



HAL
open science

Optimal fingerprinting under multiple sources of uncertainty

Alexis Hannart, Aurélien Ribes, Philippe Naveau

► **To cite this version:**

Alexis Hannart, Aurélien Ribes, Philippe Naveau. Optimal fingerprinting under multiple sources of uncertainty. *Geophysical Research Letters*, 2014, 41 (4), pp.1261-1268. 10.1002/2013GL058653 . hal-03211770

HAL Id: hal-03211770

<https://hal.science/hal-03211770>

Submitted on 30 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH LETTER

10.1002/2013GL058653

Key Points:

- Introducing an inference procedure for EIV OF models
- Clarifying the situation with respect to OF models and procedures
- Enhancing the benefit of the procedure

Correspondence to:

A. Hannart,
alexis.hannart@cima.fcen.uba.ar

Citation:

Hannart, A., A. Ribes, and P. Naveau (2014), Optimal fingerprinting under multiple sources of uncertainty, *Geophys. Res. Lett.*, *41*, 1261–1268, doi:10.1002/2013GL058653.

Received 14 NOV 2013

Accepted 6 JAN 2014

Accepted article online 8 JAN 2014

Published online 19 FEB 2014

Optimal fingerprinting under multiple sources of uncertainty

Alexis Hannart¹, Aurélien Ribes², and Philippe Naveau³

¹IFAECE, CNRS/CONICET/UBA, Buenos Aires, Argentina, ²CNRM/GAME, Météo France/CNRS, Toulouse, France, ³LSCE, IPSL, CNRS/CEA, Saclay, France

Abstract Detection and attribution studies routinely use linear regression methods referred to as optimal fingerprinting. Within the latter methodological paradigm, it is usually recognized that multiple sources of uncertainty affect both the observations and the simulated climate responses used as regressors. These include for instance internal variability, climate model error, or observational error. When all errors share the same covariance, the statistical inference is usually performed with the so-called total least squares procedure, but to date no inference procedure is readily available in the climate literature to treat the general case where this assumption does not hold. Here we address this deficiency. After a brief outlook on the error-in-variable models literature, we describe an inference procedure based on likelihood maximization, inspired by a recent article dealing with a similar situation in geodesy. We evaluate the performance of our approach via an idealized test bed. We find the procedure to outperform existing procedures when the latter wrongly neglect some sources of uncertainty.

1. Introduction

An important goal of climate research is to determine the causes of past global warming in general and the responsibility of human activity in particular; this question has thus emerged as a research topic known as detection and attribution (D & A). Over the past two decades, the results produced in this area have been instrumental in delivering far-reaching statements and in raising awareness of anthropogenic climate change (Intergovernmental Panel on Climate Change's Fourth Assessment Report: *Hegerl et al.* [2007]). From a methodological standpoint, D & A studies routinely use linear regression methods referred to as optimal fingerprinting, whereby an observed climate change y is regarded as a linear combination of externally forced signals x added to an internal climate variability term ϵ which is treated as a noise [*Hegerl and Zwiers*, 2011]. The regressors x —or “fingerprints”—consist here of spatial, temporal, or space-time patterns of response to external forcings as anticipated by one or several climate models. Practically, the results of the inference and in particular the magnitude of the uncertainty range on the regression coefficients determine whether these fingerprints are present in the observations and whether or not the observed change is attributable to a given cause.

Over the years, this approach has been formulated using a suite of multivariate linear regression models that were progressively refined, involving more and more general formulations together with inference procedures of increasing complexity. For instance, the standard general linear model and its usual inference treatment—i.e., ordinary least square (OLS)—were first introduced explicitly by *Hegerl et al.* [1996] and were then refined by *Allen and Tett* [1999]. In *Allen and Stott* [2003], an important further advance was introduced by recognizing that the forcing responses x are actually not known to certainty: they are indeed tainted by a remanent noise originating from the internal variability which is also present in climate model simulations, this, even after averaging over the simulation ensemble and because the size of the latter is usually small. Technically, the inference scheme described in *Allen and Stott* [2003] takes advantage of the fact that the error on the observed response y and on the regressors x are both assumed to be driven by internal variability (i.e., actual and simulated, respectively) and consequently share a common covariance structure. A joint “prewhitening” of y and x is thus possible, allowing next for a total least squares (TLS) inference [*Van Huffel and Vanderwaal*, 1994].

This innovation was taken one step further by *Huntingford et al.* [2006; hereinafter HSAL06] who introduced the more general hypothesis that the forcing responses x are not only tainted by remanent internal variability but also by some error associated to the climate model. From a model formulation standpoint, this contribution was an important milestone, foremost not only because it incorporated model uncertainty

into D & A studies for the first time but also because it set the ground for taking into account several other sources of error (e.g., observational error) that may affect D & A. Nonetheless, HSAL06 presents two shortcomings from a model inference standpoint. First, its more general assumptions imply that the errors on y and x no longer have the same covariance structure; thus, the TLS algorithm cannot be applied: a different inference scheme is needed. While HSAL06 recognized this need, it came short of any detailed description of the inference technique used and instead merely refer to the procedure described in *Nounou et al.* [2002] in the context of chemical analysis. Thus, the paper is not self-contained in this respect, and the method is not very easily accessible to D & A practitioners. Second, the procedure developed by *Nounou et al.* [2002] appears to treat a statistical model which—to the best of our understanding—is different from the one formulated by HSAL06 and is neither applicable nor implementable in the situation at stake (see section 2).

