Soil Organic Carbon Stocks in Forests of Switzerland

Update of soil organic carbon stock estimation for the national greenhouse gas inventory

Dr. Madlene Nussbaum, Prof. Dr. Stéphane Burgos
18 August 2021 (revised version)

Bern University of Applied Sciences (BFH)
School for Agricultural, Forestry and Food Sciences (HAFL)

Contents

Abstract

The Swiss greenhouse gas inventory (GHGI), submitted under the United Nations Framework Convention on Climate Change and under the Kyoto Protocol, reports annually on soil organic carbon (SOC) stock in forests. First SOC stock estimates were published for Swiss forests by Perruchoud et al. (2000) and later by Nussbaum et al. (2012, 2014a) based on the extensive soil data base of the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL). Since then additional 1 060 sites were either sampled by WSL or soil profile data of comparable quality were assembled from projects of Cantons. The substantially larger dataset allowed to update the SOC stock currently used by GHGI.

Based on the new dataset, we modelled SOC stock in the mineral soil for two depth compartments: 1) mineral soil 0–30 cm and 2) mineral soil 0–100 cm. We used a comprehensive set of 177 covariates representing terrain, climate, vegetation and geological variation throughout the forests of Switzerland. Relations between SOC stock and covariates were established by a machine learning procedure called random forest (RF). A previous study showed good predictive performance of RF for other soil properties from the same forest soil dataset. RF is a procedure that combines a large number of regression trees and is capable of handling correlated covariates and of building non-linear dependencies. We used a covariate selection approach that removed covariates with low importance and high correlation to other important covariates. The final models retained 40 and 45 covariates for SOC stock in 0–30 cm and 0–100 cm, respectively. These models were used to compute SOC stock prediction maps for the forested area at a resolution of 25 m pixel width.

Model prediction accuracy compared to the same independent validation dataset of 175 sites used in 2012 remained about the same. Coefficient of determination R^2 was 0.30 (0–30 cm) and 0.40 (0–100 cm) indicating moderate model performance.

GHGI reports regional mean SOC stock estimates along with standard errors for altitude stratified production regions of the National Forest Inventory (NFI). Theses regions were further split by the categories productive and unproductive forests of Swiss Land Use Statistic (AREA). Besides, each 100x100 m block cell of the AREA forest categories were also used.

SOC stock estimates per stratified NFI region and for each AREA block cell were computed as mean of the predicted 25 m-pixels within this region. Standard errors had to be approximated because RF does not consider the local similarities of predicted values (spatial autocorrelation) within a region. We used an approximation approach that corrects the RF standard errors by a simulation of spatial autocorrelation.

The total SOC stock for Swiss forests was estimated to 81.4 t ha^{-1} (SE 3.0) in 0–30 cm and to 123.3 t ha^{-1} (SE 4.9) in the first 100 cm of the mineral soil. These estimates compared well to the previous study of 2012. Estimates for the altitude stratified NFI regions differed where there were a large number of newly sampled sites (Central Plateau) or where large SOC stock and simultaneously low observation density was found (Southern Alps).

1 Introduction

The Swiss greenhouse gas inventory (GHGI), submitted under the United Nations Framework Convention on Climate Change and under the Kyoto Protocol, reports annually on changes of organic carbon (OC) stock in forests. The National Forest Inventory (NFI) offers comprehensive data to quantify the living and dead forest biomass and its change in time. Estimating stock of soil organic carbon (SOC) in forests is more difficult because the variables needed to quantify stock vary strongly in space and precise quantification of some of them is very costly. First SOC stock estimates were published for the Swiss forests by Perruchoud et al. (2000) and later by Nussbaum et al. (2012, 2014a) based on the extensive soil data base of the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL).

Since then additional sites were sampled by WSL and soil profile data of comparable quality from projects of Cantons have been assembled. Additional 1 060 locations became available to allow an update of the calculations by Nussbaum et al. (2012, Table 5 based on 1 033 locations) currently used in the GHGI.

The project BOWA-CH (Walthert et al., 2015; Hertzog, 2017; Nussbaum et al., 2019; Baltensweiler et al., 2021, submitted) used this data set to create overview maps of a wide range of soil properties at a resolution of 25 m pixel width. The soil properties relevant to calculate SOC stock were: SOC content, gravel content and density of the fine soil fraction ($\leq$ 2 mm) all predicted at depth intervals of 0–5, 5–15, 15–30, 30–60 and 60–100 cm.

The spatial prediction were finally computed by machine learning (ML) approaches. The used ML methods were able to handle large sets of correlated covariates and allowed for batch processing for many responses. Moreover, superior model performance was expected (e.g. Nussbaum et al., 2018) compared to the external-drift kriging used by Nussbaum et al. (2012). Baltensweiler et al. (2021, submitted) compared six ML approaches and the combination of these in a model averaging scheme. Among these six approaches was an automatic model selection procedure for the same type of external-drift kriging as used in 2012. The algorithm random forest (RF) overall outperformed the other approaches somewhat, hence RF models were used to computed the predictive maps.

RF allowed directly to quantify local uncertainty by using quantile regression forest, a generalization of RF (Meinshausen, 2006). Baltensweiler et al. (2021, submitted) computed 90 %-prediction intervals for each predicted grid node (pixel) of the final map. GHGI reporting needs regional aggregation of these pointwise uncertainties. Nussbaum et al. (2012) applied block kriging that integrated local variances predicted by external-drift kriging. Block kriging mean and standard error estimates allowed to consider the spatial autocorrelation within the regions.

Opposed to kriging approaches, RF is not *per se* a spatial method and does not directly model spatial autocorrelation. Each prediction and its predictive distribution is generated independently from nearby locations. To overcome the lack of modelled spatial autocorrelation an approximation to estimate regional standard errors had to be found.

The current study aimed to

1. create spatial predictions for SOC stock for 0–30 and 0–100 cm for the Swiss forested area following the approach of Baltensweiler et al. (2021, submitted) and by using the soil dataset used therein,

2. aggregate mean predictions of the following regions

   – production regions of the national forest inventory (NFI) stratified by altitude as in Nussbaum et al. (2012),
   – NFI production regions subdivided into the GHGI reporting classes 'productive forests' (CC12) and 'unproductive forests' (CC13),[1]
   – each 100 m-raster cell of the Swiss Land Use Statistic (AREA) assigned to GHGI classes 'reforestations' (CC11), 'productive forests' (CC12) and 'unproductive forests' (CC13),

3. approximate standard errors to the regional mean predictions from quantile regression forest predictive distributions for the same regions as in (2).

In the subsequent text we refer to the previous GHGI SOC stock estimates described in the technical report by Nussbaum et al. (2012), then published as peer-reviewed article (Nussbaum et al., 2014b) and eventually included in a PhD thesis (Nussbaum, 2017, chapter 5) by *study or estimates from 2012*.

---

[1]The GHGI reporting class CC11 (reforestations) was dropped due to its small area covered.

2 Materials and Methods

2.1 Mapping target area and 100x100 m AREA block cells

The study area represented the forested area of Switzerland as defined by the National Forest Inventory (NFI, Brändli et al., 2020) covering about 30 % of the surface of Switzerland. A brief pedological description of the study area can be found in Nussbaum et al. (2012, 2014b) and Baltensweiler et al. (2021, submitted).

Spatial predictions were calculated on a 25 m grid with the centres of these grid nodes being inside the NFI forest definition. The aggregation on 100x100 m block cells was done based on the Swiss Land Use Statistic Version AREA04/09 for the Kyoto09-Codes CC11 (919 cells), CC12 (1 132 603 cells) and CC13 (106 130 cells, data delivery Martin Lindenmann, Sigmaplan, 02.09.2020, filename *AREA_sample_Befliegung3_Juli_2020.csv*). As the forest definition of NFI and AREA differ, not each AREA cell contained predictions, hence no mean SOC stock could be calculated (2.8 % of AREA cells).

