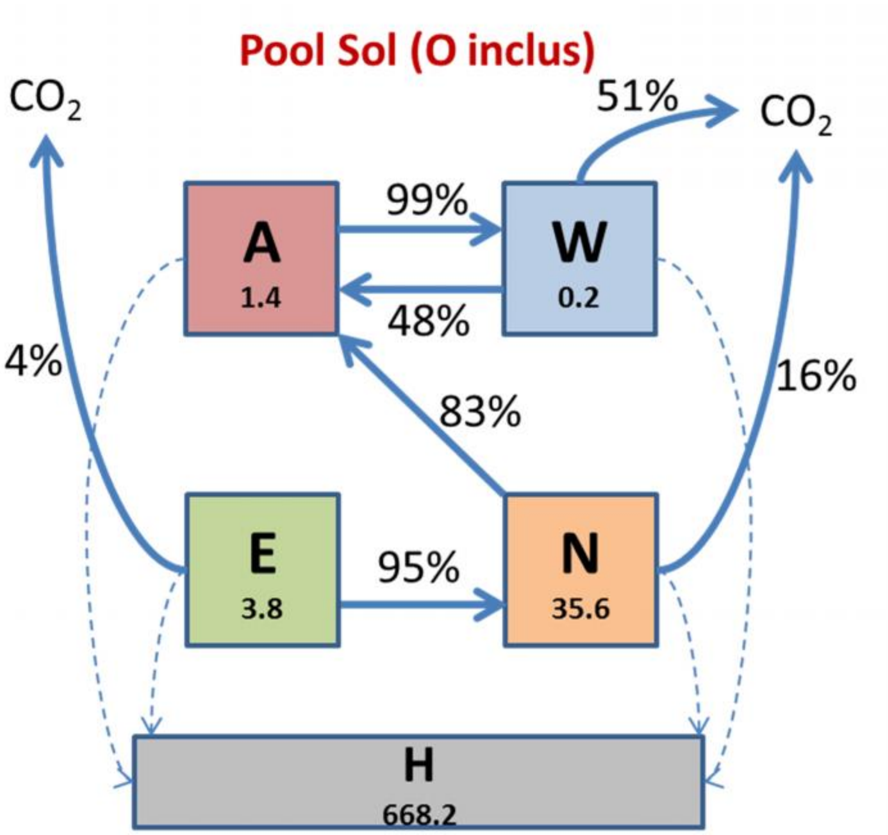


Figure S1 Partitioning of soil carbon pools in Yasso07 (after Tuomi et al., (2011b)) Letters: A: hydrolysable in Acid; W: soluble in Water; E: soluble Ethanol; N: Non-soluble; H: recalcitrant Humus. Solid arrows represented the carbon flows that are statistically significant from zero. Dashed arrows refer to the carbon flows toward H. Values in each pool is an example inverse of mean residence time (*1/k*, in year) estimated using Yasso07 parameters.

