Anonymous referee #2

Received and published 6 February 2020

The authors present a model development to simulate the concentrations, transformation and transfer fluxes of dissolved inorganic carbon (DIC), dissolved organic carbon(DOC) and terrestrial and autochthonous particulate organic carbon (POC) from head-waters to river mouth, and which is supposed to be useful for global simulations and future scenarios. I find the manuscript is poorly written, files in supporting documents are corrupted and results are not convincing, neither for model validation nor its potential global application. Therefore, I do not find that this work would merit publication.

A1: We agree we did not substantiate our global claim in this manuscript. See also reviewer 1, #1. We changed our claim and title in the revised version and provide further context in the discussion section. We selected one river, the Rhine, because of the large number of monitoring stations with ample data for model validation. For a global application, more rivers for model validation are needed. However, we do not agree that the results and model validation are not convincing. Although we see room for many improvements, we think these results are very encouraging. The framework opens new ways to investigate biogeochemistry in rivers on basin scale. Furthermore, as we expand upon in the discussion section of the revised manuscript, the potential for global application is certainly there as all used data sources are available on global scale.

Abstract: there is no reported value of globally estimated in-stream fluxes which is consistent with their claimed global model title.

A2: We agree. The paper is meant as a model description and it includes a proof of concept for one watershed. Actual application on the global scale is presented in other papers. We changed our claim and title in the revised version and provide further context in the discussion section as the potential for global application is certainly there since all used data sources are available on global scale.

Introduction: is very weakly written and mostly written after Cole et al 2007 paper. There is definitely a lack of discussion over the freshwater carbon cycle processes and need for further references. Furthermore, the authors need to discuss the existing models and any sufficiently striking advance over past publications.

A3: Processes are mentioned and referenced in section 40-46, existing models, references and their shortcomings are discussed in lines 48-60.

Throughout the text, almost all equations need to be revised and written more clearly with better abbreviations and definitions. The text is very unorganized.

A4: The equations are clearly ordered, sufficiently abbreviated and are accompanied with explanatory text. In the revision, an overview table of all equations used is found in section G in the supplementary material, in response to reviewer 1 (A20).

Results and validations: In general I do not find results convincing for the validation of model. For instance processing data from your supporting document csv for Figure6 shows all your model simulations for the latitude of 47-50 and longitude of 7-8 is overestimated. Keeping the simulations from only this lat-lon will lower your r2 to 0.18.This is the same trend in your TOC and DIC results, just the model has underestimated at all points. You need to better calibrate your parameters. Another issue is that the model is not capable of capturing any seasonality of measured data of any kind (e.g.Figure 7). As mentioned above, you need to calibrate your model parameters better.

A5: We find the results very encouraging. We agree there is substantial room for improvement here, but we believe that given the low resolution nature of the input data and hydrology and the complexity of river systems, the model simulates acceptable concentrations. Simulating the right order of magnitude in a coupled transport model, that resolves internal C processing at the same time is an achievement that deserves recognition. We do not calibrate our models, which does not lead to a better understanding, but instead, try to improve the model description of processes in case of disagreement of simulations with monitoring data. Transparency about this mismatch is a crucial start.

Regarding the seasonality, we refer to reviewer 1, R1. The model is able to capture seasonality (see figure 4). We agree that we do not observe this in our comparison with measurements.

- Source codes:
1) I tried to run the model but when I execute the commands for abiotic,respiration or biology I get the error: "ModuleNotFoundError' is not defined." Check the files you uploaded and try to include all necessary files for the model. I believe this is due to missing/not properly linked mocsy module.

A6: The README.txt accompanies the source code and shows the required installations of external modules. If mocsy is not installed it won't be able import the module. How to install modules in Python3.5 is described in the provided url's in README.txt. Also, given that you get the error '*ModuleNotFoundError is not defined*', I suspect you use Python2. DISC-CARBON is recommended to be ran in Python3.5.

2) I took a look at your sourcecode/carbon/code/reactions.py:

Thank you for taking a considerate look at the source code.

Firstly: What is your model time step? Considering Line126 in text I believed it is 1 month. But apparently in your reaction routine all you values are daily and you use 365 to convert to yearly. What is the case?

