Supplementary material

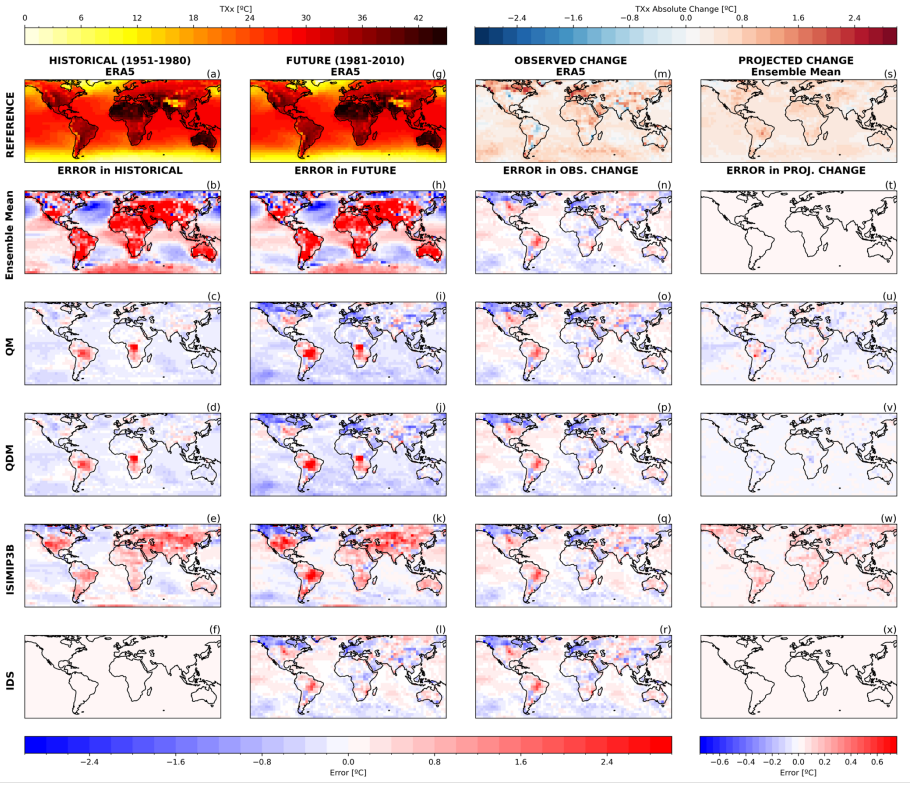


Figure S1. Spatial evaluation of the TXx indicator [°C] for historical (1951–1980) and future (1981–2010) periods. The top row shows ERA5 reference historical (a) and future (g) representations together with the corresponding observed changes (m) and raw multimodel ensemble mean model projected changes (s). Corresponding absolute differences [ºC] for raw CMCC-ESM2 model (b,h, n, t), QM (c, i, o, u), QDM (d, j, p, v), ISMIP3B (e, k, q, w) and IDS (f, l, r, x) bias-adjusted outputs relative to respective reference historical bias are shown.

