We greatly thank all the reviewers for their thorough and helpful comments which contribute to improving the manuscript.

Please find our point-by-point responses (in blue) in the following, and changes in the manuscript are *in italic* here.

RC3: Anonymous Referee #3

The manuscript describes a model that simulates transport and biogeochemistry in various estuaries in France. The 1-D grid is oriented along the flow path, which is constrained at the marine side by results from an ocean model and at the upstream river side by time-series data. The calibration of the model parameterization is based on data from 1 year and then validated based on the ability to reproduce data from the previous and following year. For most locations the model captures the trend of various biogeochemical parameters well, despite major simplifications, which include 1) a 1-D grid that cannot account for the depth-dependency of biogeochemical processes and cannot fully resolve the spatial variability in residence times; and 2) the biogeochemistry in sediments is ignored (it only accounts for organic matter burial).

RC3.1: My major criticism is that the role of biogeochemical processes is not well teased out. The reaction rates that the model explicitly resolves are not shown in any figure and the paper would benefit from a more rigorous comparison of the simulated rates with those reported in literature.

AC3.1: We understand the comment of the reviewer, which was also expressed by the other reviewers (see answers **AC1.3, AC2.1, AC2.6**). While we carefully referenced the work of Volta et al. in which an extensive description of the model is provided, we acknowledge we did not provide enough information in the present manuscript for the reader to get a good feel for the processes that are and are not included in the model nor how they are formulated. We added in the supplementary material the complete list of the processes (Table S-1) and their formulations as well as the equation controlling the production and consumption of all the state variables included in the model. We believe will help the reader interpret our results with full knowledge of the inner workings of our biogeochemical module.

RC3.2: It is unclear to what extent the trends in the modeled output are driven by biogeochemical processes and to what extent by the constantly changing boundary conditions. The influence of the boundary conditions could be especially large for sites where the location for model calibration/validation is near the inlet or outlet (such as in smaller estuaries). The reported retention rates are often low, meaning that the influx and outflux are nearly equal. This could indicate that the respective chemicals do not react much (making model results more trivial), but there could also be a dynamic balance between sinks and sources. The authors may want to improve the manuscript by describing in more detail the cycling of elements within the model domain. A sensitivity analysis would be useful to show to which biogeochemical rate constants the model results are most sensitive. Additionally, a simulation with the biogeochemical reactions turned off would be informative. Without these analyses it is hard to assess if the model captures the biogeochemical dynamics adequately.

AC3.2: We understand the reviewer's concern regarding the ability of our simulations to accurately calculate the retention within estuaries because of the potential influence of boundary conditions on our results. The numerical scheme of C-GEM and its short integration time step (300 s) allows for accurate calculations of the lateral fluxes of all state variables between each grid point. It would like to stress out that, while the temporal resolution of the model's output is 4 hours, the calculation of the cumulative transport fluxes are not obtained by the post-processing of those result files but are updated every 300 seconds in order to ensure mass conservation (which was checked prior to performing our simulations). We are thus confident that the exchange at both boundaries, which we use to calculate our nutrients and carbon budgets are accurate even if the retentions we obtain are only of the order of a few percent in estuaries with short residence times. We would also like to stress out that the length of the simulations (3 years, which roughly equals 2000 tidal cycles) prevent transient effects associated with tidally induced changes in concentrations to skew the nutrients and carbon budgets.

Finally, in the updated version of the manuscript, following a request from reviewer 1 (see **AC1.18**), we calculated the retentions rates for nutrients and carbon during the dry season (i.e. Figure 7 and associated text) for the different estuaries investigated. These retentions calculated over a shorter period with lower discharges yield significantly larger retentions and thus illustrate the ability of the model to account for the effect of biogeochemical processing on nutrients and carbon budgets within the systems simulated.

