Reviewer #1:

The authors represent a modelling development that makes it possible to have a more complete look at the global carbon cycle. They combine the well-established JULES model with a newly developed model for DOC, including soil carbon processes and leaching. This manuscript is therefore an important step towards a full carbon cycle understanding.
In general I think that the manuscript is well structured and the figures are helpful to understand the outcomes. However, there are some changes needed to make it more convincing. While I see some issues that should be clarified/solved first, I recommend publishing the manuscript in GMD after revision.

Thank you very much for your careful comments. We improved the manuscript following the reviewer's suggestions. Details are given below.

1. Please explain why the authors did not include the production of OC in soils and rivers (i.e. aquatic photosynthesis)?

Let us clarify what the model does and doesn't include.
The developments presented are made on the JULES, the Joint UK Land-Environment Simulator. The standard version of JULES, described in Clark et al. (2011), simulates vegetation and soil water and carbon processes, including the production of soil organic carbon following plant mortality, litter decomposition, respiration etc.
Our manuscript describes new developments in JULES in order to represent DOC cycling in soils, including DOC production is soils, DOC decomposition and respiration, as well as leaching of DOC from the soil to the aquatic environments.

However; the reviewer is right to say that we do not include fate of OC in rivers, this is beyond the scope of this work, as it would need to represent the full biogeochemistry of DOC in rivers (biological activity, sedimentation, gas transfer, etc.). Our developments lay the corner stone for a future representation of C cycling along the land-to-ocean continuum, including transport and transformation in the river network, as highlighted in our introduction and outlook. This is clarified now in the introduction.

2. Explain the additional value of adding Turkey Point and Guandaushi. To me it seems that the data from these two sides do not add much information, due to their much shorter time coverage. Also in the discussion, it is stated that e.g. Turkey Point is not really useful because it's located at a site that was agriculturally used until 1989. Possibly it would help to remove the Level 2-sites.

There is always a balance to find between very few sites that have well documented long-term measurements, and more sites, with shorter record and/or more complex history. In this work, we found that focusing on only 3 sites, all from Western European forests, would not suffice to gain confidence in our model, hence we decided to also include Turkey Point-89 and Guandaushi sites that we define as "level-2 sites", in order to explore broader climate/ecosystems domains. Turkey Point-89 provides a high productive forest ecosystem on a previously agricultural land. This is a typical system in North America where marginal agricultural land which have been afforested in recent decades. Also, we are keen to keep Turkey point as it is the only site that provides DOC measurements down to 100 cm, providing additional constraints on our model.
We find that our model has some difficulties in capturing dynamics at Turkey Point site. However, while we overestimate the total DOC in Turkey Point, we are able to reproduce the vertical profile of DOC in the soil.

3. Discuss the different methods in the study side description. Are there fundamental differences between 'suction cups' (p.8 l.22) and 'tension lysimeters' (p.8 l.32)? Are the observed values comparable?

First of all, we agree with the reviewer #1 that this is an important point which needs further clarification. For all sites, samples were collected using one of the two in situ soil solution extraction methods, namely suction cups or suction plates. We have, therefore, removed the term "lysimeter", which is often used incorrectly, and have replaced it with one of the two afore mentioned extraction methods. Where appropriate we have also included information on the material of the sampling device as well as the suction applied. This will allow the reader to better estimate the potential effects of different sampling methods on DOC concentrations obtained (Weihermüller et al. 2007). This is clarified in revised text:
For Hainich:
"At this site, soil solution samples were taken at three depths (5, 10 and 20 cm) using ceramic suction plates positioned at four different plots within the site. Samples were obtained by applying a tension of 100 hPa after each bi-weekly sampling occasion"

For Carlow:
"Samples were obtained by applying a tension of 400 hPa after each bi-weekly sampling occasion (Walmsley 2009)."
For Brasschaat:
"DOC samples were collected at three horizons of Al/Ap, A/E and Cg (Soil Classification Working Group 1998) referred to 10,35 and 75cm depth, by means of ceramic suction cups on a biweekly interval. Two days prior to sample collection a tension of 600 hPa was applied to each suction cup. Samples were collected at three locations and pooled into one composite sample per layer for analysis (Gielen et al. 2011)"
For Turkey Point-89:
"DOC sampling was attempted in monthly intervals at three depths of 25, 50 and 100 cm by means of porous suction cups..."
For Gunadaushi:
"DOC samples were collected at three depths of 15, 30 and 60 cm in three locations at bi-weekly interval by means of ceramic suction cups"

As samples were obtained by means of suction from a conical or plate-like porous sampling device we consider the differences due to methodological issues to be small compared to other soil solution collection methods such as passive samplers (pan lysimeters, wick samplers, resin boxes) and in particular extraction methods (centrifugation, chemical extractants such as KCl or CaCl2). The main difference between suction cups and suction plates is that the latter exhibit a larger sampling area and due to the 2D surface of the plates, the origin of the sampled solution is better defined which is important for mass balance estimations. It has also been shown that ceramic devices can absorb dissolved organic matter. As a result, concentrations in all but Carlow (where more suitable non-absorbing porous glass suction cups were used) may be underestimated. In addition, the suction applied will have an effect on the source of the soil solution collected (macro-, meso-, micropores). We are unable to quantify these effects in the current study. However, as mentioned above, we believe that the effect of different sampling methods on DOC concentrations are relatively small compared to potential differences due to sampling effects. Finally, as we are comparing observed values with the model simulation of Soil DOC concentration (not comparing observed values with each other), the difference in the sampling methods should not have a significant impact on the comparison of the model vs observed values.

When did this sampling take place (e.g. Turkey Point – 'samples could only be retrieved for 5 separate days'; all in summer? or winter?) (see also comment above (2.))

In Turkey point-89, due to the highly drained soil because of its sandy and dry condition, the sampling was taking place in spring and autumn when the soil was wet enough. Hence the dates of sampling were: 29-Nov-2004, 3-May-2005, 16-Jun-2005, 29-Jun-2005, 14-Oct-2005 (for more detail please see (Peichl et al. 2007)).
This is clarified in the revised manuscript:
"however, due to the dry sandy soils, samples could only be retrieved for 5 separate days of sampling after heavy rain fall events on 29-Nov-2004, 3-May-2005, 16-Jun-2005, 29-Jun-2005, 14-Oct-2005 (Peichl et al. 2007)."

4. Elaborate on the model input in more detail:
a) Have the data of FLUXNET and WATCH been somewhat corrected to be comparable?

No. These two data are not comparable. The FLUXNET data are the site-specific observed data. Hence, whenever we had access to the site measurements, we used them. The only site which we had to use WATCH data was Guandaushi, where we could not find any on-site measurements. Both forcing data were checked for any missing data and it was gap filled by linear interpolation. This is clarified in the revised manuscript:
"The FLUXNET database provides on-site meteorological data for each site that could be used as forcing for simulations in JULES, However, we had to use the global WATCH dataset (Weedon et al. 2010) as forcing for Guandaushi site where no on-site data was available. However, both forcing data were checked for any missing data and it was gap filled by linear interpolation."

b) Mention the names for the parameters as it's used in the tables and equations consistently (e.g. p.9 l.28: bulk density and clay content)

Corrected.

5. Please adapt the figures in a way that makes them easier to understand.
a) Fig.2 The extent can be smaller to better see where the sites are.

Corrected.

b) Fig.3 Decrease the y-axis range. A maximum of about 20 should be sufficient and the differences a better to see.

Corrected.

c) Fig .4 Decrease the y-axis range to a maximum of 40.

Corrected.

6. Combine both parts of the discussion to one. This would avoid the repetitions and can clearly combine all information/discussion on each of the sites.

Done as suggested, thank you.

Minor comments:

a) Define SOC at its first occurrence (p-2 l.28)

SOC first occurrence was defined in p-2 l.8.

b) Please use consistent and not confusing naming for the variables/parameters in the equations (e.g. R can mean 'run-off' and 'respiration')

Corrected. Runoff modified to Roff instead of R.

c) Connecting the equations to the arrows in the flow chart (Fig.1) would help in understanding the calculations. What do the numbers (1) to (12) in Fig.1 mean? They don't seem to match with the equations.

Sorry, this was unclear. Numbers were corresponding to process not to equations. As all arrows are defined within the text at the end of each process described (i.e. p.5 l.10: "(F_P; arrows 1-4 Fig. 1)" indicating the process linked to the arrows in the model figure.)
In order to avoid potential confusion, we replaced the numbers in fig.1 with letters.

d) What is the spatial resolution?

The evaluation of model was performed on plot-scale, using climate forcing data, soil and land cover for one specific point, so no horizontal spatial dimension was involved. The model is however capable of running at each spatial resolution for which forcing data are available.
This is clarified in the revised manuscript (p.10, l.15):
"The evaluation of model was performed on plot-scale, using climate forcing data, soil and land cover consistent with the site, no horizontal spatial dimension was involved."

e) Make unit naming consistent (e.g. Kg C m-2 day-1 vs. kg C m-2 day-1, p.5 l.10 vs. l.15)

Agreed. We now use "kg" throughout the text.

f) I suggest to rename 'Carbon concentration and fluxes' to 'Validation of carbon concentration and fluxes'

Corrected.

Reviewer #2:

The authors represent a model which calculates the DOC concentration to inland waters. They extended the JULES model for DOC, including soil carbon processes and leaching. This manuscript is a step towards a carbon model for aquatic systems and their export to the oceans.
In general, I think that the manuscript is well structured, but the description of the model needs some improvement. I recommend publishing the manuscript in GMD after revision.

Thank you very much for your careful comments. We improved the manuscript following the reviewer's suggestions. Details are given below.

Before I start with my comments, I must point out that I am giving feedback from a modelling point of view. My work on the global aquatic C cycle has just started, but I have a lot of expertise on global modelling.
From that point of view, I was very happy that both the abstract and the introduction start with sentences about global transport to the oceans. The importance of lateral transport is emphasized, but the model description itself does not contain a word on lateral transport. The model is actually a 1-D model and the outcome could be used to transport in the river network.
The second remark is the mentioning of the C cycle in the abstract "A model that represents the whole continuum from atmosphere to land and into the ocean would provide better understanding of the Earth's C cycle and hence more reliable historical or future projection" and introduction "Hence we need to move towards a boundless C cycle model which accounts for lateral fluxes". Why did you choose, after emphasizing C (as in total C) transport, to represent DOC only instead of modelling other species like POC, SOC and DIC as well?

We emphasize the importance of lateral transfers and C cycling along the land-to-ocean aquatic continuum in the abstract and the introduction, as this marks the ultimate goal of our model developments. Nevertheless, our manuscript represents just a first step in that development. At later stages of the overall model development, other processes and C species will be dealt with as well. This is now clarified in the abstract:

"A first and critical step in that direction is to include processes representing production and export of dissolved organic carbon in soils. Here we present an original representation of Dissolved Organic C (DOC) processes in the Joint UK Land Environment Simulator (JULES-DOCM) that integrates a representation of DOC production in terrestrial ecosystems based on incomplete decomposition of organic matter, DOC decomposition within the soil column, and DOC export to the river network via leaching."

As for representing DOC only, instead of POC, SOC and DIC, the reviewer is right, that ideally, we should represent carbon exports in all forms.
According to Meybeck (1982), DOC exports to the coast represent about 37% of C taken up on land from the atmosphere and being laterally exported along the river network. DIC is also a large source of carbon to rivers (potentially larger than DOC), but DIC sources are driven by very different processes such as rock erosion, that are not directly connected to soil organic carbon and the terrestrial carbon cycle. As a first priority, we then decided to focus the JULES developments on DOC processes. Different forms of C will need different processes to be represented in future steps of implementing the land-to-ocean aquatic continuum into the representation of the global C cycle. For instance, the simulation of POC transports would require the representation of erosion, sediment transport and autotrophic production. The representation of DIC would require the representation of weathering processes and water-air gas exchanges. This is now clarified in the introduction:

"Other forms of C need different processes to be represented to fully represent the land-to-ocean aquatic continuum of the global C cycle. Hence future work should include DIC and POC export from soils as well as the fate of all exported carbon in the river system.''

Comments/Questions

Abstract line 29-30: I think that part of the leaching to the riverine system is explained by this model. The flux from groundwater or other sources going to the riverine network are not explained by this model. You could shortly elaborate on the relevance or importance of groundwater and give a short explanation on why you ignore it for now. The model comparison is done in the soil and not in the river network.

As in most global land surface models, a ground water aquifer is not directly represented in JULES-DOCM. Runoff from soils is simply represented as two components, a surface runoff and a subsurface runoff. The subsurface runoff includes the drainage from the bottom of the 3m soil column, and thus somehow mimics the ground water base flow, in terms of water as well as in terms of DOC exports. This information is now added to the leaching section in model (p.7, line16-17):
"However subsurface runoff is also representing the drainage from the bottom of the 3m soil column, and thus mimics the groundwater base flow, in terms of water as well as in terms of DOC exports"

Please, also note that in this manuscript, we focus on DOC cycling within the soil column, and we do not yet represent C fluxes in the rivers. For that reason, we compare our simulation results against observed DOC concentrations in the soil solution. Carbon fluxes in rivers are in addition affected by decomposition of DOC, additional sources of DOC from the decomposition of POC and the evasion of CO_2 to the atmosphere. Comparing the simulated leaching of DOC to the river against observed DOC concentrations at some downstream sampling location would not be valid because of the non-conservative behaviour of DOC in the river.

Introduction: Should be more clear on the objective/aim of the model study. Same for abstract.

We agree with reviewer #2 and revised abstract and introduction accordingly.

Page 3, line 12: Why 3 meters deep?

JULES default soil depth is set at 3 meters. The root profile, the soil C stocks and the soil hydrology are all simulated over that 3-meter soil profile, which we used here for the representation of DOC. Moreover, soil depth was not always available at measurement sites. Therefore, we decided to keep the default values for which the JULES model was developed. This information is now added to revised text (p3, l.11):
"The aim of this study is to include a representation of DOC produced in terrestrial soils down to 3 meters (as soil hydrology and Carbon are simulated over 3 meter soil profile in JULES)"

Page 3, line 24: 9 PFTs at global scale. What about crops? They are mentioned in Figure S1.
Names and the number of PFT do not match with Figure S1.

We thank the reviewer for this comment. Table S1 is giving Z0 for the PFTs as described in Jobbágy & Jackson data, not the JULES PFTs, sorry for the confusion. We added another table (Table. S2) giving Z0 for the JULES PFTs.

As for crops, in this version of JULES crops are classified as C3 and C4 grasslands. Note that there is a separate version of JULES with improved representation of crops (Osborne et al., 2015), not used here as our main focus is on natural ecosystems

Page 4, eq 2: I think dz should be without subscript (2x). I don't see why it is important to calculate x? Remove eq 2?

We corrected the notation of dz.
Please note that x is the ration of SOC content within the first 1 meter of soil relative to the 3-meter profile for different biomes as given by Jobbágy & Jackson (in their Table. 3) (Jobbágy & Jackson 2000) which is used to extrapolate a profile of soil C concentrations. This is clarified in the revised manuscript.

Page 4, eq 3: I think z=1 and z=4 should be replaced by i=1 and i=4. This 1 and 4 is not explained yet (I think they are the four soil layers that will be used).

Corrected. We replaced z with i. This is indicating the normalized weighting factors for all four soil layers (i).

Page 4, line 30-31: These lines do not say anything. Which measurements? When and where taken? Why this remark here? The DOC is not mentioned here. Why are there continuous lines for measurements? For the modelled results? Eq. 3 only gives four outcomes…..

We largely rewrote section 2.2.1, clarifying the approach to distribute organic carbon (calculated as a bulk stock) in the vertical to serve as input for the DOC model.

Page 5, line 3: In figure 1 I see four carbon pools added (two for lock and two for free).

Corrected.

Page 5, eq 4: k is indicator for labile or recalcitrant. But none of the other parameters is dependent on k, so why is k included?

We thank reviewer #2 comment on this equation. We added a subscript k for the soil C stocks Sc (now Sc_k), as the soil carbon pool defines which amounts of DOC produced go to labile and refractory DOC, respectively.

Page 5, line 15: add i subscript in F_S(S) and F_T(T_soil)

Corrected.

Page 5, line 17: RothC formulations. Reference needed.

Added.

Page 5, eq 5: What is the unit of silt and clay?

Fraction. Now it is added. We also changed the values from % to fraction in table 4.

Page 5, eq 6,7,8,9: What is S_CARB,DPM? Why twice substracting R_DPM? What is F_DOC,DPM? Please make the parameter names consistent. This system is solved for each soil layer, so why "i" is not in the equations? These formulas are not clear.

We thank reviewer #2 comment on these equations.

There was some typo in Equation5. This is corrected now. Variables names have also been checked and made consistent. We also corrected the equations adding "i" to the updates of pools based on the sum of DOC processes in all layers.

Page 5, eq 8,9: I don't understand why part of respiration (R_s) is flowing to BIO and HUM?
Can you take another parameter name for beta_R? Confusing. What is F_BIO,IN?

In RothC model the assumption is that part of decomposed carbon (B_r) is released to the atmosphere and the remaining fractrion (1- B_r) is feeding microorganism in biomass (BIO) pool or is stored in the soil as the recalcitrant form, humus (HUM) with a slowest decomposition rate. These terms are fully described in JULES description model (Clark et al. 2011).
These are clarified in the revised manuscript:
"where in RothC model fraction (f_{DPM}) of litter fall (Λ_c) is directed to DPM and RPM depending on vegetation type. C pools are subjected to decomposition. Part of decomposed C as a fraction (1-B_R) of total respiration ($R_s = R_{DPM}+ R_{RPM}+ R_{BIO}+ R_{HUM}+ R_{DOC}$) is partially feeding microorganisms in soil (BIO) and partially stored as recalcitrant C in soil (HUM) depending on soil texture and the rest (B_R) is released to the atmosphere."

