Submitted to the Annals of Applied Statistics

VALID CONFIDENCE INTERVALS FOR DETECTION AND ATTRIBUTION IN CLIMATE CHANGE STUDIES

BY HANYUE CHEN[1,a], SONG XI CHEN [2,b] AND JINGKUN QIU[3,c]

[1]*Center for Statistical Science, Peking University* , [a]*ypchy@pku.edu.cn*

[2]*Department of Statistics and Data Science, Tsinghua University* , [b]*sxchen@tsinghua.edu.cn*

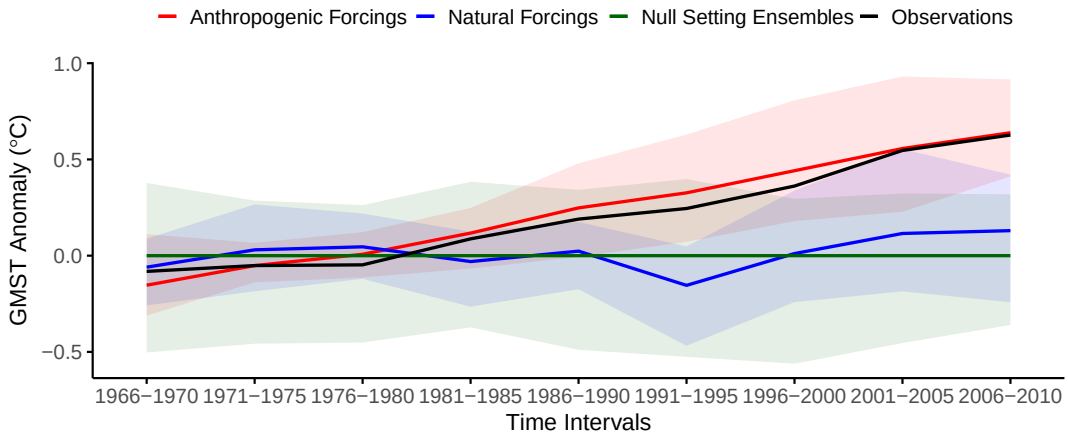
[3]*Guanghua School of Management, Peking University* , [c]*jkqiu@stu.pku.edu.cn*

This paper evaluates the popular optimal fingerprinting approach proposed by Allen and Stott (2003), which is based on errors-in-variables regression and is the benchmark method for detection and attribution analyses in climate change studies. We have found the cause of the widely observed under-coverage of the Allen–Stott confidence intervals, which is due to the use of an erroneous asymptotic variance of the weighted total least squares estimator. We extend the scope of the optimal fingerprinting approach to allow spatial and temporal dependence, as geophysical data usually exhibit such dependence, and high dimensionality, as the dimensionality of the covariance matrix of the residual is tied to the sample size in climate change studies. The proposed approach is used to evaluate the impacts of climate change on mean surface temperature.

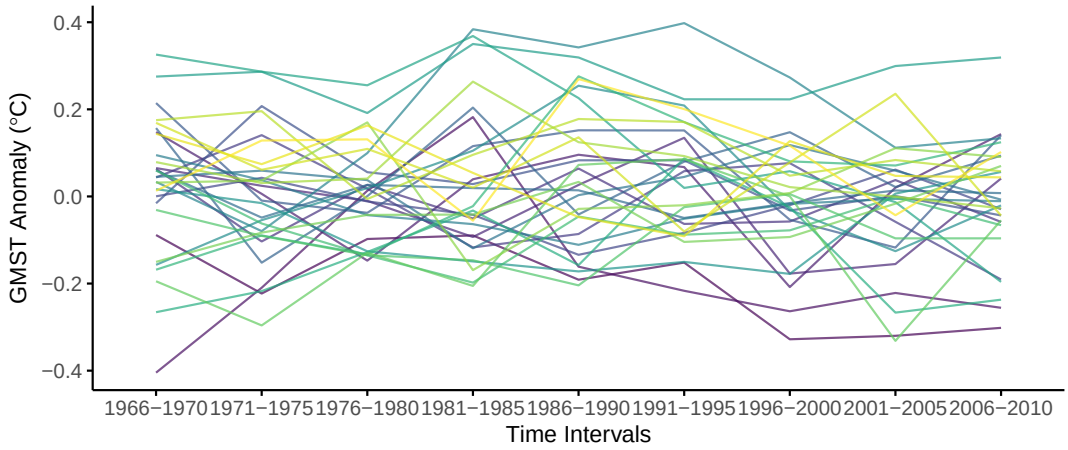
1. Introduction. The optimal fingerprinting method, developed in a series of works represented by Hasselmann (1979), North et al. (1995), Allen and Tett (1999) and Allen and Stott (2003), is the key method for detecting and attributing factors that represent the human effects on the climate system. It has been used to produce scientific evidence in the very influential assessment reports by the Intergovernmental Panel on Climate Change (IPCC, 1990, 1996, 2001, 2007, 2013, 2021). The optimal fingerprinting approach is based on the statistical linear regression that attempts to detect or attribute observed climate variables to either human-induced or naturally induced causes via the significance of the estimated linear regression coefficients. Unlike standard statistical regression, the covariates, which are called fingerprints, are generated from climate models to represent anthropogenic and natural forcings to the observed response variable, and the residual covariance matrix is estimated from simulated ensembles from a null climate model (a model that lacks external forcings on the climate system to reflect the internal variation) to facilitate the weighted least squares estimation. Figure 1a provides a comparison of 1966–2010 global five-year mean temperature evolution in observations and ensemble means generated from the aforementioned climate models. Figure 1b further presents the same evolution in each ensemble generated from the null climate model. Data for both plots are taken from Li et al. (2021) and will be explained in details in Section 7. Note that there are 223 ensembles for the null setting and only the first 30 ensembles are presented in Figure 1b for illustrative purpose.

The estimated regression coefficients, called the scaling factors, scale the fingerprints to best match the changes in the observed climate variables. The detection and attribution analyses in climate change studies can be regarded as statistical estimation and hypothesis testing tasks, and rely on both consistent estimators and valid confidence intervals of the scaling factors. The statistical inference on the regression coefficients informs the extent to which the climate system has been affected by human activity as opposed to natural variation. Chen,

Keywords and phrases: Climate Change, Confidence Intervals, Errors-in-Variables Regression, Optimal Fingerprints, Total Least Squares.



(a) Comparison of 1966–2010 global five-year mean temperature evolution.



(b) 1966–2010 global five-year mean temperature evolution in the first 30 null setting ensembles.

Figure 1: (a) Comparison of 1966–2010 global five-year mean temperature evolution in observations (black line) and ensemble means simulated from climate models representing anthropogenic forcing (red line), natural forcing (blue line) and null setting (green line). Each solid curve shows the evolution of observations or ensemble means while each shading region shows the range of the ensembles due to the climate system uncertainty. (b) 1966–2010 global five-year mean temperature evolution in the first 30 ensembles generated from the null climate model. Each curve represents the evolution of each simulated ensemble. Data of both plots are taken from Li et al. (2021) and will be explained in details in Section 7.

Chen and Mu (2024) reviewed the optimal fingerprinting approach as proposed in Allen and Tett (1999) in the light of the severe critic of McKitrick (2022), and provided conditions under which the estimated scaling factors are statistically consistent and optimal in the absence of the measurement errors.

Allen and Stott (2003) extended Allen and Tett (1999) to allow measurement errors in the fingerprints (covariates) as they are acquired from simulation of climate models and are subject to measurement errors. This brought the context to the errors-in-variables regression. Allen and Stott (2003) used the total least squares estimators (Gleser, 1981; Fuller, 1987; van Huffel and Vandewalle, 1991) for the scaling factor estimation, and had provided confidence

intervals for the scaling factors. See also a recent commentary of us (Chen, Chen and Qiu, 2025) to McKitrick (2025) on Allen and Stott (2003).

Having valid confidence intervals is vitally important for the detection and attribution analysis as a detection of a cause of climate change for a climate variable is declared if the 90% confidence interval excludes 0 and an attribution is made if the 90% confidence intervals includes 1. These are well adapted in the IPCC reports. In the absence of measurement errors, Allen and Tett (1999) constructed the confidence intervals based on the Wald type formulation of the weighted least squares estimates that is asymptotically χ^2 distributed. In the total least squares framework, Allen and Stott (2003) used a similar Wald-type formulation with the total least squares estimator to construct confidence regions or intervals. The Allen–Stott construction has been widely adopted in the detection and attribution analyses in climate change studies as reflected in Ribes, Azaïs and Planton (2009), Polson et al. (2013), and the Earth System Model Evaluation Tool (Righi et al., 2020), and has provided the technical support to findings in the IPCC reports (Bock et al., 2020; Gillett et al., 2021).

However, many climate change studies (Ribes, Planton and Terray, 2013; DelSole et al., 2019; Li et al., 2021) have shown that the Allen–Stott confidence intervals exhibited severe under-coverage, which can hamper the effort of detecting and attribution analyses in climate change studies. There have been studies in the climate change literature trying to understand the Allen–Stott formulation and find the cause of the under-coverage. Ribes, Planton and Terray (2013) raised up the issue of high dimensionality in the estimation of the residual covariance matrix on the coverage, and employed the regularized high dimensional covariance estimator of Ledoit and Wolf (2004). DelSole et al. (2019) tried to understand the formulation by linking to likelihood ratio theory, and gave a biased-corrected likelihood-based confidence interval which is equivalent to the claim in Allen and Stott (2003). DelSole et al. (2019) proposed a non-parametric bootstrap method to account for the weak-signal regime. To restore the coverage, Li et al. (2021) proposed a parametric bootstrap method to estimate an inflation factor to enlarge the confidence intervals provided by the asymptotically normal distribution of the total least squares estimator in Gleser (1981).

A major finding of this work is the discovery that the Allen–Stott formulation is not χ^2-distributed as it missed a term in the asymptotic variance of the total least squares estimator as well as a multiplicative factor. This explains the widely observed under-coverage of the Allen–Stott confidence intervals. We also analyze the likelihood ratio formulation in DelSole et al. (2019) and show that it would not be able to validate the Allen–Stott procedure.

A key aspect of the optimal fingerprinting approach is that the sample size ℓ of the observations $\mathbf{y}$ is the same as the dimension of the regression error covariance matrix $\mathbf{\Sigma}_\varepsilon$. While an increase of the sample size ℓ ensures a large sample guarantee of the total least squares estimation accuracy, such increase means the regression error covariance matrix $\mathbf{\Sigma}_\varepsilon$ is high dimensional, which is known to cause problems for statistical inference for approaches based on the finite dimensions (Bai and Yin, 1993; Chen, Zhang and Zhong, 2010). The high dimensionality issue of $\mathbf{\Sigma}_\varepsilon$ is aggravated by the rather limited number of null ensemble runs used to estimate $\mathbf{\Sigma}_\varepsilon$ due to high cost of the simulation runs of big climate models. The high dimensionality issue was noted by Ribes, Azaïs and Planton (2009) and Ribes, Planton and Terray (2013), which proposed a shrinkage estimation of $\mathbf{\Sigma}_\varepsilon$. Li et al. (2023) and Li and Li (2025) employed the shrinkage error covariance estimators of Ledoit and Wolf (2004) to counter the rank deficiency of the conventional sample covariance matrix in high dimension. The shrinkage estimators were used to rotate the errors-in-variables regression models as advocated in Allen and Stott (2003) (AS03) to attain the weighted total least squares (WTLS) estimators of the scaling factor β. Li et al. (2023) considered non-linear shrinkage and showed its advantage over the linear shrinkage of Ledoit and Wolf (2004) in the estimation of β, and Li and Li (2025) focused on the linear shrinkage formulation with a goal of minimizing the the

asymptotic covariance of the WTLS. Qiu, Chen and Chen (2026) adopted a more general framework and established a new perturbation bound for the generic WTLS estimator with respect to the weighting matrix.

We propose using the localization estimator of Sun, Chen and Qiu (2026) for the high dimensional error covariance matrix to address the high dimensionality in the fingerprint studies, which is consistent under suitable conditions, and is well suited for climate data whose dependence can be modeled via the spatial and temporal distances. We have conducted simulation experiments and empirical analysis on real climate data, which showed the promise of the high dimensional proposal in producing accurate confidence intervals.

Another development of this work is in extending the error-in-variable regression framework to spatial and temporal dependence in the observation which is the most natural for climate change and earth science studies. However, the spatial or temporal dependence among the observations in climate change studies has been largely ignored in the optimal fingerprinting approach. Indeed, Gleser (1981) to whom van Huffel and Vandewalle (1991) and Allen and Stott (2003) had drawn methodology support, was for study designs with independent replications, which does not suit the typical geophysics setting which tends to have spatial and temporal dependence.

We propose a valid χ^2 formulation for the total least squares estimates, which is shown to be able to repair the aforementioned coverage deficiency issue and can accommodate the spatial and temporal dependence as well as the high dimensionality. The proposed valid confidence intervals are employed to conduct detection and attribution analysis for global mean temperature from 1951–2010 which demonstrate the advantages of our proposal over the existing methods.

The rest of the paper is structured as follows. Section 2 summarizes Allen–Stott procedure of Allen and Stott (2003) with the broader optimal fingerprinting method for the errors-in-variables regression. Section 3 provides a theoretical understanding to the under-coverage phenomenon of the Allen–Stott confidence intervals. Section 4 establishes a rigorous asymptotic theory of the total least squares estimator suitable for spatial and temporal dependent data, which leads to valid confidence regions of the scaling factor in Section 5 with a localization estimator of $\mathbf{\Sigma}_\varepsilon$ to counter the high dimensionality. Sections 6 reports simulation results to confirm the findings in the earlier sections. Section 7 provides detection and attribution analysis for global mean temperature study using the proposed methodology. A discussion is made in Section 8.

2. Allen–Stott procedure. This section provides an overview on the optimal fingerprinting approach advocated in Allen and Stott (2003). The climate change detection and attribution analysis involves linking observed climate variables with potential anthropogenic and natural sources believed to contribute to changes in the observed climate variables.

Let $\mathbf{y}$ be an ℓ-dimensional vector of observed climate variables collected over a period of time in a spatial region, and $\mathbf{X}$ be an $\ell \times m$ matrix with column vectors $\{\mathbf{x}_k\}_{k=1}^m$, which collect the expected climate responses to external forcings corresponding to the climate variables in $\mathbf{y}$, but generated by simulation runs of certain climate models.

