



HAL
open science

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

► To cite this version:

Peter Finke, Emmanuel Opolot, Jérôme Balesdent, Asmeret Asefaw Berhe, Pascal Boeckx, et al.. Can SOC modelling be improved by accounting for pedogenesis?. *Geoderma*, 2019, 338, pp.513-524. 10.1016/j.geoderma.2018.10.018 . hal-01904580

HAL Id: hal-01904580

<https://hal.science/hal-01904580>

Submitted on 7 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Title**

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

4

5 **Authors**

6 Peter Finke^{1,*}, Emmanuel Opolot^{1,2}, Jérôme Balesdent³, Asmeret Asefaw Berhe⁴, Pascal Boeckx⁵,
7 Sophie Cornu³, Jennifer Harden⁶, Christine Hatté³, Elizabeth Williams⁴, Sebastian Doetterl⁷

8

9 **Abstract**

10 Recent research results suggest that soil organic carbon mineralization and stabilization depend to a
11 substantial degree on the soil geochemistry and the degree of weathering. We hypothesized that
12 this dependence can be translated to decay rate modifiers in a model context, and used data from
13 the Merced chronosequence (CA, U.S.A., 100 yr - 3 Myr), representing a weathering sequence, to
14 test, on a 1000 year time scale, a simple soil organic carbon (SOC) model based on the RothC26.3
15 model concepts. Four information levels were identified: (1) known decay rates per model SOC pool
16 at individual chronosequence locations, obtained by calibrating the model to measured SOC-
17 fractions and measured fresh OC-inputs; (2) average decay rates per SOC-pool, corrected per
18 location with rate modifiers based on geochemical proxies; (3) uncorrected average decay rates per
19 SOC-pool; (4) uncorrected average decay rates per SOC-pool and averaged OC-inputs. A lumped root
20 mean square error (RMSE) statistic was calculated per information level. We found that using local
21 measurements of fresh OC-input led to a decrease in RMSE of near 15% relative to information level
22 (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.

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

30

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

32

33 **1. INTRODUCTION**

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

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

54 We test this hypothesis on soils from a chronosequence of soils developed on alluvial sediments
55 near Merced, California (Harden 1982, 1987) and use the SoilGen implementation (Finke and
56 Hutson, 2008) of the RothC-concept (Jenkinson and Coleman, 1994; Coleman and Jenkinson, 2005).
57 Essentially, RothC dynamically re-allocates organic matter over different pools having unique
58 decomposition rates (c.f. section 2.2). The implementation in SoilGen follows the RothC-concept,
59 but is discretized over soil depth. RothC was developed to describe SOC-evolution in agricultural
60 field trials covering ca. 150 years and in young parent materials, thus the development domain
61 (time, vegetation and parent materials) is not the same as the application domain in this study. As
62 the temporal range of this chronosequence is far greater than the range for which RothC-pools were
63 defined, the Inert Organic Matter (IOM) pool used in the RothC-model may not be truly “inert”.
64 Radiocarbon studies (Sanderman et al., 2016) have shown that measured ages of soil carbon do not
65 support the presence of truly “inert” organic matter. More generally, the mere existence of Inert
66 Organic Matter has been questioned by the biogeochemistry community (Schmidt et al., 2011). To
67 allow evaluation of the usefulness of IOM as a model pool, we use two versions of this model: with
68 and without IOM. On the Merced chronosequence, SOC-pools have been measured by Doettrl et al.
69 (2018) and mapped onto the pools of C identified in RothC: Resistant Plant Material (RPM),
70 Decomposable Plant Material (DPM), Biomass (BIO), Humus (HUM) and Inert Organic Matter (IOM);
71 described below.

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

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

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

77 2) we correlate these rates to soil characteristics at these sites to find good candidate
78 (“influencing”) soil properties to be used as rate-modifiers;

79 3) we predict the rate-modifiers with linear regression from the measured soil characteristics
80 at the sites. Linear regression may not capture all possible relationships, but as the data set
81 is small, exploring non-linear methods was not feasible. The predicted rate-modifiers are
82 applied to SoilGen and by comparing simulated to measured pools in extra 1000-year runs,
83 with all other inputs equal to step 1, we evaluate how well the measured DPM, RPM, BIO
84 and HUM pools can be reproduced. This tells us how well rate modifiers perform when local
85 measurements of soil variables are available;

86 4) we evaluate the effect of having site-specific (and likely soil-dependent) organic matter C-
87 inputs to simulation quality.

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

92

93 **2. MATERIALS AND METHODS**

94 **2.1. Merced Chronosequence data summary**

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

101 1981; Harden, 1987: p. A8). Parent materials are felsic, magmatic and metamorphic alluvial
102 sediments with fine sandy and silty texture, which were field described and analyzed for basic soil
103 properties (texture, color, structure, consistence) allowing calculation of the profile development
104 index (Harden, 1982; Supplementary Information: Table S1). Only the oldest site has substantial
105 amounts of quartz-rich pebbles. Age differences correspond to different weathering intensities, as
106 shown by element ratios, pedogenic iron contents and type of clay minerals (Harden, 1987; Doetterl
107 et al., 2018; Supplementary Information: Table S2). Over the chosen simulation time of 1000 years,
108 we consider the sites to be geomorphologically stable: The sites are on level terrain and surface
109 runoff and water erosion are negligible. Wind erosion is also negligible due to the vegetation cover.
110 Perched water tables of short duration may occur at site PM24II. Over longer time spans this
111 geomorphological stability may not be the case, however. Data used for simulation of soil formation
112 in general and the evolution of soil organic carbon (SOC) are:

- 113 (i) for model initialization: depth patterns of texture fractions, bulk density, (clay) mineralogy,
114 Cation Exchange Capacity (CEC), root density distribution;
- 115 (ii) for model calibration and verification: depth patterns of SOC and its fractions (see below)
116 and percentage Modern Carbon (pMC). The pMC is defined as $pMC = 100 * \frac{A_{SN}}{A_{ON}}$, where A_{SN}
117 is the specific activity (decay counts per minute) of the sample normalized to $\delta^{13}C$, and A_{ON}
118 is the specific activity of the oxalic acid standard normalized to $\delta^{13}C$;
- 119 (iii) to satisfy model boundary conditions: evolution of precipitation, potential evaporation,
120 temperature, litter-C input, bioturbation.

