REFEREE #1: JUSTIN SMALL

Dear Dr. Justin Small,

we would like to thank you for the careful reading of the manuscript and the constructive comments that substantially helped to improve and clarify the paper. Answers to all your comments are detailed hereafter. Corrections to the English grammar were adopted in the revised manuscript according to the reviewer's recommendations, but they are not reported or discussed here. All authors agree with the modifications made to the manuscript. The comments by the referee are reported in bold followed by our responses (in blue). The text added to the revised manuscript is reported in italic font. The revised manuscript that includes track changes and line numbers is provided in pdf format.

Please note that, in this document, we use 'Figure' to identify the figures in the updated manuscript, while we use 'Plot' to identify the figures in this document.

The name of the experiments have been improved (following suggestions by the referees) as reported in Table 1. New names are used in this document and in the updated manuscript.

Experiment name	sea surface temperature used (T_s)	computation of C_D	computation of C_E and C_H	convective gustiness
COARE_S	SSTskin	COARE3.6	COARE3.6	Yes
ECMWF_S	SSTskin	ECMWF	ECMWF	Yes
NCAR	SST	NCAR	NCAR	No
ECMWF_NS	SST	ECMWF	ECMWF	Yes
CdNC_CeEC_NS	SST	NCAR	ECMWF	Yes
ECMWF_NS_NG	SST	ECMWF	ECMWF	No

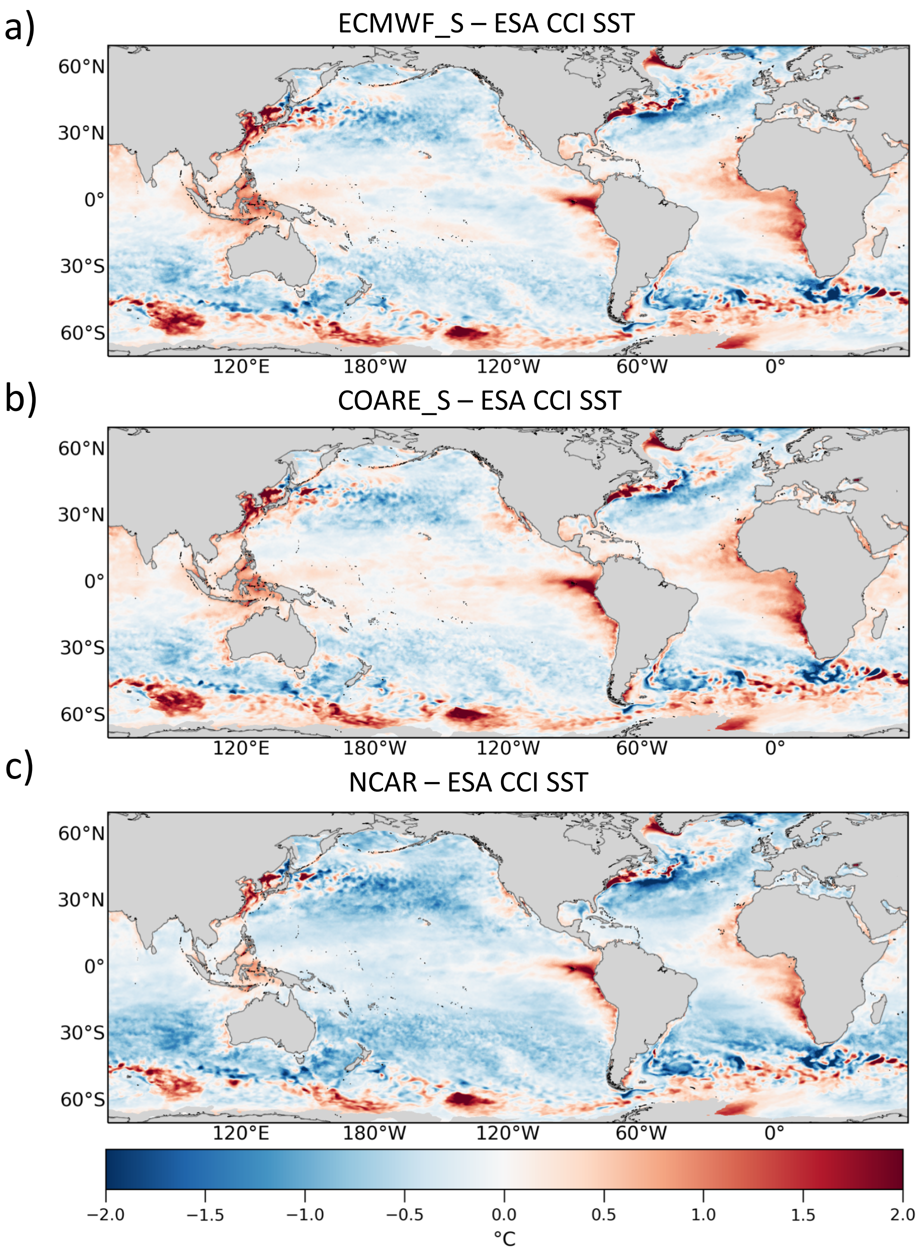
Table 1. Summary of the numerical experiments.

Major Comments

1) The paper shows differences of certain fields (SST, heat fluxes, momentum fluxes etc.) between the experiments. Is it possible to say whether any of the cases are more realistic than other, compared to observations, or is it complicated by competing and possibly cancelling effects of other parameterizations or processes? Could you look at other ocean variables (like the surface flow) to help with this?

We agree with the referee that a detailed comparison against observation can benefit the manuscript. Focusing on the sea surface temperature (SST), we compared the annual mean SST from the "control experiments" against the European Space Agency (ESA) Climate Change Initiative (CCI) SST (Merchant et al. 2019). Results are shown in the following Plot 1 (included as Figure 2 in Section 3.1 of the revised manuscript). Text has been added from lines 193 to 201: "*We compare the SST simulated by the*

ECMWF_S, COARE_S and NCAR control experiments with the European Space Agency (ESA) Climate Change Initiative (CCI) SST dataset v2.0 (hereinafter ESA CCI SST dataset) which consists of daily-averaged global maps of SST on a 0.05° x 0.05° regular grid, covering the period from September 1981 to December 2016 (Merchant et al., 2019). All the control experiments present a warm bias in the Eastern Pacific, in the Eastern Boundary Upwelling systems (EBUS), in the Western Boundary Currents (WBCs) and in the Antarctic Circumpolar Current (ACC) region. The SST reproduced by COARE_S and ECMWF_S shows a cold bias of about -1°C in the North Atlantic open ocean at mid- latitudes, and a warm bias of about 0.5°C in the Indian Ocean and the Western Pacific (Figure 2a,b); NCAR SST is also colder than observations, with a larger bias of about -2°C in the North Atlantic (Figure 2c). The bias is generally higher compared with other two experiments and covers wider areas. "



Plot 1: Annual mean SST differences between a) ECMWF_S, b) COARE_S, and c)NCAR against ESA CCI SST.

