Supporting Information for "Model biases in the atmosphere-ocean partitioning of poleward heat transport are persistent across three CMIP generations"

A. Donohoe[1,3], R. Fajber[2], T. Cox [3], K.C. Armour[3,4], D.S. Battisti [3]and

G.H. Roe [4]

[1]Polar Science Center, Applied Physics Laboratory

University of Washington

Seattle, Washington 98195, USA.

[2] Department of Atmospheric and Oceanic Sciences, McGill University

[3] Department of Atmospheric and Oceanic Sciences, University of Washington

[4]School of Oceanography, University of Washington

[5]Department of Earth and Space Sciences, University of Washington

September 29, 2023, 2:12am

1. Coupled models analyzed

We analyze pre-industrial (PI) control simulations in coupled climate models that represent the equilibrium response to fixed green house gas concentrations. We analyze 66 model simulations from three different generations of the coupled climate model inter-comparison project (CMIP): CMIP3 (Meehl et al., 2007) which ran from 2005-2006 (14 simulations); CMIP5 (Taylor et al., 2012) which ran from 2010-2014 (20 simulations) and; CMIP6 (Eyring et al., 2016) which ran from 2014-2020 (32 simulations). All calculations discussed here use annual mean long term climatologies calculated from the last 50 of available years of the PI simulation. We additionally analyze 12 CMIP5 historical simulations to evaluate the differences between the MHT/AHT/OHT in the PI simulations and historical era which may impact the observational-model comparison.

2. Observational datasets used

2.1. Top of atmosphere radiation

Observational MHT is primarily calculated using satellite derived RAD_{TOA} from the Clouds and Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) product version 4.0 (Loeb & Coauthors, 2018). This product is a gridded retrieval of net longwave and shortwave radiation at the TOA derived from instruments on the Aqua and Terra satellites. The retrieved RAD_{TOA} is subsequently adjusted to satisfy Earth's global energy imbalance of 0.71 ± 0.10 W m^{-2} constrained by long-term changes in global ocean heat content changes (Johnson et al., 2016). This adjustment is accomplished via modification of uncertain parameters in the retrieval algorithm (e.g. radiative transfer model) used to produce the gridded product and primarily involves adjustment

September 29, 2023, 2:12am

of the absolute calibration of the shortwave and longwave fluxes which have a combined uncertainty (95% confidence interval) of 4.2 W m^{-2} (Loeb et al., 2009). We also analyze unadjusted gridded CERES data from single scanner footprints (SSF) to diagnose the impact of the EBAF adjustment on MHT. The average of four (FM1 and FM2 on Terra and FM3 and FM4 and Aqua) SSF RAD_{TOA} data sets is analyzed. The climatological average RAD_{TOA} over the 3/2001-12/2018 period is used to calculate MHT from all CERES products with the exception of the Aqua SSF data which begin in 7/2002. We also use RAD_{TOA} from the Earth Radiation Budget Experiment (ERBE Barkstrom & Hall, 1982). Climatological ERBE RAD_{TOA} over the 11/1984-3/1990 period is used to calculate an additional observational estimate of MHT.

Given that the global mean net TOA radiative imbalance ranges from 7.0 W m^{-2} (3.6 PW globally) in the unadjusted CERES dataset to 4.9 W m^{-2} (2.5 PW) in ERBE dataset (see table 1 of Loeb et al., 2009) to 0.7 W m^{-2} (0.4 PW) in the CERES EBAF dataset (Johnson et al., 2016), it is perhaps surprising that the calculated MHT only differs by of order 0.1 PW across these data sets. We interpret this result to imply that the largest differences between the TOA radiation data sets is the absolute calibration (addition of a spatially invariant constant) of the shortwave and longwave fluxes which are the stated largest source of uncertainty in the data sets (Loeb & Coauthors, 2018) and make no impact on the derived MHT calculated here via removal of the global mean value. Stated otherwise, the spatial gradients in net TOA radiation are less uncertain (or at least consistent between datasets) as compared to the global means.