We changed beta_R to B_r, to avoid confusion with the beta we use in equation 3.

F BIO,IN is CUE fraction of decomposed DOC which is going back to biomass pool (described in eq 11)

Page 6, line 3: R_s neglects R_DOC but it is called total respiration?

R_s in code is indeed the total respiration including the R_DOC it in code. This was a typo that we corrected to:
R_S = R_DPM+R_RPM+R_BIO+R_HUM+R_DOC

Page 6, line 12: add i subscript in S_DOC and k subscript in K_DOC

Corrected.

Page 6, line 13: add i subscript in F_T(T_soil). Is this the same parameter as mentioned on the previous page?

Corrected. Yes, it is the same parameter. This is added to revised text:

"$F_T(T_{soil})_i$ is the soil temperature rate modifier within each soil layer (i) same as in eq.4"

Page 6, eq 11 and 12: Should there not be a sum over k (labile and recalcitrant) in these formulas?

Corrected. We added the sum sign indicating that at the end the BIO_IN flux will be the sum of both labile and recalcitrant decomposed DOC.

Page 6, line 24-26: "The assumption … (k)." I don't know what you trying to say here….
Page 6, eq 14: I don't understand. The size of the labile DOC pool is the old value minus a flux plus total size of the adsorbed pool??
Page 7, eq 15: I don't understand. This means that size of adsorbed pool is equal to F_AD_i???

We thank reviewer #2 for the comment on the adsorption/desorption.
We revised the manuscript:

“For adsorption/desorption, a constant sorption equilibrium distribution coefficient (K_D) is used to partition DOC in dissolved and adsorbed phases. The assumption is that DOC in the labile or recalcitrant pool is proportionally distributed between adsorbed DOC (S_{DOCad}) and dissolved DOC pools (S_{DOC} in soluble phase) depending on K_D from each soil layer(i) and DOC pool (k). Hence if the potentially adsorbed DOC fraction (AD_pot_i) compared to the size of the actually adsorbed DOC ($S_{DOC_{ad_{k,i}}}$) is positive then this fraction will be adsorbed and added to the adsorbed DOC pool, and if it is negative then this fraction will be desorbed and added to dissolved DOC pool per model time step.
These terms for DOC labile and recalcitrant pools in JULES-DOCM are as follow (arrow: i and j, Fig. 1):

$$AD\_pot_i = S_{DOC_{k,i}} \times K_D \times \frac{BK}{\theta v_i} \quad \textit{(eq.13)}$$

$$S_{DOC\ k,i} = S_{DOC_{k,i}} - \left(AD\_pot_i - S_{DOC_{ad_{k,i}}}\right) \quad \textit{(eq.14)}$$

$$S_{DOC_{ad_{k,i}}} = S_{DOC_{ad_{k,i}}} + \left(AD\_pot_i - S_{DOC_{ad_{k,i}}}\right) \quad \textit{(eq.15)”}$$

Also in order to make it easier to read we replaced “locked DOC” with “adsorbed DOC” and “free DOC” with “dissolved DOC”.

Page 7, line 4: add i subscript teta_v.

Corrected.

Page 7, line 7: add k,i subscript C_DOC.

Corrected.

Page 7, line 8: do you mean distance between midpoints of the soil layers?

Yes. We replaced “the distance (z_i) between every two soil depths” to “distance (z_i) between midpoints of the soil layers”

Page 7, eq 16: add i subscript in the formula (C_DOC and z).

Corrected subscript.

Do I miss which direction the diffusion goes. Does it always go from layer 2 to 1 or layer 2 to layer 3? Then there should be a subscript i,j or something….

Agreed. We changed the subscript to: subscript i for downward flow, and j for upward flow of diffusion.

Page 7, eq 17 and 18: add subscript k and i to the formulas and in the text.

Corrected. We also added that top soil is the sum of first and second soil layer, and bottom soil is sum of thirds and fourth soil layer.

Page 7, eq 17 an 18: It is confusing to have another teta with another unit in these formulas.

We changed teta_s with T_s.

Page 7, eq 19: I should expect that F_P is negative?

F_P is the production of DOC, it is never below 0.

Add all the k and i subscripts to this equation.

Corrected.

Page 7, line 29: What do you mean by main DOC model parameters? In what sense?

Default model parameters. We changed “main” to “default”.

Page 8, line 32: explain Al/Ap, A/E and Cg

Reference added for soil horizons.

Page 9, line 29: analytical spin-up? What does that mean? Why the assumption that it must be a steady state?

We removed the terms “analytical spin-up”. In order to have the present-day C, we did the spin-up looping 300 times over each site until we reached the steady state for C in soil. This is revised in the manuscript now.
“The model was first spun-up looping over period 1996 to 2014 until all the soil variables reached a steady state.”

Page 10, line 1: HWSD global data. Reference needed.

Added.

Page 10, line 6: “test the sensitivity of DOC related model parameters” On what? DOC leaching?

Sorry, this was unclear. We tested them on the DOC concentration in different depths of the soil profile. We added this information in the revised manuscript:

“In order to test the sensitivity of DOC related model parameters on the DOC concentration in different depths of the soil profile, simulations were performed with varying values for z_0, τ_z and DOC controlling parameters such as $K_{DOC(labile)}$, $K_{DOC\ (recalcitrant)}$, D_f, CUE, K_D and D (Table 1).”

Page 10, line 6-7: Why are these parameters chosen? These parameters can say something about the inner-sensitivity of the model. But how about the inputs like for example assumptions on PFT or precipitation, temperature, and so on? What about choosing different number of soil layers?

We ran the sensitivity analyses on the rate constants which we took from the literature and which could be subject to a recalibration. Simulation results may as well be sensitive to forcing data used, but that is not the point of this model development study where we used on-site observations of climate instead of global forcing data which would be subject to more uncertainties. In JULES-DOCM, the soil profile depth and number of layers is fixed and cannot be changed, because of the dependence on the representation of soil hydrology.

Page 10, line 6: How can you change beta_z? It is calculated in eq. 3. But that is a normalization?? Should you not change z_0? And what are you changing? Beta_z for each layer?

Reviewer #2 is absolutely right. We did indeed changed z_0 and based on that got the new normalized beta_z for each layer. We clarified this in the revised text

Page 10, line 8: Remark. The method of changing one parameter at the time. This is a popular method. However, it renders no information on the effects of interactions of the parameters and that it covers only a limited part of the entire parameter space.

We agree with reviewer #2 that there could be some interactions between sensitivity of different parameters, but testing this was beyond the scope of this study.

Page 10, line 8: Why 50%? 10 or 5% was also enough to say something about the sensitivity around the default values.

Since the derived model parameters from literature already had their own level of uncertainty, for instance CUE which has more than 50% or K_{DOC} with 5-40% of uncertainty, we took the 50% of change to test all the parameters at the reasonable degree.

Page 20: Figure 1 is confusing. All 8 boxes are defined for all the four soil layers, but the diffusion and soil depth give the impression that for example DOC_lock_labile only are defined in the deeper soils. Suggestion: split up the figure with the 8 pools (left hand side) and the righthand side the diffusion. Leaching and soil depth and 4 boxes (which are of the lefthand side type).

Corrected as suggested.

Table 2 should be updated with the right subscripts and so on. "i" has not the unit m….

"i" is corrected.
Subscripts are defined within text. In Table 2, we omitted subscripts for reasons of readability.

Table S1: Z_0 should be with lower case characters

Corrected.

Table S4: What are the numbers in the matrix? Leaching? Unit?

Table S4 lists the values for Figure 8 (Relative change in simulated DOC (%) for a +50% and -50% changes in model parameters). We added this now to Table S4 description.

General

Page 15, line 7: "Hence, it is important to introduce a depth-dependence decay rate for these parameters.". Now the sensitivity is used to draw this conclusion. But why not show the contribution of the DOC leaching of each soil layer? That should give a clear picture. I miss a kind of analysis of the importance of the different processes. Because the model that is proposed here has a lot of parameters. Is it possible to reduce the number of assumptions/soil layers/flows between the different pools? A broader sensitivity analysis could help.

Please note that we are not representing the DOC leaching from each soil layer. As we described in leaching of model, the first and second layer together are considered as the top soil and leaching is taken from it. The update

of these two layers will be based on the proportion of leached DOC compared to the DOC concentration in each of these layers. The same applied to the third and fourth layer as the sub soil leaching.
We are representing only the key processes for DOC including production, decomposition, adsorption/desorption, diffusion and leaching. Hence, we do not feel we could significantly reduce our model assumptions here. That being said, we find that adsorption/desorption was not making any significant change to our simulation and is probably of second order in the estimate of DOC soil concentration and export. We revised the sensitivity section in discussion:
"The sensitivity tests indicate that the parameters controlling SOC concentrations in the soil profile (Z_0 and τ_z) and the recalcitrant DOC residence time ($K_{DOC\ \mathrm{(recalcitrant)}}$) have the most significant effect on soil DOC concentration, which indicates the importance of factors controlling DOC sources."
Regarding the change of soil layers/flows please look at answer to your comment on page 10, line 6-7.

What is the used spatial distribution? And the temporal resolution of the input?

As mentioned in our response to reviewer #1, the evaluation of model was performed at plot-scale using 1 dimensional climate forcing, thus no spatial resolution. Temporal resolution of the input is 30 minutes. This is now added to manuscript (p.10, l.6):
"The meteorological forcing is provided at the measurement site level (no explicit spatial resolution) and includes the downward shortwave and longwave radiation at the surface (W m-2), rainfall (kg m-2 s-1), snowfall (kg m-2 s-1), wind speed (m s-1), atmospheric temperature (K), atmospheric specific humidity (kg kg-1) and air pressure at the surface (Pa) at half an hour time step"

Representation of dissolved organic carbon in the JULES land surface model (vn4.4_JULES-DOCM)

Mahdi Nakhavali[1], Pierre Friedlingstein[1], Ronny Lauerwald[1], Jing Tang[2,3], Sarah Chadburn[1,4], Marta Camino-Serrano[5], Bertrand Guenet[6], Anna Harper[1], David Walmsley[7], Matthias Peichl[8], Bert Gielen[9]

[1]University of Exeter, Exeter, EX4 4QE, United Kingdom
[2] Terrestrial Ecology Section, Department of Biology, University of Copenhagen, Copenhagen, Denmark
[3] Centre for Permafrost, University of Copenhagen, Copenhagen, Denmark
[4]University of Leeds, School of Earth and Environment, Leeds, United Kingdom
[5]CREAF, Barcelona, Catalonia
[6]Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France
[7]Leuphana University Lüneburg, Germany
[8]Swedish University of Agricultural Sciences, Department of Forest Ecology and management, Umeå, Sweden
[9]University of Antwerp, Antwerp, Belgium

Correspondence to: Mahdi Nakhavali (m.nakhavali@exeter.ac.uk)

Abstract. Current global models of the carbon (C) cycle consider only vertical gas exchanges between terrestrial or oceanic reservoirs and the atmosphere, thus not considering lateral transport of carbon from the continents to the oceans. Therefore, those models implicitly consider that all the C which is not respired to the atmosphere is stored on land, hence overestimating the land C sink capability. A model that represents the whole continuum from atmosphere to land and into the ocean would provide better understanding of the Earth's C cycle and hence more reliable historical or future projections. A first and critical step in that direction is to include processes representing production and export of dissolved organic carbon in soils. Here we present an original representation of Dissolved Organic C (DOC) processes in the Joint UK Land Environment Simulator (JULES-DOCM) that integrates a representation of DOC production in terrestrial ecosystems based on incomplete decomposition of organic matter, DOC decomposition within the soil column, and DOC export to the river network via leaching. The model performance is evaluated in five specific sites for which observations of soil DOC concentration are available. Results show that the model is able to reproduce the DOC concentration and controlling processes including leaching to the riverine system which is fundamental for integrating terrestrial and aquatic ecosystems. Future work should include the fate of exported DOC in the river system as well as DIC and POC export from soil.

1 Introduction

An estimated 1.9 Pg C yr^{-1} is exported from soils through the river network to the oceans, which represents a significant flux in global carbon (C) cycle (Cole et al. 2007; Regnier et al. 2013) and can affect biological and chemical properties of both aquatic (Aitkenhead & Mcdowell 2000) and terrestrial ecosystems (Kalbitz 2000). In land surface models that are part of Earth system models, only vertical fluxes of carbon between land and atmosphere are considered whilst lateral export fluxes are not included. This leads to an overestimation of soil organic C (SOC) sequestration and terrestrial C sinks (Janssens et al. 2003; Jackson et al. 2002). Hence we need to move towards a boundless C cycle model which accounts for lateral fluxes and thus produces more accurate projections of atmospheric CO_2 concentrations and C stocks (Battin et al. 2009). One of the lateral fluxes that has been neglected is the transfer of carbon from terrestrial to aquatic ecosystems in the form of dissolved organic C (DOC), which has been shown to be increased by anthropogenic perturbation such as land use change such as deforestation and increased atmospheric CO_2 concentrations (Regnier et al. 2013). DOC contributes about 37% of the global riverine carbon exports to the coast (Meybeck 1993) and adds to the net-heterotrophy of inland waters and related CO_2 emission fluxes to the atmosphere.

The main sources of DOC in terrestrial ecosystems are plant residues (Khomutova et al. 2000), humus and root exudates (Kalbitz et al. 2000; Van den berg et al. 2012; Marschner 1995). DOC within the soil can be the product of in-situ production or be brought in by advective fluxes with soil water transport. It has been hypothesized that loss of the carbon from the soil by leaching has to be taken into account to reasonably re-assess the terrestrial C budget of Europe (Siemens 2003). The fate of this DOC within inland water networks, i.e. the proportion transported to the coast or respired and emitted to the atmosphere, is the key to understanding the link to the other compartments of the Earth system (Cole et al. 2007; Battin et al. 2009). Nevertheless, it is a difficult task to link riverine and terrestrial fluxes by empirical methods, because 1) riverine fluxes are integrating fluxes from different land use systems (Kindler et al. 2011; Boyer & Groffman 1996) with different leaching rates and DOC quality, 2) in-stream transformation makes it difficult to trace back terrestrial DOC sources, and 3) the difficulty to separate natural and anthropogenic perturbation fluxes (Schelker et al. 2013; Regnier et al. 2013).

A physical-based modelling approach explicitly representing different terrestrial sources and processes involved in DOC cycling within the soil column and DOC leaching from the soil can help overcome these difficulties. Representation of DOC cycling within the soil column is also a major step toward simulating deep soil SOC formation (Rumpel & Kögel-Knabner 2011). Physical-based models help to understand the processes involved in soil DOC cycling and leaching as well as biogeochemistry of SOC in general. So far several models have been developed that simulate DOC with different temporal and spatial resolution, from 15 minutes as in SOLVEG-II (Ota et al. 2013) to monthly as in ECOSSE (Smith et al. 2010) or RivCM (Langerwisch et al. 2015) and from site scale as in DyDOC (Michalzik et al. 2003) to global scale as in TEM (Kicklighter et al. 2013). Some of these models represent DOC leaching, whereas others do not. Each model has its own particular definition for carbon pools (including DOC) and DOC production processes which can be based on turnover time, as in TERRAFLUX (Neff & Asner 2001), or based on chemical composition as in the DyDOC model (Michalzik et al.

2003). Although all these models have been evaluated, with the exception of the TEM model which was tested for arctic rivers, none of them has demonstrated its ability of representing the DOC production, processing and transport at the global scale.

In general, most of the models containing decomposition are based on first-order kinetics (Olson 1963). Frequently, models tend to represent the top soil layer as the major source for DOC production and export (Koven et al. 2013), other studies (Rumpel & Kögel-Knabner 2011; Braakhekke et al. 2013) highlight the importance of DOC for SOC production in deeper soil layers.

Here we present an original representation of DOC processes in the Joint UK Land Environment Simulator (JULES-DOCM) that integrates a representation of DOC production in terrestrial ecosystems based on incomplete decomposition of organic matter, DOC decomposition within the soil column, and DOC export to the river network via leaching. JULES has been used to evaluate the global C cycle (e.g. LeQuéré, et al. 2015, Sitch et al. 2015) and its role in the Earth system, but to date lacks the critical processes of DOC production and export. The aim of this study is to include a representation of DOC produced in terrestrial soils down to 3 meters (as soil hydrology and Carbon are simulated over 3 meter soil profile in JULES), assuming an incomplete decomposition of organic matter and its subsequent fate as DOC including i) DOC decomposition and release as CO_2 to the atmosphere, and ii) DOC export to the riverine system via leaching; to test the new model in different ecosystems and to evaluate it against specific sites where soil DOC measurements were available. Other forms of C need different processes to be represented to fully represent the land-to-ocean aquatic continuum of the global C cycle. Hence future work should include DIC and POC export from soils as well as the fate of all exported carbon in the river system.

2 Material and Methods

2.1 JULES model

JULES is a process-based model which represents energy, water and C cycling between vegetation, soil and atmosphere as described in Best et al. (2011) and Clark et al. (2011). Vegetation processes in JULES are represented in a dynamic vegetation model (TRIFFID), distinguishing 9 plant function types (PFTs) at the global scale: tropical and temperate broadleaf evergreen trees, broadleaf deciduous trees, needle-leaf evergreen trees and deciduous trees, C3 and C4 grasses, and evergreen and deciduous shrubs (Harper et al. 2016).

