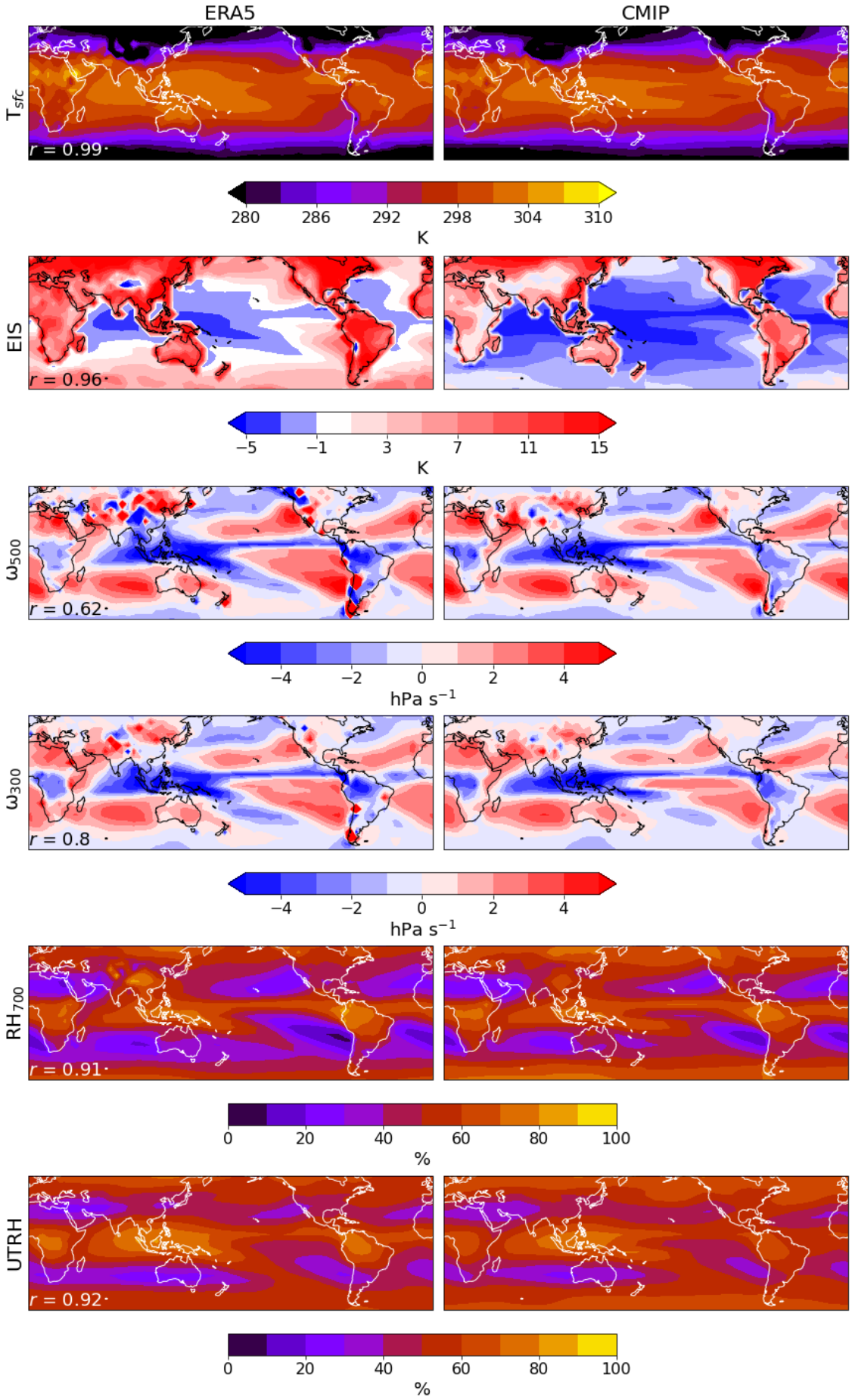
S1. CMIP models

Eighteen CMIP5/6 models are available having run the ISCCP simulator. We use the first available, historical (AMIP) realisations for CCFs and cloud radiative anomalies for the years 1981 – 2000 from the following models:

- **CMIP5**: CanESM2, CNRM-EM2-1*, HadGEM2-ES*, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC5, MIROC-ESM, MPI-ESM-LR, MRI-CGCM3,
- **CMIP6**: CanESM5, CNRM-CM6-1*, GFDL-CM4, HadGEM3-GC31-LL, IPSL-CM6A-LR*, MIROC6, MIROC-ES2L, MRI-ESM2-0, UKESM1-0-LL.

*Models where daily air temperature and/or humidity at standard pressure levels is not available, therefore no CAPE or CIN data has been calculated.


