1 Supplementary Figures

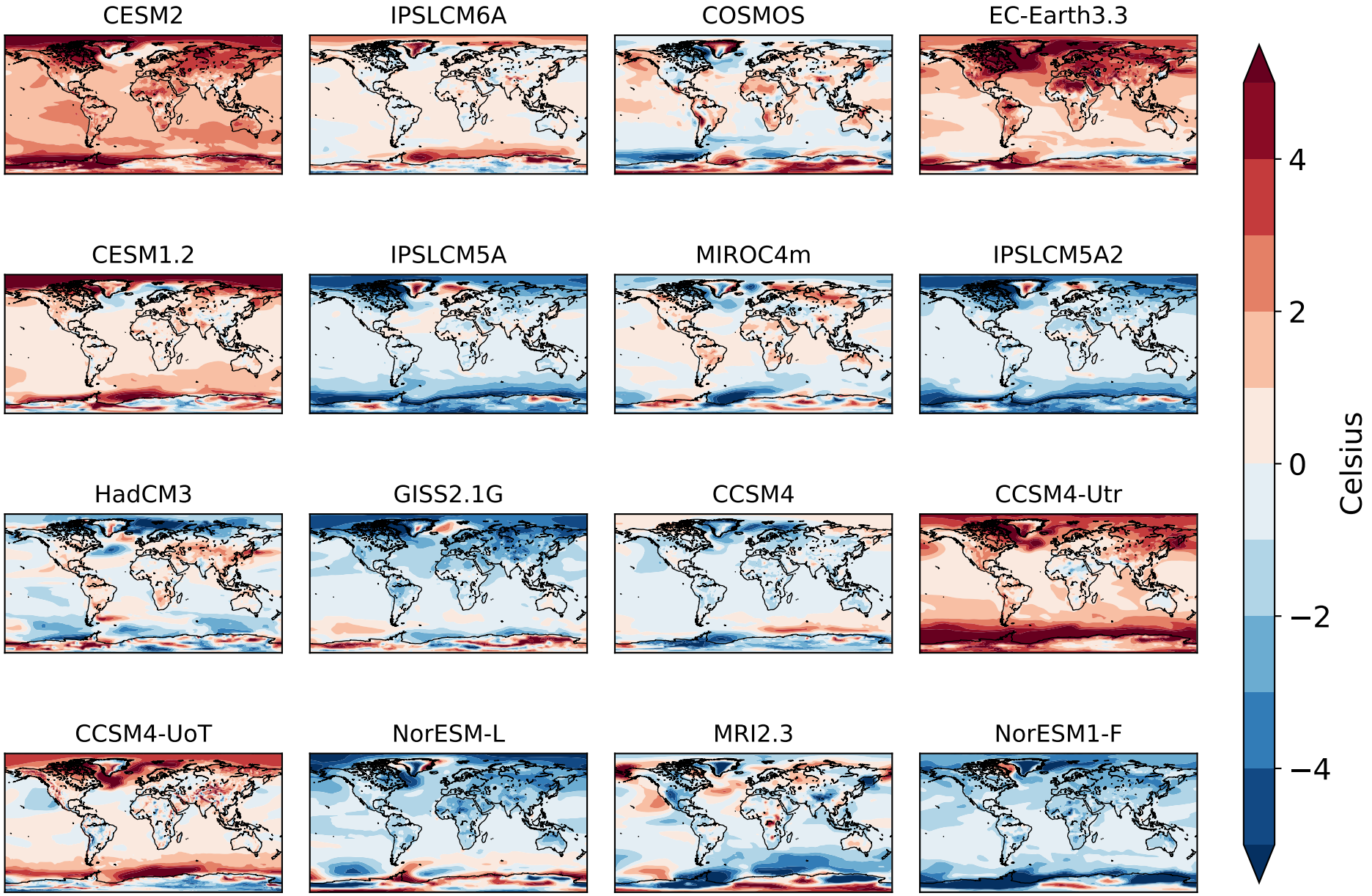

Figure 1: Near Surface Air Temperature anomaly $(Plio_{Core}$ - $PI_{Ctrl})$ from each model minus the multimodel mean Near Surface Air Temperature anomaly

