First, we would like to thank referee 2 for his careful evaluation of our manuscript and his interesting comments which we believe will help us to improve it. Please, find hereafter our response to these comments.

1) Firstly, I would say that it does not fit the scope of GMD, which is there to present new developments in models. While the companion paper, with its presentation of a mixotroph compartment, meets this criterium, the main new thing in this manuscript is a diagnostic relation between TA and salinity, which may improve the results, but conceptually is a fairly small step and has been used in many different models so far. From this side I would rather recommend publication in a different journal, where the focus is more on the considered system itself, i.e. a model for the BoM.

We understand your concern, however we believe that this manuscript has its place in GMD. Eco3M_MIX-CarbOx is a new model which has been developed to consider both mixotrophs and carbonate system. We decided to present Eco3M_MIX-CarbOx in two parts to show both sides of the model distinctly. It allowed us to propose clear studies, easier to read than one which would have been longer, and at the same time to highlight the two main developments which, together, constitute the originality of our model. However, it is important to keep in mind that both parts of the study aim to present this new model. We believe that the study, as a whole (both parts), fits well the scope of GMD and especially, meet the criteria of 'model description papers'.

Moreover, both studies are strongly linked. In the first part, we focused on the planktonic ecosystem description, especially on mixotrophs. We detailed their implementation in the model and study their dynamics in the area. In the second part, we focused on the carbonate system which is barely mentioned in the first part and detailed its representation in the model. This second study is based on the first one as we present a representation of carbonate system by considering the impact of mixotrophs (photosynthesis, respiration…) on these variables. We think that the strong connection between both studies also justify their publication in the same journal.

2) The switch to a salinity-TA relationship is motivated by the desire to represent the episodic intrusion of freshwater from the nearby Rhone into the BoM, and also the influence of evaporation and precipitation. My first question here is: If these freshwater fluxes affect the TA balance so strongly, should they not also influence DIC?

First, we would like to stress the fact that "an excess of alkalinity" which likely reflects alkalinity inputs to coastal areas has been described for the entire Mediterranean Sea (Schneider et al., 2007). This study, at the global scale, has forged our conviction that, in a coastal area close to the Rhone River, alkalinity inputs from Rhone River needed to be considered even in this 0D configuration.

The switch to a salinity-TA relationship is possible thanks to the fact that in the bay of Marseille, TA variations are mainly the results of rivers contributions, particularly the Rhône River one. It was demonstrated by the lack of variations observed when we modelled TA only based on biogeochemical processes which take place in the box. For DIC, a different reasoning must be adopted, mainly because the processes which impact DIC dynamics are very different than the one which impact TA dynamic. As we consider a surface layer, DIC dynamics is mostly the results of temperature and salinity changes (which are considered by the model) and biogeochemical processes (especially air-sea CO_2 exchanges) (Hassoun et al., 2015). The Rhône River can bring DIC to the BoM, these inputs are diluted (far from 2877 µmol kg^{-1}, the value observed in the Rhône River, Table S1) and, due to the action of other processes (solubility effects and biogeochemical processes) on DIC dynamics, which is more pronounced in this case, have a less significant impact than on TA dynamics.

Moreover, it is important to note that, in Eco3M_MIX-CarbOx, TA is the main driver of carbonate system. In other words, a change in TA results in significant changes in DIC, pH_T and pCO_2 as demonstrated by the Figure 4 of the manuscript. By representing the contribution of the Rhône River on TA we then indirectly apply it to the three other variables of the carbonate system. We are aware that the contribution of the Rhône River to the DIC considered in that way (through TA) is only an indirect effect. It is important to highlight that in the case of our 0D configuration, it is not always possible to consider allochthonous contributions. So far, we experimented two ways to consider external contributions in Eco3M_MIX-CarbOx, which gave satisfactory results:

- By using a VAR=f(S) relation (done in this paper for TA),
- By forcing the variable with a file which include an interpolation of the measurements performed at SOLEMIO for this variable (done in our first paper for nutrient concentrations)

Using a DIC=f(S) relation did not produce the expected results. In fact, it is not recommended to use such a formulation to represent DIC dynamics, especially in surface water due to the numerous processes from which its dynamic results (Hassoun et al., 2015). Using an interpolation of SOLEMIO DIC measurements force us to represent the variable dynamics as a forcing, which had no real interest here given that the representation of DIC obtained was already rather correct when based on biogeochemical processes.

Table S1: Salinity-DIC couples for LSE events measured at SOLEMIO between 6 June 2016 and 26 June 2019 (last data available).

Salinity	DIC (µmol.kg^{-1})
37.11	2321.3
37.78	2280.1
37.30	2259.3
36.82	2323.8
37.62	2288.7
37.18	2260.3
37.66	2269.8
37.32	2249.0

and nutrients as well?

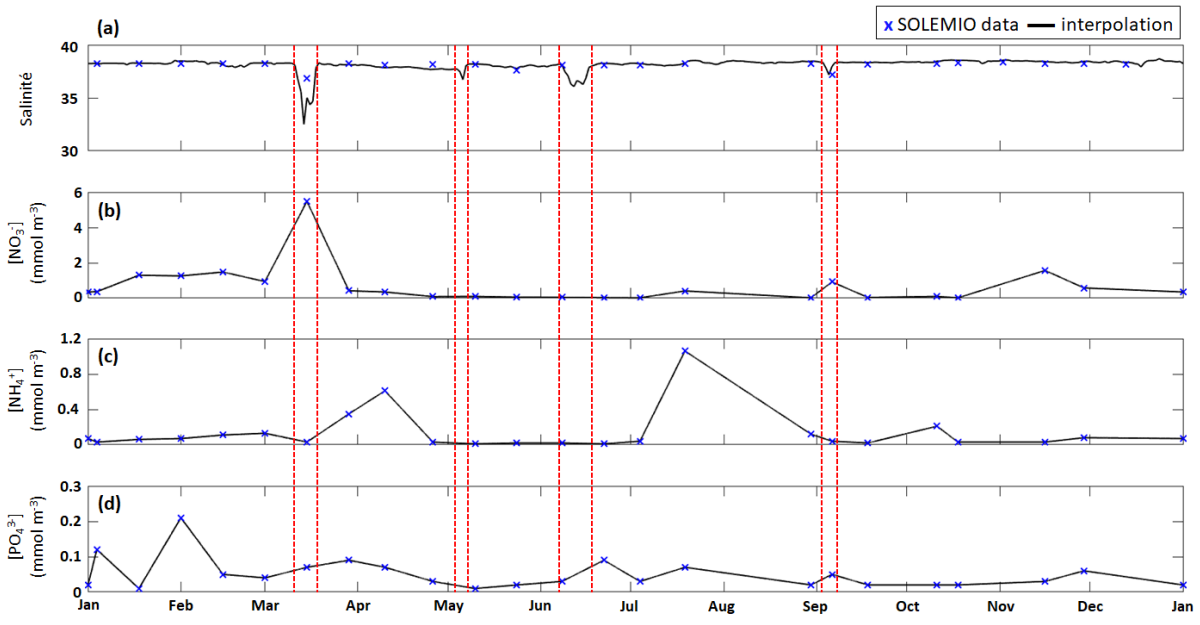

Figure S1: Time series of surface (a) salinity (CARRY measurements), and interpolated (b) NO_3^- concentration, (c) NH_4^+ concentration and (d) PO_4^{3-} concentration at SOLEMIO station. SOLEMIO data are represented by blue markers and the four LSE are indicated by the red dotted lines.

