Organic Carbon Stocks of Swiss Forest Soils

Final Report

Organic Carbon Stocks of Swiss Forest Soils

Final Report

Project Grants Nos 06.0091.PZ / J483-4723 and 06.0091.PZ / K444-2034
Federal Office for the Environment FOEN, Forest Division, 3003 Bern, Switzerland

20 September 2012

Madlene Nussbaum & Andreas Papritz
ETH Zurich, Institute of Terrestrial Ecosystems
Universitätstrasse 16, 8092 Zürich, Switzerland

Andri Baltensweiler & Lorenz Walthert
Swiss Federal Institute for Forest, Snow and Landscape Research WSL
Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

WSL

Citation information:
Nussbaum, M., Papritz, A., Baltensweiler, A., and Walthert, L. (2012). Organic Carbon Stocks of Swiss Forest Soils. Final Report. Institute of Terrestrial Ecosystems, ETH Zürich and Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Zürich and Birmensdorf.

Summary

The Swiss greenhouse gas inventory (GGI), submitted under the United Nations Framework Convention on Climate Change and under the Kyoto Protocol, reports annually on changes of organic carbon (OC) stocks in forests. The National Forest Inventory (NFI) offers comprehensive data to quantify the living and dead forest biomass and its change in time. Estimating stocks of soil organic carbon (SOC) in forests is more difficult because the variables needed to quantify stocks 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.* in 2000. Since then additional SOC measurements, new spatial data (high resolution terrain model, remotely sensed spectral data) became available. Furthermore, new methods for robust geostatistical analyses of spatial data were recently developed, which seem well suited for analysing environmental datasets. Based on these developments, we modelled OC stocks in the forest soils for three depth compartments: (i) forest floor, (ii) mineral soil 0–30 cm and (iii) mineral soil 0–100 cm. We applied a novel robust restricted maximum likelihood method to fit a linear regression model with spatially correlated errors. For the regression analysis we used a broad range of covariates derived from climate data (e.g. precipitation, temperature), two elevation models (resolutions 25 and 2 m) with their terrain attributes and spectral reflectance data representing vegetation. Furthermore, the main cartographic categories of an overview soil map and a small-scale geological map were used to represent the parent material. Based on the fitted models we mapped the spatial distribution of the SOC stocks in the three soil compartments across the forest area of Switzerland by robust lognormal kriging with a resolution of 100 m. For the same compartments, we computed also lognormal block kriging predictions of the mean SOC stocks per unit area for the five NFI production regions stratified by three altitude classes.

During the analyses we discovered that the geostatistical model of forest floor stocks fitted the data only poorly. Therefore, we computed as an alternative design-based estimates of these stocks for the NFI production regions stratified by altitude. The design-based estimate of the mean OC stock in the forest floor of Switzerland amounted to 16.7 t ha^{-1} (standard error 0.83 t ha^{-1}). Results for topsoil (0–30 cm) SOC stocks showed weak but significant residual autocorrelation. Precipitation, spectral reflectance of the vegetation at near-infrared wavelength, a topographic position index and aggregated soil and geological map information were the only significant covariates. The predictive power of the fitted model, evaluated by comparing predictions with independent validation data, was moderate (robust coefficient of determination 0.34). The predicted mean SOC stock to a depth of 30 cm amounted to 79.9 t ha^{-1} for whole Switzerland (block kriging standard error 1.52 t ha^{-1}). The model for the SOC stocks stored in 0–100 cm showed the same weak residual autocorrelation. Also a similar set of covariates was chosen: March precipitation, spectral reflectance of the vegetation at near-infrared wavelength, slope angle and aggregated soil map information. The predictive power of the fitted model, evaluated again with independent validation data, was slightly better than for the topsoil (robust coefficient of determination 0.40). On average 64 % of the SOC stock is stored in the topsoil as the prediction of the mean SOC stock to a depth of 100 cm was equal to 125.8 t ha^{-1} (block kriging standard error 2.41 t ha^{-1}).

The estimated mean stock agreed for the forest floor with a previous estimate by Moeri (2007). However, the new predictions of the mean topsoil and mean total stocks down to 1 m were significantly larger than the previous estimates by Perruchoud *et al.* (2000). The differences were small for the topsoil (4.0 t ha^{-1}, p-value: 0.004), but very pronounced (27.6 t ha^{-1}, p-value: $< 10^{-12}$) for the total stocks. Since the present study relies on much more comprehensive data, uses sophisticated geostatistical techniques for computing predictions and demonstrated that the predictions were unbiased for independent validation data, one can safely conclude that SOC stocks of mineral soils have in the past been underestimated for Swiss forests, in particular for the subsoils.

Contents

List of Figures

List of Tables

Abbreviations

asl above sea level

C carbon

CRPS continuous ranked probability score

FOEN Swiss Federal Office for the Environment

GGI greenhouse gas inventory

IHS intensity-hue-saturation color space

KABO Cantonal Soil Monitoring Network (Kantonale Bodenbeobachtung)

LASSO least absolute shrinkage and selection operator

MAD median absolute deviation

ML maximum likelihood

NABO Swiss Soil Monitoring Network (Nationale Bodenbeobachtung)

NDVI Normalized Differenced Vegetation Index

NFI National Forest Inventory

OC organic carbon

R^2 coefficient of determination

REML restricted maximum likelihood

RMSE root mean squared error

robRMSE robust root mean squared error

SE standard error

SOC soil organic carbon

STDV standard deviation

TCCCA transparency, consistency, comparability, completeness, accuracy (guidelines)

UNFCCC United Nations Framework Convention on Climate Change

WSL Swiss Federal Institute for Forest, Snow and Landscape Research

1 Introduction

1.1 Estimation of carbon stocks in Swiss forest soils

The Swiss greenhouse gas inventory (GGI), submitted under the United Nations Framework Convention on Climate Change (UNFCCC) and under the Kyoto Protocol (FOEN, 2012), reports annually on changes of organic carbon (OC) in forests. The National Forest Inventory (NFI, Brassel and Lischke, 2001) provides comprehensive data to quantify the carbon (C) stored in the living and dead forest biomass and its change in time. Estimating stocks of soil organic carbon (SOC) in forests is more challenging because SOC concentration, gravel content, bulk density and depth[1] of soils vary strongly in space, measurements of these quantities are expensive, and, in consequence, only limited information is commonly available to quantify SOC stocks.

To date, the Swiss Federal Office for the Environment (FOEN) used for Switzerland's GGI estimates of SOC stock in forests, published by Perruchoud *et al.* (2000) more than a decade ago. As SOC data was then rather scanty (data of merely 168 sites were at the time available) these authors reported estimates of the mean SOC stocks either for whole Switzerland, or for the Southern Alps, where C rich soils prevail, and the remainder of the country[2].

In the past ten years, additional SOC measurements were made by the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) at additional 865 forest sites. Furthermore, new spatial data (high-resolution terrain model, remotely sensed spectral information) became available that may serve as covariates in statistical modelling of SOC stock data. In view of the notably improved data basis, our study aims at providing (i) spatially explicit and (ii) more precise estimates of SOC stocks stored in Swiss forest soils than are currently available. We used a new, robust geostatistical approach (Künsch *et al.*, 2011, in preparation) for this.

More specifically, we conducted in this study a robust geostatistical analysis of data on OC stocks stored per unit area in three depth compartments of forest soils:

- forest floor (L, F and H horizons without deadwood),
- mineral soil 0–30 cm depth, and
- mineral soil 0–100 cm depth (or from 0 cm to depth of bedrock on shallower soils).

We mapped the spatial distribution of these three quantities across the forest area of Switzerland by robust kriging. For the same soil compartments, we also report block kriging predictions of the mean SOC stocks per unit area for the main production regions of Switzerland (Jura, Central Plateau, Pre-Alps, Alps, Southern Alps, cf. Brassel and Brändli, 1999), further stratified by altitude according to

- $\leq$ 600 m asl,
- $>$ 600 m to $\leq$ 1200 m asl, and
- $>$ 1200 m asl.

When developing and fitting the geostatistical models, we used only a subset of the sites, for which SOC stock data were available (calibration set), and used the data of the remaining sites (validation set) to check the precision of the kriging predictions. Further validation data were available from the national and two cantonal soil monitoring networks. We further compared our predictions with initial stock estimates calculated by Weggler *et al.* (2012) with the C turn-over model Yasso07 (Tuomi *et al.*, 2009, 2011).

Our contract work for FOEN is partly based on a MSc thesis (Nussbaum, 2011, in German), in which many covariates for the geostatistical modelling were generated and data on the SOC

[1] The latter three quantities are required in addition to the C concentration to estimate the SOC stock per unit area, cf. section 2.2.1.

[2] WSL later computed with the same data estimates of the mean SOC stocks for the NFI production regions Jura, Central Plateau, Pre-Alps, Alps, Southern Alps stratified by altitude (E. Thürig, personal communication).

concentration and stock of the topsoil (0–30 cm depth, including forest floor) were geostatistically analysed. In the current study, we extended Nussbaum's work by quantifying the stocks separately for the forest floor and two compartments of the mineral soil and by estimating the mean stocks for the NFI production regions stratified by altitude.

1.2 TCCCA: specific information for UNFCCC-reviewers

The present report reflects the TCCCA (transparency, consistency, comparability, completeness, accuracy) criteria as follows:

(i) Transparency The report documents what data were used for the analyses and describes the adopted statistical approach. The origin of the SOC data and the derivation of SOC stock values from the data sampled by soil horizon are explained in section 2.2.1. Table 2 gives an overview of the input data from which covariates of the statistical models were selected and refers to the orginal data sources where further information can be found. The structure of the geostatistical model and the robust parameter estimation method are described with relevant references in sections 2.3.1 and 2.3.2. Models of SOC stocks in three depth compartments of the soil (forest floor, mineral soil 0–30 cm and 0–100 cm) were developed in several steps that are detailed in section 2.3.3. Kriging predictions were computed as described in section 2.3.4 and Appendix B. Further details about kriging are documented in Appendices C–E.

The geostatistical model of the OC stock in the forest floor turned out to have too little predictive power in the validation (cf. section 3.1 and F.2). Therefore, we could not reliably map the spatial distribution of this quantity across Switzerland, and we employed a design-based procedure for estimating the mean OC stocks in the forest floor for the NFI production regions and for the whole country. The respective estimation methods are described in section 2.4. A description of the best-fit geostatistical models of the stocks in the mineral soil can be found for the depth compartment 0–30 cm in section 3.2.1 and for the compartment 0–100 cm in section 3.3.1. Our estimates of the mean stocks for the NFI production regions and for Switzerland as a whole are compared in section 4 with previous SOC stock estimates by Perruchoud *et al.* (2000) and Moeri (2007) that were both computed with less SOC data and a less comprehensive set of potential covariates for statistical modelling. There were no circumstances preventing transparent data supply with regard to SOC and input data.

(ii) Consistency The soil data had been collected over a longer period of time (cf. Figure 2). Consistency of the SOC stock data was ensured by using for all the soil samples the same methods to determine OC and bulk density. Samples stored in a soil archive were re-analysed if other OC measurement or carbonate removal methods had been initially used. Spatial consistency was assured by using only those input datasets as potential covariates for statistical models (cf. Table 2) that were available from the same source for the total Swiss forested area.

(iii) Comparability The present study modelled the OC stocks of three depth compartments of forest soils that were defined in section 1. Our SOC stock predictions are compared in section 4 with estimates for other countries or regions that used the same compartments. Comparability with other countries and regions is therefore established.

(iv) Completeness The OC predictions were computed for the forest floor and the mineral soil. Completeness therefore can only be stated for the respective stocks. Complete geographical coverage for the forest area is given as for every production region stratified by altitude (cf. section 1) soil samples were available (cf. Table 3, Table 5). Hence, predictions of OC stocks could be computed for the total Swiss forest area (cf. Table 3, Figure 7 and 11).

(v) Accuracy The accuracy of the predictions was evaluated by cross-validation and by comparison with independent validation sets of OC measurements (for method description, see section 2.3.5, for evaluation of model performance sections 3.1, 3.2.2, 3.3.2 and F.2). Uncertainties of OC and input data are discussed in detail in section 4.

1.3 Review of studies estimating SOC stock in European forests

Several studies recently estimated OC stocks of forest soils on a regional or national scale in Europe. Integrating soil data from different sources, possibly collected over a longer period of time (Krogh *et al.*, 2003), and imputing missing compulsory data, in particular bulk densities of soils (Weiss *et al.*, 2000), were common challenges in these endeavours. Some studies provided only estimates of topsoil OC content or stocks (usually 0–30 cm, e.g. Oehmichen *et al.*, 2011; Meersmans *et al.*, 2012), estimates of stocks stored in the top 50 cm of the solum (Xu *et al.*, 2011), or estimates of stocks down to one meter depth of the soil (Krogh *et al.*, 2003; Lettens *et al.*, 2005; Wiesmeier *et al.*, 2012). Commonly, stocks of forest soils were not estimated separately, but jointly with stocks of other land uses (grassland, arable land, wetland, etc.). Information on land use was often an important covariate in models of SOC stocks (e.g. Wiesmeier *et al.*, 2012).

Only few studies could rely on measured bulk densities (Krogh *et al.*, 2003; Wiesmeier *et al.*, 2012). Commonly, this parameter was estimated by pedotransfer functions from other measured soil properties, such as SOC concentration (Meersmans *et al.*, 2008, 2012; Weiss, 2000) or texture classes (Lettens *et al.*, 2005). Other studies assigned densities of similar soil types (Tomlinson and Milne, 2006) or predicted missing densities by a regression model calibrated with a small set of sites where bulk density had been measured (Xu *et al.*, 2011). Oehmichen *et al.* (2011) completed their bulk density data based on ordinal field estimates by experts as in the present study (cf. section 2.2.1).

Various modelling approaches, using a variety of covariates (soil maps, climate data, land use information, etc.) were used to estimate SOC stocks. A popular technique was to assign average SOC stocks to map units of existing soil or land use maps or to their intersections (Krogh *et al.*, 2003; Lettens *et al.*, 2005; Meersmans *et al.*, 2008; Tomlinson and Milne, 2006; Oehmichen *et al.*, 2011). Other studies did not distinguish between soil or land use classes and used an average SOC stock for the whole study area (Weiss *et al.*, 2000; Wördehoff *et al.*, 2011, 2012). Furthermore, SOC data were fitted to land use information and climate data by linear regression models, either ignoring spatial auto-correlation (e.g. Meersmans *et al.*, 2012) or taking it into account by fitting a linear model with auto-correlated errors (Wiesmeier *et al.*, 2012). Xu *et al.* (2011) predicted subsoil SOC stocks based on a kriged topsoil OC content map using a depth distribution function. Also nonlinear models were used to relate SOC stocks to environmental covariates: Apart from a geostatistical model, Schröder *et al.* (2009) used regression trees for modelling the spatial distribution of OC stocks of forest soils in Northwestern Germany, and Martin *et al.* (2011) used boosted regression trees for spatial modelling of SOC stocks across France.

Most studies did not quantify the uncertainty of the SOC stock estimates (Schröder *et al.*, 2009; Tomlinson and Milne, 2006; Xu *et al.*, 2011; Wiesmeier *et al.*, 2012; Wördehoff *et al.*, 2011, 2012). Others reported 95 %-confidence intervals or standard errors of the mean SOC stock estimates per land use class and/or total study areas (Lettens *et al.*, 2005; Meersmans *et al.*, 2008; Oehmichen *et al.*, 2011). Krogh *et al.* (2003) estimated the sampling distribution of the mean SOC stock estimate by non-parametric bootstrapping and were able to calculate the 95 %-confidence interval. Meersmans *et al.* (2012) used Monte Carlo simulations to propagate the uncertainty in the input data of the employed pedotransfer function and the regression model to the predicted SOC content. Among the studies on SOC stocks in European forests we found, only Schröder *et al.* (2009), Martin *et al.* (2011) and Meersmans *et al.* (2012) evaluated the model performance by cross-validation. We did not find any study that validated stock estimates with

independent validation data.

None of cited studies used statistically robust estimation and prediction methods, although outliers are quite common in environmental datasets. Moreover, most studies used only a limited set of environmental covariates that did not completely characterize the conditions relevant for soil formation. Thus, the usage of comprehensive climatic, topographical and spectral information as potential covariates for statistical models that are furthermore fitted robustly, as in our study, seems to be novel. Also the uncertainty of the estimates and of the model performance were not generally quantified. Hence, our study is also in this respect different from previous work.

2 Method and materials

2.1 Study area

This study focused on the forest area of Switzerland (Figure 1). "Forest" was defined as in Giamboni (2008): The area classified by the VECTOR25 forest categories (Swisstopo, 2011) with the addition of the areas devastated by the hurricanes Vivian and Lothar. From this area the intersection with the National Inventory of Mires (BFS, 2001) was subtracted. 67 % of the forested area is dominated by coniferous trees, deciduous forests prevail only at lower altitude in the Southern Alps (BFS, 2001). Forests extend in Switzerland over altitudes ranging from 190 m to 2 390 m asl. The climatic conditions vary notably over the forested area: Annual precipitation ranges from 600 mm to 2 900 mm and mean annual temperature from -1 to 13 °C. Considerable variation is also found in the geologic parent material for soil formation, which includes among others limestones in the Jura mountains and in parts of the Pre-Alps, fluvio-glacial sediments of several quaternary glaciations and of the Tertiary on the Central Plateau and igneous and metamorphic rocks in the Alps and Southern Alps. This large variation of pedogenetic factors is reflected by diverse developments of soils that resulted in varying conditions for accumulation or degradation of SOC (Walthert *et al.*, 2004).

2.2 Data

2.2.1 SOC data

Over the past 30 years, WSL opened and described more than 1 000 forest soil profiles and collected and analysed soil samples from them. The sites were visited in several projects, mostly between 1990 and 2000 (Figure 2). For the purpose of this study, one can assume that the SOC data represent one time point because we did not find any significant linear time trends when we modelled the data. The positions of the soil profiles were recorded in the field on paper maps only. Therefore, the recorded locations are expected to differ from the real positions by about ± 25 m (Nussbaum, 2011). SOC was measured for 1 033 soil profiles (Figure 1). The sites had not been selected by randomized designs, but they are nevertheless quite evenly distributed over Switzerland, so there is good reason to assume that the SOC data represent OC stock in Swiss forest soils fairly well. Three sites, situated within the area of the National Mire Inventory and belonging to the validation set (profile identifiers 61, 63 and 67), were not used to fit the final model (cf. section 2.3.3, item (v)) and to compute the kriging predictions (cf. section 2.3.4).