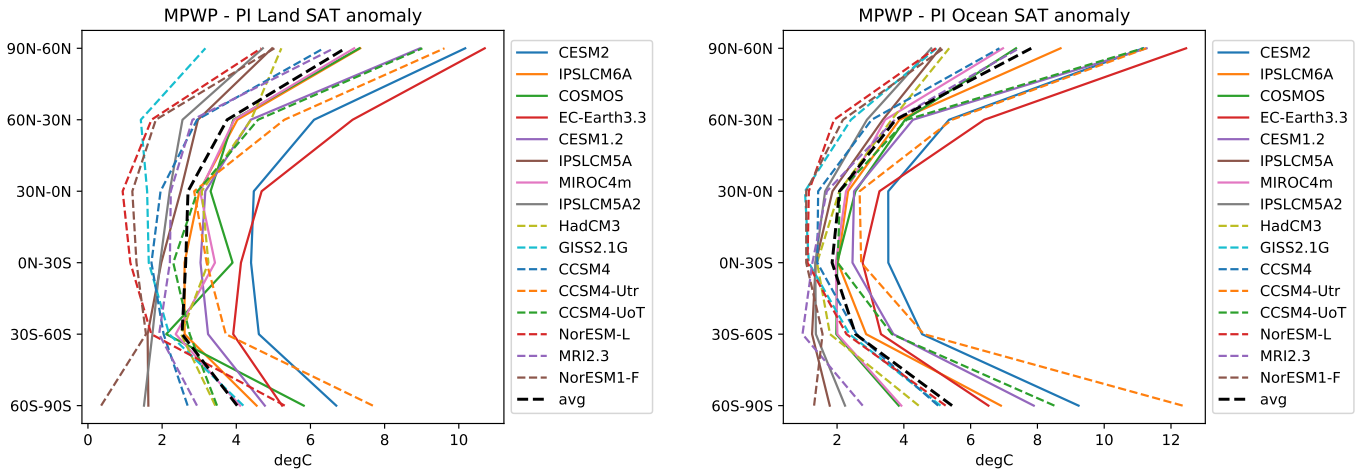

Figure 2: The average temperature anomaly for each model at different latitude bands over the land (left) and the ocean (right)