Rhône River intrusion events are associated with an increase of nutrient concentrations in the area, especially nitrate and phosphate (Fraysse et al., 2014). However, in our case, the four low salinity events are not systematically associated with a nutrient increase at the station. In fact, only the first and last events (15 March and 5 September respectively) have an impact on nutrient concentrations at SOLEMIO with the first event being the most significant (Fig. S1). This pattern can be explained by the salinity data used by the model. The measurements are performed at CARRY station (near the Côte Bleue, see figure 1 of the manuscript for location) which is more significantly impacted by the Rhône River plume as it is closer to the river mouth than SOLEMIO station. In consequence, decreases of salinity measured at CARRY are not systematically observed or can be less significant at SOLEMIO. Salinity measurements are also performed at SOLEMIO, however their temporal resolution is low (fortnightly measurements) compared to the CARRY one (hourly measurements). In fact, Rhône River intrusion events duration is variable and can be less than 15 days (ex: short-lived intrusions, Fraysse et al., 2014), therefor it is important to consider the highest temporal resolution possible to better catch them with measurements which is why we chose to work with CARRY measurements instead of SOLEMIO measurements (Fig. S1a).

That this may lead to biases is discussed in lines 477 to 484; but given the extremely high DIC concentration in Rhone water quoted on line 482, I wonder whether this inconsistency may not invalidate the main results.

We understand your concerns, however, as indicated in the previous point, the strong DIC value (2877 µmol.kg-1) observed in the Rhône River never reaches SOLEMIO as the Rhône River plume is quickly diluted and the DIC values which really reaches SOLEMIO station are rarely higher than 2300 µmol.kg-1 (mean value of 2281.5 µmol.kg-1, Table S1).

Moreover, as indicated in the previous point, it is important to note that in Eco3M_MIX-CarbOx, TA is the main driver of carbonate system. In other words, a change in TA results in significant changes in DIC, pH and pCO2 as demonstrated by the Figure 4 of the manuscript. By representing the contribution of the Rhône River on TA we then indirectly apply it to the three other variables of the carbonate system.

3) This leads me to a more conceptual difficulty with the approach. The model concept is that of an arbitrary one cubic metre volume at the surface of the bay, and that the model just represents fluxes within this volume. Spatial fluxes are excluded (except for CO2 flux, more on that below). This only allows either to model a variable as purely forced from what is happening inside the box, or to prescribe it, e.g. as a function of salinity. For a proper modelling of how external fluxes (e.g. in mol/s) change concentrations (mol/m^3/s) inside the modelled region, one would have to define the volume that is affected by these fluxes. A reasonable choice might be to model a column of water within the mixed layer, as was done in many zero-dimensional models, e.g. Fasham et al, 1990 or Hurtt and Armstrong 1999. That would allow a consistent treatment of the effects of mixing on TA, DIC, nutrients.

This difficulty becomes especially clear when the authors discuss the possible reasons for their low net annual air-sea flux of CO2, which is in contrast to observation-based estimates. Here they state that "aeration is is simulated by applying Eq (5) to 1 m^3 of surface water at the SOLEMIO station, which tends to overestimate the effect of aeration processes on DIC..." (line 534 ff). Indeed: if the control volume is that shallow, it will be lead to a too fast approach of DIC towards equilibrium, and hence an underestimate of fluxes.

We understand your concern. We think that it is important to note that, in this study, we relied on Eco3M-CarbOx (Lajaunie-Salla et al., 2021) for the calculation of carbonate variables. So, Eco3M-CarbOx was our starting point to implement carbonate system variables in Eco3M_MIX-CarbOx. With our study, we aim to bring answers to the concerns raised by this previous study, we then use the same concept and try to improve the representation of carbonate cycle variables (by adding mixotrophs organisms processes to the state equations and switching TA formulation by a newly implemented allochthonous formulation). In this way, we were able to compare both models, then knowing how our modifications impact the carbonate system variables representation. Even if we did not manage to obtain a realistic representation of air-sea CO_2 fluxes, we provide some improvements to the initial concept and give some examples of suitable and unsuitable use of it (first part and second part of this study respectively), then confirming that the only way to obtain realistic fluxes is to consider a larger layer. We could have done it the way you propose but, considering that it required a complete review of our 0D configuration, we decided to focus directly on the coupling of Eco3M_MIX-CarbOx, in 3D (which is still in test phase).

4) And finally, while the diagnostic TA leads to an improvement in model results, as evidenced by decreases in %BIAS, RMSD and a 'cost function' presented in Table 3, the improvements are fairly modest. Indeed I would be interested in knowing whether a model with TA prescribed constant at the average of observations would not have fared at least similarly good as the two presented model cases.

As suggested, we run a simulation with a constant TA (mean of SOLEMIO measurements for 2017 = 2591.2 µmol kg^{-1}). In Table S2, we presented the calculation of the statistical indicators presented in the manuscript. As we decided to modify them, based on your suggestions, we also provided average absolute error (AAE), and average error (AE) calculated as in Stow et al. (2009) except that, to be consistent with calculations of statistical indicators used previously (Allen et al., 2007) the difference is applied between observations and model which means that for %BIAS and AE, if a positive value is obtained the model underestimates the observations.

Table S2: Statistical indicators calculation for the simulation with a constant TA (TA = 2591.2 µmol kg^{-1}). Mean, SD, AE, AAE and RMSD are in the same unit than the considered variable, i.e.: µmol kg^{-1} for TA and DIC and µatm for pCO_2. CF and %BIAS are without unit.

		TA	**DIC**	***p*CO₂**	**pH$_T$**
N	Observation	20	20	20	20
Mean ± SD	Observation	2591.2 ± 19.4	2294.9 ± 24.0	391.0 ± 31.0	8.09 ± 0.03
Mean ± SD	Model	2591.2 ± 0.22	2305.7 ± 26.1	418.0 ± 28.9	8.07 ± 0.03
CF	Model	0.85	0.82	1.14	1.14
%BIAS	Model	-0.002	-0.50	-5.79	0.26
RMSD	Model	18.90	26.14	38.45	0.03
AAE	Model	16.5	19.7	35.5	0.03
AE	Model	-0.06	-11.5	-22.6	0.02

We also represented the daily mean values of TA, DIC, pH_T and pCO_2 for the simulations SIMC0, SIMC1 and constant TA (Fig. S2) to compare the three simulations carbonate system variables representation.

