Checking for model consistency in optimal fingerprinting: a comment

Ross McKitrick[1]

© The Author(s) 2021

Abstract

Allen and Tett (1999, herein AT99) introduced a Generalized Least Squares (GLS) regression methodology for decomposing patterns of climate change for attribution purposes and proposed the "Residual Consistency Test" (RCT) to check the GLS specification. Their methodology has been widely used and highly influential ever since, in part because subsequent authors have relied upon their claim that their GLS model satisfies the conditions of the Gauss-Markov (GM) Theorem, thereby yielding unbiased and efficient estimators. But AT99 stated the GM Theorem incorrectly, omitting a critical condition altogether, their GLS method cannot satisfy the GM conditions, and their variance estimator is inconsistent by construction. Additionally, they did not formally state the null hypothesis of the RCT nor identify which of the GM conditions it tests, nor did they prove its distribution and critical values, rendering it uninformative as a specification test. The continuing influence of AT99 two decades later means these issues should be corrected. I identify 6 conditions needing to be shown for the AT99 method to be valid.

Keywords Climate change · Optimal fingerprinting · Residual consistency test · Gauss-Markov Theorem · Regression misspecification

1 Introduction

In a highly influential paper published more than 20 years ago, Allen and Tett (1999, herein AT99) introduced a Generalized Least Squares (GLS) regression methodology for decomposing patterns of climate change across forcings for the purpose of making causal inferences about anthropogenic drivers of climate change, and proposed a "Residual Consistency Test" (RCT) to allow researchers to check the validity of the regression model. They claimed their method satisfies the conditions of the Gauss-Markov (GM) theorem, thereby yielding best (as in minimum variance) linear unbiased coefficient estimates, commonly denoted "BLUE", a claim that has been relied upon subsequently by other authors (e.g. Allen and Stott 2003; Hegerl and Zwiers 2011). The AT99 methodology was highlighted in the subsequent Intergovernmental Panel on Climate Change Third Assessment Report (Mitchell et al. 2001) and has since then been widely used to provide statistical support for attribution of

observed climatic changes to greenhouse gases and other anthropogenic factors.

AT99 made some errors in their summary of the GM conditions and made claims about the properties of their estimator that, at best, were never proven and in general are not true. Their GLS method does not satisfy the GM conditions and violates an important sufficient condition for unbiasedness. Additionally, AT99 provided no formal null hypothesis of the RCT nor did they prove its asymptotic distribution, making non-rejection against χ^2 critical values uninformative for the purpose of model specification testing.

2 Gauss-Markov conditions and optimal fingerprinting

To anchor the discussion in a common application framework, suppose a climate model is available that can generate simulated annual temperatures A_{itf} in locations (or grid cells) $i = 1, \ldots n$ over the time interval $t = 1, \ldots, T$ under the assumption of no forcings ($f = 0$), only natural forcings ($f = 1$), or only anthropogenic forcings ($f = 2$). Annual observed temperature anomalies A_{it}^o are also available in the same grid cells. The n-length vector $\boldsymbol{y}$ provides a measure of

✉ Ross McKitrick
ross.mckitrick@uoguelph.ca

[1] Department of Economics and Finance, University of Guelph, Guelph, ON N1G 2W1, Canada

the cell-specific changes (e.g. trends or differences in means) in A_{it}^o over the time period of interest. A matrix X is formed with g columns consisting of "signal vectors" which are created by running the model and generating A_{itf} for $f = 1, 2$; then computing changes in the same way as for y. All data are zero-centered so a constant term is not needed.[1]

A linear model regressing the observed spatial pattern of changes against the model-simulated signals is written

$$y = Xb + u \tag{1}$$

where u is the regression error and the coefficient vector b measures the scaling factors necessary for the signals to explain the observed climate patterns. A significant value of an element of b implies the associated signal is "detected".

