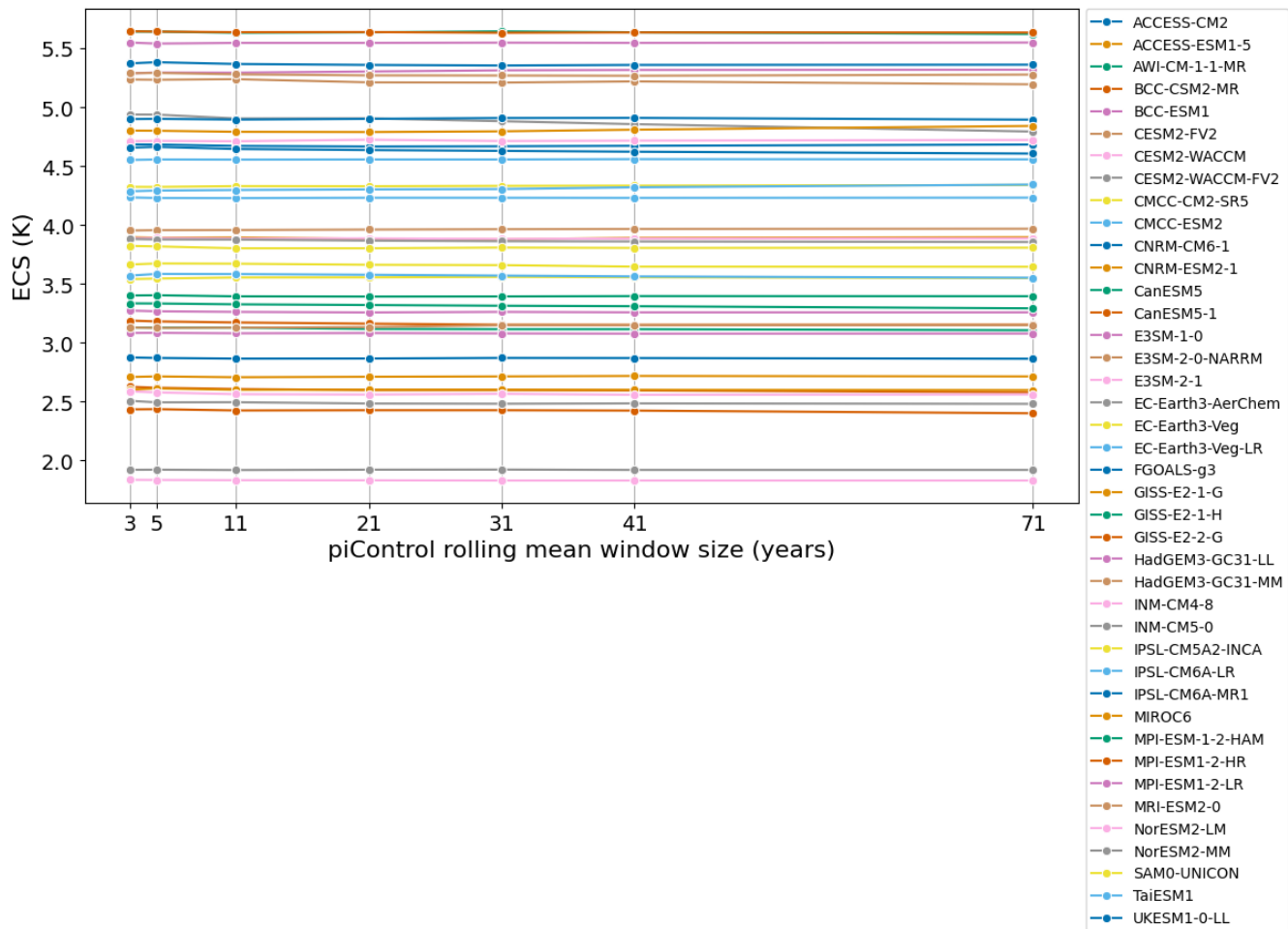


*Supplement of*

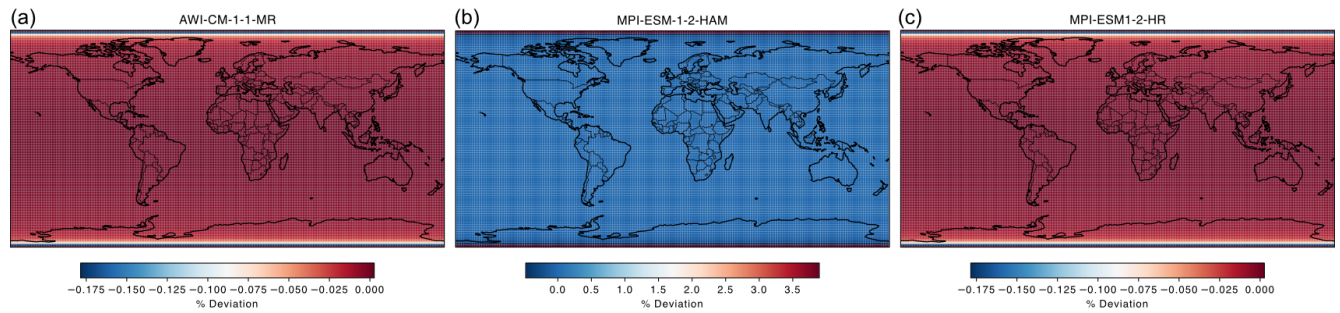
# **Standardising the “Gregory method” for calculating equilibrium climate sensitivity**