The representation of SOC in JULES, follows the formulation of the RothC soil carbon scheme (Jenkinson et al. 1990; Jenkinson & Coleman 2008), distinguishing four carbon pools: decomposable plant material (DPM), resistant plant material (RPM), heterotrophic microbial biomass (BIO) and long-lived humified material (HUM). DPM and RPM pools receive litter inputs directly from the vegetation due to defoliation, mortality and disturbance, the allocation to DPM or RPM depending on the PFT characteristics with higher fraction of decomposable litter provided from grasses and higher fraction of resistant litter provided from trees (Clark et al. 2011). HUM and BIO each receive inputs from the other two soil carbon pools, as a fraction of the decomposition that is not respired to the atmosphere.

2.2 JULES-DOCM model new features

JULES-DOCM is an extension of JULES based on version 4.4 (vn4.4 documentation in http://jules-lsm.github.io/vn4.4) , which explicitly represents DOC cycling in soils and considers DOC leaching from the soil profile. The following section deals with the representation of DOC fluxes and processes in more details.

2.2.1 Soil carbon profile

SOC is specified as the main source of DOC in JULES-DOCM. In JULES v4.4, each of the four SOC pools is treated as a single box down to 3 m, without any representation of its vertical distribution. This absence of vertical distribution has consequences in terms of simulating DOC fluxes, but also potential impacts on soil CO_2 fluxes, considering vertical variations of soil temperature and moisture. In JULES-DOCM, we introduce a vertical distribution of SOC for each soil carbon pool assuming an exponential decay with depth, with a weighting factor β_0 :

$$\beta_{0_i} = e^{-\frac{z_i}{z_0}} \times dz_i \qquad (eq.1)$$

Here, z_0 is the e-folding depth of C content within 1 meter of soil (i.e. depth at which SOC decreases by a factor of *e* relative to the surface), z_i is the soil depth of layer *i*, and dz_i is the thickness of the soil layer. In order to estimate z_0, we used the soil data from Jobbágy & Jackson (2000) that provides the vertical distribution of SOC within a 3 m soil profile based on the observed soil carbon profiles across several biomes. Jobbágy & Jackson provides soil C content in the first meter [0-1m] and for the first 3 meters [0-3m], allowing us to estimate the fraction in the first meter and derive z_0 accordingly:

$$\int_0^1 e^{-\frac{z}{z_0}} d_z = x \int_0^3 e^{-\frac{z}{z_0}} d_z \qquad (eq.2)$$

where x is the ration of SOC content within the first 1 meter of soil relative to the 3-meter profile for different biomes as given by Jobbágy & Jackson (in their Table. 3) (Jobbágy & Jackson 2000). Jobbágy & Jackson provide data for 11 PFTs. Here we first estimate z_0 for each of those PFTs, then regrouped them into the 9 JULES PFTs (see tables S1 and S2).

In order to calculate the fraction of SOC that is used as input for DOC production in each layer of the DOC model, (see equation 4 below) the weighting factors are normalised (β_{zi}):

$$\beta_{z_i} = \frac{\beta_{0_i}}{\sum_{i=1}^{i=4} \beta_{0_i}} \qquad (eq.3)$$

2.2.2 DOC fluxes and processes

In JULES-DOCM, four new DOC carbon pools have been added. First the model accounts for a labile and a recalcitrant DOC pool based on their decomposition rate (Aguilar & Thibodeaux 2005; Thibodeaux & Aguilar 2005). The labile pool is readily available for decomposition in soil solution at all times and the recalcitrant pool is subject to slower decomposition rate (Smith et al. 2010). DOC produced from plant material pools (DPM and RPM) and microbial biomass (BIO) is directed to the labile pool, while DOC from humus (HUM) is directed to the recalcitrant pool. Second, both the labile and the recalcitrant DOC pools have a dissolved and an adsorbed form, with only the dissolved pool being subjected to decomposition and leaching.

DOC production (F_P) follows first-order kinetics (Olson 1963) and the flux of carbon from SOC to DOC pools (k for labile or recalcitrant) in each soil layer (i) in kg C m^{-2} day^{-1}(F_P; arrows a-d Fig. 1) is calculated as:

$$F_{P_{k,i}} = \beta_{z_i} \times S_{C_k} \times \left(1 - e^{\left(-K_P \times F_S(S)_i \times F_T(T_{soil})_i \times F_v(v) \times D_f\right)}\right) \times e^{-\tau_{z_i}} \qquad (eq.4)$$

where S_{C_k} is amount of carbon in the soil organic pool (DPM/RPM/BIO for DOC labile pool and HUM for recalcitrant pool) in kg C m^{-2} in whole soil, K_P is DOC production rate in day^{-1}, $F_S(s)_i$ and $F_T(T_{soil})_i$ are respectively the rate modifiers due to moisture and temperature, which are controlling decomposition in each soil layer (i), $F_v(v)$ is the fraction of the vegetation. All units are given in Table 2. The moisture and temperature rate modifiers are based on the RothC formulations (Coleman & Jenkinson 2014). τ_z is the empirical factor for decrease of C decomposition rates with soil depth, as recently introduced in JULES (Burke et al. 2016).

The DOC production rate is further modified by D_f, which considers the decrease of SOC decomposition rate as increase of silt plus clay content given in fraction (Parton et al. 1987):

$$D_f = 1 - (0.75 \times (clay + silt)) \qquad (eq.5)$$

After decomposition, carbon pools (S_C) are updated by the changes in each time step (daily) as follow:

$$\frac{\Delta S_{C_{DPM}}}{\Delta t} = f_{DPM}\Lambda_c - R_{DPM} - \sum_{i=1}^{i=4} F_{P_{DPM_i}} \qquad (eq.6)$$

$$\frac{\Delta S_{C_{RPM}}}{\Delta t} = (1 - f_{DPM})\Lambda_c - R_{RPM} - \sum_{i=1}^{i=4} F_{P_{RPM_i}} \qquad (eq.7)$$

$$\frac{\Delta S_{C_{BIO}}}{\Delta t} = 0.46\,(1 - B_R)\, R_S - R_{BIO} - \sum_{i=1}^{i=4} F_{P_{BIO_i}} + \sum_{i=1}^{i=4} F_{BIO_{IN_i}} \qquad (eq.8)$$

$$\frac{\Delta S_{C_{HUM}}}{\Delta t} = 0.54\,(1 - B_R)\,R_S - R_{HUM} - \sum_{i=1}^{i=4} F_{P_{HUM\,i}} \qquad (eq.9)$$

where in RothC model fraction (f_{DPM}) of litter fall (Λ_c) is directed to DPM and RPM depending on vegetation type. C pools are subjected to decomposition. Part of decomposed C as a fraction (1-B_R) of total respiration ($R_s = R_{DPM}+ R_{RPM}+ R_{BIO}+ R_{HUM}+ R_{DOC}$) is partially feeding microorganisms in soil (BIO) and partially stored as recalcitrant C in soil (HUM) depending on soil texture and the rest (B_R) is released to the atmosphere. These parameters were already present in JULES (Clark et al. 2011). In JULES-DOCM the update of carbon pools after DOC production was added (last term of each equation, F_P…, defined in equation 4 above) as well as $F_{BIO_{IN}}$ the input flux from DOC to BIO pool, described below.

We assume that the decomposition of DOC pools (F_D) (kg C m^{-2} day^{-1}) also follows first-order kinetics depending on temperature and labile and recalcitrant DOC pool size as follow (arrows e-f Fig. 1):

$$F_{D_{k,i}} = S_{DOC_{k,i}} \times (1 - e^{(-K_{DOC_k} \times F_T(T_{soil})_i)}) \qquad (eq.10)$$

where S_{DOCi} is the DOC pool size (k for labile or recalcitrant) in kg C m^{-2} and K_{DOCk} *is* the basal decomposition rate of the dissolved DOC (k for labile or recalcitrant pool) (in day^{-1}) and $F_T(T_{soil})_i$ is the soil temperature rate modifier within each soil layer (i) same as in eq.4.

Part of decomposed DOC is respired (R_{DOC} in kg C m^{-2} day^{-1}, arrow g Fig. 1) and the rest returns to the BIO carbon pool ($F_{BIO_{IN}}$ in kg C m^{-2} day^{-1}, arrow h Fig. 1) from each soil layer (i) and DOC pools (k). This proportion is controlled by a CUE parameter (Kalbitz et al. 2003) which is set to 0.5 as a default as in Manzoni et al. (2012).

Hence distribution of decomposed DOC to the BIO pool and respiration will be:

$$F_{BIO_{IN_i}} = (1 - CUE) \times \sum F_{D_{k,i}} \qquad (eq.11)$$

$$R_{DOC_{k,i}} = CUE \times \sum F_{D_{k,i}} \qquad (eq.12)$$

For adsorption/desorption, a constant sorption equilibrium distribution coefficient (K_D) is used to partition DOC in dissolved and adsorbed phases. The assumption is that DOC in the labile or recalcitrant pool is proportionally distributed between adsorbed DOC (S_{DOCad}) and dissolved DOC pools (S_{DOC} in soluble phase) depending on K_D from each soil layer(i) and DOC pool (k). Hence if the potentially adsorbed DOC fraction (AD_pot_i) compared to the size of the actually adsorbed DOC ($S_{DOCad_{k,i}}$) is positive then this fraction will be adsorbed and added to the adsorbed DOC pool, and if it is negative then this fraction will be desorbed and added to dissolved DOC pool per model time step.

These terms for DOC labile and recalcitrant pools in JULES-DOCM are as follow (arrow: i and j, Fig. 1):

$$AD\_pot_i = S_{DOC_{k,i}} \times K_D \times \frac{BK}{\theta v_i} \qquad (eq.13)$$

$$S_{DOC\ k,i} = S_{DOC_{k,i}} - \left(AD\_pot_i - S_{DOC_{ad_{k,i}}}\right) \qquad (eq.14)$$

$$S_{DOC_{ad_{k,i}}} = S_{DOC_{ad_{k,i}}} + \left(AD\_pot_i - S_{DOC_{ad_{k,i}}}\right) \qquad (eq.15)$$

where $S_{DOC_{k,i}}$ is dissolved labile and recalcitrant DOC pools in kg C m^{-2}, K_D is the distribution factor (m^3 water kg^{-1} soil), BK is bulk density (kg soil m^{-3}) and θ_{vi} is the volumetric soil moisture (m^3 m^{-3}) and it is considered to be same for DOC labile and recalcitrant pools.

DOC diffusion ($F_{Diff\ i,j}$) in kg C m^{-2} day^{-1} between the layers is based on Fick's second law and it is the function of the diffusion coefficient (D) in m^2 day^{-1}, concentration of labile or recalcitrant DOC at different soil depths ($C_{DOCk,i,j}$) in kg C m^{-2} and the distance ($z_{i,j}$) between midpoints of soil layers (i: downward flow; j: upward flow) in m (arrow k, Fig. 1):

$$F_{Diff\ i,j} = \mathrm{D} \times \frac{\partial^2 C_{DOC_{k,i,j}}}{\partial {z_{i,j}}^2} \qquad (eq.16)$$

Leaching of the DOC is considered to occur from all 4 DOC soil layers. The top DOC is defined as the first two layers representing the first 35 cm of the soil. The lower two DOC layers represent the sub-soil from 35 cm down to 3 m. Soil leaching at the top DOC layer is dependent on the surface runoff whereas subsurface leaching is dependent on the subsurface runoff. However subsurface runoff is also representing the drainage from the bottom of the 3m soil column, and thus mimics the groundwater base flow, in terms of water as well as in terms of DOC exports. More information on the hydrology of model is given in Gedney & Cox (2003); Clark & Gedney 2008). Both DOC layers leaching fluxes are based on the concentration of dissolved DOC in the soil water. Hence leaching of DOC (L) from the dissolved labile and recalcitrant pool within the top (sum of first and second soil layer) - and sub-soil (sum of third and fourth soil layer) (T and S) in kg C m^{-2} day^{-1} is calculated as follows (arrow l, Fig.1):

$$L_T = S_{DOC_{k,h}} \times \frac{Roff_{surf}}{T_{s_i}} \qquad (eq.17)$$

$$L_S = S_{DOC_{k,h}} \times \frac{Roff_{sub}}{T_{s_i}} \qquad (eq.18)$$

where $S_{DOC_{k,h}}$ is the DOC quantity in the dissolved labile and recalcitrant pool (*h* for top or sub soil), $Roff_{surf}$ is the surface runoff, $Roff_{sub}$ is the subsurface runoff (both kg m^{-2} day^{-1}) and T_{Si} (defined in code as θ_S) is the soil moisture in each soil layer (i) (kg m^{-2}).

Hence dissolved and adsorbed DOC pools are updated as follow:

$$\frac{\Delta S_{DOC_k}}{\Delta t} = F_{P_{k,i}} + AD\_pot_i + F_{Diff_i} - F_{D_{k,i}} - L_T - L_S \qquad (eq.19)$$

Values of the default DOC model parameters are given in Table 1.

2.3 Sites description

Two data levels were provided in order to test the model performance. Level 1, including Hainich, Carlow and Brasschaat which included the carbon fluxes and continuous DOC measurements from soil water from 3 to 10 years period, and level 2, including Turkey Point 89 (TP89) and Guandaushi with fewer C fluxes measurements and discontinuous DOC measurements (Table 3). Location of sites are given in Figure 2.

2.3.1 Hainich

The site "Hainich", located in Germany – National park Hainich, (51°04′ 45″N, 10°27′07″E), is covered by an old-growth deciduous forest dominated by *Fagus sylvatica* and intermixed with *Fraxinus excelsior* and *Acer pseudoplatanus* (Mund et al. 2010). The soil class at this site is Eutric Cambisol with a high clay content and high biological activity, as illustrated by a mull or F-Mull organic layer (Table 4). The mean annual air temperature is 7.5-8°C and the annual precipitation is in the range of 750-800 mm yr^{-1} (Kutsch et al. 2010). At this site, soil solution samples were taken at three depths (5, 10 and 20 cm) using ceramic suction plates positioned at four different plots within the site. Samples were obtained by applying a tension of 100 hPa after each bi-weekly sampling occasion.

2.3.2 Carlow

The site "Carlow" is located in Ireland – County Carlow, (52° 52'N, 6° 54'W). The land cover is grassland, the soil class is Calcic Luvisol. This sandy loamy soil has a uniform profile and is well-drained (Table 4). The climate is characterized by a mean annual air temperature of 9.3°C and a mean annual precipitation of 823 mm yr^{-1} (Walmsley et al. 2011). DOC samples were collected from at two locations separated 150 m from each other, using 20 suction cups per location, with ten of these cups installed directly beneath the rooting zone and the other ten at a depth of 0.7 m. Samples were obtained by applying a tension of 400 hPa after each bi-weekly sampling occasion (Walmsley 2009).

2.3.3 Brasschaat

The site "Brasschaat" is located in Belgium and covered by mixed coniferous/deciduous (De Inslag) forest, (51°18'33" N, 4°31'14" E) with stands of old Scots Pine (*Pinus sylvestris)* (Janssens et al. 1999). The temperate maritime climate is characterized by a mean annual air temperature of 11.1°C and a mean annual precipitation of 824 mm yr^{-1} (Gielen et al. 2010). The soil class was defined as Albic Hypoluvic Arenosol (Table 4). The profile usually exhibits a high soil moisture, but due to the sandy texture and rapid hydraulic conductivity in upper horizons, it is rarely saturated (Gielen et al. 2011). DOC samples were collected at three horizons of Al/Ap, A/E and Cg (Soil Classification Working Group 1998) referred to 10,35 and 75cm depth, by means of ceramic suction cups on a biweekly interval. Two days prior to sample collection a tension of 600 hPa was applied to each suction cup. Samples were collected at three locations and pooled into one composite sample per layer for analysis (Gielen et al. 2011).

2.3.4 Turkey Point 89

The site "Turkey Point 89 (TP89)", located in southern Ontario – Canada, (42°77′57″N, 80°45′09″E), is covered by an evergreen needleleaf forest dominated by Eastern White pine *(Pinus strobus* L.*)* mixed with few stands of Oak, Paper birch, Wild black cherry and Red pine (Peichl & Arain 2006) established in 1989 on agricultural lands (Peichl, Brodeur, et al. 2010). The mean annual air temperature is 8.1°C and mean annual precipitation is 832 mm yr^{-1} (Peichl & Arain 2006). The soil class at this site is Gleyed Brunisolic Luvisol and due to the high sand content, it is well drained and has a low to moderated water holding capacity (Peichl, Brodeur, et al. 2010; Presant, E.W., Acton 1984). DOC sampling was attempted in monthly intervals at three depths of 25, 50 and 100 cm by means of porous suction cups, however, due to the dry sandy soils, samples could only be retrieved for 5 separate days of sampling after heavy rain fall events on 29-Nov-2004, 3-May-2005, 16-Jun-2005, 29-Jun-2005, 14-Oct-2005 (Peichl et al. 2007).

2.3.5 Guandaushi

The site "Guandaushi" is located in central Taiwan, (23° 8'N, 120° 8'E). The climate is characterized by distinct rainy and dry seasons and a mean annual air temperature of 22.4°C and annual precipitation in the range of 2300 to 2700 mm yr^{-1}. The land cover is subtropical mixed hardwood forest including three stands of natural hardwood and secondary hardwood on light loam textured soil and Chinese fir (*Cunninghamia lanceolate)* on heavy clay textured soil. DOC samples were collected at three depths of 15, 30 and 60 cm in three locations at bi-weekly interval by means of ceramic suction cups.

