Reply to Anonymous Referee #1

This manuscript aims to test the soil carbon model Yasso20 against a national scale dataset of soil carbon stock in Swiss forests. The paper is well written with clear objectives and generally sound methods (but see below). The core result, that is the soil mineral characteristics are primary drivers of soil carbon stock and Yasso20's predictive bias, is interesting for the whole modelling community.

Author's response

We thank the Referee for the constructive comments and suggestions, which helped to improve the quality and clarity of our manuscript. We have considered each of your remarks and modified the manuscript accordingly. You can find below our responses to your comments.

Reviewer comment

The main weakness of the manuscript comes to the lack of details about the forest context, litter and soil carbon estimation and modelling procedures:

More detailed information on forest management (LN87) and disturbance history (LN88) would be appreciated. For example, how were tree branches harvested? (Models usually ask for this kind of size thresholds that can play a role in decomposition rates). Western Europe, including Switzerland, is known to have suffered from storms in the past, including the Lothar in 1999 (see https://www.wsl.ch/en/news/25-years-after-lothar-how-the-windstorm-rebuilt-the-forest/). So none of these forest plots were attacked?

Author's response

Thank you for your remarks. Our study focuses on forests older than 120 years (old-growth forests, identified by the use of historic maps; Gosheva et al. (2017)). As the soil sampling was performed on undisturbed sites, none of the sampled 556 sites was impacted by the storms Vivian (in 1990) or Lothar (in 1999). Although most of the Swiss forests are regularly exploited (62% of the forests, up 79% in the Swiss Plateau), the harvesting practices normally consist of planned loggings of single trees with harvesting residues (typically the smaller branches and tree top, foliage and belowground part with trunk to approx. 30 cm aboveground) normally left in the stand (Brändli et al., 2020). This keeps the disturbance by management interventions at a minimum. We are aware that natural disturbances from windthrows (i.e. Vivian and Lothar) in Swiss forests led to SOC losses especially in the organic layers of forests at high elevation (1 to 2 kg C m^{-2}) (Mayer et al., 2023), but none of the sampled sites were impacted.

We have now added this information at lines 87-88 in the manuscript:

"*In combination with the typically small-scale forest management in Switzerland, which include planned loggings of single trees with harvesting residues (e.g. fine branches, foliage, belowground part with trunk to approx. 30 cm aboveground) normally left in the stand (Brändli et al., 2020), the focus on old forests ensures that there have been only reduced* *disturbances in the forest cover over the past decades. Although natural disturbances such as windthrows occurred in Swiss forests (e.g. Vivian in 1990 and Lothar in 1999) with possible impact on SOC stocks (Mayer et al., 2023), none of the forest sites was affected as sampling excluded windthrow sites."*

Reviewer comment

More detailed information should be provided on how the NPP (LN150-1) was obtained (please avoid only citing other papers) and, especially, how it was converted into the estimate of annual litter input. Since litter input is derived from model outputs, so what is the suitability and prediction quality of the model (Terra and Aqua MODIS-satellite) for Swiss forests? This is not mentioned in LN152-153. As a result, one might ask about the reliability of the input, especially when reading LN410-411.

Author's response

Thank you for your comment. Since no stand inventory or litter input measurements were available at the investigated soil sites, we obtained the NPP from Terra and Aqua MODIS satellite (500-m resolution). We considered the NPP as a proxy of long-term litter C input to the soil, which was used to simulate SOC stocks at steady state with Yasso (see line 150). This is consistent with the approach of published SOC model applications (Abramoff et al., 2022; Pierson et al., 2022), and with the current Yasso20 calibration (Viskari et al., 2022), where the litter input was similarly obtained from the GPP product of Terra and Aqua MODIS satellite, setting a maximum NPP to GPP ratio to 0.5. In our study, we also considered a maximum NPP to GPP ratio of 0.5, which agrees well with the NPP to GPP ratio (0.42-0.49) observed at two contrasting forest ecosystems in Switzerland (Etzold et al., 2011).

