



An improved projection of climate observations for detection and attribution

Alexis Hannart^{1,2}

¹Ouranos, Montreal, Canada

²IFAECI, CNRS-CONICET-UBA, Buenos Aires, Argentina

Correspondence: Alexis Hannart (hannart.alexis@ouranos.ca)

Received: 14 March 2016 – Revised: 23 April 2018 – Accepted: 26 May 2018 – Published: 26 November 2019

Abstract. An important goal of climate research is to determine the causal contribution of human activity to observed changes in the climate system. Methodologically speaking, most climatic causal studies to date have been formulating attribution as a linear regression inference problem. Under this formulation, the inference is often obtained by using the generalized least squares (GLS) estimator after projecting the data on the r leading eigenvectors of the covariance associated with internal variability, which are evaluated from numerical climate models. In this paper, we revisit the problem of obtaining a GLS estimator adapted to this particular situation, in which only the leading eigenvectors of the noise’s covariance are assumed to be known. After noting that the eigenvectors associated with the lowest eigenvalues are in general more valuable for inference purposes, we introduce an alternative estimator. Our proposed estimator is shown to outperform the conventional estimator, when using a simulation test bed that represents the 20th century temperature evolution.

1 Introduction

1.1 Context

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 (Hegerl et al., 2007); this question has thus emerged as a research topic known as detection and attribution (D&A). From a methodological standpoint, D&A studies are usually based on linear regression methods, often referred to in this particular context as optimal fingerprinting, whereby an observed climate change is regarded as a linear combination of several externally forced signals added to internal climate variability (Hasselmann, 1979; Bell, 1986; North et al., 1982; Allen and Tett, 1999; Hegerl and Zwiers, 2011). On the one hand, the latter signals consist of spatial, temporal or space–time patterns of response to external forcings as anticipated by one or several climate models. Internal climate variability, on the other hand, is usually represented by a centred multivariate Gaussian noise. Its covariance Σ , which is also estimated from model simulations, thus describes the detailed space–time features of internal variability. In particular, its eigenvectors

can often be interpreted as the modes of variability, whose relative magnitudes are reflected by the associated eigenvalues.

Denoting y the n -dimensional vector of observed climate change, $\mathbf{x} = (x_1, \dots, x_p)$ the $n \times p$ matrix concatenating the p externally forced signals and \mathbf{v} the internal climate variability noise, the regression equation is as follows:

$$y = \mathbf{x}\beta + \mathbf{v}. \quad (1)$$

The results of the inference on the vector of regression coefficients β , and the magnitude of its confidence intervals, determine whether the external signals “are present in the observations” (Hegerl and Zwiers, 2011) and whether or not the observed change is attributable to each forcing.

Under the assumption that Σ is known, the above inference problem can be solved using the standard generalized least squares (GLS) setting (e.g. Amemiya, 1985). Under this framework, the expression of the best linear unbiased estimator (BLUE) and its variance are given by the following:

$$\hat{\beta} = (\mathbf{x}'\Sigma^{-1}\mathbf{x})^{-1}(\mathbf{x}'\Sigma^{-1}y), \quad \text{Var}(\hat{\beta}) = (\mathbf{x}'\Sigma^{-1}\mathbf{x})^{-1}. \quad (2)$$

The GLS estimator of Eq. (2) can be obtained as the ordinary least squares (OLS) estimator for the linearly trans-

formed regression equation:

$$\mathbf{T}\mathbf{y} = \mathbf{T}\mathbf{x}\boldsymbol{\beta} + \mathbf{T}\mathbf{v}, \quad (3)$$

where the $n \times n$ transformation matrix \mathbf{T} is such that the components of the transformed noise $\mathbf{T}\mathbf{v}$ are independent and identically distributed (IID). This condition yields $\mathbf{T}\boldsymbol{\Sigma}\mathbf{T}' = \mathbf{I}$ and an admissible choice for \mathbf{T} is thus as follows:

$$\mathbf{T} = \mathbf{V}\boldsymbol{\Delta}^{-\frac{1}{2}}\mathbf{V}', \quad (4)$$

where $\boldsymbol{\Delta} = \text{diag}(\lambda_1, \dots, \lambda_n)$ with $\lambda_1 \geq \dots \geq \lambda_n > 0$ is the matrix of eigenvalues of $\boldsymbol{\Sigma}$ and $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ is its matrix of column eigenvectors, i.e. $\boldsymbol{\Sigma} = \mathbf{V}\boldsymbol{\Delta}\mathbf{V}'$. In Eq. (3), the data are thus projected on the eigenvectors of the covariance matrix scaled by the inverse square root of their eigenvalues.

In most applications, the covariance $\boldsymbol{\Sigma}$ is not known and Eq. (2) cannot be implemented. This has given rise to many theoretical developments (e.g. Poloni and Sbrana, 2014; Naderahmadian et al., 2015) which often rely on further assumptions regarding the covariance $\boldsymbol{\Sigma}$, in order to allow for its estimation simultaneously with $\boldsymbol{\beta}$. In the applicative context of D&A considered here, the situation may be viewed as an in-between between known and unknown $\boldsymbol{\Sigma}$ because we have access to an ensemble of unforced climate model simulations. This ensemble can be used to extract valuable information about $\boldsymbol{\Sigma}$. However, the resulting knowledge of $\boldsymbol{\Sigma}$ is imperfect for two reasons. First, the size of the ensemble is often very small, typically a few dozens, because of the high cost of numerical experiments requiring intensive computational resources. For the range of dimensions n usually considered, typically a few hundreds or thousands, the empirical covariance \mathbf{S} may thus be non-invertible. Second, climate model simulations are known to inadequately represent the modes of internal variability at the smaller scale. Those smaller scale modes have lower magnitude and are expected to be associated with eigenvectors with low eigenvalues. Furthermore, the latter eigenvectors with low eigenvalues are also poorly sampled by construction and therefore are difficult to estimate reliably (North et al., 1982). For these two reasons, it is often assumed that only a small number r of leading eigenvectors of $\boldsymbol{\Sigma}$, which are denoted $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ and r is typically on the order of 10, can be estimated from the control runs with enough reliability. Accordingly, it is assumed that the remaining $n - r$ eigenvectors and eigenvalues are unknown.

Under this assumption, the inference on $\boldsymbol{\beta}$ must therefore be obtained based on these r leading eigenvectors only. The conventional approach presently used in D&A studies to tackle this problem follows from a simple and straightforward idea: adapting Eq. (4) to this situation by restricting the image of the transformation matrix \mathbf{T} to the subspace generated by the r leading eigenvectors, thus leading to a truncated transformation having the following expression:

$$\mathbf{T}_r = \mathbf{V}_r \boldsymbol{\Delta}_r^{-\frac{1}{2}} \mathbf{V}_r', \quad (5)$$

where $\boldsymbol{\Delta}_r = \text{diag}(\lambda_1, \dots, \lambda_r)$ and $\mathbf{V}_r = (\mathbf{v}_1, \dots, \mathbf{v}_r)$. A new estimator thus follows by similarly applying the OLS estimator to the transformed data $(\mathbf{T}_r \mathbf{x}, \mathbf{T}_r \mathbf{y})$:

$$\hat{\boldsymbol{\beta}}_r = (\mathbf{x}' \boldsymbol{\Sigma}_r^{-1} \mathbf{x})^{-1} (\mathbf{x}' \boldsymbol{\Sigma}_r^{-1} \mathbf{y}), \quad (6)$$

where $\boldsymbol{\Sigma}_r^{-1} = \mathbf{V}_r \boldsymbol{\Delta}_r^{-1} \mathbf{V}_r'$ is the pseudo-inverse of $\mathbf{V}_r \boldsymbol{\Delta}_r \mathbf{V}_r'$. Finally, from a mere terminology standpoint, applying the truncated transformation \mathbf{T}_r to the data is often referred to in D&A as “projecting” the data onto the subspace spanned by the leading eigenvectors. However, \mathbf{T}_r is in fact not a projection from a mathematical standpoint insofar as $\mathbf{T}_r^2 \neq \mathbf{T}_r$ hence we prefer to use the word “transformation”.