September 29, 2023, 2:12am

2.2. Atmospheric reanalysis

AHT is derived from the time average of the vertical and zonal integral of the meridional flux of moist static energy calculated from high spatial-temporal resolution atmospheric reanalysis. Our analysis primarily focuses on AHT estimates calculated from the European Center for Medium Range Forecasting's (ECMWF) ERA5 reanalysis (Hersbach et al., 2020). We use instantaneous 6-hourly ERA5 data on 37 pressure levels and a horizontal resolution of 0.5°. Additional AHT calculations are performed and analyzed using two other sets of 6-hourly instantaneous atmospheric reanalysis: 1. ECMWF's ERA-interim reanalysis which has 37 vertical levels and horizontal resolution of 1.5° (Dee et al., 2011) and; 2. the National Center for Atmospheric Research's (NCEP) reanalysis which has 17 vertical levels and a horizontal spectral resolution of T62.

The following four-dimensional (pressure level, latitude,longitude, time) atmospheric fields are used to calculate AHT; meridional velocity (V), temperature (T), specific humidity (Q) and geopotential height (Z). The climatological surface pressure is used to set the bounds of the vertical integration. AHT calculations are preformed for each month then the results are averaged to produce a long-term average climatology. AHT climatologies are computed over the corresponding time period of the radiation data: 3/2001-12/2018 when used in conjunction with CERES data and 11/1984-2/1990 when used in conjunction with ERBE data.

3. Estimating the impact of ocean energy storage on 'implied' OHT

We first calculate the latitudinal structure of the observed long-term trend in ocean heat storage (STORAGE) over the CERES period (2000-2018) from potential temperature data in the UK Hadley Center EN4 objective ocean analysis (Good et al., 2013).

STORAGE is equal to the linear trend in zonal-mean, vertically (mass-weighted) integrated (deseasonalized) ocean potential temperature. The result is the rate of ocean heat uptake (STORAGE) in W m^{-2} at each latitude averaged over the CERES era. We convert this to an implied OHT due to ocean heat storage (OHT$_{STORAGE}$) by spatially integrating the local departure STORAGE from the global mean (indicated by an *) over the polar cap:

$$OHT(\Theta)_{STORAGE} = 2\pi a^2 \int_{\Theta}^{90} -STORAGE^* cos(\theta)d\theta. \qquad (1)$$

OHT$_{STORAGE}$ is the 'implied' OHT that would be calculated from the surface heat fluxes needed to balance the local storage *in the absence of lateral ocean transport*. We remove the OHT$_{STOARGE}$ from the 'implied' observational OHT (=MHT-AHT) to isolate the 'dynamic' OHT that would need to be transported laterally in the ocean to balance the ocean energy budget (the sum of STORAGE and energy lost from the surface of the ocean to the atmosphere). If ocean heat uptake is preferentially in the high latitudes (as is observed), the associated downward extratropical surface fluxes would be diagnosed as an equatorward 'implied' OHT and our observational based estimate of poleward OHT from the inferred surface fluxes would be biased *low* relative to an equilibrium climate system with no STORAGE. Thus, the observed high latitude ocean heat uptake *reduces* our observational estimate of OHT and therefore the model biases toward too little poleward OHT are larger in magnitude than reported here even if the magnitude of ocean heat uptake was underestimated by EN4.

Consistent with the reasoning above, model biases toward too little OHT (relative to observations) are stronger in magnitude when comparing historical simulations to (historical) observations than found in the present work which compares pre-industrial (PI)

simulations with (historical) observations. Historical simulations have slightly weaker poleward OHT into the Southern Ocean compared to their PI counterparts (c.f. the dashed and solid lines in Supplemental Fig. 3) – which is consistent with the expectations discussed above based on preferential STORAGE in the Southern Ocean – and enhanced poleward AHT in the SH as one would expect from down-gradient energy transport under delayed Southern Ocean warming (Armour et al., 2019). In addition to the differences in the AHT/OHT partitioning between the pre-industrial and historical simulations being small in magnitude (relative to the model biases) these results suggest that the model bias toward too much poleward AHT and too little poleward OHT in the SH would be larger in magnitude if observations over the historical period were compared to the historical (as opposed to PI) simulations.

