Supplement of The Cryosphere, 20, 2317–2329, 2026
https://doi.org/10.5194/tc-20-2317-2026-supplement


The Cryosphere
Open Access

*Supplement of*

# Changes in 1958–2019 Greenland surface mass balance are attributable to both greenhouse gases and anthropogenic aerosols

**Yan-Ning Kuo et al.**

*Correspondence to:* Yan-Ning Kuo (yk545@cornell.edu)

**S1 Bayesian Total Least Square (TLS) regression set-up**

In this paper, we use PyMC library to implement a Markov Chain Monte Carlo (MCMC) Bayesian TLS regression for detection and attribution (D&A) for Greenland surface mass balance (SMB), precipitation (P), and runoff (R) changes. A latent variable $x_{latemt}$, $x_{latent} \sim N(x_{obs}, \sigma_{x_{obs}}^{2})$ , is introduced in the algorithm to account for the uncertainty in observing the true $x_{obs}$ to perform a Bayesian TLS regression. Table S1 summarizes the algorithm, and the priors of parameters to be estimated.

An idealized numerical experiment is designed to compare the Bayesian TLS regression coefficients (with PyMC) against the deterministic TLS regression coefficients. First, we create toy data sets where $x_{true}$ is an evenly spaced data from 1 to 10 with spacing 0.5, $x_{obs} = x_{true} + \varepsilon_{x}$ to mimic the uncertainty when observing $x$. We define the $y_{obs} = \beta x_{true} + \varepsilon_{y}$ and $\beta = 1.0$. $\varepsilon_{x} \sim N(0, 0.3^{2})$ and $\varepsilon_{y} \sim N(0, 0.5^{2})$. We generate 50-ensemble random numbers for $\varepsilon_{x}$, $\varepsilon_{y}$ to create 50-pair of $x_{obs}$, $y_{obs}$ to solve the regression coefficient β with a deterministic Ordinary Least Square (OLS) regression, a deterministic TLS regression (through Singular Value Decomposition, SVD), and a Bayesian TLS regression (with PyMC) to compare their regression coefficients. Comparing with the true regression line (black dashed line in Figure S1), the 50-member mean regression line based on β from the Bayesian TLS regression (green line in Figure S1a) is less biased upward compared with the deterministic TLS regression line (blue line in Figure S1a). In addition, comparing the distribution of the solutions of β from these 50-pair of $x_{obs}$, $y_{obs}$, we show Bayesian TLS regression is systematically less upward biased than the deterministic TLS regression (Figure S1b).

As $x$ represents the forced response, the priors of $\sigma_{x}$ for detection and attribution (D&A) in the main manuscript are estimated as the standard error for the ensemble mean simulated by CESM2 (for example, the standard error of 50-member mean of simulated SMB from ALL is calculated as the standard deviation of the 50-member SMB divided by a square root of 50) and $x_{latent}$ is with a prior centered of the true ensemble mean with a spread of $\sigma_{x}$. The prior of $y$ is estimated by ERA-RACMO simulation, where the $\sigma_{y}$ is the standard deviation of the time series of y.