2.2 SOC stock data

The 1 033 locations used in Nussbaum et al. (2012) were complemented by 1 060 locations (Hertzog, 2017; Baltensweiler et al., 2021, submitted). About 45 % of new sites (475) were located on the Central Plateau below 600 m asl (Figure 1, Table 4). Another 80–100 sites were each at medium altitudes on the Central Plateau, lower altitudes in the Jura mountains and above 1 200 m asl in the Alps. Although a majority of new locations was on lower altitudes the dataset had been complemented throughout the forested area of Switzerland including many remote locations.

On average, there were about 16.5 observations per 100 km^2 forested area. Soil sampling locations are, however, clustered on the Central Plateau while for example Western parts of Jura and upper valleys in Ticino remain underrepresented in the dataset.

SOC stock was again calculated for each soil profile location as described in Nussbaum et al. (2012, section 2.2.1). For a description of the soil properties SOC content, density of fine soil fraction and gravel content we refer to Baltensweiler et al. (2021, submitted). Organic horizons (L, O, F, H) were excluded from the SOC stock calculations. Only few new samples with SOC measurements were available to allow for a renewal of stock estimates compared to 2012. Sites with peat layers (T horizons) were rare as well (12 sites) and were excluded.

Table 1 shows summary statistics for the validation and calibration sites of each response. As in 2012 the calibration datasets had a wider data range compared to the SOC stock used for model validation. The mean of the calibration set was smaller compared to 2012. Stocks in both depth were strongly correlated (Spearman correlation r = 0.90). Both responses were positively skewed (skewness 2.2 and 3.1, respectively).

We refrained from response transformation with e.g. natural logarithm. Random forest (RF, subsection 2.5) does not assume Gaussian error distributions and is less sensitive to skewed data than linear regression based approaches. Tests on similar data sets showed no advantage of log-transformation (Spiess, 2016) and correct unbiased back-transformation remains unclear for RF predictions.

Table 1: Descriptive statistics of SOC stock calculated for the mineral topsoil (0–30 cm) and the mineral soil to 100 cm depth [t ha^{-1}] (vs: validation data set n = 175, cs: calibration data set n = 1 918).

	SOC 0–30 cm		SOC 0–100 cm	
	cs	vs	cs	vs
minimum	0.00	26.56	0.00	27.95
maximum	394.05	226.01	963.92	569.68
mean	63.52	80.91	92.91	132.92
median	52.89	73.01	76.47	111.84
standard deviation	41.54	41.23	67.79	83.30

2.3 Selection of validation set

Baltensweiler et al. (2021, submitted) used the validation dataset formed by Hertzog (2017). The validation sites were chosen from the enlarged dataset with the same approach as in Nussbaum et al. (2012). Isolated profile locations had higher probability to be selected as opposed to spatially clustered soil profile sites.
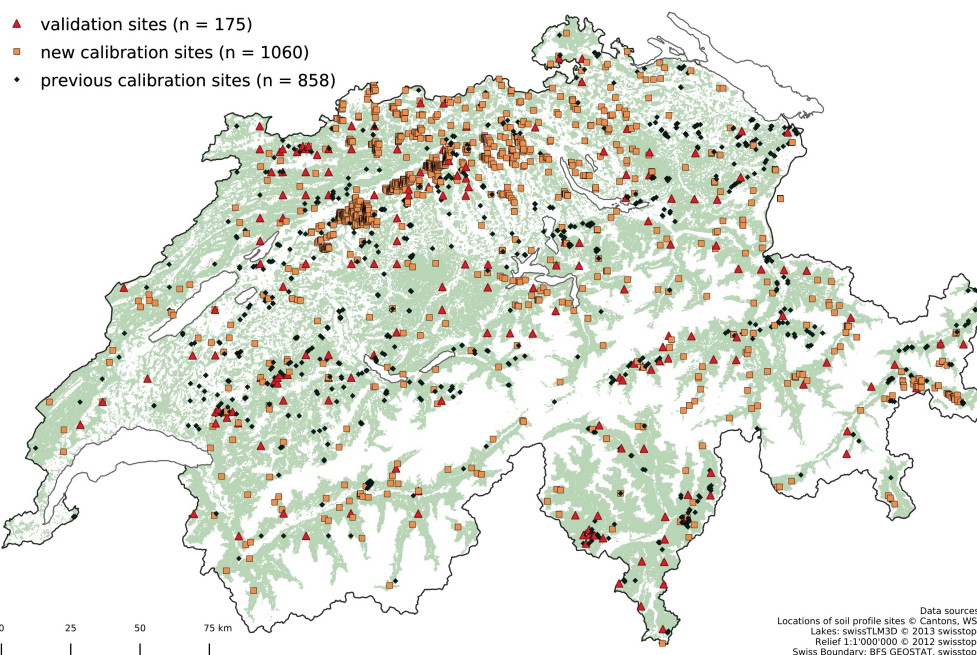
Figure 1: Locations of the 2 093 soil profiles and Swiss forest area (green) subdivided into previous and new calibration and validation sets.

To allow for a direct comparison of the model results to the previous estimates for the GHGI we decided to use the same validation set of 175 locations as employed in the study of 2012. The dataset for model calibration comprised the remaining n = 1 918 sites.

In 2012 the model selection was based on the calibration sites. Once the model structure had been found the model parameter (regression coefficients and variogram parameters) were re-estimated with the combined data of calibration and validation sites. We refrained from using the validation sites for parameter re-estimation in the present study. RF selected the covariates and built the model structure directly from the calibration data. Using the combined dataset would result in different non-validated model structure. For two subregions in the Southern Alps there were less calibration sites available than in 2012 (Table 4) by omitting the validation sites for model re-calibration.

2.4 Environmental covariates

A comprehensive set of environmental covariates to describe the most important soil forming factors was assembled by Baltensweiler et al. (2021, submitted). We used the exact same covariates as model input as Baltensweiler et al. (2021, submitted) and refer to this article for more details.

The numerous derivates calculated led to 88 covariates representing topography, 52 with climate information, 18 characterizing vegetation and 14 representing parent material, soil legacy information and landscape types. With including rotated coordinate axis and the sampling year we started model building with 177 covariates.

All covariates were prepared as raster datasets at a resolution of 5x5 m, regardless of the original resolution. The predictions were then calculated for every fifth pixel resulting in a 25 m-grid allowing to retain the detail of the high resolution covariates (e.g. terrain model).

2.5 Model fitting

Baltensweiler et al. (2021, submitted) predicted the input soil properties to calculate SOC stock. We refrained from combining these maps and compute SOC stock at each pixel. We calculated SOC stock at each soil profile location and fitted models directly to SOC stock for the following reasons: 1) Due to the different validation set these maps were calibrated on a different dataset. 2) The combination of non-parametric error distributions computed by quantile regression forest for each soil property is not straightforward and requires a simulation approach. Direct estimation of predictive distributions of SOC stock for each prediction grid node (pixel) is preferred. Hence, models were directly fitted to two responses calculated at each profile location: 1) SOC stock in 0–30 cm soil depth and 2) SOC stock in 0–100 cm soil depth.

In Baltensweiler et al. (2021, submitted) six different statistical prediction methods were tested on the same soil data set. Moreover, a model averaging scheme was implemented that combined the predictions of these six approaches. Model averaging often outperforms predictions from a single method (e.g. Nussbaum et al., 2018). Baltensweiler et al. (2021, submitted) did not find overall evidence for model averaging being an advantage and decided to use predictions computed by the best single method which was random forest (RF). We followed this conclusion and decided to use random forest to compute predictions for the present study.