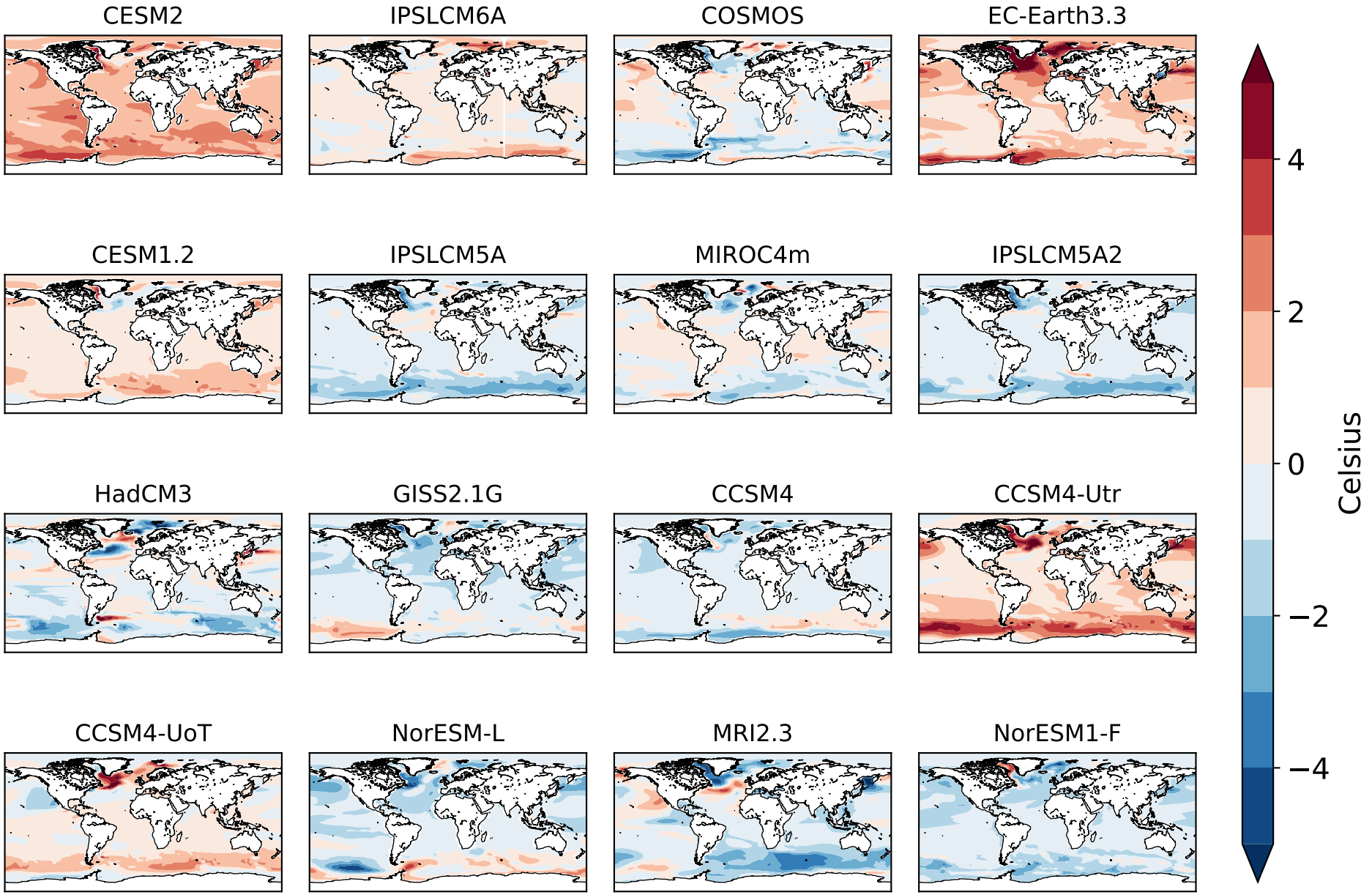

Figure 3: SST anomaly *(Plio$_{Core}$ - PI$_{Ctrl}$)* from each model minus the multimodel mean SST anomaly