The thickness of all soil horizons was recorded in the field, along with estimates of the volumetric content of gravel (particles with size >2 mm) in mineral soil horizons. Bulked soil samples were collected and analysed by soil horizons. For L horizons only thickness was recorded, but no samples were taken. For all other horizons, soil samples were collected and OC concentration was measured. When the pH of a soil sample was larger than 6.0 then carbonates were

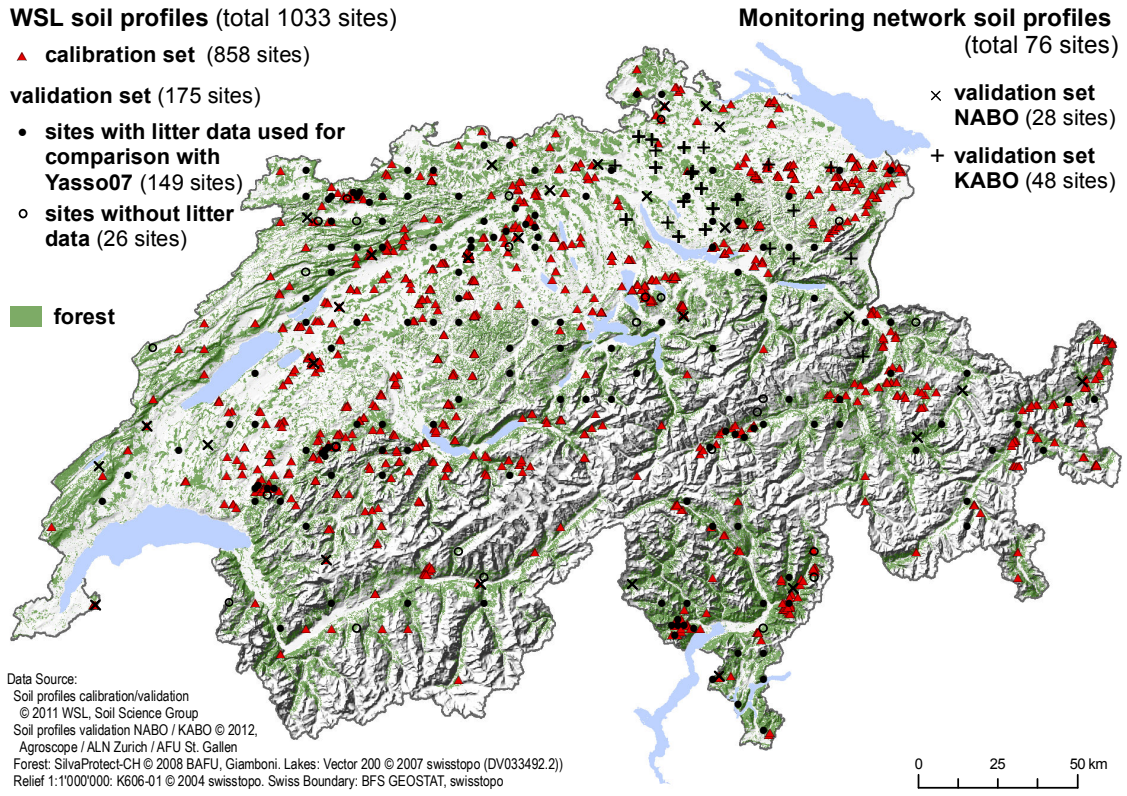

Figure 1: Map of the forest area of Switzerland (cf. section 2.1) with the positions of the 1 033 WSL soil profiles (subdivision into calibration and validation sets according to section 2.3.5) and the position of the validation sites of the national (NABO) and two cantonal (KABO) soil monitoring networks.

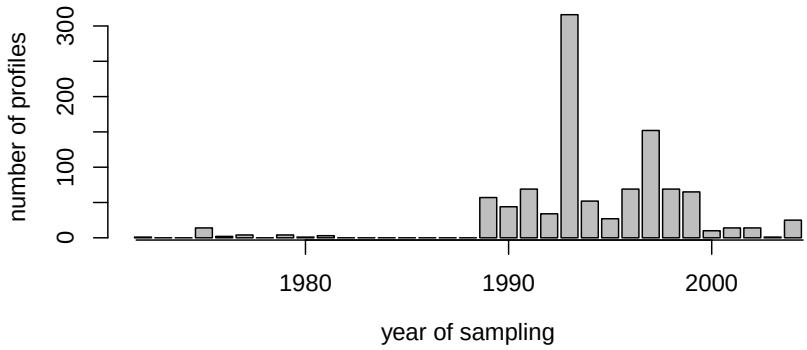

Figure 2: Number of soil profiles sampled per year.

removed by steaming hydrochloric acid prior to the measurement of the C content by an elemental C/N-analyser (combustion at 1 000 °C, Walthert *et al.*, 2010). Below this pH, carbonates were assumed to be absent. The missing information on the C content of L horizons was taken from Moeri (2007) and Wohlgemuth *et al.* (1995). From the study of Moeri complete data of 29 sites (with 4–8 replicates per site) and from Wohlgemuth *et al.* data of 18 "control sites" (8 replicates per site) was available. The median C content of these soil samples was equal to 44.3510^{-3} kg kg^{-1} and was assigned as C content to all L horizons.

Table 1: Descriptive statistics of OC measurements, computed for the forest floor, the mineral topsoil (0–30 cm), the mineral soil to 100 cm depth and the subsoil (30–100 cm, cs: calibration set, vs: validation set, stdv: standard deviation, mad: median absolute deviation).

	OC forest floor		SOC 0–30 cm		SOC 0–100 cm		SOC 30–100 cm	
	cs [t ha^{-1}]	vs [t ha^{-1}]	cs [t ha^{-1}]	vs [t ha^{-1}]	cs [t ha^{-1}]	vs [t ha^{-1}]	cs [t ha^{-1}]	vs [t ha^{-1}]
mean	20.06	16.31	70.90	80.91	108.91	132.92	38.02	52.01
median	6.65	6.65	60.89	73.01	89.44	111.84	27.89	35.62
stdv	64.49	25.05	41.00	41.23	72.50	83.30	40.59	51.56
mad	7.40	4.93	30.29	34.53	45.30	60.47	21.44	30.31

The bulk density was measured for 440 out of the about 5 000 mineral soil horizons sampled. A field estimate of the bulk density was available for all mineral soil horizons (ordinal variable with five levels). The measured bulk densities were classified by the field estimate, and the median density was computed for each level of the field estimate. The respective median was then assigned to all mineral soil horizons without bulk density measurements. An adjustment of the assigned value was necessary only for about 0.2 % of all mineral soil horizons. The plausibility of the assigned values was checked by an expert. The bulk density of the forest floor horizons were taken from Moeri (2007). The median density of the L horizons amounted to 75 kg m^{-3}, of the F horizons to 140 kg m^{-3} and of the H horizons to 220 kg m^{-3}.

The SOC stock S_i stored in horizon i per unit area [kg ha^{-1}] was calculated from the total volume V_i of the horizon per unit area [m^3 ha^{-1}], its volumetric gravel content G_i [m^3 m^{-3}], bulk density ρ_i [kg m^{-3}] and its gravimetric SOC content C_i [kg kg^{-1}] according to

$$S_i = V_i\,(1 - G_i)\,\rho_i\,C_i, \tag{1}$$

from which the stock S_c stored in a given depth compartment, was computed by

$$S_c = \sum_{i=1}^{h} w_i\,S_i, \tag{2}$$

where h is the number of horizons present at a given site and w_i is the "contribution" of horizon i to the stock in the compartment, which is equal to one if the horizon is fully contained in the compartment, equal to zero if no part of the horizon belongs to it and equal to

$$w_i = (z_c - \sum_{j=1}^{i-1} z_j)/z_i$$

if the compartment with thickness z_c only partially contains horizon i (z_i and z_j denote here the thickness of horizon i and j, respectively).

The frequency distributions of the SOC stocks stored in the three compartments forest floor, mineral soil 0–30 cm and 30–100 cm, respectively, are summarized by box-plots in Figure 3. On average, the largest stocks are stored in the mineral topsoil 0–30 cm (Table 1).

2.2.2 Input data

Table 2 gives an overview of the input data used to derive the covariates for the regression analysis, which is an important part of any geostatistical analysis. The last column lists the respective covariates. In general, the covariates could be computed for the whole study area (cf. section 2.1).

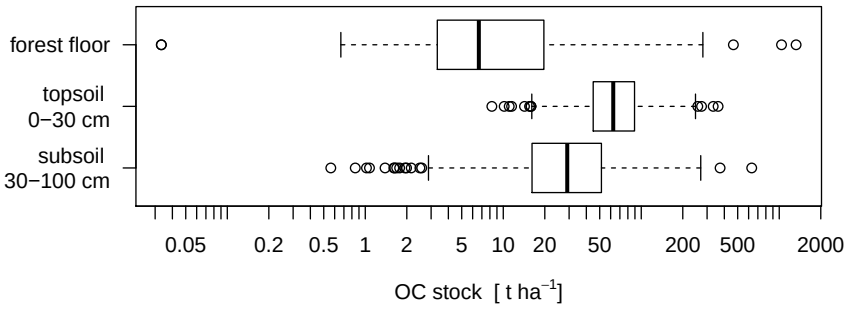

Figure 3: Box-plots of estimated SOC stocks for the three soil compartments forest floor, mineral soil 0–30 cm and 30–100 cm.

However, some input data was incomplete, e.g. values were missing for the SPOT5 mosaic where clouds covered the land surface. Nussbaum (2011) reviews the quality of the input data, and the references listed in Table 2 contain further information.

In view of the uncertainty of the recorded locations of the soil profiles (cf. section 2.2.1), the values of the covariates were averaged within circular local neighbourhoods, centred on the recorded profile locations and having radii either equal to 13 m, 19 m or 26 m. Depending on the type of the data, different statistics were computed: Arithmetic means for real numbers, medians for integers and the most frequent category for nominal or ordinal variables. For computing the kriging predictions (cf. section 2.3.4), the covariates of the regression models (cf. section 2.3.3) were extracted in the same way for the nodes of a square grid with spacing of 100 m.

2.3 Geostatistical analyses

2.3.1 Model

Our geostatistical analysis was based on the assumption that the area-specific SOC stock $S_c(\boldsymbol{s})$ at a given location $\boldsymbol{s}$, possibly transformed by some function $f(\cdot)$, can be modelled by

$$f(\, S_c(\boldsymbol{s})\,) = \boldsymbol{x}(\boldsymbol{s})^T\boldsymbol{\beta} + Z(\boldsymbol{s}) + \varepsilon(\boldsymbol{s}), \tag{3}$$

where $\boldsymbol{x}(\boldsymbol{s})^T\boldsymbol{\beta}$ stands for a linear regression model that accounts for the large-scale spatial trend in the data ($\boldsymbol{x}(\boldsymbol{s})$ is a vector with values of the covariates for location $\boldsymbol{s}$ [cf. section 2.2.2], $\boldsymbol{\beta}$ is a vector of unknown regression coefficients that must be determined from the data and ^{T} denotes the transpose). $Z(\boldsymbol{s})$ models the small-scale variation that is not accounted for by the regression model, but that is still spatially structured in the sense that $Z(\boldsymbol{s})$ and $Z(\boldsymbol{t})$ at two nearby locations $\boldsymbol{s}$ and $\boldsymbol{t}$ do not differ much and that the contrast between $Z(\boldsymbol{s})$ and $Z(\boldsymbol{t})$ increases on average with increasing distance between $\boldsymbol{s}$ and $\boldsymbol{t}$. Hence, $Z(\boldsymbol{s})$ stands for a normally distributed, autocorrelated random variable that has zero mean and that is characterized by a variogram function $\gamma(\boldsymbol{s} - \boldsymbol{t}; \boldsymbol{\theta})$ that depends on the lag distance $\boldsymbol{s} - \boldsymbol{t}$ and on a vector of unknown parameters $\boldsymbol{\theta}$ that must be determined again from the data. Lastly, $\varepsilon(\boldsymbol{s})$ is a spatially uncorrelated random variable that models the remaining spatially unstructured variation. In our robust geostatistical approach, $\varepsilon(\boldsymbol{s})$ need not be a normally distributed variable, but can be any random variable, symmetrically distributed around zero, and possibly having longer tails than a normal distribution, allowing thereby for (spatially unstructured) outliers in a dataset.

Table 2: Overview of the input data sets, their spatial resolution and (partial) list of the derived covariates (r: grid resolution).

Category	Original data set	r [m]	Derived covariates
Soil / Parent Material	Soil map 1:200 000 (BFS, 2001) Geological map 1:500 000 (Swisstopo, 2011a) Geotechnical map 1:200 000 (BFS, 2001)		soil map (physiographic units) amended with information from geological map; median soil properties (pH WSL, 2011), mass of soil with particles ≤ 2 mm, cation exchange capacity, calcium and iron stocks assigned to soil and geotechnical map units
Climate	MeteoSwiss 1961–1990 (Zimmermann and Kienast, 1999) MeteoTest 1975–2010 (Remund *et al.*, 2011) MeteoSwiss norm data 1961–1990 (MeteoSwiss, 2011)	25/100 250 2000	annual/monthly temperature and precipitation, cloud cover, sunshine duration, radiation, temperature degree days, continentality, temperature variation, ratio actual/potential evapotranspiration, site water balance
	Air pollutants (BAFU, 2011)		nitrogen wet and dry deposition, NH_3 concentration in air
Vegetation	Percentage of coniferous trees (BFS, 2001)	25	4 categories
	Percentage of coniferous trees (unpublished data, A. Baltensweiler (WSL), personal communication)	25	continuous values between 0 and 100 %
	Biogeographical regions (Gonseth *et al.*, 2001)	25	6 / 10 regions
	SPOT5 radcor mosaic of Switzerland (Swisstopo and NPOC, 2009)	10	reflection in spectral bands (red, green, near infrared), ratios of reflection, IHS transformed reflections, vegetation index, NDVI
	Digital surface model DOM (Swisstopo, 2011b)	2	height of vegetation
Relief	Digital elevation model DHM25 (Swisstopo, 2011b) Digital terrain model DTM-AV (Swisstopo, 2011b)	25 2	for both elevation models: height, slope angle, aspect, northness and eastness, (smoothed) curvatures (total/planar/profile), topographic position indices (Zimmermann, 2000; Jenness, 2006) with radii from 6 m to 2 km, flow accumulation area, topographic indices (single and multiflow, cf. Tarboton, 1997)

2.3.2 Robust estimation of model parameters

As mentioned in the previous section, the unknown parameters of the model (equation 3) are the regression coefficients $\boldsymbol{\beta}$ and the variogram parameters $\boldsymbol{\theta}$ (nugget σ_0^2, sill σ_1^2, range α) that must be estimated from the SOC data and the values of the covariates. We used a novel robust restricted maximum likelihood (REML) method for this. In short, the applied new method, which is based on robustifying the Gaussian estimating equations and is related to the Maximum Likelihood (ML) II Proposal of Richardson and Welsh (1995), automatically identifies observations in a dataset that do not follow the dominant patterns of the majority of observations and gives those outlying observations a small weight when estimating the model parameters. The interested reader can find more details about our robust REML method in Künsch *et al.* (2011, in preparation).

Given the SOC stock measurements $S_c(\boldsymbol{s}_i)$ and the values of the covariates $\boldsymbol{x}(\boldsymbol{s}_i)$ for a set of soil profile locations $\boldsymbol{s}_i$, $i = 1, 2, \ldots, n$ the estimation algorithm provides estimates $\hat{\boldsymbol{\beta}}$ of the regression coefficients along with an approximate estimate of their covariance matrix, estimates $\hat{\boldsymbol{\theta}}$ of the variogram parameters, and predictions $\hat{Z}(\boldsymbol{s}_i)$ of the normally distributed variable for the locations of the soil profiles, along with the associated kriging variances (minimized mean squared prediction errors).

2.3.3 Model building

Preliminary remark: For the analyses described in the sequel we used unless annotated otherwise only the 858 observations of the calibration subset (cf. section 2.3.5 for details).

Given a set of covariates, the robust estimation method sketched in the previous section provides estimates of the regression coefficients along with approximate standard errors (SE), so that (approximate) significance testing of a particular set of regression coefficients is possible (based on a Wald-Test). However, we had a large number of covariates, of which some were mutually strongly correlated. Furthermore, based on the results obtained by Nussbaum (2011), we had good reasons to believe that many covariates would not be statistically significant. Thus, we faced a covariate selection problem with many non-significant and partly collinear covariates. In addition, some nominal or ordinal covariates distinguished many categories, with very few OC stock measurements for some categories. Hence, the model building process involved

(i) selecting an appropriate transformation $f(\cdot)$ for the response variables $S_c(\boldsymbol{s})$),
(ii) the selection of a parsimonious set of relevant covariates,
(iii) the aggregation of sparsely populated and mutually not differing levels of categorical covariates,
(iv) the choice of a particular parametric variogram function and the optimal value of the tuning constant of our robust estimation algorithm, and
(v) fitting of a final model with the merged calibration and validation sets for computing the kriging predictions.

(i) Transformation of response variables Figure 3 shows that the empirical distributions of the stock measurements were positively skewed. Ignoring the auto-correlation and possible outliers in the dataset and using all covariates, we fitted the parameter of a Box-Cox-Transformation (Box and Cox, 1964) by ML. For all considered depth compartments, the estimated parameters of the transformations were close to zero, so that the eventually chosen transformation by the natural logarithm approximated a Box-Cox-Transformation in all instances well.

(ii) Selection of relevant covariates About 300 potential covariates were derived from the input data and were available for regression modelling (cf. section 2.2.2). The distribution of the

values of some terrain attributes and climate covariates were positively skewed, and the covariates were transformed either by the square root or the natural logarithm. Relevant covariates were then selected by a procedure consisting of three steps:

1. An initial set of covariates, from which in the subsequent steps the relevant ones were selected, was chosen based on correlation-biplots (Gabriel, 1981). For this, we used the values of the covariates for all 1 033 WSL soil profiles. Correlation-biplots illustrate multiple linear dependencies among a set of variables well. By this step we could eliminate a considerable number of covariates that were strong correlated with others (cf. Figure 19 in Appendix G).

2. The least absolute shrinkage and selection operator LASSO (Hastie *et al.*, 2009, sec. 3.4) is a covariate selection tool that minimizes a penalized sum of the squared residuals. LASSO uses as penalty the sum of the moduli of the regression coefficients. The algorithm likely excludes correlated and non-relevant covariates. LASSO was applied (a) to the full set of covariates, (b) to the subset resulting from step 1 and (c) to these sets enriched by interactions terms, i.e., new covariates obtained by multiplying two covariates used in (a) and (b). We considered in particular products between climate and relief attributes on the one hand and dummy indicator variables for biogeographical regions, soil and geotechnical map units on the other. The LASSO covariate subset with the smallest mean squared error was chosen for the following step 3.

3. Starting with the best LASSO covariate subset from step 2, the parameters of the geostatistical model (equation 3) were then estimated as described in section 2.3.2. First, we removed step by step non-relevant covariates from the LASSO subset by tenfold cross-validation (e.g. Hastie *et al.*, 2009, sec. 7.10). For this, we divided the calibration set randomly into 10 subsets, excluded one subset in turn and fitted the model by robust REML to the remaining data, and predicted the log-transformed SOC stock of the excluded subset by robust kriging (cf. section 2.3.4). We used the continuous ranked probability score (CRPS, Gneiting *et al.*, 2007) as main criterion for the model selection.

(iii) Merging levels of categorical covariates In step 3 described above, the categories of nominal and ordinal covariates were merged based on partial residual plots (e.g. Faraway, 2005, sec. 4.2), and the benefit of the aggregation was evaluated again by tenfold cross-validation.

(iv) Choice of variogram function and robustness tuning constant After aggregating the levels of categorical covariates, a range of parametric variogram functions was tested by cross-validation. For the best-fit variogram model and the final set of covariates, an optimal value of the tuning constant c, controlling the sensitivity of the robust REML procedure to outliers, was similarly selected by tenfold cross-validation.

(v) Fitting of final model For computing the kriging predictions, we fitted a final model with the same structure as in (ii) and (iii) to the merged calibration and validation sets (excluding 3 profiles situated within the area of the National Mire Inventory, cf. section 2.2.1), using the tuned variogram function and the tuned robustness constant as in (iv).

2.3.4 Predicting SOC stock at unsampled locations by robust kriging

The log-transformed SOC stocks were predicted for the nodes of a grid with spacing equal to 100 m and for the sites of the validation sets by robust kriging. One needs for this the estimated parameters of the best-fit model and for each prediction location the values of the respective

covariates (cf. equation 9 in Appendix B). Outlying observations automatically receive small weight, and our robust predictions are therefore insensitive to outliers in the SOC stock dataset. We used the standard unbiased back-transformation for lognormal kriging (e.g. Cressie, 2006) to transform the predictions of the log-transformed stocks back to the original scale of the measurements. The mean stocks per NFI production region (and for whole Switzerland), stratified by altitude, along with respective mean squared prediction errors (block kriging variances), were computed from the lognormally back-transformed predictions at the nodes of the 1-ha grid by block kriging (cf. equation 13 and 14 in Appendix B for details).

2.3.5 Evaluation of model performance

The predictive power performance of the fitted geostatistical models was evaluated in three ways:

(i) by tenfold cross-validation of the final model with the calibration dataset (cf. section 2.3.3),
(ii) by predicting the SOC stocks for the independent WSL validation dataset and by comparing the predictions with the respective measurements, and
(iii) by predicting and comparing similarly SOC stocks measured on forest sites of the national (Desaules *et al.*, 2006) and two cantonal soil monitoring networks (Gasser *et al.*, 2009; AFU, 2012).