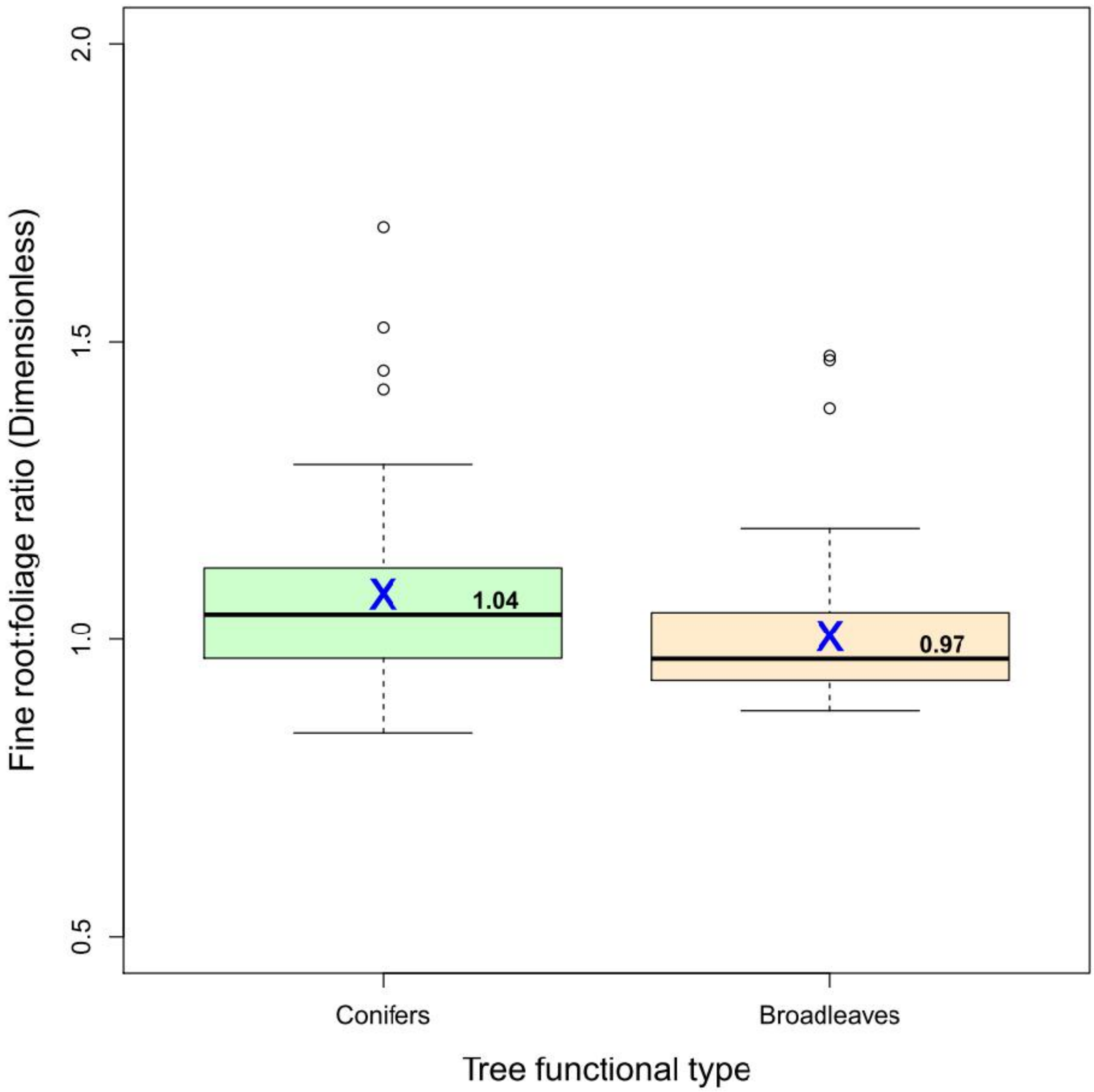


Figure S2~~3~~ Distributions of fine root:foliage ratio of litter input in different tree functional types calculated using the equation of Raich and Nadelhoffer (1989), see Jonard et al., (2017). For each boxplot, the lower and top edge of the box corresponds to the 25th and 75th percentile data points; lower and top bars the line within the box represents the median and the hollow points indicate outliers. Median values are shown beside median lines. “X” indicates mean values: 1.08 ± 0.02 (mean ± standard error) for sites dominated by conifers and 1.01 ± 0.02 for sites dominated by broadleaves.

