We thank anonymous reviewer #2 for his availability to review our manuscript and his suggestions will be taken into account. In what follows, our responses and the reviewer's comments are represented by blue and black text, respectively.

1°) I don't understand what the added value of this paper is. The authors should emphasize this point. To me, most of the results have already been shown previously. The originality of this paper should be clearly addressed.

Response 1:

It is true that our results on driving processes of mixed layer heat and salt budgets in the tropical Atlantic are mostly consistent with previous results, but we believe that the originality of our study lies in several other points:
1- While the PREFCLIM mixed layer climatology was previously published as a dataset (Dengler and Rath, 2015; Rath et al., 2016), it is the first study that scientifically exploits this dataset.
2- To our knowledge, heat/salt advective terms have been previously estimated by model, observations or both (reanalyses) but independent estimations from model on one side, and from observations on the other side, have rarely been confronted (if ever).
3- The important role of mesoscale variability (and the subsequent need for subsampling to compare modelled and observed advection) had not been previously quantified in the region of study, and this was done here in the frame of projects related to new/future satellite missions (SWOT, SMOS-HR) which should help to capture this mesoscale variability.
While we believe point 3 was already clearly stated (see last paragraph of part 4). The first 2 points are now strengthened in the introduction (see lines 115-118 "page 4").

2°) Why don't you forced the model with the same heat flux as in PFREFCLIM (or vice-versa) since it allows to do that. This will allow, at least, to overcome the problem found in the inconsistency of heat flux in the interpretation of results.

Response 2:

This difference in the heat flux comes from the fact that the PREFCLIM climatology and the NEMO tropical Atlantic simulation compared here were independently produced by two different research groups (German GEOMAR for the PREFCLIM climatology and French LEGOS for the NEMO simulations), who collaborate in this paper. We appreciate the suggestion of the reviewer to overcome the inconsistency problem of the heat flux. But unfortunately it is not so easy to update the PREFCLIM product or NEMO simulation. In particular we do not have computational time available at the moment in our laboratory to rerun the simulation with different heat forcing fluxes, and wonder if it would be worth the carbon cost. Actually, only the solar flux (SWR) is used for forcing the model, while non-solar fluxes are computed through bulk formulas, allowing feedback from simulated oceanic conditions (this now clarified at lines 285-286 "page 13" in the paper), and differences between NEMO and PREFCLIM are larger for the latent heat flux (LHF) than for the SWR (see figure R2.1). Also, the heat flux term for PREFCLIM and NEMO depends also on the mixed layer of each product. Comparing Figs R2.1 below and Fig. 4a-b, it appears that, although SWR is stronger in the model (particularly in the south-west corner), the net heat flux is generally weaker in the model, so the different solar fluxes used for NEMO and PREFCLIM has no direct influence. Moreover, the differences in solar flux forcing does not affect the differences in advective terms which are the main focus in the paper.

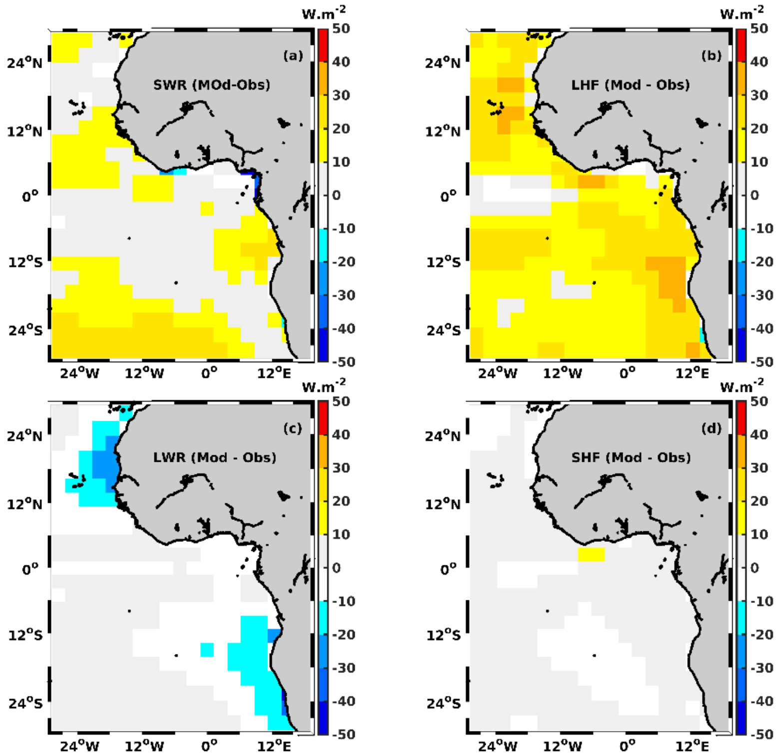

Figure R2.1: Difference of components of heat flux between model and observations: short wave radiation (a), latent heat flux (b), long wave radiation (c) and sensible heat flux (d). Except of SWR, the difference of other is computed from absolute values: abs(model) - abs(observation).

3°) The seasonal variations of the salinity tendency are quite different from Obs and Model. Can you explain this?