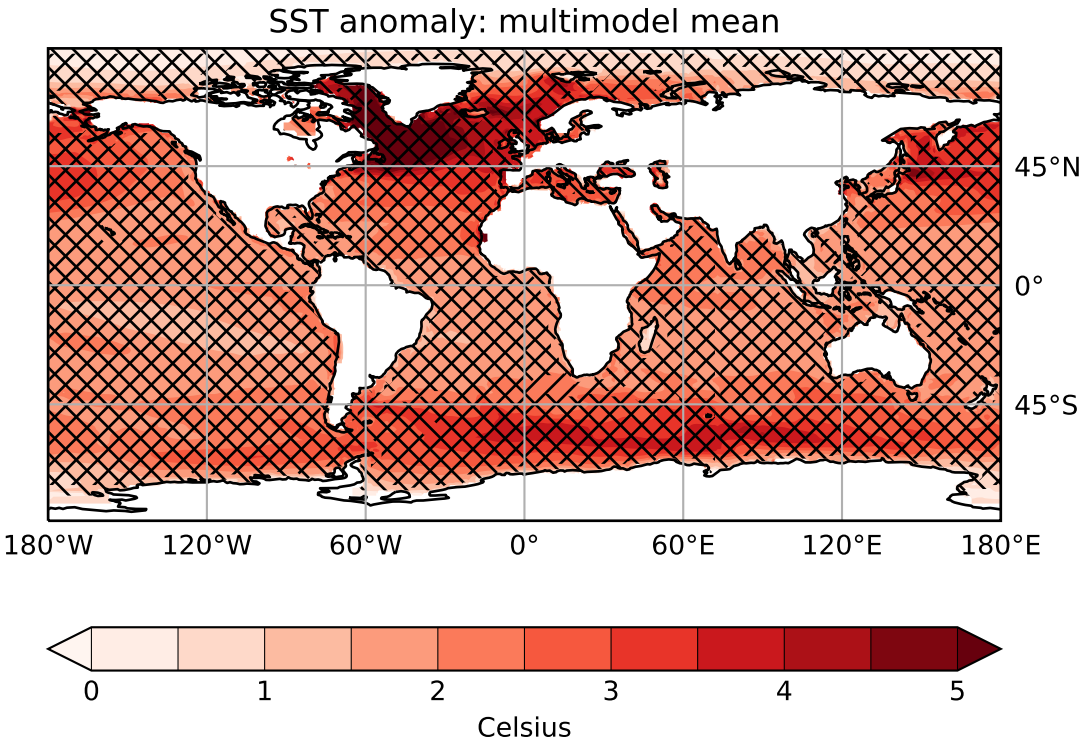

Figure 4: Multimodel mean *(Plio$_{Core}$ - PI$_{Ctrl}$)* SST anomalies (colours). Regions which have at least 80% of the models agreeing on the sign of the change are marked '/'. Regions which have the ratio of the multimodel mean SST change to the *PI$_{Ctrl}$* intermodel standard deviation greater than 1 are marked '\'. Regions which fulfil both these conditions are defined as being 'robust' across the ensemble.

