



Estimating dissolved carbon concentrations in global soils: a global database and model

Joep Langeveld^{1,2} · Alexander F. Bouwman^{1,2,3} · Wim Joost van Hoek¹ · Lauriane Vilmin^{1,5} · Arthur H. W. Beusen^{1,2} · José M. Mogollón^{1,4} · Jack J. Middelburg¹

Received: 9 December 2019 / Accepted: 3 August 2020 / Published online: 4 September 2020
© The Author(s) 2020

Abstract

Dissolved carbon (C) leaching in and from soils plays an important role in C transport along the terrestrial-aquatic continuum. However, a global overview and analysis of dissolved carbon in soil solutions, covering a wide range of vegetation types and climates, is lacking. We compiled a global database on annual average dissolved organic carbon (DOC) and dissolved inorganic carbon (DIC) in soil solutions, including potential governing factors, with 762 entries from 351 different sites covering a range of climate zones, land cover types and soil classes. Using this database we develop regression models to calculate topsoil concentrations, and concentrations versus depth in the subsoil at the global scale. For DIC, the lack of a proportional globally distributed cover inhibits analysis on a global scale. For DOC, annual average concentrations range from 1.7 to 88.3 (median = 25.27) mg C/L for topsoils ($n = 255$) and from 0.42 to 372.1 (median = 5.50) mg C/L for subsoils ($n = 285$, excluding lab incubations). Highest topsoil values occur in forests of cooler, humid zones. In topsoils, multiple regression showed that precipitation is the most significant factor. Our global topsoil DOC model ($R^2 = 0.36$) uses precipitation, soil class, climate zone and land cover type as model factors. Our global subsoil model describes DOC concentrations vs. depth for different USDA soil classes (overall $R^2 = 0.45$). Highest subsoil DOC concentrations are calculated for Histosols.

Keywords Carbon leaching · Soil carbon · Vadose zone · DOC · DIC · Soil solution · Database · Global model

1 Introduction

1.1 Scientific problem

Terrestrial leached carbon (C) is commonly ignored in global C-cycle studies [1–4] leading to an inaccurate quantification of the terrestrial C budget [3, 5]. Meanwhile, several studies show that leaching and transport through groundwater of dissolved C is a major C source to surface waters [6–9]. These freshwaters are important active systems in the global C cycle, linking the land and

the ocean [10]. Global inland waters are estimated to process between 1.9 and 5.1 Pg C year⁻¹ [11], indicating the magnitude of the C flux into streams, rivers, lakes and reservoirs is fraught with uncertainties [11–14]. This C flux from global terrestrial systems comprises litterfall from vegetation in floodplains and riparian zones of primarily headwaters, surface runoff, leaching from soils and processing via groundwater to surface waters. Several of these pathways exporting C from the land commence in soils [15, 16]. The current study concerns C in this first part of

✉ Joep Langeveld, j.j.langeveld@uu.nl | ¹Department of Earth Sciences - Geochemistry, Faculty of Geosciences, Utrecht University, P.O. Box 80021, 3508 TA Utrecht, The Netherlands. ²PBL Netherlands Environmental Assessment Agency, P.O. Box 30314, 2500 GH The Hague, The Netherlands. ³Laboratory of Marine Chemistry Theory and Technology, Ministry of Education, Ocean University of China, Qingdao 266100, People's Republic of China. ⁴Department of Industrial Ecology, Leiden University, P.O. Box 9518, 2300 RA Leiden, The Netherlands. ⁵Deltares Delft, P.O. Box 177, 2600 MH Delft, The Netherlands.



the terrestrial-aquatic continuum: dissolved C in soil solution (soil pore water).

A global spatially distributed estimate of dissolved C in soil solution may aid in the development of process-based global models (e.g. [17–19]) to constrain the C flux from terrestrial systems into inland waters [11, 20, 21]. However currently, global maps of dissolved C in soil solution are yet lacking. In the light of this knowledge gap, we address the following research questions:

- How are dissolved C concentrations in soils spatially distributed on a global scale?
- How can we make a spatially explicit first estimate of dissolved C concentrations in soils on a global scale, in order to constrain large-scale C fluxes from soils?

Dissolved C in soil solution includes both organic (DOC) and inorganic C (DIC). In sampling studies, a range of factors has been identified as potential drivers of DOC and DIC concentrations in soil solutions. However landscape heterogeneity often implies that drivers at one location cannot be identified at another sampling plot. For example, while in several plot studies the soil carbon-to-nitrogen ratio (C/N) is found to relate positively to DOC concentrations [2, 22, 23], this is not recognised on the ecosystem level [24], while other studies identified even negative correlations [25, 26]. Inventories of dissolved C in soil solutions, which could provide insight into spatial patterns of dissolved C, generally focus only on a specific region or environment like grasslands [27], temperate forests [24, 28], specific forest types [29, 30] or other vegetation types [23, 31]. These studies therefore only represent a specific environment [32–34]. Similarly, modelling studies on dissolved C in soils focus on a specific site, region or specific environment [35–40], although a few models have been developed at the scale of large river basins, countries or parts of continents [17, 41–43]. [34] presented a first continent-wide data inventory analysis and regression model for DOC in soil solutions for mainly European forests [34, 44].

However, a global scale data inventory and analysis of DOC and DIC in soil solutions covering a wide range of vegetation types and climates is currently yet lacking, inhibiting identifying a set of drivers explaining dissolved C concentrations at that scale [33, 34, 45]. In this study we present the first global open-access database on annual average DOC and DIC concentrations and fluxes in soil solution (mainly DOC). The database covers different environmental conditions and includes a range of potential drivers. Using this database we developed global (multi-) regression models to calculate topsoil DOC concentrations, and DOC concentrations versus depth in the subsoil. Our models enable making a first estimate of spatially

distributed dissolved DOC concentrations in soil solution at the global scale. In a follow-up study (Langeveld et al. in prep.), these models will be included in the Integrated Model to Assess the Global Environment (IMAGE) Dynamic Global Nutrient Model (DGNM) [19] to model the global DOC fluxes from soils and also constrain the C fluxes from the terrestrial to the aquatic system.

1.2 Introducing dissolved C biogeochemistry in soils

DOC and DIC may be removed from soils through leaching and surface runoff [15]. In the unsaturated (or vadose) zone, dissolved C in soil pore water between the surface and the saturated groundwater zone is considered, where DIC and DOC concentrations generally change with depth during transport to aquifers [24, 46, 47]. A range of factors has been identified in studies as potential drivers of DOC and DIC concentrations in soil solution, including hydrology, land cover or vegetation type, climate, temperature, terrestrial C fluxes, and soil class (Table 2). In addition, a range of chemical and physical soil properties has been identified as controls of C leaching, such as soil organic carbon-to-nitrogen (C/N) ratio, soil organic carbon (SOC) content, pH, and texture (Table 2).

DOC in soil solution can increase through direct leaching from a higher layer, decomposition of organic matter, or desorption from soil particles [48]. DOC concentrations are highest in the upper organic soil horizon, declining by 10–50% in the subsoil [31, 46]. In most soils, apart from removal due to transport, prime controls of DOC concentrations are biodegradation in the topsoil and adsorption and biodegradation in the subsoil [24, 48–50].

Dissolved CO₂, bicarbonate and carbonate together make up DIC. In the upper horizons or topsoil these carbon species are constrained by the local atmospheric CO₂ concentration and soil pH. DIC stems from weathering, CO₂ dissolution and respiration by plants, animals and microbes using DOC and particulate organic carbon (POC) [8, 47, 51]. DIC concentrations typically increase with depth due to organic C biodegradation. Oxygen strongly impacts biogeochemical processes like decomposition in the unsaturated zone [46, 52, 53].

2 Methods

The first part of the method Sect. 2.1 describes the construction of the database. The intermediate Sect. 2.2 addresses how some of the limitations of the database are approached while preparing the data for a global analysis and subsequent global model construction. This section necessarily already takes some results of the database into

account. The final Sect. 2.3 addresses the statistical analysis and model construction. The database is not provided in the article but can be found in the Supplementary Information (SI), S11.

2.1 Database construction

We collected published studies reporting measurements of dissolved C in soil solutions in the unsaturated zone, covering 762 entries from 351 sampling sites distributed over the main climate zones (Fig. 1 and Table 1). From

these studies, all individual measurements (i.e., per location and depth) were included in the database, together with ancillary information, such as climate, soil and land cover. Studies were selected when reporting on DOC or DIC concentrations. Additionally, a few studies only reporting on DOC or DIC fluxes were added. When available, measurements on related forms such as biogenic DIC or alkalinity were also included. To be included in the database, data had to be easily available and shareable to the scientific community, so the database can stay open-source. For the database to be representative of large-scale

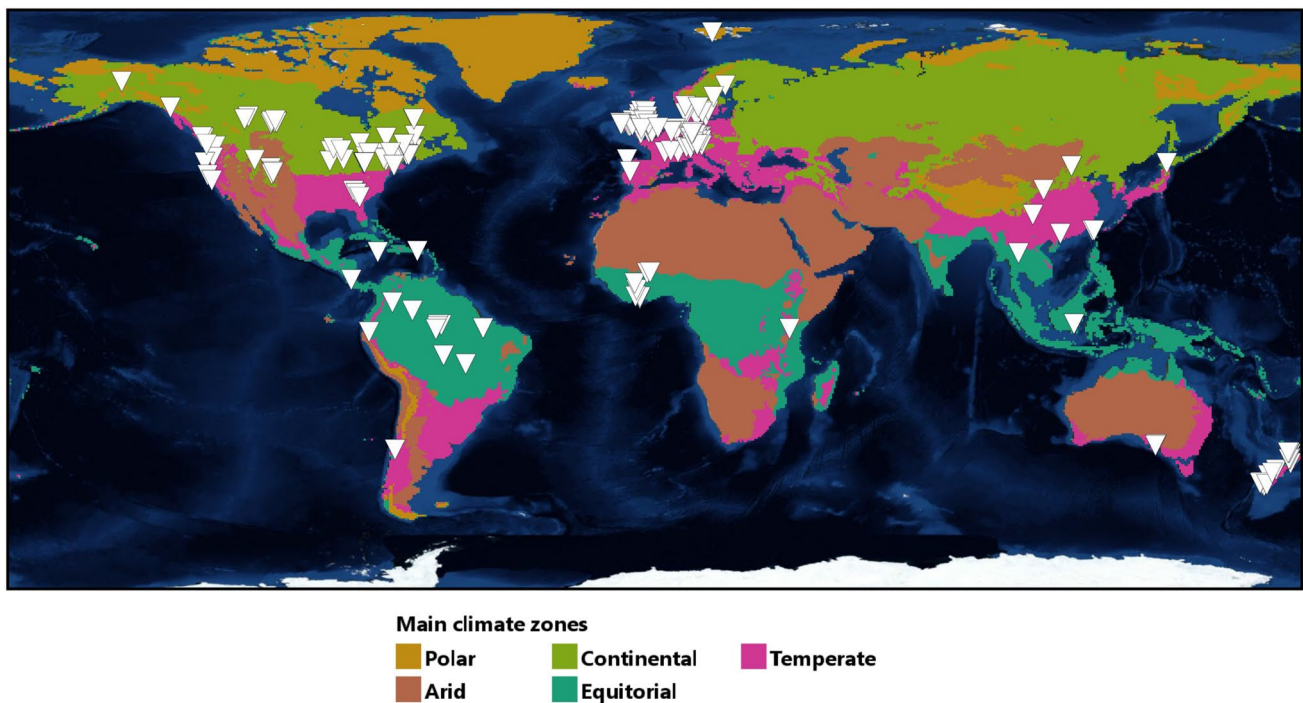


Fig. 1 Sampling sites reported in the database, distributed over main Köppen–Geiger climate zones [54]

Table 1 Number of entries and sampling sites the in database for all carbon leaching variables, and their coverage over all main climate zones

Dissolved carbon variable	Number of samples/entries	Number of sites	Number of main climate zones covered
DOC concentration (mg L^{-1})	550	229	5
DOC flux ($\text{g C m}^{-2} \text{ year}^{-1}$)	280	126	3 (continental, temperate, equatorial)
(Total) DIC concentration (mg L^{-1})	40	18	3 (continental, temperate, equatorial)
(Total) DIC flux ($\text{g C m}^{-2} \text{ year}^{-1}$)	29	14	2 (continental and temperate)
Biogenic DIC concentration (mg L^{-1})	4	4	1 (equatorial)
Biogenic DIC flux ($\text{g C m}^{-2} \text{ year}^{-1}$)	11	11	1 (temperate)
Alkalinity concentration (mg L^{-1})	82	33	4 (arid, temperate, equatorial, polar)
Alkalinity flux ($\text{g C m}^{-2} \text{ year}^{-1}$)	2	2	1 (equatorial)
DOC concentration—not yearly avg (mg L^{-1})	93	51	5
DOC flux—not yearly avg ($\text{g C m}^{-2} \text{ period}^{-1}$)	39	28	2 (continental and temperate)

Table 2 Meta-data overview, database selection of yearly DOC concentrations (550/762 entries)

