First, it is a really long paper and has been largely focused on the comparison with different observation data. The introduced new processes and the used parameters are staying as its original sources (except one parameter)? There is no sensitivity testing showing how different process/parameter interactively influence DOC concentration/fluxes. Is it real that water routing (or maybe some other hydrological) parameters introduced from Amazon basins (Lauerwald et al. 2017) can directly work on European basins? I am a bit surprised to see that without any systematic testing of these “new” parameters, the authors can already produce reasonable runoff and also DOC fluxes, as claimed.

We agree that our paper is on the long side because it includes a part on model evaluation and another, central part on model application targeting the spatio-temporal patterns of pan-European DOC leaching fluxes and their underlying drivers.

Regarding model parameter adjustment, ORCHILEAK (and its parent branch ORCHIDEE) is a model that has ultimately been developed for global-scale applications. Therefore, the overall philosophy of our modeling approach is indeed to try to minimize the number of parameters that need adjustments before reaching reasonable results.

Importantly, most of the hydrological parameters have actually been calibrated and evaluated against globally observed runoff (Ringeval et al., 2012; Yang et al., 2015) rather than only for the Amazon basin. That is why the model performs well when it is applied to different basins and this aspect will be clarified in the revised ms. Furthermore, to improve the model performance for European river networks, we have updated ORCHILEAK with the representation of hydrological processes from a more recent version of the ORCHIDEE LSM, which has been found to perform well for temperate regions in the US (MacBean et al. 2020). Therefore, a recalibration from our side was not necessary. Regarding the C fluxes, we adopted most of the configuration used in the Amazon and Congo River basins, but a few parameter adjustments were required. For instance, we adjusted two parameters which control the leaching of DOC from the top soils to
the headwater streams (see $r_{\text{con}}$ and $r_{\text{gen}}$ in our response to comment #2 of reviewer #1). Given the length of the paper, it was decided to exclude the results showing how these parameter adjustments influence the simulations because the main aim of the paper is on the application, and not on the model itself.

Secondly, as also pointed out by authors that DOC is only tiny fraction of NPP, and the evaluation of NPP with the observation did not seem to be that important for me comparing to understanding DOC-related processes in regulating the spatial and temporal patterns. I would rather see how the modelled soil temperature and moisture compare with the observations or maybe also soil CO2 fluxes to indicate potential microbial activities. Then, I think the manuscript need to look into how these parameters used in simulating DOC influencing the simulated DOC fluxes and concentration. The evaluation of NPP and SOC should probably not be in the main text (or at least reduce the length largely);

We believe that an evaluation of the NPP and SOC stocks is useful because a main aim of our research is to assess the extent to which the leaching fluxes influence the European terrestrial C budget across scales, from the basin to continental scale. Although indirect, terrestrial NPP is also found to be a key control factor of DOC leaching fluxes. However, we agree with the reviewer that other controlling factors of the DOC leaching fluxes come into play. Following her/his recommendation, our revised ms. will include a comparison of simulated temperature against the European Centre for Medium-Range Weather Forecasts reanalysis ERA5 dataset. For soil respiration and moisture, we’ll use SRBD data (Soil Respiration Database, Jian et al., 2021).

Thirdly, the modelled data has been averaged at different spatial domains, i.e., climate zones, catchments, European continent and also at specific points. What are the motivation of presenting data at these many levels of scales? This is one of the reasons I think why the manuscript becomes so long and the current presentation is not super clear about the key message.

We agree with the reviewer, but unfortunately the model evaluation relies on observational data or previously published studies that have typically addressed different spatial scales. For the results of our study, we work on the cell scale and to visualize better the spatial distribution and the possible drivers, we choose to work with climate zones. In addition, each main spatial scale primarily serves a specific purpose: the grid and basin scales are the most relevant for the model evaluation; the climate zone scale is of interest for our research objective to identify the drivers of the spatio-temporal dynamics of the leaching fluxes, and the pan-European scale provides the basis for the implications of our results for the overall C budget. This rationale will be better explained in the revised version of our ms. and we will also improve the presentation to make sure that our multi-scale approach does not lead to confusion.

