Underestimation of boreal forest soil carbon stocks related to soil classification and drainage

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Canadian Journal of Forest Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>cjfr-2015-0466.R1</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Article</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>04-May-2016</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Dalsgaard, Lise; Norwegian Institute of Bioeconomy Research, Lange, Holger; Norwegian Institute of Bioeconomy Research Strand, Line; Norwegian University of Life Sciences Callesen, Ingeborg; University of Copenhagen Borgen, Signe; Norwegian Institute of Bioeconomy Research Liski, Jari; Finnish Environment Institute Astrup, Rasmus; Norwegian Institute of Bioeconomy Research</td>
</tr>
<tr>
<td>Keyword:</td>
<td>soil classification, soil forming factors, podzolization, moisture regime, Yasso07 SOC simulation</td>
</tr>
</tbody>
</table>
Underestimation of boreal forest soil carbon stocks related to soil classification and drainage

Lise Dalsgaard (Lise.Dalsgaard@nibio.no)¹*
Holger Lange (Holger.Lange@nibio.no)¹
Line Tau Strand (Line.Strand@nmbu.no)²
Ingeborg Callesen (ica@ign.ku.dk)³
Signe Kynding Borgen (Signe.Borgen@nibio.no)¹
Jari Liski (Jari.Liski@ymparisto.fi)⁴
Rasmus Astrup (Rasmus.Astrup@nibio.no)¹

Affiliations:

¹Norwegian Institute of Bioeconomy Research (NIBIO), Box 115, 1431 Ås, Norway.
²Norwegian University of Life Sciences (NMBU), Department of Environmental Sciences, Box 5003, 1430 Ås, Norway.
³University of Copenhagen, Department of Geosciences and Natural Resource Management, Rolighedsvej 23, 1958 Frederiksberg C.
⁴Finnish Environment Institute (SYKE), P.O. Box 140, FI-00251 Helsinki, Finland.

* Corresponding author (+47 97480373)
Abstract

Soil organic carbon (C), accumulated over millennia, comprise more than half the C stored in boreal and temperate forest landscapes. We used the Norwegian forest inventory- and soil survey network (n=719, no deep organic soils) to explore the validity of a deterministic model representation of this pool (Yasso07). We statistically compared simulated and measured soil C stocks and related differences (measured – simulated) to site factors (drainage, topography, climate, vegetation, C:N ratio and soil classification). Median C stocks (kg C m$^{-2}$) were 5.0 (model) and 14.5 (measurements). Soil C differences related to site factors ($r^2$: 0.16 – 0.37). For Brunisols, Gleysols and wet Organic soils, differences related primarily to topographic wetness. For Regosols, Podzols and Dystric Eluviated Brunisols, they related to climate, profile depth, and, in some cases, drainage class and site index. We argue that soil moisture regimes in our study area overrule tree productivity effects in the determination of soil C stocks and present conditions for soil formation that the model cannot (and does not explicitly) account for. These are processes such as humification and podsolization that involve eluviation and illuviation of DOC with sesquioxides to form spodic B-horizons and carbon enrichment due to hampered decomposition in frequently anoxic conditions.

Key words: Drainage, NFI, SOC simulation, soil classification, soil forming factors, soil moisture regime, soil survey, podsolization, Yasso07.
Introduction

Of the total carbon (C) stock in the world’s forests, 32% (272 Pg) is stored in the boreal region and only about 20% of this is stored in the living vegetation as biomass. The remaining stocks are found in soil (60%) or as litter and dead wood (Pan et al. 2011). While large C stocks are found in wet organic soils with up to 100-150 kg C m\(^{-2}\) when measured to soil depths > 1 m, (Rapalee et al. 1998), mineral soil types cover the majority of the area and typically store up to a maximum of 26 and 17 kg C m\(^{-2}\) (Gleysols and Podzols to 1 m depth, respectively) or, in shallow soils, as low as <1 kg C m\(^{-2}\) (DeLuca and Boisvenue 2012). The variability in soil C stocks across forest landscapes is large (Saby et al. 2008) resulting from the variability in soil formation i.e. topography and parent material and the long term interactions with organic matter input, organisms, climate and disturbance. The result of soil formation can be summarized by classification into soil types based on the morphology and properties of the soil pedon. While the soil forming processes are largely known, the prediction of forest soil C stocks is challenging. Losses of C from a substantial soil C pool in boreal forests could potentially contribute to positive feedback in the climate system though uncertainty is large due to insufficient knowledge as well as a lack of integration, in models, of the underlying processes known to influence soil C dynamics (Conant et al. 2011; Davidson and Janssens 2006; Todd-Brown et al. 2014). The spatial distribution of forest soil C stocks in the landscape interacts with drivers such as climate, management and vegetation dynamics and thus is an important aspect of model predictions of soil C. Global as well as regional models have been found to lack the ability to predict the spatial distribution of soil C indicating that key processes are missing (Shaw et al. 2014; Todd-Brown et al. 2013). On a national scale, models have been evaluated to improve their prediction of spatially distributed soil C stocks and to identify model deficiencies related to processes occurring at a wide range of environmental and topographical conditions (Black et al. 2014; Tupek et al. 2016).
model Yasso07, with its focus on litter input, litter chemistry and climate controlled mineralization rates (Tuomi et al. 2011a; Tuomi et al. 2009), has been applied at the global scale showing a good correspondence to measured soil C stocks (Thum et al. 2011). In its application in relatively dry boreal forest landscapes in Finland (Rantakari et al. 2012) and Sweden (Ortiz et al. 2013), regional and national scale simulations corresponded with observed soil C stocks. Using the earlier version of this model (Liski et al. 2005), it was suggested that for regions with high precipitation (P) and/or acid bedrock, the simulated soil C stocks of only 40-50% of observed stocks as found in regions of Switzerland (Thürig et al. 2005) and in Norway (de Wit et al. 2006), was due to the lack of representation of the effects of water saturation on soil aeration and the transport and stabilization of organic matter in the soil pedon over time. Norwegian mineral forest soils contain on average 13-14 kg C m\(^{-2}\) to the depth of the soil profile (de Wit and Kvindesland 1999; Strand et al. 2016), nearly twice the amount found in boreal forest mineral soils of Sweden and Finland (Rantakari et al. 2012; Stendahl et al. 2010). In Scandinavia, soil C stocks are found to increase with mean annual precipitation (MAP) and mean annual temperature (MAT) for well-drained soils (Callesen et al. 2003) and with increasingly poor drainage (Olsson et al. 2009).

The climatic range in Norway is wide (MAP 500-4000 mm, MAT -4 to 8 °C). The litter data used for the parameterization of Yasso07 covers both these ranges but high precipitation values only from climates warmer than the Norwegian (Tuomi et al. 2009). In this study we evaluate the Yasso07 models climate dependency in relatively extreme climatic situations with the aim to describe the limits of applicability in climate and drainage gradients. The steep and irregular topography found in Norway combined with many shallow soils and a wet and cool climate, promotes a high frequency of hydromorphic soils (32% based on soil survey, (Strand et al. 2016)). This, in addition to a predominant soil process of podsolization on acid soils in regions of high precipitation, will challenge a soil model where effects of
anoxic conditions and transport and stabilization of organic matter on the accumulation of C
are not specifically addressed. Our objective is to evaluate the ability of Yasso07 to predict
forest soil C stocks and its spatial distribution in a wet and topographically heterogeneous
boreal landscape and contribute to the identification of circumstances where the model fails
and thus point to model issues in need of improvement. We expect that the use of soil
classification as demonstrated in (Shaw et al. 2015) can highlight soil genesis that involves
humification processes, and that it may be useful to identify the limits of application of the
soil C model. In this study we compare simulated and measured forest soil C stocks on a plot
level (n=719). We study the soil C differences (measured – simulated) and hypothesize that
they are related to i) climate, ii) soil drainage class, iii) soil type and iv) stand properties.
Further, we hypothesize that we can identify site properties for which the model is successful
in estimating soil C stocks and those for which it fails.

Materials and methods

The comparison of measured and simulated soil C stocks was done on the spatial scale of
individual NFI (national forest inventory) plots where data were available in a systematic grid
covering the Norwegian forest landscape. On these plots, soil survey data (ICP Forest Level I,
International Co-operative Programme on Assessment and Monitoring of Air Pollution
Effects on Forests) offer a set of field based (measured) soil C stock estimates. The soil C
decomposition model Yasso07 was used to simulate soil C stocks at a subset of plots (n=719).
The subset was the maximum number of plots matched between the soil survey and the
current NFI grid avoiding cases of inconsistent plot identification in early inventories.
Differences (measured - simulated stocks) were analyzed in relation to hypotheses i)-iv) using
linear regressions and analysis of variance.