Because of the general relevance of HSAL06 model formulation in D & A, it is fundamental in our view to insure that a sound and easily implementable statistical inference method is available for this model. The main objective of this paper is to lay out such an inference method. Section 2 recalls the model, introduces a few notations, and then proposes a brief outlook on the statistical literature available for this type of regression models enhancing the peculiarity of the model studied. Section 3 lays out our inference scheme, and section 4 illustrates it and evaluates its performance on simulated series. Section 5 discusses results and concludes.

2. Defining and Discussing Error-in-Variable Models

We propose to study the following class of error-in-variables (EIV) models:

$$\begin{cases} y = \mathbf{x}^* \beta + \epsilon \\ \mathbf{x} = \mathbf{x}^* + \mathbf{v} \end{cases} \quad (1)$$

where the observation vector y of length n and the simulation ensemble average response matrix \mathbf{x} of size $n \times p$ are known. In contrast, the matrix of actual regressors \mathbf{x}^* of size $n \times p$ is not known with certainty but the known matrix \mathbf{x} is assumed to be a noised version of it. The vector of regression coefficient β of length p is unknown and represents the main focus of the inference, as in any classical linear regression. Concerning the noise dependence structure, we assume that the observation noise vector ϵ and the column vectors v_1, \dots, v_p are all independent, Gaussian distributed with zero-mean and covariance matrices $\Sigma, \Omega_1, \dots, \Omega_p$, respectively.

It should be underlined that in both the D & A and the EIV literature, the model structural equation (1) is often formulated under the reduced form $y = (\mathbf{x} - \mathbf{v})\beta + \epsilon$ in which the latent variable \mathbf{x}^* was eliminated by using $\mathbf{x}^* = \mathbf{x} - \mathbf{v}$. However, it can be argued that this form is somewhat ambiguous (see Appendix A); hence, we prefer the formulation of equation (1).

It should also be noted that the regression models mentioned in section 1 are all special cases of the general formulation of equation (1). Denoting Σ the covariance associated to internal variability, then *Allen and Tett* [1999] correspond to $\Omega_i = 0$, *Allen and Stott* [2003] to $\Omega_i = \Sigma/n_r$, with n_r the size of the ensemble, and HSAL06 to $\Omega_i = \Sigma/n_r + \Lambda$ where Λ is the covariance associated to the climate model error. In practice, these matrices are usually estimated beforehand from ensemble runs [*Allen and Tett*, 1999; *Ribes et al.*, 2012]. While this estimation step has been shown to critically influence the end result, it is in general handled preliminarily and rather independently from the regression inference which is at stake here. It is thus omitted for concision in the present paper where Σ, Ω_i are assumed to be known.

An extensive body of literature, dating back as far as *Adcock* [1878], is available on EIV linear regression models, presenting a variety of model formulations for applications in different fields (e.g., see the book of *Fuller* [1987] or the review article of *Gillard* [2010]). Reviewing it is beyond our scope, instead we highlight a cleaving aspect that makes the D & A setup a somewhat peculiar situation with respect to most EIV models found in literature. The main peculiarity of our model resides in the assumed structure of noise dependence (see Appendix B). In a nutshell, our setup assumes dependence between lines and independence between columns (i.e., $n \times n$ covariances), whereas most existing EIV models consider the symmetric case of independence between lines and dependence between columns (i.e., $p \times p$ covariance). An example of contribution falling into the latter category is *Nounou et al.* [2002], wherein all covariance matrices are $p \times p$. Note that an overlap does exist between these two contrasted categories, but it is reduced to the case where all the noise components are independent. Such a “white” setup is a commonplace one, referred to as the Deming

regression when $p = 1$ [Deming, 1943], that can be handled with the TLS algorithm when noise variances are known and is the focus of many articles in the more problematic case where variances are known only partially Gillard [2010]. While the model of Allen and Stott [2003] is not a “white” setup strictly speaking, it can easily be connected to this case by “whitening” the data. In contrast, the general case of HSAL06 at stake here is neither “whitenable” nor does it have any overlap with the traditional EIV setup of line independence and column dependence. The method of Nounou *et al.* [2002]—as most EIV methods found in literature—is hence not able to treat the inference problem at stake here and is simply not implementable in the first place because of the mismatch in matrix sizes.