ERA5
CMIP
T_{sfc}
r = 0.99
280
286
292
298
304
310
K
EIS
r = 0.96
−5
−1
3
7
11
15
K
ω_{500}
r = 0.62
−4
−2
0
2
4
hPa s^{-1}
ω_{300}
r = 0.8
−4
−2
0
2
4
hPa s^{-1}
RH_{700}
r = 0.91
0
20
40
60
80
100
%
UTRH
r = 0.92
0
20
40
60
80
100
%

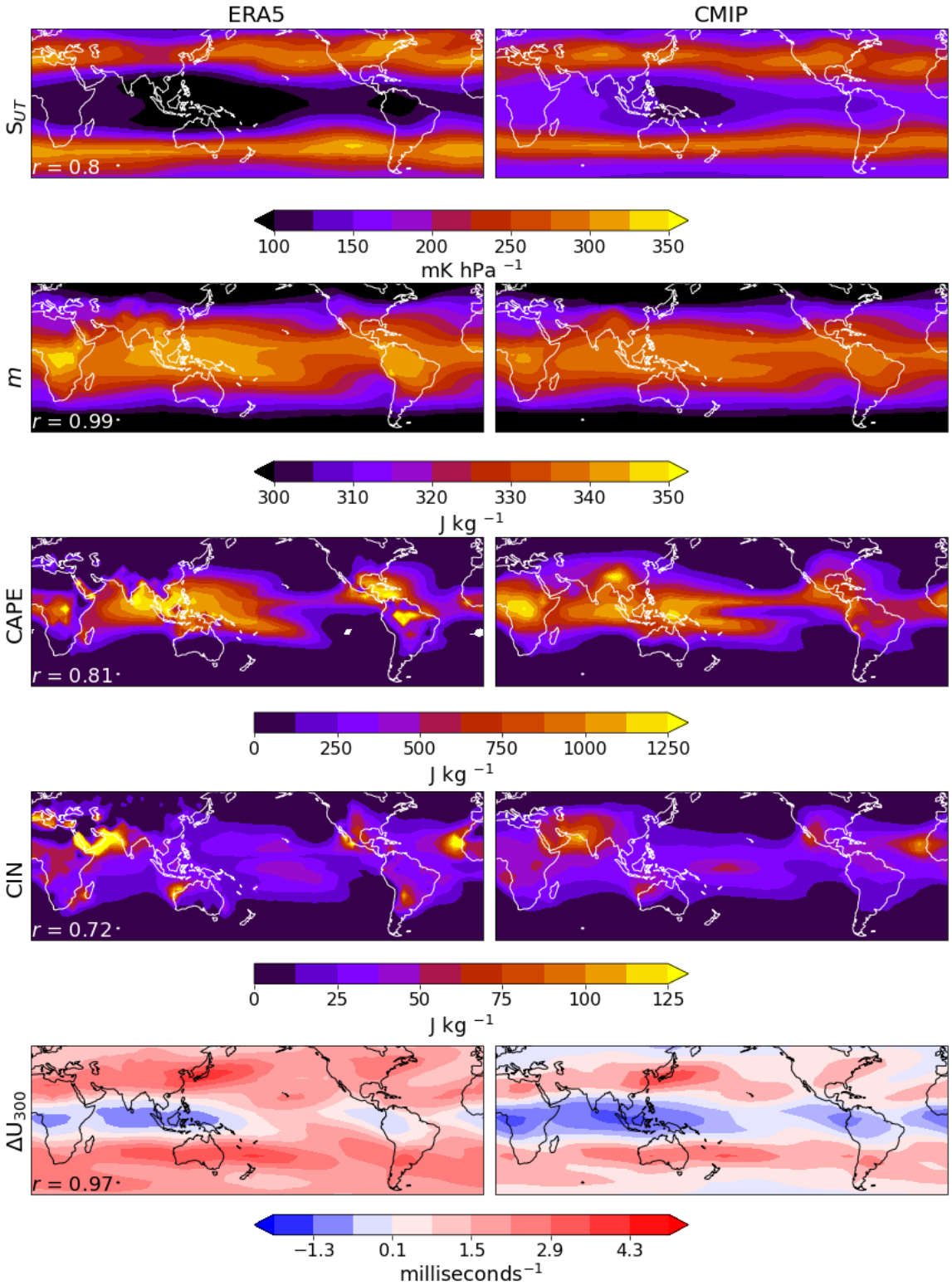


Figure S1. ERA5 (left column) and multi-model mean CMIP (right column) monthly climatological mean for the CN21 and candidte CCFs (from top to bottom, surface temperature, estimated inversion strength, vertical velocity at 500hPa and 300hPa, relative humidity at 700hPa and in the upper troposphere, static stability in the upper troposphere, boundary layer moist static energy, convective available potential energy, convective inhibition, and easterly wind shear. Note different scales and units on the colourbars for each variable. The Pearson r correlation coefficient for the spatial distribution of ERA5 and multi-model CMIP mean values is shown in the bottom left of each ERA5 figure.