The Norwegian NFI and ICP forest soil database

https://mc06.manuscriptcentral.com/cjfr-pubs
The basic structure of the NFI is a 3×3 km grid of permanent circular sample plots (radius 8.92 m) covering Norway (Tomter et al. 2010). Stand, plot and tree characteristics are quantified on each NFI plot every 5 years i.e. the whole country is fully covered each 5 years. On a subset of the NFI plots (in a systematic 9 x 9 km grid), soil profile description, soil sampling and chemical analyses were carried out on 1040 profiles between 1988 and 1992 (Esser 1994; Esser and Nyborg 1992; Lorenz 1995). 719 of the soil survey plots were matched with the current NFI grid and are fully located within forest (NFI data exclude soils with > 0.4 m deep organic layers/peat). Due to the distribution of the forest area, most soil data are located in the south-eastern parts of the country compared to the central mountain regions, the north and along the west coast. Soil profiles were systematically placed, described according to standard procedures and soil samples were collected by soil horizon in the soil profile as well as in auger samples from across the plot; see details on soil sampling and chemical analyses elsewhere in this issue (Strand et al. 2016). Registrations included general area data, topography, landform, parent material and drainage. The Canadian soil classification system was applied (ACECSS 1987).

Calculation of field based soil C stocks

We applied the C stock calculation approach as described in (Strand et al. 2016). The C:N ratio was calculated for the organic horizons or, if missing, the surface A horizon only. Total profile C stock was calculated down to the full depth of each profile. Norwegian forest soils are generally shallow and profile depths are limited by either bedrock or stoniness, however, where the profiles were deep, a cut-off depth at one meter from the soil surface was used ($C_{prof}$). All data needed for C stock calculations were present in the database except for bulk density. Bulk density was estimated from a pedo-transfer function (Baritz et al. 2010) or from separate data (de Wit and Kvinesland 1999; Strand et al. 2016). Fifteen percent of the profiles had no information on the exact thickness of the deepest horizon, for these horizons,
we set a thickness of 0.2 m, less if this caused the total profile depth to exceed one meter. For
mineral soil horizons missing information on stone content (6% of horizons) and gravel
content (15% of horizons), we estimated these fractions either by using the average of the
above and below horizons, or using the median of these fractions from soils of the same
parent material (Strand et al. 2016). Error estimates in (Strand et al. 2016) suggest that the
largest errors were found for the hydromorphic soils (1.0 – 6.2 kg C m\(^{-2}\)) and less for the self-
drained mineral soils (0.7 - 3.2 kg C m\(^{-2}\)).

Definition of soil types

Based on soil profile classification, five soil types were defined for the purpose of this study:
i) Podzols and Eluviated Dystric Brunisols (PO-BRe) – zonal soils where climate and
vegetation are the most influential soil forming factors; ii) Gleysols and wet Organic soils
(GL-ORw) – intra zonal soils where local factors such as topography or parent material have a
major influence on soil characteristics overriding the effects of climate and vegetation; iii)
Non-soils and Folisols (NO-FO, azonal) where soil development is restricted by shallow soil
deposits; iv) Regosol (RE, azonal) where soil development is restricted mainly by young age;
and v) the remaining profiles classified as Brunisols (BR) - these show a wide range of soil
characteristics similar to, but not qualifying to, other soil orders.

The Yasso07 model

Yasso07 simulates soil C stocks based on litter decomposition where soil C is distributed to
pools that are unrelated to soil depth and does thereby not rely on the soil pedon concept.
Yasso07 comprise a set of first order differential equations and it relies on three main
assumptions on the litter decomposition process (Tuomi et al. 2011b; Tuomi et al. 2009).
According to the first assumption, litter consists of four chemically different compound
groups distinguished by solubility in solvents or hydrolysability in acid, and each of these
groups has a mass loss rate independent of litter origin. These AWENH groups are E: soluble in a non-polar solvent (ethanol or dichloromethane), W: soluble in water, A: hydrolysable in acid or N: neither soluble nor hydrolysable. According to the second assumption, the mass loss rate of each compound group depends on MAP and MAT. The climate dependence of decomposition is assumed (of necessity) in the model to be identical for all compound groups (Tuomi et al. 2009). Dependence factors control the decomposition rates and in the climatic range relevant to this study, increase asymptotically with MAP (saturating at 1500 – 2000 mm annually) and exponentially or linearly with MAT (Dalsgaard et al. 2016). According to the third assumption, decomposition of the four compound groups found in litter results in 1) the formation of a more recalcitrant compound group (humus, H), 2) mass flow among the compound groups and 3) mass loss from the system. Large litter elements go through a process specific for woody litter decomposition in addition to the process of non-woody litter decomposition and their decomposition takes longer compared to non-woody litter (Tuomi et al. 2011a). The parameter values of Yasso07 were found using a Markov chain Monte Carlo method and a large dataset of litter decomposition and soil C. The dataset had litterbag mass loss measurements from Europe, North America and Central America, dead wood decomposition measurements from Finland, Estonia and Russia (Tuomi et al. 2011a), soil C stock measurements from Finland (Liski and Westman 1995) and soil C accumulation measurements up to 4 kg C m$^{-2}$ from a 5500-year-long soil chronosequence on the Finnish west coast (Liski et al. 1998). The soil C stock measurements were particularly important for determining the C flows to the most recalcitrant N and H compartments of Yasso07 and consisted of paired estimates of litter input and soil C stocks in the range 4-12 kg C m$^{-2}$ for six sites differing in forest productivity class and tree species (Table A1). Although these data were only from Finland, Yasso07 has been shown be suitable for estimating soil C stocks at a range of spatial scales, and also at global scale with some exceptions (Goll et al. 2015) (see
Discussion). Each of the five compound groups (AWENH) represents a model soil C pool; these are not allocated to specific soil layers.

*Model input*

Yasso07 takes input on plant and tree litter, MAT and MAP. The annual litter input is estimated for five different sources using biomass models and turnover rates for components of tree and ground vegetation (Dalsgaard et al. 2016; Miljødirektoratet 2015). Tree biomass was estimated from Swedish models for aboveground (Marklund 1988) and belowground (Petersson and Ståhl 2006) parts. Ground vegetation biomass was based on Finnish models for aboveground parts (Muukkonen and Mäkipää 2006) and a review of the relationship to belowground parts (Peltoniemi et al. 2004). The five sources of litter were i) the living trees on each NFI plot using individual tree data for Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.) and deciduous species, primarily birch (*Betula pendula* Roth. and *Betula pubescens* Ehr.); ii) ground vegetation distributed to lichens, moss, shrubs, herbs and grasses with biomass depending on dominating tree species and stand age; iii) natural mortality; iv) residues from non-commercial wood extraction and v) residues from commercial thinning and final harvest using individual tree data and NFI harvest observations. The sources iii), and iv) were estimated as a percentage of the living tree biomass based on tables of average biomass change observed in two consecutive NFI’s (2000-2004 and 2005-2009) and stratified according to site index, stand age and tree species composition (Antón-Fernández and Astrup 2012). Turnover rates for the living tree and plant components and the studies and compilations from which they originate are listed in (Dalsgaard et al. 2016). They varied across tree species and ground vegetation layer except for tree fine roots where a value of 0.6 year\(^{-1}\) is applied for all tree species and for ground vegetation belowground turnover where a value of 0.33 year\(^{-1}\) was used for dwarf shrubs and for herbs and grasses. The distribution to compound groups (litter chemistry) for the litter
components was based on a global database (Liski et al. 2009; Miljødirektoratet 2015; Tuomi et al. 2009) and for ground vegetation on (Peltoniemi et al. 2004). Climate data was down-scaled by the Norwegian Meteorological Institute to the exact position (and height above sea level) of the NFI plots (Solberg et al. 2015).

Model simulation

Initial soil C stocks in the five model pools (AWENH) were found for each plot by spin-up to a steady-state i.e. running the model for 5000 time steps (mimicking 5000 years) using constant litter input. Litter inputs used in the spin-up should ideally reflect the history and litter production of the site prior to the simulation time series. To represent stand history, information available in the first permanent sample plot NFI (1986-1993) was used to create a number of rules (Table 1) for stand development starting in 1956-1960. Using these rules, a historical litter input time series was created for each plot where, in the majority of the cases, this was a matter of estimating previous stand dynamics in the same stand. Model input for the plot specific spin-up was the 30-year mean of plot climate and the mean litter input estimated for all plots in the first permanent sample plot NFI (ca. 11 000 plots), stratified by site index (9 levels) and dominant tree species (3 levels) . The litter input for the spin-up was estimated following the same methodology as for the historic and current time series and includes litter from the living trees, ground vegetation, non-commercial wood extraction, thinning, final harvest and natural mortality (Table 2). After the spin-up, the historical litter time series for each plot prior to the first NFI was used (pre-simulation) followed by the litter input time series based on the actual permanent plot NFI time series (the current litter time series). The use of spin-up initialization of the soil C pools indicates that we assumed plots to have been under forest cover and that the soil C pools were in a steady-state at the beginning of the simulations. However, after the pre-simulation the soil C pools are no longer in a steady-state. A similar approach was used to initialize model runs for soil C development in
Swedish forests (Ortiz et al. 2013). Yasso07 was applied in parallel runs for non-woody, fine-woody and coarse-woody litter respectively (size parameter 0, 2 and 10 cm; Miljødirektoratet 2015). Climate for the spin-up calculation was the most recent climate normal (1961-1990) and for the litter time series (historical and current) it was the mean for 1991-2008. Output from the simulations was expressed in kg C m$^{-2}$. The model application was limited to plots where NFI data show an organic soil horizon < 0.4 m (regardless of litter-derived organic material or peat). Input calculation and model application was done in R (R Development Core Team 2010).

