Can SOC modelling be improved by accounting for pedogenesis?

Peter Finke, Emmanuel Opolot, Jérôme Balesdent, Asmeret Asefaw Berhe, Pascal Boeckx, Sophie S. Cornu, Jennifer Harden, Christine Hatté, Elizabeth Williams, Sebastian Doetterl

Title

Can SOC modelling be improved by accounting for pedogenesis and weathering in a chronosequence?

Authors

Peter Finke[1,*], Emmanuel Opolot[1,2], Jerôme Balesdent[3], Asmeret Asefaw Berhe[4], Pascal Boeckx[5], Sophie Cornu[3], Jennifer Harden[6], Christine Hatté[3], Elizabeth Williams[4], Sebastian Doettrl[7]

Abstract

Recent research results suggest that soil organic carbon mineralization and stabilization depend to a substantial degree on the soil geochemistry and the degree of weathering. We hypothesized that this dependence can be translated to decay rate modifiers in a model context, and used data from the Merced chronosequence (CA, U.S.A., 100 yr - 3 Myr), representing a weathering sequence, to test, on a 1000 year time scale, a simple soil organic carbon (SOC) model based on the RothC26.3 model concepts. Four information levels were identified: (1) known decay rates per model SOC pool at individual chronosequence locations, obtained by calibrating the model to measured SOC-fractions and measured fresh OC-inputs; (2) average decay rates per SOC-pool, corrected per location with rate modifiers based on geochemical proxies; (3) uncorrected average decay rates per SOC-pool; (4) uncorrected average decay rates per SOC-pool and averaged OC-inputs. A lumped root mean square error (RMSE) statistic was calculated per information level. We found that using local measurements of fresh OC-input led to a decrease in RMSE of near 15% relative to information level (4). Applying geochemical rate modifiers led to a further reduction of 20%. Thus, we conclude that

[1] Department of Soil Management, Ghent University, B-9000 Ghent Belgium; [2] Department of Agricultural Production, College of Agricultural and Environmental Sciences, Makerere University, Kampala, Uganda; [3] Aix-Marseille Univ, CNRS, INRA, IRD, Coll de France, CEREGE, 13545 Aix-en-Provence, France; [4] University of California, CA 95343 Merced U.S.A. [5] Department of Applied Analytical and Physical Chemistry, Ghent University, B-9000 Ghent Belgium; [6] U.S. Geological Survey, CA 94025 Menlo Park U.S.A. [7] Institut für Geographie, Universität Augsburg, 86135 Augsburg Germany.

there is a benefit of including geochemical rate modifiers in this SOC-model. We repeated this analysis for a five-pool and a four-pool SOC model that either included or excluded an inert organic matter pool. In terms of the lumped RMSE both models performed similarly, but by comparing measured and simulated percentage Modern Carbon (pMC) for bulk SOC we concluded that measured pMC was best approximated using a four-pool SOC model (without an Inert Organic Matter pool), and that it is likely that a five-pool model including a very slowly decaying pool would further improve model performance.

Keywords: Soil Organic Carbon; Modelling; Pedogenesis; Weathering; Chronosequence

1. INTRODUCTION

Most earth system models (e.g., Goosse et al., 2010; Kaplan et al., 2011) consider both the input of plant litter in the soil and the decay of Soil Organic Carbon (SOC) as functions of climate, vegetation and land use. It was found (Doetterl et al. 2015; Lawrence et al. 2015; Mathieu et al., 2015) that, under similar vegetation, soil organic carbon mineralization and stabilization is mostly dependent on the soil geochemistry and the degree of weathering (i.e., reactive surface area, texture, Total Reserve of Bases, Si, Fe and Al contents, etc.). However, it is not known how these soil age related attributes quantitatively relate to the age structure (age per type of SOC) of soil organic matter. According to these studies, climate is just one of the drivers. Climate, vegetation as well as geochemical soil composition, resulting from parent material as well as climatic history, may interactively affect the soil microbiological community composition, which are the actual SOC decomposers (e.g. Doetterl et al., 2018). For these experimental findings to be applicable, the effects of mineral weathering and resulting geochemical composition and clay mineralogy on SOC decomposition need to be translated into a model context. This necessarily involves the identification of influential geochemical variables and linking these to model parameters. This study is an attempt to do so in a chronosequence with known gradients in weathering.

Working hypothesis is that decomposition rates of SOC-pools are different along a chronosequence because of different degrees of weathering of primary minerals and formation of secondary mineral products along this chronosequence, and that these differences can be partly explained by quantitative expressions of weathering and soil geochemistry, and partly by site-specific C-input to the soil.

We test this hypothesis on soils from a chronosequence of soils developed on alluvial sediments near Merced, California (Harden 1982, 1987) and use the SoilGen implementation (Finke and Hutson, 2008) of the RothC-concept (Jenkinson and Coleman, 1994; Coleman and Jenkinson, 2005). Essentially, RothC dynamically re-allocates organic matter over different pools having unique decomposition rates (c.f. section 2.2). The implementation in SoilGen follows the RothC-concept, but is discretized over soil depth. RothC was developed to describe SOC-evolution in agricultural field trials covering ca. 150 years and in young parent materials, thus the development domain (time, vegetation and parent materials) is not the same as the application domain in this study. As the temporal range of this chronosequence is far greater than the range for which RothC-pools were defined, the Inert Organic Matter (IOM) pool used in the RothC-model may not be truly "inert". Radiocarbon studies (Sanderman et al., 2016) have shown that measured ages of soil carbon do not support the presence of truly "inert" organic matter. More generally, the mere existence of Inert Organic Matter has been questioned by the biogeochemistry community (Schmidt et al., 2011). To allow evaluation of the usefulness of IOM as a model pool, we use two versions of this model: with and without IOM. On the Merced chronosequence, SOC-pools have been measured by Doettrl et al. (2018) and mapped onto the pools of C identified in RothC: Resistant Plant Material (RPM), Decomposable Plant Material (DPM), Biomass (BIO), Humus (HUM) and Inert Organic Matter (IOM); described below.

The working hypothesis is tested in a four-step approach:

1) we apply parameter estimation onto the SOC-model, containing RPM, DPM, BIO, HUM and IOM-pools, at 5 sites of varying age along the chronosequence to obtain site-specific rate

coefficients. The SOC-model is run for 1000 years on each site to obtain independence of the result on initial pool values;

2) we correlate these rates to soil characteristics at these sites to find good candidate ("influencing") soil properties to be used as rate-modifiers;
3) we predict the rate-modifiers with linear regression from the measured soil characteristics at the sites. Linear regression may not capture all possible relationships, but as the data set is small, exploring non-linear methods was not feasible. The predicted rate-modifiers are applied to SoilGen and by comparing simulated to measured pools in extra 1000-year runs, with all other inputs equal to step 1, we evaluate how well the measured DPM, RPM, BIO and HUM pools can be reproduced. This tells us how well rate modifiers perform when local measurements of soil variables are available;
4) we evaluate the effect of having site-specific (and likely soil-dependent) organic matter C-inputs to simulation quality.

As an independent check, we compare measured and calculated radiocarbon (percentage modern carbon) of the calibrated SOC-model for bulk SOC, using the site-specific rate modifiers. This is done for longer time periods (several millenniums) to include the potential effect of old SOC on pMC for those sites in older parts of the chronosequence.

2. MATERIALS AND METHODS

2.1. Merced Chronosequence data summary

Soil data were partly collected in 2013, partly taken from previous studies (Harden, 1982, 1987) on a chronosequence of five terrace levels near Merced, California (Table 1) with BSk-climate according to the Köppen-Geiger classification, with average annual temperature of 16.3 °C and 315 mm annual rainfall. The chronosequence spans 0.1 to 3000 ka, where age constraints and associated uncertainties were obtained by stratigraphic correlation, fossils of Hemphillian fauna, ^{14}C, U-series and K-Ar dating and the occurrence of the Gilbert-Gauss magnetic reversal (Marchand and Allwardt,

1981; Harden, 1987: p. A8). Parent materials are felsic, magmatic and metamorphic alluvial sediments with fine sandy and silty texture, which were field described and analyzed for basic soil properties (texture, color, structure, consistence) allowing calculation of the profile development index (Harden, 1982; Supplementary Information: Table S1). Only the oldest site has substantial amounts of quartz-rich pebbles. Age differences correspond to different weathering intensities, as shown by element ratios, pedogenic iron contents and type of clay minerals (Harden, 1987; Doetterl et al., 2018; Supplementary Information: Table S2). Over the chosen simulation time of 1000 years, we consider the sites to be geomorphologically stable: The sites are on level terrain and surface runoff and water erosion are negligible. Wind erosion is also negligible due to the vegetation cover. Perched water tables of short duration may occur at site PM24II. Over longer time spans this geomorphological stability may not be the case, however. Data used for simulation of soil formation in general and the evolution of soil organic carbon (SOC) are:

(i) for model initialization: depth patterns of texture fractions, bulk density, (clay) mineralogy, Cation Exchange Capacity (CEC), root density distribution;

(ii) for model calibration and verification: depth patterns of SOC and its fractions (see below) and percentage Modern Carbon (pMC). The pMC is defined as $pMC = 100 * \frac{A_{SN}}{A_{ON}}$, where A_{SN} is the specific activity (decay counts per minute) of the sample normalized to δ^{13}C, and A_{ON} is the specific activity of the oxalic acid standard normalized to δ^{13}C;

(iii) to satisfy model boundary conditions: evolution of precipitation, potential evaporation, temperature, litter-C input, bioturbation.

Data from categories (i) and (ii) are taken from Harden (1987) and Doetterl et al. (2018). In category (iii), weather data were taken from http://www.usclimatedata.com for Merced and the year 2010 (precipitation, daily minimum and maximum temperature) and the daily potential evaporation was calculated using the Hargreaves equation (Hargreaves and Samani, 1985). Average rain water composition for 2010 was taken from the National Atmospheric Deposition Program (http://nadp.sws.uiuc.edu/) for Yosemite National Park-Hodgdon Meadow (site CA99). For the

simulation period (1000 years), we assumed a constant climate, constant vegetation type and constant litter-C input. This model spin-up period and the stable model boundary inputs have the function to obtain independency of the size of the SOC-pools at the end of the simulation from the (unknown) initial situation.

To obtain C-input, first the aboveground net primary production (ANPP) was estimated using the Normalized Difference Vegetation Index from 2010-2015 derived from the MOD13Q1 Modis (TERRA) vegetation index from NASA. Carbon biomass was then derived from reference values from literature for grassland in the study region (Berhe, 2012). The ANPP-C was used to estimate the belowground litter production by a standard value for grassland (86% of litter production belowground) as in Kononova (1975). This root litter production was then distributed over the root profile using the root density fractions per depth interval, while the aboveground litter production contributes to the ectorganic layer.