1.2 Objectives

The solution recalled above in Eq. (6) is easy to use and may seem to be a natural choice. It emerged as a popular approach in D&A arguably for these reasons. Nevertheless, this approach has two important limitations. Firstly, as pointed out by Hegerl et al. (1996), the leading eigenvectors sometimes fail to capture the signals \mathbf{x} . Indeed, the leading eigenvectors capture variability well, but not necessarily the signal, merely because the corresponding physical processes may differ and be associated with different patterns. Secondly, by construction, the subspace spanned by the leading eigenvectors maximizes the amplitude of the noise, thereby minimizing the ratio of signal to noise and affecting the performance of the inference on $\boldsymbol{\beta}$.

The first objective of this article is to highlight in detail these two limitations: both will be formalized and discussed immediately below in Sect. 2. Its second objective is to circumvent these limitations by building a new estimator of $\boldsymbol{\beta}$ that performs better than the conventional estimator $\hat{\boldsymbol{\beta}}_r$ of Eq. (6) while still making the same assumptions and using the same data. Most importantly, the central assumption here is that only the r leading eigenvectors of $\boldsymbol{\Sigma}$ can be estimated with enough accuracy from the control runs. Furthermore, we assume that r is given and is small compared to n (e.g. r is on the order of 10), and that N control runs are available to estimate the r leading eigenvectors with $N \geq r$.

Section 2 highlights the limitations of the conventional estimator $\hat{\boldsymbol{\beta}}_r$. Section 3 introduces our proposed estimator $\hat{\boldsymbol{\beta}}_r^*$. Section 4 evaluates the benefit of our approach based on a realistic test bed, and illustrates it on actual observations of surface temperature. Section 5 concludes the paper.

2 Limitations of the conventional estimator

2.1 General considerations

It is useful to start this discussion by returning for a moment to the case where $\boldsymbol{\Sigma}$ is assumed to be fully known, i.e. its n eigenvectors are all known perfectly, as well as their eigenvalues. Under this assumption, let us evaluate the benefit of a given eigenvector \mathbf{v}_k , $k = 1, \dots, n$, for the inference

of β . For this purpose, we propose to compare the variance $\text{Var}(\hat{\beta}) = (\mathbf{x}'\Sigma^{-1}\mathbf{x})^{-1}$ of the BLUE estimator of Eq. (2) obtained by using the transformation matrix $\mathbf{T} = \mathbf{V}\Delta^{-\frac{1}{2}}\mathbf{V}'$, to the variance of the estimator $\hat{\beta}_{-k}$ obtained by using the transformation matrix $\mathbf{T}_{-k} = \mathbf{V}_{-k}\Delta_{-k}^{-\frac{1}{2}}\mathbf{V}_{-k}'$ where \mathbf{V}_{-k} and Δ_{-k} exclude the k th eigenvector and the k th eigenvalue respectively. From Eq. (2), these variances can be rewritten after some algebra:

$$\begin{aligned}\text{Var}(\hat{\beta}) &= \left(\sum_{i=1}^n \mathbf{x}'\mathbf{v}_i\mathbf{v}_i'\mathbf{x}/\lambda_i\right)^{-1} \text{Var}(\hat{\beta}_{-k}) \\ &= \left(\sum_{i=1, \dots, n, i \neq k} \mathbf{x}'\mathbf{v}_i\mathbf{v}_i'\mathbf{x}/\lambda_i\right)^{-1}.\end{aligned}\quad (7)$$

The Sherman–Morrison–Woodbury formula is a well known formula in linear algebra, which shows that the inverse of a rank-one correction of some matrix, can be conveniently computed by doing a rank one correction to the inverse of the original matrix. By applying the Sherman–Morrison–Woodbury formula to Eq. (7), we obtain the following:

$$\text{Var}(\hat{\beta}_{-k}) = \text{Var}(\hat{\beta}) + \frac{\text{Var}(\hat{\beta}_{-k})\mathbf{x}'\mathbf{v}_k\mathbf{v}_k'\mathbf{x}\text{Var}(\hat{\beta}_{-k})}{\lambda_k + \mathbf{v}_k'\mathbf{x}\text{Var}(\hat{\beta}_{-k})\mathbf{x}'\mathbf{v}_k}.\quad (8)$$

Unsurprisingly, an immediate implication of Eq. (8) is that removing any eigenvector \mathbf{v}_k results in an increase of the variance of the estimator. Indeed, the matrix in the second term of the right-hand side of Eq. (8) is positive definite; it can thus be interpreted as the cost in variance increase of excluding \mathbf{v}_k or conversely as the benefit in variance reduction of including \mathbf{v}_k . Moreover, the most valuable eigenvectors for inferring β in the above-defined sense are those combining a small eigenvalue (low λ_k) and a good alignment with the signals \mathbf{x} (high $\mathbf{x}'\mathbf{v}_k$). Accordingly, for any eigenvector \mathbf{v}_k , its benefit is a decreasing function of λ_k , everything else being constant.

A similar conclusion can be reached by using a different approach. Consider this time the transformation $\mathbf{T}^* = \mathbf{x}'\Sigma^{-1}$. Since \mathbf{T}^* is a $p \times n$ matrix, the data dimension is thus drastically reduced from n to p in the transformed Eq. (3). Furthermore, noting that the noise components $\mathbf{T}^*\mathbf{v}$ have covariance $\mathbf{T}^*\Sigma\mathbf{T}^{*'}$, and applying the GLS estimator to the transformed data $(\mathbf{T}^*\mathbf{y}, \mathbf{T}^*\mathbf{x})$, we obtain an estimator $\hat{\beta}^*$ which has the following expression:

$$\hat{\beta}^* = (\mathbf{x}'\mathbf{T}^*(\mathbf{T}^*\Sigma\mathbf{T}^{*'})^{-1}\mathbf{T}^{*'}\mathbf{y})^{-1}(\mathbf{x}'\mathbf{T}^*(\mathbf{T}^*\Sigma\mathbf{T}^{*'})^{-1}\mathbf{T}^{*'}\mathbf{y}).\quad (9)$$

After some algebra, the above expression of $\hat{\beta}^*$ simplifies to the following:

$$\hat{\beta}^* = (\mathbf{x}'\Sigma^{-1}\mathbf{x})^{-1}(\mathbf{x}'\Sigma^{-1}\mathbf{y}),\quad (10)$$

which is exactly the expression of the estimator of Eq. (2). Therefore, remarkably, the BLUE estimator obtained from

the full data (\mathbf{y}, \mathbf{x}) is entirely recovered from the transformed data $(\mathbf{T}^*\mathbf{y}, \mathbf{T}^*\mathbf{x})$ when using the subspace spanned by \mathbf{T}^* . The latter subspace can thus be considered as optimal, in the sense that it reduces the data dimension to its bare minimum p while still preserving all the available information for inferring β . The p directions of the subspace spanned by \mathbf{T}^* are denoted (u_1^*, \dots, u_p^*) . They have the following expression:

$$u_k^* = \sum_{i=1}^n \frac{\mathbf{v}_i'x_k}{\lambda_i} \mathbf{v}_i.\quad (11)$$

Consistently with Eq. (8), Eq. (11) shows that for each $k = 1, \dots, p$, the optimal direction u_k^* emphasizes the eigenvectors that best trade off the smallest possible eigenvalue (i.e. small noise) with the best possible alignment between the eigenvector and the signal x_k (i.e. strong signal). It is interesting to notice that the foremost importance of small eigenvalues, which is underlined here, is also quite obvious in the formulation of optimal detection proposed by North and Stevens (1998). However, the importance of small eigenvalues was lost later on, as truncation on leading eigenvectors became a popular practice.