A7: All parameter values in mon_abio_default.ini / mon_resp_default.ini / mon_bio_default.ini are defined per day and need to be multiplied with 365

Also what is fT? temperature controlling effect on reactions? If so it needs to be explained more. What is sigma and temperature correction parameter?

A8: Description of this part appears to be missing in the manuscript. Added in the revised version.

- added: with a temperature response function *f(T)* as follows:

$$f(T) = e^{\frac{(T_{opt}-T)^2}{\sigma^2}} \quad (22b)$$

This response function is applied to respiration, excretion and mortality. Parameter values T_{opt} and σ are after (Garnier et al., 2000) and noted for *ALG* and ALG_{benth} in table 2.

Line 508-522 of code: why are your codes representing respiration, morality and excretion identical? Am I missing something here?

A9: The lines are identical, parameter values are different.

- Model exports (Netcdf files): As a test I just analysed the raw output data and took a look at DOC and DIC concentration randomly. Why do you have a negative value for DOC concentration in model? And why do you have a value of 7589.186 mg L-1 for DOC concentration? Why are there some very high values such as 37316.32 as DIC concentration? Why didn't you elaborate on the abnormal model results, as you are aiming to have a global version of your model and these extreme or negative values will affect your future possible global estimation? This shows there may be some errors in your model that you have not fixed and it may be giving nonsense outputs for some cases.

A7: We agree that the extremely high values are beyond the range being observed in the real world. However, these are not the result of errors in the model framework; they are a feature of the model. This is what happens when models data are combined; apparently an approach that is not valid for all gridcells. Concentrations are very sensitive to the magnitude of C delivery fluxes or water body volume, both of them are very prone to uncertainty. We do not aim to identify all these uncertainties and to discuss the results of individual gridcells. With this manuscript we show that we are able to simulate representative concentrations for most waterbodies with a process-based, coupled model. In the revised manuscript, we discuss the suspicious output.

- Table 1. Where did you find 1% fraction of DOC from total soil carbon ? I have never seen this fraction in any study. The DOC fraction is estimated between (20-40%)(Meybeck, 1982). Moreover, DOC transport from soil to river system is not only through surface runoff but also through subsurface runoff and drainage.

A8: We agree, the DOC fraction in SOC might be much higher than 1%; nobody knows. The study by Meybeck does not provide any insight whether this DOC originates from soil runoff or from in-stream production, so it is a irrelevant reference in this context. A better parameterization for DOC delivery to streams and rivers is available in a future version of DISC.

L41- Wrong reference. I could not find any statement regarding the terrestrial or aquatic sources

A9: reference is fine. Literally in this reference, "Autochthonous– Organic matter that was produced in the system of interest. Contrast it to allochthonous, which refers to organic matter that was produced elsewhere (for example in the watershed) and imported into the system of interest."

L42- Again, Wrong reference. Exactly which part of Cole et al, 2007 paper is referring to POC as a form of delivered C to surface waters?

A10: We agree that there is no explicit reference to POC from terrestrial sources. The used reference is incorrect for the statement. The reference belongs to the next statement. Corrected in the revised manuscript

L44- Give a reference for the mentioned processes

A11: The missing reference ended up wrongly after the previous statement. Corrected in the revised manuscript.

L50- Name some of these perturbations which are in line with your context?

A12: Added specific perturbations described in the mentioned studies.
Line 51
- added: such as reservoir building, eutrophication, land use change, climate change

L54- what do you mean by "informed projections"?

A13: informed projections are more than extrapolations on the basis of regression models. This is better specified in the revised manuscript.
Line 54-55
-added: There is a need for a modelling method that is beyond regression to make projections for the future.

L58-60- Indeed we need a model which includes all these missing compartments. But your model does not include any of these. Then what is the striking advance in your model compared to the existing models?

A14: DISC-CARBON does include all the mentioned perturbations; they are included in the hydrological and IMAGE model simulations that are used as inputs.

L60- You need to include in your introduction more detail on existing models and what the improvement is in your model. e.g. compared to RIVERSTRAHLER, ORCHILEAK,DLEM, etc.