2) Can the results be put in context by comparing with known sensitivities to other well-known parameterizations or processes? Do the results have any impact on meridional heat transport?

The reviewer's suggestion is really interesting, but we think it goes somehow beyond the scope of the present work. The impact of bulk parameterizations on air-sea processes has been put in a more general context in a new paragraph added in the introduction (lines 39-46): *" It is worth mentioning that the online prognostic approach does only partially close the air-sea feedback. Surface winds and clouds are affected by the SST structure on daily time-scale which, in turn, affect the SST and the TASFs (Desbiolles et al., 2021; de Szoeke et al., 2021; Gaube et al., 2019; Li and Carbone, 2012; Small et al., 2008). The closed air-sea feedback (hereinafter coupled approach) in the system might substantially impact the turbulent fluxes (Lemarié et al., 2021; Small et al., 2008), but the coupled approach is still not yet mature in the ocean model community. Recently Lemarié et al. (2021) implemented a first attempt of a simplified atmospheric boundary layer model (ABL) to improve the representation of air-sea interactions in NEMOv4.2. However, the online prognostic SST approach is still largely used by the ocean modeling community in a variety of applications."*

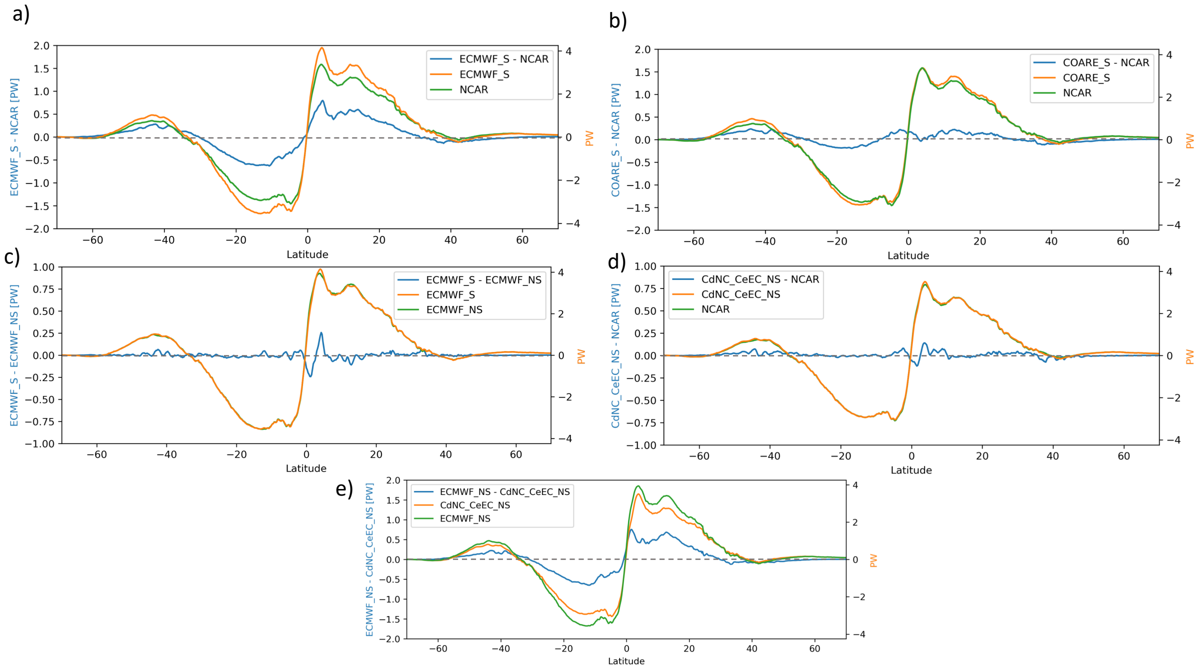
To quantify the impact of modified formulations on the meridional heat transport (MHT) , we computed it in the upper 100m of the global ocean and analyzed the differences among experiments (in Plot 2). The MHT in the ECMWF experiments is always higher compared to experiments that employ C_D from NCAR formulation. ECMWF (with and without skin temperature) and NCAR-based experiments present the largest differences (Plot 2 a,e) that peak generally in the tropical band (about 0.8 PW, 20% of NCAR absolute value) where ECMWF wind stress is stronger than NCAR one. Global MHT in COARE_S and NCAR runs is comparable (Plot 2b), with differences lower than 0.3 PW. The transport differs only in the tropical latitude band in all the experiments that used the same C_D (i.e. ECMWF_S and ECMWF_NS; NCAR and CdNC_CeEC_NS), and they are quite small (about 0.1 PW, Plot 2 c ,d).

Plot 2 (a,b) has been included in Section 3.3 of the revised manuscript, as Figure 4c,d; Plot 2e is in Section 3.4 as Figure 11c.

The following text was added to describe Figure 4 (lines 219-225): "*Changes on the simulated SST can reflect on the temperature profile in the upper ocean and the distribution of heat on global scales. We have computed the global ocean heat transport in the upper 100 meters and compared it among experiments. Figure 4 (c,d) presents the meridional heat transport (MHT) as a function of latitude. The MHT is larger in ECMWF_S compared to NCAR mostly at all latitudes (Figure 4c), with the largest differences (about 0.8 PW, 20% of NCAR absolute value) in the tropical band* where ECMWF wind stress is stronger than NCAR one *(Figure 3a). COARE_S and NCAR compare well, with differences* lower than 0.3 PW *(Figure 4d). Then, we will focus only on the differences between ECMWF_S and NCAR to analyze in detail the relationship between TASFs and SST. We show differences in MHT only when relevant.*"

The following text added in Section 3.4 (lines 313-317) describes Figure 11c: *" It is important to highlight that the differences in the wind stress are also responsible for the changes in the meridional heat transport. MHT differences between ECMWF_NS and CdNC_CeEC_NS resemble the differences between ECMWF_S and NCAR (compare Figure 4c and Figure 11c), with a higher transport in ECMWF_NS at all latitudes. The largest differences are located in the tropical region (up to 0.6 PW, about 18% of NCAR mean value), where the differences in meridional transport (linked to the equatorial upwelling) between the two experiments are likely maxima"*

In addition, we included the following text in the conclusion section (lines 384-385): “*Stronger wind stress results in an increase of the poleward heat transport in the upper ocean, which a more pronounced increase in the ±20 latitude band.*”



Plot 2: Global Meridional heat transport in the upper 100m ocean (values on the right y axis) and differences (values on the left y axis) between a) ECMF_S and NCAR, b) COARE_S and NCAR, c)ECMWF_S and ECMWF_NS, d) CdNC_CdEC_NS and NCAR, and e) ECMWF_NS and CdNC_CdEC_NS.