Bioturbation, the soil mixing activity by soil meso- and macrofauna or treefalls, is of importance because of the associated vertical transport of soil matter. It was shown to be an important factor in soil horizonation in both field- and simulation studies (e.g. Phillips, 2007; Finke, 2012). Bioturbation also creates macro-pores, which influence CO_2-transport (Singer et al., 2001). No bioturbation data is available for the study area. However, Wilkinson et al. (2009) reported values between 10 -50 Mg ha^{-1} y^{-1} and Gobat (2004) reported values in the range of 13 to 75 Mg ha^{-1} y^{-1} for Kansas (Tallgrass Prairie, temperature between 6-19 °C). We decided to take 30 Mg ha^{-1} y^{-1} of bioturbation for all chronosequence soils, equally distributed over the root zone compartments and assumed bioturbation independent of soil fertility levels as major bioturbators in the area are mammals (ground squirrels and pocket gophers) that do not consume soil materials.

At the sites, from a mass perspective, dust input is minor over the simulation period of 1000 years. There are some indications (Aciego et al., 2017) that Phosphorus (P) inputs by dust may be important. P-input would affect plant biomass production, but as litter production was a simulation input based on measured ANPP, this fertilization effect by P was implicitly taken into account.

Table 1- General data for Merced chronosequence (Harden, 1982, 1987), WRB2014 classification by the authors.

Site name and code	WGS84 N / WGS84 W	Altitude (m)	Sampled depths (cm) topsoil / subsoil	Age and uncertainty range (ky)	WRB, 2014	USDA, 1975 (great group)	Vegetation
Post Modesto PM24II	37.62072 / -120.61555	57	0-9 / 19-30	0.1 [0-1]	Eutric Gleyic Fluvisols	Typic Xerorthents	Grassland with scattered oak (Quercus lobata)
Post Modesto PM22	37.48854 / -120.52946	56	0-13 / 13-35	3 [1-8.3]	Eutric Cambisols		
Modesto M1	37.52861 / -120.40640	90	0-5 / 22-53	35 [20-70]	Haplic Luvisols	Typic Haploxeralfs	
Riverbank R	37.52449 / -120.45434	94	0-12 / 12-39	300 [250-570]	Chromic Luvisols		
China Hat CH	37.46767 / -120.36948	224	2-12 / 12-30	3000 [730-4000]	Rhodic Luvisols	Typic Palexeralfs	

SOC fractions were measured in the bulk soil in topsoil and subsoil with the purpose to comply with the four conceptual fractions proposed by Stewart et al. (2008): (i) free particulate organic matter [POM], considered as unprotected; (ii) microaggregate-associated [S+A] SOC, considered as physically protected; (iii) silt- and clay-associated [s+c] SOC, considered as mineral associated, geochemically protected and (iv) nonhydrolyzable [rSOC] SOC equivalent to biochemically protected. The chronosequence soils do not contain carbonates. Physical separation (sieving) techniques (Doetterl et al., 2015) resulted in estimates of particulate organic matter (POM; >250 µm), C associated to stable microaggregates (S+A; 53-250 µm) and C associated to non-aggregated clay and silt (s+c; <53 µm). Additionally, a <63 µm fraction was ultrasonically dispersed (energy level of 22 J/mL) from a separate bulk sample and the dispersed solution was then wet sieved over a 63 µm sieve. The suspension < 63 µm was filtered through a 0.45 µm nylon mesh and material >0.45 µm was dried at 40 °C and weighted. This fraction was hydrolyzed with 25mL 6N HCl at room

168 temperature for 24 h. The hydrolysis residue after washing was used as estimate of resistant organic
169 carbon (rSOC). The amount of Dissolved Organic Carbon (DOC) was considered negligible under the
170 semi-arid climate at Merced and was not measured. The measured SOC fractions were converted to
171 the C-pools (DPM+RPM), (BIO+HUM) and IOM using the protocol of Zimmermann et al. (2007), c.f.
172 Figure 1. Percentage Modern Carbon (pMC) was measured for topsoil and subsoil in bulk SOC and in
173 the fractions from the physical fractionation. The water used during the fractionation might contain
174 some dissolved organic matter. This would be measured as part of the smallest fraction (s+c; Figure
175 1). DPM/RPM and BIO/HUM ratios for temperate grassland under equilibrium conditions
176 (Zimmerman et al., 2007: table 1: DPM/RPM=0.1271 and BIO/HUM=0.0259) were used to split the
177 pools (DPM+RPM) and (BIO+HUM) obtained from measurements into the RothC-pools.
178 For a more comprehensive description of the collected data we refer to Harden (1987) and Doetterl
179 et al. (2018); for a data summary see Supplementary Information: Table S3.
180

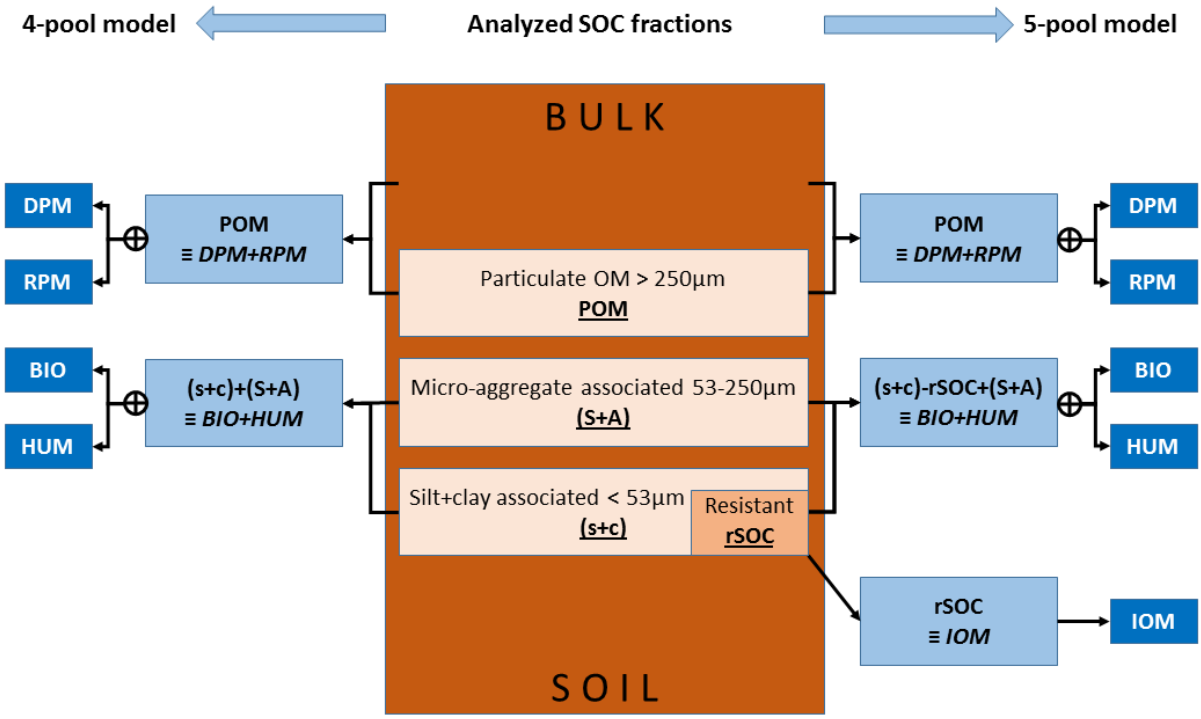

Figure 1- Analyzed SOC fractions (Stewart et al., 2008; Doetterl et al., 2015) and conversion to four-pool and five-pool RothC-models. ⊕ indicates distribution over RothC-pools according to Zimmerman et al. (2007) for temperate grassland. RothC-pools are DPM=Decomposable Plant Material; RPM=Resistant Plant Material; BIO=Biomass; HUM=Humus and IOM=Inert Organic Matter. Analytical SOC-fractions are POM=particulate organic Matter; (S+A)= C associated to stable microaggregates (53-250 µm); (s+c)= C associated to non-aggregated clay and silt (<53 µm); and rSOC= resistant organic carbon measured after cold acid hydrolysis (<63 µm).

2.2. Soil model

The effect of mineralogy, climate and vegetation on the depth distribution of SOC over time was simulated with the SoilGen2.15 model (Finke and Hutson, 2008; Opolot and Finke, 2015). SoilGen2.15 simulates flows of water, heat, solutes and CO_2 in unconsolidated geomaterials by numerically solving partial differential equations (the Richards equation, heat flow equation, advection-dispersion equation and CO_2-diffusion equation respectively), where the column is vertically discretized in equal compartments of five cm thickness in this study. Additionally, transport of solid matter occurs in the form of clay migration (a leaching process) and as a consequence of

bioturbation (a mixing process). Besides the flow of matter, for each soil compartment various sink and source terms are calculated: (i) Soil texture is modified by physical weathering driven by temperature fluctuations; (ii) soil mineralogical composition is modified by chemical weathering of 15 predefined and two user-defined silicate minerals; (iii) equilibriums of calcite and gypsum control the precipitation or dissolution of these salts; (iv) amounts of SOC in the pools RPM, DPM, BIO and HUM are dynamically calculated (see below). These combined processesmimic soil formation (Figure 2). The time step for water flow calculations is less than one hour (depending on the rainfall dynamics), for heat flow and physical weathering it is one hour, for solute transport and dissolution chemistry it is less than one day (depending on water flow dynamics) and for the SOC-cycle and chemical weathering it is one day. Bioturbation is calculated for annual time steps. These time steps are matching the dynamics of the individual processes, so that highly dynamic processes (e.g. water flow) as well as relatively slow processes (e.g. weathering of minerals) are simulated efficiently. Several studies confronted outputs of the SoilGen model in multi-millennium simulations to measurements: Finke (2012), Sauer et al. (2012), Yu et al. (2013) Zwertvaegher et al. (2013), Finke et al. (2015) and Keyvanshokouhi et al. (2016) quantified model accuracy at more than 100 locations for SOC, texture, Cation Exchange Capacity, Calcite content, Base Saturation and pH, and concluded fair to good performance. Keyvanshokouhi et al. (2016) concluded that the SoilGen model is suitable for global change effect studies on soils.