We also tested the suitability of Terra and Aqua MODIS-satellite to estimate NPP at Swiss forests by comparing the NPP obtained by Terra and Aqua MODIS-satellite with NPP derived by stand inventory data at 18 Swiss forest sites of the long-term forest ecosystem research LWF (Etzold et al., 2014). On average for the period 2001-2010, the mean NPP estimated by the satellite approach was 0.49±0.04 kg C m^{-2} yr^{-1}, while a mean NPP of 0.46±0.05 kg C m^{-2} yr^{-1} was estimated by the terrestrial approach for the same period (Etzold et al., 2014). In addition, the satellite NPP proves to be a reasonable proxy of large range of forest productivity across Swiss forests (see new **Fig. S1a** below) that is consistent with the gross volume increment trends across Swiss forest regions shown by the Swiss NFI (Brändli et al., 2020).

Taken together, these results confirm the suitability of the NPP obtained Terra and Aqua MODIS-satellite as a proxy of long-term litter inputs for Swiss forest conditions.

We have now expanded and clarified in the manuscript the limitations of the approaches (satellite and terrestrial) to estimate litter inputs at line 413 of the discussion, and added Figure S1a to the supplement:

"*Satellite-derived NPP is here used as input for Yasso simulations of SOC stocks at steady state. Since direct measurements of forest stands and detailed information of soil C inputs are* *often lacking at larger scales - as in this study - the use of NPP as a proxy of long-term litter C input to the soil is consistent with SOC model applications at the regional and global scales (Abramoff et al., 2022; Pierson et al., 2022), as well as in the calibration of Yasso20 (Viskari et al., 2022). We are aware that litter input derived from satellites can be uncertain, thus potentially contributing to the observed discrepancies between simulated and measured SOC stocks at the site level. In fact, the fine scale variability in litter inputs cannot be captured by satellite-derived NPP estimates given (1) the larger pixel size of MODIS (500 m x 500 m) compared to the site scale of the soil sampling, and (2) the partitioning into tree components using average allocation factors, due to the lack of data at the site level. Satellite-derived NPP may have resulted in an overestimation of the litter input in regions with intensive forest management as in the Plateau, since small-scale disturbances such as thinning are not well detected by satellite estimates (Neumann et al., 2015; Park et al., 2021). Lastly, forests allocate a portion of NPP not only to fast-cycling components that are annually returned to the soil (i.e. fine roots and foliage) but also to components with slower turnover time such as stems and branches. Nevertheless, the satellite NPP approach proves to be a reasonable proxy of the large range of forest productivity across Swiss forests, i.e. ranging from 0.3 kg C m^{-2} yr^{-1} in the Alps to 0.8 kg C m^{-2} yr^{-1} at the warmest sites (see **Fig. S1a**), which is consistent with the gross volume increment shown across different regions in the Swiss NFI (Brändli et al., 2020). Moreover, at the 18 sites of the long-term forest monitoring program LWF, the mean NPP over the period 2001-2010 based on the satellite approach amounted to 0.49±0.04 kg C m^{-2} yr^{-1} as compared to 0.46±0.05 kg C m^{-2} yr^{-1} estimated by the terrestrial approach for the same period (Etzold et al., 2014).*

Similarly, terrestrial methods based on forest inventories may lead to uncertain estimates of litter inputs. These uncertainties mostly relate to (1) country-specific allometries and expansion factors used to estimate tree biomass, (2) turnover times applied to obtain the litter inputs, and (3) failing to appropriately estimate inputs from fine roots and understory vegetation, which remain severely unconstrained despite their major contribution to forest soil C inputs (Didion, 2020; Neumann et al., 2020).”

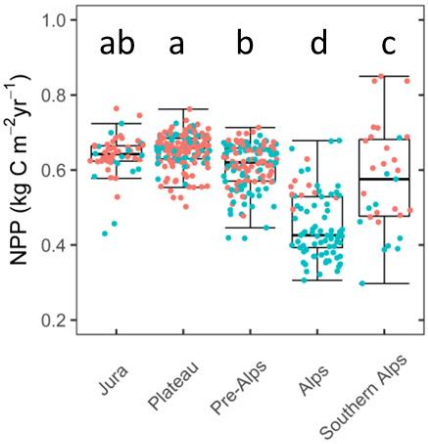


***Fig. S1a**. Net primary production (NPP) across Swiss forest regions, excluding waterlogged soils. Total n sites = 468. Letters indicate significantly different across regions, based on ANOVA followed by Tukey’s test with $P < 0.05$.*

Reviewer comment

Then, how was tree litter chemical quality split? It is not very clear in the text (LN158-160). Was it partitioned at the level of tree species or at the level of tree functional type (broadleaved versus coniferous)?

