Supporting Information for "Increased Variability of Biomass Burning Emissions in CMIP6 Amplifies Hydrologic Cycle in the CESM2 Large Ensemble"

Kyle B. Heyblom[1], Hansi A. Singh[1], Philip J. Rasch[2,3], Patricia DeRepentigny[4]

[1]University of Victoria, School of Earth and Ocean Sciences, Victoria, BC, Canada

[2]Pacific Northwest National Laboratory, Atmospheric Sciences and Global Change Division, U.S. DOE Office of Science, Richland, WA, USA

[3]University of Washington, Department of Atmospheric Science, Seattle, WA, USA

[4]National Center for Atmospheric Research, Climate and Global Dynamics Laboratory, Boulder, CO, USA

Contents of this file

1. Text S1
2. Figures S1 to S7

Introduction

Here, we present the methods we used to evaluate statistical significance (Text S1-S2), as well as supplemental figures that further our findings on differences between CMIP6 (HiVarBB) and smoothed (SmoothBB) biomass burning emission scenarios in the CESM2 Large Ensemble. These figures show: statistical significance of area-averaged differences of metrics shown in Figure 2 (Figure S1), seasonal differences in relative liquid precipitation (Figure S2), statistical significance of differences in precipitation efficiency and ocean heat content (Figure S3), differences in atmospheric and cloud properties (Figure S4), differences in the moist and dry atmospheric energy transport components (Figure S5), and differences in the Atlantic meridional overturning circulation (AMOC; Figure S6).

Text S1. Evaluating spatial statistical significance. We assess spatial (i.e., grid point, zonally-averaged, and vertical profile) statistical significance using a Welch's t-test. We additionally limit significance determinations for false discoveries using the recommendations made by Wilks (2016). We use an α_{FDR} of 0.10 to approximate a global significance level of 0.05.

Text S2. Evaluating area-averaged statistical significance. We use a non-parametric bootstrapping approach to determine the statistical significance of the area-average differences between fields in HiVarBB and SmoothBB ensemble member sets. We conduct this test by randomly dividing all 80 members into two groups and determining the difference in the means of each group. We repeat this random selection a hundred thousand times to develop a distribution of random differences. We determine significance if the mean difference between the HiVarBB and SmoothBB ensemble member sets is outside of the 2.5 and 97.5 percentile range, signifying the two-tail 95% confidence interval of the distribution of differences between randomly divided members. This test allows us to determine whether, with 95% confidence, the mean difference between the HiVarBB and

SmoothBB ensemble member sets is greater than what could be generated by chance if the mean difference was only influenced by internal variability. To verify that significant differences are unique to the GFED period, we also conduct sensitivity tests by running the test over multiple time periods, both before and after the GFED period.

References

Wilks, D. S. (2016). "The stippling shows statistically significant grid points": How research results are routinely overstated and overinterpreted, and what to do about it. *Bulletin of the American Meteorological Society*, *97*(12), 2263–2273. doi: 10.1175/BAMS-D-15-00267.1

