We sincerely appreciate the reviewer for her/his constructive comments on the manuscript. Our responses are listed as follows in blue. Text are revised accordingly.

Review from Referee #2

Here are my comments.

Lines 150, 152 and 157 Van -> Van Roekel

Response: As suggested, the text is revised.

Line 169, section 2.3.2: the Stokes drift should also be used in the advection of any tracer, including temperature and also in the calculation of the vertical velocity in difference/convergence term NEMO4.

Response: We agree. According to the work of Couvelard et al. (2020), we calculate the Stokes drift profile and add it to the corresponding advection terms and convergence terms in MOM4. The text is revised, and the results are updated.

Line 179, section 2.4: there is an inconsistency in considering the impact of the surface current and of the surface Stokes drift on the momentum flux in the atmosphere model, but not in the wave model. With ST4 (and ST3), the surface momentum balance is re-evaluated in order to determine the friction velocity that is then used as part of the source terms calculation, and hence the evolution of the wave field. To be consistent, WW3 should be forced not with the absolute 10m wind, with the relative 10m wind with respect to the surface current.

Response: As suggested, the relative 10-m wind with respect to the surface current is applied in the wave model as well. Figure R1 and text are updated accordingly.

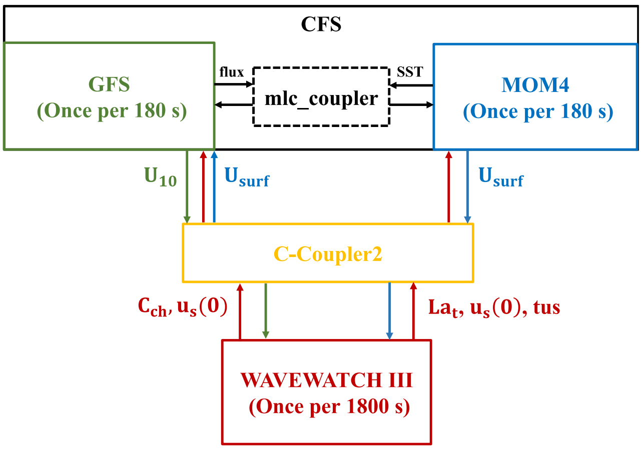

Figure R1. A schematic diagram of the atmosphere-ocean-wave coupled modeling system. The arrows indicate the coupled variables that are passed between the model components. In the diagram, C_{ch}, La_t, $u_s(0)$, tus, U_{10}, and U_{surf} are Charnock parameter (red arrows), turbulent Langmuir number (red arrows), surface Stokes drift velocity (red arrows), Stokes drift transport (red arrows), 10-m wind (green arrows) and surface current (blue arrows), respectively.

Line 216: as far as I can understand from the text, the ST4-Fan scheme is used for z0 in the atmospheric model only, and not in WW3. This is not consistent and should be made clearer that z0 *and hence u* inside WW3 will still be based on a Charnock determined from a modified version of Janssen wave induced stress (Ardhuin et al. 2010).*

Response: To clarify, we revise Section 2.5.2 to address that the new z0 scheme (previously ST4-Fan and now revised to ST4-M04) is used in CFS, while the z0 in WW3 is calculated by the ST4 source term (Ardhuin et al., 2010) with the method of wave-induced kinematic stress (Janssen 1989, 1991).

Also, WW3 can be run with a cap on z0, and hence on the Charnock it could return to the atmosphere. See z0max in table 2.6 in the WW3 manual. It is indeed set to a large value for TEST473f, however, TEST500 has z0max=0.002 for instance. I believe, it is worth mentioning as ST4-Fan is not the only way to limit Charnock for high winds. One should also mention some recent developments on modifying ST4 for high winds (Bidlot et al. 2020 and Li et al. 2021), without a very awkward parameterization of Fan et al. (sorry I notice that it is mentioned later (line 488), but it might too late).

Response: As suggested, the introduction of recent developments on modifying ST4 for high winds (Bidlot et al. 2020 and Li et al. 2021) is moved from Section 5 to Section 2.5.2.

I have serious problem with ST4-Fan. It will indeed limit Charnock for large winds, but from figure S2, it does not seem to make sense that Charnock is largest in the Tropics. The Charnock parameter was introduced to represent the impact of waves on the momentum transfer at the sea surface. It was recognized that young wind-sea should extract more momentum than older more mature old sea. So why is Charnock largest in the Tropics where the sea state should be dominated by old wind-sea and swell?