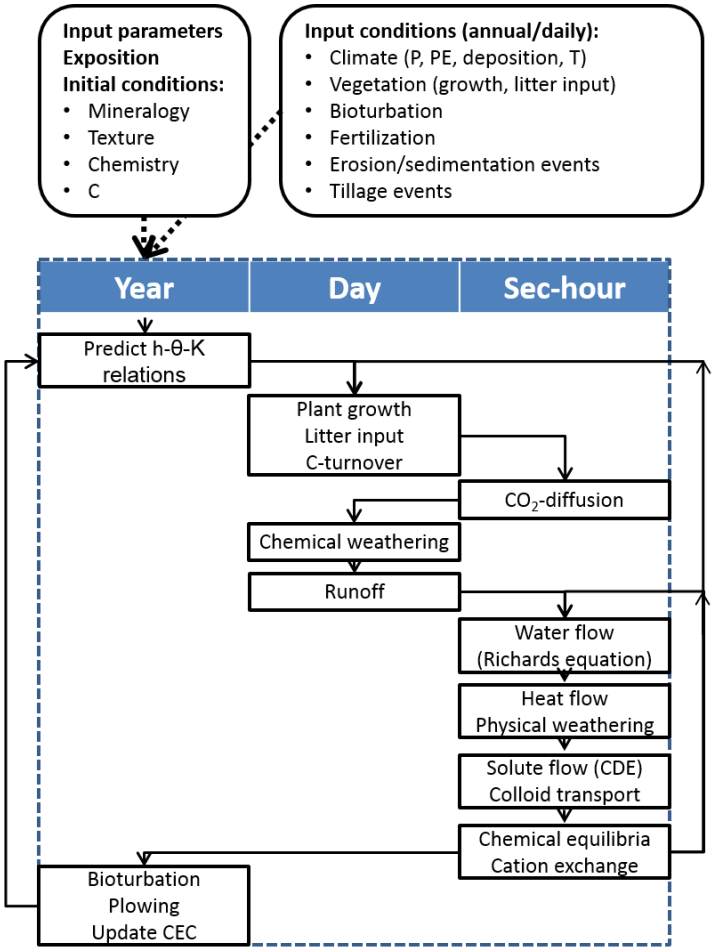

Figure 2- Simplified process flowchart of SoilGen model. Copied from Minasny et al. (2015). P=precipitation, PE=evaporation, T=temperature, CDE=convection diffusion equation, CEC=cation exchange capacity, h-θ-K relations describe the relations between soil water pressure head, soil water content and hydraulic conductivity.

For the SOC-cycle, the concepts and pools of the RothC26.3 model (Jenkinson and Coleman, 1994; Coleman and Jenkinson, 2005) are implemented for each one of the vertical compartments (in this case, of five cm thickness) and using daily timesteps:

- Litter arrives at the soil surface as leaf litter or in belowground compartments as root litter. SoilGen uses the simulated root mass density depth pattern (an exponential function declining with depth) to distribute the total root litter over the compartments. The total amount of root and leaf litter is user input per vegetation type.

- All incoming litter is divided by a fixed DPM/RPM ratio (0.67; Jenkinson and Coleman, 1994) into Resistant Plant Material (RPM) and Decomposable Plant Material (DPM) pools. Both these pools decay with rates k_{RPM} and k_{DPM}
- The resulting decay products from RPM and DPM are split over the Humus (HUM), Biomass (BIO) and mineralized (CO_2) pools using the clay content and a fixed BIO/HUM ratio (46/54; Jenkinson and Coleman, 1994).
- HUM and BIO decay with rates k_{HUM} and k_{BIO} into HUM, BIO and CO_2. This cycling of SOC mimics a food web.
- All rates are modified by factors calculated using temperature and moisture deficit in each soil compartment, as in Jenkinson and Coleman (1994). Both temperature and moisture content are simulated at sub-daily timesteps. The moisture deficit is distributed over the root compartments, using the air-filled porosity per compartment as a proxy for the relative water stress. We assume no other depth dependencies of rates besides those related to differences in temperature and moisture deficit, even though there are indications that these do exist (see Mathieu et al., 2015, but also see Solly et al., 2015 with a contrasting conclusion for Beech forest soils). To mimic effect of the rhizosphere on abundance of micro-organisms, no SOC-decay is assumed to occur below the root zone. SOC-pools may be redistributed over depth as a consequence of bioturbation.
- The Inert Organic Matter pool (IOM) accounts for stable (non-decaying) organic matter, present at the start of the simulation period and inert during the period covered by the simulation.

All SOC-pools in SoilGen contain C, but also the cations and anions taken up by the vegetation via the transpiration stream migrate through the pools, thus finally not only CO_2 but also these ions are released (in the soil solution). Parallel to the C-cycle, a ^{14}C cycle is simulated. Litter inputs use the $\Delta^{14}C$ of the atmosphere at the year of input t (Hua et al., 2013; Reimer et al., 2013) to construct an extra ^{14}C-related input pool of new litter by $LitterC_t * (1+ \Delta^{14}C_t/1000)$, where $LitterC_t$ is the litter-C input in year t. The ^{14}C follows the same pathway as C, thus the ^{14}C pools degrade as the SOC-pools,

but additionally radioactive decay takes place, which results in dynamic pools representing radiocarbon coined as RPMx, DPMx, BIOx, HUMx and the static pool IOMx. This allows the calculation of $\Delta^{14}C$ of total SOC by:

$$\Delta^{14}C = \left(\frac{RMPx+DPMx+BIOx+HUMx+IOMx}{RPM+DPM+BIO+HUM+IOM}-1\right) * 1000 \quad \text{(eq.1)}$$

in any subsequent year. For comparison to measurements, pMC is calculated by:

$$pMC = 100 * ((\Delta^{14}C/1000) + 1) * e^{((y-1950)/(5730/LN(2)))} \quad \text{(eq.2)},$$

where y is the year of sampling.

We apply two variants of this SOC-model (Figure 1): (i) the full five-pool model and (ii) a four-pool model in which IOM is absent and all SOC is subject to decay. The major unknowns that are used in the SOC-module are the decay rates for the four pools.

2.3. Research Layout and parameter fitting protocol

The available data permit calculation of the size of the individual SOC-pools at the five chronosequence sites. The litter-C input and various geochemical properties are also measured at these sites (Doettrl et al., 2018). This data-rich environment permits an analysis of SOC-decay by site-specific model parametrization and calibration, but does not correspond to more common situations where no individual SOC-pools can be obtained from measurements. We therefore consider various information levels, from data-rich to relatively data-poor, to evaluate the quality lost by decreasing data-richness. Figure 3 summarizes the corresponding research layout.

- At the highest information level (info level 1), a site-specific calibration of the decay rate factors for RPM, DPM, HUM and BIO is performed, using the local profile data to parametrize SoilGen including local estimates of the SOC-input by plant litter. The calibration protocol follows a downstream scheme, which means that the pathway of SOC-decay via RPM and DPM to HUM and BIO determines the calibration order of the decay rate factors. Additionally, rate speeds are considered (slower rates are calibrated first). This results in the calibration sequence k_{RPM}-k_{DPM}-k_{HUM}-k_{BIO}. For each rate, eight equidistant values are taken in a range reported in

previous calibration studies (e.g. Yu et al., 2013). The value of the best performing rate factor was obtained by (i) selecting the two consecutive simulated rates with one positive and one negative deviation between measurements and simulations, and (ii) interpolating between these rates to find the rate with an error of near zero. This approach assumes a monotonously decreasing or increasing relation between model error and rate value, which was checked to be true by graphical analysis of the simulations for five sites, eight different values for all four rate coefficients for both the four- and five-pool models, thus for 320 cases. A still better match between simulated and measured pools could have been obtained by an iterative approach to search the optimal parameter value (e.g. by a bisection procedure with a convergence criterion), but this would have added little precision to this study (c.f. Fig. 4) and would have greatly increased computation time. The quality of a calibration run for each pool was expressed via the (absolute) difference between the measured and simulated SOC-pools (Mg ha^{-1} cm^{-1} soil) over the same depth intervals combined over topsoil and subsoil. As an example, if measurements apply to 0-9 cm for topsoil and to 19-30 cm for subsoil, values for the same depth intervals were taken from the simulations. SOC is expressed per cm soil to allow comparisons between soils that had unequal sampling layer thickness.

- The rates obtained by calibration were correlated to geochemical and mineralogical data from the same profiles, to test if these data provide proxies for physico-chemical protection of SOC. We used profile-averages for these proxy-data to avoid weighting for unequal layer thicknesses and because one of the proxies was at the profile-scale. Strong and significant correlations indicate candidate proxies for a next step where rate modifiers are calculated by linear regression. The rate modifiers are applied onto each pool to calculate the loss from that pool by

$$loss = Y^{*}(1-e^{-x1^{*}x2^{*}\overline{k_p}^{*}t}) \quad \text{(eq.3)},$$

where *Y* is the size of a pool (RPM, DPM, BIO, HUM) at the start of the (daily) time step for the C-cycle sub-model (Mg ha^{-1} y^{-1}), $\overline{k_p}$ is the average rate constant (y^{-1}) obtained from the calibrations for pool *p*; $x1$ is a (dynamic) rate modifier for the combined effect of moisture, soil cover and

temperature, fluctuating over time and depth; $x2$ is an additional rate modifier representing the physico-chemical protection; t is the period of decay (1/365 year). The rate modifier $x2$ gradually changes over time because of weathering processes affecting geochemical soil properties. The value of $x2$ is calculated for each site i by:

$$x2_i = \frac{\beta_0 + \beta_1 * G_i}{\bar{k}}, \qquad \text{(eq.4)},$$

where β_o and β_1 are regression coefficients, G_i is the value of the geochemical proxy at i and $\bar{k}$ is the average of the calibrated rate constants (y^{-1}) over all locations. In long-term simulations, G_i would be time-dependent to reflect the effect of geochemical weathering processes. In addition, local estimates of the SOC-input by plant litter are input to SoilGen. These simulations applying the rate modifier (one per site) correspond to a second, lower information level (info level 2) where no site-specific rate constants are known but the mineralogy and geochemistry is known and can be used to modify rates.

- A next lower information level (info level 3) corresponds to the usage of averaged rate constants in combination with local estimates of the SOC-input by plant litter. We distinguished two variants: the literature value for the rate constants (Coleman & Jenkinson, 2005, c.f. Table 2) and the average of the calibrated rate constants. The former is expected to perform worse than the latter as the averages of calibrated constants would give unbiased (average of zero) simulation errors in the study area, but the default rates would likely yield biased (non-zero) simulation errors.
- The lowest information level (info level 4) corresponds to the usage of average rate constants and of the same SOC-litter inputs at all sites, assuming net primary production does not depend on soil but only on climate. Again, we distinguished two variants: the rates from literature and the averaged rates from calibration.