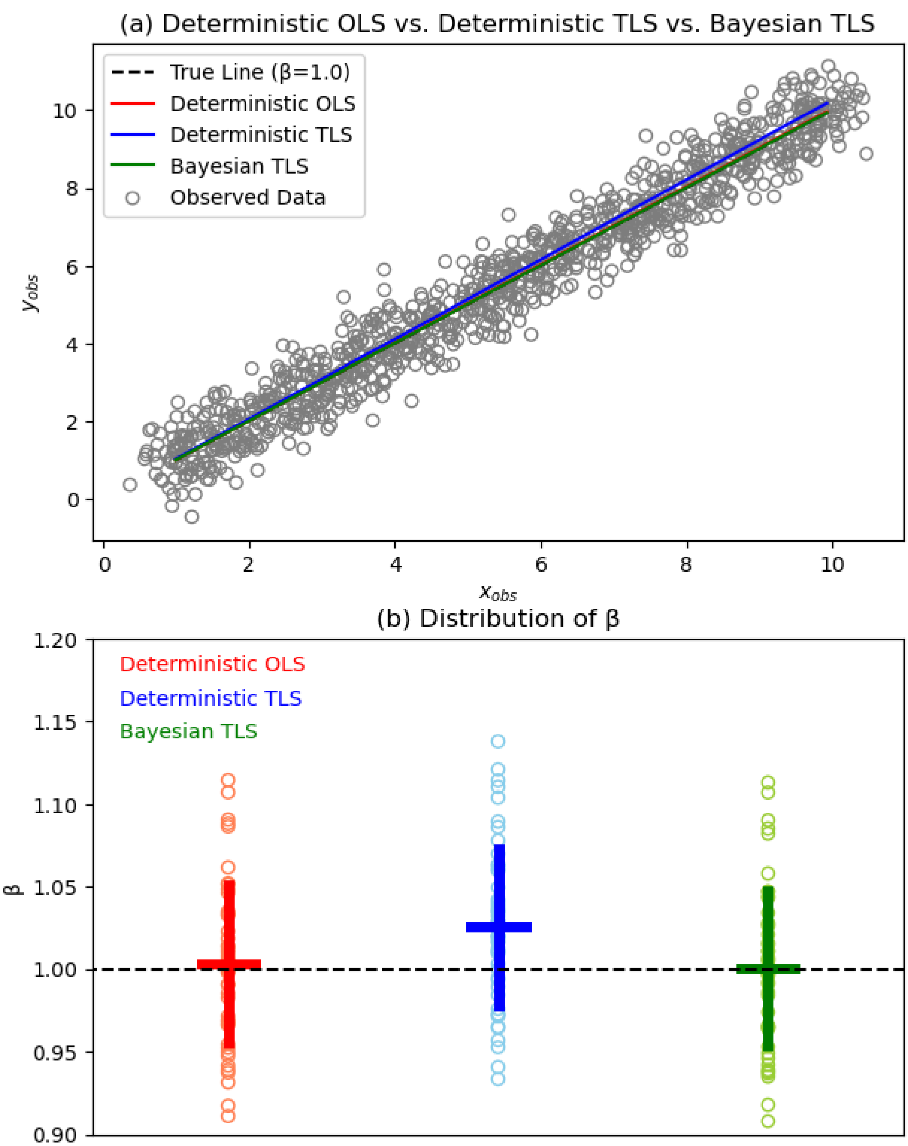


**Figure S1.** 50-member idealized numerical experiments with toy data to compare the performances of a deterministic Ordinary Least Square (OLS) regression, a deterministic Total Least Square (TLS) regression, and a Bayesian TLS regression. (a) Mean solutions from the three different regressions for the regression slope. (b) The distributions of the regression coefficients from these three regressions. The horizontal bars indicate the 50-member mean regression coefficients and the vertical bars show the range $\pm1$ standard deviation of the regression coefficients around the 30-member mean.

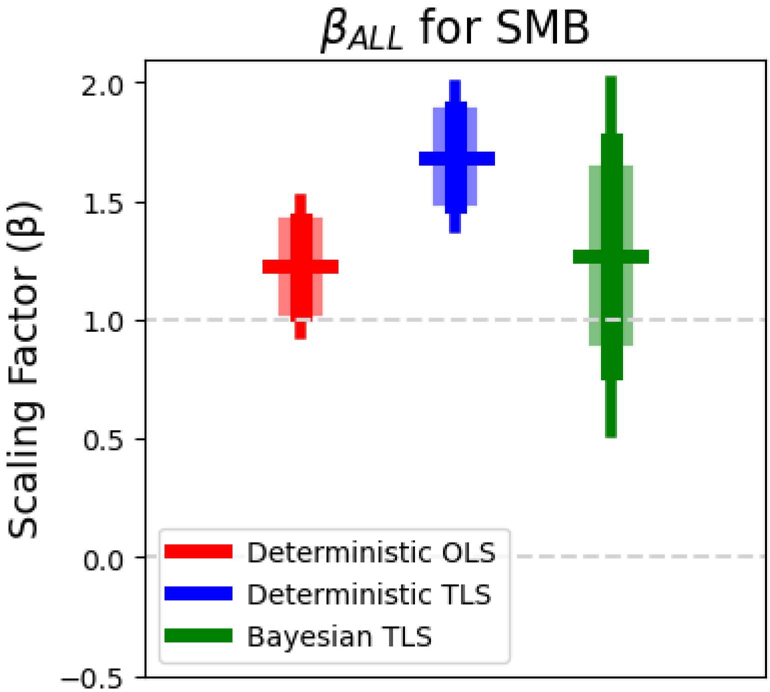


**Figure S2.** The comparison of scaling factor (β) attributing GrIS SMB changes to ALL historical forcings with CESM2-LE with a deterministic Ordinary Least Square (OLS) regression, a deterministic Total Least Square (TLS) regression, and a Bayesian TLS regression. The light shading/thick vertical line/thin vertical line is $\pm 1/1.64/2.57$ standard deviation of the 10,000 posterior samples of β, indicating the 66%/ 90%/ 99% confidence intervals in Detection and Attribution (D&A) (*likely*/ *very likely*/ *virtually certain* based on IPCC guideline).

