Reviewer's comments are in black. Authors responses are in blue. *Changes to the text are in italic.*

The authors provide a new perspective on the contribution of reduced shipping sulfate emissions to the anomalous observed warming in 2023. While short-term climate forcers are mostly from continental emission sources, this work focuses on a maritime source. Further explanation is needed to make the attribution analysis more comprehensive and accessible to a broader audience. I have several concerns I would like the authors to consider:

We thank the reviewer for their comments, which we address below.

1. Temperature response to shipping emission reduction contributes considerably to observed temperature anomaly in 2023 (Fig. 2). Given that the CMIP6 SSP 3-7.0 scenario has been suggested to fail to capture East and South Asian anthropogenic aerosol emissions since 2006 (Zhang et al., 2019; Ramachandran et al., 2020), and has shown to have both local and remote impact (Wang et al., 2021; Xie et al., 2023). CESM2 LENS driven by SSP 3-7.0 underestimates the observed both net short-wave radiative flux and net short-wave minus long-wave radiative flux at the top of atmosphere since 2015 (Fig. 1) and, however, CESM2 LENS seems to capture the observed temperature anomaly in 2022 in Fig.2. Considering the low bias in radiative flux at the top of atmosphere in CESM2 LENS, it suggests other factors, such as internal variability, also contribute. Additionally, Fig. 3 shows anomalous warming in the tropical Pacific resembling an El Niño pattern. I recommend the authors discuss the impact of bias in emission scenarios and the potential impact of other factors.

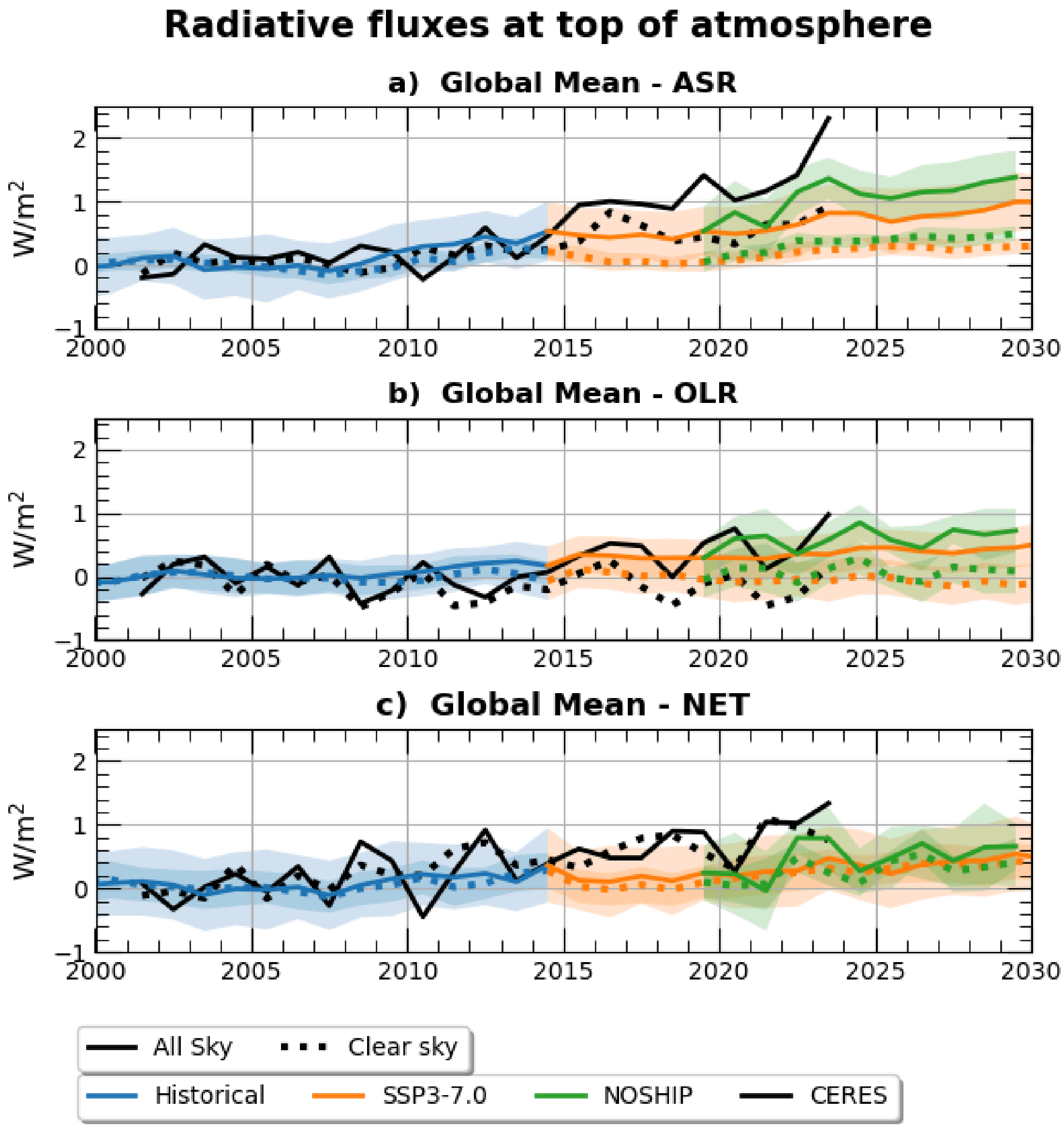
   Our discussion has been updated, given also the input from reviewer 1. We have added some comments in Section 2.1 related to the useful references the reviewer provided: "*Our results are also consistent with previous studies in which aerosol emissions have already been shown to present a source of bias for CESM2 compared to observations (Zhang et al., 2019; Ramachandran et al., 2020) in other regions.*"

