Climate Dynamics (2025) 63:261
https://doi.org/10.1007/s00382-025-07750-4

ORIGINAL ARTICLE

Comments on “Consistent climate fingerprinting"by McKitrick (2025)

Hanyue Chen[1] · Song Xi Chen[2] · Jingkun Qiu[3]

Received: 15 May 2025 / Accepted: 2 June 2025 / Published online: 16 June 2025

Abstract
We provide comments on McKitrick (Consistent climate fingerprinting. Manuscript, 2025) Consistent Climate Fingerprinting, outlining our views on some keys points of the paper.

Keywords Instrumental variables · Optimal fingerprints · Residual consistent test · Climate change

We are envy of Professor McKitrick for having another paper on the climate fingerprints (OF) published in *Climate Dynamics*, as we know so well the challenges in having only one paper published in the journal. In the rest of the comments, we refer to McKitrick as the author, the paper (McKitrick 2025) under comment as M25, and AT99 for Allen and Tett (1999). Like M25, we also use the notations of Chen et al. (2024) which is referred to as CCM24 rather than C23 in M25.

M25 claimed that *“Chen et al. (2023) (C23) concur on the key points of M22"in the abstract, and “C23 restated and reaffirmed the main critique in M22 on AT99” in Section 2.1*. These accounts are not accurate as in fact the review of CCM24 was more understanding to Allen and Tett (1999), which was sharply different from the critiques of McKitrick (2022) (M22). Furthermore, CCM24 found many of M22’s critiques were either over-stating or unnecessary.

We have to remember Allen and Tett were not statisticians. We were and still are marveled by their endeavor to formulate more accurate estimates of the scaling factors β with the help of the Gauss-Markov Theorem in statistics and climate models. Their advocate for using separate null simulation to estimate the covariance matrix $\mathbf{C}_N$ of the climate noise $\mathbf{u}$ was very stimulating to us, as statisticians like us or econometricians like Professor McKitrick tended to be confined on the estimated residuals of the regression fits and build statistical models on them. This is understandable as earth scientists are believers of earth system models.

The rest of our comments on the M25 are focused on five aspects, which are structured by first presenting the views of M25 (in italic), followed by our comments.

1 Unbiasedness and consistency of $\hat{\boldsymbol{\beta}}$

A key point in CCM24 is that assuming $E(\mathbf{u} \mid \mathbf{X}, \mathbf{P}) = E(\mathbf{u} \mid \mathbf{X}) = 0$ would effectively counter much of the critiques of M22 on AT99 regarding unbiasness and consistency of the so-called fingerprint estimator $\hat{\boldsymbol{\beta}}$.

1.1. The author questions $E(\mathbf{u} \mid \mathbf{X}, \mathbf{P}) = 0$ by raising the possible dependence between the climate simulations which generates the fingerprints $\mathbf{X}$ and the null ensembles $\mathbf{Y}_N$ leading to $\mathbf{P}$ for estimating $\mathbf{C}_N$, as stated in Section 2.2 *“We can question whether repeated runs from the same model are truly independent, and since different models share a common physical representation of the climate system even runs from different models may not be fully independent."*

We disagree with this entirely. The independence we refer to is statistical independence rather than physical independence. Generating statistically independent ensembles are commonly used in many scientific fields in Monte Carlo simulations and in Statistics, in climate simulations and in Ensemble Kalman Filters in particular, based on generating pseudo-independence random numbers. Certainly, physical features would be repeated among the statistically independence ensembles, which has no bearing on the discussion here. Indeed, generating independent Gaussian distributed

✉ Song Xi Chen
sxchen@tsinghua.edu.cn

1 Center for Statistical Science, Peking University, Beijing, China

2 Department of Statistics and Data Science, Tsinghua University, Beijing, China

3 Guanghua School of Management, Peking University, Beijing, China

data samples does not mean the simulation should not bear the feature of the Gaussian distribution. In fact, they should be so.