In light of these considerations, the two main limitations of the conventional estimator $\hat{\beta}_r$ of Eq. (6), mentioned in Sect. 1, appear more clearly. Firstly, in agreement with Hegerl et al. (1996), there is no guarantee that the leading eigenvectors will align well with the signals at stake. Secondly, the leading eigenvectors are by definition those with the largest eigenvalues. Therefore, a straightforward truncation onto the leading eigenvectors, such as the one used in $\hat{\beta}_r$, maximizes the noise instead of minimizing it. While the first limitation may be an important pitfall in some cases, it will not be an issue whenever the signals align well with the leading eigenvectors, which nothing prevents in general. In contrast, the second limitation always manifests whenever $\hat{\beta}_r$ is used and is thus arguably a more serious problem of the conventional approach. In the present paper, we thus choose to focus restrictively on this second limitation, and to not address the first one.

2.2 Illustration

In order to illustrate the second issue more concretely, we use surface temperature data from Ribes and Terray (2012), described in more detail in Section 4. The signals \mathbf{x} considered here consist of the responses to anthropogenic and natural forcings, with $p = 2$ and $n = 275$. Both \mathbf{x} and the covariance Σ are obtained from climate model simulations. The r leading eigenvectors $\mathbf{V}_r = (\mathbf{v}_1, \dots, \mathbf{v}_r)$ are obtained from Σ and are assumed to be known. The proportion of the noise's total variance retained by projecting the data on \mathbf{V}_r is denoted $\sigma_n^2(r)$ while the proportion of the signal's total variance associated with the same projection is denoted $\sigma_s^2(r)$. Since the projection matrix associated with projecting on the subspace

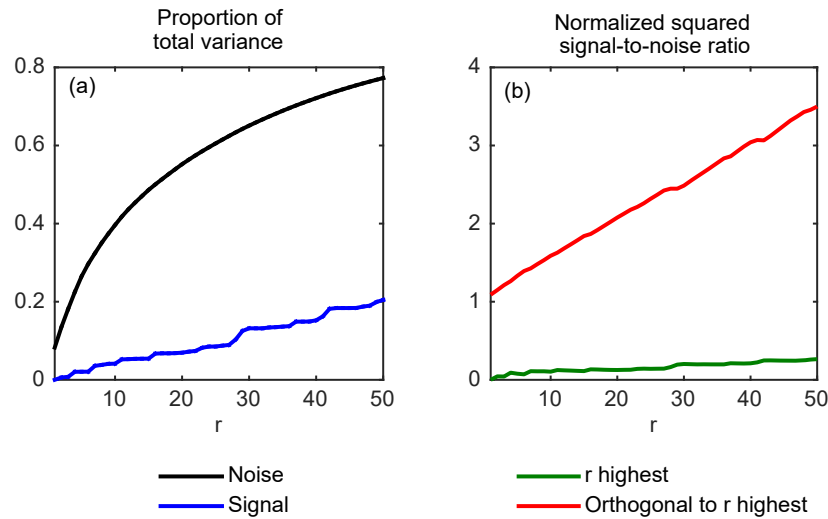


Figure 1. (a) Proportion of the total variance of the projected anthropogenic signal $\mathbf{V}_r' \mathbf{x}_1$ (thick blue line) and of the projected noise $\mathbf{V}_r' \mathbf{v}$ (thick black line) as functions of r ; uniform increase r/n (light black line). (b) Signal-to-noise ratio of the projected data when using the projection subspace \mathbf{V}_r consisting of the r leading eigenvectors (thick green line) and when using the orthogonal subspace \mathbf{V}_r^\perp (thick red line) as functions of r . Both values are normalized to the signal-to-noise ratio of the full data.

\mathbf{V}_r is equal to $\mathbf{V}_r \mathbf{V}_r'$, we have the following:

$$\begin{aligned} \sigma_n^2(r) &= \mathbb{E}(\widehat{\text{Var}}(\mathbf{V}_r \mathbf{V}_r' \mathbf{v})) / \mathbb{E}(\widehat{\text{Var}}(\mathbf{v})) \\ &= \mathbb{E}(\mathbf{v}' \mathbf{V}_r \mathbf{V}_r' \mathbf{v}) / \mathbb{E}(\mathbf{v}' \mathbf{v}) = \sum_{i=1}^r \lambda_i / \sum_{i=1}^n \lambda_i, \end{aligned} \quad (12)$$

$$\begin{aligned} \sigma_s^2(r) &= \mathbb{E}(\widehat{\text{Var}}(\mathbf{V}_r \mathbf{V}_r' \mathbf{x}_i)) / \mathbb{E}(\widehat{\text{Var}}(\mathbf{x}_i)) \\ &= \mathbf{x}_i' \mathbf{V}_r \mathbf{V}_r' \mathbf{x}_i / \mathbf{x}_i' \mathbf{x}_i \quad \text{for } i = 1, \dots, p, \end{aligned} \quad (13)$$

where for any vector \mathbf{z} , $\widehat{\text{Var}}(\mathbf{z})$ denotes the empirical variance of \mathbf{z} . Figure 1a plots $\sigma_n^2(r)$ and $\sigma_s^2(r)$ computed from Eq. (12) applied to the above-mentioned data. For simplicity, $\sigma_s^2(r)$ is shown only for the anthropogenic signal x_1 . As expected from the expression of $\sigma_n^2(r)$ given in Eq. (12), it can be seen from this plot that $\sigma_n^2(r)$ increases with r much faster than r/n , since by definition the leading eigenvectors concentrate most of the noise's variance. In contrast, $\sigma_s^2(r)$ increases with r roughly like r/n : this shows that, in this example, the signal projects nearly uniformly on each eigenvector. For instance, when $r = 20$, \mathbf{V}_r involves 7 % of the n eigenvectors, and accordingly, projecting on \mathbf{V}_r retains 8 % of the signal's total variance. In contrast, as much as 55 % of the noise's total variance is retained by projecting on \mathbf{V}_r . Consequently, the signal-to-noise ratio of the projected data (see Fig. 1b) is reduced by a factor of 7 compared to the signal-to-noise ratio of the full data. Hence, it is fair to say that, in this case, the projection on the leading eigenvectors is greatly suboptimal, insofar as it drastically reduces the signal-to-noise ratio.

Let us now consider the quantities $\sigma_n^{2*}(r)$ and $\sigma_s^{2*}(r)$ defined as above, but projecting this time on the subspace or-

thogonal to \mathbf{V}_r . By construction,

$$\begin{aligned} \sigma_n^{2*}(r) &= 1 - \sigma_n^2(r), \\ \sigma_s^{2*}(r) &= 1 - \sigma_s^2(r). \end{aligned} \quad (14)$$

As a direct consequence, for $r = 20$, the projection on the subspace orthogonal to \mathbf{V}_r captures 92 % = 100 % – 8 % of the total variance of the signal, while the variance of the projected noise is only 20 % = 100 % – 80 % of the total. Therefore, the signal-to-noise ratio of the projected data (see Fig. 1b) is this time magnified by a factor of 2 compared to the signal-to-noise ratio of the full data, and by a factor of 14 when comparing to the signal-to-noise ratio of the data projected on \mathbf{V}_r . It can thus be seen that projecting the data on the orthogonal to \mathbf{V}_r greatly magnifies the signal-to-noise ratio by filtering out a large fraction of the noise, while retaining most of the signal. These characteristics arguably make it a much more relevant projection subspace for the purpose of estimating β than the subspace of the leading eigenvectors. We further elaborate on this idea in Sect. 3.

3 An orthogonal estimator

3.1 Description

With these preliminary considerations in hand, we return to the situation of direct interest here: inferring β when only the r leading eigenvectors are available. In an attempt to improve the conventional estimator $\hat{\beta}_r$, we propose the following new estimator:

$$\hat{\beta}_r^* = (\mathbf{x}' \mathbf{x} - \mathbf{x}' \mathbf{V}_r \mathbf{V}_r' \mathbf{x})^{-1} (\mathbf{x}' \mathbf{y} - \mathbf{x}' \mathbf{V}_r \mathbf{V}_r' \mathbf{y}). \quad (15)$$

Like the conventional estimator $\hat{\beta}_r$, the proposed estimator $\hat{\beta}_r^*$ only uses the data (\mathbf{y}, \mathbf{x}) and the r leading eigenvectors $\mathbf{V}_r = (\mathbf{v}_1, \dots, \mathbf{v}_r)$. Its formulation stems from the above considerations regarding the suboptimality of the subspace \mathbf{T}_r of Eq. (5) spanned by the leading eigenvectors, as we now highlight.

