

# Integrated Optimal Fingerprinting: Method Description and Illustration

ALEXIS HANNART

*Institut Franco-Argentin d'Etudes du Climat et ses Impacts, CNRS, Buenos Aires, Argentina*

(Manuscript received 7 February 2014, in final form 13 August 2015)

## ABSTRACT

The present paper introduces and illustrates methodological developments intended for so-called optimal fingerprinting methods, which are of frequent use in detection and attribution studies. These methods used to involve three independent steps: preliminary reduction of the dimension of the data, estimation of the covariance associated to internal climate variability, and, finally, linear regression inference with associated uncertainty assessment. It is argued that such a compartmentalized treatment presents several issues; an integrated method is thus introduced to address them. The suggested approach is based on a single-piece statistical model that represents both linear regression and control runs. The unknown covariance is treated as a nuisance parameter that is eliminated by integration. This allows for the introduction of regularization assumptions. Point estimates and confidence intervals follow from the integrated likelihood. Further, it is shown that preliminary dimension reduction is not required for implementability and that computational issues associated to using the raw, high-dimensional, spatiotemporal data can be resolved quite easily. Results on simulated data show improved performance compared to existing methods w.r.t. both estimation error and accuracy of confidence intervals and also highlight the need for further improvements regarding the latter. The method is illustrated on twentieth-century precipitation and surface temperature, suggesting a potentially high informational benefit of using the raw, nondimension-reduced data in detection and attribution (D&A), provided model error is appropriately built into the inference.

## 1. Context and motivations

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 [IPCC's Fourth Assessment Report (Hegerl et al. 2007a)]. From a methodological standpoint, D&A studies involve a range of methods (e.g., correlation comparison, regression-based decomposition, and Bayesian inference) but are frequently based on linear regression methods referred to as optimal fingerprinting, whereby an observed climate change  $\mathbf{y}$  is regarded as a linear combination of  $p$  externally forced signals  $\mathbf{x}_1, \dots, \mathbf{x}_p$  added to an internal climate variability term  $\boldsymbol{\nu}$  (Hegerl and Zwiers 2011). The regressors  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_p)$ —or fingerprints—consist here

of spatial, temporal, or space–time patterns of response to external forcings as anticipated by one or several climate models, whereas internal climate variability  $\varepsilon$  is treated as a multivariate Gaussian noise with mean zero and covariance  $\boldsymbol{\Sigma}$ . Practically, the results of the inference and, in particular, the magnitude of the uncertainty range on the regression coefficients  $\boldsymbol{\beta}$ , 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 by using a suite of multivariate linear regression models that were progressively adapted. The standard general linear model and its usual inference treatment [i.e., ordinary least squares (OLS)] was first introduced explicitly by Hegerl et al. (1996) and was then refined by Allen and Tett (1999):

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

The expression of the maximum likelihood estimator (MLE) of  $\boldsymbol{\beta}$  in this model is given by

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

Corresponding author address: Alexis Hannart, IFAECI, CIMA, Ciudad Universitaria, Pab. II, Piso 2, Buenos Aires 1428, Argentina.  
E-mail: alexis.hannart@cima.fcen.uba.ar

where the so-called *precision matrix*, defined as the inverse  $\Sigma^{-1}$  of the covariance  $\Sigma$ , is here assumed to be known. In Allen and Stott (2003), a further advance was introduced by recognizing that the true forcing responses  $\mathbf{x}^*$  are actually not known to certainty. Instead, we have access to  $\mathbf{x}$ , which is equal to  $\mathbf{x}^*$  plus a remnant noise  $\boldsymbol{\eta}$  originating from the internal variability, which is present in climate model simulations even after averaging:

$$\mathbf{y} = \mathbf{x}\boldsymbol{\beta} + \boldsymbol{\nu} \quad \text{and} \quad \mathbf{x} = \mathbf{x}^* + \boldsymbol{\eta}. \quad (3)$$

The column vectors of  $\boldsymbol{\eta}$  are independent and identically distributed (iid) according to a centered multivariate Gaussian with covariance  $\Sigma/m$ , where  $m$  is the size of the ensemble. The so-called total least squares (TLS) inference scheme used in this case (Van Huffel and Vandewalle 1994) also assumes the precision matrix to be known and yields a closed-form expression for  $\boldsymbol{\beta}$  as a function  $\mathbf{x}$ ,  $\mathbf{y}$ , and  $\Sigma^{-1}$ , which is a generalization of Eq. (2) and is detailed further in section 2.

In any of the above two linear regression inferences, the knowledge of the covariance matrix  $\Sigma$  associated to internal climate variability and of the precision matrix  $\Sigma^{-1}$  is critical for inference on  $\boldsymbol{\beta}$ . In optimal fingerprinting,  $\Sigma^{-1}$  is evaluated prior to—and independently from—the regression procedure, based on a sample of  $r$  simulated, unforced variations  $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_r)$ , are referred to as “control runs.” A trivial option to estimate  $\Sigma$  out of  $\boldsymbol{\varepsilon}$  would be to use the so-called *empirical covariance*:

$$\hat{\Sigma} = \frac{1}{r} \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}' = \mathbf{S}, \quad (4)$$

but this estimate is known to perform poorly here, because control runs are costly, and thus the size of the sample  $r$  is often small as compared to the dimension of the data  $n$ . Further,  $\mathbf{S}$  has rank  $r$  and is thus noninvertible when  $r < n$ , impeding a straightforward inversion of  $\mathbf{S}$  into an estimator of the precision matrix. The classic and perhaps most common solution to this problem in D&A to date was proposed by Hegerl et al. (1996) and consists in using a truncated pseudoinverse of  $\mathbf{S}$  as an estimator of the precision matrix, defined by

$$\hat{\Sigma}^{-1} = \mathbf{P}_k \mathbf{D}_k^{-1} \mathbf{P}_k', \quad (5)$$

where  $\mathbf{P} \mathbf{D} \mathbf{P}'$  is the diagonalized factorization of  $\mathbf{S}$  with  $\mathbf{P}$  its matrix of eigenvectors; where  $\mathbf{D} = \text{diag}(d_1, \dots, d_r, 0, \dots, 0)$  with  $d_1 > \dots > d_r > 0$  is its matrix of eigenvalues; and where only the  $k$  leading eigenvectors of the empirical covariance  $\mathbf{S}$  are retained ( $k \leq r$ ). Note that in Eq. (5), as in the remainder of this article,  $\hat{\Sigma}^{-1}$  denotes an estimator of the precision matrix  $\Sigma^{-1}$ , as opposed to the inverse of an estimator of the covariance matrix  $\Sigma$ . This

approach is generally referred to as *covariance truncation*. It is straightforward to show that, from the perspective of linear regression, covariance truncation is equivalent to projecting the data onto the  $k$  leading eigenvectors of  $\mathbf{S}$  and then to scaling each obtained new coordinate based on the inverse square root of the corresponding eigenvalue. The choice of an optimal truncation order  $k$  in this approach is recognized to be difficult and remains partly arbitrary thus far but has been shown to significantly influence inference results. This motivated Ribes et al. (2009, 2012) and Ribes and Terray (2012) to estimate  $\Sigma$  and  $\Sigma^{-1}$  based on a more recent estimator, the so-called Ledoit and Wolf (LW) estimator (Ledoit and Wolf 2004). The latter consists of a linear shrinkage toward a target matrix proportional to the identity with an optimal shrinkage intensity defined in order to minimize the quadratic error of estimation of  $\Sigma$ . It was initially introduced in the context of an application in finance (Ledoit and Wolf 2003) and has been successfully used for numerous and diverse applications since then. It is defined by

$$\hat{\Sigma} = \alpha \lambda \mathbf{I} + (1 - \alpha) \mathbf{S}, \quad \text{and} \quad \hat{\Sigma}^{-1} = [\alpha \lambda \mathbf{I} + (1 - \alpha) \mathbf{S}]^{-1}, \quad (6)$$

where  $\lambda = \text{Tr}(\mathbf{S})/n$  and  $\alpha = \sum_{i=1}^r \|\mathbf{S} - \varepsilon_i \varepsilon_i'\|^2 / r^2 [\text{Tr}(\mathbf{S}^2) - \text{Tr}^2(\mathbf{S})/n] > 0$ . The estimator  $\hat{\Sigma}^{-1}$  of the precision matrix can this time be obtained by straightforward inversion of  $\hat{\Sigma}$ , which is now invertible by construction. It is straightforward to show that this approach is equivalent to regularizing the eigenvalues ( $d_1, \dots, d_n$ ) of  $\mathbf{S}$  by scaling them toward their mean  $\lambda$  [i.e.,  $\tilde{d}_i = (1 - \alpha)d_i + \lambda\alpha$ ] while keeping the eigenvectors unchanged.

For uncertainty analysis, the confidence intervals around  $\hat{\boldsymbol{\beta}}$  that are readily available in the above regression models under a known, fixed covariance [e.g.,  $\mathbf{V}(\hat{\boldsymbol{\beta}}) = (\mathbf{x}' \hat{\Sigma}^{-1} \mathbf{x})^{-1}$  in the OLS model] would be an option. However, it is usually recognized that since  $\Sigma$  is not actually a known constant but an estimate, applying such formulas straightforwardly may yield erroneous uncertainty assessment. To circumvent this issue, in any two above approaches, control runs are split beforehand into two samples of size  $r_1$  and  $r_2$ , respectively, where  $r_1$  and  $r_2$  are arbitrarily chosen under the constraint  $r_1 + r_2 = r$ . Equations (5) or (6) are then applied on both samples to derive two independent estimates of  $\Sigma$  and  $\Sigma^{-1}$ . The first estimate is used for inference of the value of  $\boldsymbol{\beta}$ , whereas the second one is used for uncertainty analysis on  $\hat{\boldsymbol{\beta}}$ . For illustration, in an OLS context, this approach yields

$$\hat{\boldsymbol{\beta}} = (\mathbf{x}' \hat{\Sigma}_1^{-1} \mathbf{x})^{-1} (\mathbf{x}' \hat{\Sigma}_1^{-1} \mathbf{y}) \quad \text{and} \quad \hat{\mathbf{V}}(\hat{\boldsymbol{\beta}}) = (\mathbf{x}' \hat{\Sigma}_1^{-1} \mathbf{x})^{-1} (\mathbf{x}' \hat{\Sigma}_1^{-1} \hat{\Sigma}_2 \hat{\Sigma}_1^{-1} \mathbf{x}) (\mathbf{x}' \hat{\Sigma}_1^{-1} \mathbf{x})^{-1}. \quad (7)$$

On the other hand, the dimension  $n$  of raw simulations and observations is usually considered to be too high for a straightforward application of the above procedures to the raw data. This reserve is justified in particular by the fact that covariance and precision matrix estimates are anticipated to perform poorly when  $n$  gets large while  $r$  remains fixed (i.e.,  $r$  is smaller and smaller relative to  $n$ ). This situation leads most authors to opt for an initial dimension reduction step, which is preliminary to the above-described treatments. The reduction of dimension from a few millions to a few hundreds typically is performed based on a variety of methods. Projecting the raw data onto an arbitrarily chosen number of leading spherical harmonics can be viewed as the most commonplace method thus far for global studies. The implications of the choices related to the reduction method on the reliability of inferences are not well understood in general, although Ribes et al. (2012) have begun to study some aspects of the quality of inferences.

To summarize, usual optimal fingerprinting procedures can be described generically by the following four steps:

- 1) Reduction of the dimension of the raw data  $\mathbf{x}$ ,  $\mathbf{y}$ , and  $\boldsymbol{\varepsilon}$ .
- 2) Split of  $\boldsymbol{\varepsilon}$  into two subsamples; derivation of two independent estimates of the covariance and precision matrices.
- 3) Inference of the value of  $\boldsymbol{\beta}$  based on the first estimate.
- 4) Inference of the confidence intervals around  $\hat{\boldsymbol{\beta}}$  based on the second estimate.

While the approaches mentioned above have successfully produced numerous D&A results thus far [e.g., Stott et al. 2004; Zhang et al. 2007; Hegerl et al. 2007; Min et al. 2008, 2011], it is argued that they share an important shortcoming at a structural level, thus creating an opportunity for potential improvement. This shortcoming lies in our view in the sequential and disconnected nature of the treatment: the four above-mentioned steps certainly have their logic with respect to the end purpose of performing inference on  $\boldsymbol{\beta}$ , but they are disconnected from one another in a rather troublesome way from a theoretical perspective (Fig. 1a). Indeed, the theorems usually invoked in statistics to guarantee the theoretical soundness of an inferential treatment—whether w.r.t. the performance of the resulting estimators or on the analysis of their uncertainty—are, in general, established and hold only for single statistical models. They do not hold a priori for sequential application of separate models and estimators, unless validity could be established explicitly. The reason for this is quite intuitive: by imposing independent, nonintegrated steps in which a given step's outcome is the input for the next step, it makes it quite difficult to keep track of the flow of uncertainty as well as

potential biases between quantities of interest within a rigorous mathematical framework. There are thus few guarantees that the above estimators are optimal, nor that the confidence intervals are actually correct—the latter aspect being of particular concern in a D&A context, because the level of certainty accompanying most statements delivered by D&A studies is, in itself, an important component thereof.

The objective of this paper is to propose an approach to the D&A problem that avoids a sequential analysis approach and thereby attempts to improve the quantification of uncertainty and the reliability of inference. The general features of our proposal are presented in section 2a; section 2b formalizes the approach in detail for the most simple situation of interest, that of the linear regression model of Eq. (1); section 2c explains how and why the approach can be extended almost identically to the more complex errors-in-variables model of Eq. (3); section 2d briefly treats the case wherein, for each individual forcing, the full ensemble of available runs, instead of the single ensemble-average run, is used for inference. Next, section 3 shows that, by virtue of a classic linear algebra simplification, dealing with high-dimensional variables in the context of this model does not imply greater computational complexity than the dimension reduction initial step in the existing sequential approach. Section 4 shows results of the procedure on simulated data and compares its performance to existing ones. Section 5 applies the procedure on twentieth-century data. Section 6 discusses results and concludes.