Random forest combines a large number of regression trees. Two mechanisms are used to de-correlate trees and, consequently, reduce the variance of the predictions: 1) bootstrap sampling (bagging) creates a different responses datasets for each tree by repeatedly drawing a random sample with replacement from the calibration set and, 2) at each node only a predefined number of covariates (m_{try}) randomly selected of all covariates (p) are tested as candidates for binary splitting of the calibration data. The regression trees are fully grown up to $n_{nodesize}$ observations left in each leaf. Predictions are simple means of all n_{tree} fitted regression trees (Gareth et al., 2017).

To reduce the large number of covariates we applied sequential recursive backward elimination (Brungard et al., 2015; Hertzog, 2017) based on node-impurity covariate importance (Hastie et al., 2009b).

This approach evaluates the covariate importance for each RF fit and covariates with lowest importance are removed for the next smaller model fit. To speed up computation, we removed 5 to 10 covariates at each step fitting models with 167, 157, ..., 107, 102, ..., 7, 2 covariates. Optimal number of covariates was found by evaluating out-of-bag (OOB, see subsection 2.6) model error for each of these RF fits and using the model with lowest error. To additionally reduce the covariates we applied an ad-hoc decorrelation procedure that was performing similarly to a more complex approach (Hertzog, 2017).

The main tuning parameter of the algorithm is the number m_{try} of covariates tested at each split. For covariate removal we used as a default $m_{try} = p/3$. To find optimal m_{try} for the final covariate set we minimized OOB RMSE by iterating through $m_{try} = 1, 2, .., p$.

The implementation of RF differed in one aspect from Baltensweiler et al. (2021, submitted): we used the computational more efficient R package *ranger* instead of *randomForest* and *quantregForest*. In meantime, *ranger* supported the estimation of prediction intervals by quantile regression forest (Meinshausen, 2006).

2.6 Validation statistics

We used Bias (Nussbaum et al., 2018, formula 3), root-mean squared error (RMSE, Nussbaum et al., 2018, formula 4), the squared Pearson's correlation coefficient (r^2) and the coefficient of determination computed as mean squared error skill score (R^2, equals SS_{mse} in Nussbaum et al., 2018, formula 5) to evaluate model performance. Nussbaum et al. (2012) calculated R^2 as the squared Pearson's correlation coefficient (r^2) which report the linear relationship while R^2 reflects the relation to the 1:1-line indicating a perfect fit. The larger the difference between R^2 and r^2, the larger the bias.

The statistics were computed for predictions calculated at the validation sites and for so called out-of-bag (OOB) predictions. OOB predictions originate from the bootstrap procedure that is part of RF algorithm. For each tree about 30 % of the observations are omitted (out-of-bag) because of the repeated random sampling (with replacement) of the original calibration data. This tree is then used to compute predictions for the omitted 30 % of the sites. OOB predictions are comparable to cross-validation predictions Hastie et al. (2009a) reported in 2012.

Additionally, we computed the coverage of the 90 %-prediction intervals compared to the observed SOC stock at the validation sites. Ideally 90 % of the SOC stock of the validation dataset were inside the intervals (e.g. Jiang et al., 2008).

2.7 Regional means and standard errors

Regional SOC stock mean estimates were calculated as in block kriging (Nussbaum et al., 2012) by averaging the values of each pixel within the specified region. Regions were either the NFI production regions stratified by altitude, these regions additionally stratified by productive and unproductive forests (CC12, CC13) or the 100x100 m AREA block cells.

We approximated regional standard errors to the mean estimates under the assumption of spatially varying model errors as in Wadoux et al. (2018). The pointwise standard deviations derived from quantile regression forest were scaled by a factor generated by a geostatistical simulation. The simulation resulted in a map representing the scaling factor for each response and took the spatial autocorrelation into account. Depending on the degree of autocorrel-

ation the pointwise standard deviation was then reduced accordingly and was aggregated over the region needed.

Based on a recommendation from G. Heuvelink (Wageningen University Research WUR, Emails January 2021) we used the following steps to approximate standard errors to the regional mean estimates from quantile regression forest:

1. We established full predictive distributions for each predicted location by computing 100 quantiles from $\alpha = 0$ to 1 by steps of 0.01. From these distributions we computed a standard deviation map and standard deviations for all calibration sites.

2. We derived residuals for the calibration sites by subtracting the difference of the RF prediction from the observed SOC stock value. Next, we standardized the residuals by dividing by the standard deviation computed in step (1).

3. An ordinary kriging model (Webster and Oliver, 2007) was fitted to the standardized residuals (step 2) with restricted maximum likelihood parameter estimation (REML, R Package georob). The standardized residuals (step 2) had nearly Gaussian distribution, thus no transformation by e.g. logarithm was performed. Spatial autocorrelation of standardized residuals was weak. For both responses the geostatistical model fit resulted in variogram close to *pure-nugget*.

4. The estimated variogram parameters were used to perform a Gaussian simulation (Webster and Oliver, 2007) with 50 realisations resulting each in a map of simulated standardized residuals. To speed up computations the neighbourhood was limited to 300 observations to be used to create the simulated predictions (R Package gstat, function krige).

5. The 50 realisations of simulated standardized residuals were averaged and multiplied with the quantile regression forest standard deviation maps (step 1). The simulated spatial autocorrelation of the standardized residuals is so used to scale the local standard deviations.

6. For each region the mean of absolute values of scaled standard deviations (step 5) within this region was computed as standard error approximation.

3 Results and Discussion

3.1 Models for SOC stock in 0–30 cm and 0–100 cm depth

The recursive backward elimination and decorrelation procedure removed a large amount of covariates. The models comprised a remainder of 40 and 45 covariates for SOC stock in 0–30 cm and 0–100 cm, respectively. The optimal number of covariates tested at each split (tuning parameter m_{try}) was set to 3 and 5 for the two models as these values resulted in the lowest OOB model errors.