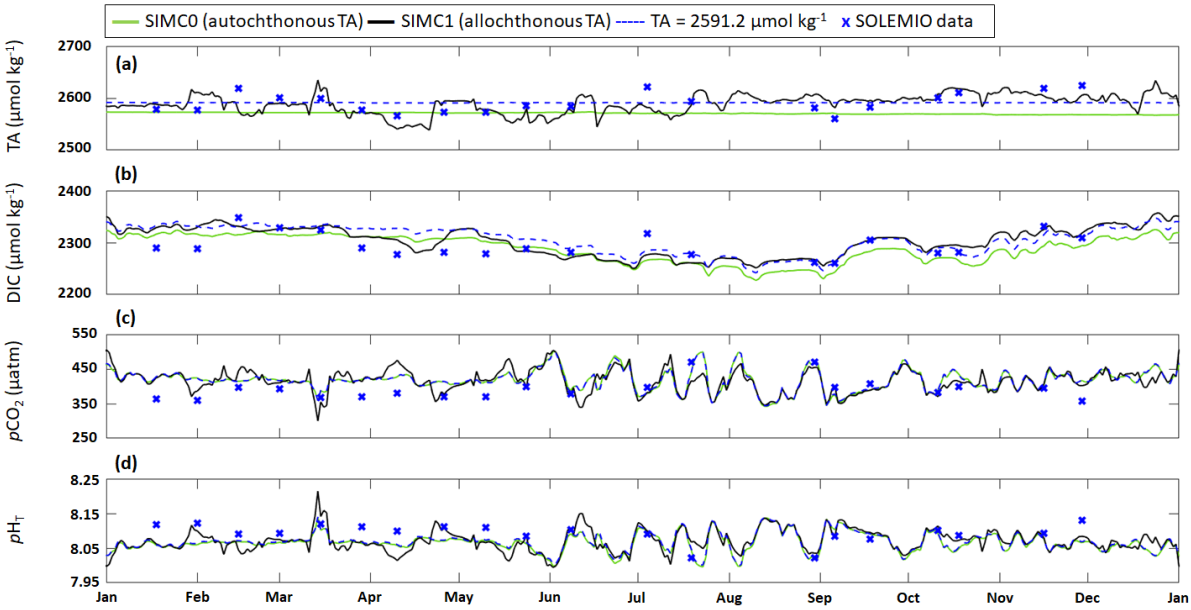

Figure S2: Comparison of model outputs from SIMC0 (autochthonous formulation, Table 2 of the manuscript), SIMC1 (allochthonous formulation, Table 2 of the manuscript), and constant TA simulation, model runs showing daily average (a) TA, (b) DIC, (c) pCO_2 and, (d) pH_T for 2017. SOLEMIO data are represented by blue markers.

We are aware that improvements seem fairly modest, especially when studying the statistical indicators (Table S2). However, values are generally slightly better for the simulation SIMC1, especially for the three other carbonate system variables. In fact, the major improvement bring by the switch to an allochthonous formulation for TA, is that TA variations are represented which seems more realistic and tend to improve the representation of the other three carbonate system variables.

Major comments :

Figure 1 is identical to the one on the companion paper and definitively isn't needed should this paper be published in GMD.

Done. We replaced the map by the following sentence: **[A map of the study area showing the location of stations where measurements were carried, and places of interest can be found in Barré et al. (2023a).]** and added at the end of the 2.1.

State equation for TA, Eq. (1): The terms in the equation are not properly defined. The definition of the terms is given in the Appendix (Table A1), but the table is not referenced here.

Thank you for pointing this out, we added it l.144: [where i represents the number of organisms. **Processes description can be found in Table A1 (Appendix A) and formulations are available in Barré et al. (2023a).** In this formulation, TA only depends on biogeochemical processes (i.e., TA riverine inputs are excluded).] and l.170: [where i represents the number of organisms. **Processes description can be found in Table A1 (Appendix A) and formulations are available in Barré et al. (2023a).** As an additional modification, we use a more recent version of the gas transfer velocity calculation introduced by Wanninkhof (2014).].

The two linear S-TA relations, presented on page 7, which are valid below and above a salinity threshold of 37.8 are discontinuous at S=37.8. This should lead to sudden jumps in the TA value if this threshold is crossed. Are there any effects of this discontinuity visible in the results?

Thank you for this interesting comment. Indeed, using the allochthonous formulation leads to sudden jump in TA when the thresholds is crossed. These sudden jumps are also observed in pH_T and pCO_2.

When we implemented the TA allochthonous formulation, we used two points: a first one which represents the Rhône River water at the river mouth (S = 0, TA = 2885 µmol kg^{-1}) and a second one which represents the Rhône River water which reach SOLEMIO during a LSE (S = 36.82, TA = 2600.6 µmol kg^{-1}). We chose the second point as it was the most significative LSE on the period covered by the SOLEMIO measurements (2017 and 26 June 2019). We then consider it as representative of the Rhône River water which reach the BoM. Even though TA values associated with LSE are variable and highly depend on the period of the year (Fig. S3). TA values equal or above 2600 µmol kg^{-1} do not seem the most representative (Table S3) and LSE seem associated with TA highly dependent on Rhône River seasonality (mean value = 2575 µmol kg^{-1}).

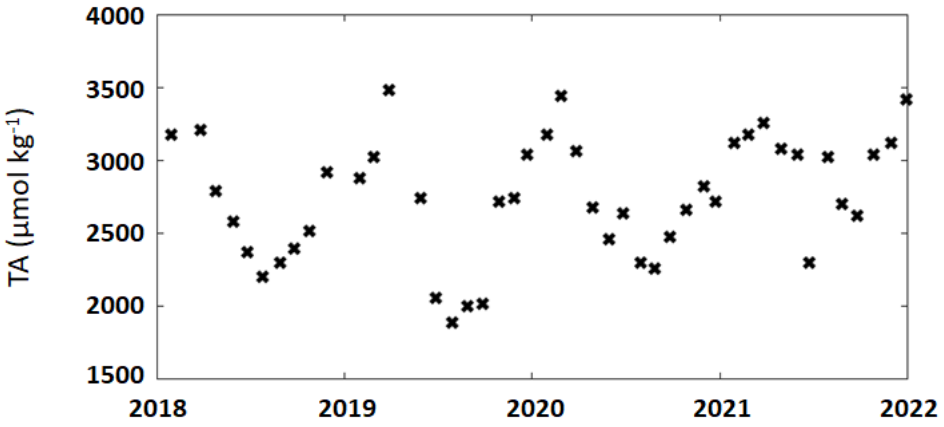

Figure S3: TA measurements in the Rhône River (data: Naïades, https://naiades.eaufrance.fr, first data available: January 2018).

Table S3: Salinity-TA couples for LSE events measured at SOLEMIO between 6 June 2016 and 26 June 2019 (last data available).

Salinity	TA (µmol kg-1)
37.11	2603.0
37.78	2579.6
37.30	2585.5
36.82	2600.6
37.62	2585.8
37.18	2560.8
37.66	2568.4
37.32	2520.7

We think that it might be interesting to improve our allochthonous formulation to better manage the threshold crossing case. To avoid (or at least reduce these instabilities) it could be interesting to take into account Rhône River seasonality in the allochthonous formulation and we plan to do this in a future work.

Page 8, line 175: In principle the model equations would not change had you chosen to assume the effected layer to be deeper, except that then the flux would then be distributed over a larger volume. Why not take at least H as the annual average mixed layer depth> Taking it as 1m is equivalent to speeding up the gas exchange by a factor H_real, the real affected layer.

As suggested, we ran a simulation with a modified function for aeration process (Eq. S1, SIMR1 in the following). We considered a mean annual value of 30.5 m for mixed layer depth (mean of winter value

= 41 m and summer value = 20 m (Wimart-Rousseau et al., 2020)). Daily average of modelled carbonate system variables and air-sea CO_2 fluxes are represented in figure S4.