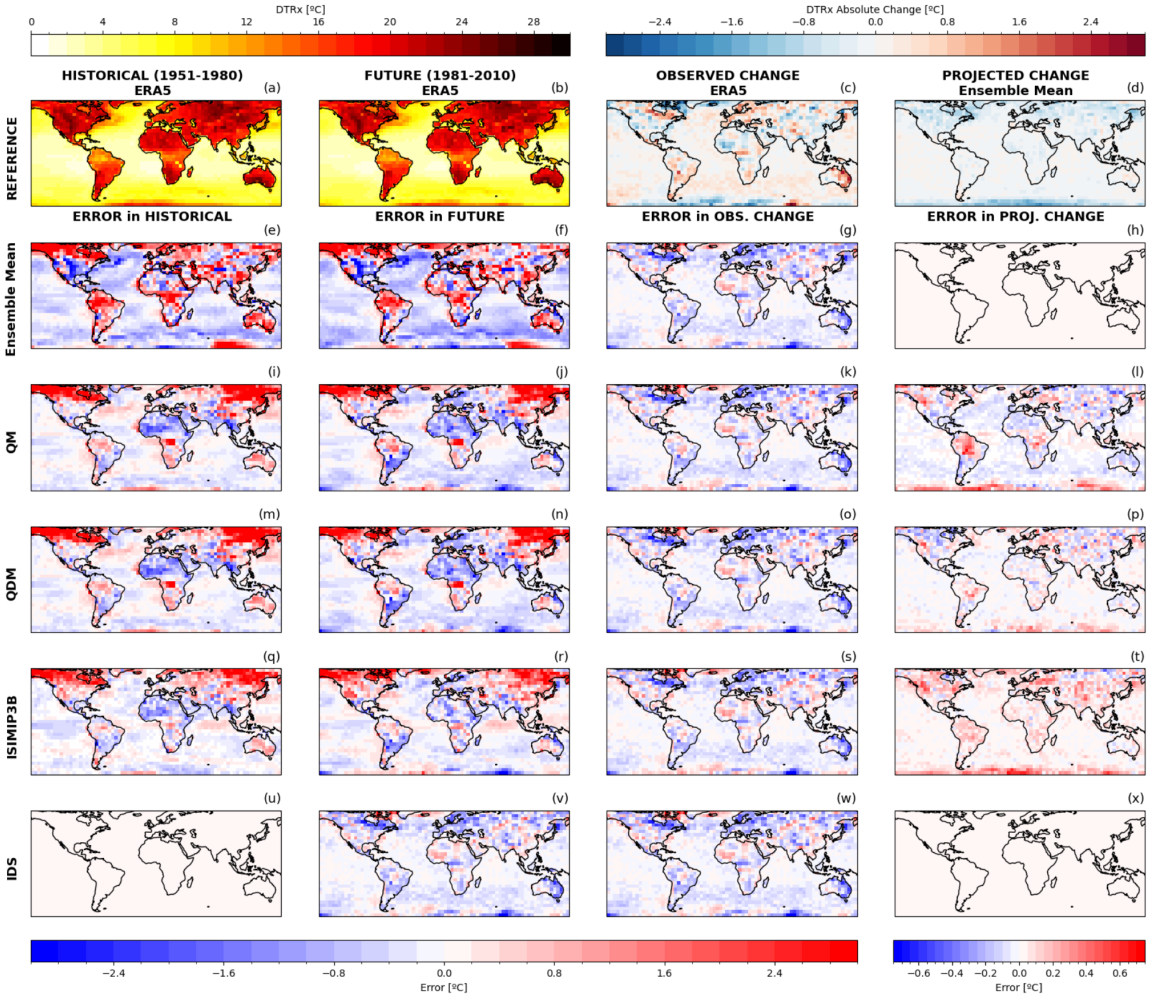
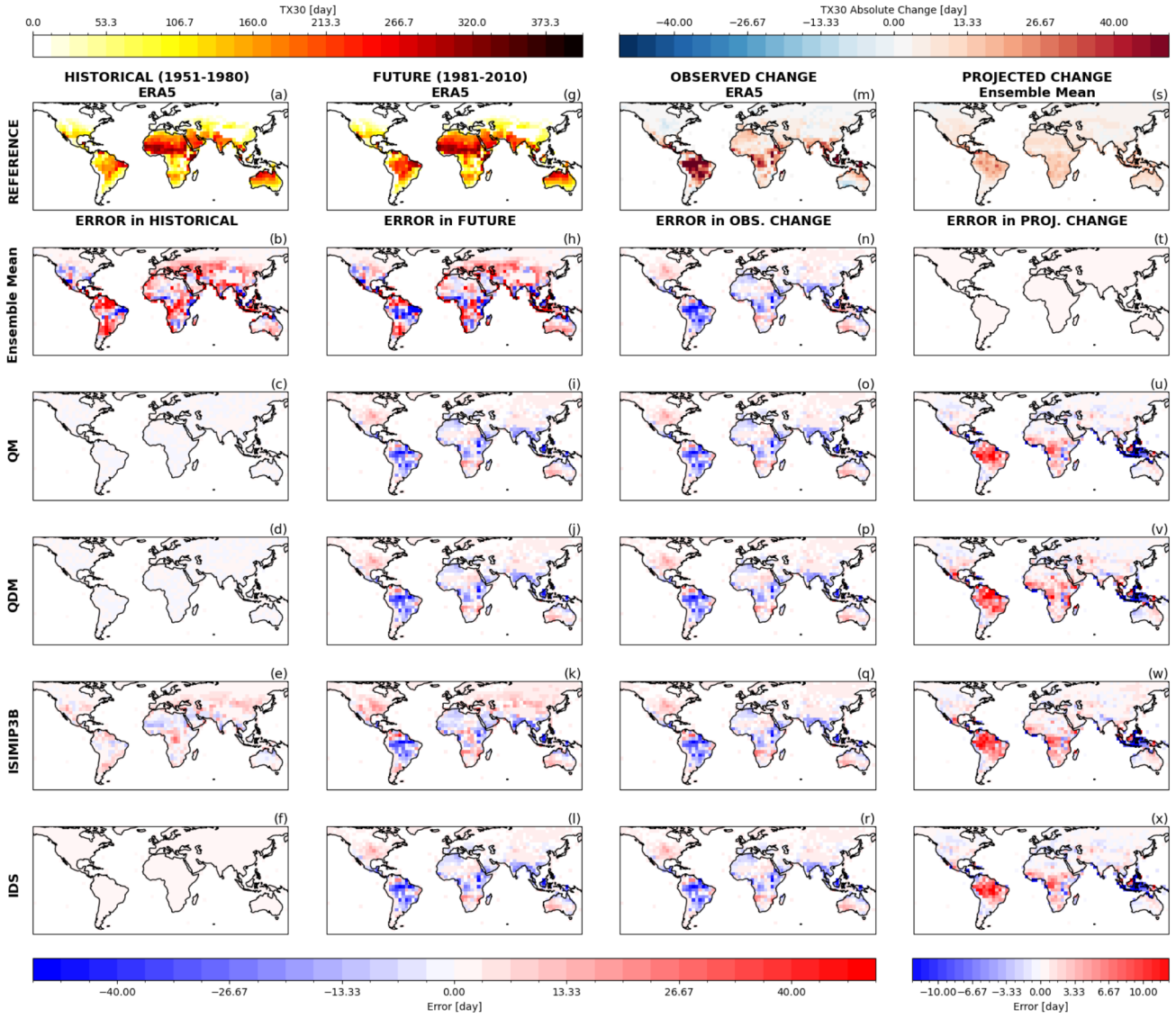