As $\mathbf{X}$ are results of simulated runs of climate models, they are perturbation versions of the unobservable forcing variables $\mathbf{X}_0 = [\mathbf{x}_{0,1}, \ldots, \mathbf{x}_{0,m}]$ such that

$$\mathbf{X} = \mathbf{X}_0 + \boldsymbol{\varepsilon}_x, \quad \boldsymbol{\varepsilon}_x = [\boldsymbol{\varepsilon}_{x,1}, \ldots, \boldsymbol{\varepsilon}_{x,m}], \tag{1}$$

where $\{\boldsymbol{\varepsilon}_{x,k}\}_{k=1}^m$ are ℓ-dimensional random vectors with mean zero and covariance matrices $\{\mathbf{\Sigma}_{x,k}\}_{k=1}^m$. The errors-in-variables regression model considered in Allen and Stott (2003) is

$$\mathbf{y} = \mathbf{X}_0\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{2}$$

where $\boldsymbol{\varepsilon}$ is a vector of residuals (errors) representing internal climate variability with zero means and covariance matrix $\boldsymbol{\Sigma}_\varepsilon$, and $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_m]^T$ are the scaling factors. Assume that $\{\mathbf{x}_k\}_{k=1}^m$ have the same covariance structure with the internal error $\boldsymbol{\varepsilon}$ so that $\boldsymbol{\Sigma}_{x,k} = \boldsymbol{\Sigma}_\varepsilon$ for $k \in \{1, \cdots, m\}$, which is needed for the identification purpose of the errors-in-variables regression, and can be argued if the models used to generate the ensembles of $\mathbf{X}_k$ mimic the internal climate variability as well.

We would like to emphasize that the assumption of $\boldsymbol{\Sigma}_{x,k} = \boldsymbol{\Sigma}_\varepsilon$ is for ease of exposition and can be generalized as $\boldsymbol{\Sigma}_{x,k} = c_k \boldsymbol{\Sigma}_\varepsilon$ with c_k known as adopted in Li et al. (2021, 2023). With known c_k's, the scaling factors c_k can be standardized to 1 via rescaling each fingerprint by $1/\sqrt{c_k}$. Hence, without loss of generality, we take the simplified case of $\boldsymbol{\Sigma}_{x,k} = \boldsymbol{\Sigma}_\varepsilon$.

Models (1) and (2) constitute the errors-in-variables regression model which updated the Allen and Tett (1999) setting without measurement errors in the fingerprints $\{\mathbf{x}_k\}_{k=1}^m$. The main tasks of the fingerprint detection and attribution analysis under the errors-in-variables model are twofold: (i) estimation of the regression coefficients $\{\beta_i\}_{i=1}^m$ and (ii) constructing the 90% confidence intervals for β_i's. If the confidence interval for β_i excludes 0, the influence of the i-th associated external forcing is said to be detected since the null hypothesis $H_0 : \beta_i = 0$ would be rejected at a 10% significance level. If the confidence interval includes 1, then the observed climate changes in $\mathbf{y}$ may be attributed to the i-th forcing since $H_0 : \beta_i = 1$ cannot be rejected.

There are a range of statistical methods for estimating β_i's for errors-in-variables regression, which include the orthogonal regression (Pearson, 1901) for the univariate case, the maximum likelihood estimation (MLE) under the joint Gaussian distribution of the measurement errors $\boldsymbol{\varepsilon}_x$ and $\boldsymbol{\varepsilon}$ (Kendall and Stuart, 1977; Fuller, 1987). Allen and Stott (2003) employed the total least squares estimator in van Huffel and Vandewalle (1991), which referred to Gleser (1981) for theoretical properties.

A key assumption of Gleser (1981) for the total least squares is

(3) the rows of the error matrix $[\boldsymbol{\varepsilon}_x, \boldsymbol{\varepsilon}]$ are independent and identically distributed (IID) with mean zero and the covariance matrix proportional to the identity matrix of $(m+1)$ dimensions, namely $\sigma^2 \mathbf{I}_{m+1}$.

The specific covariance structure in (3) ensures the identification of β_i's with the total least squares estimation and is reasonable in climate change studies, because of (i) the errors $\{\boldsymbol{\varepsilon}_{x,k}\}_{k=1}^m$ and $\boldsymbol{\varepsilon}$ associated with independent simulations of the internal climate models, and (ii) the early assumption that $\boldsymbol{\Sigma}_{x,k} = \boldsymbol{\Sigma}_\varepsilon$.

Under (3), the total least squares estimator is

$$\hat{\boldsymbol{\beta}}_{TLS} = (\mathbf{X}^T\mathbf{X} - \lambda_{m+1}([\mathbf{X}, \mathbf{y}]^T[\mathbf{X}, \mathbf{y}])\mathbf{I}_m)^{-1}\mathbf{X}^T\mathbf{y}, \tag{4}$$

where $\lambda_{m+1}(\mathbf{A})$ denotes the $(m+1)$-th largest eigenvalue of a symmetric matrix $\mathbf{A}$; see Golub and van Loan (1980) for a detailed derivation.

As in Allen and Tett (1999), to leverage on the Gauss–Markov theorem, Allen and Stott (2003) proposed to pre-whiten the errors-in-variables model via an inverse square root of the covariance matrix $\boldsymbol{\Sigma}_\varepsilon$ obtained by ensemble simulations of a climate model under the null setting. Since $\boldsymbol{\Sigma}_\varepsilon$ is unknown, an estimator $\hat{\boldsymbol{\Sigma}}_\varepsilon$ is obtained by n ensemble runs under the null setting of climate models with no external forcings as $\hat{\boldsymbol{\Sigma}}_\varepsilon = n^{-1}\mathbf{Y}_N\mathbf{Y}_N^T$ where the columns of $\mathbf{Y}_N$ represent n vectors of "pseudo-observations" of climate internal noise. Although it is common to assume that the null simulations are statistically independent of the observed physical world so that $\hat{\boldsymbol{\Sigma}}_\varepsilon$ is independent of $\mathbf{X}$ and $\mathbf{y}$ in the literature (Li et al., 2021, 2023; Chen, Chen and Mu, 2024), we emphasize that this assumption can be relaxed in our framework, as noted by Qiu, Chen and Chen (2026), provided the estimation of $\hat{\boldsymbol{\Sigma}}_\varepsilon$ has sufficient

rate of convergence. This brings flexibility for using a broader class of null simulations in practice.

The pre-whitening operator $\hat{\mathbf{P}}$ is obtained by the spectral decomposition of $\hat{\mathbf{\Sigma}}_\varepsilon$ and hence $\hat{\mathbf{\Sigma}}_\varepsilon^{-1}$ such that $\hat{\mathbf{\Sigma}}_\varepsilon^{-1} = \hat{\mathbf{P}}^T\hat{\mathbf{P}}$. The total least squares estimator, we name as the weighted total least squares estimator, after the pre-whitening is

$$\hat{\boldsymbol{\beta}}_{WTLS} = (\mathbf{X}^T\hat{\mathbf{\Sigma}}_\varepsilon^{-1}\mathbf{X} - \lambda_{m+1}([\mathbf{X},\mathbf{y}]^T\hat{\mathbf{\Sigma}}_\varepsilon^{-1}[\mathbf{X},\mathbf{y}])\mathbf{I}_m)^{-1}\mathbf{X}^T\hat{\mathbf{\Sigma}}_\varepsilon^{-1}\mathbf{y}. \tag{5}$$

Allen and Stott (2003) considered the perfectly pre-whitened data such that $\hat{\mathbf{\Sigma}}_\varepsilon^{-1} = \mathbf{\Sigma}_\varepsilon^{-1}$ and a quantity

$$\Delta s^2(\boldsymbol{\beta}) = [\boldsymbol{\beta}^T, -1]([\mathbf{X},\mathbf{y}]^T\mathbf{\Sigma}_\varepsilon^{-1}[\mathbf{X},\mathbf{y}] - \lambda_{m+1}([\mathbf{X},\mathbf{y}]^T\mathbf{\Sigma}_\varepsilon^{-1}[\mathbf{X},\mathbf{y}])\mathbf{I}_{m+1})[\boldsymbol{\beta}^T, -1]^T,$$

and claimed that

$$\Delta s^2(\boldsymbol{\beta}) \sim \chi^2_m. \tag{6}$$

The equivalence between the expression of $\Delta s^2(\boldsymbol{\beta})$ above and the original form in Equation (28) of Allen and Stott (2003) is given in Section S1 of the supplementary material (SM).

Under the claim, the 90% confidence region for $\boldsymbol{\beta}$ is $\{\boldsymbol{\beta} \mid \Delta s^2(\boldsymbol{\beta}) \le \chi^2_{m,0.1}\}$, and those for β_i can be readily obtained. In order to take into account the uncertainty in estimating $\mathbf{\Sigma}_\varepsilon$, Allen and Stott (2003) proposed to use the sample-splitting approach, which split the null setting control runs into two parts. It uses one part to obtain $\hat{\mathbf{\Sigma}}_{\varepsilon,1}$ for the weighted total least squares estimate for $\boldsymbol{\beta}$, and the other part for another estimate $\hat{\mathbf{\Sigma}}_{\varepsilon,2}$ to be used in $\Delta s^2(\boldsymbol{\beta})$ for constructing the confidence region. The quantity $\Delta s^2(\boldsymbol{\beta})$ with the sample splitting is claimed to follow the $mF_{m,\nu}$ distribution, where ν is the rank of $\hat{\mathbf{\Sigma}}_{\varepsilon,2}$. The above confidence region is the foundation for confidence intervals for each scaling factor β_i used in the IPCC reports, for instance in Chapter 3 of IPCC (2021).

Let $\hat{\boldsymbol{\beta}}_{GTLS}$ be a version of the weighted total least squares estimator in (5) but with $\hat{\mathbf{\Sigma}}_\varepsilon^{-1}$ replaced by the true $\mathbf{\Sigma}_\varepsilon^{-1}$ in (5), which is called the generalized total least squares estimator, which only serve for the theoretical purposes. It seems to us that the Claim (6) by Allen and Stott (2003) comes from reasoning based on the Wald-type statistic associated with the generalized total least squares estimator $\hat{\boldsymbol{\beta}}_{GTLS}$. It is shown in Section S1 of the SM that

$$\Delta s^2(\boldsymbol{\beta}) = \ell(\hat{\boldsymbol{\beta}}_{GTLS} - \boldsymbol{\beta})^T\hat{\mathbf{V}}_{AS}^{-1}(\hat{\boldsymbol{\beta}}_{GTLS} - \boldsymbol{\beta}), \tag{7}$$

where $\hat{\mathbf{V}}_{AS} = \ell(\mathbf{X}^T{\mathbf{\Sigma}_\varepsilon}^{-1}\mathbf{X} - \lambda_{m+1}([\mathbf{X},\mathbf{y}]^T\mathbf{\Sigma}_\varepsilon^{-1}[\mathbf{X},\mathbf{y}])\mathbf{I}_m)^{-1}$. However, $\Delta s^2(\boldsymbol{\beta})$ is not χ^2_m distributed as $\hat{\mathbf{V}}_{AS}$ is not a consistent estimator of the asymptotic variance of $\hat{\boldsymbol{\beta}}_{GTLS}$, as detailed in Section 3. The invalidity of Claim (6) is the cause of the observed under-coverage of the confidence intervals.

Trying to understand Claim (6), DelSole et al. (2019) considered a log-likelihood ratio (LR) statistic under an extra assumption

$$\text{the rows of the error matrix } [\boldsymbol{\varepsilon}_x, \boldsymbol{\varepsilon}] \text{ are IID } N(0, \sigma^2\mathbf{I}_{m+1}) \text{ distributed,} \tag{8}$$

which was assumed in Gleser (1981) when establishing an equivalence between the total least squares estimator and the MLE. The log-LR in DelSole et al. (2019) was

$$D(\boldsymbol{\beta}) = \ell(m+1)\log(\frac{[\boldsymbol{\beta}^T, -1][\mathbf{X},\mathbf{y}]^T\mathbf{\Sigma}_\varepsilon^{-1}[\mathbf{X},\mathbf{y}][\boldsymbol{\beta}^T, -1]^T}{\lambda_{m+1}([\mathbf{X},\mathbf{y}]^T\mathbf{\Sigma}_\varepsilon^{-1}[\mathbf{X},\mathbf{y}]) \cdot (1 + \boldsymbol{\beta}^T\boldsymbol{\beta})}). \tag{9}$$

Writing the logarithm of the LR as the difference of the two log-terms and used the approximation $\log(x) \approx x - 1$ for x close to 1 would lead to $\Delta s^2(\boldsymbol{\beta})$ up to a constant. DelSole et al. (2019) then tried to argue via Wilks' theorem that $D(\boldsymbol{\beta}) \sim \chi^2_m$, which led to the 90%

confidence region for $\boldsymbol{\beta}$ as $\{\boldsymbol{\beta} \mid D(\boldsymbol{\beta}) \leq \chi^2_{m,0.1}\}$. Due to the inconsistency of the MLE, DelSole et al. (2019) further proposed a bias-correction procedure to $D(\boldsymbol{\beta})$ that gave rise to a $D_{dof}(\boldsymbol{\beta})$ with confidence intervals effectively equivalent to those of Claim (6). Despite the interpretation, DelSole et al. (2019) found the confidence intervals prescribed by (6) still had significant under-coverage.

There are several issues with the above interpretation. One is that the $\log(x) \approx x - 1$ approximation may not be accurate as the x here (the ratio of the numerator to the denominator of the LR) may not be close to 1. Another is that the Wilks' theorem is not suitable for the total least squares problem where $\mathbf{X}_0$ has to be treated as a high dimensional nuisance parameter as $\ell \to \infty$. The high-dimensionality of X_0 invalidates the application of Wilks' theorem, which implies that likelihood ratio-based argument may not provide a support to the Allen–Stott formulation.

Li et al. (2021) regarded the cause of the under-coverage as insufficient uncertainty consideration in estimating $\boldsymbol{\Sigma}_\varepsilon$ in (6), and proposed to inflate the covariance by an inflation factor and used a bootstrap method to calibrate the factor for better coverage. Our analysis in the next section will provide a theoretical insight on Claim (6) that reveals the cause of the under-coverage.