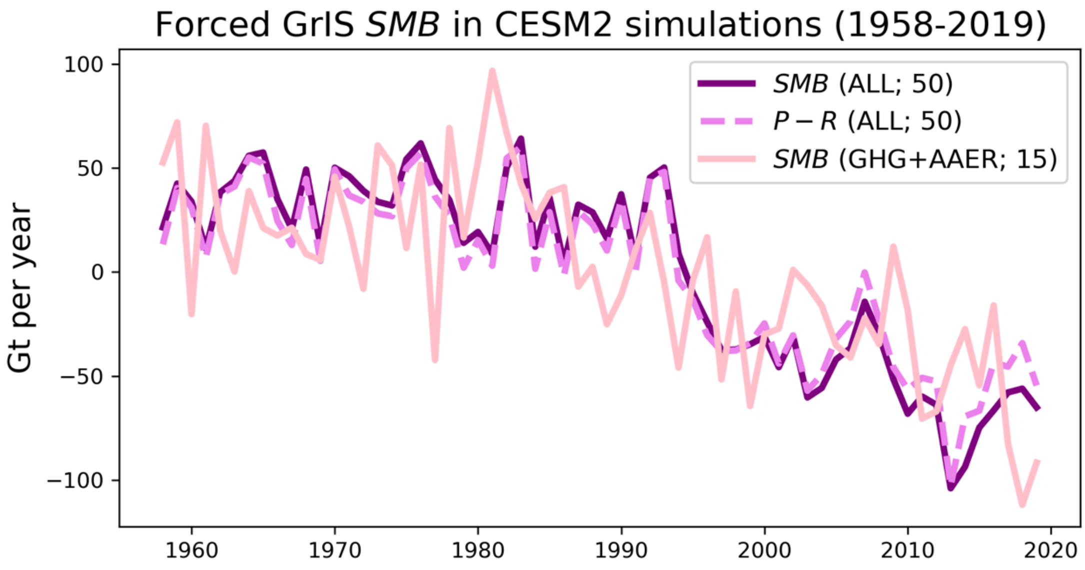


**Figure S3.** Comparison between the 50-member ensemble mean all forcings (ALL)-forced surface mass balance (SMB) (purple solid line) against the 50-member ensemble mean ALL-forced preciptiation-minus-runoff (*P*-*R*) (light purple dashed line) and the sum of 15-member ensemble mean greenhouse gases (GHG)-forced and 15-member ensemble mean anthropogenic aerosols (AAER)-forced SMB (GHG+AAER) (pink) for 1958–2019.