**Statistical analyses**

Three statistical analyses were carried out: simple linear regression of measured vs. simulated soil C stocks, relative importance analyses to find the power of continuous predictors to explain the difference between measured and simulated stocks, and an analysis of variance to find the significance of categorical predictors for explaining the difference between measured and simulated stocks. The latter two analyses were carried out for all plots as one ensemble as well as for each soil type. To ensure compatibility in scale between measured and simulated soil C stocks (to represent the average soil C stock for a 250 m$^2$ plot), the measured stocks were adjusted for the percentage of bare rock found in the field survey (i.e. plot area where no soil is found). Thus, on plots where only a fraction of the area contained soil, the adjustment of the measured soil C stocks resulted in a lower value. For the majority of plots (79%), this adjustment was less than 5% and for only 3% of the plots was the adjustment 50% or more. However, with the adjustment, an additional source of randomness was introduced.

Further, to represent the sensitivity of the model estimates to errors/uncertainty in litter input rates, two simulations were carried out; i) a standard simulation (Dalsgaard et al. 2016) and ii) a simulation where living fine root turnover for trees and ground vegetation was doubled.
(turnover rates (year$^{-1}$) for trees 0.6 vs. 1.2 and for herbs, grasses and shrubs 0.33 vs. 0.66).

Thus, two litter input datasets were produced differing only in the amount of fine root litter.

Fine root turnover rates are challenging to measure and an uncertainty of 100% is not unrealistic (Brunner et al. 2013). The model output from the year of the soil survey as well as the spin-up C stock was used in the analyses to evaluate the stability of results. The continuous predictors included in the statistical analyses were MAP and MAT (1961-1990), the Yasso07-specific climate dependence factor derived from them (equation 1), the profile depth (as measured in the field), height above sea level, topographic wetness index (twi), and the C:N ratio of the organic layer and inclination (Table 3). Twi was calculated with the national Norwegian 10-meter resolution digital elevation model as input for the multiflow algorithm (Toma et al. 2002) in the GRASS GIS software. High twi values correspond to areas receiving water from the surrounding terrain. The climate dependence factor (cdf) describes the combined effect of precipitation and temperature as formulated in Yasso07:

$$cdf = \exp(\alpha MAT + \beta MAT^2)(1 - \exp(-\gamma MAP))$$

(1)

where values for $\alpha, \beta, \gamma$ (fitted parameters) are from earlier studies (Thum et al. 2011).

Although being a nonlinear deterministic function of MAP and MAT, the cdf turns out to be a more suitable predictor than either MAT or MAP alone, or in combination. This is due to the model property that decomposition is directly dependent on cdf. Soil texture was initially included but showed non-significant effects of %clay, %silt and %sand in simple linear regression as well as multiple regression including the continuous predictors mentioned above. The soil profile data are dominated by sandy soils (median %clay 5 and median %sand 63) and clay content in previous studies did not appear to influence the soil C difference nor the soil C stocks (Strand et al. 2016). The categorical predictors were (Table 4): dominant tree species, site index (stand height (m) at age 40 years), vegetation type (combining three levels of nutrient status with three levels of moisture status) and drainage class (A=excessively
through S=very poorly, 16 levels). Drainage class was based on a combination of internal (fx. observed color and depth of mottles in soil horizons) and external (fx topography) factors. Nearly 50% of the profiles are found in spruce dominated forests which typically have a higher site index than pine and birch forests, and 60% of the profiles belong to the zonal soil type primarily with good to medium/poor drainage. Each analysis is detailed below.

Regression analysis of measured vs. simulated soil C stocks

A simple linear regression analysis was performed to investigate the relationship between the measured and the simulated C stock. Different variations of both measured and simulated C stocks were available (i.e. with or without correction for bare rock surface; standard or double fine root litter input). For measured soil C stocks \( (C_{\text{prof}}) \) two variations were tested: with and without bare rock correction. For simulated C stocks, four variations were tested: standard and double litter input runs at time of the soil survey, standard and double litter input run using spin-up values. Correlation coefficients and p-values for the slope estimates are reported.

Analysis of the relative importance of continuous predictors

First, the soil C difference (measured – simulated) was calculated for the two variations of measured soil C stock and the four variations in simulated soil C stock as described above. The soil C difference was regressed against all eight continuous predictors simultaneously. Second, the relative importance and the associated direction of the slope (+ or -) among continuous predictors in the explanation of soil C difference were found for i) all plots regardless of soil type and ii) for each soil type separately. Relative importance provides a natural break-down of the total explanatory value of a statistical model to contributions from each predictor. The metric used for the break-down in our case is the averaging sequential sums of squares over orderings of regressors \( (\text{Lindeman 1980}) \); available as the “lmg” method in the “relaimpo” package \( (\text{Groemping 2006}) \) in R. The differences (or similarities) in results
emerging when applying different variations of measured and simulated C stocks serve to test the stability of results, particularly with respect to the significance of individual predictors. For simplicity, we report results mainly from the standard model run, $C_{\text{prof}}$ corrected for bare rock surface.

**Analysis of variance - significance of class variables**

Building on the importance analyses we investigated the significance of the categorical predictors for explaining soil C differences using $C_{\text{prof}}$ corrected for bare rock surface. The number of separate levels was: three for tree species; nine for site index; 11 for vegetation type; and 17 for drainage class (one of these representing “no data”). The relative importance analysis served as a selection tool for the statistical model, i.e. for each variant of observations, simulations and soil type (one of the five, or the whole dataset), we included the 2-3 most important continuous predictors (requiring that this subset explained a minimum of 60% of the total $r^2$ achieved). Then, an analysis of variance model was fitted, using the four categorical predictors. As far as the number of observations allowed, up to fourth order interactions were kept, both between continuous and categorical predictors and across these. The results were ranked according to their p-value. All $r^2$-values reported are adjusted.

**Results**

Simulated soil C stocks (mean as well as median values) were substantially lower than measured stocks (Figure 1, Figure 2) with a mean (SEM=standard error of the mean) and median of 5.3 (0.1) and 5.0, respectively, kg C m$^{-2}$ for the standard model application. The mean (SEM) and median of measured soil C stocks was 15.8 (0.3) and 14.5 kg C m$^{-2}$. While simulated soil C stocks visually decline with increasing MAP and MAT, this trend is not seen for the measured soil C stocks (Figure 2). Measurements are characterized by a more pronounced tail of extreme (high) values which is not seen in the simulations (Figure 3). Soil
C differences (measured – simulated) were in general large and positive. Small and negative
differences were found in well drained conditions and on poorly developed soils (soil types
NO-FO, RE, BR) as compared to conditions characterized by poor drainage, intra zonal soils
(GL-ORw) and highly developed soils (PO-BRe) (Figure 4). On average, simulated stocks
increased by 28% when fine root turnover rates were doubled. Measured stocks were slightly
decreased when adjusted for the coverage of bare rock surfaces; mean 15.0, median 13.7 kg C m$^{-2}$.

Simple linear regression analysis showed that there was no significant relationship between
measured and simulated soil C stocks regardless of bare rock correction or not ($r^2$-values <
0.01). The correction for bare rock further increased the residual sum of squares. When
simulated soil C stocks were equal to the spin-up stocks, and the uncorrected $C_{prof}$ were used,
the slope was significant ($p<0.05$), but negative.

The relative importance of the eight continuous predictors for the soil C differences showed
variability among the soil types, but stability across the different C stock variations and
against correction for bare rock. The most important predictors across the different soil types
were cdf, MAP (all soil types except GL-ORw), profile depth (soil types PO-BRe, RE) and
twi (soil types BR, GL-ORw) (Table 5). Only for the GL-ORw soil type was the C:N ratio
ranked important. Slopes for the most important predictors were positive except for MAP in
the case of PO-BRe and RE. Generally, the group of important predictors for the soil C
difference stayed the same independent of simulation- or measurements variant, as long as the
soil type was kept constant. An example of this is the PO-BRe soil type exhibiting the largest
number of samples. The sequence of importance for this soil type, as shown in Table 5, was
valid for both the simulated stock from spin-up runs and the simulated stocks at the time of
the soil survey. In the latter case, for doubled fine root turnover, the profile depth and MAP
exchanged the places two and three in importance. For GL-ORw and BR soils, the regression
was clearly dominated by the twi and for RE soils, the soil C difference is mainly explained by profile depth. Due to insufficient data, no regression was developed for the group of NO-FO. The \( r^2 \)-values were moderate (0.18-0.37); lowest for GL-ORw soils and highest for PO-BRe.