Last, I have listed a lot of detailed comments on the method and model description section. I think authors should definitely clarify in many parts and justify the reasons behind. In general, I would like to see testing/understanding of model process interactions, parameter values (and their sources) before digging into any spatio-temporal patterns.

It was decided not to focus on testing/understanding model process interactions and show how parameter adjustments influence the simulations because the main aim of the paper is on the application, and not on the model itself. We agree that parameters values and their sources should be written clearer in the main text.
Detailed comments:

Through the text, please consider to specify either soil DOC or river DOC fluxes/concentration, not using unspecified DOC fluxes/concentration. It becomes a bit unclear in some parts of text if it is describing soil or river processes.

Agreed. We will carefully check that this important point is fully addressed in our revised version.

L93-94, need studies to support the negligible point sources.

Meybeck (1986) showed that DOC from sewage is very labile and does only affects concentration within short distances downstream of water processing plants.

L151-152, how is soil carbon distributed in these 11 layers? In Fig 1, it says in each layer there are all of these 11 carbon pools? If so, in deeper soil, how these litter pools look like? A few comments about Fig. 1, I saw there is C feedbacks from passive SOC à active or slow SOC à active, what do these arrows mean?

For the litter pool we distinguish the metabolic and structural litter pool. Active, slow and passive represent three states of SOC that can decompose according to the soil moisture and soil temperature. The vertical soil profile is described by an 11-layer discretization of a 2 m soil profile, with geometrically increasing layer thickness from top to bottom (Fig. 1 and Lauerwald et al., 2017). In each soil layer, ORCHILEAK explicitly simulates the fresh litter input (depending on the simulated vertical root distribution), decomposition of each organic matter pool (including litter and SOC), C transformation between different organic matter pools (showed by arrows between different pools in Fig. 1), C transport and diffusion between neighboring soil layers, and the loss of DOC due to leaching. The organic matter decomposition scheme in ORCHILEAK follows the CENTURY model (Parton et al., 1987). For a specific organic C pool at each time step, only a fraction of the decayed C is respired as CO₂ to the atmosphere (mimics the microbial respiration), the remaining will be transferred to other organic pools (mimics the microbial growth and mortality). The arrows between different SOC pools denotes the transformation between different SOC pools during the microbial decomposition processes. Please see Parton et al. (1987) for detailed explanation on the biogeochemical mechanisms of the CENTURY model.

L156, Lateral leaching of DOC is not clearly marked in Fig.1, as advection is only markedly up and down, so where is runoff?

Agreed, this was confusing. In the revised ms., fig.1 (left panel) will be improved to better represent the lateral C fluxes associated to runoff and drainage. We will also show more clearly that the RHS of the figure is actually a zoom of the processes occurring within a given layer.

L157, “drainage” could often mean surface runoff, consider to call it subsurface runoff.

To avoid confusion, we will define precisely what each of these terms precisely mean: drainage or subsurface runoff is the water flux from the last layer soil (~2m) layer and runoff is the water flux from the topsoil surface.

L158-159, sorption-desorption also contribute to the soil DOC dynamics.

Agreed. The sentence will be modified as follows: “The DOC dynamics in the soil
are controlled by production, decomposition, transport and sorption-desorption processes..."

L164-165, where did you get the diffusion coefficient and what is the value you used? It needs to be clear here.

The diffusion coefficient is from Ota et al. (2013) and is equal to \(1.06 \times 10^{-5} \text{ m}^2 \text{ d}^{-1}\).

L168-169, need to explain what is Fickian-type transport and how did you implement it in the model to represent bioturbation?

The effect of mixing by bioturbation on the distribution of soil properties is commonly represented in models as a diffusion-like process (e.g. Camino et al., 2018). In this case, the intensity of mixing of each variable (including solid species) is proportional to its concentration gradient, and the proportionality constant (the "bioturbation coefficient") was set to \(2.74 \times 10^{-7} \text{ m}^2 \text{ d}^{-1}\), according to Koven et al. (2013). This description will be added to our revised ms.

L171, which time scale (daily?) do soil DOC concentrations reach equilibrium based on diffusion? And then is there an order how different processes (production, decomposition, bioturbation, diffusion, sorption-desorption, advection) were implemented?