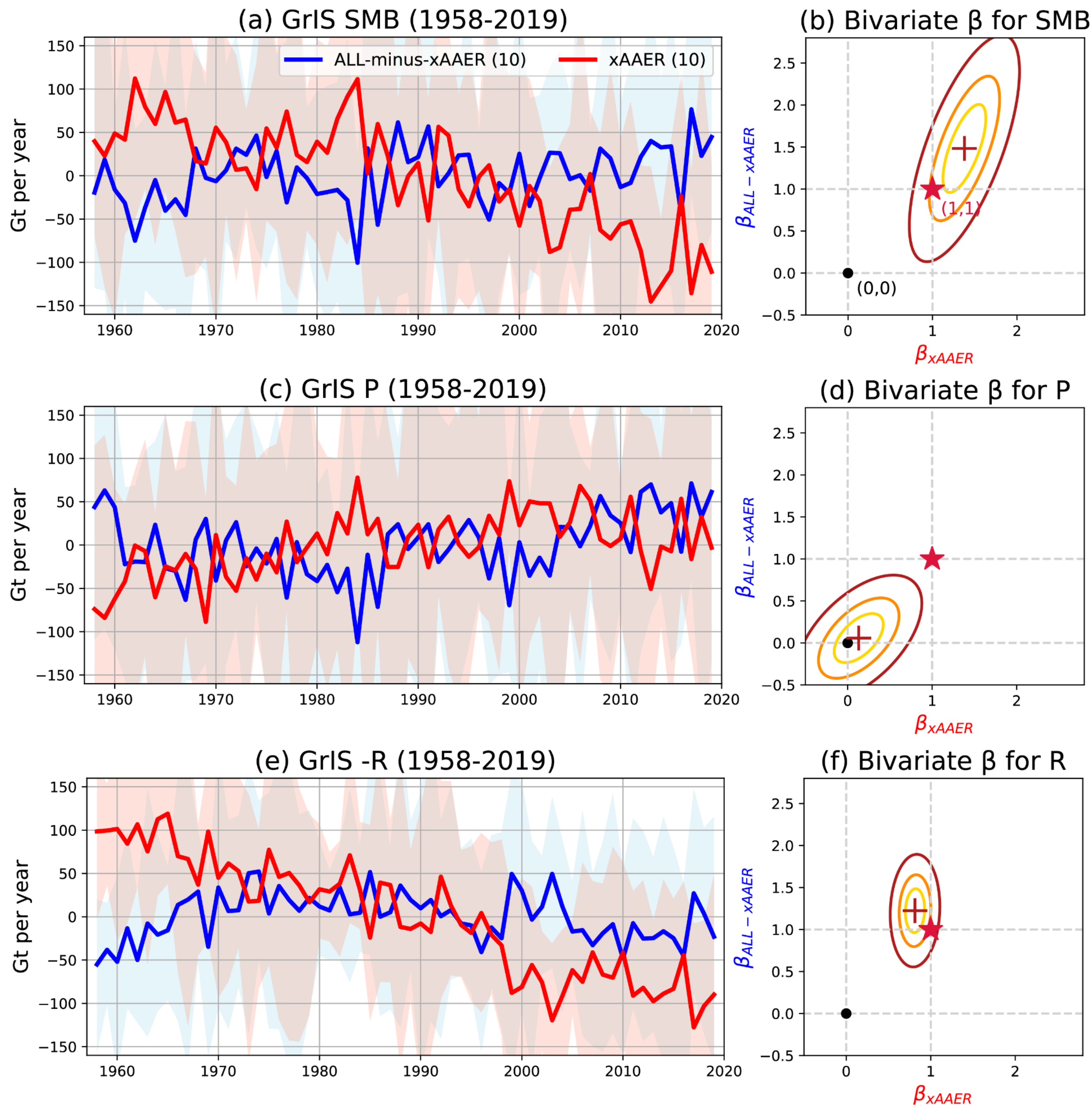


**Figure S4.** Same as Figure 3 but with ALL-minus-xAAER and xAAER for the bivariate detection and attribution (D&A) analyses for 1958–2019. The uncertainty ellipses consider the covariance between two β and different sets of simulations included for the Bayesian TLS regression might affect the orientation of the ellipses. The relative spread in the x and y directions is largely the same as Figure 3, which is the main focus for the comparisons in our study.