## 2. Integrated optimal fingerprinting models and inference

### a. General approach

Our proposal consists in restoring unity in the structural formulation underpinning optimal fingerprinting methods—hence the proposed name of “integrated optimal fingerprinting.” It is argued that optimal fingerprinting can indeed be formulated and treated quite easily within a single statistical model (Fig. 1b). In the latter model, all data to be used for inference are dealt with simultaneously: simulated response runs  $\mathbf{x}$ , control runs  $\boldsymbol{\varepsilon}$ , and climate observations  $\mathbf{y}$ —all of them in their raw, high-dimensional form—share the same status as observations from a statistical modeling perspective. On the other hand, the model deals similarly with the unknown quantities (i.e., the coefficients  $\boldsymbol{\beta}$ , the covariance  $\boldsymbol{\Sigma}$ , and the smooth responses  $\mathbf{x}^*$ ), which are all treated as parameters. Nevertheless, we introduce a key distinction between those parameters by recognizing that in the present linear regression context, the end purpose is to infer the coefficients  $\boldsymbol{\beta}$ , not the covariance  $\boldsymbol{\Sigma}$ , nor the

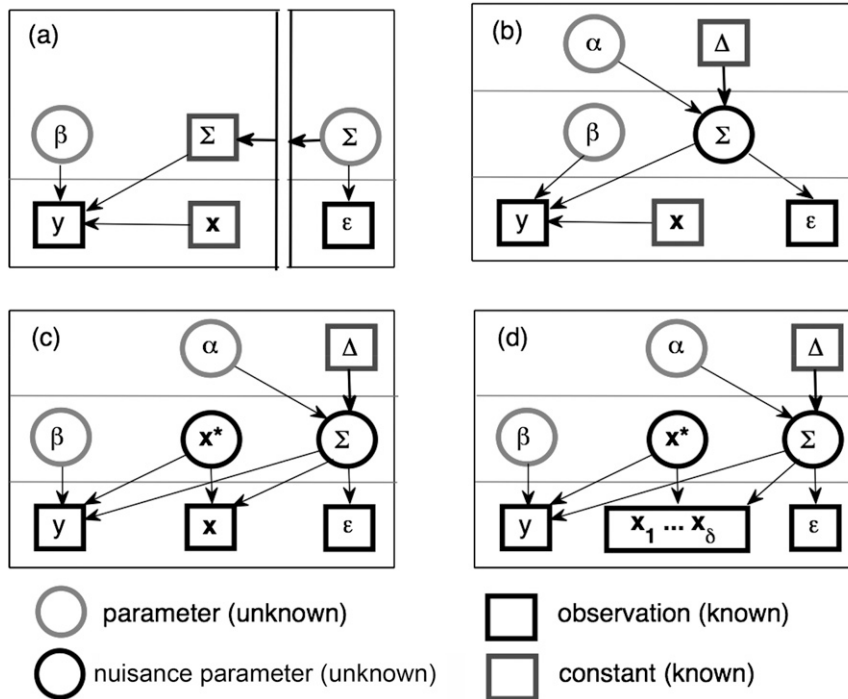


FIG. 1. Structural charts of the described statistical models, showing the status of each quantity (squares–circles) and their conditional dependence (arrows). (a) Sequential approach used traditionally in optimal fingerprinting. (b) Integrated approach, OLS model. (c) Integrated approach, TLS model. (d) Integrated approach, TLS model, all runs.

deterministic responses to forcings  $\mathbf{x}^*$ . The latter two parameters are thus treated here as nuisance parameters (i.e., parameters that are not of immediate interest but must still be accounted for in inference).

The standard way to deal with nuisance parameters in statistics is first to eliminate them from the complete likelihood of the model before proceeding as usual with likelihood maximization and uncertainty analysis on the remaining parameters of interest. Eliminating nuisance parameters has long been a central problem in statistical inference and has been extensively studied in many different contexts (Basu 1977). There exist several approaches for elimination. A first strategy consists in deriving the so-called *profile* likelihood by maximizing the complete likelihood function over the entire domain of values of the nuisance parameter for every value of the parameter of interest. However, this straightforward elimination procedure has its limitations: the complete likelihood may, for instance, not be bounded, in which case a profile likelihood cannot be obtained. A second strategy, which may, in particular, be suitable when such an issue arises, consists of integrating out the nuisance parameters in the complete likelihood instead of maximizing them out, in order to obtain the so-called *integrated* likelihood. This approach is described extensively, for

instance, in Berger et al. (1999). As discussed therein, a uniform integration of the complete likelihood over the entire domain is most straightforward, but an interesting alternative option consists in introducing a prior distribution on the nuisance parameters in order to constrain their values beforehand, which allows for the introduction of some regularization. If any, the parameters of the prior distribution may themselves be subsequently evaluated by maximizing the integrated likelihood. In this case, eliminating the nuisance parameters comes at the cost of introducing new parameters associated to the prior. Yet this trade-off may be beneficial and may reduce the number of degrees of freedom in case the dimension of the former exceeds that of the latter, as will be shown below.

Let us consider for illustration the most frequently used linear regression model with iid residuals, which corresponds to the above setting of Eq. (1) under the particular case  $\Sigma = \sigma^2 \mathbf{I}$ , where  $\sigma^2$  is unknown. This model is classically solved by treating the unknown variance of errors as a nuisance parameter and by eliminating it from the likelihood function following the first strategy above: the complete likelihood  $\ell_c(\boldsymbol{\beta}, \sigma^2)$  is maximized in  $\sigma^2$  for any  $\boldsymbol{\beta}$ , thus obtaining a point estimate  $\hat{\sigma}_\beta^2 = (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{x}\boldsymbol{\beta})/n$  and the profile likelihood  $\ell(\boldsymbol{\beta}) = \ell_c(\boldsymbol{\beta}, \hat{\sigma}_\beta^2)$ . The latter profile likelihood  $\ell(\boldsymbol{\beta})$  is

subsequently maximized in  $\beta$  to obtain a point estimate, and profile likelihood ratios are derived for uncertainty analysis on  $\beta$ . But in the more general case of interest here, where  $\Sigma$  is entirely unknown, the matrix  $\hat{\Sigma}_\beta = (\mathbf{y} - \mathbf{x}\beta)(\mathbf{y} - \mathbf{x}\beta)'/n$  is not of full rank and is, of course, not an acceptable covariance estimate. As long as  $r < n$ , this problem remains even when the  $r$  control runs are built into the inference. It is thus often not possible to obtain a profile likelihood in this case. Instead, our proposal consists in eliminating the nuisance parameter  $\Sigma$  by integration of the complete likelihood and to introduce a parameterized prior on  $\Sigma$  in doing so. The latter corresponds to a general strategy referred to as shrinkage estimation (e.g., [Chen 1979](#); [Ledoit and Wolf 2003](#); [Hannart and Naveau 2014](#)), which is used in D&A by the regularized optimal fingerprinting (ROF) method of [Ribes et al. \(2009\)](#). The prior aims at weakly constraining the structure of  $\Sigma$  based on a general preset target  $\Delta$ . As discussed in [Hannart and Naveau \(2014\)](#) and in the above references on shrinkage estimation, the latter constraint is justified by, and aims to compensate for, the huge informational imbalance that exists between an  $n \times n$  unknown matrix to be evaluated out of an  $n \times r$  matrix of known data, but seeks to do so without restrictively imposing a predefined structure on  $\Sigma$ . Instead, an unspecified degree of resemblance between  $\Sigma$  and the predefined structure  $\Delta$  is merely assumed by using an appropriate prior on  $\Sigma$ . The level of resemblance is parameterized with  $\alpha \in [0, 1]$ , and its value is inferred from observations. Inference can then be handled by standard maximization of the integrated likelihood, which here yields closed formula for estimators and confidence intervals that are easy and fast to implement.

### b. Ordinary least squares model

As sketched above, the proposed model treats similarly  $\mathbf{y}$  and  $\varepsilon$  as observations;  $\beta$  and  $\alpha$  as unknown parameters, the values of which need to be estimated;  $\Sigma$  as a nuisance parameter that influences all observations but the estimation of which is not of interest per se; and  $\mathbf{x}$  and  $\Delta$  as known, fixed constants. This integrated formulation is represented in the structural chart of [Fig. 1b](#), in contrast to the usual sequential formulation shown in [Fig. 1a](#). As a starting point, the joint distribution of observations  $(\mathbf{y}, \varepsilon)$  is derived. The former can be factorized into

$$p(\mathbf{y}, \varepsilon | \Sigma, \beta) = p(\mathbf{y} | \Sigma, \beta) \times p(\varepsilon | \Sigma). \quad (8)$$

Next, both individual terms in the right-hand side of Eq. (8) are derived. First, the structural assumption inherent to the linear regression model of Eq. (1) implies

$$p(\mathbf{y} | \Sigma, \beta) = 2\pi^{-n/2} |\Sigma|^{-1/2} \exp \left[ -\frac{1}{2} (\mathbf{y} - \mathbf{x}\beta)' \Sigma^{-1} (\mathbf{y} - \mathbf{x}\beta) \right]. \quad (9)$$

On the other hand,  $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_r)$  is assumed to be an iid sample of vectors distributed according to a multivariate Gaussian PDF with mean zero and covariance matrix  $\Sigma$ . The probability density function (PDF) of  $\varepsilon$  is thus

$$p(\varepsilon | \Sigma) = 2\pi^{-nr/2} |\Sigma|^{-r/2} \exp \left[ -\frac{1}{2} \text{Tr}(\mathbf{r}\Sigma^{-1}) \right], \quad (10)$$

where  $\mathbf{S} = \varepsilon\varepsilon'/r$  is the empirical covariance matrix. Combining Eqs. (9) and (10), the complete likelihood  $\ell_c(\beta, \Sigma)$  of the model is

$$\ell_c(\beta, \Sigma) = 2\pi^{-[n(r+1)/2]} |\Sigma|^{-[(r+1)/2]} \exp \left[ -\frac{1}{2} (\mathbf{y} - \mathbf{x}\beta)' \Sigma^{-1} (\mathbf{y} - \mathbf{x}\beta) - \frac{1}{2} \text{Tr}(\mathbf{r}\Sigma^{-1}) \right]. \quad (11)$$

The conjugate distribution on  $\Sigma$ , the so-called inverse Wishart distribution, is chosen for convenience as the prior distribution on  $\Sigma$ :

$$\begin{aligned} p(\Sigma | \xi, \Omega) \\ = 2^{-(\xi n/2)} \Gamma_n \left( \frac{\xi}{2} \right)^{-1} |\Omega|^{\xi/2} |\Sigma|^{-[(\xi+n+1)/2]} \exp \left[ -\frac{1}{2} \text{Tr}(\Omega \Sigma^{-1}) \right] \end{aligned} \quad (12)$$

where  $\xi > n + 1$  and  $\Omega \in \mathcal{S}^{+*}$  are two hyperparameters, and  $\Gamma_n(\cdot)$  is the multivariate Gamma function. For further convenience, a different parameterization of the above inverse Wishart distribution is introduced by using  $(\alpha, \Delta)$  instead of  $(\xi, \Omega)$ , where metaparameter  $\Delta = \Omega/(\xi - n - 1)$  is more directly interpretable as the

mean of the distribution, whereas, on the other hand, metaparameter  $\alpha = (\xi - n - 1)/(r + \xi - n - 1)$  drives the variance of the elements of  $\Sigma$  with variance zero for  $\alpha = 1$  and infinite variance for  $\alpha = 0$  ([appendix A](#)).

Next, the nuisance parameter  $\Sigma$  is integrated out in Eq. (8) to obtain the integrated likelihood  $\ell(\beta, \alpha)$ , which no longer depends on  $\Sigma$  but only on the parameter of interest  $\beta$  and on the regularization parameter  $\alpha$ :

$$\ell(\beta, \alpha) = \int_{\Sigma} \ell_c(\beta, \Sigma) p(\Sigma | \alpha) d\Sigma. \quad (13)$$

This integration can be performed exactly because of conjugacy ([appendix B](#)). The following expression is obtained after some algebra:



$-2\log\ell(\boldsymbol{\beta}, \alpha)$

$$= \left( \frac{r}{1-\alpha} + n + 2 \right) \log \left[ 1 + \frac{1-\alpha}{r} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \boldsymbol{\Sigma}_\alpha^{-1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta}) \right] \quad (14)$$

up to a constant that does not depend on  $\boldsymbol{\beta}$ , which will be clarified below, and where the covariance matrix  $\boldsymbol{\Sigma}_\alpha$ , defined by

$$\boldsymbol{\Sigma}_\alpha = \alpha \boldsymbol{\Delta} + (1 - \alpha) \mathbf{S}, \quad (15)$$

has been introduced. Even if estimating  $\boldsymbol{\Sigma}$  is not of primary interest here, the matrix  $\boldsymbol{\Sigma}_\alpha$  is an intermediate

quantity interpretable as a linear shrinkage estimator of the covariance, where  $\alpha$  is the shrinkage intensity and  $\boldsymbol{\Delta}$  is the shrinkage target. By maximizing the likelihood of Eq. (14) in  $\boldsymbol{\beta}$  and  $\alpha$ , the desired MLE of  $\boldsymbol{\beta}$  is obtained:

$$\hat{\alpha} = \arg\max_{\alpha \in [0,1]} [\log\ell(\alpha)] \quad \text{and} \quad \hat{\boldsymbol{\beta}} = (\mathbf{x}' \boldsymbol{\Sigma}_{\hat{\alpha}}^{-1} \mathbf{x})^{-1} (\mathbf{x}' \boldsymbol{\Sigma}_{\hat{\alpha}}^{-1} \mathbf{y}), \quad (16)$$

where the expression of the integrated likelihood  $\ell(\alpha)$  is given by

$$\begin{aligned} -2\log\ell(\alpha) = & \phi\left(\frac{r}{1-\alpha} + 1\right) - \phi\left(\frac{\alpha r}{1-\alpha}\right) - n\left(\frac{r}{1-\alpha} + n + 2\right) \log\left(\frac{1-\alpha}{r} + 1\right) \\ & + \left(\frac{r}{1-\alpha} + n + 2\right) \log|\boldsymbol{\Sigma}_\alpha| - \left(\frac{\alpha r}{1-\alpha} + n + 1\right) \log|\boldsymbol{\Delta}| \\ & + \left(\frac{r}{1-\alpha} + n + 2\right) \log\left[1 + \frac{(1-\alpha)n}{r} F_\alpha\right]. \end{aligned} \quad (17)$$

For clarity, the function and the intermediate quantities  $F - F_\alpha = (1/n) (\mathbf{y} - \mathbf{x}\boldsymbol{\beta}_\alpha)' \boldsymbol{\Sigma}_\alpha^{-1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta}_\alpha)$ —and  $\boldsymbol{\beta}_\alpha = (\mathbf{x}' \boldsymbol{\Sigma}_\alpha^{-1} \mathbf{x})^{-1} (\mathbf{x}' \boldsymbol{\Sigma}_\alpha^{-1} \mathbf{y})$  are introduced in Eq. (17). The maximization of  $\ell(\alpha)$  for  $\alpha \in [0, 1]$  based on the closed-form expression in Eq. (17) may be obtained by application of any maximization procedure [for instance, the Newton–Raphson algorithm (Press et al. 2007)].

It should be noted that the expression of  $\hat{\boldsymbol{\beta}}$  in Eq. (16) is identical to that of the standard OLS estimator under a model in which the estimated covariance  $\boldsymbol{\Sigma}_\alpha$  is treated as if it was known and constant (i.e., sequential approach). It is also similar to the expression of the estimator of Ribes et al. (2012) under the special case  $\boldsymbol{\Delta} = \text{Tr}(\mathbf{S})/n\mathbf{I}$  and  $\alpha_{\text{lw}} = \hat{\alpha}$ . Nevertheless, the similarity with Ribes et al. (2012) is limited to the aspect of the expression in Eq. (16). Indeed, in practice, as will be shown in section 4, the obtained values of  $\alpha_{\text{lw}}$  and  $\hat{\alpha}$  differ, yielding different estimates of  $\boldsymbol{\beta}$  even in the special case  $\boldsymbol{\Delta} = \text{Tr}(\mathbf{S})/n\mathbf{I}$ . Further, Eq. (16) is applicable with any target  $\boldsymbol{\Delta}$  (i.e., not restrictively to a target proportional to the identity). Finally, the confidence intervals obtained in our approach differ substantially, in both theory and practice, from those obtained under the approach of Ribes et al. (2012), as is now detailed.