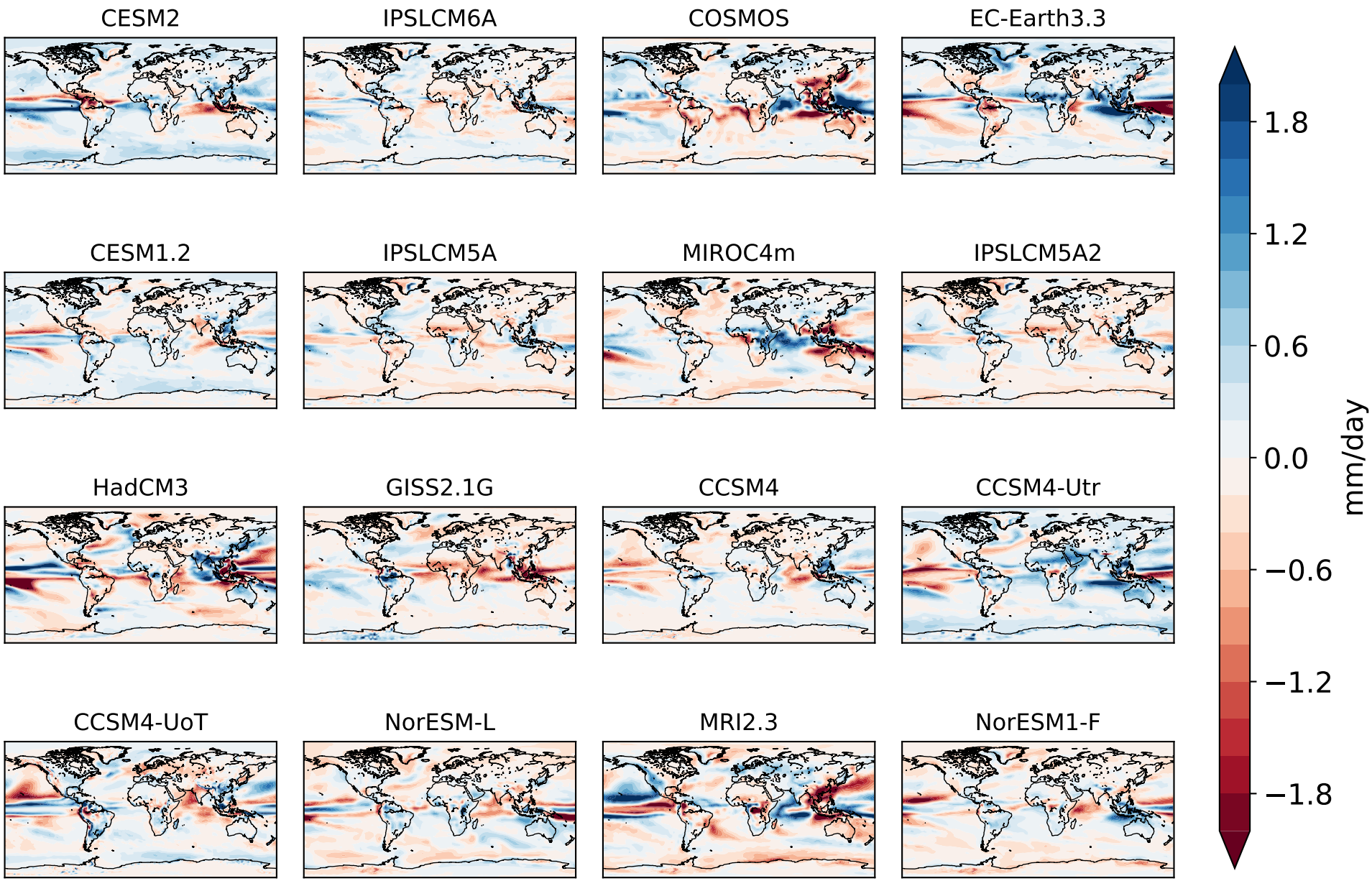

Figure 5: Precipitation anomaly *(Plio$_{Core}$ - PI$_{Ctrl}$)* from each model minus the multimodel mean Precipitation anomaly