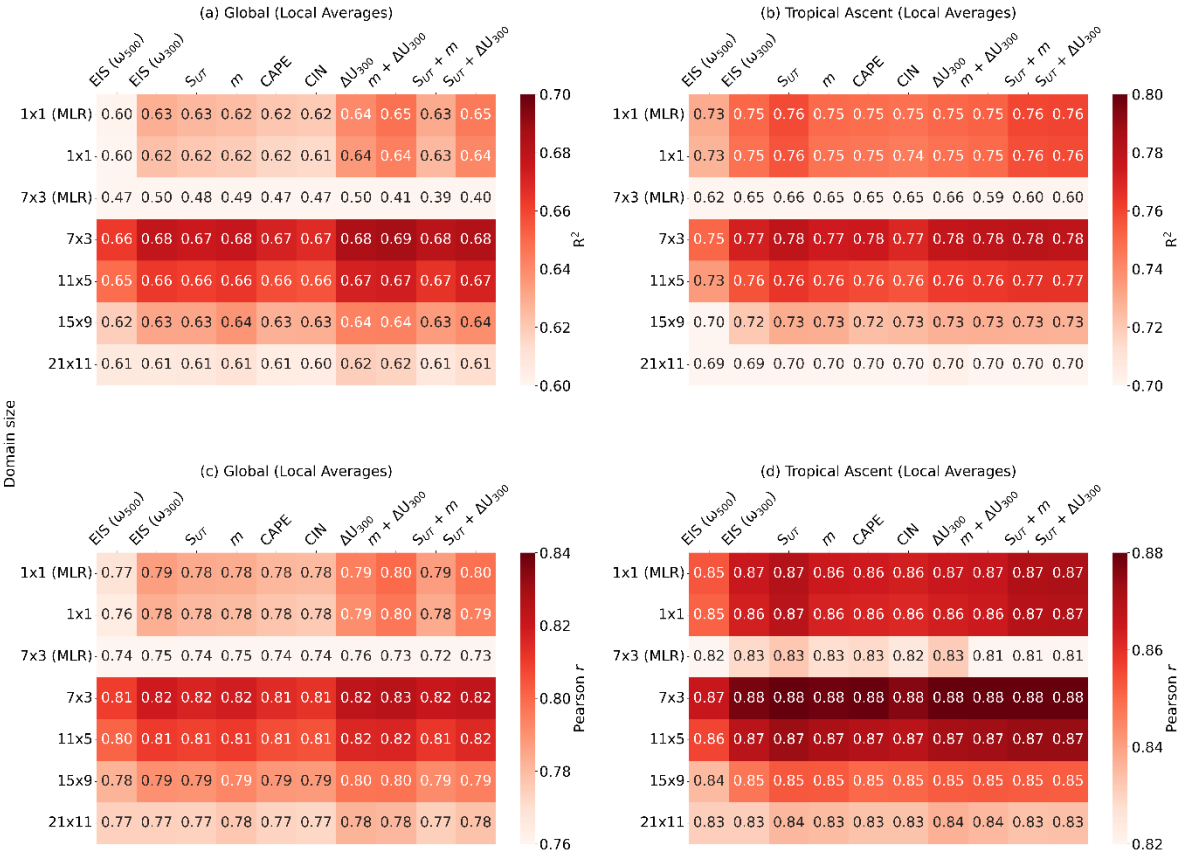


Figure S2. Matrices showing skill metrics for local predictions made for the observed R_{LW} time series at each domain size using different "CCF configurations". A "CCF configuration" refers to the selection of cloud controlling factors used to predict R_{LW}. Each configuration uses T_{sfc}, RH_{700}, UTRH and ω_{300} (with the exception of the first column, where ω_{500} is used instead) and a candidate CCF(s) (e.g., S_{UT}), which is used to label each column. R^2 (top panels) and Pearson *r* scores (bottom panels) are shown for predictions made locally, with metrics averaged (a and c) globally and (b and d) in tropical ascent regions, defined as grid-cells with observed climatological EIS < 1 K, ω_{500} < 0 hPa s^{-1}. Metrics are weighted by the cosine of latitude and monthly standard deviation of R_{LW} of each grid-cell. All predictions are made using ridge regression, except for row 1x1 (MLR) and 7x3 (MLR), which are made using multiple linear regression. This is analogous to (a) and (b) in Fig. 2 in the main text.