Figure 2 shows the remaining covariates for SOC stock in 0–100 cm depth. For SOC stock in 0–30 cm similar covariates have been selected. As in 2012 precipitation and the physiographic units of the overview soil map (BEK) were amongst the most important covariates. Precipitation showed again a mostly linear positive effect on SOC stock (Figure 3). RF was able to handle large amount of covariates without becoming unstable, therefore many other climate covariates were included in the model. Terrain attributes rank somewhat less important, but were still selected to represent a wide terrain variation at different scales. In 2012 a near-infrared band was selected while only one Landsat derivate (average summer normalized differenced vegetation index) remained among the relevant covariates now. The spatial autocorrelation was attributed to rotated spatial coordinate axis. Three of these axis were among the 10 most important covariates (Figure 2) and showed strong non-linear effects (Figure 3).
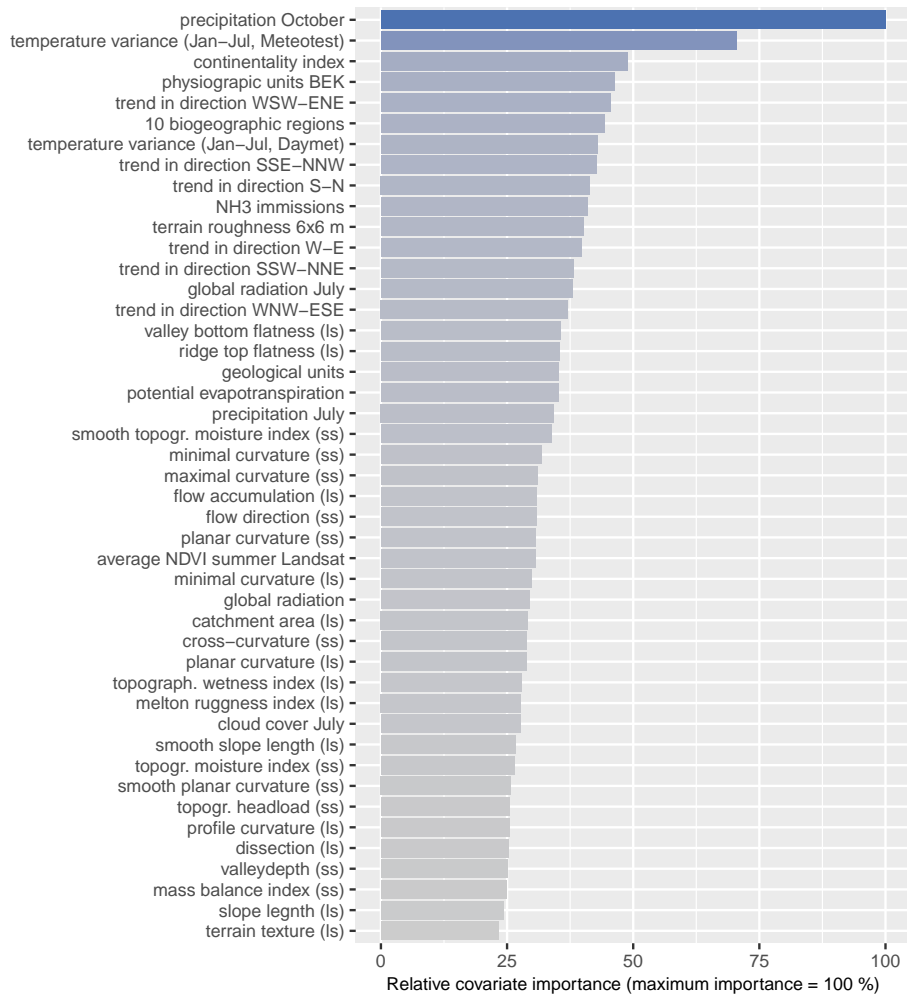
Figure 2: Covariates ordered according to their covariate (impurity) importance of 45 covariates selected by random forest for the response SOC stock in 0–100 cm soil depth (ss: terrain attribute calculated at small scale, ls: large scale terrain attribute; trend in directions based on rotated spatial coordinate axis by 30° and 60°, e.g. WSW-ENE: cardinal direction west-south-west to east-north-east).
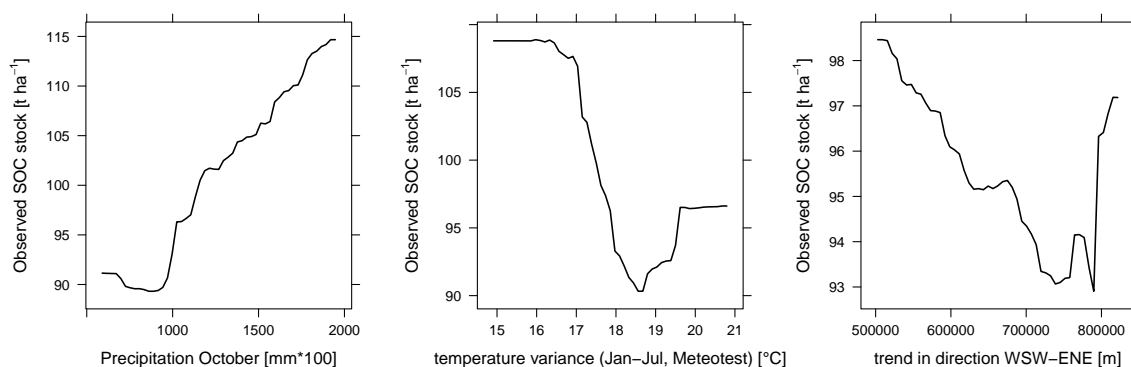
Figure 3: Partial dependence plots for three selected covariates to illustrate the non-linear response-covariate relationship established by random forest for SOC stock 0–100 cm. The solid curve is created by moving through the values of a covariate to be analysed while keeping all other covariate effects fixed. The larger the range of the SOC stock on y-axis the larger the capacity of a covariate to differentiate the SOC values. The large decrease at temperature variance from 17 to 18.5 °C (centre panel) indicates that most data splits of the random forest regression trees have been made in this range when this covariate was used for splitting the calibration data.

3.2 Evaluation of model performance

Model performance remained moderate. r^2 reported in 2012 were 0.34 and 0.40 for SOC stock 0–30 cm and 0–100 cm, respectively, and were hence similar to the performance achieved now by a much larger data set (Table 2). Predicted plotted against observed SOC stock values formed a considerable scatter around the 1:1-line indicating perfect prediction (Figure 4). The prediction error computations on the validation data set reflected an overall accuracy for the very variable Swiss forested area. The validation set stressed isolated locations (subsection 2.3). Increased estimation accuracy for areas with larger point densities is expected.

Table 2: Statistics of prediction errors of soil organic carbon stock (SOC) for two depth compartments for out-of-bag predictions (OOB, only calibration set of n = 1 918) and the validation set (n = 175)

		Bias [t ha^{-1}]	RMSE [t ha^{-1}]	r^2	R^2
OOB predictions	0–30 cm	0.89	35.21	0.28	0.28
	0–100 cm	2.06	57.72	0.28	0.28
predictions for validation set	0-30 cm	0.72	34.49	0.31	0.30
	0–100 cm	-10.24	64.52	0.43	0.40

Internal OOB prediction R^2 were lower than external validation R^2. This indicates that the models did not over-fit the data. Marginal bias was very small for SOC stock 0–30 cm, however, SOC stock in 0–100 cm was on average underestimated by 10 t ha^{-1} at the validation sites. The presence of Bias in model validation was only partly reflected by the overall SOC stock estimate for the total Swiss forested area. The SOC stocks in 0–100 cm were estimated to be lower by only 2.5 t ha^{-1} compared to the estimate of 2012 where only a small marginal Bias was reported (Table 4). The overall mean SOC stock in 0–100 cm was estimated to be 2.5 t ha^{-1} smaller than the previous estimate (Table 4).

The 90 %-prediction intervals computed by quantile regression forest slightly overestimated the magnitude of prediction errors (Figure 5). Of the expected 10 % of validation sites only 4 % and 5.7 % were outside the intervals. In 2012 kriging variances were also somewhat too large. RF was not able to fully correct for this overestimation.
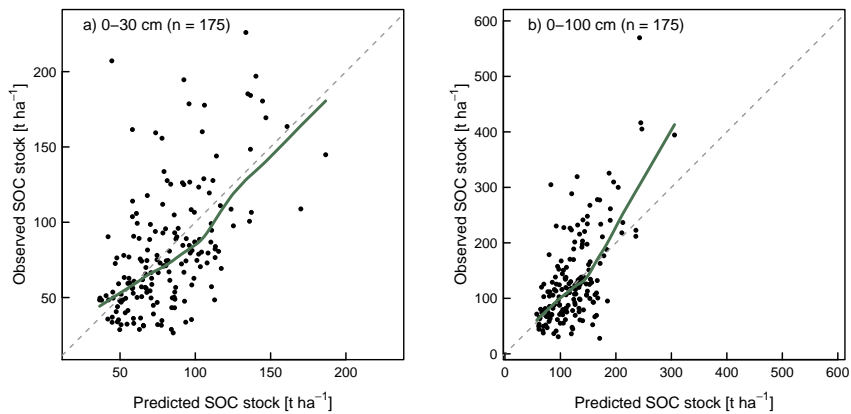
Figure 4: Scatter plots of observed against predicted soil organic carbon (SOC) stock in a) 0–30 cm and b) 0–100 cm of the mineral soil computed for the sites of the validation data set (solid line: lowess scatterplot smoothers, dashed line: 1:1-line, n: number of sites)
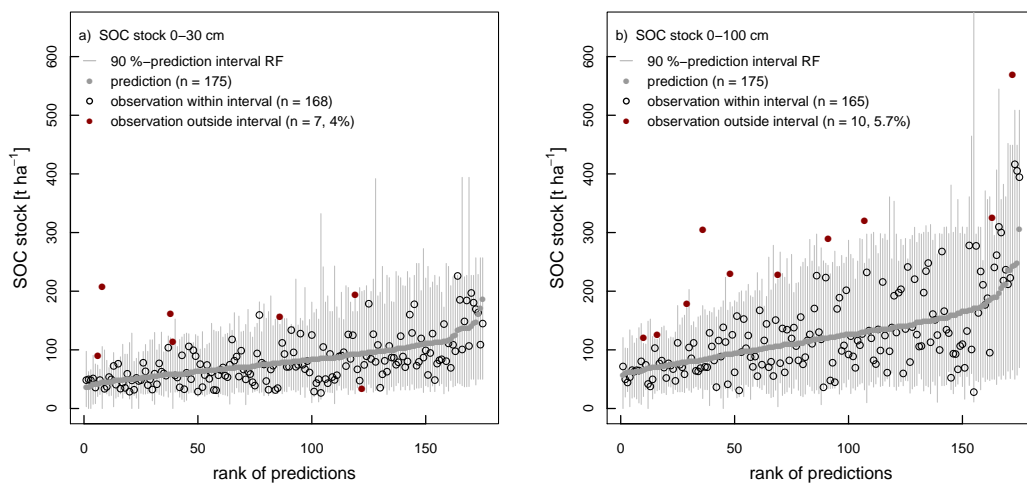