2.4 Model input and setting

Model performance was tested against observed data from Guanduashi and four FLUXNET sites (Hainich, Carlow, Brasschaat and Turkey Point-89). The FLUXNET database provides on-site meteorological data for each site that could be used as forcing for simulations in JULES, However, we had to use the global WATCH dataset (Weedon et al. 2010) as forcing for Guandaushi site where no on-site data was available. However, both forcing data were checked for any missing data and it was gap filled by linear interpolation. The meteorological forcing is provided at the measurement site level (no explicit spatial resolution) and includes the downward shortwave and longwave radiation at the surface (W m-2), rainfall (kg m-2 s-1), snowfall (kg m-2 s-1), wind speed (m s-1), atmospheric temperature (K), atmospheric specific humidity (kg kg-1) and air pressure at the surface (Pa) at half an hour time step (Best et al. 2011).

For Brasschaat, additional model parameters such as BK and clay were taken from Janssens et al. (1999). The model was first spun-up looping over period 1996 to 2014 until all the soil variables reached a steady state. For Hainich, site parameters were taken from Kutsch et al. (2010). The spin-up was run looping 300 times over the years 2004 to 2014. For Carlow, site parameters were taken from Walmsley (2009) and Kindler & Siemens (2010). The spin-up was run looping 300 times over the years 2004-2009. For Turkey Point-89, site parameters were taken from Peichl & Arain (2006) and spin-up was run looping 300 times over the years 2002-2007. For Guandaushi, site vegetation parameters were taken from Liu & Sheu (2003) and soil parameters from HWSD global data (Nachtergaele et al. 2010) and spin-up was run looping 300 times over years 1990 to 2000. The evaluation of model was performed on plot-scale, using climate forcing data, soil and land cover consistent with the site, no horizontal spatial dimension was involved.

2.5 Sensitivity test

In order to test the sensitivity of DOC related model parameters on the DOC concentration in different depths of the soil profile, simulations were performed with varying values for z_0, τ_z and DOC controlling parameters such as $K_{DOC\ (labile)}$, $K_{DOC\ (recalcitrant)}$, D_f, CUE, K_D and D (Table 1).

In total, 16 runs were performed by modifying each parameter once by increasing it 50% and once by decreasing it by 50%. In order to do the comparison with measurements, runs were performed for 3 meters soil depth for the periods that measurements were available. Hence, Brasschaat runs were performed for the years 2006-2010, Hainich runs for the years 2005-2014 and Carlow runs for the years 2006-2008.

2.6 Statistical analysis

In order to test the model performance, with regard to simulated C stock and fluxes, we used an ANOVA (Analysis of variance) test to compare the model results from the default set of parameters against measurements. In order to test the parameter impact on the simulated DOC concentrations, we computed the RMSE values from each set of model parameter configurations.

3 Results

3.1 Validation of carbon concentration and fluxes

To examine the performance of soil DOC simulations, it is first necessary to explore other carbon fluxes which link to soil DOC pools. The first flux to be validated is the gross primary production (GPP), for which we have observed values (Table 3). The modelled mean GPP for Brasschaat and Carlow was significantly lower than measurements with 867±25 g C m^{-2} $year^{-1}$ compared to 1173.3±91 g C m^{-2} $year^{-1}$and 903.2 g C m^{-2} $year^{-1}$ compared to 1165.3 g C m^{-2} $year^{-1}$ ($p < 0.05$, Table S3), respectively. For Turkey Point 89 and Hainich, the measured GPP was in line with our model results with 1731.5±108 g C m^{-2} $year^{-1}$ and 1606.74±101 g C m^{-2} $year^{-1}$ compared to 1635.1± 62 g C m^{-2} $year^{-1}$ and 1455±167 g C $m^{-2}$$year^{-1}$ ($p = 0.162$, Table S3). The modelled NPP was higher than observed values for Hainich and for Turkey Point-89, while it was lower than observed values for Brasschaat (Table 5).

Total soil respiration measurements were available for Brasschaat, Hainich and Turkey Point-89 (Table 3) and were compared with the modelled outputs. The simulated values were close to observed values at Hainich, while the modelled values for Brasschaat were significantly higher (p-value < 0.05, Table S3) and for Turkey Point-89 higher (p-value = 0.0896, Table S3), than the observed values (Table 5).

Finally, we compared the SOC in measurements and model outputs, where the measurements from Brasschaat for 100 cm, Hainich for 60 cm, Carlow for 50 cm and Turkey Point-89 for 15cm (A-horizon) of soil were available. The modelled SOC stock for Brasschaat in the first 100 cm and for Hainich down to 60 cm were slightly lower than the observations, while for Carlow the simulated stocks down to 50 cm and for Turkey Point-89 the simulated stocks down to 15 cm were higher than the observed stocks (Table 5).

3.2 DOC simulations

In general, JULES-DOCM was capable of reproducing the DOC concentrations at all the tested sites using the default set of parameters (Table 1) chosen as representative for the top soil (Fig. 3 Level 1 sites, Fig. 4 level 2 sites). For Hainich, the simulated average values and value range were close to observed values at 10 cm and 20 cm (Table 5, RMSE values for 10 cm and 20 cm are 3.0 and 2.5 mg L^{-1} respectively). For Brasschaat, the simulation underestimated DOC concentrations at all depths, but with an increasing underestimation with soil depth (Table 5, RMSE values for 10, 35 and 75 cm are 22.9, 18.4 and 16.8 mg L^{-1} respectively). For Carlow, the modelled and measured values were close at depths of 10 cm and 77 cm, but strongly underestimated at the intermediate depth of 28 cm (Table 5, RMSE values for 10, 10-38 and 28-77 cm are 3, 10.2 and 1.5 mg L^{-1} respectively). At Turkey Point-89, the modelled and observed values were close at 25cm depth, but the DOC concentration average over the profile down to 100 cm was overestimated (Table 5). For Guandaushi, DOC measurements from three different stands (Natural hardwood, secondary hardwood and Chinese fir) values were compared with modelled values. The model values for a depth of 15 cm were closer to observed values for Chinese fir than for natural hardwood or secondary hardwood sites. For 30cm depth, the simulated DOC concentration was substantially lower than the measured DOC averaging over three stands in Guandaushi (Table 5).

Overall, the model was capable of reproducing the seasonality of DOC concentrations for the European sites where long-term observation data are available (Fig. 5). However, at Braschaat the simulated DOC peaked from April-July while observed DOC peaked from July-September.

We also examined the hydrology of the model and its interaction with DOC concentration and leaching (e.g. Hainich - Fig. 6; other sites are plotted in Fig. S3). It can be seen for the period 2005 to 2014 that during heavy precipitation, high runoff was produced which caused the higher leaching, and the consequence was a drop in the DOC concentration in 3 meters of soil.

3.3 Sensitivity tests

Sensitivity to model parameters was tested on the three European sites where a representative time-series of observed DOC concentrations was available (e.g. Hainich-10cm, Fig. 7). The results indicate that among all the parameters in all three sites, the model shows the highest sensitivity to SOC vertical profile, controlled by parameter z_0 (eq. 1), and the changing of SOC decomposition rate with soil depth, parameter, τ_z (eq. 4) (p-values < 0.05, Table S6). Among the DOC controlling parameters, the model shows the highest sensitivity to the basal decomposition rate of recalcitrant DOC ($K_{DOC\ (\mathrm{recalcitrant})}$) (eq.10), which is the inverse of the residence time of DOC in the recalcitrant pool.

The sensitivity of the model to each of these parameters was different at each site. For Hainich, the highest sensitivity was assigned to τ_z. Here, a change in τ_z by 50% leads to a 36% change in the mean DOC within 3 m, while a 50% change in $K_{DOC\ (\mathrm{recalcitrant})}$ leads to a 29% change and a 50% change in z_0 leads to a 25% change in simulated DOC concentrations (Fig 8a). The closest value for the mean DOC in 10 cm in Hainich (8.8 mg L^{-1}) to the measurement was produced by the default set (8.9 mg L^{-1}), while the highest value for DOC was reached with the 50% increase in τ_z (12.7 mg L^{-1}) and the lowest DOC value was produced with 50% decrease in τ_z (4.7 mg L^{-1}). In contrast to that, at a depth of 20 cm, the closest value to the mean of measured DOC (5.6 mg L^{-1}) was produced by 50% decrease in $K_{DOC\ (recalcitrant)}$ (4.9 mg L^{-1}) (Fig. 9-a).

In Brasschaat, the highest sensitivity was to z_0, closely followed by τ_z and $K_{DOC\ (recalcitrant)}$. A 50% change in each of these parameters led to a 36-40% change in DOC concentration over the 3 meters of soil profile (Fig. 8-b). At 10 cm, the closest value to measurements mean (39.4 mg L^{-1}) was produced by 50% increase in τ_z (39.2 mg L^{-1}). At 35 cm depth, the closest value to mean measurement (29.3 mg L^{-1}) was calculated by 50% increase in $K_{DOC\ (recalcitrant)}$ (16.2 mg L^{-1}) which was also the highest simulated value as well. At 75 cm, the closest value to mean of DOC measurement (22.0 mg L^{-1}) was produced by 50% increase in $K_{DOC\ (recalcitrant)}$ (8.1 mg L^{-1}) as it was the highest of the simulated values (Fig. 9-b).

For Carlow, the most sensitive parameters were τ_z and $K_{DOC\ (recalcitrant)}$: a 50% change in those parameters leads to a 31.5% and 27.4% in simulate DOC. A 50% change in z_0 leads to a low but still significant change of 6.5% change in simulated DOC within 3 meters of soil (p-value <0.05, Table S6) (Fig. 8-c). In 10cm, the closest modelled value to the mean measurement (5.7 mg L^{-1}) was produced by default parameter set (5.8 mg L^{-1}). Between 10 to 28 cm all the parameter sets underrepresented the DOC concentration mean measurement (13.1 mg L^{-1}) and the closest and highest value was produced by 50% in τ_z (3.8 mg L^{-1}). For 28 to 77cm, the closest value to the measurement (4.8 mg L^{-1}) was calculated by increasing τ_z by 50% (4.5 mg L^{-1}) (Fig. 9-c).

4 Discussion

4.1 Measurements versus model simulations

Overall, JULES-DOCM reproduced the range of GPP for most of our sites to an acceptable degree. At some sites, due to over/underestimated autotrophic respiration, the NPP and total respiration values were slightly different than measurements. Consequently, the modelled carbon stocks were different from the measurements in most of the sites, but yet capable of representing the general patterns that were observed in the measurements.

In Brasschaat, the modelled SOC was lower than the measurements, which could be due to the underestimated NPP (Table 5) and, as a consequence, the underestimated litter input, but also due to the overestimated soil respiration and SOC decomposition rates. The underestimation of SOC as a source of DOC led to a general underestimation of DOC. Nevertheless, the decrease of relative DOC concentration through soil is consistent with the observations.

In Hainich, a slightly overestimated NPP partly counter-balanced the overestimated soil respiration. Nevertheless, the SOC concentration simulated down to 60 cm was lower than the measurement at this depth. As we did not have observations of SOC down to 3 meters, we cannot certainly say if the simulated total SOC stock (13.7 kg C m^{-2}) over the whole soil column is close to the reality or not. Some of the controlling parameters like DOC basal decomposition rates are kept constant over the soil profile in our simulation, while they are maybe not constant with depth in the real world, perhaps due to priming effects (Guenet et al. 2010). That could explain why at Hainich, the simulated and observed DOC concentrations are very close at 10 cm depth, while they differ more at 20 cm depth.

In Carlow, the slight overestimation of GPP led to the overestimated SOC concentrations down to 50 cm, whilst again we cannot say with certainty that the whole SOC stock is overestimated, as the SOC stock has not been measured down to three meters. Some sources suggest that the SOC in Carlow grassland could be higher than the reported value in our reference, if we calculate the C in soil based on the fraction of loss of ignition (LOI) (Walmsley 2009; Hoogsteen et al. 2015). As Carlow is our only grassland biome site, additional data from different study sites would be valuable to achieve a more representative parametrization of soil carbon processes under grassland. One of the parameters to be optimized for such sites could be CUE which has a strong impact on the stocks and fluxes. Also, since the measured values for NPP or soil respiration for this site were not available to us, we were unable to assess whether we over- or underestimated these fluxes and if this could have potentially biased our SOC stock simulations. DOC measurements were provided from two plots which were placed on different terrain positions. The measurements from plot 2 (150 meter in south-westerly direction from plot 1) at 10 to 28 cm depth had a higher DOC concentration than plot 1 at the 10 cm (Walmsley 2009). This could be the result of small scale variations related to terrain position, which can be related to different soil moisture regimes and lateral import of DOC. It is not possible to represent such small-scale variation in global models like JULES-DOCM.

At Turkey Point-89, the simulated GPP is close to the observations, while NPP is slightly overestimated. The simulated soil respiration and decomposition rates are higher than observed values. The overestimated SOC concentration in the top soil could be the result of an overestimated depth gradient in SOC concentration, which in our simulations is derived from global data (Jobbágy & Jackson 2000). Also, we simulated the steady state SOC profile for forest vegetation, whereas the forest stand at the site is relatively young and succeeded agricultural land use in 1989, and thus, the SOC profile is likely not representative for a forest site. The overestimated DOC concentration for 100 cm depth may be due to this change in land use, which was not taken into account during simulations, providing more C input for DOC production. At this site, the observed higher soil moisture in the deeper profile could indicate a potentially high advection of DOC to the lower layers (Peichl, Arain, et al. 2010). This could be another reason for the lower DOC in 100 cm from measured compared to the modelled results.

In Gundaushi due to the lack of SOC, or vegetation carbon fluxes measurements from the site, we have no information on SOC concentrations and stocks. The lower values of DOC from our model compared to the measurements could be due to: Firstly, the high temporal variability of observed concentrations (large standard deviation for all the depths from the three stands). Second, the high value of DOC input from rainfall, which is not represented in JULES-DOCM (Liu & Sheu 2003). Recent studies have indicated that including this flux in models can have a significant impact on the DOC in soil (Lauerwald et al. 2017).

As there are no measurements of lateral leaching of DOC from soil to the river, our evaluation of this flux is based on the simulated DOC concentration and runoff. Hence as the simulated hydrology of the JULES model has been evaluated previously (Gedney, N. , Cox 2003; Clark & Gedney 2008), in this study, we assume that we will get robust estimates of DOC leaching by multiplying simulated concentration by runoff, as long as simulated DOC concentrations can be validated.

Overall, besides over/underestimation of DOC at some sites, the model was capable of representing the trend of DOC concentration at different depths when comparing to the measurements at all the sites.

4.2 Sensitivity analysis

The sensitivity tests indicate that the parameters controlling SOC concentrations in the soil profile (Z_0 and τ_z) and the recalcitrant DOC residence time ($K_{DOC\ \mathrm{(recalcitrant)}}$) have the most significant effect on soil DOC concentration, which indicates the importance of factors controlling DOC sources. Nevertheless, DOC related model parameters such as basal DOC decomposition rate are constant over different depths, which could be the reason for the difference between the modelled and measured values, especially in the deeper soil layers. Hence, it is important to introduce a depth-dependence decay rate for these parameters.

A limitation in our simulation is that we use a single, calibrated value for recalcitrant DOC residence time, which is the most sensitive DOC controlling parameter. It has been shown that this parameter can vary with biodegradability of SOC and litter under different PFTs and at different sites (Kalbitz et al. 2003; Turgeon 2008). However, more detailed data for different biomes is needed for calibrating different residence times for different PFTs.

5 Conclusion

Applying a carbon cycle model that integrates the whole continuum from land to ocean to atmosphere provides a better understanding of the Earth's carbon cycle and makes more reliable future projections. In this study, we presented DOC related processes in JULES, JULES-DOCM, which includes the DOC produced in the soil down to three meters and its subsequent fate including its decomposition and release as CO_2 to the atmosphere, and its export to the river network via leaching in different ecosystems. Results show that the model is capable of representing the DOC stocks, processes and its export to the riverine systems from different ecosystems. In future, our developments in the representation of DOC leaching will lead to a model approach integrating terrestrial and aquatic C cycling. However, more field data are still required to improve the model parametrization and performance.

Code availability

The code written for this version of JULES can be found at:

https://code.metoffice.gov.uk/svn/jules/main/branches/dev/mahdinakhavali/vn4.4_JULES_DOCM/ (registration required)

Acknowledgments: The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 643052 (C-CASCADES project). We want to thank Altaf Arain, Tim Moore and Gerd Glexiner for providing the DOC measurements. Ronny Lauerwald received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 703813 for the Marie Sklodowska-Curie European Individual Fellowship "C-Leak". Jing Tang is financed by Marie Sklodowska-Curie Action Individual Fellowship (MABVOC: 707187) and supported by Danish National Research Foundation (CENPERM DNRF100). Marta Camino-Serrano acknowledges funding from the European Research Council Synergy grant ERC-2013-SyG-610028 IMBALANCE-P.

References:

Aguilar, L. & Thibodeaux, L.J., 2005. Kinetics of peat soil dissolved organic carbon release from bed sediment to water. Part 1. Laboratory simulation. *Chemosphere*, 58(10), pp.1309–1318.