3. Analysis on the χ^2_m Claim (6). This section analyzes Claim (6) of Allen and Stott (2003), and prepares for an alternative procedure for confidence regions for the scaling factor $\boldsymbol{\beta}$. The theoretical underpinning of Allen and Stott (2003) for the total least squares estimator was Gleser (1981). Gleser (1981) assumes that there are ℓ independent replications under the errors-in-variables model as in Assumption (3) which requires the rows of $[\boldsymbol{\varepsilon}_x, \boldsymbol{\varepsilon}]$ to be IID. However, this may not be the case for the optimal fingerprinting studies as observations over different components of ℓ-dimensional $\mathbf{y}$ may not be independent as they would be either spatially or temporally dependent, depending how $\mathbf{y}$ is collected.

We show in the following that even with independent replications, Claim (6) is still not valid. We first outline the properties of the total least squares estimator given in Gleser (1981), which assumes that, in additional to (3),

$$\boldsymbol{\Delta}_I = \lim_{\ell \to \infty} \ell^{-1} \mathbf{X}_0^T \mathbf{X}_0 \text{ exists and is a positive definite matrix.} \tag{10}$$

After establishing the strong consistency (denoted as $\xrightarrow{\text{a.s.}}$) of the total least squares estimator such that $\hat{\boldsymbol{\beta}}_{TLS} \xrightarrow{\text{a.s.}} \boldsymbol{\beta}$ as $\ell \to \infty$, Gleser (1981) concluded that under Conditions (3) and (10), when the rows of $[\boldsymbol{\varepsilon}_x, \boldsymbol{\varepsilon}]$ have finite fourth moments, $\sqrt{\ell}(\hat{\boldsymbol{\beta}}_{TLS} - \boldsymbol{\beta})$ has asymptotic m-variate normal distribution with zero means and covariance matrix determined by up to the fourth moments of the rows of $[\boldsymbol{\varepsilon}_x, \boldsymbol{\varepsilon}]$. Specifically, if

$$\text{up to the fourth moments of } [\boldsymbol{\varepsilon}_x, \boldsymbol{\varepsilon}] \text{ are the same as those of the normal distribution } N(0, \sigma^2 \mathbf{I}_{m+1}), \tag{11}$$

$$\ell(\hat{\boldsymbol{\beta}}_{TLS} - \boldsymbol{\beta})^T \mathbf{V}_{TLS}^{-1} (\hat{\boldsymbol{\beta}}_{TLS} - \boldsymbol{\beta}) \xrightarrow{d} \chi^2_m \text{ as } \ell \to \infty, \tag{12}$$

where the asymptotic variance of $\hat{\boldsymbol{\beta}}_{TLS}$ is

$$\mathbf{V}_{TLS} = (1 + \boldsymbol{\beta}^T \boldsymbol{\beta}) \boldsymbol{\Delta}_I^{-1} \left\{ \sigma^2 \boldsymbol{\Delta}_I + \sigma^4 (\mathbf{I}_m + \boldsymbol{\beta}\boldsymbol{\beta}^T)^{-1} \right\} \boldsymbol{\Delta}_I^{-1}. \tag{13}$$

To compare Claim (6) with (12), we assume $\boldsymbol{\Sigma}_\varepsilon = \mathbf{I}_\ell$ with $\sigma^2 = 1$. In this case, the data would meet the IID condition, without the need for the pre-whitening. Derivations in Section S1 of the SM shows that

$$\Delta s^2(\boldsymbol{\beta}) = \ell(\hat{\boldsymbol{\beta}}_{TLS} - \boldsymbol{\beta})^T \hat{\mathbf{V}}_{AS}^{-1} (\hat{\boldsymbol{\beta}}_{TLS} - \boldsymbol{\beta}), \tag{14}$$

where $\hat{\mathbf{V}}_{AS} = (\ell^{-1}\mathbf{X}^T\mathbf{X} - \hat{\sigma}^2_{TLS}\mathbf{I}_m)^{-1}$ and $\hat{\sigma}^2_{TLS} = \ell^{-1}\lambda_{m+1}([\mathbf{X},\mathbf{y}]^T[\mathbf{X},\mathbf{y}])$. It is shown in Section S2 of the SM that $\hat{\mathbf{\Delta}}_I := \ell^{-1}\mathbf{X}^T\mathbf{X} - \hat{\sigma}^2_{TLS}\mathbf{I}_m \xrightarrow{a.s.} \mathbf{\Delta}_I$, then

$$\begin{aligned}(1+\boldsymbol{\beta}^T\boldsymbol{\beta})^{-1}\mathbf{V}_{TLS} - \hat{\mathbf{V}}_{AS} &\xrightarrow{a.s.} (1+\boldsymbol{\beta}^T\boldsymbol{\beta})^{-1}\mathbf{V}_{TLS} - \mathbf{\Delta}_I^{-1} \\ &= \left\{\mathbf{\Delta}_I^{-1} + \mathbf{\Delta}_I^{-1}(\mathbf{I}_m + \boldsymbol{\beta}\boldsymbol{\beta}^T)^{-1}\mathbf{\Delta}_I^{-1}\right\} - \mathbf{\Delta}_I^{-1} \\ &= \mathbf{\Delta}_I^{-1}(\mathbf{I}_m + \boldsymbol{\beta}\boldsymbol{\beta}^T)^{-1}\mathbf{\Delta}_I^{-1},\end{aligned} \tag{15}$$

which is positive definite.

Thus, Claim (6) is not valid, as $\hat{\mathbf{V}}_{AS}$ misses the factor $(1+\boldsymbol{\beta}^T\boldsymbol{\beta})$ and exhibits a positive definite gap from $\mathbf{V}_{TLS}$. The positive definite gap provide a theoretical understanding to the widely observed under-coverage phenomenon of confidence intervals based on Claim (6).

DelSole et al. (2019) observed that confidence intervals based on Claim (6) had more severe under-coverage in the weak-signal case than those with stronger ones. This is indeed consistent with our revelation since it can be explained by (15) by noting that (i) $\mathbf{\Delta}_I$ as the limit of $\ell^{-1}\mathbf{X}_0^T\mathbf{X}_0$ measures the overall signal strength; and (ii) the smaller the signals are, the smaller $\mathbf{\Delta}_I$ is, and the larger the gap in (15) becomes.

To reach a confidence region effectively equivalent to those in Allen and Stott (2003), DelSole et al. (2019) derived a bias-corrected LR, $D_{dof}(\boldsymbol{\beta})$, based on the $D(\boldsymbol{\beta})$ in Section 2 and claimed that $D_{dof}(\boldsymbol{\beta}) \sim \chi^2_m$. We have shown in Section S1 of the SM that for large ℓ

$$D_{dof}(\boldsymbol{\beta}) \approx \ell(\hat{\boldsymbol{\beta}}_{TLS} - \boldsymbol{\beta})^T\hat{\mathbf{V}}_{LR}^{-1}(\hat{\boldsymbol{\beta}}_{TLS} - \boldsymbol{\beta}) \tag{16}$$

where $\hat{\mathbf{V}}_{LR} = (1+\boldsymbol{\beta}^T\boldsymbol{\beta})\hat{\mathbf{\Delta}}_I^{-1}$.

It can be shown that

$$\begin{aligned}(1+\boldsymbol{\beta}^T\boldsymbol{\beta})^{-1}(\mathbf{V}_{TLS} - \hat{\mathbf{V}}_{LR}) &\xrightarrow{a.s.} \left\{\mathbf{\Delta}_I^{-1} + \mathbf{\Delta}_I^{-1}(\mathbf{I}_m + \boldsymbol{\beta}\boldsymbol{\beta}^T)^{-1}\mathbf{\Delta}_I^{-1}\right\} - \mathbf{\Delta}_I^{-1} \\ &= \mathbf{\Delta}_I^{-1}(\mathbf{I}_m + \boldsymbol{\beta}\boldsymbol{\beta}^T)^{-1}\mathbf{\Delta}_I^{-1},\end{aligned} \tag{17}$$

with the gap between $(1+\boldsymbol{\beta}^T\boldsymbol{\beta})^{-1}\mathbf{V}_{TLS}$ and $(1+\boldsymbol{\beta}^T\boldsymbol{\beta})^{-1}\hat{\mathbf{V}}_{LR}$ being the same as that between $(1+\boldsymbol{\beta}^T\boldsymbol{\beta})^{-1}\mathbf{V}_{TLS}$ and $\hat{\mathbf{V}}_{AS}$. This explains the under-coverage that occurs for confidence intervals based on $D_{dof}(\boldsymbol{\beta})$, as observed empirically by DelSole et al. (2019).

Li et al. (2021) attributed the under-coverage to insufficient uncertainty consideration in the estimation of $\mathbf{\Sigma}_\varepsilon$, and proposed a parametric bootstrap procedure to calibrate the standard errors of the normal approximation of $\boldsymbol{\beta}$ by inflating the estimates of the asymptotic variance with a factor estimated by the standard errors of the bootstrap sample. This effectively touched on the issue of high dimensionality of $\mathbf{\Sigma}_\varepsilon$, the issue we will address later.

4. Total least squares for spatially dependent data. As having been stated, the climate observations $\mathbf{y}$ and the fingerprints $\mathbf{X}$ from simulation of the climate models are likely to be spatially or temporally dependent. This means that Gleser (1981), that assumed the independent replication design, is not applicable for the climate change studies. We establish the theoretical properties for the total least squares estimator $\hat{\boldsymbol{\beta}}_{TLS}$ in (4) and the weighted total least squares estimator $\hat{\boldsymbol{\beta}}_{WTLS}$ in (5) which allows spatial dependence so that valid confidence intervals can be constructed for the detection and attribution analysis.

To quantify the spatial dependence, we introduce the notion of α-mixing that describes a form of spatial dependence such that the dependence between regions weakens as the distance between them increases. For two regions I and J, let $\rho(I,J)$ be the shortest distance between them. Let σ_I be the information set generated by the data $\{(\mathbf{X}_i, y_i)\}$ in region I, which contains all data information in region I. The α-mixing coefficients that describes the dependence between regions I and J which are k-distance apart is

$$\alpha(k) = \sup_{\rho(I,J)\geq k} \sup_{A\in\sigma_I, B\in\sigma_J} |\, P(A\cap B) - P(A)P(B)\,|\,. \tag{18}$$

Note that $\mid P(A \cap B) - P(A)P(B) \mid$ is a measure of dependence between an event A from region I and an event B from region J, as the difference would be zero if A and B are independent. The mixing coefficient $\alpha(k)$ is the largest possible dependence among all pairs of such sets from the two regions respectively. See Rosenblatt (1956), Cressie (2015) and Doukhan (1994) for more details.

The observations $\{(\mathbf{X}_i, y_i)\}_{i=1}^{\ell}$ is said to be α-mixing if $\alpha(k) \to 0$ as $k \to \infty$. Let d be the underlying dimension of (regular) spatial grids, for instance $d = 1$ for time series data or data along a latitude, $d = 2$ for longitude–latitude lattice data over a plane. Let $\delta > 0$ be a constant to be defined in (22). We impose the following condition

$$\sum_{k=1}^{\infty} k^{(1+2/\delta)d-1}\alpha(k) < \infty, \tag{19}$$

which prescribes a rate for $\alpha(k)$ to diminish to zero. Similar conditions are commonly used in spatial statistics (Jenish and Prucha, 2009; Cressie, 2015).

We generalize Gleser (1981)'s results for the total least squares estimator $\hat{\boldsymbol{\beta}}_{TLS}$ in (4). In the presence of spatial dependence, we typically have $\boldsymbol{\Sigma}_\varepsilon \neq \sigma^2\mathbf{I}_\ell$ that contradicts Gleser (1981)'s assumption (11). Therefore, we need the following conditions as a spatial generalization of (10) and (11) in Gleser (1981):

$$\boldsymbol{\Delta}_\varepsilon = \lim_{\ell\to\infty} \ell^{-1}\mathbf{X}_0^T\boldsymbol{\Sigma}_\varepsilon\mathbf{X}_0 \text{ exists and is a positive definite matrix,} \tag{20}$$

$$\text{the trace limits } \lim_{\ell\to\infty} \ell^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_\varepsilon) \text{ and } \tau = \lim_{\ell\to\infty} \ell^{-1}\mathrm{tr}(\boldsymbol{\Sigma}_\varepsilon^2) \text{ exist, and} \tag{21}$$

(22) the columns of $[\boldsymbol{\varepsilon}_x, \boldsymbol{\varepsilon}]$ are independent, $L_{4+2\delta}$ uniformly integrable for some $\delta > 0$, and the first four moments of each column are the same as those of $N(0, \boldsymbol{\Sigma}_\varepsilon)$.

Note that in Condition (22), the uniform integrability condition is just slightly stronger than a finite forth moment condition, and the Gaussian moments match condition is imposed to obtain an explicit expression for the asymptotic variance. The asymptotic normality would still hold without matching the Gaussian moments, but the corresponding asymptotic variance would take a very complicated form. To gauge on the robustness of the latter assumption, we will conduct a simulation experiment with T_ν distributed errors in Section 6.

Based on Conditions (10), (19)–(22), as shown in Qiu, Chen and Chen (2026), the total least squares estimator (4) is a consistent estimator, namely

$$\hat{\boldsymbol{\beta}}_{TLS} \xrightarrow{p} \boldsymbol{\beta} \text{ as } \ell \to \infty. \tag{23}$$

Furthermore, we have

$$\ell(\hat{\boldsymbol{\beta}}_{TLS} - \boldsymbol{\beta})^T\tilde{\mathbf{V}}_{TLS}^{-1}(\hat{\boldsymbol{\beta}}_{TLS} - \boldsymbol{\beta}) \xrightarrow{d} \chi_m^2 \text{ as } \ell \to \infty, \tag{24}$$

where the asymptotic variance of $\hat{\boldsymbol{\beta}}_{TLS}$ is

$$\tilde{\mathbf{V}}_{TLS} = (1 + \boldsymbol{\beta}^T\boldsymbol{\beta})\boldsymbol{\Delta}_I^{-1}\{\boldsymbol{\Delta}_\varepsilon + \tau(\mathbf{I}_m + \boldsymbol{\beta}\boldsymbol{\beta}^T)^{-1}\}\boldsymbol{\Delta}_I^{-1}. \tag{25}$$

The asymptotic analysis and mathematical proofs of these results are beyond the scope of this paper and can be found in Qiu, Chen and Chen (2026).