The analysis of variance with soil C difference as the dependent variable, run irrespective of soil type, suggested a large number (21) of significant effects (not shown). Drainage class, site index and vegetation type were significant single effects. The analysis, when performed for each soil type (Table 6), showed that for GL-ORw soils, there was a marginally significant interaction effect between twi, drainage, and vegetation type and for PO-BRe soils, there were 7 significant effects, including site index and drainage class as single effects. For RE and BR soil types, none of the class variables led to significant effects. For soil type NO-FO, the number of samples was too limited to perform the analysis.

The effects of drainage class and, for PO-BRe soils, drainage, site index, profile depth and vegetation type are illustrated in the supplement (Figure S1-S9). Increasingly poor drainage was paralleled by decreasing simulated soil C but in general increasing measured soil C (Figure S1, S2). Where, for PO-BRe soils, simulated soil C increased with site index in response to a higher litter input, this was not evident for the measured soil C (Figure S3). Where measured soil C tended to increase with profile depth in general, simulated soil C tended to increase with profile depth only in well-drained conditions (Figure S4). The effect of climate (cdf) was illustrated in Figures S6 – S8 indicating that the systematic bias (as in Figures S1 - S3) was smaller for a subset restricted to low MAP and MAT. Vegetation type did not show a consistent picture (Figure S5, S9).

NO-FO was clearly the soil type with the lowest soil C difference across the whole dataset but we did not have enough profiles to perform the regressions as for the other soil types. Figure
S1 indicates that also for NO-FO soils does the model underestimate stocks in poorly drained plots.

No significant effect of tree species was found and though lower simulated soil C stocks for deciduous forest plots than for coniferous plots are indicated (Figures S2) an effect on the soil C difference was not found to be significant.

**Discussion**

For the majority of cases, model simulations of soil C stocks for 719 forest plots across Norway were considerably lower (on average 8-10 kg C m$^{-2}$) than measured soil C stocks. The difference was most severe for plots characterized by poor drainage and for soil types where excess moisture and/or eluviation/illuviation processes dominate soil development. The lack of a significant (or in one case a negative) relationship between measured and simulated soil C stocks confirms that crucial processes determining Norwegian forest soil C stocks seem to be lacking in the simulations.

*Factors influencing model performance – is soil type important?*

Our dataset is characterized by large climate gradients, in particular for precipitation, and the soil C difference clearly increased with increasing cdf as based on the importance analyses and confirmed visually in that the differences increase with increasing MAT as well as with MAP. Plot specific comparisons of measured and simulated soil C stocks in Swedish forest also revealed that models (Yasso07 as one) underestimated measured stocks, and that the underestimation was most severe at nutrient rich and moist sites (Tupek et al. 2016). The study from Sweden exhibits a smaller gradient in both temperature and precipitation than our study (as well as an expected higher frequency of nutrient rich sites) which may be why the conclusions are slightly different between the two studies.
We expected that the most important predictors for the soil C differences would vary among soil types due to differences in the prevailing soil forming processes and this was confirmed. C stocks are to a large extent reflected in different soil types due to differences in soil forming processes ([De Vos et al. 2015; Strand et al. 2016]) where only a small range of the soil forming processes are represented in the decomposition model that we apply. For BR and GL-ORw, moisture, expressed through the twi, was singled out as the most important variable to explain the structure of the soil C differences. Twi reflects the topography and not the hydrology or absolute moisture content of a plot; however, in the Norwegian climate low lying areas will typically become water-logged. Frequent anoxic conditions will slow down decomposition and promote accumulation of soil C particularly in organic layers ([Olsson et al. 2009]). A positive correlation between C:N ratios and soil C differences in the GL-ORw soil type most likely indicates a low degree of humification at high C:N ratios in these soils. Also, DOC transported from the surrounding areas may reprecipitate in soils of low lying areas. An increase in soil C difference with MAP (as found in the importance analyses and in the anova), in addition to the effect of cdf, was found only for BR soils. The effect exemplifies that not all climatic effects are aligned with cdf (i.e. the decomposition rate of the model), and was, in the case of BR soils, likely due to lower MAP and MAT (and lower cdf) than for any of the other soil types. Soil C difference increased with profile depth in RE and PO-BRe soils, which made up the majority of soil profiles in the dataset, and climate (cdf) was a strong predictor. Whereas the addition, decomposition losses and transformation is largely included in Yasso07, the transport (and loss) of DOC from surface horizons to the lower (B) soil horizon are not considered. Leaching and reprecipitation of DOC in the podsolization process is favored by mild winters and a large precipitation surplus. In areas where these processes are significant for soil development, and soil processes have been ongoing in a rather stable vegetation regime for thousands of years, even minute amounts of
pedologically stabilized C added every year to a B horizon may be of significance at the
millennial time scale, but also A-horizons are strongly humified and contribute to C stocks in
warm and wet parts of Norway. DOC fluxes within the soil profile i) increase with MAP
(Michalzik et al. 2001), and ii) with soil organic C stock of the topsoil (Gielen et al. 2011),
and iii) are higher in conifers than in broadleaves (Froberg et al. 2011). All three conditions
are characteristics of a large portion of the Norwegian forest soils. In Japan, DOC flux rates
from the O-horizon were measured to 6-9 g C m$^{-2}$ year$^{-1}$ and of this flux 0.9 – 2.8 g C m$^{-2}$ year$^{-1}$
percolated into the B-horizon (Fujii et al. 2011). In the Fichtelgebirge, Bavaria, Germany, it
was found that 2 - 5.5 kg C m$^{-2}$ (19%-50% of soil organic C) was sorbed, stabilized DOC in
the soil profile (Kaiser and Kalbitz 2012; Kalbitz and Kaiser 2008). If 1-3 g of DOC is
precipitated every year for 3000 years in a continuously vegetated ecosystem, it will
potentially amount to 3-9 kg C m$^{-2}$ when the B horizon in that soil profile is sampled. This is a
simplification of the ongoing processes, but illustrates our point.

Twi alone was not a significant predictor for PO-BRe but effects were found for site index
and for drainage. Where twi is calculated from topographic information, the drainage classes
are determined during profile description based on interpretation of morphological features
related to soil moisture. These drainage classes result from a combination of topographical
placement and soil properties determining the permeability and water holding capacity. Thus,
soil hydrology as a combination of precipitation, evapotranspiration, topography and soil
conductivity is a better predictor for the soil C differences in these soils. Local hydrological
effects seem to overrule the importance of litter input in Norwegian podzols and eluviated
brunisols. Soil texture was not an important predictor and the data set included only few clay-
rich profiles. Thus, in general, high soil C accumulation in poorly drained soils does not seem
to be related to low soil conductivity due to high clay content but possibly to compacted or
cemented soil layer. Soil texture, particularly clay content and activity may control soil C
stocks (Richter et al. 1999; Torn et al. 1997) due to its ability to stabilize mineral soil C but this does not emerge as a strong predictor in our case. Including soil texture did not improve model representation of soil C stocks in Swedish forests (Tupek et al. 2016). Our observation that soil moisture regime is important for the model representation of soil C accumulation is not a new insight. Soil moisture and anoxic soil conditions were related to model underestimation of soil C stocks for regional (Tupek et al. 2016) and for global (Goll et al. 2015) model applications. Comparing model and measured C pools in forest across Canada revealed a poor correlation for the soil pool and incorporation of information and processes related to forest and soil types were expected to be necessary in future model development (Shaw et al. 2015; Shaw et al. 2014).

In Norwegian forests, tree species are confounded with site index (Spruce > Pine, deciduous) which is likely the reason for a lack of tree species effects. Also, the balance of litter input and decomposition produced in the model reflects the experience on C accumulation particularly in organic layers of coniferous and deciduous forests (Vesterdal et al. 2013). Thus, the model ranks tree species as expected for soils where tree species and litter input is a major influencing factor in soil development.

In earlier national applications of Yasso07, measured (soil inventory) and simulated soil C average stock was similar in Sweden for mineral soil types with 69 and 73 t C ha$^{-1}$ (year 2000) respectively (Ortiz et al. 2013) and in southern Finland with 67 vs. 67 t C ha$^{-1}$ (Rantakari et al. 2012). Both applications represent a dryer climate than the Norwegian one and the Swedish study included measured (partly interpolated) soil C only to 0.5 m depth. If, in our study, we limit the observations to those approximately similar in climate to Sweden (MAP < 1000 mm, MAT < 7 °C) the standard simulated stock mean is 6.5 kg C m$^{-2}$ and 8.4 kg C m$^{-2}$ using double fine root litter input. These amounts are comparable to those found in Sweden, however, the bias is still large (median of measured stocks for this subset of data is
13.0 kg m\(^{-2}\)). We suspect that the potentially deep profile depth (unless restricted) and a wide range of soil types in the Norwegian soil data set, including a high proportion of hydromorphic soils (Strand et al. 2016), enhances the expression of a variety of soil processes that are not explicitly accounted for in Yasso07 and therefore the resulting differences between measured and simulated C stocks become large.