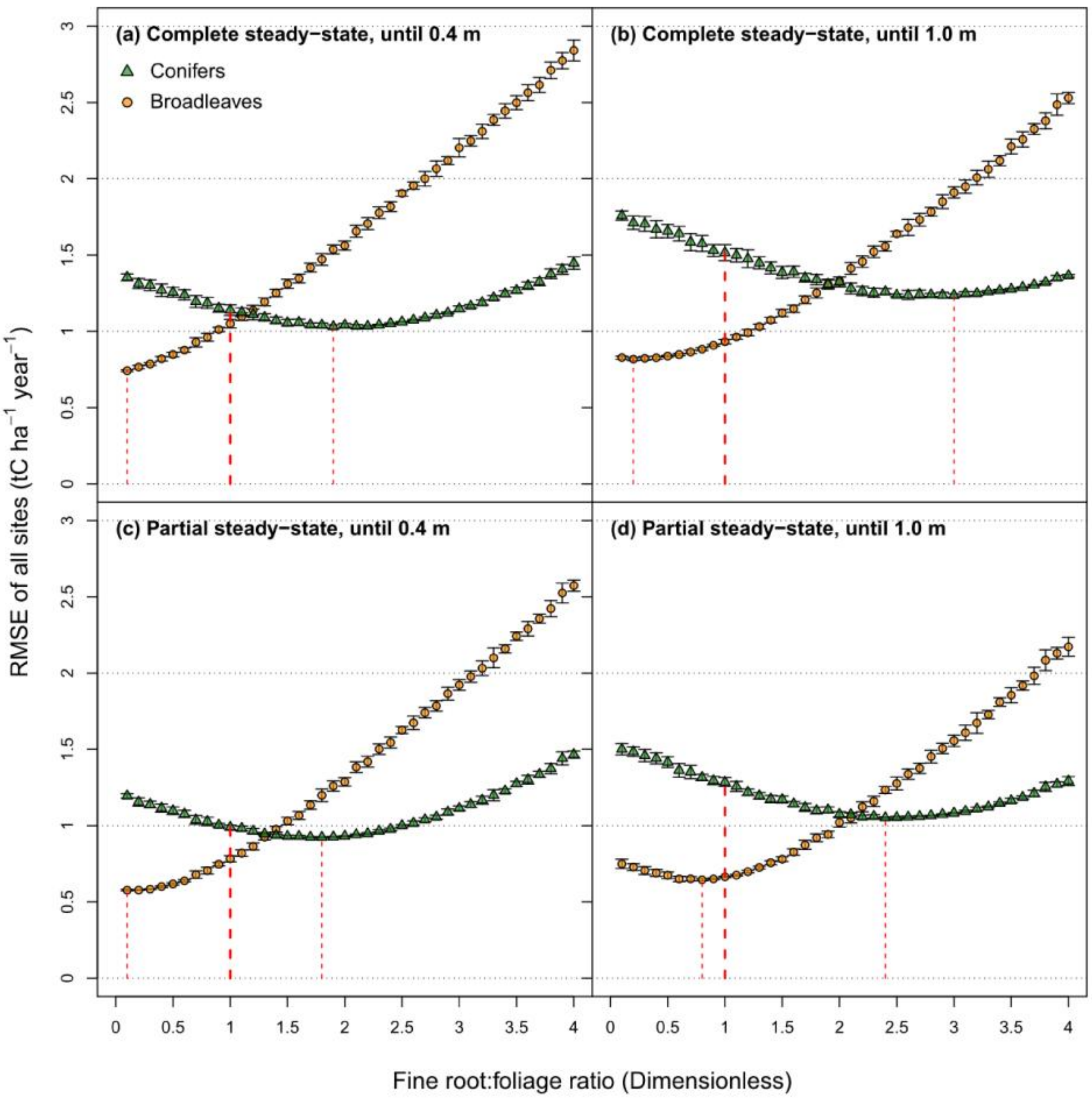


Figure S3~~2~~ Influence of the choice of model initialization method for soil carbon quantity (stock until 0.4 m versus stock until 1.0 m) and quality (complete versus partial steady-state assumption) and the choice of ~~-~~fine root:foliage ratio of litter input (from 0.1 to 4.0) on the performance of Yasso07 toward the French RENECOFOR data. RMSE – root mean square error; Error bars are standard deviations of 10 simulations differing in parameters which were randomly chosen. Red dash lines perpendicular to *x*-axis: the two thin ones showing the values of fine root:foliage ratio for the minima of RMSE for broadleaves ~~(0.1)~~ and conifers ~~(1.9)~~, respectively; the thick red dash line at 1.0 (i.e., the ratio used for result presentation) ~~showing that the RMSE of broadleaves and conifers are slightly higher than the minima, but still acceptable~~. The case in (c) gives the best model fit (lowest RMSE), but the case in (d) was preferentially chosen, as Yasso07 is validated by and predicts soil carbon data until 1.0 m.