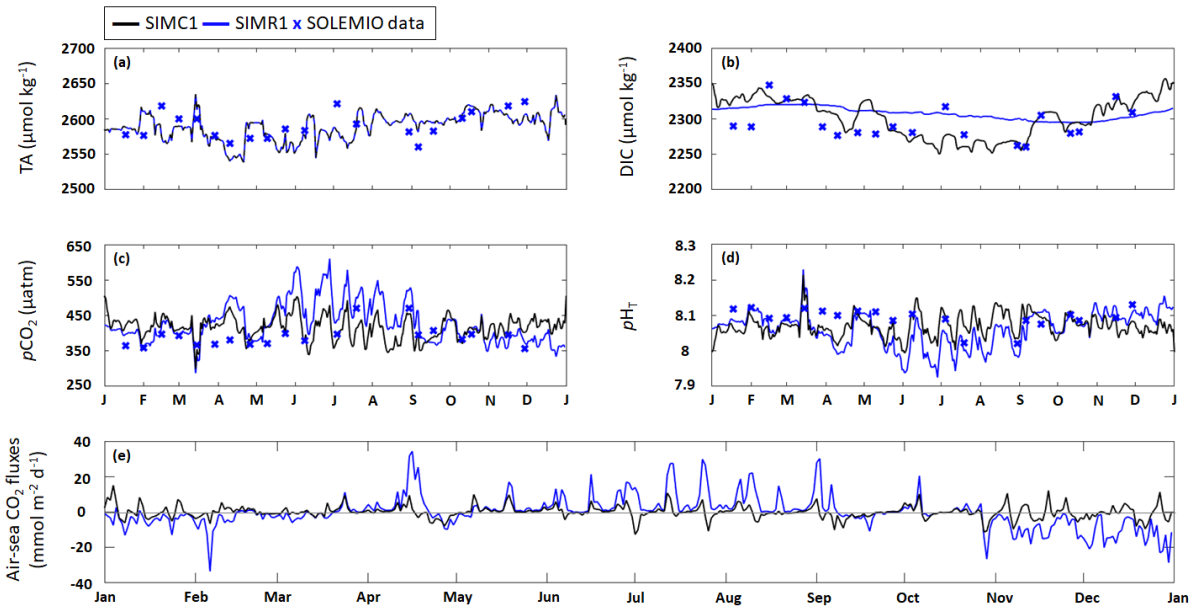

Figure S4: Comparison of model outputs from SIMC1 (aeration process apply on a 1 m layer, Table 2 of the manuscript) and SIMR1 (aeration process apply on 30.5 m layer, model runs showing daily average (a) TA, (b) DIC, (c) pCO_2, (d) pH_T, and Air-sea CO_2 fluxes for 2017. SOLEMIO data are represented by blue markers.

We obtained a mean annual value for air-sea CO2 fluxes of -113.6 mmol m^{-2} yr^{-1} which is better than the one obtained previously (-0.21 mmol m^{-2} yr^{-1}) but still lower than the value suggested by Wimart-Rousseau et al. (2020) (-803 mmol m^{-2} yr^{-1}). These results are interesting as we can see that, by considering a larger layer, we better represent the seasonality of air-sea CO_2 fluxes (sink in winter and source in summer) (Fig. 4e) which is mainly explained by the fact that we are able to represent the undersaturation observed for pCO_2 in winter, especially at the end of the year (Fig. S4c). In this simulation, we used a constant MLD, however, we believe that using a variable MLD (deeper in winter than in summer) could emphasize this result. These results also show that by considering a larger layer to apply the aeration fluxes, we significantly modify DIC representation (Fig. S4b). Thus, the seasonality well modelled previously, is no longer visible. DIC values are also much less variable, then far from the dynamics described by observations.

To conclude, we believe that more than modifying the thickness of the layer impacted by aeration process, we need to move to a 3D configuration to better represent air-sea CO_2 fluxes without impacting the representation of DIC which is already, rather correct.

Page 8 and Appendix B, pH and pCO2 calculation: It is good to see that the pH scale differences are taken properly into account, and fugacity has been calculated correctly. But much of this is fairly standard, e.g. the iterative calculation of pH, described in Figure B1. This could be left away.

We considered your comment, however, as we provide some corrections to Lajaunie-Salla et al. (2021) Appendix A, we think that it is necessary to keep Appendix B in the manuscript.

Page 9, Figure 3: The quality of the Figure is awful. But also it does not convey much information, I would leave it away.

We considered your suggestion. However, we think that figure 3 allows to better visualize the model calculation steps and how these can be modified by the choice of TA formulation. We decided to modified figure 3 to improve its quality:

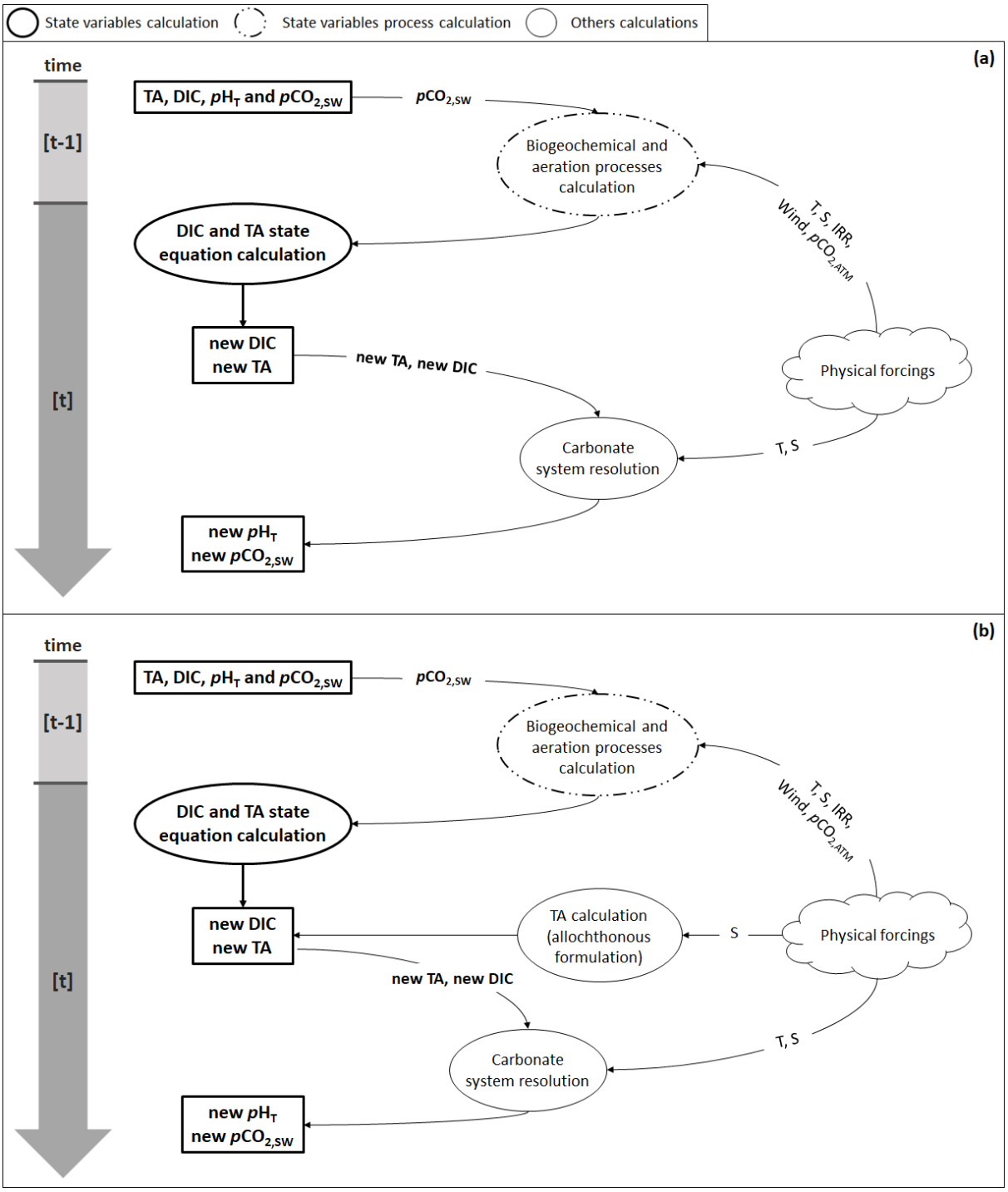

Figure 3: Flow diagram illustrating the steps needed to calculate pHT and pCO_2 (a) using the autochthonous formulation (Eq. 1) and (b) with the allochthonous formulation (Eq. 2 and 3). Physical forcings include temperature (T), salinity (S), solar irradiance (IRR), wind speed (Wind) and atmospheric pCO_2 ($pCO_{2,ATM}$).

We hope it is clearer this way.

Page 10, lines 218-220: It is not clear to me how the salinity-normalized nTA and nDIC are exactly defined, by a linear correlation with salinity with zweo intercept? If so, why do that if the observed S-TA relation in the oceanographic region is different?

Salinity-normalised changes in nTA and nDIC were calculated by dividing by in situ salinity and multiplying by mean salinity.

To clarify, we added (l.219): [Though, we isolate the changes of TA and DIC due to variations in freshwater inputs using the salinity normalised TA (nTA) and DIC (nDIC) **which are obtained by dividing the considered variable by in situ salinity and by multiplying the result by mean in situ salinity,** and adding another term to regroup them.]

page 11, definition of the statistical indicators: while the definition of RSMD and %BIAS is rather clear, that of the cost function is less clear: Typically, a cost function aggregates model-data-disaggreement for different variables, possibly with different units, into a single scalar variable (Stow et al, 2009). But what exactly the variables are that enter the CF, and how the different variables are nondimenionalized and aggregated into one CF should be properly defined.

We based our CF calculation on Allen et al. (2007). In this work, CF is defined as: [The cost function gives a non-dimensional value which is indicative of the "goodness of fit" between two sets of data; it quantifies the difference between model results and measurement data (see OSPAR Commission, 1998). It is a measure of ratio of the model data misfit to a measure of the variance of the data; the closer the value is to zero the better the model.], and is calculated as follow:

$$CF = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{|O_i - M_i|}{\sigma_O}\right)$$

Where O represents the observations, M the model results and σO is the standard deviation of the observations.

page 11, interpretation of statistical indicators: Whether a CF<1 is considered very good, would probably depend on the definition of CF, and cannot be stated as generally as on line 255-256. If the individual cost function terms e.g. consist of the squared model-data difference scaled by the variance in the individual variables, and are then added together, the expected height of the CF would depend on how many different variables are finally added together. Also, I don't think one can generally say (line 252-253) that a %BIAS<10% is excellent; I would think that depends on the ratio of natural variability to the mean of the variable in question. For TA, with a high background value, a 10% BIAS is rather large.

To interpret CF and %BIAS, we used the interpretation of Radach & Moll (2006) and Marechal (2004) respectively. We understand that these interpretations seem not strict enough. Moreover, CF indicator seems rather insensitive as almost all variables for the tested simulations show CF value lower than 1. To improve our statistical analysis of model results, we proposed to consider two other indicators: AE and AAE, to replace CF. We based their calculations on Stow et al. (2009) except that, to be consistent with the calculation of RMSD used previously (Allen et al., 2007), the difference is applied between observations and model which means that for AE, if a positive value is obtained the model underestimates the observations. Moreover, we added the formulation of each statistical indicators used to avoid confusion.

We modified the 2.4 accordingly:

(l.246): [We used three statistical indicators for the comparison between simulation and SOLEMIO data: the percent bias (%BIAS), the cost function (CF) and the root mean square deviation (RMSD). These indicators were used with two Eco3M_MIX-CarbOx simulations (SIMC0 and SIMC1) and the reference Eco3M-CarbOx simulation (Lajaunie-Salla et al., 2021). %BIAS is calculated according to Allen et al. (2007) and allows to quantify the model's tendency to under- or overestimate the observations. In our case, a positive %BIAS means that the model underestimated the in situ observations and vice versa. %BIAS is interpreted according to Marechal (2004). We use the absolute values of %BIAS, to assess the overall agreement between the model results and observations. The agreement is considered: excellent if %BIAS < 10 %, very good if 10 % ≤ %BIAS < 20 %, good if 20 % ≤ %BIAS < 40 % and poor otherwise. The cost function is calculated based on Allen et al. (2007). It is a dimensionless indicator that quantifies the goodness of fit between the model and observations. According to Radach and Moll (2006), CF < 1 is considered very good, 1 ≤ CF < 2 is good, 2 ≤ CF < 3 is reasonable, while CF ≥ 3 is poor. RMSD quantifies the difference between model results and observations (Allen et al., 2007). The closer RMSD is to 0, the more reliable the model. All statistical indicators are calculated using surface SOLEMIO data from 2017. The model data is averaged using the mean of the output from the date in question ± five days. Using temporal mean and standard deviation of model results allowed us to better account of variability at SOLEMIO station. By comparing the statistical indicators obtained for SIMC0, SIMC1 and Eco3M-CarbOx we also obtained an indication of how changes in the carbonate formulation affected the results.]

to: [**We used four statistical indicators for the comparison between simulation and SOLEMIO data: the percentage bias (%BIAS), the average error (AE), the average absolute error (AAE) and the root mean square deviation (RMSD, also refer as root mean square error in the literature - RMSE). They were used with two Eco3M_MIX-CarbOx simulations (SIMC0 and SIMC1) and the reference Eco3M-CarbOx simulation (Lajaunie-Salla et al., 2021). The %BIAS is calculated as follow:**

$$\%\mathbf{BIAS} = \frac{\sum_{i=1}^{N}(O_i - M_i)}{\sum_{i=1}^{N} O_i} * 100 \quad \textbf{(16)}$$

where O represents the observations and M the model results (Allen et al., 2007). This indicator allows to quantify the model's tendency to under- or overestimate the observations. The closer the value is to 0, the better the model. Here, a positive %BIAS means that the model underestimated the in situ observations and vice versa. On an indicative basis, the %BIAS can be interpreted according to Marechal (2004): Absolute values of %BIAS allow to assess the overall agreement between the model results and observations and the agreement is considered: excellent if %BIAS < 10 %, very good if 10 % ≤ %BIAS < 20 %, good if 20 % ≤ %BIAS < 40 % and poor otherwise.