As a notable exception within the above-described EIV landscape, Schaffrin and Wieser [2008; hereinafter SW08] treated in the context of an application in geodesy the more general case where both lines and columns are simultaneously dependent. This contribution assumes that the dependence within the noise matrix \mathbf{v} has a block structure described by the Kronecker product $\mathbf{Q} \otimes \mathbf{\Omega}$ where \mathbf{Q} (respectively $\mathbf{\Omega}$) is a fixed known $p \times p$ (respectively $n \times n$) covariance matrix describing the dependence among columns (respectively lines), which means that $E(v_{ij}v_{kl}) = \mathbf{Q}_{ji} \times \mathbf{\Omega}_{ik}$ for any (i, j, k, l) . Then, it proposes an inference procedure based on minimizing a weighted squared error criterion, which coincides with the negative loglikelihood of the model under a Gaussian setting. Minimization is performed by mean of an iterative algorithm coined “weighted total least squares” (WTLS), obtained by (i) deriving the first-order nonlinear conditions in β and \mathbf{x}^* (or, equivalently in the paper, in β and the residuals, under the model structural constraint), (ii) expressing their solution as a function of itself (i.e., “fixed point” formulation), and (iii) iterating the fixed point equations to convergence. Results show fast convergence toward the exact solution (when known).

By choosing $\mathbf{Q} = \mathbf{I}_p$, the framework of SW08 does correspond to our model, provided all the covariance matrices $\mathbf{\Omega}_i$ thereof are equal. Therefore, running the above algorithm would be a valid solution to solve our model in this case. However, this assumption may be restrictive, and on the other hand, the algorithm does not produce confidence intervals but merely a point estimate, an important shortcoming in a D & A context. These two limitations motivates us to design a new procedure, which is adapted from the one of SW08 and follows a similar approach.

3. Inference of the EIV Model

We wish to build an inference procedure for β , i.e., to derive a point estimate and confidence intervals that depend only of the two known variables \mathbf{x} and y . For this purpose, we follow a maximum likelihood estimation (MLE) approach as is commonplace in EIV regression (e.g., SW08) and in optimal fingerprinting [e.g., Allen and Stott, 2003]. The loglikelihood function of model (1) is

$$\ell(\beta, \mathbf{x}^* | y, \mathbf{x}) = -\frac{1}{2}(y - \mathbf{x}^* \beta)' \mathbf{\Sigma}^{-1} (y - \mathbf{x}^* \beta) - \frac{1}{2} \sum_{i=1}^p (x_i - x_i^*)' \mathbf{\Omega}_i^{-1} (x_i - x_i^*). \quad (2)$$

We wish to maximize (2) in (β, \mathbf{x}^*) . The first-order conditions are

$$\frac{\partial \ell}{\partial \beta} = -\mathbf{x}^* \mathbf{\Sigma}^{-1} \mathbf{x}^* \beta + \mathbf{x}^* \mathbf{\Sigma}^{-1} y = 0 \quad (3)$$

$$\frac{\partial \ell}{\partial x_i^*} = -(\mathbf{\Omega}_i^{-1} + \beta_i^2 \mathbf{\Sigma}^{-1}) x_i^* + (\beta_i \mathbf{\Sigma}^{-1} \bar{y}_i + \mathbf{\Omega}_i^{-1} x_i) = 0 \quad \text{for } i = 1, \dots, p \quad (4)$$

where x_i and x_i^* are the i th column vector of \mathbf{x} and \mathbf{x}^* , respectively, and $\bar{y}_i = y - \beta_1 x_1^* - \dots - \beta_{i-1} x_{i-1}^* - \beta_{i+1} x_{i+1}^* - \dots - \beta_p x_p^*$. The system of $p + 1$ equations corresponding to equations (3) and (4) must now be solved in (β, \mathbf{x}^*) , but the solution cannot be obtained under a closed form because of the nonlinearity of the system implied by the product terms $\beta_i x_i^*$. Following SW08, we can thus reformulate expressions (3) and (4) under a fixed point equation to be solved in (β, \mathbf{x}^*) :

$$\beta = (\mathbf{x}^* \mathbf{\Sigma}^{-1} \mathbf{x}^*)^{-1} (\mathbf{x}^* \mathbf{\Sigma}^{-1} y) \quad (5)$$

$$x_i^* = (\mathbf{\Omega}_i^{-1} + \beta_i^2 \mathbf{\Sigma}^{-1})^{-1} (\beta_i \mathbf{\Sigma}^{-1} \bar{y}_i + \mathbf{\Omega}_i^{-1} x_i) \quad \text{for } i = 1, \dots, p \quad (6)$$

