Climate Change Fingerprinting

with help from chatGPT

May 13, 2025

Contents

1 Classical Climate Change Fingerprinting

1.1 Set up the linear model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

- $\mathbf{y}$ — N-vector of observed anomalies (e.g. gridded temperatures concatenated over space $\times$ time).
- $\mathbf{X}$ — $N \times J$ design matrix whose J columns are the *fingerprints* (model-simulated responses) of the external forcings of interest: greenhouse gases, aerosols, solar, volcano, ...
- $\boldsymbol{\beta}$ — unknown *scaling factors* (one per forcing) to estimate and test.
- $\boldsymbol{\varepsilon}$ — internal variability, assumed mean 0 with covariance $\boldsymbol{\Sigma} = \mathrm{Cov}[\boldsymbol{\varepsilon}]$.

Because $\mathbf{y}$ has strong space–time correlations, ordinary least squares (OLS) would underestimate uncertainty. Generalized least squares (GLS) fixes this by explicitly using $\boldsymbol{\Sigma}$.

1.2 Estimate the noise covariance matrix

Take $\boldsymbol{\Sigma}$ from long control-run simulations, filtered exactly like the observations:

$$\hat{\boldsymbol{\Sigma}} \;=\; \frac{1}{M}\sum_{m=1}^{M}\big(\mathbf{c}^{(m)}-\bar{\mathbf{c}}\big)\big(\mathbf{c}^{(m)}-\bar{\mathbf{c}}\big)^{\top}. \tag{2}$$

Large N makes $\hat{\boldsymbol{\Sigma}}$ ill-conditioned, so in practice we

- rotate into the leading EOFs of $\hat{\boldsymbol{\Sigma}}$ and truncate, and/or
- apply shrinkage or ridge regularisation.

1.3 Pre-whiten

Compute the symmetric square-root inverse $\hat{\boldsymbol{\Sigma}}^{-1/2}$ (in the retained EOF subspace) and premultiply:

$$\tilde{\mathbf{y}} = \hat{\boldsymbol{\Sigma}}^{-1/2}\mathbf{y}, \qquad \tilde{\mathbf{X}} = \hat{\boldsymbol{\Sigma}}^{-1/2}\mathbf{X}.$$

After whitening the noise is (approximately) i.i.d. and we may use OLS on $(\tilde{\mathbf{y}}, \tilde{\mathbf{X}})$. In compact form

$$\boxed{\hat{\boldsymbol{\beta}}_{\mathrm{GLS}} \;=\; \big(\mathbf{X}^{\top}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{X}\big)^{-1}\mathbf{X}^{\top}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{y}} \tag{3}$$

with $\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \big(\mathbf{X}^{\top}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{X}\big)^{-1}$.

1.4 Detection and attribution

1) **Detection**: For each forcing j, test $H_0\colon \beta_j = 0$. Rejecting (e.g. at 5%) means the fingerprint is *detected*.

2) **Attribution / consistency**: Check whether β_j is statistically consistent with 1. If 1 lies in the confidence interval, the model response amplitude is consistent with observations.

3) **Residual consistency**:

$$Q = \frac{\mathbf{r}^\top \hat{\boldsymbol{\Sigma}}^{-1}\mathbf{r}}{N-J}, \qquad \mathbf{r} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}.$$

Compare Q with the corresponding χ^2 quantile; failure suggests a missing forcing, non-linear response, or mis-specified $\hat{\boldsymbol{\Sigma}}$.

1.5 Practical refinements

Issue	**Typical remedy**
Errors in fingerprints $\mathbf{X}$ (finite ensemble size)	Total-least-squares (TLS), *generalised* TLS, or recent estimating-equation approaches.
Multicollinearity between forcings	Dimensionality reduction (eigen-fingerprints), ridge regression, Bayesian priors.
High-dimensional $\boldsymbol{\Sigma}$	Shrinkage estimators, covariance tapering, or projection onto a reduced EOF basis before GLS.
Low signal-to-noise	Optimal filtering: weight fingerprints toward low-noise directions (the original "optimal fingerprint" idea).

1.6 Readings

- Hegerl, G. C. *et al.* (1996), *Journal of Climate*.
- Hannart, A. (2019), *ASCMO* — explanation of projection choices and S/N optimization.

1.7 Key take-aways

- Classical fingerprinting is *just* a GLS regression whose design matrix is built from model-simulated responses, and whose noise covariance comes from control runs.

- The estimated *scaling factors* quantify how strongly each forcing contributes to the observed change; their confidence intervals provide the detection/attribution statements.
- Everything hinges on a realistic $\mathbf{\Sigma}$; dimension reduction and regularisation are therefore essential in real applications.

1.8 Covariance Matrix Estimation

Accurate estimation of the internal variability covariance $\mathbf{\Sigma}$ is crucial especially when it needs to be inverted. First, the Sample Covariance Matrix is calculated as

$$C = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{z}_i - \bar{\mathbf{z}})(\mathbf{z}_i - \bar{\mathbf{z}})^T.$$

This matrix is often unstable or singular when the number of variables (grid points) p is large relative to the sample size (number of months) N. To address this, we can use a Shrinkage Estimator,

$$\widehat{\mathbf{\Sigma}} = \lambda\mathbf{T} + (1-\lambda)C, \quad \lambda \in [0,1],$$

where

- $\mathbf{T}$: stable target (e.g. scaled identity or diagonal).
- λ: shrinkage intensity.

This guarantees positive-definiteness, reduces estimation error in high dimensions. One specific variant is the Ledoit–Wolf Optimal Shrinkage, where,

$$\lambda^* = \frac{\mathbb{E}\big[\|\mathbf{S} - \mathbf{\Sigma}\|_F^2\big]}{\mathbb{E}\big[\|\mathbf{T} - \mathbf{\Sigma}\|_F^2 + \|\mathbf{S} - \mathbf{T}\|_F^2\big]},$$

estimated with plug-in values from the data.

Other Estimators are Tapering/Banding which impose distance-based decay of covariances, Factor Models, which use a low-rank signal plus white noise, or Sparse Precision Estimators (graphical lasso) which estimate a sparse inverse covariance (conditional dependencies).