4. Impact of spatial resolution on calculated AHT

Given that the ERA5 reanalysis is the highest spatial resolution considered here and produces the largest poleward AHT, the reader may be suspicious of whether the reanalysis are of sufficient spatial and temporal resolution (on the model output grid) to capture the processes responsible for AHT. We address the potential limitation of the 6-hourly instantaneous temporal resolution of the data first. Instantaneous data does *not* alias the variance (or co-variance) at any frequency with the exception of the discrete harmonics of the sampling period (periods of 6 hours, 3 hours, 1.5 hours, etc) which should be negligible in a continuous spectra. To test this conclusion, we sub-sampled random (white noise) 1 minute data at 6 hourly intervals and found the variance was reduced by less than 0.01% over 100,000 Monte-Carlo realizations. To evaluate the potential limitation of the horizontal resolution of the reanalysis, we calculate the cross-spectra of meridional velocity

and temperature/humidity at 40N, 700 hPa during DJF, the location and season of global maximum climatological poleward AHT (supplemental Fig. 3). Both moist and dry AHT are primarily accomplished by wavenumbers less than 15 with negligible contributions from wavenumbers greater than 90 (corresponding to the smallest resolved wave at 2° longitude grid spacing). Therefore, reducing the resolution of the reanalysis from 0.5 degrees to 2 degrees is equivalent to spectrally truncating the co-spectra at wavenumber 90 which results in a loss of covariance (AHT) of 0.009 % for the dry AHT and 0.021 % for the moisture transport. Stated otherwise, the enhanced horizontal resolution of the ERA5 reanalysis (relative to the resolution of the NCEP reanalysis) makes a negligible contribution to the derived AHT. This analysis does not preclude the possibility that spatial structures smaller than the 0.5° resolution of the ERA5 reanalysis contribute to AHT but does suggest that the enhanced resolution of the ERA5 reanalysis relative to the NCEP reanalysis makes a negligible contribution to the calculated AHT. This conclusion is consistent with the near equivalence of two different AHT calculations in the NCEP CESM simulation shown in Section 2.3; the AHT calculated (dynamically) from the vertical and zonal integral of the product of meridional velocity and temperature/humidity on the 1.25° and 30 vertical level output grid matches that inferred from (energy conservation) of TOA radiation and surface fluxes (Fig. 1 of main text).

References

Armour, K., Siler, N., Donohoe, A., & Roe, G. (2019). Meridional atmospheric heat transport constrained by energetics and mediated by large-scale diffusion. *J. Climate*, *32*(12), 3655-3680. doi: 10.1175/JCLI-D-18-0563.1

Barkstrom, B. R., & Hall, J. B. (1982). Earth radiation budget experiment (erbe)-an

overview. *J. Energy*, *6*, 141-146.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., P., K., ...
van de Berg, A. B. (2011). The ERA-Interim reanalysis: Configu-
ration and performance of the data assimilation system. *Quart. J. Roy.
Meteor. Soc.*, *137*, 553-597. Retrieved from `https://www.ecmwf.int/en/`
`forecasts/datasets/reanalysis-datasets/era-interim` ([accessed 05-January-
2018, https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim])

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor,
K. E. (2016). Overview of the coupled model intercomparison project phase 6 (cmip6)
experimental design and organization. *Geoscientific Model Development (Online)*, *9*(5).
doi: 10.5194/gmd-9-1937-2016

Good, S. A., Martin, M., & Rayner, N. (2013). En4: quality controlled ocean temperature
and salinity profiles and monthly objective analyses with uncertainty estimates. *J.
Geophys. Res.-Oceans*, *118*, 6704-6716.