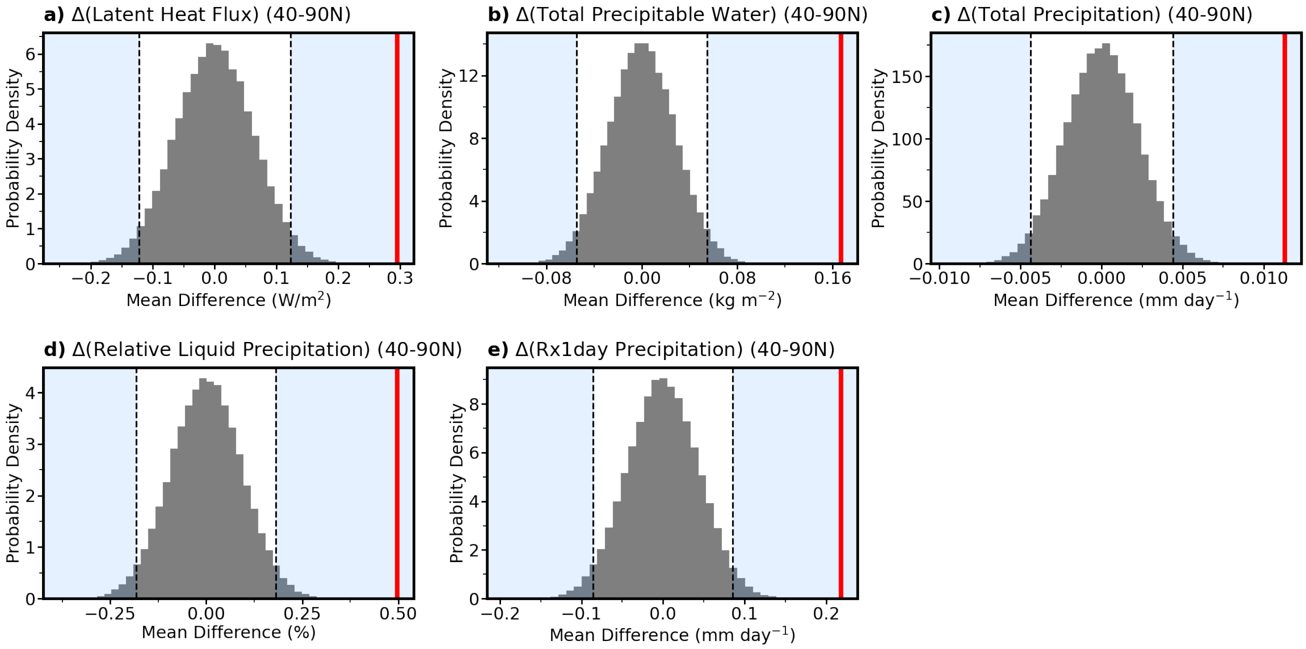

Figure S1. Statistical significance of area-averaged differences in the atmospheric hydrologic cycle. (a) latent heat flux, in W m^{-2}; (b) column-integrated precipitable water, in kg m^{-2}; (c) total precipitation, in mm day^{-1}; (d) percentage of precipitation that is liquid; and (e) annual maximum daily precipitation (Rx1day) in mm day^{-1}, all from 40-90°N over the GFED period (1997–2014). The gray histogram shows a probability density distribution of means derived from a non-parametric bootstrapping test (see Text S2), and the blue shading indicates the region outside of the (two-sided) 95% confidence intervals; the difference between HiVarBB and SmoothBB ensemble means (red line) is statistically significant (at the 95% level) if it falls within the blue shaded region.