1.2. The author claimed in lines 146 to 149 that *"The consistency proof in C23 involves three additional conditions labeled C.13(i-iii). The second is $\hat{\alpha}_\kappa^{-2} = O_p(1)$ which means, in words, that the inverse of the squared κ th (smallest) eigenvalue of $\hat{\mathbf{C}}_N$ is bounded in probability above or, equivalently, that the eigenvalue does not go to zero even asymptotically as $\ell, n \to \infty$. Likewise the third requires $\hat{\alpha}_\kappa^{-4}$ not go to zero."*

This is a misunderstanding as κ is the number of non-zero eigenvalues of $\hat{\mathbf{C}}_N$. Hence, the eigenvalue conditions C.13 (ii) and (iii) in CCM24 only restrict the κ-th largest eigenvalue of $\hat{\mathbf{C}}_N$ rather than the smallest eigenvalue, which means that the concern of the author is unfounded. Moreover, C.13 (iii) in CCM24 is $E(\hat{\lambda}_1^2\hat{\alpha}_\kappa^{-4}) = O(1)$, which is not equivalent to "$\hat{\alpha}_\kappa^{-4}$ not go to zero"as the author thought.

2 Optimality of $\hat{\boldsymbol{\beta}}$, and $\hat{\mathbf{C}}_N$, $\tilde{\mathbf{C}}_N$ and $\mathbf{C}_N$

Another insight in CCM24 was made by (i) introducing $\tilde{\mathbf{C}}_N$ as the underlying covariance matrix prescribed by the null climate model and (ii) $\tilde{\mathbf{C}}_N$ may be different from $\mathbf{C}_N$, the covariance of the climate noise $\mathbf{u}$. These allow imperfect null simulation such that $\tilde{\mathbf{C}}_N \neq \mathbf{C}_N$, when $\hat{\boldsymbol{\beta}}$ is only restricted BLUE rather than a full BLUE (best linear unbiased estimator). We hope that with the advance of the modeling of the climate system, the gap between $\tilde{\mathbf{C}}_N$ and $\mathbf{C}_N$ would be reduce as the science advances and thus, the estimator $\hat{\boldsymbol{\beta}}$ in the AT99 would be more accurate with more accurate climate models.

2.1. *The author pointed out several typos in CCM24 that "They refer to a matrix C_N (not bold-faced) which has rank κ' and is formed using the first κ' eigenvectors of $\boldsymbol{C}_N$. Somewhat confusingly they occasionally treat C_N and $\boldsymbol{C}_N$ as interchangeable (such as in Remark 7).... Also in Remark 2 and elsewhere they use P (not bold-faced) interchangeably with **P** (bold-faced)."(lines 89–97)*

The notations C_N and P (not bold-faced) in CCM24 were typos, and should be $\mathbf{C}_N$ and $\mathbf{P}$ (bold-faced), which should explain much of the complaints of M25 above. CCM24 defined the rank of $\mathbf{C}_N$ as κ' ($\kappa' \leq \ell$), which covers the general cases where $\mathbf{C}_N$ can be reduced rank. As ℓ increases with the increase of the resolution of observations, it is natural to consider that the covariance of the climate internal error may not have full rank.

2.2. *The author was quite concerned with the reduced rank of $\hat{\boldsymbol{C}}_N$ as shown in "CCM24 point out that we must assume the availability of a full rank estimator $\hat{\boldsymbol{C}}_N$ even as ℓ becomes arbitrarily large."(lines 149–151)*

This should not be an issue since CCM24 was quite clear in defining $\tilde{\mathbf{C}}_N$ as the probability limit of $\hat{\mathbf{C}}_N$ as $n \to \infty$. Hence, there is no contradiction between $\hat{\mathbf{C}}_N$ being rank deficient, due to the number of ensemble n being less than ℓ, and $\tilde{\mathbf{C}}_N$ being full rank if $\kappa' = \ell$.

3 Residual consistency test

3.1. *The author wrote "They (CCM24) propose $H_0 : \tilde{\boldsymbol{C}}_N = \boldsymbol{C}_N$. However this requires that $\tilde{\boldsymbol{C}}_N$ is full rank and that the rank-deficient $\hat{\boldsymbol{C}}_N$ consistently estimates it, which is a contradiction."(lines 163–164).*

It should be emphasized that CCM24 did not require $\kappa' = \ell$ for the RCT, which means that $\mathbf{C}_N$ can be reduced rank such that $\kappa' < \ell$. Therefore, under the null hypothesis, $\tilde{\mathbf{C}}_N$ can have reduced rank and will not contradict with the rank-deficient of $\hat{\mathbf{C}}_N$. Even if $\kappa' = \ell$, which leads to the full rank $\tilde{\mathbf{C}}_N$, there is still no contradiction as stated in 2.2.

