Comments to gmd-2021-436
“The bulk parameterizations of turbulent air-sea fluxes in NEMO4: the origin of Sea Surface Temperature differences in a global model study” by Bonino et al.

The authors exploit a state-of-the-art ocean model to evaluate the differences in three algorithms of bulk air-sea turbulent fluxes. The use of an ocean model enables to have a dynamical SST that modulates the fluxes, so that the air-sea coupling is partially represented. The atmospheric response to this SST forcing is, in fact, still missing.

The work is surely of interest, timely and well structured. Apart from some English editing and some improvements of the figures, to increase their readability, there are few major points (see the ‘Major comments’ section) that should be addressed before being accepted. In particular, some more analyses of the oceanic response to the different forcing should be included, to fully take advantage of the information provided by these heavy simulations. For these reasons, I suggest major revisions.

Major comments

A lot of effort is put in explaining quite successfully why there is a cold bias in the equatorial eastern Pacific and the EBUS regions between ECMWF and NCAR bulk schemes. However, very little attention is given to the strong differences in heat fluxes found over western boundary current (WBC) systems, which are known to be areas with strong air-sea interactions (See figure 2 of the manuscript). I invite the authors to dig a bit more in this direction and explore if there is a meaning in the spatial and temporal variability of this difference. If it was only noise, I would expect it to average to zero in an annual mean, but, indeed, it is visible in both ECMWF-NCAR and COARE-NCAR differences (figure 3).

What are the limitations in running simulations that last one year only? Is there any dependence on the specific year (e.g. in terms of ENSO phase, or any other climatic mode)? What about the spinup of the model? Which year has been considered?

With respect to Brodeau et al. (2017, B17 hereafter), the fluxes are computed using a dynamical SST field that responds to the atmospheric forcing. However, the atmospheric dynamics is known to respond to the SST even on daily and sub-daily time scales (see the review of Small et al., 2008, and some examples of applications in different areas of the world such as Li and Carbone, 2012; Gaube et al., 2019; Desbiolles et al., 2021). It would be interesting to discuss a fully coupled approach, as it has been shown that surface winds and clouds are affected by the SST structure on daily time-scales which, then, affect the SST and the surface turbulent fluxes back. This is only mentioned at the end of the manuscript and it should probably be included in the Introduction, as well. Moreover, the closed loop of this kind of ocean-atmosphere interactions has been proposed to be responsible for a three-to-six day oscillation (Strobach et al., 2020): I wonder if these oscillations are also observed here and whether they depend on the flux parameterizations.

In general, the fact that full ocean simulations are performed seems a bit underexploited. I think that much more information could be extracted, for example when discussing the role of different wind stress and wind stress curl in controlling the surface cooling in the EBUS and equatorial regions. Would it be possible to disentangle the role of upwelling and the role of entrainment in this surface cooling? What about doing some heat budget in the oceanic mixed layer to understand what processes are mostly modified by the different bulk algorithms?

The statistical significance of the differences between the experiments should be assessed. If the distributions are Gaussian, a t-test should be enough.

There are various differences with respect to the estimates shown in B17. In particular:

1. the authors consider a single year, whereas B17 consider the period 1982-2014;
2. the authors use ERA5 data to force NEMO, whereas B17 use ERA-Interim data;
3. COARE 3.5 is used here, and COARE 3.0 is used in B17;
4. Different versions of the ECMWF model are considered (cycle 40 and 41).

For these reasons, I would be more cautious in comparing the present results with those presented in B17. Would it be possible, for example, to compute the heat fluxes using the local midnight SST throughout the day, to mimic the fixed-SST approach, as in B17, and compare these fluxes to the prognostic-SST ones? This would avoid all the limitations highlighted before, as the original data would be the same.

Minor comments

There are typos throughout the text: I suggest a careful reading of the manuscript.

Many maps are hard to read, because the contour lines often mask the color shading. I suggest:

- enlarging the maps (as currently done in Fig 5, at least);
- reducing the number of contour lines (or removing them, if not necessary);
- verifying that the contours are properly plotted and not broken at 180° or 0° longitude;
- removing the word ‘exp =’ in the titles, as it is redundant.