Also, so a fixed set of allocation factors were used to estimate tree component (LN155) without distinguishing between tree functional type?

Author's response

We appreciate your remarks and suggestions. The litter chemical quality was split into AWEN fractions based on the percentage of broadleaf and conifer species at each site recorded by field assessments (line 115-118). First "*we partitioned the NPP into broadleaf and conifer species, multiplying the NPP by the percentage of broadleaf and conifer species recorded by field assessments at each site*" (line 153). Then the NPP – after being partitioned into broadleaf-derived and conifer-derived NPP - was split into AWEN fractions using the functions and approach used in the Swiss GHG inventory (Didion, 2023) based on measured fractions in the long-term forest monitoring program LWF (Didion et al., 2014).

We have now clarified the methodology used at lines 155 and 158-160 in the manuscript:

"*The C inputs for each pool were separated into the four AWEN components based on the percentage of broadleaf and conifer species at each site, using the functions and approach adopted in the Swiss GHG inventory (Didion, 2023) based on measured fractions at sites of the long-term forest monitoring program LWF (Didion et al., 2014).*"

In this study, the estimates of NPP allocation into tree components rely on long-term forest ecosystem data from LWF network (Etzold et al., 2014), see lines 154-157. We did not distinguish between tree species or tree functional types, since we did not observe major differences between different tree functional types in C allocation to different tree compartments. In the future, this approach can be further refined by establishing species-specific C allocation factors, which should also consider the effect of climatic conditions on C allocation.

Reviewer comment

I find the categorization (LN119) risky. In this case, stands of 40-49% broadleaf and stands of 51-60% broadleaf which are actually quite similar stands in term of species mixture, will belong to two contrasting categories. I wonder if a new category of "mixed stand" should be proposed and compared with more pure stands? Also, how was the chemical quality of the litter in these mixed stands partitioned? As a reader, I am also curious to know how big a difference that the effect of species functional type can theoretically make in Yasso20 simulations in your case? (According to Fig. 3b, there is limited difference in the model's predicts). Would it be possible to do an uncertainty analysis on how the 0%, 50% and 100% of broadleaves at a given site (with given climatic inputs) can change the predictions? This can be done for all sites or only for those where both broadleaves and conifers coexist, enabling us to better understand the sensitivity of the species related parameters used in your case to the model predictions.

Author's response

Thank you for your suggestions. In the manuscript, we have separated forest sites into broadleaf and coniferous forests only for visualization purposes in the Figures (e.g. Figure 2-4). The percentage of broadleaf species recorded at each site has already been considered for partitioning NPP input at each site, which is then used as model input for Yasso20 simulations. We have specified this at line 152-154: "*We partitioned the NPP into broadleaf and conifer species, multiplying the NPP by the percentage of broadleaf and conifer species recorded by field assessments at each site*".

We have now reformulated the sentence at line 118-119:

"*Only for visualization purposes, the forest sites were subdivided into two types based on the broadleaf percentage: coniferous (0-50%) and broadleaf (51-100%) forests.*"

Reviewer comment

Finally, in terms of data representation and visualization, the authors chose to present simulated versus observed C using bar plots (Fig. 3), box plots (Fig. S1) or scatter plots (deviation versus factors as in Fig. S3). While it is good, it would be more informative to directly show the scatter plots of simulated C versus observed C at the site level with a 1:1 line to see if and how the model captures the observations according to different tree functional types (and/or regions). A new (or supplementary) plot at the site level would be desirable as a complement to Fig. 3.

Author's response

We presented the simulated vs observed C stocks using bar plots to better visualize in which regions the simulated SOC stocks deviate more from the measured stocks, in order to infer on the main processes driving mineral stabilization.