Let us denote \mathbf{V}_r^\perp a $n \times (n-r)$ matrix consisting of $n-r$ orthonormal column vectors spanning the subspace orthogonal to $(\mathbf{v}_1, \dots, \mathbf{v}_r)$. In other words, we have the following by construction:

$$\begin{aligned}\mathbf{V}_r \mathbf{V}_r' + \mathbf{V}_r^\perp \mathbf{V}_r^{\perp'} &= \mathbf{I}, \\ \mathbf{V}_r^{\perp'} \mathbf{V}_r^\perp &= \mathbf{I}.\end{aligned}\quad (16)$$

Note that infinitely many matrices \mathbf{V}_r^\perp satisfy Eq. (16), and that it is straightforward to construct one such matrix from the r leading eigenvectors $(\mathbf{v}_1, \dots, \mathbf{v}_r)$, for instance by using the Gram–Schmidt algorithm. Furthermore, it is also important to note that, since \mathbf{V}_r^\perp is orthogonal to \mathbf{V}_r , it spans the same subspace as the subspace spanned by the $n-r$ eigenvectors $(\mathbf{v}_{r+1}, \dots, \mathbf{v}_n)$ associated with the smallest eigenvalues, because the latter are also orthogonal to $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ by definition. Therefore, it is not necessary to have knowledge of these $n-r$ eigenvectors $(\mathbf{v}_{r+1}, \dots, \mathbf{v}_n)$ in order to be able to project the data onto the latter subspace: only the knowledge of $(\mathbf{v}_1, \dots, \mathbf{v}_r)$ is required for that purpose. Thus, with the knowledge of the r leading eigenvectors only, we are able to project the data onto the orthogonal subspace spanned by \mathbf{V}_r^\perp using the transformation matrix $\mathbf{T}_r^\perp = \mathbf{V}_r^\perp \mathbf{V}_r^{\perp'} = \mathbf{I} - \mathbf{V}_r \mathbf{V}_r'$, thus obtaining $(\mathbf{T}_r^\perp \mathbf{y}, \mathbf{T}_r^\perp \mathbf{x})$. The estimator proposed in Eq. (15) consists of using the OLS estimator corresponding to these transformed data:

$$\begin{aligned}\hat{\beta}_r^* &= (\mathbf{x}' \mathbf{V}_r^\perp \mathbf{V}_r^{\perp'} \mathbf{x})^{-1} (\mathbf{x}' \mathbf{V}_r^\perp \mathbf{V}_r^{\perp'} \mathbf{y}) \\ &= (\mathbf{x}' \mathbf{x} - \mathbf{x}' \mathbf{V}_r \mathbf{V}_r' \mathbf{x})^{-1} (\mathbf{x}' \mathbf{y} - \mathbf{x}' \mathbf{V}_r \mathbf{V}_r' \mathbf{y}).\end{aligned}\quad (17)$$

It is interesting to note that the orthogonal projection subspace \mathbf{V}_r^\perp does not actually need to be obtained to derive $\hat{\beta}_r^*$ because only $\mathbf{V}_r^\perp \mathbf{V}_r^{\perp'}$ matters for this purpose, and this matrix happens to be directly known from \mathbf{V}_r through the orthogonality Eq. (16). This yields the right-most term of Eq. (17), which corresponds to the expression of $\hat{\beta}_r^*$ given above in Eq. (15).

However, the choice of applying the OLS estimator to the projected data as implied by Eq. (17), rather than applying the GLS estimator, may seem surprising at first. The rationale for this choice is the following. Since we only assume knowledge of the r leading eigenvectors, the projection subspace \mathbf{V}_r^\perp is thus by construction unknown from the standpoint of the covariance of the projected noise. In other words, while the projected noise $\mathbf{T}_r^\perp \mathbf{v}$ in the projected regression equation remains Gaussian centred, we do not know anything about its covariance. Thus, we use by default the assumption that $\mathbf{T}_r^\perp \mathbf{v}$ is IID with unknown variance σ^2 , which is a commonplace assumption used in regression analysis when nothing

is known about the noise's covariance. An immediate implication of this assumption is that the associated log likelihood is the following:

$$-2 \log \ell(\boldsymbol{\beta}, \sigma^2) = n \log \sigma^2 + \frac{1}{\sigma^2} (\mathbf{y} - \mathbf{x} \boldsymbol{\beta})' \mathbf{V}_r^\perp \mathbf{V}_r^{\perp'} (\mathbf{y} - \mathbf{x} \boldsymbol{\beta}). \quad (18)$$

The complete likelihood of Eq. (18) can be concentrated in $\boldsymbol{\beta}$ by maximizing out σ^2 to obtain the concentrated likelihood $\ell_c(\boldsymbol{\beta})$:

$$\begin{aligned}-2 \log \ell_c(\boldsymbol{\beta}) &= n \log \left\{ (\mathbf{y} - \mathbf{x} \boldsymbol{\beta})' \mathbf{V}_r^\perp \mathbf{V}_r^{\perp'} (\mathbf{y} - \mathbf{x} \boldsymbol{\beta}) \right\} \\ &= n \log \left\{ (\mathbf{y} - \mathbf{x} \boldsymbol{\beta})' (\mathbf{I} - \mathbf{V}_r \mathbf{V}_r') (\mathbf{y} - \mathbf{x} \boldsymbol{\beta}) \right\}.\end{aligned}\quad (19)$$

Our proposed OLS estimator $\hat{\beta}_r^*$ results from the maximization of the above likelihood. The above considerations also have some implications with respect to the uncertainty analysis on the proposed estimator $\hat{\beta}_r^*$. Indeed, in the present context, it is straightforward to obtain confidence intervals around $\hat{\beta}_r^*$ based on the ratio of likelihoods $\ell_c(\boldsymbol{\beta})/\ell_c(\hat{\beta}_r^*)$, following a classic approach in statistics that is often used in D&A (e.g. Hannart et al., 2014).

It is important to underline that the latter approach to uncertainty quantification relies on the assumption that the projected noise $\mathbf{T}_r^\perp \mathbf{v}$ is IID. Since the unknown true covariance of $\mathbf{T}_r^\perp \mathbf{v}$ is in fact equal to $\boldsymbol{\Sigma}_r^\perp = \mathbf{V}_r^\perp \boldsymbol{\Delta}_r^\perp \mathbf{V}_r^{\perp'}$, this assumption is incorrect and hence the resulting confidence intervals are prone to be incorrect as well. An initial assessment under test bed conditions described in Section 4 suggested that the proposed approach is conservative in the sense that it leads to an overestimation of uncertainty (not shown). However, a more detailed assessment of the performance of this uncertainty quantification procedure is needed, but is beyond the scope of the present paper.

3.2 Extension to total least squares

The signals \mathbf{x} are estimated as the empirical averages of several ensembles of finite size m , which introduces sampling uncertainty. In order to account for the latter, it is commonplace in D&A to reformulate the linear regression model of Eq. (1) as an error-in-variables (EIV) model (Allen and Stott, 2003):

$$\begin{aligned}\mathbf{y} &= \mathbf{x}^* \boldsymbol{\beta} + \mathbf{v}, \\ \mathbf{x} &= \mathbf{x}^* + \boldsymbol{\varepsilon},\end{aligned}\quad (20)$$

where \mathbf{x}^* is now an unobserved latent variable, \mathbf{x} is the ensemble average estimator of the true responses \mathbf{x}^* and $\boldsymbol{\varepsilon}$ is a Gaussian centred noise term with covariance $\boldsymbol{\Sigma}/m$.