The effects of the information levels on the quality of SOC-simulations over all pools at an individual site is calculated by the site Root Mean Square Error:

330 $$RMSE = \sqrt{\frac{1}{4}\sum_{p=1}^{4}(S_p - O_p)^2}\,,$$ (eq.5),

331 where S_p and O_p are simulated and observed SOC-amounts (Mg ha^{-1} cm^{-1} soil) in pool p. We do not
332 consider the difference between measured and simulated IOM because both are either zero (four-
333 pool model) or equal (five-pool model) since IOM does not decay (Figure 1).
334 All model simulations covered 1000 years to obtain a stable distribution of SOC over the pools and to
335 avoid an effect of the (unknown) initial SOC-content on final SOC-pools. We compared 100 and
336 1000-year simulations for the young soil PM24II (0.1 ka) and found a minor effect of simulation
337 duration on pool sizes. The SOC-content at the start of the simulations reflected the measured rSOC.
338 In the five-pool model, rSOC was set equal to the IOM-pool, in the four-pool model it was set equal
339 to the initial HUM pool.
340 An over-all index of simulation quality over the five sites is calculated by:

341 $$\widetilde{RMSE} = \sqrt{\frac{1}{5}\sum_{i=1}^{5}(RMSE_i)^2}$$ (eq.6)

342 We acknowledge that other accuracy indices could also have been applied, but as we are interested
343 in the decrease of accuracy as a function of information level rather than accuracy itself, we
344 restricted ourselves to RMSE. Adding a penalty for extra parameters needed (at information level 2)
345 could be done by using a statistic such as the Akaike Information Criterion, but this would not inform
346 on the additional cost, which may strongly vary per soil (or weathering) parameter.

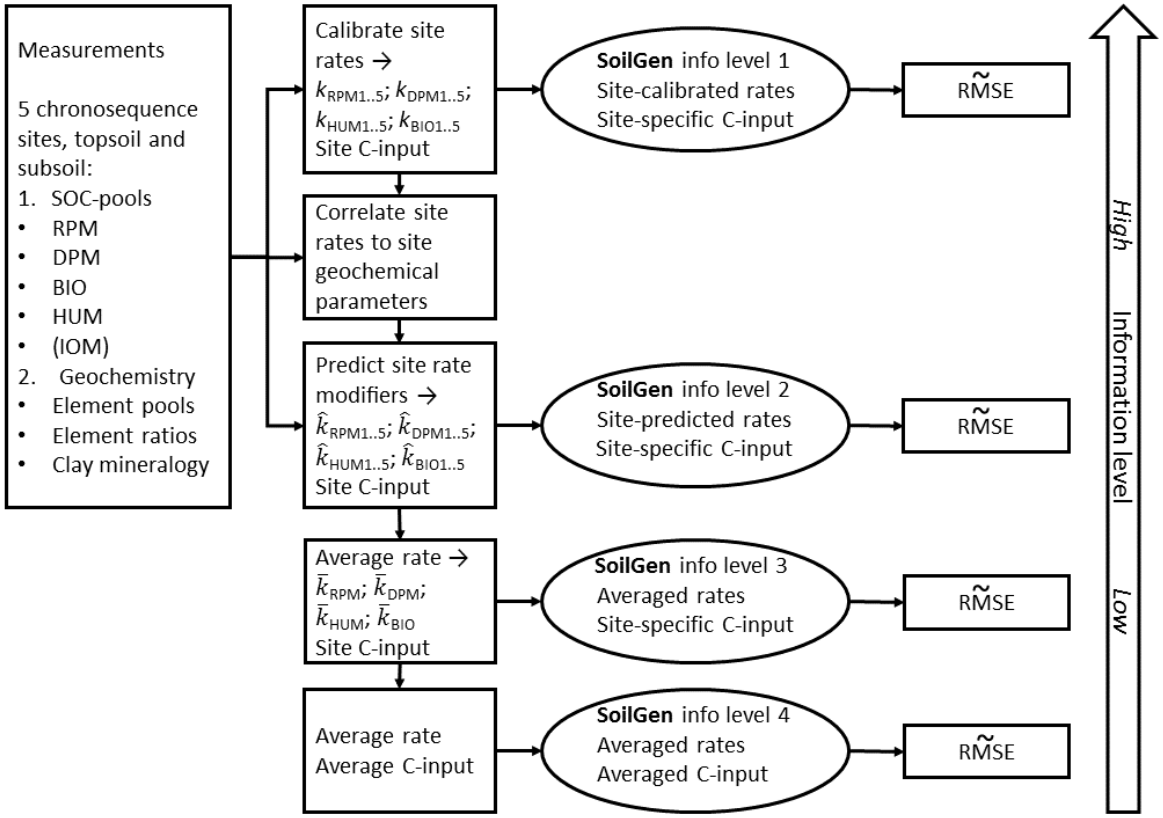

Figure 3- Research layout, applied onto both the 4-pool and the 5-pool SOC-model. k refers to rate constants; RPM, DPM, HUM, BIO and IOM are Soil Organic Carbon (SOC-) pools (described in Fig.1); $\widetilde{RMSE}$=Root Mean Square Error over 4 pools and 5 sites. Sampling depths for topsoil and subsoil in Table 1. Information levels 1 to 4 are described in section 2.3.*

3. RESULTS AND DISCUSSION

3.1. Information level 1: Site-specific calibration

Primary purpose of the site-specific calibration was to see if rate constant values would differ by location along the chronosequence. This was found to be the case (Table 2). Figure 4 shows that the pool sizes are reproduced accurately. RMSE- and $\widetilde{RMSE}$-values in Table 2 indicate that the five-pool model was calibrated slightly more accurately than the four-pool model, but the difference is minimal. In the four-pool model, the k_{HUM} and k_{BIO} indicate slower decay than in the five-pool model, which might be expected because the recalcitrant OC is part of the HUM pool in this model version. k_{DPM} and k_{RPM} did not change because measurements of DPM and RPM are the same for both the

four- and five-pool models. At 1000 simulation years, an IOM pool (five-pool model) apparently hardly influences the simulation quality when compared to the four-pool model, even when the IOM equals between 30 and 50% of SOC in the five-pool model simulation results (Figure 4) and the BIO and HUM-pools are consequently smaller in the five-pool model than in the four-pool model. The fastest decomposition rates for fresh organic matter (k_{RPM} and k_{DPM}) are found at the oldest site CH. This can be explained by a relatively high amount of 1:1 (low activity) clay minerals (Supplementary Information: Table S2) that have a weaker capacity to form organo-mineral assemblages and thus result in less protection of relatively fresh (RPM, DPM) SOC than by soils with higher amounts of 2:1 minerals (Baldock & Skjemstad, 2000). Slightly faster decay of the HUM and BIO pools is found in the older site R, but only in the site CH with the five-pool model. Finally, it should be noted that differences in total SOC between sites are partly explained by different Litter-C inputs (Table 2; Supplementary Information Table S3; c.f. section 3.3).

Table 2- Litter-C input, calibrated rate constants (k, expressed in y^{-1}) and RMSE per site for the five-pool and the four-pool model. Italic values in brackets are calibration domain within which the optimal value for k was found. RPM, DPM, HUM, BIO are pools described in Fig.1. C-inputs are based on Doetterl et al., 2018. Default rates RothC-model are taken from Coleman & Jenkinson (2005).

Site	**C-input**	**5-pool model**					**4-pool model**				
	Mg C ha^{-1} y^{-1}	**k_{RPM}** *[0.075; 0.525]*	**k_{DPM}** *[0.100; 1.600]*	**k_{HUM}** *[0.005; 0.035]*	**k_{BIO}** *[0.100; 0.800]*	**RMSE**	**k_{RPM}** *[0.075; 0.525]*	**k_{DPM}** *[0.100; 1.600]*	**k_{HUM}** *[0.005; 0.035]*	**k_{BIO}** *[0.100; 0.800]*	**RMSE**
		y^{-1}				**Mg ha^{-1} cm^{-1}**	**y^{-1}**				**Mg ha^{-1} cm^{-1}**
PM24II	1.3	0.070	0.370	0.008	0.333	0.0246	0.070	0.370	0.005	0.200	0.0090
PM22	1.3	0.059	0.286	0.009	0.370	0.0452	0.059	0.286	0.005	0.217	0.0429
M1	1.0	0.068	0.345	0.009	0.385	0.0093	0.068	0.345	0.010	0.303	0.0260
R	1.2	0.060	0.286	0.023	1.000	0.0265	0.060	0.286	0.012	0.526	0.0283
CH	0.8	0.185	1.000	0.011	0.526	0.0063	0.185	1.000	0.007	0.294	0.0197
						$\widetilde{RMSE}$					$\widetilde{RMSE}$

Average rate $\bar{k}$	0.089	0.455	0.012	0.526	*0.0264*	0.089	0.455	0.008	0.303	*0.0275*
Default k RothC-model	0.30	10.0	0.02	0.66	*0.2721*	0.30	10.0	0.02	0.66	*0.4213*

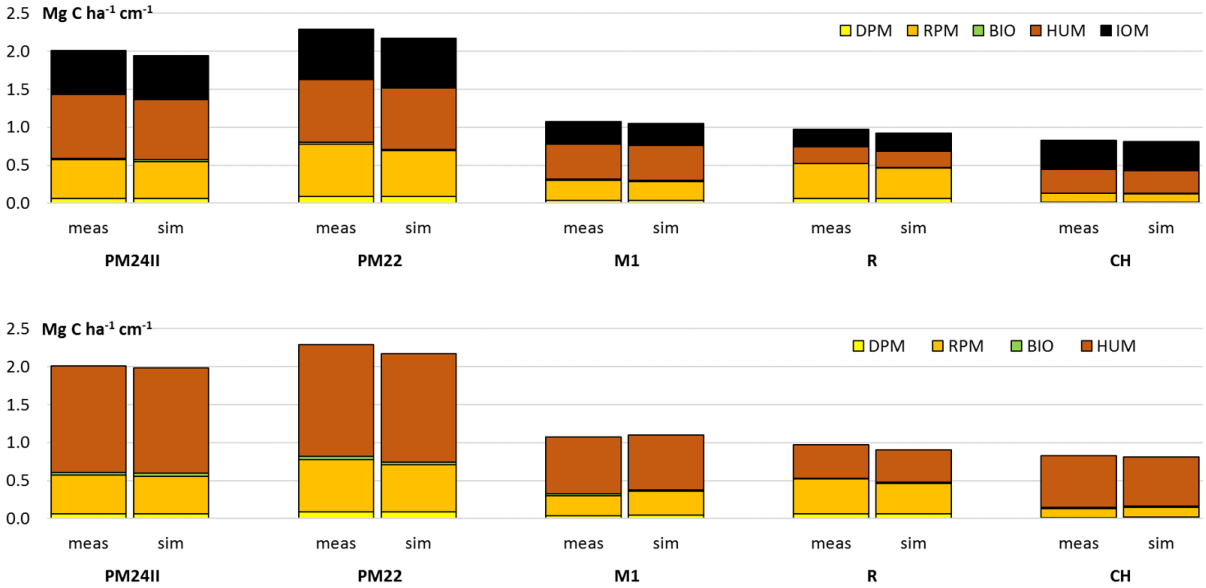

Figure 4- Pool sizes derived from converted measurements (meas; c.f. section 2.1) and from simulations (sim) along the chronosequence for the five-pool model (top) and the four-pool model (bottom) expressed per sampled cm depth. Sampled depth intervals in Table 1. RPM, DPM, HUM, BIO and IOM are pools described in Fig.1.