Response: We agree that the calculation of Charnock parameter in ST4-Fan is problematic for the sea state dominated by old wind-sea and swell. To solve this problem, we revise ST4-Fan based on Moon et al. (2004). According to Fan et al. (2012), the equations of Charnock parameter are derived based on the observations in Moon et al. (2004). In Moon et al. (2004), the Charnock parameter decreases with the increase of wave age at low-middle winds (<30 m/s), but levels off or increases at high winds (>30 m/s; Fig. 3 of Moon et al. 2004). Moon et al. (2004) proposed Eqn. R1 to estimate the Charnock parameter, and gave different values of a and b every 5 m/s ranging from 10 m/s to 50 m/s (Table 1 of Moon et al. 2004). Based on this, Fan et al. (2012) proposed Eqn. R2 to calculate a and b. Because b is always positive (Eqn. R2), the Charnock from Eqn. R1 increases with wave age even at low wind speed, which generates large Charnock in the tropics.

$$C_{ch} = a\left(\frac{c_{pi}}{u_*}\right)^b, \quad \text{(R1)}$$

$$a = \frac{0.023}{1.0568^{U_{10}}}, b = 0.012U_{10}, \quad \text{(R2)}$$

Therefore, we have re-derived the relationship (Eqn. R3) between a/b and 10-m wind speed U_{10} by fitting the values in Table 1 of Moon et al. (2004) for U_{10}>15 m/s,

$$a = \frac{1}{0.1477{U_{10}}^2 - 0.7395U_{10} - 10.9995},$$
$$b = 1.5661E^{-5}{U_{10}}^3 - 0.002{U_{10}}^2 + 0.1017U_{10} - 1.6182. \quad \text{(R3)}$$

Because the observations in Moon et al. (2004) are obtained under tropical cyclones, there is no reliable data for wind $\leq$ 15 m/s. In Eqn. R3, b is negative (positive) from relatively small (large) wind speed. So Eqn. R3 is used for U_{10}>15 m/s, whereas the original Charnock relationship of WW3 ST4 scheme (Janssen 1989, 1991) is used for $U_{10} \leq$15 m/s. The revised parameterization is called ST4-M04. The comparison of z_0 among ST4-M04, ST4-Fan, ST4 and GFS is shown in Fig. R2, indicating that the new ST4-M04 is close to ST4-Fan at high wind speed and can improve z_0 at low wind speed. Consequently, the Charnock parameter in ST4-M04 is low in the Tropics as expected (Fig. R3).

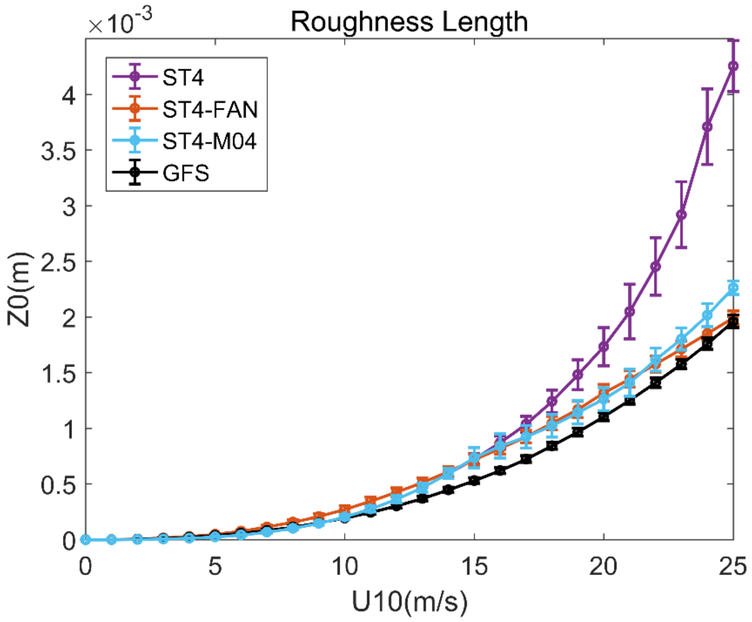

Figure R2. Relationships between momentum roughness length z_0 (m) in the coupled system and 10-m wind speed (m/s); error bars indicate twice the standard deviations for each point.