In the present integrated model, the usual profile likelihood-based method of Venzon and Moolgavkar (1988) can be applied straightforwardly to derive confidence intervals around  $\hat{\boldsymbol{\beta}}$ . For a given confidence level  $\eta$  (0.95, say) the former approach consists in solving  $2\log\ell(\boldsymbol{\beta}, \hat{\alpha}) - 2\log\ell(\hat{\boldsymbol{\beta}}, \hat{\alpha}) = c_\eta$  in  $\boldsymbol{\beta}$ . After some algebra, it can be shown to be equivalent to solving in

$\boldsymbol{\beta}$ ,  $(\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \boldsymbol{\Sigma}_{\hat{\alpha}}^{-1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta}) = c_\eta^* - nF_{\hat{\alpha}}$ , where the transformed confidence threshold level  $c_\eta^*$  is defined by

$$\begin{aligned} c_\eta^* = & \left\{ \exp\left[ c_\eta / \left( \frac{r}{1-\hat{\alpha}} + n + 2 \right) \right] - 1 \right\} \left( \frac{r}{1-\hat{\alpha}} + nF_{\hat{\alpha}} \right) \\ & \simeq c_\eta \left[ 1 + \frac{n}{r} (F_{\hat{\alpha}} - 1)(1 - \hat{\alpha}) \right], \end{aligned} \quad (18)$$

where the approximation holds for  $n$  large. Under a model in which the estimated covariance  $\boldsymbol{\Sigma}_\alpha$  is treated as if it were known and constant [i.e., sequential approach, as in Ribes et al. (2012)], one obtains confidence intervals by solving  $(\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \boldsymbol{\Sigma}_\alpha^{-1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta}) = c_\eta - nF_\alpha$ . Confidence intervals obtained under this assumption would therefore differ from those obtained under our approach, since  $c_\eta^* \neq c_\eta$  in general. Further, the relative position of  $c_\eta^*$  and  $c_\eta$  is driven (approximately for large  $n$ ) by the position of  $F_{\hat{\alpha}}$  w.r.t. 1. Therefore, the confidence intervals in our approach will be larger than those obtained by assuming that the estimated covariance is known when the model fit is bad and, conversely, smaller when the model fit is good. In both cases, the gap in confidence intervals between both approaches will be emphasized when  $r/n$  and/or  $\hat{\alpha}$  are small. Figures 2a and 2b show confidence regions obtained under both situations for  $p = 2$ .

### c. Total least squares model

This approach is now extended to the errors in variables introduced by Allen and Stott (2003) and used in

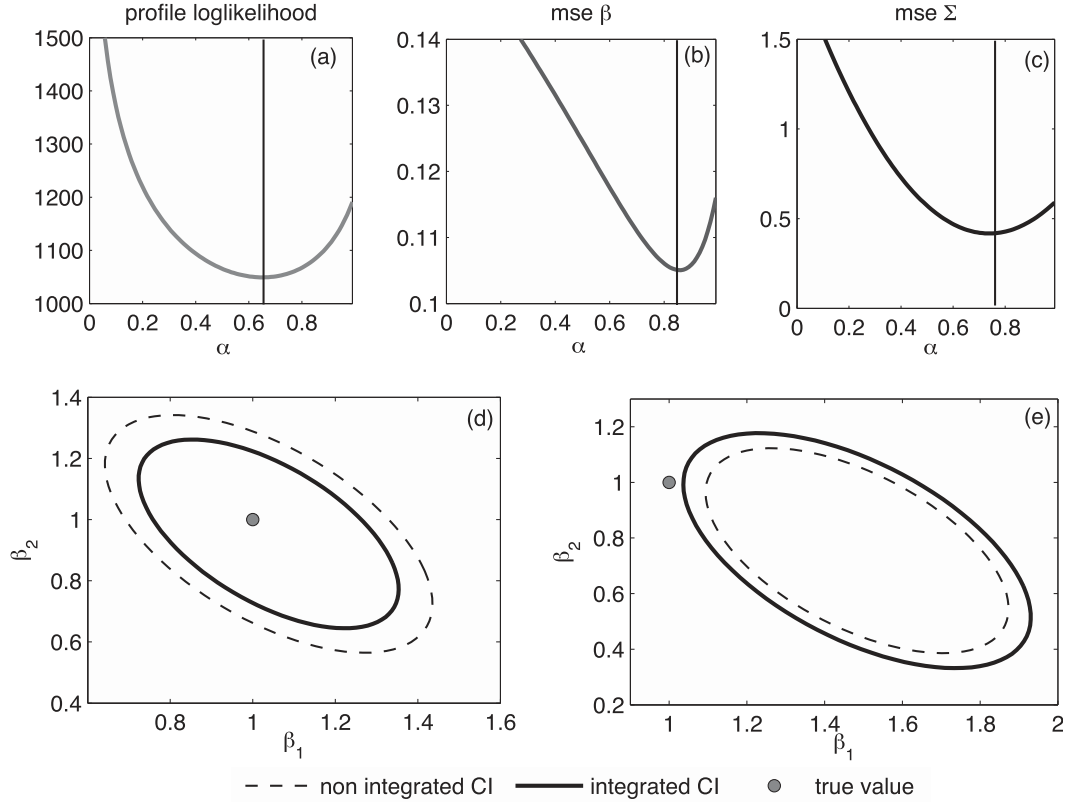


FIG. 2. Illustration of IOF for one single simulation. (a) Profile likelihood  $\ell(\alpha)$  for  $\alpha \in [0, 1]$ . (b) MSE of  $\beta_\alpha$  for  $\alpha \in [0, 1]$ . (c) MSE of  $\Sigma_\alpha$  for  $\alpha \in [0, 1]$ . (d) Actual value of  $\beta$  (dot), 90% confidence region under IOF (thick line), and under the assumption that  $\Sigma_\alpha$  is a constant (dotted line). (e) As in (d), but under a different simulation.

various studies (Lott et al. 2013), which is often called the total least squares model. The following PDF prevails:

$$p(\mathbf{y}, \mathbf{x}, \varepsilon | \Sigma, \mathbf{x}^*, \beta) = p(\mathbf{y} | \Sigma, \beta, \mathbf{x}^*) \times p(\mathbf{x} | \mathbf{x}^*, \Sigma) \times p(\varepsilon | \Sigma), \quad (19)$$

where  $\mathbf{x}$  is now treated as an observed version of the actual design matrix  $\mathbf{x}^*$  that contains noise. In this case,  $\mathbf{x}^*$  is not known; thus, it is treated as a nuisance parameter. This means that

$$p(\mathbf{y} | \Sigma, \beta, \mathbf{x}^*) = 2\pi^{-(n/2)} |\Sigma|^{-(1/2)} \exp \left[ -\frac{1}{2} (\mathbf{y} - \mathbf{x}^* \beta)' \Sigma^{-1} (\mathbf{y} - \mathbf{x}^* \beta) \right], \quad (20)$$

and the PDF of each column vector  $\mathbf{x}_i$  of  $\mathbf{x}$  is assumed to be distributed according to a Gaussian PDF with mean  $\mathbf{x}_i^*$  and variance  $\Sigma/m$ , where  $m \geq 1$  denotes the size of the ensemble that was averaged to obtain each  $\mathbf{x}_i$ :

$$p(\mathbf{x} | \mathbf{x}^*, \Sigma) = \prod_{i=1}^p 2\pi m^{-(n/2)} |\Sigma|^{-(1/2)} \exp \left[ -\frac{1}{2} m (\mathbf{x}_i - \mathbf{x}_i^*)' \Sigma^{-1} (\mathbf{x}_i - \mathbf{x}_i^*) \right]. \quad (21)$$

The same sequence as in model (8) is then followed in order to obtain an integrated likelihood in  $\beta$ . This means that  $\Sigma$  is first integrated out in Eq. (20) in order to obtain a marginalized likelihood. Next, it is concentrated in  $\beta$  and  $\alpha$  by maximizing in  $\mathbf{x}^*$ , as would be done in the standard model with known covariance. In contrast with  $\Sigma$ , the nuisance parameter  $\mathbf{x}^*$  is thus not

integrated out: it is instead eliminated by mean of maximization to obtain the so-called profile likelihood (Basu 1977). In applying this sequence, the integration in  $\Sigma$  can be performed exactly, as well as the maximization in  $\mathbf{x}^*$  after some algebra (appendix C). These two steps yield an expression for the likelihood  $\ell(\beta, \alpha)$  that is quite similar to Eq. (14):

$$-2\log\ell(\boldsymbol{\beta}, \alpha) = \left(\frac{r}{1-\alpha} + p + n + 2\right) \times \log \left[1 + \lambda_{\boldsymbol{\beta}} \frac{1-\alpha}{r} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \boldsymbol{\Sigma}_{\alpha}^{-1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})\right] \quad (22)$$

up to a constant that depends only on  $\alpha$ , and where the scaling factor  $\lambda_{\boldsymbol{\beta}} = (1 + \boldsymbol{\beta}'\boldsymbol{\beta}/m)^{-1}$  has been introduced for convenience. The MLEs of  $\alpha$  and  $\boldsymbol{\beta}$  are this time

$$\hat{\alpha} = \operatorname{argmax}_{\alpha \in [0,1]} [\log\ell(\alpha)] \quad \text{and} \\ \hat{\boldsymbol{\beta}} = \operatorname{argmax}_{\boldsymbol{\beta} \in \mathbb{R}^p} [\lambda_{\boldsymbol{\beta}} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \boldsymbol{\Sigma}_{\hat{\alpha}}^{-1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})] = -\sqrt{m} \mathbf{v}_{\hat{\alpha}}^* / v_{\hat{\alpha}}. \quad (23)$$

Note that  $\hat{\boldsymbol{\beta}}$ , defined to be the value maximizing  $\lambda_{\boldsymbol{\beta}} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \boldsymbol{\Sigma}_{\hat{\alpha}}^{-1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})$ , differs from Eq. (16) because

of the presence of the factor  $\lambda_{\boldsymbol{\beta}}$ , which depends on  $\boldsymbol{\beta}$  and synthesizes the influence of the errors-in-variables assumption on the model likelihood. Yet it also has a convenient closed-form expression,  $\hat{\boldsymbol{\beta}} = -m \mathbf{v}_{\hat{\alpha}}^* / v_{\hat{\alpha}}$ , obtained as the whitened TLS estimator under covariance  $\boldsymbol{\Sigma}_{\hat{\alpha}}$ . It is briefly recalled for completeness that, for any  $\alpha \in [0, 1]$ ,  $\mathbf{v}_{\alpha}^*$  is a  $p$ -dimensional vector, and  $v_{\alpha}$  is a scalar such that the  $(p+1)$ -dimensional vector  $[\mathbf{v}_{\alpha}^*, v_{\alpha}]'$  is the unitary eigenvector associated to the smallest eigenvalue of the  $(p+1) \times (p+1)$  positive definite matrix  $(\sqrt{m} \mathbf{x} \mathbf{y})' \boldsymbol{\Sigma}_{\alpha}^{-1} (\sqrt{m} \mathbf{x} \mathbf{y})$ . Therefore, as in section 2a, the expression for  $\boldsymbol{\beta}$  matches exactly that obtained under a TLS model in which the estimated covariance  $\boldsymbol{\Sigma}_{\hat{\alpha}}$  would be treated as if it were known and constant (i.e., sequential approach).

Finally, the expression for the concentrated likelihood  $\ell(\alpha)$  is given by

$$\begin{aligned} -2\log\ell(\alpha) &= \phi\left(\frac{r}{1-\alpha} + p + 1\right) - \phi\left(\frac{\alpha r}{1-\alpha}\right) - n\left(\frac{r}{1-\alpha} + n + p + 2\right) \log\left[\frac{1-\alpha}{r}(p+1) + 1\right] \\ &\quad + \left(\frac{r}{1-\alpha} + n + p + 2\right) \log|\boldsymbol{\Sigma}_{\alpha}| - \left(\frac{\alpha r}{1-\alpha} + n + 1\right) \log|\boldsymbol{\Delta}| \\ &\quad + \left(\frac{r}{1-\alpha} + n + p + 2\right) \log\left[1 + \lambda_{\alpha} \frac{(1-\alpha)n}{r} F_{\alpha}\right], \end{aligned} \quad (24)$$

where  $\boldsymbol{\beta}_{\alpha} = -m \mathbf{v}_{\alpha}^* / v_{\alpha}$  and  $\lambda_{\alpha} = [1 + (\boldsymbol{\beta}_{\alpha}' \boldsymbol{\beta}_{\alpha} / m)]^{-1}$ . The concentrated likelihood is thus slightly modified as compared to Eq. (17).

#### d. All runs total least squares model

In practice, matrix  $\mathbf{x}$  in the above model was obtained by averaging an ensemble of  $m$  matrices of simulated responses  $(\mathbf{x}_1, \dots, \mathbf{x}_m)$  (i.e.,  $\mathbf{x} = \sum_{j=1}^m \mathbf{x}_j / m$ ). Such averaging reduces variance by a factor of  $m$ , but some worthwhile information regarding  $\boldsymbol{\Sigma}$  is lost in the process. To make the most of this information and integrate it into the inference, the model is reformulated by accounting explicitly for  $(\mathbf{x}_1, \dots, \mathbf{x}_m)$  instead of its average  $\mathbf{x}$ :

$$\begin{aligned} p(\mathbf{y}, \mathbf{x}_1, \dots, \mathbf{x}_m, \boldsymbol{\varepsilon} | \boldsymbol{\Sigma}, \mathbf{x}^*, \boldsymbol{\beta}, \alpha) \\ = p(\mathbf{y} | \boldsymbol{\Sigma}, \boldsymbol{\beta}, \mathbf{x}^*) \times \prod_{j=1}^m p(\mathbf{x}_j | \mathbf{x}^*, \boldsymbol{\Sigma}) \times p(\boldsymbol{\varepsilon} | \boldsymbol{\Sigma}), \end{aligned} \quad (25)$$

where  $\mathbf{x}_1, \dots, \mathbf{x}_m$  are now treated as  $m$  observed, independently noise-contaminated versions of the unknown design matrix  $\mathbf{x}^*$ , and are thus assumed to be iid Gaussian with mean  $\mathbf{x}^*$  and covariance  $\boldsymbol{\Sigma}$ . This means that variations among ensemble members are assumed to be due to internal variability only and that models are assumed to correctly simulate the latter, which are two strong assumptions. After some calculations,  $\prod_{j=1}^m p(\mathbf{x}_j | \mathbf{x}^*, \boldsymbol{\Sigma})$  can be simplified into

$$\begin{aligned} \prod_{j=1}^m p(\mathbf{x}_j | \mathbf{x}^*, \boldsymbol{\Sigma}) &\propto |\boldsymbol{\Sigma}|^{-[p(m-1)/2]} \exp\left[-\frac{1}{2} \operatorname{Tr}(pm \bar{\mathbf{S}} \boldsymbol{\Sigma}^{-1})\right] \\ &\quad \times p(\mathbf{x} | \mathbf{x}^*, \boldsymbol{\Sigma}), \end{aligned} \quad (26)$$

where  $\bar{\mathbf{S}} = \sum_{i=1}^p [\sum_{j=1}^m (\mathbf{x}_{ij} - \mathbf{x}_i)(\mathbf{x}_{ij} - \mathbf{x}_i)' / m] / p$  is the sample covariance matrix of the ensemble of simulated response runs, and  $\mathbf{x}_{ij}$  denotes the  $i$ th column vector of matrix  $\mathbf{x}_j$ .

An interesting consequence of the factorization in Eq. (26) is that our modified model is actually formally identical to the model of section 2b if substituting  $\mathbf{S}$  and  $r$  by  $(r\mathbf{S} + pm\bar{\mathbf{S}}) / (r + pm)$  and  $r + pm$ , respectively. Practically, this means that, after subtraction of the ensemble mean, each obtained response run must be treated identically to an internal variability control run. This simple calculation shows that the new model is equivalent to, and basically boils down to, a potentially significant increase in the size of the control ensemble from  $r$  to  $r + pm$ , which will improve the performance of statistical inference.