3.2. Identification of influencing soil properties

Calibrated rate constants were correlated to various measured soil properties to identify candidate properties for rate modifiers (Table 3). Despite the small number of sites, strong and significant correlations were found. Results for the five-pool and four-pool models are similar. It should be noted that the results for the much larger pools RPM and HUM are more relevant than those of the volatile pools of DPM and BIO, which may be more responsive to climate fluctuations.

First, important observations are the high and significant correlations between k_{RPM} (and k_{DPM}) and age and the profile development index (PDI, c.f. Supplementary Information; Harden, 1982). The PDI is assessed using basic field soil data, which makes it an easily available rate-modifier in other areas as well.

Good geochemical candidates for rate modifiers in the five-pool model are 1:1 clay mineral content and 2:1 clay mineral content. The Fe/Si-ratio is a reasonable candidate. These properties are model outputs of SoilGen. Age is of less relevance because it is not a model output, but an input and is not easily estimated for field sites which makes it a less attractive predictor for rate modifiers.

The 1:1 clay mineral content correlates positively with all four rates, thus decay of all four pools is faster with higher 1:1 clay mineral content. This is mainly of significance and importance for the large RPM-pool. The positive correlations may be explained by the weak bonding between OM and these low-activity clay minerals and results in less protection of the OM in organo-mineral assemblages. The 2:1 clay mineral content correlates negatively with all four rates, thus decay of all four pools is slower with higher 2:1 clay mineral content. This is explained by the stronger bonding between these high-activity clay minerals and OM, resulting in protection against decomposers, as also stated by others (e.g. Barré et al., 2014; Cuadros, 2017).

Fe/Si as well as Al/Si ratios correlate negatively with all four rates, thus soils with higher Fe- or Al-content and lower Si-content had slower decay. This holds especially for BIO and to a lesser degree for HUM, and may be explained by OM-stabilization due to organo-mineral interactions at metal surfaces (Kögel-Knabner et al., 2008). Similarly, higher Si contents are positively correlated (though not at high significance) with decay rates, most importantly that of RPM and DPM.

In the four-pool model, a significant negative correlation between silt content and k_{HUM} was found. Additionally, BS% correlates negatively (but not at high significance) to k_{RPM}, indicating that soils with a high base saturation have a slower decay of RPM. This might indicate that a higher BS, stimulating biological activity of macrofauna, results in stronger aggregation of the soil and thus better physical protection of RPM from decomposers (Ewing et al., 2006). This effect inverts for k_{HUM} and k_{BIO} (no significant correlations), which may indicate that microfauna can access the HUM and BIO pools. In B-horizons (below the sampled layers in this profile) with stable aggregates, this pattern may inverse to slow decay of SOC, as found by Ewing et al., 2006 for Californian soils. A non-significant correlation between K_{HUM} and clay content, additional to the effect already part of the

RothC26.3 correction for rates, was found for the four-pool model, which suggests that the RothC26.3 correction for clay content may be improved.

Table 3- Pearson correlation coefficients between calibrated rate constants (k_RPM, k_DPM, k_HUM and k_BIO, in y^{-1}) and measured soil properties at five sites. Underlined: significant at α=10%; double underlined: significant at 5%; yellow-marked

field: correlation switched from non-significant to significant or vice versa on change from five-pool to four-pool model.

RPM, DPM, HUM, BIO and IOM are defined in Fig.1 and in the introduction.

	five-pool model					four-pool model			
Pearson correlations	k_RPM	k_DPM	k_HUM	k_BIO	IOM	k_RPM	k_DPM	k_HUM	k_BIO
k_{RPM} (y^{-1})	1.00					1.00			
k_{DPM} (y^{-1})	1.00	1.00				1.00	1.00		
k_{HUM} (y^{-1})	-0.12	-0.14	1.00			-0.23	-0.25	1.00	
k_{BIO} (y^{-1})	-0.05	-0.07	1.00	1.00		-0.09	-0.11	0.88	1.00
IOM (Mg ha^{-1} cm^{-1} soil)	-0.14	-0.13	-0.67	-0.67	1.00				
Age (y)	0.99	0.98	0.02	0.09	-0.21	0.99	0.98	-0.13	0.04
Log Age (y)	0.63	0.61	0.52	0.56	-0.70	0.63	0.61	0.54	0.61
Profile Development Index (-)	0.96	0.95	0.13	0.20	-0.37	0.96	0.95	0.04	0.18
C-input (Mg ha^{-1} y^{-1})	-0.75	-0.73	-0.56	-0.62	0.63	-0.75	-0.73	-0.38	-0.59
Clay (%)	-0.04	-0.06	0.18	0.15	-0.76	-0.04	-0.06	0.76	0.40
Silt (%)	0.63	0.64	-0.55	-0.49	0.66	0.63	0.64	-0.82	-0.64
Sand (%)	-0.62	-0.62	0.54	0.50	0.02	-0.62	-0.62	0.16	0.39
Total Reserve of Bases ($cmol_c$ kg^{-1} cm^{-1} soil)	-0.57	-0.54	-0.40	-0.44	0.55	-0.57	-0.54	-0.47	-0.50
Specific Surface Area treated (m^2 g^{-1})	-0.36	-0.46	-0.41	-0.47	0.02	-0.46	-0.46	0.18	-0.26
Specific Surface Area untreated (m^2 g^{-1})	-0.35	-0.35	-0.29	-0.34	-0.25	-0.35	-0.35	0.38	-0.09
1:1 clay mineral (%)	0.89	0.88	0.32	0.39	-0.43	0.89	0.88	0.14	0.34
2:1 clay mineral (%)	-0.92	-0.91	-0.26	-0.33	0.40	-0.92	-0.91	-0.08	-0.29
Si (g/kg)	0.77	0.75	0.45	0.50	-0.51	0.77	0.75	0.32	0.49
Fe/Si	-0.47	-0.45	-0.77	-0.81	0.77	-0.47	-0.45	-0.66	-0.82
Al/Si	-0.77	-0.74	-0.37	-0.43	0.39	-0.77	-0.74	-0.22	-0.39
P (mg/kg)	-0.18	-0.14	-0.65	-0.66	0.67	-0.18	-0.14	-0.74	-0.74
Base Saturation (%)	-0.65	-0.68	0.52	0.48	0.06	-0.65	-0.68	0.33	0.43
Cation Exchange Capacity ($cmol_c$/kg soil)	-0.46	-0.45	-0.60	-0.65	0.10	-0.46	-0.45	-0.02	-0.46

3.3. Simulation quality at different information levels

Soil properties with significant correlations to one or more calibrated rate constants (Table 3) were considered suitable candidates to estimate rate modifiers $x2$ (eq. 3) by linear regression. Thus, using the contents of 1:1 clay minerals and 2:1 clay minerals, the Fe/Si ratio and (for the four-pool model) the silt content, rate modifiers were estimated, the model was run and simulated pools were compared to measured pools (information level 2). Interaction effects of suitable candidate modifiers were also investigated (Fe/Si x 1:1 clay minerals and silt content x 1:1 clay minerals). Results (Figure 5) clearly show that the $\widetilde{\mathrm{RMSE}}$ reduces (and quality increases) with addition of information. The best results were obtained by local calibration (information level 1), and the worst when rates as well as C-inputs were assumed the same at all sites (information level 4). We scaled the $\widetilde{\mathrm{RMSE}}$ at information level 4a (average of measured rates and average of measured C-inputs) to 100% in the discussion below.

Estimating rate modifiers by geochemical parameters (information level 2) increased $\widetilde{\mathrm{RMSE}}$ by 45-50% (compared to information level 1) for the best performing geochemical proxy, the Fe/Si ratio. This ratio is a proxy for bonding in pedogenic hydroxides. Slightly worse results were obtained by using 1:1 or 2:1 clay mineral content to modify rates. Interaction terms of clay mineralogy and Fe/Si ratio did not give better (lower) values of $\widetilde{\mathrm{RMSE}}$. From the Fe/Si based rate modifiers to the usage of average rates (information level 3a) the $\widetilde{\mathrm{RMSE}}$ increased by about 20%, which illustrates the added value of including geochemical properties in SOC-modelling, and the importance of incorporating the degree of weathering in SOC-modelling. It should be noted that using the default RothC-rates (information levels 4b and 3b) led to a much worse performance in terms of the $\widetilde{\mathrm{RMSE}}$ (30% to 70% higher than at information level 4a), as these default rates added a bias component to the error relative to the usage of locally-averaged rates. Thus, the 20% is a conservative estimate of the gain in quality that can be realized by using rate modifiers, and usage of the default rate constants of RothC introduces large errors in the simulations of the individual pools. Finally, at the

4th information level (4a) another substantial increase of ca. 15% in the $\widetilde{\mathrm{RMSE}}$ occurred when, in addition to averaged rates, average values for Litter-C inputs are used instead of local values, a situation corresponding to the usage of literature values. Again, the scenario 4b based on default RothC-rates performs much worse than scenario 4a based on the average of the calibrated rates. The decrease in $\widetilde{\mathrm{RMSE}}$ over the information levels from 4 to 1 show that using local estimates of C-input contributes positively to simulation quality, though less than when geochemical proxies are used as rate modifiers. C-inputs were found to decrease with greater soil age (Table 2), which is, under the same climate, likely related to less available nutrients with increased degree of weathering. This confirms that the SOC-model should include adequate C-inputs by either accounting for soil fertility in simulating biomass production or by using local measurements.