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

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

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

139 Bioturbation, the soil mixing activity by soil meso- and macrofauna or treefalls, is of importance
140 because of the associated vertical transport of soil matter. It was shown to be an important factor in
141 soil horizonation in both field- and simulation studies (e.g. Phillips, 2007; Finke, 2012). Bioturbation
142 also creates macro-pores, which influence CO₂-transport (Singer et al., 2001). No bioturbation data is
143 available for the study area. However, Wilkinson et al. (2009) reported values between 10 -50
144 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
145 Prairie, temperature between 6-19 °C). We decided to take 30 Mg ha⁻¹ y⁻¹ of bioturbation for all
146 chronosequence soils, equally distributed over the root zone compartments and assumed
147 bioturbation independent of soil fertility levels as major bioturbators in the area are mammals
148 (ground squirrels and pocket gophers) that do not consume soil materials.

149 At the sites, from a mass perspective, dust input is minor over the simulation period of 1000 years.

150 There are some indications (Aciego et al., 2017) that Phosphorus (P) inputs by dust may be
151 important. P-input would affect plant biomass production, but as litter production was a simulation
152 input based on measured ANPP, this fertilization effect by P was implicitly taken into account.

153

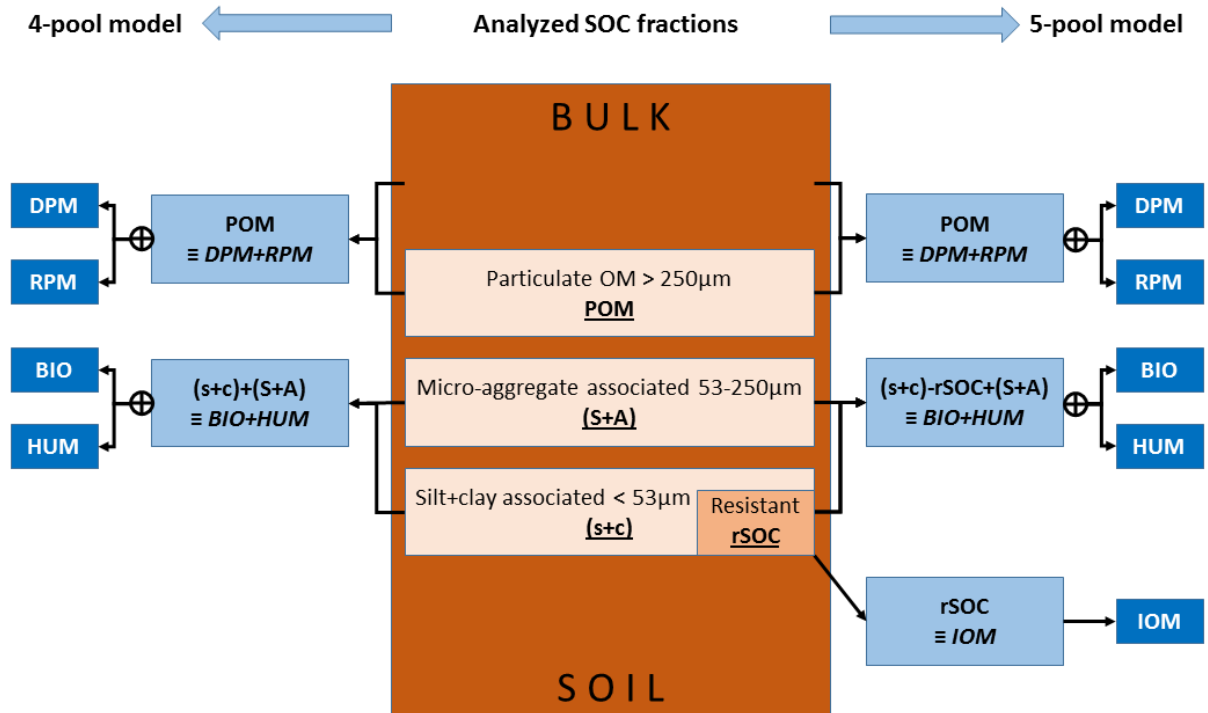
154 *Table 1- General data for Merced chronosequence (Harden, 1982, 1987), WRB2014 classification by the authors.*

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 (<i>Quercus lobata</i>)
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	

155

156 SOC fractions were measured in the bulk soil in topsoil and subsoil with the purpose to comply with
157 the four conceptual fractions proposed by Stewart et al. (2008): (i) free particulate organic matter
158 [POM], considered as unprotected; (ii) microaggregate-associated [S+A] SOC, considered as
159 physically protected; (iii) silt- and clay-associated [s+c] SOC, considered as mineral associated,
160 geochemically protected and (iv) nonhydrolyzable [rSOC] SOC equivalent to biochemically protected.
161 The chronosequence soils do not contain carbonates. Physical separation (sieving) techniques
162 (Doetterl et al., 2015) resulted in estimates of particulate organic matter (POM; >250 μm), C
163 associated to stable microaggregates (S+A; 53-250 μm) and C associated to non-aggregated clay and
164 silt (s+c; <53 μm). Additionally, a <63 μm fraction was ultrasonically dispersed (energy level of 22
165 J/mL) from a separate bulk sample and the dispersed solution was then wet sieved over a 63 μm
166 sieve. The suspension < 63 μm was filtered through a 0.45 μm nylon mesh and material >0.45 μm
167 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