Technical comments

The authors use both the words ‘parameterization’ and ‘parametrization’. I suggest choosing the former and keeping it consistent throughout the text. Note that also the verb ‘parameterise’ is used in some places.

Abstract: specify the NEMO acronym.

L28 (and elsewhere): I would avoid the word ‘state’, as it has a specific thermodynamic connotation. I would simply refer to wind, air temperature and humidity as ‘surface atmospheric variables’.

L32: One might want to add the reference to the latest version of the COARE 3.5 algorithm, namely Edson et al. (2013).

L50: It is true that with the prognostic SST approach, there is a negative feedback between the heat fluxes and the SST, but having a dynamical ocean can also modify the heat fluxes in the other direction. Namely, the heat fluxes can be strengthened (in absolute value) with the upper ocean mixing. Is there a way to disentangle these two contributions?

L76: ‘Laplacian’ should be capitalized.

L77: ‘s-1’ should be ‘s^{-1}’.

L84 (and L148): The reference to ERA5 should be updated to Hersbach et al. (2020).

Eq (1a) should contain *U* instead of *u*.

L97: Q_T is dominated by Q_L because Q_L is much larger than Q_S. I wonder if the buoyancy flux, in which the sensible and the latent heat flux terms are comparable, is a more appropriate variable to consider. A recent example of its dynamical importance is the work by De Szoeke et al. (2021), where the buoyancy flux is shown to control the low-level cloud formation in the tropical Indian Ocean. This, then, has a significant influence on the surface fluxes.

L100: Replace ‘increased’ with ‘increase’.

L103: An article ‘the’ has been misspelled.

LL106-107: The explanation of the cool-skin effect is not very clear.

L115: As the atmospheric component is not dynamically coupled to the ocean model I would modify this sentence: the air-sea feedback loop is partially closed.

L120: Is the word ‘divergence’ used to indicate the difference between the model formulations? If so, please replace it because it has specific meaning in vectorial calculus, which is a bit misleading.

L124: Edson et al. (2013) introduced COARE 3.5 and not 3.6. This should be corrected throughout the manuscript.

Fig1: Instead of using thin lines for the moisture transfer coefficients, thick dashed lines would be more visible and easier to distinguish from the drag coefficients. I would also suggest reducing the range of wind speed in panel (b) up to 22 or 25 m/s, as the focus is on the left side of the panel.

L134: Repetition of the word ‘of’.

LL134-144: There are several typos or missing articles and commas in this paragraph, please revise.

L148: Remove ‘Weather’.

L151: The word ‘uses’ is chopped.

LL150-163: I would suggest removing the bullet points and use plain text, instead, to remove the repetitions and enable a smoother reading. Table 1 is already giving a schematic recap of the experiment setup. It is also not clear what is the difference between the parameterizations that use the absolute wind speed (as in experiments ECMWF_S, COARE_S, ECMWF_NS and CdNCAR_CeEC) and the parameterization that does not include the current correction (as in the NCAR experiment). It seems that the ocean surface currents are never used in this set of experiments (L169). Thus, it can simply be stated once, as this is not a parameter that changes. I also find the names of the experiments very confusing: what about making them more explicit with something like: CdEC_CeEC (instead of ECMWF_S), CdCO_CeCO (COARE_S), CdNC_CeNC_NS (NCAR), CdEC_CeEC_NS (ECMWF_NS), CdNC_CeEC_NS (CdNCAR_CeEC)? In this way, the differences among them are readily available in their names.

Table 1: What about adding ‘Experiment name’ in the first row of the first column? A column indicating whether the gustiness in the computation of the wind stress is included would be useful.

Figure 2 is not described in the main text, please do. What is its link with the SST bias? The name of the experiments should be kept consistent throughout the text. Here, for example, ‘COARE3.6_S’ should be replaced with ‘COARE_S’, or its correct name. It should be clarified (and motivated) whether the annual mean of the percentage difference or the percentage difference of the annual means is computed.