Figure S2. Spatial evaluation of the DTRx indicator [°C] for historical (1951–1980) and future (1981–2010) periods. The top row shows ERA5 reference historical (a) and future (g) representations together with the corresponding observed changes (m) and raw multimodel ensemble mean model projected changes (s). Corresponding absolute differences [ºC] for raw CMCC-ESM2 model (b,h, n, t), QM (c, i, o, u), QDM (d, j, p, v), ISMIP3B (e, k, q, w) and IDS (f, l, r, x) bias-adjusted outputs relative to respective reference historical bias are shown.



16

17 Figure S3. Spatial evaluation of the TX30 indicator [°C] for historical (1951–1980) and future (1981–2010) periods. The top
18 row shows ERA5 reference historical (a) and future (g) representations together with the corresponding observed changes (m)
19 and raw multimodel ensemble mean model projected changes (s). Corresponding absolute differences [ºC] for raw CMCC-
20 ESM2 model (b,h, n, t), QM (c, i, o, u), QDM (d, j, p, v), ISMIP3B (e, k, q, w) and IDS (f, l, r, x) bias-adjusted outputs relative
21 to respective reference historical bias are shown. For this indicator, pixels on land are only considered.

22

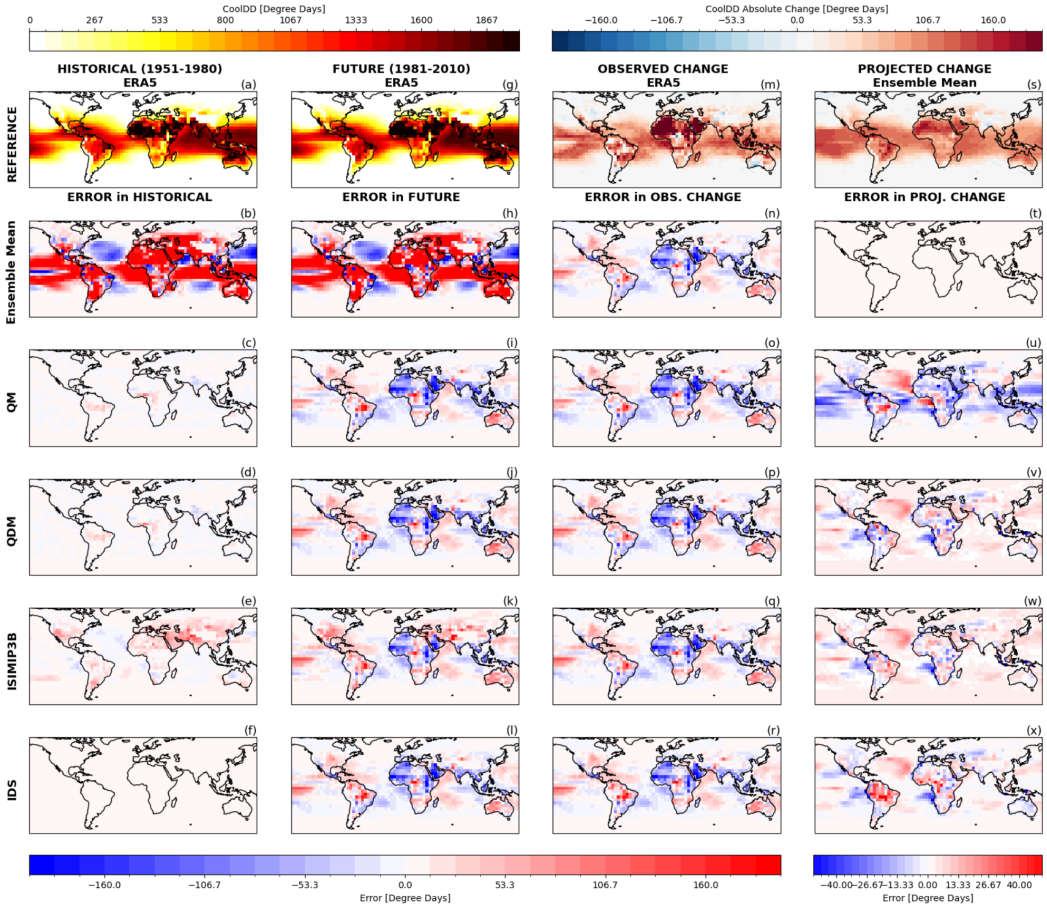


Figure S4. Spatial evaluation of the CoolDD indicator [Degree Days] for historical (1951–1980) and future (1981–2010) periods. The top row shows ERA5 reference historical (a) and future (g) representations together with the corresponding observed changes (m) and raw multimodel ensemble mean model projected changes (s). Corresponding absolute differences [Degree Days] for raw CMCC-ESM2 model (b,h, n, t), QM (c, i, o, u), QDM (d, j, p, v), ISMIP3B (e, k, q, w) and IDS (f, l, r, x) bias-adjusted outputs relative to respective reference historical bias are shown. For this indicator, pixels on land are only considered.