181

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

188

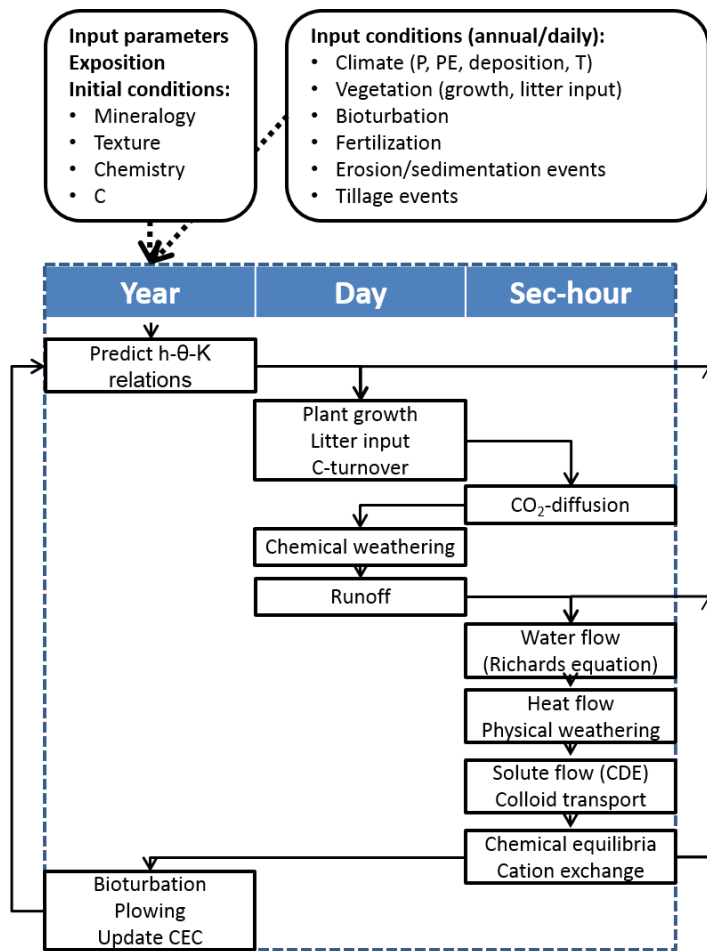
189 **2.2. Soil model**

190 The effect of mineralogy, climate and vegetation on the depth distribution of SOC over time was
 191 simulated with the SoilGen2.15 model (Finke and Hutson, 2008; Opolot and Finke, 2015).

192 SoilGen2.15 simulates flows of water, heat, solutes and CO₂ in unconsolidated geomaterials by
 193 numerically solving partial differential equations (the Richards equation, heat flow equation,
 194 advection-dispersion equation and CO₂-diffusion equation respectively), where the column is
 195 vertically discretized in equal compartments of five cm thickness in this study. Additionally, transport
 196 of solid matter occurs in the form of clay migration (a leaching process) and as a consequence of

197 bioturbation (a mixing process). Besides the flow of matter, for each soil compartment various sink
198 and source terms are calculated: (i) Soil texture is modified by physical weathering driven by
199 temperature fluctuations; (ii) soil mineralogical composition is modified by chemical weathering of
200 15 predefined and two user-defined silicate minerals; (iii) equilibriums of calcite and gypsum control
201 the precipitation or dissolution of these salts; (iv) amounts of SOC in the pools RPM, DPM, BIO and
202 HUM are dynamically calculated (see below). These combined processes mimic soil formation (Figure
203 2). The time step for water flow calculations is less than one hour (depending on the rainfall
204 dynamics), for heat flow and physical weathering it is one hour, for solute transport and dissolution
205 chemistry it is less than one day (depending on water flow dynamics) and for the SOC-cycle and
206 chemical weathering it is one day. Bioturbation is calculated for annual time steps. These time steps
207 are matching the dynamics of the individual processes, so that highly dynamic processes (e.g. water
208 flow) as well as relatively slow processes (e.g. weathering of minerals) are simulated efficiently.
209 Several studies confronted outputs of the SoilGen model in multi-millennium simulations to
210 measurements: Finke (2012), Sauer et al. (2012), Yu et al. (2013) Zwertvaegher et al. (2013), Finke et
211 al. (2015) and Keyvanshokouhi et al. (2016) quantified model accuracy at more than 100 locations
212 for SOC, texture, Cation Exchange Capacity, Calcite content, Base Saturation and pH, and concluded
213 fair to good performance. Keyvanshokouhi et al. (2016) concluded that the SoilGen model is suitable
214 for global change effect studies on soils.