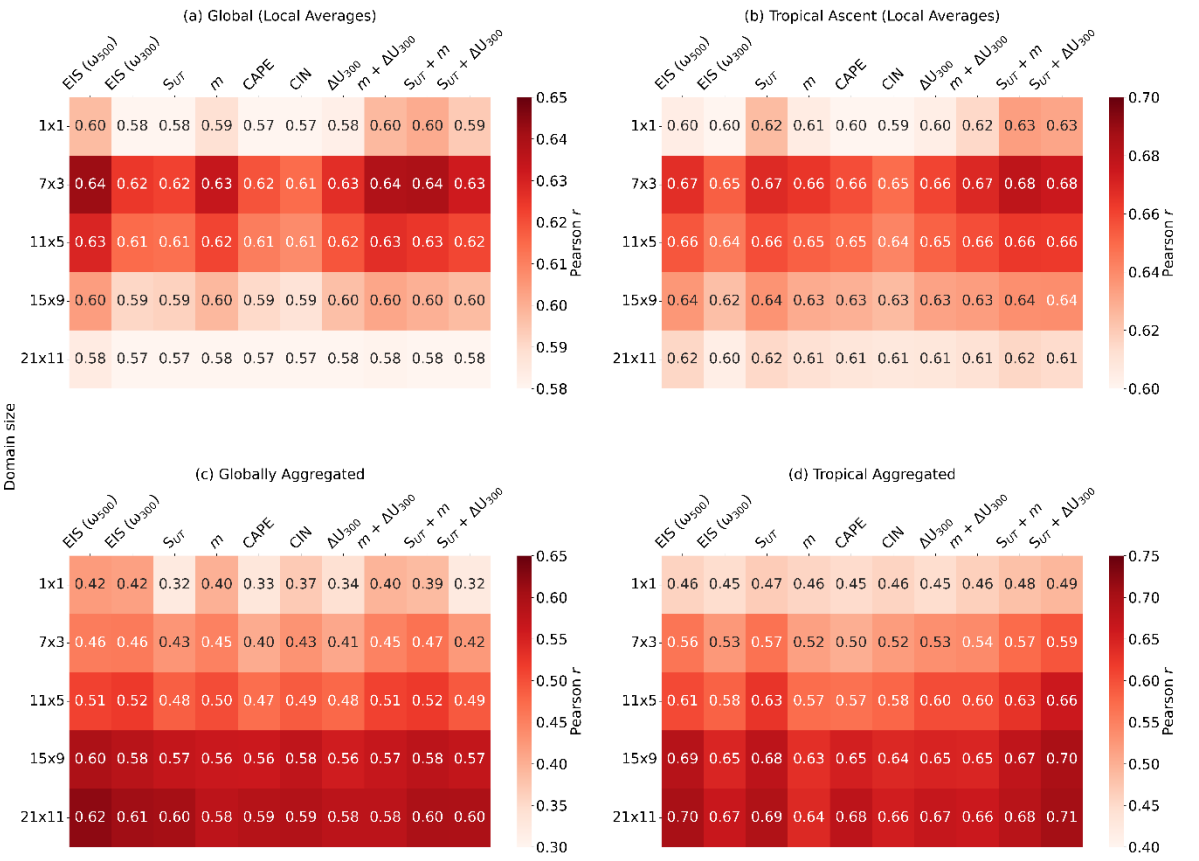


Figure S3. Matrices showing skill metrics for predictions made for the observed R_{NET} time series at each domain size using different "CCF configurations". A "CCF configuration" refers to the selection of cloud controlling factors used to predict R_{NET}. Each configuration uses T_{sfc}, RH_{700}, UTRH and ω_{300} (with the exception of the first column, where ω_{500} is used instead) and a candidate CCF(s) (e.g., S_{UT}), which is used to label each column. Predictions are made locally, with the Pearson r averaged (a) globally and (b) in tropical ascent regions defined as grid-cells with observed climatological EIS < 1 K, $\omega_{500} < 0$ hPa s^{-1}. Metrics are weighted by the cosine of latitude and monthly standard deviation of R_{LW} of each grid-cell. Pearson r is also shown for aggregated predictions, (c) globally and (d) in the tropical ascent regions, and compared to similarly aggregated observations. Note different scales for each colorbar.