Denote the (unobservable) covariance matrix of u as Ω, thus $E(uu^T) = \Omega$. Note Ω is a symmetric nonstochastic matrix with full rank. The GM Theorem states that if Ω is conditionally (on X) homoscedastic and has zero off-diagonal elements, meaning it can be represented as a scalar multiple of the $n \times n$ identity matrix I_n, which we write

$$E(uu^T|X) = \Omega = \sigma^2 I_n \tag{2}$$

where σ^2 is the variance of the errors u, and if the expectation of the errors conditional on X is zero:

$$E(u|X) = 0, \tag{3}$$

then the OLS estimator

$$\hat{b}_{OLS} = (X^TX)^{-1}X^Ty$$

is unbiased and has the minimum variance among all linear (in y) unbiased estimators (Davidson and MacKinnon 2004). In this way, an estimator of b that satisfies the GM conditions (2) and (3) can be said to be BLUE.

If GM condition (2) fails OLS coefficient estimates will still be unbiased but potentially inefficient, whereas if GM condition (3) fails OLS coefficients will be biased and inconsistent. If condition (2) fails the common remedies are feasible GLS (where 'feasibility' refers to use of a full rank estimator $\hat{\Omega}$ to provide weights for the regression model (1)) or use of White's (1980) heteroskedasticity-consistent (HC) variance estimator.

In response to concerns that detection of greenhouse gas signals would be thwarted by heteroskedasticity, AT99

proposed the following variant on feasible GLS. They call Ω the "climate noise" matrix and denote it C_N[2]:

$$E(uu^T) = C_N. \tag{4}$$

Rather than define it as the outer product of the error terms in (1) they assume it can be computed using the spatial covariances from the preindustrial (unforced) control run of a climate model. Denote the matrix root of the inverse of C_N as P, so $P^TP = C_N^{-1}$. While C_N is obtainable from a climate model, due to the limited degrees of freedom in climate models it is in practice rank deficient and its inverse does not exist. Suppose the rank of C_N is $K < n$. They propose an estimator $\hat{C}_N$ defined using a $K \times n$ matrix $P^{(K)}$ consisting of the first K eigenvectors of C_N weighted by their inverse singular values, which implies $P^{(K)}\hat{C}_N P^{(K)T} = I_K$ (the rank-K identity matrix) and $(P^{(K)T}P^{(K)})^+ = \hat{C}_N$ where $+$ denotes that it is the Moore–Penrose pseudo-inverse. AT99 assert that the regression

$$P^{(K)}y = P^{(K)}Xb + P^{(K)}u \tag{5}$$

yields an unbiased estimator $\tilde{b} = (X^TP^{(K)T}P^{(K)}X)^{-1}X^TP^{(K)T}P^{(K)}y$ which satisfies the GM conditions as long as $P^{(K)}$ exists, regardless of the value of K, because, by definition,

$$E(Puu^TP^T) = PC_NP^T = I_n. \tag{6}$$

Equation (6) corresponds to Eq. (3) in AT99.

The assertion that (6) implies $\tilde{b}$ is BLUE depends on 5 assumptions, not all of which AT99 state: (i) P is nonstochastic; (ii) it does not matter that $I_K \neq I_n$; (iii) there are no necessary conditions other than (6) for the GM Theorem to hold; (iv) the non-existence of C_N^{-1} has no implications for the properties of $\tilde{b}$ and (v) test statistics associated with $\tilde{b}$ do not depend on the assumption that $C_N = \Omega$. All five assumptions are incorrect.

(i) As was observed in AT99, C_N (and P) are functions of random natural forcings so they are random matrices. Ignore for the moment the distinction between P and $P^{(K)}$. AT99 set the randomness issue aside but the observation that the climate model generates "noisy" or random vectors is the motivation behind proposed refinements in Allen and Stott (2003) and elsewhere. Since u is random, if P is also random Eq. (6) is incorrect. Neither can it be the case that $C_N = \Omega$ since one is random and the other is not. Using a standard decomposition, $\tilde{b} = b + (X^TP^TPX)^{-1}X^TP^TPu$. To prove consist-

---

[1] This set-up assumes we are focused only on explaining the spatial variation of temperature change. It would be possible to pool spatial and temporal dimensions but this would raise additional issues related to the need to test time series misspecifications so we will omit this aspect.