215



216

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

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

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

- 227 • All incoming litter is divided by a fixed DPM/RPM ratio (0.67; Jenkinson and Coleman, 1994) into
 228 Resistant Plant Material (RPM) and Decomposable Plant Material (DPM) pools. Both these pools
 229 decay with rates k_{RPM} and k_{DPM}
- 230 • The resulting decay products from RPM and DPM are split over the Humus (HUM), Biomass (BIO)
 231 and mineralized (CO_2) pools using the clay content and a fixed BIO/HUM ratio (46/54; Jenkinson
 232 and Coleman, 1994).
- 233 • HUM and BIO decay with rates k_{HUM} and k_{BIO} into HUM, BIO and CO_2 . This cycling of SOC mimics a
 234 food web.
- 235 • All rates are modified by factors calculated using temperature and moisture deficit in each soil
 236 compartment, as in Jenkinson and Coleman (1994). Both temperature and moisture content are
 237 simulated at sub-daily timesteps. The moisture deficit is distributed over the root
 238 compartments, using the air-filled porosity per compartment as a proxy for the relative water
 239 stress. We assume no other depth dependencies of rates besides those related to differences in
 240 temperature and moisture deficit, even though there are indications that these do exist (see
 241 Mathieu et al., 2015, but also see Solly et al., 2015 with a contrasting conclusion for Beech forest
 242 soils). To mimic effect of the rhizosphere on abundance of micro-organisms, no SOC-decay is
 243 assumed to occur below the root zone. SOC-pools may be redistributed over depth as a
 244 consequence of bioturbation.
- 245 • The Inert Organic Matter pool (IOM) accounts for stable (non-decaying) organic matter, present
 246 at the start of the simulation period and inert during the period covered by the simulation.
- 247 All SOC-pools in SoilGen contain C, but also the cations and anions taken up by the vegetation via the
 248 transpiration stream migrate through the pools, thus finally not only CO_2 but also these ions are
 249 released (in the soil solution). Parallel to the C-cycle, a ^{14}C cycle is simulated. Litter inputs use the
 250 $\Delta^{14}C$ of the atmosphere at the year of input t (Hua et al., 2013; Reimer et al., 2013) to construct an
 251 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
 252 input in year t . The ^{14}C follows the same pathway as C, thus the ^{14}C pools degrade as the SOC-pools,

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

$$256 \quad \Delta^{14}C = \left(\frac{RMPx + DPMx + BIOx + HUMx + IOMx}{RPM + DPM + BIO + HUM + IOM} - 1 \right) * 1000 \quad (\text{eq.1})$$

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

$$258 \quad pMC = 100 * ((\Delta^{14}C / 1000) + 1) * e^{((y-1950)/(5730/LN(2)))} \quad (\text{eq.2}),$$

259 where y is the year of sampling.

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

263

264 **2.3. Research Layout and parameter fitting protocol**

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

- 272 • At the highest information level (info level 1), a site-specific calibration of the decay rate factors
273 for RPM, DPM, HUM and BIO is performed, using the local profile data to parametrize SoilGen
274 including local estimates of the SOC-input by plant litter. The calibration protocol follows a
275 downstream scheme, which means that the pathway of SOC-decay via RPM and DPM to HUM
276 and BIO determines the calibration order of the decay rate factors. Additionally, rate speeds are
277 considered (slower rates are calibrated first). This results in the calibration sequence
278 $k_{RPM} - k_{DPM} - k_{HUM} - k_{BIO}$. For each rate, eight equidistant values are taken in a range reported in

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

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

$$301 \text{ loss} = Y * (1 - e^{-xI * x2 * \overline{k_p} * t}) \quad (\text{eq.3}),$$

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

305 temperature, fluctuating over time and depth; x_2 is an additional rate modifier representing the
306 physico-chemical protection; t is the period of decay (1/365 year). The rate modifier x_2 gradually
307 changes over time because of weathering processes affecting geochemical soil properties. The
308 value of x_2 is calculated for each site i by:

$$309 \quad x_{2i} = \frac{\beta_0 + \beta_1 * G_i}{\bar{k}}, \quad (\text{eq.4}),$$

310 where β_0 and β_1 are regression coefficients, G_i is the value of the geochemical proxy at i and \bar{k} is
311 the average of the calibrated rate constants (y^{-1}) over all locations. In long-term simulations, G_i
312 would be time-dependent to reflect the effect of geochemical weathering processes. In
313 addition, local estimates of the SOC-input by plant litter are input to SoilGen. These simulations
314 applying the rate modifier (one per site) correspond to a second, lower information level (info
315 level 2) where no site-specific rate constants are known but the mineralogy and geochemistry is
316 known and can be used to modify rates.