Aitkenhead, J.A. & Mcdowell, W.H., 2000. Aitkenhead JA, McDowell WH. 2000. Soil C:N ratio as a predictor of annual riverine DOC flux at local and global scales. Global Biogeochem Cycles. *Global Biogeochemical Cycles*, 14(1), pp.127–138.

Battin, T.J. et al., 2009. The boundless carbon cycle. *Nature Geoscience*, 2(9), pp.598–600. Available at: http://dx.doi.org/10.1038/ngeo618.

Van den berg, L.J.L., Shotbolt, L. & Ashmore, M.R., 2012. Dissolved organic carbon (DOC) concentrations in UK soils and the influence of soil, vegetation type and seasonality. *Science of the Total Environment*, 427–428, pp.269–276. Available at: http://dx.doi.org/10.1016/j.scitotenv.2012.03.069.

Best, M.J. et al., 2011. The Joint UK Land Environment Simulator (JULES), model description – Part 1: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, 4(3), pp.701–722. Available at: http://nora.nerc.ac.uk/15031/%5Cnhttp://www.geosci-model-dev.net/4/701/2011/.

Boyer, J.N. & Groffman, P.M., 1996. Bioavailability of water extractable organic carbon fractions in forest and agricultural soil profiles. *Soil Biology and Biochemistry*, 28(6), pp.783–790.

Braakhekke, M.C. et al., 2013. Modeling the vertical soil organic matter profile using Bayesian parameter estimation. *Biogeosciences*, 10(1), pp.399–420.

Burke, E.J., Chadburn, S.E. & Ekici, A., 2016. A vertical representation of soil carbon in the JULES land surface scheme with a focus on permafrost regions. *Geoscientific Model Development Discussions*, (September), pp.1–30. Available at: http://www.geosci-model-dev-discuss.net/gmd-2016-235/.

Clark, D.B. et al., 2011. The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, 4(3), pp.701–722. Available at: http://nora.nerc.ac.uk/15031/%5Cnhttp://www.geosci-model-dev.net/4/701/2011/.

Clark, D.B. & Gedney, N., 2008. Representing the effects of subgrid variability of soil moisture on runoff generation in a land surface model. , 113(May 2007), pp.1–13.

Cole, J.J. et al., 2007. Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. *Ecosystems*, 10(1), pp.171–184.

Coleman, K. & Jenkinson, D.., 2014. RothC - A Model for the Turnover of Carbon in Soil. Model description and windows user guide. *Evaluation of Soil Organic Matter Models: Using Existing Long-Term Datasets*, I(June), pp.237–246.

Gedney, N. , Cox, P.M., 2003. The Sensitivity of Global Climate Model Simulations to the Representation of Soil Moisture Heterogeneity. , (2000), pp.1265–1275.

Gielen, B. et al., 2010. Decadal water balance of a temperate Scots pine forest (Pinus sylvestris L.) based on measurements and modelling. *Biogeosciences*, 7(October), pp.1247–1261. Available at: http://www.biogeosciences.net/7/1247/2010/.

Gielen, B. et al., 2011. The importance of dissolved organic carbon fluxes for the carbon balance of a temperate Scots pine forest. *Agricultural and Forest Meteorology*, 151(3), pp.270–278.

Guenet, B. et al., 2010. Priming effect: Bridging the gap between terrestrial and aquatic ecology. *Ecology*, 91(10), pp.2850–2861.

Harper, A.B. et al., 2016. Improved representation of plant functional types and physiology in the Joint UK Land Environment Simulator (JULES v4.2) using plant trait information. *Geoscientific Model Development*, 9(7), pp.2415–2440.

Hoogsteen, M.J.J. et al., 2015. Estimating soil organic carbon through loss on ignition: Effects of ignition conditions and structural water loss. *European Journal of Soil Science*, 66(2), pp.320–328.

Jackson, R., Banner, J. & Jobbágy, E., 2002. Ecosystem carbon loss with woody plant invasion of grasslands. *Nature*, 277(July), pp.623–627. Available at: http://www.nature.com/nature/journal/v418/n6898/abs/nature00910.html.

Janssens, I.A. et al., 1999. Above- and belowground phytomass and carbon storage in a Belgian Scots pine stand. *Annals of Forest Science*, 56(2), pp.81–90.

Janssens, I.A. et al., 2003. Europe ' s Terrestrial Biosphere Anthropogenic CO 2 Emissions. *Science*, 300(June), pp.1538–1542.

Jenkinson, D.S. et al., 1990. The turnover of organic carbon and nitrogen in soil. *The Royal Society*, 329(1255).

Jenkinson, D.S. & Coleman, K., 2008. The turnover of organic carbon in subsoils. Part 2. Modelling carbon turnover. *European Journal of Soil Science*, 59(2), pp.400–413.

Jobbágy, E.G. & Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological Applications*, 10(2), pp.423–436.

Kalbitz, K. et al., 2003. Biodegradation of soil-derived dissolved organic matter as related to its properties. *Geoderma*, 113, pp.273–291.

Kalbitz, K. et al., 2000. Controls on the Dynamics of Dissolved Organic Matter in Soils A Review.pdf. *Soil Science*, 165(4), pp.277–304.

Kalbitz, K. ym., 2000. Controls on the dynamics of dissolved organic matter in soils: a review [Soils Issues]. *Soil Science*, 165(April), pp.277–304.

Khomutova, T.E. et al., 2000. Mobilization of DOC from sandy loamy soils under different land use (Lower Saxony, Germany). *Plant and Soil*, pp.13–19.

Kicklighter, D.W. et al., 2013. Insights and issues with simulating terrestrial DOC loading of Arctic river networks. *Ecological Applications*, 23(8), pp.1817–1836.

Kindler, R. et al., 2011. Dissolved carbon leaching from soil is a crucial component of the net ecosystem carbon balance. *Global Change Biology*, 17(2), pp.1167–1185. Available at: http://doi.wiley.com/10.1111/j.1365-2486.2010.02282.x.

Kindler, R. & Siemens, J., 2010. Dissolved carbon leaching from soil is a crucial component of the net ecosystem carbon balance. *Global Change ...*, pp.1167–1185. Available at: http://onlinelibrary.wiley.com/doi/10.1111/j.1365-

2486.2010.02282.x/full.

Koven, C.D. et al., 2013. The effect of vertically resolved soil biogeochemistry and alternate soil C and N models on C dynamics of CLM4. *Biogeosciences*, 10(11), pp.7109–7131.

Kutsch, W.L. et al., 2010. Heterotrophic soil respiration and soil carbon dynamics in the deciduous Hainich forest obtained by three approaches. *Biogeochemistry*, 100(1–3), pp.167–183. Available at: http://link.springer.com/10.1007/s10533-010-9414-9.

Langerwisch, F. et al., 2015. Deforestation in Amazonia impacts riverine carbon dynamics. *Earth System Dynamics Discussions*, 6(2), pp.2101–2136. Available at: http://www.earth-syst-dynam-discuss.net/6/2101/2015/.

Lauerwald, R. et al., 2017. ORCHILEAK: A new model branch to simulate carbon transfers along the terrestrial-aquatic continuum of the Amazon basin. *Geoscientific Model Development Discussions*, (April), pp.1–58. Available at: http://www.geosci-model-dev-discuss.net/gmd-2017-79/.

Liu, C.P. & Sheu, B.H., 2003. Dissolved organic carbon in precipitation, throughfall, stemflow, soil solution, and stream water at the Guandaushi subtropical forest in Taiwan. *Forest Ecology and Management*, 172(2–3), pp.315–325.

Manzoni, S. et al., 2012. Environmental and stoichiometric controls on microbial carbon-use efficiency in soils. *New Phytologist*, 196(1), pp.79–91.

Marschner, H., 1995. *Mineral Nutrition of Higher Plants*, Academic Press. Available at: https://books.google.co.uk/books?id=phnp-H1XeBkC.

Meybeck, M., 1993. Riverine transport of atmospheric carbon: Sources, global typology and budget. *Water, Air, & Soil Pollution*, 70(1–4), pp.443–463.

Michalzik, B. et al., 2003. Modelling the production and transport of Dissolved Organic Carbon in forest soils. *Biogeochemistry*, 66, pp.241–264. Available at: http://link.springer.com/10.1023/B:BIOG.0000005329.68861.27.

Mund, M. et al., 2010. The influence of climate and fructification on the inter-annual variability of stem growth and net primary productivity in an old-growth, mixed beech forest. *Tree Physiology*, 30(6), pp.689–704.

Nachtergaele, F. et al., 2010. The Harmonized World Soil Database. *Proceedings of the 19th World Congress of Soil Science, Soil Solutions for a Changing World, Brisbane, Australia, 1-6 August 2010*, pp.34–37.

Neff, J.C. & Asner, G.P., 2001. Dissolved organic carbon in terrestrial ecosystems: Synthesis and a model. *Ecosystems*, 4(1), pp.29–48.

Olson, J.S., 1963. Energy Storage and the Balance of Producers and Decomposers in Ecological Systems. , 44(2), pp.322–331.

Ota, M., Nagai, H. & Koarashi, J., 2013. Root and dissolved organic carbon controls on subsurface soil carbon dynamics: A model approach. *Journal of Geophysical Research: Biogeosciences*, 118(4), pp.1646–1659.

Parton, W, J., Schimel, D.S. & Ojima, C.V.C.D.S., 1987. Analysis of Factors Controlling Soil Organic Matter Levels in Great Plains Grasslands. *Soil Science Society of America Journal*, 51(i), pp.1173–1179.

Peichl, M., Brodeur, J.J., et al., 2010. Biometric and eddy-covariance based estimates of carbon fluxes in an age-sequence of temperate pine forests. *Agricultural and Forest Meteorology*, 150(7–8), pp.952–965. Available at: http://dx.doi.org/10.1016/j.agrformet.2010.03.002.

Peichl, M. et al., 2007. Concentrations and fluxes of dissolved organic carbon in an age-sequence of white pine forests in Southern Ontario, Canada. *Biogeochemistry*, 86(1), pp.1–17.

Peichl, M. & Arain, M.A., 2006. Above- and belowground ecosystem biomass and carbon pools in an age-sequence of temperate pine plantation forests. *Agricultural and Forest Meteorology*, 140(1–4), pp.51–63.

Peichl, M., Arain, M.A. & Brodeur, J.J., 2010. Age effects on carbon fluxes in temperate pine forests. *Agricultural and Forest Meteorology*, 150(7–8), pp.1090–1101. Available at: http://dx.doi.org/10.1016/j.agrformet.2010.04.008.

Presant, E.W., Acton, C.J., 1984. The soils of the regional municipality of Haldimand-Norfolk. *Agricul- ture Canada, Ministry of Agriculture and Food*, Volume 1(Report No. 57), p.100.

Quéré, Le, C. et al., 2015. Global Carbon Budget 2015. *Earth System Science Data*, 7(2), pp.349–396.

Regnier, P. et al., 2013. Anthropogenic perturbation of the carbon fluxes from land to ocean. *Nature Geoscience*, 6(8), pp.597–607. Available at: http://www.nature.com/doifinder/10.1038/ngeo1830.

Rumpel, C. & Kögel-Knabner, I., 2011. Deep soil organic matter-a key but poorly understood component of terrestrial C cycle. *Plant and Soil*, 338(1), pp.143–158.

Schelker, J. et al., 2013. Drivers of increased organic carbon concentrations in stream water following forest disturbance: Separating effects of changes in flow pathways and soil warming. *Journal of Geophysical Research: Biogeosciences*, 118(4), pp.1814–1827.

Schrumpf, M. et al., 2011. How accurately can soil organic carbon stocks and stock changes be quantified by soil inventories? *Biogeosciences*, 8(5), pp.1193–1212.

Siemens, J., 2003. The European Carbon Budget: A Gap. *Science (New York, N.Y.)*, 302(5651), p.1681.

Sitch, S. et al., 2015. Recent trends and drivers of regional sources and sinks of carbon dioxide. *Biogeosciences*, 12(3), pp.653–679.

Smith, J. et al., 2010. Estimating changes in national soil carbon stocks using ECOSSE - a new model that includes upland organic soils. Part II. Application in Scotland. , (February 2016), pp.1–35.

Smith, J. et al., 2010. Model to Estimate Carbon in Organic Soils – Sequestration and Emissions (ECOSSE). *Carbon*, 44(August), pp.1–73.

Soil Classification Working Group, 1998. The Canadian System of Soil Classification. *The Canadian System of Soil Classification, 3rd ed. Agriculture and Agri-Food Canada Publication 1646*, p.187.

Thibodeaux, L.J. & Aguilar, L., 2005. Kinetics of peat soil dissolved organic carbon release to surface water. Part 2. A chemodynamic process model. *Chemosphere*, 60(9), pp.1190–1196.

Walmsley, D.C. et al., 2011. Dissolved carbon leaching from an Irish cropland soil is increased by reduced tillage and cover cropping. *Agriculture, Ecosystems and Environment*, 142(3–4), pp.393–402. Available at: http://dx.doi.org/10.1016/j.agee.2011.06.011.

Walmsley, D.C., 2009. *Quantifying Dissolved Carbon and Nitrogen Losses from Soils Subjected to Different Land-use and Management Practices*, University College Dublin. Available at: https://books.google.fr/books?id=HCz2ZwEACAAJ.

Xia, J.Y. et al., 2012. A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state. *Geoscientific Model Development*, 5(5), pp.1259–1271.