[2] Note the N subscript stands for "noise" not the sample size, which is n.

ency of the AT99 optimal fingerprinting method with random P it needs to be shown (not merely assumed) that the second term on the right converges in distribution to $N(0, I_g)$, which requires proving the necessary conditions

[N1] $\quad plim\left(\left(X^T P^T P X\right)^{-1} X^T P^T P u\right) = 0_g$

and

[N2] $\quad plim\left(\left(X^T P^T P X\right)^{-1/2} X^T P^T P \Omega P^T \right.$

$PX\left(X^T P^T P X\right)^{-1/2}\left.\right) = I_g$

where the right-hand side of [N1] is a vector of g zeroes.[3] Here the probability limit is over $n \to \infty$ but note that the rank of $P^{(K)}$ will be constrained to K regardless of sample size, which would make proving [N1] and [N2] particularly challenging. AT99 side-stepped any discussion of these requirements by assuming $P^{(K)}$ is nonstochastic but in later literature when it was assumed to be random this issue was not revisited, so these conditions still need to be proven. This is true also of methods using regularized variance matrix estimators (see e.g. Hannart 2016) in which invertibility is achieved by using a weighted sum of C_N and the n-dimensional identity matrix.

(ii) If $P^{(K)}$ is used in Eq. (6) the resulting identity matrix will have rank less than n so GM condition (2) automatically fails and application of the AT99 method cannot yield a BLUE estimate of b. Even if P is nonstochastic, the pseudo-inverse $\left(P^{(K)T} P^{(K)}\right)^+$ cannot equal Ω regardless of sample size so $\hat{C}_N$ is a biased and inconsistent estimator of Ω.[4] Using an inconsistent estimator of Ω does not necessarily prevent obtaining consistent coefficient covariances. White (1980) famously showed that under certain assumptions, if the estimator $\hat{\Omega}$ consists of a diagonal matrix with the squared OLS residuals along the main diagonal it is inconsistent but $n^{-1} X^T \hat{\Omega} X$ is nonetheless a consistent estimator of $n^{-1} X^T \Omega X$ (Davidson and MacKinnon 2004 p. 198). This is the so-called HC method and is a popular alternative to feasible GLS in statistics and econometrics. To obtain a comparable result, AT99 needed to show[5]

[N3] $\quad plim\left(n^{-1} X^T \hat{C}_N X\right) = \lim_{n \to \infty} n^{-1} X^T E\left(u u^T\right) X.$

But even if this could be proven, and if its convergence properties could be shown to rival that of White's method or the variants that have been proposed since then, all it would show is that the AT99 optimal fingerprinting method has the same consistency property as White's method, which is much easier to use and does not introduce other forms of bias.

(iii) AT99 failed to state GM condition (3) and to test whether it holds or not. Failure of conditional invariance yields biased and inconsistent slope coefficients so it is critically important to detect and remedy potential violations. Adding P to the model does not imply conditional invariance of u or σ^2 to X is no longer needed, instead it adds the requirement of invariance to P. Hence AT99 should have listed the analogue to Eq. (3) which in the nonstochastic P case implies they needed to show

[N4] $\quad E(Pu|X, P) = 0$ (7)

Also Eq. (6) (Eq. (3) in AT99) as written was incorrect and should have been stated as

$E\left(P u u^T P^T | X, P\right) = I_K.$ (8)

AT99 and all who have followed in the fingerprinting literature omitted condition (3) (or (7)) in their discussion of the GM theorem. Equation (7) does not follow from Eq. (8) and needs to be tested separately. Violations of (7) lead to biased and inconsistent coefficient estimates and are typically more difficult to remedy than violations of (8). There is a voluminous econometrics and statistics literature pertaining to testing Eq. (2) (e.g. Davidson and MacKinnon 2004; Wooldridge 2019 etc.) which could be extended to the case of Eq. (7). Since this has never been done, there is no assurance that any applications of the AT99 method have yielded unbiased coefficient estimates.