Factor or variable reported	Topsoil		Subsoil		All		Studies that identify this factor as a main control of soil solutions DOC concentration ⁱ	Studies that identify this factor as a secondary or unimportant control of soil solutions DOC concentration ⁱ
	N _e	N _s	N _e	N _s	N _e	N _s		
General								
Entry ID	265	188	285	193	550	229		
Reference	265	188	285	193	550	229		
Location ID	265	188	285	193	550	229		
Sampling location	265	188	285	193	550	229		
Coordinates (reported or looked up) ^a	265	188	285	193	550	229		
First sampling year	265	188	285	193	550	229		
Last sampling year	265	188	285	193	550	229		
Sampling period/incubation time for lab studies	219	153	249	162	468	192		
Reported sampling method	256	179	277	185	533	218	No driver ^j	No driver ^j
Sampling method classification ^b	256	179	277	185	533	218	No driver ^j	No driver ^j
Measuring frequency	170	111	191	133	361	146		
Samples with more entries for this location ^c	265	188	285	193	550	229		
Environmental properties								
Temperature (°C) ^d	202	144	228	154	430	175	[24, 26, 45, 79–81]	[23]
Water drainage flux (mm year ⁻¹)	27	19	65	48	92	49	[24, 37, 73]	[4, 45]
Precipitation (mm year ⁻¹) ^d	191	139	221	142	412	164	[26]	
Reported climate zone	76	59	93	60	169	69	[68, 72]	
Climate zone classification ^b	265	188	285	193	550	229	[68, 72]	
Reported biome/vegetation	265	188	285	193	550	229	[32, 69–71, 82]	[24, 26, 83]
Land cover classification ^b	265	188	285	193	550	229	[32, 69–71, 82]	[24, 26, 83]
Soil properties								
Topsoil or subsoil (old definition)	265	188	285	193	550	229	See ‘Sampling depth’	
Topsoil or subsoil (updated definition)	265	188	285	193	550	229	See ‘Sampling depth’	
Reported soil type (class)	265	188	285	193	550	229	[4, 68, 85, 86]	
USDA soil classification ^b	265	188	285	193	550	229	[4, 68, 85–87]	
Other soil properties (texture)	147	104	166	109	313	123	[4]	
Soil texture classification ^b	141	98	162	105	303	117	[4]	
Reported horizon	224	160	214	151	438	188	See ‘Sampling depth’	
Horizon classification ^b	236	163	236	168	472	204	See ‘Sampling depth’	
Sampling depth (cm) ^e	249	172	267	180	516	206	[24, 31, 46, 48, 49, 88]	
C/N ratio soil ^f	92	67	68	49	160	79	[2, 22–24, 89] (positive) [25, 26] (negative)	[2, 24, 84]
Soil organic carbon content (%) ^g	67	55	89	60	156	71	[31, 90–92]	[2, 24, 26]
pH soil (occasionally in solution) ^h	227	160	202	143	429	185	[24, 93–95] (positive) [81, 96, 97] (negative)	[45, 73, 81]
Fe, Al in soil or soil solutions	83	65	69	59	152	72	[2, 31, 90, 95, 98–101]	[46, 95, 102]
Fe, Al in soil solutions (mg C L ⁻¹)	27	26	23	21	50	26		
Fe, Al in soil (g kg ⁻¹)	46	29	30	22	76	29		
Fe, Al in soil (%)	10	10	10	10	20	11		
Terrestrial C budget elements								
Heterotrophic respiration (g C m ⁻² year ⁻¹)	2	2	0	0	2	2	[23, 31, 37, 103, 104]	
NPP (g C m ⁻² year ⁻¹)	2	1	3	1	5	1		

Table 2 (continued)

Factor or variable reported	Topsoil		Subsoil		All		Studies that identify this factor as a main control of soil solutions DOC concentration ⁱ	Studies that identify this factor as a secondary or unimportant control of soil solutions DOC concentration ⁱ
	N _e	N _s	N _e	N _s	N _e	N _s		
NEP (g C m ⁻² year ⁻¹)	2	1	3	1	5	1		
NEE (g C m ⁻² year ⁻¹)	6	6	17	12	23	12		
NBP or NECB (g C m ⁻² year ⁻¹)	6	6	18	13	24	13		
Dissolved C (selected only yearly avg concentrations)								
Total DIC concentration (mg C L ⁻¹)	13	11	27	18	40	18		
DOC concentration (mg C L ⁻¹)	265	188	285	193	550	229		
Biogenic DIC concentration (mg C L ⁻¹)	2	2	2	2	4	2		
Alkalinity concentration (mg C L ⁻¹)	35	30	45	30	80	32		

N_e=number of entries; N_s=number of studies

NPP net primary production, NEP net ecosystem production, NEE net ecosystem exchange, NBP or NECB net biome production or net ecosystem carbon balance, without leaching. See also SI 1

^aLocation coordinates as reported in the study, or obtained from using the location description in the study

^bSee classification tables in SI

^cWhere measured at one sampling site at different depths, every observations is a single entry with the same sampling ID

^dMeasured on the site or from weather station data

^eSampling depth is reported when available. Shallow or surface level observations with actual depth not reported, while sampling depths are given for other soil layers, are assumed to be 2.5 cm

^fFor 3 entries C/N ratio in soil solutions is used instead

^gStudies report SOC using three different units. When converting from kg m⁻² to % we assumed a layer thickness of 10 cm and a fixed bulk density of 1.0 (topsoil) and 1.3 (subsoil), except for Andisols or Histosols (0.7 kg dm⁻³)

^hSoil pH measured in wet soil. Some measurements in CaCl₂ were transformed using formula 1 by [105]. A number of studies reports pH measured in solution instead of in soil (14% all pH data). As these are strongly related [85, 95], they are included when no soil pH is available

ⁱMentioned studies sometimes consider relations in only top- or a subsoil

^jThough not a driver, several studies discuss the impact of the sampling method

spatial trends in dissolved C content, without seasonal or temporary effects [4, 31, 45, 55], we included only annual average values. These annual averages were included in the database as reported, or calculated from at least four measurements representing shorter time periods covering at least 1 year. A number of measurements were found not to fulfill the latter criterion and are thus reported under a different category as non-yearly averages (See Table 1). Similar to [24] we included both flux-weighted and straight average annual concentrations, as on a yearly scale the difference between both methods is limited [56]. All selected data are from measurements in the unsaturated zone unless specified differently, for example for peat soils (Histosols) with a high water table.

Entries in the database for the same sampling site but for different depths have the same unique sampling ID. Thus, entries with the same sampling ID can be analysed for attenuation profiles over depth. All entries were initially reported as a topsoil if reported as such or, when unknown, when sampled shallower than 10 cm (when sampling depth available). Other soils were classified as

subsoils. As in studies the classification of either top or subsoils was not consistent, an alternative classification following [57] was also included in the database, classifying all samples measured within 20 cm as topsoils and all samples from below 20 cm as subsoils. This second classification enables comparison of the data within consistent soil layer classes, and is therefore used in our analysis.

Information on general, environmental and soil characteristics, soil properties or data on the terrestrial C budget were included when available (see Table 2 for the available meta-data for DOC concentrations, or the database in SI 1). Based on the descriptions in the literature, we identified classifications for soil horizons, United States Department of Agriculture (USDA) soil class [58, 59], vegetation and, when available, soil texture (see SI 1). Several climate classifications were extracted from global climate maps representing the second half of the 20th century according to [54] SI 1 and Fig. 1).

DOC concentrations in the subsoil were calculated as relative concentrations compared to those in the topsoil. This approach allows to include data from laboratory

experiments, since we assume biases inherent to incubation experiments are consistent in topsoil and subsoil. Relative concentrations were calculated for locations with data for various depths, provided that the shallowest depth is < 10 cm (with 10 cm taken as the median of the topsoil according to the classification following [57]).

In addition to the values reported in literature, data were extracted for 19 different parameters from International Soil Reference and Information Centre (ISRIC) soil data of the Harmonized World Soil Database (HWSD) [57, 60] for the corresponding grid cell coordinates of every sampling site in the database. We included these 30 second-resolution data ($1^\circ/120^\circ$), as well as different types of aggregations to 30 min (0.5°), in the topsoil analysis. Further, temperature and precipitation data [61] (30 min, long-term average + trend; data available from the climate database of the IMAGE model [62]) were extracted for every sampling site in the database for the corresponding grid cell. The database was compiled in Microsoft Excel, data were analysed with codes written in R or Python.

2.2 Limitations on a global scale

The distribution of data on potential drivers of DOC concentrations appeared unbalanced; the choice of factors included in sampling studies varies (Table 2). This is a problem earlier recognized on a smaller scale [48, 63]. As a result of these data gaps, for many factors, analysis is only possible on a limited part of the DOC concentration data. Even more, including several factors strongly reduces the amount of data involved (e.g. for topsoils, including pH, SOC and CN cuts the amount of data from 255 to 40), thereby impeding analysis at a global scale. Indeed, significant relations may be found for sub-sets of the data, but these relations only represent a certain limited environment. For example, for data exclusively from Spodosols (analysis including all soil types gave no significant results), a (relatively weak) relation can be found for C/N ratio versus DOC concentrations (SI 6). However, a model like this would only depend on a limited amount of data and is only valid for this soil type, based on data almost exclusively found in temperate forests. Such a model does not fit with the objective of our study. Data from soil databases may be an alternative; However, soil data from [60] (30 second-resolution data, as well as the dominant and mean for aggregated data to 30 min) yielded a poor correlation with observed topsoil DOC concentrations or other factors. For example for SOC, the Pearson correlation coefficient was not higher than 0.07 for the three extracted datasets.

We aim to constrain DOC concentrations on a global scale and thus not site-specific. A general aspect is that on such temporal scales integrated C cycle fluxes may be strongly related to 'average' biophysical conditions like in

many other large-scale models such as Century [64] and LPJmL [65–67]. This makes empirical relationships between DOC concentrations and environmental conditions useful for bridging the gap between site and landscape scales. Thus, from the database we selected those factors with a global data coverage and therefore had to ignore some soil properties and environmental factors with potential predictability. Instead we use, beside climate and soil data extracted from grid, factors such as soil class, which often capture overall environmental conditions that determine soil forming factors and the resulting soil physical and chemical characteristics [4, 68], and land cover/vegetation [32, 37, 69–71], as a proxy for the soil C cycle, and climate zones representing temperature and moisture conditions [68, 72]. In other studies similar generalized controlling factors have been proposed, such as "biological activity" [49], "microbial metabolism" [73] or a "physico-chemically dominated control" [45]. In that respect a C model using 'average' biophysical conditions is consistent with other global C models that use biome type as a basis for modelling global soil respiration [74] or climatic life zones for estimating global soil organic carbon pools [75].

2.3 Statistical analysis and model construction

All numerical variables were plotted against each other and studied using simple descriptive methods (pearson R, spearman R, mean, median, standard deviation), for all classifications such as land cover type, top/sub soil or soil type. The effect of the type of sampling method is also analysed. Several transformations were examined for all C data (ln, log10, square root, square, reciprocal). DOC concentration data by far have the largest global coverage (Table 1). Therefore, only yearly DOC concentration data were selected for the construction of the global model. We used Box plots, QQ-plots and histograms to present and analyse frequency distributions and to identify outliers. Because most data are not normally distributed, the Mann–Whitney U test was used to compare differences between subgroups. We constructed different models to calculate the annual average DOC concentrations in topsoil and subsoil.

The topsoil model is constructed in a combined forward/backward multiple-linear regression analysis using the 'step' function in R [76], correcting with the Akaike information criterion (AIC) for the number of factors included. Covariate non-linearity was examined using partial residual plots (PRP), taking first the variables that show the strongest non-linearity, following [77]. Added variable plots (AVP) were used to identify outliers (isolated largest residuals or largest partial leverages) [78] (See SI 7). The model performance was expressed by the coefficient of determination (R^2) and the root mean square error (RMSE).

The subsoil model was constructed as a relative attenuation function with depth starting from the topsoil concentration. Codes for data processing, regression analysis and model construction were developed in R or Python.

3 Results and Discussion

This chapter contains four sections. In the first Sect. 3.1, the general results of the database are presented and summarized. The Sects. 3.2 and 3.3 contain the results and analysis for topsoils and subsoils. Both sections first address the database results for the respective soil layer (Sects. 3.2.1 and 3.3.1). Subsequently, the models for topsoil (Sect. 3.2.2) and subsoil (Sect. 3.3.2) are presented and discussed. The final section addresses the limitations and application of the models.

3.1 General aspects and data selection

The database contains mainly data for annual average DOC concentrations and fluxes (72% and 37% of the database entries), with only limited data for DIC (< 11%) (Table 1). DIC data generally show concentration/flux profiles as expected, increasing with depth [8, 47, 51, 106] due

to less efficient exchange of CO₂ with the atmosphere [107, 108]. However, the DIC data are scant, which inhibits analysis and model development on a global scale. Therefore, we selected only DOC data for further analysis.

The number of DOC concentration data clearly exceed that of DOC flux data. More importantly, the global coverage of DOC concentration data is much larger as it covers all main climate zones, unlike the DOC flux data. This is of key importance, as we aim to assess dissolved C on a global scale. Therefore our further analysis focuses on DOC concentrations (Table 1).