Next, we take advantage of equations (5) and (6) to implement the following iterative procedure:

$$\left\{ \begin{array}{l} \bullet \text{ initialization: } \mathbf{x}^{*(0)} = \mathbf{x} \text{ and } \beta^{(0)} = (\mathbf{x}'\Sigma^{-1}\mathbf{x})^{-1}(\mathbf{x}'\Sigma^{-1}\mathbf{y}). \\ \bullet \text{ iterationstep1: } x_i^{*(t+1)} = (\Omega_i^{-1} + \beta_i^{(t)2}\Sigma^{-1})^{-1}(\beta_i^{(t)}\Sigma^{-1}\bar{y}_i^{(t)} + \Omega_i^{-1}x_i) \text{ for each } i. \\ \bullet \text{ iterationstep2: } \beta^{(t+1)} = (\mathbf{x}^{*(t+1)'}\Sigma^{-1}\mathbf{x}^{*(t+1)})^{-1}(\mathbf{x}^{*(t+1)'}\Sigma^{-1}\mathbf{y}). \\ \bullet \text{ stopping: repeat iterations until } \|\beta^{(t+1)} - \beta^{(t)}\|/\|\beta^{(t)}\| < \delta_0. \end{array} \right. \quad (7)$$

where $\delta_0 > 0$ is the requested precision level. Scheme (7) yields a value $(\hat{\beta}, \hat{\mathbf{x}}^*)$ which is by construction a fixed point of the function $\phi(\beta, \mathbf{x}^*)$ defined by equations (5) and (6) and thereby satisfies to the first-order condition.

Defining the *partial* MLE as the value that maximizes the likelihood in one given variable when all others are hold fixed, the right-hand terms of equations (5) and (6) also give the expressions of the so-called partial MLEs $\hat{\beta}_{\mathbf{x}^*}$ and $\hat{\mathbf{x}}_{\beta}^*$. Scheme (7) may thus also be interpreted as a *partial iterative maximization*. Such a procedure is not uncommon in statistics [e.g., Lauritzen, 1996], and it is similar to the widely used expectation-maximization (EM) procedure [Dempster et al., 1977]. Under the EM procedure, $\hat{\beta}$ would indeed also be obtained by iterative maximization in β (“M” step) but this maximization would be performed on a likelihood function obtained by averaging over \mathbf{x}^* (“E” step), whereas in (7) it is obtained by maximizing over \mathbf{x}^* (step 1). The partial iterative maximization approach is thus sometimes referred to as “hard EM” because of this similarity. Finally, scheme (7) may also be paralleled with the Gibbs algorithm widely used in Bayesian statistics Geman and Geman [1984], which similarly consists of a “partial iterative” simulation of β and \mathbf{x}^* conditionally on one another. Rephrased in this context, scheme (7) basically retains at each step the most likely value of the conditional distribution, as opposed to a random realization thereof.

With a point estimate $\hat{\beta}$ in hand, we now turn to confidence intervals. A substantial benefit of working with a MLE approach is that commonplace and widely applicable results are available to approximate confidence intervals. In particular, if we define the profile likelihood for each individual coefficient β_i as

$$\ell_i(\beta_i | y, \mathbf{x}) = \max_{(\beta_{-i}, \mathbf{x}^*)} \ell(\beta, \mathbf{x}^* | y, \mathbf{x})$$

where $\beta_{-i} = (\beta_1, \dots, \beta_{i-1}, \beta_{i+1}, \dots, \beta_p)$, then the corresponding profile likelihood ratio test statistic is asymptotically distributed according to the χ^2 distribution with 1° of freedom. This means that an approximate $(1 - \alpha)\%$ confidence region for β_i is the set of values: $\{\beta_i \in \mathbb{R} \mid 2(\ell_i(\hat{\beta}_i) - \ell_i(\beta_i)) \leq c_{1-\alpha}\}$ where $c_{1-\alpha}$ is the $(1 - \alpha)$ th quantile of the χ^2 distribution with 1° of freedom (see Venzon and Moolgavkar [1988] for more details). The confidence interval is thus obtained by solving for $\ell_i(\beta_i) = \ell_i(\hat{\beta}_i) - c_{1-\alpha}/2$ in β_i , which is done here using the Newton-Raphson algorithm [Press et al., 2007]. The profile likelihood $\ell_i(\beta_i | y, \mathbf{x})$ is derived every time needed by application of the same partial iterative maximization procedure, but with β_i hold constant. More specifically, step 1 is run identically, but step 2 is replaced by

$$\beta_{-i}^{(t+1)} = (\mathbf{x}_{-i}^{*(t+1)'}\Sigma^{-1}\mathbf{x}_{-i}^{*(t+1)})^{-1}\mathbf{x}_{-i}^{*(t+1)'}\Sigma^{-1}(\mathbf{y} - \beta_i\mathbf{x}_i^{*(t+1)})$$

where $\mathbf{x}_{-i}^* = (x_1^*, \dots, x_{i-1}^*, x_{i+1}^*, \dots, x_p^*)$.

Finally, one important element in D & A studies consists in deriving a goodness-of-fit metric after the inference was performed, in order to check the consistency of the assumed model. A classic approach to such a consistency test within the proposed EIV linear regression model would be for instance to use the weighted sum of squared residuals obtained from the maximized loglikelihood $\ell(\hat{\beta}, \hat{\mathbf{x}}^* | y, \mathbf{x})$. This quantity conveniently follows a χ^2 distribution under the present assumption of known covariances; hence, a critical value could straightforwardly be obtained for application of the test—in theory. However, in practice, covariances are not known but estimated which has been shown to strongly influence the test’s distribution. How to deal with the latter remains an open question at present (see Ribes et al. [2012] for a detailed explanation) which is beyond the scope of this letter.