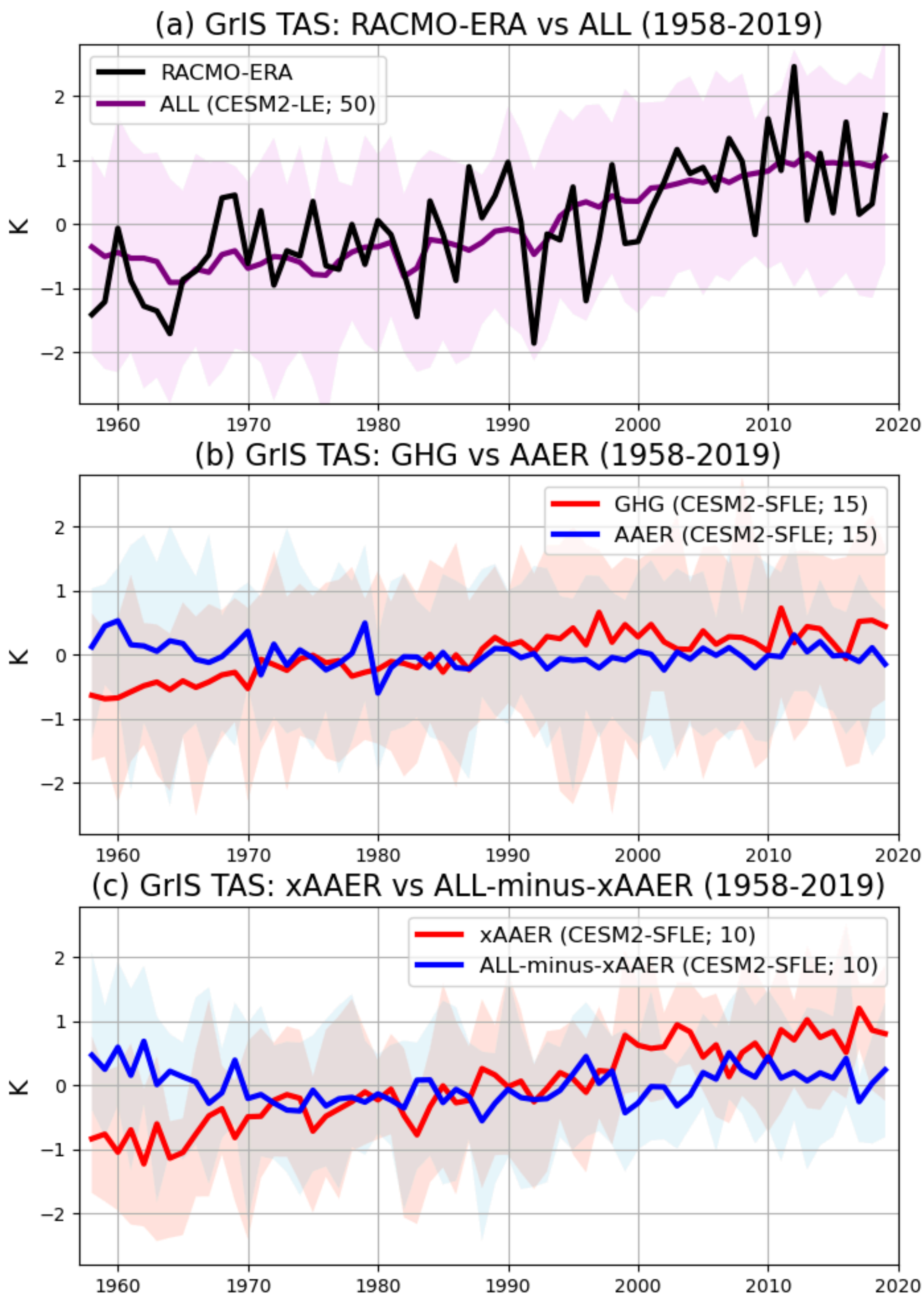


**Figure S5.** Time series of GrIS TAS since 1958 from (a) RACMO-ERA (black) and ensemble mean of CESM2-LE (purple), (b) ensemble means of GHG and AAER from CESM2-SFLE, and (c) ensemble means of xAAER and ALL-minus-xAAER from CESM2-SFLE. The shadings are the full range across ensemble members (i.e., maxima to minima).

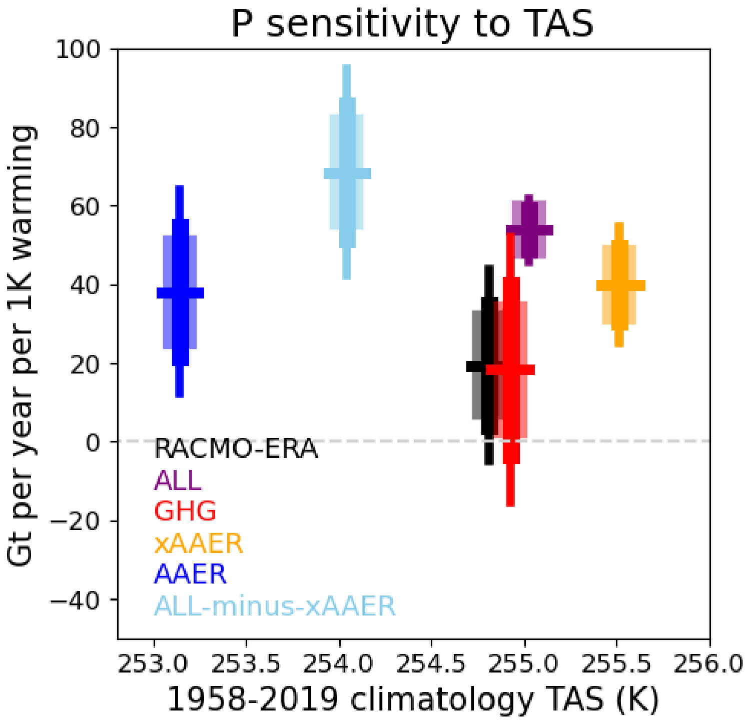


**Figure S6.** The observed (RACMO-ERA; black) and forced (ALL, GHG, xAAER, AAER, ALL-minus-xAAER; purple, red, orange, blue, light blue respectively) GrIS precipitation (*P*) sensitivity to near-surface temperature (TAS) during 1958-2019 sorted by the climatological mean near-surface temperature.