We analysed the data according to the consistent topsoil–subsoil classification following [57]. Annual average DOC concentrations range from 1.7 to 88.3 (median = 24.43) mg C L⁻¹ for topsoils and from 0.42 to 372.1 (median = 6.65) mg C L⁻¹ for subsoils (excluding laboratory incubations). When excluding data from Histosols, median values are 25.27 and 5.50 mg C L⁻¹ for topsoil and subsoil respectively. DOC concentrations decrease with depth from the organic rich topsoils down to the subsoils (Fig. 2). Earlier, smaller dataset compilations from forests by Dalva and Moore [46] and Michalzik et al. [24] show similar frequency distribution profiles and concentration ranges. Both the whole data set and horizon sub-sets are positively skewed distributed (Fig. 2 and SI 3), similar to

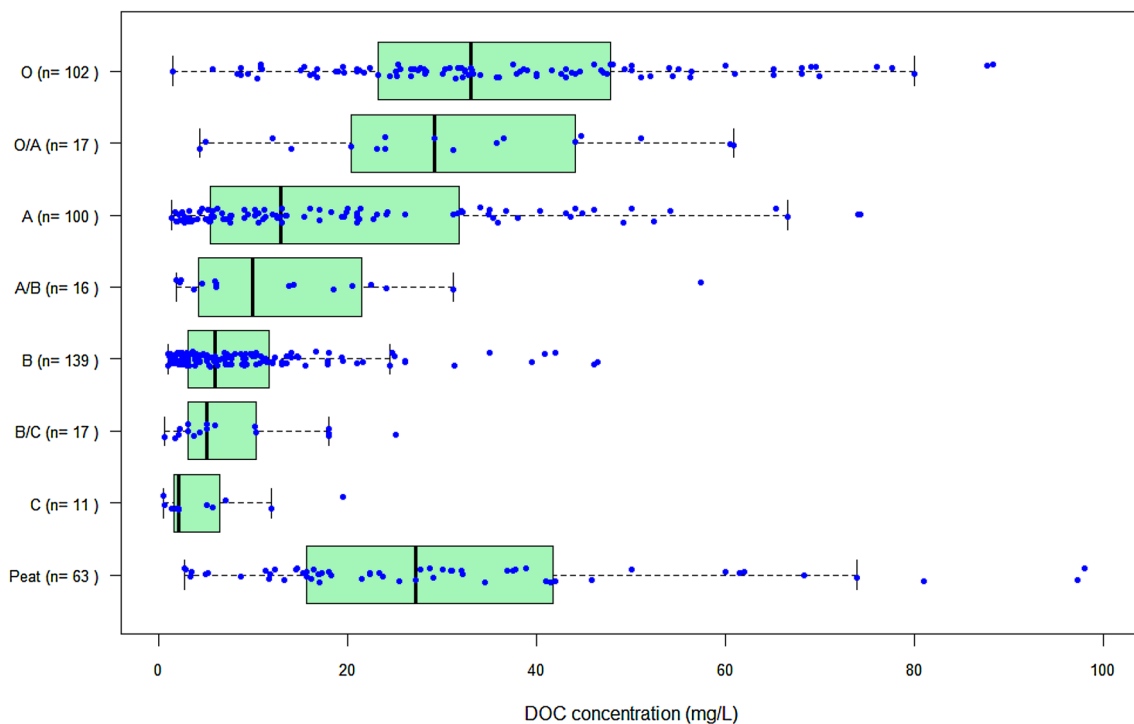


Fig. 2 DOC concentration (mg C L⁻¹) data distributed over horizon classes (n = 465, 85% of all DOC concentration data, laboratory incubation data are excluded). Boxplot bars are medians. Whiskers extend up to 1.5 times the interquartile range (IQR) from the lower

and upper quartile, unless exceeding the minimum or maximum. Individual data are jittered for visualization purposes. Five high non-lab-incubation values from peat samples (up to 372 mg C L⁻¹) are outside the figure range, but included in the distribution

Fig. 3 Top and subsoil DOC concentration (mg C L⁻¹) data distributed per measuring method (n = 550). Boxplot bars are medians. Whiskers extend up to 1.5 times the interquartile range (IQR) from the lower and upper quartile, unless exceeding the minimum or maximum. Circles are values exceeding the whiskers. Two subsoil values (up to 372 mg C L⁻¹) from Histosols are not shown, but included in the distribution. 'Passive device' is a shorter name for the group 'Passive bottle/well/tray/ditch/gauge/plate'

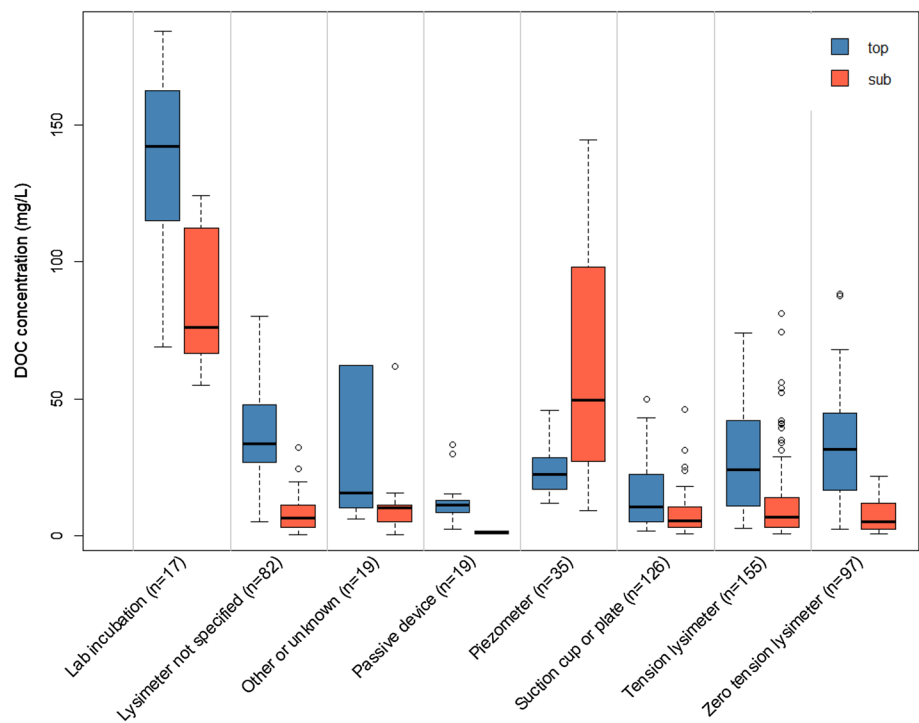
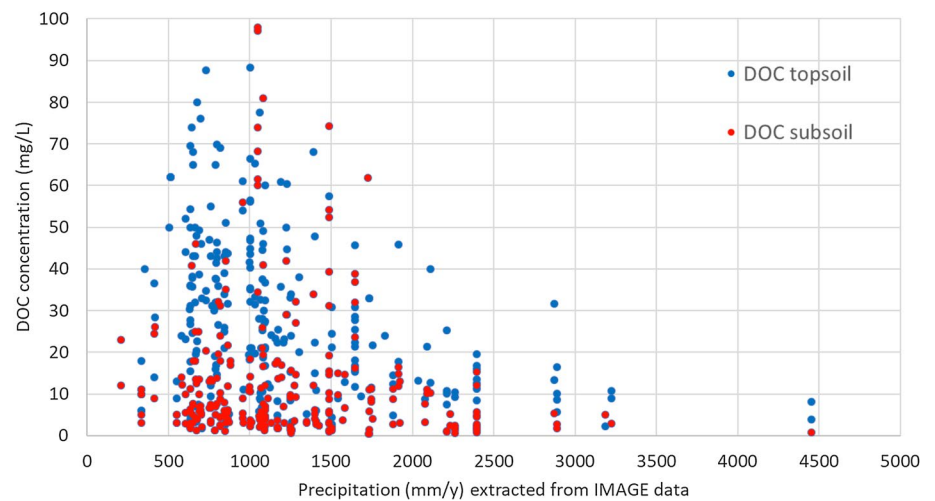


Fig. 4 Precipitation data from [61] versus DOC concentrations (mg C/L) (n = 533). Laboratory incubation data are excluded. Five DOC subsoil peat values (111.25–372.1) are not shown, but included in the regression. R² values for all data, topsoil data and subsoil data are respectively 0.0325, 0.129 and 0.0073, only significant for all data and topsoil data ($p < 0.0001$)



the distribution of several soil properties in global datasets [57, 60].

A range of different techniques has been used for measuring DOC concentrations. Laboratory experiments yield significantly higher concentrations in both topsoils ($p < 0.001$, Fig. 3) and subsoils ($p < 0.001$) than other methods, with five times higher median concentrations in topsoils. This can be attributed to soil sample disturbance and different conditions, that are often not representative for field conditions [32, 109–113]. In a review study, [48]

conclude that relations identified in laboratory studies can often not be confirmed in the field, in particular for subsoils [48]. Therefore, we used data from the laboratory experiments only for analysing the relative concentration changes with depth in subsoils, for which we use the retained fraction compared to the topsoil.

In subsoils, piezometer-based concentrations are much higher than for all other sampling methods ($p < 0.001$, Fig. 3), as these samples are almost exclusively taken in Histosols, where DOC degradation is inhibited by high

groundwater tables [85]. Where Zsolnay [114] and Buckingham et al. [56] found different concentrations between zero-tension and tension lysimeters [27], we did not observe this in neither top- ($p > 0.1$) or subsoil data ($p > 0.1$). Thus, both methods are combined in the analysis [30, 115].

For either top or subsoils, no single potential numerical driving factor showed a clear correlation with DOC concentrations in a simple (linear) regression (maximum $R^2 = 0.13$; see for example Fig. 4). This suggests that DOC concentrations are not driven by one uniform first-order governing factor, but by a set of interrelated drivers and controls that are spatially variable [32, 91]. For example [23] showed for temperate forest soils that both DOC production and heterotrophic respiration (HR) increase with temperature and also soil C/N ratios [23, 116].

Measured temperature and precipitation show a good correlation to corresponding climate data (R^2 of 0.94 and 0.49). As the measured data are not available for all measurements, we used the temperature and precipitation data from [61] instead of the measured values in our further analysis.

Concentrations of DOC in top and subsoils differ by soil class, land cover type and climate zone (Fig. 5). In contrast to other soil classes, Histosols generally contain higher DOC concentrations in the subsoil compared to the topsoil [53, 85, 117]. This has a clear confounding effect on some of the classified data, e.g. for the land cover type 'grass agriculture', which seem to have higher median concentrations in subsoils than topsoils (Fig. 5). However, when we exclude the Histosols, a more consistent pattern is shown (SI 3), with concentrations clearly decreasing with depth.

3.2 Topsoils

3.2.1 Database

For the topsoil (classification following [57]), 255 entries on DOC concentrations are included in the database (excluding 10 laboratory incubations). The specific horizon within the topsoil being studied can impact the amount of DOC measured. [23] identifies that in a range of forest floor samples, 74% of the DOC is from the organic horizon, 12% from litter and 13% from the deeper roots, emphasizing the role of the upper layers in DOC production. For our topsoil data, O, A and O/A horizons are dominantly present, with lowest DOC concentrations in the A horizon (Fig. 6). Despite being often from saturated soils, values from Histosols do not differ significantly from the other topsoil data ($p > 0.1$), so do not have to be filtered out in a further topsoil data analysis. Since B horizons are commonly subsoils we excluded them from the analysis of topsoil DOC concentrations.

All main climate zones are covered in the data, although there is a bias towards temperate and continental climates (Fig. 5), an issue recognized in earlier studies [32, 34]. Highest concentrations occur in humid continental climates and lowest in semi-arid climates. Also, oceanic climates have higher values than the three tropical climates ($p < 0.001$). Generally, we observe higher median concentrations in more moderate climates, with some of the warmer regions having lower concentrations (Fig. 5 and SI 3). This could be explained by the lower decomposition rates in temperate zones, caused by sub-optimal conditions for microbial degradation [23, 49, 72, 73]. Further, a limited or absent litter layer in semi-arid, tundra, and savannah climates restricts the amount of organic C available for decomposition. Where some studies identified different DOC concentrations between coniferous and deciduous forests [70, 71], this was not observed at the global scale ($p > 0.1$), despite the large number of data entries (Fig. 5), consistent with observations by Michalzik et al. [24]. We therefore aggregated the forest data into one main class (SI 1), as we can use climate zones to account for the climatic impact on the tropical and montane forest data.

3.2.2 Model

The topsoil multi-regression model was constructed involving the variables which are available for all DOC database entries. Therefore, factors like soil texture, C/N ratio or pH could not be included. On the basis of the data entries in the database (244 entries used for the topsoil model) and using the AIC to select the best fitting model, four factors are included in the model for calculating DOC concentration in the topsoil soil solutions:

$$DOC_{top} = \text{MAX}(0.0, 1.623 + coef_{CZ} + coef_{SC} + coef_{LU} - 0.008207P_{ann}) \quad (1)$$

where DOC_{top} is the DOC concentration in topsoil soil solutions (mg C L^{-1}), $coef_{CZ}$, $coef_{SC}$ and $coef_{LU}$ are the coefficients for respectively main climate zones, soil class and main land cover groups (Table 3), and P_{ann} is the annual average precipitation (mm year^{-1}) [61]. 31% of the variation is explained by the model (RMSE = 14.9, RSE = 15.5, 242 degrees of freedom). Neither including sub-classes for climate and land cover, nor transforming factors to non-linear functions yielded better model performance. Including soil data extracted from soil databases in the multi-regression analysis did neither yield a higher model performance.