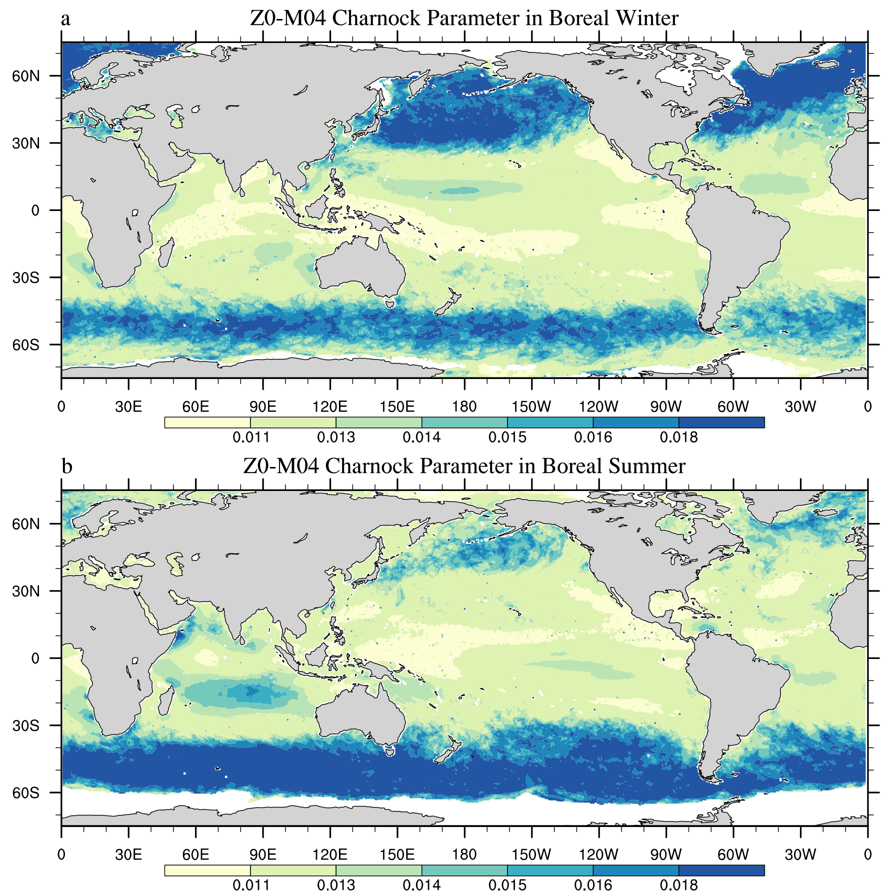

Figure R3. The Charnock parameter C_{ch} obtained by ST4-M04 in boreal winter (a) and summer (b).

Line 231: it is more than just the fetch, but also the development stage of the sea state (young sea, old sea…)

Response: As suggested, the text is revised to "Since the z0 is determined only by wind-sea conditions in ST4 and ST4-M04 scheme, the STD at a given wind speed is owing to variations in wind fetch and development stage of sea state. The reduced STDs in ST4-M04 scheme, compared to ST4, imply less sensitivity of z0 to fetch and sea state".

Line 233, before section 2.6: Could you say what is done in WW3 when sea ice is present. It seems that there is a large impact near or within the sea ice, so it is important to know how the different fields were produced by WW3. Quite often, WW3 is set-up so that the wave spectra are reset to 0 every time step for all areas with a sea ice cover above a certain threshold. If it is the case, this will mean that the estimate for Stokes drift and wave age would be those for very young windsea. It can then be debated whether these estimates are correct, as it is known that in the presence of sea ice, the high frequency waves are heavily damped.

Response: In WW3, the sea ice concentrations are from CFSR data (Saha et al., 2014). The ice blocking IC0 source term with the critical ice concentration of 50% is applied. So for ice concentration in the range of 0-50%, the wave spectra are not zero, while for ice concentration larger than 50% the wave spectra are zero. Indeed, the estimate for Stokes drift and wave age would be those for young wind-sea, inconsistent with the phenomenon that high-frequency waves are severely damped in the presence of sea ice. To diminish the influence, since the estimate of roughness and fluxes in the presence of sea ice are different with those in open oceans in CFS, we turn off the coupling from WW3 to CFS to avoid any conflicts with sea ice. Therefore, the effect of wave-state in these areas on air-sea flux is not considered. Text is revised to clarify.