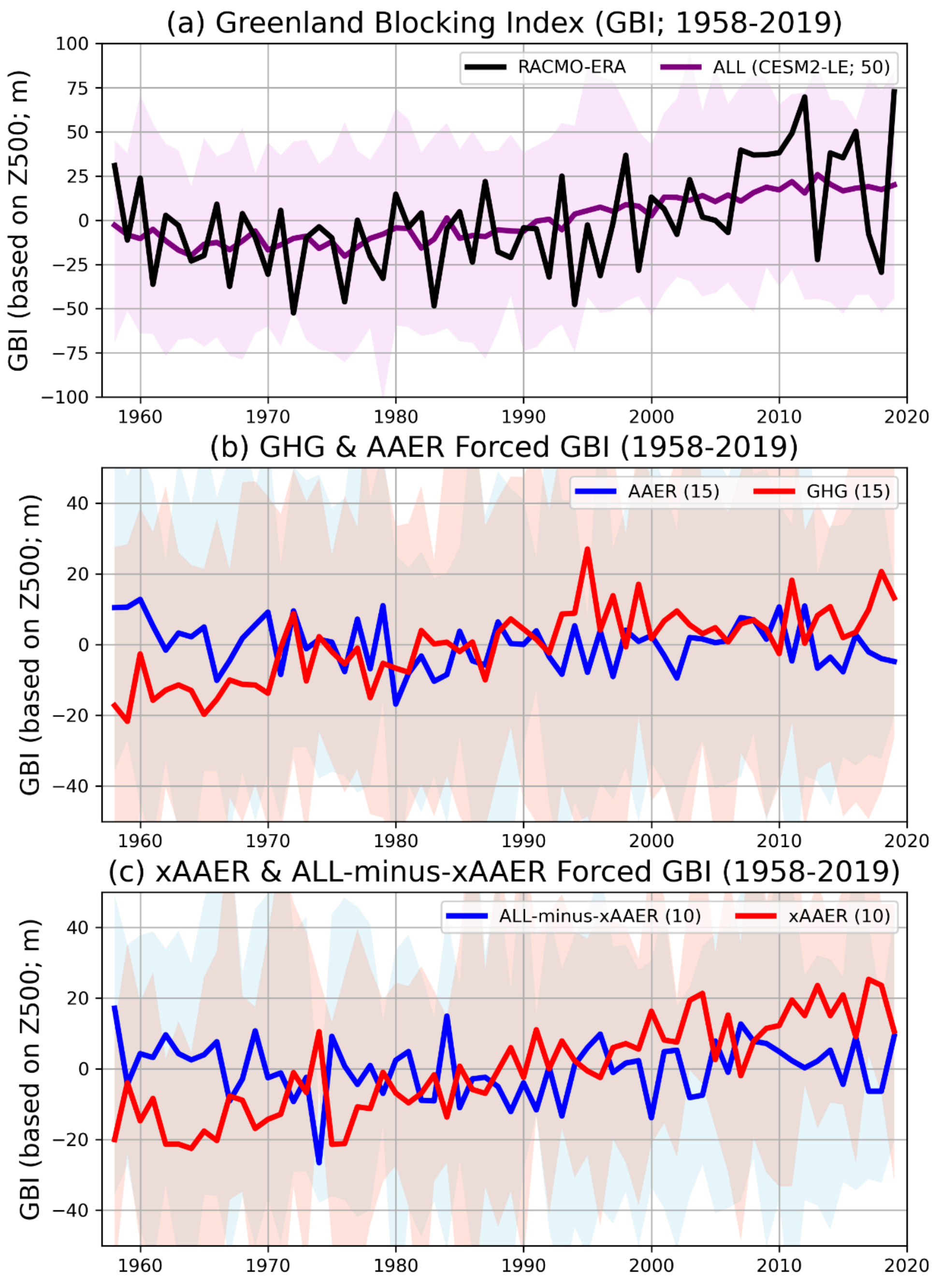


**Figure S7.** Time series of Greenland Blocking Index (GBI) since 1958 from (a) RACMO-ERA (black) and and ensemble mean of CESM2-LE (purple), (b) ensemble means of GHG and AAER from CESM2-SFLE, and (c) ensemble means of xAAER and ALL-minus-xAAER from CESM2-SFLE. The shadings are the full range across ensemble members (i.e., maxima to minima). GBI is defined as an area-weighted average of $Z_{500}$ over 60-80°N, 20-80°W, calculated from ERA5.