The scatter plot of measured versus calculated DOC concentrations (Fig. 7) shows that the model describes the main trend reasonable ($R^2 = 0.36$). However, the

Fig. 5 Top and subsoil DOC concentration (mg C L^{-1}) data distributed over **a** USDA soil classes, **b** sub-climate zones according to [54] and **c** land cover groups. Laboratory incubation data are excluded, $n = 533$. Boxplot bars are medians. Whiskers extend up to 1.5 times the interquartile range (IQR) from the lower and upper quartile, unless exceeding the minimum or maximum. Circles are values exceeding the whiskers. Two subsoil values (up to 372 mg C L^{-1}) from Histosols are not shown, but included in the distribution. In **a**, data for soils with a double or mixed classification is represented in both single classes. See SI 3 for similar distributions as for **b**, **c**, but excluding Histosols

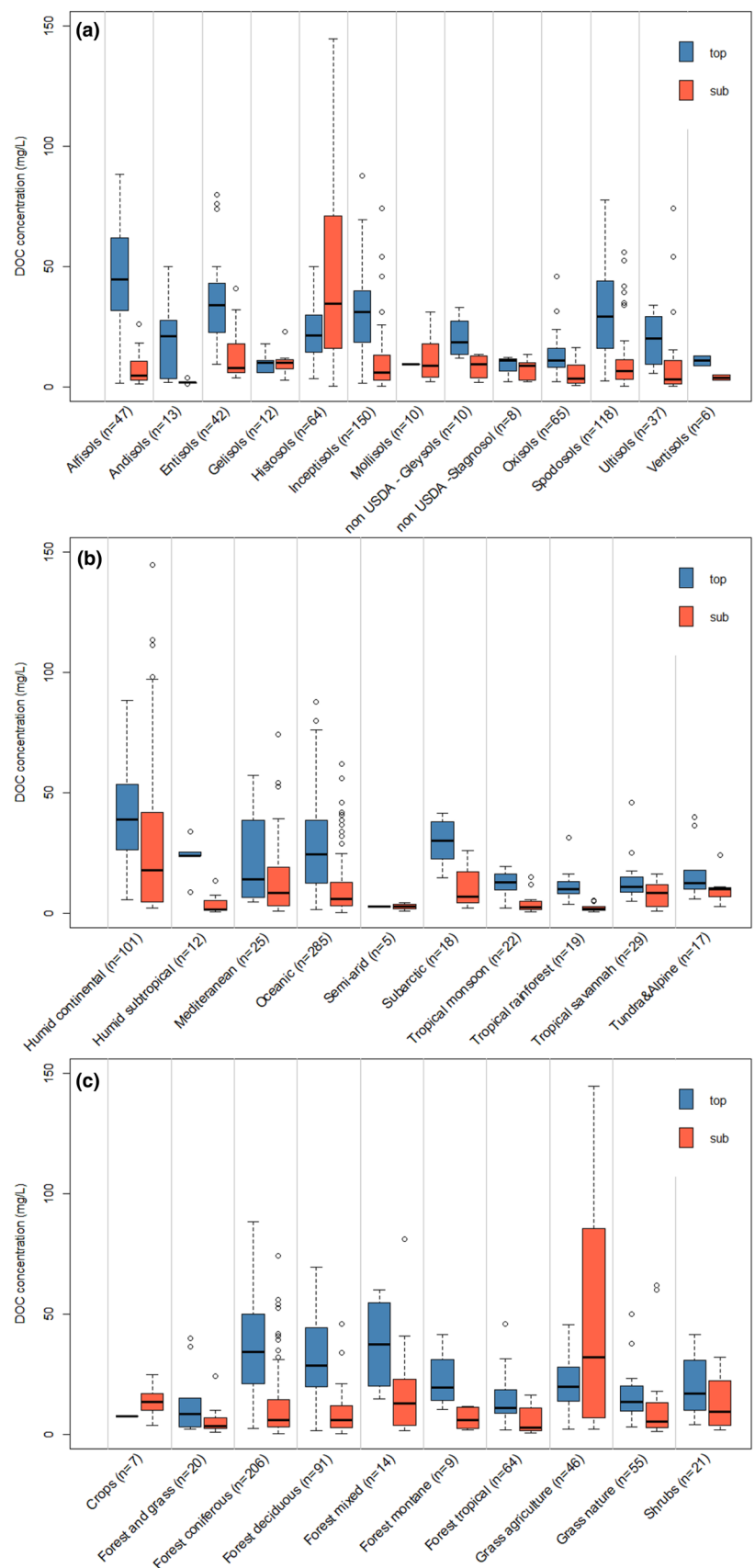
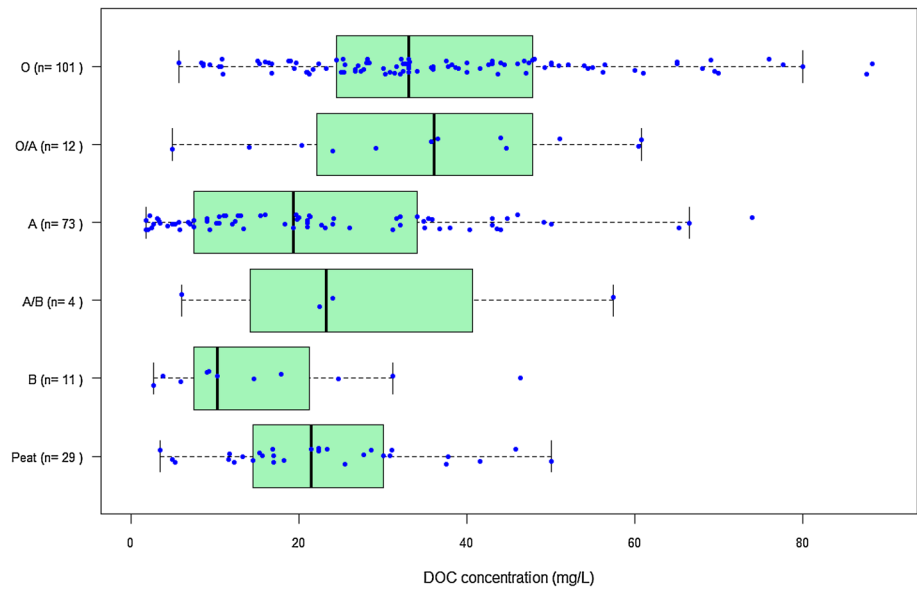


Fig. 6 Topsoil DOC concentration (mg C L^{-1}) data distributed over horizon classes ($n = 230$, 90% of all topsoil DOC concentration data). Laboratory incubation data are excluded. Boxplot bars are medians. Whiskers extend up to 1.5 times the interquartile range (IQR) from the lower and upper quartile, unless exceeding the minimum or maximum. Individual data are jittered for visualization purposes



model tends to overestimate lower concentrations for temperate and continental zones, while higher concentrations ($> 50 \text{ mg C L}^{-1}$) are consistently underestimated. The main reason is that a wide range of DOC concentrations occurs within a limited range of annual precipitation. Precipitation is negatively correlated to DOC concentrations ($p < 0.001$; Fig. 4), though the large residuals indicate considerable uncertainty. For a P_{ann} in the range

of $700\text{--}1000 \text{ mm year}^{-1}$ (mainly continental and temperate climates), DOC concentrations range from 1.7 to 87.6 mg C L^{-1} with high concentrations corresponding to the high values in continental and temperature climates shown in Fig. 7. The uncertainty may be due to the absence of information on the seasonality of precipitation and the multiple ways in which precipitation may influence concentrations, e.g. temporary accumulation

Table 3 Regression coefficients for the topsoil model in equation 1

USDA soil classes	$coef_{SC}$	Climate zones ^a	$coef_{CZ}$	Land use class	$coef_{LU}$
Alfisols	0.00	Arid (and semiarid)	0.00	Crops	0.00
Andisols	- 15.27	Continental	27.90	Forest	24.55
Entisols	- 2.07	Equatorial (tropical)	16.67	Forest grass shrubs	16.14
Gelisols	- 32.61	Polar (and alpine, montane)	29.36	Grass agriculture	25.08
Histosols	- 11.52	Temperate	21.03	Grass nature	12.20
Inceptisols	- 9.54				
Mollisols	- 19.21				
Oxisols	- 7.94				
Spodosols	- 6.44				
Ultisols	- 8.25				
Vertisols	- 25.76				
Aridisols	n.a. ^b				
Other ^c					
non USDA—Gleysols	- 14.28				
non USDA—Stagnosol	- 29.41				

^aMain Köppen–Geiger climate zones [54]

^bAridisols are not included in the database. As these soils occur in arid to semi-arid regions with sparse vegetation and very low microbial activity, we assume the DOC concentration to be equal to that in precipitation. For the concentration in precipitation, we calculated the median value from the overview by Aitkenheadsps-Peterson et al. [118]; 1.55 mg C L^{-1} . For comparison, this is up to 50% lower than the database DOC values, available for a steppe climate, but in oxisols. A fixed value of 1.55 mg C L^{-1} is also used for ice cover (land use or soil class maps), rock cover, shifting sands, salt plains (soil class maps) or hot deserts (land use maps)

^cFor some DOC data, only classification based on groundwater behavior (non-USDA) was possible

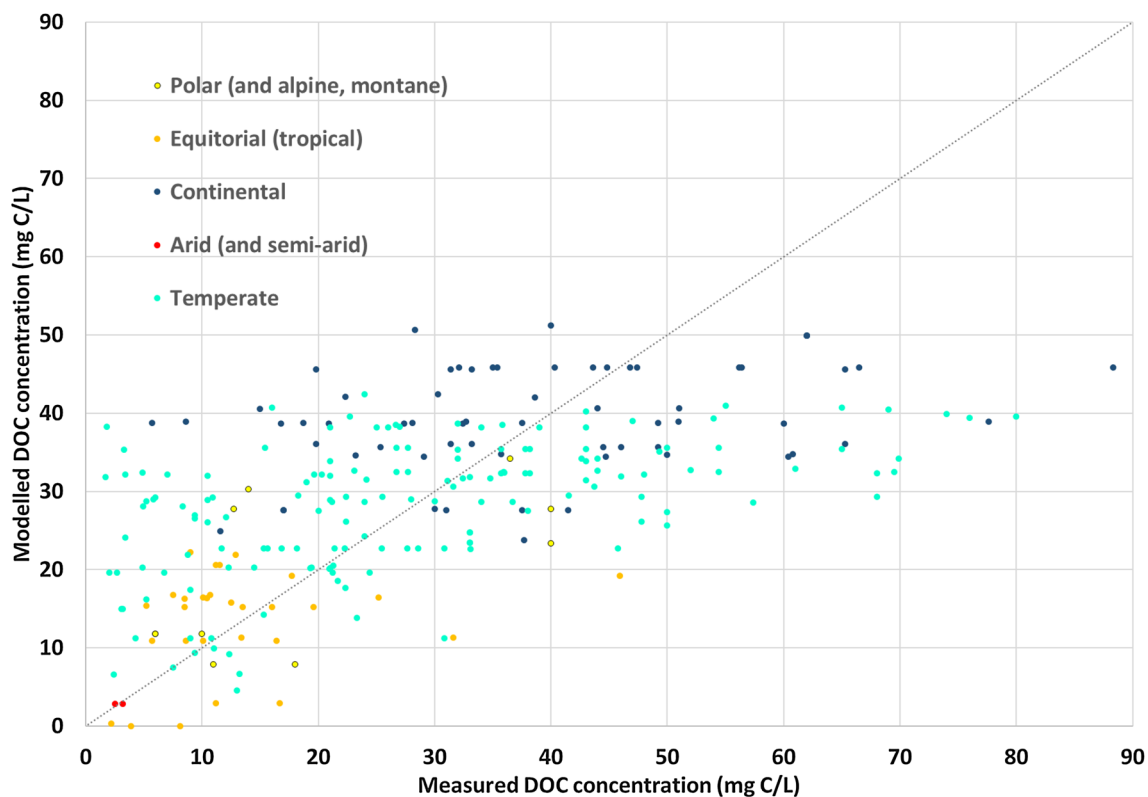


Fig. 7 Modelled topsoil DOC concentration (mg C L^{-1}) versus observed values ($n = 264$), distributed over the five main climate zones [54]. $R^2 = 0.36$, dotted line is 1:1 line. Entries with reported double/mixed USDA soil classes ($n = 20$) are duplicated to enable

including soil classes in the regression analysis. Laboratory incubation and B horizon data are excluded (see text). Entry #220 is also excluded as identified as a high leverage point in a partial regression analysis using an added variable plot (AVP)

in dry periods and flushing in wet seasons, dilution [4, 48, 119] or via soil moisture content [72, 116, 120]. Another cause of uncertainty is the poor spatial coverage of some governing factors, with the result that they could not be used for spatially dependent relationships. For example, C/N ratio may be an important factor in the case of Spodosols in temperate/continental climate zones ($R^2 = 0.36$), consistent with several observations [2, 22–24, 89].

Finally, we calculated the global distribution of topsoil DOC concentrations using grid data on land cover [62], precipitation [61], USDA soil classes from [59] and climate zones by Kottek et al. [54] (Fig. 8). For areas with Aridisols (soil class), ice cover (land use or soil class maps), rock cover, shifting sands, salt plains (soil class maps) or hot deserts (land use maps) we assume the DOC concentration to be equal to that in precipitation, using the median value from the overview by Aitkenheadsp-Peterson et al.