**Anna Zehrung et al.**

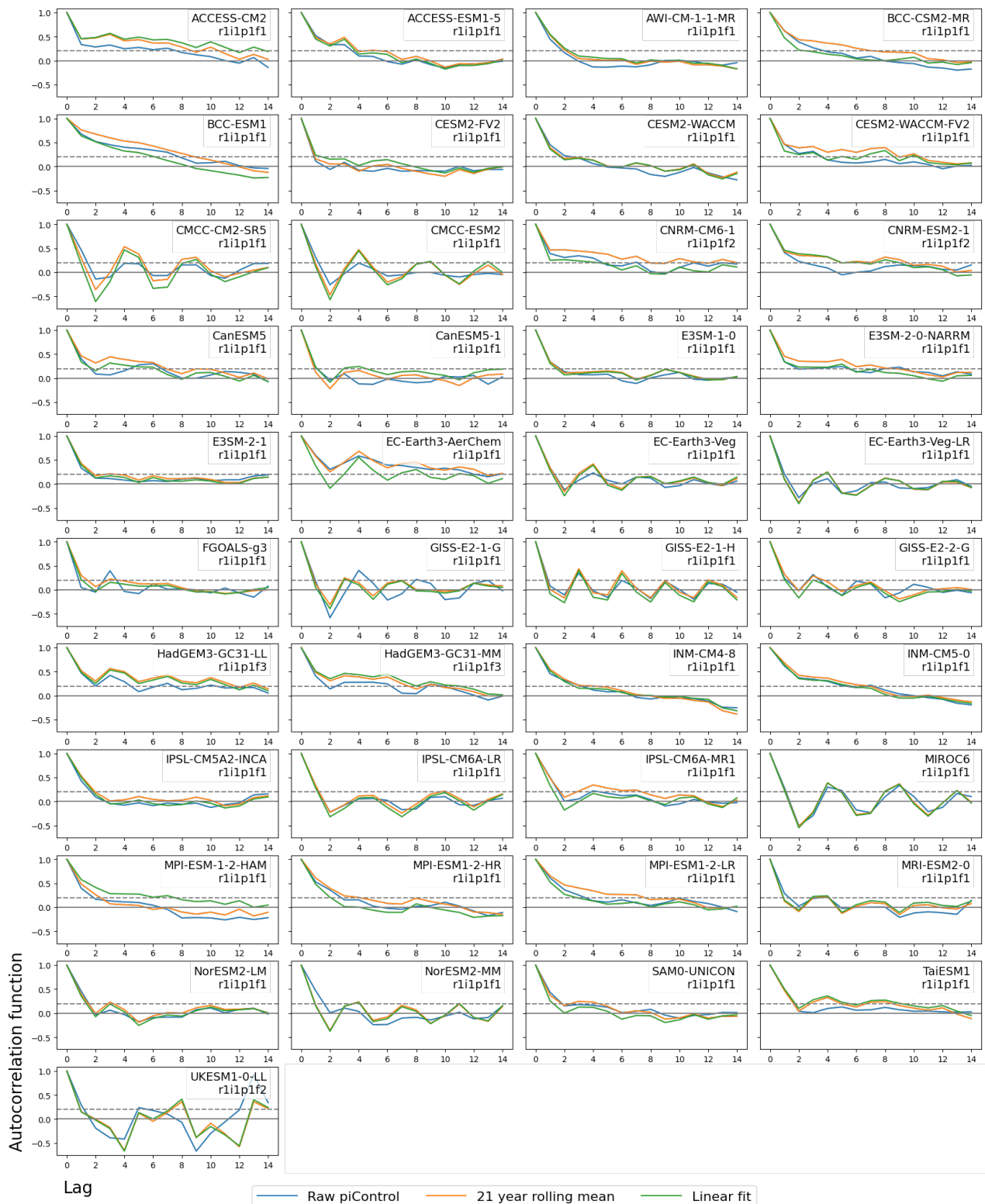
*Correspondence to:* Anna Zehrung (azehrung@student.unimelb.edu.au)



**Figure S1.** The ECS estimates using ordinary least squares regression calculated using different rolling mean window sizes over the piControl during the anomaly calculation.