We had to use the WSL dataset also for a comparison of measured SOC stocks with initial stock estimates obtained by the Yasso07 model (Weggler *et al.*, 2012). Yasso07 needs data on litter input into the soil. Such data were available only for 269 WSL soil profile sites. The dataset could therefore not be split randomly into calibration and validation sets. Rather, we selected those 134 out of the 269 sites that were evenly arranged on a 8×8-km-grid over Switzerland, and another 41 sites, selected such that we achieved more or less a balanced representation of the various soil map units and of the forest types in both the validation and calibration sets. Nussbaum (2011) describes in more detail how the dataset was split. In summary, 858 observation were assigned to the calibration and 175 to the validation set (Figure 1).

The following statistics of the relative (cross-validation) prediction errors were computed to characterize the marginal bias and the overall precision of the predictions:

$$\text{BIAS} = -\frac{1}{n}\sum_{i=1}^{n}\frac{(S_c(\boldsymbol{s}_i)-\hat{S}_c(\boldsymbol{s}_i))}{S_c(\boldsymbol{s}_i)}, \tag{4}$$

$$\text{robBIAS} = -\operatorname*{median}_i\left(\frac{(S_c(\boldsymbol{s}_i)-\hat{S}_c(\boldsymbol{s}_i))}{S_c(\boldsymbol{s}_i)}\right), \tag{5}$$

$$\text{RMSE} = \left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{S_c(\boldsymbol{s}_i)-\hat{S}_c(\boldsymbol{s}_i)}{S_c(\boldsymbol{s}_i)}\right)^2\right)^{1/2}, \tag{6}$$

$$\text{robRMSE} = \operatorname*{MAD}_i\left(\frac{S_c(\boldsymbol{s}_i)-\hat{S}_c(\boldsymbol{s}_i)}{S_c(\boldsymbol{s}_i)}\right). \tag{7}$$

The symbol n denotes the number of observations in the (cross-)validation set and MAD stands for median absolute deviation. In addition, CRPS and a robust estimate of the coefficient of determination R^2 (Croux and Dehon, 2003)[3] was computed by

$$R^2 = 1 - \left(\frac{\sum_{i=1}^{n}|S_c(\boldsymbol{s}_i)-\hat{S}_c(\boldsymbol{s}_i)|}{\sum_{i=1}^{n}|S_c(\boldsymbol{s}_i)-\operatorname*{median}_i(S_c(\boldsymbol{s}_i))|}\right)^2. \tag{8}$$

[3]This criterion is tailored for a robust regression analysis where the sum of the moduli of the residuals is minimized (L1-regression). We did not use such an approach, but nevertheless found this criterion well-suited to summarize the predictive power of the fitted models.

For computing CRPS, we assumed that the prediction errors of the log-transformed stocks were normally distributed with mean zero and variance equal to the robust kriging variance of the log-transformed stock.

2.4 Design-based estimation of OC stock of forest floor

In the course of the geostatistical analyses we discovered that the model fitted to the OC stock data of the forest floor predicted the WSL validation data only poorly (cf. section F.2 in Appendix). As an alternative to the geostatistical analysis, we computed therefore design-based[4] estimates of the mean stocks for the NFI production regions and for the whole country. Three soil profiles (profile identifiers 357, 361 and 387) with peat topsoil and very large stocks (1 325, 465 and 1 038 t ha^{-1}, respectively) were omitted for design-based estimation. Furthermore, the only site situated on the Central Plateau above 1200 m asl (profile identifier 579) was assigned to region "Jura > 1200 m" because it was close to this region. Hence, there remained no data for the region "Central Plateau > 1200 m", and no design-based estimate could be computed. This is not a serious problem, because this region covers an area of only about 25 km^2 (Table 3).

To compute the design-based estimates, we assumed that the soil profile sites had been chosen by a simple random sample without replacement, which is obviously not true. Furthermore, we supposed that each soil profile was representative for a surrounding area of 1 ha and that the forested area is equal to 1 283 800 ha. This defined the finite population size $N = 1\,283\,800$, from which the $n = 1\,030$ soil profiles were supposedly sampled[5].

The design-based estimate of the mean stock in the forest floor for whole Switzerland is simply the arithmetic mean, $\bar{S}_c$, of the measurement, and its standard error can be estimated by

$$\mathrm{SE}[\bar{S}_c] = \left(\frac{(N-n)}{N} \frac{s^2}{n} \right)^{1/2},$$

where s^2 is the customary sample variance of the measurements. For the NFI production regions, stratified by altitude, the mean stocks along with the associated standard errors were estimated by a ratio domain estimate (cf. Särndal *et al.*, 1992, sec. 5.8).

[4] For the difference between model- and design-based estimation we refer the reader to de Gruijter *et al.* (2006, pp. 15).

[5] The assumed size of the forest area is not critical because the finite population correction $1 - n/N$ was close to one.

3 Results

3.1 OC stock in forest floor

We report here the block kriging predictions of the mean stocks per NFI productions region stratified by altitude, along with the respective design-based estimates (Table 3). Since the predictive performance of the fitted geostatistical model turned out to be weak in the validation, we have not much confidence in the geostatistical predictions, and we report the remaining results of the geostatistical analysis of the forest floor OC data in Appendix F.

The design-based estimates of the mean stocks varied from 3 t ha^{-1} (Alps ≤ 600 m) to 33 t ha^{-1} (Alps > 1200 m), with a mean estimate for the whole country of 16.7 t ha^{-1} (SE: 0.8 t ha^{-1}; 95 %-confidence interval: [15.1, 18.3] t ha^{-1}). For most production regions and for the whole country the block-kriging predictions of the stocks were larger than the design-based estimates, although the 95 %-confidence regions (estimate or prediction ± 1.96 SE) frequently overlapped, in particular also for the average stock of Switzerland. The comparison of the kriging predictions with the measurements of the WSL validation set indicated that kriging overestimated the stocks in the forest floor systematically (median estimate of relative bias: 99 %, cf. Table 4). Hence, one might therefore prefer the design-based estimates over the kriging predictions.

In the Pre-Alps, Alps and Southern Alps both types of estimates showed increasing stocks with increasing altitude. This was less clear for the Jura region and the Central Plateau, where

Table 3: Design-based estimates and robust block kriging predictions of the mean OC stocks in the forest floor of the five production regions of the NFI, stratified by altitude, and of Switzerland as a whole (A_{forest}: forest area, n: number of soil profiles used for computing the predictions and for design-based estimation, $\bar{S}_c$: design-based estimate, $\hat{S}_c$: block kriging prediction; SE: standard error).

Region	Altitude asl [m]	A_{forest}[a] [ha]	n	$\bar{S}_c$ [t ha^{-1}]	$\hat{S}_c$ [t ha^{-1}]	SE[$\bar{S}_c$] [t ha^{-1}]	SE[$\hat{S}_c$] [t ha^{-1}]
Jura	≤ 600	51 647	22	9.5	10.1	1.57	1.00
	(600–1200]	116 363	68	7.5	13.2	0.70	1.37
	> 1200	23 507	5[b]	7.8	20.7	1.74	4.21
Central Plateau	≤ 600	125 578	199	8.7	10.5	0.68	0.82
	(600–1200]	91 196	105	11.4	12.0	1.45	0.92
	> 1200	2 480	1[c]	–	19.7	–	6.46
Pre-Alps	≤ 600	8 619	50	7.5	9.6	1.25	0.97
	(600–1200]	129 997	189[d]	16.3	17.1	1.55	1.26
	> 1200	59 571	68[e]	26.2	22.9	4.77	2.17
Alps	≤ 600	7 901	4	3.1	9.7	0.47	1.58
	(600–1200]	95 101	64	20.0	17.0	2.64	1.59
	> 1200	212 260	164	33.4	32.0	3.53	3.46
Southern Alps	≤ 600	18 460	22	8.2	13.2	1.62	1.92
	(600–1200]	51 485	40	11.0	19.7	2.11	2.40
	> 1200	57 330	32	30.8	36.6	5.43	4.95
Switzerland		1 051 495	1 033[f]	16.7	19.5	0.83	1.22

[a] number of nodes of 1-ha grid intersected with five production regions and three altitude classes less the nodes, where values of covariates were missing (cf. section 2.2.2)
[b] 6 profiles for design-based estimation
[c] this profile was assigned to region the "Jura > 1200 m" for design-based estimation
[d] 188 profiles for design-based estimation
[e] 66 profiles for design-based estimation
[f] 1 030 profiles for design-based estimation

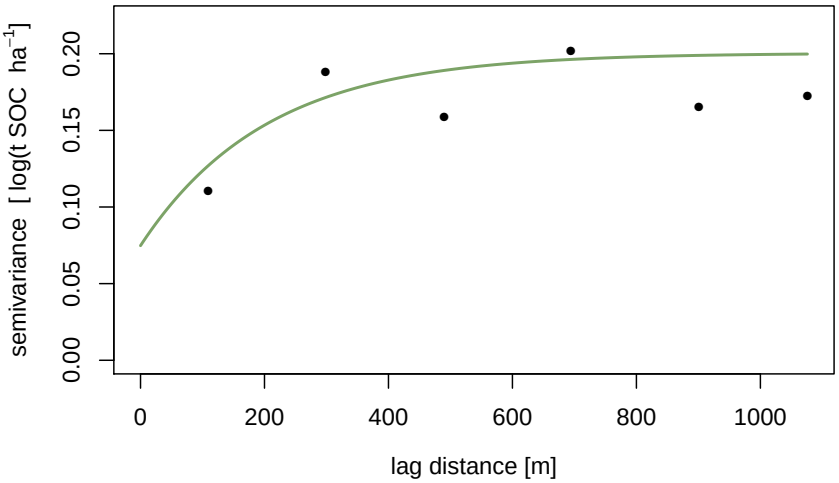

Figure 4: Fitted exponential variogram of the log-transformed SOC stock in the top 30 cm of the mineral soil (line: robust REML estimate computed with covariates listed in section 3.2.1 and tuning constant $c = 2$, dots: method-of-moments estimate of sample variogram of robust regression residuals).

only the block kriging prediction suggested a positive dependence of stocks on altitude. However, only few profiles were available for the stratum "> 1200 m" there, and the standard errors of the block kriging estimates were large. Contrary to what one would expect from a model-based procedure, the standard errors of the block kriging predictions were not clearly smaller than the standard errors of the design-based estimates. Considering the positive bias and the poor precision of kriging in the validation for the forest floor stock (Table 4 and section F.2 in Appendix) we have more confidence in the design-based estimates than in the block kriging predictions.

3.2 SOC stock in mineral soil 0–30 cm

3.2.1 Model

The model building procedure selected just five covariates for the model of the log-transformed SOC stock in the top 30 cm of the mineral soil (cf. residual plots in Appendix G):

- square root of mean annual precipitation,
- reflection in the near-infrared band of the SPOT5 mosaic,
- topographic position index with radius 500 m according to Jenness (2006), with a separate coefficient for soils rich and poor in clay,
- median mass of soil particles with diameter < 2 mm assigned to geotechnical map units, and
- categorical covariate distinguishing five aggregated soil map units (for aggregation, see residual plots in Appendix G).

The residuals of the linear regression with the above covariates were spatially correlated with a nugget/sill ratio of 0.37 and an effective range of about 600 m (cf. Figure 4). The correlation was stronger here than in the model of the forest floor OC stock (section F.1 in Appendix). The optimal value of the robust tuning constant was $c = 2$. This indicates that the robustly estimated model fitted the data better than a customary REML estimate.

Table 4: Statistics of relative OC stock prediction errors, computed for the measurements of the WSL and of the soil monitoring datasets (NABO/KABO) with the respective final models (statistics defined in section 2.3.5, cv: cross-validation, vs: validation set).

	WSL forest floor		WSL 0–30 cm		WSL 0–100 cm		NABO/KABO vs	
	cv	vs	cv	vs	cv	vs	0–30 cm	0–100 cm
BIAS	0.135	2.634	0.207	0.135	0.232	0.152	0.198	0.117
robBIAS	0.891	0.990	0.096	0.070	0.124	0.066	0.107	0.070
RMSE	22.91	8.387	0.568	0.488	0.646	0.556	0.508	0.492
robRMSE	1.351	1.467	0.410	0.388	0.436	0.420	0.426	0.429
CRPS	0.791	0.647	0.238	0.221	0.252	0.247	0.238	0.249
R^2	–	–	0.309	0.337	0.305	0.403	0.216	0.166

3.2.2 Model validation

Figure 5 shows the measurements of the stock in the top 30 cm of the mineral soil, plotted against the respective predictions for the calibration set (cross-validation results, prediction of log-transformed stock) and for the WSL validation set (independent validation data, back-transformed predictions and measurements plotted on log-scale). The lines of the loess smoother (Cleveland *et al.*, 1992) are in both panels close to the 1:1-line, indicating absence of noticeable (conditional) bias. This is confirmed by the BIAS and robBIAS statistics in Table 4: Irrespective how the statistic was computed, the relative (marginal) bias was always less than 20 %. However, the scatter of the points around the 1:1-line is quite large in both panels of Figure 5, and this was mirrored in the large root mean squared relative errors: The MAD of the relative prediction errors was about 40 %, irrespective of the data set, and the non-robust estimates were even larger (49 % and 57 %, respectively). According to the robust estimate of the coefficient of determination the model explained only about a third of the variation in the measurements, and the robust estimate of the linear correlation between predictions and observations was equal to 0.55 in cross-validation and 0.58 for the validation set. The differences in these statistics between cross-validation and the independent WSL validation set were small. Hence, there was no indication that the model over-fitted the data.

The kriging variances slightly over-estimated the magnitude of the prediction errors, as only 6 of the 175 validation observations lay outside of the 95 %-prediction intervals (expected: 9 observations, Figure 6).

3.2.3 Prediction

For computing the predictions, the parameters of the final model were estimated with the data of 1 022 sites, based on the set of covariates listed in section 3.2.1. For 8 sites, the information on the reflectance in the near-infrared band was missing, and 3 soil profiles, situated within the area of the National Mire Inventory were excluded (cf. section 2.2.1). The robust lognormal kriging predictions are mapped in Figure 7 for the nodes of the 1-ha grid, and the block kriging predictions of the mean stocks for the NFI production regions and for whole Switzerland are listed in Table 5.

The largest stocks (> 120 t ha^{-1}) are stored at higher elevation in the Jura, in parts of the Pre-Alps and in the southern part of the Canton of Ticino. Small stocks (< 60 t ha^{-1}) are found at low elevation on the Central Plateau and in the inner-alpine valleys. A large part of the Jura region and of the Pre-Alps have intermediate stocks.

The predictions of the mean stocks varied for the NFI production regions between 55 t ha^{-1} (Central Plateau ≤ 600 m) and about 120 t ha^{-1} (Central Plateau and Jura > 1200 m). For the

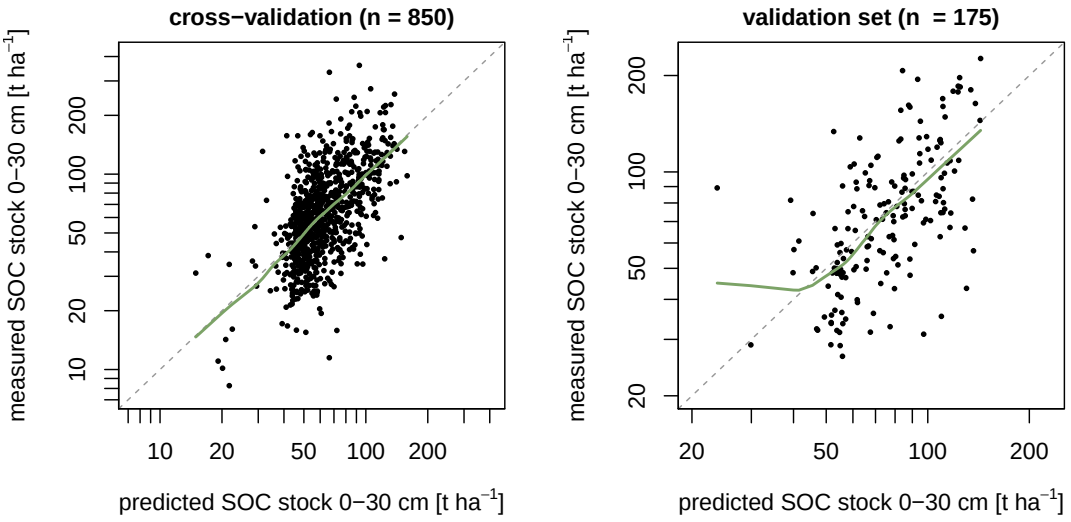

Figure 5: Scatterplots of measured against predicted SOC stock in the top 30 cm of the mineral soil. Cross-validation predictions of log-transformed stocks for calibration set (left panel) and lognormally back-transformed predictions computed with the calibration data for the sites of the WSL validation set (right panel, green lines: loess scatterplot smoothers).

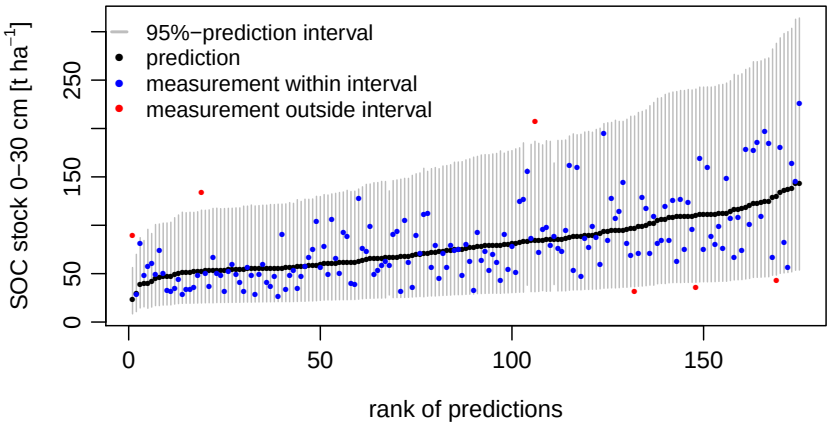

Figure 6: Ranked predictions of the SOC stock in the top 30 cm of the mineral soil for the WSL validation set (black) with 95 %-prediction intervals (vertical grey lines). Measurements inside the intervals are shown in blue, those outside in red.

whole country, a mean value of 80 t ha^{-1} was predicted with a standard error of 1.5 t ha^{-1}, yielding a 95 %-prediction interval of [76.9, 82.9] t ha^{-1}. Except for the Southern Alps, the predictions showed again the tendency of increasing stocks with increasing altitude, in particular for the Jura and the Pre-Alps, where the 95 %-tolerance regions of the predictions (prediction ± 1.96 SE) generally did not overlap for the different elevation strata. In spite of the rather poor precision of the point predictions, demonstrated in the previous section for the WSL validation set, the standard errors of the predicted mean stocks in the NFI regions were small (2–4 t ha^{-1}) in comparison with the stocks themselves (60–120 t ha^{-1}). This gain in precision is due to the "bulking" effect of block kriging: By predicting mean values, the small-scale variation is averaged out.