[118] (Table 3). Model results yield topsoil concentrations up to 51 mg C L^{-1} , with higher values generally in higher latitude continental zones with abundant forests. Lower concentrations occur in the arid zones with limited vegetation. DOC concentrations in equatorial regions, such as rainforests, are generally below $\sim 25 \text{ mg C L}^{-1}$. In an earlier overview [34] also found DOC concentrations in tropical topsoils to be significantly lower than those in boreal or temperate topsoils (see Fig. 1d [34]). This confirms that the effect of temperature on DOC is site-specific and possibly indirect [89], unlike results from some early forest studies [45, 80]. Probably, low tropical topsoil DOC concentrations are caused by optimal conditions for microbial degradation, reducing the amount of DOC before it leaches into the soil [23, 49, 73]. Topsoil concentrations in Histosols do not exceed those in mineral soils (Figs. 6 and 8), similar to earlier findings [34].

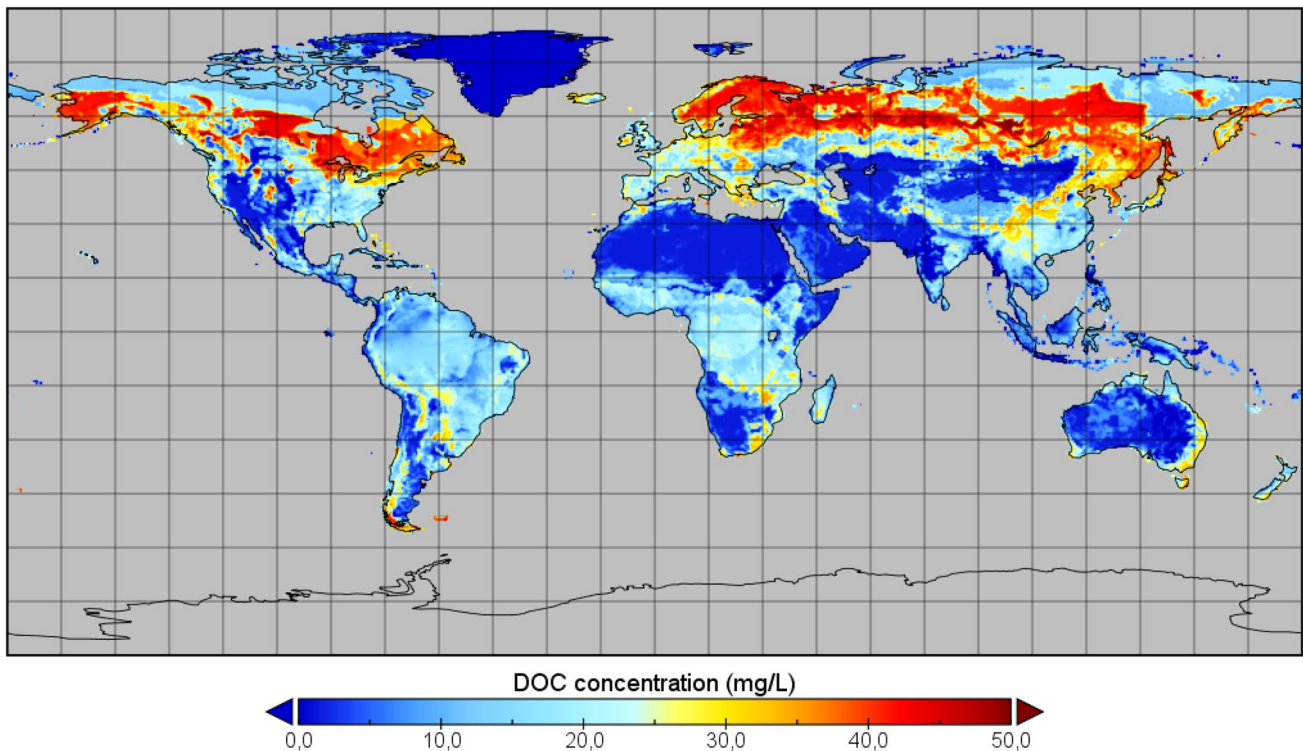


Fig. 8 Modelled global topsoil DOC concentrations in soil solutions (mg C L^{-1}) for the year 2000 with a 30 min resolution

3.3 Subsoils

3.3.1 Database

For the subsoil, 285 entries on annual average DOC concentrations are included in the database (including data measured in laboratory incubations). Of the reported horizons (not present for all DOC data; Fig. 9), B horizons are dominant, with only few data for O/A horizons present. Except for the limited amount of O/A and C horizon values, the data in the horizon classes have similar distributions and medians. Unlike for topsoils, subsoil DOC concentrations in Histosols strongly differ from other subsoils. Low biodegradation rates due to anoxic conditions, caused by high water tables, as well as the high organic matter content in Histosols result in high DOC concentrations at larger depths [85, 117].

When excluding Histosols, higher subsoil concentrations generally relate to higher topsoil DOC concentrations (Fig. 5), whereby the gradients are smaller for areas with low input due to low microbial activity, such as shrublands, tundra or permafrost (Gelisols) [23, 49, 72]. Soils under crops are a special case, with higher DOC concentrations in subsoils than in topsoils (Fig. 5). A possible explanation is DOC excretion by deeper plant roots [121]. However,

data for soils under crops are scarce and may not represent actual conditions for croplands, similar to [83].

3.3.2 Model

DOC concentrations in subsoils are regulated through biodegradation and the balance of adsorption–resorption processes [24, 48, 50]. Evidence of a strong influence of specific driving factors is mainly derived from laboratory experiments and less evident than for topsoils [48]. As physical–chemical factors are a main control of subsoil DOC concentrations [31, 48, 50], soil class can therefore be used as a proxy for physical–chemical conditions [100]. In addition, soil classes also partly account for differences in hydrological flow paths [86]. Several field studies identified differences in subsoil DOC concentrations between USDA soil classes [4, 68, 85, 86].

We calculate DOC concentrations in the subsoil as relative concentrations compared to the topsoil. In other words, the relative concentration in the subsoil is the quotient of a concentration at depth x relative to the topsoil concentration. Assuming a constant attenuation with depth within a soil class, an exponential decay function can be fitted for each soil class. Figure 10 shows the example for Spodosols, with the corresponding coefficients for all classes in Table 4. In

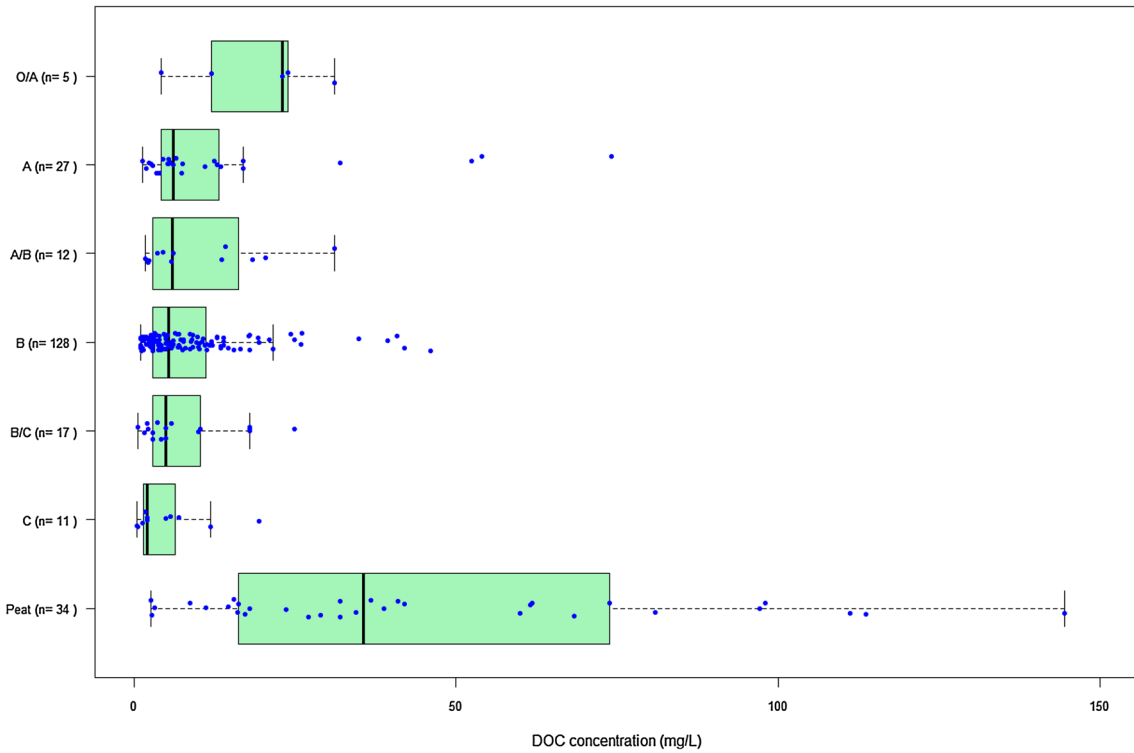


Fig. 9 Subsoil DOC concentration (mg C L^{-1}) data distributed over horizon classes ($n = 234$, 82% of all subsoil DOC concentration data). Laboratory incubation data are excluded ($n = 7$). Values for classes 'O' ($n = 1$) and 'groundwater' ($n = 1$). Two high peat values (372.1 and $222.5 \text{ mg C L}^{-1}$) are not shown but included in the distri-

buton. Boxplot bars are medians. Whiskers extend up to 1.5 times the interquartile range (IQR) from the lower and upper quartile, unless exceeding the minimum or maximum. Individual data are jittered for visualization purposes

Fig. 10 Data and model for DOC concentrations with depth, relative to topsoil concentrations. Example for Spodosol soil class data ($n = 81$). $R^2 = 0.74$ (exponential regression). Two values above 1 (triangles) are excluded from the regression

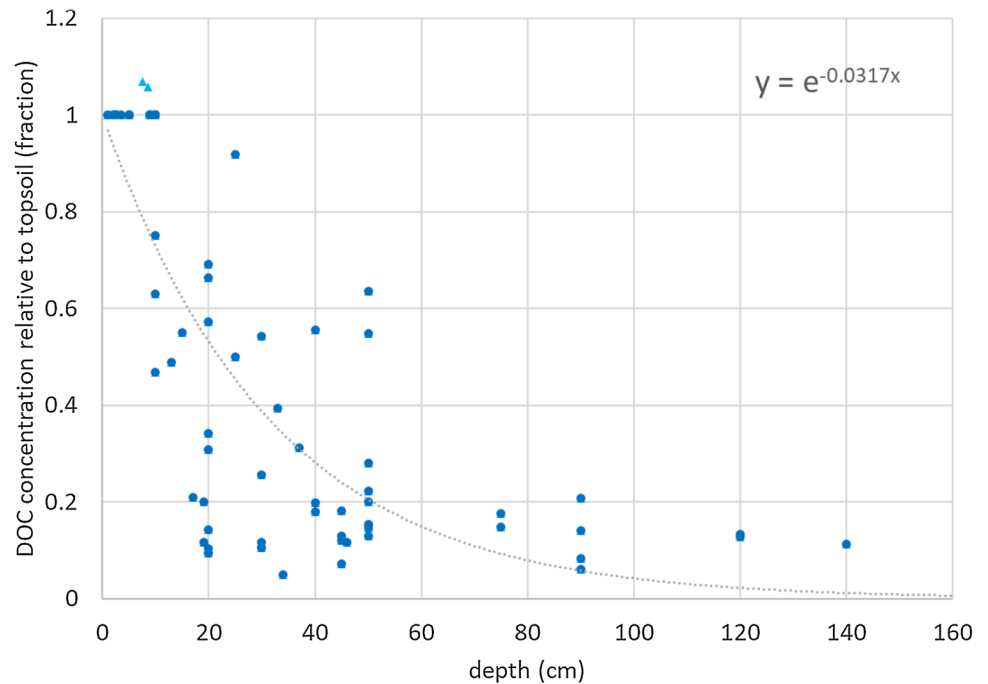


Table 4 Regression coefficients for the subsoil model in equation 2, based on the nonlinear (weighted) least-squares estimates

USDA soil class	k-value	n ^a	RMSE ^a	R ^b
All	-0.0267	368	0.39	0.35
Average: all except obvious Histosols ^b	-0.0276	355	0.32	0.45
Alfisols	-0.0436	34	0.26	0.74
Andisols	-0.0273	17	0.27	0.51
Entisols	-0.0186	42	0.25	0.58
Gelisols	-0.0098	10	0.14	0.43
Histosols ^c	0.0059	26	0.74	0.01
Inceptisols	-0.0298	115	0.37	0.41
non USDA—Gleysols	-0.0230	8	0.30	0.51
Spodosols	-0.0317	81	0.21	0.74
Ultisols	-0.0353	19	0.16	0.85
Vertisols	-0.0111	6	0.08	0.96
Oxisols	-0.0275	52	0.43	0.23
Mollisols—average value ^d	-0.0276	n.d.	n.d.	n.d.
other—average value ^e	-0.0276	n.d.	n.d.	n.d.