LL185-192: This paragraph is rather general and could be moved backward in the manuscript. One would expect here to find a reasoning on figures 2 and 3, such as why such patterns are observed, which specific reasons could explain them and their relationship, etc.

Fig4: The skin SST effect has a component at the daily scale. I wonder if, by considering the annual mean, the signal averages to zero. What about computing the temporal standard deviation of the difference SSTskin-SST?

LL202-207: The link between the figure and the text is not fully clear. By looking at the figure one might think that, on the annual average, there is an increase of SST when using the

SSTskin correction (is the sign of the difference ECMWF_S-ECMWF_NS correct?), because of a dominant diurnal warming effect. This is in contrast with the statement that the cool skin effect is dominant over the warm layer one. Is the annual mean computed using hourly outputs? How is the mean warming interpreted? What about its spatial structure?

L216: Are you referring here to the ECMWF_S or the ECMWF_NS experiment? Because, from the titles of the figures, the experiment ECMWF_S is considered in figure 3a, whereas the experiment ECMWF_NS is considered in figure 5a. Please clarify.

L218: The sentence is not grammatically correct and there is a repetition of the word 'differences'.

LL229-230: Is it 'higher heat absorption' or 'weaker heat loss'? The logical link between the latent heat considerations and the fact that it is the wind stress to be responsible for the observed cold SST pattern difference between ECMWF_S and NCAR is not clear.

L233: 'Computing'?

Fig6: Panel b) is not showing time series: the caption of the figure should be modified. From this figure one would not say that the mean excess QT is 10W/m2, as stated at line 221: where does this amount come from?

L239: Up to now, it is not very clear which parameterizations use the gustiness correction in the computation of the wind stress. As noted above, this information could be included in Table 1 and some more details on how the gustiness is included in the scheme should be given.

L240: 'Caused by the gustiness correction'.

Fig8: I suspect that the gustiness correction is highly variable in time on daily or even sub-daily scale. As for the CSWL correction, thus, I am not sure that showing the annual mean of such variables is enough. Wouldn't it be of interest to show the variance or the RMSE of the two model setup to better display where this highly variable correction is relevant?

L254: There seems to be an extra '(' at the end of the line.

L260: As the outputs of the ocean model are available, would it be possible to quantify the contribution of the surface cooling due to the modified Ekman suction between the configurations? Can a scaling between the anomalous wind curl and the anomalous SST cooling be derived?

LL277-289: By looking at figure 2 one would expect stronger differences in the SST in the COARE-NCAR comparison, and not in the ECMWF-NCAR one. What about showing the mean difference of the wind stress (not in percentage) and, maybe, the mean difference in wind stress curl, as it relates to the upwelling? Then, is there a contribution to the surface cooling from an increased entrainment of cold waters in the OML (oceanic mixed layer) because of a stronger wind stress?

Fig9: Panel (a) is it the annual mean of the percentage variation or the percentage variation of the annual mean?

L342: ‘favorite’.

Additional references

- Desbiolles et al. (2021) Links between sea surface temperature structures, clouds and rainfall: Case study of the Mediterranean Sea, *Geophysical Research Letters,* https://doi.org/10.1029/2020GL091839
- De Szoeke et al. (2021) Diurnal ocean surface warming drives convective turbulence and clouds in the atmosphere, *Geophysical Research Letters*, https://doi.org/10.1029/2020GL091299
- Gaube et al. (2019) Satellite observations of SST-induced wind speed perturbation at the oceanic mesoscale, *Geophysical Research Letters*, https://doi.org/10.1029/2018GL080807
- Li and Carbone (2012) Excitation of rainfall over the tropical western Pacific, *Journal of the Atmospheric Sciences*, https://doi.org/10.1175/JAS-D-11-0245.1
- Hersbach et al. (2020) The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, https://doi.org/10.1002/qj.3803
- Small et al. (2008) Air-sea interactions over ocean fronts and eddies, *Dynamics of Atmospheres and Oceans*, https://doi.org/10.1016/j.dynatmoce.2008.01.001
- Strobach et al. (2020) Three-to-six-day air-sea oscillation in models and observations, *Geophysical Research Letters,* https://doi.org/10.1029/2019GL085837