3.2. *The author raised questions on RCT with concerns on the convergence of $\hat{\boldsymbol{C}}_N$: "CCM24 show that under the null the RCT is asymptotically χ^2 where convergence is governed by the speed at which $\hat{\boldsymbol{C}}_N \xrightarrow{p} \tilde{\boldsymbol{C}}_N$ which in turn is limited by n, which as discussed is typically small and difficult to accurately estimate. Thus we have no way of knowing how far the sample critical values depart from the asymptotic χ^2 levels."(lines 166–170)*

CCM24 showed that the distribution of the RCT statistic r^2 under the null hypothesis $H_0 : \tilde{\mathbf{C}}_N = \mathbf{C}_N$ is only asymptotically $\chi^2_{\kappa-m}$ distributed rather than exactly distributed. This means that the χ^2 cut-off value for the RCT may encounter finite sample approximation error, which would be reduced as n increased. Many of the statistical tests are based on the asymptotic approximation of the critical values. The bootstrap resampling method has been showed in many situations that is able to speed up the convergence and improve the finite sample performance.

3.3. The author commented the RCT under the alternative in lines 170–172 *"C23 (CCM24) also make some useful observations on power analysis, however under the alternative the RCT is no longer distributed χ^2 and the distribution would be samples specific so they can do little here except point to the challenges should anyone attempt it."*

We disagree on this view as statistical power of a test is generated from different distributions of the test statistic under the null and the alternative. CCM24 showed that under the alternative, the RCT statistic r^2 is asymptotically distributed as a sum of weighted independent χ_1^2, which deviates from the $\chi^2_{\kappa-m}$ under the H_0 and hence can produce power for the RCT.

3.4. *The author stated in lines 173–175 "One of the points raised in M22a was that the RCT does not provide a test of specification failures in equation [1] so even if the critical values are assumed known and the test score does not reject, that does not establish the regression coefficients are unbiased or consistent."*

The null distribution of the RCT proposed by AT99 was under the linear specification of the regression model. In the case of mis-specified regression specification, the power of the RCT would be affected by both the mis-specification and the deviation between $\mathbf{C}_N$ and $\tilde{\mathbf{C}}_N$. The mis-specified regression may lead to more power.

4 On the total least square (TLS)

When the fingerprints are subject to the measurement errors, Allen and Stott (2003) developed a version of the OF approach via the statistical error-in-variable regression.

The author stated in lines 64 that *"McKitrick* (2023) *showed that... TLS cannot be used to test the null because the estimator is undefined when* $\boldsymbol{\beta} = 0$."*In* McKitrick (2023), *the same statement was made by referring to* Fuller (1987, page 32).

This interpretation is not accurate. Fuller (1987, page 32) did never claim that the TLS estimator is undefined when $\boldsymbol{\beta} = 0$. His case where the TLS estimator may not exist happens only if $\mathbf{X}^T\mathbf{Y} = 0$ and $\mathbf{Y}^T\mathbf{Y} > \delta\mathbf{X}^T\mathbf{X}$, where δ is the ratio of the variance of the noise in $\mathbf{Y}$ to that of the noise in $\mathbf{X}$. For Fuller (1987)'s simple setting with independent data, it can be immediately shown that the joint probability of $\mathbf{X}^T\mathbf{Y} = 0$ and $\mathbf{Y}^T\mathbf{Y} > \delta\mathbf{X}^T\mathbf{X}$ approaches to zero as $\ell \to \infty$. Qiu et al. (2025), a technical report attached to Chen et al. (2025), has extended the results in Gleser (1981, Lemma 3.3) to further show that the non-existence of TLS estimator is also an event with probability approaching to zero for spatial dependent data; see (9) in Qiu et al. (2025) for details. As the non-existence of the TLS estimator is an asymptotically negligible event, it suffices to take $\hat{\boldsymbol{\beta}} = 0$ even if that happens.