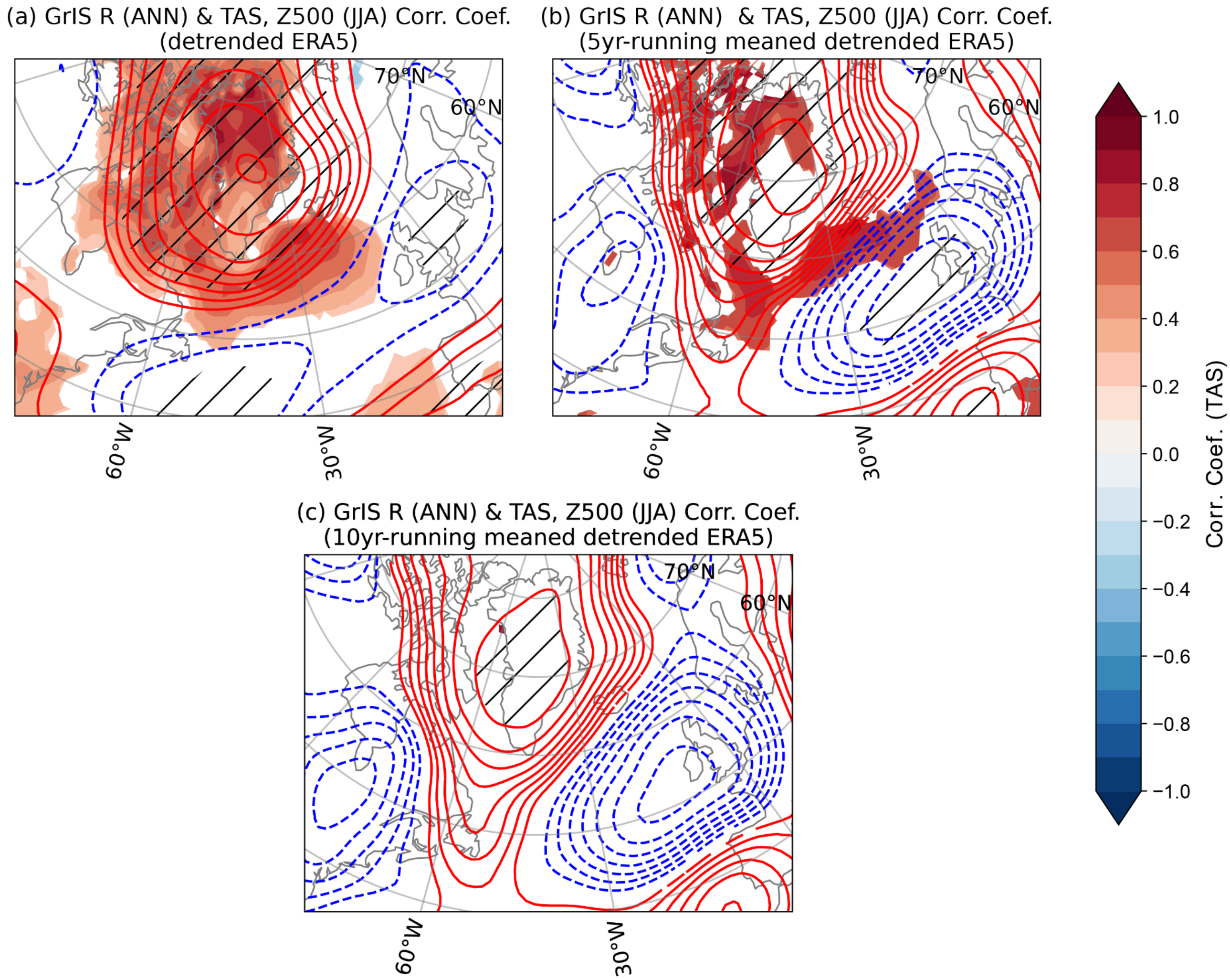


**Figure S8.** Correlation maps of GrIS $R$ to 500 hPa geopotential height ($Z_{500}$; contours) and near-surface air temperature (TAS; shading) during 1958-2019 in ERA5 reanalysis after (a) detrending, (b) detrending and applying a 5-year running mean, and (c) detrending and applying a 10-year running mean. Regions with insignificant TAS correlations are masked out and regions with significant $Z_{500}$ correlations (95% confidence) are hatched. Contours start from $\pm 0.1$ with contour spacings of 0.1. The significance level of correlation in the running mean correlation maps considers the effective degree of freedom (eDOF) as the total years of data (1958-2019, 62 years) divided by the running mean window (e.g., for 5-year running mean correlation, eDOF is 62/5).