4. Illustration and Simulation Results

This section illustrates and evaluates the performance of our WTLS procedure by applying it to simulated values of \mathbf{x} and y . The use of simulated rather than real data aims at verifying that our inference procedure performs correctly, a goal which requires the actual values of β and \mathbf{x}^* to be known. Our idealized

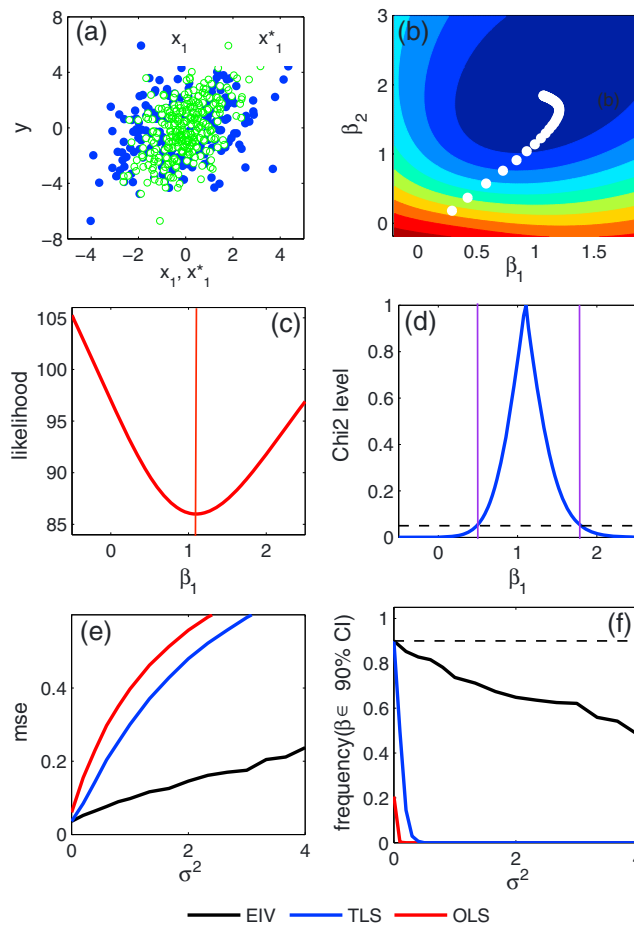


Figure 1. (a–d) Illustration of the inference procedure for a simulation $n = 275$, $p = 2$, and $\sigma = \sigma_0$ and (e and f) performance results. Data scatterplot (x_1, y) (blue dots) and (x_1^*, y) (green circles) shown in Figure 1a. Contour plot of the negative profile loglikelihood $-\ell_i(\beta)$ and trajectory of $\beta^{(t)}$ showing convergence to the minimum shown in Figure 1b. Plot of the negative profile loglikelihood $-\ell_i(\beta_1)$ shown in Figure 1c. Plot of the χ^2 probability level and confidence interval shown in Figure 1d. Average mean squared error of the estimator obtained with our procedure (EIV, black line), TLS (blue line), and OLS (red line) shown in Figure 1e. Frequency of the actual value of β falling within the 90% confidence interval for our procedure, TLS, and OLS shown in Figure 1f.

data simulation assumptions are based on Ribes *et al.* [2012]. They aim at replicating realistically a global data set of twentieth century temperature, as obtained after dimension reduction, and is described in detail in Appendix C together with assumptions regarding Σ . Regarding Ω_i , we assumed as in HSAL06 that $\Omega_i = \Sigma/n_i + \Lambda$ for all i with $n_i = 5$ and $\Lambda = \sigma^2 \mathbf{I}$ representing the climate model error covariance—as a default choice in the absence of a sound estimate of the covariance associated to model error. We ran the inference procedure on ten samples of size $N = 1000$ corresponding to ten values of σ^2 ranging from 0 to $4 \times \sigma_0^2$, where we use $\sigma_0^2 = \text{Tr}(\Sigma)/n$ as a reference error amplitude.