Figure 5: Predictions of SOC stock for the validation sites ordered increasingly with 90 %-prediction intervals (vertical grey lines). Observed SOC stocks inside the intervals were plotted by open circles, those outside by red filled symbols.

3.3 Predictions of SOC stock

Figure 6, 7 and 10 show an overview of the predicted SOC stock at the 1 ha AREA resolution. The overall spatial pattern of SOC stock for 0–30 and 0–100 cm remains similar to 2012.

The regional mean estimates for the stratified NFI production regions (Table 3) as well remained in the same range. The total SOC stock for Swiss forests was estimated to 81.4 t ha^{-1} (SE 3.0) in 0–30 cm and to 123.3 t ha^{-1} (SE 4.9) in the first 100 cm of the mineral soil. In 2012 these estimates were 79.9 t ha^{-1} (SE 1.5) and 125.8 t ha^{-1} (SE 2.4), respectively.

Differences become evident in areas where the calibration sites were strongly complemented. For example, in the eastern Jura foothills $\leq$ 600 m SOC stock seems to have been overestimated in 2012 by the low number of soil profiles available. This is further reflected in the large difference between the regional estimates of 2012 and 2021 displayed in Table 4. Further, considerable differences were observed in the Southern Alps where large SOC stocks are found and few calibration sites were available for both studies.

The new regional mean SOC stock estimates were beyond an interval established by the 2012 mean estimate $\pm$ 1.96*SE (95 % interval) for seven regions for 0–30 cm. Only for Jura $\leq$ 600 m and Southern Alps $\leq$ 600 m the new mean estimates were clearly outside this interval. For SOC stock in 0–100 cm there were four new estimates outside an 95 %-interval of $\pm$ 1.96*SE of 2012. Well outside were Alps $\leq$ 600 m and Southern Alps $>$ 1 200 m. For both regions SOC stock was difficult to estimate because of few observed sites. The first subregion covers only a small but variable area and in the latter generally large magnitude of SOC stocks are observed.

The first was complemented with about 100 new soil profiles and in the latter region generally large SOC stock were difficult to estimate because of few observed sites.

The new approximations of standard errors range similar to the standard error estimates of 2012. There is no general trend for either larger nor smaller standard errors to the new mean estimates. The standard errors for whole Switzerland were estimated to be slightly larger compared to the previous estimates (Table 4). The standard errors calculated for each 100x100 m AREA block cell (Figure 8 and 9) reflect the magnitude of SOC stock and the low point density of calibration sites.

Table 5 reports the mean SOC stocks and associated standard errors for the same altitude stratified NFI regions as in Table 3, but additionally split by productive (CC12) and unproductive (CC13) forests. Depending on environmental factors for each of the CC12 and CC13 sub-region the mean estimates differ compared to the full regions. Standard errors are similar; they are slightly smaller if the split in CC12 and CC13 resulted in a more homogenous region or larger for more heterogeneous regions in relation to their size.

Although there are only 13 and 9 calibration sites for 0–30 cm and 0–100 cm, respectively, we were able compute estimations for the unproductive forest areas (CC13). These estimates are only valid under the assumption that productive vs. unproductive forests are not the main driving factor for the spatial distribution of SOC stocks. It is assumed that the main factors explaining SOC stocks are represented by the environmental covariates included in the model and that such models allow for spatial prediction of CC13 areas.

We would like to point out that a validation sampling design targeting CC13 areas would be needed to verify this assumption (currently only one validation site within CC13). Moreover, sampling of new locations within CC13 chosen to optimally calibrate models would be desirable to support estimates in Table 5.
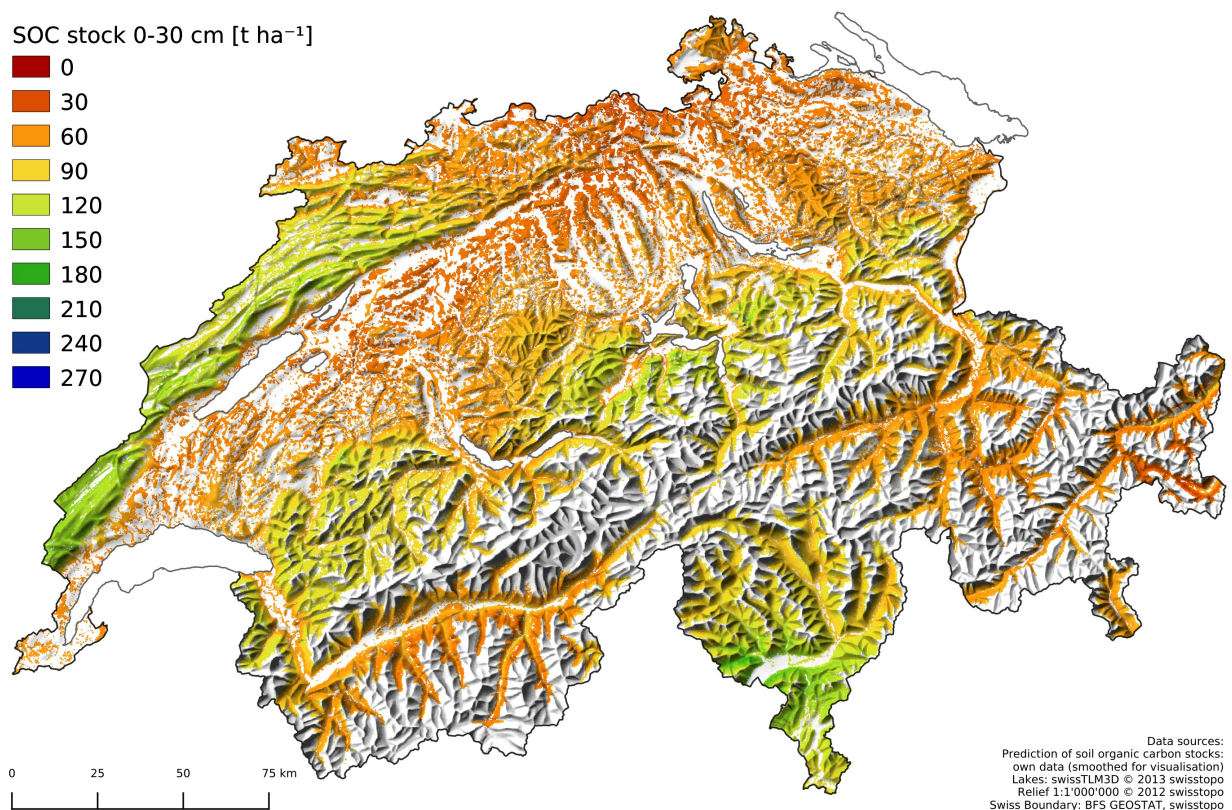
Figure 6: Mean random forest predictions for each 100x100 m AREA block cell for SOC stock in 0–30 cm of the mineral soil of Swiss forest soils (smoothed for small scale visualisation with mean per 3x3 pixel width = 300 m).