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

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

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

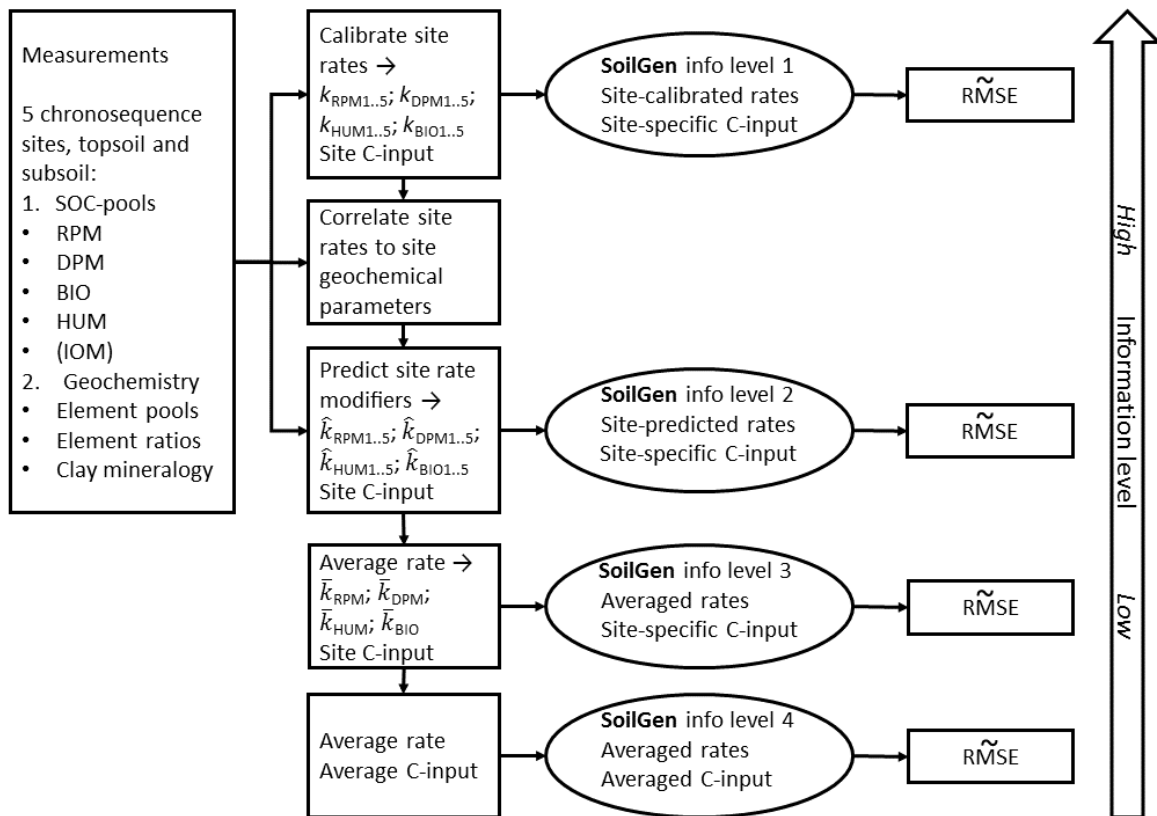
331 where S_p and O_p are simulated and observed SOC-amounts ($\text{Mg ha}^{-1} \text{cm}^{-1}$ soil) in pool p . We do not
332 consider the difference between measured and simulated IOM because both are either zero (four-
333 pool model) or equal (five-pool model) since IOM does not decay (Figure 1).

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

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.



347

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

351

352

353 3. RESULTS AND DISCUSSION

354 3.1. Information level 1: Site-specific calibration

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

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

374

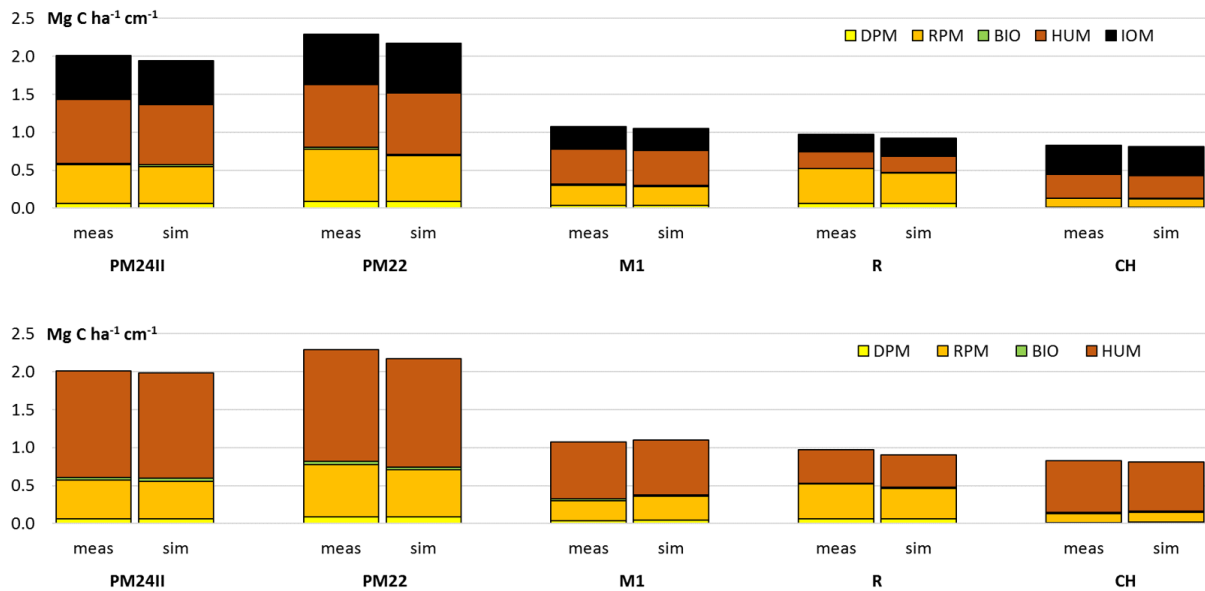
375 *Table 2- Litter-C input, calibrated rate constants (k, expressed in y^{-1}) and RMSE per site for the five-pool and the four-pool*
 376 *model. Italic values in brackets are calibration domain within which the optimal value for k was found. RPM, DPM, HUM,*
 377 *BIO are pools described in Fig.1.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

Site	C-input $Mg\ C\ ha^{-1}\ y^{-1}$	5-pool model					4-pool model				
		k_{RPM}	k_{DPM}	k_{HUM}	k_{BIO}	RMSE	k_{RPM}	k_{DPM}	k_{HUM}	k_{BIO}	RMSE
		<i>[0.075;</i> <i>0.525]</i>	<i>[0.100;</i> <i>1.600]</i>	<i>[0.005;</i> <i>0.035]</i>	<i>[0.100;</i> <i>0.800]</i>		<i>[0.075;</i> <i>0.525]</i>	<i>[0.100;</i> <i>1.600]</i>	<i>[0.005;</i> <i>0.035]</i>	<i>[0.100;</i> <i>0.800]</i>	
		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
						RMSE					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

379

380



381

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

385

386 3.2. Identification of influencing soil properties

387 Calibrated rate constants were correlated to various measured soil properties to identify candidate
 388 properties for rate modifiers (Table 3). Despite the small number of sites, strong and significant
 389 correlations were found. Results for the five-pool and four-pool models are similar. It should be
 390 noted that the results for the much larger pools RPM and HUM are more relevant than those of the
 391 volatile pools of DPM and BIO, which may be more responsive to climate fluctuations.
 392 First, important observations are the high and significant correlations between k_{RPM} (and k_{DPM}) and
 393 age and the profile development index (PDI, c.f. Supplementary Information; Harden, 1982). The PDI
 394 is assessed using basic field soil data, which makes it an easily available rate-modifier in other areas
 395 as well.