Comparing with Gleser (1981), we note that the term $\{\sigma^2\boldsymbol{\Delta}_I + \sigma^4(\mathbf{I}_m + \boldsymbol{\beta}\boldsymbol{\beta}^T)^{-1}\}$ in (13) becomes $\{\boldsymbol{\Delta}_\varepsilon + \tau(\mathbf{I}_m + \boldsymbol{\beta}\boldsymbol{\beta}^T)^{-1}\}$ in (25). The two terms coincide when $\boldsymbol{\Sigma}_\varepsilon = \sigma^2\mathbf{I}_\ell$, suggesting that Gleser (1981)'s result is a special case of the finding here. However, when $\boldsymbol{\Sigma}_\varepsilon \neq \sigma^2\mathbf{I}_\ell$ in the presence of spatial dependence, $\boldsymbol{\Delta}_\varepsilon = \lim_{\ell\to\infty} \ell^{-1}\mathbf{X}_0^T\boldsymbol{\Sigma}_\varepsilon\mathbf{X}_0 =$

$\lim_{\ell\to\infty}\mathrm{Var}(\ell^{-1}\sum_{i=1}^{\ell}\mathbf{X}_{0i}\varepsilon_i)$ can be interpreted as the long-run covariance matrix of $\mathbf{X}_{0i}\varepsilon_i$ while $\tau=\lim_{\ell\to\infty}\ell^{-1}\mathrm{tr}(\mathbf{\Sigma}_\varepsilon^2)=\lim_{\ell\to\infty}\ell^{-1}\sum_{i=1}^{\ell}\sum_{j=1}^{\ell}\{\mathrm{Cov}(\varepsilon_i,\varepsilon_j)\}^2$ reflects the accumulation of spatial dependence of the random errors.

Under the spatial dependence, we can also analyze the weighted total least squares estimator $\hat{\boldsymbol{\beta}}_{WTLS}$ in (5), which is obtained after pre-whitening the data matrix by an ensemble estimator $\hat{\mathbf{\Sigma}}_\varepsilon$ for the error covariance matrix $\mathbf{\Sigma}_\varepsilon$, where the pre-whitening operator $\hat{\mathbf{P}}$ may be obtained by the spectral decomposition of $\hat{\mathbf{\Sigma}}_\varepsilon$ so that $\hat{\mathbf{\Sigma}}_\varepsilon^{-1}=\hat{\mathbf{P}}^T\hat{\mathbf{P}}$.

Suppose that $\hat{\mathbf{\Sigma}}_\varepsilon\xrightarrow{p}\tilde{\mathbf{\Sigma}}_\varepsilon$ and $\tilde{\mathbf{\Sigma}}_\varepsilon^{-1}=\tilde{\mathbf{P}}^T\tilde{\mathbf{P}}$. Let $\tilde{\alpha}(k)$ be the α-mixing coefficients for the pre-whitened data matrix $(\tilde{\mathbf{P}}\mathbf{X},\tilde{\mathbf{P}}\mathbf{y})$ with a similar definition as that in (18). To facilitate the analysis, we require the following conditions modified from (19)–(21):

$$\sum_{k=1}^{\infty}k^{(1+2/\delta)d-1}\tilde{\alpha}(k)<\infty, \tag{26a}$$

$$\tilde{\mathbf{\Delta}}_p=\lim_{\ell\to\infty}\ell^{-1}\mathbf{X}_0^T\tilde{\mathbf{\Sigma}}_\varepsilon^{-1}\mathbf{X}_0 \text{ exists and is positive definite,} \tag{26b}$$

$$\tilde{\tilde{\mathbf{\Delta}}}_p=\lim_{\ell\to\infty}\ell^{-1}\mathbf{X}_0^T\tilde{\mathbf{\Sigma}}_\varepsilon^{-1}\mathbf{\Sigma}_\varepsilon\tilde{\mathbf{\Sigma}}_\varepsilon^{-1}\mathbf{X}_0 \text{ exists and is positive definite, and} \tag{26c}$$

$$\lim_{\ell\to\infty}\ell^{-1}\mathrm{tr}(\mathbf{\Sigma}_\varepsilon\tilde{\mathbf{\Sigma}}_\varepsilon^{-1}) \text{ and } \tilde{\tau}=\lim_{\ell\to\infty}\ell^{-1}\mathrm{tr}(\mathbf{\Sigma}_\varepsilon\tilde{\mathbf{\Sigma}}_\varepsilon^{-1}\mathbf{\Sigma}_\varepsilon\tilde{\mathbf{\Sigma}}_\varepsilon^{-1}) \text{ exist} \tag{26d}$$

Under Conditions (22), (26a)–(26d) and $\sqrt{\ell}\|\tilde{\mathbf{\Sigma}}_\varepsilon\|\|\hat{\mathbf{\Sigma}}_\varepsilon^{-1}-\tilde{\mathbf{\Sigma}}_\varepsilon^{-1}\|\xrightarrow{p}0$ where $\|\cdot\|$ denotes the matrix spectral norm, according to Qiu, Chen and Chen (2026), $\hat{\boldsymbol{\beta}}_{WTLS}$ is a weakly consistent estimator for $\boldsymbol{\beta}$, namely $\hat{\boldsymbol{\beta}}_{WTLS}\xrightarrow{p}\boldsymbol{\beta}$ as $\ell\to\infty$, and that

$$\ell(\hat{\boldsymbol{\beta}}_{WTLS}-\boldsymbol{\beta})^T\mathbf{V}_{WTLS}^{-1}(\hat{\boldsymbol{\beta}}_{WTLS}-\boldsymbol{\beta})\xrightarrow{d}\chi_m^2 \quad \text{as } \ell\to\infty, \tag{27}$$

where the asymptotic variance of $\hat{\boldsymbol{\beta}}_{WTLS}$ is

$$\mathbf{V}_{WTLS}=(1+\boldsymbol{\beta}^T\boldsymbol{\beta})\tilde{\mathbf{\Delta}}_p^{-1}\{\tilde{\tilde{\mathbf{\Delta}}}_p+\tilde{\tau}(\mathbf{I}_m+\boldsymbol{\beta}\boldsymbol{\beta}^T)^{-1}\}\tilde{\mathbf{\Delta}}_p^{-1}. \tag{28}$$

We emphasize that the asymptotic variance $\mathbf{V}_{WTLS}$ in (28) depends on $\mathbf{\Sigma}_\varepsilon$, whose accurate estimation depends on (i) the limit of the ensemble estimator from the null simulation of climate model, (ii) proper handling of the high dimensionality of $\hat{\mathbf{\Sigma}}_\varepsilon$ due to ℓ having to be large to ensure the asymptotic normality.

It is noted that a similar asymptotic covariance matrix expression as (28) has been derived by Li et al. (2023) and Li and Li (2025) who employed the shrinkage error covariance estimators to counter the rank deficiency of the conventional sample covariance matrix in high dimension. However, there are several differences between the results presented in this paper and those in Li et al. (2023) and Li and Li (2025). First, the theory of Li et al. (2023) and Li and Li (2025) was largely based on the independent errors-in-variables framework of Gleser (1981). Consequently, while their results allow for spatial dependence, they required a strictly Gaussian assumption to safely apply the eigen decomposition as if the data were independent. In contrast, we allow for spatially dependent replications in the errors-in-variables regression without the normality assumption, which is better suited for climate change studies. Another difference is in using difference covariance estimators to Σ_ϵ in that Li et al. (2023) and Li and Li (2025) formulated the shrinkage estimator while we directly estimate it via the localization as will be shown in Section 5. Moreover, a key result of our work is in finding the source of the widely observed under-coverage issue of the confidence intervals proposed by AS03 and the following up studies in Section 3. The under-coverage issue has been a mystery in the optimal fingerprinting approach with a set of high profile following up works for instance in Ribes, Azaïs and Planton (2009) and DelSole et al. (2019). Our finding helps explain the widely observed under-coverage issue.

For easy referencing, Table 1 below summarizes recurring symbols used in the paper.

TABLE 1
Main notations used in the paper.

Symbol	Meaning
$\boldsymbol{\Sigma}_\varepsilon$	Covariance matrix of the internal variability error $\boldsymbol{\epsilon}$.
$\hat{\boldsymbol{\Sigma}}_\varepsilon$	Generic estimator of $\boldsymbol{\Sigma}_\varepsilon$ from control runs.
$\tilde{\boldsymbol{\Sigma}}_\varepsilon$	Limit of $\hat{\boldsymbol{\Sigma}}_\varepsilon$.
$\hat{\boldsymbol{\Sigma}}_S$	Shrinkage estimator of $\boldsymbol{\Sigma}_\varepsilon$.
$\hat{\boldsymbol{\Sigma}}_L$	Localization estimator of $\boldsymbol{\Sigma}_\varepsilon$.
$\boldsymbol{\Delta}_I$	$\lim_{\ell\to\infty} \ell^{-1}\mathbf{X}_0^\top\mathbf{X}_0$.
$\boldsymbol{\Delta}_\varepsilon$	$\lim_{\ell\to\infty} \ell^{-1}\mathbf{X}_0^\top\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{X}_0$.
$\boldsymbol{\Delta}_p$	$\lim_{\ell\to\infty} \ell^{-1}\mathbf{X}_0^\top\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{X}_0$.
$\hat{\boldsymbol{\Delta}}_p$	$\lim_{\ell\to\infty} \ell^{-1}\mathbf{X}_0^\top\tilde{\boldsymbol{\Sigma}}_\varepsilon^{-1}\mathbf{X}_0$.
$\tilde{\hat{\boldsymbol{\Delta}}}_p$	$\lim_{\ell\to\infty} \ell^{-1}\mathbf{X}_0^\top\tilde{\boldsymbol{\Sigma}}_\varepsilon^{-1}\boldsymbol{\Sigma}_\varepsilon\tilde{\boldsymbol{\Sigma}}_\varepsilon^{-1}\mathbf{X}_0$.
$\hat{\boldsymbol{\Delta}}_p$	Estimator of $\boldsymbol{\Delta}_p$.

5. Weighted total least squares confidence intervals. While Section 4 provides a general asymptotic framework of the WTLS estimator under potentially imperfect pre-whitening ($\tilde{\boldsymbol{\Sigma}}_\varepsilon \neq \boldsymbol{\Sigma}_\varepsilon$) for theoretical completeness, constructing valid and practical confidence intervals requires a consistent estimator of $\boldsymbol{\Sigma}_\varepsilon$. Therefore, in this section, we propose constructing the confidence region of $\boldsymbol{\beta}$ from the asymptotic distribution of $\hat{\boldsymbol{\beta}}_{WTLS}$ in (27) based on the consistent error covariance matrix estimator $\hat{\boldsymbol{\Sigma}}_\varepsilon \xrightarrow{p} \tilde{\boldsymbol{\Sigma}}_\varepsilon$. We assume that $\tilde{\boldsymbol{\Sigma}}_\varepsilon = \boldsymbol{\Sigma}_\varepsilon$, in which case the climate model that conducts the null simulation can fully capture the covariance structure of internal climate variability. In this case, (27)–(28) becomes

$$\ell(\hat{\boldsymbol{\beta}}_{WTLS} - \boldsymbol{\beta})^T\mathbf{V}_{GTLS}^{-1}(\hat{\boldsymbol{\beta}}_{WTLS} - \boldsymbol{\beta}) \xrightarrow{d} \chi_m^2 \text{ as } \ell\to\infty, \tag{29}$$

where the asymptotic variance $\mathbf{V}_{WTLS}$ is simplified to

$$\mathbf{V}_{GTLS} = (1+\boldsymbol{\beta}^T\boldsymbol{\beta})\boldsymbol{\Delta}_p^{-1}\{\boldsymbol{\Delta}_p + (\mathbf{I}_m + \boldsymbol{\beta}\boldsymbol{\beta}^T)^{-1}\}\boldsymbol{\Delta}_p^{-1}, \tag{30}$$

and $\boldsymbol{\Delta}_p = \lim_{\ell\to\infty}\ell^{-1}\mathbf{X}_0^T\boldsymbol{\Sigma}_\varepsilon^{-1}\mathbf{X}_0$.

To construct confidence regions, we need to estimate $\mathbf{V}_{GTLS}$. Under Conditions (22), (26a)–(26d) and that $\sqrt{\ell}\|\boldsymbol{\Sigma}_\varepsilon\|\|\hat{\boldsymbol{\Sigma}}_\varepsilon^{-1} - \boldsymbol{\Sigma}_\varepsilon^{-1}\| \xrightarrow{p} 0$, the analysis in Qiu, Chen and Chen (2026) shows that

$$\hat{\boldsymbol{\Delta}}_p = \ell^{-1}\mathbf{X}^T\hat{\boldsymbol{\Sigma}}_\varepsilon^{-1}\mathbf{X} - \hat{\sigma}^2\mathbf{I}_m \xrightarrow{p} \boldsymbol{\Delta}_p, \tag{31}$$

where $\hat{\sigma}^2 = \ell^{-1}\lambda_{m+1}([\mathbf{X},\mathbf{y}]^T\hat{\boldsymbol{\Sigma}}_\varepsilon^{-1}[\mathbf{X},\mathbf{y}]) \xrightarrow{p} \sigma^2 = 1$. Let

$$\hat{\mathbf{V}}_{GTLS} = (1+\hat{\boldsymbol{\beta}}_{WTLS}^T\hat{\boldsymbol{\beta}}_{WTLS})\left\{\hat{\sigma}^2\hat{\boldsymbol{\Delta}}_p^{-1} + \hat{\sigma}^4\hat{\boldsymbol{\Delta}}_p^{-1}(\mathbf{I}_m + \hat{\boldsymbol{\beta}}_{WTLS}\hat{\boldsymbol{\beta}}_{WTLS}^T)^{-1}\hat{\boldsymbol{\Delta}}_p^{-1}\right\}, \tag{32}$$

which is a consistent estimator of $\mathbf{V}_{GTLS}$.