SOC stock 0-100 cm [t ha⁻¹]

SOC stock 0-100 cm [t ha^{-1}]

- 0
- 30
- 60
- 90
- 120
- 150
- 180
- 210
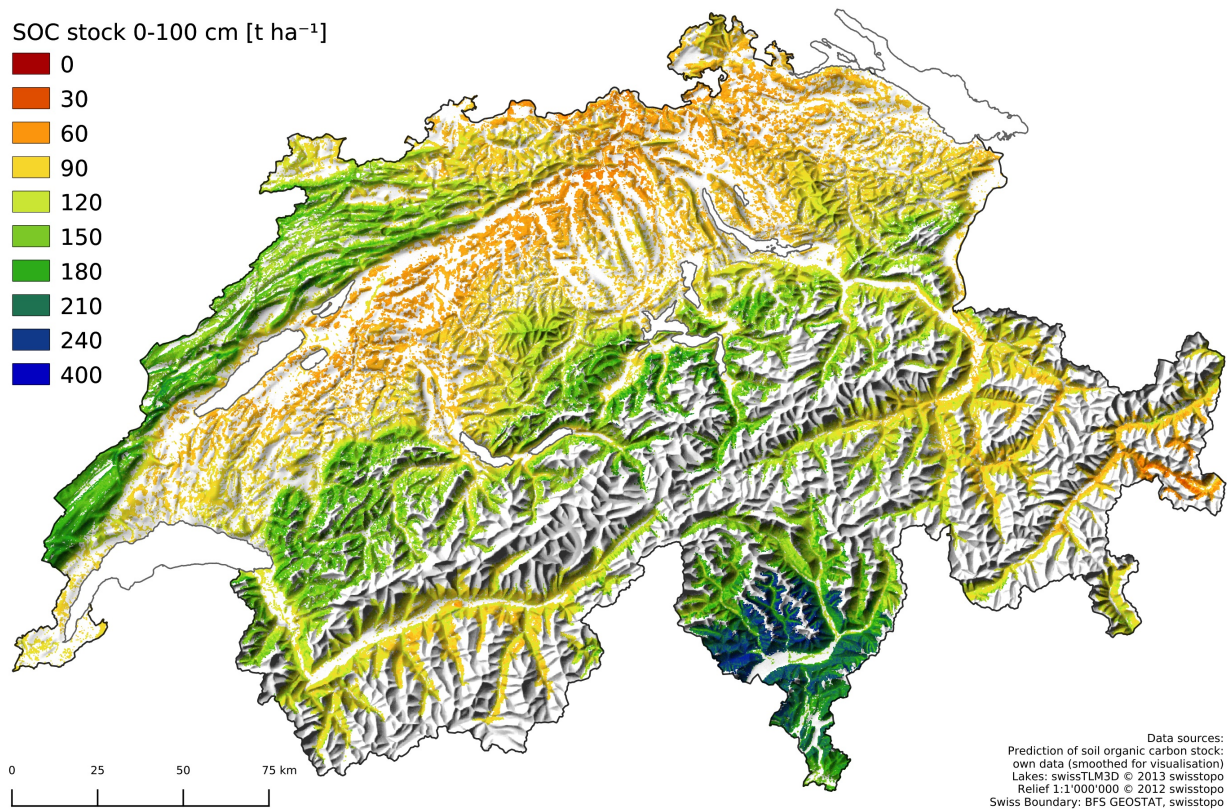- 240
- 400

0   25   50   75 km

Figure 7: Mean random forest predictions for each 100x100 m AREA block cell for the soil organic carbon (SOC) stock in 0–100 cm of the mineral soil of Swiss forest soils (smoothed for small scale visualisation with mean per 3x3 pixel width = 300 m).
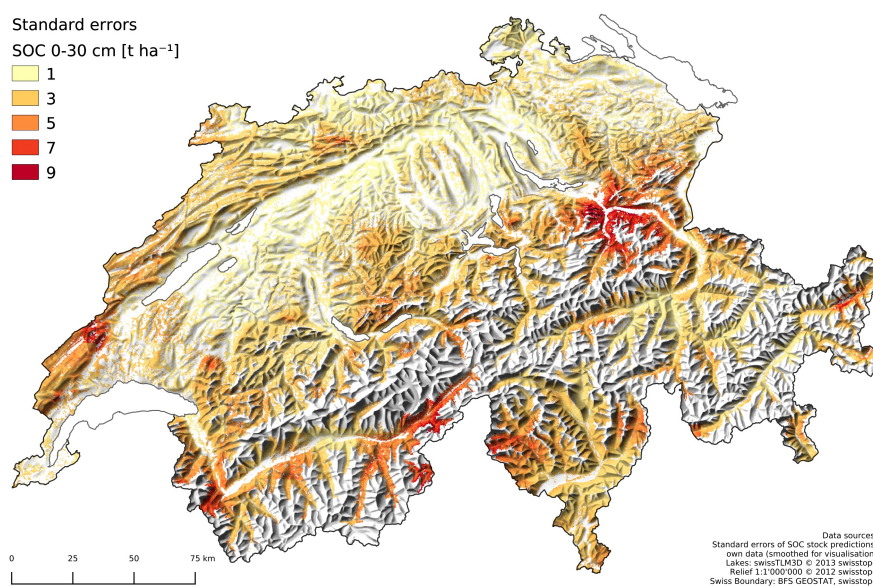
Standard errors
SOC 0-30 cm [t ha^{-1}]

- 1
- 3
- 5
- 7
- 9

0   25   50   75 km

Figure 8: Standard errors for each 100x100 m AREA block cell for SOC stock in 0–30 cm soil depth.
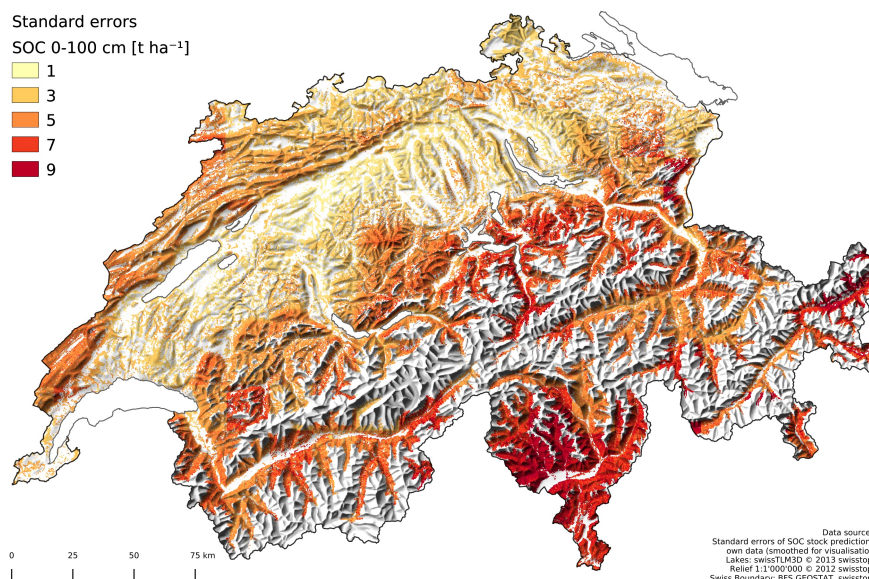
Figure 9: Standard errors for each 100x100 m AREA block cell for SOC stock in 0–100 cm soil depth.

Table 3: Random forest predictions of the mean SOC stock in the mineral topsoil (0–30 cm soil depth) and in the mineral soil form 0–100 cm depth of the five production regions of the National Forest Inventory, stratified by altitude, and of Switzerland as a whole (n: number of soil profiles in cs: calibration dataset and vs: validation data set; $\hat{S}_c$: mean random forest predictions; $SE[\hat{S}_c]$: standard error approximation).