We based our calculation of AE, AAE and RMSD on Stow et al. (2009). Together, these three statistical indicators provide an indication of model prediction accuracy.

$$\mathbf{AE} = \frac{\sum_{i=1}^{N}(O_i - M_i)}{n} \quad \textbf{(17)}$$

$$\mathbf{AAE} = \frac{\sum_{i=1}^{N}(|O_i - M_i|)}{n} \quad \textbf{(18)}$$

$$\mathbf{RMSD} = \sqrt{\frac{\sum_{i=1}^{N}(O_i - M_i)^2}{N}}$$

(19)

The three of them aim to measure the size of the discrepancies between model results and observations, the closer the value is to 0, the better the agreement between model results and observations. However, when interpreting AE, it is important to note that value near zero can be misleading because negative and positive discrepancies can cancel each other. That is why it is important to calculate, in addition to AE, AAE and RMSD which allow to overcome this effect (Stow et al., 2009). Such as %BIAS, a positive value of AE means that the model underestimated the in situ observations and vice versa.

The model data is averaged using the mean of the output from the date in question ± five days. Using temporal mean and standard deviation of model results allowed us to better account of variability at SOLEMIO station. By comparing the statistical indicators obtained for SIMC0, SIMC1 and Eco3M-CarbOx we also obtained an indication of how changes in the carbonate formulation affected the results.]

We changed the Table 3 (Page 14):

[**Table 3: Comparing the different model results to surface observations at SOLEMIO station for TA, DIC, seawater pCO_2, and pH_T. N represents the number of observations. Mean, SD, AE, AAE and RMSD are in the same unit than the considered variable, i.e.: µmol kg^{-1} for TA and DIC and µatm for pCO_2. %BIAS is without unit.**

		TA	DIC	pCO_2	pH_T
N	Observations	20	20	20	20
Mean ± SD	Observations	2591.2 ± 19.4	2294.9 ± 24.0	391.0 ± 31.0	8.09 ± 0.030
Mean ± SD	SIMC0	2576.1 ± 1.5	2293.6 ± 25.1	413.5 ± 16.5	8.07 ± 0.015
	SIMC1	2588.6 ± 16.4	2301.1 ± 24.5	409.1 ± 21.4	8.07 ± 0.020
	CarbOx	2574.5 ± 3.6	2292.5 ± 26.0	413.9 ± 15.9	8.07 ± 0.010
%BIAS	SIMC0	0.58	0.05	-5.75	0.29
	SIMC1	0.09	-0.27	-4.61	0.21
	CarbOx	0.64	0.1	-5.86	0.29
AE	SIMC0	15.12	1.25	-22.5	0.02
	SIMC1	2.57	-6.2	-18.02	0.02
	CarbOx	16.7	2.4	-22.9	0.02
AAE	SIMC0	18.7	20.4	35.9	0.03
	SIMC1	16.3	17.2	34.7	0.03
	CarbOx	20.1	21.2	35.3	0.03
RMSD	SIMC0	24.90	24.26	38.75	0.04
	SIMC1	20.03	21.83	40.27	0.04
	CarbOx	26.56	24.90	38.29	0.04

]

Considering a suggestion from referee 1, we also added appendix C, the calculation of statistical indicator for H^+ concentration (Kwiatkowski & Orr, 2018).

[**Appendix C: Statistic indicators calculation for H^+ concentration**

Table C1: Comparing the different model results to surface observations at SOLEMIO station for H^+ concentration. N represents the number of observations. Mean, SD, AE, AAE and RMSD are in the same unit than the considered variable, i.e.: mmol m^{-3} for H^+ concentrations. % BIAS is without unit.

		[H+]
N	Observations	20
Mean ± SD	Observations	$8.08 \times 10^{-9} \pm 5.52 \times 10^{-10}$
Mean ± SD	SIMC0	$8.89 \times 10^{-9} \pm 2.91 \times 10^{-10}$
	SIMC1	$8.39 \times 10^{-9} \pm 4.06 \times 10^{-10}$
	CarbOx	$8.52\ 10^{-9}\ \ 2.80\ 10^{-10}$
%BIAS	SIMC0	-5.33
	SIMC1	-3.91
	CarbOx	-5.47
AE	SIMC0	-4.30×10^{-10}
	SIMC1	-3.15×10^{-10}
	CarbOx	$-4.42\ 10^{-10}$
AAE	SIMC0	6.45×10^{-10}
	SIMC1	6.05×10^{-10}
	CarbOx	$6.36\ 10^{-10}$
RMSD	SIMC0	6.98×10^{-10}
	SIMC1	7.14×10^{-10}
	CarbOx	$6.93\ 10^{-10}$

]

and modified the results part (3.1) accordingly (l.301):

[Regarding the coast function, simulations yielded CF < 2 for all variables which is considered very good (CF < 1) or good (1 ≤ CF < 2) (Table 3). The %BIAS parameter yielded "excellent" results for all variables (using the interpretation form Marechal, 2004, i.e., %BIAS < 10 %). The highest values for %BIAS (in absolute terms) were obtained for pCO2 with ~6 % while the remaining variables had values < 1 %. Similarly, pCO2 had the highest RMSD which suggests that this parameter is not as well represented in the model as the other variables. Furthermore, SIMC1 produced the best TA representation yielding the lowest values for CF, %BIAS and RMSD (Table 3). Moreover, SIMC1 produced an annual mean-TA that was closest to the observations. While the SIMC0 and Eco3m-CarbOx results are fairly similar. SIMC0 produced a slightly better representation of TA compared to Eco3m-CarbOx. Similar conclusions can be drawn for pHT where SIMC1 also outperformed SIMC0 based on CF and %BIAS (Table 3). For studying DIC and pCO2, the situation is less clear as the simulations performed differently for different indicators, making it difficult to pick a clear winner. Still SIMC1 shows the best CF and RMSD values for DIC, and the best CF and %BIAS for pCO2. In conclusion, SIMC1 shows the best overall indicator values for the examined variables (more specifically, it outperformed the other simulations in 9 of 12 indicator comparisons).]

to: [**For statistical indicators, %BIAS values are systematically lower than 10 %, with the highest values obtained for pCO_2 with ~6 % while the remaining variables had values < 1 %. Similarly, pCO_2 had the highest RMSD, AAE and AE which suggests that this parameter is not as well represented in the model as the other variables. Furthermore, SIMC1 produced the best TA representation resulting in the lowest values for %BIAS, AE, AAE and RMSD (Table 3). Moreover, SIMC1 produced an annual mean-TA that was closest to the observations. While the SIMC0 and Eco3m-CarbOx results are fairly similar. SIMC0 produced a slightly better representation of TA compared to Eco3m-CarbOx (%BIAS, AE, AAE and RMMSD slightly lower). For pH_T, SIMC1 outperformed SIMC0 based on %BIAS (Table 3), however, AE, AAE and RMSD values are similar for the three simulations. We then performed the calculation**

of statistical indicators on H+ concentration as, according to some authors (Kwiatkowski & Orr, 2018), comparing H^+ concentrations is a better practice than comparing pH. Results are available in Appendix C. Based on Table C1, SIMC1 also outperformed SIMC0 based on AE and AAE. For studying DIC and pCO_2, the situation is less clear as the simulations performed differently for different indicators, making it difficult to pick a clear winner. Still SIMC1 shows the best AAE and RMSD values for DIC, and the best %BIAS, AE, and AAE for pCO_2. In conclusion, SIMC1 shows the best overall indicator values for the examined variables (more specifically, it outperformed the other simulations in 13 of 20 indicator comparisons when including H+ concentrations comparison).]