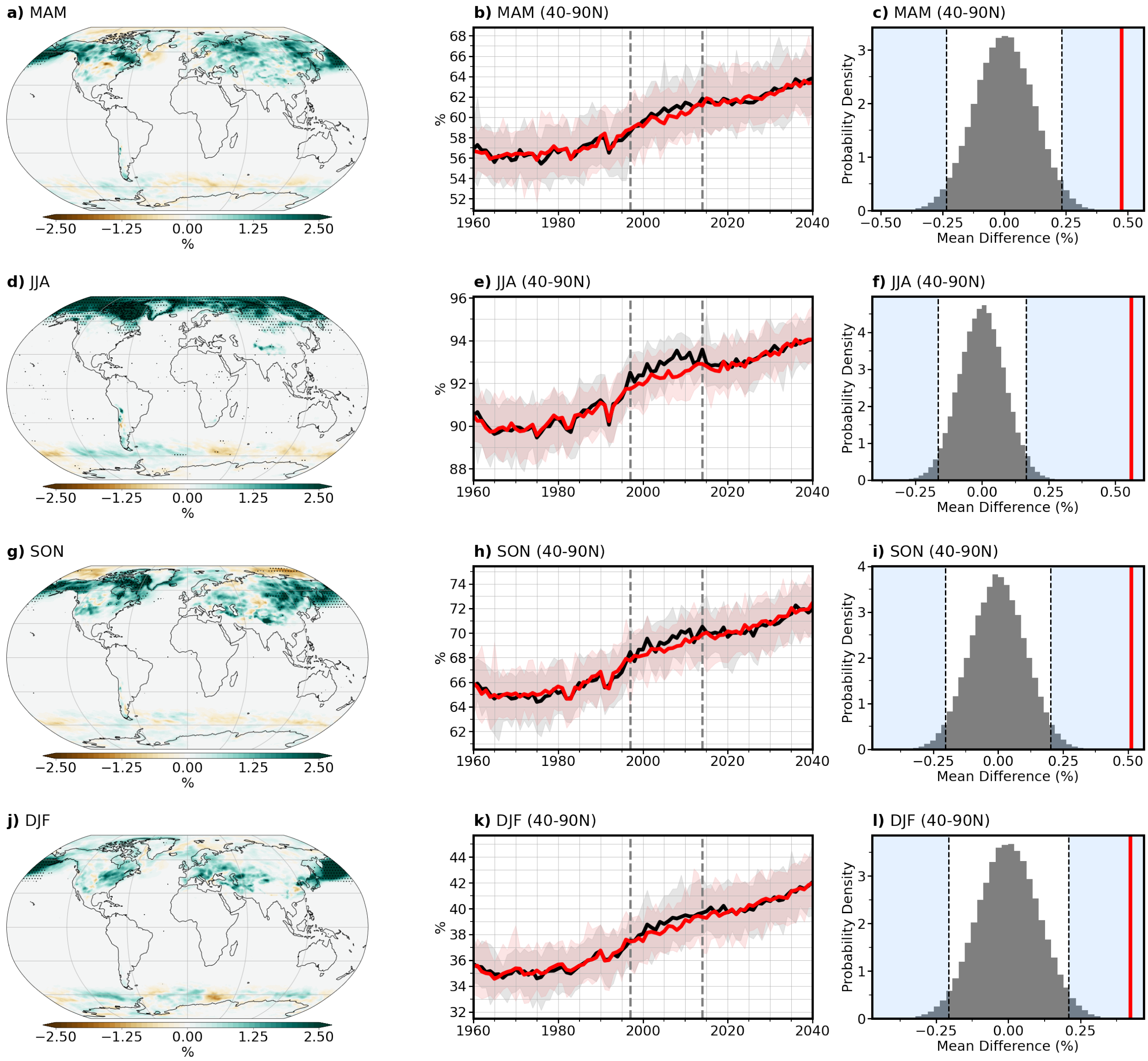

Figure S2. Differences in seasonal relative liquid precipitation. The left and middle columns are the same as in Figure 2, but showing differences in percentage of precipitation that is liquid in (a-c) March-May (MAM), (d-f) June-August (JJA), (g-i) September-November (SON), (j-l) and December-February (DJF). The right column shows the statistical significance of the difference in HiVarBB and SmoothBB ensemble means from 40-90°N over the GFED period. The gray histogram shows a probability density distribution of means derived from a non-parametric bootstrapping test (see Text S2), and the blue shading indicates the region outside of the (two-sided) 95% confidence intervals. The difference between HiVarBB and SmoothBB ensemble means (red line) is statistically significant (at the 95% level) if it falls within the blue shaded region.

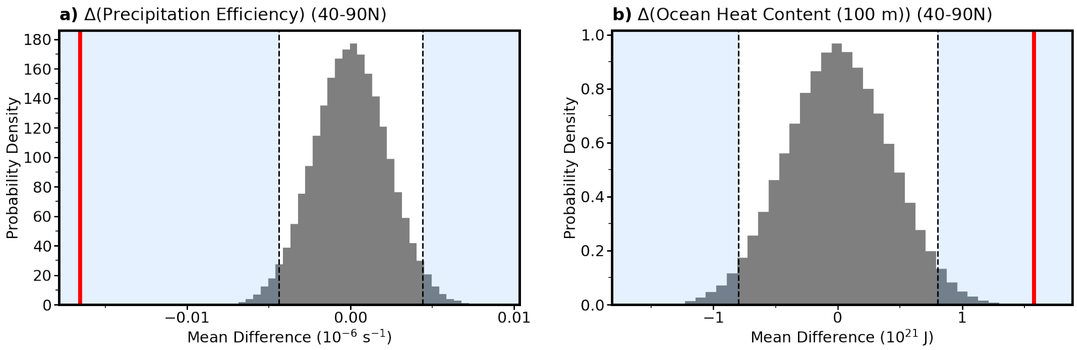

Figure S3. Statistical significance of area-averaged differences in moderating factors. (a) precipitation efficiency, and (b) upper (top 100 m) ocean heat content from 40-90°N during the GFED period (1997–2014). The gray histogram shows a probability density distribution of means derived from a non-parametric bootstrapping test (see Text S2), and the blue shading indicates the region outside of the (two-sided) 95% confidence intervals. The difference between HiVarBB and SmoothBB ensemble means (red line) is statistically significant (at the 95% level) if it falls within the blue shaded region.