### 3. Implementation in large dimension

For small values of  $n$ —say in the order of a few hundred—computational aspects are not an issue: basic routines of a standard software package run on a desktop computer can cope with a straightforward implementation of



Eqs. (14)–(23). But the straightforward approach is no longer suitable for large values of  $n$ , because it would imply inverting and/or diagonalizing  $n \times n$  matrices, thus requiring large computational resources. The purpose of this section is to show that, under a rather general assumption, by merely reformulating our equations through some algebraic manipulation, it is possible to drastically reduce their computational complexity.

The main assumption here is that the inverse square root  $\Delta^{-1/2}$  of the target matrix  $\Delta$ , can be computed at a very low computational cost no matter what  $n$  is. This assumption is not very restrictive because, by definition, a target matrix is a general structure that is likely to have a sufficiently simple mathematical expression (e.g., a diagonalized expression, or see, for instance, section 5) to allow for a simple derivation of its inverse square root. Note that it clearly holds for the default target matrix proportional to the identity matrix. Under this assumption, the following transformed variables are derived as a starting point:

$$\tilde{\mathbf{e}} = \Delta^{-1/2} \mathbf{e}, \quad \tilde{\mathbf{x}} = \Delta^{-1/2} \mathbf{x}, \quad \text{and} \quad \tilde{\mathbf{y}} = \Delta^{-1/2} \mathbf{y}. \quad (27)$$

Next, the following six dimension-reduced statistics are derived:

$$\begin{aligned} \mathbf{S}_{\mathbf{xx}} &= \tilde{\mathbf{x}} \tilde{\mathbf{x}}^T (p \times p), & \mathbf{S}_{\mathbf{xy}} &= \tilde{\mathbf{x}} \tilde{\mathbf{y}}^T (p \times 1), & \mathbf{S}_{\mathbf{yy}} &= \tilde{\mathbf{y}} \tilde{\mathbf{y}}^T (1 \times 1), \\ \mathbf{S}_{\mathbf{ex}} &= \tilde{\mathbf{e}} \tilde{\mathbf{x}}^T (r \times p), & \mathbf{S}_{\mathbf{ey}} &= \tilde{\mathbf{e}} \tilde{\mathbf{y}}^T (r \times 1), & \text{and} \\ \mathbf{S}_{\mathbf{ee}} &= \tilde{\mathbf{e}} \tilde{\mathbf{e}}^T (r \times r). \end{aligned} \quad (28)$$

After implementation of Eqs. (27) and (28), the dimension of the obtained statistics is merely driven by  $r$  and  $p$  but no longer by  $n$ . If our end result could be obtained based on these elementary statistics only, then the large dimensionality issue would be resolved.

At this stage, it is important to note that all the calculations required for the computation of  $\hat{\beta}$  and  $\hat{\alpha}$  can be expressed based on four basic quantities:  $\mathbf{x}' \Sigma_{\alpha}^{-1} \mathbf{x}$ ,  $\mathbf{x}' \Sigma_{\alpha}^{-1} \mathbf{y}$ ,  $\mathbf{y}' \Sigma_{\alpha}^{-1} \mathbf{y}$ , and  $\log |\Sigma_{\alpha}|$  that all depend on  $\alpha$ . Our goal thus boils down to deriving closed expressions for these four quantities as functions of  $\alpha$  and the six elementary statistics of Eq. (28) only. For this purpose, let us consider the matrix  $\Delta^{-1/2} \mathbf{S} \Delta^{-1/2}$  and its diagonal factorization  $\mathbf{P} \mathbf{W} \mathbf{P}'$ , where  $\mathbf{P}$  is orthonormal and  $\mathbf{W}$  is diagonal. Because  $\Delta^{-1/2} \mathbf{S} \Delta^{-1/2} = \tilde{\mathbf{e}} \tilde{\mathbf{e}}' / r$  has rank  $r$ ,  $\mathbf{W}$  has  $r$  positive eigenvalues that are grouped under the  $r \times 1$  vector  $\mathbf{w}$ , with corresponding eigenvectors grouped under the  $n \times r$  matrix  $\mathbf{P}^*$ . Similarly,  $\mathbf{P}_{\perp}^*$  denotes an  $n \times (n-r)$  matrix of unitary eigenvectors corresponding to the eigenvalue zero: that is,  $\mathbf{P} = [\mathbf{P}^* \mathbf{P}_{\perp}^*]$ . A classical simplification for obtaining  $\mathbf{w}$  and  $\mathbf{P}^*$  without diagonalizing the  $n \times n$

matrix  $\tilde{\mathbf{e}} \tilde{\mathbf{e}}' / r$  is to diagonalize instead the  $r \times r$  matrix  $\mathbf{S}_{\mathbf{ee}} / r$ . Indeed, it can immediately be shown that both have the same  $r$  positive eigenvalues and that, if  $\mathbf{Q}$  denotes its  $r \times r$  orthonormal matrix of eigenvectors, then  $\mathbf{P}^* = \tilde{\mathbf{e}} \mathbf{Q}^*$ , with  $\mathbf{Q}^* = \mathbf{Q} \Lambda^{-1/2}$  and  $\Lambda = \mathbf{Q}' \mathbf{S}_{\mathbf{ee}} \mathbf{Q}$  a diagonal matrix used for normalization. On the other hand, the computation of  $\mathbf{P}_{\perp}^*$  would be a daunting task for large  $n$ , but fortunately—as will be clarified in Eq. (29)— $\mathbf{P}^* \mathbf{P}_{\perp}^{*'} is of interest, rather than  $\mathbf{P}_{\perp}^*$ . The former quantity can be obtained easily by  $\mathbf{P}^* \mathbf{P}_{\perp}^{*'} = \mathbf{I} - \mathbf{P}^* \mathbf{P}^{*'}$  as a consequence of the fact that  $\mathbf{P} \mathbf{P}' = \mathbf{I}$ . With the above notation and properties$

$$\begin{aligned} \Delta^{-1/2} \Sigma_{\alpha}^{-1} \Delta^{-1/2} &= \mathbf{P} [\alpha \mathbf{I} + (1 - \alpha) \mathbf{W}]^{-1} \mathbf{P}' = \mathbf{P} \begin{bmatrix} \mathbf{U}_{\alpha} & 0 \\ 0 & \frac{1}{\alpha} \mathbf{I} \end{bmatrix} \mathbf{P}' \\ &= \mathbf{P}^* \mathbf{U}_{\alpha} \mathbf{P}^{*'} + \frac{1}{\alpha} \mathbf{P}_{\perp}^* \mathbf{P}_{\perp}^{*'} = \frac{1}{\alpha} \mathbf{I} + \mathbf{P}^* \mathbf{V}_{\alpha} \mathbf{P}^{*'}, \end{aligned} \quad (29)$$

where  $\mathbf{U}_{\alpha} = \text{diag}\langle\{[\alpha + (1 - \alpha)w_i]^{-1}\}_{i=1}^r\rangle$  and  $\mathbf{V}_{\alpha} = \text{diag}\langle\{[\alpha + (1 - \alpha)w_i]^{-1} - \alpha^{-1}\}_{i=1}^r\rangle$  are two  $r \times r$  diagonal matrices that are fast to compute for any  $\alpha$ . Finally, some algebra yields

$$\mathbf{x}' \Sigma_{\alpha}^{-1} \mathbf{x} = \frac{1}{\alpha} \mathbf{S}_{\mathbf{xx}} + \mathbf{S}_{\mathbf{ex}}' \mathbf{Q}^* \mathbf{V}_{\alpha} \mathbf{Q}^{*'} \mathbf{S}_{\mathbf{ex}}, \quad (30)$$

which is the desired reduced-form expression. The other quantities of interest are obtained similarly:

$$\begin{aligned} \mathbf{x}' \Sigma_{\alpha}^{-1} \mathbf{y} &= \frac{1}{\alpha} \mathbf{S}_{\mathbf{xy}} + \mathbf{S}_{\mathbf{ex}}' \mathbf{Q}^* \mathbf{V}_{\alpha} \mathbf{Q}^{*'} \mathbf{S}_{\mathbf{ey}}, \\ \mathbf{y}' \Sigma_{\alpha}^{-1} \mathbf{y} &= \frac{1}{\alpha} \mathbf{S}_{\mathbf{yy}} + \mathbf{S}_{\mathbf{ey}}' \mathbf{Q}^* \mathbf{V}_{\alpha} \mathbf{Q}^{*'} \mathbf{S}_{\mathbf{ey}}, \quad \text{and} \end{aligned}$$

$$\log |\Sigma_{\alpha}| = (n - r) \log \alpha + \sum_{i=1}^r \log [\alpha + (1 - \alpha)w_i]. \quad (31)$$

Therefore, under a rather general setting and after some initial calculations, the complexity of the procedure is reduced from  $O(n)$  down to  $O(r)$ , thus unlocking the possibility of treating very large datasets on a desktop computer and without any need for a preliminary, independent dimension reduction.

#### 4. Numerical simulations

This section illustrates and evaluates the performance of our integrated procedure by applying it to simulated values of  $\mathbf{x}$ ,  $\mathbf{y}$ , and  $\mathbf{e}$ . The initial use of simulated rather than real data aims at verifying that our inference procedure performs correctly, a goal that requires the actual values of  $\beta$  to be known. It also allows performance comparison with existing procedures.

### a. Description of the setup

Our idealized data simulations are based to a large extent on the numerical experiments setup of Ribes et al. (2012) and aim at realistically replicating a global dataset of twentieth-century temperature. Four pools of experiments were generated by combining the two models introduced in section 2 (i.e., OLS and TLS) together with two distinct sets of assumptions w.r.t. the size and structure of the covariance matrix  $\Sigma$ . Regarding the latter, a first set of assumptions stick exactly to those proposed by Ribes et al. (2012) by using  $n = 275$  and a covariance matrix  $\Sigma$  estimated beforehand by these authors out of various CMIP5 control runs after a preliminary dimension reduction of the data based on spherical harmonics. In contrast to this classical dimension-reduced situation (DR), our second set of assumptions attempts to represent the non-classical situation described in section 3, where a raw, nondimension-reduced dataset (NDR) is treated directly. For this purpose, it is assumed that all variables are obtained on a regular rectangular  $100 \times 100$  spatial grid (i.e.,  $n = 10\,000$  grid points with spatial coordinates  $s_1, \dots, s_n$ ). Note that, differently from the above DR setup, which is spatiotemporal, the temporal dimension of fingerprints is thus not considered in the NDR setup for simplicity; the variables of interest may correspond, say, to temperature change over the twentieth century at every location  $i = 1, \dots, n$ . The covariance matrix  $\Sigma$  is given a spatial structure that is defined in the following way. Consider the isotropic exponential covariance matrix  $\mathbf{E}$  defined by  $E_{ij} = \exp(-\|s_i - s_j\|/d_0)$ , with parameter  $d_0$  such that  $\sum_{i,j} E_{ij}/n = 0.2$ , as well as its diagonalized form. To obtain a spatial structure that is simultaneously less smooth and more general than  $\mathbf{E}$ , its eigenvalues are modified by simply multiplying them using  $n$  i.i.d. random factors taken from the uniform distribution on  $[0.5, 1.5]$ , while maintaining its eigenvectors unchanged. All covariance structures obtained using this simple procedure share with  $\mathbf{E}$  the feature of concentrating a large fraction of the correlation in the vicinity of any given grid point while, at the same time, allowing for a wide variety of shapes that strongly differ from  $\mathbf{E}$ . In particular, in strong contrast with  $\mathbf{E}$ , the covariance matrices  $\Sigma$  derived in this way are nonisotropic (Fig. 3c). They also appear to be able to represent such a complex structure as, for instance, a remote teleconnection dipole (Fig. 3d).

The matrix of unobserved regressors  $\mathbf{x}^*$  under the TLS model was obtained by simulating each of its  $n$  rows independently from a  $p$ -variate Gaussian distribution with fixed correlation  $\rho = 0.3$  and fixed variance  $\sigma_{\mathbf{x}^*}^2$ , where  $\sigma_{\mathbf{x}^*}^2$  is chosen to obtain a realistic signal-to-noise ratio. The matrix of noise-contaminated, observed regressors  $\mathbf{x}$  was then obtained from  $\mathbf{x} = \mathbf{x}^* + \mathbf{v}$  after simulating the

$p$  columns of  $\mathbf{v}$  independently from an  $n$ -variate Gaussian distribution with covariance  $\Sigma/m$ , where the ensemble size  $m$  is assumed to be equal to five. The vector of observations  $\mathbf{y}$  was obtained from  $\mathbf{y} = \mathbf{x}^*\boldsymbol{\beta} + \boldsymbol{\varepsilon}$  after simulating  $\boldsymbol{\varepsilon}$  from an  $n$ -variate Gaussian distribution with covariance  $\Sigma$  and  $\boldsymbol{\beta} = (1, \dots, 1)'$ . Simulations of  $\mathbf{x}$  and  $\mathbf{y}$  under the OLS model were obtained as a particular case of the TLS simulations setup (i.e., using  $m = \infty$  and  $\mathbf{v} \equiv 0$ ). Finally, under both OLS and TLS models, the matrix of control runs  $\boldsymbol{\varepsilon}$  was obtained by simulating the  $r$  runs independently from an  $n$ -variate Gaussian distribution with covariance  $\Sigma$ . The critical influence of the number of available control runs  $r$  on inference performance was assessed by repeating the entire process for 10 values of  $r$  ranging from  $r = 5$  to  $r = 150$ .

Implementing our integrated optimal fingerprinting (IOF) procedure requires the use of a target matrix  $\Delta$ , which describes the information available beforehand on the general structure of  $\Sigma$ . Four different targets  $\Delta_0$ ,  $\Delta_1$ ,  $\Delta_2$ , and  $\Delta_3$  were tested for each simulation, yielding four different estimates in each case. The identity matrix scaled to  $\Sigma$  [i.e.,  $\Delta_0 = \text{Tr}(\Sigma/n)\mathbf{I}$ ] was our default target. It allows for a direct comparison to Ribes et al. (2012), which is also based on this generic target. Yet the performance improvement achieved under a more specifically relevant target is also of interest. In the first, DR case with  $n = 275$ ,  $\Delta_3$  is a random perturbation of  $\Sigma$  obtained using the inverse Wishart distribution. In the second, NDR case with  $n = 10\,000$ ,  $\Delta_3$  is the isotropic exponential covariance matrix  $\mathbf{E}$ . In both cases, two intermediate targets,  $\Delta_k = w_k\Delta_0 + (1 - w_k)\Delta_3$  for  $k = 1, 2$ , with  $w_1 = 0.33$  and  $w_2 = 0.66$ , were introduced.

For each of the 40 experiments described above ( $= 2$  models  $\times 2$  setups  $\times 10$  values of  $r$ ),  $N = 1000$  simulations were generated. Inference was performed on each simulation using our IOF procedure successively with the four targets defined above. For the sake of comparison, inference was also performed based on the shrinkage procedure of Ribes et al. (2012) (referred to as ROF in Figs. 4–8) as well as the so-called *oracle* procedure (e.g., Ledoit and Wolf 2004). The oracle procedure consists in setting the value of  $\alpha$  to  $\alpha_o$ , where  $\alpha_o$  is obtained by minimizing the estimation error on  $\boldsymbol{\beta}$  viewed as a function of  $\alpha$ . The oracle estimator is thus, of course, not implementable in a real application where the true value of the parameter is not known. Its interest is rather that it defines an upper bound on the performance of any shrinkage estimator (i.e., ROF or IOF) and is thereby a useful benchmark.