Table 3, Page 14: If the variance of the observed TA and DIC values is on the order of 20 micromol/kg (note, units should be given in the table), then I'd say a RMSD of about the same order of magnitude is not an excellent agreement. It is not terrible either, though. A similar remark holds for %BIAS.

We hope that modifications mentioned above allow to clarify this point. We indicated units in the caption of Table 3.

Figure 4, page 12: The time-series of the difference between the model runs (right panel) does not convey much new information, I would remove them.

We represented the differences between simulations as we think that it allows to better visualize them and then emphasize the fact that modifying TA formulation yields different model outputs for DIC, pCO_2 and pH_T. Considering that, we decided to keep the figure 4 as is.

Also, I have a question to the data (crosses in Figure 4): to me it is not clear whether all four carbon system variables were measured independently, or whether e.g. DIC and TA were measured, and pH and pCO2 calculated from them. If they were measured independently, how consistent are they with respect top each other, given the used set of carbon system equations?

DIC and TA are measured, and we calculate pH_T and pCO_2 by using CO2SYSv3 (Sharp et al., 2020, originally developed by Lewis and Wallas (1998)) on MATLAB. The set of constants used is the same than the one used to perform the calculation of pH_T and pCO_2 in the model.

Figure 5, page 15: The subpanels on the right are simply a cutout of the panels on the left for the summer period. What is the purpose of this duplicated information?

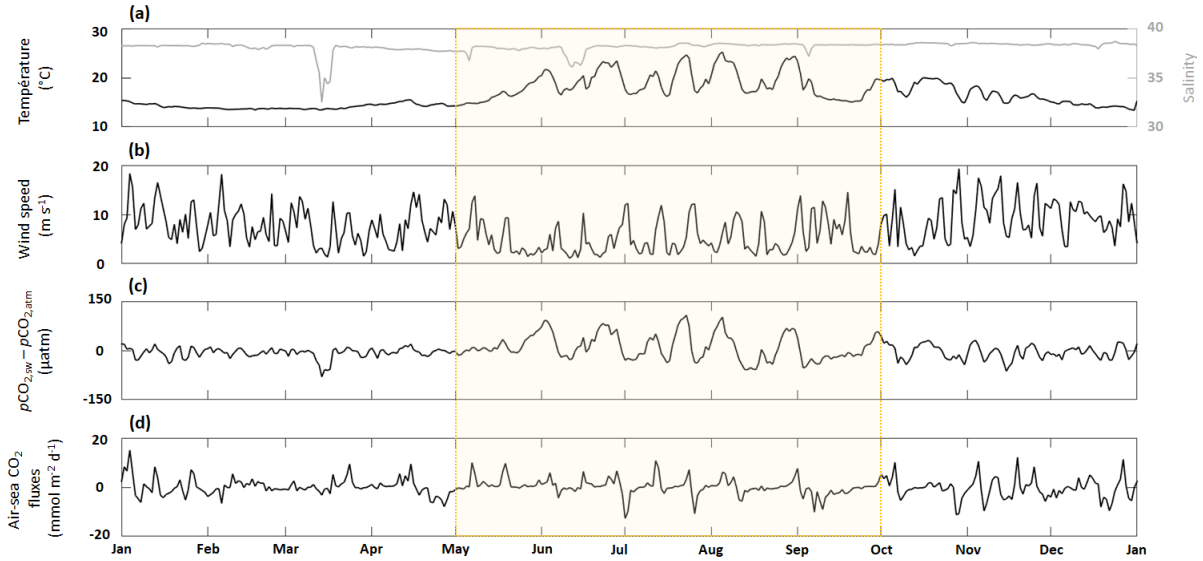

Figure 5: Time series of (a) in situ daily average sea surface temperature (black line) and salinity (grey line) (b) SIMC1 daily average wind speed (c) the difference between SIMC1 daily average seawater

pCO_2 and in situ daily average atmospheric pCO_2 (d) SIMC1 daily average air-sea CO_2 fluxes (aeration process). The summer upwelling period (from 1 May to 1 October) is highlighted in yellow.

We wanted to highlight the SUP, we agree that these panels do not bring more information and that left panel are pretty clear so we modified Figure 5 by deleting the right panels and highlighted the SUP in yellow on left panels.

Page 21, Lines 506 ff: Would including the DIC and nutrient input from upwelling improve the model-data agreement, or the converse?

As indicated in point 2), in our configuration, we can consider these inputs in two ways: by using a relation with salinity or by using an interpolation of SOLEMIO measurements which is then read by the model. For DIC, both methods do not give satisfactory results. For nutrients, we used the second methods in Barré et al. (2023a). In both cases, it is difficult to directly test (through a simulation which take them into account) their effect on model-data agreement. However, to answer your question, we propose the followings hypotheses based on DIC and nutrients measurements study. We represent a linear interpolation of SOLEMIO measurements for these variables on figure S5:

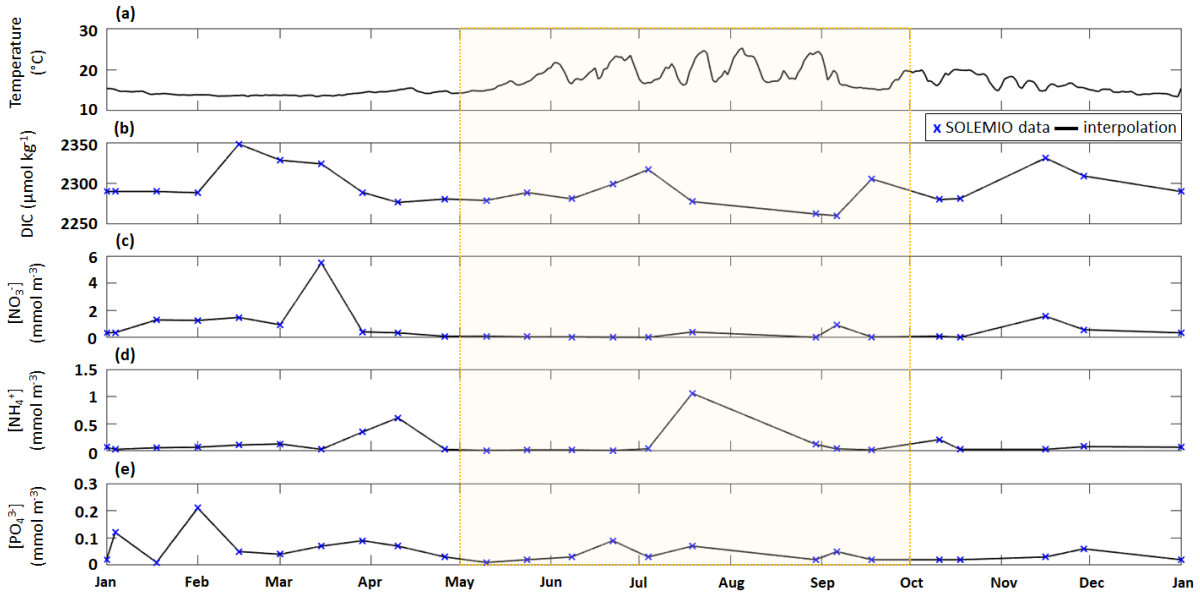

Figure S5: Time series of surface (a) temperature (PLANIER measurements), and interpolated (b) DIC, (c) NO_3^- concentration, (d) NH_4^+ concentration and (e) PO_4^{3-} concentration at SOLEMIO station. SOLEMIO data are represented by blue markers and the SUP is shaded in yellow.