We have now added the 1:1 plot of simulated vs observed C stocks in the supplemental material (see Fig. S2 below). As expected, the Yasso20-simulated values show only a smaller variability in SOC stocks compared to the measured SOC across sites, although the average simulated SOC stocks match well the average SOC stocks across Swiss forest soils. Possible reasons for this are: (1) the uncertain litter input estimation, since satellite-derived NPP was used as a proxy of site productivity, (2) high spatial variability of soil properties which drives very high SOC stocks at some sites (e.g. high Fe and Al), and (3) anaerobic conditions at the site scale that retard the SOC mineralization rates and lead to locally higher SOC stocks.

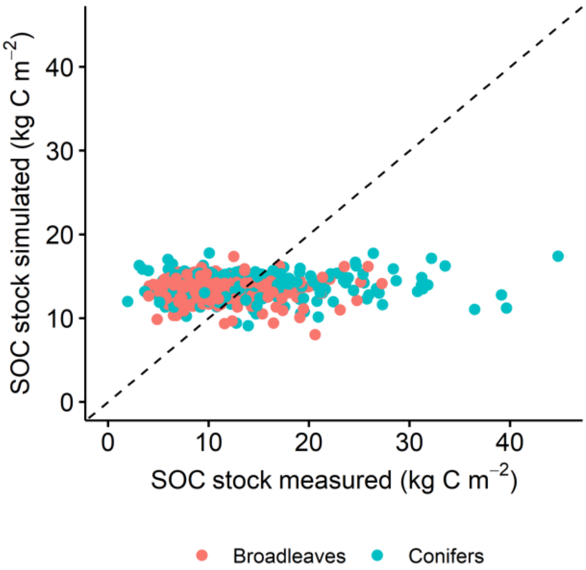


Fig. S2. Comparison between Yasso20-simulated and measured SOC stocks by forest types. Total n sites = 468, excluding waterlogged soils.

Specific points:

Reviewer comment

LN88: "minimal disturbances" was mentioned, but latter we read "frequent fires" (LN354). Fires are an important disturbance in the forest system. Is this a contradiction? What about storms (see above)?

Author's response

Thank you for the comment. We have included in the manuscript your remark that disturbances such as forest management or windthrows may affect soil C in Swiss forests at line 88 (*see previous Author's response*), although none of the sampled site was directly affected by windthrows. In Switzerland, forest fires - including wildfires during the dry season and deliberate fires common especially in the Southern Alps for over 10,000 years – contributed to increase SOC storage due to increased resistance of charred material to decomposition (Eckmeier et al., 2010). High amounts of charred material persisted in Fe- and Al-rich soils for millennia (Eckmeier et al., 2010). Here, we cannot exclude that the sampled sites have not been affected by forest fires over millennia time scales. Thus, we already discussed in the manuscript at line 353-358 how fire-derived black C may potentially affect soil C stock measured in the Southern Alps, showing that the main results remained unchanged even by excluding this region from the analysis.

Reviewer comment

LN125: how were the percentages of C content of the litter measured?

Author's response

The C content of ground organic materials (including the following horizons: undecomposed litter L, fermentation F, and humified H) and mineral soil were measured by dry combustion using an elemental analyzer (NC 2500, CE Instruments, Italy). This is reported at line 105-106.

To improve clarity, we have added in the manuscript at line 125 the following sentence:

"*multiplied by the percentage of C content obtained by dry combustion (see section 2.2)*."

Reviewer comment

LN129: it would be good to explain how the volumetric stone content was estimated.

Author's response

Thank you for the comment. **We have added this information in the manuscript at line 129**:

"volumetric stone content visually estimated from the soil profile (Richard et al., 1978)"

Reviewer comment

LN183: also with a standard deviation of 1 ?

Author's response

We have centered all independent variables before analysis to a mean of 0. We did not scale the variables to standard deviation of 1. This allows to provide in Table 2 and 3 (column Est.) meaningful model estimates, which can be interpreted in the same measurement units of the predictor and response variables.