RC3.3: In larger estuaries the retention of nutrients is higher, which indicates that biogeochemical processes play a more important role. At these locations early diagenetic processes could also potentially have a larger effect on the water quality. The model contains denitrification in the water-column, but as the water remains oxygenated, it is probably unimportant. Benthic denitrification is likely more important, but not modeled/parameterized. Sediments could also act as a source of nitrogen by releasing DIN derived from remineralized organic matter. The model does neither account for benthic PO4 dynamics nor benthic O2 consumption. The manuscript does not make a compelling argument for why these benthic processes can be ignored.

AC3.3: A detailed answer regarding our justification for the lack of a benthic module in the current version of C-GEM is also provided in answer **AC1.18**. We agree that our current set-up does not entirely represent the nutrient dynamics in the location where diagenetic processes may play a large role in N and P cycling. In addition to the justifications provided in answers AC1.17, **AC2.2, AC2.11 &12**, we would also like to mention that, because we performed our calibration, water column processes in C-GEM actually implicitly also account for benthic ones in our calibration. We agree that this is not ideal nor mechanistically accurate but we believe this also limits (for denitrification for instance) the drawback of not explicitly representing benthic processes. Indeed, such a task would require, for the calibration of the new benthic module, additional data such as benthic remineralization rates, denitrification rates that are not necessarily all available in the systems we investigated. Moreover, the addition of a full diagenetic module at each grid cell of our model would increase its calculation time by one order of magnitude and require a very long spin-up to generate initial conditions for the benthic species, which would be difficult to compare with the very scare measurements available. There exist simpler benthic modules of lower complexity, which would limit the computation cost of adding an explicit

representation of benthic processes (see Soetaert et al., 2005 for example) but those nonetheless require more data than we currently have to be applied with a satisfying level of confidence. While we believe this is the way forward in the long run, we would be afraid of adding more uncertainty than actual mechanistic understanding to the model by including too many processes that we cannot properly constrain or evaluate. This rationale to justify the current level of complexity of our biogeochemical model is now further discussed in section 4.1 of the updated manuscript.

"*In its current setup, the biogeochemical module of C-GEM considers some of the most essential biogeochemical processes and reactions (i.e. primary production, organic matter degradation, denitrification…). In spite of generally good ability of the model to capture the main spatial and temporal biogeochemical dynamics of the different systems studied (i.e. longitudinal, seasonal and amplitude of the variations of nutrients carbon and oxygen fields), several potentially important processes contributing to the N and P cycling in estuarine environments in particular are still ignored or largely simplified. These include benthic-pelagic exchanges, sorption–desorption of phosphorus, mineral precipitation or a more complex representation of the biological planktonic/benthic compartments (such as grazing by higher trophic levels, or multiple reactive organic carbon pools for instance). This limits the depth of mechanistic understanding the model can provide of nutrient cycling, particularly regarding interactions between pelagic and benthic compartments which can significantly influence the intensity but also the timing of nutrient and organic matter cycling in estuaries (Laruelle et al., 2009). The addition of a full diagenetic module at each grid cell of our model would be possible but would also increase its calculation time by one order of magnitude and require a very long spin-up to generate initial conditions for the benthic species. There exist simpler benthic modules of lower complexity, which would limit the computation cost of adding an explicit representation of benthic processes (see Soetaert et al., 2005) but those would nonetheless significantly increase the data demand of the model to be properly calibrated. Thus, while we believe the inclusion of an explicit benthic compartment to our model is the way forward on the long run, such an increase in complexity without sufficient data for a proper calibration and evaluation might introduce more uncertainty than actual mechanistic understanding to the model. In the present study, a simple representation of particulate matter burial was nonetheless implemented and applied to phytoplankton and TOC to provide a first-order representation of the process, which is necessary to evaluate the retention of carbon and nutrients within the system. We believe this addition, coupled with denitrification provides a first insight on the main pathways removing nutrients from estuaries.*"

Please see also **AC1.17**, **AC2.2**, **AC2.11** and **AC2.12, AC3.9** for new inputs on N,P cycling and TP:TN ratio.