Line 256 and table S4: NDBC website reports its own buoy data, as well as buoy data from other NOAA agencies, such as the data from the TAO array, and a few other buoy data providers along the US. But it also reports surface observations from many coastal stations (i.e. on land, piers or towers) that are definitely not buoy. They also report observations from the oil and gas industry in the Gulf of Mexico. In table S4, anything that does not have a 5-digit identifier is probably not a buoy, and the oil and gas data from the Gulf of Mexico have an identifier like 423xx, 428xx. All those non buoy data should not be use in this analysis.

Response: As suggested, all non-buoy data are removed. The text is revised accordingly.

Line 261: section 4.1: I am surprised by the lack of impact of the FLUX experiment on the SST, in particular in the equatorial Pacific. My experience is that including the winds relative the surface current in the momentum flux (surface stress) used to force the ocean model has quite an impact on the SST around the equatorial Pacific. To be sure, could you confirm, as indicated that MOM4 is indeed forced with the surface stress as shown in equation (7).

Response: We checked the code as suggested, and the MOM4 is indeed forced with the surface stress as Eqn. 7 in the text. Previously, we showed the SST changes by a percentage relative differences (PRD $= \frac{|\widehat{y_s}-y|-|\widehat{y_c}-y|}{|y|} \times 100\%$, where y is OISST, $\widehat{y_c}$ is simulated SST in CTRL and $\widehat{y_s}$ is simulated SST in other experiments) in Fig. 3&4. Since the absolute value of OISST on the denominator is relatively large, the PRD looks low in lower latitude. In the revised text, all the maps of PRD are replaced by SST differences between sensitive experiments and CTRL (Fig. R4b&d). The SST differences between FLUX and CTRL around the equatorial Pacific are manifest.

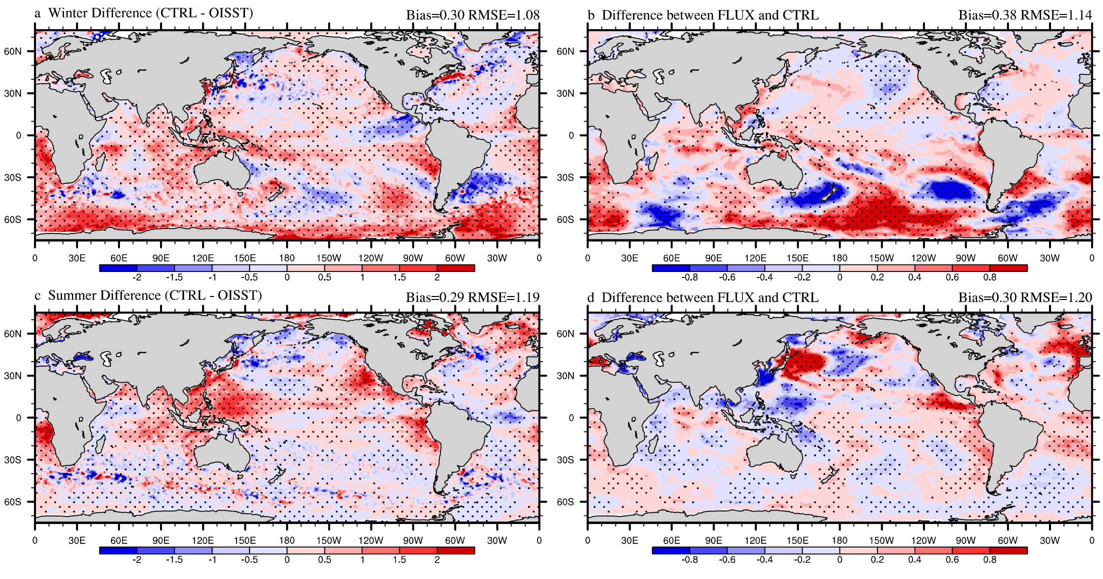

Figure R4. The 53-day average SST (°C) bias in CTRL and the SST difference between FLUX and CTRL: (a) the SST bias between CTRL and OISST (CTRL minus OISST) in Jan-Feb, 2017, (b) the difference between FLUX and CTRL (FLUX minus CTRL) in Jan-Feb, 2017, (c) the SST bias between CTRL and OISST (CTRL minus OISST) in Aug-Sep, 2018, and (d) the difference between FLUX and CTRL (FLUX minus CTRL) in Aug-Sep, 2018. Dotted areas are statistically significant at 95% confidence level.