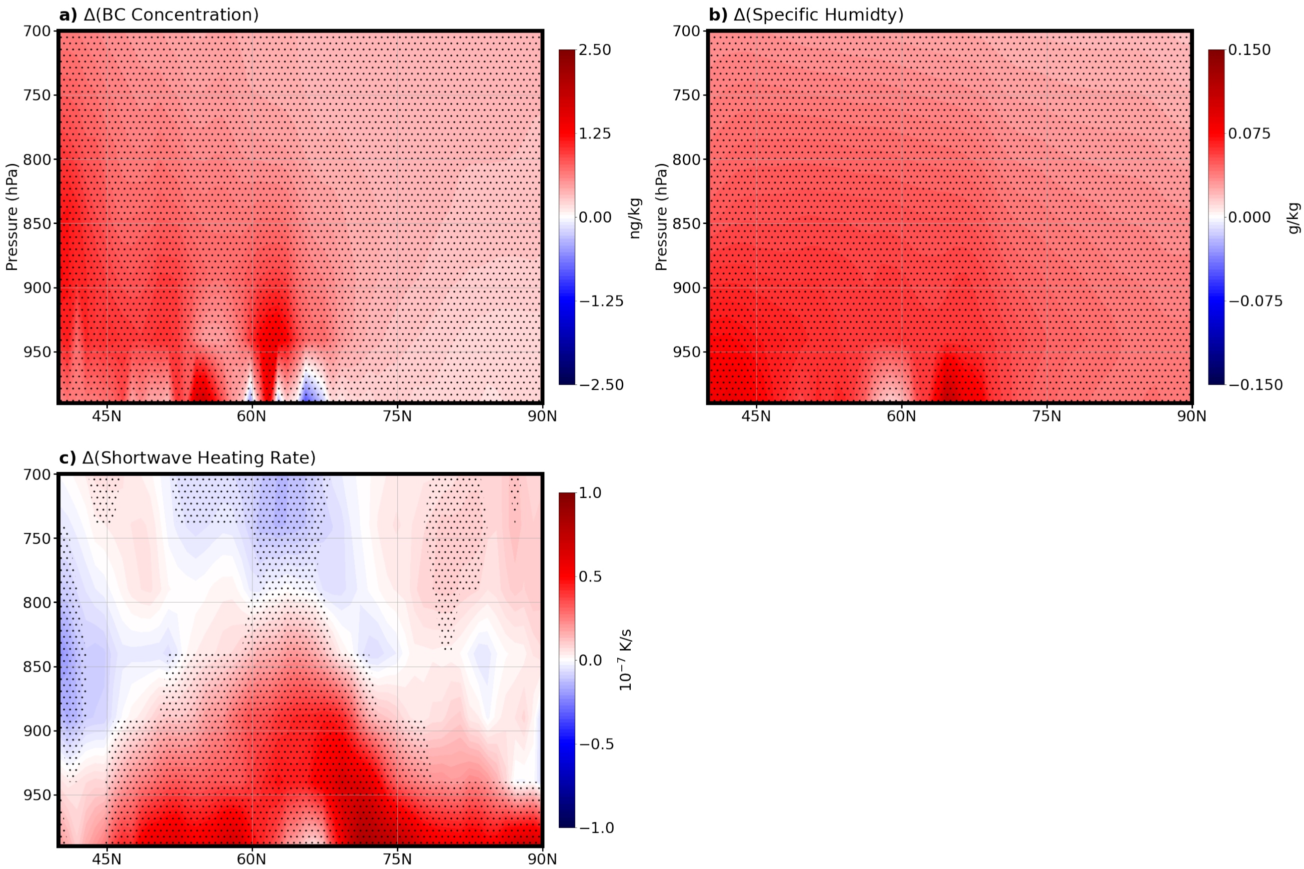

Figure S4. Zonal-mean ensemble mean difference of mechanisms affecting the precipitation efficiency. (a) black carbon concentration (in ng/kg), (b) specific humidity (in g/kg), (c) shortwave heating rate (in 10^{-7} K/s) from 40-90°N. Ensemble mean differences are computed as as the average of HiVarBB ensemble members minus the average of SmoothBB ensemble members during the GFED period (1997–2014). Stippling signifies 95% confidence in the significance of the difference between ensemble member sets (see Text S1).