Hersbach, H. B. B., Berrisford, P., Hirahara, S., Horanyi, A., Munoz?Sabater, J., Nicolas,
J., & Peubey, C. (2020). The era5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*. doi:
10.1002/qj.3803

Johnson, G., Lyman, J., & Loeb, N. (2016). Improving estimates of Earth's energy
imbalance. *Nat. Clim. Chang.*, *6*(7), 639-640.

Loeb, N. G., & Coauthors. (2018). Clouds and the Earth's radiant energy system
(CERES) energy balanced and filled (EBAF) top-of-atmosphere (TOA) edition 4.0 data
product. *J. Climate*, *31*(2), 895-918.

Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., ...

Wong, T. (2009). Towards optimal closure of the Earth's top-of-atmosphere radiation budget. *J. Climate*, *22*, 748-766.

Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., . . . Taylor, K. E. (2007). The WCRP CMIP3 multi-model dataset: A new era in climate change research. *Bull. Amer. Meteor. Soc.*, *88*, 1383-1394.

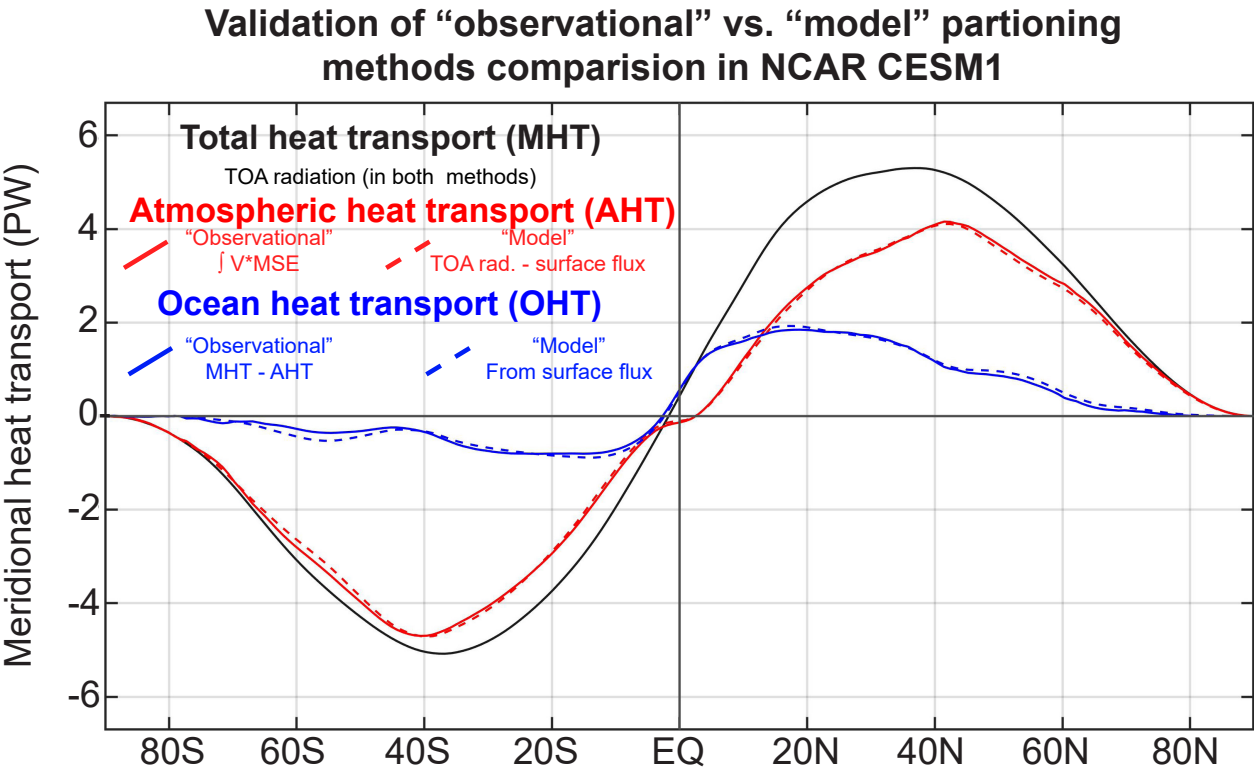Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, *93*, 485-498.