Region	Altitude	SOC 0–30 cm				SOC 0–100 cm			
		cs	vs	$\hat{S}_c$	$SE[\hat{S}_c]$	cs	vs	$\hat{S}_c$	$SE[\hat{S}_c]$
	[m]	n	n	[t ha^{-1}]	[t ha^{-1}]	n	n	[t ha^{-1}]	[t ha^{-1}]
Jura	≤600	112	9	56.55	1.80	112	9	92.90	2.99
	600–1200	108	20	101.29	2.95	108	20	139.49	4.23
	>1200	6	0	128.44	3.47	6	0	164.10	4.82
Central Plateau	≤600	674	18	50.67	1.58	673	18	76.40	2.68
	600–1200	187	18	64.11	2.19	187	18	90.70	3.07
	>1200	1	0	127.56	4.20	1	0	165.23	5.32
Pre-Alps	≤600	59	1	63.34	2.44	59	1	96.24	3.94
	600–1200	222	24	79.88	3.00	222	24	117.09	4.59
	>1200	80	18	103.34	3.42	81	18	149.84	5.45
Alps	≤600	6	0	69.61	3.57	6	0	109.69	4.93
	600–1200	109	11	77.41	3.48	109	11	118.28	5.41
	>1200	261	33	75.29	3.50	261	33	115.34	5.73
Southern Alps	≤600	16	8	118.48	3.15	16	8	186.44	7.88
	600–1200	38	10	111.66	3.17	38	10	188.86	7.61
	>1200	39	5	98.10	3.37	39	5	173.18	7.16
Switzerland		1918	175	81.35	2.95	1918	175	123.28	4.93
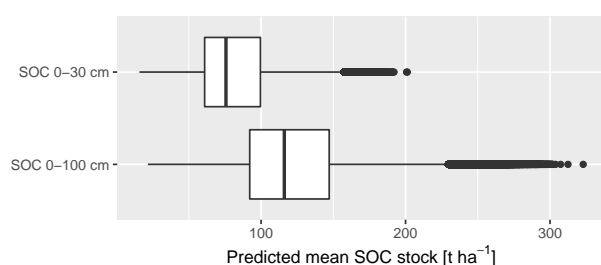
Figure 10: Boxplots to illustrate the distribution of mean SOC stock of all 100x100 m AREA block cells (the box corresponds to the interquartile range with the lower end of the box being the first quartile [25 % percentile] and the upper end the third quartile [75 % percentile], the solid line covers 1.5 times the interquartile range. Values beyond the solid line are considered outliers).

Table 4: Block kriging predictions reported by Nussbaum et al. (2012) and differences to predictions reported in Table 3. There were less calibration sites (cs) in 2021 for certain regions because the validation set was not used for re-estimation of model parameters (Δ = 2021 minus 2012, for abbreviations see Table 3).

Region	Altitude	SOC 0–30 cm					SOC 0–100 cm				
		2012			Δ 2021		2012			Δ 2021	
		cs	$\hat{S}_c$	$SE[\hat{S}_c]$	cs	$\hat{S}_c$	cs	$\hat{S}_c$	$SE[\hat{S}_c]$	cs	$\hat{S}_c$
	[m]	n	[t ha^{-1}]	[t ha^{-1}]	n	[t ha^{-1}]	n	[t ha^{-1}]	[t ha^{-1}]	n	[t ha^{-1}]
Jura	≤600	22	82.65	3.34	90	-26.10	22	104.77	6.51	90	-11.87
	600–1200	68	102.03	3.57	40	-0.74	68	145.89	5.02	40	-6.40
	>1200	5	121.34	5.39	1	7.10	5	168.10	7.52	1	-4.00
Central Plateau	≤600	199	55.40	1.55	475	-4.73	199	81.36	2.11	474	-4.96
	600–1200	102	62.12	1.68	85	1.99	102	92.22	2.11	85	-1.52
	>1200	1	122.00	7.07	0	5.56	1	171.04	10.29	0	-5.81
Pre-Alps	≤600	50	66.10	2.06	9	-2.76	50	93.24	2.38	9	3.00
	600–1200	184	75.91	2.00	38	3.97	184	112.92	2.95	38	4.17
	>1200	66	95.78	3.27	14	7.56	66	153.78	4.98	15	-3.94
Alps	≤600	4	66.47	2.44	2	3.14	4	99.58	2.86	2	10.11
	600–1200	64	74.39	2.42	45	3.02	64	120.89	3.62	45	-2.61
	>1200	163	69.48	1.85	98	5.81	163	115.30	3.19	98	0.04
Southern Alps	≤600	22	102.37	4.07	-6	16.11	22	196.26	9.84	-6	-9.82
	600–1200	40	108.99	4.09	-2	2.67	40	209.50	9.83	-2	-20.64
	>1200	32	107.08	4.11	7	-8.98	32	192.37	8.04	7	-19.19
Switzerland		1 022	79.9	1.52	896	1.45	1 022	125.8	2.41	896	-2.52

Table 5: Random forest predictions of the mean SOC stock in the mineral topsoil (0–30 cm soil depth) and in the mineral soil form 0–100 cm depth of the five production regions of the National Forest Inventory, stratified by altitude and additionally by productive (CC12) and unproductive forests (CC13) according to Swiss Land Use Statistic (AREA, n: number of soil profiles in cs: calibration dataset and vs: validation data set, the smaller *n* are caused by the different forest definition of NFI and AREA; $\hat{S}_c$: mean random forest predictions; $SE[\hat{S}_c]$: standard error approximation).

Region	Altitude	CC	SOC 0–30 cm				SOC 0–100 cm			
			cs	vs	$\hat{S}_c$	$SE[\hat{S}_c]$	cs	vs	$\hat{S}_c$	$SE[\hat{S}_c]$
	[m]		n	n	[t ha^{-1}]	[t ha^{-1}]	n	n	[t ha^{-1}]	[t ha^{-1}]
Jura	≤600	12	106	9	56.54	1.79	106	9	92.67	2.97
		13	0	0	57.08	1.74	0	0	93.71	3.05
	600–1200	12	101	20	101.16	2.95	101	20	139.98	4.22
		13	0	0	106.91	2.90	0	0	138.66	4.31
	>1200	12	6	0	129.46	3.50	6	0	165.96	4.85
		13	0	0	123.61	3.35	0	0	155.14	4.72
Central Plateau	≤600	12	642	18	50.30	1.56	641	18	75.74	2.64
		13	1	0	52.49	1.74	1	0	80.36	3.01
	600–1200	12	183	18	64.13	2.19	183	18	90.58	3.06
		13	0	0	65.35	2.31	0	0	94.01	3.35
	>1200	12	1	0	128.23	4.17	1	0	166.22	5.32
		13	0	0	125.59	4.26	0	0	161.68	5.17
Pre-Alps	≤600	12	54	1	63.42	2.42	54	1	96.31	3.93
		13	1	0	61.82	2.84	1	0	95.15	4.14
	600–1200	12	209	23	80.18	3.01	209	23	117.52	4.61
		13	1	0	80.81	3.07	1	0	117.88	4.68
	>1200	12	75	17	103.28	3.42	75	17	149.26	5.44
		13	2	0	103.24	3.41	3	0	149.93	5.39
Alps	≤600	12	5	0	69.96	3.60	5	0	110.91	4.93
		13	0	0	70.90	3.86	0	0	110.36	5.48
	600–1200	12	102	11	77.51	3.47	102	11	118.56	5.39
		13	1	0	77.52	3.65	1	0	116.90	5.58
	>1200	12	235	32	73.68	3.44	235	32	112.51	5.67
		13	6	0	82.29	3.68	6	0	127.89	5.95
Southern Alps	≤600	12	13	8	119.86	3.14	13	8	189.65	7.94
		13	0	0	100.90	2.79	0	0	170.74	6.55
	600–1200	12	36	9	112.91	3.15	36	9	190.59	7.63
		13	0	1	99.71	3.22	0	1	173.34	7.30
	>1200	12	36	5	97.05	3.26	36	5	170.91	7.02
		13	1	0	102.48	3.68	1	0	184.37	7.65

4 Conclusion

This study provided new SOC stock estimates for the GHGI based on 1918 calibration sites compared to 858 calibration sites used for building the model structure and validation in 2012. The overall model performance computed on the same validation sites did however not increase. Since the model fits were based on a much broader data source we assume overall more reliable estimates. Generally, regional estimates compared well to the previous study of 2012. Estimates for the altitude stratified NFI regions differed to 2012 and were likely improved where there were a large number of newly sampled sites (Central Plateau) or where large SOC stock and simultaneously low observation density was found (Southern Alps).