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

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

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

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

422 RothC26.3 correction for rates, was found for the four-pool model, which suggests that the

423 RothC26.3 correction for clay content may be improved.

424

425 *Table 3- Pearson correlation coefficients between calibrated rate constants (k_{RPM} , k_{DPM} , k_{HUM} and k_{BIO} , in y^{-1}) and*

426 *measured soil properties at five sites. Underlined: significant at $\alpha=10\%$; double underlined: significant at 5%; yellow-marked*

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

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

Pearson correlations	five-pool model					four-pool model			
	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)	<u>1.00</u>	1.00				<u>1.00</u>	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	<u>1.00</u>	1.00		-0.09	-0.11	<u>0.88</u>	1.00
IOM (Mg ha ⁻¹ cm ⁻¹ soil)	-0.14	-0.13	-0.67	-0.67	1.00				
Age (y)	<u>0.99</u>	<u>0.98</u>	0.02	0.09	-0.21	<u>0.99</u>	<u>0.98</u>	-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 (-)	<u>0.96</u>	<u>0.95</u>	0.13	0.20	-0.37	<u>0.96</u>	<u>0.95</u>	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 (%)	<u>0.89</u>	<u>0.88</u>	0.32	0.39	-0.43	<u>0.89</u>	<u>0.88</u>	0.14	0.34
2:1 clay mineral (%)	<u>-0.92</u>	<u>-0.91</u>	-0.26	-0.33	0.40	<u>-0.92</u>	<u>-0.91</u>	-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	<u>-0.81</u>	0.77	-0.47	-0.45	-0.66	<u>-0.82</u>
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