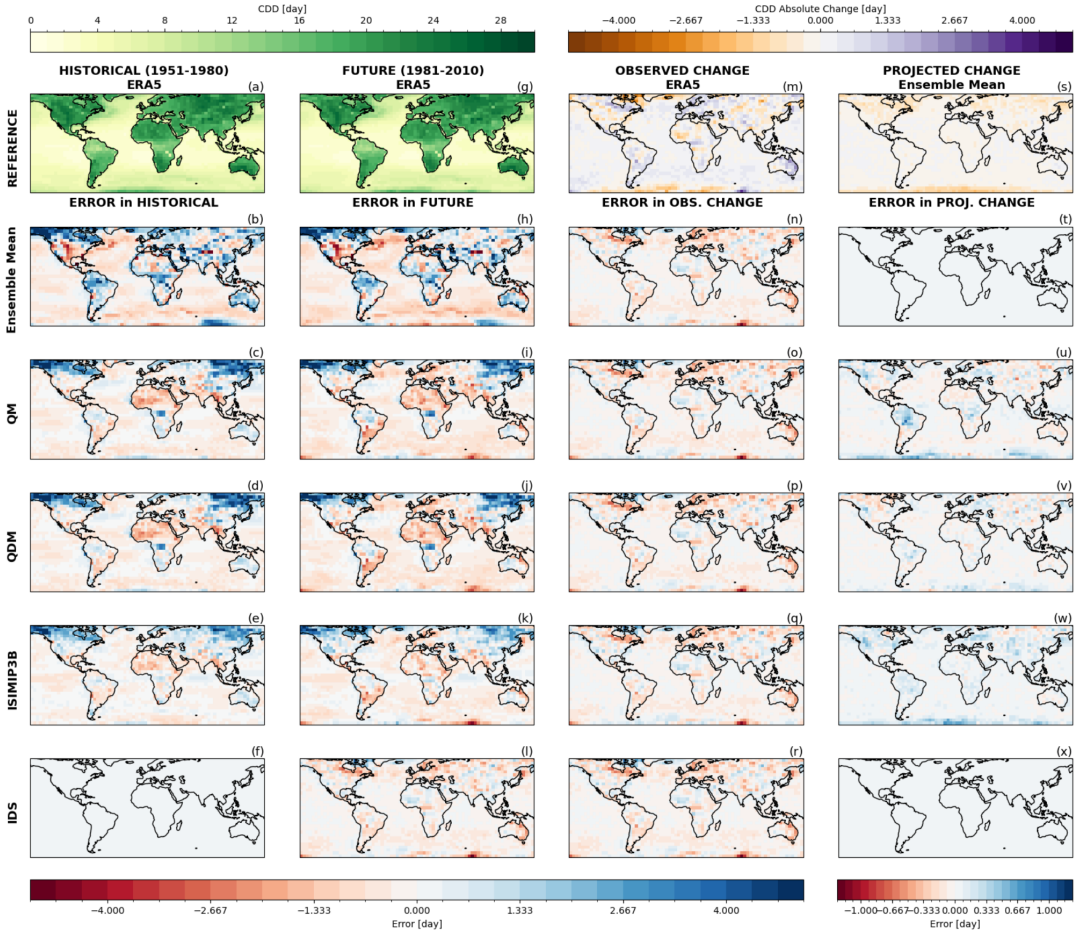


Figure S5. Spatial evaluation of the CDD indicator [Day] for historical (1951–1980) and future (1981–2010) periods. The top row shows ERA5 reference historical (a) and future (g) representations together with the corresponding observed changes (m) and raw multimodel ensemble mean model projected changes (s). Corresponding absolute differences [Day] for raw CMCC-ESM2 model (b,h, n, t), QM (c, i, o, u), QDM (d, j, p, v), ISMIP3B (e, k, q, w) and IDS (f, l, r, x) bias-adjusted outputs relative to respective reference historical bias are shown. For this indicator, pixels on land are only considered.