It is straightforward to generalize the approach exposed in Sect. 2.3 to this error-in-variables regression model. Likewise, the data (\mathbf{y}, \mathbf{x}) , the latent variable \mathbf{x}^* , and the noise terms \mathbf{v} and $\boldsymbol{\varepsilon}$ are projected on the orthogonal subspace using \mathbf{V}_r^\perp . The projected residuals are then also assumed IID with

unknown variance σ^2 . This yields the following complete likelihood:

$$-2\log \ell(\boldsymbol{\beta}, \mathbf{x}^*, \sigma^2) = n(p+1)\log \sigma^2 + \frac{1}{\sigma^2}(\mathbf{y} - \mathbf{x}^* \boldsymbol{\beta})' \mathbf{V}_r^\perp \mathbf{V}_r^{\perp'} (\mathbf{y} - \mathbf{x}^* \boldsymbol{\beta}) + \frac{m}{\sigma^2} \text{Tr} \left\{ (\mathbf{x} - \mathbf{x}^*)' \mathbf{V}_r^\perp \mathbf{V}_r^{\perp'} (\mathbf{x} - \mathbf{x}^*) \right\}, \quad (21)$$

which can be concentrated in $\boldsymbol{\beta}$ by maximizing out σ^2 and \mathbf{x}^* to obtain the concentrated likelihood $\ell_c(\boldsymbol{\beta})$:

$$-2\log \ell_c(\boldsymbol{\beta}) = n(p+1)\log \left\{ \frac{(\mathbf{y} - \mathbf{x} \boldsymbol{\beta})' (\mathbf{I} - \mathbf{V}_r \mathbf{V}_r') (\mathbf{y} - \mathbf{x} \boldsymbol{\beta})}{1 + \frac{1}{m} \boldsymbol{\beta}' \boldsymbol{\beta}} \right\}. \quad (22)$$

The so-called total least squares (TLS) estimator results from the maximization of the above concentrated likelihood in $\boldsymbol{\beta}$, and is available under a closed form expression:

$$\hat{\boldsymbol{\beta}}_r^* = -\frac{\sqrt{m}}{z_{p+1}} \mathbf{z}, \quad (23)$$

where $\mathbf{z} = (z_1, \dots, z_p)'$ is a p -dimensional vector and z_{p+1} is a scalar, such that $(z_1, \dots, z_{p+1})'$ is the eigenvector associated with the smallest eigenvalue of the positive definite matrix $\mathbf{S} = [\sqrt{m} \mathbf{x} \mathbf{y}]' (\mathbf{I} - \mathbf{V}_r \mathbf{V}_r') [\sqrt{m} \mathbf{x} \mathbf{y}]$ of size $p+1$. The estimator of Eq. (23) is a generalization of the OLS estimator of Eq. (17), the latter is indeed a special case of the former for $m = +\infty$. Furthermore, in the particular case $p = 1$ which is often referred to as the Deming regression, the estimator of Eq. (23) takes the following expression:

$$\hat{\beta}_r^* = \frac{\sqrt{(m S_{xx} - S_{yy})^2 + 4m S_{xy}^2} - (m S_{xx} - S_{yy})}{2 S_{xy}}, \quad (24)$$

where $S_{xx} = \mathbf{x}'(\mathbf{I} - \mathbf{V}_r \mathbf{V}_r') \mathbf{x}$, $S_{yy} = \mathbf{y}'(\mathbf{I} - \mathbf{V}_r \mathbf{V}_r') \mathbf{y}$ and $S_{xy} = \mathbf{x}'(\mathbf{I} - \mathbf{V}_r \mathbf{V}_r') \mathbf{y}$. Next, following the same approach as in Sect. 2.3, it would be tempting to derive confidence intervals around $\hat{\beta}_r^*$ based on the ratio of likelihoods $\ell_c(\boldsymbol{\beta})/\ell_c(\hat{\beta}_r^*)$. While this approach performs reasonably well in the context of EIV regression with known variance (Hannart et al., 2014), it is known to be problematic in the present context of EIV regression with unknown variance (Casella and Berger, 2008; Sect. 12.2). In this case, more suitable alternatives are available, e.g. an estimator of the variance of $\hat{\beta}_r^*$ for $p = 1$ is given by Fuller (1987) Sect. 1.3; it has the following expression:

$$\widehat{V}(\hat{\beta}_r^*) = \frac{1}{n-r} \left(\hat{\beta}_r^* \frac{S_{xx}}{S_{xy}} - 1 \right) \left(m \hat{\beta}_r^* \frac{S_{xx}}{S_{xy}} + \hat{\beta}_r^{*2} \right), \quad (25)$$

from which confidence intervals can be obtained. When $p > 1$ the above variance estimator is applied successively to each component i of $\hat{\beta}_r^*$ ($i = 1, \dots, p$) by replacing \mathbf{y} with $\mathbf{y} - \sum_{k \neq i} \hat{\beta}_{r,k}^* x_k$ in Eq. (25). It is worthwhile noting that various approaches are available in the literature to tackle the

general problem of deriving confidence intervals for $\boldsymbol{\beta}$ under the present EIV regression model with unknown variance. While it would be worthwhile to investigate and to compare their performance in practice in a D&A context, such investigation is beyond the scope of this paper.

Summarizing, the proposed solution here to deal with the case of error in variables thus consists of using Eq. (23) to derive an estimate of $\boldsymbol{\beta}$ and Eq. (25) to derive a confidence interval.

3.3 Continuity considerations

The motivation of this subsection is merely to provide a more in-depth mathematical grounding to the proposed estimator, beyond the general considerations of Sect. 2.1. This subsection can safely be skipped without affecting the understanding of the remainder of this article, in so far as it does not provide any additional results or properties regarding the estimator itself.

The idea here is that, in addition to the general considerations of Sect. 2.1, the proposed estimator $\hat{\beta}_r^*$ can also be justified by following a continuity extension argument. For any given \mathbf{x} and \mathbf{y} , let us consider the BLUE estimator of Eq. (2) as a function of the known covariance $\boldsymbol{\Sigma}$. For this purpose, we define the function f on the set of positive definite matrices of dimension n , denoted $\mathcal{S}_n^{++}(\mathbb{R})$, as follows:

$$f : \mathcal{S}_n^{++}(\mathbb{R}) \rightarrow \mathbb{R}^p, \quad f(\mathbf{A}) = (\mathbf{x}' \mathbf{A}^{-1} \mathbf{x})^{-1} (\mathbf{x}' \mathbf{A}^{-1} \mathbf{y}), \quad (26)$$

so that we have $\hat{\beta} = f(\boldsymbol{\Sigma})$.

In the present context, we assumed that $\boldsymbol{\Sigma}$ is unknown but that $\boldsymbol{\Sigma}_r = \mathbf{V}_r \boldsymbol{\Delta}_r \mathbf{V}_r'$ is known. While it is tempting to apply f in $\boldsymbol{\Sigma}_r$, unfortunately $\boldsymbol{\Sigma}_r$ is not positive definite so that f is not defined in $\boldsymbol{\Sigma}_r$. Hence, an estimate of $\boldsymbol{\beta}$ cannot be obtained by straightforwardly applying the function f in $\boldsymbol{\Sigma}_r$. Nevertheless, $\boldsymbol{\Sigma}_r$ belongs to the set of positive semi-definite matrices, denoted $\mathcal{S}_n^+(\mathbb{R})$, which has the following interesting topological property:

$$\mathcal{S}_n^+(\mathbb{R}) = \overline{\mathcal{S}_n^{++}(\mathbb{R})}, \quad (27)$$

where for any subset E , the notation \overline{E} used above in Eq. (27) denotes the *adherence* of E . Equation (27) thus implies that $\mathcal{S}_n^{++}(\mathbb{R})$ is dense in $\mathcal{S}_n^+(\mathbb{R})$, which means that for any positive semi-definite matrix \mathbf{A} there exists a sequence of positive definite matrices converging to \mathbf{A} . Equation (27) therefore suggests the possibility to define a function \bar{f} which continuously extends f to the set of positive semi-definite matrices. For this purpose, let us consider the matrix $\mathbf{A} + \varepsilon \mathbf{I}$ with $\varepsilon > 0$. It is immediate to show that for any $\mathbf{A} \in \mathcal{S}_n^+(\mathbb{R})$ and for any arbitrarily small but positive ε , we have $\mathbf{A} + \varepsilon \mathbf{I} \in \mathcal{S}_n^{++}(\mathbb{R})$. Taking advantage of this remark, we define \bar{f} by $\bar{f}(\mathbf{A}) = \lim_{\varepsilon \rightarrow 0^+} f(\mathbf{A} + \varepsilon \mathbf{I})$. With this definition, \bar{f} can be obtained