2. The authors estimate an increase in radiative forcing of 0.2 W/m² due to reduced shipping sulfate emissions, and aerosol-cloud interaction plays an important role. This is a significant magnitude, and I would expect a strong surface temperature response. There is no further explanation on how cloud response indirectly contributes to change in radiative forcing. The authors conclude that there is a three-year lag in surface temperature response due to ocean. I recommend showing results starting from 2020 and including the range of ensemble members in Fig 2. Additionally, a spatial map showing the geographic pattern of surface temperature

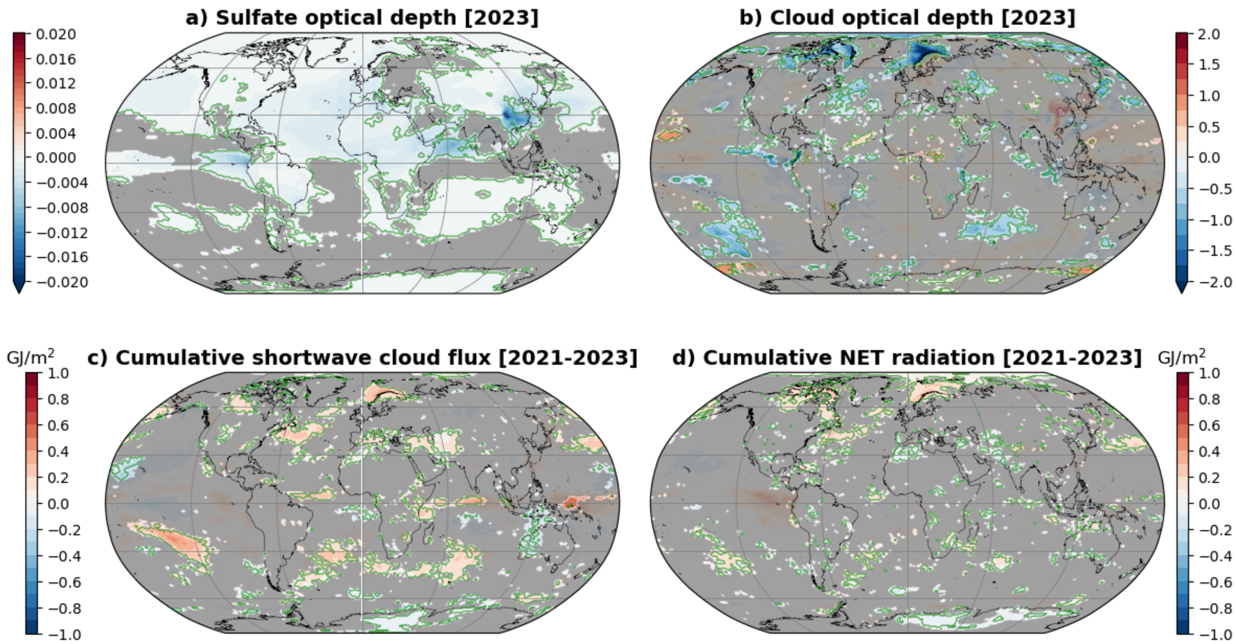
response, as well as full-sky radiative forcing, would be helpful, especially as shipping reductions are most pronounced in the North Atlantic.