Response 3:

The differences in the seasonal variations of the salinity tendency between Obs and Model directly derive from the differences in the mixed-layer salinity seasonal variations (Fig. 8a/d/g/j), which have generally a larger amplitude and lower minima in the model, as also seen in Fig. 3. This can be due to several causes. First, although the PREFCLIM product benefited from newly available hydrographic data in the Senegal, Angola and Namibia coastal waters, the data density is still low in the equatorial Gulf of Guinea (Dengler and Rath, 2015), where lie the freshest waters associated with heavy rains and the large Congo and Niger river plumes, hence the largest difference in seasonal salinity variations in the Equatorial box (Fig. 8d). Poor data density can also be associated with a seasonal bias that may prevent capturing the full seasonal cycle. Second, hydrographic profiles, notably those from Argo floats, do not sample the salinity minimum found in the upper few meters of the ocean in regions highly stratified by rain and river plumes, which induces SSS estimations higher than those observed from satellite (Boutin et al., 2016; Houndegnonto et al., 2021). This leads to overestimation of mixed-layer salinity too. Third, while the NEMO model configuration has high vertical resolution in the upper few meters and homogeneous spatial coverage, the way it reproduces mixed-layer salinity highly depends on its freshwater forcing, including river runoff, and its own dynamics that are of course not perfect. This discussion has been added in the paper (see lines 584-596 “page 31”).

4°) Is the use of PREFCLIM suitable for this kind of study since it provides a coarse temporal and spatial resolution?

Response 4:

Yes, we believe it was worth using PREFCLIM in this study. Indeed, the coarse temporal and spatial resolution of the PREFCLIM climatology (for heat/salt budget terms) does not allow the consideration of nonlinear terms related to mesoscale activity. This problem could be solved with the arrival of new satellite missions for high frequency variability such as SWOT, SMOS-HR, as noted in the paper (see lines 656-658 “page 33”). But the comparison of PREFCLIM with offline NEMO also shows that there is still uncertainty in other budget terms, and it is always important to have some observations to validate a model.

5°) What is the criterion used to delimitate the boxes? You should clarify this in the text.

Response 5:

The criteria are partly subjective, but we now tried to precise them, and also note that there is a trade-off between SST and SSS characteristics (see lines 262-270 “page 12”).

Line 37: are you sure that paper of Giordani and Caniaux, 2001 deals about the altantic cold tongue?

Response 6:

Thanks for pointing at that. It was a mistake, it has been replaced by Caniaux et al. (2011) in the revised manuscript.

Line 257: Please can you explain why there a discrepancy between the model and the Obs only at the equator?

Response 7:

As pointed out above, this disagreement is partly associated with the different data sources used for heat fluxes in the observations (TropFlux) and as forcing in the model (DFS5.2). However, there is in places more disagreement in non-solar fluxes calculated by the model through bulk formulas (see figure R2.2). These causes for the differences are cited in the paper (see lines 284-286 “page 13”). A sensitivity test based on the weak correlations (see Figure 4.d of the manuscript) between heat fluxes from observations and the model west of 12°W in the equatorial band (3°S-3°N, West of 12°W) shows most important differences in the solar short-wave flux and latent heat flux between observations and the model. We think it is beyond the scope of the paper to investigate further reasons for these differences.