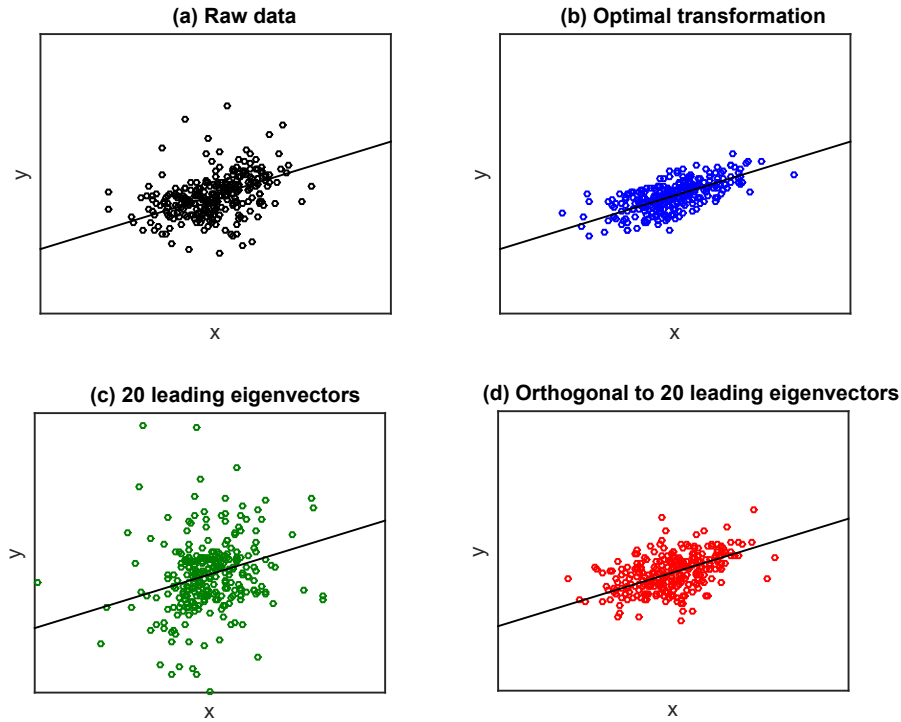


Figure 2. (a) Scatter plot of one realization of the raw data (y, x) using the simulation test bed of Section 4: the observation vector y is plotted on the vertical axis and the first column vector x_1 of the matrix x is plotted on the horizontal axis. (b) Same as (a) for the transformed data Ty and Tx using the optimal transformation matrix $T = V\Delta^{-\frac{1}{2}}V'$. (c) Same as (b) when using the truncated transformation matrix $T_r = V_r\Delta_r^{-\frac{1}{2}}V_r'$ on the leading eigenvectors, with $r = 20$. (d) Same as (b) when using the proposed transformation matrix $T_r^\perp = V_r^\perp V_r^{\perp'}$ with $r = 20$.

under a closed form expression:

$$\begin{aligned} \bar{f} : S_n^+(\mathbb{R}) &\rightarrow \mathbb{R}^p, \\ \bar{f}(A) &= (x'x - x'V_A V_A' x)^{-1} (x'y - x'V_A V_A' y). \end{aligned} \quad (28)$$

The proof of Eq. (28) is detailed in Appendix A. The proposed estimator $\hat{\beta}_r^*$ of Eq. (15) is therefore obtained by applying the extended function \bar{f} defined above, to the known, positive semi-definite matrix Σ_r :

$$\hat{\beta}_r^* = \bar{f}(\Sigma_r) = (x'x - x'V_r V_r' x)^{-1} (x'y - x'V_r V_r' y). \quad (29)$$

4 Simulations and illustration

4.1 Data

We illustrate the method by applying it to surface temperature data and climate model simulations over the 1901–2000 period and over the entire surface of the Earth. For this purpose, we use the data described and studied in Ribes and Terray (2012). The temperature observations are based on the HadCRUT3 merged land–sea temperature data set (Brohan et al., 2006). The simulated temperature data used to estimate the response patterns are provided by results from the CMIP5 archive arising from simulations performed with

the CNRM-CM5 model of Météo France under NAT and ANT experiments. For both experiments, the size of the ensembles is $m = 6$. However, estimates of internal variability are based on intra-ensemble variability from 10 CMIP5 models (CNRM-CM5, CanESM2, HadGEM2-ES, GISS-E2-R, GISS-E2-H, CSIRO-Mk3-6-0, IPSL-CM5A-LR, BCC-CSM1-1, NorESM1-M, FGOALS-g2) under NAT and ANT experiments as well as pre-industrial simulations from both the CMIP3 and CMIP5 archives, leading to a large ensemble of 374 members. Standard pre-processing is applied to all observed and simulated data. Data are first interpolated on the $5^\circ \times 5^\circ$ observational grid and then the spatio-temporal observational mask is applied. The dimension of the data set is then further reduced by computing decadal means and projecting the resulting spatial patterns onto spherical harmonics with resolution T4 to obtain vectors of size $n = 275$ (see Ribes et al., 2012 for more details about ensembles and pre-processing). The ANT and NAT individual responses ($p = 2$) are estimated by averaging the corresponding two ensembles to obtain the matrix of regressors x . The covariance matrix Σ is estimated as the empirical covariance of the large ensemble of control runs.

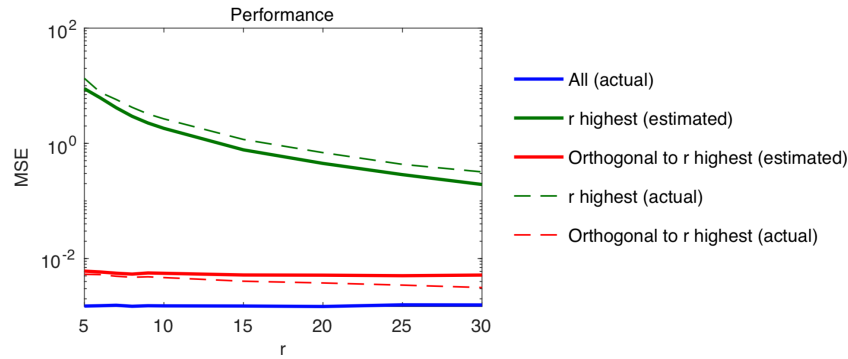


Figure 3. Performance results on simulated data: MSE of several estimators of β (vertical axis) as functions of the number of eigenvectors r (horizontal axis) assumed to be reliably estimable from control runs: conventional estimator $\hat{\beta}_r$ computed based on estimated eigenvectors (thick green line) with $N = 30$, and based on the true eigenvectors (light green line); proposed orthogonal estimator $\hat{\beta}_r^*$ computed based on estimated eigenvectors (thick red line) with $N = 30$, and based on the true eigenvectors (light red line); GLS estimator computed based on the true value of the covariance Σ (thick blue line).

4.2 Performance on simulated data

This section evaluates the performance of the estimators described above, by applying it to simulated values of y obtained from the linear regression equation $y = x\beta + v$ assumed by the model of Eq. (1), where the noise v is simulated from a multivariate Gaussian distribution with covariance Σ and where $\beta = (1, 1)'$ is used. This setting thus assumes that x is known exactly, therefore it does not include the EIV situation prevailing in Eq. (20) and in Sect. 3.3 where x is measured with an error.

The use of simulated rather than real data aims at verifying that our proposed estimator performs correctly, and at comparing its performance with the conventional procedure, a goal which requires the actual values of β to be known. A sample of $N = 30$ control runs was also simulated from a multivariate Gaussian distribution with covariance Σ , and the r leading eigenvectors and eigenvalues were estimated from these control runs.