By Slutsky's theorem, it can be derived from (29) that

$$Q(\boldsymbol{\beta}) = \ell(\hat{\boldsymbol{\beta}}_{WTLS} - \boldsymbol{\beta})^T\hat{\mathbf{V}}_{GTLS}^{-1}(\hat{\boldsymbol{\beta}}_{WTLS} - \boldsymbol{\beta}) \xrightarrow{d} \chi_m^2. \tag{33}$$

Then, an asymptotic 90% confidence region for $\boldsymbol{\beta}$ can be constructed as

$$\{\boldsymbol{\beta} \mid Q(\boldsymbol{\beta}) \le \chi_{m,0.1}^2\}, \tag{34}$$

which is a joint confidence region for $\boldsymbol{\beta}$. In the simulation and the real-data studies, we follow the convention in the optimal fingerprinting literature to use the derived marginal confidence intervals for individual scaling factors, constructed based on the diagonal entries of $\hat{\mathbf{V}}_{GTLS}$.

Specifically, for a particular scaling factor $\beta_i (i = 1, \cdots, m)$, an asymptotic 90% confidence interval for β_i is $(\hat{L}_i, \hat{U}_i)$ where

$$\hat{L}_i = \hat{\beta}_{WTLS,i} - \sqrt{\ell^{-1}\hat{V}_{ii}}\phi_{0.05} \quad \text{and} \quad \hat{U}_i = \hat{\beta}_{WTLS,i} + \sqrt{\ell^{-1}\hat{V}_{ii}}\phi_{0.05},$$

$\hat{V}_{ii}$ is the (i,i) entry of $\hat{\mathbf{V}}_{GTLS}$, and $\hat{\beta}_{WTLS,i}$ is the i-th element of $\hat{\boldsymbol{\beta}}_{WTLS}$, and $\phi_{0.05}$ is the 5% upper-quantile of the standard normal distribution.

As revealed by (27) and (28), the key factor that guarantees the convergence of $\hat{\mathbf{V}}_{GTLS}$ is that $\hat{\boldsymbol{\Sigma}}_\varepsilon \xrightarrow{p} \boldsymbol{\Sigma}_\varepsilon$. If $\hat{\boldsymbol{\Sigma}}_\varepsilon$ is not a consistent estimator for $\boldsymbol{\Sigma}_\varepsilon$, then $\hat{\mathbf{V}}_{GTLS}$ does not converge to $\mathbf{V}_{GTLS}$ and $Q(\boldsymbol{\beta})$ is no longer asymptotically χ^2_m distributed, which can contribute to inaccurate coverage for the confidence intervals.

A major cause for $\hat{\boldsymbol{\Sigma}}_\varepsilon$ not converging to $\boldsymbol{\Sigma}_\varepsilon$ is the high dimensionality of $\boldsymbol{\Sigma}_\varepsilon$. We note that (i) $\boldsymbol{\Sigma}_\varepsilon$ is $\ell \times \ell$ which grows as the sample size $\ell \to \infty$, and (ii) $\hat{\boldsymbol{\Sigma}}_\varepsilon$ is based on n ensemble runs of a climate model under the null setting, where the number of ensembles n may be less than ℓ due to high cost of running the ensembles. This makes the estimation of $\boldsymbol{\Sigma}_\varepsilon$ a high dimensional statistical problem, as noticed by Ribes, Planton and Terray (2013) and Chen, Chen and Mu (2024).

The high dimensionality is particularly pronounced for high resolution geophysical problems, which makes ℓ large. If $\hat{\boldsymbol{\Sigma}}_\varepsilon$ is the sample covariance matrix based on n ensembles, according to Bai and Yin (1993), $\hat{\boldsymbol{\Sigma}}_\varepsilon$ is no longer consistent to $\boldsymbol{\Sigma}_\varepsilon$ if $\ell/n \not\to 0$. Ribes, Azaïs and Planton (2009) employed the shrinkage estimator proposed by Ledoit and Wolf (2004) to counter the high dimensionality. A major development was in Bickel and Levina (2008a) that proposed the first generation of consistent estimator of high dimensional covariance matrix by a banding estimator for high dimensional covariance matrices with a univariate bandable structure. Other estimators include the L_q type sparse class that leads to the thresholding estimators in Bickel and Levina (2008b) and Cai, Liu and Luo (2011), the Kronecker-structured bandable class in Zhang, Shen and Kong (2023), and the general localization class in Sun, Chen and Qiu (2026). The localization class implies that the correlation between two locations is a decreasing function of the geographic distance, which is more general than the bandable class. In this paper, we adopt the localization approach, as it is well suited to the spatio-temporal data in climate change studies.

Let $\mathbf{Y}_N \in \mathbb{R}^{\ell \times n}$ be the control ensemble data matrix under the null climate setting, where n is the ensemble sample size and ℓ is the sample size of spatio-temporal replication. Denote the sample covariance matrix of $\mathbf{Y}_N$ as $\mathbf{S}_n = \mathbf{Y}_N \mathbf{Y}_N^T/(n-1) =: (\hat{s}_{ij})_{1 \le i,j \le \ell}$, where $\hat{s}_{ij}$ is the sample covariance between the i-th and the j-th components. Let d_{ij} denote the geographic distance between the location of two components i and j, and let t_i denote the time index associated with the i-th component. Given the spatial cutoff b_s and the temporal cutoff b_t, the localization covariance estimator is

$$\hat{\boldsymbol{\Sigma}}_{\mathrm{L}} = (\hat{s}_{ij} \cdot 1(d_{ij} \le b_s) \cdot 1(|\, t_i - t_j \,| \le b_t))_{1 \le i,j \le \ell} \tag{35}$$

where $1(\cdot)$ is the indicator function.

To implement the localization estimator, we need to choose the spatial and the temporal localization cutoffs b_s and b_t. We follow the sample-splitting idea of Bickel and Levina (2008a). We repeatedly split the control runs contained in $\mathbf{Y}_N$ randomly into two disjoint subsamples. Let $\mathcal{I}_{1,i}$ and $\mathcal{I}_{2,i}$ denote the training and validation subsamples for the ith split of sizes n_1 and $n_2 = n - n_1$, respectively. We took $n_1 = \lfloor n/3 \rfloor$ and repeated the random splitting for $M = 50$ times. Let $\hat{\boldsymbol{\Sigma}}_{1,i}(b_s, b_t)$ be the localized covariance estimator computed from the i-th split of the training subsample $\mathcal{I}_{1,i}$, and $\hat{\boldsymbol{\Sigma}}_{2,i}$ be the sample covariance matrix

computed from the i-th validation subsample $\mathcal{I}_{2,i}$. A penalized cross-validation (CV) score for the localization cutoff parameters is

$$\hat{R}(b_s, b_t) = \frac{1}{M}\sum_{i=1}^{M}\left\|\hat{\boldsymbol{\Sigma}}_{1,i}(b_s, b_t) - \hat{\boldsymbol{\Sigma}}_{2,i}\right\| + \eta\frac{\log(n)}{n}\frac{p(b_s, b_t)}{\ell}, \tag{36}$$

where $\|\cdot\|$ is the spectral matrix norm, η is a penalty parameter, and $p(b_s, b_t)$ is the number of free covariance parameters induced by the cutoff pair (b_s, b_t). The selected localization parameters are chosen to minimize the penalized CV score over a range of candidate values. Next two sections contain practical results on the selection of the two tuning parameters in the simulation and case studies.

6. Simulation study. To evaluate the performance of different estimators for the scaling factors and the associated coverage properties of the confidence intervals, we consider four estimators of $\boldsymbol{\beta}$: the generalized total least squares estimator with the true $\boldsymbol{\Sigma}_\varepsilon$ for data rotation, the total least squares estimator in (4) without data rotation, and the weighted total least squares estimators in (5) with two estimated $\hat{\boldsymbol{\Sigma}}_\varepsilon^{-1}$ for data rotation: the shrinkage estimator (Ledoit and Wolf, 2004) utilized in climate change studies by Ribes, Azaïs and Planton (2009) that combined the sample covariance matrix with the identity matrix and the localization estimator (Sun, Chen and Qiu, 2026) which assumes the localization covariance structure. The three rotation-based estimators (GTLS and the two WTLS estimators) were compared to demonstrate the effect of estimating $\boldsymbol{\Sigma}_\varepsilon$. Furthermore, we evaluated the coverage of confidence intervals based on Claim (6) in Allen and Stott (2003), the likelihood ratio formulation in DelSole et al. (2019) and the asymptotic distribution of the GTLS and WTLS estimators given in (33).

The key ingredients of the EIV Model (2) in the simulation were constructed with reference to the annual mean near-surface air temperature anomaly data generated by models participating in the Coupled Model Intercomparison Project Phase 5 (CMIP5), which included 35 and 46 ensemble members under certain ANT and NAT forcings, respectively, for the period of 1850–2020 at $2.8^\circ \times 2.8^\circ$ resolution over North America ($-130 \sim -50^\circ E, 30 \sim 60^\circ N$) (Swart et al., 2019). For the ANT and NAT forcings, each ensemble had identical external forcings with different initial conditions. We considered the temperature data used in Li et al. (2021), which partitioned the North America area into 48 spatial grids of $10^\circ \times 5^\circ$ and selected the 60-year time span from 1951–2010 divided into 5-year averaged temperature anomalies relative to the 1961–1990 baseline. Following the pre-processing steps of Ma et al. (2023) and Li et al. (2021), the period 1961–1965 and missing observations were excluded, yielding a total of $\ell = 512$ data points. To gain aspect of "asymptotic", namely to compare performance across different number of spatial replication ℓ, we further retained only the latter half of the processed space-time vector, resulting in a sample of $\ell = 253$.

The underlying error covariance matrix $\boldsymbol{\Sigma}_\varepsilon$ in Model (2) was assigned based on the null setting control runs of models in CMIP5 provided by Li et al. (2021). Specifically, for a given ℓ at a spatio-temporal resolution, from the 223 ensemble runs under the null setting control runs in CMIP5, we obtained the sample covariance matrix $\mathbf{S}_0$. We then applied the localization operation (35) with a spatial cutoff distance of 1000 km and a temporal cutoff of 5 years on $\mathbf{S}_0$, and truncated the negative eigenvalues to zero to obtain a non-negative definite $\boldsymbol{\Sigma}_0$. We fixed the spatial and temporal cutoffs as 1000 km and 5 years to represent a setting with both spatial and temporal dependence. We have also discussed other combinations of cutoff values and the results are presented in Section S8 of the SM. We finally set $\boldsymbol{\Sigma}_\varepsilon = \boldsymbol{\Sigma}_0 + 0.3\mathbf{I}_\ell$ to guarantee the positive definiteness.

Regarding the fingerprints, let $\hat{\mathbf{X}}_1$ and $\hat{\mathbf{X}}_2$ be the sample means of $a_1 = 35$ and $a_2 = 46$ ensemble runs under the ANT and NAT forcings, respectively. The underlying fingerprints

$\mathbf{x}_{0,i}$, for $i = 1$ and 2, were set to be $k\hat{\mathbf{X}}_i$, where $k = 0.5$ and 1 represented the weak and strong signal strengths, respectively, following Li et al. (2021). It is noted that the signal-to-noise ratio defined in DelSole et al. (2019) is monotone increasing with respect to k. According to Model (1), the observed fingerprints with measurement errors, $\mathbf{X}_1$ and $\mathbf{X}_2$, were generated from $N_\ell(\mathbf{x}_{0,i}, \mathbf{\Sigma}_\varepsilon)$, for $i = 1, 2$, respectively. We chose $\boldsymbol{\beta} = (\beta_{ANT}, \beta_{NAT})^T = (1, 0)^T$. The observations $\mathbf{y}$ were generated from $N_\ell(\mathbf{X}_0\boldsymbol{\beta}, \mathbf{\Sigma}_\varepsilon)$ following Model (2).

To implement the weighted total least squares, one needs to estimate $\mathbf{\Sigma}_\varepsilon$ in order to rotate the models and observations. Specifically, $\mathbf{Y}_N$ was generated from $n = 100, 200, 400$ and 600 ensemble runs, where each run was made from the multivariate normal distribution $N_\ell(\mathbf{0}, \mathbf{\Sigma}_\varepsilon)$. Clearly, the number of ensemble runs n was small relative to ℓ, the dimension of the "observation" $\mathbf{y}$, which brought up the high dimensionality that we have to attend to.

We considered two high dimensional estimators of $\mathbf{\Sigma}_\varepsilon$. One was the shrinkage estimator $\hat{\mathbf{\Sigma}}_S = \rho_1\mathbf{I}_\ell + \rho_2\mathbf{S}_n$ (Ribes, Azaïs and Planton, 2009) for estimating $\mathbf{\Sigma}_\varepsilon$ based on the sample covariance $\mathbf{S}_n = \mathbf{Y}_N\mathbf{Y}_N^T/n$, where ρ_1 and $\rho_2 \in (0, 1)$ were determined by the method in Ledoit and Wolf (2004). Specifically,

$$\rho_1 = \hat{\kappa}\hat{\mu}, \qquad \rho_2 = 1 - \hat{\kappa}, \tag{37}$$

where $\hat{\mu} = \ell^{-1}\mathrm{tr}(\mathbf{S}_n), \hat{\kappa} = \min\{\hat{\pi}/\|\mathbf{S}_n - \hat{\mu}\mathbf{I}_\ell\|_F^2, 1\}$, and $\hat{\pi} = \frac{1}{n}\sum_{j=1}^n \|\mathbf{Y}_{N,j}\mathbf{Y}_{N,j}^T - \mathbf{S}_n\|_F^2$. Then, $\mathbf{\Sigma}_\varepsilon^{-1}$ was estimated by the Moore–Penrose pseudoinverse of $\hat{\mathbf{\Sigma}}_S$ (Ribes, Azaïs and Planton, 2009), which may be regarded as a high dimensional refinement of Allen and Stott (2003). The other method was the localization estimator $\hat{\mathbf{\Sigma}}_L$, constructed according to (35) from the sample covariance $\mathbf{S}_n$. To select the cutoff values in the localization estimator, we used the penalized CV score given in Section 5 based on $M = 50$ repeated sample splitting. We considered three penalty levels of $\eta = 0.1, 0.2$ and 0.3 and examined ensemble sizes of $n = 100, 200, 400,$ and 600 and each experimental setting was repeated 100 times. Figure S1 of Section S3 in the SM displays the mean penalized CV score curves of the simulation experiments with respect to spatial and temporal cutoff values.The results showed that the cutoff pair (1000 km, 5 years) was selected in all cases with the ensemble size larger than 100. Therefore, we used (1000 km, 5 years) as the cutoff values in constructing $\hat{\mathbf{\Sigma}}_L$ in the simulation experiments, and $\hat{\mathbf{\Sigma}}_L^{-1}$ was the Moore-Penrose pseudo-inverse. We also experimented the oracle generalized total least squares that used the true $\mathbf{\Sigma}_\varepsilon$ in the pre-whitening operation to serve as a benchmark.