Forests in Switzerland span over a large variety of soil forming factors with a wide range of parent materials and climatic conditions. The sampled locations have not been determined in view of this large variability, but have been chosen by purposive sampling on a project by project basis. A analysis of model performance by region and relating performance results to sampling densities was beyond this study. We nevertheless assume an increased map accuracy for regions now more densely populated with soil profile sites.

We did not perform a detailed analysis of the overall representativeness of the soil profile sites for the main factors determining SOC stock. But, Figure 1 shows areas with diverse topography and very few samples. The Southern Alps – having in general large SOC stock – seem undersampled. Currently, WSL in collaboration with BFH-HAFL conducts an analysis to complement the dataset with soil profile sites. The new sampling locations will be chosen based on soil property uncertainty maps (Baltensweiler et al., 2021, submitted) and their potential to complete the coverage of soil forming factors.

The non-spatial nature of RF was circumvented by introducing rotated coordinate axis as covariates. The approximation of standard error to regional mean estimates used a geostatistical approach able to integrate relative location of each site. The standard errors resulted in a similar value range as block kriging standard errors published in 2012. Although the approximation was a workaround we conclude that the standard errors are of similar accuracy.

Above we presented new SOC stock estimates for altitude stratified NFI regions based on a larger soil sampling basis. In addition we provided a further regional subdivision into productive and unproductive forests. The larger soil sample did not improve model performance measures, but we still attribute larger reliability and recommend to use the present estimates for GHGI reporting until an even broader data source becomes available.

Acknowledgments

List of Figures

List of Tables

Literature

Baltensweiler, A., Walthert, L., Hanewinkel, M., Zimmermann, S., and Nussbaum, M.: Machine learning based soil maps for the Swiss forest area for a wide range of soil properties, Geoderma Regional, 2021, submitted.

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

Brungard, C. W., Boettinger, J. L., Duniway, M. C., Wills, S. A., and Edwards Jr., T. C.: Machine learning for predicting soil classes in three semi-arid landscapes, Geoderma, 239–240, 68–83, doi:10.1016/j.geoderma.2014.09.019, 2015.

Gareth, J., Witten, D., Hastie, T., and Tibshirani, R.: An introduction to statistical learning: With applications in R, Springer texts in statistics, Springer, New York and Heidelberg and Dordrecht and London, corrected at 8th printing edn., 2017.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, second edition edn., URL https://web.stanford.edu/~hastie/Papers/ESLII.pdf, 2009a.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning; Data Mining, Inference and Prediction, Springer, New York, 2 edn., 2009b.

Hertzog, M.: Modelling soil attributes with the Random Forest method for the Swiss forest area, 2017.

Jiang, W., Varma, S., and Simon, R.: Calculating Confidence Intervals for Prediction Error in Microarray Classification Using Resampling, Statistical Applications in Genetics and Molecular Biology, 7, 0–19, 2008.

Meinshausen, N.: Quantile regression forests, Journal of Machine Learning Research, 7, 983–999, 2006.

Nussbaum, M.: Digital Soil Mapping for Switzerland Evaluation of statistical approaches and mapping of soil properties, 2017.

Nussbaum, M., Papritz, A., Baltensweiler, A., and Walthert, L.: Organic Carbon Stocks of Swiss Forest Soils, URL http://e-collection.library.ethz.ch/eserv/eth:6027/eth-6027-01.pdf, 2012.

Nussbaum, M., Papritz, A., Baltensweiler, A., and Walthert, L.: Digitale Bodeneigenschaftskarten mit Boosted Structured Regression am Beispiel des organischen Kohlenstoffvorrats im Schweizer Wald, 2014a.

Nussbaum, M., Papritz, A., Baltensweiler, A., and Walthert, L.: Estimating soil organic carbon stocks of Swiss forest soils by robust external-drift kriging, Geoscientific Model Development, 7, 1197–1210, doi:10.5194/gmd-7-1197-2014, URL http://www.geosci-model-dev.net/7/1197/2014/, 2014b.

Nussbaum, M., Spiess, K., Baltensweiler, A., Grob, U., Keller, A., Greiner, L., Schaepman, M. E., and Papritz, A.: Evaluation of digital soil mapping approaches with large sets of environmental covariates, SOIL, pp. 1–22, doi:10.5194/soil-4-1-2018, URL https://doi.org/10.5194/soil-4-1-2018, 2018.

Nussbaum, M., Burgos, S., Baltensweiler, A., and Walthert, L.: Digital Soil Maps of Swiss forest soils: Mapping of basic soil properties at six depth by machine learning based model averaging, Internal Project Report, Zollikofen, 2019.

Perruchoud, D., Walthert, L., Zimmermann, S., and Lüscher, P.: Contemporay Carbon Stocks of Mineral Forest Soils in the Swiss Alps, Biogeochemistry, 50, 111–136, 2000.

Spiess, K.: Vorhersage von Bodeneigenschaften mit Quantile Regression Forest, Validierung und Vergleich mit den Vorhersagen aus geoadditiven Modellen, Ph.D. thesis, Departement für Umweltsystemwissenschaften der ETH Zürich, Zürich, 2016.

Wadoux, A. M.-C., Brus, D. J., and Heuvelink, G. B. M.: Accounting for non-stationary variance in geostatistical mapping of soil properties, Geoderma, 324, 138–147, 2018.

Walthert, L., Scherler, M., Stähli, M., Huber, M., Baltensweiler, A., Ramirez Lopez, L., and Papritz, A.: Böden und Wasserhaushalt von Wäldern und Waldstandorten der Schweiz unter heutigem und zukünftigem Klima (BOWA-CH), doi:10.3929/ethz-a-010658682, URL http://e-collection.library.ethz.ch/view/eth:49159, 2015.

Webster, R. and Oliver, M. A.: Geostatistics for Environmental Scientists, John Wiley & Sons, New York, 2 edn., 2007.

Appendix

Table 6: List of files in data delivery.

Filename	File formats	Description
SOC_stocks_nfi_region_alti_update2021	csv, xlsx, gpkg	Table 3 as data tables and as Geopackage containing the vector geometries of the altitude stratified forest production regions.
SOC_stocks_nfi_regions_CC12-13	csv, xlsx	Table 5 as data tables.
SOC_stocks_AREA_100m	csv, tif	Table with a row for each 100x100 m AREA block cell of the codes CC11, CC12 and CC13. If a cell is empty, then the block mean could not be calculated mainly because of missing predictions. Predictions were only calculated for the forested area as defined according to subsection 2.1. *columns:* *x, y*: spatial coordinates in LV03 / EPSG:21781 *kyoto09*: CC code *SOC30_mean* and *SOC100_mean*: block cell mean SOC stock in 0–30 cm and in 0–100 cm *SOC30_se* and *SOC100_se*: block cell standard error of SOC stock in 0–30 cm and in 0–100 cm. . Corresponding spatial raster files (GeoTiff): *SOC30_mean_AREA.tif*, *SOC30_se_AREA.tif*, *SOC100_mean_AREA.tif*, *SOC100_se_AREA.tif*