Line 474: it not true, the SST bias in the Tropics was not reduced due to drag from swell. Breivik et al. (2015) considered the impact of sea state dependent Charnock, using the formulation from Janssen on the air-side surface stress, which was modulated further considering the momentum flux balance between wind input and wave breaking to determine the ocean-side stress that is then used to force the ocean. Moreover, the impact of wave breaking was also considered as input to the upper ocean mixing scheme (TKE). All these effects had an impact on the SST. What was not discussed in Breivik et al. (2015) is the impact of surface currents on the SST response in the Tropical Pacific as that effect has already been introduced in the ECMWF system years before. As I mentioned earlier, I am surprised by the lack of sensitivity on the CFS system to surface currents in the Tropics.
Response: As suggested, the text is revised to "In Breivik et al. (2015), considered the surface waves effects including wave-related Charnock parameter, modification of ocean-side stress, wave dissipation-related turbulent kinetic energy flux, Stokes-Coriolis force and Langmuir mixing, the bias of SST simulation in the tropics is reduced". Figure R4 shows that surface current and Stokes drift can influence SST in tropics. The values can be up to more than 0.5℃, comparable with the effect of modified ocean-side stress in Breivik et al. (2015; Fig.1). Text is revised to clarify the effects of surface current and Stokes drift on SST.

References
Ardhuin, F., Rogers, E., Babanin, A. V., Filipot, J., Magne, R., Roland, A., Der Westhuysen, A. V., Queffeulou, P., Lefevre, J. M., and Aouf, L.: Semiempirical Dissipation Source Functions for Ocean Waves. Part I: Definition, Calibration, and Validation, Journal of Physical Oceanography, 40, 1917-1941, http://dx.doi.org/10.1175/2010JPO4324.1, 2010.
Breivik, O., Mogensen, K., Bidlot, J., Balmaseda, M., and Janssen, P. A. E. M.: Surface wave effects in the NEMO ocean model: Forced and coupled experiments, Journal of Geophysical Research, 120, 2973-2992, http://dx.doi.org/10.1002/2014JC010565, 2015.
Couvelard, X., Lemarié, F., Samson, G., Redelsperger, J.-L., Ardhuin, F., Benshila, R., and Madec, G.: Development of a two-way-coupled ocean-wave model: assessment on a global NEMO(v3.6)-WW3(v6.02) coupled configuration, Geosci. Model Dev., 13, 3067-3090, https://doi.org/10.5194/gmd-13-3067-2020, 2020.
Fan, Y., Lin, S., Held, I. M., Yu, Z., and Tolman, H. L.: Global Ocean Surface Wave Simulation Using a Coupled Atmosphere–Wave Model, Journal of Climate, 25, 6233-6252, http://dx.doi.org/10.1175/JCLI-D-11-00621.1, 2012.
Janssen, P. A. E. M.: The Quasi-linear theory of wind wave generation applied to wave forecasting. J. Phys. Oceanogr. 21, 1631-1642. https://doi.org/10.1175/1520-0485(1991)021<1631:QLTOWW>2.0.CO;2, 1991.
Janssen, P. A. E. M.: Wave-induced stress and the drag of air flow over sea waves. J. Phys. Oceanogr. 19(6), 745-754. https://doi.org/10.1175/1520-1310 0485(1989)019<0745:WISATD>2.0.CO;2, 1989.
Moon, I., Ginis, I., and Hara, T.: Effect of surface waves on Charnock coefficient under tropical cyclones, Geophysical Research Letters, 31, http://dx.doi.org/10.1029/2004GL020988, 2004.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y., Chuang, H., and Iredell, M. D.: The NCEP Climate Forecast System Version 2, Journal of Climate, 27, 2185-2208, http://dx.doi.org/10.1175/JCLI-D-12-00823.1, 2014.