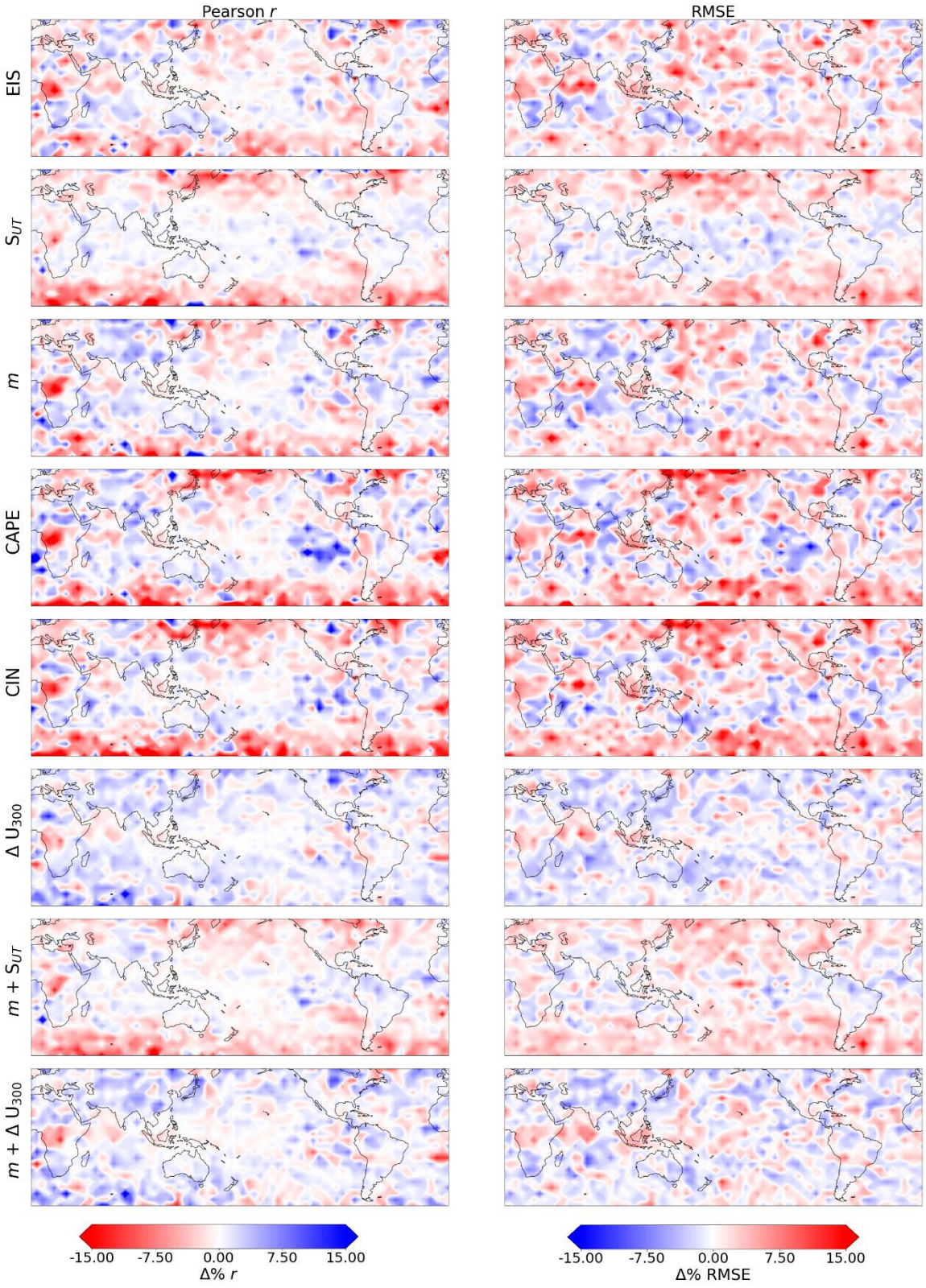


Figure S4. The % difference in local Pearson r (left column) and RMSE (right column) scores for predictions made of observed R_{LW} compared to configuration $S_{UT} + \Delta U_{300}$ (also with $T_{sfc,}$ RH_{700}, UTRH and ω_{300}), using a 21x11 domain for each configuration. A "CCF configuration" refers to the selection of cloud controlling factors used to predict R_{LW}. Each configuration uses T_{sfc}, RH_{700}, UTRH and ω_{300} and a candidate CCF(s) (e.g., S_{UT}), which is used to label each row. See the local skill metrics for $S_{UT} + \Delta U_{300}$ in Figure 3 of the main text. Note the reversed colorbars for each column. This is to indicate where performance is improved, corresponding to an increase in Pearson r or decrease in RMSE.

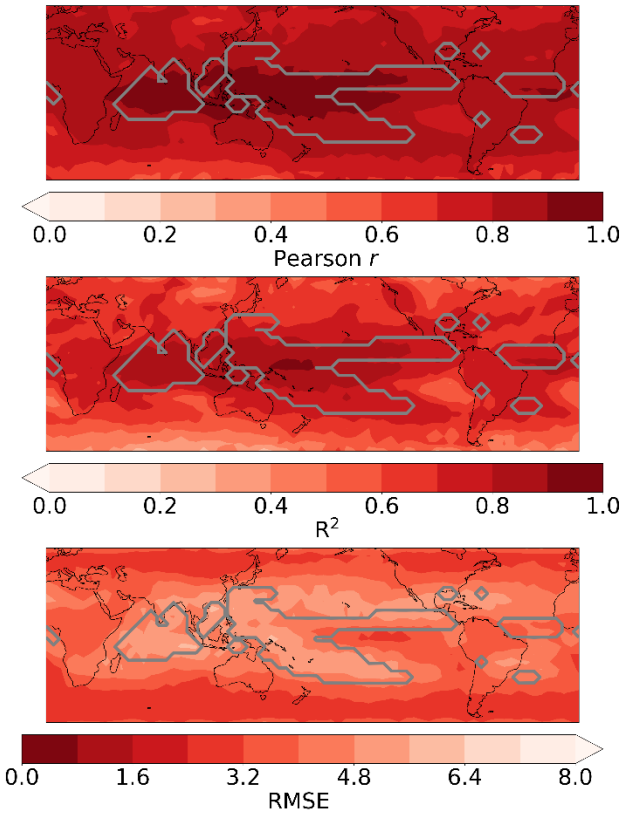


Figure S5. Multi-model median skill metrics (from top to bottom: Pearson r, R^2 score, and RMSE) for predictions of R_{LW} time series using the configuration $S_{UT} + \Delta U_{300}$ (also with T_{sfc}, RH_{700}, UTRH and ω_{300}), with CCFs within a spatial domain of 21x11. Grey contours show the tropical ascent regions. This figure is analogous to Figure 3 in the main text, where skill metrics for the observations are shown instead.

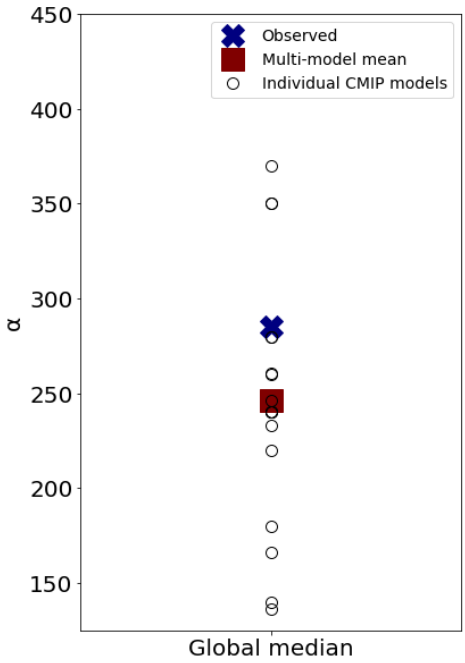


Figure S6. The global median regularization parameter, α, for each of the CMIP models, with the multi-model mean (square) and global median α for observed predictions also shown. Configuration $S_{UT} + \Delta U_{300}$ (also with T_{sfc}, RH_{700}, UTRH and ω_{300}) and a 21x11 domain has been used.