Spatial scale of Yasso07 application

Yasso07 is used for greenhouse gas reporting and other national or regional scale applications (Ortiz et al. 2013; Rantakari et al. 2012). However, the validity of Yasso07 results has been tested at a range of spatial scales. At the global scale, the distribution of organic soil C to different biomes calculated using Yasso07 was “in good agreement” with a Harmonized World Soil Data Base, though for tree dominated biomes and biomes affected by anaerobic decomposition and permafrost the model showed a consistent underestimation (Goll et al. 2015). Simulated soil C stock in next-to-cropland forests and croplands was similar to Finnish measurements (Akujärvi et al. 2014) and C stock and stock changes were estimated “reasonably well” in a region of Loss Plateau in China (Wu et al. 2015). A number of studies have made comparisons on the site scale; simulated soil C stocks were about 20 % smaller than measured at two sites in Switzerland (11.2 vs. 13.9 kg m\(^{-2}\) at one site, 11.1 vs. 14.2 kg m\(^{-2}\) at another) (Rühr and Eugster 2009). Good agreement between measured and simulated stocks and stock changes associated with afforestation and deforestation were found in Finland (Karhu et al. 2011) and the Loess Plateau in China (Lu et al. 2013). According to these validity tests, we considered the Yasso07 model and our approach suitable for the present study. Based on earlier studies using Yasso (de Wit et al. 2006), a considerable underestimation by the model was expected. Summing up the plot scale results, as in the current study, was a sound way of obtaining regional and national estimates. From earlier we found that the spatial scale in itself (i.e. the spatial resolution of climatic input data) has a
relatively minor effect on the simulated stocks (Dalsgaard et al. 2016). The plot scale simulations made it possible to identify the most important reasons for the mismatch between the model-calculated estimates and the measurements. This information is very useful when developing the model further, and such information would not be obtained if the model calculations were performed on large regional scales only. Model complexity implies that model parameters are required which are not routinely available at the site level, however, to model soil C stocks in a wide range of moisture conditions, as in our study area, it seems that the representation of important processes must be included, possibly using soil type as a short-cut to represent soil genesis and at the same time minimizing uncertainty (Shaw et al. 2015). Though the estimation of soil C changes is not part of this study, results will be valuable for the design of such a study and indicate where and whether simulated changes can be measured in an experimental layout with the measurement uncertainties that exist.

Potential bias in model input data (litter) and validation data (measured soil C)

Simulated soil C stocks clearly reflect the productivity (litter input) for the plots in accordance with the original data in Table A1. Litter input estimates are based on a chain of sub-models most of which originate from outside Norway (biomass models) and/or are associated with a substantial uncertainty (fx. natural mortality and turnover rates for plant components). The mean annual litter input (for the years 2008-2012) for the standard application of Yasso07 to forests in Norway (whole NFI, ca. 11 000 plots) was 0.221 kg C m$^{-2}$ year$^{-1}$ (Dalsgaard et al. 2016). Annual litter input estimates (kg C m$^{-2}$ year$^{-1}$) as generated from the four most northern Swedish NFI regions were 0.080-0.180 plus an additional 0.030 from ground vegetation in pine forests, and 0.180-0.300 and ca. 0.055 from ground vegetation for spruce forest (Ortiz et al. 2013). For the Swedish study, the two tree species represent approximately equal amounts of total biomass stock. While an exact country average for Sweden cannot be calculated from the information available in the Swedish study, numbers indicate that estimated litter input on
average is higher than for the standard application for Norway. Whereas the Swedish litter estimates are representative for productive coniferous forests (80% of the productive forest in terms of standing biomass), the Norwegian litter estimates include all forest types where 30% of the forest area is unproductive forest, thus lower average litter estimates for Norway as compared to the Swedish study, would be expected. The estimated litter input for Southern Finland was ca. 0.290 kg C m$^{-2}$ year$^{-1}$ (Rantakari et al. 2012). Further, different geographical aggregation levels as well as possible differences in forest age class distribution across countries makes a comparison between studies very difficult. While the numbers given for the three countries seem to be in the same range, the lack of Norwegian biomass and turnover-models, including those for ground vegetation, is a source of uncertainty. Where the soil environment is characterized by frequently changing water levels, high root mortality can be expected (Rewald et al. 2011) and thus turnover rates and root litter input could be different from values measured at well drained sites. We do not have data available to describe such effects on a large scale. Differences in litter input estimates among countries may be due to the applied biomass models and turnover rates (e.g. for foliage and fine roots, where turnover rates are latitude dependent for the Swedish study and fixed, but different, values for the Finnish and Norwegian studies are used). Considerable uncertainty exists in the calculation of annual litter input and this uncertainty is a major cause of uncertainty in the simulation of soil C stocks (Ortiz et al. 2013). In our study, doubling the fine root litter input resulted in a 28% increase of simulated soil C stocks, but this did not change the pattern of an overall large difference between measured and simulated soil C, and a systematically higher difference for some sites (depending on climate, soil type, drainage).

Calculating soil C stocks from measurements is associated with a range of uncertainties and theoretically, a strong bias in the measured stocks could account for the large observed soil C difference. The standard error on profile soil C stock estimates from uncertainty in bulk
density (10% SE), coarse fraction and C content data for the soil survey data depended on soil type and was less than 1.8 kg C m$^{-2}$ for an individual upland mineral soils profile (Strand et al. 2016), thus, not large enough to explain the differences between simulated and measured soil C observed in our study. From a smaller but more complete Nordic forest soil database (Nordsoil, well-drained profiles), it was shown that soil C increases with MAT and MAP (highest soil C stock 22 kg C m$^{-2}$; few observations > 1400 mm MAP) (Callesen et al. 2003). This empirical model (Callesen et al. 2003) was applied to the Norwegian soil survey data used in our study. Excessively to moderately well drained profiles deeper than 0.5 m showed a positive significant relationship ($r^2$=0.69, slope = 0.8). This relationship was not significant when including hydromorphic soils. This indicates that forest soils in Norway have a strong representation of soils with hydromorphic properties for which the formation and the fate of the soil C stock cannot be simulated in any model that conceptually includes only aerobic processes. The empirical model (Callesen et al. 2003) was in this case extrapolated outside its precipitation range, which did not seem to cause a strong bias for the well-drained soils. A preliminary comparison of soil C stocks in regions along the border between Norway and Sweden showed no differences in soil depth and soil C concentrations and thus C stocks (Callesen et al., unpublished). While an average forest soil C stock of ca. 15 kg C m$^{-2}$, as found in our soil survey data, is high compared to measured values at the regional or national scale in Sweden and Finland, similar magnitudes have been reported for other regions of the world. Across Canada forest floor and mineral soil C (to depth of 0.55 m) was 15 kg C m$^{-2}$ or higher (by forest type) for approximately 50% of the 270 plots with a mean of 16.2 kg C m$^{-2}$ in the most dominating forest type (Black spruce (Picea mariana (Mill.)), n=69) (Shaw et al. 2014). Across 22 countries in Europe (not including Norway), a mean estimate for soil C stocks (to one meter depth) based on soil survey data was summarized by soil taxonomical group and for Gleysols, Podzols and Regosols (WRB classification) estimates were 18.2, 10.4
and 8.7 kg C m\(^{-2}\), respectively (De Vos et al. 2015). In Denmark mean forest soil C stocks (to one meter depth) was measured by soil taxonomical group (WRB) and excluding Histosols and soils with aquic moisture regime, they ranged from 9.6 kg C m\(^{-2}\) (Arenosols) to 16.1 kg C m\(^{-2}\) (Cambisols) and by far the most important group (Podzols, 59 of 105 profiles) had a mean C stock of 15.0 kg C m\(^{-2}\) (Callesen et al. 2015). Although clearly uncertainties in the calculations of profile soil C stocks as well as in the sampling intensity (Muukkonen et al. 2009) are present, we do not have any reason to believe that measured soil C is biased to the degree that it would explain the observed discrepancies between measured and simulated soil C stocks.