We have replaced the word "standardized" with "centered" in the sentence at line 183 to improve clarity:

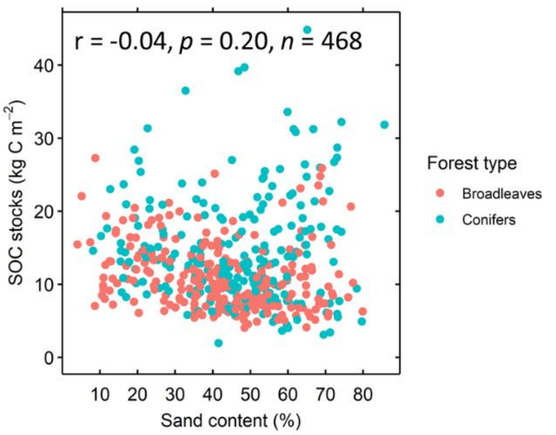
"*All explanatory variables were centered to a mean of 0*"

Reviewer comment

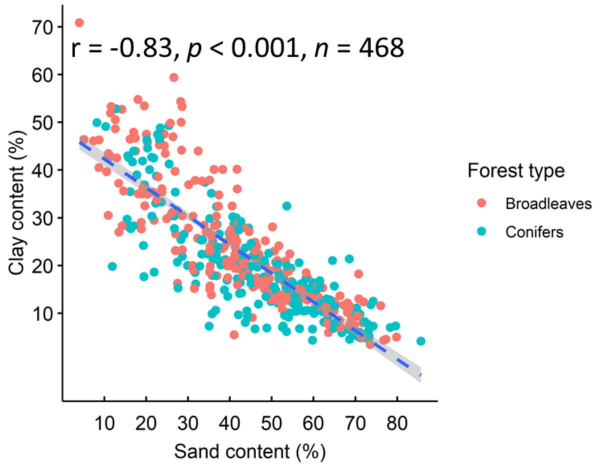
LN 320: very interesting result for clay. Have you also tried the sand content (which does not necessarily correlate with the clay content)?

Author's response

We tested whether the sand content would be significantly correlated to the SOC stocks, which we plot here below for your information. No significant correlation was found (where r is the Person correlation coefficient), similarly to the clay content (Fig. 4 in the MS):



In addition, a strong correlation of sand and clay content was observed in this study:



For these reasons, we have decided to include only the clay content (%) in the analysis presented in the manuscript.

Reviewer comment

Table 2 (and also other Tables that summarize statistics), I suggest using a MAP unit of 100 mm rather than mm to give a more meaningful sense for each increment of 1 unit.

Author's response

Thank you for your suggestion. **We will modify the Tables using a MAP unit of 100 mm instead of 1 mm.**

Reviewer comment

We also often see the important factor of "slope", but it is rarely mentioned. And why is that?

Author's response

We agree that the slope can be an important driver of SOC stocks. In fact, we have included slope as a predictor in our statistical analysis. In the MS at line 272-274, we report that slope has a negative effect on SOC stocks. To not increase further the length of our manuscript and considering that slope did not appear a dominant driver of SOC stocks in the different biogeographic regions (Table S5), we do not discuss about slope in more detail. However, topography represents an important factor influencing SOC stocks, with soil erosion occurring on steep slopes while waterlogging in flat areas (Fernández-Romero et al., 2014) which can strongly impact the variability of SOC stocks at the site-scale.

Reviewer comment

Figure 2: Why was a polynomial regression also used? Would it be possible to plot the litter input data in addition to NPP?

Author's response

We have shown a polynomial regression in Figure 2a (in addition to the linear correlation), since this better captures the productivity trend across elevation with NPP peaking at an elevation of about 700 m a.s.l., which then decreased with increasing elevation (line 210).

Since we do not have the litter input at the sites, we considered the NPP as a proxy of long-term litter carbon input to the soil, see previous *Author's response*. We have now plotted the NPP data in the new **Fig. S1a**, which show the variability in NPP across Swiss forest regions.

Reviewer comment

Figure 3: I have the impression that the Yasso20 estimates for all five regions are not very different and possibly not very significant in most regions. In (b) this may also be the case: does Yasso20 predict significantly different stocks between broadleaved and coniferous stands? It would also be interesting to read statistical diagnoses somewhere among the simulated (or observed) groups.

Author's response

We have now added different letters to the updated **Fig. 3a,b** (see below) which indicate significantly different means (1) among the measured SOC stocks (capital letters), and (2) among the Yasso20-simulated SOC stocks (lowercase letters).

Here, we show that Yasso20-simulated SOC stocks are statistically different across regions (**Fig. 3a**) and between forest types (see **Fig. 3b**), but differences are much smaller compared to differences observed for measured SOC stocks. Yasso20-simulated SOC stocks across regions show a different pattern compared to observed SOC stocks, with Jura and Pre-Alps > Southern Alps. We regard the smaller variation of SOC stocks modelled by Yasso20 as an indication for the lacking representation of region-specific soil characteristics such as mineralogy.