Figure 2 shows scatter plots of one realization of the data (x, y) under this test bed, for the raw data (x, y) as well as for the transformed data (Tx, Ty) , when using the three main transformations considered above: the optimal transformation $T = V\Delta^{-\frac{1}{2}}V'$ which assumes a fully known covariance Σ ; the truncated transformation matrix $T_r = V_r\Delta_r^{-\frac{1}{2}}V_r'$ on the r leading eigenvectors, with $r = 20$, and the proposed projection matrix $T_r^\perp = V_r^\perp V_r^{\perp'}$ on the orthogonal to the r leading eigenvectors, again with $r = 20$. Figure 2 allows the behavior of these transformations to be visualized. Firstly, the benefit of the optimal transformation $T = V\Delta^{-\frac{1}{2}}V'$ appears clearly, as it greatly enhances the signal. Secondly, it is also apparent that the standard truncated transformation using the r leading eigenvectors fails to enhance the signal, actually leading to a cloud of points which is more noisy than the original raw data. In contrast, the proposed transformation on the orthogonal to the r leading eigenvectors accu-

rately captures part of the signal enhancement produced by the optimal transformation.

The conventional estimator $\hat{\beta}_r$ and the proposed estimator $\hat{\beta}_r^*$ of Eq. (15) were both derived based on the r leading estimators estimated from the N control runs. Then, these two estimators were derived again, this time using the true values of the r leading eigenvectors, for comparison. Finally, the GLS estimator of Equation (1) was derived based on the true value of the covariance Σ . The performance of each five estimators of β thus obtained was evaluated based on average mean squared error (MSE) $\sum_{j=1}^J (\hat{\beta}_j - \beta)'(\hat{\beta}_j - \beta)/N$, where $j = 1, \dots, J$ denotes the simulation number and $J = 1e4$. These calculations were repeated for several values of r ranging from 5 to 30. Results are plotted in Fig. 3. They show an important gap in MSE between the proposed estimator and the conventional approach with the former exceeding the latter by a factor of 1000 for $r = 5$, decreasing to 80 for $r = 30$. This gap suggests a substantial benefit of using $\hat{\beta}_r^*$ and strongly emphasizes the relevance of this approach compared to the conventional one.

When comparing the performance of $\hat{\beta}_r^*$ using estimated vs. actual values of the r leading eigenvectors, a slight benefit of using actual values is found. This benefit increases with r , as eigenvectors with lower eigenvalues tend to be more difficult to estimate. In contrast, when comparing the performance of $\hat{\beta}_r$ using estimated vs. actual values of the r leading eigenvectors, a slight benefit of using the estimated values is found. This superiority of the estimated, and thus incorrect, values may seem surprising at first, but is in fact natural and in line with our main point. Indeed, it can be explained by the fact that the estimated leading eigenvectors \hat{V}_r match to a large extent with V_r but also span to some extent the true orthogonal subspace V_r^\perp , which, according to our main point here, is much more suited for estimating β . Unsurprisingly, the estimator based on the estimated leading eigenvectors \hat{V}_r

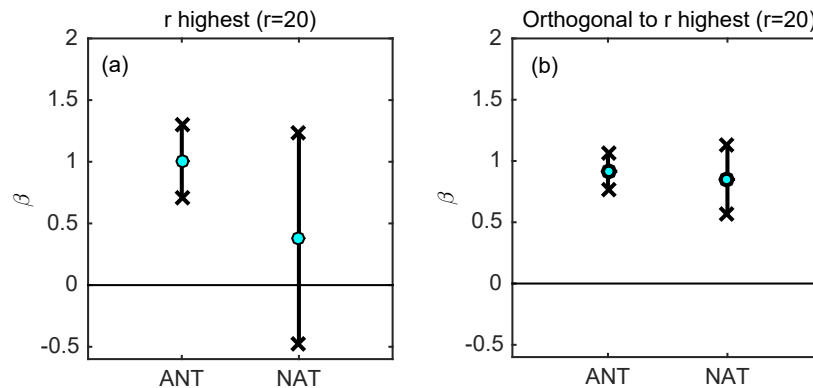


Figure 4. Illustration of temperature data with $r = 20$. (a) Regression coefficients β_{ANT} and β_{NAT} estimated using the conventional estimator $\hat{\beta}_r$: value of the estimate (green dots) and 95 % confidence interval (thick black lines). (b) Same as (a) under the proposed estimator $\hat{\beta}_r^*$.

thus performs better than its counterpart based on actual values.

4.3 Illustration on real data

The method is finally illustrated by applying it to actual observations of surface temperature y , with the same regressors x and covariance Σ as described above. The observed temperature data y is based on the HadCRUT3 merged land–sea temperature data set (Brohan et al., 2006). The method is applied for $r = 20$. Further, we use a number of control runs N that is small compared to n , which is a common situation, even though we have access under the present test bed to 374 control runs. Thus, $N = 30$ control runs are randomly selected from the 374 control runs available and the eigenvectors are estimated from these 30 runs. For the inference of β and confidence intervals, we use the solution described in Sect. 3.2 which consists of using Eq. (23) to derive an estimate of β and Eq. (25) to derive a confidence interval. Results are shown in Fig. 4. For the scaling factor corresponding to anthropogenic forcings, both methods yield an estimate which is close to 1 and significantly positive. However, the level of uncertainty obtained with the proposed estimator is 3 times smaller than when using the conventional method. An even wider uncertainty gap is observed for the coefficient corresponding to natural forcings. In this case, the confidence interval includes zero under the conventional approach, but it does not under the proposed orthogonal approach. However, as already stated in Sect. 3, a caveat in the latter result is that the uncertainty quantification procedure used here has not been evaluated in detail, and thus we do not know for sure that the obtained uncertainty estimate is well calibrated.

5 Conclusions

We have introduced a new estimator of the vector β of linear regression coefficients, adapted to the context where only the r leading eigenvectors of the noise's covariance Σ are

known, which is an assumption often relevant in climate change attribution studies. General considerations have first shown that when the covariance is known, its most relevant eigenvectors for projection are those trading off the smallest possible eigenvalue with the best possible alignment to the signal. The optimal direction is thus associated in general with the smallest eigenvalues, not with the highest ones. Our proposed estimator builds upon this finding and is based on projecting the data on a subspace which is orthogonal to the r leading eigenvectors.

When applied on a simulation test bed that is realistic with respect to D&A applications, we find that the proposed estimator outperforms the conventional estimator by several orders of magnitude for small values of r . Our proposal is thus prone to significantly affect D&A results by decreasing the estimated level of uncertainty.

Substantial further work is needed to evaluate the performance of the proposed uncertainty quantification, in particular in an EIV context. Such an evaluation was beyond the scope of the present work. Its primary focus was to demonstrate that the choice of the leading eigenvectors as a projection subspace is a vastly suboptimal one for D&A purposes; and that projecting on a subspace which is orthogonal to the r leading eigenvectors yields a significant improvement.

Data availability. The data used for illustration throughout this article was first described and studied by Ribes and Terray (2013). As mentioned in the latter study, it can be accessed at the following url: <http://www.cnrm-game.fr/spip.php?article23&lang=en> (last access: 19 December 2018).