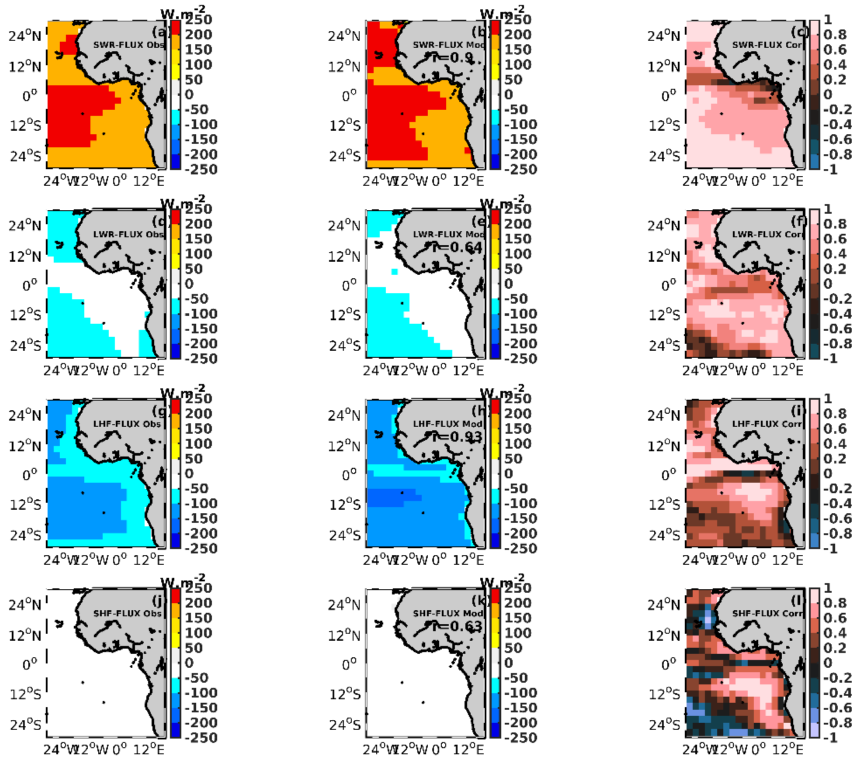

Figure R2.2: Mean of different components of heat flux from observations (a,d,g,j) and model re-sampled at 2.5° resolution (b,e,h,k): short wave radiation (a,d), long wave radiation (d,e), latent heat flux (g,h) and sensible heat flux (j,k). (c,f,i,l) represent the seasonal correlation of these components between observations and model, which is 95% significant when r > 0.58. Over (b,e,h,k), r indicates the spatial correlation between observations and model, which is 95% significant when r > 0.12.

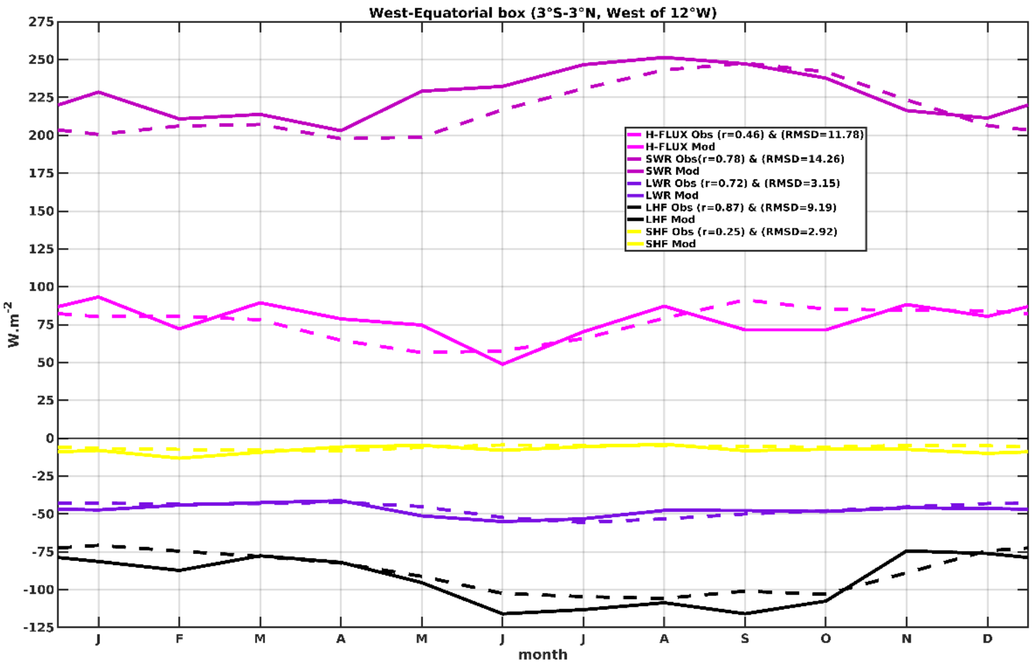

Figure R2.3: Seasonal cycle of heat flux and this decomposition from observations (dashed line) and model (full line) in selected region (3°S-3°N, West of 12°W for sensitivity test). All terms are in °C per month. r indicates the seasonal correlation between observations and model, which is 95% significant when r > 0.58.

Reference

Boutin, J., Chao, Y., Asher, W. E., Delcroix, T., Drucker, R., Drushka, K., Kolodziejczyk, N., Lee, T., Reul, N., Reverdin, G., Schanze, J., Soloviev, A., Yu, L., Anderson, J., Brucker, L., Dinnat, E., Santos-Garcia, A., Jones, W. L., Maes, C., Meissner, T., Tang, W., Vinogradova, N., and Ward, B.: Satellite and in situ salinity understanding near-surface stratification and subfootprint variability, Bull. Am. Meteorol. Soc., 97, 1391–1407, https://doi.org/10.1175/BAMS-D-15-00032.1, 2016.

Dengler, M. and Rath, W.: Seasonal heat and fresh water mixed-layer balance climatology, 1–15 pp., 2015.

Houndegnonto, O. J., Kolodziejczyk, N., Maes, C., Bourlès, B., Da-Allada, C. Y., and Reul, N.: Seasonal Variability of Freshwater Plumes in the Eastern Gulf of Guinea as Inferred From Satellite Measurements, J. Geophys. Res. Ocean., 126, 1–27, https://doi.org/10.1029/2020JC017041, 2021.

Rath, W., Dengler, M., Lüdke, J., Schmidtko, S., Schlundt, M., Brandt, P., Bumke, K., Ostrowski, M., van der Plas, A., Junker, T., Mohrholz, V., Sarre, A., Tchipalanga, P. C. M., and Coelho, P.: PREFCLIM: A high-resolution mixed-layer climatology of the eastern tropical Atlantic, https://doi.org/10.1594/PANGAEA.868927, 2016.