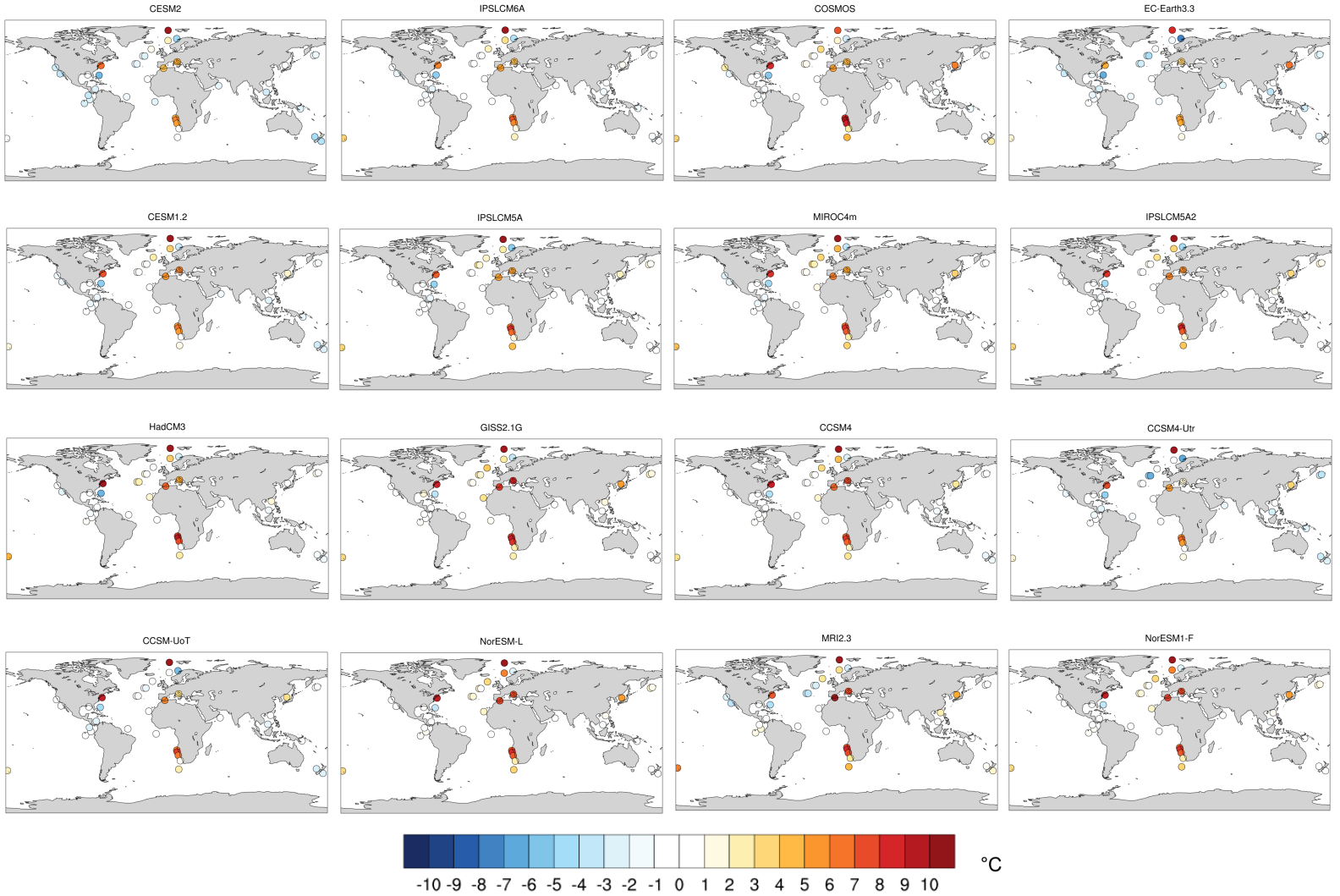

Figure 6: The difference between the SST anomaly derived from the each model and the SST anomaly obtained from the data (PRISM4 - NOAA ERSSTv5)

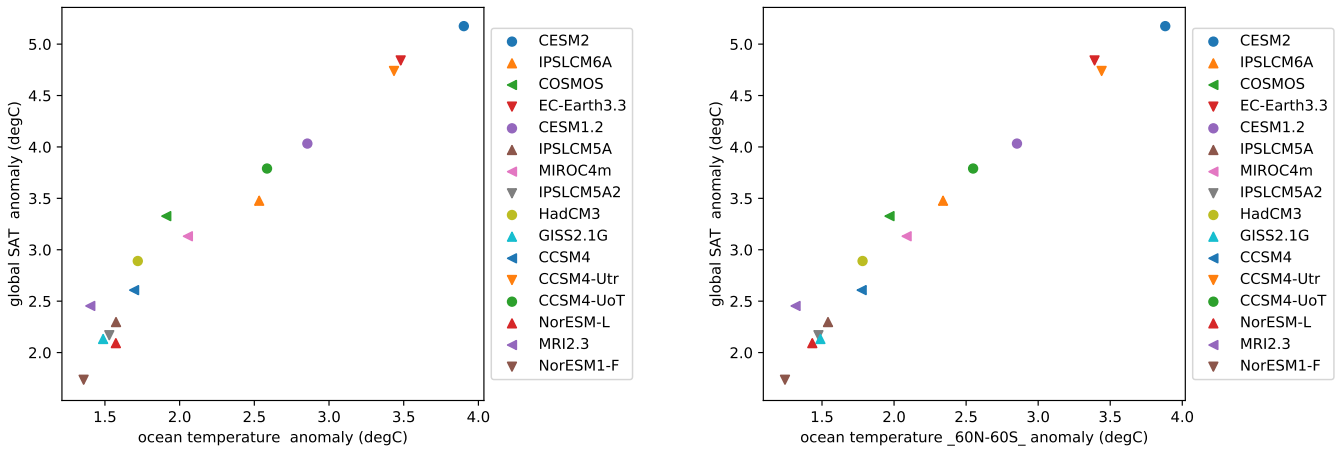

Figure 7: The relationship between the global mean *(Plio$_{Core}$ - PI$_{Ctrl}$)* SAT anomaly and the *(Plio$_{Core}$ - PI$_{Ctrl}$)* SST anomaly. Left: the SST anomaly was calculated over the whole ocean and the regression equation is SAT$_{anom}$ = (1.18 × SSTA) + 0.66, with R_{sq} = 0.97. Right: the SSTA was calculated over 60°N - 60° S, and the regression equation is SAT$_{anom}$ = (1.16 × OTA) + 0.74, with R_{sq}=0.97.