Minor Comments

1) Line 20. Isn’t surface radiative flux also highly important???

Thank you for the comment. We included the radiative flux, and the new sentence at line 19 was modified in: ”*These transfers of energy are primarily driven by surface radiative flux and turbulent air–sea fluxes (TASFs), which include wind stress and the turbulent heat flux components (THFs, latent and sensible heat fluxes).* ”

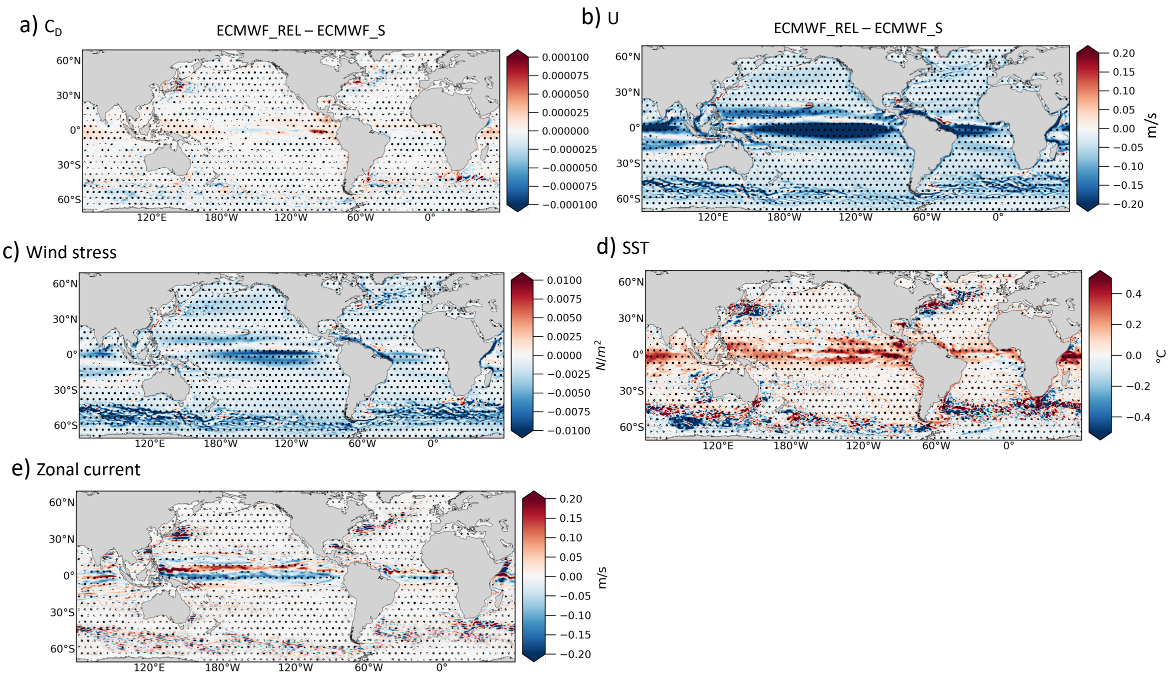
2) Line 78. Re “Marsaleix et al.” – based on the title of this paper, it does not obviously mention TKE, but it does mention energetics. Please confirm it is the correct reference. Sorry, I have not read it.

Following this suggestion, we checked the references and the correct one is indeed Blanke and Delecluse, 1993. We modified the manuscript accordingly.

3) At this point, the reviewer might anticipate experiments to look at the effect of including surface currents in stress. Your paper does not do this, which is fine, but you may want to refer to the extensive literature on the subject (e.g. Renault et al., Sun et al, and many others).

We thank the reviewer for this comment. As requested by Referee#3, we performed an extra experiment, 1 year long, in which we applied the relative wind, instead of absolute wind, in the

ECMWF_S bulk parameterization. We refer to the new experiment as ECMWF_REL . Plot 3 presents the results. As expected, the wind stress is reduced by the inclusion of the surface ocean velocity in the bulk formula, with respect to the absolute wind simulation: the wind speed in ECMWF_REL is weaker (up to -0.2 m/s) than ECMWF_S in the equatorial band (Plot 3b). As expected from the dependencies between C_D and the wind speed (Figure 1b of the manuscript), we find higher values of C_D in ECMWF_REL in the area of calm wind conditions and weaker values elsewhere. Differences of C_D and U between experiments are reflected onto the resulting wind stress field after bulk calculation (Plot 3c): the ECMWF_REL wind stress is weaker than ECMWF_S, especially where the U differences are higher (e.g. equatorial band). This difference in wind stress also leads to the SST differences (Plot 3d), hence ECMWF_REL results are warmer than ECMWF almost everywhere. Changes in wind stress also affect the current (Plot 3e): due to the weaker wind stress along the equator, the ECMWF_REL zonal currents are weaker than ECMWF ones. Even though the results provide insight into the effects tha bulk modifications can have in the upper ocean , we think that the current-stress negative feedback needs more and longer experiments (i.e. one for each bulk parameterization) to be properly assessed. We do not include a proper analysis in the manuscript, but we consider the effect of relative vs. absolute wind in the manuscript. Text was added in section 2.2 (lines 122-126): " *The effect of the ocean current interaction/feedback in the bulk formulation has been widely explored in the literature (e.g. Renault et al., 2019a, b; Sun et al., 2019). Although many previous studies highlighted the substantial difference in the surface input to the ocean between calculations that use absolute vs. relative wind, we have preferred to leave this aspect to further work since the implementation of this correction does substantially depend on the characteristics of the forcing fields (Renault et al., 2020).*"



Plot 3: Annual mean differences of a) drag coefficient (C_D), b) wind speed (U), c) Wind stress, d) SST and e) zonal current between ECMWF_S and ECMWF_REL.

4) List starting Line 110. I would add:

3. Effect of including ocean current in stress

4. The form of the exchange coefficients

Then you can mention which of these effects you look at. Am I correct in thinking you do not explicitly look at the effect of convective gustiness? See comment later.