A15: We think that the introduction sufficiently summarizes the current state of the science. The introduction is not a place to repeat other findings. Other globally applied watershed models are mentioned and their shortcomings are discussed in lines 53-59.

L64- What do you mean by designed for global application? You have neither tested your model at the global scale nor presented any global results. The only evaluation which I can find is based on the Rhine basin.

A16: We agree that the statement that it is designed for global is not merited with this manuscript.

title:
-from: CARBON-DISC 1.0 - A coupled, process-based model of global in-stream carbon biogeochemistry
-to: CARBON-DISC 1.0 - A spatio-temporal, process-based watershed model of in-stream carbon biogeochemistry

Line 23:
-from: considering that this is a global model
-to: considering the low spatial resolution of the model constraints

Line 64-65:
-from: This new model is specifically designed for global applications and describes the spatial and temporal variability of carbon concentration and fluxes based on the river basin hydrology from headwaters to mouth and carbon cycling processes.
-to: This new model describes the spatial and temporal variability of carbon concentration and fluxes based on the river basin hydrology from headwaters to mouth and carbon cycling processes.

Line 66-67:
-from: Here, we present the main features of the CARBON-DISC module, we apply the module in the stream network of the Rhine basin and evaluate its characteristics with a sensitivity analysis.
-to: Here, we present the main features of the CARBON-DISC module, and as a proof of concept, we apply the module in the stream network of the Rhine basin and evaluate its characteristics with a sensitivity analysis.

L72- what do you mean by the period 1900-2000? Models can simulate C, N and P depending on the forcing data period.

A17: We mean that the forcing data period is available from 1900 to 2000. We made this clearer in the revised manuscript.
line 72-75:
- from: The IMAGE-DGNM model framework integrates the PCRaster Global Waterbalance (PCR-GLOBWB) dynamic global hydrology model (Sutanudjaja et al., 2018) with the IMAGE model (Stehfest et al., 2014) that provides data for C, nitrogen (N) and phosphorus (P) delivery to inland waters (streams, rivers, lakes, reservoirs, floodplains) for the period 1900-2000.
- to: The IMAGE-DGNM model framework integrates the PCRaster Global Waterbalance (PCR-GLOBWB) dynamic global hydrology model (Sutanudjaja et al., 2018) with the IMAGE model (Stehfest et al., 2014) that provides forcing data for the DISC module for C, nitrogen (N) and phosphorus (P) delivery to inland waters (streams, rivers, lakes, reservoirs, floodplains) for the period 1900-2000.

L73- biogeochemistry *processes

A18: corrected as suggested in revised manuscript in line 75

L76- Only air temperature is taken from CRU?

A19: correct

L77- What is the reason for focusing on monthly scale?

A20: It is the highest temporal resolution available for forcing data. Made this clearer in the revised manuscript.
Line 79-81
Added: Minimum timestep is limited by hydrological constraints and biogeochemical data. In the presented study, hydrological constraints and biogeochemical data provide monthly data; hence, results are presented in a monthly timestep. The framework however, is able to process any timestep.

L79-81: These line are not clear. Need a revision. I do not see the link between IMAGE and soil organic carbon in Fig.1.

A21: These lines provide a general, clear and accurate description of data that is being used in the DISC module. Before diving in all the details, we describe the general types of data being used.

L81- Add more information of the hydrology parameterization.

A22: Repetition of methodologies that are already presented by other studies doesn't have added value. Also, the manuscript is first and foremost a description of the C components in the aquatic biogeochemistry module DISC.

L85- Add these fluxes to arrows from IMAGE and SOC to CARBON-DISC box on Fig.1.

A23: these fluxes are depicted in figure 3, which is a more detailed representation of delivery and interaction fluxes of individual C forms. The objective of figure 1 is to provide the context of the DISC module and to provide general information on how other models/frameworks are involved.

Figure1. Confusing. You need to revise this figure. Make it consistent with the explanation from L79 to 87.

A24: In the revised text, we link the figure description more clearly to the figure. We applied same words in the text as in the figure in the revised manuscript. Also, we made the link with table 1 to appropriately describe the C forcing data.