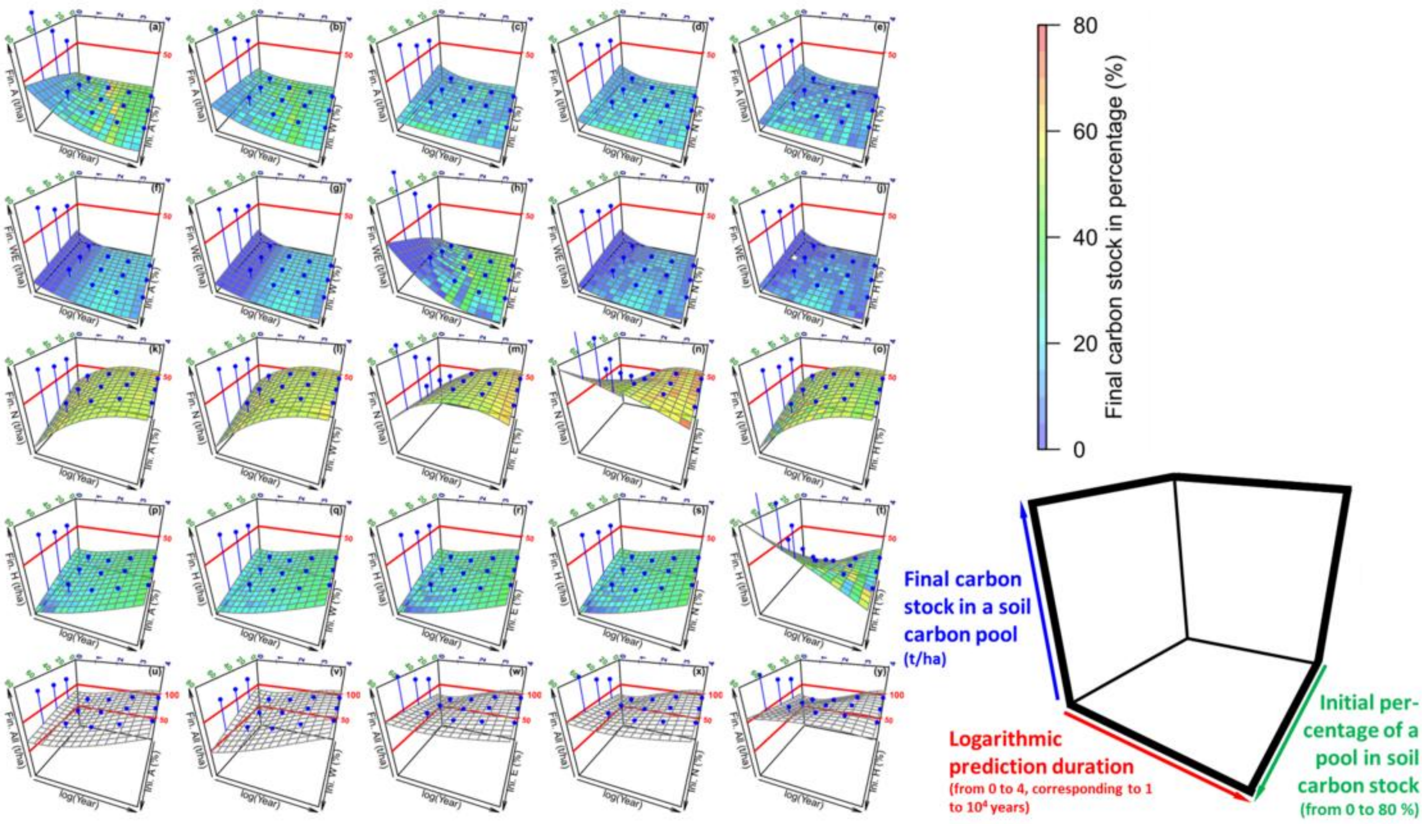


Figure S4~~8~~ Sensitivity analysis of the impact of carbon pool composition of initial soil C stock (*x*-axis (↙), in %) and simulation length (*y*-axis (→), in logarithmic years) on final soil carbon stock (*z*-axis (↑), in tC ha^{-1}). Here, the results are generated using the mean broadleaved litter input quantity and quality of the RENECOFOR sites. Initial soil carbon stock was fixed to 100 tC ha^{-1}. Subplots in each row show the final stock evolution of one type of soil carbon pools (i.e. A, W, E, N and H). Particularly, in the 2nd row W and E were combined due to their weak quantities in most of cases. Subplots in each column show the effect of one type of soil chemical groups on the final stocks of the five soil carbon pools (each of them for the first four and the last one is the total stock). In each subplot, a membrane (with grids for three-dimensional effect) represents the loess fit (polynomial equation) to *z* (in tC ha^{-1}) as a function of *x* and *y*; the color of the membrane represent the relative value of *z* (in %), i.e. the proportion of one soil carbon pool within the total soil carbon stock. No color is assigned to the membranes in the last row, because the relative value is 100 %. Blue lollipops denote the standard deviations of the simulated mean z (on the membrane surface) given each (x, y) locations, which follow a systematic distribution.