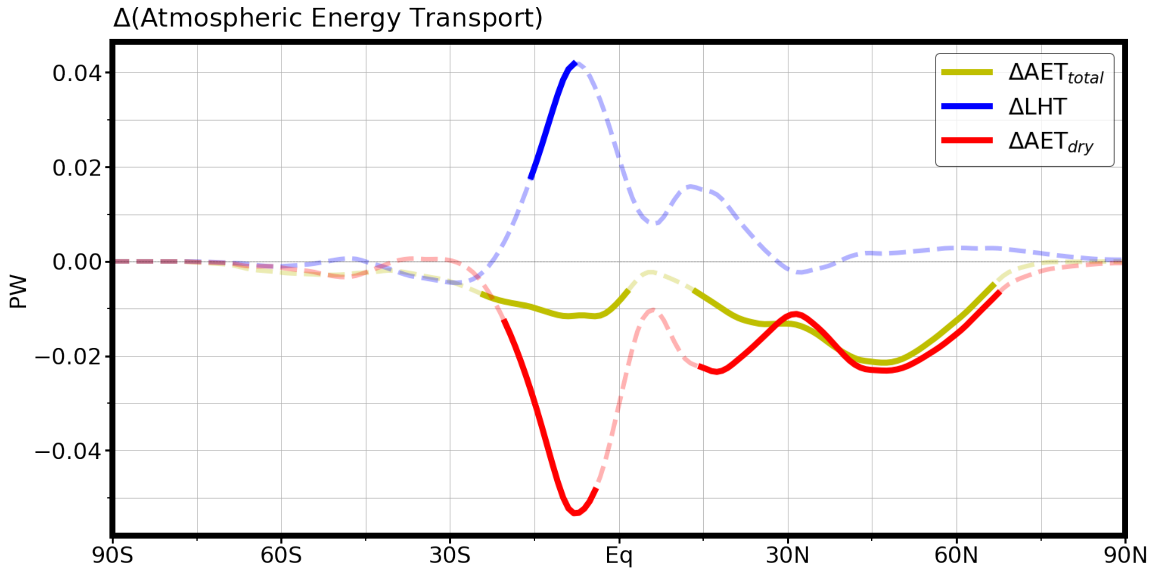

Figure S5. Differences in meridional atmospheric energy transport components. Ensemble mean difference (average of HiVarBB ensemble members minus average of SmoothBB ensemble members) in the total atmospheric energy transport (ΔAET$_{total}$, yellow line), latent heat transport (ΔLHT, blue line), and dry static energy transport (ΔAET$_{dry}$, red line) during the GFED period (1997–2014), in PW. Solid lines signify 95% confidence in the significance of the difference between HiVarBB and SmoothBB ensemble member sets (see Text S1).

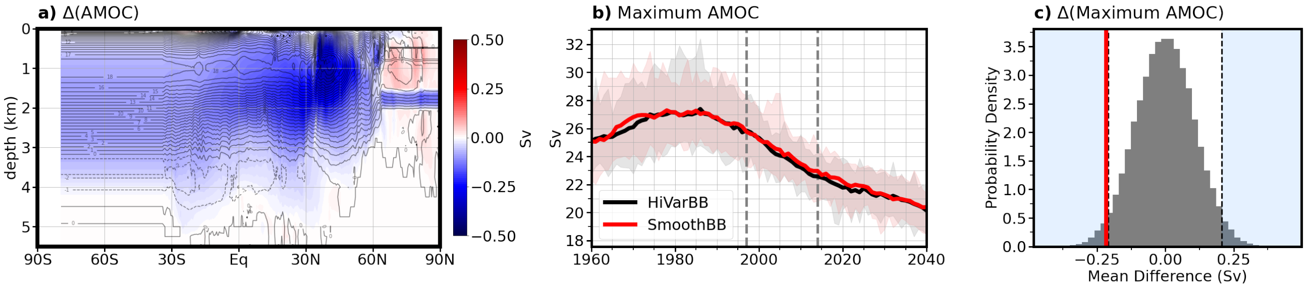

Figure S6. Differences in Atlantic meridional overturning circulation (AMOC). (a) zonal-mean ensemble mean difference (average of HiVarBB ensemble members minus average of the SmoothBB ensemble members), (b) annual mean Atlantic meridional overturning maximum from HiVarBB (black curve) and SmoothBB (red curve) ensemble members; thick lines denote the ensemble mean, shading denotes one standard deviation of each ensemble member set, and horizontal gray dotted lines delineate the GFED period (1997–2014), and (c) statistical significance of the difference in Atlantic meridional overturning maximum ensemble means during the GFED period. The gray histogram shows a probability density distribution of means derived from a non-parametric bootstrapping test (see Text S2), and the blue shading indicates the region outside of the (two-sided) 95% confidence intervals. The difference between HiVarBB and SmoothBB ensemble means (red line) is statistically significant (at the 95% level) if it falls within the blue shaded region.