We will provide a review on the OF approach with the TLS proposed by Allen and Stott (2003) in a fresh work (Chen et al. 2025).

5 Instrumental variable method

The author's use of the IV (also known as the two-stage least squares) deviates from the typically practice adopted in empirical studies, for instance that in a well-known empirical paper Acemoglu et al (2001, AJR01), written by three authors who were awarded 2024 Nobel Prize in economics, and that in the standard econometrics textbook Stock and Watson (2015, SW15). As agreed by the author, there are two key conditions for a valid IV to ensure the consistency. One is that the IV must be correlated with the covariates $\mathbf{X}$ (relevance condition). The other is that the IV must be uncorrelated with the error term $\mathbf{u}$ (exogeneity condition). However, the author has not taken sufficient effort to justify why the IVs chosen in the paper satisfy the two conditions. This makes the empirical results unreliable as violation of the two conditions may lead to misleading results.

5.1. <u>Relevance condition.</u> As noted in M25 (line 260), the relevance condition can be checked by a weak instrument test, which is usually in the form of reporting the first stage regression results (namely regressing each covariate $\mathbf{X}$ on the IVs) so that the influence of the IVs on $\mathbf{X}$ can be presented, for instance Tables 4–7 in AJR01 or Table 12.1 in SW15.

Although the author has noticed the importance of a weak instrument diagnostics, there is no such analysis reported in the work.

5.2. <u>Exogeneity condition.</u> The IVs chosen in M25 are geographic variables such as being in the northern hemisphere, latitude, and longitude. It is hard to justify that the climate noise would not be influenced by these geographic features used as the IVs by the author. For instance, the internal climate variability in the tropics should be different from that in the Antarctic, which is what $\mathbf{C}_N$ tries to account. Therefore, it is reasonable to believe that the geographic variables cannot be uncorrelated with the error term $\mathbf{u}$, which would violate the exogeneity condition, and invalidate them as the IVs. Furthermore, these variables are typically treated as the control variables instead of IVs even for economic empirical studies (AJR01).

Although M25 has provided the p-value of the Sargan over-identification test to provide a statistical evidence of the validity of the exogeneity condition, the author did not fully explain it from the point of views of the Earth Science. As SW15 (page 447) stated "*Assessing whether the instruments are exogenous necessarily requires making an expert judgment based on personal knowledge of the application*". Hence, such a careful reasoning with expert domain knowledge consideration is typically the core and the most important part of an empirical studies. However, this is not found in M25.

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

Data availability Not applicable.

Code Availability Not applicable.

Declarations

Conflict of interest There is no conflict of interest.

References

Acemoglu D, Johnson S, Robinson JA (2001) The colonial origins of comparative development: an empirical investigation. Am Econ Rev 91(5):1369–1401

Allen M, Stott P (2003) Estimating signal amplitudes in optimal fingerprinting, part i: theory. Clim Dyn 21:477–491

Allen M, Tett S (1999) Checking for model consistency in optimal fingerprinting. Clim Dyn 15:419–434

Chen H, Chen SX, Mu M (2024) A statistical review on the optimal fingerprinting approach in climate change studies. Clim Dyn 62:1439–1446

Chen H, Chen SX, Qiu J (2025) On the optimal fingerprinting approach with errors in fingerprints. Manuscript

Fuller W (1987) Measurement error models. Wiley Series in Probability and Statistics, Wiley

Gleser LJ (1981) Estimation in a multivariate "errors in variables'' regression model: large sample results. Ann Stat 9(1):24–44

McKitrick R (2022) Checking for model consistency in optimal fingerprinting: a comment. Clim Dyn 58:405–411

McKitrick R (2023) On the choice of TLS versus OLS in climate signal detection regression. Clim Dyn 60:359–374

McKitrick R (2025) Consistent climate fingerprinting. Manuscript

Qiu J, Chen H, Chen SX (2025) Total least squares for spatial dependent data. Technical Report

Stock JH, Watson MW (2015) Introduction to econometrics, 3rd edn. The Pearson Series in Economics, Pearson