^an= amount of database values of this class for which a relative reduction can be calculated. RMSE = root-mean-square error, measure of the differences between values predicted and observed. $p < 0.01$, except for Gelisols (< 0.1) and Histosols (see ^c)

^bAverage coefficient of regression on the data, excluding values in peat. Mixed soil classes including a Histosol are included

^cOut of all subsoil Histosol data for which a relative quotient with depth can be calculated, 28% has a relative reduction below 1.0. We therefore conducted the regression over all data, resulting in a very low positive coefficient, causing a quasi-linear positive relation. The p value is quite high (0.11) due to the large spread of both positive and negative data

^dInsufficient subsoil data in Mollisols is available in the database; the average coefficient without Histosols is therefore used. n.d. means no data in the database available

^eThe USDA soil class map contains also the classes salt, shifting sand, rock and ice. Rock and ice are assumed to have no subsoil (NaN). For shifting sand and salt, the average coefficient without Histosols is used. n.d. means no data in the database available

Histosols, DOC concentrations in subsoils can be both smaller or larger than those in topsoils (Table 4). With a known topsoil concentration for a corresponding soil class, the DOC subsoil concentration at depth x is calculated as follows:

$$DOC_{sub} = DOC_{top} e^{(x \cdot coef_{SCsub})} \quad (2)$$

where DOC_{sub} is the DOC concentration at depth x in the unsaturated zone, DOC_{top} is the topsoil concentration as measured or modelled by equation 1, and $coef_{SCsub}$ is the soil class-dependent decay coefficient (Table 4).

For mineral soils, R^2 values of the regressions range from 0.23 (Oxisols) to 0.74 (Alfisols and Spodosols), up to 0.96 (Vertisols, but limited data) (Table 4). For some soil classes, in particular Vertisols, Gelisols and Andisols, the analysis is

based on a limited dataset (Table 4) reflecting the small area covered by these soils (Vertisols 2%, Andisols 1%) or the permanently frozen state of the soil (Gelisols) [58, 59]. Except for Histosols and Oxisols, the functions describe patterns of the relative concentrations vs. depth quite well with no consistent overestimation nor underestimation for undep soils. For deeper subsoils (> 1.0 m) the model seems to consistently underestimate DOC concentrations, for example due to subsoil biodegradation [48, 50], DOC production or exudation by plant roots [121] or DOC release due to desorption from soil colloids [27, 31, 99, 122]. In addition, the degradation of DOC could be overestimated by ignoring increasing recalcitrance of the organic C with depth [123].

Data and profiles for Oxisols and Histosols clearly differ from the other soil classes (SI 5) because in the strongly weathered Oxisols relatively high concentrations occur at greater depths due to high water percolation rates [86]. In Histosols, relative concentrations are observed to be both above (72% of the data) and below (28%) 1.0, with a very low R^2 (0.01), depending on the depth of the water table (see footnote c in Table 4).

Using Eq. 2, the modelled topsoil DOC concentrations (Fig. 8), k-values (Table 4) for USDA soil classes from [59] we calculated the global distribution of subsoil DOC concentrations at a depth of 1 m (Fig. 11). Modelled DOC concentrations at this depth range up to 266 mg C L⁻¹, though values above 20 mg C L⁻¹ only occur in areas with abundant Histosols. Highest subsoil concentrations in non-Histosols are between 5 and 10 mg C L⁻¹ and mainly found in permafrost areas (Fig. 8).

The high concentrations in subsoils of Histosols are consistent with observations by Camino-Serrano et al. [34], who attributed the high DOC concentrations in subsoils to the low hydraulic conductivity in these soils. Histosol concentrations, in particular in subsoils, dominantly depend on the water table depth [53] and therefore the thickness of the oxic zone, and can thus both increase or decrease with depth [85, 117]. High concentrations in subsoil Gelisols are similar to observations of [124, 125] and caused by reduced DOC degradation due to low microbial activity in the permafrost [126].

Our subsoil model uses only a single factor, soil classes to estimate subsoil DOC on a global scale. Although it is of course only a simple model aimed to make a first estimate, one should note that other factors such as climate effects and vegetation are inherently included in the model. For example, Oxisols and Ultisols are typical subtropical and tropical soils [86], while Aridisols are typical for (semi) deserts with generally low vegetation coverage [59, 62]. Thus, the soil type not solely represents soil characteristics,

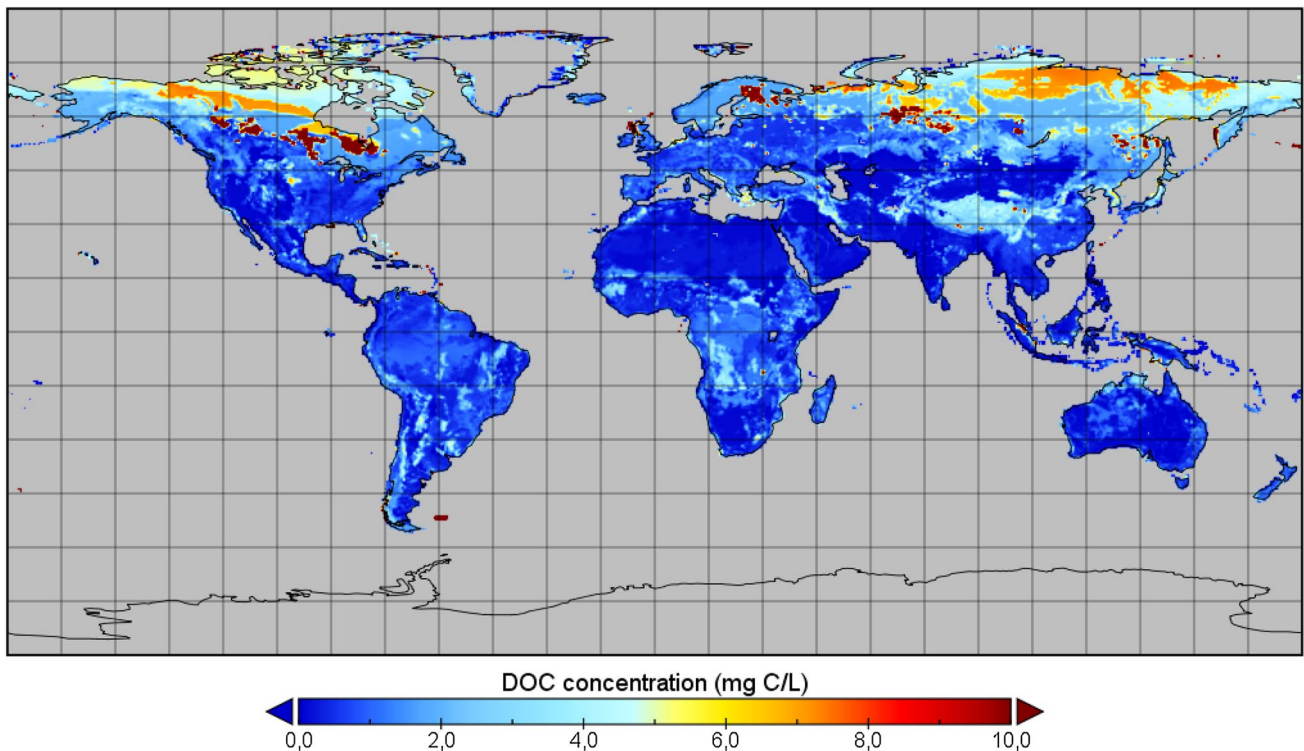


Fig. 11 Modelled global subsoil DOC concentrations in soil solutions (mg C L^{-1}) at a depth of one meter. Based on equation 2 using topsoil data from Fig. 8

but may indirectly also incorporate the effects of environmental and climatic conditions on DOC. As such, the subsoil model, like the topsoil model, may be viewed as a simplified but effective first step in estimating DOC concentrations on a global scale.

3.4 Limitations, application and perspective

Our global database on dissolved C in soil solutions does not yet include some other factors that have been identified in the literature as potential controls of DOC. For example, cation exchange capacity (CEC) [127], terrestrial acid deposition [128], anion deficit [97, 129], soil specific surface area [122] or the composition of the DOC [52, 130]. Still, only few studies incorporated in our database include a broad range of parameters, measured in a similar way. To enable further in-depth analysis and to include additional process controls on a global scale, a standardized set of ancillary data and uniform sampling method is required. The ICP Forests program set up such a framework for monitoring in European forests [131], which enabled recent in-depth analysis and model construction on this sub-continental scale [e.g., 34, 44, 43, 128, 132].

Our model simulates DOC concentrations at the global scale, which is a major step forward compared to current, recent large-scale (country, large basin, sub-continental region) models [e.g. 17, 41, 42, 43]. Though the temporal scale is 1 year, temporal downscaling of the yearly average modelled DOC concentrations could be based on relative seasonal variability [e.g. 128].

This model will be included (Langeveld et al. in prep.) in the Integrated Model to Assess the Global Environment (IMAGE) Dynamic Global Nutrient Model (DGNM) [19]. When hydrology [e.g. 133, 134] is known, we can attempt to model the global DOC fluxes and also constrain the C fluxes from terrestrial to aquatic systems on a global scale.

4 Conclusions

We present the first global database on annual average DOC and DIC in soil solutions, covering all main climate zones. As data on DIC are scant, we conducted our analysis and model construction on annual average DOC concentrations. Highest topsoil DOC concentrations occur in forests in humid continental climates, while topsoils of Histosols do not have higher DOC concentrations than other

soil classes. In contrast, highest concentrations in subsoils occur in Histosols. Our analysis shows that DOC concentrations are controlled by a complex of processes that vary in space. DOC concentrations from laboratory experiments are consistently higher than values found in the field.

We identified a set of four indirect controls of global topsoil DOC concentrations, i.e. precipitation, climate zones, vegetation types and soil classes. Further, our analysis showed that global subsoil DOC concentrations vs. depth can be modelled for all USDA soil classes. Here, soil class represents generalized physico-chemical properties that are not represented when only climate or land cover are used. Future sampling studies on DOC should be conducted in regions with land cover types currently under-represented, such as crops, preferably over different soil classes. A standardized set of ancillary data and uniform sampling method would enable further constraining of global dissolved C concentrations in soil solutions.

5 Supplementary material and data availability

Supplementary material, including the database, is publicly available on PANGAEA (<https://www.pangaea.de/>).

Acknowledgements This research was funded by the NWO New Delta 2014 project no. 869.15.015. W.J.H was funded by the NWO New Delta 2014 project no. 869.15.014. A.F.B. and A.H.W.B. received support from the PBL Netherlands Environmental Assessment Agency through in-kind contributions to The New Delta 2014 ALW projects no. 869.15.015 and 869.15.014. L.V. was funded by the Earth and life sciences (ALW) Open Programme 2016 project no. ALWOP.230 We are grateful to Marta Camino-Serrano (Centre for Ecological Research and Forestry Applications, CREAM) for providing us with a list of several studies to include in the database, which we acknowledge her by mentioning it in the database references concerned. We further thank Peter Janssen (PBL) and Maarten Zeylmans Van Emmichoven (Physical Geography - UU) for assisting in the analysis design and data extraction from maps respectively.

Author contribution JL, AFB, AHWB and JJM designed the study. Database compilation was conducted by JL. Code writing and data extraction from grid was done by JL, AHWB, JMM and LV. All authors contributed to the analysis and discussion of the results. JL wrote the manuscript. AFB, LV, JJM contributed to the revision of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare no conflict interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate

if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Clark DA, Brown S, Kicklighter D, Chambers J, John RT, Ni J (2001) Measuring net primary production in forests: concepts and field methods. *Ecol Appl* 11(2):356–370
- Kindler R, Siemens JAN, Kaiser K, Walmsley DC, Bernhofer C, Buchmann N, Cellier P, Eugster W, Gleixner G, Grünwald T (2011) Dissolved carbon leaching from soil is a crucial component of the net ecosystem carbon balance. *Glob Change Biol* 17(2):1167–1185
- Luyssaert S, Ciais P, Piao SL, Schulze E-D, Jung M, Zaehle S, Schelhaas MJ, Reichstein M, Churkina G, Papale D (2010) The European carbon balance. Part 3: forests. *Glob Change Biol* 16(5):1429–1450
- Don A, Schulze E-D (2008) Controls on fluxes and export of dissolved organic carbon in grasslands with contrasting soil types. *Biogeochemistry* 91(2–3):117–131
- Kling GW, Kipphut GW (1991) Arctic lakes and streams as gas conduits to the atmosphere: implications for tundra carbon budgets. *Science* 251(4991):298–301
- Marín-Spiotta E, Gruley KE, Crawford J, Atkinson EE, Miesel JR, Greene S, Cardona-Correa C, Spencer RGM (2014) Paradigm shifts in soil organic matter research affect interpretations of aquatic carbon cycling: transcending disciplinary and ecosystem boundaries. *Biogeochemistry* 117(2–3):279–297
- Hotchkiss ER, Hall Jr RO, Sponseller RA, Butman D, Klaminder J, Laudon H, Rosvall M, Karlsson J (2015) Sources of and processes controlling CO₂ emissions change with the size of streams and rivers. *Nat Geosci* 8(9):696–699
- Winterdahl M, Wallin MB, Karlson RH, Laudon H, Guist M, Lyon SW (2016) Decoupling of carbon dioxide and dissolved organic carbon in boreal headwater streams. *J Geophys Res Biogeosci* 121(10):2630–2651
- Rasilo T, Hutchins RHS, Ruiz-González C, Giorgio PA (2017) Transport and transformation of soil-derived CO₂, CH₄ and DOC sustain CO₂ supersaturation in small boreal streams. *Sci Total Environ* 579:902–912
- Cole JJ, Prairie YT, Caraco NF, McDowell WH, Tranvik LJ, Striegl RG, Duarte CM, Kortelainen P, Downing JA, Middelburg JJ, Melack J (2007) Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems* 10(1):172–185. <https://doi.org/10.1007/s10021-006-9013-8>
- Drake TW, Raymond PA, Spencer RGM (2018) Terrestrial carbon inputs to inland waters: a current synthesis of estimates and uncertainty. *Limnol Oceanogr Lett* 3(3):132–142
- Aufdenkampe AK, Mayorga E, Raymond PA, Melack JM, Doney SC, Alin SR, Aalto RE, Yoo K (2011) Riverine coupling of biogeochemical cycles between land, oceans, and atmosphere. *Front Ecol Environ* 9(1):53–60
- Ciais P, Sabine C, Bala G, Bopp L, Brovkin V, Canadell J, Chhabra A, DeFries R, Galloway J, Heimann M (2013) Carbon and other biogeochemical cycles. In: *Climate Change 2013 the physical science basis: working group I contribution to the fifth assessment report of the intergovernmental panel on climate*