Appendix A: Proof of equation (28)

Let $\mathbf{A} \in S^+(\mathbb{R})$ be a non-invertible positive matrix. Denote Δ_A the matrix of non-zero eigenvalues of \mathbf{A} , \mathbf{V}_A the matrix of eigenvectors associated with Δ_A and \mathbf{V}_A^\perp the matrix of eigenvectors associated with the eigenvalue zero. For $\varepsilon > 0$, we have the following:

$$\begin{aligned} (\mathbf{A} + \varepsilon \mathbf{I})^{-1} &= (\mathbf{V}_A \Delta_A \mathbf{V}_A' + \varepsilon \mathbf{I})^{-1} \\ &= (\mathbf{V}_A \Delta_A \mathbf{V}_A' + \varepsilon \mathbf{V}_A^\perp \mathbf{V}_A^{\perp'} + \varepsilon \mathbf{V}_A \mathbf{V}_A')^{-1} \\ &= \left(\mathbf{V}_A (\Delta_A + \varepsilon \mathbf{I}) \mathbf{V}_A' + \varepsilon \mathbf{V}_A^\perp \mathbf{V}_A^{\perp'} \right)^{-1} \quad (\text{A1}) \\ &= \mathbf{V}_A (\Delta_A + \varepsilon \mathbf{I})^{-1} \mathbf{V}_A' + \frac{1}{\varepsilon} \mathbf{V}_A^\perp \mathbf{V}_A^{\perp'}. \end{aligned}$$

Therefore,

$$\begin{aligned} f(\mathbf{A} + \varepsilon \mathbf{I}) &= \left(\mathbf{x}' \mathbf{V}_A (\Delta_A + \varepsilon \mathbf{I})^{-1} \mathbf{V}_A' \mathbf{x} + \frac{1}{\varepsilon} \mathbf{x}' \mathbf{V}_A^\perp \mathbf{V}_A^{\perp'} \mathbf{x} \right)^{-1} \\ &\quad \left(\mathbf{x}' \mathbf{V}_A (\Delta_A + \varepsilon \mathbf{I})^{-1} \mathbf{V}_A' \mathbf{y} + \frac{1}{\varepsilon} \mathbf{x}' \mathbf{V}_A^\perp \mathbf{V}_A^{\perp'} \mathbf{y} \right) \\ &= \left(\mathbf{x}' \mathbf{V}_A \mathbf{B}_\varepsilon \mathbf{V}_A' \mathbf{x} + \mathbf{x}' \mathbf{V}_A^\perp \mathbf{V}_A^{\perp'} \mathbf{x} \right)^{-1} \\ &\quad \left(\mathbf{x}' \mathbf{V}_A \mathbf{B}_\varepsilon \mathbf{V}_A' \mathbf{y} + \mathbf{x}' \mathbf{V}_A^\perp \mathbf{V}_A^{\perp'} \mathbf{y} \right), \quad (\text{A2}) \end{aligned}$$

with $\mathbf{B}_\varepsilon = \varepsilon (\Delta_A + \varepsilon \mathbf{I})^{-1}$. We have $\lim_{\varepsilon \rightarrow 0^+} \mathbf{B}_\varepsilon = 0$, hence,

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} f(\mathbf{A} + \varepsilon \mathbf{I}) &= \left(\mathbf{x}' \mathbf{V}_A^\perp \mathbf{V}_A^{\perp'} \mathbf{x} \right)^{-1} \left(\mathbf{x}' \mathbf{V}_A^\perp \mathbf{V}_A^{\perp'} \mathbf{y} \right) \\ &= (\mathbf{x}' \mathbf{x} - \mathbf{x}' \mathbf{V}_A \mathbf{V}_A' \mathbf{x})^{-1} (\mathbf{x}' \mathbf{y} - \mathbf{x}' \mathbf{V}_A \mathbf{V}_A' \mathbf{y}), \quad (\text{A3}) \end{aligned}$$

which proves Eq. (28).

Competing interests. The author declares that there is no conflict of interest.

Acknowledgements. This work was supported by the ANR grant DADA and by the LEFE grant MESDA. I thank Dáithí Stone, Philippe Naveau and two anonymous reviewers for very useful comments and discussions that helped improve considerably this article.

Edited by: Christopher Paciorek

Reviewed by: two anonymous referees

References

- Allen, M. and Stott, P.: Estimating signal amplitudes in optimal fingerprinting, part i: theory, *Clim. Dynam.*, 21, 477–491, 2003.
- Allen, M. and Tett, S.: Checking for model consistency in optimal fingerprinting, *Clim. Dynam.*, 15, 419–434, 1999.
- Amemiya, T.: *Advanced Econometrics*, Harvard University Press, 1985.
- Bell, T. P.: Theory of optimal weighting of data to detect climate change, *J. Atmos. Sci.*, 43, 1694–1710, 1986.
- Brohan, P., Kennedy, J., Harris, I., Tett, S., and Jones, P.: Uncertainty estimates in regional and global observed temperature changes: a new data set from 1850, *J. Geophys. Res.*, 111, D12106, <https://doi.org/10.1029/2005JD006548>, 2006.
- Casella, G. and Berger, R. L.: *Statistical Inference*, Duxbury, 2nd Edition, 2008.
- Fuller, W. A.: *Measurement Error Models*, John Wiley, New York, 1987.
- Hannart, A.: Integrated Optimal Fingerprinting: Method description and illustration, *J. Climate*, 29, 1977–1998, 2016.
- Hannart, A., Ribes, A., and Naveau, P.: Optimal fingerprinting under multiple sources of uncertainty, *Geophys. Res. Lett.*, 41, 1261–1268, <https://doi.org/10.1002/2013GL058653>, 2014.
- Hasselmann, K.: On the signal-to-noise problem in atmospheric response studies, in: *Meteorology of tropical oceans*, edited by: Shaw, D. B., 251–259, Royal Meteorological Society, London, 1979.
- Hegerl, G. C. and North, G. R.: Comparison of statistically optimal methods for detecting anthropogenic climate change, *J. Climate*, 10, 1125–1133, 1997.
- Hegerl, G. and Zwiers, F.: Use of models in detection and attribution of climate change, *WIREs Clim Change*, 2, 570–591, <https://doi.org/10.1002/wcc.121>, 2011.
- Hegerl, G. C., von Storch, H., Hasselmann, K., Santer, B. D., Cubasch, U., and Jones, P. D.: Detecting greenhouse gas induced climate change with an optimal fingerprint method, *J. Climate*, 9, 2281–2306, 1996.
- Hegerl, G. C., Zwiers, F. W., Braconnot, P., Gillett, N. P., Luo, Y., Marengo Orsini, J. A., Nicholls, N., Penner, J. E., and Stott, P. A.: 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: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.
- Lott, F. C., Stott, P. A., Mitchell, D. M., Christidis, N., Gillett, N. P., Haimberger, L., Perlwitz, J., and Thorne, P. W.: Models versus radiosondes in the free atmosphere: A new detection and attribution analysis of temperature, *J. Geophys. Res.-Atmos.*, 118, 2609–2619, <https://doi.org/10.1002/jgrd.50255>, 2013.
- Min, S., Zhang, X., and Zwiers, F. W.: Human-Induced Arctic Moistening, *Science*, 320, 518–520, 2008.
- Naderahmadian, Y., Tinati, M. A., and Beheshti, S.: Generalized adaptive weighted recursive least squares dictionary learning, *Signal Process.*, 118, 89–96, 2015.
- North, G. R. and Stevens, M. J.: Detecting climate signals in the surface temperature record, *J. Climate*, 11, 563–577, 1998.
- North, G. R., Bell, T. L., Cahalan, R. F., and Loeng, F. J.: Sampling errors in the estimation of empirical orthogonal functions, *Mon. Weather Rev.*, 110, 600–606, 1982.
- Poloni, F. and Sbrana, G.: Feasible generalized least squares estimation of multivariate GARCH(1,1) models, *J. Multivar. Anal.*, 129, 151–159, 2014.
- Ribes, A. and Terray, L.: Regularised Optimal Fingerprint for attribution. Part II: application to global near-surface temperature based on CMIP5 simulations, *Clim. Dynam.*, 41, 2837–2853, <https://doi.org/10.1007/s00382-013-1736-6>, 2012.
- Ribes, A., Planton, S., and Terray, L.: Regularised Optimal Fingerprint for attribution. Part I: method, properties and idealised analysis, *Clim. Dynam.*, 41, 2817–2836, <https://doi.org/10.1007/s00382-013-1735-7>, 2012.
- Zhang, X., Zwiers, F. W., Hegerl, G. C., Lambert, F. H., Gillett, N. P., Solomon, S., Stott, P. A., and Nozawa, T.: Detection of human influence on twentieth-century precipitation trends, *Nature*, 448, 461–465, 2007.