The performance of each estimator of  $\boldsymbol{\beta}$  was evaluated based on average mean squared error (MSE)  $1/N \sum_{k=1}^N \|\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}\|^2 / \|\boldsymbol{\beta}\|^2$ , where  $k = 1, \dots, N$  denotes the simulation number, normalized to the mean squared error of the estimator obtained under known covariance. The accuracy of confidence intervals was evaluated by

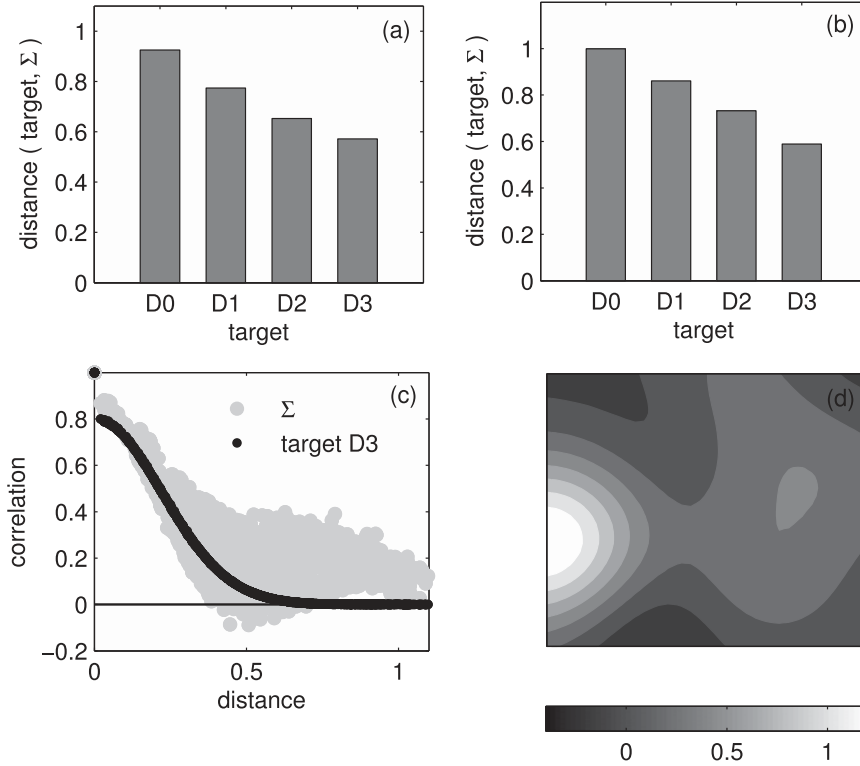


FIG. 3. Description of target matrices  $\Delta$  used in section 3. (a) Euclidian distance between  $\Delta_k$ ;  $k = 0, 1, 2$ , and 3; and  $\Sigma$  normalized to  $\|\Sigma\|$  in the dimension-reduced case. (b) As in (a), but in the nondimension-reduced case. (c) Scatterplot of correlation and distance  $\|s_i - s_j\|$  for all grid points in the nondimension-reduced case:  $\Sigma$  (gray dots) and exponential covariance  $E$  (black line). (d) Contour plot of correlation with one given grid point (located at maximum) for  $\Sigma$ .

comparing the empirical frequency  $1/N \sum_{k=1}^N 1_{\beta \in I_k}$  of the confidence regions, including the actual  $\beta$ , to its theoretical value of 90%.

### b. Results

As a first general remark, it appears that, in all cases and rather expectedly so, the MSE of the estimator of  $\beta$  is a monotonically decreasing function of  $r$  that gets steeper as  $r$  gets smaller and finally diverges to infinity as  $r \rightarrow 0$ . Concerning the OLS model (Figs. 4, 5), the MSE of our procedure under the target  $\Delta_0$  (IOF0) is indistinguishable from that of the ROF procedure, both being also indistinguishable from the MSE of the oracle estimator (Figs. 4a, 5a). This means that both ROF and IOF0 are able to come up with the best possible value of  $\alpha$ . The accuracy of confidence intervals is excellent in all three cases in the DR case, with the empirical frequency matching the theoretical value 90% (Fig. 4b). However, a pronounced discrepancy prevails w.r.t. the accuracy of confidence intervals in the NDR case: while the accuracy of IOF0 remains excellent, that of ROF degrades to about a 10% negative mismatch for small values of  $r$  (i.e., overconfidence) worsening to as high as a 40% mismatch for  $r = 150$  (Fig. 4b). Although it

was not computed, further worsening appears to be likely to occur for larger values of  $r$  given the consistently decreasing shape of the empirical frequency curve for ROF.

On the other hand, introducing more relevant targets shows a clear benefit in an IOF context. The MSE indeed sharply decreases when moving from IOF0 to IOF1, IOF2, and IOF3 (Figs. 4c, 5c). The performance gain appears to be higher when  $r$  is small; indeed, the better information on  $\Sigma$  included by the improved target naturally has a stronger impact in this situation, since less information is then required from control runs. When an improved target is used and  $r$  is small, the inference procedure is thus able to capture it by yielding a higher weight  $\alpha$  given to the target (not shown). The MSE decrease is particularly sharp in the NDR case: performance nearing perfection (i.e., nearing performance under known  $\Sigma$ ) is obtained even for small values of  $r$  and even though the target differs substantially from  $\Sigma$  in terms of Frobenius distance (Fig. 3b). The accuracy of confidence intervals remains excellent in IOF1, IOF2, and IOF3 (Figs. 4d, 5d), which means that the decrease in MSE is accurately captured by a decrease in the amplitude of the confidence region, on average.

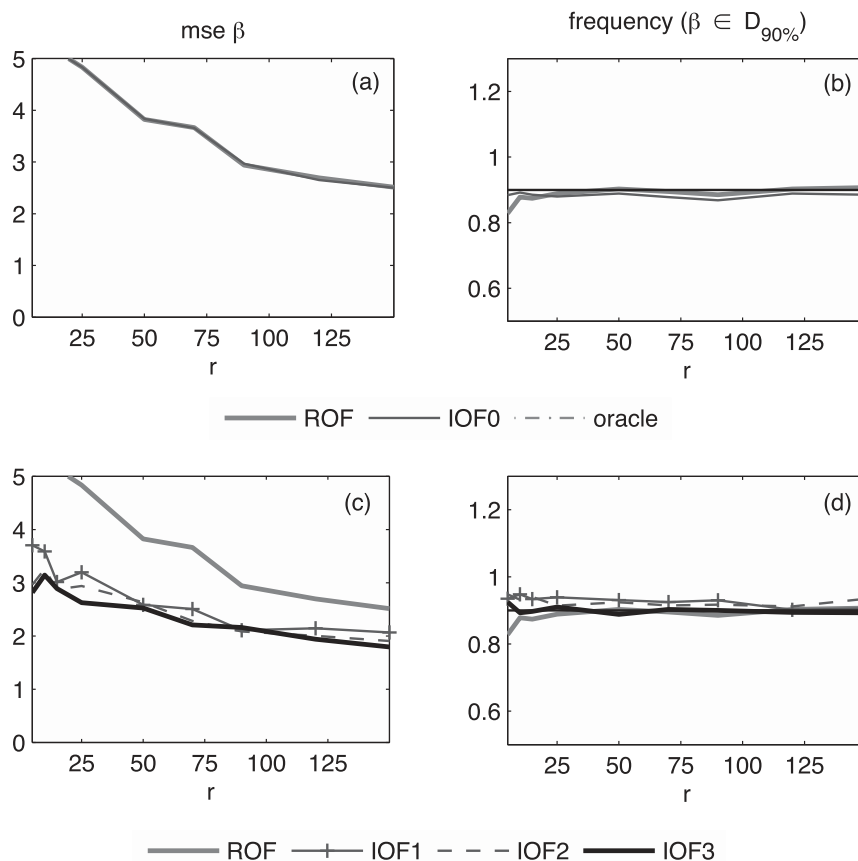


FIG. 4. Simulation results for the OLS model, dimension-reduced case. (a),(c) Average MSE of estimators of  $\beta$  as a function of the number of control runs  $r$ : IOF0 (light line), IOF1 (light line with crosses), IOF2 (dashed line), IOF3 (thick black line), ROF (thick gray line), and oracle (light gray line). (b),(d) Empirical frequency of  $\beta$  falling into the 90% confidence region (same legend).

The above findings obtained under the OLS model are maintained under the TLS model (Figs. 6, 7), except for a few changes that are highlighted. First, note that TLS results in lower estimation performance level as compared to OLS for all experiments and for all procedures. This holds here, not because TLS as a method performs worse in general, but because the design matrix is simulated with no noise under the OLS setting, whereas it is simulated with noise in the TLS setting. The gap in performance between OLS and TLS thus corresponds to the effect on estimation performance of noise in the design matrix when the correct method is used. Second, the accuracy of confidence intervals is strongly degraded for both ROF and IOF under the DR case (Figs. 6b,d). Under the NDR case, the degradation also occurs for ROF but not for IOF which maintains perfect accuracy (Figs. 7b,d). Third, the MSE benefit of using more relevant targets appears to be more pronounced under the TLS model than under the OLS model. It also appears to have a benefit in terms of confidence intervals in the DR

case, as their accuracy is partially restored by using more relevant targets (Fig. 6d).

## 5. Illustration

### a. Data

The purpose of this section is to illustrate the method by applying it in the context of a D&A study based on actual observations and GCM model simulations. The study focuses on two variables: surface temperature and precipitation, considered over the 1905–2005 period and over the entire surface of Earth, when observations are available. While numerous previous D&A studies are already available based on these variables (e.g., Hegerl et al. 2007b; Zhang et al. 2007; Noake et al. 2012; Polson et al. 2013; Lott et al. 2013; Marvel and Bonfils 2013), a novelty of this illustration is to highlight the effect on D&A results of using straightforwardly the variables of interest in their raw temporal and spatial form, and thus in potentially high dimension, without any prior processing for dimension

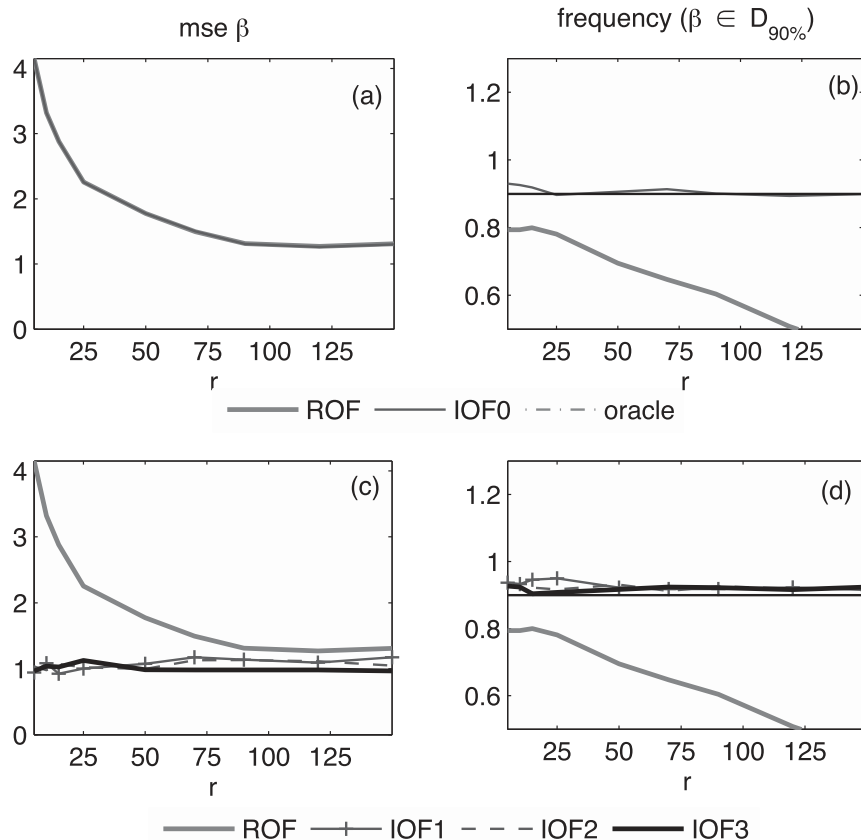


FIG. 5. As in Fig. 4, but for the OLS model, nondimension-reduced case.

reduction. Second, the benefit of more relevant and more constraining targets for the covariance matrix of internal variability than the identity matrix is enhanced.

Observations were obtained, on the one hand, from the HADCRUT4 monthly temperature dataset (Morice et al. 2012) over the period 1901–2000 and, on the other hand, from a combination of the GPCP v2.2 monthly precipitation dataset (Adler et al. 2003) over the period 1979–2000 with the GPCC v6 dataset (Schneider et al. 2013) over the period 1901–78. GCM model simulations were obtained from the IPSL CM5A-LR (Dufresne et al. 2012), the runs of which were downloaded from the CMIP5 database. An ensemble of runs consisting of two sets of forcings (i.e.,  $p = 2$ ) was used: the natural set of forcings (NAT) and the anthropogenic set of forcings (ANT), for which three runs are available in each case over the period of interest (i.e.,  $m = 3$ ) and from which an ensemble average was derived. On the other hand, a single preindustrial control run of 1000 yr is available and was thus split into 10 individual control runs of 100 yr. Next, following section 2d, the six residual runs consisting of the difference between individual response runs and the corresponding ensemble average were used as if they were additional control runs (i.e.,  $r = 16$ ).

Precipitation units in simulations were converted from kilograms per meter squared per second to millimeters per month to match the units of the observational dataset (scaling factor of  $2.63 \times 10^6$ ). Both precipitation and temperature in both observations and simulations were converted to anomalies by subtracting the time average over the reference period 1961–90.

The data were averaged temporally and spatially using time–space windows of different resolutions. The commonplace temporal resolution of 10 yr was used for all windows, and five different spatial resolutions on a latitude per longitude grid:  $1 \times 1$ ,  $2 \times 3$ ,  $5 \times 10$ ,  $10 \times 18$ , and  $21 \times 36$  corresponding to grid cells of size  $180^\circ \times 360^\circ$ ,  $90^\circ \times 120^\circ$ ,  $36^\circ \times 36^\circ$ ,  $18^\circ \times 20^\circ$ , and  $8.5^\circ \times 10^\circ$  and to theoretical values of  $n$  of 10, 60, 500, 1800, and 7560. Averaging over windows was performed for both observations and simulations by using only values for which observations were nonmissing for a like-to-like comparison between observations and simulations. Windows with more than 80% of missing observations were suppressed, yielding final values of  $n$  of 10, 60, 470, 1530, and 5700 for temperature and 10, 30, 270, 790, and 2960 for precipitation.