Overall I find the manuscript well written, but it could be further improved. The results section contains many interpretations, which do not belong in this section.

Specific Comments:

RC3.4: Line 240-246: Sinks and sources in the model are not only related to the inlet and the outlet, but also to exchanges with the sediment and atmosphere. It is specified for Flux_out that it refers only to the outlet. For Flux_in it is stated that it accounts for all inputs. Does this refer only to the river inlet or also inputs from the atmosphere?

AC3.4: In our simulations, Flux_in only refer to the river inlet and tributaries. Considering the relatively limited surface areas of the studied estuaries, we do not account for atmospheric deposition of N or P, so the only exchange with atmosphere left would be denitrification, which is considered as part of the retention within the system.

RC3.5: The calculated residence time of water is only based on river discharge, but does not account for inputs from the ocean. For the theoretical case that the river influx goes to 0, the water residence time would approach infinity. Obviously, the real residence time of water also depends on the exchange with the ocean and the reported residence times are less meaningful near the ocean boundary. It may be good to point this out.

AC3.5: Thanks for this suggestion. We made it clearer by calling it "*fresh water residence time"* in the paper.

RC3.6: The estuaries are described to be macrotidal, but I believe the text does not describe the vertical stratification. Perhaps the vertical stratification can be ignored in these macrotidal estuaries. Readers may appreciate a better characterization of the flow in these estuaries to have a better idea of the implications of the 1-D approach.

AC3.6: We understand the reviewer's concern and acknowledge that a 1-D model such as C-GEM cannot resolve vertical stratification. However, in macro-tidal estuaries, it can generally be assumed that the water column is well mixed vertically considering the scale of the tidal energy dissipated over relatively shallow water columns. This is reflected in the work by Savenije (2012) for example and we introduced a sentence attesting the overall mixed nature of e.g. the Seine estuary at each tide cycle (Brion et al., 2000) and other of the Atlantic Coast of Europe (Middelburg and Herman, 2007).

"A 1-D model can be considered as well adapted to these shallow macrotidal estuaries that mixed at each tide cycle as shown by Brion et al. (2000) on the Seine River and Middelburg and Herman (2007) for other estuaries of the Atlantic Coast of Europe, which was also supported by Savenije (2012)."

RC3.7: Line 410-412: The regulation of outflow by the dam is an interesting point. Is this regulation explicitly modeled or averaged over time?

AC3.7: This regulation by the dam at the entrance of the Vilaine estuary is not explicitly modelled. The gauging station is located upstream of the dam, so that the model does not take into account the water regulation. Despite this weakness, except SPM, the biogeochemical variables are in a good order of magnitude.

RC3.8: Line 451-453: The text states that the grid resolution (2 km) may be too coarse for small estuaries (< 30) km. If the model grid contains less than 15 grid cells, simulations can probably be run quickly. Why did the authors not run simulations with a finer grid resolution?

AC3.8: We understand the reviewer's remark. However, our initial aim was to investigate several estuaries with different physical and biogeochemical properties with an identical set-up, both in terms of hydrodynamics (temporal and spatial resolution included) and biogeochemistry (identical reaction rate constants). We thus did not set a finer grid resolution for the small estuaries because we wanted to inter-compare the model performances. We however agree with the reviewer that modifying the grid could be done in future work to improve the model's performances in small systems.

RC3.9: Line 457-460: "In the present… denitrification for nitrogen". The text ignores that nutrients are released during remineralization and can be transported back into the overlying water. The statement that denitrification is "the only other potential nutrient removal term" missing in the model is wrong. There are other benthic processes that can act as a sink for nitrogen, sediments can sequester PO4, also Si can be taken up and released by various benthic processes, and there are many other nutrients (e.g. iron) not accounted for in the model. Also, the model does not account at all for the uptake of O2 by the sediment. These benthic processes are often parameterized in ocean circulation models. The benthic remineralization rate could be estimated based on the current burial flux and then be used to approximate the exchange fluxes of O2, P and N. The authors should either implement these processes or explain why these fluxes can be ignored.