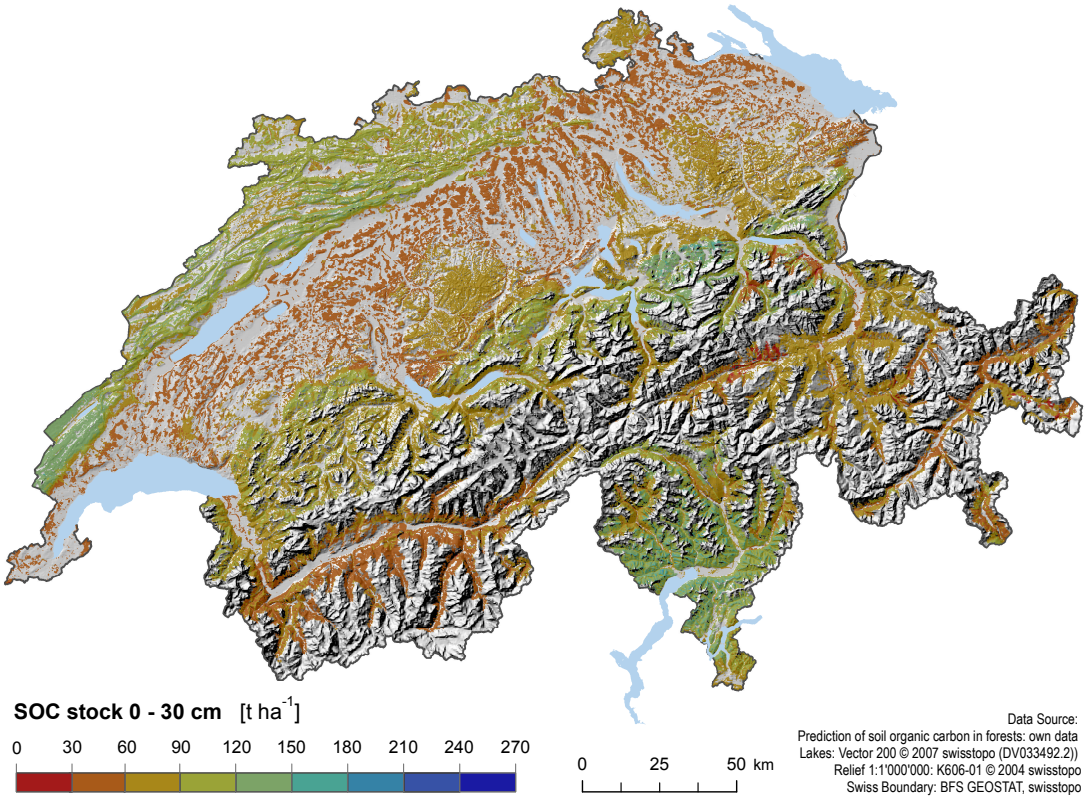

Figure 7: Robust lognormal kriging prediction of the SOC stock in the top 30 cm of the mineral topsoil of Swiss forests (computed with best-fit model with covariates according to section 3.2.1 and tuning constant $c = 2$, smoothed with a focal mean with a radius of 1 pixel = 100 m).

3.3 SOC stock in mineral soil 0–100 cm

3.3.1 Model

Compared with the model of the topsoil stock (section 3.2.1) very similar covariates were chosen for the model of the SOC stock stored in the mineral soil in 0–100 cm depth:

- square root of March precipitation,
- reflection in the near-infrared band of the SPOT5 mosaic,
- slope angle computed from high resolution (2 m) terrain model, and
- categorical covariate distinguishing nine soil map units (for aggregation, see residual plots in Appendix G).

The residuals of the linear model were to the same degree spatially auto-correlated as the residuals of the topsoil model (nugget/sill ratio 0.41, effective range 660 m, Figure 8). According to the cross-validation results, robust REML with a tuning constant $c = 2$ fitted the data again better than a non-robust REML fit.

3.3.2 Model validation

Figure 9 shows the measurements plotted against the predictions of the stocks. As noticed already for the topsoil, the cross-validation predictions and the predictions for the WSL validation set showed only little bias (loess smoother lines close to 1:1-line, robBIAS $\leq$ 12 %). Compared to the predictions of the topsoil stock, those of the stock stored to 100 cm depth were less precise

Table 5: Robust block kriging prediction of the mean SOC stocks in the mineral topsoil (0–30 cm depth) and in the mineral soil from 0–100 cm of the five production regions of the National Forest Inventory, stratified by altitude, and of Switzerland as a whole (n: number of soil profiles used for computing predictions, $\hat{S}_c$: block kriging prediction; SE: standard error).

Region	Altitude [m]	0–30 cm n	0–30 cm $\hat{S}_c$ [t ha^{-1}]	0–30 cm SE[$\hat{S}_c$] [t ha^{-1}]	0–100 cm n	0–100 cm $\hat{S}_c$ [t ha^{-1}]	0–100 cm SE[$\hat{S}_c$] [t ha^{-1}]
Jura	≤ 600	22	82.6	3.34	22	104.8	6.51
	600–1200	68	102.0	3.56	68	145.9	5.02
	> 1200	5	121.3	5.39	5	168.1	7.52
Central Plateau	≤ 600	199	55.4	1.55	199	81.4	2.11
	600–1200	102	62.1	1.68	102	92.2	2.11
	> 1200	1	122.0	7.07	1	171.0	10.29
Pre-Alps	≤ 600	50	66.1	2.06	50	93.2	2.38
	600–1200	184	75.9	2.00	184	112.9	2.95
	> 1200	66	95.8	3.27	66	153.8	4.98
Alps	≤ 600	4	66.5	2.44	4	99.6	2.86
	600–1200	64	74.4	2.42	64	120.9	3.62
	> 1200	163	69.5	1.85	163	115.3	3.19
Southern Alps	≤ 600	22	102.4	4.07	22	196.3	9.84
	600–1200	40	109.0	4.09	40	209.5	9.83
	> 1200	32	107.1	4.11	32	192.4	8.04
Switzerland		1 022[a]	79.9	1.52	1 022[a]	125.8	2.41

[a] data of near-infrared reflectance was missing for 8 out of 1 033 WSL profile sites because of clouds in the SPOT5 mosaic and 3 sites laying within the National Mire Inventory were excluded (cf. section 2.2.1).

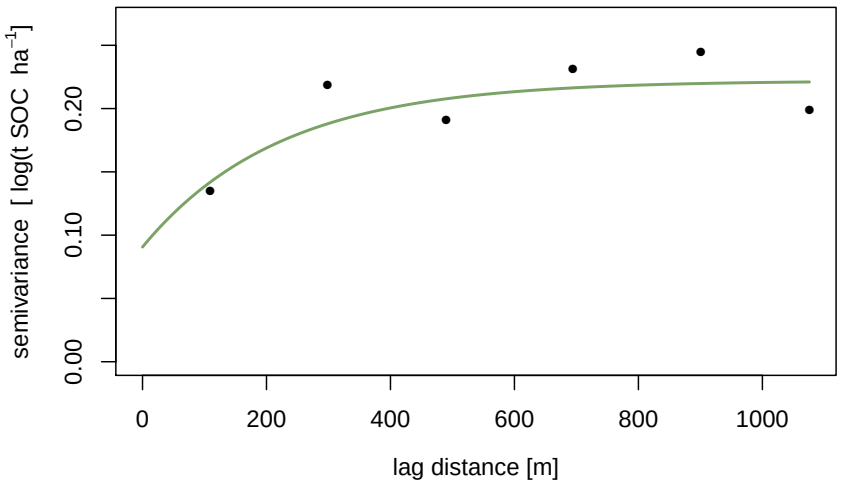

Figure 8: Fitted exponential variogram of the log-transformed SOC stock in 0–100 cm depth of the mineral soil (line: robust REML estimate computed with covariates listed in section 3.3.1 and tuning constant $c = 2$, dots: method-of-moments estimate of sample variogram of robust regression residuals).

(robRMSE: 44 % [cross-validation 0–100 cm] and 42 % [validation set 0–100 cm] vs. 41 % [cross-validation 0–30 cm] and 39 % [validation set 0–30 cm]). The larger values of CRPS

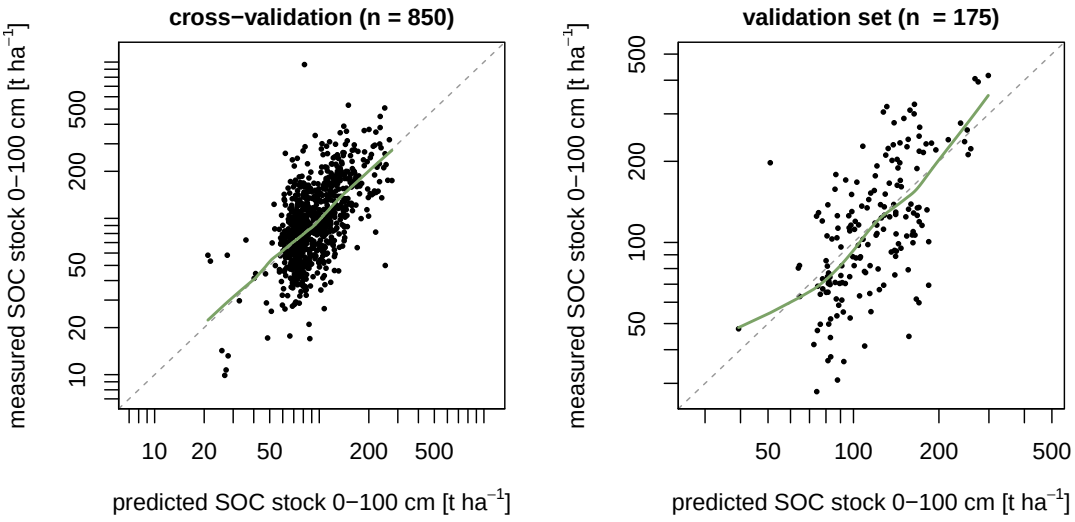

Figure 9: Scatterplots of measured against predicted SOC stock in 0–100 cm depth of the mineral soil. Cross-validation predictions of log-transformed stocks for calibration set (left panel) and lognormally back-transformed predictions computed with the calibration data for the sites of the WSL validation set (right panel, green lines: loess scatterplot smoothers).

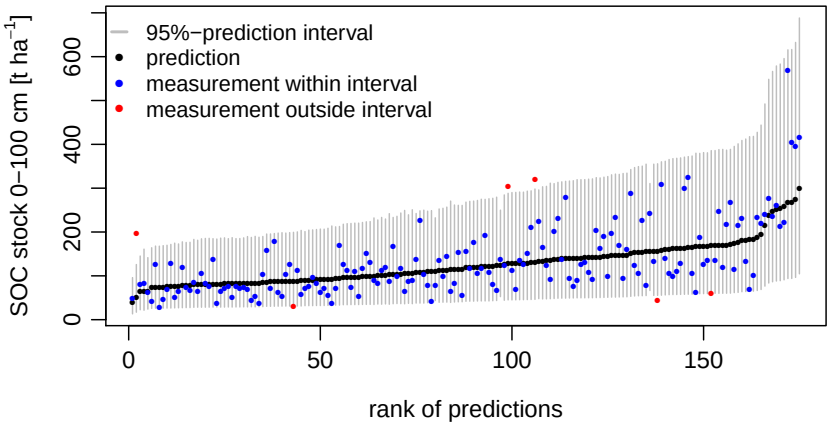

Figure 10: Ranked predictions of the SOC stock in 0–100 cm depth of the mineral soil for the WSL validation set (black) with 95 %-prediction intervals (vertical grey lines). Measurements inside the intervals are shown in blue, those outside in red.

(0.25 vs. 0.22–0.24) confirm that the total stock in the mineral soil were slightly less precisely predicted than the topsoil stock, although the robustly estimated R^2-statistic (40 %) was for the stock 0–100 cm of the validation set a bit larger than for the topsoil stock (34 %). The robustly estimated correlation between predictions and measurements was equal to 0.63 for the validation data and again equal to 0.55 in cross-validation. Since the predictive power of the model for the validation set was about the same as in cross-validation, we conclude again that the model did not over-fit the data.

Similar to the topsoil stock, the kriging variance were again somewhat too large, as only 6 instead of the expected 9 observations fell outside of the 95 %-prediction intervals (Figure 10).

3.3.3 Prediction

For computing the predictions, the parameters of the final model were estimated with the data of 1 022 sites, based on the set of covariates listed in section 3.3.1. Eleven observations were omitted for the same reasons as mentioned in section 3.2.3 for the topsoil stocks. The robust lognormal kriging predictions of the stocks stored to 100 cm depth are mapped in Figure 11 for the nodes of the 1-ha grid, and the block kriging predictions of the mean stocks for the NFI production regions and for whole Switzerland are listed in Table 5.

The spatial patterns in Figure 11 are broadly the same as in the map of the topsoil stocks (0–30 cm), except that the predicted total stocks were much larger in parts of Ticino where a combination of large precipitation, forest fire, and metaphorphic parent material lead to an accumulation of organic matter in deeper soil horizons (cf. section 4 for a more detailed discussion). Large stocks up to 400 t ha^{-}1 were also found in the parts of the Eastern Pre-Alps, where precipitation is large and water-logged soils developed on Flysch parent material. Very small total stocks in the mineral soil were predicted for Permian Verrucano sand stones in parts of the Eastern Pre-Alps and Alps.

For the whole country, about 126 t ha^{-1} are stored in the mineral soil down to 100 cm. The block kriging standard error was equal to 2.4, yielding a 95 %-prediction interval of [121.1, 130.5] t ha^{-1}. Thus, about 45 t ha^{-1} SOC are stored in the subsoils of Swiss forests. The largest stocks were predicted for the Southern Alps (196–210 t ha^{-1}), irrespective of altitude. The smallest stocks (80–90 t ha^{-1}) were predicted for the two lower elevation strata of the Central Plateau and the lowest elevation stratum of the Pre-Alps. Except for the Southern Alps, the stocks at higher elevation were on average again larger, in particular, when one compares the strata < 600 m and 600–1200 m.

3.4 Model validation with data from soil monitoring networks

The predictive power of the models was evaluated in sections 3.2.2 and 3.3.2 with the WSL data. Here we present an additional validation of the fitted models with data of the national and two cantonal soil monitoring networks. For the top 30 cm of the mineral soil, SOC stock data of 75 soil profiles were available, and for the mineral soil down to 100 cm depth we had complete SOC stock measurements of 51 soil profiles. Other cantonal monitoring networks did not measure SOC stocks of forest soil profiles, especially bulk density measurements were missing, and no further validation data was available for the forest floor. To assess the quality of the predictions, we used the same plots and criteria as before (cf. section 2.3.5).

As can be seen from Figure 12, except for two sites, the measurements of the national soil monitoring network (NABO) were predicted well. The NABO sites were distributed evenly over the whole study area (cf. Figure 1), representing thereby a large part of the relevant covariate space. The predictions for the sites of the soil monitoring networks of St. Gallen and Zurich (KABO) spanned a narrower range than the respective measurements and the predictions for the NABO sites. The sites of the two networks mostly lie in Northeastern Switzerland in an area with similar pedogenetic conditions.

The validation statistics were calculated for the merged set of national and cantonal monitoring data (cf. Table 4). According to the BIAS statistics the systematic error was small. However, R^2 with values 0.22 and 0.17 was smaller than in the other validations, where a range between 0.3 to 0.4 had been found. This reflects probably the fact that soil sampling and analysis likely differed from the methods employed at WSL.

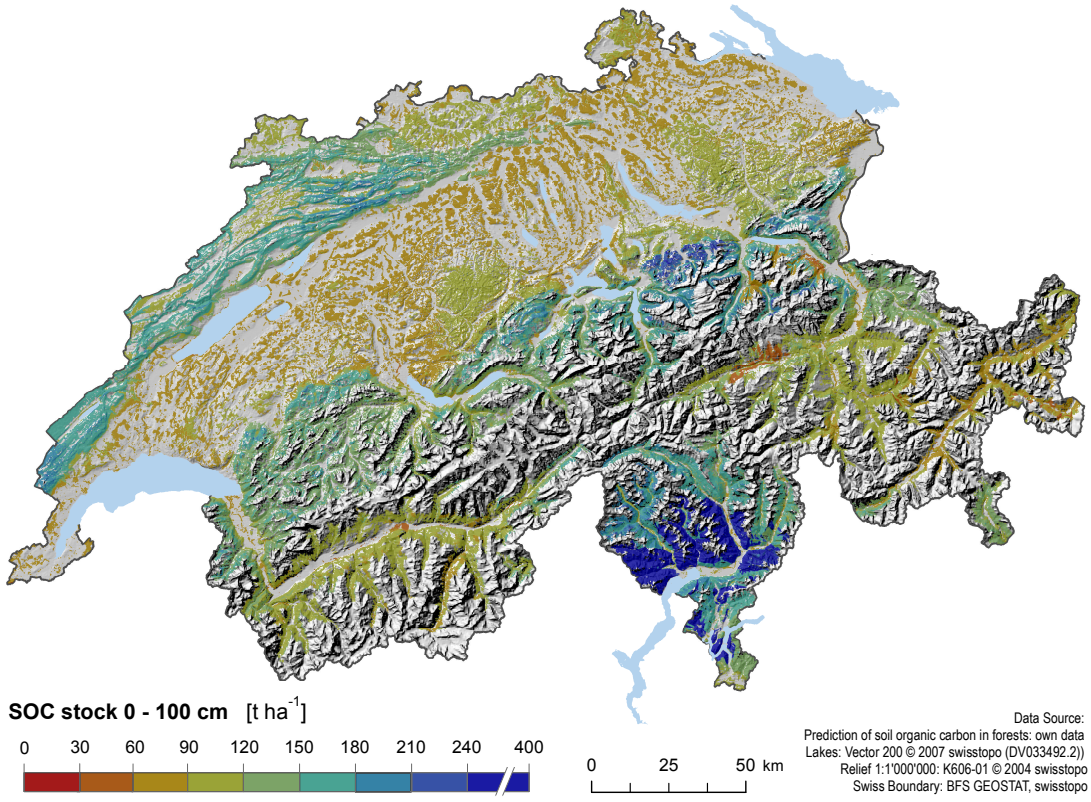

Figure 11: Robust lognormal kriging prediction of the SOC stock in 0–100 cm of the mineral soil of Swiss forests (computed with best-fit model with covariates according to section 3.3.1 and tuning constant $c = 2$, smoothed as Figure 7)

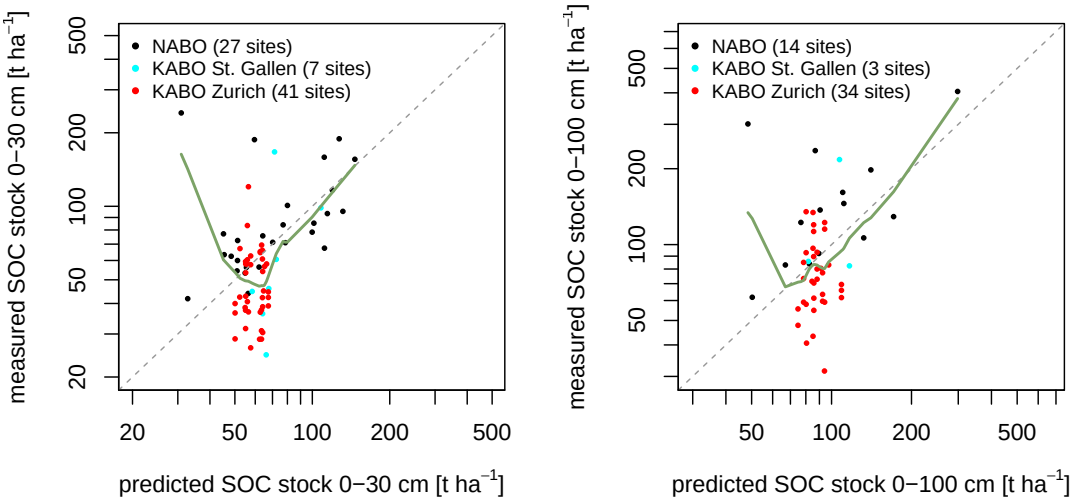

Figure 12: Measured SOC stocks in 0–30 cm and 0–100 cm depth at sites of three soil monitoring networks plotted against the kriging predictions computed from the models of the respective soil compartments (cf. sections 3.2.1 and 3.3.1, green lines: loess scatterplot smoothers).