Validation of "observational" vs. "model" partioning methods comparision in NCAR CESM1

Figure S1. Comparison of the MHT/AHT/OHT partitioning method used for the observations versus that used for the models in an NCAR CESM1 simulation in which the atmospheric fields used to calculate AHT were exported akin to the atmospheric reanalysis. MHT (black) is calculated from the TOA radiation integrated over the polar cap in both methods. AHT (red) is calculated from the time averaged vertical and zonal integral of the product of atmospheric MSE and meridional velocity in the observational approach (solid) and from the spatial integral over the polar cap of TOa radiation minus the surface flux in the model approach (dashed). OHT (blue) is calculated from the residual of MHT and AHT in the observational methodology (solid) and from the spatial integral over the polar cap of the surface heat flux in the model methodology (dashed).
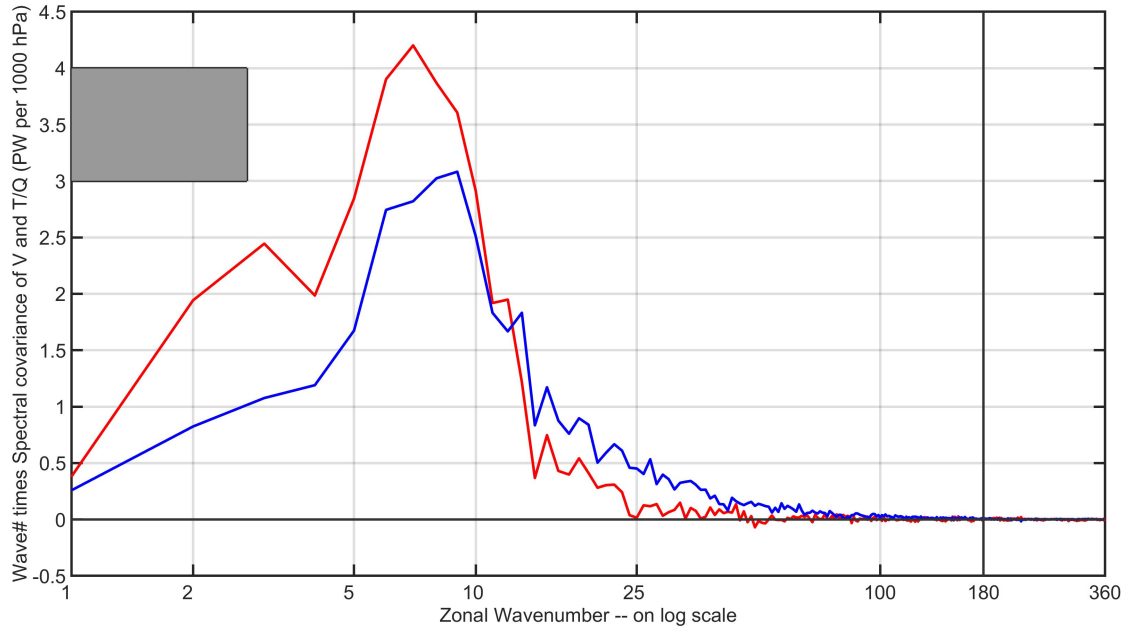
Figure S2. The spectra of atmospheric heat transport at 40N and 700 hPa. The red line shows the spectral co-variance of meridional velocity and temperature (time the specific heat of dry air) and the blue line shows the spectral covariance of meridional velocity and specific humidity (times the latent heat of vaporization of water). The co-spectra are calculated from the product of the spectral power of meridional velocity and temperature/humidity at each instant times the cosine of the spatial phase (wavenumber specific) then time averaged. The wavenumbers on the x-axis are presented on a log scale such that the independent spectral realizations are more densely packed on the right hand side of the plot and the spectral co-variances on the y-axis are multiplied by wavenumber in order to preserve the interpretation of the area under the curve representing the heat transport. The gray shaded box shown an area equal to one PW of zonally and vertically integrated AHT if the spectral co-variance at 700 hPa was realized throughout the atmospheric column. The vertical black line shows the spectral truncation of 4 degrees longitude grid spacing.