In order to clarify the reviewer's point, we updated Figure A2 to highlight the difference between the forcing in All and Clear sky condition: we believe these analyses, in both CESM2 and CERES, depict clearly the contribution deriving from cloud-induced changes. In Fig. 3, we already show surface temperatures and the main component of radiative forcing of interest (in that case, we integrate absorbed solar radiation over the 3 years as a measure of absorbed energy, which more closely ties to oceanic temperatures). We have now added in the supplementary material (new Figure A3) also a map of aerosol and cloud visible aerosol optical depth anomalies in 2023 (note the difference in scale), and the cumulative SW cloud flux and net radiation anomalies, which we believe better support Figure 3.

For his other point, we note that we showed results starting in 2020 for both radiative forcing (Fig. 1) and global temperatures (Fig. A3), but we think Fig. 2 is clearer as is. The aim of the figure is to show the likelihood of occurring temperature based on observations with and without the impact of shipping emission estimated from the average and higher temperature anomalies from NOSHIP (that we calculate as the average and the average plus one standard deviation on the ensemble) and not showing the single member contribution with respect to the observations. We elaborated more about the method in the main text to make it clearer.



New Fig. A2: Time series of annual mean deviation from the 2000-2007 period for globally averaged (a) Absorbed Solar Radiation (ASR, defined as incoming minus outgoing shortwave), (b) Outgoing Longwave Radiation (OLR) and (c) ASR minus OLR (NET) radiative flux at the top of the atmosphere (TOA), in all sky (solid lines) and clear sky (dotted lines) conditions. The shaded area for CESM2-LENS2 simulations (Historical, SSP3-7.0, NOSHIP) represents one standard deviation calculated on ensemble members.



New Fig. A3: Maps of changes in sulfate and cloud visible optical depth in 2023 (a and b, respectively), and cumulative shortwave cloud flux and NET radiation over 2021-2023 (c and d, respectively) due to reduction in shipping emissions in CESM2-LENS2. Shaded areas indicate regions where the differences are not statistically significant at the 10% level, green contours indicate regions where the differences are not statistically significant at the 5% level.

I recommend including one or two more observational datasets or reanalyses to account for uncertainty. As ERA5 has shown bias in capturing observed top-of-atmosphere radiative fluxes, it is unclear why its results are included in the main figure and Fig. A5. The authors should clarify this choice. Additionally, the description of the Berkeley dataset is missing. Providing a brief description in the main text would be beneficial.

We added two more temperature datasets, the Global Surface Temperature (NOAAGlobalTemp) and the Met Office Hadley Centre/Climatic Research Unit global surface Temperature (HadCRUT5), which are now in the new Figures A4 and A5. Between January 2015 and May 2024 the new two datasets lie very closely between the lower estimate from Berkeley (black solid line) and higher estimate values from ERA5 (dark brown solid line), as shown in the revised Figure A5, and therefore they do not change our conclusions.

We believe our addition of ERA5 is useful here in comparison to observational datasets, especially as the temperature estimates are quite close to Berkeley and other measurements whereas the forcing presents such a large bias, pointing to the need to better include aerosol representation in reanalyses, which we think fits with the main point of our paper. We did explore the use of other reanalysis products for radiative forcing, but we didn’t find any others that performed better than ERA5, which we also chose based on its widespread use.

We now include a section about the various datasets in the *Methods appendix,* under *Observations and reanalysis*:

“Sulfur shipping emissions are from the Community Emission Data System (CEDS) which provides estimates of emissions of anthropogenic greenhouse gases, reactive gases and aerosols, from 1750 to nowadays, based on existing emission inventories, emission factors, and activity/driver data (Hoesly and Smith, 2024).