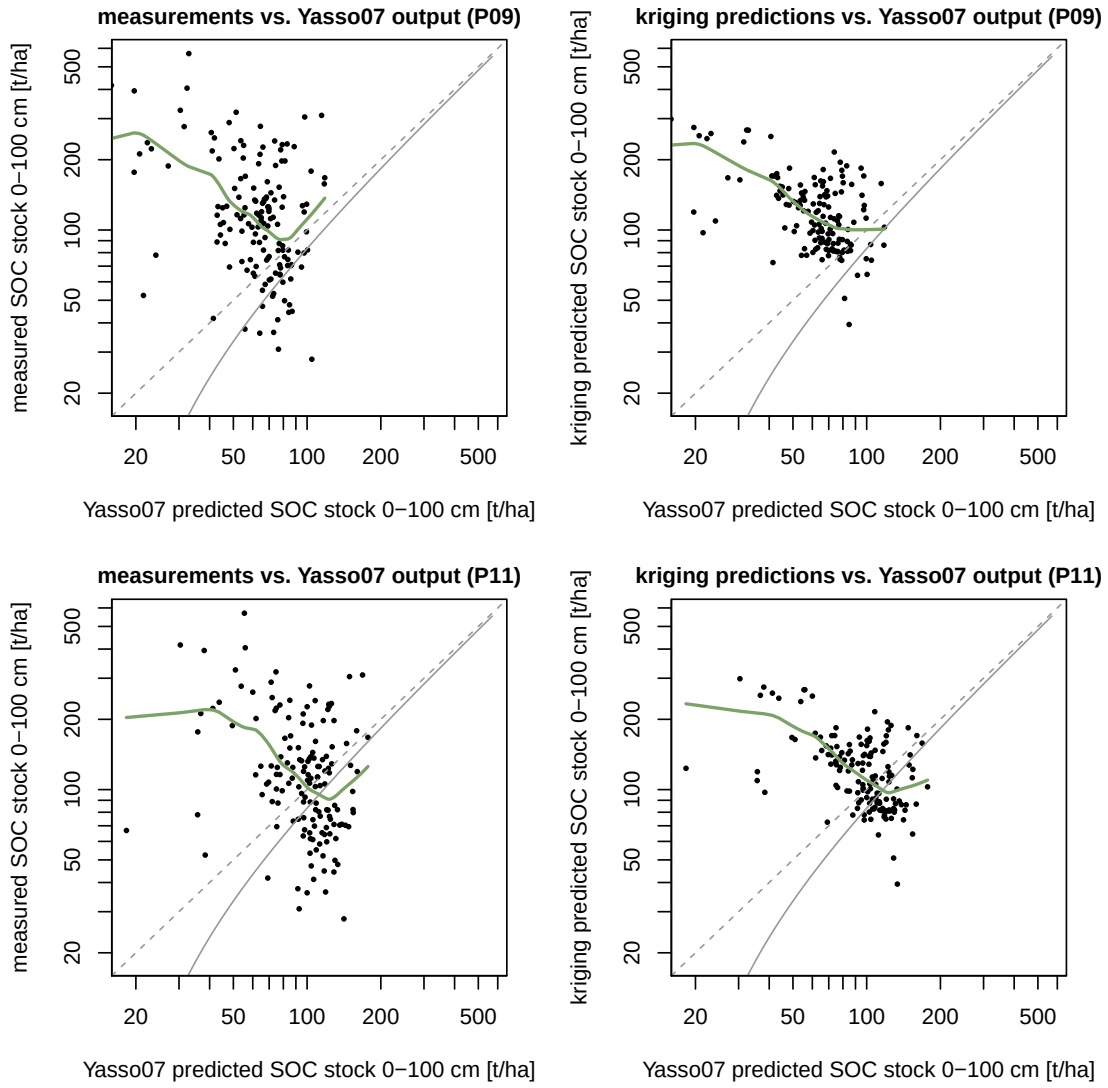

Figure 13: Comparison of Yasso07 estimates of OC stock (sum of OC in deadwood, litter and mineral soil 0–100 cm) with measured stocks (left panels) or kriging predictions (right panels), based on model described in section 3.3.1. Yasso07 estimates were computed either with the 2009 parameter set (top row, Tuomi *et al.*, 2009) or the 2011 set (bottom row, Tuomi *et al.*, 2011), cf. Weggler *et al.* (2012, secs 2.3.1 and 4.6) for details (green lines: loess scatterplot smoothers; grey solid lines: 1:1-lines adjusted for average forest floor OC stock in Switzerland [16.7 t ha^{-1}], cf. Table 3; grey dashed lines: unadjusted 1:1-line).

3.5 Comparison of kriging predictions with stock estimates by Yasso07 model

Yasso07 (Tuomi *et al.*, 2009, 2011) and its predecessor Yasso (Liski *et al.*, 2005) model the accumulation and degradation of OC in the soil. The models were developed for general application and therefore require only little input data and few parameters. Initial OC stocks are estimated by Yasso07 in a "spin-up" procedure based on long-term mean annual temperature and precipitation and C input into the soil (woody and non-woody litterfall). This is to ensure that the model is in equilibrium with climate, prior to simulating scenarios of changes in litterfall inputs and climate. Yasso07 outputs the combined OC stored in deadwood, litter and soil. Weggler *et al.* (2012) validated Yasso07 output of initial stocks and changes for conditions in Switzerland.

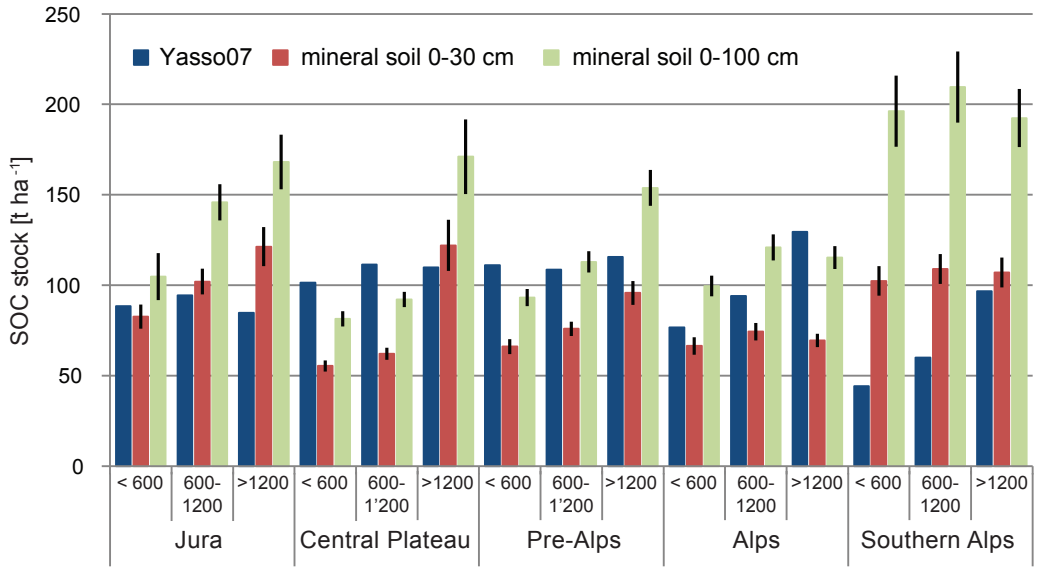

Figure 14: Comparison of initial OC stock estimates (including deadwood, forest floor and mineral soil 0–100 cm) obtained by the model Yasso07 with block kriging predictions (with standard errors) SOC stocks in mineral soil 0–30 cm and 0–100 cm. No error estimates were available for the Yasso07 estimates.

Figure 13 contrasts the Yasso07 stock estimates, computed with two different parameter sets, with the measured SOC stocks in 0–100 cm and our kriging predictions. Evidently, there is hardly any relation, neither to the measured stocks, nor to our predictions. Stock estimates computed with the 2009 parameter set strongly underestimated the measured stocks, in particular if one takes into account that the Yasso07 estimates include OC in the forest floor (and in deadwood), cf. solid grey lines in Figure 13. Stock estimates computed with the 2011 parameter set were marginally less biased, but the conditional bias remained severe, confirming findings by Weggler *et al.* (2012, sec. 4.6).

Figure 14 compares our block kriging OC stock predictions with the Yasso07 estimates by NFI production region stratified by altitude classes. The Yasso07 estimates were computed with the 2009 parameter set (Tuomi *et al.*, 2009). Not surprisingly, the Yasso07 estimates are again negatively biased, in particular for the Southern Alps. Furthermore, except for the Alps (and the Southern Alps), Yasso07 estimates did not clearly increase with altitude, unlike our block kriging predictions.

4 Discussion

SOC stocks in mineral soil For the SOC stocks in the mineral soil of Swiss forests, the model building procedure (cf. section 2.3.3) effectively reduced the large number of potential covariates and their first-order interactions to a small and meaningful set. Precipitation (mean yearly and mean March, respectively) remained — with positive regression coefficients (cf. Figures 21 and 23 in Appendix G) — in both models. The same result had been obtained before by Perruchoud *et al.* (2000). Martin *et al.* (2011) and Meersmans *et al.* (2012) similarly found that precipitation was an important covariate for modelling the spatial distribution of topsoil (0–30 cm) OC stocks across France: Wet and cool climates favours SOC accumulation. But as precipitation and temperature are quite strongly correlated by their dependence on elevation, one cannot easily separate the effects of the two covariates. Near infrared reflectance was also chosen for both models and thus appears to be an important covariate. This variable discriminates between conifers and deciduous trees, the former having smaller reflectance for wavelengths from 750 to 1300 nm (Cipar *et al.*, 2004). The negative regression coefficients (Figures 21 and 23) for this covariate thus imply that SOC stocks are larger under conifers than under deciduous trees, which is in agreement with common understanding of OC accumulation in forests. Furthermore, merged (and partly amended) units of the soil map were informative predictors for modelling mineral SOC stocks. They contain information about the parent material, which is evidently important for soil formation.

The residual spatial auto-correlation remained rather weak in both models. This suggests that the regional patterns in SOC stock data could be reasonably well modelled by the fitted regression models. The short-ranged spatial dependence has consequences for computing the kriging predictions: Only about 5 % of the nodes of the grid, for which kriging predictions were computed, lie within a distance of the effective ranges of the two variograms (600 and 660 m, respectively). The kriging predictions deviate only within these zones from predictions obtained by the fitted regression models alone[6]. Hence, for most parts of Switzerland, only the fitted regression models mattered for kriging, and nearby observations did not directly influence the predictions.

The evaluation of the predictive performance of the two models by cross-validation and with independent validation data showed that the predictions were neither conditionally nor marginally biased. This was not a priori clear because the profile locations of the calibration set had not been selected by a randomized design[7]. But random dispersion of the prediction errors remained large as the robustly estimated relative root mean squared errors with values between 39 % and 44 % demonstrate. This was also reflected in the rather small, robustly estimated R^2, which reached only values between 0.3 to 0.4. R^2-values of the validations with data of national and cantonal soil monitoring data were even smaller. This can likely be attributed to differences in soil sample collection and analytical procedures. Since there were no indications that our models over-fitted the data — the validation statistics matched the respective cross-validation statistics well — we can be quite confident that the evaluated statistics provide a fair picture of the predictive performance of the fitted models when predicting stocks in the mineral soil for the whole country.

Several causes are likely responsible for the moderate precision of the predictions:

- Bulk density had been merely measured for about 8 % of all mineral soil horizons, and for the vast majority of horizons only a field estimate of the density was available (cf.

[6]The elements of the covariance vector $\boldsymbol{\gamma}(\boldsymbol{s}_0)$ in equation 9 (Appendix B) are close to zero for distances between prediction $\boldsymbol{s}_0$ and data locations $\boldsymbol{s}_i$ larger than the effective range of the variogram. The second term of the right-hand side of equation 9 thus vanishes if all $\boldsymbol{s}_i$ are farther away from $\boldsymbol{s}_0$ than the effective range.

[7]The WSL validation data provide nevertheless spatially representative information about the predictive performance of the models because the majority of its profile locations were evenly distributed on a 8×8-km-grid across the whole country (cf. section 2.3.5).

section 2.2.1). Since measurement errors of SOC concentration hardly matter (Walthert *et al.*, 2010), the quality of the SOC stock data was mainly controlled by the quality of the bulk density estimates and the field estimates of volumetric gravel content. We believe that respective errors were the most important cause of the limited precision of the predictions.

- The soil samples had been collected over a longer period of time (Figure 2). Although we found no significant linear change of SOC stocks in time, it is nevertheless likely that this had added some extra-variation to the data.
- The uncertainty in the recorded coordinates of the soil profile locations (±25 m) was larger than the resolution (2 m) of the spatially best resolved input data. Although the values of these covariates have been averaged over circular windows with various radii, the errors in the recorded locations nevertheless weakened the predictive power of the covariates.
- Poor quality of original data from which the input data were derived (e.g. errors in digital elevation and terrain models, clouds covering forest canopy in the SPOT5 mosaic) resulted in distorted or missing values for some covariates (see Nussbaum, 2011, for more details).
- Covariates characterizing OC input into the soil (e.g. accurate and highly resolved information on tree species composition of forests) and factors relevant for OC turnover in soils (e.g. pH, soil texture) were largely lacking in the input data, because such information is currently not available contiguously for whole Switzerland. Similarly, only coarse information about parent material was available from the soil and geotechnical maps (both at scales of 1:200 000). We tried to extrapolate soil information, measured at the 1 033 soil profile sites, to whole Switzerland by assigning median values of abundance of particle size classes, pH and soil mass to the units of the soil or geotechnical maps. This had some success as both clay content (in an interaction with a topographic position index) and mass of soil particles < 2 mm were covariates of the best-fit model of topsoil mineral SOC stocks. However, it is quite clear that availability of covariates with spatially highly resolved, accurate information on parent material, C input into the soil and on factors controlling OC turnover would have resulted in models with far better predictive power.

Other studies found cross-validation statistics of similar magnitude: Martin *et al.* (2011) obtained by cross-validation a (non-robust) R^2 of 0.36 for predicting topsoil OC stocks of French forests. Schröder *et al.* (2009) reported a mean relative error of 35 %. As an aside: Perruchoud *et al.* (2000) obtained *goodness-of-fit* R^2-values of 0.14 for the top (0–20 cm) mineral soil stocks and 0.24 for the total stocks down to the bedrock, which puts our apparently small (cross-)*validation* R^2-values somewhat into perspective.

According to the cross-validation results, moderate robustification of the parameter estimation procedure (with tuning constant $c = 2$) increased the predictive power of the fitted models slightly in comparison with customary REML estimation and kriging. The residual diagnostic plots in Figures 20 and 22 (Appendix G) revealed that there were indeed a few observations in the dataset that did not follow the patterns of the majority of the data and therefore appeared as outliers in the residual plots. Using robust procedures then ensures that these observations do not upset statistical analyses.

OC stock in forest floor We did not trust the results of the geostatistical analysis of the forest floor OC stock data because the fitted model predicted the WSL validation data poorly (cf. sections 2.4 and F.2). We estimated the mean stocks for the NFI production regions and for the whole country therefore by design-based estimates, although the underlying assumption (simple random sample) was clearly not met. Beside Horvitz-Thompson estimators, calibrated estimators (with mean annual temperature and merged soil map units as auxilary variables, cf. Deville and Särndal, 1992) were used to estimate the mean stocks. However, calibrated estimation did not clearly reduce the standard errors of the estimates, and we reported here therefore only the

Horvitz-Thompson estimates. Since we had to use all the data for computing design-based estimates and could furthermore not compute point predictions, we could neither validate estimated stocks in the forest floor nor the quality of the estimated standard errors of the stocks.

Besides the reasons listed above, the following causes were likely decisive for the failure of our geostatistical analysis of the forest floor data:

- Parts of the forest floor (L horizon) are spatially and seasonally very variable (Moeri, 2007).
- Only the thickness of the L horizons had been measured for all soil profiles, along with OC concentrations of F and H horizons. OC concentrations of L horizons and bulk densities of all forest floor horizons were measured only for a small number of sites and respective medians were assigned to horizons with missing data. Hence, the data on forest floor stocks were likely subject to (considerable) errors that could however not be quantified.

Comparison with Yasso07 stock estimates We have shown in section 3.5 that the initial stock estimates obtained by Yasso07 for 149 out of 175 sites of the WSL validation set were, unlike our kriging predictions, both marginally and conditionally strongly negatively biased (Figure 13). The aggregated Yasso07 stock estimates differed also for the NFI production regions strongly from our block kriging predictions (Figure 14). Yasso07 was validated in more depth by Weggler *et al.* (2012) for Swiss conditions. While modelled litter decomposition (with the parameter set 2009) agreed well with measured data, measured SOC stocks were underestimated. Since Yasso07 models OC stocks solely based on OC inputs and climate without any information on soils, some deviation from the measured stock for single sites is to be expected. The expectation is that spatially aggregated stock estimates will be more accurate. However, this is doubtful because spatial averaging ("bulking") reduces only the variance (random variation) but not the bias (systematic error), and it is therefore unlikely that Yasso07's inital stocks estimates for the NFI production regions are much better than its predictions for single sites. The particularly large discrepancy for the Southern Alps (Figure 14) was likely due to two reasons:

- The probable accumulation of pyrogenic C could not be reproduced by Yasso07 without soil information.
- Yasso07 uses annual time steps and thus likely overestimated decomposition in this region, which is characterized by large annual precipitation, but considerable inter-annual variation of precipitation.

The tendency of increasing SOC stocks with increasing altitude (e.g. Hagedorn *et al.*, 2010) was more apparent in our block kriging predictions than in Yasso07's initial stock estimates. In Yasso07 this trend may not be reproduced to the same degree. The mean annual precipitation in Switzerland is in the upper range of the current Yasso07 model parametrization what probably reduces the effect of increased precipitation with altitude on the Yasso07 stock estimates. The "spin-up" procedure that was used to derive initial OC stocks by Yasso07 may be yet another reason for the discrepancy between its stock estimates and our block kriging predictions: The assumption that the initial stocks accumulated under current climate (mean of period 1961–1990) might be too coarse an approximation.

Comparison with previous SOC stock estimates for Switzerland To date, no maps of SOC stocks have been published for Switzerland that could be used for a comparison with our results. Nevertheless, several patterns in our mineral soil SOC stock maps (Figures 7 and 11) matched qualitatively our expectations as soil scientist:

- Small SOC stocks were predicted for lower altitude on the Central Plateau where acid soils prevail and in the dry inner Alpine valleys (Valais in particular).

- The large SOC stocks predicted for the Jura region are likely related to OC stabilisation by calcium (Walthert *et al.*, 2004).
- Very large SOC stocks in the Southern Alps, particularly in the subsoils, are likely caused by stabilisation through geogenetic elements (Blaser *et al.*, 2005, chap. 3).
- The block kriging predictions of the mean stocks per elevation-stratified NFI production region nicely reflected the gradient described by Hagedorn *et al.* (2010) of increasing SOC stocks with increasing altitude, in particular for the strata of ≤600 m and (600–1200] m asl (Figure 14).

For the total forest area of Switzerland, our design-based estimate of the mean forest floor OC stock was equal 16.7 t ha^{-1} (95 %-confidence interval [15.1, 18.3] t ha^{-1}, Table 3), and the block kriging predictions of the mean mineral soil SOC stocks amounted to 79.9 t ha^{-1} (95 %-prediction interval [76.9, 82.9] t ha^{-1}) for the top 30 cm of the solum and to 125.8 t ha^{-1} (95 %-prediction interval [121.1, 130.5] t ha^{-1}) for 0–100 cm depth (Table 5).

Based on Perruchoud *et al.* (2000) and Moeri (2007), the Swiss GGI used so far SOC stock estimates of 17.5 t ha^{-1} for the forest floor, 75.9 t ha^{-1} for the the top 30 cm of the mineral soil and 98.2 t ha^{-1} for the mineral soil down to 100 cm depth. Unlike the previous stock estimate for the forest floor, which falls into our confidence interval, the previous estimates for mineral soil stocks are significantly smaller than the values predicted in this study (p-values of one-sided z-tests: 0.004 and $< 10^{-12}$, respectively). According to Perruchoud *et al.* (2000), about 77 % of the SOC stock is found in the mineral topsoil (0–30 cm), whereas we predicted here a proportion of only 64 %, which matches the 64.3 % computed directly from the SOC data (n = 1 030) very well. We have currently no explanation for these discrepancies, except that our analysis relies on a substantially larger database than was available to Perruchoud *et al.*. As we have good reasons to trust our stock estimates for the mineral soil, we conclude that SOC stocks in the topsoil had been sligtly and those in the subsoil strongly underestimated in the past.