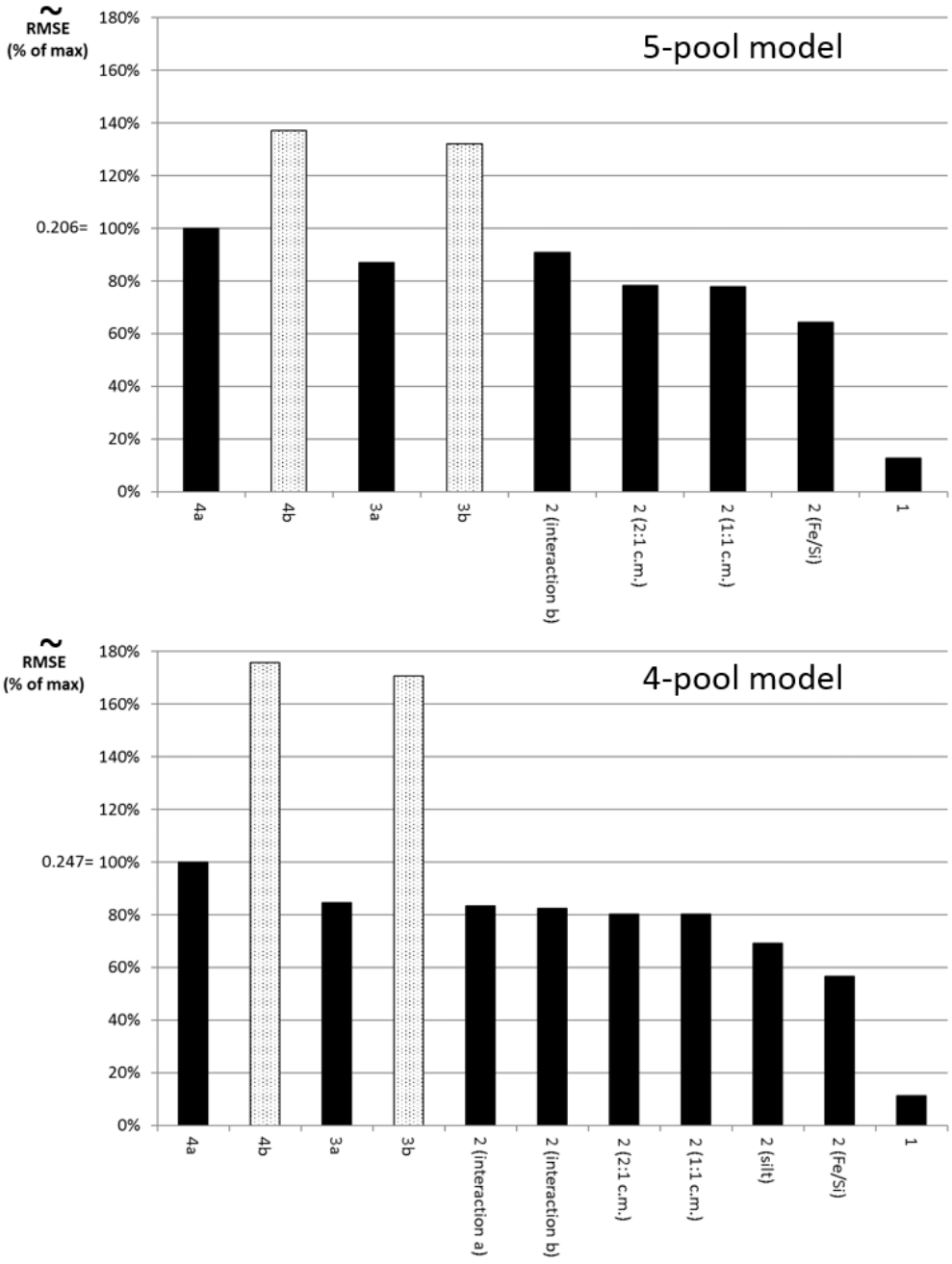

Figure 5- $\widetilde{RMSE}$ at different information levels: 1=site calibration; 2=site rate modifiers (BS=Base Saturation, c.m.=clay mineral content, interaction a= silt x 1:1 clay mineral content, interaction b= Fe/Si x 1:1 clay mineral content); 3a=average rates; 3b (dotted bars)=default rates RothC26.3; 4a=average rates with average litter-C input; 4b (dotted bars)=default rates RothC26.3 with average C-input. High $\widetilde{RMSE}$ denotes low model performance. Values for $\widetilde{RMSE}$ at information level 4a are given at 100% scale.

The quality of simulations at all information levels depends on the root depth distribution. Since this is used to distribute belowground litter inputs it will determine the litter input in the sampled soil

474 depth compartments. Additionally, the simulation quality depends on the bioturbation. As simulated
475 bioturbation mixes the SOC-pools over the rooted profile, it will decrease the sensitivity of simulated
476 SOC-pools over depth to the root profile. Because the rooting profile and the bioturbation settings
477 were the same for all simulations per site, we assume with confidence that the differences between
478 the information levels are hardly dependent on assumptions on rooting profile and bioturbation.
479 Rate constants (before modification by *x1* and *x2*) in this model study do not vary with depth. Rate
480 modifier *x1* depends on temperature and moisture deficit and these do vary with depth. Rate
481 modifier *x2* could be estimated per depth compartment using a geochemical proxy, but in this study,
482 we estimated *x2* for the whole profile. Since not all geochemical proxies (e.g. newformed clay
483 minerals) can be simulated with SoilGen we used measured values. Although this study shows the
484 value of including geochemical proxies to correct rate constants, the actual gain in quality over long
485 time periods will depend on the accuracy of simulations of these geochemical proxies. For short and
486 recent periods (a few millenniums, c.f. the Fe/Si ratios in Table S2), measured proxy values will
487 suffice.
488 In SoilGen, the uncorrected rate constant is set to zero below the rooting depth, which is below the
489 sampled subsoil layers in this study. Both a constant *x2* and no-decay below the rooting zone are
490 over-simplifications and may contribute to the RMSE at all information levels when deeper profiles
491 would be considered. For instance, Mathieu et al. (2015) inventoried turnover rates of deeper
492 carbon (below 20 cm depth) and found different rates for different soil types. Balesdent et al. (2017)
493 measured turnover rates in cultivated soils using ^{13}C labeling techniques, and found turnover rates in
494 subsoils (below 30 cm depth) to be about four times slower than in topsoils. Thus, future models
495 might benefit from exploring the candidate mechanisms that govern dynamic depth adaptation of
496 rate constants. Associations of microbial-mineralogical-root dynamics are likely key to these
497 mechanisms.
498
499 **3.4. Radiocarbon**

Values for pMC were measured in topsoil and subsoil bulk samples (Doetterl et al., 2018; Supplementary Information; Table 3) and were also calculated from additional simulations, using the best performing geochemical rate modifiers at information level 2 (i.e. Fe/Si ratio). Simulation periods for these additional simulations were 0.1 ka for site PM24II, 3 ka for PM22 and 10 ka for the other (older) sites. The measured values of pMC (Figure 6) are generally close to or above 100% which indicates that most of the SOC is young. The reason for the exceptional measured value for pMC-bulk soil for the subsoil of site R (300 ka) is that the sample had a large POM content of young age, which was likely due to some root fragments that were, unintendedly, still in the sample in combination with a low total SOC content.

Results show that simulations with the four-pool model better match the pMC measurements (RMSE of 10.9%) than those with the five-pool model (RMSE of 30.3%). Topsoil pMC-simulation results are better than those from the subsoil. Assuming an IOM-pool that exists for 10 ka time extent gives a poor estimation of the pMC, which suggests that this SOC cannot be considered inert over several millennia. This was also concluded by Sanderman et al. (2016). Considering IOM to be older than 10 ka years leads to worse results. A theoretical limiting case with all IOM as non-radioactive IOM, calculating by setting *IOMx* in eq.1 equal to zero, is depicted in Figure 6. Thus, it can be concluded that probably the four-pool model is more appropriate for the chronosequence than the five-pool model with inert SOC. However, a modified five-pool model including slow decay of SOC (in a pool replacing IOM) might lead to more accurate results than the results that we obtained. However, the pathways towards such pool then still need to be defined.

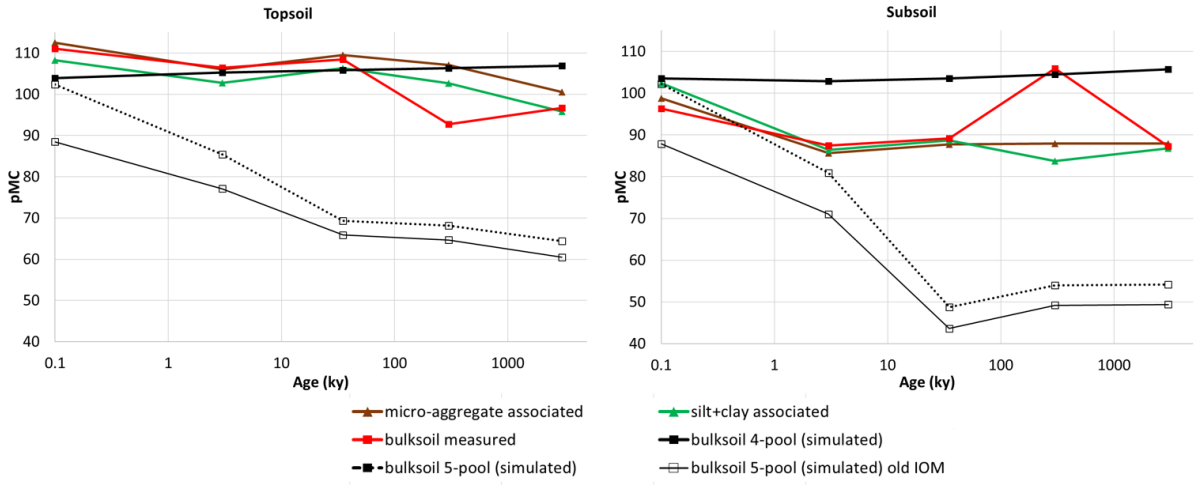

Figure 6- Percentage Modern Carbon (pMC) obtained from measurements in bulk soil (solid red) and fractions (solid brown and green) and by simulation with the four-pool model (solid black line, solid rectangles) and five-pool model (dashed black line, open rectangles) for topsoil (left) and subsoil (right). Solid black line with open markers indicate a limiting case without ^{14}C in IOM.