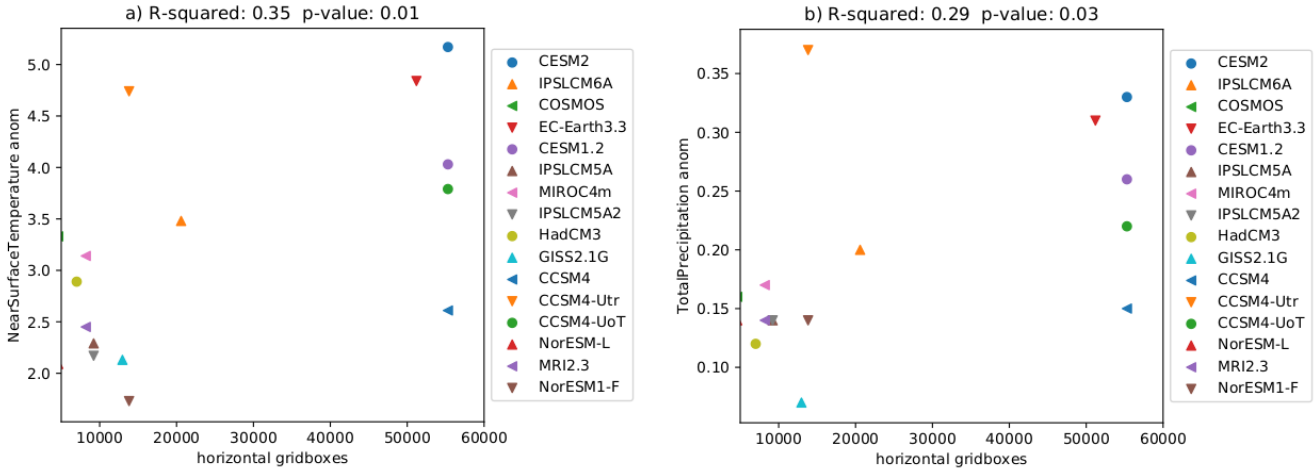

Figure 8: a) the correlation between the $Plio_{Core}$ - PI_{Ctrl} SAT anomaly and the number of atmopsheric gridboxes. (b) as (a) but for precipitation