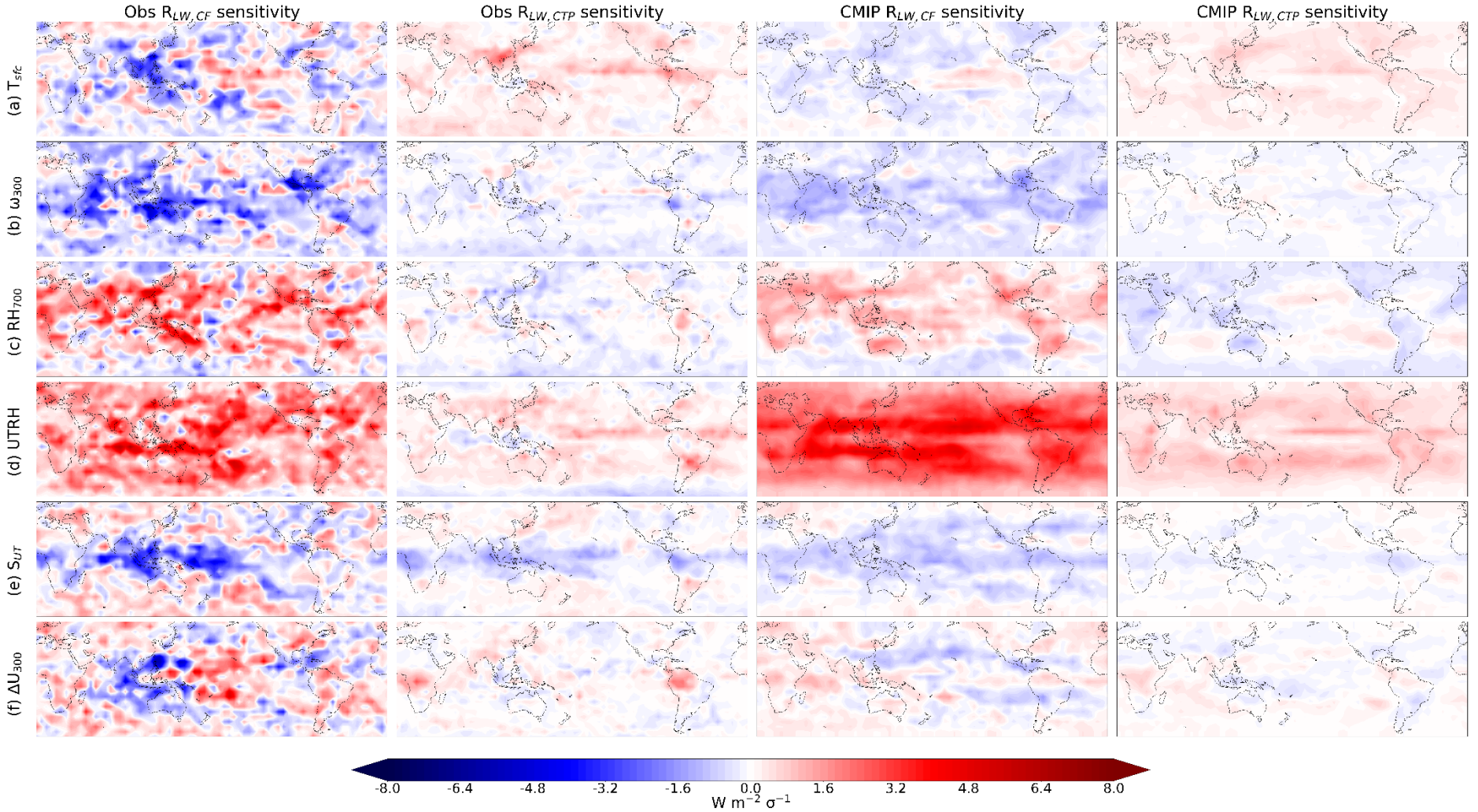


Figure S7. Sensitivities ($\sum \boldsymbol{\theta_i}$) to the cloud controlling factors in configuration $S_{UT} + \Delta U_{300}$(also with T_{sfc}, RH_{700}, UTRH and ω_{300}) for radiative anomalies caused by changes in cloud fraction (panels a and c) and cloud top pressure (panels b and d), derived using a 21x11 domain and defined for a one-standard deviation anomaly in each CCF. To produce the maps, we sum all elements of the sensitivity vectors at each point r. Sensitivities are shown for the observations (first two columns) and the multi-model mean (last two panels). The same colorbar has been used throughout to show the relative strengths of the sensitivities.