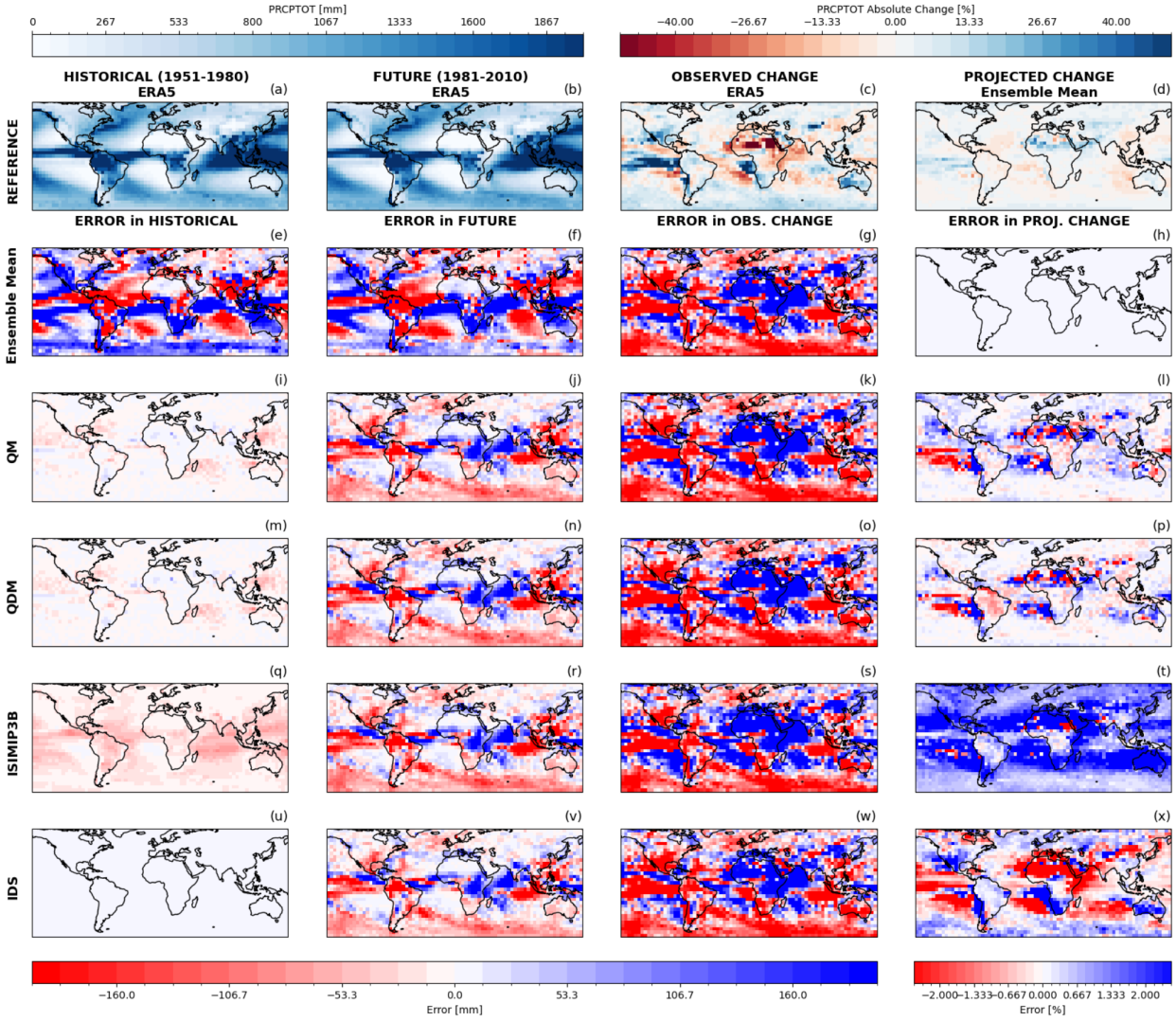
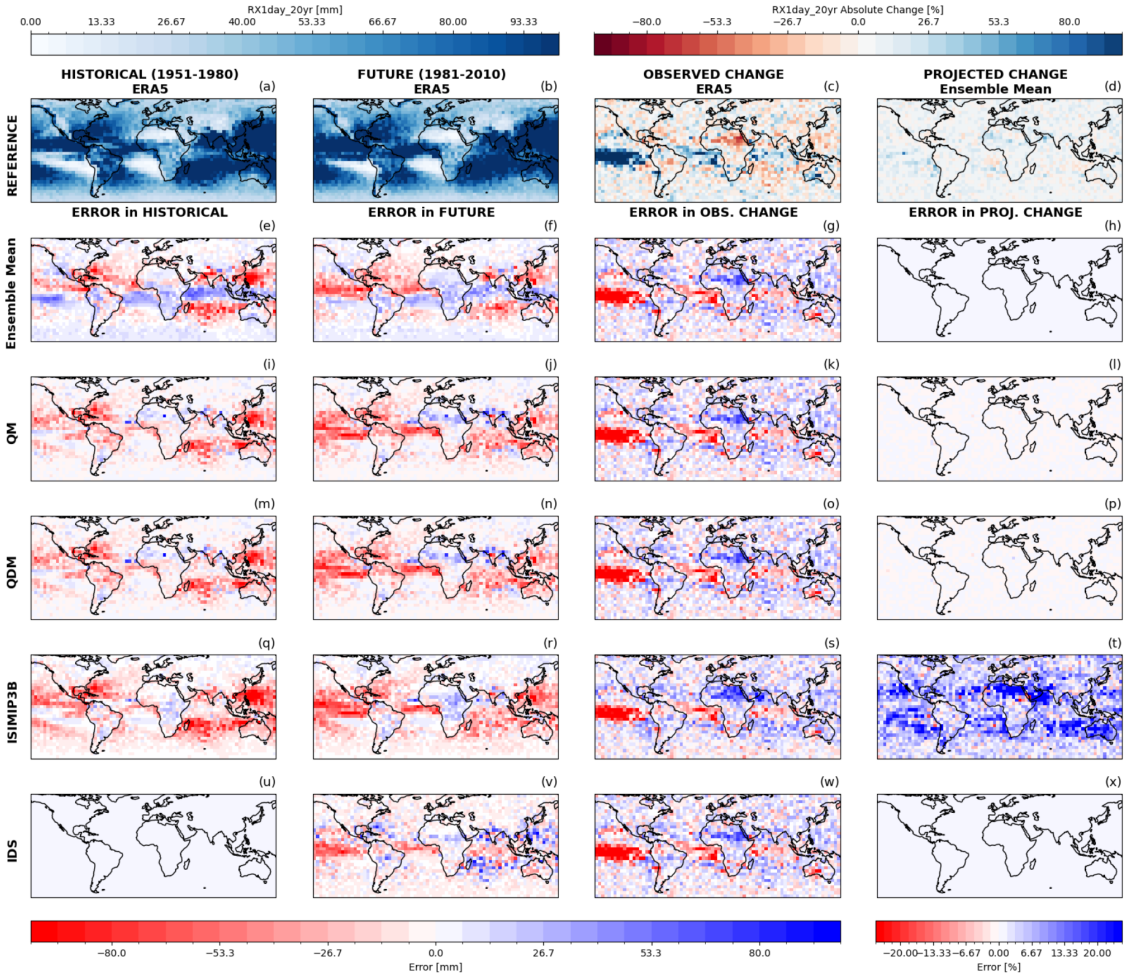