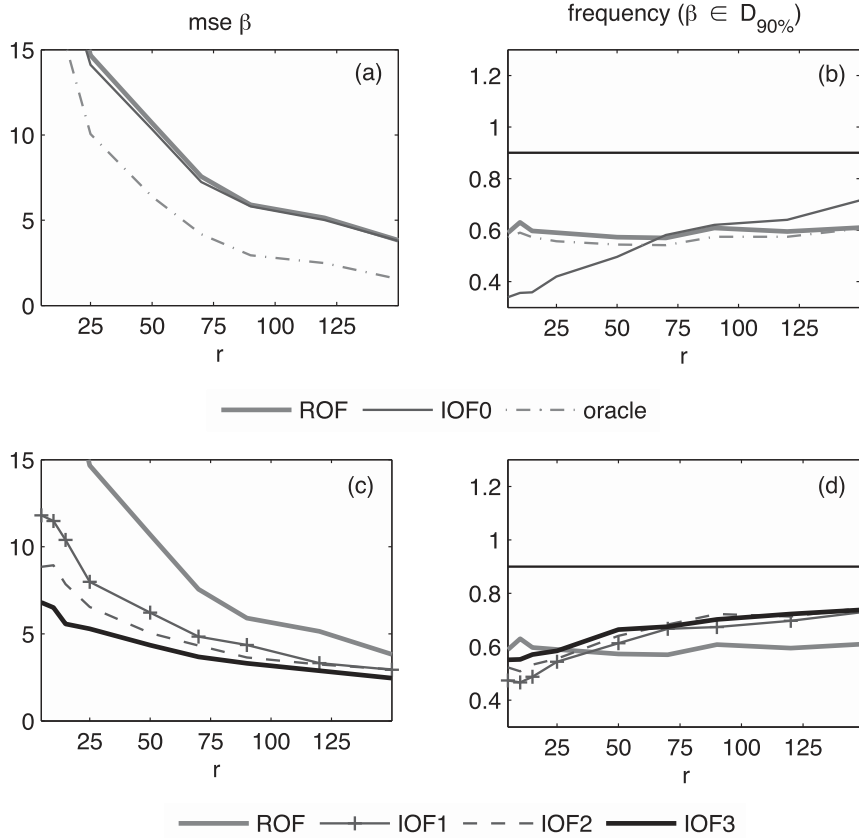


FIG. 6. As in Fig. 4, but for the TLS model, dimension-reduced case.

### b. Target matrix

The main hypothesis required for implementation of the described method consists in the structure of the target  $\Delta$ . The following assumption is used:

$$\Delta = \mathbf{R}_t \otimes \Sigma_s^{1/2} \mathbf{R}_s \Sigma_s^{1/2}, \quad (32)$$

where  $\mathbf{R}_t$  is a correlation matrix of size  $n_t$  reflecting the temporal dependence and is assumed to be exponential with parameter  $\theta_t$ ;  $\mathbf{R}_s$  is a correlation matrix of size  $n_s$  reflecting the spatial dependence and is assumed to be exponential with parameter  $\theta_s$ ;  $\Sigma_s = \text{diag}(\sigma_1^2, \dots, \sigma_{n_s}^2)$  is the diagonal matrix of variances at each location  $s$ ; and  $\otimes$  denotes the Kronecker product. The latter Kronecker product simply reflects the fact that it is assumed that the dependence in time, described by  $\mathbf{R}_t$ , is independent of the location  $s$  and, reciprocally, that the dependence in space, described by  $\mathbf{R}_s$ , is independent of time  $t$ . Note that the Kronecker product does not commute: Eq. (32) implies that the data is ordered first by date and second by location (i.e., the space–time data vectors of size  $n_t n_s$  consist of  $n_t$  spatial vectors of size  $n_s$ , which are concatenated by date). The assumption of time–space

independence has an important practical implication for implementation, as  $\Delta^{-1/2}$  can be obtained from

$$\Delta^{-1/2} = \mathbf{R}_t^{-1/2} \otimes \Sigma_s^{-1/4} \mathbf{R}_s^{-1/2} \Sigma_s^{-1/4}, \quad (33)$$

which requires us to diagonalize two matrices of size  $n_t$  and  $n_s$  instead of one matrix of size  $n_t n_s$ . The parametric form assumed for  $\Delta$  is required to estimate beforehand the local variances  $(\sigma_1^2, \dots, \sigma_{n_s}^2)$  and the exponential parameters  $(\theta_t, \theta_s)$  driving the range of the space–time dependence for temperature and precipitation. Local variances were obtained directly from the empirical covariance  $\mathbf{S}$  by  $\sigma_i^2 = \sum_{k=0}^{n_t-1} \mathbf{S}_{ks+i}/n_t$  for  $i = 1, \dots, n_s$ . With the local variances known, the exponential parameters were obtained by minimizing the mean squared error  $\|\mathbf{S} - \Delta\|^2$  viewed as a function of  $(\theta_t, \theta_s)$  (Fig. 8). Finally, the method was also implemented on two particular cases of Eq. (32) for the sake of comparison: (i) the diagonal target  $\Delta = \mathbf{I} \otimes \Sigma_s$  obtained by setting correlation matrices  $\mathbf{R}_s$  and  $\mathbf{R}_t$  to identity matrices with the same local variances; and (ii) the default target  $\Delta = \lambda \mathbf{I}$  with  $\lambda = \Sigma_{i=1}^{n_s} \sigma_i^2 / n_s = \text{Tr}(\mathbf{S}) / (n_t n_s)$ . By order of increasing complexity, the three targets defined above are denoted herein as  $\Delta_0$ ,  $\Delta_1$ , and  $\Delta_2$ .



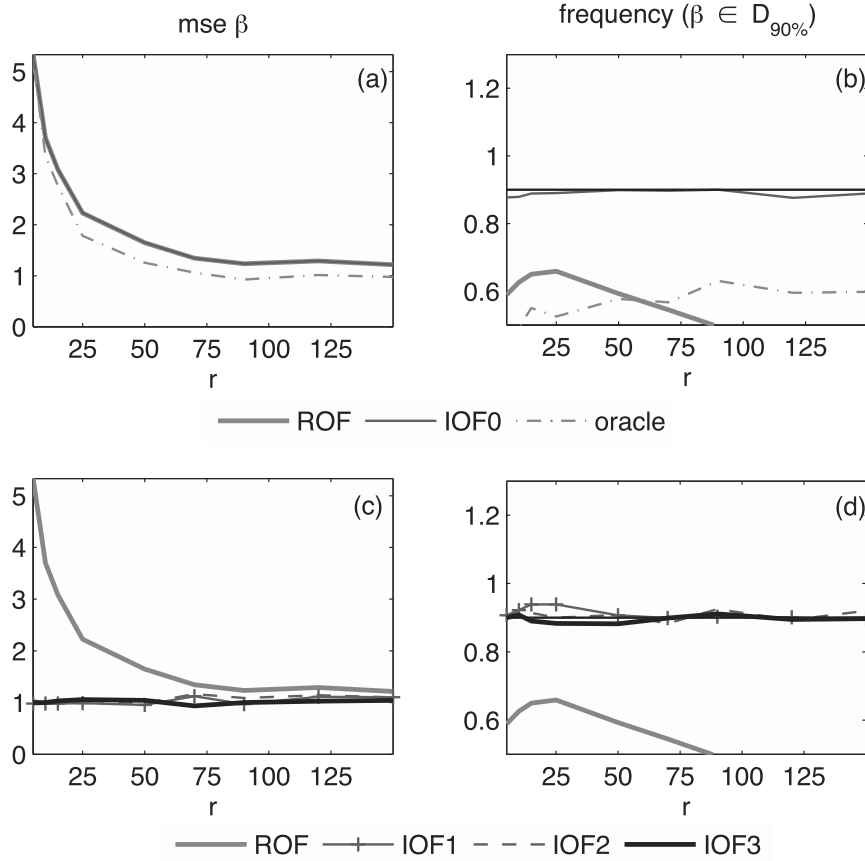


FIG. 7. As in Fig. 4, but for the TLS model, nondimension-reduced case.

### c. Model error

A well-known limitation of model Eq. (19) consists in assuming that the error terms  $\nu$  and  $\eta$  in the model Eq. (3) only originate from internal variability. This assumption means that the climate model used to simulate the responses to individual forcings is assumed to be perfect, that the model correctly simulates internal variability, and that observations are free of measurement error. When using real observations and simulations, these assumptions can be challenged. These missing error terms are likely to affect the performance of the estimation of  $\beta$  and accuracy of confidence intervals, because parameter  $\beta$  may then adjust to compensate for the missing error instead of fitting patterns  $\mathbf{x}^*$  to observations  $y$ , hence yielding values that are physically unrealistic.

While this problem is in our view an important matter in D&A, addressing it completely would imply substantially modifying our model assumptions and, in particular, introducing covariance matrices describing model and observational error. Such modifications would, in turn, raise inference difficulties that are beyond the scope of this article. The latter difficulties are nevertheless

discussed further in section 6, but for the sake of this illustration, we merely attempted to address this limitation in a simplified manner by introducing a slight modification to our model that does not affect its implementability. For this purpose, the covariance matrix  $\Sigma$  associated with  $\nu$  is simply inflated by a factor  $\tau \geq 1$  in order to reflect additional error terms. This inherently assumes that those extra error terms share the same covariance structure up to a scaling factor, which is not a particularly realistic assumption but is preferable in our view to letting  $\beta$  adjust for missing errors instead, as mentioned above. The value of  $\tau$  is chosen so that the inflated covariance provides as good a fit as possible to the vector of residuals  $\hat{\nu} = \mathbf{y} - \hat{\mathbf{x}}^* \beta_0$  with  $\beta_0 = (1, \dots, 1)'$ . This approach yields

$$\hat{\tau} = \frac{1}{n} \lambda_{\beta_0}^2 (\mathbf{y} - \mathbf{x} \beta_0)' \Sigma_{\alpha_0}^{-1} (\mathbf{y} - \mathbf{x} \beta_0), \quad (34)$$

where  $\alpha_0$  accordingly maximizes the integrated likelihood of Eq. (24) under  $\beta = \beta_0$ . Then expressions of the likelihood, of the MLEs  $\hat{\alpha}$  and  $\hat{\beta}$ , and of their corresponding confidence intervals in model Eq. (19) must hence be modified to include this new parameter. After some calculations, this boils down to changing  $\lambda_{\beta}$

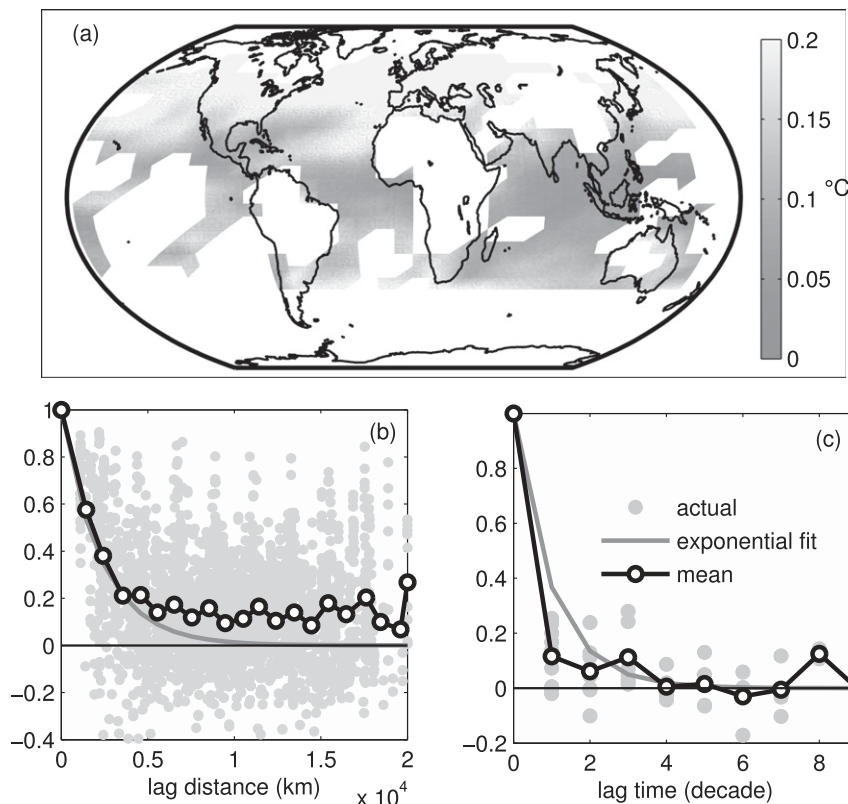


FIG. 8. Estimation of the parameters of the target matrix  $\Delta$  at resolution  $8.5^\circ \times 10^\circ$  for temperature data. (a) standard deviations ( $\sigma_1, \dots, \sigma_{n_x}$ ). (b) Spatial correlation as a function of distance: actual values (gray dots), mean of actual values by distance decile (black line with dotted circles), and fitted exponential model (bold gray line). (c) As in (b), but for temporal correlation.

into  $(\hat{\tau} + \beta' \beta / m)^{-1}$  in Eqs. (22) and (23);  $\lambda_\alpha$  into  $(\hat{\tau} + \beta' \beta / m)^{-1}$  in Eq. (24); and  $m$  to  $\hat{\tau} m$  in the calculation of  $\beta_\alpha$  and  $\beta$ . Note that, as expected, all estimators are thus left unchanged when using  $\hat{\tau} = 1$ . On the other hand, it is interesting to notice that, from the point of view of  $\beta$ , introducing such an error inflation factor  $\hat{\tau}$  is equivalent to artificially inflating the size of the ensemble  $m$  by the same factor  $\hat{\tau}$ . This is consistent with the results of Ribes and Terray (2012), who found that, in the context of a D&A study based on temperature data, artificially increasing the ensemble size in the TLS model led to improved estimates of  $\beta$ . The scaled error TLS model introduced above sheds light on this finding, as it shows that artificially increasing the ensemble size inherently corresponds to accounting for some missing error.

#### d. Results

The proposed method yields estimates of  $\beta_{\text{ANT}}$  that remain between 0 and 1 for all five resolutions, all three targets, and for both variables (Fig. 9). Estimated values of  $\beta_{\text{NAT}}$  are positive except occasionally at high resolution for both variables; they depart from one in a more

pronounced way in this case. Confidence intervals are much larger for precipitation than for temperature. Confidence intervals are much larger for  $\beta_{\text{NAT}}$  than for  $\beta_{\text{ANT}}$ . Confidence intervals exclude the value 0 for  $\beta_{\text{ANT}}$ , allowing detection to be performed, in nearly all situations. This holds as well for  $\beta_{\text{NAT}}$  for target  $\Delta_2$  or globally averaged variables but otherwise not.

For both forcings and both variables, the amplitude of confidence intervals is found to decrease as resolution increases (Fig. 10a).

The choice of the target does not appear to have a strong impact on the amplitude of confidence intervals on  $\beta$  but it does appear to influence significantly the value of its point estimate. In any case, the goodness of fit of the model—measured herein by  $\ell(\beta, \hat{\alpha})/n$ —is directly related to the complexity of the target: it reaches its maximum for target  $\Delta_2$  and its minimum for target  $\Delta_0$  (Fig. 10c).