The DOC profile is time-variant and does reach equilibrium (or steady-state). Its temporal evolution is simulated at the daily time-step, in the following order: First, production and decomposition fluxes are calculated, and the DOC stocks per soil layer and poll are updated accordingly. Then follow the simulation of DOC advection in the soil column, followed by diffusion of DOC. Finally, the export of DOC through leaching from top- and bottom soil with runoff and drainage, respectively, are calculated. We will add those details to the model description in the method section.

L199-202, how did you separate solid and particulate litter and SOC pools? What do you mean "the production of DOC is dependent on leaching rates"? this is not what provided in equation 8. Please clarify this.

This sentence was confusing and will be changed to: “While preserving the basic structure of ORCHIDEE-SOM, we thus adapted the model in a way that organic C exchange occurs mainly among the litter and SOC pools, similar to the original Century model, while the production of DOC as side product of this C exchange is dependent on production rates as used in ECOSSE.”

L211, As far as I know, this linear equilibrium is mainly tested for peatland soils, see Yurova et al., (2008), which is not considered in this study. For mineral soils, it often uses Langmuir isotherm, see Lilienfein et al. (2004), Kothawala et al., (2008), and Tang et al., (2018). I would like to see some studies supporting of using linear equilibrium partition equations for mineral soils. Furthermore, equation 9 ignores the initial desorption term when the DOC concentration is zero. For sorption-desorption, which time step in the model can the equilibrium reach?

We used a linear adsorption isotherm as in Neff and Asner (2001) and Wu et al. (2014). We assume that equilibrium between the dissolved and absorbed phases is instantaneous. Moreover, the work by Kothawala et al. 2008 showed that linear equation also performed fairly well and were parametrized on more observations than Langmuir equations.
Equation 4-11, there are a lot of constant used and also temperature and soil moisture dependent equations. I think it is important to present these parameter values (and the sources where these values are from). I would like to see how these $k_L$, $k_{SOC}$, $k_{DOC}$ are modelled.

In the revised ms. a supplementary table with all parameter values along with appropriate references will be added.

L222-224, L225-226, There are many constants introduced in this section, I did not track if these cited studies are also from the same region, and it could be good to know. Depending on how valid these fractions are, the conclusion on manure impacts on DOC could vary, right?

Indeed, the type of manure input in different regions can be very different, and the physicochemical properties (e.g. C:N ratio and the ratio of dissolved and particulate organic matter) depends strongly on the specific type of manure input. However, current forcing data of manure only provide the amount of annual manure inputs, but not the specific composition and physicochemical properties of the manure. Therefore, following ORCHIDEE-CNP (Goll et al., 2017), we set the C:N ratio and DOC:POC ratio of manure to the mean value of previous observations.

L237-239, This sentence is not clear. Where is surface runoff? As it says water budget, it should also consider water content changes between two timesteps.

Of course. The paragraph will be altered as: "The water budget within the soil is determined by the infiltration rate, the evaporation and transpiration from the soil, runoff from the top soil and drainage at the bottom of the soil column. An imbalance between these fluxes leads to changes in water content within the soil."

Table 1. suggest to list the covering periods for some of the dataset. What do you mean by "After HSWD V1.1"? Then why not use soil texture data from HSWD, same as other soil properties?

We will delete ‘After’, and simply write the name of the source ‘HSWD V1.1’. The ‘After’ was referring to the fact that the original data that has a higher resolution was aggregated to a 0.5 degree raster which is consistent with the other forcing data used. The fact that soil texture is taken from another source may seem odd, but is consistent with older applications of ORCHILEAK (Lauerwald et al. 2017, Hastie et al. 2019, 2021). We kept the Reynolds et al. (1999) as source for soil texture, as the soil hydrology model was calibrated to work with the soil classes used in that dataset.

L280-282, It definitely needs more details about which hydrological parameter values were based on Lauerwald et al., (2017). For me, I cannot imagine the tuned hydrological parameters for Amazon basin can "directly" work for different basins in Europe. Please justify this! Then where can we see how the calibration of surface roughness works, what values were used before and after calibration, and which discharge datasets used for this calibration? Was this discharge dataset used for calibration also included in the evaluation later in figures? Need clarification.