Comparison with SOC stock estimates for other European countries Several studies estimated SOC stocks in forest soils of other European countries for the same depth compartments as used in our study: For the forest floor, estimates are available for Austria 15 t ha^{-1} (Weiss *et al.*, 2000), Lower Saxony 25 t ha^{-1} (Wördehoff *et al.*, 2011), Schleswig-Holstein 35 t ha^{-1} (Wördehoff *et al.*, 2012), 19 t ha^{-1} (Schröder *et al.*, 2009) and Germany as a whole 19.8 ± 0.6 t ha^{-1} (Oehmichen *et al.*, 2011). Meersmans *et al.* (2012) estimated for the mineral topsoil (0–30 cm) of French forests a SOC stock of 94–95 t ha^{-1} depending on the used model and Lettens *et al.* (2005) estimated a value range of 87–93 t ha^{-1} (standard errors 4–14 t ha^{-1}) for different forest types. Much larger stocks of 163 t ha^{-1} were estimated for the same depth compartments of Irish forests (Xu *et al.*, 2011), but also much smaller stocks were found for German forests of 59 and 68 t ha^{-1} for two sampling periods (Oehmichen *et al.*, 2011).

The estimates of SOC stocks in 0–100 cm of the the mineral soil amounted for Danish forests to 169 t ha^{-1} (95 %-confidence interval 148–188 t ha^{-1}, Krogh *et al.*, 2003) and for Belgian forest to 148–155 t ha^{-1} (standard errors 12–26 t ha^{-1}, Lettens *et al.*, 2005) and for Bavarian forest to 98 t ha^{-1} (Wiesmeier *et al.*, 2012). For the stocks stored in the entire forest soil profiles, Wördehoff *et al.* (2011) reported for Lower Saxony estimates of 72 and 102 t ha^{-1} for deciduous and coniferous stands, respectively, and Schröder *et al.* (2009) estimated for North Rhine-Westphalia total stocks of 91 t ha^{-1}.

Some or these estimates are in good agreement with our predictions, other differ quite strongly from our figures. It is very difficult to account for the observed discrepancies: The number of soil profiles, from which data were available, varied strongly between the studies, along with type of available soil data (SOC content, bulk density, gravel content). Furthermore diverse estimation procedures were employed, ranging from a simple arithmetic mean to spatial regression analyses, comparable to the ones used in our study (cf. section 1.3).

It is rather remarkable that only two studies made efforts to validate their SOC stock estimation procedures, which should be a "must" in the context of greenhouse gas reporting, as it is very difficult to assess the reliability of reported estimates without such information. Only Martin *et al.* (2011) and Schröder *et al.* (2009) (cross-)validated their stock predictions, but none of the studies validated predictions with independent validation data. Our study seems to be quite unique in its attempts to thoroughly validate both the levels of the predicted SOC stocks and the modelled prediction uncertainty. Our analysis of the forest floor data exemplifies that this is essential for qualifying the results, even if this eventually leads to complete rejection of a modelling approach.

Acknowledgements

The authors would like to thank Hans Rudolf Künsch for advice on the approximation of the block kriging variances and Nele Rogiers, Andreas Schellenberg and Markus Didion for their valuable comments on an earlier version of this report. Furthermore we would like to thank Daniela Marugg, Ubald Gasser and Armin Keller for readily providing soil data for the validation of the predictions. This study was funded by the Federal Office for the Environment FOEN.

References

AFU (2012). Cantonal monitoring network St. Gallen. Database extract 19 June 2012 done by Daniela Marugg.

BAFU (2011). Luftbelastung: Karten Jahreswerte. Ammoniak und Stickstoffdeposition, Jahresmittel 2007 (modelliert durch METEOTEST). http://www.bafu.admin.ch/luft/luftbelastung/schadstoffkarten.

BFS (2001). *GEOSTAT Benützerhandbuch*. Bundesamt für Statistik, Bern.

Blaser, P., Zimmermann, S., Luster, J., and Walthert, L., L. P. (2005). *Waldböden der Schweiz. Band 2. Regionen Alpen und Alpensüdseite*. Eidg. Forschungsanstalt WSL and Hep Verlag, Birmensdorf and Bern.

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society Series B*, **26**, 211–243.

Brassel, P. and Brändli, U.-B., editors (1999). *Schweizerisches Landesforstinventar. Ergebnisse der Zweitaufnahme 1993-1995*. Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft WSL, Birmensdorf, Bundesamt für Umwelt, Wald und Landschaft BUWAL, Bern, Birmensdorf, Bern.

Brassel, P. and Lischke, H., editors (2001). *Swiss National Forest Inventory: Methods and Models of the Second Assessment*. Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf.

Cipar, J., Cooley, T., Lockwood, R., and Grigsby, P. (2004). Distinguishing between coniferous and deciduous forests using hyperspectral imagery. In *Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International*, volume 4, pages 2348–2351.

Cleveland, W. S., Grosse, E., and Shyu, W. M. (1992). Local regression models. In J. M. Chambers and T. J. Hastie, editors, *Statistical Models in S*, chapter 8. Wadsworht & Brooks/Cole.

Cressie, N. (2006). Block kriging for lognormal spatial processes. *Mathematical Geology*, **38**(4), 413–443.

Croux, C. and Dehon, C. (2003). Estimators of the multiple correlation coefficient: Local robustness and confidence intervals. *Statistical Papers*, **44**, 315–334.

de Gruijter, J. J., Brus, D. J., Bierkens, M. F. P., and Knotters, M. (2006). *Sampling for Natural Resouces Monitoring*. Springer-Verlag, Berlin.

Desaules, A., Schwab, P., Keller, A., Ammann, S., Paul, J., and Bachmann, H. J. (2006). Anorganische Schadstoffgehalte in Böden der Schweiz und Veränderungen nach 10 Jahren. Ergebnisse der Nationalen Bodenbeobachtung 1985–1999. Bericht, Agroscope FAL Reckenholz, Eidgenössische Forschungsanstalt für Agrarökologie und Landbau, Zürich Reckenholz.

Deville, J. C. and Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, **87**, 376–382.

Faraway, J. J. (2005). *Linear Models with R*, volume 63 of *Texts in Statistical Science*. Chapman & Hall/CRC, Boca Raton.

FOEN (2012). Switzerland's greenhouse gas inventory 1990–2010. National Inventory Report 2012, Submission of 13 April 2012 under the United Nations Framework Convention on Climate Change and under the Kyoto Protocol, Federal Office for the Environment FOEN, Climate Division, Bern.

Gabriel, K. R. (1981). Biplot display of multivariate matrices for inspection of data and diagnostics. In V. Barnett, editor, *Interpreting Multivariate Data*, Chichester. John Wiley & Sons. Proceedings of the Conference Entitled "Looking at Multivariate Data"; Sheffield, 24-27 March 1980.

Gasser, U., Stutz, H., and Bouquet, F. (2009). Bodenversauerung und Bodenfruchtbarkeit im Zürcher Wald. *Zürcher Wald*, **2009/1**, 4–9.

Gneiting, T., Balabdaoui, F., and Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society Series B*, **69**(2), 243–268.

Gonseth, Y., Wohlgemuth, T., Sansonnens, B., and Buttler, A. (2001). Die biogeographischen Regionen der Schweiz. Erläuterungen und Einteilungsstandard. *Umwelt-Materialien Nr. 137, BUWAL, Bundesamt für Umwelt, Wald und Landschaft*.

Gurtz, J., Baltensweiler, A., and Lang, H. (1999). Spatially distributed hydrotope based modelling of evapotranspiration and runoff in mountainous basins. *Hydrological Processes*, **13**(17), 2751–2768.

Hagedorn, F., Moeri, A., Walthert, L., and Zimmermann, S. (2010). Kohlenstoff in Schweizer Waldböden - bei Klimaerwärmung eine potenzielle CO_2-Quelle – Soil organic carbon in Swiss forest soils - a potential CO_2 source in a warming climate. *Schweizerische Zeitschrift für Forstwesen*, **161**, 530–535.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning; Data Mining, Inference and Prediction*. Springer, New York, second edition.

Jenness, J. (2006). Topographic position index (TPI) v. 1.2. http://www. jennessent.com.

Krogh, L., Noergaard, A., Hermansen, M., Greve, M. H., Balstroem, T., and Breuning-Madsen, H. (2003). Preliminary estimates of contemporary soil organic carbon stocks in Denmark using multiple datasets and four scaling-up methods. *Agriculture, Ecosystems & Environment*, **96**(1-3), 19–28.

Künsch, H. R., Papritz, A., Schwierz, C., and Stahel, W. A. (2011). Robust estimation of the external drift and the variogram of spatial data. In *Proceedings of the ISI 58th World Statistics Congress of the International Statistical Institute*, volume http://isi2011.congressplanner.eu/pdfs/650268.pdf, Dublin.

Künsch, H. R., Papritz, A., Schwierz, C., and Stahel, W. A. (in preparation). Robust geostatistics.

Lettens, S., Van Orshoven, J., van Wesemael, B., De Vos, B., and Muys, B. (2005). Stocks and fluxes of soil organic carbon for landscape units in Belgium derived from heterogeneous data sets for 1990 and 2000. *Geoderma*, **127**(1-2), 11–23.

Liski, J., Palosuo, T., Peltoniemi, M., and Sievänen, R. (2005). Carbon and decomposition model Yasso for forest soils. *Ecological Modelling*, **189**(1-2), 168–182.

Martin, M. P., Wattenbach, M., Smith, P., Meersmans, J., Jolivet, C., Boulonne, L., and Arrouays, D. (2011). Spatial distribution of soil organic carbon stocks in France. *Biogeosciences*, **8**(5), 1053–1065.

Meersmans, J., De Ridder, F., Canters, F., De Baets, S., and Van Molle, M. (2008). A multiple regression approach to assess the spatial distribution of soil organic carbon (SOC) at the regional scale (Flanders, Belgium). *Geoderma*, **143**(1-2), 1–13.

Meersmans, J., Martin, M. P., Lacarce, E., De Baets, S., Jolivet, C., Boulonne, L., Lehmann, S., Saby, N. P. A., Bispo, A., and Arrouays, D. (2012). A high resolution map of French soil organic carbon. *Agronomy for Sustainable Development*, **32**(4), 841–851.

MeteoSwiss (2011). Mean monthly and yearly mean norm values of precipitation, temperature and relative sunshine duration (1961-1990). http://www.meteoschweiz.admin.ch/web/en/services/ data_portal/grided_datasets.html.

Moeri, A. C. (2007). Kohlenstoffvorräte in Schweizer Waldböden mit besonderer Berücksichtigung der organischen Auflage. Diplomarbeit, Geographisches Institut der Universität Zürich.

Nussbaum, M. (2011). *Modellierung des organischen Kohlenstoffgehalt und -vorrats in Schweizer Waldböden*. Masterarbeit (korrigierte Version vom 5. Dezember 2011), Universität Zürich, Geographisches Institut, Zürich.

Oehmichen, K., Demant, B., Dunger, K., Grüneberg, E., Hennig, P., Kroiher, F., Neubauer, M., Polley, H., Riedel, T., Rock, J., Schwitzgebel, F., Stümer, W., Wellbrock, N., Ziche, D., and Bolte, A. (2011). Inventurstudie 2008 und Treibhausgasinventar Wald. *Johann Heinrich von Thünen-Institut (vTI), Landbauforschung, Sonderheft*, **343**.

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proc. Roy. Soc. London A.*, **193**, 120–145.

Perruchoud, D., Walthert, L., Zimmermann, S., and Lüscher, P. (2000). Contemporay carbon stocks of mineral forest soils in the Swiss Alps. *Biogeochemistry*, **50**, 111–136.

Remund, J., Frehner, M., Walthert, L., Kägi, M., and Rihm, B. (2011). Schätzung standortspezifischer Trockenstressrisiken in Schweizer Wäldern.

Richardson, A. M. and Welsh, A. H. (1995). Robust restricted maximum likelihood in mixed linear models. *Biometrics*, **51**, 1429–1439.

Särndal, C.-E., Swensson, B., and Wretman, J. (1992). *Model Assisted Survey Sampling*. Springer, New York.

Schröder, W., Pesch, R., and Schmidt, G. (2009). Grossräumige Regionalisierung der Kohlenstoffbindung in Wäldern Nordrhein-Westfalens. *Umweltwissenschaften und Schadstoffforschung*, **21**(6), 516–526.

Swisstopo (2011). VECTOR25. http://www.swisstopo.admin.ch/internet/swisstopo/de/home/products/landscape/vector25.html.

Swisstopo (2011a). Geologische Karte der Schweiz 1:500000. http://www.swisstopo.admin.ch/internet/swisstopo/de/home/products/maps/ geology/geomaps/gm500.html.

Swisstopo (2011b). Höhenmodelle. http://www.swisstopo.admin.ch/internet/swisstopo/de/home/products/height.html.

Swisstopo and NPOC (2009). SPOT5 Radcor Mosaic of Switzerland.

Tarboton, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, **33**(2), PP. 309–319.

Tomlinson, R. and Milne, R. (2006). Soil carbon stocks and land cover in Northern Ireland from 1939 to 2000. *Applied Geography*, **26**(1), 18–39.

Tuomi, M., Thum, T., Järvinen, H., Fronzek, S., Berg, B., Harmon, M., Trofymow, J. A., Sevanto, S., and Liski, J. (2009). Leaf litter decomposition — estimates of global variability based on Yasso07 model. *Ecological Modelling*, **220**, 3362–3371.

Tuomi, M., Rasinmäki, J., Repo, A., Vanhala, P., and Liski, J. (2011). Soil carbon model Yasso07 graphical user interface. *Environmental Modeling and Software*, **26**(11), 1358–1362.

Walthert, L., Zimmermann, S., Blaser, P., Luster, J., and Lüscher, P. (2004). *Waldböden der Schweiz. Band 1. Grundlagen und Region Jura*. Eidg. Forschungsanstalt WSL and Hep Verlag, Birmensdorf and Bern.

Walthert, L., Graf, U., Kammer, A., Luster, J., Pezzotta, D., Zimmermann, S., and Hagedorn, F. (2010). Determination of organic and inorganic carbon, δ13c, and nitrogen in soils containing darbonates after acid fumigation with HCl. *Journal of Plant Nutrition and Soil Science*, **173**(2), 207–216.

Weggler, K., Steinmann, K., Kaufmann, E., and Thürig, E. (2012). Modeling soil carbon stock including dead wood and l, f, h-horizon using the models Massimo and Yasso07. Final report bafu contract-nr. 04.1140.pj / k283-0072, Federal Institute for Forest, Snow and Landscape Research WSL.

Weiss, A. K. (2000). *Anthropogenic and Dynamic Contributions to Ozone Trends in the Swiss Total Ozone, Umkehr and Balloon Sounding Series*. Ph.D. thesis, Swiss Federal Institute of Technolgy Zurich. Dissertation ETH No. 13635.

Weiss, P., Schieler, K., Schadauer, K., Radunsky, K., and Englisch, M. (2000). *Die Kohlenstoffbilanz des Österreichischen Waldes und Betrachtungen zum Kyoto-Protokoll*, volume 106. Umweltbundesamt.

Wiesmeier, M., Spörlein, P., Geuss, U., Hangen, E., Haug, S., Reischl, A., Schilling, B., Lützow, M., and Kögel-Knabner, I. (2012). Soil organic carbon stocks in southeast Germany (Bavaria) as affected by land use, soil type and sampling depth. *Global Change Biology*, **18**(7), 2233–2245.

Wohlgemuth, T., Kuhn, N., Lüscher, P., Kull, P., and Wüthrich, H. (1995). Vegetations- und Bodendynamik auf rezenten Windwurfflächen in den Schweizer Nordalpen. *Schweizer Zeitschrift des Forstwesens*, **146**(11), 873–891.

Wördehoff, R., Spellmann, H., Evers, J., and J., N. (2011). Kohlenstoffstudie Forst und Holz Niedersachsen. Beiträge aus der Nordwestdeutschen Forstlichen Versuchsanstalt 6, Nordwestdeutsche Forstliche Versuchsanstalt (NW-FVA), Göttingen.

Wördehoff, R., Spellmann, H., Evers, J., Cihan, T. A., and J., N. (2012). Kohlenstoffstudie Forst und Holz Schleswig-Holstein. Bericht, Nordwestdeutsche Forstliche Versuchsanstalt (NW-FVA), Göttingen.

WSL (2011). Swiss National Forest Inventory NFI. Database extract 1 December 2011 done by Andri Baltensweiler.

Xu, X., Liu, W., Zhang, C., and Kiely, G. (2011). Estimation of soil organic carbon stock and its spatial distribution in the Republic of Ireland. *Soil Use and Management*, **27**(2), 156–162.

Zimmermann, N. E. (2000). Calculation of topographic position. http://www.wsl.ch/staff/niklaus.zimmermann/programs/aml4_1.html.

Zimmermann, N. E. and Kienast, F. (1999). Predictive mapping of alpine grasslands in Switzerland: Species versus community approach. *Journal of Vegetation Science*, **10**(4), 469–482.

Appendices

A List of covariates (digital)

The digital appendix includes the list of covariates (in German) with data description (resolution, entity or categories, source) as Excel and PDF file with the following file names:

liste_der_kovariablen.xlsx

liste_der_kovariablen.pdf

B Robust kriging predictions

Point kriging We predicted the log-transformed SOC stocks, say $Y_c(\boldsymbol{s}) = \log(S_c(\boldsymbol{s}))$, at the nodes of a grid with spacing equal to 100 m and for the soil profile locations of the validation sets by robust kriging. Based on the set of covariates (known $\boldsymbol{x}(\boldsymbol{s}_0)$) and the estimated parameters $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\theta}}^{\mathrm{T}} = (\hat{\sigma}_0^2, \hat{\sigma}_1^2, \hat{\alpha})$ of the final model, the robust kriging predictions were computed for a location $\boldsymbol{s}_0$ without measurement by

$$\hat{Y}_c(\boldsymbol{s}_0) = \boldsymbol{x}(\boldsymbol{s}_0)^{\mathrm{T}}\hat{\boldsymbol{\beta}} + \boldsymbol{\gamma}(\boldsymbol{s}_0)^{\mathrm{T}}\,\boldsymbol{\Gamma}^{-1}\,\hat{\boldsymbol{Z}}, \tag{9}$$

where $\hat{\boldsymbol{Z}}^{\mathrm{T}}$ is the vector with the robust kriging predictions of $\boldsymbol{Z}^{\mathrm{T}} = (Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2), \ldots, Z(\boldsymbol{s}_n))$ obtained by the robust REML algorithm for the n soil profile locations; $\boldsymbol{\gamma}(\boldsymbol{s}_0)^{\mathrm{T}} = (\gamma(\boldsymbol{s}_0 - \boldsymbol{s}_1; \hat{\boldsymbol{\theta}}), \gamma(\boldsymbol{s}_0 - \boldsymbol{s}_2; \hat{\boldsymbol{\theta}}), \ldots, \gamma(\boldsymbol{s}_0 - \boldsymbol{s}_n; \hat{\boldsymbol{\theta}}))$ is the vector with the estimated covariances between $\boldsymbol{Z}^{\mathrm{T}}$ and $Z(\boldsymbol{s}_0)$; and $\boldsymbol{\Gamma}$ is the $n \times n$-matrix of the covariances between all pairs $Z(\boldsymbol{s}_i)$ and $Z(\boldsymbol{s}_j)$, i.e. the (i, j)-th element of $\boldsymbol{\Gamma}$ is equal to $\gamma(\boldsymbol{s}_i - \boldsymbol{s}_j; \hat{\boldsymbol{\theta}})$. Since outliers receive small weight when computing $\hat{\boldsymbol{Z}}$ by the robust REML algorithm, the prediction at $\boldsymbol{s}_0$ by equation 9 is also insensitive to outlying observations.