On average, the scheme (7) converged in 24 iterations which took 0.05 s using a desktop computer and the algorithmic optimization described in Appendix D. Computing confidence intervals took an extra 0.1 s, which overall makes the procedure applicable identically for higher values of both p and n (within the same order of magnitude). Figures 1a–1d show an illustration for a randomly selected simulation under $\sigma = \sigma_0$. The performance of the MLE $\hat{\beta}$ was evaluated based on average mean squared error $1/N \sum_{k=1}^N \|\hat{\beta}^{(k)} - \beta^{(k)}\|^2 / \|\beta^{(k)}\|^2$, where $k = 1, \dots, N$ denotes the simulation index. The accuracy of the confidence interval $I_{0.9}$ was evaluated by comparing the empirical frequency $1/N \sum_{k=1}^N \mathbf{1}_{\beta_i^{(k)} \in I_{0.9}}$ of each actual scalar coefficient $\beta_i^{(k)}$ falling into the interval, to its theoretical value 0.9. For the sake of comparison to existing procedures, both performance metrics were also systematically derived on each simulation by conducting an OLS and a TLS inference. In the context of this test bed, it is clear that the latter two procedures do make an incorrect assumption on Ω_i —i.e., $\Omega_i = \mathbf{0}$ and $\Omega_i = \Sigma/n_i$, respectively—yet they provide a useful benchmark.

Our simulation results are summarized in Figures 1e and 1f. They show that the procedure is able to estimate the true value of β with a moderate dispersion, no bias (not shown) and a low mean square error (MSE) overall. As expected, the MSE performance of the estimator decreases when σ increases since the signal to noise ratio then degrades. Expectedly as well, this degradation is lower than the degradation in the performance of the TLS and OLS estimators, because for the latter two, performance is further degraded by the increasing discrepancy between the model assumed for inference and the correct model. Therefore, our results also enhance the important benefit of using the correct covariance in the presence of multiple sources of error. Finally, the procedure also appears to produce relatively accurate confidence intervals. Our approach works well for small and moderate noise levels, but it appears to have its limits for large σ (Figures 1e and 1f). The performance decline occurs in the direction of an underestimation of actual uncertainty, i.e., overconfidence. Additional research is needed to treat the case of σ large.

5. Discussion and Conclusion

We have proposed here an inference method which is suitable for the optimal fingerprinting case in which multiple sources of error having each their particular dependence structure are assumed to be present in both observations and simulated responses. The method is applicable in particular to the situation studied by HSAL06 where two sources of error (i.e., internal variability and climate model error) are assumed to be at play, but it could also cope with different and/or additional error terms. Observational error would be a natural candidate, taking advantage of recent observational data sets which give a precise quantification of observational uncertainty and its dependence structure.

Technically, the method relies on an iterative likelihood minimization scheme that is similar to the one introduced in SW08 and may be viewed as an extension of the classic TLS algorithm and hence was coined “weighted total least square” (WTLs) by these authors. It appears to perform well when applied to idealized simulation data, with respect to both point estimation and confidence intervals. Yet the method has a few limitations that open opportunities for further improvements and require discussion. First, while in practice the algorithm always converged toward a global maximum of the likelihood function when applied to simulated data, its convergence was neither mathematically established nor was the fact that the limit—if reached—is the desired global maximum. This limitation might be overcome by further theoretical work based for instance on *De Leeuw* [1994] who established a few general conditions of convergence for a similar partial maximization iterative algorithm. As disturbing as this situation may feel, it is yet in our view fair to say that it is far from unusual. Indeed, convergence conditions of even widely used algorithms in statistics (e.g., k-means) are not always very rigorously understood yet make sense heuristically and in practice work—as can be argued to be the case of our procedure and that of SW08 from which it is inspired. On the other hand, we derived confidence intervals based on an approximation that holds asymptotically, but the associated convergence conditions that determine its validity domain were not investigated in detail. Actually, as suggested by our results under high model error σ (Figure 1f), it would be reasonable to expect that convergence is not reached for n as low as a few hundreds when the signal-to-noise ratio is too weak, causing overconfidence in the present situation. Several options are available to address this issue, in particular, the Bayesian framework is adapted to deal with situations where frequentist, asymptotic results are prone to fail. In any case, as problematic as it might be, this limitation applies equally to most existing D & A methods—in particular to the aforementioned TLS procedure—in which confidence intervals are derived based on the same asymptotic approximation, thus justifying in our view a specific methodological focus in D & A on this question.

Despite of these limitations, the benefit of using our method instead of the OLS or TLS algorithm, which under our simulation assumptions rely on an incorrect model, appears to be material. Indeed, our results show that neglecting error sources that actually significantly affect the data studied is heavily castigated: first, by a considerably higher error on the regression coefficient β (Figure 1e) and second, by an important underestimation of associated confidence intervals (Figure 1f). Due to this ambiguous combination of an additional error on β together with an overconfident uncertainty quantification, it is difficult to speculate whether our procedure would in general yield confidence intervals of smaller or larger amplitude than those obtained under an OLS or TLS treatment applied to the same data and henceforth whether it would lead to more or less assertive D & A statements. Settling this question would be important in our view and would require using our method to revisit one or several previous case studies—for instance that of HSAL06 later reproduced in *Hegerl et al.* [2007] would be a natural option. Such a comparison would certainly be relevant

and instructive but was beyond the scope and length constraint of the present letter, which limited itself to methodological aspects.