AC3.9: This is right and we agree that the current level of complexity of our biogeochemical module ignores potentially important processes contributing to the transformations of Si, P and N in estuaries. We do acknowledge these limitations in the text and tried in its updated version to make these statements clearer. We do for instance agree that our statement regarding the burial being the only benthic process able to remove N from the system, other than denitrification is an oversimplification that we paid attention to correct in the updated manuscript.

We are indeed convinced that C-GEM should explicitly include a benthic representation of N, P and Si remineralization as well as O2 uptake. While it is true that such fluxes are included in some models (in particular ocean circulation models as pointed out by the reviewer) and that there exist estimates of some of the reaction rates controlling those fluxes, implementing the latter in our estuarine model would be a massive undertaking which added value remains unclear considering the numerous uncertainties associated with these processes in estuarine settings. Moreover, the implementation of an explicit benthic module may require adding a diagenetic model at each grid cell of the model, which is a massive undertaking. Such a task would require numerous data that do not exist for all the systems we investigated, would largely increase the uncertainty of our results if each new process is not sufficiently constrained, and would likely increase the calculation time by one order of magnitude.

Luckily, there exist alternative solutions with benthic modules of lower complexity and we are working towards the integration of such type of benthic module but those still require ample data collection, and calibration before we can be confident that it would introduce more insight than uncertainty to our results. At current, we believe the current level of complexity of C-GEM already provides valuable insight through this collection of fully transient multi-annual simulations over 7 estuaries. We believe our work is a necessary step towards the implementation of more complex models in the future. We tried to update our manuscript to both make the scope

of our study clearer and better justify the current level of complexity of our model in our simulations.

Please see also our answer **AC3.3**.

RC3.10: The fits for the Charente estuary are not so good. Remarkably the trends in simulated PO4 and DO values are opposite to those in the measurements in the calibrated year. Also NH4 concentrations are generally too low. After reading the manuscript it is not clear to me why the model could not reproduce these trends. If the model fails to reproduce the data, should the results be presented in the manuscript?

AC3.10: It is true that the model didn't catch the peak and bottom values for the Charente at this station during summer. In order to provide a better feel for the model performance, we added a figure of the seasonal variations for the stations closest to the marine boundary within the salinity gradient in the supplementary material (Figure S-3). It can be observed that, in the cross-section in Figure S-3, PO4 and DO values were better simulated than at the station presented in the main text although we acknowledge that NH4 was still too low.

In spite of the fact that the performances of the model are not as convincing in the Charente as they are in other systems, we believe it was nonetheless relevant to include them in our manuscript for two main reasons:

1-The performance of our model is the Charente, although not fully accurate, still falls within the range of values of the observations (see Figure 4) for most of the variables. Ammonium is indeed not well simulated by the model but its level is low compared to nitrate, which is well represented. We also acknowledge that the strong depletion in oxygen is not reproduced by the model. Overall it seems that the model can nonetheless provide information on the main biogeochemical characteristics of an estuary, even when data are scarce.

2-We would like to stress out the results of our simulations are generated by the application of the same biogeochemical parametrization to all systems. Rather than calibrating our model to each system individually, which would certainly yield better results but would essentially consist of developing 7 individual models. Instead, we wanted to evaluate how much insight can be provided by the application of a generic model and parametrization at the regional scale as a first step towards the application of such modeling approach to all the estuaries of a continuous stretch of coast, regardless of their data coverage. We thus believe that the arguably 'mediocre' performance of our model in the Charente and the ability of our simulations to capture the overall dynamics of 7 different systems with a single set-up is representative of the level of confidence we could have in results obtained for estuaries totally devoid of data. In this regard, we believe it is important to include the results of all our simulated systems and that this level of accuracy is very encouraging in the perspective of future regional simulations.

RC3.11: Many statements in the results section contain an opinion or assessment:

AC3.11: Thanks for the suggestions. Several statements were reconsidered.