429

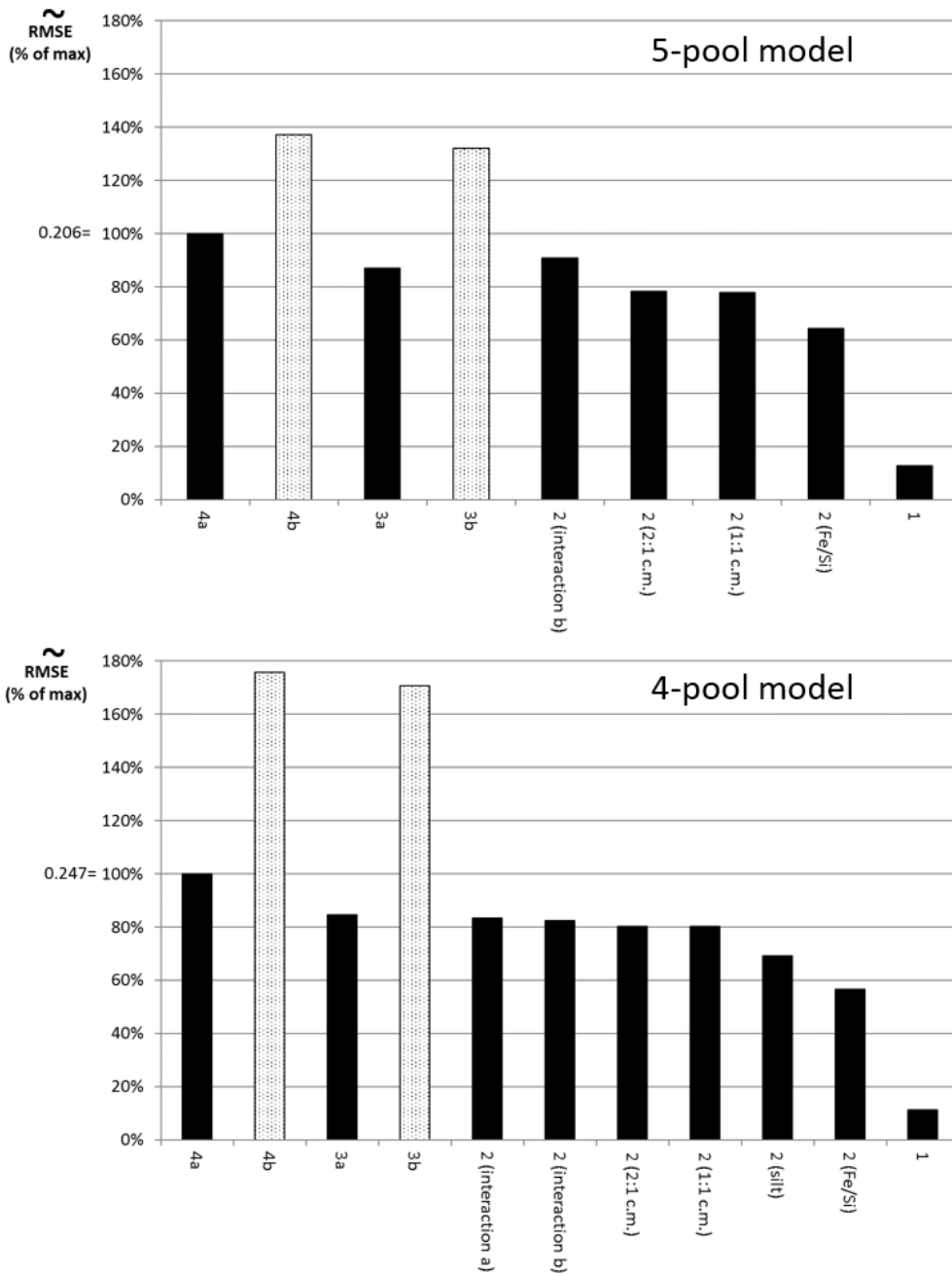
430

431 3.3. Simulation quality at different information levels

432 Soil properties with significant correlations to one or more calibrated rate constants (Table 3) were
433 considered suitable candidates to estimate rate modifiers x_2 (eq. 3) by linear regression. Thus, using
434 the contents of 1:1 clay minerals and 2:1 clay minerals, the Fe/Si ratio and (for the four-pool model)
435 the silt content, rate modifiers were estimated, the model was run and simulated pools were
436 compared to measured pools (information level 2). Interaction effects of suitable candidate
437 modifiers were also investigated (Fe/Si x 1:1 clay minerals and silt content x 1:1 clay minerals).
438 Results (Figure 5) clearly show that the \overline{RMSE} reduces (and quality increases) with addition of
439 information. The best results were obtained by local calibration (information level 1), and the worst
440 when rates as well as C-inputs were assumed the same at all sites (information level 4). We scaled
441 the \overline{RMSE} at information level 4a (average of measured rates and average of measured C-inputs) to
442 100% in the discussion below.

443 Estimating rate modifiers by geochemical parameters (information level 2) increased \overline{RMSE} by 45-
444 50% (compared to information level 1) for the best performing geochemical proxy, the Fe/Si ratio.
445 This ratio is a proxy for bonding in pedogenic hydroxides. Slightly worse results were obtained by
446 using 1:1 or 2:1 clay mineral content to modify rates. Interaction terms of clay mineralogy and Fe/Si
447 ratio did not give better (lower) values of \overline{RMSE} . From the Fe/Si based rate modifiers to the usage
448 of average rates (information level 3a) the \overline{RMSE} increased by about 20%, which illustrates the
449 added value of including geochemical properties in SOC-modelling, and the importance of
450 incorporating the degree of weathering in SOC-modelling. It should be noted that using the default
451 RothC-rates (information levels 4b and 3b) led to a much worse performance in terms of the \overline{RMSE}
452 (30% to 70% higher than at information level 4a), as these default rates added a bias component to
453 the error relative to the usage of locally-averaged rates. Thus, the 20% is a conservative estimate of
454 the gain in quality that can be realized by using rate modifiers, and usage of the default rate
455 constants of RothC introduces large errors in the simulations of the individual pools. Finally, at the

456 4th information level (4a) another substantial increase of ca. 15% in the \overline{RMSE} occurred when, in
457 addition to averaged rates, average values for Litter-C inputs are used instead of local values, a
458 situation corresponding to the usage of literature values. Again, the scenario 4b based on default
459 RothC-rates performs much worse than scenario 4a based on the average of the calibrated rates.
460 The decrease in \overline{RMSE} over the information levels from 4 to 1 show that using local estimates of C-
461 input contributes positively to simulation quality, though less than when geochemical proxies are
462 used as rate modifiers. C-inputs were found to decrease with greater soil age (Table 2), which is,
463 under the same climate, likely related to less available nutrients with increased degree of
464 weathering. This confirms that the SOC-model should include adequate C-inputs by either
465 accounting for soil fertility in simulating biomass production or by using local measurements.



466

467 *Figure 5- \widetilde{RMSE} at different information levels: 1=site calibration; 2=site rate modifiers (BS=Base Saturation, c.m.=clay*
 468 *mineral content, interaction a= silt x 1:1 clay mineral content, interaction b= Fe/Si x 1:1 clay mineral content); 3a=average*
 469 *rates; 3b (dotted bars)=default rates RothC26.3; 4a=average rates with average litter-C input; 4b (dotted bars)=default*
 470 *rates RothC26.3 with average C-input. High \widetilde{RMSE} denotes low model performance. Values for \widetilde{RMSE} at information level*
 471 *4a are given at 100% scale.*

472 The quality of simulations at all information levels depends on the root depth distribution. Since this
 473 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.