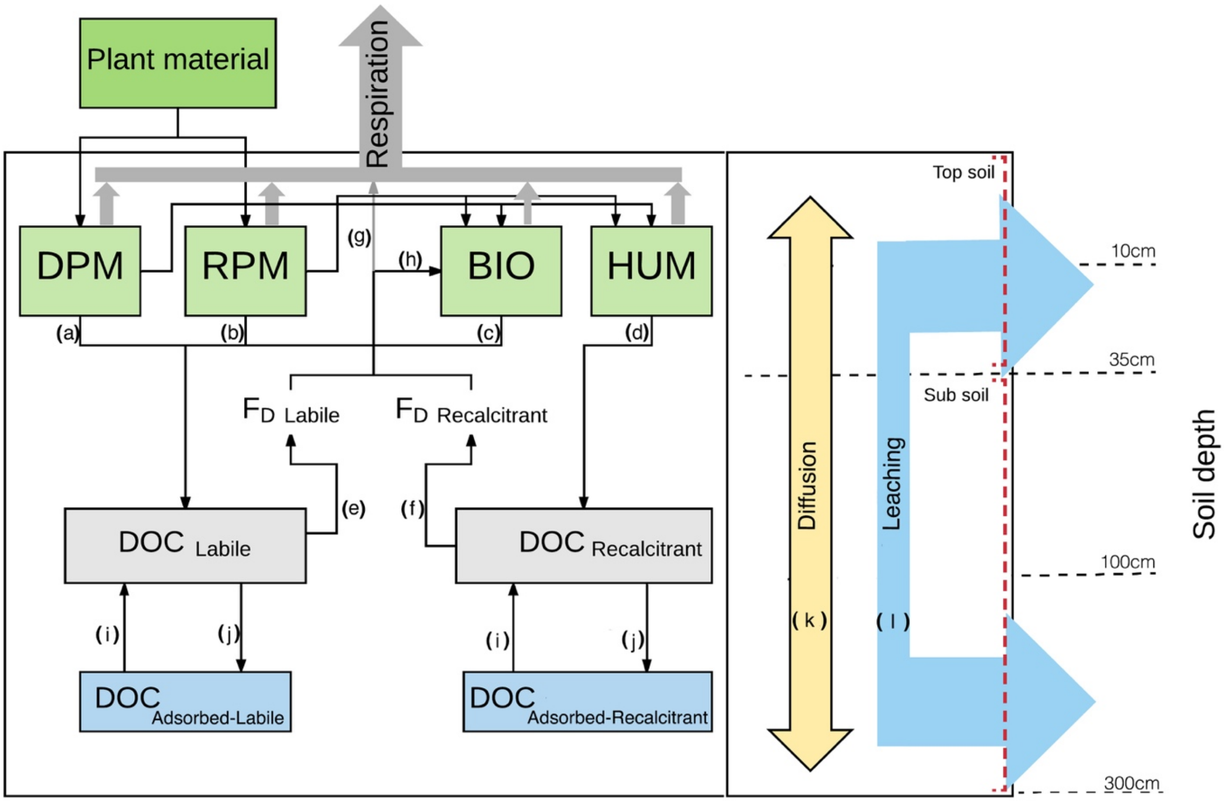

Figure 1. JULES-DOCM model structure

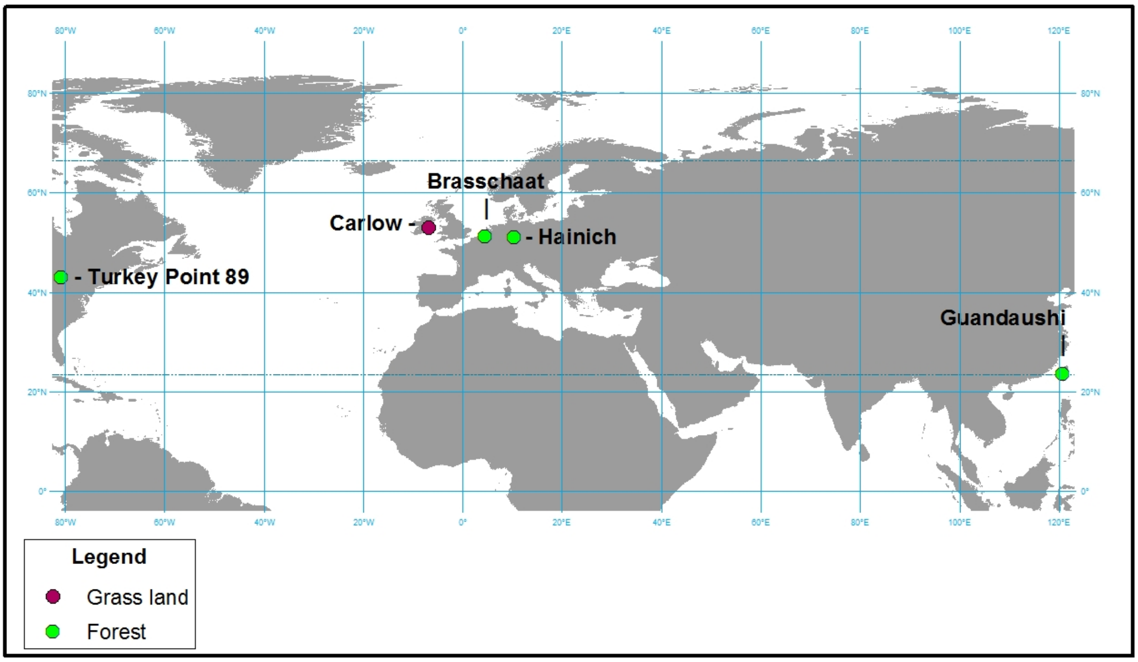

Figure 2. Study sites

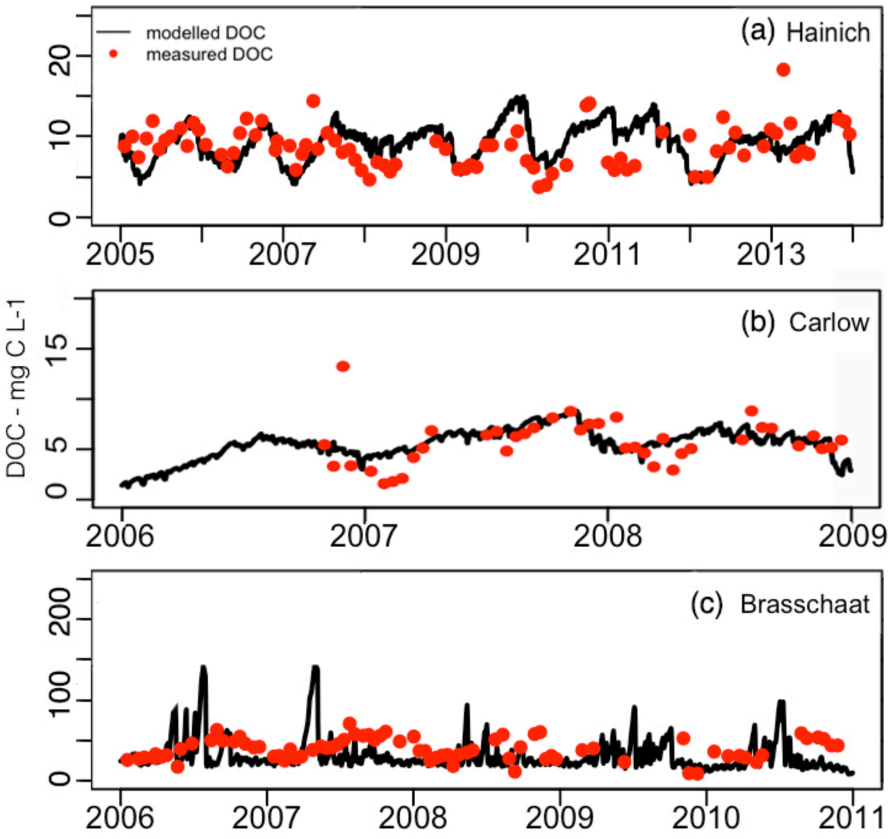

Figure 3. DOC concentration (mg C L^{-1}) at 10 cm depth measured (red dots) and simulated (black lines) for (a) Hainich, (b) Carlow, and (c) Brasschaat. Results for other depths are given in Figure S2.

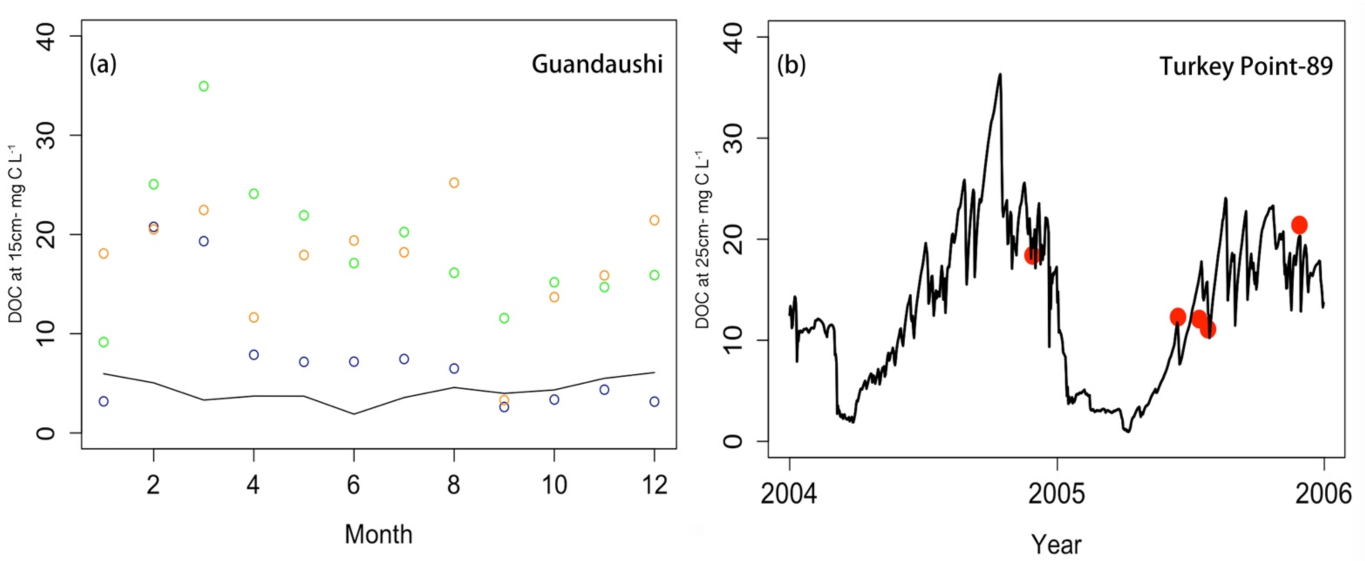

Figure 4. DOC concentration (mg C L^{-1}) for (a) Guandaushi at 15 cm measured (black circle: Chinese Fir, green circle: natural hardwood, orange circle: secondary wood) and simulated (black lines) and for (b) Turkey Point 89 at 25 cm measured (red dots) and simulated (black lines). Results for other depths are given in Figure S2.

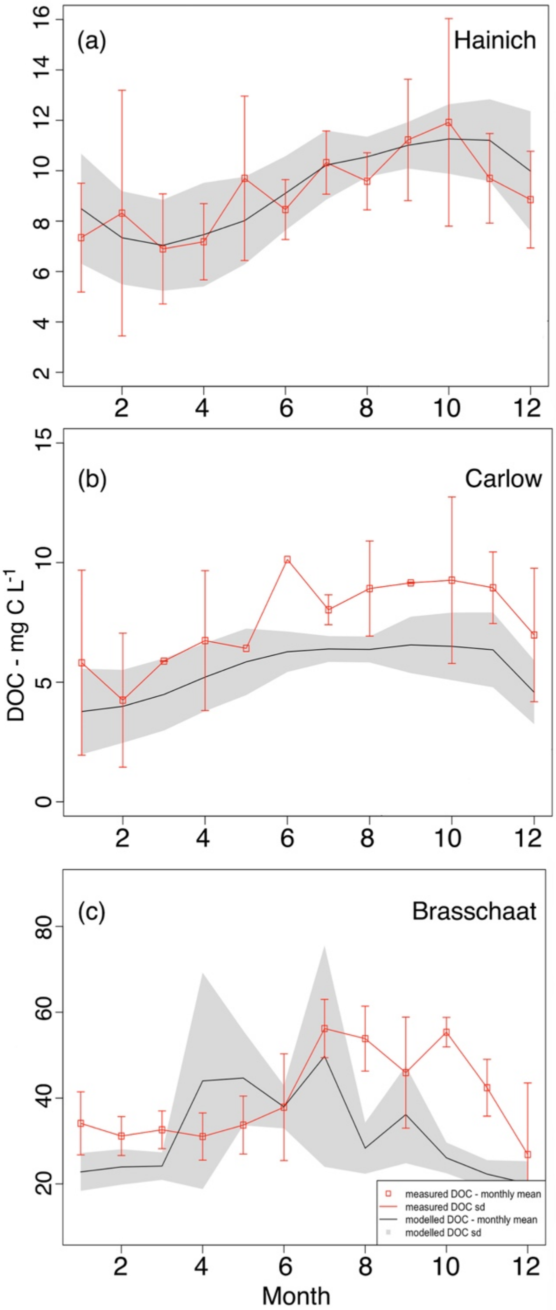

Figure 5. a) Monthly DOC (mg C L^{-1}) at 10 cm in Level 1- sites modelled (black line: mean, grey line: standard deviation) versus measured (red square: mean, red line: standard deviation) for studied period (a) Hainich averaging from 2005-2014 (b) Carlow, averaging from 2006-2008 (c) Brasschaat, averaging from 2006 –2010. Results for other depths are given in Figure S6.

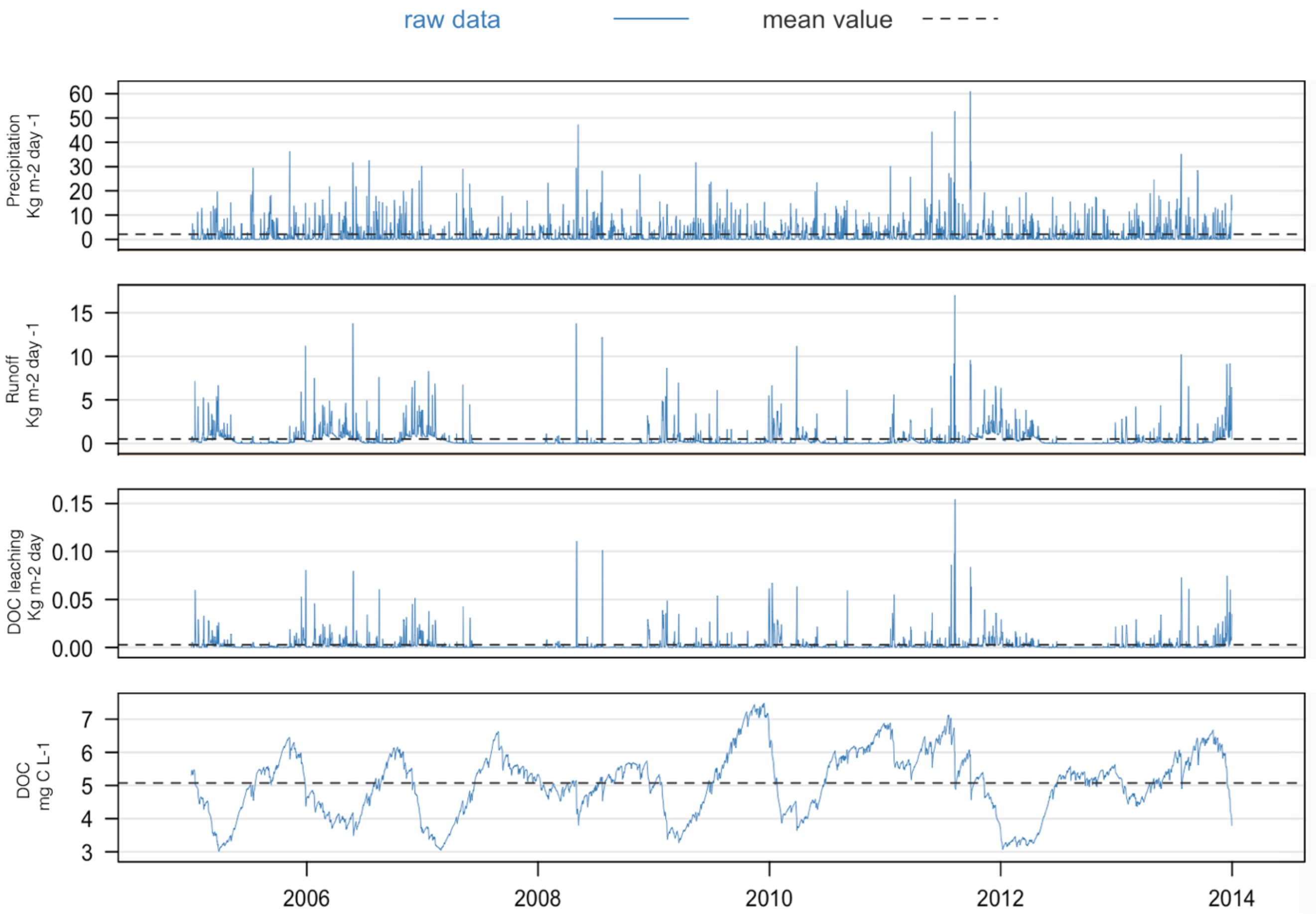

Figure 6. Observed precipitation, simulated runoff, DOC leaching and DOC concentration in Hainich from 2006 to 2013 indicating the relation between the averaged DOC concentrations at 3 m of soil with leaching as a result of runoff that follows large precipitation events.

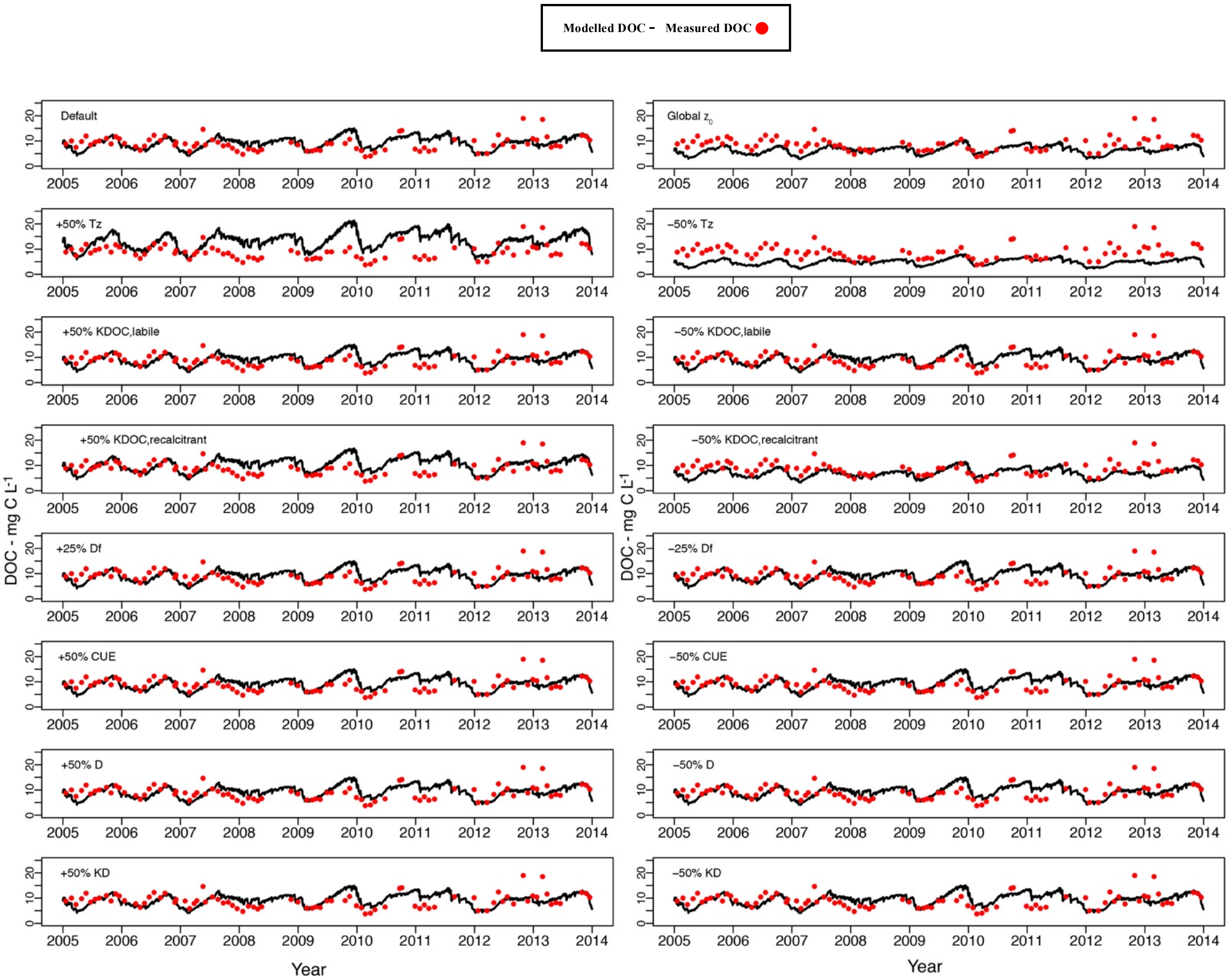

Figure 7. DOC concentration (mg C L^{-1}) simulated with sensitivity parameter sets (black line) versus measured (red dot) at 10 cm depth in Hainich for period 2004-2013. Parameter sets description and values are given in Table 1. Results for other sites are given in Figure S1.