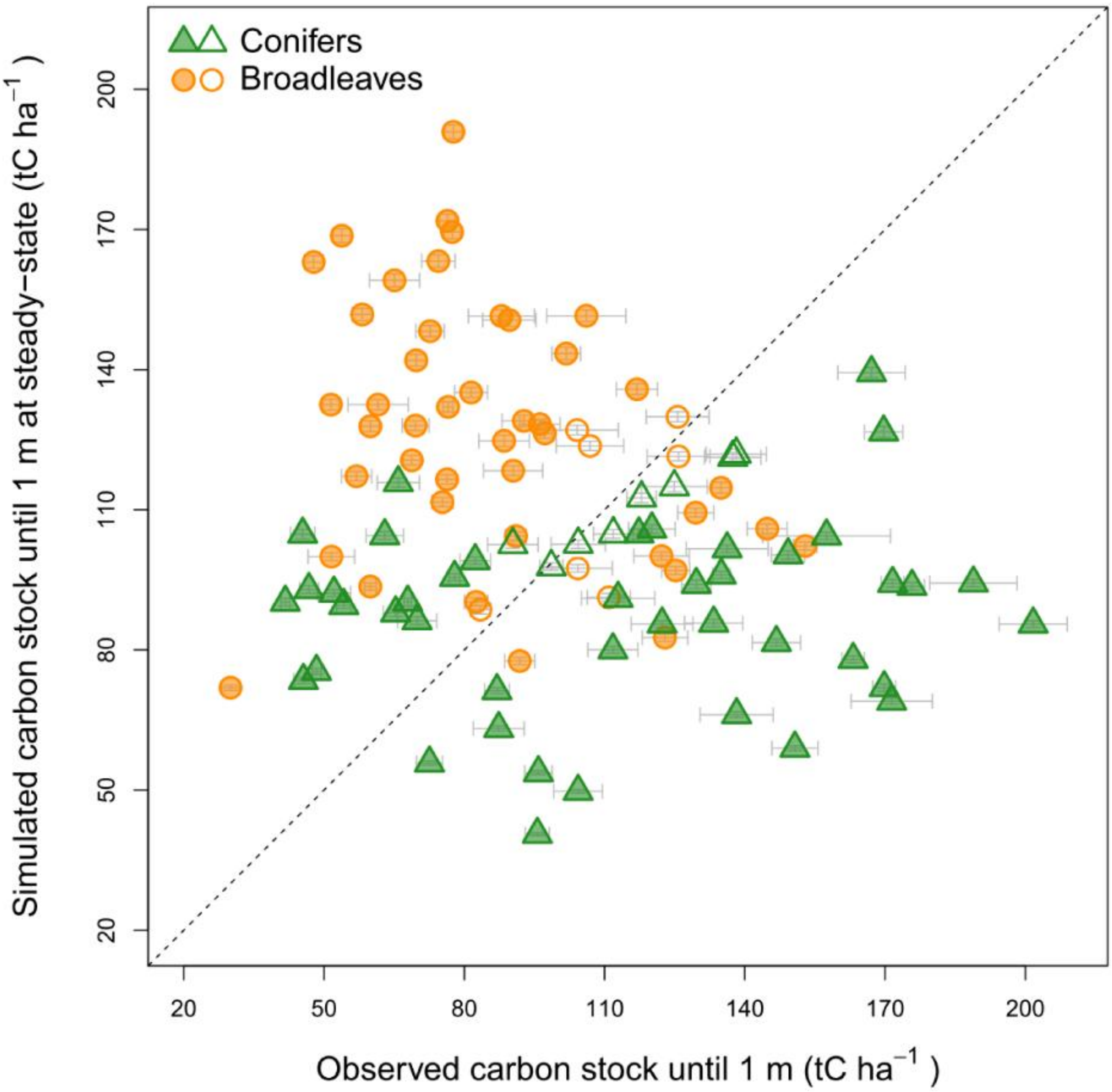


Figure ~~S4~~ S5 Comparing the ~~ison between simulated and observed annual~~ steady-state carbon stock until 1 m (CS, in tC ha^{-1}) with the observed carbon stock until 1 m at t1, which were used for model input. Round and triangle symbols represent sites dominated by broadleaves and conifers, respectively. The chosen fine root:foliage ratio for broadleaves and conifers is 1.0. Error bars represent standard errors; hollow and filled points represent non-significant and significant differences between simulated and observed ACC according to t-test (at 95% confidence level).