We added the two bullet points and the following text (lines 121-122): "*In this study, we attempt to disentangle the effects of the first two aspects on SST (section 3.2,3.3 and 3.4), and we discuss the effect of the inclusion of convective gustiness in the wind stress computation (section 3.4) .*"

5) Line 128 . NCAR scheme ... minimum wind speed of 0.5 m/s ... This is interesting, and I just confirmed this is also done in the CESM scheme. Note that the Large and Yeager drag coefficient actually goes to infinity as you approach zero wind speed (your Fig. 1b). So even if the wind speed gets very low, the momentum flux remains significant.

Yes, we confirm that the description here refers to the NCAR formulation as introduced in NEMO where the minimum wind speed is used.

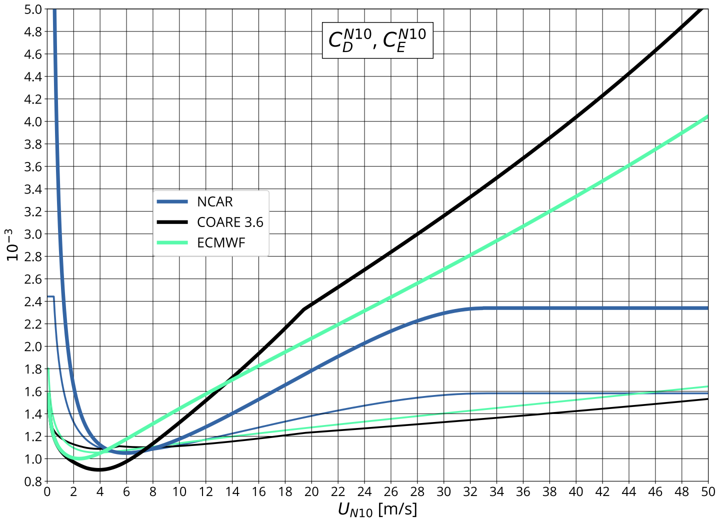
6) Lines 129-134. I would say that the Large and Yeager scheme also uses MOST. It combines MOST theory with a semi-empirical form of drag coefficient.

We clarified this point in the text. The sentence has been modified (lines 144-145): "*The NCAR parameterization uses a combination of the MOST theory with a semi-empirical form of drag coefficient in which the BTCs are computed as function of … "*

7) Fig. 1b. I understand that you do not focus on high/extreme wind speeds, but I am curious to know what happens above 35m/s. There is some discussion on this topic in Fu et al. (2021), their sections 3.2 and 4.1. (Note that their paper employs the original Large and Yeager (2004) form of drag coefficient, without reduction at high wind speeds.) Note also that ERA5 is a high-resolution dataset and will include extreme wind events. Reference: Fu, Dan et al. 2021: Introducing the new Regional Community Earth System Model, R-CESM. B. Amer. Meteor. Soc., 102, E1821-E1843, https://doi.org/10.1175/BAMS-D-20- 0024.1

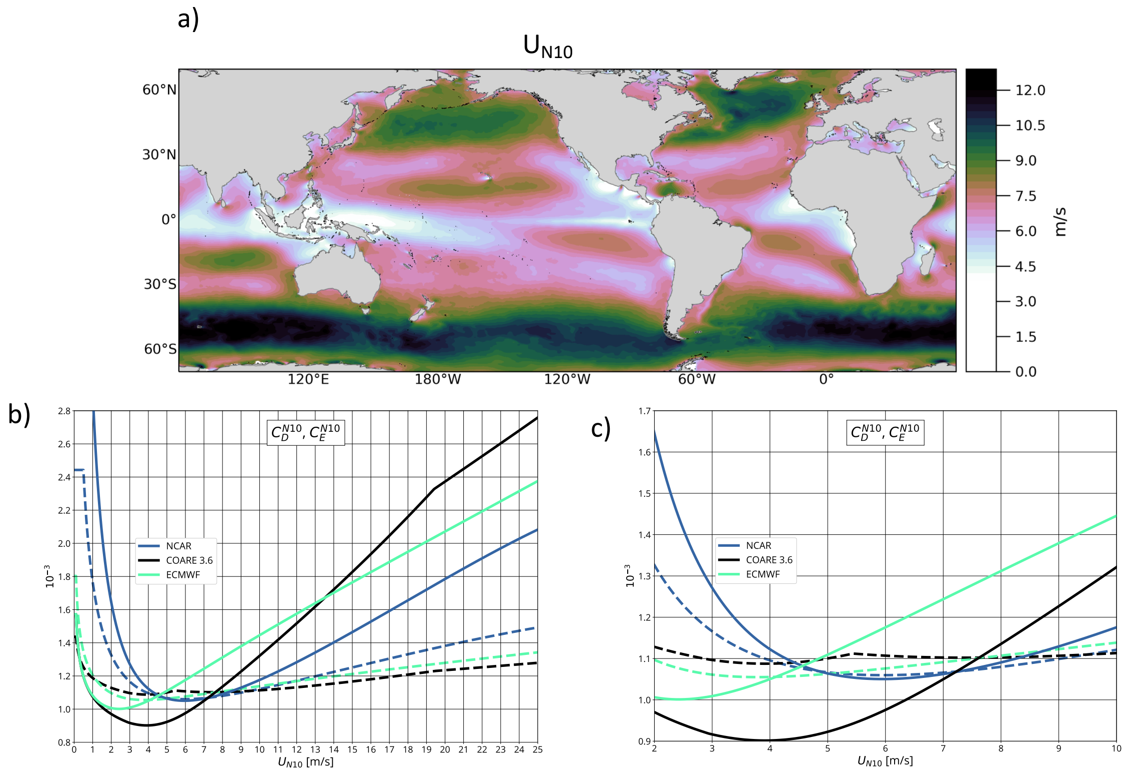
Lines 134-145. It would be useful to show a zoomed-in plot of Fig .1b for winds 10m/s or less.

We thank the reviewer for suggesting these interesting papers. Here is how neutral drag and moisture transfer coefficients (thick and thin lines, respectively) vary for wind stronger than 35 m/s:



Plot 4: Neutral drag and moisture transfer coefficients for COARE_S (black), NCAR (blue), and ECMWF_S (green) bulk parameterizations (thick and thin lines, respectively), as functions of the neutral wind speed at 10m

Our study e does not focus on extreme wind events, then we decided to keep the original plot (in Figure 1 of the manuscript)in the paper with a 0-25m/s range (as suggested by Referee#2) to which we added a zoomed-in subplot for winds lower than 10m/s. The complete Figure 1 of the revised manuscript is:



Plot 5: a) Annual mean of U_{N10} from NCAR parameterization b) and c) Neutral drag and moisture transfer coefficients for COARE (black), NCAR (blue), and ECMWF (green) bulk parameterizations (solid and dashed lines, respectively), as functions of the neutral wind speed at 10m.