The variance of the prediction errors of the log-transformed stocks, i.e. the robust kriging variances, were computed by

$$\begin{aligned} \sigma_K^2(\boldsymbol{s}_0) \;=\; & \mathrm{Var}[Y_c(\boldsymbol{s}_0) - \hat{Y}_c(\boldsymbol{s}_0)] = \hat{\sigma}_0^2 + \hat{\sigma}_1^2 - \boldsymbol{\gamma}(\boldsymbol{s}_0)^{\mathrm{T}}\,\boldsymbol{\Gamma}^{-1}\,\boldsymbol{\gamma}(\boldsymbol{s}_0) + \\ & \left(\boldsymbol{\gamma}(\boldsymbol{s}_0)^{\mathrm{T}}\,\boldsymbol{\Gamma}^{-1}, \boldsymbol{x}(\boldsymbol{s}_0)^{\mathrm{T}}\right)\,\mathrm{Cov}\left[\begin{pmatrix} \boldsymbol{Z} - \hat{\boldsymbol{Z}} \\ -\hat{\boldsymbol{\beta}} \end{pmatrix}, \left((\boldsymbol{Z} - \hat{\boldsymbol{Z}})^{\mathrm{T}}, -\hat{\boldsymbol{\beta}}^{\mathrm{T}}\right)\right]\begin{pmatrix} \boldsymbol{\Gamma}^{-1}\,\boldsymbol{\gamma}(\boldsymbol{s}_0) \\ \boldsymbol{x}(\boldsymbol{s}_0) \end{pmatrix}. \end{aligned} \tag{10}$$

The robust REML algorithm provides an approximation of the joint covariance matrix of $\hat{\boldsymbol{Z}}$ and $\hat{\boldsymbol{\beta}}$ (cf. Künsch *et al.*, 2011, eq. 19). From this we can compute the joint covariance matrix of the n prediction errors at the soil profile locations, $\boldsymbol{Z} - \hat{\boldsymbol{Z}}$, and of $-\hat{\boldsymbol{\beta}}$, required to evaluate the above expression, by

$$\begin{aligned} \mathrm{Cov}\left[\begin{pmatrix} \boldsymbol{Z} - \hat{\boldsymbol{Z}} \\ -\hat{\boldsymbol{\beta}} \end{pmatrix}, \left((\boldsymbol{Z} - \hat{\boldsymbol{Z}})^{\mathrm{T}}, -\hat{\boldsymbol{\beta}}^{\mathrm{T}}\right)\right] = \mathrm{Cov}\left[\begin{pmatrix} \hat{\boldsymbol{Z}} \\ \hat{\boldsymbol{\beta}} \end{pmatrix}, \left(\hat{\boldsymbol{Z}}^{\mathrm{T}}, \hat{\boldsymbol{\beta}}^{\mathrm{T}}\right)\right] + \\ \begin{bmatrix} \boldsymbol{\Gamma} & \boldsymbol{0}_{n\times p} \\ \boldsymbol{0}_{p\times n} & \boldsymbol{0}_{p\times p} \end{bmatrix} - b\,\boldsymbol{M}^{-1}\begin{bmatrix} \boldsymbol{\Gamma} & \boldsymbol{0}_{n\times p} \\ \boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Gamma} & \boldsymbol{0}_{p\times p} \end{bmatrix} - b\begin{bmatrix} \boldsymbol{\Gamma} & \boldsymbol{\Gamma}\boldsymbol{X} \\ \boldsymbol{0}_{p\times n} & \boldsymbol{0}_{p\times p} \end{bmatrix}\boldsymbol{M}^{-1}, \end{aligned} \tag{11}$$

where b, $\boldsymbol{X}$, $\boldsymbol{M}$ are as in Künsch *et al.* (2011), p is the number of regression coefficients in $\boldsymbol{\beta}$, and $\boldsymbol{0}_{r\times c}$ is a matrix with r rows and c columns and all entries equal to zero.

To transform the predictions of the log-transformed stocks back to the original scale of the measurements, we used the standard unbiased back-transformation for lognormal kriging (e.g. Cressie, 2006)

$$\hat{S}_c(\boldsymbol{s}_0) = \exp\left(\hat{Y}_c(\boldsymbol{s}_0) + 1/2\,(\hat{\sigma}_0^2 + \hat{\sigma}_1^2 - \mathrm{Var}[\hat{Y}_c(\boldsymbol{s}_0)])\right). \tag{12}$$

The variance of $\hat{Y}_c(\boldsymbol{s}_0)$ was computed by equation 15 given in Appendix C (one has to set $\boldsymbol{s}_i = \boldsymbol{s}_j = \boldsymbol{s}_0$ for this).

Block kriging The mean SOC stocks in the NFI production regions (and for whole Switzerland), stratified by altitude (cf. section 1), were computed from the lognormally back-transformed robust point predictions at the nodes of the 1-ha grid by block kriging. In more detail, we approximated the mean stock, say $S_c(B_k)$, of region B_k by

$$\hat{S}_c(B_k) = 1/N_k \sum_{\boldsymbol{s}_i \in B_k} \hat{S}_c(\boldsymbol{s}_i), \tag{13}$$

where the notation $\sum_{\boldsymbol{s}_i \in B_k}$ means summation over the N_k nodes of the 1-ha grid that fall into region B_k. The variance of the prediction error, $S_c(B_k) - \hat{S}_c(B_k)$, is then equal to

$$\mathrm{Var}[S_c(B_k) - \hat{S}_c(B_k)] = \frac{1}{N_k^2} \sum_{\boldsymbol{s}_i \in B_k} \sum_{\boldsymbol{s}_j \in B_k} \mathrm{Cov}[S_c(\boldsymbol{s}_i) - \hat{S}_c(\boldsymbol{s}_i), S_c(\boldsymbol{s}_j) - \hat{S}_c(\boldsymbol{s}_j)], \tag{14}$$

where the covariances of the prediction errors at locations $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ are given by equation 17 in Appendix D. However, N_k was too large (order of magnitude 10^4–10^5) to evaluate the double sum in equation 14 in acceptable computing time. We used therefore the Monte-Carlo approximation, detailed in Appendix E, which approximated $\mathrm{Var}[S_c(B_k) - \hat{S}_c(B_k)]$ well.

Due to the central limit theorem, the distribution of the prediction errors of the mean stocks per NFI production region or for the whole country can be approximated by a normal distribution, in spite of the fact that the prediction errors for the nodes of the 1-ha grid followed lognormal laws.

C Covariances of robust kriging predictions and covariances of robust predictions with observations

The covariance of the robust kriging predictions of the log-transformed SOC stocks at two unsampled locations $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ is given by

$$\mathrm{Cov}[\hat{Y}_c(\boldsymbol{s}_i), \hat{Y}_c(\boldsymbol{s}_j)] = \left(\boldsymbol{\gamma}(\boldsymbol{s}_i)^T \boldsymbol{\Gamma}^{-1},\, \boldsymbol{x}(\boldsymbol{s}_i)^T\right) \mathrm{Cov}\left[\begin{pmatrix} \hat{\boldsymbol{Z}} \\ \hat{\boldsymbol{\beta}} \end{pmatrix}, \left(\hat{\boldsymbol{Z}}^T, \hat{\boldsymbol{\beta}}^T\right)\right] \begin{pmatrix} \boldsymbol{\Gamma}^{-1}\boldsymbol{\gamma}(\boldsymbol{s}_j) \\ \boldsymbol{x}(\boldsymbol{s}_j) \end{pmatrix}, \tag{15}$$

and the covariance of $\hat{Y}_c(\boldsymbol{s}_i)$ with the true log-transformed stock $Y_c(\boldsymbol{s}_j)$ at location $\boldsymbol{s}_j$ is equal to

$$\begin{aligned} \mathrm{Cov}[\hat{Y}_c(\boldsymbol{s}_i), Y_c(\boldsymbol{s}_j)] &= \left(\boldsymbol{\gamma}(\boldsymbol{s}_i)^T \boldsymbol{\Gamma}^{-1},\, \boldsymbol{x}(\boldsymbol{s}_i)^T\right) \mathrm{Cov}\left[\begin{pmatrix} \hat{\boldsymbol{Z}} \\ \hat{\boldsymbol{\beta}} \end{pmatrix}, Y_c(\boldsymbol{s}_j)\right] \\ &= b\left(\boldsymbol{\gamma}(\boldsymbol{s}_i)^T \boldsymbol{\Gamma}^{-1},\, \boldsymbol{x}(\boldsymbol{s}_i)^T\right) \boldsymbol{M}^{-1} \begin{pmatrix} \boldsymbol{\gamma}(\boldsymbol{s}_j) \\ \boldsymbol{X}^T\boldsymbol{\gamma}(\boldsymbol{s}_j) \end{pmatrix}, \end{aligned} \tag{16}$$

where b, $\boldsymbol{X}$, $\boldsymbol{M}$ are again as in Künsch *et al.* (2011). The robust REML algorithm approximates the joint covariance matrix of $(\hat{\boldsymbol{Z}}^T, \hat{\boldsymbol{\beta}}^T)$ (cf. Künsch *et al.*, 2011, eq. 19) for the evaluation of equation 15.

D Covariances of lognormal kriging prediction errors

The covariance between the lognormally back-transformed prediction errors at two locations $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ is equal to

$$\begin{aligned} \mathrm{Cov}[\mathrm{S_c}(\boldsymbol{s}_\mathrm{i}) - \hat{\mathrm{S}}_c(\boldsymbol{s}_\mathrm{i}), \mathrm{S_c}(\boldsymbol{s}_\mathrm{j}) - \hat{\mathrm{S}}_c(\boldsymbol{s}_\mathrm{j})] = \mu(\boldsymbol{s}_i)\,\mu(\boldsymbol{s}_j) \Big\{ &\exp\left(\mathrm{Cov}[Y_c(\boldsymbol{s}_i), Y_c(\boldsymbol{s}_j)]\right) \\ &- \exp\left(\mathrm{Cov}[Y_c(\boldsymbol{s}_i), \hat{Y}_c(\boldsymbol{s}_j)]\right) - \exp\left(\mathrm{Cov}[\hat{Y}_c(\boldsymbol{s}_i), Y_c(\boldsymbol{s}_j)]\right) \\ &+ \exp\left(\mathrm{Cov}[\hat{Y}_c(\boldsymbol{s}_i), \hat{Y}_c(\boldsymbol{s}_j)]\right)\Big\} \end{aligned} \tag{17}$$

with $\mu(\boldsymbol{s}_i)$ (and $\mu(\boldsymbol{s}_j)$ analogously) approximated by

$$\mu(\boldsymbol{s}_i) \approx \exp\left(\boldsymbol{x}(\boldsymbol{s}_i)^{\mathrm{T}}\hat{\boldsymbol{\beta}} + 1/2\,(\hat{\sigma}_0^2 + \hat{\sigma}_1^2)\right).$$

$\mathrm{Cov}[Y_{\mathrm{c}}(\boldsymbol{s}_i), Y_{\mathrm{c}}(\boldsymbol{s}_j)]$ can be computed from the estimated variogram, exploiting the well-known relation between a weakly stationary variogram and an auto-covariance function, and the other covariance terms are given by equations 15 and 16 in Appendix C.

E Approximation of the prediction errors of the mean stocks per NFI production region and for whole Switzerland

To approximate $\mathrm{Var}[S_{\mathrm{c}}(B_k) - \hat{S}_{\mathrm{c}}(B_k)]$ (equation 14), we selected n_k of

(a) the N_k nodes of the 1-ha grid falling into region B_k, or
(b) all N_k grid nodes discretizing the forest area of Switzerland

randomly without replacement and computed for each such sample the approximation

$$\begin{aligned}\mathrm{Var}[S_{\mathrm{c}}(B_k) - \hat{S}_{\mathrm{c}}(B_k)] \approx{}& \frac{1}{N_k^2}\sum_{\boldsymbol{s}_i \in B_k}\mathrm{Var}[S_{\mathrm{c}}(\boldsymbol{s}_i) - \hat{S}_{\mathrm{c}}(\boldsymbol{s}_i)] \\ &+ \frac{N_k - 1}{N_k\, n_k\,(n_k - 1)}\sum_{\boldsymbol{s}_i \in \text{sample}}\ \sum_{\boldsymbol{s}_j \in \text{sample},\, \boldsymbol{s}_j \neq \boldsymbol{s}_i}\mathrm{Cov}[S_{\mathrm{c}}(\boldsymbol{s}_i) - \hat{S}_{\mathrm{c}}(\boldsymbol{s}_i), S_{\mathrm{c}}(\boldsymbol{s}_j) - \hat{S}_{\mathrm{c}}(\boldsymbol{s}_j)].\end{aligned} \tag{18}$$

For the NFI production regions, we evaluated the above expression for 1 000 independently chosen samples, each sample consisting of $\max(0.01\,N_k, 500)$ nodes in B_k, and approximated $\mathrm{Var}[S_{\mathrm{c}}(B_k) - \hat{S}_{\mathrm{c}}(B_k)]$ by their mean. To predict the mean stocks for whole Switzerland, we averaged the approximations for 2 000 samples, each sample consisting of about 5 500 randomly chosen grid nodes.

F Geostatistical analysis of OC stock in forest floor

F.1 Model

Since we used a robust estimation procedure, we did not exclude the three soil profiles with peat topsoil (cf. section 2.4), but used all 858 observations of the calibration set for estimating the model parameters. The multi-step model building procedure (cf. section 2.3.3) selected for the final model of the log-transformed OC stock in the forest floor the following set of covariates:

- spatial minimum of potential direct solar radiation in March in a local neighbourhood with radius 75 m,
- potential direct solar radiation in July,
- ratio of actual to potential evapotranspiration, the latter according to Penman (Penman, 1948),
- square root of combination of planar and profile curvature,
- square root of planar curvature,
- pH assigned to soil map units,
- categorical covariate distinguishing three aggregated soil map units,
- indicator variable discriminating coniferous and deciduous forest,
- single flow topographic wetness index based on 25 m elevation model (DHM25), with a separate regression coefficient for coniferous and deciduous forest.

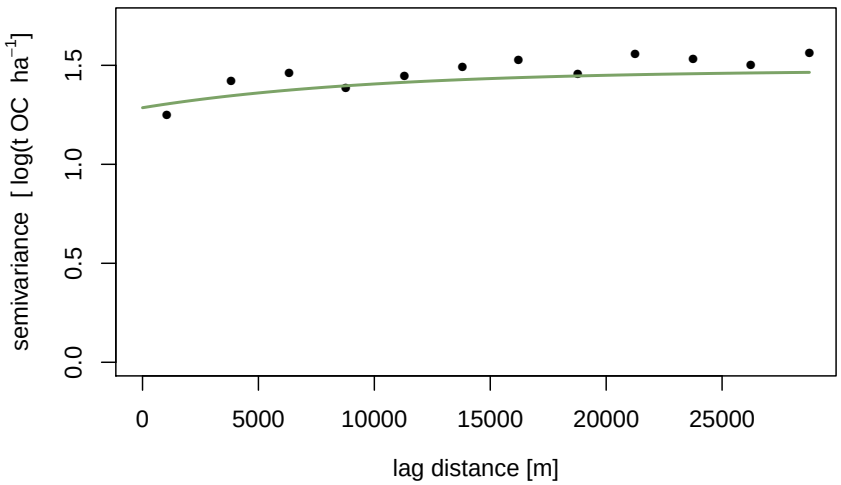

Figure 15: Fitted exponential variogram of the log-transformed OC stock in forest floor (line: robust REML estimate computed with covariates listed in section F.1 and tuning constant $c = 2$, dots: method-of-moments estimate of sample variogram of robust regression residuals).

The residuals of the linear regression with above covariates were only weakly auto-correlated (Figure 15), but with a very wide range of 30 km. Although the nugget/sill-ratio of the fitted exponential variogram was large (0.87), this model nevertheless outperformed a model with a pure-nugget variogram in the cross-validation. The optimal value of the robust tuning constant was $c = 2$. This indicates that the robustly estimated model fitted the data better than a customary REML estimate.

F.2 Model validation

Figure 16 shows the measurements plotted against the respective predictions for the calibration set (cross-validation results, prediction of log-transformed stock) and for the WSL validation set (independent validation data, back-transformed predictions and measurements plotted on log-scale). The predictions of the validation data were quite scattered, and the loess smoother line indicated a substantial positive conditional bias. This was confirmed by the large positive marginal bias: for nearly 50 % of the data the predictions were at least twice as large as the measurements (Table 4). In the cross-validation of the calibration data, both the marginal and conditional biases were smaller and did not differ much from the results obtained for the mineral soil.

The root mean squared relative errors were very large (robRMSE: 135 % [cross-validation] 147 % [validation]; RMSE: 2 290 % [cross-validation] 839 % [validation]). The large difference between the non-robust and robust measures of overall precision indicates that some data were very poorly predicted by the model. Furthermore, the large CRPS statistic and a larger variation of the prediction errors (mean absolute error: 16.4 for validation set) than the stock measurements (mean absolute difference to median: 12.4, again validation set), which would result in a negative coefficient of determination according to equation 8, confirmed that the predictive power of the geostatistical model was poor for the forest floor stocks.

According to Figure 17 of 175 validation sites 8 lay outside the 95 %-prediction intervals (expected: 9 observations), which means that the magnitude of the prediction errors was modelled well.

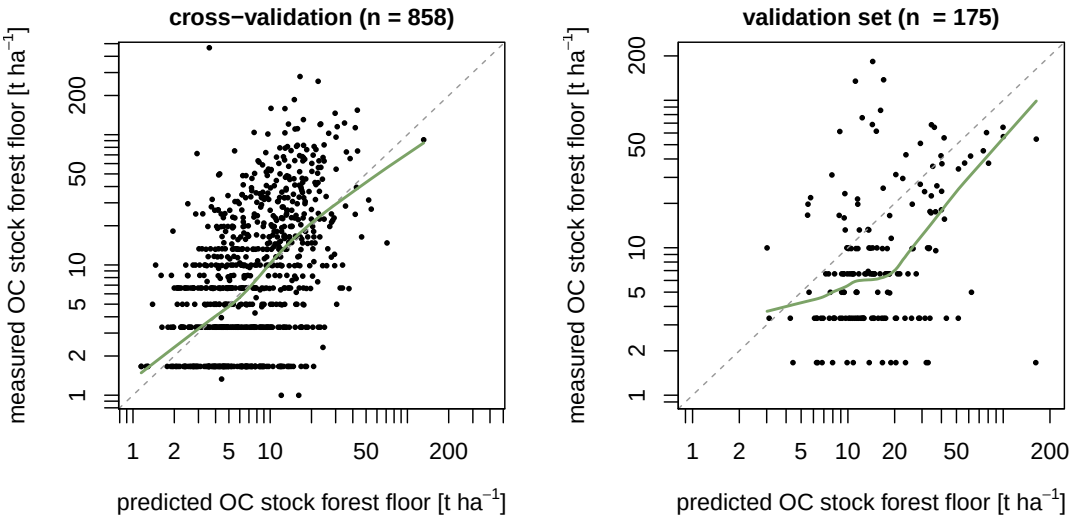

Figure 16: Scatterplots of measured against predicted OC stocks in the forest floor. Cross-validation predictions of log-transformed stocks for calibration set (left panel) and lognormally back-transformed predictions computed with the calibration data for the sites of the WSL validation set (right panel, green lines: loess scatterplot smoothers).

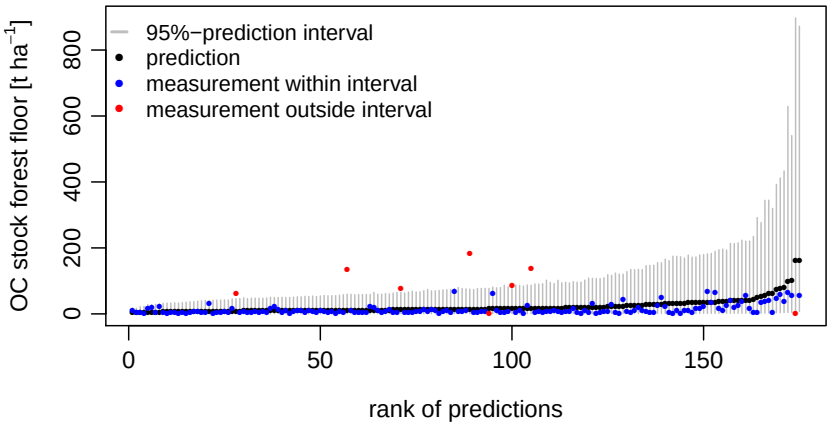

Figure 17: Ranked predictions of the OC stock in the forest floor for the WSL validation set (black) with 95 %-prediction intervals (vertical grey lines). Measurements inside the intervals are shown in blue, those outside in red.