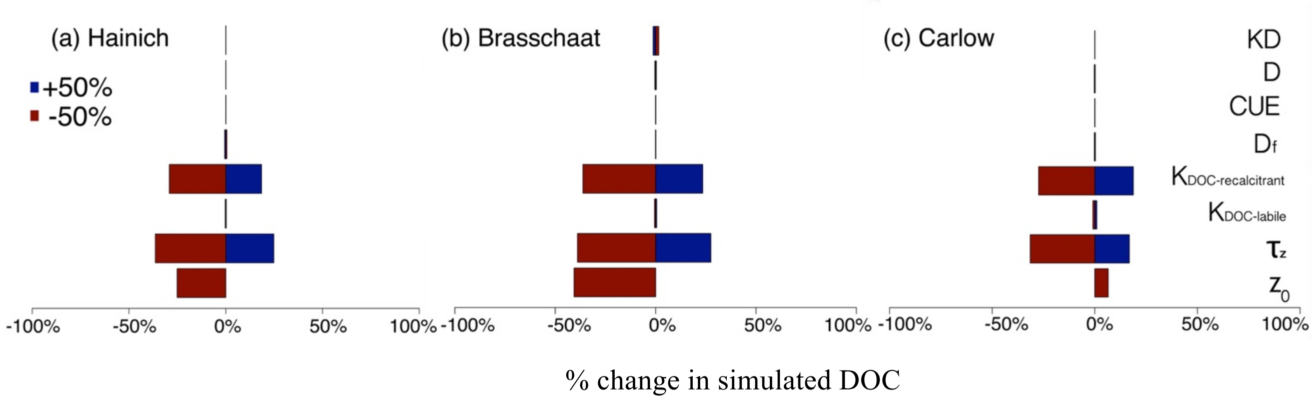

Figure 8. Relative change in simulated DOC (%) for a +50% (blue) and -50% (red) change in each parameter for level 1- sites: (a) Hainich, (b) Brasschaat and (c) Carlow. Values are given in Table S5.

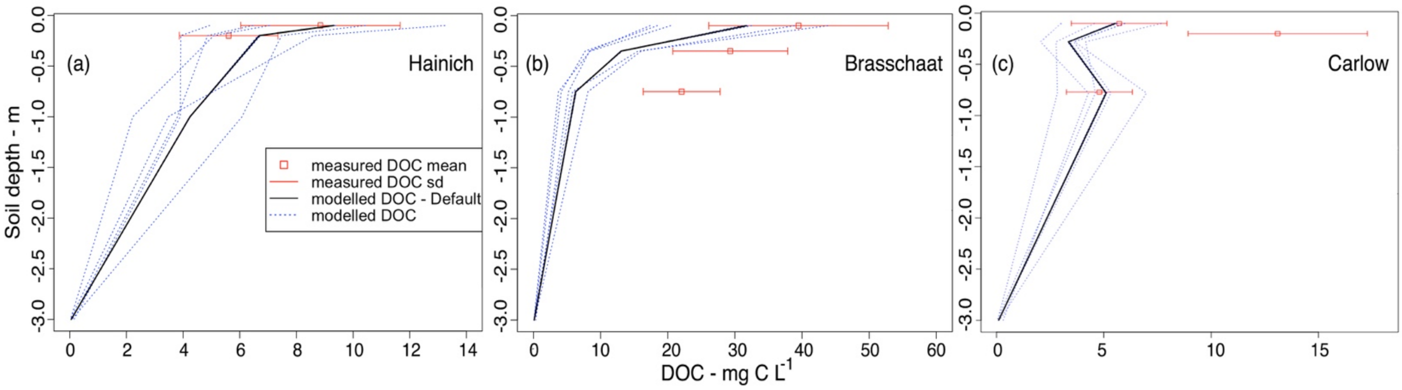

Figure 9. DOC concentration (mg C L^{-1}) in 3 m soil depth at level 1-sites modelled (black line: default parameter set; blue dashed line: sensitivity test parameter set) vs. measured (red square: mean; red line: standard deviation) for (a) Hainich (b) Brasschaat (c) Carlow. Plot of each parameter in 3 m soil depth in Figure S5.

Table 1. DOC relevant parameters in JULES-DOCM model

Parameter	Description	Value	Unit	Sensitivity test values (+/-)	
Carbon parameters					
z_0	e-folding depth of carbon content within 1 meter of soil [1]	Values range (65.68 - 167.13)	m^{-1}	PFT based	109.55
τ_z	Decay of Carbon decomposition with depth (z)[2]	2	m^{-1}	3.0	1.0
DOC parameters					
K_P	Rate constant for DOC production specific to each carbon pool[3]	1e-4, 5e-6, 5e-5, 2e-6	day^{-1}	-	-
$K_{DOC\ (labile)}$	Basal decomposition rate of dissolved DOC labile pool[4]	3.0 Value range (0.46-100)	days	4.5	1.5
$K_{DOC\ (recalcitrant)}$	Basal decomposition rate of dissolved DOC recalcitrant pool[5]	600.0 Value range (66-5000)	days	900.0	300.0
D_f	DOC production/decomposition modifier depending on clay and silt fraction[6]	0.75	-	1.0	0.5
CUE	Carbon use efficiency[7]	0.5	-	0.75	0.25
K_D	Distribution coefficient of adsorbed DOC [8]	8.05e-6	m^3 water Kg^{-1} soil	1.207e-4	4.025e-6
D	DOC diffusion coefficient [9]	1.062e-05	$m^2\ day^{-1}$	1.594e-05	5.313e-06

[1] Jobbágy & Jackson 2000
[2] Koven et al. 2013; Burke et al. 2016
[3] Jo Smith et al. 2010
[4] Kalbitz et al. 2003; Turgeon 2008
[5] Kalbitz et al. 2003; Turgeon 2008
[6] Parton, W et al. 1987
[7] Manzoni et al. 2012
[8] Moore et al. 1992
[9] Ota et al. 2013

Table 2. Symbols definition and units

Symbol	Units	Definition
BK	kg m^{-3}	Bulk density
$\boldsymbol{B_R}$		Fraction of soil respiration which is respired
$\boldsymbol{\beta_z}$	m^{-1}	Carbon distribution with depth, depending on biome
CUE		Carbon use efficiency
C_{DOC}	kg C m^{-2}	Amount of DOC subjected to transport by diffusion
D	m^{2} day^{-1}	DOC diffusion coefficient
D_f		DOC production / decomposition modifier depending on clay and silt fraction
dz	m	Soil layer thickness
$\Delta S_{C_{BIO}}$	kg C m^{-2} day^{-1}	Biomass carbon pool update
$\Delta S_{C_{DPM}}$	kg C m^{-2} day^{-1}	Decomposable plant material carbon pool update
$\Delta S_{C_{HUM}}$	kg C m^{-2} day^{-1}	Humus carbon pool update
$\Delta S_{C_{RPM}}$	kg C m^{-2} day^{-1}	Resistant plant material carbon pool update
ΔS_{DOC}	kg C m^{-2} day^{-1}	Labile and recalcitrant DOC pools update
AD_pot	kg C m^{-2} day^{-1}	Adsorbed/desorbed DOC from labile and recalcitrant pools
$F_{BIO_{IN}}$	kg C m^{-2} day^{-1}	Decomposed DOC flux from labile and recalcitrant pool into biomass pool
F_D	kg C m^{-2} day^{-1}	Labile and recalcitrant decomposed DOC flux
F_{Diff}	kg C m^{-2} day^{-1}	Flux of DOC transported by diffusion
$F_{P_{BIO}}$	kg C m^{-2} day^{-1}	DOC flux originated from biomass carbon pool
$F_{P_{DPM}}$	kg C m^{-2} day^{-1}	DOC flux originated from decomposable plant material carbon pool

Symbol	Units	Definition
$F_{P_{HUM}}$	kg C m^{-2} day^{-1}	DOC flux originated from humus carbon pool
$F_{P_{RPM}}$	kg C m^{-2} day^{-1}	DOC flux originated from resistant plant material carbon pool
$F_S(s)$	kg m^{-2}	Soil moisture rate modifier
$F_T(T_{soil})$	K	Soil temperature rate modifier
$F_v(v)$		Fractional coverage of a vegetation type
f_{dpm}		Fraction of litter that is decomposable plant material
K_P	day^{-1}	Rate constant for DOC production specific to the pool
K_{DOC}	days	Basal decomposition rate of dissolved DOC labile and recalcitrant pools
K_D	m^3 water Kg^{-1} soil	Distribution coefficient of adsorbed DOC
Λ_c	kg C m^{-2} day^{-1}	Litterfall rate
L_T	kg m^{-2} day^{-1}	Leaching from labile and recalcitrant DOC pools in top soil
L_S	kg m^{-2} day^{-1}	Leaching from labile and recalcitrant DOC pools in sub soil
m		DOC decomposition rate type (labile or recalcitrant)
R_{BIO}	kg C m^{-2} day^{-1}	Respiration from biomass carbon pool
R_{DPM}	kg C m^{-2} day^{-1}	Respiration from decomposable plant material carbon pool
R_{DOC}	kg C m^{-2} day^{-1}	Respiration from labile and recalcitrant DOC pools
R_{HUM}	kg C m^{-2} day^{-1}	Respiration from humus carbon pool
R_{RPM}	kg C m^{-2} day^{-1}	Respiration from resistant plant material carbon pool
$Roff_{surf}$	kg m^{-2} day^{-1}	Surface Runoff

Symbol	Units	Definition
$Roff_{sub}$	kg m^{-2} day^{-1}	Sub-Surface Runoff
S_C	kg C m^{-2}	Soil carbon storage
$\boldsymbol{T_s}$	kg m^{-2}	Soil moisture content
S_{DOC}	kg C m^{-2}	Labile and recalcitrant DOC storages
$S_{DOC_{ad}}$	kg C m^{-2}	Adsorbed labile and recalcitrant DOC storages
$\boldsymbol{\theta_v}$	kg m^{-3}	Volumetric Soil moisture content
τ_z	m^{-1}	Decay of Carbon decomposition with depth
z	m	Soil depth
z_0	m	e-folding depth of carbon content within 1 meter of soil

Table 3. Data availability for model evaluation at different

Sites	Brasschaat[1]	Carlow[1]	Guandaushi[2]	Hainich[1]	Turkey-point89[2]
			Carbon fluxes		
GPP	2000-2006	2008		2000-2012	2005-2008
NPP	2000			2000-2007	2005-2008
Soil respiration	2000-2006			2000-2007	2005-2008
C content	1995-1998	2006-2009		2000-2007	2004-2006
			DOC measurements		
1 year			1999		
1 to 5 years		2006-2009			2004-2005
5 to 10 years	2000-2008			2001-2014	

1. level 1 site 2. Level 2 site

Table 4. Evaluation Level 1-sites characteristics

	Site		
	Brasschaat	**Carlow**	**Hainich**
		Characteristics	
Ecosystem	Evergreen forest	Grassland	Deciduous forest
Soil classification	Arenosol	Luvisol	Cambisols
BK (kg m-3)	1.4	1.07	1.2
Clay (fraction)	0.034	0.22	0.589
Sand (fraction)	0.8912	0.51	0.031
Silt (fraction)	0.0748	0.27	0.38
		Measurement depth (cm)	
Carbon content	100[1]	50[2]	60[3]
DOC concentration	10,35,75	5,10,20	10-77
		FLUXNET meteorological observations	
Period	1996-2014	2004-2014	2004-2009

1. Janssens et al. 1999 2. Kindler & Siemens 2010 3. Schrumpf et al. 2011

Table 5. The measured (Obs.) vs. the modelled (Mod.) carbon fluxes, SOC concentration and soil DOC concentration at different soil depths in five study sites.

Variables	**Level-1**						**Level-2**			
	Brasschaat		**Carlow**		**Hainich**		**Turkey Point-89**		**Guandaushi**	
	Obs.	**Mod.**	**Obs.**	**Mod.**	**Obs.**	**Mod.**	**Obs.**	**Mod.**	**Obs.**	**Mod.**
	Carbon fluxes (g C m^{-2} yr^{-1}) and SOC (kg C m^{-2})									
GPP	1173±92	867±25	903	1165	1606±102	1455±168	1732±108	1635±63	-	-
NPP	850	596.1	-	-	673±33	833±153	814±51	1013±92	-	-
Soil Res*	411±34	625±54	-	-	883±206	909±66	693±16	1006±142	-	-
SOC	11.47	8.01	2.3	4.17	11.75	8.63	1.85	3.39	-	-
	DOC concentration (mg C L^{-1})									
10 cm	39±15	28±13	7±3	6±1	9±3	9±2	-	-	-	-
15 cm	-	-	-	-	-	-	-	-	nh: 19±12 sh:17±12 cf: 8±15	4±1
20 cm	-	-	-	-	6±2	7±2	-	-	-	-
25 cm	-	-	-	-	-	-	15±4.5	16±4	-	-
10-28 cm	-	-	13±4	4±1	-	-	-	-	-	-
30 cm	-	-	-	-	-	-	-	-	nh: 9±7 sh: 15±8 cf: 7±17	3±1
35 cm	29±2	13±9	-	-	-	-	-	-	-	-
75 cm	22±1	6±6	-	-	-	-	-	-	-	-
28 to 77	-	-	5±2	5±0.2	-	-	-	-	-	-
100 cm	-	-	-	-	-	-	2.2±0.2	7.9±2	-	-

*: soil respiration, nh: Natural Hardwood; sh: Secondary hardwood; cf: Chinese fir

Representation of dissolved organic carbon in the JULES land surface model (vn4.4_JULES-DOCM)

Supporting information

Table S1. Z_0 values for each PFT

PFT	Z_0
Boreal forest	0.775625
Crops	1.13717
Deserts	1.67113
Sclerophyllous shrubs	1.22839
Temperate deciduous forest	0.725914
Temperate evergreen forest	0.857235
Temperate grassland	1.22839
Tropical deciduous forest	1.67113
Tropical evergreen forest	1.0188
Temperate grassland/savana	1.45185
Tundra	0.656898

Table S2. Z_0 values for each PFT in JULES-DOCM

JULES PFT	Z_0
Tropical broadleaf evergreen forest	1.0188
Temperate broadleaf evergreen forest	0.857235
broadleaf deciduous forest	0.725914
needle-leaf evergreen forest	0.857235
needle-leaf deciduous forest	0.725914
C3 grass	1.13717
C4 grass	1.13717
Evergreen shrubs	1.22839
Deciduous shrubs	1.22839

Table S3. Anova test results for Carbon fluxes (Df: Degree of freedom, Sum sq: sum of squares, Mean sq: mean of squares, Pr: p-value)

ANOVA

	Df	Sum sq	Mean Sq	F value	Pr(>F)
GPP					
Hainich	1	102955	102955	5.352	0.0343
	16	307785	19237		
Brasschaat	1	280924	280924	62.27	1.33E-05
	10	45117	4512		
Turkey Point-89	1	9295	9295	4.714	0.162
	2	3943	1972		
NPP					
Hainich	1	88254	88254	7.222	0.0362
	6	73323	12220		
Turkey Point-89	1	39632	39632	7.154	0.116
	2	11080	5530		
SOIL RESPIRATION					
Hainich	1	1400	1400	0.06	0.815
	6	140896	23483		
Brasschaat	1	160497	160947	77.44	1.40E-06
	12	24870	2073		
Turkey Point-89	1	98114	98114	9.687	0.0896
	2	20256	10128		

We include all the sensitivity runs for Level-1 sites: Hainich, Brasschaat and Carlow for all the depths where the measurements were available. Red points are indicating measurements where black points are values from model (Fig. S2). Also representing Level-2 sites: Turkey Point 89 and Guandaushi comparison of modelled DOC versus measured in deeper soil depths (Fig. S3).