- change, volume 9781107057. Cambridge University Press, pp 465–570. <https://doi.org/10.1017/CBO9781107415324.015>
14. Regnier P, Friedlingstein P, Ciais P, Mackenzie FT, Gruber N, Janssens IA, Laruelle GG, Lauerwald R, Luysaert S, Andersson AJ (2013) Anthropogenic perturbation of the carbon fluxes from land to ocean. *Nat Geosci* 6(8):597–607
 15. Schiff SL, Aravena R, Trumbore SE, Dillon PJ (1990) Dissolved organic carbon cycling in forested watersheds: a carbon isotope approach. *Water Resour Res* 26(12):2949–2957
 16. Abril G, Borges A (2019) Ideas and perspectives: carbon leaks from flooded land: do we need to replumb the inland water active pipe? *Biogeosciences* 16(3):769–784
 17. Tian H, Yang Q, Najjar RG, Ren W, Friedrichs MAMM, Hopkinson CS, Pan S (2015) Anthropogenic and climatic influences on carbon fluxes from eastern North America to the Atlantic Ocean: A process-based modeling study. *J Geophys Res Biogeosci* 120(4):757–772. <https://doi.org/10.1002/2014JG002760> Received
 18. Lauerwald R, Regnier P, Camino-Serrano M, Guenet B, Guimberteau M, Ducharne A, Polcher J, Ciais P (2017) ORCHILEAK (revision 3875): a new model branch to simulate carbon transfers along the terrestrial-aquatic continuum of the Amazon basin. *Geosci Model Dev* 10:3821–3859
 19. Vilmin L, Mogollón JM, Beusen AHW, van Hoek WJ, Liu X, Middelburg JJ, Bouwman AF (2020) Modeling process-based biogeochemical dynamics in surface freshwaters of large watersheds with the IMAGE-DGNM framework. *J Adv Model Earth Syst.* <https://doi.org/10.1029/2019MS001796>
 20. Tank SE, Fellman JB, Hood E, Krutzberg ES (2018) Beyond respiration: controls on lateral carbon fluxes across the terrestrial-aquatic interface. *Limnol Oceanogr Lett* 3(3):76–88
 21. Bouwman AF, Bierkens MFP, Griffioen J, Hefting MM, Middelburg JJ, Middelkoop H, Slomp CP (2013) Nutrient dynamics, transfer and retention along the aquatic continuum from land to ocean: towards integration of ecological and biogeochemical models. *Biogeosciences* 10(1):1–22
 22. Inger AR, Jan M (1999) Dissolved organic matter (DOM) in acid forest soils at Gådsjön (Sweden): natural variabilities and effects of increased input of nitrogen and of reversal of acidification. *Water Air Soil Pollut* 114(1–2):199–219
 23. Aitkenhead-Peterson JA (2000) Source, production and export of dissolved organic carbon and nitrogen. Doctoral Dissertations. 2112. University of New Hampshire
 24. Michalzik B, Kalbitz K, Park J-H, Solinger S, Matzner E (2001) Fluxes and concentrations of dissolved organic carbon and nitrogen—a synthesis for temperate forests. *Biogeochemistry* 52(2):173–205
 25. Fröberg M, Berggren D, Bergkvist B, Bryant C, Mulder J (2006) Concentration and fluxes of dissolved organic carbon (DOC) in three Norway spruce stands along a climatic gradient in Sweden. *Biogeochemistry* 77(1):1–23
 26. Sawicka K (2014) Spatial and temporal patterns in soil water: dissolved organic carbon—new insight into long-term British monitoring plots. URL <http://www.ecn.ac.uk/events/ecn-20-years-symposium/presentations/mon-06-sawicka>
 27. Sparling GP, Chibnall EJ, Pronger J, Rutledge S, Wall AM, Campbell DI, Schipper LA (2016) Estimates of annual leaching losses of dissolved organic carbon from pastures on allophanic soils grazed by dairy cattle, Waikato, New Zealand. *N Z J Agric Res* 59(1):32–49
 28. Borken W, Ahrens B, Schulz C, Zimmermann L (2011) Site-to-site variability and temporal trends of DOC concentrations and fluxes in temperate forest soils. *Glob Change Biol* 17(7):2428–2443
 29. Liu CP, Sheu BH (2003) Dissolved organic carbon in precipitation, throughfall, stemflow, soil solution, and stream water at the Guandaushi subtropical forest in Taiwan. *For Ecol Manag* 172(2–3):315–325
 30. Schrumpf M, Zech W, Lehmann J (2006) TOC, TON, TOS and TOP in rainfall, throughfall, litter percolate and soil solution of a montane rainforest succession at Mt. Kilimanjaro, Tanzania. *Biogeochemistry* 78(3):361–387
 31. Neff JC, Asner GP (2001) Dissolved organic carbon in terrestrial ecosystems: synthesis and a model. *Ecosystems* 4(1):29–48
 32. Chantigny MH (2003) Dissolved and water-extractable organic matter in soils: a review on the influence of land use and management practices. *Geoderma* 113(3–4):357–380
 33. Deb SK, Shukla MK (2011) A review of dissolved organic matter transport processes affecting soil and environmental quality. *J Environ Anal Toxicol* 1(2):106
 34. Camino-Serrano M, Gielen B, Luysaert S, Ciais P, Vicca S, Guenet B, Vos BD, Cools N, Ahrens B, Arain MA (2014) Linking variability in soil solution dissolved organic carbon to climate, soil type, and vegetation type. *Glob Biogeochem Cycles* 28(5):497–509
 35. Ian C IC (1991) A model of dissolved organic carbon concentrations in soil and stream waters. *Hydrol Process* 5(3):301–307
 36. Boyer EW, Hornberger GM, Bencala KE, McKnight D (1996) Overview of a simple model describing variation of dissolved organic carbon in an upland catchment. *Ecol Model* 86(2–3):183–188
 37. Currie WS, Aber JD (1997) Modeling leaching as a decomposition process in humid montane forests. *Ecology* 78(6):1844–1860
 38. Michalzik B, Tipping E, Mulder J, Lancho JFG, Matzner E, Bryant CL, Clarke N, Lofts S, Esteban MAV (2003) Modelling the production and transport of dissolved organic carbon in forest soils. *Biogeochemistry* 66(3):241–264
 39. Wu H, Peng C, Moore TR, Hua D, Li C, Zhu Q, Peichl M, Arain MA, Guo Z (2014) Modeling dissolved organic carbon in temperate forest soils: TRIPLEX-DOC model development and validation. *Geosci Model Dev* 7(3):867–881
 40. Dick JJ, Tetzlaff D, Birkel C, Soulsby C (2015) Modelling landscape controls on dissolved organic carbon sources and fluxes to streams. *Biogeochemistry* 122(2–3):361–374
 41. Rowe EC, Tipping E, Posch M, Oulehle F, Cooper DM, Jones TG, Burden A, Hall J, Evans CD (2014) Predicting nitrogen and acidity effects on long-term dynamics of dissolved organic matter. *Environ Pollut* 184:271–282
 42. Stergiadi M, Perk MVD, Nijs TD, Bierkens MFP (2016) Effects of climate change and land management on soil organic carbon dynamics and carbon leaching in northwestern Europe. *Biogeosciences* 13(5):1519–1536
 43. Sawicka K, Rowe EC, Evans CD, Monteith DT, Vanguelova EI, Wade AJ, Clark JM (2017) Modelling impacts of atmospheric deposition and temperature on long-term DOC trends. *Sci Total Environ* 578:323–336
 44. Serrano MC, Pannatier EG, Vicca S, Luysaert S, Jonard M, Ciais P, Guenet B, Gielen B, nuelas JP, Sardans J (2016) Trends in soil solution dissolved organic carbon (DOC) concentrations across European forests. *Biogeosciences* 13:5567–5585
 45. Michalzik B, Matzner E (1999) Dynamics of dissolved organic nitrogen and carbon in a Central European Norway spruce ecosystem. *Eur J Soil Sci* 50(4):579–590
 46. Dalva M, Moore TR (1991) Sources and sinks of dissolved organic carbon in a forested swamp catchment. *Biogeochemistry* 15(1):1–19
 47. Shibata H, Mitsuhashi H, Miyake Y, Nakano S (2001) Dissolved and particulate carbon dynamics in a cool-temperate forested basin in northern Japan. *Hydrol Process* 15(10):1817–1828
 48. Kalbitz K, Solinger S, Park J-H, Michalzik B, Matzner E (2000) Controls on the dynamics of dissolved organic matter in soils: a review. *Soil Sci* 165(4):277–304

49. Dai KH, David MB, Vance GF (1996) Characterization of solid and dissolved carbon in a spruce-fir Spodosol. *Biogeochemistry* 35(2):339–365
50. Sanderman J, Baldock JA, Amundson R (2008) Dissolved organic carbon chemistry and dynamics in contrasting forest and grassland soils. *Biogeochemistry* 89(2):181–198
51. Davidson EA, Trumbore SE (1995) Gas diffusivity and production of CO₂ in deep soils of the eastern Amazon. *Tellus B Chem Phys Meteorol* 47(5):550–565
52. McLaughlin JW, Liu G, Jurgensen MF, Gale MR (1996) Organic carbon characteristics in a spruce swamp five years after harvesting. *Soil Sci Soc Am J* 60(4):1228–1236
53. Trettin CC, Jurgensen MF (2003) Carbon cycling in wetland forest soils. Lewis Publishers, Boca Raton
54. Kottek M, Grieser J, Beck C, Rudolf B, Rubel F (2006) World map of the Köppen–Geiger climate classification updated. *Meteorol Z* 15(3):259–263
55. Neff JC, Hobbie SE, Vitousek PM (2000) Nutrient and mineralogical control on dissolved organic C, N and P fluxes and stoichiometry in Hawaiian soils. *Biogeochemistry* 51(3):283–302
56. Buckingham S, Tipping E, Hamilton-Taylor J (2008) Dissolved organic carbon in soil solutions: a comparison of collection methods. *Soil Use Manag* 24(1):29–36
57. Batjes NH (2016) Harmonized soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks. *Geoderma* 269:61–68
58. Bouwman, AF (1990) Global distribution of the major soils and land cover types. In: *Soils and the greenhouse effect*. John Wiley and Sons, Chichester, UK, pp 33–59
59. USDA-NRCS (2005) Global soil regions map | NRCS Soils. URL https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/use/?cid=nrcs142p2_054013
60. Batjes NH (2015) World soil property estimates for broad-scale modelling (WISE30sec). Technical report
61. Harris IPDJ, Jones PD, Osborn TJ, Lister DH (2013) Updated high-resolution grids of monthly climatic observations—the CRU TS3. 10 Dataset. *Int J Climatol* 34(3):623–642
62. Stehfest E, van Vuuren D, Bouwman L, Kram T (2014) Integrated assessment of global environmental change with IMAGE 3.0: model description and policy applications. Netherlands Environmental Assessment Agency (PBL)
63. Evans CD, Monteith DT, Cooper DM (2005) Long-term increases in surface water dissolved organic carbon: observations, possible causes and environmental impacts. *Environ Pollut* 137(1):55–71
64. Parton WJ (1996) The century model. In: Evaluation of soil organic matter models. Springer, New York, pp 283–291
65. Sitch S, Smith B, Prentice IC, Arneeth A, Bondeau A, Cramer W, Kaplan JO, Levis S, Lucht W, Sykes MT (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Glob Change Biol* 9(2):161–185
66. Bondeau A, Smith PC, Zaehle S, Schaphoff S, Lucht W, Cramer W, Gerten D, Lotze-Campen H, Müller C, Reichstein M, Smith B (2007) Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Glob Change Biol* 13(3):679–706. <https://doi.org/10.1111/j.1365-2486.2006.01305.x>
67. von Bloh W, Schaphoff S, Müller C, Rolinski S, Waha K, Zaehle S (2018) Implementing the nitrogen cycle into the dynamic global vegetation, hydrology, and crop growth model LPJmL (version 5.0). *Geosci Model Dev* 11(7):2789–2812
68. Tipping E, Woolf C, Rigg E, Harrison AF, Ineson P, Taylor K, Benham D, Poskitt J, Rowland AP, Bol R (1999) Climatic influences on the leaching of dissolved organic matter from upland UK moorland soils, investigated by a field manipulation experiment. *Environ Int* 25(1):83–95
69. Aber JD, Nadelhoffer KJ, Steudler P, Melillo JM (1989) Nitrogen saturation in northern forest ecosystems. *BioScience* 39(6):286–378
70. Currie WS, Aber JD, McDowell WH, Boone RD, Magill AH (1996) Vertical transport of dissolved organic C and N under long-term N amendments in pine and hardwood forests. *Biogeochemistry* 35(3):471–505
71. Fernández-Sanjurjo MJ, Vega VF, Garcia-Rodeja E (1997) Atmospheric deposition and ionic concentration in soils under pine and deciduous forests in the river Sor catchment (Galicia, NW Spain). *Sci Total Environ* 204(2):125–134
72. Litaor MI (1988) Soil solution chemistry in an alpine watershed, Front Range, Colorado, USA. *Arctic Alpine Res* 20(4):485–491
73. McDowell WH, Likens GE (1988) Origin, composition, and flux of dissolved organic carbon in the Hubbard Brook Valley. *Ecol Monogr* 58(3):177–195
74. Raich JW, Schlesinger WH (1992) The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate. *Tellus B* 44(2):81–99
75. PostWilfred M, Emanuel WR, Zinke PJ, Stangenberger AG (1982) Soil carbon pools and world life zones. *Nature* 298(5870):156
76. The R Stats Package. Choose a model by AIC in a Stepwise Algorithm (2020). URL <http://search.r-project.org/R/library/stats/html/step.html>
77. Fox J (2015) Applied regression analysis and generalized linear models. Sage Publications, New York
78. Velleman PF, Welsch RE (1981) Efficient computing of regression diagnostics. *Am Stat* 35(4):234–242
79. Guggenberger G (1992) Eigenschaften und Dynamik gelöster organischer Substanzen (DOM) auf unterschiedlich immissionsbelasteten Fichtenstandorten. Lehrstuhl für Bodenkunde und Bodengeographie der Univ
80. Guggenberger G, Zech W (1993) Dissolved organic carbon control in acid forest soils of the Fichtelgebirge (Germany) as revealed by distribution patterns and structural composition analyses. *Geoderma* 59(1–4):109–129
81. Cronan CS (1985) Comparative effects of precipitation acidity on three forest soils: carbon cycling responses. *Plant Soil* 88(1):101–112
82. McDowell WH, Currie WS, Aber JD, Yang Y (1998) Effects of chronic nitrogen amendments on production of dissolved organic carbon and nitrogen in forest soils. *Water Air Soil Pollut* 105(1–2):175–182
83. Salazar O, Balboa L, Peralta K, Rossi M, Casanova M, Tapia Y, Singh R, Quemada M (2019) Effect of cover crops on leaching of dissolved organic nitrogen and carbon in a maize-cover crop rotation in Mediterranean Central Chile. *Agric Water Manag* 212:399–406
84. Moore TR, Paré D, Boutin R (2008) Production of dissolved organic carbon in Canadian forest soils. *Ecosystems* 11(5):740–751
85. Easthouse KB, Mulder J, Christophersen N, Seip HM (1992) Dissolved organic carbon fractions in soil and stream water during variable hydrological conditions at Birkenes, southern Norway. *Water Resour Res* 28(6):1585–1596
86. Johnson MS, Johannes L, Guimaraes CE, Filho JPN, Riha SJ (2006) DOC and DIC in flowpaths of Amazonian headwater catchments with hydrologically contrasting soils. *Biogeochemistry* 81(1):45–57
87. Strobel BW, Hansen HCB, Borggaard OK, Andersen MK, Raulund-Rasmussen K (2001) Composition and reactivity of DOC in forest floor soil solutions in relation to tree species and soil type. *Biogeochemistry* 56(1):1–26