September 29, 2023, 2:12am

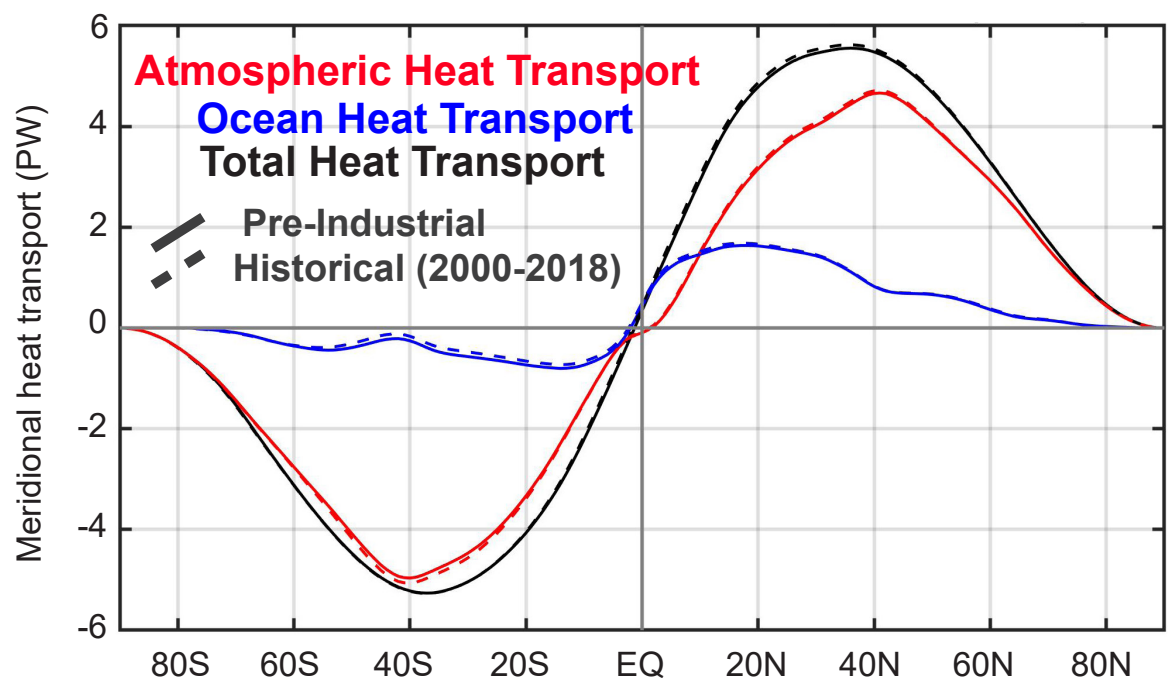# Historical versus pre-industrial MHT in CMIP5