For DIC, during the SUP, measurements show values around 2283 µmol kg^{-1} (mean of DIC measurements during the SUP). All upwellings of the period do not necessarily impact DIC. Only two events are noticeable: at the beginning of July and mid-September. These two events do not seem to be correlated with LSE, or Cortiou water inputs (generally associated with high NH_4^+ concentrations). The first event is not reproduced by the model, we then assume that this DIC increase could be associated with an upwelling event. However, the second one is well reproduced by the model which means that it is not resulting from an upwelling input (as, for now, we do not take these inputs into account in the model) (Fig. 4 of the manuscript). Considering these results, we believe that adding DIC inputs from upwelling could improve the realism of our representation, and consequently the data-model agreement.

For nutrients, it clearly appears that during the SUP, their dynamics are only slightly affected by upwelling events as nutrients concentrations remain close to 0 for most of the time. Only two nutrient inputs are noticeable during the SUP: in July and September. However, these events do not correspond to upwellings as the first one is associated with Cortiou water which reaches SOLEMIO (high NH_4^+ concentration) and the second one is, as showed in the manuscript associated with a Rhône River intrusion. This low impact can be explained by the fact that, when the upwelling takes place, nutrients which are upwelled are quickly consumed by the phytoplankton present in the area, then not reaching the station. Considering these results, we suppose that taking into account nutrients inputs associated with upwelling events could improve the model data agreement as it might bring some more realism to our representation, but not enough to consider them here, as their impact at the station is quite limited.

To conclude, we think that considering upwelling inputs could be a great addition to improve the realism of our representation and then the model-data agreement, especially for DIC, however, considering that, in the present configuration, these contributions can hardly be taken into account, we believe that switching to a 3D configuration will be the most appropriate way to confirm this.

Page 22, Line 530ff: Can one give a conjecture why the model overestimates pCO2 during winter?

We already explained why the model overestimates pCO_2 during winter l.535: [Seawater pCO_2, air-sea CO_2 fluxes and DIC are closely connected (Appendix B, Fig. 3). In Eco3M_MIX-CarbOx, aeration is simulated by applying Eq. (5) to 1 m^3 of surface water at SOLEMIO station which tends to overestimate the impact of aeration process on DIC and, due to the close link between DIC and pCO_2, also on pCO_2].

As seawater pCO_2 calculation is closely linked to air-sea CO_2 fluxes and DIC, we can assume that, when we tend to overestimate the impact of aeration process on DIC, we then impact pCO_2 and also overestimate it in winter.

Technical comments:

Thank you for this, we take them into account.

References:

Allen, J. I., Holt, J. T., Blackford, J. and Proctor, R.: Error quantification of a high-resolution coupled hydrodynamic ecosystem coastal-ocean model: Part 2. Chlorophyll-a, nutrients and SPM, Journal of Marine Systems, 68, 381-404, https://doi.org/10.1016/j.jmarsys.2007.01.005, 2007.

Barré, L., Diaz, F., Wagener, T., Van Wambeke, F., Mazoyer, C., Yohia, C. and Pinazo, C.: Implementation and assessment of a model including mixotrophs and the carbonate cycle (Eco3M_MIX-CarbOx v1.0) in a highly dynamic Mediterranean coastal environment (Bay of Marseille, France) (Part I): Evolution of ecosystem composition under limited light and nutrient conditions, https://doi.org/10.5194/gmd-16-6701-2023, 2023.

Fraysse, M., Pairaud, I., Ross, O. N., Faure, V. M. and Pinazo, C.: Intrusion of Rhone River diluted water into the Bay of Marseille: Generation processes and impacts on ecosystem functioning, Journal of Geophysical Research: Oceans, 119, https://doi.org/10.1002/2014JC010022, 2014.

Kwiatkowski, L. and Orr, J. C.: Diverging seasonal extremes for ocean acidification during the twenty-first century, Nature Climate Change, 8(2), 141-145, 2018.

Lajaunie-Salla, K., Diaz, F., Wimart-Rousseau, C., Wagener, T., Lefevre, D., Yohia, C., Xueref-Remy, I., Nathan, B., Armengaud, A., and Pinazo, C.: Implementation and assessment of a carbonate system

model (Eco3m-CarbOx v1.1) in a highly dynamic Mediterranean coastal site (Bay of Marseille, France), Geoscience Model Developpment, 14, 295–321, https://doi.org/10.5194/gmd-14-295-2021, 2021.

Lewis, E. and Wallace, D. W. R.: Program developed for CO_2 system calculations, 1998.

Maréchal, D.: A soil-based approach to rainfall-runoff modelling in ungauged catchments for England and Wales, Ph.D. Thesis, Cranfield University, 157pp, 2004.

Radach, G. and Moll, A.: Review of three-dimensional ecological modelling related to the North Sea shelf system. Part II: model validation and data needs, Oceanography and Marine Biology, 44, 1–60, 2006.

Sharp, J. D., Pierrot, D., Humpbreys, M. P., Epitalon, J. -M., Orr, J. C., Lewis, E. R. and Wallace, D. W. R.: CO2SYSv3 for MATLAB (v3.0.1). Zenodo [code] https://doi.org/10.5281/zenodo.3952803, 2020.

Schneider, A., Douglas, W. R. W. and Körtzinger, A.: Alkalinity of the Mediterranean Sea, Geophysical Research Letters, 34, https://doi.org/10.1029/2006GL028842, 2007.

Stow, C. A., Jolliff, J., McGillicuddy Jr, D. J., Doney, S. C., Allen, J. I., Friedrichs, M. A., Rose, K. A. and Wallhead, P.: Skill assessment for coupled biological/physical models of marine systems, Journal of Marine Systems, 76(1-2), 4-15, 2009.

Wimart-Rousseau, C., Lajaunie-Salla, K., Marrec, P., Wagener, T., Raimbault, P., Lagadec, V., Lafont, M., Garcia, N., Diaz, F., Pinazo, C., Yohia, C., Garcia, F., Xueref-Remy, I., Blanc, P. E., Armengaud, A., and Lefèvre, D.: Temporal variability of the carbonate system and air–sea CO_2 exchanges in a Mediterranean human-impacted coastal site, Estuar. Coast. Shelf S., 236, https://doi.org/10.1016/j.ecss.2020.106641, 2020.