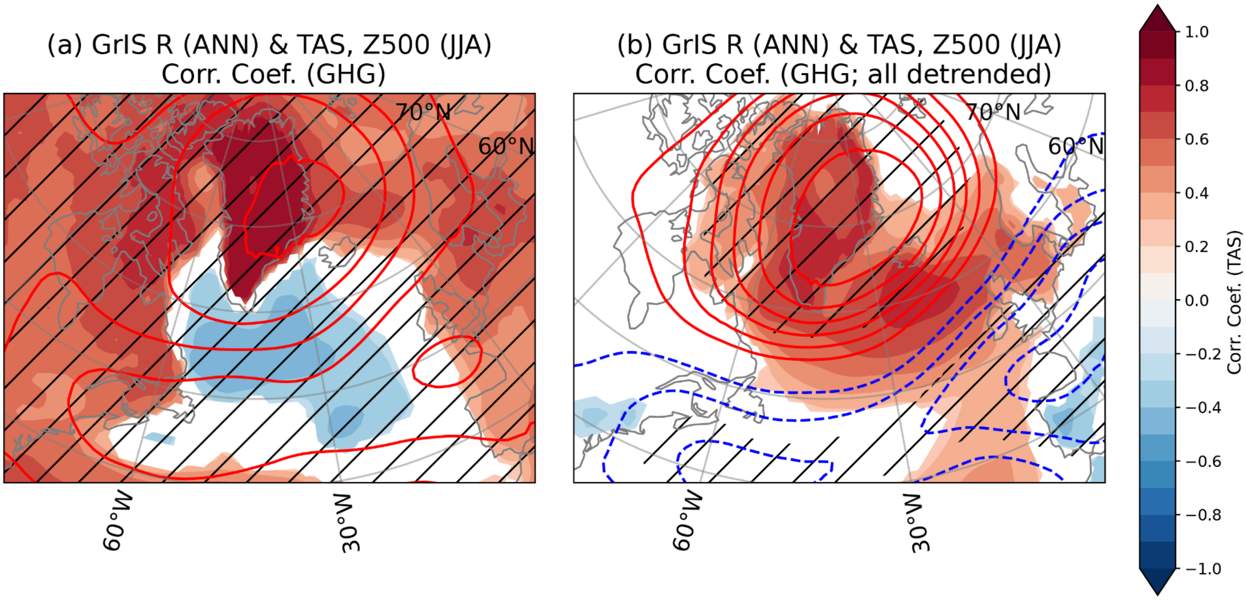


**Figure S9.** Correlation maps of GrIS *R* to 500 hPa geopotential height ($Z_{500}$; contours) and near-surface air temperature (TAS; shading) during 1958-2019 in (a) GHG and (b) detrended GHG. Regions with insignificant TAS correlations are masked out and regions with significant $Z_{500}$ correlations (95% confidence) are hatched. Contours start from $\pm 0.1$ with contour spacings of 0.1.

**Table S1. Bayesian Total Least Square (TLS) regression details.**

**Algorithm steps**

1. Define the observation of $x_{obs}$
2. Define the priors of $\sigma_x$, $\sigma_y$, $x_{latent}$, $slope(\beta)$ (see below)
3. Define the likelihood function of $y$ as $y \sim N(\beta x_{latemt},\ \sigma_y^2)$ with the observation of $y_{obs}$

Inference for posterior samples based on the priors and likelihood function of all the parameters (with 4000 burn-out samples and 10000 samples for analyses in this study)

| **Parameters and their priors** | **Priors (idealized experiments)** | **Priors (main manuscript D&A)** | **Priors (main manuscript temperature sensitivity)** |
|---|---|---|---|
| $\sigma_x \sim H(\sigma_{x_{obs}}^2)$ | $\sigma_x \sim H(1)$ | $\sigma_x \sim H(SE(x_{ensmean})^2)$ | $\sigma_{TAS} \sim H(SE(x_{TAS_{ensmean}})^2)$ |
| $\sigma_y \sim H(\sigma_y^2)$ | $\sigma_y \sim H(1)$ | $\sigma_y \sim H(\sigma_{ERA-RACMO}^2)$ | $\sigma_{SMB/R} \sim H(SE(x_{SMB/R_{ensmean}})^2)$ |
| $x_{latemt} \sim N(x_{obs}, \sigma_x^2)$ | $x_{latent} \sim N(x_{obs}, \sigma_x^2)$ | $x_{latent} \sim N(x_{ensmean}, \sigma_x^2)$ | $x_{latent} \sim N(x_{TAS_{ensmean}}, \sigma_x^2)$ |
| $\beta \sim N(0,\ \sigma_y^2)$ | $\beta \sim N(0,\ \sigma_y^2)$ | $\beta \sim N(0,\ \sigma_y^2)$ | $\beta \sim N(0,\ \sigma_y^2)$ |