Estimates of the covariance inflation parameter  $\tau$  range between 1.6 and 2.1 for temperature and between 1.9 and 6.5 for precipitation (Fig. 10b). In both cases, the minimum value is reached at low resolution  $2 \times 3$ . It then increases as resolution gets higher, which means that missing errors (i.e., model and observational error) are

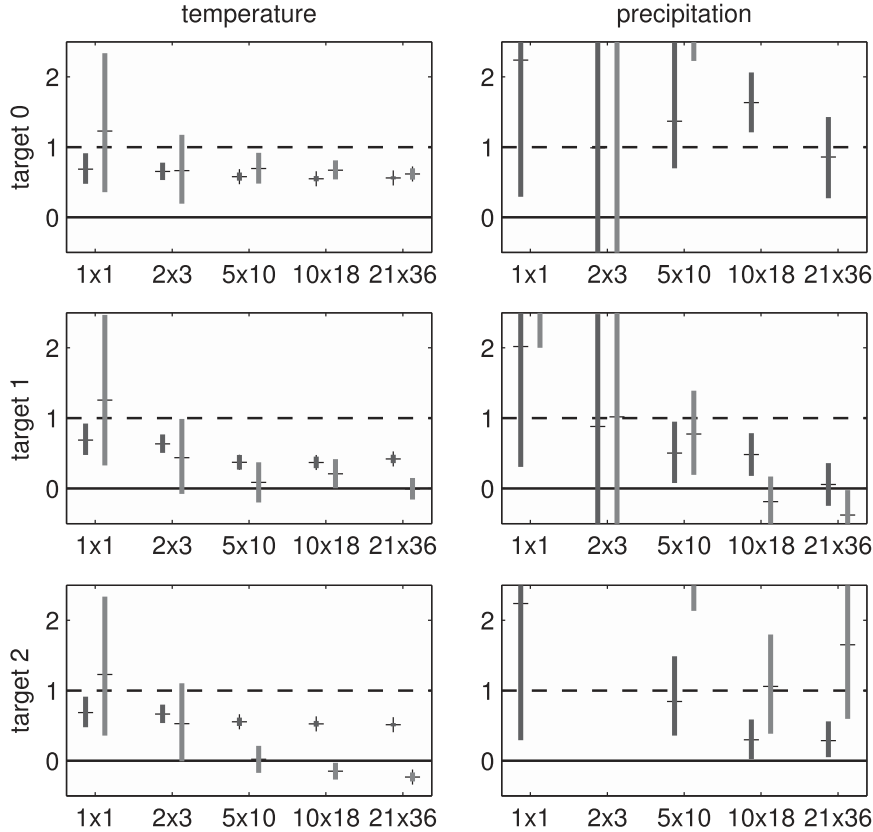


FIG. 9. Inference results. Vertical axis: estimated values and confidence intervals for  $\beta_{\text{NAT}}$  and  $\beta_{\text{ANT}}$  (gray and dark gray vertical lines). Horizontal axis: spatial resolution (left to right) lowest to highest. Labels  $1 \times 1$ ,  $2 \times 3$ ,  $5 \times 10$ ,  $10 \times 18$ , and  $21 \times 36$  correspond respectively to grid cells of size  $180^\circ \times 360^\circ$ ,  $90^\circ \times 120^\circ$ ,  $36^\circ \times 36^\circ$ ,  $18^\circ \times 20^\circ$ , and  $8.5^\circ \times 10^\circ$ . (left) Temperature data; (right) precipitation data. (top) Inference based on target matrix  $\Delta_0$ . (middle) Inference based on target matrix  $\Delta_1$ . (bottom) Inference based on target matrix  $\Delta_2$ .

more important at higher resolutions. Introducing an extra parameter necessarily improves goodness of fit by construction. Yet, in order to evaluate whether or not the gain in goodness of fit justifies the introduction of the extra parameter under scrutiny, the classic Bayes information criterion (BIC; Schwarz 1978) can be used. In this present case, this implies that the parameter  $\tau$  is worth including whenever the gain in goodness of fit  $\ell(\hat{\beta}, \hat{\alpha})/n$  is greater than  $\log n$ , or  $\Delta\ell(\hat{\beta}, \hat{\alpha})/n \log n \geq 1$ . The latter condition is met for all targets, variables, and resolutions. Further, Fig. 10d shows that  $\Delta\ell(\hat{\beta}, \hat{\alpha})/n \log n \geq 1$  steeply increases with resolution. This finding is not surprising, since missing errors also do, thus amplifying the benefit of including an error inflation parameter.

## 6. Conclusions

An integrated optimal fingerprinting method has been introduced. In contrast with existing approaches, the method treats within one statistical model all the

available data (i.e., climate observations, model response simulations, and control simulations) in a spatiotemporal, high-dimensional format. This means that no reduction of the dimension of the data is required initially. The critical covariance estimation step is built into the integrated inference procedure, thereby offering a more coherent treatment of all uncertainty sources. Inference of the regression coefficients vector  $\beta$  is obtained under closed-form expressions, which are implementable at a low computational cost even in high dimension.

A key idea underpinning the method consists in regularizing the covariance by mean of a hierarchical formulation. Closed-form expressions can then be obtained because of a convenient choice of conjugate distributions (i.e., Gaussian and inverse Wishart distributions), yielding a so-called linear shrinkage estimator toward the predefined target structure  $\Delta$ . One strength of the method as compared to that of Ribes et al. (2012) is that the target can be of any structure, thus opening access

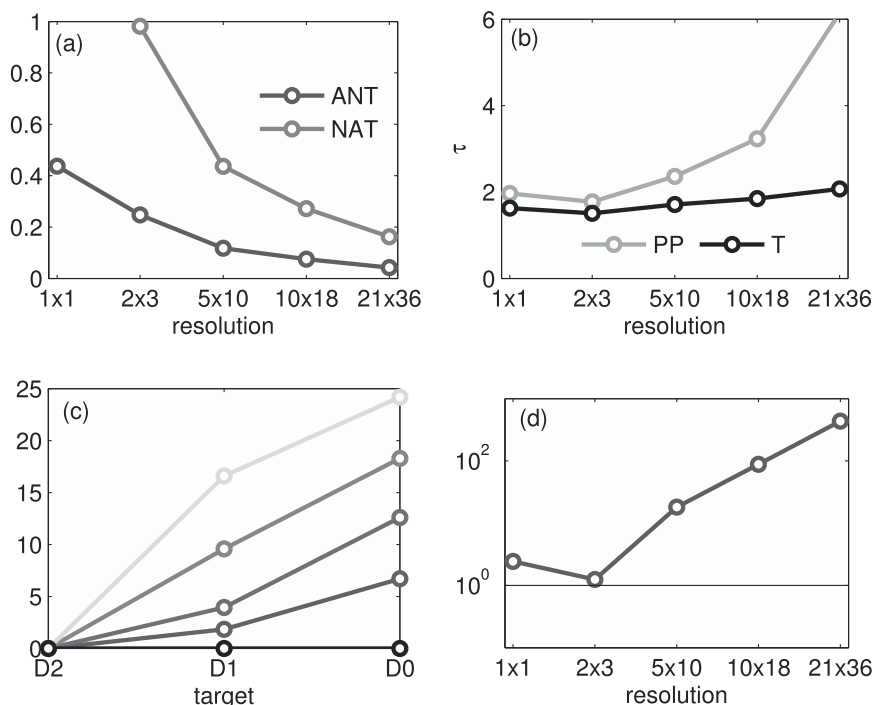


FIG. 10. Inference results. (a) Magnitude of confidence intervals for  $\beta_{\text{NAT}}$  and  $\beta_{\text{ANT}}$  (gray and dark lines) as a function of spatial resolution. (b) Average value of estimated covariance inflation factor  $\tau$  as a function of spatial resolution for temperature (black line) and precipitation (gray line). (c) Model negative log likelihood for each target (horizontal axis) and each spatial resolution (black line: lowest resolution; light gray line: highest resolution) for temperature data. (d) Average difference in model log likelihood between model with and without covariance inflation parameter  $\tau$  (bold line) and selection threshold implied by BIC (light line) for temperature data.

to a wide range of structures relevant for regularization, especially in a spatiotemporal context, where numerous covariance models are available. A second distinctive advantage is that the uncertainty associated with the latter linear shrinkage estimation of the covariance is consistently built into the inference of the parameter of interest  $\beta$  and of its confidence intervals, which indeed was the initial purpose of using an integrated model.

The hierarchical formulation of our model is actually equivalent in a Bayesian context to introducing an informative, conjugate prior distribution for  $\Sigma$  and a noninformative, improper, uniform prior distribution for  $\beta$  and  $\mathbf{x}^*$ . Rephrased in Bayesian terms, the proposed maximum likelihood estimator of  $\beta$  corresponds to the mode of its marginal posterior distribution; the confidence region corresponds to the quantiles of the same distribution, while estimation of  $\alpha$  is based on a so-called *empirical Bayesian* approach, meaning that a hyperparameter is chosen so as to maximize the marginal probability of the data. One interesting aspect of this approach is that it may be extended quite naturally by using informative priors on

$\beta$  and  $\mathbf{x}^*$  as well as on  $\Sigma$ , thereby introducing further regularization into D&A studies. One motivation to introduce such further regularization on  $\beta$  and  $\mathbf{x}^*$  is the disturbing fact that, in their current form and inclusive of the present method, TLS inferences often yield physically unrealistic values of both the coefficients  $\beta$ , which are sometimes inferred at values well above one or below zero, and of the patterns  $\mathbf{x}^*$ , which are often inferred to be less smooth than  $\mathbf{x}$ . It can be hoped that such a comprehensive regularization approach based on introducing informative priors, may yield more physically meaningful inference results. In any case, our proposal represents an effort to move toward that direction.

One limitation of our method is that it relies on the assumption that covariance matrices are proportional. To our knowledge, this limitation is actually shared by all optimal fingerprinting methods and studies to date, except those of [Huntingford et al. \(2006\)](#), who relaxed this assumption for the first time in the context of D&A, and [Hannart et al. \(2014\)](#), who proposed an inference procedure to deal with this situation. When this assumption is

relaxed (i.e., in the more general errors-in-variables regression model, where the covariance associated with the column vectors of  $\mathbf{x}$  is not proportional to  $\Sigma$ ) closed-form expressions are no longer valid, and likelihood maximization becomes mathematically intractable. Yet the proposed hierarchical model formulation and the regularization it allows remain, in our view, relevant in this case as well. Solving for the same model extended to this more general situation would thus be useful in our view and would imply designing an ad hoc numerical inference procedure. The latter might be inspired by the scheme of Hannart et al. (2014). The assumption that the noise is Gaussian is also a limitation of the method when applied at high resolution, because we can no longer rely on spatial averaging for the Gaussian assumption to hold.

The proposed method is found to outperform alternative inference methods of Allen and Tett (1999), Allen and Stott (2003), and Ribes et al. (2012) based on a simulation study w.r.t. both the error of estimation of  $\beta$  and the accuracy of the confidence intervals. In particular, the gap in performance with Ribes et al. (2012) is found to be significant even when the same target is used (i.e.,  $\Delta \propto \mathbf{I}$ ) and becomes wider when a more specific target is used relevantly or when the number of control runs  $r$  is small. However, despite these gains in performance, simulation results show that the method still yields imperfect confidence intervals in the TLS case and needs to be further refined to solve this issue.

Finally, the method was illustrated by applying it to global temperature and precipitation observational datasets at several increasing spatial resolutions, yielding estimates of  $\beta$  of increasing precision (uncertainty  $\sim n^{-0.35}$ ). Thus, a novelty of this study is to highlight the fact that dimension reduction may result in a loss of worthwhile information from a D&A standpoint. From this perspective, it appears to be essential to build climate model error into the inference, as the former tends to increase at higher resolution. It also appears to be relevant to focus further on the design of D&A tools, which, like the present method, are able to handle high-dimensional data, in an attempt to make the most of available observations to detect and attribute climate change.

*Acknowledgments.* I gratefully acknowledge Aurélien Ribes and Philippe Naveau for stimulating discussions while working on this study. I would also like to thank Francis Zwiers and an anonymous reviewer for many suggestions that resulted in a substantially improved manuscript. This work has been supported by the LEFE-INSU Grant MESDA and the ANR (Agence Nationale de la Recherche) Grant DADA.

## APPENDIX A

### Inverse Wishart Parameterization Choice

The inverse Wishart distribution is usually parameterized with  $\xi > n + 1$  and  $\Omega \in \mathcal{S}^{+*}$ :

$$p(\Sigma | \xi, \Omega) = 2^{-(\xi n/2)} \Gamma_n\left(\frac{\xi}{2}\right)^{-1} |\Omega|^{\xi/2} |\Sigma|^{-[(\xi+n+1)/2]} \times \exp\left[-\frac{1}{2} \text{Tr}(\Omega \Sigma^{-1})\right]. \quad (\text{A1})$$

The following bijective parameter change is used:

$$(\alpha, \Delta) = \left( \frac{\xi - n - 1}{r + \xi - n - 1}, \frac{\Omega}{\xi - n - 1} \right) \Leftrightarrow (\xi, \Omega) \\ = \left( \frac{\alpha r}{1 - \alpha} - n - 1, \frac{\alpha r}{1 - \alpha} \Delta \right), \quad (\text{A2})$$

with  $\alpha \in [0, 1]$  and  $\Delta \in \mathcal{S}^{+*}$ . With this new choice of metaparameters, the expression of the inverse Wishart PDF becomes

$$p(\Sigma | \alpha, \Delta) \propto |\Sigma|^{-[\alpha n/2(1-\alpha)]} \exp\left[-\frac{\alpha n}{2(1-\alpha)} \text{Tr}(\Delta \Sigma^{-1})\right], \quad (\text{A3})$$

where the normalizing constant is omitted for simplicity. Metaparameter  $\Delta$  can now be interpreted straightforwardly as the distribution mean, and metaparameter  $\alpha$  drives the magnitude of the variance of the elements of  $\Sigma$ :

$$\text{Var}(\Sigma_{ij} | \alpha, \Delta) \simeq \frac{1 - \alpha}{\alpha n} (\Delta_{ij}^2 + \Delta_{ii} \Delta_{jj}), \quad (\text{A4})$$

with zero variance for  $\alpha = 1$  and infinite variance for  $\alpha = 0$ .