Figure S3. Comparison of the MHT/AHT/OHT partitioning between CMIP historical simulations (dashed lines) and pre-industrial simulation (solid lines). The MHT is shown in black. The AHT is shown in red. The OHT is shown in blue. Both lines are the ensemble mean of the 12 models that have sufficient output for the historical simulations. Historical simulations are averaged over the 2000-2018 period with no adjustment made for ocean heat storage to mimic the observational methodology.
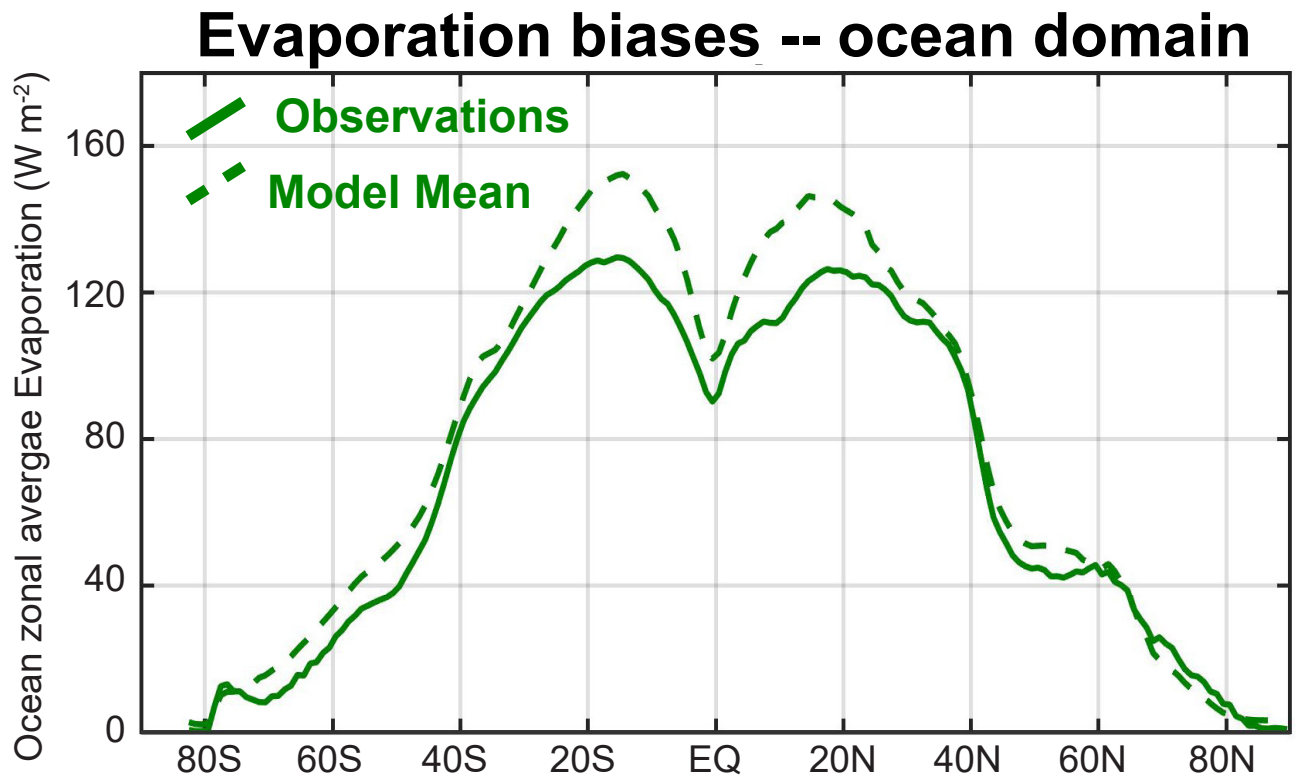
Evaporation biases -- ocean domain