L90- The whole definition of IMAGE and IMAGE-DGNM is very confusing and not clear. After several times back and forth I could understand the model structure. You need to revise this whole section. Moreover, the SPM is missing on the arrow from SOC to surface runoff. All the particle forms that you are discussing from L90 to 96 should be also in Fig.1 to make it easier to understand.

A25: figure 1 clearly shows the relation between IMAGE and IMAGE-DGNM. We intentionally did not add individual C forms to the figure as to keep the general overview more general. Individual species interactions are shown in figure 3.

L99- You missed mentioning the temperature and radiation when explaining PCR-GLOBWB.

A26: radiation is not part of the PCR-GLOBWB model. Temperature is an external and properly referenced dataset, which is used in PCR-GLOBWB and also consistently applied in the DISC framework. Temperature is referenced at line 28. Solar radiation is an additionally developed module described in lines 320 to 355.

L99- Change solar irradiance to radiation to be consistent with Fig.1

A27: changed in the revised manuscript

L100-104- poorly written. needs revision.

A28: comments are too unspecific to make change this part of the manuscript.

L104- Which equations?

A29: added the equations 17 and 18 at line 109 in revised manuscript.

L106- missing "," after environments

A30: added in the revised manuscript

L116- What is the rationale to compare the monthly simulations with bi-weekly measurements?

A31: It is for both datasets (observations & simulations) the highest resolution available

Figure 2. caption: add what the yellow circles are

A32: added in the caption

L125-129- Needs revision. You need to change the abbreviations for C. It is confusing as it is now. For instance, you can change the Ci to Ctot, Chyd and so on.

A33: We do not agree with the suggested change. Equation 1 represents the central governing differential equation describing the change of total C. Total C is generally altered by either biogeochemical processes or by hydrological transport. Also, all abbreviations are described right next to the equation.

L130- I do not see the description of bgc.

A34: bgc is described in line 128 as the biogeochemical component of changed amounts

L135- This has to come before the description of total C species definitions.

A35: We do not agree with this suggestion. Equation 2 is one of the differential equations that is part of equation 1. We choose to start with the general equation and then dive deeper in the details.

L137-138- Unclear. Is this referring to eq.1? Which hydrological model data?

A36: this is referring to PCR-GLOBWB, the used hydrological model as was mentioned in the first line of the method section.

L155- Here and for all the equations after this one definitions should be much shorter and better written. Use proper abbreviations.

A37: Although we agree that the equations take a lot of space, we do not agree with this suggestion. Shorter written equations won't make this clearer. We don't see the need for abbreviations.

L210- As an example I am pointing this out. You are bringing SEDOC here (eq11) and then I need to scroll down to eq21a and b to understand it. Very poorly written and needs better flow through the text. Bring the explanation of processes involved in each equation to appear right after it. Until further revision of the methods, I skip the rest of the method and equations.

A38: We strongly disagree with the sentiment in this comment. The structure to describe the differential equations is intentional. We first present the general differential equation for the form of C and then describe each component of the differential equation. We believe this is clear and consistently written and rather than poorly written.

L331- What do you mean by scheme?

A39: We understand that use of the word 'scheme' is confusing. We mean numerical scheme (variables and their interactions). The numerical scheme applied is explained in very general terms in lines 93-101 and in more details in section 2.2, the model description. The schematic representation of the scheme is found in figure 3.

L344- What was the rationale to do 750 model spin-ups? Was that enough to reach equilibrium? The presentation of your model mass balance is missing. You need to add it as well.

A40: For the sensitivity analysis, we've run the model 750 times, using the same 1 spin-up for all these runs. The carbon budget is presented in figure 4.

L335- Need further explanation of the rationale for these two set-ups.

A41: a more extensive rationale for these two set-ups is added to the revised manuscript in lines 161 to 171.

L340- I did not find any "scheme" in your supplementary material files.

A42: This line ended up unintentionally in the manuscript. the representations of the abiotic and respiration scheme ended up not being included in the supplementary information. They are added in the revised supplementary information in section H.