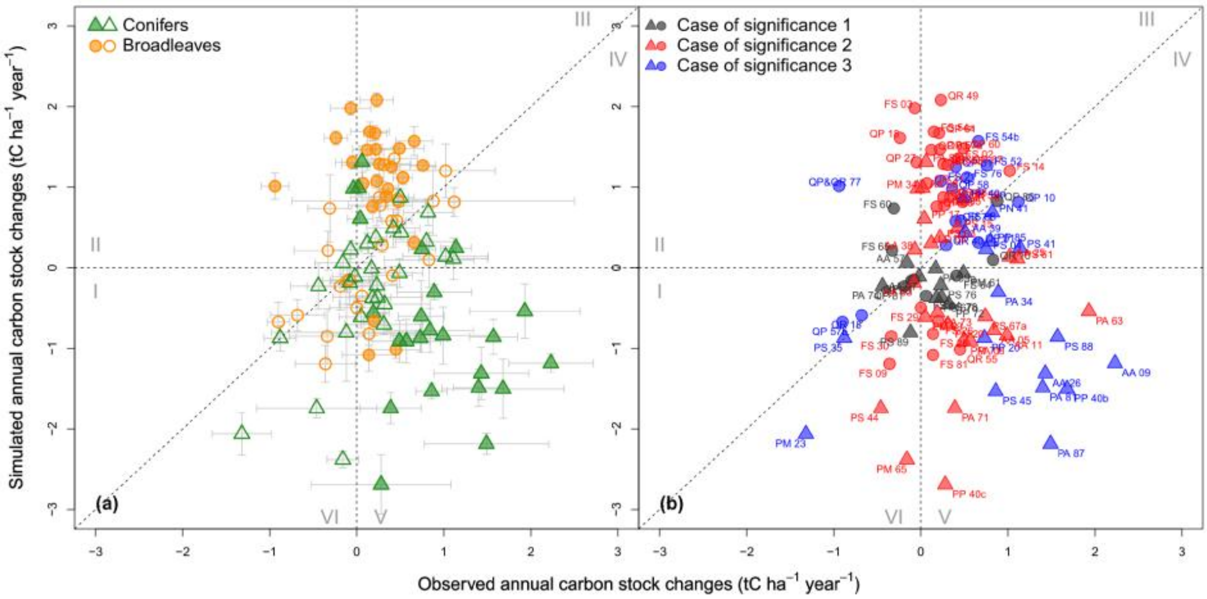


Figure S6 Comparison between simulated and observed annual carbon stock changes (ACC, in tC ha^{-1} year^{-1}). Round and triangle symbols represent sites dominated by broadleaves and conifers, respectively. The complete steady-state assumption was used for initializing carbon quality of the stock until 1.0 m. The chosen fine root:foliage ratio for broadleaves and conifers is 1.0. To facilitate discussions, we set Roman numbers (I-VI) denoting the six zones in which data points are distributed. In (a), error bars represent standard errors; hollow and filled points represent non-significant and significant differences between simulated and observed ACC according to t-test (at 95% confidence level). In (b), case of significance: 1 – no significant difference from 0 for neither observed nor simulated ACC; 2 - a significant difference from 0 for either observed or simulated ACC and 3: - a significant difference from 0 for both observed and simulated ACC.