8) Line 221. Fig. 6a shows very small QT (~1 W/m2) over most of the Globe, only small regions reach 10W/m2.

Thank you, it was a mistake. We corrected it.

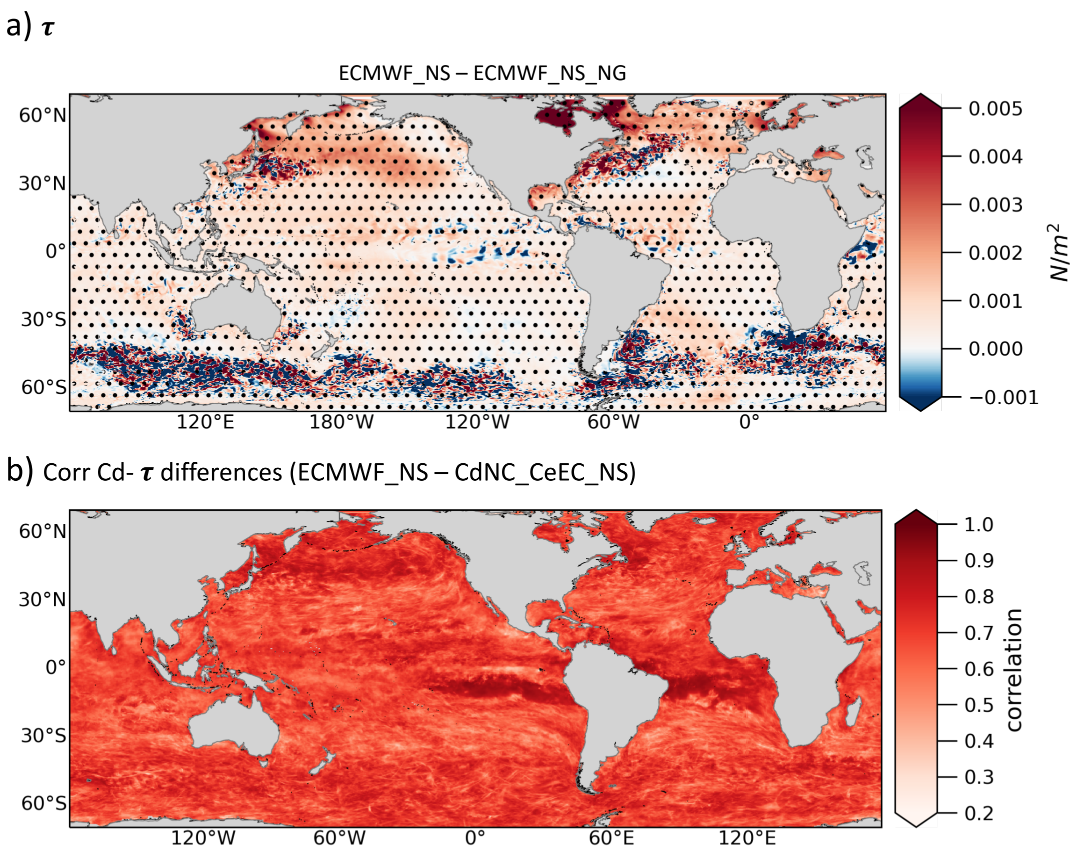
9) Line 258-259. I believe you do not explicitly look at the sensitivity to convective gustiness parameterization. So your inferences here are solely based on the fact that CD differences are small in these regions? You can consider running an extra sensitivity experiment with convective gustiness switched off in ECMWF.

Following this comment, we performed an extra ECMWF run where convective gustiness is switched off (it is named ECMWF_NS_NG). Plot 5a compares the wind stress in the "original" ECMWF_NS and the new ECMWF_NS_NG. Results confirm that wind stress is stronger almost everywhere when the convective gustiness is included in the U calculation. Nevertheless, it is worth underling that the differences in C_D between experiments are highly variable in time and that could hide the relationship between U, C_D and $\boldsymbol{\tau}$.

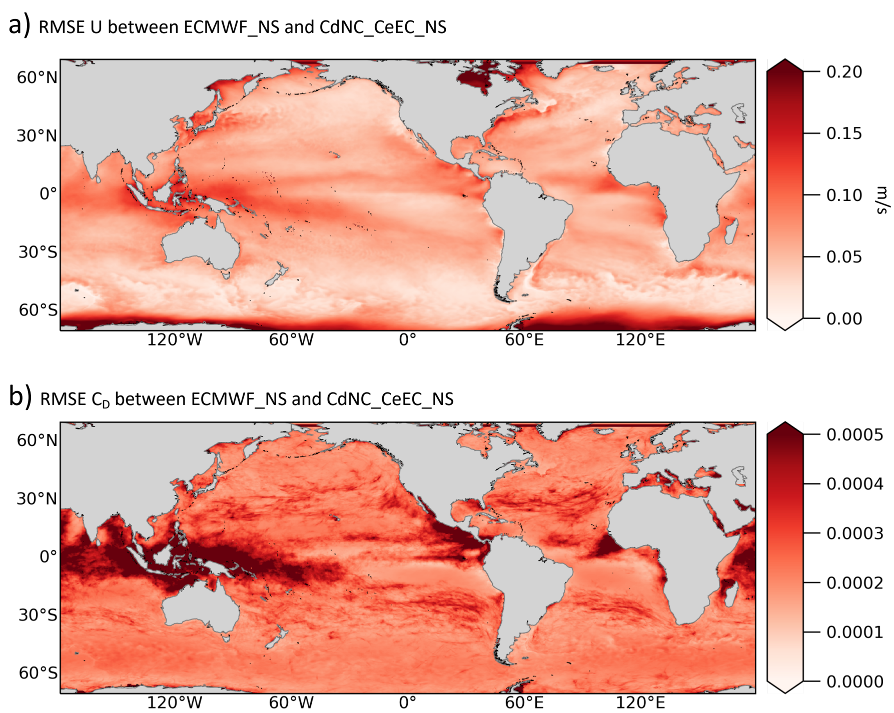
Following also suggestion by the Referee#2, we calculated the RMSE of the differences in the wind speed (U) and drag coefficient (C_D) between the ECMWF_NS and CdNC_CeEC_NS experiments (compare Plot 7a below with Figure 9a,c). The spatial distribution of the RMSE of U (Plot 7a) resembles the annual mean differences of U between the two experiments (shown in Figure 9a). The RMSE of C_D (Plot 7b) is large where the annual mean differences between ECMWF_NS and CdNC_CeEC_NS are negative (Figure 9b,c).