(iv) AT99 stated (p. 423) "we do not actually require C_N^{-1} for $[\tilde{b}]$ to be BLUE" but this is inaccurate. Note that C_N is the estimator of the unobservable Ω, and since it is singular it must be approximated using $\hat{C}_N = \left(P^{(K)T} P^{(K)}\right)^+$ or through application of a regularization method. In feasible GLS the existence of the inverse of the error covariance matrix estimator is one of the sufficient conditions for proving $E(\tilde{b}) = b$ (Kmenta 1986 p. 615) so in the AT99 framework whenever recourse must be made to $\hat{C}_N$ it perforce cannot be shown that the coefficient estimates are unbiased. Estimating Ω using $\hat{C}_N$ solves a computational problem but creates a theoretical one. Even if C_N were invertible

---

[3] To see why [N2] holds, define $A \equiv X^T P^T P X$. Under the null $P^T P = \Omega^{-1}$. Then [N2] becomes $A^{-1/2} A \left(A^{-1/2}\right)^T = I$. Premultiplying by $\left(A^{-1/2}\right)^T$ confirms the result.

[4] AT99 (p. 423) stated their assumption that $\hat{C}_N$ is a "reliable" estimate of Ω but did not define the concept of reliability in terms of consistency or unbiasedness.

[5] A similar result would be needed for methods based on Ledoit and Wolf (2004) that replace C_N with a regularized form instead of a reduced rank form (reviewed in Hannart 2016).

however, that would only be the start of the matter since consistency of $\tilde{b}$ depends on specific assumptions about the structure of $\hat{\Omega}$ and that of the other variables in the regression model. Amemiya (1973) derived sufficient conditions for consistency and efficiency of a GLS estimator when the errors are known to follow a mixed autoregressive moving average process. But the proof depends on 5 assumptions about the numerical properties of X and u that might not hold in all applications. In the case of AT99 the reader is not told what assumptions about X and $\hat{C}_N$ are necessary for $\tilde{b}$ to converge in distribution to $N\left(b, (X^T\hat{C}_N^{-1}X)^{-1}\right)$, the authors merely assert (p. 422) that it does. Especially in light of how small K is in many applications, the non-existence of C_N^{-1} and the absence of proof of consistency means claims about the unbiasedness of optimal fingerprinting regression coefficients based on the AT99 method are conjectural. What is needed is a proof of the statement

[N5] $\sqrt{n}\left(\tilde{b} - E(\tilde{b})\right) \to^d N(0, V)$

where $V = plim\left(n^{-1}X^T\hat{C}_N^{-1}X\right)^{-1}$. If u is assumed to be Normal it would suffice to prove [N1] and [N2], otherwise a central limit theorem is required.

(v)   The distribution of a test statistic under the null cannot be conditional on the null being false. Reliance on the assumption that $\hat{C}_N \cong C_N = \Omega$ creates a risk of spurious inference in the AT99 framework since it assumes that the climate model is a perfect representation of the real climate. The climate model embeds the assumption that greenhouse gases have a significant effect on the climate along with other assumptions about the magnitude and effects of natural forcings. In a typical optimal fingerprinting application, the researcher seeks to compute the distribution of a test statistic under the null hypothesis that greenhouse gases have no effect on the climate. No statistic can be constructed in the AT99 framework that maintains this assumption. Use of preindustrial control runs to generate C_N, or combining data from different climate models, does not remedy this issue since all such models, even in their preindustrial era simulations, embed the assumption that elevated greenhouse gases (if present) would have a large effect relative to those of natural forcings.

3 Residual consistency test

The foregoing has established that AT99 erred in claiming their method satisfies the GM conditions thereby yielding unbiased and efficient estimates. Violations of GM

conditions are typically detected using standard specification tests as explained in any regression textbook. These have been missing in the optimal fingerprinting literature and instead practitioners have relied more or less exclusively on the RCT introduced in AT99. It entails obtaining an error covariance matrix C_N' where the $'$ denotes that it is from a different climate model than the one used to generate $P^{(K)}$ in Eq. (5), then using the quadratic form

$$\text{RCT} = u_1^T P'^T P' u_1 \tag{9}$$

where P' is computed from C_N' in the manner described above and u_1 is a set of residuals formed by plugging the estimated coefficients from Eq. (5) into Eq. (1). If RCT is less than the 5% critical value from the centered $\chi^2(K - g)$ distribution, where g is the number of columns of X, the null hypothesis is not rejected.