Appendix A: Notations and Causality

The causal relationship between the true unknown regressors \mathbf{x}^* and their observed version \mathbf{x} matters. Two contrasted situations prevail in this respect. In our setup as in the vast majority of cases, the observed variable \mathbf{x} is a noised version of the latent variable \mathbf{x}^* , i.e., \mathbf{x} is caused by \mathbf{x}^* . By contrast, the symmetric case where it is the latent variable \mathbf{x}^* which this time is a noised version of the observed variable \mathbf{x} , i.e., where \mathbf{x}^* is caused by \mathbf{x} , may be treated as well in theory. In nonrigorous language, the error in the variable \mathbf{x} is thus “additive” in the latter case and “subtractive” in the former. This interpretation inspires the following two notations: $y = (\mathbf{x} + \mathbf{v})\beta + \epsilon$ and $y = (\mathbf{x} - \mathbf{v})\beta + \epsilon$, respectively; the latter being often used in D & A and in EIV literature in lieu of equation (1). Nevertheless, these two notations are confusing because they are equivalent strictly speaking—one may indeed algebraically substitute \mathbf{v} with $-\tilde{\mathbf{v}}$ in the second equation to obtain the first—yet they aim at representing two different situations. We are faced here with the well-known, extensively discussed deficiency of algebraic equations in representing causal relationships, due to the fact that the former are symmetrical objects, whereas the latter are directional ones [Pearl, 2009]. The consensual way to cope with this deficiency is to augment the equation with a so-called “path diagram” in which arrows are drawn from causes to effects [Wright, 1921]. Under this convention, the path diagram $\mathbf{x}^* \rightarrow \mathbf{x}$ should be added to equation (1) to reflect our “subtractive noise” situation, and $\mathbf{x} \rightarrow \mathbf{x}^*$ should be added for the “additive noise” situation. Finally, it is worth emphasizing that the causal distinction discussed here is not a merely theoretical one; it does strongly influence inference results (not shown) so we can definitely not afford the confusion.

Appendix B: Noise-Dependence Structure

In our model, we assume dependence between lines and independence between columns instead of independence between lines and dependence between columns as in most existing EIV models. Line dependence corresponds here to dependence in space and time. Internal variability, for instance, uses to exhibit large-scale modes of variability (e.g., El Niño–Southern Oscillation and Atlantic Multidecadal Oscillation (AMO)) so that the value observed at a given location or instant is not independent from the value observed nearby. Here we are using the same type of assumption to describe the model response error (e.g., too sensitive models may simulate an overestimated greenhouse gas (GHG) warming at most locations, or error in large-scale patterns). Column independence corresponds to independence between the column vectors of the noise \mathbf{v} . The internal variability component of \mathbf{v} can be considered independent because each model run represents an independent realization of internal variability by construction. On the other hand, it is fair to say that—at least theoretically—one may expect some degree of column dependence on the model error component of noise \mathbf{v} , for two or more forcings that affect the same physical processes. However, in practice in D & A applications, the forcings usually considered (e.g., GHGs, aerosols) involve distinct physical processes yielding contrasted responses in order for attribution to be possible, which tends to lessen this issue.

Appendix C: Simulation Test Bed Assumptions

The model patterns \mathbf{x}^* correspond to a simulation ensemble average from the Centre National de Recherches Météorologiques–Climate Model version 5 under natural and anthropogenic forcings, respectively (i.e., $p = 2$). Dimension reduction was based on spherical harmonics for spatial reduction and averaging by decade for temporal reduction, eventually yielding $n = 275$. The covariance matrix Σ associated to internal variability was estimated on an ensemble of control simulations of the same model by mean of linear shrinkage. From the assumed covariances and patterns, we finally simulate successively the noise vectors ϵ and $(v_i)_{i=1,2}$ and the data to be used for inference $x_i = x_i^* + v_i$ and $y = \mathbf{x}^*\beta + \epsilon$ with $\beta = (1, 1)'$.

Appendix D: Implementation Details

The computational cost of the above calculations is driven by the inversion of the $\mathbf{n} \times \mathbf{n}$ matrix $(\mathbf{\Omega}_i^{-1} + \beta_i^2 \mathbf{\Sigma}^{-1})$, its most intensive piece, which must be performed at every iteration of step 1. If the latter were to be performed naively, the algorithm would be prohibitively costly when n is large. This computational hassle can be significantly lightened by deriving beforehand $\mathbf{\Omega}_i^{-\frac{1}{2}}$, then form the product $\mathbf{\Omega}_i^{-\frac{1}{2}} \mathbf{\Sigma} \mathbf{\Omega}_i^{-\frac{1}{2}}$ and diagonalize it into $\mathbf{\Omega}_i^{-\frac{1}{2}} \mathbf{\Sigma} \mathbf{\Omega}_i^{-\frac{1}{2}} = \mathbf{P}_i' \mathbf{\Delta}_i \mathbf{P}_i$ with $\mathbf{P}_i' \mathbf{P}_i = \mathbf{I}$ and $\mathbf{\Delta}_i$ diagonal, for $i = 1, \dots, p$. Equation (6) can then be modified into