For this reason, we computed the correlation between the C_D differences and the $\boldsymbol{\tau}$ differences for the ECMWF_NS and CdNC_CeEC_NS runs. The correlation is always significant with positive values, (Plot 6b). The higher the difference in C_D, the stronger the difference in wind stress.

The following text was added in manuscript (lines 289-295): "*In regions where the differences in* C_D *and wind stress are opposite (e.g. the north-west and south-west Pacific and Atlantic ocean, Indian ocean, Baja California), the high time-variability of the* C_D *differences (not shown) could hide the relation between* C_D *and* $\boldsymbol{\tau}$*. In addition, including the convective gustiness in U calculation strengthens the wind stress in ECMWF_NS. Both hypotheses are verified, the* ECMWF_*NS experiment presents a stronger wind stress almost everywhere over the global ocean compared to a twin experiment where the convective gustiness is not used in the computation (Figure S2) and the correlation between* C_D *differences and wind stress differences is always significant and positive (not shown). The higher the difference in* C_D*, the stronger the differences in wind stress.*"



Plot 6: a) Annual mean differences of wind stress between ECMWF_NS and ECMWF_NS_NG, b) correlation between C_D differences and τ differences between ECMWF_NS and CdNC_CdEC_NS. Hatching indicates significant values (95% confidence level).

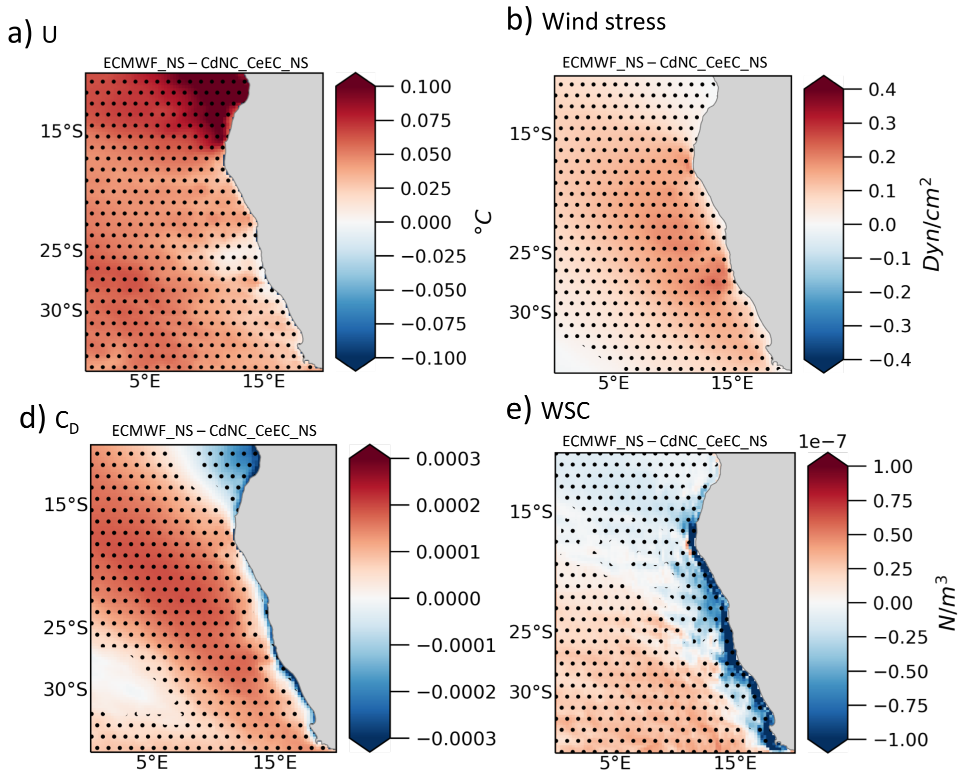


Plot 7: Root Mean Square Error of a) wind speed (U) b) drag coefficient (C_D) differences between *ECMWF_NS and CdNC_CeEC_NS.*

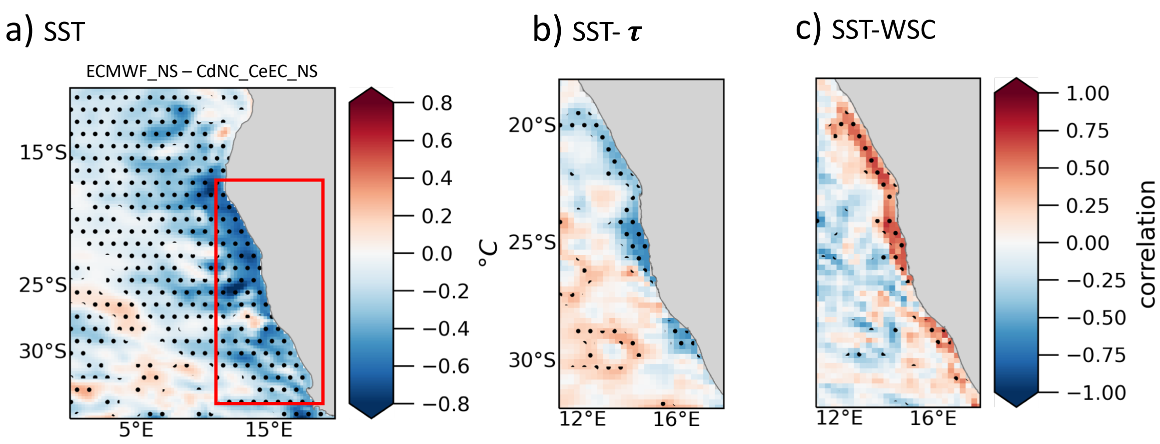
10) Lines 259 onwards. It is not obvious to me from Fig. 8 that the EBUS will be notable regions of enhanced stress and WSC. Perhaps zoom in on an example EBUS and show the causal links more clearly between U, CD, TAU and WSC.

We zoomed the results over the Benguela upwelling system (Plot 8). The ECMWF_NS experiment shows a notable increase of wind stress and wind curl along the Benguela coast. As previously commented, the wind stress is not affected by the type of first-order feedback at play for the NSHFs (SST-QT negative feedback in section 2.2). The wind stress is stronger when the C_D is larger in ECMWF_NS than CdNC_CeEC_NS C_D. The wind stress is slightly weaker when the C_D is lower in CdEC_CeEC_NS than CdNC_CeEC_NS.

It is worth noting that these cross-shore differences of wind stress lead to stronger wind stress curl in ECMWF_NS with respect to CdNC_CeEC_NS.