But what is the null hypothesis? A difficulty in interpreting the RCT is that AT99 never stated it formally. They wrote (p. 424):

> Our null-hypothesis, $\mathcal{H}_0$, is that the control simulation of climate variability is an adequate representation of variability in the real world in the truncated statespace which we are using for the analysis, i.e. the subspace defined by the first K EOFs of the control run does not include patterns which contain unrealistically low (or high) variance in the control simulation of climate variability. Because the effects of errors in observations are not represented in the climate model, $\mathcal{H}_0$ also encompasses the statement that observational error is negligible in the truncated state-space (on the spatio-temporal scales) used for detection. A test of $\mathcal{H}_0$, therefore, is also a test of the validity of this assumption. If we are unable to reject $\mathcal{H}_0$, then we have no explicit reason to distrust uncertainty estimates based on our analysis.

The first sentence is quite imprecise: what makes a control simulation "adequate"? It implies a bound on the bias matrix $D = \hat{C}_N - \Omega$ with no guidance as to how to quantify it or whether it approaches 0 as $n \to \infty$. The next two sentences imply that the null hypothesis also involves a condition on σ^2, namely that it is "negligible", with no further clarification. The final sentence is meaningless at best, and in fact is untrue.

The first difficulty of interpreting the RCT is that AT99 do not formally connect it to GM conditions (7) or (8), so readers do not know which one is rejected if the RCT value is large, or which if either can be retained if the RCT is small. And since (7) and (8) are independent conditions, a small value of any single test statistic would not suffice to accept both. Moreover there can be many reasons conditions (7) and (8) fail so hypothesis tests need to be constructed against

specific alternatives. Even if we assume the regression errors are Normally distributed, alternatives not ruled out by $\mathcal{H}_0$ (for example non-linearity of the true regression equation) can yield residuals that follow a non-central χ^2 distribution (see examples in Ramsey 1969) in which case use of the critical values from the central χ^2 tables would be invalid. There is simply no one universal specification test. That is why standard regression textbooks recommend applying a battery of tests to check the GM conditions. Imbers et al. (2014) recommended using a somewhat wider range of specification tests but the advice has yet to be generally applied. Failure to reject the RCT has been interpreted as comprehensive support for fingerprinting model specifications, an entirely unwarranted assumption.

Thus it is not possible to assign meaning to either rejection or non-rejection of the RCT. The customary way to present a new test statistic is to state the associated null and alternative hypotheses mathematically, list all the assumptions necessary to derive its distribution, prove its distribution either exactly or asymptotically, and present evidence of its size and power either theoretically or based on simulations. A typical example (with some applicability to the RCT) is in Newey (1985). None of these steps were followed in AT99 nor in subsequent treatments of the issue. Users have thus been interpreting and relying on the RCT with no guidance as to its power, distribution or true critical values.

As a specific example of the problem, suppose that GM condition (7) does not hold and the coefficient estimates from model (5) are biased and inconsistent. This could happen if, for example, the matrix X omits one or more variables on which y significantly depends and which are partially correlated with one of the columns of X. Suppose also that the selected C_N' matrix is diagonal with every entry on the diagonal, denoted c_{ii}', varying in direct proportion to the corresponding squared element of u_1 but is always a little greater in magnitude. Then

$$\text{RCT} = \sum_{i=1}^{K} \frac{u_{1i}^2}{c_{ii}'} < K.$$

By construction C_N' satisfies $\mathcal{H}_0$ in the loose sense of being "adequate". Since the 5% critical values for $\chi^2(K-2)$ always exceed K, in the $g = 2$ case we will have [5% critical value] $> K >$ RCT hence the RCT would never reject. Other non-diagonal versions of C_N' could be constructed which are likewise "adequate" yet as long as the elements of $P'^T P' u_1$ are each a bit smaller than the corresponding values $1/u_{1i}$, RCT will not exceed K even while the GM conditions fail to hold and the estimator $\tilde{b}$ is biased and inconsistent.

In sum, it is unclear what the RCT tests. This needs to be rectified by stating $\mathcal{H}_0$ formally, with specific reference to conditions (7) and (8), providing a proper derivation of the asymptotic distribution of the RCT under the null including

a list of all the assumptions of the data generating process, specifying the assumed limiting properties of D, and presenting evidence of the RCT's power characteristics. More specifically, to validate the use that has been made of the RCT to date requires proof of the following statement:

[N6]   [7]&[8] $\Leftrightarrow u_1^T P'^T P' u_1 < \alpha_{.05}$

where $\alpha_{.05}$ is the 5% critical value from the central $\chi^2(K-m)$ distribution.