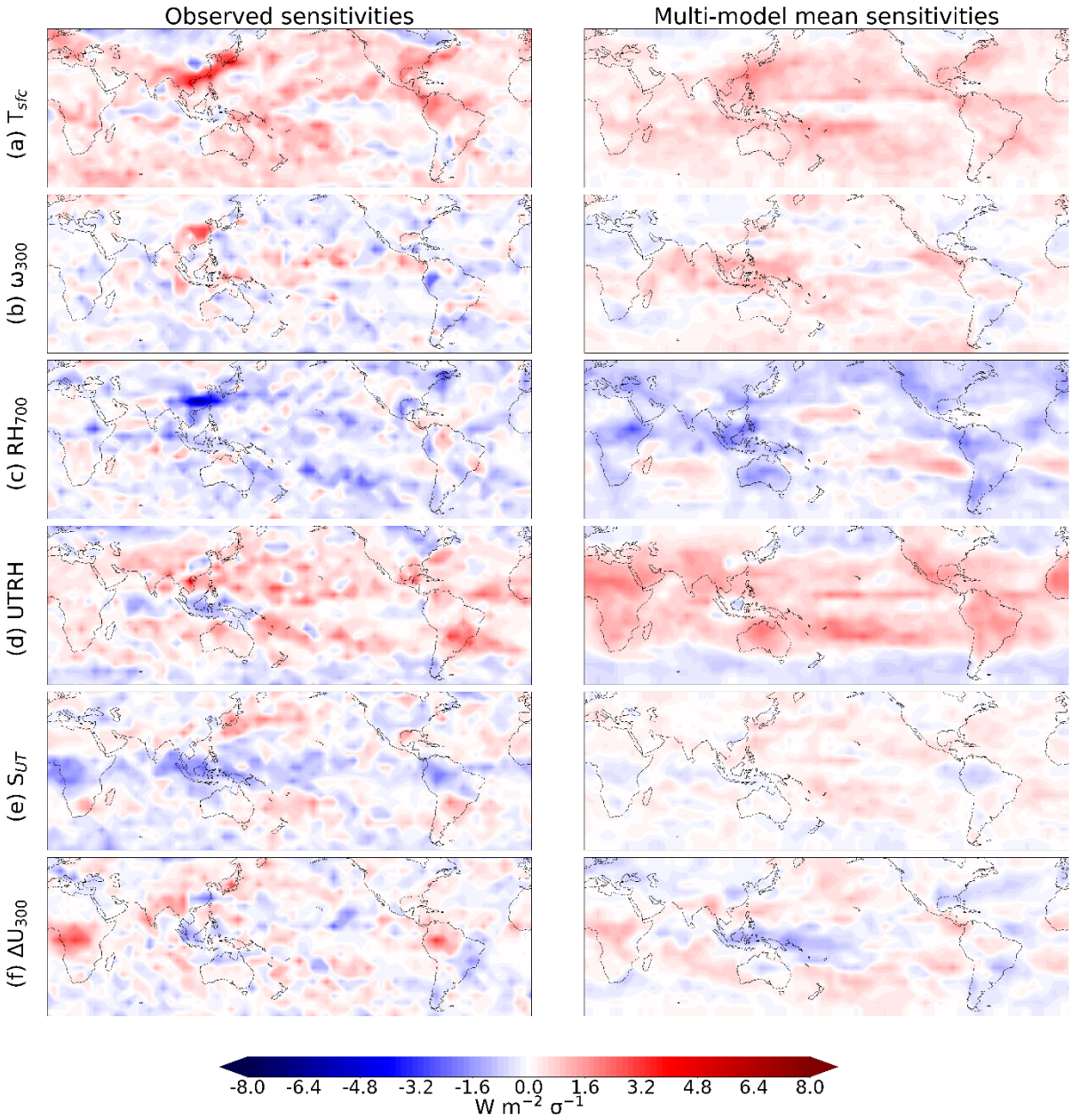


Figure S8. R_{NET} sensitivities ($\sum \boldsymbol{\Theta_i}$) to the cloud controlling factors in configuration $S_{UT} + \Delta U_{300}$ (also with T_{sfc}, RH_{700}, UTRH and ω_{300}), derived using a 21x11 domain and defined for a one-standard deviation anomaly in each CCF (scaled using ERA5 CCFs for visualisation purposes). To produce the maps, we sum all elements of the sensitivity vectors at each point r. The left column shows observed sensitivities, and the right shows the multi-model mean.

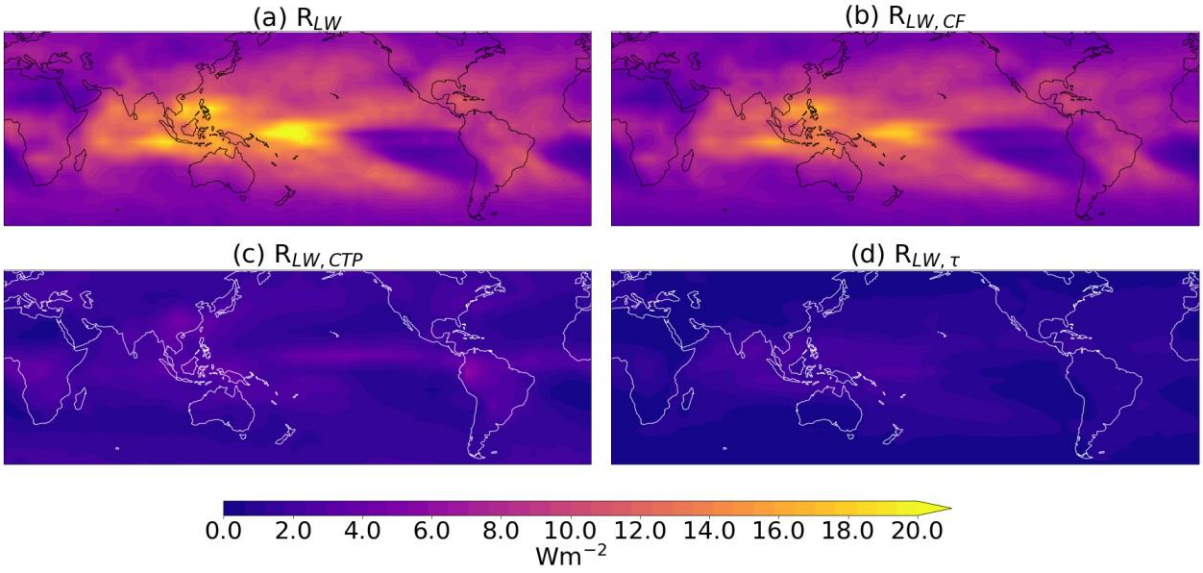


Figure S9. Observed monthly climatologies for (a) longwave high-cloud radiative anomalies and the decompositions into radiative anomalies induced by changes in (b) cloud fraction (c) cloud top pressure and (d) optical depth.