Figure S6. Spatial evaluation of the PRCPTOT indicator [mm] for historical (1951–1980) and future (1981–2010) periods. The top row shows ERA5 reference historical (a) and future (g) representations together with the corresponding observed changes (m) and raw multimodel ensemble mean model projected changes (s). Corresponding absolute differences [mm] for raw CMCC-ESM2 model (b,h, n, t), QM (c, i, o, u), QDM (d, j, p, v), ISMIP3B (e, k, q, w) and IDS (f, l, r, x) bias-adjusted outputs relative to respective reference historical bias are shown. For this indicator, pixels on land are only considered.



47

48 Figure S7. Spatial evaluation of the RX1day_20yr indicator [mm/day] for historical (1951–1980) and future (1981–2010)
49 periods. The top row shows ERA5 reference historical (a) and future (g) representations together with the corresponding
50 observed changes (m) and raw multimodel ensemble mean model projected changes (s). Corresponding absolute differences
51 [mm/day] for raw CMCC-ESM2 model (b,h, n, t), QM (c, i, o, u), QDM (d, j, p, v), ISMIP3B (e, k, q, w) and IDS (f, l, r, x)
52 bias-adjusted outputs relative to respective reference historical bias are shown. For this indicator, pixels on land are only
53 considered.

54

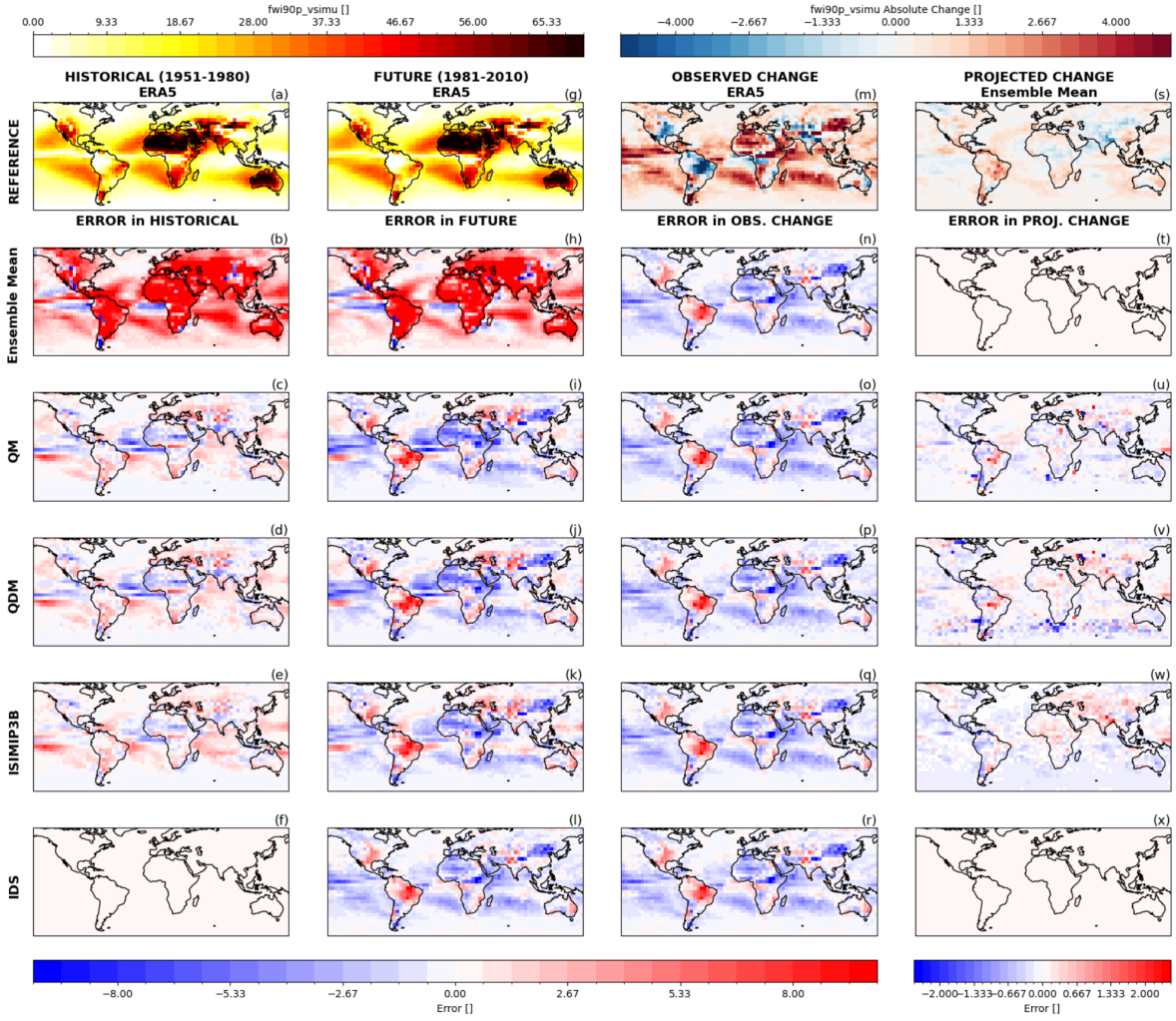