## APPENDIX B

### Integrated Likelihood in the OLS Model

As a straightforward consequence of Eq. (A1),

$$\int_{\Sigma} |\Sigma|^{-[(\xi+n+1)/2]} \exp\left[-\frac{1}{2} \text{Tr}(\mathbf{A} \Sigma^{-1})\right] d\Sigma \\ = 2^{\xi n/2} \Gamma_n\left(\frac{\xi}{2}\right) |\mathbf{A}|^{-(\xi/2)} \quad (\text{B1})$$

for any  $\mathbf{A} \in \mathcal{S}^{+*}$ . Going further, the so-called normalizing constant is denoted  $K(\xi, \mathbf{A}) = 2^{\xi n/2} \Gamma_n(\xi/2) |\mathbf{A}|^{-\xi/2}$ . On the other hand, combining Eqs. (9), (10), and (12) yields

$$\begin{aligned}
p(\mathbf{y}, \boldsymbol{\varepsilon}, \boldsymbol{\Sigma} | \boldsymbol{\beta}, \alpha) &= 2\pi^{-(n/2)} |\boldsymbol{\Sigma}|^{-(1/2)} \exp \left[ -\frac{1}{2} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta}) \right] \\
&\times 2\pi^{-(nr/2)} |\boldsymbol{\Sigma}|^{-(r/2)} \exp \left[ -\frac{1}{2} \text{Tr}(r\mathbf{S}\boldsymbol{\Sigma}^{-1}) \right] \\
&\times K(\xi, \boldsymbol{\Omega})^{-1} |\boldsymbol{\Sigma}|^{-[(\xi+n+1)/2]} \exp \left[ -\frac{1}{2} \text{Tr}(\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}) \right] \\
&= 2\pi^{-[n(r+1)/2]} K(\xi, \boldsymbol{\Omega})^{-1} |\boldsymbol{\Sigma}|^{-[(\xi+n+r+2)/2]} \exp \left[ -\frac{1}{2} \text{Tr}(\mathbf{A}\boldsymbol{\Sigma}^{-1}) \right], \quad (\text{B2})
\end{aligned}$$

with  $\mathbf{A} = \boldsymbol{\Omega} + r\mathbf{S} + (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})(\mathbf{y} - \mathbf{x}\boldsymbol{\beta})'$ . The above expression in  $\boldsymbol{\Sigma}$  can be integrated by applying Eq. (B1) to obtain the following:

$$\begin{aligned}
p(\mathbf{y}, \boldsymbol{\varepsilon} | \boldsymbol{\beta}, \alpha) &= \int_{\boldsymbol{\Sigma}} p(\mathbf{y}, \boldsymbol{\varepsilon}, \boldsymbol{\Sigma} | \boldsymbol{\beta}, \alpha) d\boldsymbol{\Sigma} \\
&= K(\xi + r + 1, \mathbf{A}) \times K(\xi, \boldsymbol{\Omega})^{-1}, \quad (\text{B3})
\end{aligned}$$

where the constant  $2\pi^{-n(r+1)/2}$  has been dropped. Next, some rearrangements of the determinant of  $\mathbf{A}$  are performed based on the parameter change of Eq. (A2):

$$\begin{aligned}
|\mathbf{A}| &= (\xi + r - n - 1)^n \left| \alpha\boldsymbol{\Delta} + (1 - \alpha)\mathbf{S} + \frac{1}{\xi + r - n - 1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})(\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \right| \\
&= (\xi + r - n - 1)^n |\boldsymbol{\Sigma}_\alpha| \left| \mathbf{I} + \frac{1}{\xi + r - n - 1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})(\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \boldsymbol{\Sigma}_\alpha^{-1} \right|. \quad (\text{B4})
\end{aligned}$$

Sylvester's determinant theorem (i.e.,  $|\mathbf{I} + \mathbf{U}\mathbf{V}'| = |\mathbf{I} + \mathbf{V}'\mathbf{U}|$  for any two rectangular matrices and  $\mathbf{V}$  of same dimension) yields

$$\begin{aligned}
|\mathbf{A}| &= (\xi + r - n - 1)^n |\boldsymbol{\Sigma}_\alpha| \\
&\times [1 + 1/(\xi + r - n - 1) (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \boldsymbol{\Sigma}_\alpha^{-1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})]. \quad (\text{B5})
\end{aligned}$$

Finally, Eq. (14) is obtained after some algebra by combining Eqs. (A2), (B1), (B3), and (B5) and taking the logarithm.

omitted). This yields the following expression of the integrated likelihood in  $\boldsymbol{\beta}$ :

$$\ell_{\text{TLS}}(\boldsymbol{\beta}) = \lambda_{\boldsymbol{\beta}} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta}), \quad (\text{C2})$$

with  $\lambda_{\boldsymbol{\beta}} = (1 + \boldsymbol{\beta}'\boldsymbol{\beta}/m)^{-1}$ .

Next, the same steps that are applied in appendix B to the likelihood of the OLS model are applied to the complete likelihood of the TLS model given in Eq. (19). After integration in  $\boldsymbol{\Sigma}$  and some simplifications, the following reduced expression of the marginal likelihood in  $(\boldsymbol{\beta}, \mathbf{x}^*, \alpha)$  is obtained:

$$\begin{aligned}
\ell(\boldsymbol{\beta}, \mathbf{x}^*, \alpha) &= \frac{r}{1 - \alpha} \log |\mathbf{I} + \mathbf{X}'\mathbf{X}| \\
\text{with: } \mathbf{X} &= \left( \frac{1 - \alpha}{r} \right)^{1/2} \boldsymbol{\Sigma}_\alpha^{-1/2} [\mathbf{y} - \mathbf{x}^*\boldsymbol{\beta}, m^{1/2}(\mathbf{x} - \mathbf{x}^*)], \quad (\text{C3})
\end{aligned}$$

where  $\mathbf{X}$  is an  $n \times (p + 1)$  matrix. Maximizing  $\ell(\boldsymbol{\beta}, \mathbf{x}^*, \alpha)$  in  $\mathbf{x}^*$  is equivalent to maximizing  $\log |\mathbf{I} + \mathbf{X}'\mathbf{X}|$  in  $\mathbf{X}$  under the constraint  $\mathbf{X}_1 - m^{-1/2}\mathbf{X}_{-1}\boldsymbol{\beta} = \mathbf{z}$ , with  $\mathbf{z} = [(1 - \alpha)/r]^{1/2} \boldsymbol{\Sigma}_\alpha^{-1/2} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})$ , where  $\mathbf{X}_1$  denotes the first column vector of  $\mathbf{X}$ , and  $\mathbf{X}_{-1}$  denotes an  $n \times p$  matrix consisting of column vectors 2, ...,  $p + 1$  of  $\mathbf{X}$ .

Introducing  $\boldsymbol{\beta}^* = [1, m^{-1/2}\boldsymbol{\beta}]'$  and the vectorial Lagrange multiplier  $\boldsymbol{\gamma}$ , the corresponding Lagrangian  $\mathcal{L}(\mathbf{X}, \boldsymbol{\gamma})$  function can be written as

## APPENDIX C

### Integrated Likelihood in the TLS Model

Let us denote  $\ell_{\text{TLS}}(\boldsymbol{\beta}, \mathbf{x}^*)$  the negative log likelihood in  $(\boldsymbol{\beta}, \mathbf{x}^*)$  of the standard TLS model with known covariance  $\boldsymbol{\Sigma}$ , defined by

$$\begin{aligned}
\ell_{\text{TLS}}(\boldsymbol{\beta}, \mathbf{x}^*) &= (\mathbf{y} - \mathbf{x}^*\boldsymbol{\beta})' \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{x}^*\boldsymbol{\beta}) \\
&+ m \sum_{i=1}^p (\mathbf{x}_i - \mathbf{x}_i^*)' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \mathbf{x}_i^*). \quad (\text{C1})
\end{aligned}$$

An explicit maximization in  $\mathbf{x}^*$  of the right-hand term of Eq. (C1) can be performed (detailed calculations are



$$\mathcal{L}(\mathbf{X}, \boldsymbol{\gamma}) = \log|\mathbf{I} + \mathbf{X}\mathbf{X}'| - \boldsymbol{\gamma}(\mathbf{X}\boldsymbol{\beta}^* - z). \quad (\text{C4})$$

The first-order condition  $\partial\mathcal{L}/\partial\mathbf{X} = 0$  implies

$$\begin{aligned} \mathbf{X}_1 &= -(\mathbf{I} + \mathbf{X}\mathbf{X}')\boldsymbol{\gamma} \quad \text{and} \\ \mathbf{X}_{-1} &= m^{-1/2}(\mathbf{I} + \mathbf{X}\mathbf{X}')\boldsymbol{\gamma}\boldsymbol{\beta}'. \end{aligned} \quad (\text{C5})$$

Introducing Eq. (C5) into the constraint yields

$$(\mathbf{I} + \mathbf{X}\mathbf{X}')\boldsymbol{\gamma} = -\lambda_{\beta}z. \quad (\text{C6})$$

Substituting Eqs. (C5) and (C6) into Eq. (C3) and then maximizing in  $\mathbf{x}^*$  based on Eq. (C2), it becomes

$$\ell(\boldsymbol{\beta}, \alpha) = \frac{r}{1-\alpha} \log \left[ 1 + \lambda_{\beta} \frac{1-\alpha}{r} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta})' \boldsymbol{\Sigma}_{\alpha}^{-1} (\mathbf{y} - \mathbf{x}\boldsymbol{\beta}) \right], \quad (\text{C7})$$

which is the expression given in the main text.

## REFERENCES

- Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.*, **4**, 1147–1167, doi:[10.1175/1525-7541\(2003\)004<1147:TVGPCP>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2).
- Allen, M., and S. Tett, 1999: Checking for model consistency in optimal fingerprinting. *Climate Dyn.*, **15**, 419–434, doi:[10.1007/s003820050291](https://doi.org/10.1007/s003820050291).
- , and P. Stott, 2003: Estimating signal amplitudes in optimal fingerprinting, part I: Theory. *Climate Dyn.*, **21**, 477–491, doi:[10.1007/s00382-003-0313-9](https://doi.org/10.1007/s00382-003-0313-9).
- Basu, D., 1977: On the elimination of nuisance parameters. *J. Amer. Stat. Assoc.*, **72**, 355–366, doi:[10.1080/01621459.1977.10481002](https://doi.org/10.1080/01621459.1977.10481002).
- Berger, J. O., L. Brunero, and L. W. Robert, 1999: Integrated likelihood methods for eliminating nuisance parameters. *Stat. Sci.*, **14**, 1–28.
- Chen, C. F., 1979: Bayesian inference for a normal dispersion matrix and its application to stochastic multiple regression analysis. *J. Roy. Stat. Soc.*, **B41**, 235–248.
- Dufresne, J.-L., and Coauthors, 2012: Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5. *Climate Dyn.*, **40**, 2123–2165, doi:[10.1007/s00382-012-1636-1](https://doi.org/10.1007/s00382-012-1636-1).
- Hannart, A., and P. Naveau, 2014: Estimating high dimensional covariance matrices: A new look at the Gaussian conjugate framework. *J. Multivar. Anal.*, **131**, 149–162, doi:[10.1016/j.jmva.2014.06.001](https://doi.org/10.1016/j.jmva.2014.06.001).
- , A. Ribes, and P. Naveau, 2014: Optimal fingerprinting under multiple sources of uncertainty. *Geophys. Res. Lett.*, **41**, 1261–1268, doi:[10.1002/2013GL058653](https://doi.org/10.1002/2013GL058653).
- Hegerl, G. C., and F. Zwiers, 2011: Use of models in detection and attribution of climate change. *Wiley Interdiscip. Rev.: Climate Change*, **2**, 570–591, doi:[10.1002/wcc.121](https://doi.org/10.1002/wcc.121).
- , H. Von Storch, K. Hasselmann, B. Santer, U. Cubash, and P. Jones, 1996: Detecting greenhouse-gas-induced climate change with an optimal fingerprint method. *J. Climate*, **9**, 2281–2306, doi:[10.1175/1520-0442\(1996\)009<2281:DGGICC>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<2281:DGGICC>2.0.CO;2).
- , and Coauthors, 2007a: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 663–746.
- , T. J. Crowley, M. Allen, W. T. Hyde, H. N. Pollack, J. Smerdon, and E. Zorita, 2007b: Detection of human influence on a new, validated 1500-year temperature reconstruction. *J. Climate*, **20**, 650–666, doi:[10.1175/JCLI4011.1](https://doi.org/10.1175/JCLI4011.1).
- 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](https://doi.org/10.1029/2005GL024831).
- Ledoit, O., and M. Wolf, 2003: Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empirical Finance*, **10**, 603–621, doi:[10.1016/S0927-5398\(03\)00007-0](https://doi.org/10.1016/S0927-5398(03)00007-0).
- , and —, 2004: A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.*, **88**, 365–411, doi:[10.1016/S0047-259X\(03\)00096-4](https://doi.org/10.1016/S0047-259X(03)00096-4).
- Lott, F. C., P. A. Stott, D. M. Mitchell, N. Christidis, N. P. Gillett, L. Haimberger, J. Perlwitz, and P. W. Thorne, 2013: Models versus radiosondes in the free atmosphere: A new detection and attribution analysis of temperature. *J. Geophys. Res. Atmos.*, **118**, 2609–2619, doi:[10.1002/jgrd.50255](https://doi.org/10.1002/jgrd.50255).
- Marvel, K., and C. Bonfils, 2013: Identifying external influences on global precipitation. *Proc. Natl. Acad. Sci. USA*, **110**, 19 301–19 306, doi:[10.1073/pnas.1314382110](https://doi.org/10.1073/pnas.1314382110).
- Min, S., X. Zhang, and F. W. Zwiers, 2008: Human-induced Arctic moistening. *Science*, **320**, 518–520, doi:[10.1126/science.1153468](https://doi.org/10.1126/science.1153468).
- , —, —, and G. C. Hegerl, 2011: Human contribution to more-intense precipitation extremes. *Nature*, **470**, 378–381, doi:[10.1038/nature09763](https://doi.org/10.1038/nature09763).
- Morice, C. P., J. J. Kennedy, N. A. Rayner, and P. D. Jones, 2012: HADCRUT4 global dataset of surface temperature. Climatic Research Unit, University of East Anglia, accessed 11 June 2014. [Available online at <https://crudata.uea.ac.uk/cru/data/temperature/HadCRUT.4.4.0.0.median.nc>.]
- Noake, K., D. Polson, G. Hegerl, and X. Zhang, 2012: Changes in seasonal land precipitation during the latter twentieth century. *Geophys. Res. Lett.*, **39**, L03706, doi:[10.1029/2011GL050405](https://doi.org/10.1029/2011GL050405).
- Polson, D., G. C. Hegerl, X. Zhang, and T. J. Osborn, 2013: Causes of robust seasonal land precipitation changes. *J. Climate*, **26**, 6679–6697, doi:[10.1175/JCLI-D-12-00474.1](https://doi.org/10.1175/JCLI-D-12-00474.1).
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 2007: *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. Cambridge University Press, 1256 pp.
- Ribes, A., and L. Terray, 2012: Application of regularised optimal fingerprinting to attribution. Part II: Application to global near-surface temperature. *Climate Dyn.*, **41**, 2837–2853, doi:[10.1007/s00382-013-1736-6](https://doi.org/10.1007/s00382-013-1736-6).
- , J.-M. Azais, and S. Planton, 2009: Adaptation of the optimal fingerprint method for climate change detection using a well-conditioned covariance matrix estimate. *Climate Dyn.*, **33**, 707–722, doi:[10.1007/s00382-009-0561-4](https://doi.org/10.1007/s00382-009-0561-4).
- , S. Planton, and L. Terray, 2012: Application of regularised optimal fingerprinting to attribution. Part I: Method, properties and idealised analysis. *Climate Dyn.*, **41**, 2817–2836, doi:[10.1007/s00382-013-1735-7](https://doi.org/10.1007/s00382-013-1735-7).

- Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf, 2013: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theor. Appl. Climatol.*, **115**, 15–40, doi:[10.1007/s00704-013-0860-x](https://doi.org/10.1007/s00704-013-0860-x).
- Schwarz, G., 1978: Estimating the dimension of a model. *Ann. Stat.*, **6**, 461–464, doi:[10.1214/aos/1176344136](https://doi.org/10.1214/aos/1176344136).
- Stott, P. A., D. A. Stone, and M. R. Allen, 2004: Human contribution to the European heatwave of 2003. *Nature*, **432**, 610–614, doi:[10.1038/nature03089](https://doi.org/10.1038/nature03089).
- Van Huffel, S., and J. Vandewalle, 1994: *The Total Least Squares Problem: Computational Aspects and Analysis*. SIAM, 288 pp., doi:[10.1137/1.9781611971002](https://doi.org/10.1137/1.9781611971002).
- Venzon, D. J., and S. H. Moolgavkar, 1988: A method for computing profile-likelihood-based confidence intervals. *Appl. Stat.*, **37**, 87–94, doi:[10.2307/2347496](https://doi.org/10.2307/2347496).
- Zhang, X., F. W. Zwiers, G. C. Hegerl, F. H. Lambert, N. P. Gillett, S. Solomon, P. A. Stott, and T. Nozawa, 2007: Detection of human influence on twentieth-century precipitation trends. *Nature*, **448**, 461–465, doi:[10.1038/nature06025](https://doi.org/10.1038/nature06025).