88. Sanderman J, Amundson R (2008) A comparative study of dissolved organic carbon transport and stabilization in California forest and grassland soils. *Biogeochemistry* 89(3):309–327
89. Aitkenhead JA, McDowell WH (2000) Soil C: N ratio as a predictor of annual riverine DOC flux at local and global scales. *Glob Biogeochem Cycles* 14(1):127–138
90. Moore TR, De Souza W, Koprivnjak JF (1992) Controls on the sorption of dissolved organic carbon by soils. *Soil Sci* 154(2):120–129
91. Ranville J (2005) Evaluating the influence of soil geochemical characteristics on metal bioavailability. US Geological Survey. Mineral Resources Program. <http://minerals.usgs.gov/mrerp/reports/Ranville-04HQGR0154.pdf>
92. Liu E, Yan C, Mei X, Zhang Y, Fan T (2013) Long-term effect of manure and fertilizer on soil organic carbon pools in dryland farming in northwest China. *PLoS ONE* 8(2):e56536
93. Brooks PD, McKnight DM, Bencala KE (1999) The relationship between soil heterotrophic activity, soil dissolved organic carbon (DOC) leachate, and catchment-scale DOC export in headwater catchments. *Water Resour Res* 35(6):1895–1902
94. Lofts S, Woof C, Tipping E, Clarke N, Mulder J (2001) Modelling pH buffering and aluminium solubility in European forest soils. *Eur J Soil Sci* 52(2):189–204
95. Lu X, Gilliam FS, Yu G, Chen H, Mo J (2013) Long-term nitrogen addition decreases carbon leaching in a nitrogen-rich forest ecosystem. *Biogeosciences* 10(6):3931–3941
96. Antweiler RC, Drever JI (1983) The weathering of a late Tertiary volcanic ash: importance of organic solutes. *Geochim Cosmochim Acta* 47(3):623–629
97. Marin LE, Kratz TK, Bowser CJ (1990) Spatial and temporal patterns in the hydrogeochemistry of a poor fen in northern Wisconsin. *Biogeochemistry* 11(1):63–76
98. Grieve IC (1990) Variations in chemical composition of the soil solution over a four-year period at an upland site in southwest Scotland. *Geoderma* 46(4):351–362
99. Koprivnjak J-F, Moore TR (1992) Sources, sinks, and fluxes of dissolved organic carbon in subarctic fen catchments. *Arctic Alpine Res* 24(3):204–210
100. Kaiser K, Guggenberger G, Zech W (1996) Sorption of DOM and DOM fractions to forest soils. *Geoderma* 74(3–4):281–303
101. Major J, Lehmann J, Rondon M, Goodale C (2010) Fate of soil-applied black carbon: downward migration, leaching and soil respiration. *Glob Change Biol* 16(4):1366–1379
102. David MB, Driscoll CT (1984) Aluminum speciation and equilibria in soil solutions of a Haplorthod in the Adirondack Mountains (New York, USA). *Geoderma* 33(4):297–318
103. Sato A, Seto M (1999) Relationship between rate of carbon dioxide evolution, microbial biomass carbon, and amount of dissolved organic carbon as affected by temperature and water content of a forest and an arable soil. *Commun Soil Sci Plant Anal* 30(19–20):2593–2605
104. Kang H, Freeman C, Ashendon TW (2001) Effects of elevated CO₂ on fen peat biogeochemistry. *Sci Total Environ* 279(1–3):45–50
105. Ahern CR, Baker DE, Aitken RL (1995) Models for relating pH measurements in water and calcium chloride for a wide range of pH, soil types and depths. In: Date RA, Grundon NJ, Rayment GE, Probert ME (eds) *Plant-soil interactions at low pH: principles and management*. Developments in plant and soil sciences, vol 64. Springer, Dordrecht
106. Trumbore SE, Davidson EA, de Camargo PB, Nepstad DC, Martinelli LA (1995) Belowground cycling of carbon in forests and pastures of Eastern Amazonia. *Glob Biogeochem Cycles* 9(4):515–528
107. Jassal RS, Black TA, Drewitt GB, Novak MD, Gaumont-Guay D, Nesic Z (2004) A model of the production and transport of CO₂ in soil: predicting soil CO₂ concentrations and CO₂ efflux from a forest floor. *Agric For Meteorol* 124(3–4):219–236
108. Drewitt GB, Black TA, Jassal RS (2005) Using measurements of soil CO₂ efflux and concentrations to infer the depth distribution of CO₂ production in a forest soil. *Can J Soil Sci* 85(2):213–221
109. Lawrence GB, David MB (1996) Chemical evaluation of soil-solution in acid forest soils. *Soil Sci* 161(5):298–313
110. Kane ES, Valentine DW, Michaelson GJ, Fox JD, Ping CL (2006) Controls over pathways of carbon efflux from soils along climate and black spruce productivity gradients in interior Alaska. *Soil Biol Biochem* 38(6):1438–1450
111. Guggenberger G, Kaiser K (2003) Dissolved organic matter in soil: challenging the paradigm of sorptive preservation. *Geoderma* 113(3–4):293–310
112. Jones DL, Willett VB (2006) Experimental evaluation of methods to quantify dissolved organic nitrogen (DON) and dissolved organic carbon (DOC) in soil. *Soil Biol Biochem* 38(5):991–999
113. Chantigny MH, Harrison-Kirk T, Curtin D, Beare M (2014) Temperature and duration of extraction affect the biochemical composition of soil water-extractable organic matter. *Soil Biol Biochem* 75:161–166
114. Zsolnay A (2003) Dissolved organic matter: artefacts, definitions, and functions. *Geoderma* 113(3–4):187–209
115. Litaor MI (1988) Review of soil solution samplers. *Water Resour Res* 24(5):727–733
116. Gödde M, David MB, Christ MJ, Kaupenjohann M, Vance GF (1996) Carbon mobilization from the forest floor under red spruce in the northeastern USA. *Soil Biol Biochem* 28(9):1181–1189
117. Moore TR, Clarkson BR (2007) Dissolved organic carbon in New Zealand peatlands. *N Z J Mar Freshw Res* 41(1):137–141
118. Aitkenhead-Peterson JA, McDowell WH, Neff JC (2003) Sources, production, and regulation of allochthonous dissolved organic matter inputs to surface waters. In: Findlay SEG, Sinsabaugh RL (eds) *Aquatic ecology, aquatic ecosystems*. Academic Press, pp 25–70
119. Meir P, Malhi Y, Hodnett M, Grace J (2004) Soil CO₂ efflux in a tropical forest in the central Amazon. *Glob Change Biol* 10(5):601–617
120. Christ MJ, David MB (1996) Temperature and moisture effects on the production of dissolved organic carbon in a spodosol. *Soil Biol Biochem* 28(9):1191–1199
121. Undurraga P, Zagal E, Sepúlveda G, Valderrama N (2009) Dissolved organic carbon and nitrogen in Andisol for six crop rotations with different soil management intensity. *Chil J Agric Res* 69(3):445–454
122. Nelson PN, Baldock JA, Oades JM (1992) Concentration and composition of dissolved organic carbon in streams in relation to catchment soil properties. *Biogeochemistry* 19(1):27–50
123. Catalán N, Marcé R, Kothawala DN, Tranvik LJ (2016) Organic carbon decomposition rates controlled by water retention time across inland waters. *Nat Geosci* 9(5):1–7. <https://doi.org/10.1038/ngeo2720>
124. Stutter MI, Billett MF (2003) Biogeochemical controls on streamwater and soil solution chemistry in a High Arctic environment. *Geoderma* 113(1–2):127–146
125. MacLean R, Oswald MW, Irons JG, McDowell WH (1999) The effect of permafrost on stream biogeochemistry: a case study of two streams in the Alaskan (USA) taiga. *Biogeochemistry* 47(3):239–267
126. Petrone KC (2005) Export of carbon, nitrogen and major solutes from a boreal forest watershed: the influence of fire and permafrost (Doctoral dissertation). University of Alaska Fairbanks

127. Kahle M, Kleber M, Jahn R (2004) Retention of dissolved organic matter by phyllosilicate and soil clay fractions in relation to mineral properties. *Org Geochem* 35(3):269–276
128. Sawicka K, Monteith DT, Vanguelova EI, Wade AJ, Clark JM (2016) Fine-scale temporal characterization of trends in soil water dissolved organic carbon and potential drivers. *Ecol Indic* 68:36–51
129. Fujii K, Funakawa S, Hayakawa C, Kosaki T (2008) Contribution of different proton sources to pedogenetic soil acidification in forested ecosystems in Japan. *Geoderma* 144(3–4):478–490
130. Fellman JB, D'Amore DV, Hood E, Boone RD (2008) Fluorescence characteristics and biodegradability of dissolved organic matter in forest and wetland soils from coastal temperate watersheds in southeast Alaska. *Biogeochemistry* 88(2):169–184
131. Nieminen T, Vos BD, Cools N, Koenig N, Fischer R, Iost S, Meessenburg H, Nicolas M, O'Dea P, Cecchini G, Ferretti M, De La Cruz A, Derome K, Lindroos A-J, Graf PE (2016) Soil solution collection and analysis, ICP forests manual. Technical report, UNECE ICP Forests Programme Co-ordinating Centre, Thünen Institute of Forest Ecosystems, Eberswalde, Germany. URL https://www.icp-forests.org/pdf/manual/2016/ICP_Manual_2016_01_part1.pdf
132. Johnson J, Pannatier EG, Carnicelli S, Cecchini G, Clarke N, Cools N, Hansen K, Meessenburg H, Nieminen TM, Pihl-Karlsson G (2018) The response of soil solution chemistry in European forests to decreasing acid deposition. *Glob Change Biol* 24(8):3603–3619
133. Van Beek LPH, Wada Y, Bierkens MFP (2011) Global monthly water stress: 1. Water balance and water availability. *Water Resour Res* 47(7):W07517-1–W07517-25
134. Inge EMG, Rens LPHB, Tom G, Nils M, Oliver S, Edwin HS, Marc FPB (2017) A global-scale two-layer transient ground water model: development and application to groundwater depletion. *Adv Water Resour* 102:53–67. ISSN 0309-1708

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.