We constructed the 90% (climate change studies tend to use such confidence level) confidence intervals for the scaling factor β_{ANT} and β_{NAT} based on the asymptotic distribution of the weighted total least squares estimators with different asymptotic variance estimation. The following estimators of the variance of $\boldsymbol{\beta}$ were considered: Claim (6) by Allen and Stott (2003) (AS), the bias-corrected likelihood ratio in (16) by DelSole et al. (2019) (LR) and the asymptotic normality of the WTLS $\hat{\boldsymbol{\beta}}_{WTLS}$ in (33) (WTLS), the asymptotic normality of the total least squares $\hat{\boldsymbol{\beta}}_{TLS}$ in (24) (TLS) without the pre-whitening and the one associated with the generalized total least squares that used the true $\mathbf{\Sigma}_\varepsilon$ to serve as a benchmark (GTLS). We also experimented with the parametric bootstrap calibration (PBC) confidence intervals proposed in Li et al. (2021) with 500 bootstrap replications. For each configuration, the simulation was repeated 1000 times with the random number generated with seed 26 in R to ensure reproducibility.

Table 2 reports the empirical biases and the root mean square errors (RMSEs) for estimating β_{ANT} associated with the anthropogenic forcing, while those for β_{NAT} with the natural forcing are reported in Section S5 of the SM, which showed a similar pattern of results. The table shows that the biases of the estimators were quite small for all methods, and both the biases and RMSEs tended to decrease as the ensemble size n and the sample size ℓ were increased, which provided empirical support for the statistical consistency of all estimators.

(a) $k = 0.5$

ℓ	GTLS	TLS	n	WTLS	
				Shrinkage	Localization
253	0.0013 (0.299)	0.0032 (0.293)	100	0.0050 (0.293)	0.0015 (0.332)
			200	0.0003 (0.298)	0.0027 (0.322)
			400	0.0019 (0.299)	-0.0010 (0.315)
			600	0.0014 (0.299)	-0.0012 (0.313)
512	-0.0001 (0.258)	-0.0017 (0.244)	100	-0.0008 (0.244)	-0.0093 (0.411)
			200	0.0015 (0.258)	0.0035 (0.280)
			400	0.0019 (0.259)	0.0038 (0.272)
			600	0.0021 (0.260)	0.0036 (0.273)

(b) $k = 1$

ℓ	GTLS	TLS	n	WTLS	
				Shrinkage	Localization
253	0.0004 (0.126)	0.0007 (0.129)	100	0.0014 (0.129)	0.0001 (0.137)
			200	-0.0009 (0.132)	-0.0003 (0.134)
			400	-0.0001 (0.131)	-0.0013 (0.132)
			600	-0.0000 (0.131)	-0.0014 (0.132)
512	0.0030 (0.125)	0.0013 (0.123)	100	0.0017 (0.124)	0.0082 (0.456)
			200	0.0013 (0.123)	0.0024 (0.129)
			400	0.0021 (0.122)	0.0051 (0.128)
			600	0.0019 (0.124)	0.0028 (0.126)

TABLE 2

Average empirical biases (RMSEs) of four estimators of $\beta_{ANT} = 1$, the generalized total least squares with the true $\mathbf{\Sigma}_\varepsilon$ for the pre-whitening (GTLS), the total least squares (TLS) estimator without the pre-whitening and the weighted total least squares (WTLS) with $\mathbf{\Sigma}_\varepsilon$ estimated by the shrinkage estimator $\hat{\mathbf{\Sigma}}_S$ (Shrinkage) and the localization estimator $\hat{\mathbf{\Sigma}}_L$(Localization) by $(\hat{d}_s, \hat{d}_t) = (1000, 5)$, with the signal strength $k = 0.5$ and 1.

As expected, the GTLS with the true $\mathbf{\Sigma}_\varepsilon$ yielded the smallest RMSEs among estimators. The TLS without the pre-whitening step also performed competitively with the GTLS, The WTLS estimator with the shrinkage for covariance estimation had RMSEs close to those of GTLS and TLS, but did not show a clear trend as the ensemble size n was increased. The WTLS estimator with the localization formulation also showed performance similar to GTLS and TLS in large n settings, although it could be more variable for smaller n. Moreover, the RMSEs of all estimators decreased as the signal strength k of the fingerprints increased, indicating that the estimation is more accurate when the signal level is stronger.

Table 3 reports the coverage rates and lengths of the 90% confidence intervals (CIs) for β_{ANT} for the CIs of Allen and Stott (2003), the modified likelihood ratio (DelSole et al., 2019), the TLS and the WTLS with the shrinkage and the proposed localization estimator in the pre-whitening according to (29) and (32), and the GTLS, as described earlier. It is noted that $\mathbf{\Sigma}_\varepsilon$ still needs to be estimated in the construction of the CIs based on the asymptotic variance for the TLS CIs, while for the GTLS CIs, the true $\mathbf{\Sigma}_\varepsilon$ will be used. A key feature of the table is that the AS CIs suffered from the most severe under-coverage for all considered settings, which confirmed our discussion on Claim (6). Although the modified likelihood ratio (LR) method improved the coverage rates of the AS method, the CIs still suffered from severe under-coverage, which confirmed our discussion after (17). The two WTLS CIs with the shrinkage and the localization estimation of $\mathbf{\Sigma}_\varepsilon$ offered improved coverage rates than those of the CIs of the AS and the corrected LR. Among the two WTLS CIs, the localization CIs tended to have more accurate coverage rates, which was largely due to the localization CIs being wider than those of the shrinkage. While the under-coverage issue of the AS and

(a) $k = 0.5$

ℓ	GTLS	n	TLS	AS	LR	WTLS	PBC
				Shrinkage			
		100	0.853 (0.831)	0.724 (0.642)	0.737 (0.658)	0.848 (0.823)	0.851 (0.972)
		200	0.842 (0.841)	0.727 (0.650)	0.734 (0.667)	0.839 (0.836)	0.852 (1.011)
		400	0.854 (0.858)	0.724 (0.662)	0.738 (0.679)	0.842 (0.854)	0.878 (1.039)
253	0.913 (1.000)	600	0.862 (0.871)	0.731 (0.672)	0.745 (0.689)	0.841 (0.868)	0.884 (1.091)
				Localization			
		100	0.920 (0.999)	0.757 (0.764)	0.774 (0.786)	0.875 (1.012)	0.977 (15.383)
		200	0.916 (1.002)	0.761 (0.765)	0.768 (0.786)	0.889 (1.019)	0.923 (1.473)
		400	0.914 (0.998)	0.774 (0.763)	0.782 (0.783)	0.884 (1.012)	0.910 (1.358)
		600	0.913 (0.997)	0.774 (0.762)	0.781 (0.781)	0.890 (1.010)	0.908 (1.929)
				Shrinkage			
		100	0.851 (0.688)	0.732 (0.550)	0.741 (0.561)	0.846 (0.684)	0.845 (0.849)
		200	0.843 (0.724)	0.737 (0.576)	0.746 (0.587)	0.832 (0.722)	0.850 (0.859)
		400	0.846 (0.733)	0.736 (0.582)	0.743 (0.594)	0.836 (0.732)	0.865 (0.884)
512	0.900 (0.853)	600	0.853 (0.742)	0.736 (0.588)	0.748 (0.600)	0.846 (0.742)	0.861 (0.904)
				Localization			
		100	0.920 (0.838)	0.734 (0.668)	0.747 (0.716)	0.860 (0.937)	0.929 (7.426)
		200	0.911 (0.877)	0.773 (0.683)	0.779 (0.697)	0.887 (0.906)	0.942 (1.574)
		400	0.910 (0.873)	0.791 (0.679)	0.797 (0.693)	0.892 (0.898)	0.903 (1.163)
		600	0.910 (0.872)	0.778 (0.678)	0.792 (0.692)	0.888 (0.896)	0.905 (1.187)

(b) $k = 1$

ℓ	GTLS	n	TLS	AS	LR	WTLS	PBC
				Shrinkage			
		100	0.825 (0.349)	0.783 (0.319)	0.793 (0.324)	0.823 (0.344)	0.835 (0.379)
		200	0.836 (0.353)	0.798 (0.322)	0.805 (0.327)	0.825 (0.348)	0.835 (0.388)
		400	0.843 (0.362)	0.794 (0.329)	0.804 (0.334)	0.829 (0.355)	0.864 (0.406)
253	0.904 (0.410)	600	0.849 (0.369)	0.803 (0.333)	0.811 (0.338)	0.835 (0.360)	0.866 (0.420)
				Localization			
		100	0.908 (0.429)	0.843 (0.379)	0.849 (0.385)	0.870 (0.412)	0.989 (6.679)
		200	0.908 (0.433)	0.841 (0.379)	0.847 (0.385)	0.873 (0.414)	0.918 (0.487)
		400	0.907 (0.432)	0.841 (0.378)	0.850 (0.384)	0.876 (0.412)	0.909 (0.466)
		600	0.907 (0.432)	0.846 (0.378)	0.853 (0.384)	0.877 (0.412)	0.911 (0.464)
				Shrinkage			
		100	0.827 (0.335)	0.771 (0.301)	0.780 (0.306)	0.828 (0.332)	0.835 (0.338)
		200	0.830 (0.337)	0.771 (0.303)	0.787 (0.308)	0.830 (0.334)	0.848 (0.342)
		400	0.836 (0.342)	0.795 (0.307)	0.805 (0.312)	0.839 (0.338)	0.852 (0.352)
512	0.903 (0.419)	600	0.840 (0.347)	0.785 (0.311)	0.795 (0.316)	0.826 (0.343)	0.858 (0.360)
				Localization			
		100	0.920 (0.423)	0.804 (0.379)	0.814 (0.401)	0.863 (0.456)	0.951 (3.046)
		200	0.920 (0.422)	0.849 (0.370)	0.853 (0.376)	0.898 (0.419)	0.952 (0.681)
		400	0.910 (0.421)	0.847 (0.368)	0.850 (0.374)	0.892 (0.416)	0.896 (0.423)
		600	0.917 (0.421)	0.857 (0.368)	0.867 (0.374)	0.899 (0.416)	0.902 (0.421)

TABLE 3

Empirical coverages (lengths) of five 90% *confidence intervals for* β_{ANT} *with respect to sample size* ℓ*, the ensemble size* n *and the signal strength* k*. The confidence intervals were based on Claim* (6) *by Allen and Stott (2003) (AS), the bias-corrected LR in* (16) *by DelSole et al. (2019) (LR), the asymptotic normality of* $\hat{\boldsymbol{\beta}}_{WTLS}$ *in* (33) *(WTLS) and the parametric bootstrap clibration by Li et al. (2021) (PBC), with the shrinkage estimator* $\hat{\boldsymbol{\Sigma}}_S$ *and the localization estimator* $\hat{\boldsymbol{\Sigma}}_L$ *by* $(\hat{d}_s, \hat{d}_t) = (1000, 5)$*, respectively, and the generalized total least squares with true* $\boldsymbol{\Sigma}_\varepsilon$ *(GTLS).*

LR CIs was severer under the weaker signal setting of $k = 0.5$, more under-coverage was presence for the stronger signal setting of $k = 1.0$ for the WTLS CIs.

It is reassuring to see that as the number of observations ℓ increased, the lengths of most CIs were reduced despite the spatial dependence, while the effect on the coverage was less pronounced. It is good to see that the coverage rates of the CIs by the GTLS and the TLS with the localization method were generally satisfactory in terms of coverage rates and lengths. That the coverage rates of the oracle GTLS CIs were quite close to the nominal 90% offered some confirmation of the asymptotic results in (29). The GTLS CIs were the oracle counterpart of the two WTLS CIs. Although slightly suffered from the under-coverage issue when $n = 100$, the WTLS CIs with the localization were closer to those of the GTLS in terms of both coverage rates and the lengths especially when the number of null ensemble n was large ($n = 600$). When using the localization estimator for weaker signals, TLS yielded coverage rates closer to the 90% nominal level than those obtained by the WTLS. However, for stronger signals and larger n, the coverage rates of the WTLS CIs with the localization estimator provided coverage rates matching or even exceeding those of the TLS CIs, while having shorter lengths. Table 3 also shows that the WTLS CIs with the bootstrap improved the coverage of the shrinkage CIs by enlarging its length. However, the improvement was less effective for the localization CIs as it tended to produce CIs which were too wide especially for the weaker signal case with fewer null ensembles, resulting in an over-coverage issue. The effects of the signal strength were largely reflected by the lengths of the CIs with the weaker signal case having wider CIs than their stronger signal counterpart. This is consistent with the finding from Table 2 regarding the RMSEs of the estimators.

To gain better insight into the sources of the coverage errors of the WTLS CIs, given the quite satisfactory performances of the GTLS CIs, we examine the estimates for three key components of the asymptotic variance expression in (32), namely $\hat{\sigma}^2\hat{\mathbf{\Delta}}_p^{-1}$, $(\mathbf{I}_2 + \hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}^T)^{-1}$ and $1 + \hat{\boldsymbol{\beta}}^T\hat{\boldsymbol{\beta}}$, by the WTLS with the localization and the shrinkage methods with the GTLS serving as the benchmarks. Table 4 summarizes the results of the examination conducted for the same simulation settings as Table 2 by reporting the average empirical ratios for $\|\hat{\sigma}^2\hat{\mathbf{\Delta}}_p^{-1}\|_F/\|\sigma^2\mathbf{\Delta}_p^{-1}\|_F$ and $\|(\mathbf{I}_2 + \hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}^T)^{-1}\|_F/\|(\mathbf{I}_2 + \boldsymbol{\beta}\boldsymbol{\beta}^T)^{-1}\|_F$ and their standard deviations, respectively, where $\|\cdot\|_F$ is the Frobenius norm. The results on the ratio $(1 + \hat{\boldsymbol{\beta}}^T\hat{\boldsymbol{\beta}})/(1 + \boldsymbol{\beta}^T\boldsymbol{\beta})$ is shown in Table S1 in Section S4 of the SM. Table 4 and Table S1 shows the estimations of $(\mathbf{I}_2 + \hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}^T)^{-1}$ and $1 + \hat{\boldsymbol{\beta}}^T\hat{\boldsymbol{\beta}}$ had good quality for both versions of the WTLS. However, the estimation of $\hat{\sigma}^2\hat{\mathbf{\Delta}}_p^{-1}$ by the WTLS with the shrinkage method incurred substantial underestimation, while the localization method had little bias and was quite comparable to the oracle GTLS, especially for larger ensemble number n. This explains why the CIs with the shrinkage estimator for the covariance matrix tended to have more coverage errors than the CIs with the localization estimator.