Line 305: "These comparisons… available observations"

AC2.11a: The sentence has been rewritten and now reads:

"The model outputs were compared to available observations in order to provide bias and RMSE, which are reported in Table 7. These statistical indicators reflect overall good performances of the model through the standards provided by Moriasi et al. (2007, −0.7 < Bias < 0.7)."

Line 335: "seasonal trends are properly captured…"

AC2.11b: The sentence has been rewritten and now reads:

"*In addition, the seasonal trends generated by the model for nutrients, DO, and Chl-a concentrations generally follow the variations reported by field measurements both in timing and amplitude."*

Line 342-344: "DO showed… high summer mineralization": Please point to results that back up this statement about the relative importance of the effect of solubility and mineralization. Figure 4 only shows peaks and troughs.

Line 344: "Conversely, phytoplankton… time scale". Please, just describe the model output.

AC2.11c: A sentence has been added, and now the paragraph is as followed:

*"DO showed a regular trend with high values in winter and low values in summer, according to its solubility, but also to its consumption with high summer mineralization. Indeed, whereas the Charente River showed DO observed values of about 11 mg L^{-1} in winter (~100% saturation) much lower DO values down to 3-4 mg O_2 L^{-1} were found in summer (i.e. ~35% saturation, well illustrating a high summer O_2 consumption (**Error! Reference source not found.**). Indeed, phytoplankton biomass (Chl a) simulations showed a shift in relation to the observations (with a right level, however), which led to short-term summer DO peaks of the model which did not fit the observations. Noteworthy is the excessively scarce phytoplankton data which cannot support the modelled pattern at such a time scale."*

Line 354-360: "Considering the suitable agreement… retention rates." The first sentence is an assessment and the remainder of the paragraph are methods.

AC2.11d: Many thanks for this remark, the following sentence has been removed from the results section and appears in method one**:** dissolved inorganic nitrogen (DIN=NH_4+NO_3), dissolved inorganic phosphorus (DIP=PO_4), dissolved silica (DSi), Phy (Chl-a, using a C/Chl-a ratio of 40 (Jakobsen and Markager, 2016)), and TOC were calculated (**Error! Reference source not found.**). In addition, the Redfield–Brzezinkski ratio C:N:P:Si = 106:16:1:15–20 (Brzezinski, 1985; Redfield et al., 1963) was used to take into account the organic fractions and estimate total nitrogen (TN), total phosphorus (TP), and total silica (TSi), whose fluxes were preferentially chosen for calculating overall retention rates (see section **Error! Reference source not found.**).

Line 412-414: "Most of the… retention itself"

AC2.11f: This sentence has also been removed from the results section: most of the retention may therefore occur in the dam reservoir, thus favoring nutrient uptake by phytoplankton and loss of N via denitrification (Garnier et al., 1999; Seitzinger et al., 2006; Yan et al., 2021) and in total reducing estuarine retention itself.

There are other examples where the tone is not neutral or the text is not limited to results. The authors could improve the text by revising the results section.

Several sentences have been shortened in the result section to ensure neutrality.

Minor comments:

RC3.12: Line 155: "Numerical Schemes". This paragraph mentions already existing models that have been used and describes the conceptual model, but not so much the numerical implementation. Therefore, consider changing the heading.

AC3.12: We followed the suggestion of the reviewer and updated the title of this section to "Model Description and Set-up".

RC3.13: Line 173 (Figure 2b): What process does the arrow from "Denitrification" to "Aerobic Degradation" represent?

AC3.13: This arrow did not correspond to any actual flux but was the result of a mistake in the design of the conceptual scheme. The figure was corrected and the new Figure 2b now accurately represents the different fluxes and state variables of the model.