Figure S4. Comparison between evaporation over the ocean in models (ensemble mean) and observations (WHOI OA flux). All values show the annual mean average over the ocean domain and are expressed as latent heat fluxes in W m^{-2}.
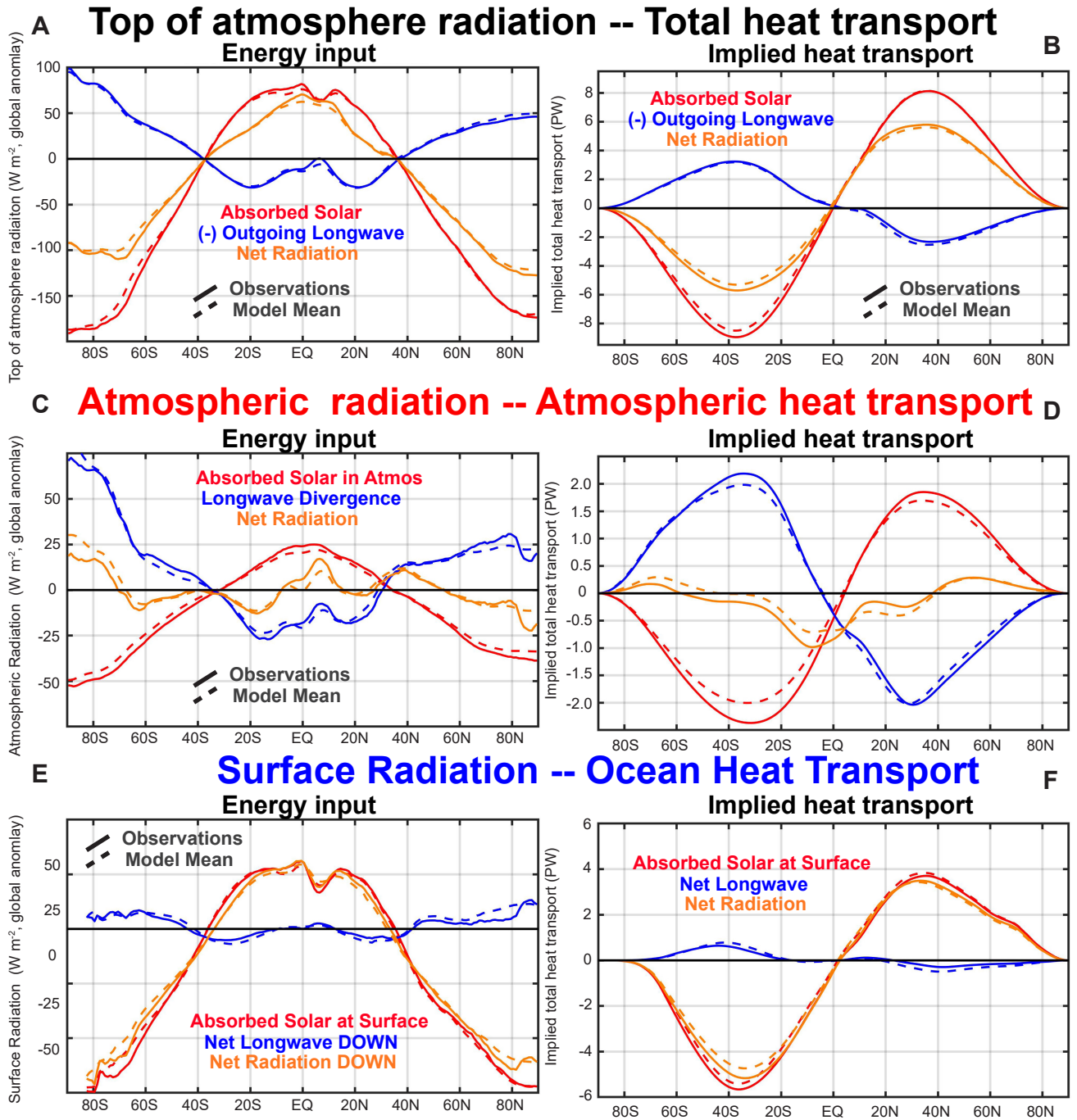
Figure S5. Comparison of radiation in observations (solid lines) and model ensemble

mean (dashed lines) at the top of atmosphere (A), in the atmospheric column (C) and at

the surface (E). Shortwave fluxes are shown in red, longwave fluxes are shown in blue and

the net radiation is shown in orange with positive values defined as a heating tendency

on the climate system, atmosphere and surface respectively. The global mean of each

term has been removed to emphasize the contribution to the spatial gradients in heating.

The right panels show the implied heat transport of the radiative components for the total

(atmosphere plus ocean) meridional heat transport (B, MHT), atmospheric heat transport

(D, AHT) and ocean heat transport (F, OHT) in PW. Note that the y-axis range differs