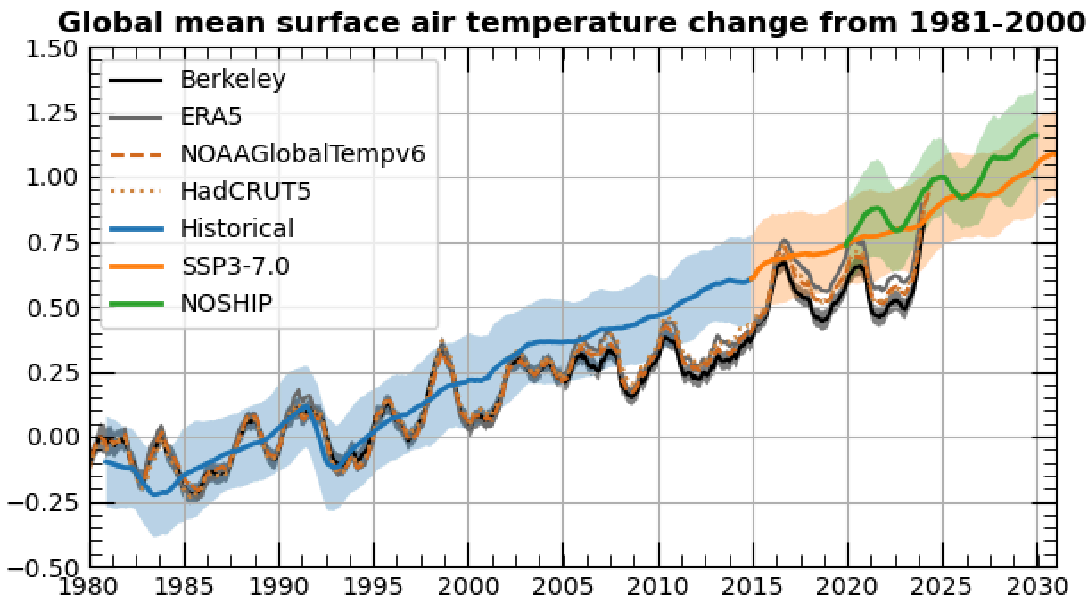
To compare simulated radiative fluxes, we used satellite data from the Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled Top-of-Atmosphere fluxes version 4.2 (CERES_EBAF_Edition4.2, NASA/LARC/SD/ASDC (2023)) and climate reanalysis data from the fifth generation European Centre for Medium-range Weather Forecasts (ECMWF) reanalysis (ERA5, Hersbach et al.).

Simulated surface temperatures are compared with ERA5, Berkeley Earth, Met Office Hadley Centre/Climatic Research Unit global surface temperature data set version 5.0.2.0 (HadCRUT5), and the NOAA Global Surface Temperature Dataset version 6.0 (NOAAGlobalTempv6). For the results in our Figure 2 we tested all datasets but ultimately only showed Berkley, as our conclusions were largely independent of the dataset chosen Fig. A5.

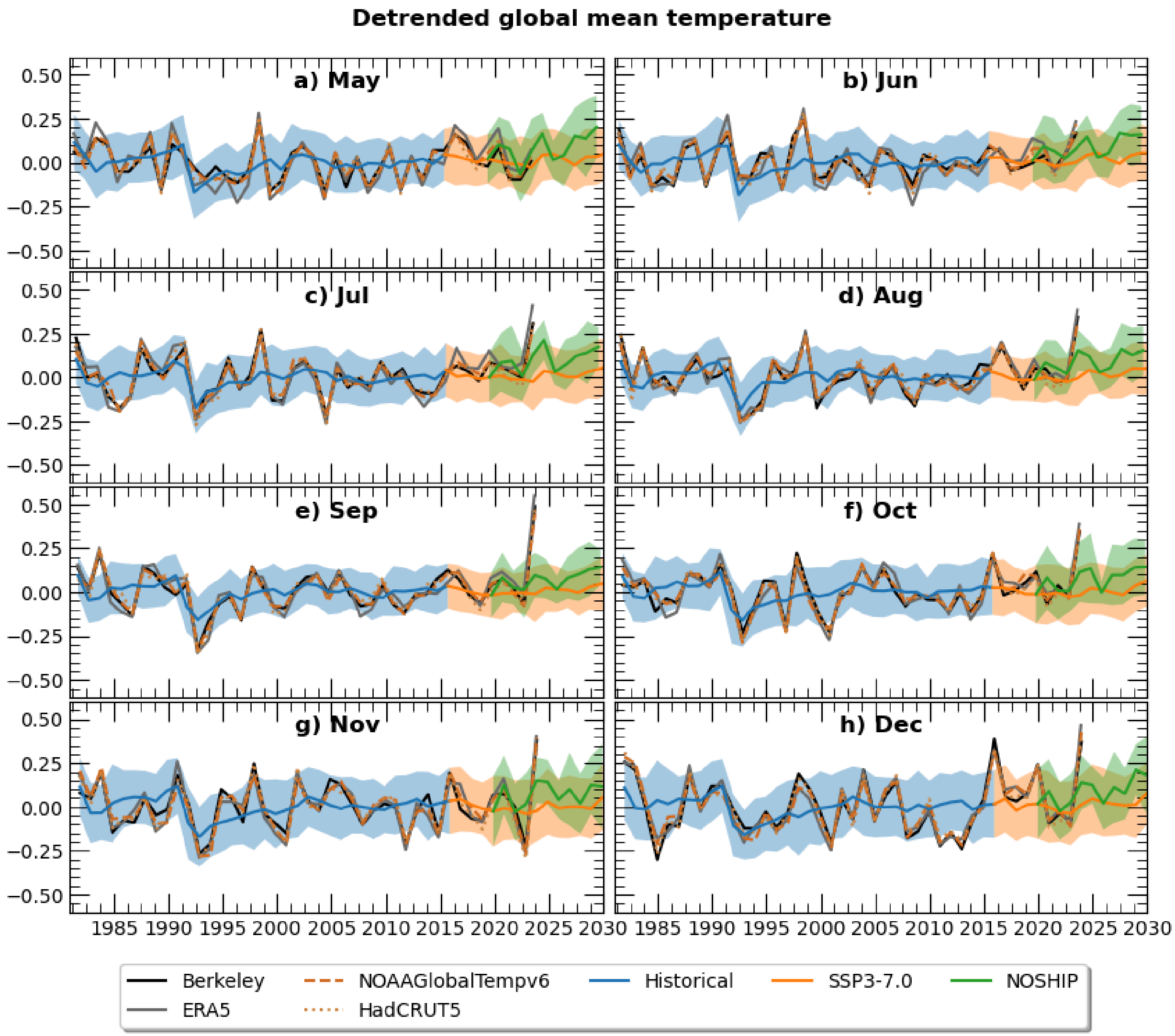
Berkeley Earth Land/Ocean Temperature Record (Rohde and Hausfather, 2020) combines the Berkeley Earth land-surface temperature field with an interpolated version of the Met Office Hadley Centre Sea Surface Temperature dataset version 4.0.0.0 (HadSST4).