Representativity of Yasso07 for Norwegian forest soil

The Yasso07 model does not consider the vertical organization of pedological horizons and the differences in physical protection, soil climate and biological activity depending on soil depth. These processes are implicitly accounted for, but only to the degree they were present in the soil data used to calibrate the model. Further, Yasso07 does not specifically represent slow decomposition rates in poorly drained and frequently anoxic conditions i.e. conditions promoting the accumulation of high C stocks particularly in the organic layer of forest soils (Olsson et al. 2009). We argue that due to a high frequency of hydromorphic soils and climatic conditions that enhance humification of A-horizons and podsolization, the Yasso07 model is not able to reproduce measured soil C stocks in the majority of Norwegian forests. Other causes of mismatch between simulations and measurements such as errors in litter input estimates and the lack of model representation of site nutrient dynamics and soil texture (Tupek et al. 2016) may be relevant but were overruled by other factors in our data set. Soil genesis may be quantitatively described by climate, organisms, relief, parent material and time. Carbon models including only production and aerobic mineralization cannot be expected to produce stocks similar to measured stocks unless soil forming processes at the
appropriate timescale (100++ yrs) are addressed and the pedon concept of a layered soil profile is adopted. To avoid the most prominent adverse conditions (anoxia), the model is to be used only on well drained upland sites with no signs of hydromorphic features (e.g. strong humus enrichment in A horizons, pseudogley or gley colours), nor strong humus enrichment in Bh- and Bhs-horizons due to podsolization. Only on sites where such conditions are avoided can the processes represented in the current parameterization of Yasso07 be assumed to make up the most important ones. For the practical purposes in Norway, our results showed that the most adverse conditions (in terms of Yasso07 application) coincide with poor drainage classification and/or a warm and humid climate.

The absence of forest fires is important for developing thick O-horizons (Wardle et al. 1997) and assumptions on historic fire regime has been shown to have a profound effect on simulated soil C stocks (Kelly et al. 2015; Liski et al. 1998). Forest fires in areas of high precipitation, like western Norway, may have been much less frequent than in Sweden and Finland.

When used to estimate short term (annual-decadal) changes, Yasso07 has shown good agreement between simulations and measured soil respiration (Rühr and Eugster 2009) and remaining C in foliage and fine root litter after 10 years and in lying dead trees after 14-21 years (Didion et al. 2014); both studies from mountain forests in Switzerland. While we do not have the necessary data to test the ability of Yasso07 to simulate soil C changes on a country scale, we expect that for conditions where the model parameterization represents the most important soil processes, the change estimates are more likely to be realistic than in condition where other processes dominate.

Conclusions
Simulated soil C stocks were, on average, 8-10 kg C m\(^{-2}\) lower than measured stocks. We found that soil-forming processes related to hydrology (climate, twi, drainage) and deep soil profiles were important in explaining the large differences and that predictors were soil type specific. High MAT and MAP (combined in the model decomposition factor cdf) was related to high model-measurement discrepancy for all soil types. For the largest group of soils, PO-BRe, it was evident that where the model predicts that soil C stocks depend on site productivity, measurements show that site productivity differences are, to a large extent, overruled by effects of an aquic soil moisture regime, i.e. frequent anoxic conditions/poor drainage and podsolization. We argue that, where these processes dominate, Yasso07 cannot be expected to predict measured soil C stocks. While uncertainties in model input and measured soil C stocks are not trivial issues, we argue that they are unlikely to explain the observed differences between measured and simulated soil C stocks. We believe that our study has challenged the Yasso07 model with its application in a wide range of hydrological conditions and contributed with information that is useful for further model development.

**Acknowledgements**

Clara Antón-Fernández produced the maps shown in Figure 1. We wish to acknowledge the Norwegian National Forest Inventory for data access and Rune Eriksen for facilitating data extraction. The research was funded by the Norwegian Research Council (grant no. 210464/O10: “Decision support models for increased harvest and climate-motivated forest policies”). Jari Liski received funding from the Ministry of Environment in Finland (project “Improvement of Yasso soil carbon model for greenhouse gas inventories”).

**References**


Tables

Table 1. Rules used to establish stand development and the historical litter time series used in the pre-simulations 30 years prior to the first permanent NFI sample plot observation.

<table>
<thead>
<tr>
<th>Rule</th>
<th>No. plots</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>Unproductive forest &lt; 30 years old; no harvest; tree species, age and forest biomass constant back in time from values in the first permanent plot inventory.</td>
</tr>
<tr>
<td>2</td>
<td>564</td>
<td>Forest &gt; 30 years old, tree species as in the first permanent plot inventory, no harvest, forest biomass development estimated from NFI&lt;sup&gt;a&lt;/sup&gt;.</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>Previous grassland or cropland, forest biomass development from NFI&lt;sup&gt;a&lt;/sup&gt;, a low constant biomass&lt;sup&gt;b&lt;/sup&gt; prior to stand establishment.</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>Previous grassland, cropland or barren land (n=4) or cases of tree species conversion to spruce in existing forest (n=9). Biomass development from NFI&lt;sup&gt;a&lt;/sup&gt;. For cases of forest conversion a harvest of a mature stand is assumed prior to the establishment of the current stand; for cases of afforestation a low constant biomass&lt;sup&gt;b&lt;/sup&gt; prior to the establishment of the current stand. Info available only for some regions of Norway.</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>As rule 5. Tree species conversion (n=2), afforestation (n=0).</td>
</tr>
<tr>
<td>8</td>
<td>136</td>
<td>Remaining plots, no conversion assumed and harvest of mature stand prior to the establishment of the current stand. Biomass development from NFI&lt;sup&gt;a&lt;/sup&gt;.</td>
</tr>
</tbody>
</table>
a From tables of average biomass change in two consecutive inventories (2000-2004 and 2005-2009) and stratified according to site index, stand age and tree species composition (Antón-Fernández and Astrup 2012).

b Biomass (t ha\(^{-1}\)) for grassland (15.5), cropland (1.4), barren lands (0.4) (Løken et al. 2012).
Table 2. Litter input (kg C m$^{-2}$ year$^{-1}$) used in the spin-up calculation for the initial soil C stocks; given for groups of site index and dominating tree species.

<table>
<thead>
<tr>
<th>SI$^a$</th>
<th>Spruce standard$^b$</th>
<th>Spruce double$^b$</th>
<th>Pine standard</th>
<th>Pine double</th>
<th>Deciduous standard</th>
<th>Deciduous double</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unproductive forest</td>
<td>0.096</td>
<td>0.126</td>
<td>0.157</td>
<td>0.207</td>
<td>0.045</td>
<td>0.056</td>
</tr>
<tr>
<td>6</td>
<td>0.167</td>
<td>0.23</td>
<td>0.184</td>
<td>0.239</td>
<td>0.071</td>
<td>0.084</td>
</tr>
<tr>
<td>8</td>
<td>0.192</td>
<td>0.262</td>
<td>0.195</td>
<td>0.251</td>
<td>0.094</td>
<td>0.109</td>
</tr>
<tr>
<td>11</td>
<td>0.210</td>
<td>0.287</td>
<td>0.219</td>
<td>0.279</td>
<td>0.132</td>
<td>0.152</td>
</tr>
<tr>
<td>14</td>
<td>0.214</td>
<td>0.292</td>
<td>0.228</td>
<td>0.291</td>
<td>0.149</td>
<td>0.175</td>
</tr>
<tr>
<td>17</td>
<td>0.26</td>
<td>0.352</td>
<td>0.238</td>
<td>0.304</td>
<td>0.168</td>
<td>0.198</td>
</tr>
<tr>
<td>20</td>
<td>0.286</td>
<td>0.384</td>
<td>0.215</td>
<td>0.275</td>
<td>0.166</td>
<td>0.196</td>
</tr>
<tr>
<td>23</td>
<td>0.313</td>
<td>0.433</td>
<td>0.255</td>
<td>0.323</td>
<td>0.187</td>
<td>0.238</td>
</tr>
<tr>
<td>26</td>
<td>0.397</td>
<td>0.547</td>
<td>0.255</td>
<td>0.323</td>
<td>0.163</td>
<td>0.187</td>
</tr>
</tbody>
</table>

$^a$ Site index: stand height (m) at age 40 years.

$^b$ “standard” uses fine root turnover rates 0.6 (trees) and 0.33 (dwarf shrubs; herbs and grasses); “double” uses fine root turnover rates 1.2 and 0.66 respectively.
Table 3. Median, minimum and maximum of continuous predictors used in the importance analyses.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>median</th>
<th>min</th>
<th>max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>C:N</td>
<td>29</td>
<td>12</td>
<td>74</td>
<td>635</td>
</tr>
<tr>
<td>twi</td>
<td>6.7</td>
<td>-3.40E+18</td>
<td>14.5</td>
<td>719</td>
</tr>
<tr>
<td>h.a.s.l. (m)</td>
<td>330</td>
<td>2</td>
<td>1005</td>
<td>719</td>
</tr>
<tr>
<td>Incl. (%)</td>
<td>16</td>
<td>0</td>
<td>85</td>
<td>692</td>
</tr>
<tr>
<td>PD (m)</td>
<td>0.55</td>
<td>0</td>
<td>1.30</td>
<td>714</td>
</tr>
<tr>
<td>MAT (°C)</td>
<td>3.0</td>
<td>-2.4</td>
<td>7.4</td>
<td>719</td>
</tr>
<tr>
<td>MAP (mm)</td>
<td>886</td>
<td>355</td>
<td>3133</td>
<td>719</td>
</tr>
</tbody>
</table>

*a*: C:N (C:N ratio), twi (topographic wetness index), h.a.s.l. (height above sea level), Incl. (inclination), MAP (mean annual precipitation), PD (profile depth), MAT (mean annual temperature).
Table 4. Distribution of soil survey profile data to site index (SI), vegetation type, dominant tree species, soil type and drainage class.