4 Conclusion

The methodology of AT99 is seminal to the optimal fingerprinting literature (see reviews in IADAG 2005; Hegerl and Zwiers 2011; Bindoff et al. 2013 and Hannart 2016). A very large number of studies has appeared over the past two decades applying the AT99 method, or extensions thereof, to global (e.g. Hegerl et al. 1997, 2000; Tett et al. 1999; Allen and Stott 2003; Allen et al. 2006) and regional (e.g. Gillett et al. 2008) temperature indicators, as well as to other climate variables like atmospheric moisture (Santer et al. 2007), snow cover (Najafi et al. 2016), forest fires (Gillett et al. 2004) and others (see Hegerl et al. 2006). The literature resting on AT99 figures prominently in the increasing confidence with which the Intergovernmental Panel on Climate Change (IPCC 2013) and others have attributed most modern climate change to anthropogenic influences, chiefly greenhouse gases, especially since 1950. Inferences based at least in part on the AT99 methodology have driven some of the most consequential policy decisions in modern times.

Confidence in the results of the AT99 methodology rests on their claim that it satisfies the GM conditions and thereby yields unbiased and efficient coefficients, and that as long as the RCT yields a small score relative to the 5% critical values of the central χ^2 distribution the variance estimator is valid. These claims are untrue. AT99 omitted a key necessary condition for the GM Theorem and misstated the others. The way the method is commonly used (with $K < n$) the GM conditions automatically fail. The omitted GM condition has never been mentioned or tested in the optimal fingerprinting literature and hence there is no assurance past fingerprinting results are unbiased or consistent. RCT values provide no assurance on this point. AT99 presented only a vaguely-worded null hypothesis and did not relate it to the GM conditions, nor did they state the necessary assumptions to derive its distribution.

Despite the long passage of time since its publication, owing to the influence of the AT99 methodology this situation should be rectified. The purpose of this comment is to encourage the authors of AT99 to do so by proving the six

necessary conditions [N1]—[N6], or alternatively to advise readers of the implications if it is not possible to do so.

Declarations

References

Allen MR, Stott PA (2003) Estimating signal amplitudes in optimal finger-printing, part I: theory. Clim Dyn 21:477–491. https://doi.org/10.1007/s00382-003-0313-9

Allen MR, Tett SFB (1999) Checking for model consistency in optimal fingerprinting. Clim Dyn 15:419–434. https://doi.org/10.1007/s003820050291

Allen MR, Gillett NP, Kettleborough JA, Hegerl G, Schnur R, Stott PA, Boer G, Covey C, Delworth TL, Jones GS, Mitchell JFB, Barnett TP (2006) Quantifying anthropogenic influence on recent near-surface temperature change. Surv Geophys 27(5):491–544. https://doi.org/10.1007/s10712-006-9011-6

Amemiya T (1973) Generalized least squares with an estimated auto-covariance matrix. Econometrica 41(4):723–732. https://doi.org/10.2307/1914092

Bindoff NL, Stott PA, AchutaRao KM, Allen MR, Gillett N, Gutzler D, Hansingo K, Hegerl G, Hu Y, Jain S, Mokhov II, Overland J, Perlwitz J, Sebbari R, Zhang X (2013) Detection and attribution of climate change: from global to regional. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Climate change 2013: the physical science basis contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge

Davidson R, MacKinnon J (2004) Econometric theory and methods. OUP, New York

Gillett NPAJ, Weaver FWZ, Flannigan MD (2004) Detecting the effect of climate change on canadian forest fires. Geophys Res Lett. https://doi.org/10.1029/2004GL020876

Gillett NP, Stone DA, Stott PA, Nozawa T, Alexey Yu, Karpechko GC, Hegerl MFW, Philip PDJ (2008) Attribution of polar warming to human influence. Nat Geosci 1:750–754. https://doi.org/10.1038/ngeo338