Plot 8: a) Annual mean differences of a) wind speed (U), b) wind stress (τ), c) drag coefficient (C_D), d) wind stress curl (WSC) between ECMWF_NS and CdNC_CeEC_NS. Hatching indicates significant values (95% confidence level).

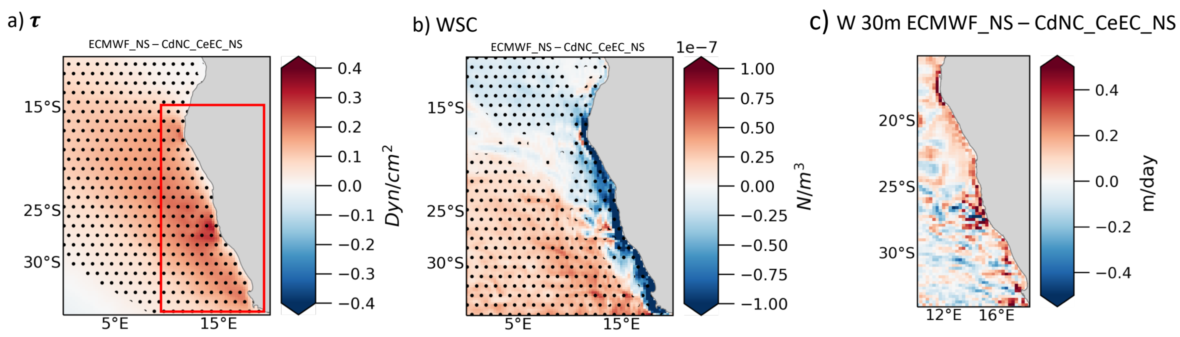


Plot 9: a) Annual mean differences of SST between ECMWF_NS and CdNC_CeEC_NS; b) correlation between SST differences and wind stress differences between ECMWF_NS and CdNC_CeEC_NS; c) same as in b) but for SST differences and wind stress curl differences. Hatching indicates significant values (95% confidence level).

To show the relationship between variables, we show the correlation of SST differences with wind stress and wind stress curl differences (Plot 9b,c). We added Plot 9 in the supplementary material as Figure S5, and we modified Section 3.4 including the following text (lines 311-312): "*..., the enhanced wind stress and negative wind stress curl in ECMWF_NS reinforce the vertical velocity with respect to CdNC_CeEC_NS (Figure S4), resulting in colder surface temperature (see correlation maps Figure S5).* "

10) On this topic, the lead author has 2 nice papers on EBUS in JRA55do and ERA-Interim-forced runs. Based on this experience, can you comment on whether the changes to TAU and WSC are realistic and whether they would make a sizable change to upwelling?

We thank the referee for this comment. Comparing Figure 4 of Bonino et al. 2018 and Plot 10, we can notice that, during upwelling season (ONDJ), the differences in wind stress and wind stress curl between JRA55do and ERA-Interim experiments show higher range of values with respect to the wind stress and wind stress curl differences between experiment that used different C_D parameterization. Since in the former pair of experiments, the atmospheric forcing is totally different, I would say that this result is quite expected. To better quantify the differences in the upwelling regions, we plotted the vertical velocity at 30m over the Benguela region (as Figure 6 in Bonino et al. 2018) for the ECMWF_NS and CdNC_CeEC_NS experiments (Plot 10c, d) and the differences between the two (Plot 9c) during upwelling season (ONDJ).



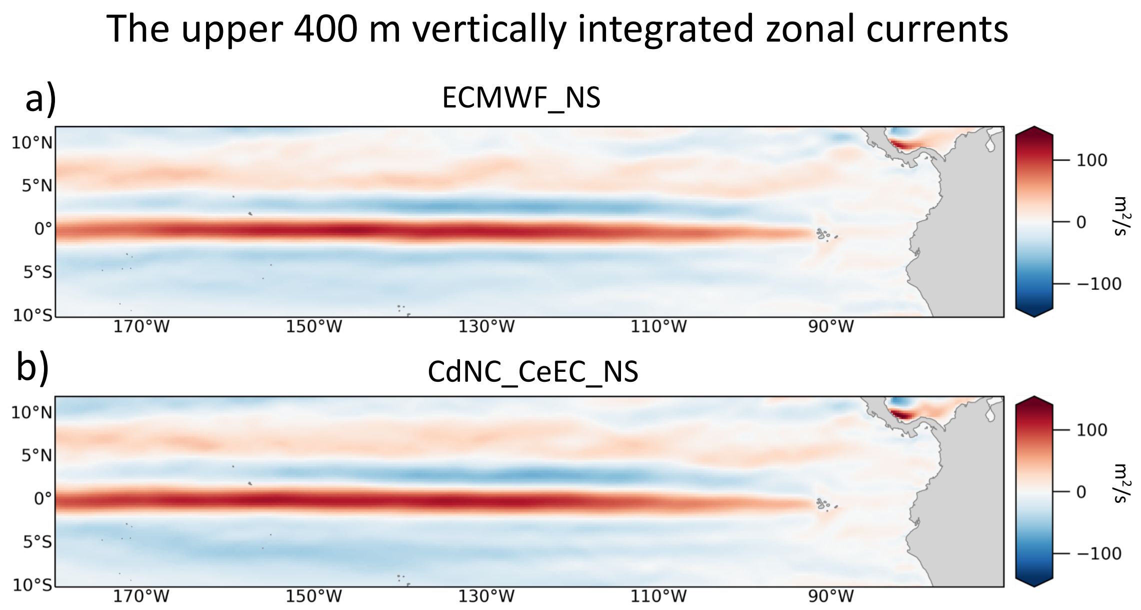
Plot 10: Seasonal mean differences (ONDJ) of a) wind stress (**τ**) and b) wind stress curl (WSC) between ECMWF_NS and CdNC_CeEC_NS; c) Differences in vertical velocity at 30m (W 30m) between the two ECMWF_NS and CdNC_CeEC_NS. Hatching indicates significant values (95% confidence level). Red square identifies the area shown in panel c).

As expected, the vertical velocity is stronger in ECMWF_NS experiments. The ECMWF_NS upwelling increases by about 30%. Comparing Figure 6 (bottom row) in Bonino et al 2018 with Plot 10, we can notice that the differences in vertical velocity are, in this study, half of the differences in Bonino et al. 2018, in terms of absolute values. It is worth noting that here the differences in wind stress and the WSC are both upwelling favorable for ECMWF_NS, while in Bonino et al 2018 wind stress differences are upwelling favorable for JRA55do and the WSC difference are upwelling favorable for ERA-Interim. Nevertheless, as expected, the differences in the wind forcing between experiments drives the difference in vertical velocity: the weaker - in terms of absolute values - the ECMWF_NS and CdNC_CeEC_NS wind forcing differences with respect to JRA55do and ERA-Interim, the weaker are the differences in vertical velocity. This suggests that the differences in wind stress and wind stress curl are comparable with results from Bonino et al 2018: the weaker (greater) the differences in wind forcing, the weaker (greater) are the differences in vertical velocity.