HadCRUT5 (Morice et al., 2021) uses a statistical infilling method to integrate sea-surface temperature data from the HadSST4 with land-surface air temperature data from the Climatic Research Unit temperature dataset version 5.0.0.0 (CRUTEM5) .

NOAAGlobalTempv6 (Huang and Zhang) combines the land-ocean surface temperature analysis from the Extended Reconstructed Sea Surface Temperature (ERSSTv5) with land surface air temperature analysis, which are is on the Global Historical Climatology Network-Monthly (GHCN-M) temperature database.”



New Fig A4: Time series of 12-months rolling mean of global mean temperature changes from 1981-2000 for CESM2-LENS2 simulations (Historical, SSP3-7.0, NOSHIP) and observations and reanalysis (Berkeley, ERA5, NOAAGlobalTempv6, HadCRUT5). The shaded area for CESM2-LENS2 simulations represents one standard deviation calculated on ensemble members.



New Fig A5: Time series of detrended global mean temperature changes from the 1981-2000, distinguished for each month of the year, for CESM2-LENS2 simulations (Historical, SSP3-7.0, NOSHIP) and observations and reanalysis (Berkeley, ERA5, NOAAGlobalTempv5, HadCRUT5). The shaded area for CESM2-LENS2 simulations represents one standard deviation calculated on ensemble members.

Reference

Ramachandran, S., Rupakheti, M. and Lawrence, M.G., 2020. Aerosol-induced atmospheric heating rate decreases over South and East Asia as a result of changing content and composition. Scientific Reports, 10(1), p.20091.

Zhang, Q., Zheng, Y., Tong, D., Shao, M., Wang, S., Zhang, Y., Xu, X., Wang, J., He, H., Liu, W. and Ding, Y., 2019. Drivers of improved PM2. 5 air quality in China from 2013 to 2017. Proceedings of the National Academy of Sciences, 116(49), pp.24463-24469.

Xie, Y., Huang, J., Wu, G., Lei, N. and Liu, Y., 2023. Enhanced Asian warming increases Arctic amplification. Environmental Research Letters, 18(3), p.034041.

Wang, Z., Lin, L., Xu, Y., Che, H., Zhang, X., Zhang, H., Dong, W., Wang, C., Gui, K. and Xie, B., 2021. Incorrect Asian aerosols affecting the attribution and projection of regional climate change in CMIP6 models. npj Climate and Atmospheric Science, 4(1), p.2.