$$\tilde{\mathbf{x}}_i^* = (\mathbf{\Delta}_i + \beta_i^2 \mathbf{I})^{-1} (\beta_i \tilde{\mathbf{y}}_i + \mathbf{\Delta}_i \tilde{\mathbf{x}}_i) \quad \text{for } i = 1, \dots, p \quad (\text{D1})$$

where $\tilde{\mathbf{u}} = \mathbf{P}_i \mathbf{\Omega}_i^{-\frac{1}{2}} \mathbf{u}$ for any vector \mathbf{u} . Under the transformed expression (6), the time consuming $n \times n$ matrix inversion of step 1 now becomes much lighter since $\mathbf{\Delta}_i + \beta_i^2 \mathbf{I}$ is diagonal.

Acknowledgements

We gratefully acknowledge the Centre National de la Recherche Scientifique (CNRS), Météo France, the Consejo Nacional de Investigacion Cientifica y Tecnologica (CONICET), and the University of Buenos Aires (UBA) for their support in this collaboration. We would also like to thank Dáithí Stone and an anonymous reviewer for helpful suggestions and insightful reviews. Part of this work has been supported by the LEFE-MESDA project, the ANR-DADA project, the EU-FP7 ACQWA project (www.acqwa.ch), the PEPER-GIS project, the ANR-MOPERA project, the ANR-McSim project, and the MIRACCLE-GICC project.

The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

References

- Adcock, R. J. (1878), A problem in least squares, *The Analyst*, Des Moines, Iowa, pp. 5–53.
- Allen, M., and P. Stott (2003), Estimating signal amplitudes in optimal fingerprinting, Part I: Theory, *Clim. Dyn.*, *21*, 477–491.
- Allen, M., and S. Tett (1999), Checking for model consistency in optimal fingerprinting, *Clim. Dyn.*, *15*(6), 419–434.
- Deming, W. E. (1943), *Statistical Adjustment of Data*, Wiley, New York.
- De Leeuw, J. (1994), Block relaxation algorithms in statistics, in *Information Systems and Data Analysis*, edited by H. H. Bock et al., pp. 308–325, Springer-Verlag, Berlin.
- Dempster, A. P., N. M. Laird, and D. Rubin (1977), Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc. Ser. B*, *39*(1), 1–38.
- Fuller, W. A. (1987), *Measurement Error Models*, John Wiley, New York.
- Geman, S., and D. Geman (1984), Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Mach. Intell.*, *6*(6), 721–741.
- Gillard, J. (2010), An overview of linear structural models in errors in variables regression, *REVSTAT Stat. J.*, *8*(1), 57–80.
- Hegerl, G., H. Von Storch, B. Santer, U. Cubash, and P. Jones (1996), Detecting greenhouse gas-induced climate change with an optimal fingerprint method, *J. Clim.*, *9*(10), 2281–2306.
- Hegerl, G. C., F. W. Zwiers, P. Braconnot, N. P. Gillett, Y. Luo, J. A. Marengo Orsini, N. Nicholls, J. E. Penner, and P. A. Stott (2007), Understanding and attributing climate change, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., Cambridge Univ. Press, Cambridge, U. K., and New York.
- Hegerl, G., and F. Zwiers (2011), Use of models in detection and attribution of climate change, *Wiley Interdiscip. Rev. Clim. Change*, *2*, 570–591, doi:10.1002/wcc.121.
- Huntingford, C., P. Stott, M. Allen, and F. Lambert (2006), Incorporating model uncertainty into attribution of observed temperature change, *Geophys. Res. Lett.*, *33*, L05710, doi:10.1029/2005GL024831.
- Pearl, J. (2009), Causal inference in statistics: An overview, *Stat. Surv.*, *3*, 96–146.
- Lauritzen, S. L. (1996), *Graphical Models*, Clarendon Press, Oxford, U. K.
- Nounou, M. N., B. R. Bakshi, P. K. Goel, and X. Shen (2002), Process modeling by Bayesian latent variable regression, *AIChE J.*, *48*, 1775–1793.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (2007), *Numerical Recipes: The Art of Scientific Computing*, 3rd ed., Cambridge Univ. Press, New York.
- Ribes, A., S. Planton, and L. Terray (2012), Regularised optimal fingerprint for attribution. Part I: Method, properties and idealised analysis, *Clim. Dyn.*, *41*(11–12), 2817–2836.
- Schaffrin, B., and A. Wieser (2008), On weighted total least-squares adjustment for linear regression, *J. Geod.*, *82*, 415–421.
- Van Huffel, S., and J. Vanderwaal (1994), *The Total Least Squares Problem: Computational Aspects and Analysis*, SIAM, Philadelphia, Pa.
- Venzon, D. J., and S. H. Moolgavkar (1988), A method for computing profile likelihood-based confidence intervals, *Appl. Stat.*, *37*, 87–94.
- Wright, S. (1921), Correlation and causation, *J. Agric. Res.*, *20*, 557–585.