CONCLUSIONS AND OUTLOOK

1. Decay rates for the RothC-pools showed significant correlations to both a soil development index (PDI) based on basic field observations as on geochemical weathering metrics. The correlation to PDI would allow rate-modification at large point data sets, as PDI is a function of commonly recorded soil data. This possibility needs to be further explored.
2. Over-all, using decay-rate modifiers based on geochemical weathering indices (clay mineralogy and element ratios) significantly improved simulations for SOC storage, but requires measuring and/or modelling these parameters
3. Both the five-pool SOC-model, including inert IOM, and the four-pool model without IOM were improved by decay rate modifiers. Comparisons of measured and simulated pMC of bulk SOC samples showed that the inert IOM-pool degrades over periods of time, which provides further evidence that inclusion of an inert pool is not appropriate at millennial timescales.

4. Local, site specific estimations of C-input generally increased simulation quality, suggesting an impact of the degree of weathering and soil development. Further research should be performed to evaluate this conclusion.
5. Major limitation of this study was the small number of sites; thus, results need re-examination at additional sites with varying age and geochemical properties.

ACKNOWLEDGEMENTS

This research was financed in the framework of the BELSPO funded project P7/24 'SOGLO'. Special thanks to C. Lawrence (US Geological Survey) for providing a USGS internal review of this manuscript, and to 3 anonymous referees.

REFERENCES

Aciego, S. M., Riebe, C. S., Hart, S. C., Blakowski, M. A., Carey, C. J., Aarons, S. M., Dove, N.C., Botthoff, J.K., Sims, K.W.W., Aronson, E. L. 2017. Dust outpaces bedrock in nutrient supply to montane forest ecosystems. Nature communications, 8, 14800. http://dx.doi.org/10.1038/ncomms14800 .

Baldock, J.A., Skjemstad, J.O. 2000. Role of the soil matrix and minerals in protecting natural organic materials against biological attack. Organic Geochemistry 31: 697-710. http://dx.doi.org/10.1016/S0146-6380(00)00049-8

Balesdent, J., Basile-Doelsch, I., Chadoeuf, J., Cornu, S., Fekiacova, Z., Fontaine, S., Guenet, B., Hatté, C., 2017. Renouvellement de carbone profond des sols cultivés : une estimation par compilation de données isotopiques. Biotechnologie, Agronomie, Société et Environnement 21 : 1-9.

Barré, P., Fernández-Ugalde, O., Virto, I., Velde, B. and Chenu, C..2014. Impact of phyllosilicate mineralogy on organic carbon stabilization in soils: Incomplete knowledge and exciting prospects. Geoderma 235-236: 382-395. http://dx.doi.org/10.1016/j.geoderma.2014.07.029

565 Berhe, A.A. 2012. Decomposition of organic substrates at eroding vs. depositional landform
566 positions. Plant Soil 350: 261-280. http://dx.doi.org/10.1007/s11104-011-0902-z
567 Coleman, K., Jenkinson, D.S., 2005. RothC-26.3: a model for the turnover of carbon in soil. Model
568 Description and Users Guide. November 1999 Issue (Modified April 2005).
569 https://www.rothamsted.ac.uk/sites/default/files/RothC_guide_WIN.pdf (accessed June
570 2017).
571 Cuadros, J., 2017. Clay minerals interaction with microorganisms: a review. Clay Minerals 52: 235–
572 261. http://dx.doi.org/10.1180/claymin.2017.052.2.05
573 Doetterl, Sebastian, Antoine Stevens, Johan Six, Roel Merckx, Kristof Van Oost, Manuel Casanova
574 Pinto, Angélica Casanova-Katny, Cristina Muñoz, Mathieu Boudin, Erick Zagal Venegas, Pascal
575 Boeckx, 2015. Soil carbon storage controlled by interactions between geochemistry and
576 climate. Nature Geoscience 8: 780–783. http://dx.doi.org/10.1038/NGEO2516
577 Doetterl, S., Berhe, A.A., Bodé, S., Fiener, P., Finke, P., Fuchslueger, L., Griepentrog, M., Harden, J.,
578 Nadeu, E., Schnecker, J., Six, J., Trumbore, S., Van Oost, K., Vogel, C., Boeckx, P. 2018. Links
579 among warming, carbon and microbial dynamics mediated by soil mineral weathering.
580 Nature Geoscience. http://dx.doi.org/10.1038/s41561-018-0168-7
581 *2018*Ewing, S.A., Sanderman, J., Baisden, W.T., Wang, Y., Amundson, R. 2006. Role of large-scale soil
582 structure in organic carbon turnover: Evidence from California grassland soils. Journal of
583 Geophysical Research:Biogeosciences 111 issue:G3.
584 http://dx.doi.org/10.1029/2006JG000174
585 Finke, P.A., Hutson, J., 2008. Modelling soil genesis in calcareous löss. Geoderma 145: 462-479.
586 http://dx.doi.org/10.1016/j.geoderma.2008.01.017
587 Finke, P.A, 2012. Modeling the genesis of Luvisols as a function of topographic position in loess
588 parent material. Quaternary International 265: 3-17.
589 http://dx.doi.org/10.1016/j.quaint.2011.10.016

590 Finke, P., Samouëlian, A., M. Suarez-Bonnet, B. Laroche and Cornu S. 2015. Assessing the usage
591 potential of SoilGen2 to predict clay translocation under forest and agricultural land uses.
592 European Journal of Soil Science 66(1): 194-205. http://dx.doi.org/10.1111/ejss.12190
593 Gobat, J-M., Aragon, M., Matthey, W., 2004. The Living Soil: Fundamentals of Soil Science and Soil
594 Biology. Science Publishers, Enfield, U.S.A. 602 pp.
595 Goosse, H., Brovkin, V., Fichefet, T., Haarsma, R., Huybrechts, P., Jongma, J., Mouchet, A., Selten, F.,
596 Barriat, P.-Y., Campin, J.-M., Deleersnijder, E., Driesschaert, E., Goelzer, H., Janssens, I.,
597 Loutre, M.-F., Morales Maqueda, M.A., Opsteegh, T., Mathieu, P.-P., Munhoven, G.,
598 Pettersson, E.J., Renssen, H., Roche, D.M., Schaeffer, M., Tartinville, B., Timmermann, A., and
599 Weber, S. L., 2010. Description of the Earth system model of intermediate complexity
600 LOVECLIM version 1.2. Geoscientific Model Development 3: 603-633.
601 https://doi.org/10.5194/gmd-3-603-2010
602 Harden, J.W., 1982. A quantitative Index of Soil Development from Field Descriptions: Examples from
603 a Chronosequence in Central California. Geoderma 28(1): 1-28.
604 https://doi.org/10.1016/0016-7061(82)90037-4
605 Harden, J.W. (editor), 1987. Soils Developed in Granitic Alluvium near Merced, California. U.S.
606 Geological Survey Bulletin 1590A. https://pubs.usgs.gov/bul/1590a/report.pdf
607 Hargreaves, G.H., Samani, Z.A. 1985. Reference crop evapotranspiration from temperature. Appl.
608 Eng. Agric. 1 (2), 96–99. http://dx.doi.org/10.13031/2013.26773
609 Hua, Q., Barbetti, M., Rakowski, A.Z., 2013. Atmospheric radiocarbon for the period 1950–2010.
610 Radiocarbon 55 (4): 2059–2072. http://dx.doi.org/10.2458/azu_js_rc.v55i2.16177
611 Jenkinson, D.S., Coleman, K., 1994. Calculating the annual input of organic matter to soil from
612 measurements of total organic carbon and radiocarbon. European Journal of Soil Science 45,
613 167–174. http://dx.doi.org/10.1111/j.1365-2389.1994.tb00498.x

614 Kaplan, J.O., Krumhardt, K.M., Ellis, E.C., Ruddiman, W.F., Lemmen, C., Klein Goldewijk, K., 2011.
615 Holocene carbon emissions as a result of anthropogenic land cover change. The Holocene 21
616 (5): 775–791. http://dx.doi.org/10.1177/0959683610386983
617 Keyvanshokouhi, S., Cornu, S., Samouëlian, A., Finke, P. 2016. Evaluating SoilGen2 as a tool for
618 projecting soil evolution induced by global change. Science of the Total Environment 571:
619 110-123. http://dx.doi.org/10.1016/j.scitotenv.2016.07.119
620 Kögel-Knabner, I., Guggenberger, G., Kleber, M., Kandeler, E., Kalbitz, K., Scheu, S., Eusterhues, K.,
621 Leinweber, P. 2008. Organo-mineral associations in temperate soils: Integrating biology,
622 mineralogy, and organic matter chemistry. J. Plant Nutr. Soil Sci. 171: 61–82.
623 http://dx.doi.org/10.1002/jpln.200700048 61
624 Kononova, M.M., 1975. Humus of virgin and cultivated soils. In: Gieseling, J.E. (Ed.), Soil components.
625 I. Organic Components. Springer, Berlin, pp. 475–526.
626 Lawrence, Corey R., Jennifer W. Harden, Xiaomei Xu, Marjorie S. Schulz, Susan E. Trumbore, 2015.
627 Long-term controls on soil organic carbon with depth and time: A case study from the
628 Cowlitz River Chronosequence, WA USA. Geoderma 247–248: 73–87.
629 http://dx.doi.org/10.1016/j.geoderma.2015.02.005
630 Marchand, D.E., and Allwardt, A., 1981, Late Cenozoic stratigraphic units in northeastern San Joaquin
631 Valley, California: U.S. Geological Survey Bulletin 170, 70 p.
632 Mathieu, J., Hatté, C., Parent, E. Balesdent, J., 2015 Deep soil carbon dynamics are driven more by
633 soil type than by climate: a worldwide meta-analysis of radiocarbon profiles. Global Change
634 Biology 21, 4278-4292. http://dx.doi.org/10.1111/gcb.13012 .
635 Minasny, B. Finke, P., Stockmann, U., Vanwalleghem, T. and McBratney, A., 2015. Resolving the
636 integral connection between pedogenesis and landscape evolution. Earth Science Reviews
637 150: 102-120. http://dx.doi.org/10.1016/j.earscirev.2015.07.004