ORCHIDEE is a global model and the hydrology had been calibrated already. Nevertheless, we updated ORCHILEAK with a more recent hydrology scheme used in the recent standard version of ORCHIDEE (from ORCHIDEE 2.0). MacBean et al. 2020 has evaluated the model for temperate systems (see also comment 1
L287-289: You had the climate forcing since 1978, why not looping period starting from 1978 for spinup? For spinup and transient periods, do you still run the model with 3-hour resolution or can the model run in coarser temporal resolution? If so, there are many other longer time series data can be used for these two periods to conquer the issues of inconsistent dataset, e.g., CO2 data from 1861 with the repeated climate forcing from 1980-2006.

There was a mistake in the ms. We’ll change the paragraph about the spin up. We spun up over only a four year period (1979-1982) because the first year of forcing was not complete. After we did the spin up, we made a partly transient simulation with changing CO2 and land cover until 1978, while looping over the years of climate forcing we had available (1979-2006). Then, starting from 1979 we did a fully transient simulation with the right climate forcing, land cover map and CO2 value for each year of simulation.

L315: suggest to provide NPP comparison maps in the appendix. I think the evaluation of NPP and SOC here and also later in the result section can be largely shortened in the main text.

We believe that this section is useful regarding the main aim of our research which consists in understanding how leaching fluxes influence the European terrestrial C budget. However, it is right that the paper is long so part of the section (comparison at the basin/climate scale) will be moved to supplementary.

L350, where were these measured fluxes measured? In soil profile? If so, I would suggest to look at soil DOC concentration instead. I assume there are more soil DOC concentration available in Europe, e.g., Camino-Serrano et al., (2016).

Our main results focus on DOC leaching flux, and we thus wanted to compare our results with observations of leaching fluxes, which to our knowledge were only measured for the EU by Kindler et al. (2010) for both top and bottom soil. We also compared DOC concentrations profiles in the soil with the study from Camino et al., 2014. Indeed, there are many more studies on DOC concentrations in the soil but we selected the one by Camino et al., 2014 because it provides a synthesis at the pan-European scale, and is thus ideal to extract “representative” concentration profiles over a sufficient large domain, compatible to the regional scope of our study. The rationale for doing the comparison with Camino et al., 2014 will be added to the revised version of our ms.

L353, you have shown in Fig 2d about the mismatch of river network between the data used in the model and the real river network. But I might miss the description about dealing with this mismatch when you do one to one comparison with discharge or other data at point level. please clarify this.

At the point level, we selected the cell from the routing scheme that corresponds to the coordinates of the gauging station. Since ORCHIDEE is a coarse resolution model, the river network is also represented at a coarse resolution. When a mismatch occurred, due to the coarse resolution, the cell corresponding to the observation was shifted by one grid.

Figure 4, the model seems doing rather good job, although hydrological parameters were originally from Amazon basin? How do you think these hydrological parameters should work across different basins? It could be nice to elaborate a bit on this.
ORCHIDEE is a global land surface model. Most of the hydrological parameters actually have been calibrated against globally observed runoff (Ringeval et al., 2012; Yang et al., 2015), rather than only in Amazon basin. That is why the model performs well when it is applied to other basins.

L442, How about soil moisture comparison?

A comparison with observations on soil moisture will be included (see comment 2 of rev 2).

L446-448, "the underestimation of SOC stocks … " If I get this correctly, the underestimation of SOC stocks were checked at catchment level, but the comparison of DOC stocks are at a few points?, if so, it is difficult to put these two comparisons together to argue too high decomposition rates, right? Please justify this.

In this paragraph, the DOC stock refers to an average for all forests, and we observe a significant overestimation in the top soil compared to the observations. Since we diagnosed previously that SOC stocks were underestimated while DOC stocks appear here overestimated, one plausible explanation would be to assume too high SOC decomposition rates.

L461-464, on L350-353, the authors mentioned the difficulty/impossibility of comparing with these local measurements, but why it is then correct/possible to compare these 17 points with the modelled average leaching fluxes over Europe. It did not seem correct for me.

