This paper aims to reveal spatio-temporal patterns and drivers of DOC over Europe. It contributes to the current literature with main focus on the large-scale DOC variability. This paper has intensively compared the modelled values with different scales of observation data and tried to demonstrate the capability of the model in representing the European-scale soil and river DOC fluxes.

There are a few major concerns about this paper:

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.

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);
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.

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.

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.

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

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?

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

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

L158-159, sorption-desorption also contribute to the soil DOC dynamics.
L164-165, where did you get the diffusion coefficient and what is the value you used? It needs to be clear here.

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

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?

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.

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?

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.

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?

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.

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?
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.

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.

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.

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).

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.

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.

L442, How about soil moisture comparison?

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.

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.

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

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?

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.

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.

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

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

L696-597, the higher leaching to NPP ratios could be also related to more organic soils and lower NPP.

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.

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.
Reference