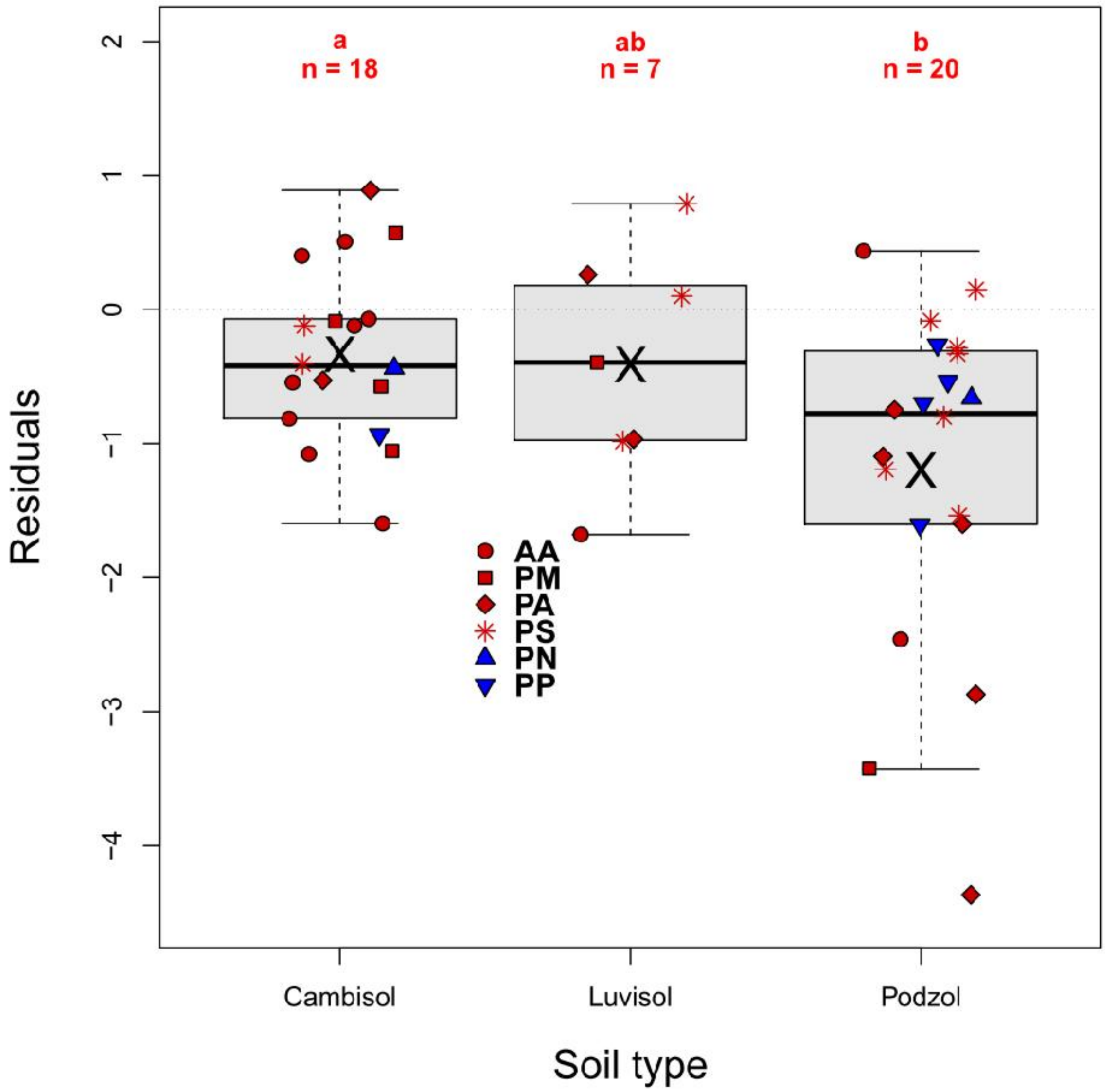


Figure S7~~5~~ Distributions of the residuals (simulated minus observed annual carbon stock changes) of Yasso07's fit for sites dominated by conifers. For each boxplot, the lower and top edge of the box corresponds to the 25th and 75th percentile data points; lower and top bars the line within the box represents the median; no outilier points in this case. "X" indicates mean values: -0.33 ± 0.15 (mean ± standard error) for cambisol,-0.41 ± 0.32 for luvisol and -1.20 ± 0.28 for podzol. Species accronyms: AA – *Abies alba*; PM – *Pseudotzuga menziesii*; PA – *Picea abies*; PS – *Pinus sylvestris*; PN – *Pinus nigra*; PP – *Pinus pinaster*. Letters above boxplots indicate diagnostics according to Tukey HSD test. Colors for different species: deep red for species that can be found for all the three types of soil; blue for species that can only found for cambisol and podzol, but not for luvisol.