We agree that comparing model results to observations at the local scale, as shown in figure 6, is difficult (and indeed large discrepancies between observations and model results at point level can be diagnosed). This is why we choose to compare the 17-point average against the model-averaged result. But we agree that even in this case, the data-averaged value should be considered with care, taken the limited amount of observations. In the revised ms., we will insist on the limits of the model evaluation for the DOC leaching fluxes.

Fig. 6. There are low DOC leaching values along the coast, such as along Italy, UK, France? Why this pattern occurs?

We normalized by the area of the whole cell (even when the cell was not entirely covered by land). Thus when the land only accounts a small fraction (e.g. 25%) of the 1°*1° grid cell, then the DOC leaching from the 25% land is averaged to the whole grid cell. So that averaging can impact the results for “coast cells”. We will clarify this point in the revised ms.

L491-492 “labile and refractory DOC”, I did not recall from the method section there is a separation of different DOC pools. please clarify it in the method how the model deal with these two pools?

Indeed, this distinction should already have been included in the method section. DOC degradation in the river considers two different pools (labile and refractory), each one having a different decay constant. The methods section will be amended to include this description.

Section 3.1.6, a bit confused here: so all the comparisons earlier has included manure implementation, right? If so, I don’t think the authors need a separate section to describe the impacts, and it breaks the flow a bit.
Agreed. We will move this section on manure impacts to the supplementary.

Section 3.2, please justify the reason behind to look at the drivers of DOC leaching at climatic zones, instead of at catchments? And here I guess DOC leaching refer to leaching from soil? Please specify.

The grid and basin scales were the most relevant for the model evaluation. However, when searching for potential drivers of soil DOC leaching, temperature, runoff and drainage (driven by precipitation) were diagnosed and, therefore, a climatologic segmentation of the European domain can capture the spatial variability of these physical variables.

L586-587, Litter and SOC are the sources for DOC production, not NPP. NPP cannot be important control on DOC leaching fluxes either.

Yes, but Litter and SOC stocks are fed by NPP and, thus, indirectly DOC leaching depends on the terrestrial NPP. This indirect effect will be clarified in the revised ms.

L587-590, Did the normalization implemented at annual scale? It cannot use NPP to normalize DOC at monthly scale.

Yes, the normalization was implemented at the annual scale. This will be clarified in the revised ms.

L596-597, the higher leaching to NPP ratios could be also related to more organic soils and lower NPP.

More SOC or litter is actually also due to the low temperature. Thus lower temperatures lead to longer turnover times of organic C in the soil.

L605- 609, please consider to move most of these part of text into the method section. How did the authors end up choosing these three variables? Why not doing a sensitivity testing on the model first? What are the benefits of only using these three variables in comparing with the whole model? not clear why we need this section.

In our opinion, the predictive equation results from an analysis of the model results and is thus not already explained in the methods section which describes the model. Regarding the variables, it was found in the result section that DOC leaching is correlated to the total runoff (figure 11). Our hypothesis was that temperature could also affect the leaching through its control on the DOC production and decomposition rates.

The main purpose of this section was to highlight that once normalized to the terrestrial NPP, surface runoff, drainage (and their ratio) alone can explain most of the spatio-temporal variability in DOC leaching fluxes, temperature only playing a subordinate role.

L641-642, “our lower value may come from xxx”, I don’t think peatlands in UK and Scandinavia area the reasons why the model underestimated the whole region averages. Please elaborate on the uncertainties of the included processes.

We believe that the exclusion of peatlands is at least part of the explanation but the reviewer is right that uncertainties in the processes included or omitted could also explain some of the discrepancy. In the revised version, a new section on model shortcomings will be added that will reflect more broadly and
comprehensively on the various sources of uncertainties. The statement in lines 641-642 will be toned down.

Reference


References


ERA5-Land Hourly Data from 1981 to Present C3S ERA5-Land reanalysis (2019); https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview


Ringeval et al., 2012 Modelling sub-grid wetland in the ORCHIDEE global land surface model: evaluation against river discharges and remotely sensed data. Geosci. Model Dev. 5, 941-962, 2012