Hannart A (2016) Integrated optimal fingerprinting: method description and illustration. J Clim. https://doi.org/10.1175/JCLI-D-14-00124.1

Hegerl G, Zwiers F (2011) Use of models in detection and attribution of climate change. Wires Clim Change. https://doi.org/10.1002/wcc.121

Hegerl GC, Hasselmann K, Cubasch U, Mitchell JFB, Roeckner E, Voss R, Waszkewitz J (1997) Multi-fingerprint detection and attribution analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change. Clim Dyn 13(8):613–634. https://doi.org/10.1007/s003820050186

Hegerl GC, Stott PA, Allen MR, Mitchell JFB, Tett SFB, Cubasch U (2000) Optimal detection and attribution of climate change: sensitivity of results to climate model differences. Clim Dyn 16(10):737–754. https://doi.org/10.1007/s003820000071

Hegerl GC, Karl TR, Allen M, Bindoff NL, Gillett N, Karoly D, Zhang X, Zwiers F (2006) Climate change detection and attribution: beyond mean temperature signals. J Clim. https://doi.org/10.1175/JCLI3900.1

Imbers J, Ana L, Chris H, Myles A (2014) Sensitivity of climate change detection and attribution to the characterization of internal climate variability. J Clim. https://doi.org/10.1175/JCLI-D-12-00622.1

International Ad Hoc Detection and Attribution Group (IADAG) (2005) Detecting and attributing external influences on the climate system: a review of recent advances. J Clim 18:1291–1314

Intergovernmental Panel on Climate Change (IPCC) (2013) Climate change 2013: the physical science basis. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Contribution of Working Group I to the fifth assessment report of the Intergovernmental panel on climate change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA

Kmenta J (1986) Elements of econometrics, 2nd edn. MacMillan, New York

Ledoit O, Wolf M (2004) A well-conditioned estimator for large dimensional covariance matrices. J Multivar Anal 88:365–411. https://doi.org/10.1016/S0047-259X(03)00096-4

Mitchell JFB, Karoly DJ, Hegerl GC, Zwiers FW, Allen MR, Marengo J, Barros, V, Berliner M, Boer G, Crowley T, Folland C, Free M, Gillett N, Groisman P, Haigh J, Hasselmann K, Jones P, Kandlikar M, Kharin V, Kheshgi H, Knutson T, MacCracken M, Mann M, North G, Risbey J, Robock A, Santer B, Schnur R, Schönwiese C, Sexton D, Stott P, Tett S, Vinnikov K, Wigley T (2001) Detection of climate change and attribution of causes. In: Intergovernmental Panel on Climate Change Third Assessment Report Climate Change 2001: The Scientific Basis, Cambridge, CUP. https://www.ipcc.ch/report/ar3/wg1/

Najafi MR, Zwiers FW, Gillett NP (2016) Attribution of the spring snow cover extent decline in the northern hemisphere, Eurasia and North America to anthropogenic influence. Clim Change 136:571–586. https://doi.org/10.1007/s10584-016-1632-2

Newey W (1985) Generalized method of moments specification testing. J Econom 29(3):229–256. https://doi.org/10.1016/0304-4076(85)90154-X

Ramsey JB (1969) Tests for specification errors in classical linear least-squares regression analysis. J R Stat Soc 31(2):350–371. https://doi.org/10.1111/j.2517-6161.1969.tb00796.x

Santer BD, Mears C, Wentz FJ, Taylor KE, Gleckler PJ, Wigley TML, Barnett TP, Boyle JS, Brüggemann W, Gillett NP, Klein SA, Meehl GA, Nozawa T, Pierce DW, Stott PA, Washington WM, Wehner MF (2007) Identification of human-induced changes in atmospheric moisture content. Proc Natl Acad Sci 104(39):15248–15253. https://doi.org/10.1073/pnas.0702872104

Tett SFB, Stott PA, Allen MR, Ingram WJ, Mitchell JFB (1999) Causes of twentieth century temperature change near the earth's surface. Nature 399:569–572. https://doi.org/10.1038/21164

White H (1980) A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48(4):817–838

Wooldridge J (2019) Introductory econometrics: a modern approach, 7th edn. Cengage, Boston