<table>
<thead>
<tr>
<th>Tree species</th>
<th>Soiltypee</th>
<th>Most frequent vegetation typea</th>
<th>SIb</th>
<th>Drainagec (frequency)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A-D</td>
</tr>
<tr>
<td>Spruce</td>
<td>BR</td>
<td>MM/RM (14)</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>GL-ORw</td>
<td>MM (17)</td>
<td>11</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>NO-FO</td>
<td>PM (5)</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>PO-BRe</td>
<td>MM (104)</td>
<td>14</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>MM (7)</td>
<td>9.5</td>
<td>11</td>
</tr>
<tr>
<td>Pine</td>
<td>BR</td>
<td>PM (12)</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>GL-ORw</td>
<td>PM (13)</td>
<td>6</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>NO-FO</td>
<td>PM (17)</td>
<td>6</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>PO-BRe</td>
<td>PM (70)</td>
<td>8</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>PM (18)</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>Deciduousc</td>
<td>BR</td>
<td>RD/RM/RW (12)</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>GL-ORw</td>
<td>RM (5)</td>
<td>11</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>NO-FO</td>
<td>MM (3)</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>PO-BRe</td>
<td>MM (21)</td>
<td>11</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>RE</td>
<td>RM (4)</td>
<td>11</td>
<td>2</td>
</tr>
</tbody>
</table>

a First letter denotes nutrient level: P (poor), M (medium), R (rich); second letter denotes moisture level: D (dry), M (medium), W (wet). No. of observations is shown in parenthesis.

b Forest stand height (m) at age 40 years. Given as the median.

c All deciduous tree species are included but birch is the most frequent.

d A (excessively drained) through S (very poorly drained), not classified (n.c.).
PO-BRe: Podzols and Eluviated Dystric Brunisols; BR: Brunisols (excluding Eluviated Dystric); GL-ORw: Gleysols and wet Organic; RE: Regosols; NO-FO: Non-soils and Folisols.
Table 5. Relative importance among continuous predictors included in multiple linear regressions for all plots regardless of soil type and for each soil type separately. Total $r^2$-value is given in brackets. Response variable is the soil C difference to 1 m depth (corrected for bare rock surface). Importance-values give the fraction of explanation for each variable within the total $r^2$. The three highest values of importance are shown in bold. The sign in front of each importance value indicates the slope direction, e.g.: for all soil types, the SOC difference increases with increasing cdf, but decreases with increasing MAP.

<table>
<thead>
<tr>
<th>Variable(^a)</th>
<th>All soil types (0.28)</th>
<th>BR(^b) (0.25)</th>
<th>GL-ORw(^b) (0.18)</th>
<th>PO-BRe(^b) (0.37)</th>
<th>RE(^b) (0.23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cdf</td>
<td>+0.242</td>
<td>+0.122</td>
<td>+0.138</td>
<td>+0.309</td>
<td>+0.118</td>
</tr>
<tr>
<td>C:N</td>
<td>+0.012</td>
<td>+0.006</td>
<td>+0.124</td>
<td>+0.011</td>
<td>+0.022</td>
</tr>
<tr>
<td>twi</td>
<td>+0.038</td>
<td>+0.493</td>
<td>+0.466</td>
<td>+0.005</td>
<td>+0.05</td>
</tr>
<tr>
<td>h.a.s.l.</td>
<td>+0.048</td>
<td>+0.044</td>
<td>+0.038</td>
<td>+0.055</td>
<td>+0.061</td>
</tr>
<tr>
<td>Incl.</td>
<td>+0.011</td>
<td>-0.065</td>
<td>+0.006</td>
<td>+0.021</td>
<td>+0.002</td>
</tr>
<tr>
<td>MAP</td>
<td>-0.165</td>
<td>+0.183</td>
<td>-0.121</td>
<td>-0.213</td>
<td>-0.092</td>
</tr>
<tr>
<td>PD</td>
<td>+0.394</td>
<td>+0.051</td>
<td>+0.016</td>
<td>+0.26</td>
<td>+0.569</td>
</tr>
<tr>
<td>MAT</td>
<td>-0.089</td>
<td>+0.035</td>
<td>-0.091</td>
<td>-0.126</td>
<td>+0.086</td>
</tr>
</tbody>
</table>

\(^a\) cdf (climate dependence factor), C:N (C:N ratio), twi (topographic wetness index), h.a.s.l. (height above sea level), Incl. (inclination), MAP (mean annual precipitation), PD (profile depth), MAT (mean annual temperature).

\(^b\) PO-BRe: Podzols and Eluviated Dystric Brunisols; BR: Brunisols (excluding Eluviated Dystric); GL-ORw: Gleysols and wet Organic; RE: Regosols; NO-FO: Non-soils and Folisols.
Table 6. Analysis of variance for each soil type. Continuous predictors: PD = profile depth, cdf = climate dependence factor, twi = topographic wetness index, MAP = mean annual precipitation. Class variables: D = drainage class, SI = site index, VT = vegetation type, p-values for significant effects (p ≤ 0.05) are shown. PO-BRe: Podzols and Eluviated Dystric Brunisols; BR: Brunisols (excluding Eluviated Dystric); GL-ORw: Gleysols and wet Organic; RE: Regosols. Response variable is the soil C difference to 1 m depth (corrected for bare rock surface).

<table>
<thead>
<tr>
<th></th>
<th>PO-BRe</th>
<th>BR</th>
<th>GL-ORw</th>
<th>RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>cdf</td>
<td>2.00E-16</td>
<td>twi</td>
<td>0.04</td>
<td>cdf</td>
</tr>
<tr>
<td>PD</td>
<td>2.75E-14</td>
<td>MAP</td>
<td>0.04</td>
<td>twi x D x VT</td>
</tr>
<tr>
<td>SI</td>
<td>1.22E-07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cdf x PD</td>
<td>1.05E-05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>0.000407</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cdf x PD x VT</td>
<td>0.020559</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PD x D</td>
<td>0.040478</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure captions

Figure 1. Maps showing the magnitude of the difference between measured and simulated soil C stocks (kg C m$^{-2}$, left), mean annual temperature (centre) and mean annual precipitation (right). Colored areas in the left map also shows forest area as opposed to unforested areas (white).

Figure 2. Soil C stocks to 1 m depth from soil survey data (measured) and from Yasso07 simulations (standard simulation or double fine root litter; see text). Estimates are presented for classes of mean annual precipitation (left) and mean annual temperature (right). Midpoint of box show the median values, box height extends to upper and lower quartiles, whiskers extent to 1.5 times height of box.

Figure 3. Frequency distribution of simulated (standard) and measured soil C stock estimates (corrected for bare rock).

Figure 4. Soil C difference (measured – simulated C stocks, kg C m$^{-2}$) with standard as well as double fine root litter input. Differences are presented for class variables: drainage class (left), soil type (center), vegetation type (right). Soil types are: NO-FO (Non-soil, Folisols), RE (Regosols), BR (Brunisols excluding Dystric Eluviated), PO-BRe (Podzols and Dystric Eluviated Brunisols), GL-ORw (Gleysols and wet Organic soils). Vegetation types are indicated with the first letter denoting nutrient level: P (poor), M (medium), R (rich) and the second letter denoting moisture level: D (dry), M (medium), W (wet). NV indicates no
vegetation type was assigned in the survey. Drainage is defined as A (excessively drained) through to S (very poorly drained). Midpoint of box show the median values, box height extends to upper and lower quartiles, whiskers extent to 1.5 times height of box.
Appendix 1

Table A1. Paired estimates (mean, minimum and maximum) of soil C stocks (organic layer plus topmost one meter mineral soil layer) and litter input (trees plus understory vegetation) used in the parameterization of Yasso07 (Liski and Westman, 1995). Given for different site types (productivity classes) and forest tree species.

<table>
<thead>
<tr>
<th>Site</th>
<th>Soil C stocks (kg C m(^{-2}))</th>
<th>Litter input (kg C m(^{-2}) year(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>min</td>
</tr>
<tr>
<td>CT_SP</td>
<td>5.8</td>
<td>4.0</td>
</tr>
<tr>
<td>VT_SP</td>
<td>5.7</td>
<td>4.6</td>
</tr>
<tr>
<td>VT_NS</td>
<td>6.9</td>
<td>6.2</td>
</tr>
<tr>
<td>MT_SP</td>
<td>6.9</td>
<td>5.6</td>
</tr>
<tr>
<td>MT_NS</td>
<td>8.6</td>
<td>7.5</td>
</tr>
<tr>
<td>OMT_NS</td>
<td>9.6</td>
<td>6.9</td>
</tr>
</tbody>
</table>

\(a\) CT: Calluna type, VT: Vaccinium type, MT: Myrtilius type, OMT: Oxalis-Myrtilius type, SP: Scots Pine, NS: Norway Spruce.