We have modified Fig. 3a,b and its caption, which now includes also the comparison among measured (capital letters), and among simulated SOC stocks (lower case letters):

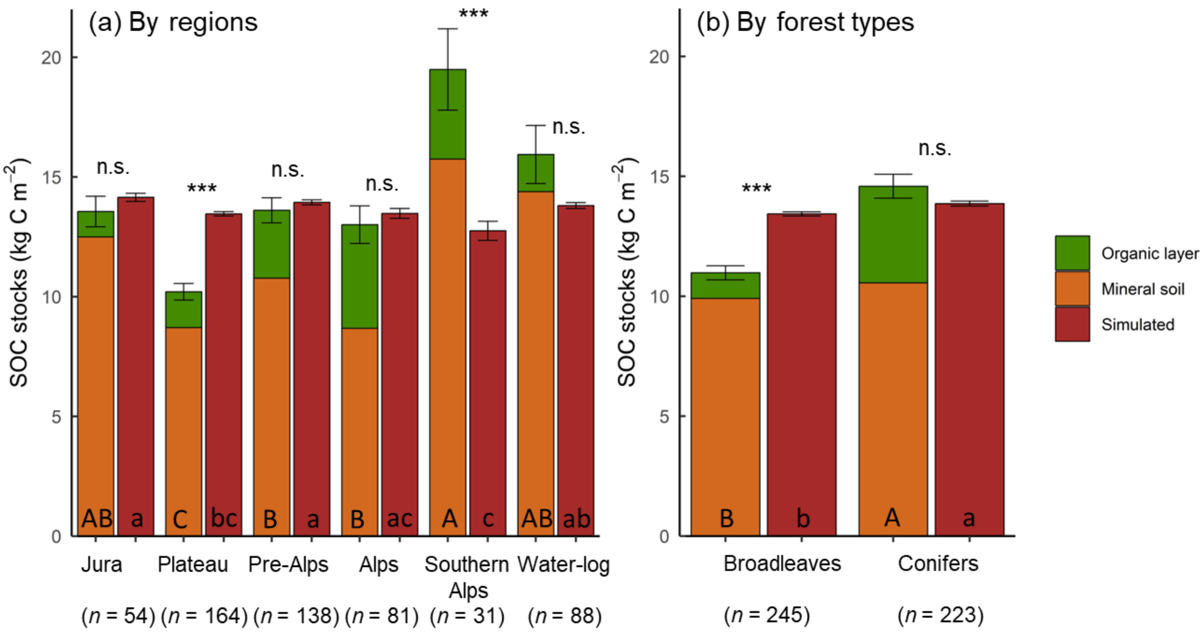


Figure 3: Comparison between measured SOC stocks vs Yasso20-simulated stocks by: (a) biogeographic regions of Switzerland (waterlogged soils shown separately), and (b) forest types, excluding waterlogged soils. The measured stocks are shown as organic layers and mineral soils to 100 cm depth (total *n* sites = 556; excluding waterlogged soils, *n* sites = 468). The simulated stocks at each site are based on the mean of 500 replicate simulations representing model parameters uncertainty. SOC stocks are represented as means ± standard errors. *P*-values are calculated with Welch's *t*-tests: $P \geq 0.05$ (n.s., not significant), $P < 0.001$ (***). Capital letters indicate significantly different means among measured SOC stocks, while lowercase letters indicate significantly different means among Yasso20-simulated SOC stocks, based on ANOVA followed by Tukey's test (Figure 3a) and Welch's t-tests (Figure 3b) with $P < 0.05$.

Reviewer comment

Fig. 4: It would be good to briefly explain why such a subjective cut-off point of pH = 5 was chosen for data analysis. It is particularly useful for non-soil experts, although it is intuitively logical for soil experts.