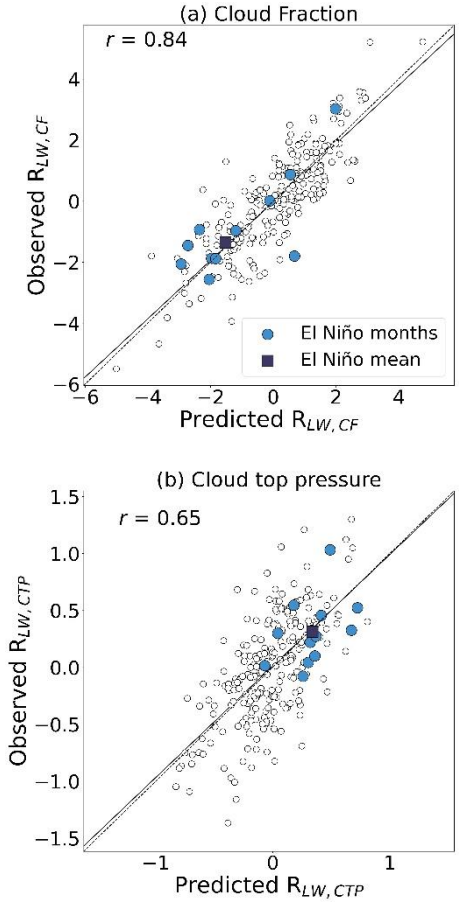


Figure S10. Scatter plot showing the correlation between observed and predicted monthly tropical ascent-aggregated (a) $R_{LW,CF}$ and (b) $R_{LW,CTP}$ time series using configuration $S_{UT} + \Delta U_{300}$ (in addition to T_{sfc}, RH_{700}, UTRH, and ω_{300}) and a 21x11 domain. El Niño months are shown using coloured circles, with the annual mean shown using a coloured square. Solid lines show y = x, and the dashed lines show the line-of-best fit through the points.

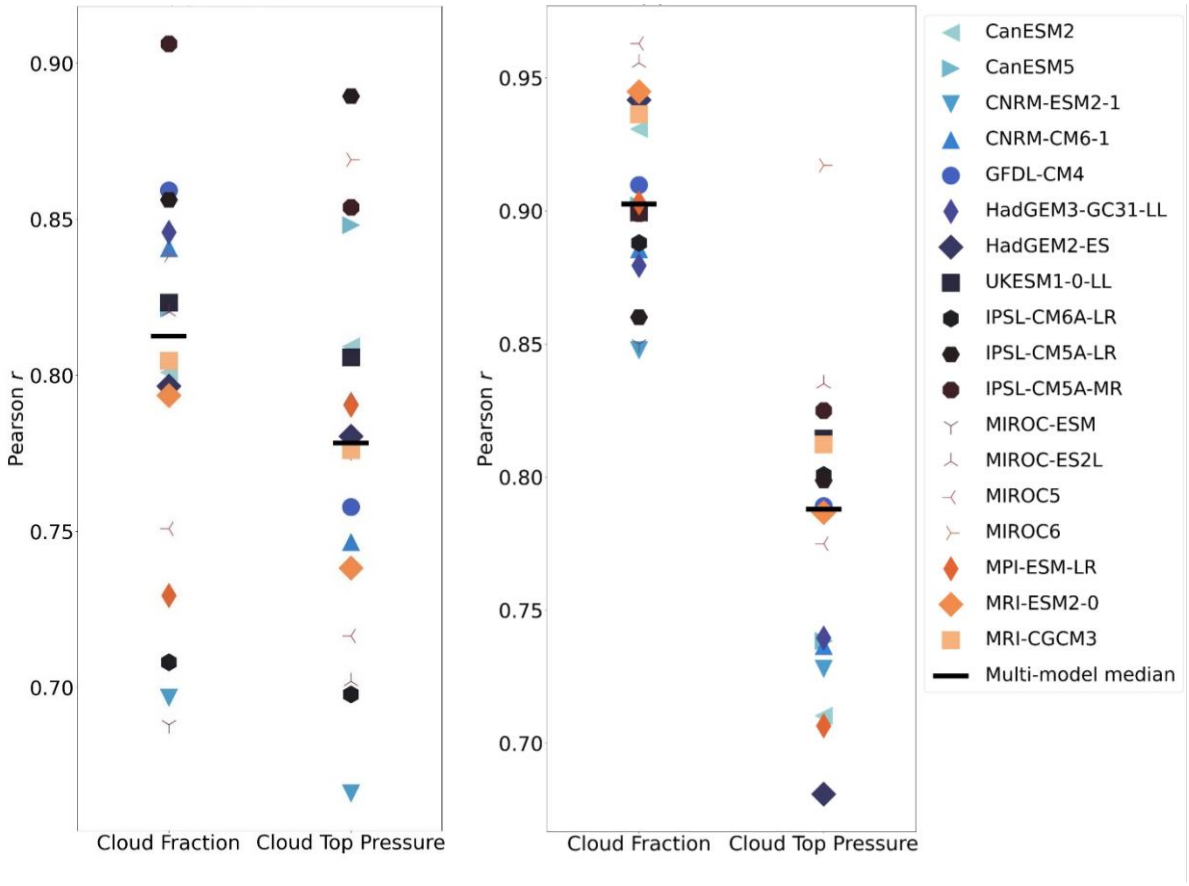


Figure S11. Scatter plot showing the correlation coefficients for the 20-year time series for the individual CMIP models and the multi-model median (left panel) globally-integrated $R_{LW,CF}$ and $R_{LW,CTP}$ (right panel) $R_{LW,CF}$ and $R_{LW,CTP}$ integrated in the tropical ascent region using configuration $S_{UT} + \Delta U_{300}$ (in addition to T_{sfc}, RH_{700}, UTRH, and ω_{300}) and a 21x11 domain.