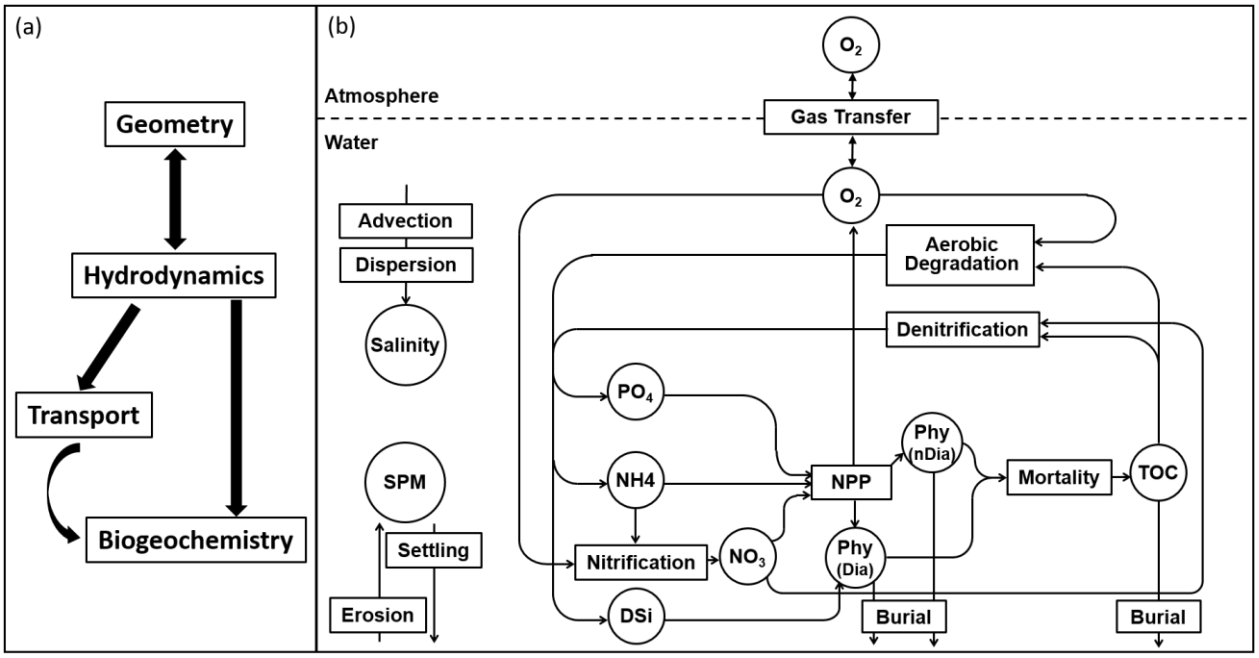

Figure 1 (a) The C-GEM concept. (b) Conceptual scheme of the biogeochemical module used in C-GEM in this study: a circle represents the state variables while a rectangle represents the processes; Dia corresponds to diatoms.

RC3.14: Line 204: “spin-up” instead of “warm-up”

AC3.14: This correction was implemented.

RC3.15: Line 204-205: “repeating the… steady-state conditions.” Here I got lost.

AC3.15: The sentence has been rewritten as follows:

“The simulation starts following a 60-day spin-up during which only the hydrodynamics and transport modules are resolved over a repeating identical tidal cycle using, which enables the system to reach a dynamic steady-state, providing realistic initial conditions for the biogeochemical module.”

In a nutshell, the 60 days model spin-up is performed using only the hydrodynamics and transport module to generate realistic longitudinal profiles for all state variables before the beginning of the actual simulation. These concentration profiles provide the biogeochemical module with realistic initial conditions constrained by the physics of the system.

RC3.16: Line 209: “Calibration was implemented based on 2015, as average…” Consider rephrasing.

AC3.16: The sentence was changed into:

“Calibration was implemented based on 2015, as it is representative of average hydrological conditions in France in recent years.”

RC3.17: Line 337-338: “showed a seasonal decrease from winter to autumn”… probably from autumn to winter is meant.

AC3.17: Yes indeed, we apologize for the mistake. The correction was made.