Author's response

Thank you for the remark. **This clarification will be added to the revised manuscript at line 254-255:**

"*We separated the dataset into soils with pH ≤ 5 and with pH > 5, since pH is recognized as a predictor of SOC stabilization mechanisms, mediated by Al- or Fe- under acidic conditions while Ca exerts a dominant influence under increasing pH (Rowley et al., 2018)."*

References - Reply to Anonymous Referee #1

Abramoff, R.Z., Guenet, B., Zhang, H., Georgiou, K., Xu, X., Rossel, R.A.V., Yuan, W., Ciais, P., 2022. Improved global-scale predictions of soil carbon stocks with Millennial Version 2. Soil Biology and Biochemistry 164, 108466.

Brändli, U.-B., Abegg, M., Allgaier, B.L., 2020. Schweizerisches Landesforstinventar: Ergebnisse der vierten Erhebung 2009-2017. WSL.

Didion, M., 2020. Extending harmonized national forest inventory herb layer vegetation cover observations to derive comprehensive biomass estimates. Forest Ecosystems 7, 1-14.

Didion, M., 2023. Data on soil carbon stock change, carbon stock and stock change in surface litter and in coarse dead wood prepared for the Swiss NIR 2024 (GHGI 1990–2022).

Didion, M., Frey, B., Rogiers, N., Thürig, E., 2014. Validating tree litter decomposition in the Yasso07 carbon model. Ecological modelling 291, 58-68.

Eckmeier, E., Egli, M., Schmidt, M., Schlumpf, N., Nötzli, M., Minikus-Stary, N., Hagedorn, F., 2010. Preservation of fire-derived carbon compounds and sorptive stabilisation promote the accumulation of organic matter in black soils of the Southern Alps. Geoderma 159, 147-155.

Etzold, S., Ruehr, N.K., Zweifel, R., Dobbertin, M., Zingg, A., Pluess, P., Häsler, R., Eugster, W., Buchmann, N., 2011. The carbon balance of two contrasting mountain forest ecosystems in Switzerland: similar annual trends, but seasonal differences. Ecosystems 14, 1289-1309.

Etzold, S., Waldner, P., Thimonier, A., Schmitt, M., Dobbertin, M., 2014. Tree growth in Swiss forests between 1995 and 2010 in relation to climate and stand conditions: Recent disturbances matter. Forest Ecology and Management 311, 41-55.

Gosheva, S., Walthert, L., Niklaus, P.A., Zimmermann, S., Gimmi, U., Hagedorn, F., 2017. Reconstruction of historic forest cover changes indicates minor effects on carbon stocks in Swiss forest soils. Ecosystems 20, 1512-1528.

Mayer, M., Rusch, S., Didion, M., Baltensweiler, A., Walthert, L., Ranft, F., Rigling, A., Zimmermann, S., Hagedorn, F., 2023. Elevation dependent response of soil organic carbon stocks to forest windthrow. Science of the Total Environment 857, 159694.

Neumann, M., Godbold, D.L., Hirano, Y., Finér, L., 2020. Improving models of fine root carbon stocks and fluxes in European forests. Journal of Ecology 108, 496-514.

Neumann, M., Zhao, M., Kindermann, G., Hasenauer, H., 2015. Comparing MODIS net primary production estimates with terrestrial national forest inventory data in Austria. Remote Sensing 7, 3878-3906.

Park, J.H., Gan, J., Park, C., 2021. Discrepancies between global forest net primary productivity estimates derived from MODIS and forest inventory data and underlying factors. Remote Sensing 13, 1441.

Pierson, D., Lohse, K.A., Wieder, W.R., Patton, N.R., Facer, J., de Graaff, M.-A., Georgiou, K., Seyfried, M.S., Flerchinger, G., Will, R., 2022. Optimizing process-based models to predict current and future soil organic carbon stocks at high-resolution. Scientific Reports 12, 10824.

Richard, F., Lüscher, P., Strobel, T., 1978. Physikalische Eigenschaften von Böden der Schweiz. Eidgenössischen Anstalt für das forstliche Versuchswesen.

Rowley, M.C., Grand, S., Verrecchia, É.P., 2018. Calcium-mediated stabilisation of soil organic carbon. Biogeochemistry 137, 27-49.

Viskari, T., Pusa, J., Fer, I., Repo, A., Vira, J., Liski, J., 2022. Calibrating the soil organic carbon model Yasso20 with multiple datasets. Geoscientific Model Development 15, 1735-1752.