F.3 Prediction

For computing the predictions of the stock, the parameters of the final model were estimated with the data of all 1 033 sites, based on the set of covariates listed in section F.1. The robust lognormal kriging predictions are mapped in Figure 18 for the nodes of the 1-ha grid. Large stocks (> 40 t ha^{-1}) were mainly predicted for higher altitudes in the Alps and Southern Alps where on igneous and metamorphic parent material podzolic soils prevail and for parts of the Pre-Alps. Small stocks (< 15 t ha^{-1}) were predicted for the Central Plateau, the alpine valley grounds and a large part of the Jura region, intermediate stocks for the foothills of the Pre-Alps and the highest part of the Jura mountains. The block-kriging predictions of the mean stocks for the NFI production regions and for whole Switzerland were listed in Table 3.

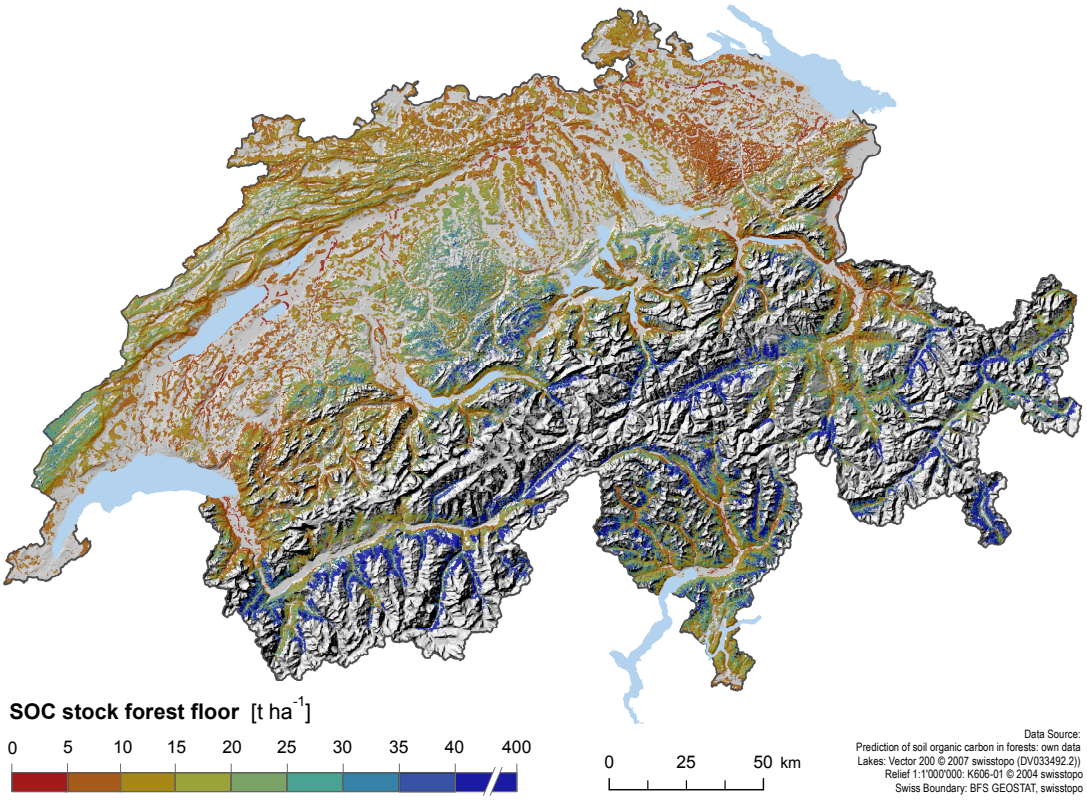

Figure 18: Robust lognormal kriging predictions of the OC stock in the forest floor of Swiss forests (computed with best-fit model with covariates according to section F.1 and tuning constant $c = 2$, smoothed as Figure 7).

G Additional plots

G.1 Correlation-biplots (example)

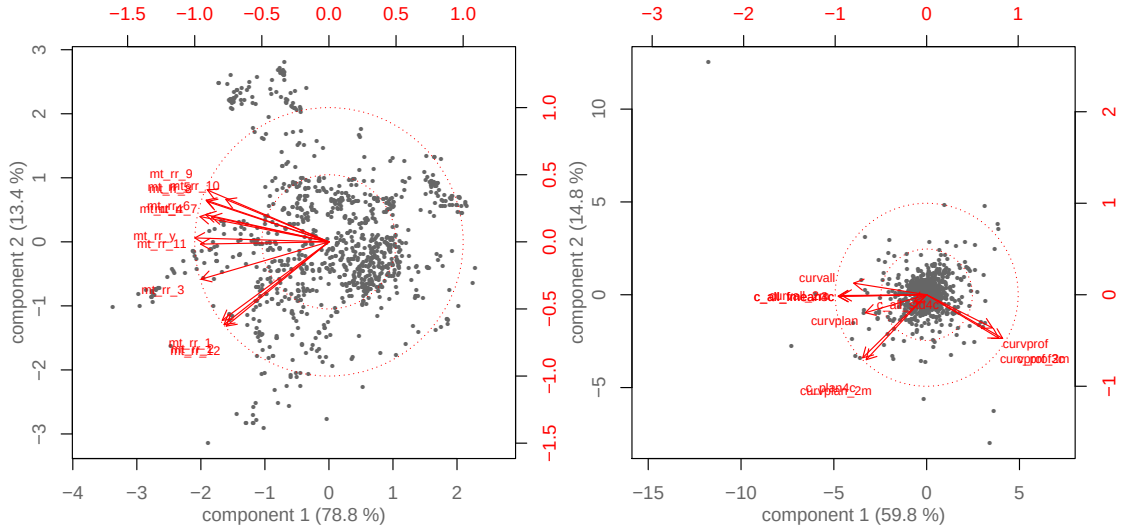

Figure 19: Correlation-biplots (Gabriel, 1981) of the first and second principal component showing for mean monthly and yearly precipitation (left panel) and planform, profile and total curvature of two terrain models (right panel) groups of strongly correlated covariates. For an explanation of the acronyms, see list of covariates in Appendix A.

G.2 Residual plots from model of SOC stock in 0–30 cm

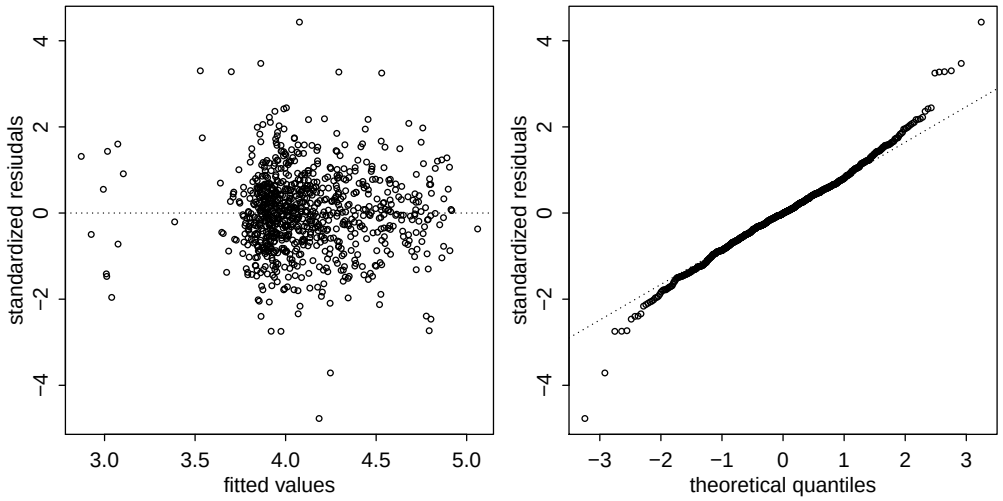

Figure 20: Tukey-Anscombe plot (left panel) and normal quantile-quantile plot (right panel) of the standardized residuals of the model for prediction of the SOC stock in the mineral soil 0–30 cm (cf. section 3.2.1). For information on the plots, see e.g. Faraway (2005, pp. 53).

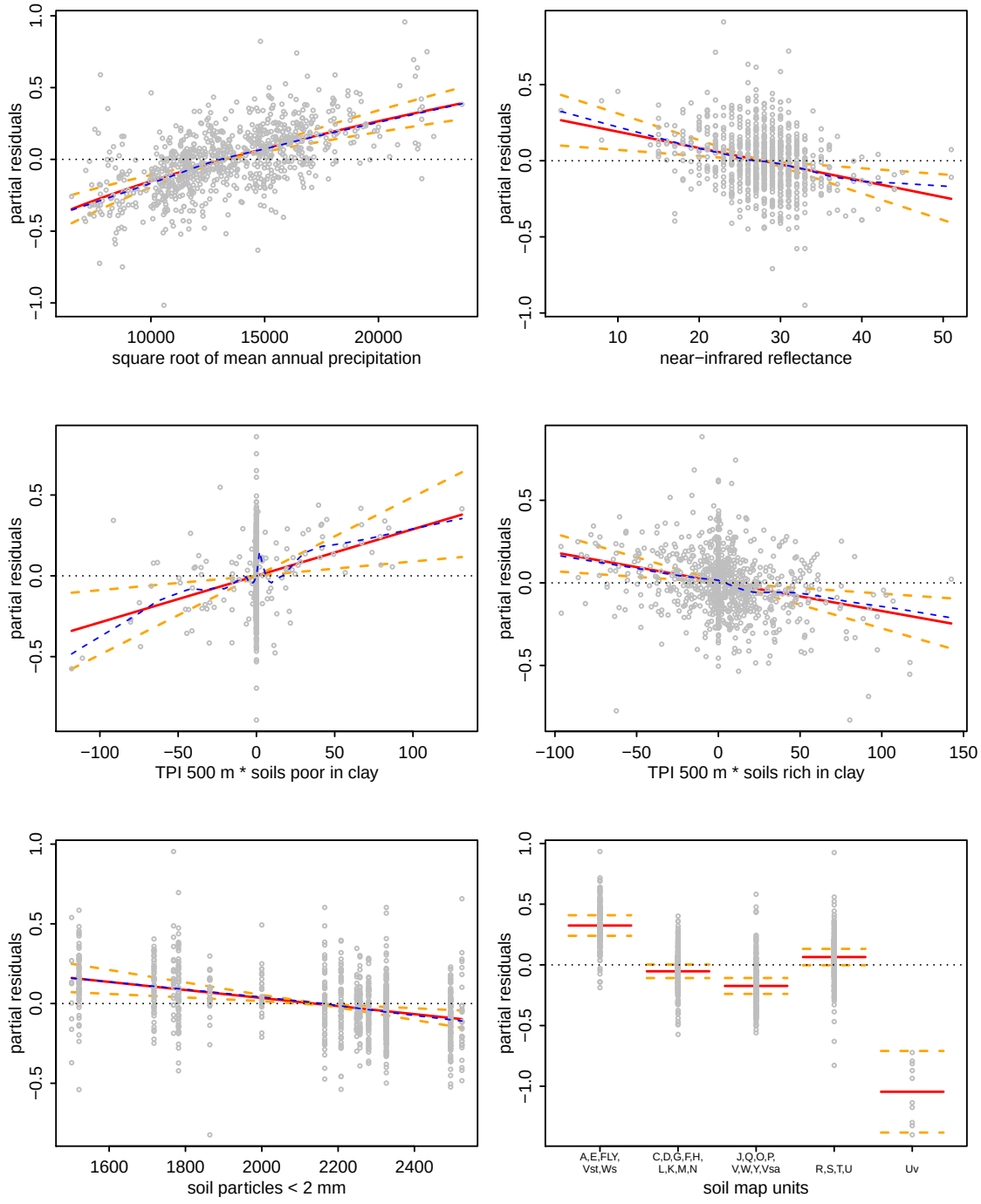

Figure 21: Partial residual plots for each covariate used in the model for prediction of the SOC stock in the mineral soil 0–30 cm (cf. section 3.2.1). For information on the plot, see e.g. (Faraway, 2005, pp. 72).

G.3 Residual plots from model of SOC stock in 0–100 cm

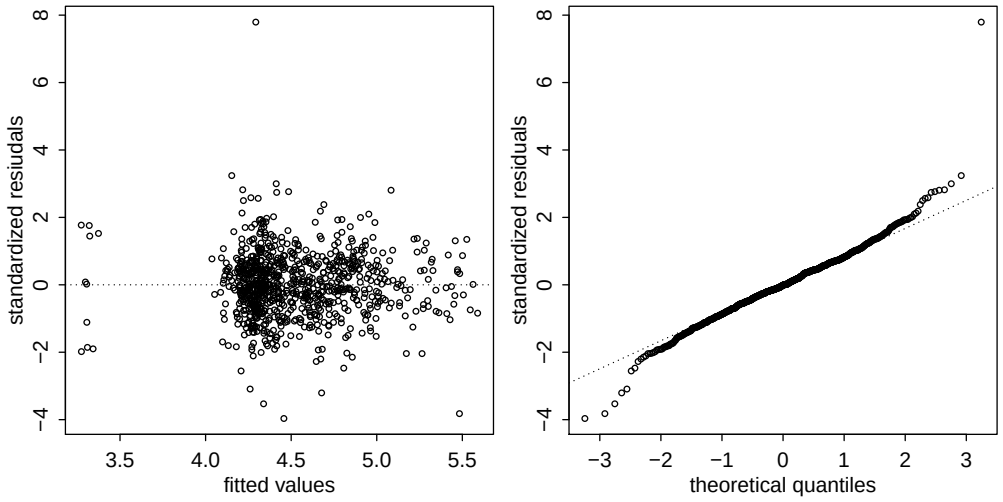

Figure 22: Tukey-Anscombe plot (left panel) and normal quantile-quantile plot (right panel) of the standardized residuals of the model for prediction of the SOC stock in the mineral soil 0–100 cm (cf. section 3.3.1).

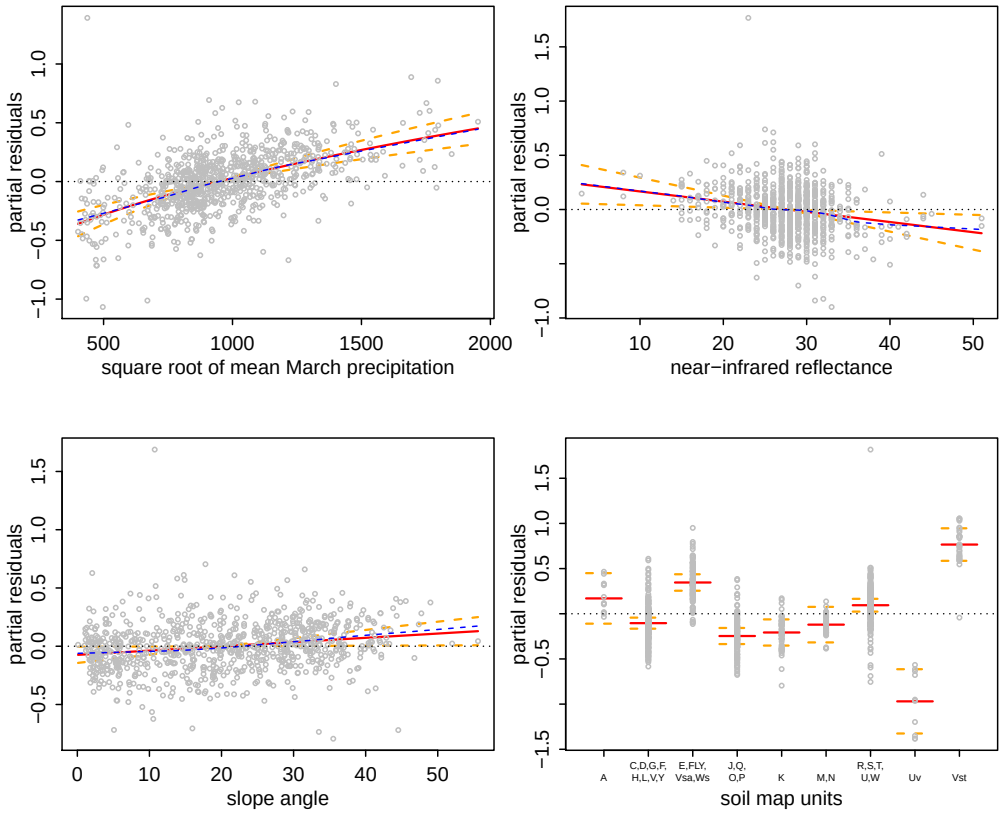

Figure 23: Partial residual plots for each covariate used in the model for prediction of the SOC stock in the mineral soil 0–100 cm (cf. section 3.3.1).

H Result tables and maps (digital)

The following files contain the main results of the project and are included in the digital appendix:

Description	Name of file
table (Excel file) containing the predictions of the OC stocks in the forest floor and in the mineral soil 0–30 cm and 0–100 cm for NFI production regions stratified by three altitude classes (Tables 3 and 5)	SOC_stocks_nfi_region_alti.xlsx
shapefile of the NFI production regions, stratified by three altitude classes, with attribute table containing the prediction results of Tables 3 and 5	shape/SOC_stocks_nfi_region_alti.shp
meta information on the shapefile	shape/readme.txt
ESRI grid (resolution: 100 m) containing the robust lognormal kriging prediction of the SOC stock in 0–30 cm of the mineral soil of Swiss forests (Figure 7)	raster/SOC_0_30cm/ha0_30n
ESRI grid (resolution: 100 m) containing the robust lognormal kriging prediction of the SOC stock in 0–100 cm of the mineral soil of Swiss forests (Figure 11)	raster/SOC_0_100cm/ha0_100n
meta information on the ESRI Grid files	raster/readme.txt
map (as PDF and JPG files) of the robust lognormal kriging prediction of the SOC stock in 0–30 cm of the mineral soil of Swiss forests (Figure 7)	figure_7_map_prediction_SOC_0_30cm.pdf/.jpg
map (as PDF and JPG files) of the robust lognormal kriging prediction of the SOC stock in 0–100 cm of the mineral soil of Swiss forests (Figure 11)	figure_11_map_prediction_SOC_0_100cm.pdf/.jpg

I R scripts (digital)

The R scripts used for the preparation of the data and their analysis are included in the digital appendix. All R code is distributed under the GNU General Public License, Version 2, June 1991. The following files are found in the respective folder:

Topic	Name of file
GNU General Public License Version 2	copying.txt
georob functions (package)	georob/georob.R georob/georob.auxiliary.funcs.R georob/georob.lgnpp.R georob/georob.methods.R georob/georob.predict.R georob/georob.xvalid.R georob/validate.predictions.R georob/variogram.R georob/georob_example.txt
preparation of data for analysis and exploratory analysis of data	1_daten_aufbereiten.R 2_datenfile_fuer_regression.R 3_explorative_analyse_erklaerende_variablen.R 4_corg_explor_variablen.R 5_fire_ph_xeromoder_analyse.R 6_regression_am_profil_version1.R 7_regression_am_profil_version2.R
design-based estimation of OC stock in the forest floor	8_design_basierte_schaetzung_vorrat_version2.R
geostatistical analysis of OC stock in the forest floor	9_corg_vorrat_auflage_lasso.R 10_corg_vorrat_auflage_georob.R 11_corg_vorrat_auflage_validate_predict.R
geostatistical analysis of SOC stock in the mineral soil 0–30 cm	12_corg_vorrat_minerde0_30_lasso.R 13_corg_vorrat_minerde0_30_georob.R 14_corg_vorrat_minerde0_30_validate_predict.R
geostatistical analysis of SOC stock in the mineral soil 0–100 cm	15_corg_vorrat_minerde0_100_lasso.R 16_corg_vorrat_minerde0_100_georob.R 17_corg_vorrat_minerde0_100_validate_predict.R
validation with monitoring data	18_validate_kabo_nabo.R
graphics for final report	19_abbildungen_reporting_bafu.R