The litter production of trees (Table A1) was calculated as described in Liski et al. (2005).

The litter production of understory vegetation (Table A1) was calculated from coverage data using a similar approach as in Ortiz et al. (2013), biomass models of Muukkonen et al. (2006) and biomass turnover rates of Peltoniemi et al. (2004).


Soil C stocks (kg C m\(^{-2}\))

Density

Simulated

Measured

Soil C stocks (kg C m\(^{-2}\))
Soil C difference (measured − simulated soil C stock, kg C m\(^{-2}\)).

Drainage: NO−FO, RE, BR, PO−BR, GL−ORw.

Soiltype: PD, MD, RD, PM, MM, RM, PW, MW, RW, NV.

Vegetation type: standard, double.

Page 47 of 58
Supplement

Figure legends:

**Figure S1.** Measured and simulated soil C (carbon) for all plots as in Table 4 (n=691 due to missing observations in one or more of the plotting variables). “Standard” and “double” refer to the two simulations run (differ in fine root turnover rates as detailed in the article). Drainage classes are: vg-g (very good - good), g-m (good - medium), m-p (medium - poor) and p-vp (poor – very poor). Midpoint of box show the median values, box height extends to upper and lower quartiles, whiskers extent to 1.5 times height of box.

**Figure S2.** Measured and simulated soil C (carbon) for all “Podzol-eluvbrunisol” plots and distributed to the three main forest tree species (tree species dominant for each plot, n=422). “Standard” and “double” refer to the two simulations run (differ in fine root turnover rates as detailed in the article). Drainage classes are: vg-g (very good - good), g-m (good - medium), m-p (medium - poor) and p-vp (poor – very poor). Midpoint of box show the median values, box height extends to upper and lower quartiles, whiskers extent to 1.5 times height of box.

**Figure S3.** Measured and simulated soil C (carbon) for all “Podzol-eluvbrunisol” plots and distributed to each drainage class and site index (stand height (m) at age 40 years). Site index 4 is not defined in the NFI but denotes unproductive forest. Data as in Figure S2. “Standard” and “double” refer to the two simulations run (differ in fine root turnover rates as detailed in the article). Drainage classes are: vg-g (very good - good), g-m (good - medium), m-p (medium - poor) and p-vp (poor – very poor). Midpoint of box show the median values, box height extends to upper and lower quartiles, whiskers extent to 1.5 times height of box.

**Figure S4.** Measured and simulated soil C (carbon) for all “Podzol-eluvbrunisol” plots and distributed to each drainage class and profile depth (cm). Data as in Figure S2. “Standard” and “double” refer to the two simulations run (differ in fine root turnover rates as detailed in the article). Drainage classes are: vg-g (very good - good), g-m (good - medium), m-p (medium - poor) and p-vp (poor – very poor). Midpoint of box show the median values, box height extends to upper and lower quartiles, whiskers extent to 1.5 times height of box.

**Figure S5.** Measured and simulated soil C (carbon) for all “Podzol-eluvbrunisol” plots and distributed to classes of profile depth (cm) and vegetation type. Data as in Figure S2. “Standard” and “double” refer to the two simulations run (differ in fine root turnover rates as detailed in the article). Nine vegetation types where first term (Poor, Medium, Rich) describes nutrient level and second term (Dry, Medium, Wet) describes moisture level. One class where no registration was available (NullVeg). Midpoint of box show the median values, box height extends to upper and lower quartiles, whiskers extent to 1.5 times height of box.
**Figure S6.** Measured and simulated soil C (carbon) for all soil types. Data included for plots with MAT < 2 °C and MAP < 800 mm (n=178). “Standard” and “double” refer to the two simulations run (differ in fine root turnover rates as detailed in the article). Drainage classes are: vg-g (very good - good), g-m (good - medium), m-p (medium - poor) and p-vp (poor – very poor). Midpoint of box show the median values, box height extends to upper and lower quartiles, whiskers extent to 1.5 times height of box.

**Figure S7.** Measured and simulated soil C (carbon) for “Podzol-eluvbrunisol” plots with MAT < 2 °C and MAP < 800 mm and distributed to the three main forest tree species (tree species dominant for each plot, n=109). “Standard” and “double” refer to the two simulations run (differ in fine root litter turnover rates as detailed in the article). Drainage classes are: vg-g (very good - good), g-m (good - medium), m-p (medium - poor) and p-vp (poor – very poor). Midpoint of box show the median values, box height extends to upper and lower quartiles, whiskers extent to 1.5 times height of box.

**Figure S8.** Measured and simulated soil C (carbon) for “Podzol-eluvbrunisol” plots with MAT < 2 °C and MAP < 800 mm and distributed to each drainage class and site index (stand height (m) at age 40 years, n=109). Site index 4 is not defined in the NFI but denotes unproductive forest. “Standard” and “double” refer to the two simulations run (differ in fine root turnover rates as detailed in the article). Drainage classes are: vg-g (very good - good), g-m (good - medium), m-p (medium - poor) and p-vp (poor – very poor). Midpoint of box show the median values, box height extends to upper and lower quartiles, whiskers extent to 1.5 times height of box.

**Figure S9.** Measured and simulated soil C (carbon) for “Podzol-eluvbrunisol” plots with MAT < 2 °C and MAP < 800 mm and distributed to classes of profile depth (cm) and vegetation type. Data as in Figures S7and S8. “Standard” and “double” refer to the two simulations run (differ in fine root turnover rates as detailed in the article). Nine vegetation types (not all represented in this subset) where first term (Poor, Medium, Rich) describes nutrient level and second term (Dry, Medium, Wet) describes moisture level. Midpoint of box show the median values, box height extends to upper and lower quartiles, whiskers extent to 1.5 times height of box.
Figure S1: All soil types, all profiles

Soil C (kg C m\(^{-2}\))

- brunisol
- gleysoil–wetorganic
- nonsoil–folisol
- podzol–eluvbrunisol
- regosol

Drainage, aggregated classes

- measured
- standard sim
- double sim
Figure S2: Podzol, eluviated Brunisol

Drainage, aggregated classes

Soil C (kg C m$^{-2}$)

Norway Spruce

Scots Pine

Deciduous

measured  
standard sim  
double sim

vg−g  g−m  m−p  p−vp  vg−g  g−m  m−p  p−vp  vg−g  g−m  m−p  p−vp

https://mc06.manuscriptcentral.com/cjfr-pubs
Figure S3: Podzol, eluviated Brunisol

Soil C (kg C m$^{-2}$)

Site index

<table>
<thead>
<tr>
<th>vg-g</th>
<th>g-m</th>
<th>m-p</th>
<th>p- vp</th>
</tr>
</thead>
</table>

measured
standard sim
double sim

https://mc06.manuscriptcentral.com/cjfr-pubs
Figure S4: Podzol, eluviated Brunisol Profile depth (cm)

<table>
<thead>
<tr>
<th>Depth Range (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0−30</td>
</tr>
<tr>
<td>30−60</td>
</tr>
<tr>
<td>60−90</td>
</tr>
<tr>
<td>90−130</td>
</tr>
</tbody>
</table>

Soil C (kg C m$^{-2}$)

- vg−g
- g−m
- m−p
- p−vp

Measured
Standard sim
Double sim
Figure S5: Podzol, eluviated Brunisol

Vegetation type

Soil C (kg C m$^{-2}$)

<table>
<thead>
<tr>
<th></th>
<th>0−30</th>
<th>30−60</th>
<th>60−90</th>
<th>90−130</th>
</tr>
</thead>
<tbody>
<tr>
<td>0−30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PoorDry</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>PoorMedium</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>PoorWet</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>MediumDry</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>MediumMedium</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>MediumWet</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>RichDry</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>RichMedium</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>RichWet</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>NullVeg</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

measured
standard sim
double sim

https://mc06.manuscriptcentral.com/cjfr-pubs

Canadian Journal of Forest Research
Figure S6: All soil types, T < 2 (°C), P < 800 (mm)

<table>
<thead>
<tr>
<th>Soil Type</th>
<th>Drainage, Aggregated Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>brunisol</td>
<td>gleysoil–wetorganic</td>
</tr>
<tr>
<td>nonsoil–folisol</td>
<td>podzol–eluvbrunisol</td>
</tr>
<tr>
<td>regosol</td>
<td></td>
</tr>
</tbody>
</table>

Soil C (kg C m⁻²)

- Measured
- Standard sim
- Double sim

Drainage classes: vg, g, m, p, vp
Figure S7: Podzol, eluviated Brunisol. T < 2 (°C), P < 800 (mm)
Figure S8: Podzol, eluviated Brunisol, T < 2 (°C), P < 800 (mm)
Figure S9: Podzols eluviated Brunisol, T < 2 (°C), P < 800 (mm)