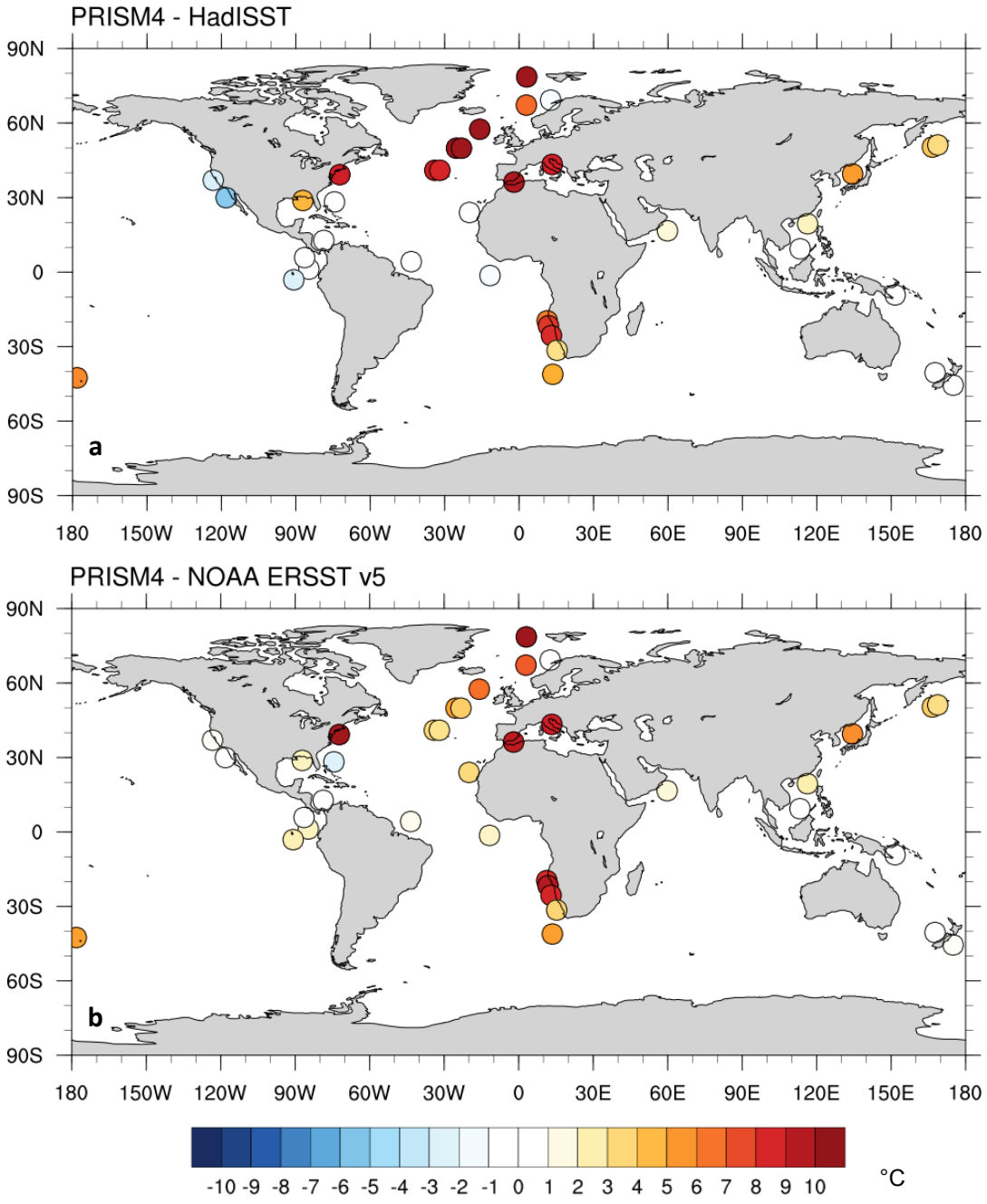

Figure 9: SST anomaly from the data using preindustrial SST from different sources. Top Panel: PRISM4 SST - HadISST. Lower Panel: PRISM4 SST - NOAA ERSSTv5

Table 1: Initial conditions and spin up. (Note the Near Surface Air Temperature anomaly drift was calculated over the final 100 years of each simulation)

Model name	initial ocean temperature	Run length years	net TOA radiation W/m^2	NSAT anomaly drift °C / 100 years
CCSM4	warm start with blended PRISM3D and CCSM4 PlioMIP1	1000	0.04	-0.05
CCSM4-UoT	Levitus	2820	0.1	0.04
CCSM4-Utr	horizontally homogeneous 15°C top - 4°C bottom (tanh function) (globally averaged 3°C above preindustrial)	2000	-0.07	0.02
CESM1.2	warm start blended PRISM3D and CCSM4 PlioMIP1	1200	0.17	-0.07
CESM2	PRISM3D	1200	0.2	-0.1
COSMOS	Levitus	2000	1.89 (diff from PI = +0.16)	-0.1
EC-Earth3.3	Levitus SST average 18.4°C	1400	0.67	-0.03
GISS2.1G	Levitus	1250	038	0.18
HadCM3	zonally-averaged 3D temperatures from a preindustrial simulation	2500	0.05	0.08
IPSLCM6A	preindustrial	1450	0.91	0.4
IPSLCM5A2.1	End of PlioMIP1 expt 2	1500	0.43	-0.09
IPSLCM5A	End of PlioMIP1 expt 2	800	0.69	-0.15
MIROC4m	preindustrial	4000	0.84	0.08
MRI-CGCM2.3	present day	1000	2.69 (diff from PI = -0.17)	0.18
NorESM1-F	end of 2000 year 400ppmv CO_2 simulation	500	-0.01	0.04
NorESM-L	end of PlioMIP1	1200	0.10	0.13