Figure S8. Spatial evaluation of the FWI90p indicator [ ] for historical (1951–1980) and future (1981–2010) periods. The top row shows ERA5 reference historical (a) and future (g) representations together with the corresponding observed changes (m) and raw multimodel ensemble mean model projected changes (s). Corresponding absolute differences [] for raw CMCC-ESM2 model (b,h, n, t), QM (c, i, o, u), QDM (d, j, p, v), ISMIP3B (e, k, q, w) and IDS (f, l, r, x) bias-adjusted outputs relative to respective reference historical bias are shown. For this indicator, pixels on land are only considered.

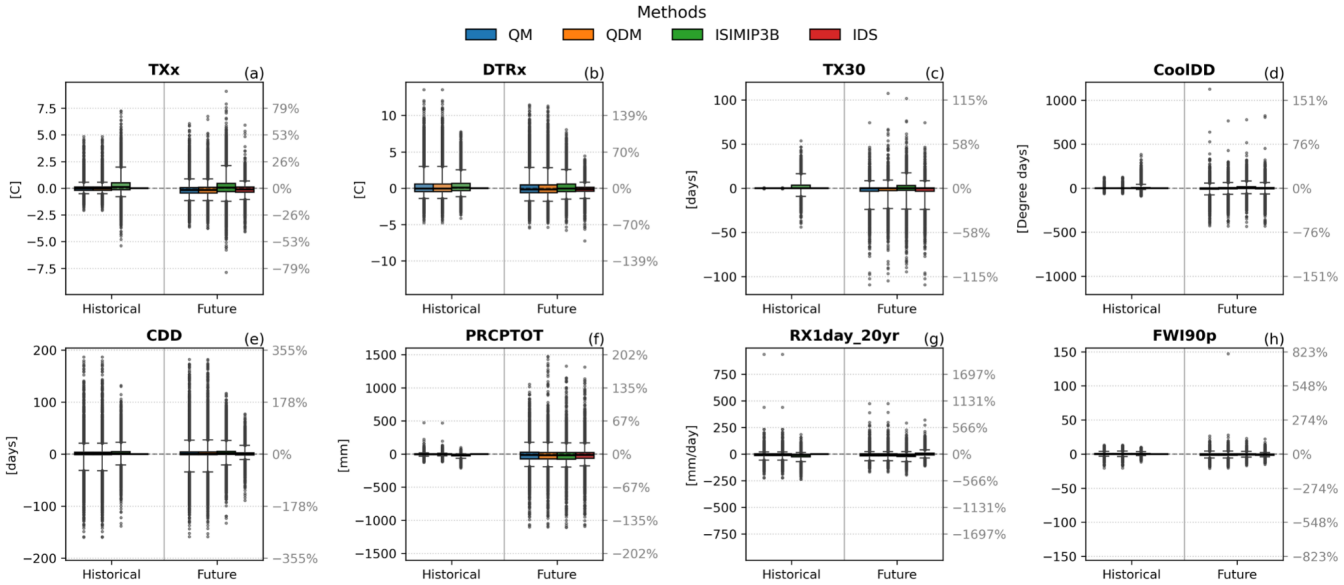


Figure S9. Absolute errors considering all locations and CMIP6 models between bias-adjusted estimates and observational reference for the four BA methods are shown for both historical (left) and future (right) periods across the eight climate indicators: TXx (a), DTR (b), TX30 (c), CoolDD (d), CDD (e), PRCPTOT (f), RX1day_20yr (g) and FWI90p (h). Boxplots indicate the 25, 50 and 75th percentiles together with whiskers covering up to the 5 and 95th percentiles of the samples Outliers are shown represented by black dots. The right vertical axis shows the normalized error obtained by dividing the absolute error by the cGSV.