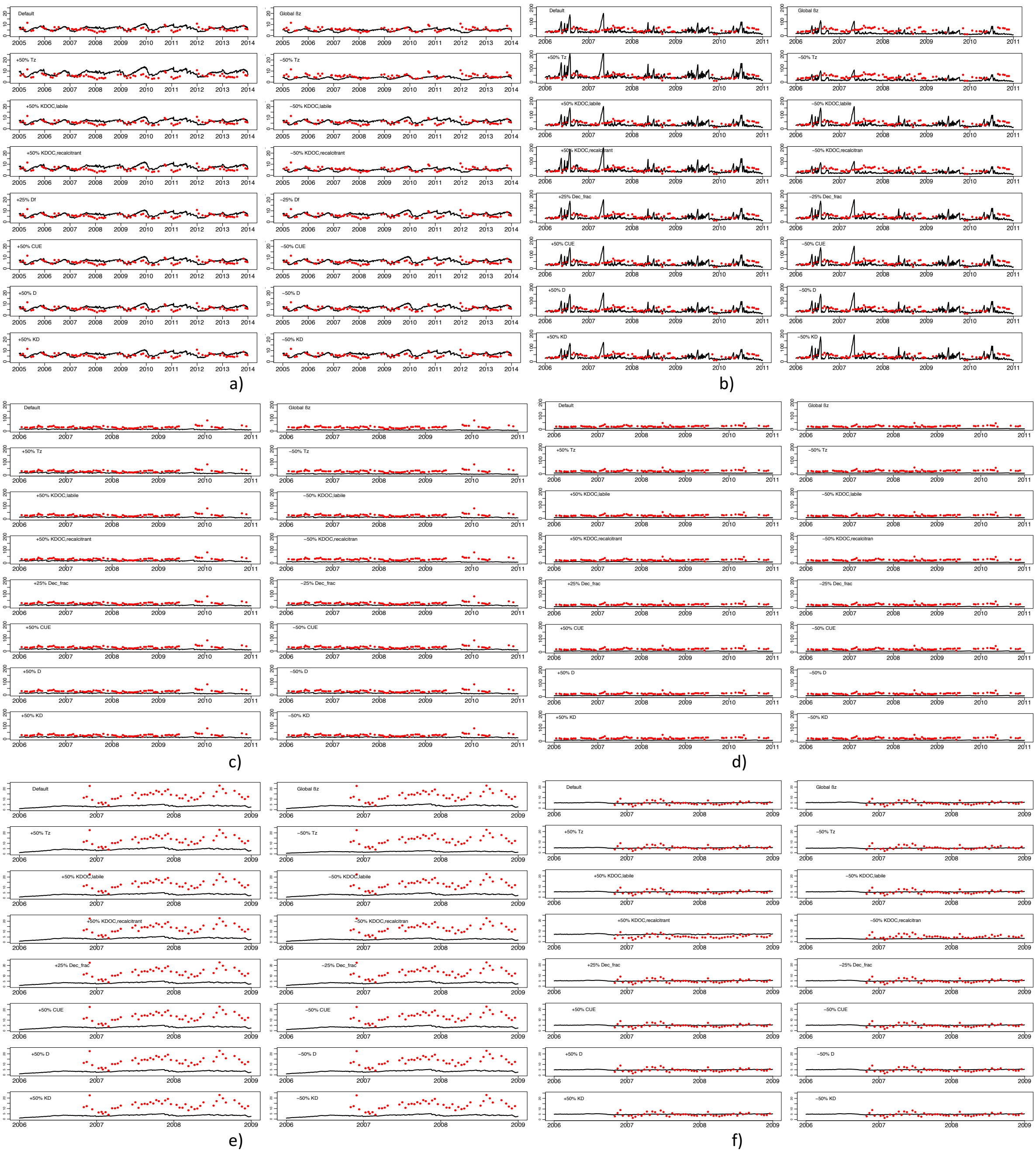

Figure S1. Sensitivity tests (black line) versus measurements (red dot) at a) 20cm depth – 2004-2013, Hainich b) 10cm depth – 2006-2010, Brasschaat c) 35cm depth – 2006-2010, Brasschaat d) 75cm depth – 2006-2010, Brasschaat e) 10 to28cm depth – 2006-2009, Carlow f) 28 to78cm depth – 2006-2009, Carlow; X axis is year and Y axis is DOC concentration in mg C L^{-1}. Parameter sets description and values in Table 1

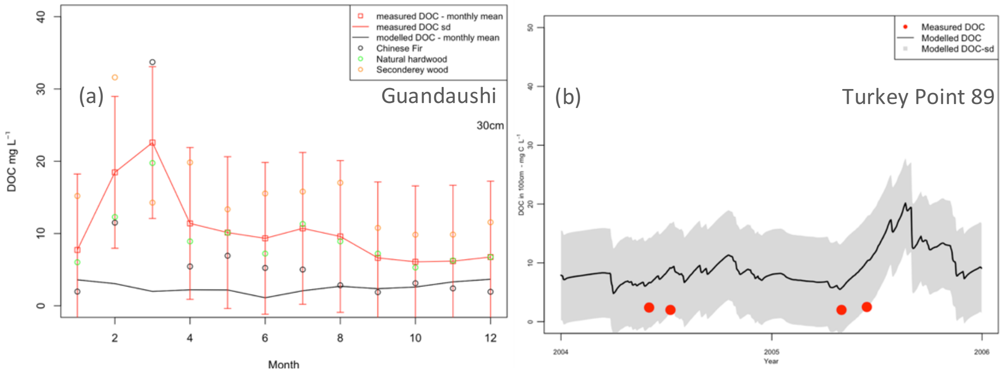

Figure S2. DOC concentration (mg C L-1) for (a) Guandaushi at 30 cm measured (black dot: Chinese Fir, green dot: natural hardwood, orange dot: secondary wood, red square: mean, red line: standard deviation) and simulated (black lines) and for (b) Turkey Point 89 at 100 cm measured (red dots) and simulated (black lines, grey line: standard deviation).

We examine the hydrology of the model and its interaction with DOC concentration and leaching for Level-1 sites: Carlow and Brasschat (Fig. S4) and overall model performance in DOC representation in all depths by comparing modelled versus measurements during study period in Hainich, Brasschaat and Carlow (Fig. S5).

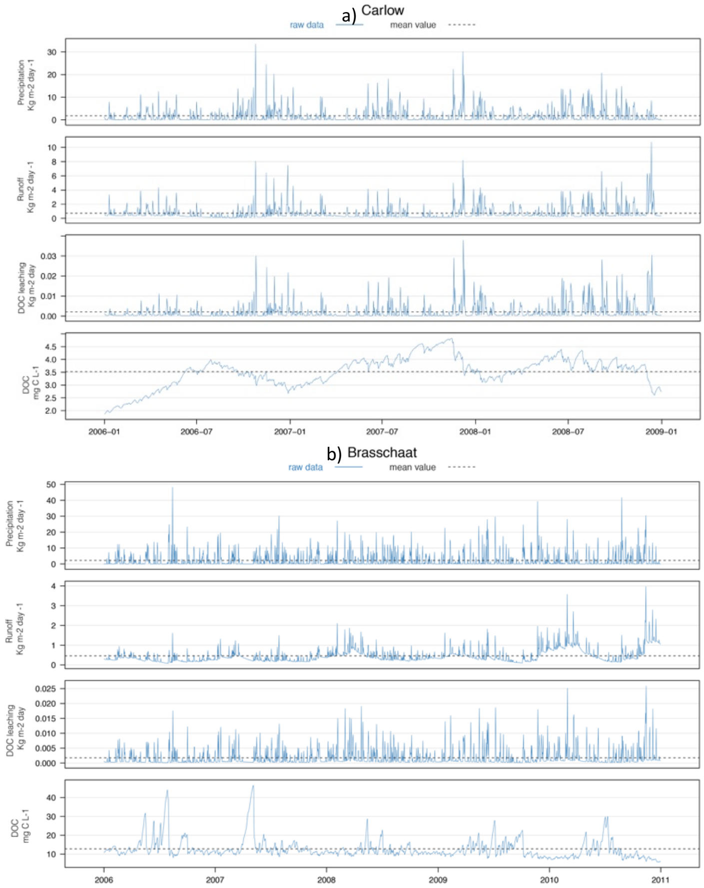

Figure S3. a) Precipitation, runoff, DOC leaching and DOC concentration in Carlow from 2006 to 2009 b) in Brasschaat from 2006 to 2010 simulated data with JULES-DOCM

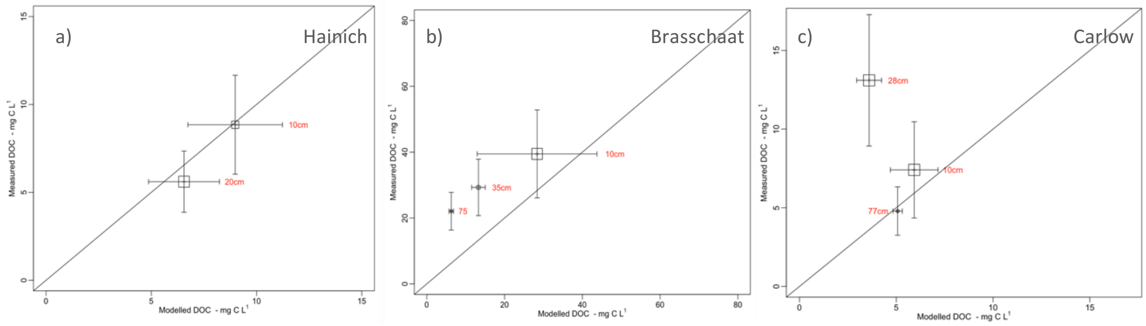

Figure S4. a) Measured vs modelled DOC (mg C L-1) with default set in Hainich from 2006 to 2013 at 10 and 20cm b) in Brasschaat from 2006 to 2010 at 10,35 and 75cm c) in Carlow from 2006 to 2009 at 10,10 to 28 and 28 to 78cm

Sensitivity of model parameters (Table S2) was tested in Level-1 Sites for the depths where the DOC measurements where available (Fig. S5) and the results were reported in percentage of change compared to default parameters set (Table S3). Anova test was used in order to determine each parameter's impact significance on DOC representation (Table S4).

Table S4. JULES-DOCM parameters set for sensitivity test

ID	*Description*
SET-1	*Default*
SET -2	*Global* β_z
SET -3	*+50%* τ_z
SET -4	*-50%* τ_z
SET -5	*+50%* $K_{DOC,\ labile}$
SET -6	*-50%* $K_{DOC,\ labile}$
SET -7	*+50%* $K_{DOC,\ recalcitrant}$
SET -8	*-50%* $K_{DOC,\ recalcitrant}$
SET -9	*+25%* D_f
SET -10	*-25%* D_f
SET -11	*+50% CUE*
SET -12	*-50% CUE*
SET -13	*+50% D*
SET -14	*-50% D*
SET -15	*+50%* K_D
SET -16	*-50%* K_D

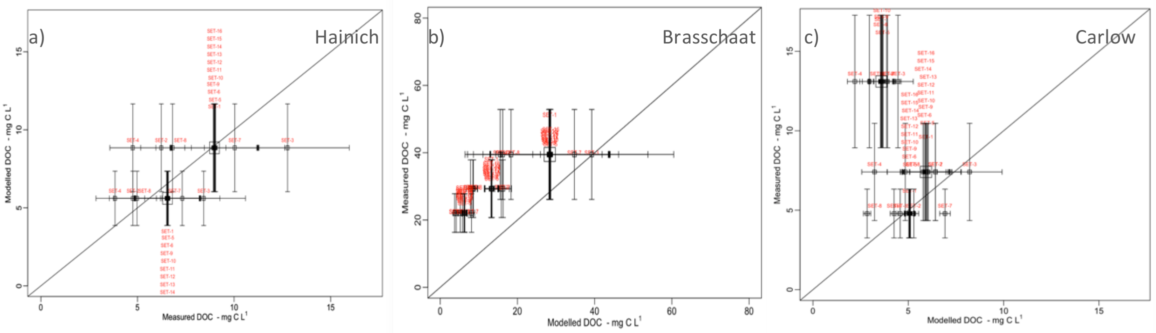

Figure S5. a) Measured (X axis) vs modelled (Y axis) DOC sensitivity runs (Table1) in Hainich from 2006 to 2013 at 10 and 20cm b) in Brasschaat from 2006 to 2010 at 10,35 and 75cm c) in Carlow from 2006 to 2009 at 10,10 to 28 and 28 to 78cm

Table S5. Relative change in simulated DOC (%) for a +50% and -50% changes in model parameters for Hainich, Carlow and Brasschaat

HAINICH	Z_0	τ_Z	$K_{DOC-LABILE}$	$K_{DOC-RECALCITRANT}$	D_F	CUE	D	K_D
50%	**-**	**24.82748**	**0.2746118**	**18.51532**	**-0.590175**	**-0.06117**	**-0.04519973**	**-0.05364935**
-50%	**-25.12175**	**-36.45213**	**-0.2729602**	**-29.18424**	**0.586697**	**0.06068464**	**0.04545651**	**0.05483722**
CARLOW								
50%	**-**	**16.81662**	**0.9795210**	**18.77268**	**-0.2522954**	**-0.08957044**	**0.1873639**	**-0.05575369**
-50%	**6.52764**	**-31.50205**	**-0.9754175**	**-27.40512**	**0.2517171**	**0.08983245**	**-0.2774403**	**0.05659863**
BRASSCHAAT								
50%	**-**	**27.52056**	**0.5294166**	**23.45682**	**-0.1300973**	**-0.1176923**	**-0.3806475**	**-1.256365**
-50%	**-40.6144**	**-38.92471**	**-0.5120930**	**-36.20752**	**0.1305755**	**0.1183834**	**0.3794148**	**1.571266**

Table S6. Anova test results for sensitivity test of Level-1 sites

DOC Hainich	Df	Sum sq	Mean Sq	F value	Pr(>F)
			ANOVA		
set-2	1	10686	10686	1066	**2.00E-16**
set-3	1	10437	10437	491.9	**2.00E-16**
set-4	1	22499	22499	2572	**2.00E-16**
set-5	1	1	1.277	0.094	0.759
set-6	1	1	1.267	0.093	0.76
set-7	1	5805	5805	382.7	**2.00E-16**
set-8	1	14422	14422	1322	**2.00E-16**
set-9	1	6	5.898	0.437	0.508
set10	1	6	5.828	0.427	0.514
set11	1	0	0.063	0.005	0.946
set12	1	0	0.062	0.005	0.946
set13	1	0	0.035	0.003	0.96
set14	1	0	0.035	0.003	0.96
set15	1	0	0.049	0.004	0.952
set16	1	0	0.049	0.004	0.952

DOC Carlow	Df	Sum sq	Mean Sq	F value	Pr(>F)
			ANOVA		
set-2	1	116	115.98	20.27	**6.81E-06**
set-3	1	770	769.7	109.73	**2.00E-16**
set-4	1	2701	2701.1	703.7	**2.00E-16**
set-5	1	3	2.611	0.486	0.486
set-6	1	3	2.59	0.49	0.484
set-7	1	959	959.2	145.6	**2.00E-16**
set-8	1	2044	2044.2	488.3	**2.00E-16**
set-9	1	0	0.173	0.033	0.857
set10	1	0	0.172	0.032	0.857
set11	1	0	0.022	0.004	0.949
set12	1	0	0.022	0.004	0.949
set13	1	0	0.096	0.018	0.894
set14	1	0	0.096	0.018	0.894
set15	1	0	0.008	0.002	0.968
set16	1	0	0.008	0.002	0.968

DOC Brasschaat	Df	Sum sq	Mean Sq	F value	Pr(>F)
			ANOVA		
set-2	1	98664	98664	5924	**2.00E-16**
set-3	1	45301	45301	126.7	**2.00E-16**
set-4	1	90625	90625	585.2	**2.00E-16**
set-5	1	17	16.76	0.069	0.792
set-6	1	16	15.69	0.065	0.798
set-7	1	32911	32911	109.1	**2.00E-16**
set-8	1	78414	78414	462.8	**2.00E-16**
set-9	1	1	1.01	0.004	0.948
set10	1	1	1.02	0.004	0.948
set11	1	1	0.83	0.003	0.953
set12	1	1	0.83	0.003	0.953
set13	1	9	8.67	0.036	0.849
set14	1	9	8.61	0.036	0.85
set15	1	94	94.41	0.415	0.519
set16	1	148	147.7	0.566	0.452

Seasonality of DOC concentration in different depths of Level-1 sites (Hainich, Carlow and Brasschaat) was tested by comparing monthly modelled DOC means versus measurements (Fig. S7).

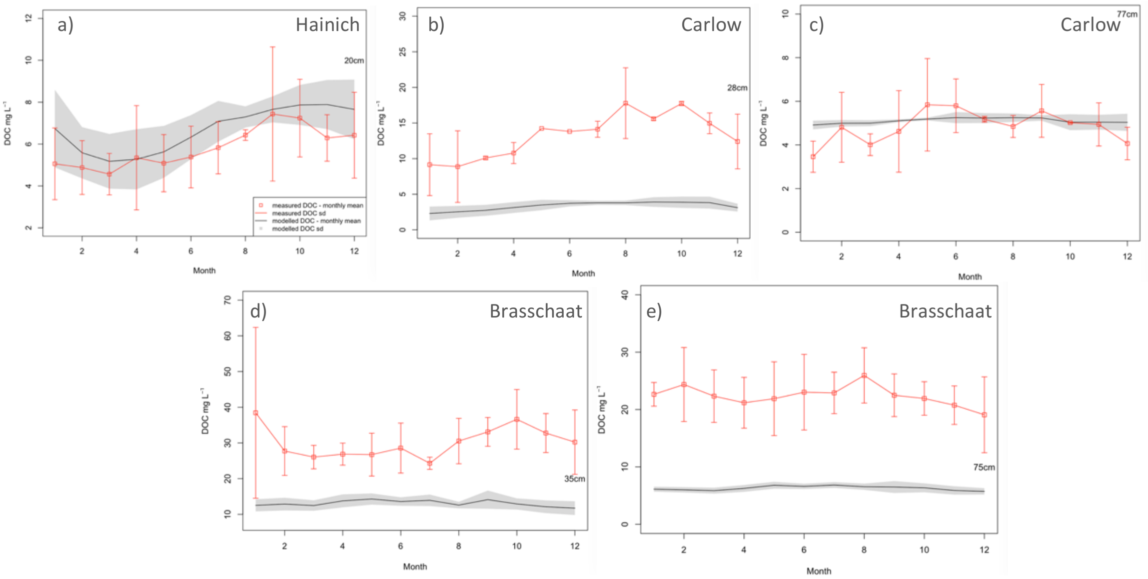

Figure S6. Monthly DOC means modelled versus measured (mg C L-1) for studied period at a) 20cm of Hainich b)10 to 28cm of Carlow c) 28 to 77 cm of Carlow d) 35cm of Brasschaat e) 75cm of Brasschaat

Reference:

Jobbágy, E.G. & Jackson, R.B., 2000. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecological Applications, 10(2), pp.423–436.