The WTLS CIs were constructed based on the assumption that the underlying error distribution has its first four moments match those of the Gaussian distribution. To gain information on the robustness of the CIs against departures from the Gaussian-type moment Condition (22) in Section 4, we conducted simulation experiments using heavier-tailed multivariate t_ν-distributions for the random errors, with degrees of freedom $\nu \in \{16, 24, 32\}$ and covariance matrix $\mathbf{\Sigma}_\varepsilon$. The rest of the simulation design was kept identical, except that only the strong signal ($k = 1$) was evaluated. Table 5 reports the empirical coverage rates for the CIs for the ANT scaling factor together with their average length based on 1000 replications. The results showed that the coverage rates of the WTLS confidence intervals using the Gaussian-matching moment asymptotic variance were sensitive to the degree of freedom, which reflects the departure from the Gaussian assumption. In particular, the coverage rates of the WTLS CIs with the shrinkage method were more affected by the heavy-tailedness than those of the localization method. When the degree of freedom was small, the error distribution deviated more from the Gaussian distribution, and the WTLS CIs exhibited larger

(a) $k = 0.5$

		$\|\hat{\sigma}^2\hat{\mathbf{\Delta}}_p^{-1}\|_F/\|\sigma^2\mathbf{\Delta}_p^{-1}\|_F$			$\|(\mathbf{I}_2+\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}^\top)^{-1}\|_F/\|(\mathbf{I}_2+\boldsymbol{\beta}\boldsymbol{\beta}^\top)^{-1}\|_F$		
ℓ	n	GTLS	WTLS		GTLS	WTLS	
			Shrinkage	Localization		Shrinkage	Localization
253	100	0.992 (0.256)	0.822 (0.205)	1.009 (0.292)	0.966 (0.038)	0.969 (0.038)	0.962 (0.040)
	200		0.879 (0.224)	1.032 (0.287)		0.966 (0.038)	0.962 (0.038)
	400		0.894 (0.226)	1.026 (0.276)		0.966 (0.038)	0.963 (0.038)
	600		0.906 (0.230)	1.024 (0.273)		0.966 (0.037)	0.964 (0.038)
512	100	1.088 (0.201)	0.899 (0.162)	1.099 (0.251)	0.977 (0.037)	0.977 (0.035)	0.967 (0.040)
	200		0.863 (0.147)	1.024 (0.180)		0.980 (0.037)	0.978 (0.038)
	400		0.872 (0.148)	1.017 (0.175)		0.980 (0.036)	0.979 (0.037)
	600		0.878 (0.149)	1.015 (0.173)		0.980 (0.036)	0.979 (0.038)

(b) $k = 1$

		$\|\hat{\sigma}^2\hat{\mathbf{\Delta}}_p^{-1}\|_F/\|\sigma^2\mathbf{\Delta}_p^{-1}\|_F$			$\|(\mathbf{I}_2+\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}^\top)^{-1}\|_F/\|(\mathbf{I}_2+\boldsymbol{\beta}\boldsymbol{\beta}^\top)^{-1}\|_F$		
ℓ	n	GTLS	WTLS		GTLS	WTLS	
			Shrinkage	Localization		Shrinkage	Localization
253	100	0.966 (0.114)	0.802 (0.097)	0.975 (0.134)	0.993 (0.020)	0.993 (0.020)	0.991 (0.021)
	200		0.856 (0.105)	0.997 (0.127)		0.992 (0.020)	0.992 (0.020)
	400		0.871 (0.105)	0.995 (0.123)		0.992 (0.020)	0.992 (0.020)
	600		0.882 (0.107)	0.994 (0.121)		0.992 (0.020)	0.992 (0.020)
512	100	1.068 (0.092)	0.885 (0.077)	1.075 (0.117)	0.996 (0.017)	0.995 (0.017)	0.993 (0.019)
	200		0.850 (0.072)	1.006 (0.086)		0.996 (0.018)	0.996 (0.018)
	400		0.858 (0.073)	1.000 (0.083)		0.996 (0.018)	0.996 (0.017)
	600		0.865 (0.073)	0.998 (0.082)		0.996 (0.018)	0.996 (0.017)

TABLE 4

Empirical averages (standard deviations) of two key quantities in the asymptotic variance of the WTLS estimator, $\|\hat{\sigma}^2\hat{\mathbf{\Delta}}_p^{-1}\|_F/\|\sigma^2\mathbf{\Delta}_p^{-1}\|_F$ and $\|(\mathbf{I}_2+\hat{\boldsymbol{\beta}}\hat{\boldsymbol{\beta}}^\top)^{-1}\|_F/\|(\mathbf{I}_2+\boldsymbol{\beta}\boldsymbol{\beta}^\top)^{-1}\|_F$ with respect to sample size ℓ, the ensemble size n and the signal strength k, where $\hat{\mathbf{\Delta}}_p$ (defined in (31)) and $\hat{\boldsymbol{\beta}}$ were estimated based on the generalized total least squares with true $\mathbf{\Sigma}_\varepsilon$ (GTLS), and two WTLS estimators with the shrinkage estimator $\hat{\mathbf{\Sigma}}_S$ and the localization estimator $\hat{\mathbf{\Sigma}}_L$ with $(\hat{d}_s, \hat{d}_t) = (1000, 5)$.

ℓ	DF	GTLS	TLS	WTLS	
				Shrinkage	Localization
253	16	0.861 (2.695)	0.896 (0.695)	0.793 (6.285)	0.841 (1.117)
	24	0.869 (0.591)	0.896 (0.478)	0.799 (0.399)	0.857 (0.524)
	32	0.875 (0.512)	0.898 (0.476)	0.801 (0.393)	0.862 (0.508)
512	16	0.820 (31.836)	0.872 (0.602)	0.771 (0.527)	0.818 (40.358)
	24	0.840 (0.973)	0.879 (0.456)	0.787 (0.376)	0.841 (1.433)
	32	0.862 (0.508)	0.889 (0.429)	0.800 (0.346)	0.857 (0.598)

TABLE 5

Empirical coverage rates (lengths) of the 90% confidence intervals for β_{ANT} under Student's t errors with degrees of freedom (DF) 16, 24 and 32, with respect to sample size ℓ and the heavy-tail level of the regression errors based on 1000 simulations with $k = 1$. The confidence intervals are constructed from the generalized total least squares with the true $\mathbf{\Sigma}_\varepsilon$ (GTLS), the total least squares estimator without pre-whitening (TLS), and the weighted total least squares (WTLS) with $\mathbf{\Sigma}_\varepsilon$ estimated by the shrinkage estimator $\hat{\mathbf{\Sigma}}_S$ and the localization estimator $\hat{\mathbf{\Sigma}}_L$ with $(\hat{b}_s, \hat{b}_t) = (1000, 5)$.

coverage errors despite having longer lengths. As the degree of freedom increased, the error distribution became closer to Gaussian, and the overall performance of the WTLS CIs improved substantially. In contrast, the TLS CIs remained the most stable and generally gave coverage rates close to the nominal 0.9 level across all degrees of freedom, which were even closer than those of the GTLS CIs.

7. Detection and attribution of changes in mean temperature. To demonstrate the proposed weighted total least squares estimation and confidence intervals for the scaling factors, we conducted detection and attribution analysis on the annual mean near-surface air temperature data from 1951 to 2010 based on the data provided by Li et al. (2021). The 60-year span was divided into 5-year averaged anomalies of the temperature relative to the average over 1961–1990, with the period 1961–1965 removed according to the data pre-processing in Ma et al. (2023). The study regions were the globe (GL), North America (NA), Central North America (CNA) and the Northern Hemisphere (NH), the Northern Hemisphere mid-latitudes (NHM).

The fingerprints $\mathbf{X}$ and the null setting control runs $\mathbf{Y}_N$ were from the CMIP5 climate model simulations, with 35 and 46 runs for the ANT and NAT forcings and 223 control runs for the null setting, respectively. The observed temperature anomalies $\mathbf{y}$ were obtained from the HadCRUT4 dataset. The observations in each region over the different time intervals were gridded and then vectorized in the order of longitude, latitude, and time. The numbers of observations ℓ after removing missing values were 572 for the GL region at $40^\circ \times 30^\circ$ resolution, 512 with $10^\circ \times 5^\circ$ resolution for NA, 176 with $5^\circ \times 5^\circ$ resolution for CNA, 297 with $40^\circ \times 30^\circ$ resolution for NH, and 396 with $40^\circ \times 10^\circ$ resolution for NHM.

We computed the weighted total least squares estimates for the scaling factors β_{ANT} and β_{NAT} on the mean temperature changes with the shrinkage estimator $\hat{\mathbf{\Sigma}}_S$ and the localization estimator $\hat{\mathbf{\Sigma}}_L$ for the pre-whitening operation. The localization cutoffs were selected by the penalized CV score given in (36) with $\eta = 0.1$ in Section 5. Figure S2 in Section S7 of the SM displays the penalized CV score curves under various penalty levels. The selected cutoffs are reported in Table 6.

Region	Abbreviation	Spatial Cutoff (km)	Temporal Cutoff (years)	Max Distance (km)
Global	GL	12000	20	20015.10
North America	NA	6200	5	7522.50
Central North America	CNA	2300	0	2782.70
Northern Hemisphere	NH	10400	20	20015.10
Northern Hemisphere $30^\circ N$ to $70^\circ N$	NHM	10400	10	13343.40

TABLE 6

Selected localization cutoff values $(\hat{b}_s, \hat{b}_t)$ by the penalized CV score for the five study regions.

Three weighted total least squares confidence intervals are obtained based on the asymptotic distribution of the WTLS estimators with different asymptotic variance estimation, which were Claim (6) by Allen and Stott (2003) (AS), the bias-corrected LR in (16) by DelSole et al. (2019) (LR), the asymptotic normality of $\hat{\boldsymbol{\beta}}_{WTLS}$ in (33) (WTLS). We also compared the total least squares estimate with two confidence intervals based on the shrinkage estimator $\hat{\mathbf{\Sigma}}_S$ and the localization estimator $\hat{\mathbf{\Sigma}}_L$. Moreover, we also implemented the parametric bootstrap calibration (PBC) of Li et al. (2021) for the shrinkage estimator and the localization estimator.

Figure 2 displays the estimated scaling factors β_{ANT} and β_{NAT} for ANT and NAT forcings, respectively, and the corresponding 90% confidence intervals in each region. Overall, the two pre-whitening strategies, the shrinkage and the localization, produced different estimates (centers of confidence intervals) of the scaling factors, with the difference being large for the NAT signal. Among those based on the WTLS with the pre-whitening scheme, the AS and the LR confidence intervals were narrower than the WTLS confidence intervals based on

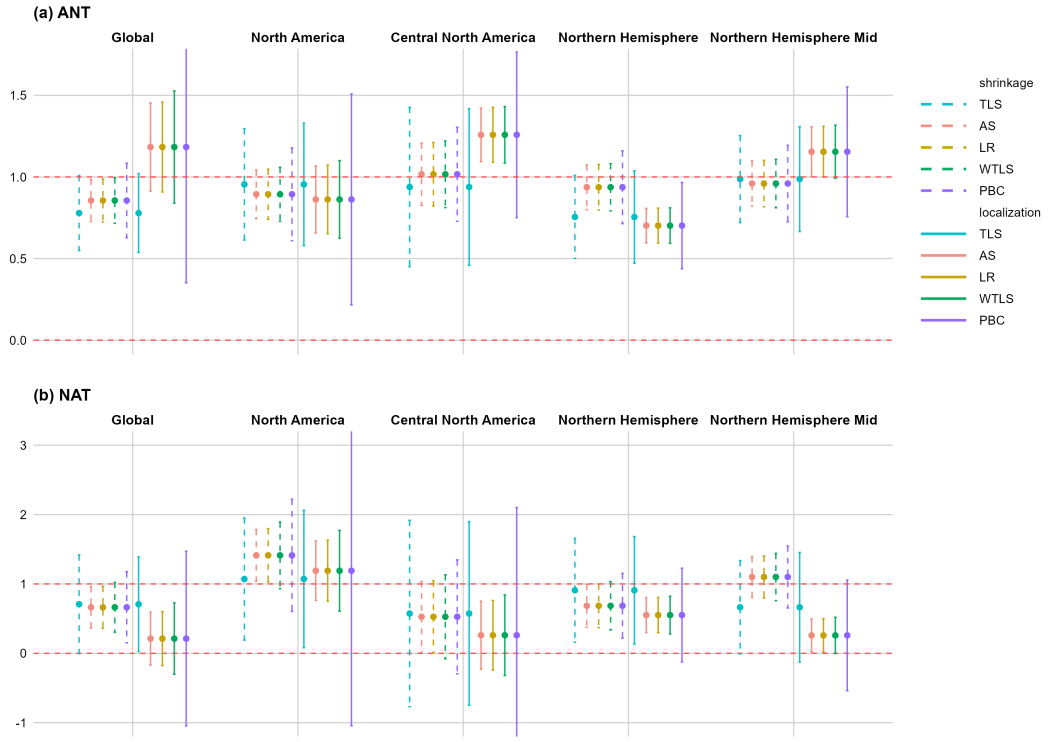


Figure 2: Estimates and the 90% confidence intervals for β_{ANT} (a) and β_{NAT} (b) for the mean temperature for five regions. The estimates are the WTLS and the TLS with the shrinkage or the localization estimation of Σ_ϵ. The confidence intervals are the WTLS based on Claim (6) (AS), the bias-corrected LR in (16) (LR) and the asymptotic normality of $\hat{\boldsymbol{\beta}}_{WTLS}$ in (33) (WTLS), and the parametric bootstrap calibration (PBC). The TLS confidence intervals are based on (24).

the correct asymptotic variance, which confirmed our theoretical findings in Section 3. The bootstrap-based PBC intervals were much longer than the other intervals.