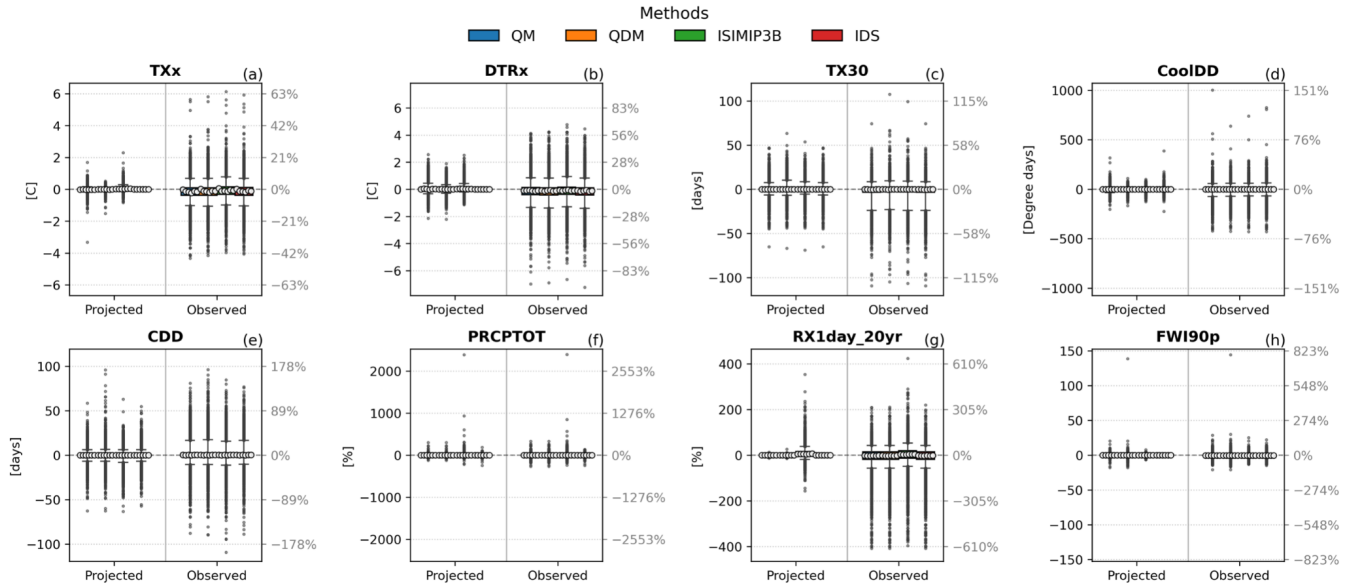


Figure S10. Absolute and relative errors [%] considering all locations and CMIP6 models between bias-adjusted estimates and model projected (left) / observed (right) trends are shown for the four BA methods across the eight climate indicators: TXx (a), DTR (b), TX30 (c), CoolDD (d), CDD (e), PRCPTOT (f), RX1day_20yr (g) and FWI90p (h). Boxplots indicate the 25, 50 and 75th percentiles together with whiskers covering up to the 5 and 95th percentiles of the samples. Outliers are shown represented by black dots. The right vertical axis shows the normalized error obtained by dividing the absolute error by the cGSV. See text for relative error normalization.

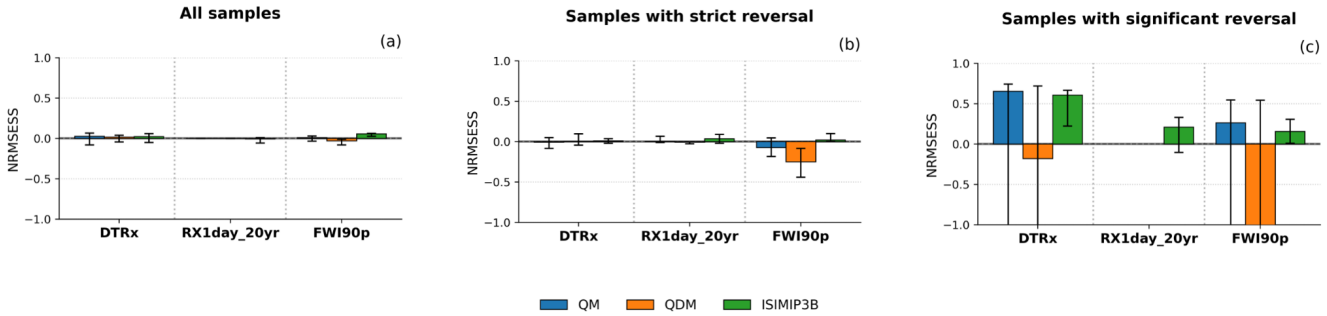


Figure S11. a) Multimodal ensemble median of the NRMSESS for the three indicators analysed in Fig. 5 including all locations for each of the BA methods (in color). Same figure but now including only those samples with strict b) and significant c) reversals of model projected change. The minimum and maximum of the ensemble is indicated with a vertical segment. Note that the larger the NRMSESS is, the more accurate the method is. Negative values are capped at -1.