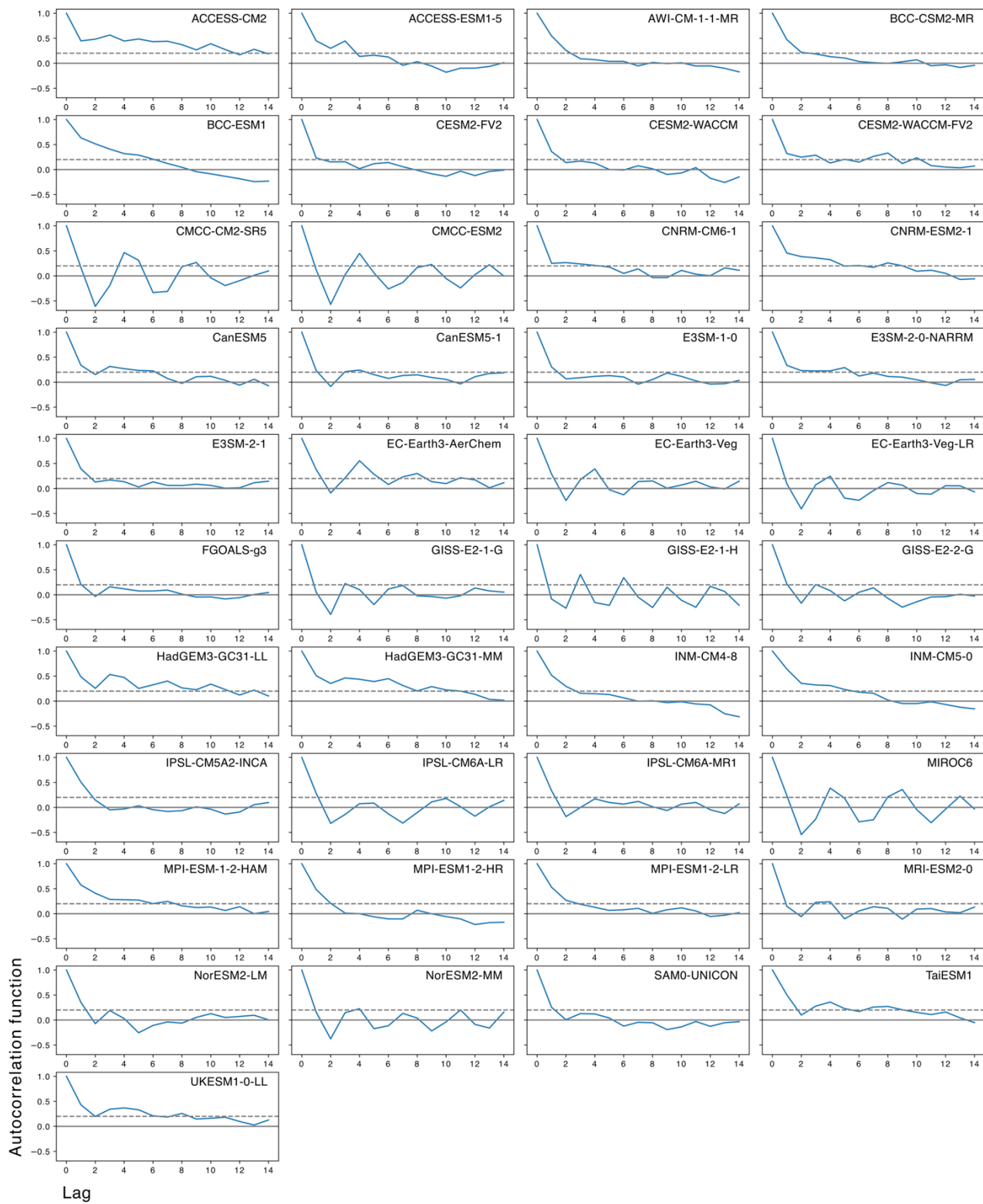


**Figure S2.** Showing cell area relative to the  $\cos(\text{lat})$  approximation for the global mean weighting outlier models. **a)** AWI-CM-1-1-MR, **b)** MPI-ESM-1-2-HAM, **c)** MPI-ESM1-2-HR. Note the differences in gradient scale for the percent deviation.