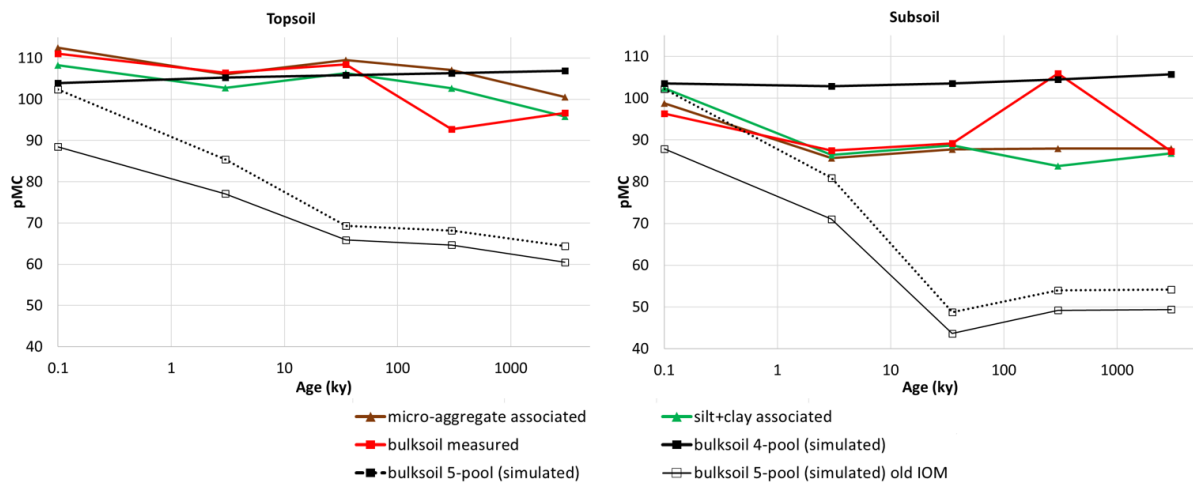
498

499 **3.4. Radiocarbon**

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

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

520



521

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

526

527 **CONCLUSIONS AND OUTLOOK**

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

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

544

545 ACKNOWLEDGEMENTS

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

549

550 REFERENCES

- 551 Aciego, S. M., Riebe, C. S., Hart, S. C., Blakowski, M. A., Carey, C. J., Aarons, S. M., Dove, N.C.,
552 Botthoff, J.K., Sims, K.W.W., Aronson, E. L. 2017. Dust outpaces bedrock in nutrient supply to
553 montane forest ecosystems. Nature communications, 8, 14800.
554 <http://dx.doi.org/10.1038/ncomms14800> .
- 555 Baldock, J.A., Skjemstad, J.O. 2000. Role of the soil matrix and minerals in protecting natural organic
556 materials against biological attack. Organic Geochemistry 31: 697-710.
557 [http://dx.doi.org/10.1016/S0146-6380\(00\)00049-8](http://dx.doi.org/10.1016/S0146-6380(00)00049-8)
- 558 Balesdent, J., Basile-Doelsch, I., Chadoeuf, J., Cornu, S., Fekiacova, Z., Fontaine, S., Guenet, B., Hatté,
559 C., 2017. Renouvellement de carbone profond des sols cultivés : une estimation par
560 compilation de données isotopiques. Biotechnologie, Agronomie, Société et Environnement
561 21 : 1-9.
- 562 Barré, P., Fernández-Ugalde, O., Virto, I., Velde, B. and Chenu, C..2014. Impact of phyllosilicate
563 mineralogy on organic carbon stabilization in soils: Incomplete knowledge and exciting
564 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 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](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., Hafliðason, 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., Schüllli-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](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

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

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

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

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

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

677 Yu, Y.Y., Finke, P.A, Guo, Z.T, Wu H. B., 2013. Sensitivity analysis and calibration of a soil carbon
678 model (SoilGen2) in two contrasting loess forest soils. *Geoscientific Model Development* 6,
679 29-44. www.geosci-model-dev.net/6/29/2013/

680 Zimmerman, M., Leifeld, J., Schmidt, M.W.I, Smith, P., and Fuhrer, J., 2007. Measured soil organic
681 matter fractions can be related to pools in the RothC model. *European Journal of Soil Science*
682 58: 658-667. <http://dx.doi.org/10.1111/j.1365-2389.2006.00855.x>

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

687
688

690 **Supplementary Information**

691

692 *Table S1: Basic soil data aggregated to the upper meter of soil to enable correlation with decay rates. Profile Development*

693 *Index calculated according to Harden (1982) for the closest characterized location in the same unit of the*

694 *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

695 Note: The Profile Development Index (Harden, 1982) combines eight soil properties estimated in the

696 field with soil thickness: (1) clay films, (2) texture plus wet consistence, (3) color hue and

697 chroma as proxy for rubification, (4) structure, (5) dry consistence, (6) moist consistence, (7)

698 color value, and (8) pH. After the eight field properties are quantified (method in Table II in

699 Harden, 1982), they are normalized onto a scale of 0-100%, where 100% indicates a strong

700 development of the property. and averaged per soil horizon. The result per horizon is

701 multiplied by the horizon thickness to give the soil profile soil development index.

702

703 *Table S2: Soil chemical data averaged for the upper meter of soil. Ka=kaolinite, Il=illite, Sm=smectite, Vm=vermiculite.*

704 *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

705 Explanation of the weathering data (see also Doetterl et al., 2018): Weathering from 2:1 to 1:1 clays

706 releases Si. We observed quartz-rich pebbles at the oldest site, which are likely

707 accumulations of Si released from weathering of the smaller particles (e.g. 2:1 to 1:1 clays).

708 Data (not shown, but see Supplementary Information for Doetterl et al. ,2018), show an

709 increase with age in pedogenic Fe, but not in total Fe. The common view is that weathering

710 releases Si and concentrates Fe and Al, but Harden (1988) concluded for these soils that this

711 does not occur at all particle size classes equally and that this explains the Si, Fe/Si and Al/Si

712 ratios observed in bulk samples.

713

714 *Table S3: Sampling depths, measured pools and calculated RothC-pools: DPM=Decomposable Plant Material;*

715 *RPM=Resistant Plant Material; BIO=Biomass; HUM=Humus and IOM=Inert Organic Matter. Analytical SOC-*

716 *fractions are POM=particulate organic Matter; (S+A)= C associated to stable microaggregates (53-250 μm); (s+c)=*

717 *C associated to non-aggregated clay and silt (<53 μm); and rSOC= resistant organic carbon measured after cold*

718 *acid hydrolysis (<63 μm). pMC=percent Modern Carbon. Analytical methods in Doetterl et al., 2018.*

Pofile	Sampled		Bulk density	Total SOC	Measured C-pools				SoilGen_RothC pools			pMC			
	Depth				POM	s+c	-rSOC	S+A	rSOC	DPM +	BIO +	IOM	Bulk SOC	s+c	S+A
	Top	Bot								RPM	HUM				
	(cm)	(cm)	(g cm ⁻³)	(g kg ⁻¹ soil)	(% of SOC)				(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 2018Harden, 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](https://doi.org/10.1016/0016-7061(82)90037-4)