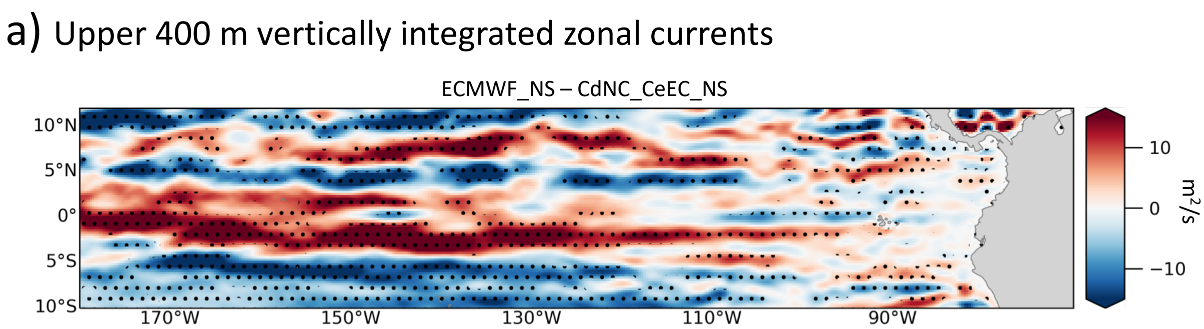
Plot 10 is included in the manuscript in the supplementary material as FigureS5. The following sentence was added in the manuscript (lines 309-313): *"These relations are confirmed along the coast of the Benguela Upwelling System (Figures S4 and S5). During the Benguela upwelling season310 (ONDJ), the enhanced wind stress and negative wind stress curl in ECMWF_NS reinforce the vertical velocity with respect to CdNC_CeEC_NS (Figure S4), resulting in colder surface temperature (see correlation maps Figure S5)."*

11) Line 268. Can you see any changes to the North Equatorial Undercurrent, which is mainly an WSC-driven system (e.g. Sun et al. 2021 and references therein, https://doi.org/10.1016/j.ocemod.2021.101876)

We thank the reviewer for the comment and for suggesting this interesting paper. We calculated the upper 400m vertically integrated zonal currents as Sun et al. 2021. Plot 11 shows the results for ECMWF_NS and CdNC_CdEC_NS experiments, while Plot 12 shows their differences. We noticed differences in the Equatorial Undercurrent more than in the North Equatorial Countercurrent. Kessler et al. (2003) suggested that the strip of positive WSC at the equator is the key to produce the north branch of Equatorial Undercurrent and, indeed, ECMWF_S shows remarkable difference in it with respect to NCAR (see Plot 12 below). This result is certainly interesting, nevertheless, we think that a deeper analysis on the equatorial currents would require further work and it would be something interesting to further investigate it in a future study. The current manuscript is now dense with new information and we prefer to not include more material for ease of reading.



Plot 11: Upper 400m vertically integrated zonal currents for a) *ECMWF_NS and b) CdNC_CeEC_NS.* Hatches display significant values (95% confidence level).

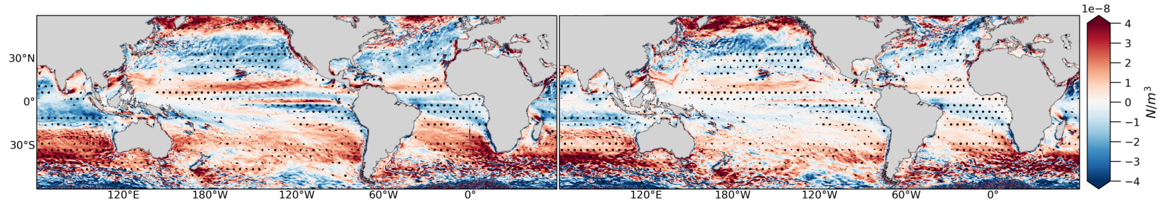


Plot 12: a) Annual mean difference of Upper 400m vertically integrated zonal currents between *ECMWF_NS and CdNC_CeEC_NS. Hatching indicates significant values (95% confidence level).*

12) Line 286. But the equator to 10deg. N difference is still large with COARE (Fig. 2a, right)

We thank the referee for the comment. We plotted the curl differences between the experiments (Plot 13), as also suggested by the Referee#2. This plot is added as Figure 3c in the revised manuscript. Results show that the wind stress curl is stronger in COARE_S than in NCAR in the north equatorial region, but this positive difference is less pronounced than ECMWF - NCAR case. We modified the paragraph in the manuscript (lines 334-338) as follows: "*As regard the equatorial upwelling, the weak increasing of the wind stress in the north equatorial region (e.g. northern equatorial cold front, Figure 3b) compared to NCAR wind stress (Figure 3a), prevents the enhancement of the positive wind stress curl in COARE_S (Figure 3c). Nevertheless, to properly identify the drivers of the pattern in the SST differences between COARE_S and NCAR extra dedicated numerical experiments should be performed.* "

c) WSC



Plot 13: wind stress curl (W SC) between ECMWF_S and NCAR experiments (left) and COARE_S and NCAR experiments (right). Hatching indicates significant values (95% confidence level).

REFERENCES

Blanke, B., & Delecluse, P. (1993). Variability of the tropical Atlantic Ocean simulated by a general circulation model with two different mixed-layer physics. *Journal of Physical Oceanography*, *23*(7), 1363-1388.

Bonino, G., Masina, S., Iovino, D., Storto, A., & Tsujino, H. (2019). Eastern Boundary Upwelling Systems response to different atmospheric forcing in a global eddy-permitting ocean model. *Journal of Marine Systems*, *197*, 103178.

Merchant, C.J., Embury, O., Bulgin, C.E., Block, T., Corlett, G., Fiedler, E., Good, S.A., Mittaz, J., Rayner, N., Berry, D. and Eastwood, S., 2019. Satellite-based time-series of sea-surface temperature since 1981 for climate applications. Nat. Sci. Data. 6:223. doi: 10.1038/s41597-019-0236-x

Kessler WS, Johnson GC, Moore DW (2003) Sverdrup and nonlinear dynamics of the Pacific equatorial currents. J Phys Oceanogr 33: 994–1008