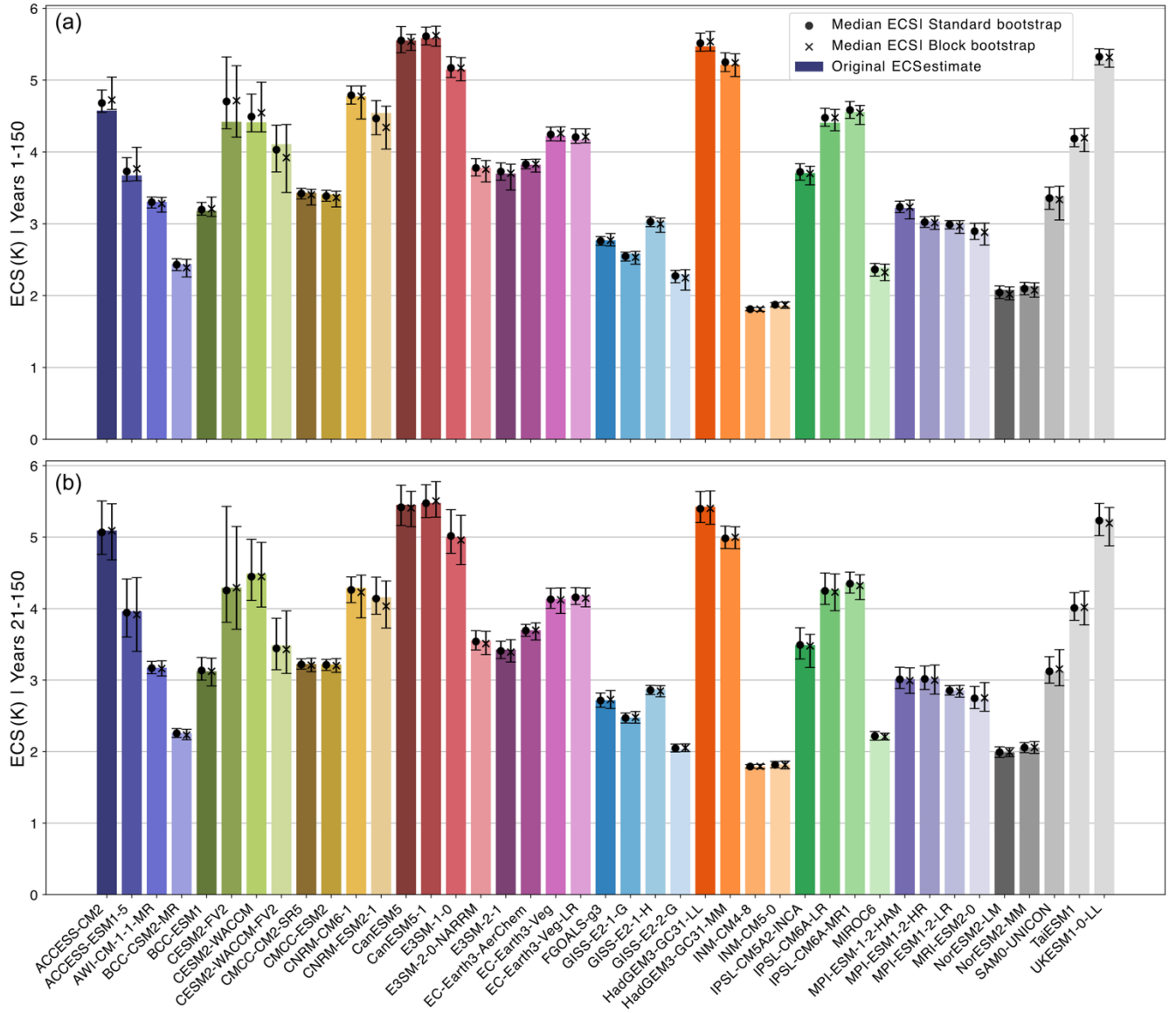
15 **Figure S3.** The autocorrelation function of surface air temperature anomaly timeseries comparing the three alternate anomaly methods (with global mean weighted by cell area), lags 1 to 15 years. The autocorrelation function shows the residuals calculated following the removal of a quadratic fit to years 21-150 of the anomaly between the piControl and abrupt-4xCO<sub>2</sub>. We exclude the first 20 years of the timeseries in the interests of optimising the fit over the data to calculate the residuals. We define an autocorrelation value of less than 0.2 as representing decorrelation.

20



**Figure S4.** The autocorrelation function of abrupt-4xCO<sub>2</sub> surface air temperature timeseries, lags 1 to 15 years. The autocorrelation function shows the residuals calculated following the removal of a quadratic fit to years 21-150. Note that this is not showing anomalies, just the raw perturbed data. We define an autocorrelation value of less than 0.2 as representing  
25 decorrelation.

30



**Figure S5.** ECS uncertainty using a total least squares regression fit. **a)** ECS estimates for each model using the baseline Gregory Method, using years 1-150. Bars represent 95% confidence intervals, with medians calculated using a simple bootstrap (solid circle) and a moving block bootstrap with a block size of 4 (cross). **b)** The same as (a), but the ECS and bootstrap uncertainties are calculated using years 21-150 of the RNDT and TAS anomaly timeseries. See Methods for details on confidence interval calculations.