For detection and attribution of the anthropogenic (ANT) signal reported in Panel (a) of Figure 2, all methods yielded detection for all five regions, since all the confidence intervals excluded 0. Attribution was also broadly supported, except the non-Bootstrap WTLS CIs with the localization for Central North America and Northern Hemisphere. Among the cases where the attribution could be made, the extent that 1 was included by the CIs varied by regions and by the methods (the shrinkage or the localization). The TLS without the pre-whitening attained detection and attribution for all five regions with either the shrinkage or the localization estimators of Σ_ϵ appeared in its asymptotic variance.

The results for the natural (NAT) forcing shown in Panel (b) of Figure 2 were quite different from those for the ANT forcing in that β_{NAT} was estimated much closer to 0 than those of β_{ANT}, which was expected. Because the NAT signal was weaker than the ANT signal, the corresponding confidence intervals were wider and the conclusion were more likely to change across methods and covariance estimators. The detection and attribution based on the WTLS with the shrinkage method were quite different from those based on the localization among those without the bootstrap implementation, for instance in the Global and Northern Hemisphere Mid region where the shrinkage gave detection and attribution but the localization based WTLS offered the opposite conclusion. The two methods were only agreeable in the North America region. In contrast both strains of the TLS CIs were less affected by the

way the Σ_ϵ was estimated for both ANT and NAT signal detection and attribution as the Σ_ϵ estimation would not alter the TLS estimates and the center of the CIs. The TLS had detection and attribution for only two of the five regions, which were for all five regions for the ANT signal.

8. Discussion. This paper offered a comprehensive review on the optimal fingerprinting approach built upon the statistical errors-in-variables regression model proposed in Allen and Stott (2003). It has identified the source of the widely observed under-coverage of the confidence intervals used in the detection and attribution of the climate change analyses. We have repaired the under-coverage problem by proposing a new weighted total least squares approach that is applicable to spatial and temporal climate data, based on a localization estimator of the high dimensional error covariance matrix.

Our study finds the weighted TLS method was quite sensitive to the estimator of the high dimensional residual covariance matrix, as demonstrated in the case study. As there is no theoretical guarantee of the benefits for doing the pre-whitening in the presence of the errors-in-variable, and the quite high computational cost under the high dimensionality, the TLS method without the pre-whening appears to be attractive. Our simulation and case studies revealed its robust performance.

We have assumed that the controlled null simulations used to estimate Σ_ϵ are statistically independent of the observation Y and the simulated fingerprints X. Ideally, the control runs used in detection and attribution are typically pre-industrial control simulations under fixed 1850 forcings, designed to isolate internally generated variability rather than to reproduce the observed trajectory. The observations and the control run outputs are largely separated in terms of simulation design (Eyring et al., 2016). Thus, the independence between the control runs and the observations should be maintained, which is also ideally needed as stated in the review paper Chen, Chen and Mu (2024). In practice, there may be some degree of information leakage may in principle arise through model development and tuning, since climate models are often calibrated against selected aspects of the observed climate. Even in the case of information leakage that renders the independence between the control runs and the observation, as noted by Qiu, Chen and Chen (2026), the theoretical results would still be valid. This is because the asymptotic normality of the WTLS estimator can be attained by using a perturbation argument, which only requires that Σ_ε can be accurately enough estimated by $\hat{\Sigma}_\varepsilon$ and thus relaxes the independent condition.

Our theory for the WTLS and the TLS estimators in Section 4 brings a perspective that covariance estimator based on the null ensembles $\widehat{\Sigma}_\varepsilon \to \widetilde{\Sigma}_\varepsilon$ which may differs from the true Σ_ε. This prospect had been raised in Chen, Chen and Mu (2024) when discussing the optimality of Allen and Tett (1999) fingerprint estimator without measurement errors. The conclusion was that if $\widetilde{\Sigma}_\varepsilon \neq \Sigma_\varepsilon$, the Allen–Tett scaling factor estimator is no longer optimal. Raising the prospect of $\widetilde{\Sigma}_\varepsilon \neq \Sigma_\varepsilon$ represents a general setting, and Section 4 establish the asymptotic theory for the WTLS estimator under such general setting. It also allows a complete theory that characterizes the impact of using an imperfect covariance estimator. However, in practice, the confidence regions of all methods have to rely on a consistent estimator of Σ_ϵ under $\hat{\Sigma}_\epsilon = \Sigma_\epsilon$, the WTLS or the TLS alike, and the shrinkage or the localization estimators alike. That is the rationale for providing the asymptotic distribution under $\widetilde{\Sigma}_\varepsilon = \Sigma_\varepsilon$ in Section 5, which allows for direct comparison between the general $\boldsymbol{V}_{WTLS}$ in (28) and $\boldsymbol{V}_{GTLS}$ in (30). It is hoped that with the advance of climate system modeling the difference between $\widetilde{\Sigma}_\varepsilon$ and Σ_ε would be reduced as our understanding on the earth system progresses.

Data availability. The data used in this study were taken from the supplementary materials of Li et al. (2021). The observational data are based on HadCRUT4 and the fin-

gerprint and control-run data are based on CMIP5 multi-model simulations. The processed dataset is available via https://figshare.com/articles/dataset/Uncertainty_in_Optimal_Fingerprinting_is_Underestimated/14981241/3.

The code for simulations and case study is publicly available at https://github.com/mirrorchy/Undercoverage-of-Optimal-Fingerprinting.

Funding. This research was supported by the National Natural Science Foundation of China Grant 12292983.

SUPPLEMENTARY MATERIAL

Supplementary Information for "Valid Confidence Intervals for Detection and Attribution in Climate Change Studies"

The supplementary material contains mathematical proofs and additional simulation results.

REFERENCES

ALLEN, M. and STOTT, P. (2003). Estimating signal amplitudes in optimal fingerprinting, part I: Theory. *Climate Dynamics* **21** 477–491.

ALLEN, M. and TETT, S. (1999). Checking for model consistency in optimal fingerprinting. *Climate Dynamics* **15** 419–434.

BAI, Z. D. and YIN, Y. Q. (1993). Limit of the smallest eigenvalue of a large dimensional sample covariance matrix. *Annals of Probability* **21** 1275–1294.

BICKEL, P. J. and LEVINA, E. (2008a). Regularized estimation of large covariance matrices. *The Annals of Statistics* **36** 199 – 227.

BICKEL, P. J. and LEVINA, E. (2008b). Covariance regularization by thresholding. *Ann. Statist.* **36** 2577–2604.

BOCK, L., LAUER, A., SCHLUND, M., BARREIRO, M., BELLOUIN, N., JONES, C., MEEHL, G. A., PREDOI, V., ROBERTS, M. J. and EYRING, V. (2020). Quantifying progress across different CMIP phases with the ESMValTool. *J. Geophys. Res.* **125**.

CAI, T., LIU, W. and LUO, X. (2011). A constrained ℓ_1 minimization approach to sparse precision matrix estimation. *J. Amer. Statist. Assoc.* **106** 594–607.

CHEN, H., CHEN, S. X. and MU, M. (2024). A statistical review on the optimal fingerprinting approach in climate change studies. *Climate Dynamics* **62** 1439–1446.

CHEN, H., CHEN, S. X. and QIU, J. (2025). Comments on "Consistent climate fingerprinting" by McKitrick (2025). *Climate Dynamics* **63** 261.

CHEN, S. X., ZHANG, L.-X. and ZHONG, P.-S. (2010). Tests for high-dimensional covariance matrices. *Journal of the American Statistical Association* **105** 810–819.

CRESSIE, N. A. C. (2015). *Statistics for Spatial Data*, revised ed. *Wiley Classics Library*. John Wiley & Sons, Inc., New York.

DELSOLE, T., TRENARY, L., YAN, X. and TIPPETT, M. K. (2019). Confidence intervals in optimal fingerprinting. *Climate Dynamics* **52** 4111–4126.

DOUKHAN, P. (1994). *Mixing: Properties and Examples*. *Lecture Notes in Statistics* **85**. Springer-Verlag, New York.

EYRING, V., BONY, S., MEEHL, G. A., SENIOR, C. A., STEVENS, B., STOUFFER, R. J. and TAYLOR, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development* **9** 1937–1958.

FULLER, W. A. (1987). *Measurement Error Models*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons, Inc., New York.

GILLETT, N. P., KIRCHMEIER-YOUNG, M., RIBES, A., SHIOGAMA, H., HEGERL, G. C., KNUTTI, R., GASTINEAU, G., JOHN, J. G., LI, L., NAZARENKO, L., ROSENBLOOM, N., SELAND, Ø., WU, T., YUKIMOTO, S. and ZIEHN, T. (2021). Constraining human contributions to observed warming since the pre-industrial period. *Nat. Clim. Chang.* **11** 207–212.

GLESER, L. J. (1981). Estimation in a multivariate "errors in variables" regression model: Large sample results. *The Annals of Statistics* **9** 24–44.

GOLUB, G. H. and VAN LOAN, C. F. (1980). An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis* **17** 883–893.

HASSELMANN, K. (1979). On the signal-to-noise problem in atmospheric response studies. *Meteorology of the Tropical Oceans* 251–259.

IPCC (1990). *Climate Change: The IPCC Scientific Assessment*. [Houghton, J.T., G. J. Jenkins, and J. J. Ephraums (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 365 pp.

IPCC (1996). *Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change*. [Houghton, J.T., L.G. Meira Filho, B.A. Callander, N. Harris, A. Kattenberg, and K. Maskell (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 572 pp.

IPCC (2001). *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. [Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 881 pp.

IPCC (2007). *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp.

IPCC (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

IPCC (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2391 pp.

JENISH, N. and PRUCHA, I. R. (2009). Central limit theorems and uniform laws of large numbers for arrays of random fields. *J. Econometrics* **150** 86–98.

KENDALL, M. and STUART, A. (1977). *The Advanced Theory of Statistics.*, 4th ed. MacMillan, New York.

LEDOIT, O. and WOLF, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis* **88** 365–411.

LI, H. and LI, Y. (2025). Regularized Fingerprinting with Linearly Optimal Weight Matrix in Detection and Attribution of Climate Change. Available at https://doi.org/10.48550/arXiv.2505.04070.

LI, Y., CHEN, K., YAN, J. and ZHANG, X. (2021). Uncertainty in optimal fingerprinting is underestimated. *Environmental Research Letters* **16** 084043.

LI, Y., CHEN, K., YAN, J. and ZHANG, X. (2023). Regularized fingerprinting in detection and attribution of climate change with weight matrix optimizing the efficiency in scaling factor estimation. *The Annals of Applied Statistics* **17** 225 – 239.

MA, S., WANG, T., YAN, J. and ZHANG, X. (2023). Optimal fingerprinting with estimating equations. *Journal of Climate* **36** 7109–7122.

MCKITRICK, R. (2022). Checking for model consistency in optimal fingerprinting: A comment. *Climate Dynamics* **58** 405–411.

MCKITRICK, R. (2025). Consistent climate fingerprinting. *Climate Dynamics*. To appear.

NORTH, G. R., KIM, K.-Y., SHEN, S. S. P. and HARDIN, J. W. (1995). Detection of forced climate signals. part 1: Filter theory. *Journal of Climate* **8** 401–408.

PEARSON, K. (1901). On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **2** 559–572.

POLSON, D., HEGERL, G., ZHANG, X. and OSBORN, T. (2013). Causes of robust seasonal land precipitation changes. *Journal of Climate* **26** 6698–6715.

QIU, J., CHEN, H. and CHEN, S. X. (2026). Errors-in-variables regression for dependent data with estimated error covariance matrix: To prewhiten or not? Available at https://doi.org/10.48550/arXiv.2601.01351.

RIBES, A., AZAÏS, J. and PLANTON, S. (2009). Adaptation of the optimal fingerprint method for climate change detection using a well-conditioned covariance matrix estimate. *Climate Dynamics* **33** 707–722.

RIBES, A., PLANTON, S. and TERRAY, L. (2013). Application of regularised optimal fingerprinting to attribution. Part I: Method, properties and idealised analysis. *Climate Dynamics* **41** 2817–2836.

RIGHI, M., ANDELA, B., EYRING, V., LAUER, A., PREDOI, V., SCHLUND, M., VEGAS-REGIDOR, J., BOCK, L., BRÖTZ, B., DE MORA, L., DIBLEN, F., DREYER, L., DROST, N., EARNSHAW, P., HASSLER, B., KOLDUNOV, N., LITTLE, B., LOOSVELDT TOMAS, S. and ZIMMERMANN, K. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 – technical overview. *Geosci. Model Dev.* **13** 1179–1199.

ROSENBLATT, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U.S.A.* **42** 43–47.

SUN, H.-X., CHEN, S. X. and QIU, Y. (2026). Localization Estimator for High-dimensional Tensor Covariance Matrices. Available at https://doi.org/10.48550/arXiv.2601.06989.

SWART, N. C., COLE, J. N. S., KHARIN, V. V., LAZARE, M., SCINOCCA, J. F., GILLETT, N. P., ANSTEY, J., ARORA, V., CHRISTIAN, J. R., HANNA, S., JIAO, Y., LEE, W. G., MAJAESS, F., SAENKO, O. A., SEILER, C., SEINEN, C., SHAO, A., SIGMOND, M., SOLHEIM, L., VON SALZEN, K., YANG, D. and WINTER, B. (2019). The Canadian Earth System Model version 5 (CanESM5.0.3). *Geoscientific Model Development* **12** 4823–4873.

VAN HUFFEL, S. and VANDEWALLE, J. (1991). *The Total Least Squares Problem: Computational Aspects and Analysis*. Society for Industrial and Applied Mathematics, Philadelphia.

ZHANG, Y., SHEN, W. and KONG, D. (2023). Covariance estimation for matrix-valued data. *J. Amer. Statist. Assoc.* **118** 2620–2631.