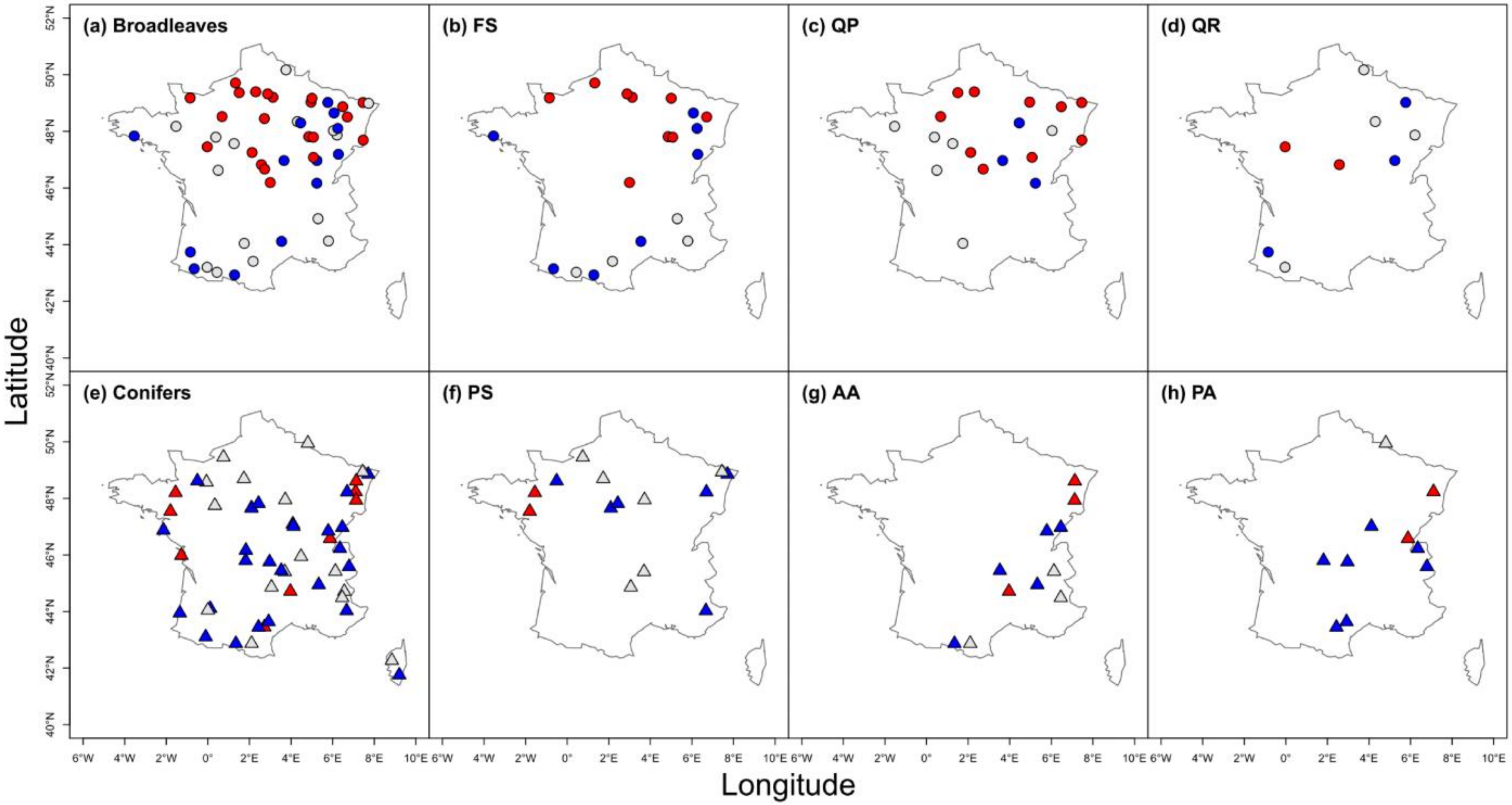


Figure ~~S6~~ S8 Spatial visualization of residuals (i.e. the difference between simulated and observed annual carbon changes) for sites dominated by broadleaves (a) and conifers (b). Colors: red – overestimation with residuals being significantly > 0; blue – underestimation with residual being significantly < 0; grey – residuals that are not significantly different from 0. Species abbreviations: FS – *Fagus sylvatica*; QP– *Quercus petraea*; QR - *Quercus robur* (*including* two mixed *Quercus* sites); PS - *Pinus sylvestris*; AA- *Abies alba*; PA - *Picea abies*.

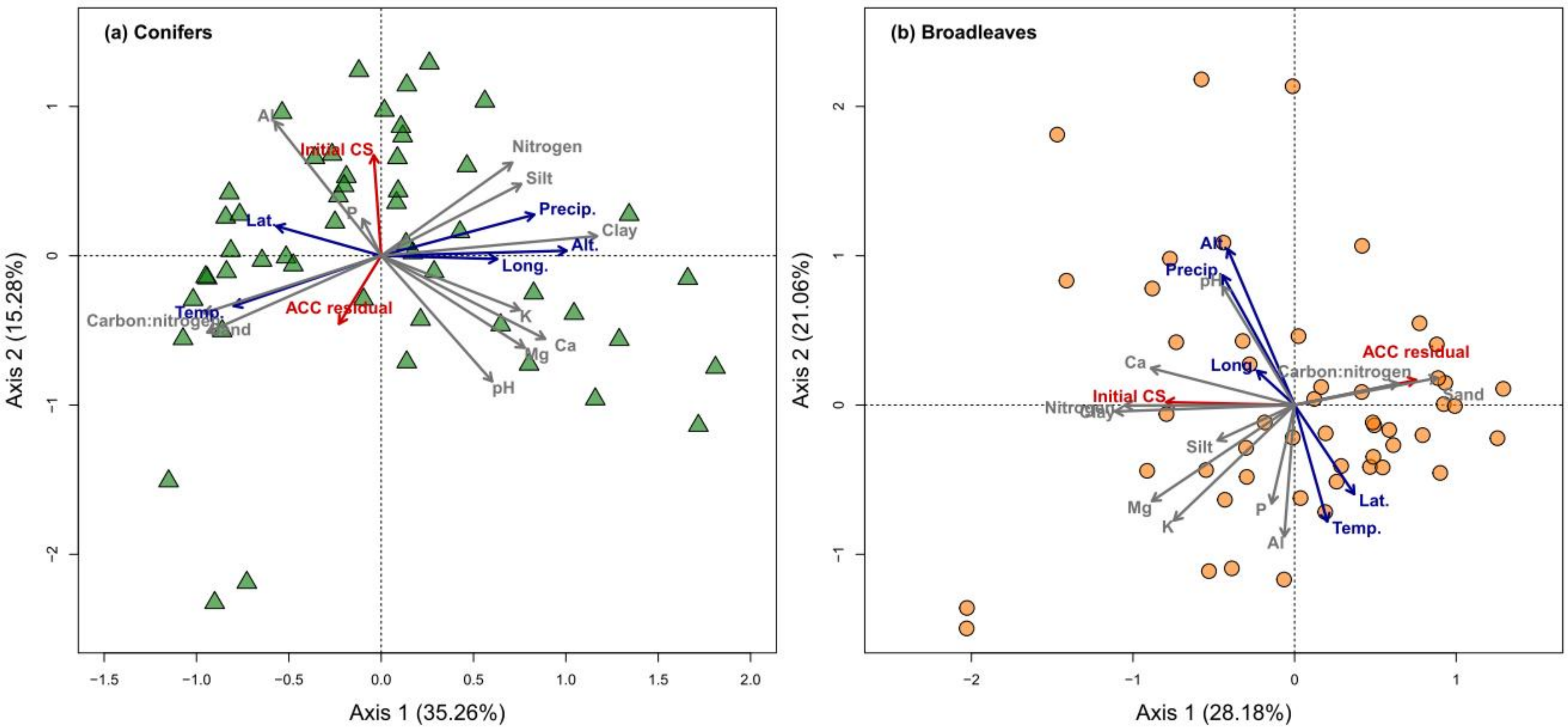


Figure S9~~7~~ Relationships among indicators of site features and model predicts using principal component analyses for sites dominated by conifers (a) and broadleaves (b), respectively. Colours of arrows: red – residuals of annual carbon change and observed initial carbon stock; grey – soil physical and chemical properties; blue –site geographical and climatic variables. Each point corresponds to one RENECOFOR site. See Table S2 for full names of soil properties.