638 Opolot, E., Finke, P.A., 2015. Evaluating sensitivity of silicate mineral dissolution rates to physical
639 weathering using a soil evolution model (SoilGen2.25). Biogeosciences 12: 6791-6808.
640 http://dx.doi.org/10.5194/bg-12-6791-2015
641 Phillips, J.D., 2007. Development of texture contrast soils by a combination of bioturbation and
642 translocation. Catena 70 (1), 92-104. http://dx.doi.org/10.1016/j.catena.2006.08.002
643 Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Bronk Ramsey, C., Buck, C.E., Cheng, H.,
644 Edwards, R.L., Friedrich, M., Grootes, P.M., Guilderson, T.P., Haflidason, H., Hajdas, I., Hatté,
645 C., Heaton, T.J., Hogg, A.G., Hughen, K.A., Kaiser, K.F., Kromer, B., Manning, S.W., Niu, M.,
646 Reimer, R.W., Richards, D.A., Scott, E.M., Southon, J.R., Turney, C.S.M., van der Plicht, J.,
647 2013. IntCal13 and MARINE13 radiocarbon age calibration curves 0-50000 years calBP.
648 Radiocarbon 55(4). http://dx.doi.org/10.2458/azu_js_rc.55.16947
649 Sanderman, J., Baisden, W.T., Fallon, S. 2016. Redefining the inert organic carbon pool. Soil Biology
650 and Biochemistry, 92, 149–152. http://dx.doi.org/10.1016/j.soilbio.2015.10.005
651 Sauer, D., Finke, P.A., Schülli-Maurer, I., Sperstad, R., Sørensen, R., Høeg, H.I., Stahr, K. 2012. Testing
652 a soil development model against southern Norway soil chronosequences. Quaternary
653 International 265: 18-31. http://dx.doi.org/10.1016/j.quaint.2011.12.018
654 Schmidt, M.W., M.S. Torn, S. Abiven, T. Dittmar, G. Guggenberger, I.A. Janssens, M. Kleber, I. Kögel-
655 Knabner, J. Lehmann, D.A.C. Manning, P. Nannipieri, D.P. Rasse, S. Weiner, S.E.P. Trumbore.
656 2011. Persistence of soil organic matter as an ecosystem property. Nature, 478 (2011), pp.
657 49-56
658 Singer, A.C., Jury, W., Luepranchai, E., Yahng, C.-S., Crowley, D.E., 2001. Contribution of earthworms
659 to PCB bioremediation. Soil Biology & Biochemistry 33, 765–776.
660 http://dx.doi.org/10.1016/S0038-0717(00)00224-8
661 Stewart, C. E., Paustian, K., Plante, A. F., Conant, R. T., Six, J., 2008. Soil carbon saturation: Linking
662 concept and measurable carbon pools. Soil Sci. Soc. Am. J. 72, 379-392

Solly, E.F., Schöning, I., Herold, N., Trumbore, S.E., Schrumpf, M., 2015. No depth-dependence of fine root litter decomposition in temperate beech forest soils. Plant and Soil 392: 273-282. https://doi.org/10.1007/s11104-015-2492-7

USDA, 1975. Soil taxonomy – a basic system of soil classification for making and interpreting soil surveys. U.S. Department of Agriculture handbook 436: 754 p.

White, A.F., Blum, A.E., Schultz, M.S., Bullen, T.D., Harden, J.W., Peterson, M.L., 1996. Chemical weathering of a soil chronosequence on granite alluvium. I. Reaction rates based on changes in soil mineralogy. Geochim. et Cosmochim. Acta, 60: 2533-2550.

Wilkinson, M.T., Richards, P.J., Humphreys, G.S., 2009. Breaking ground: Pedological, geological, and ecological implications of soil bioturbation. Earth-Science Reviews 97: 257–272. http://dx.doi.org/10.1016/j.earscirev.2009.09.005

WRB 2014. World Reference Base for Soil Resources 2014. International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Report No 106, FAO, Rome.

Yu, Y.Y., Finke, P.A, Guo, Z.T, Wu H. B., 2013. Sensitivity analysis and calibration of a soil carbon model (SoilGen2) in two contrasting loess forest soils. Geoscientific Model Development 6, 29-44. www.geosci-model-dev.net/6/29/2013/

Zimmerman, M., Leifeld, J., Schmidt, M.W.I, Smith, P., and Fuhrer, J., 2007. Measured soil organic matter fractions can be related to pools in the RothC model. European Journal of Soil Science 58: 658-667. http://dx/doi.org/10.1111/j.1365-2389.2006.00855.x

Zwertvaegher, A., Finke, P., De Smedt, Ph., Gelorini, V., Van Meirvenne, M., Bats, M., De Reu, J., Antrop, M., Bourgeois, J., De Maeyer, P., Verniers, J., Crombé, Ph. 2013. Spatio-temporal modeling of soil characteristics for soilscape reconstruction. Geoderma 207-208: 166-179. http://dx.doi.org/10.1016/j.geoderma.2013.05.013

689

Supplementary Information

Table S1: Basic soil data aggregated to the upper meter of soil to enable correlation with decay rates. Profile Development Index calculated according to Harden (1982) for the closest characterized location in the same unit of the chronosequence, sd relate to the variation inside each unit in the chronosequence.

Profile	Age (years)	Log Age	Profile Development Index (sd)	C_input	clay	silt	sand
			-	Mg ha^{-1} yr^{-1}	%	%	%
PM24II	100	2.00	6.20 (7.19)	1.27	8.7	60.7	29.1
PM22	3000	3.48	13.72 (7.19)	1.29	10.9	65.8	22.5
M1	35000	4.54	60.95 (21.13)	1.2	24.5	53.4	18.2
R	300000	5.48	85.85 (39.37)	0.97	16.1	51.8	31.6
CH	3000000	6.48	324.40 (62.68)	0.84	14.3	68.6	16.1

Note: The Profile Development Index (Harden, 1982) combines eight soil properties estimated in the field with soil thickness: (1) clay films, (2) texture plus wet consistence, (3) color hue and chroma as proxy for rubification, (4) structure, (5) dry consistence, (6) moist consistence, (7) color value, and (8) pH. After the eight field properties are quantified (method in Table II in Harden, 1982), they are normalized onto a scale of 0-100%, where 100% indicates a strong development of the property. and averaged per soil horizon. The result per horizon is multiplied by the horizon thickness to give the soil profile soil development index.

Table S2: Soil chemical data averaged for the upper meter of soil. Ka=kaolinite, Il=illite, Sm=smectite, Vm=vermiculite. Specific Surface Area is the total surface area of soil per unit of mass. Analytical methods in Doetterl et al., 2018.

Profile	Total Reserve of Bases	Specific Surface Area treated	Specific Surface Area untreated	1:1 clay minerals (Ka)	2:1 clay minerals (Il-Sm-Ve)	Si	Fe/Si	Al/Si	P	Base Saturation	Cation Exchange Capacity
	$cmol_c$ kg^{-1} cm^{-1}	m^2 g^{-1}	m^2 g^{-1}	%	%	g kg^{-1}	-	-	mg kg^{-1}	%	$cmol_c$ kg^{-1} soil

PM24II	33.76	11.38	6.10	16	81	260	0.14	0.34	1186	79.8	14.65
PM22	16.80	19.97	10.75	19	81	300	0.12	0.23	553	100.0	11.20
M1	14.37	27.42	21.69	20	78	302	0.10	0.26	433	83.1	24.16
R	11.41	10.24	6.94	30	67	343	0.05	0.19	191	100.0	2.87
CH	3.95	8.91	5.87	49	41	391	0.06	0.08	391	75.7	2.96

Explanation of the weathering data (see also Doetterl et al., 2018): Weathering from 2:1 to 1:1 clays releases Si. We observed quartz-rich pebbles at the oldest site, which are likely accumulations of Si released from weathering of the smaller particles (e.g. 2:1 to 1:1 clays). Data (not shown, but see Supplementary Information for Doetterl et al. ,2018), show an increase with age in pedogenic Fe, but not in total Fe. The common view is that weathering releases Si and concentrates Fe and Al, but Harden (1988) concluded for these soils that this does not occur at all particle size classes equally and that this explains the Si, Fe/Si and Al/Si ratios observed in bulk samples.

Table S3: Sampling depths, measured pools and calculated RothC-pools: DPM=Decomposable Plant Material; RPM=Resistant Plant Material; BIO=Biomass; HUM=Humus and IOM=Inert Organic Matter. Analytical SOC-fractions are POM=particulate organic Matter; (S+A)= C associated to stable microaggregates (53-250 µm); (s+c)= C associated to non-aggregated clay and silt (<53 µm); and rSOC= resistant organic carbon measured after cold acid hydrolysis (<63 µm). pMC=percent Modern Carbon. Analytical methods in Doetterl et al., 2018.

Pofile	**Sampled Depth**		**Bulk density**	**Total SOC**	**Measured C-pools**				**SoilGen_RothC pools**			**pMC**		
	Top	Bot			POM	s+c -rSOC	S+A	rSOC	DPM + RPM	BIO + HUM	IOM	Bulk SOC	s+c	S+A
	(cm)	(cm)	g cm^{-3}	(g kg^{-1} soil)	(% of SOC)			(% s+c)	(g kg^{-1} soil)			(%)	(%)	(%)
Topsoil														
PM24II	0	9	1.30	23.6	40	38	22	77.83	9.4	10.1	4.0	111	108	113
PM22	0	13	1.21	36.2	48	26	26	70.70	.7.4	12.2	6.7	106	103	106
M1	0	5	1.28	24.5	48	32	20	85.83	11.8	8.5	4.2	108	107	110
R	0	12	1.65	12.8	67	10	22	67.54	8.6	2.3	1.9	93	103	107

CH	2	12	1.62	6.6	33	16	52	79.67	2.2	1.7	2.7	97	96	101
Subsoil														
PM24II	19	30	1.41	8.1	3	33	64	83.82	0.2	3.5	4.3	96	103	99
PM22	13	35	1.36	7.8	0	31	69	78.84	0.0	3.6	4.2	87	87	86
M1	22	53	1.61	4.6	15	44	41	82.58	0.7	2.4	1.6	89	89	88
R	12	39	1.88	2.5	25	13	62	68.54	0.6	0.8	1.1	106	84	88
CH	12	30	1.65	4.2	0	35	64	78.56	0.0	.2.1	2.1	87	87	88

719

720 **References**

721 Doetterl, S., Berhe, A.A., Bodé, S., Fiener, P., Finke, P., Fuchslueger, L., Griepentrog, M., Harden, J.,
722 Nadeu, E., Schnecker, J., Six, J., Trumbore, S., Van Oost, K., Vogel, C., Boeckx, P. 2018. Links
723 among warming, carbon and microbial dynamics mediated by soil mineral weathering.
724 Nature Geoscience. http://dx.doi.org/10.1038/s41561-018-0168-7

725 *2018*Harden, J.W., 1988. Genetic interpretations of elemental and chemical differences in a soil
726 chronosequence, California. Geoderma 43: 179-193

727 Harden, J.W., 1982. A quantitative Index of Soil Development from Field Descriptions: Examples from
728 a Chronosequence in Central California. Geoderma 28(1): 1-28.
729 https://doi.org/10.1016/0016-7061(82)90037-4