Mathematical Statistics
Stockholm University

Towards a flexible statistical modelling by latent factors for evaluation of simulated responses to climate forcings: Part I

Ekaterina Fetisova, Gudrun Brattström,
Anders Moberg, Rolf Sundberg

Research Report 2017:12

ISSN 1650-0377

Postal address:
Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:
http://www.math.su.se

Mathematical Statistics
Stockholm University
Research Report **2017:12**,
http://www.math.su.se

Towards a flexible statistical modelling by latent factors for evaluation of simulated responses to climate forcings: Part I

Ekaterina Fetisova* Gudrun Brattström† Anders Moberg‡ Rolf Sundberg§

October 2017

Abstract

Evaluation of climate model simulations is a crucial task in climate research. In a work consisting of three parts, we propose a new statistical framework for evaluation of simulated responses to climate forcings, based on the concept of latent (unobservable) factors. Here, in Part I, we suggest several latent factor models of different complexity that can be used for evaluation of temperature data from climate model simulations against climate proxy data from the last millennium. Each factor model is developed for use with data from a single region, which can be of any size. To be able to test the hypotheses of interest, we have applied the technique of *confirmatory factor analysis*. We also elucidate the link between our factor models and the statistical methods used in Detection and Attribution (D&A) studies. In particular, we demonstrate that our factor models can be used as an alternative approach to the methods used in D&A studies. An additional advantage of their use is that they, in contrast to the commonly used D&A methods, make it, in principle, possible to investigate whether the forcings of interest act additively or if any interaction effects exist. In Part II we investigate and illustrate the expansion of factor models to structural equation models, which permits the statistical modelling of more complicated climatological relationships. The performance of some of our statistical models suggested in Part I and Part II is evaluated and compared in a numerical experiment, whose results are presented in Part III.

Keywords: Confirmatory Factor Analysis, Structural Equation models, Measurement Error models, Climate model simulations, Climate forcings, Climate proxy data, Detection and Attribution

*Department of Mathematics, Stockholm University; `katarina@math.su.se`
†Department of Mathematics, Stockholm University; `gudrun@math.su.se`
‡Department of Physical Geography, Stockholm university, Sweden; `anders.moberg@natgeo.su.se`
§Department of Mathematics, Stockholm University; `rolfs@math.su.se`

1 Introduction

Substantial efforts have been made in the past decades to understand how different climate factors contribute to climate variability, in particular regarding the roles of anthropogenic influences versus natural variability. The main findings and methods used are discussed in assessment reports by the Intergovernmental Panel on Climate Change (IPCC) ([33] - [37]). These reports review evidence for an increasing anthropogenic influence over time and also an increasing confidence in our understanding, with the latest report concluding that: "It is *extremely likely* that human influence has been the dominant cause of the observed warming since the mid-20th century". This conclusion is based on a combination of various observational evidence, theoretical knowledge, modelling studies and application of statistical analysis techniques. In the following paragraphs, we will briefly mention and explain some important concepts in order to provide a motivation for deeper discussion and continued development of some of the statistical techniques that have been applied, in particular in the context of understanding causes for climate variations in the last millennium.

Causes of climate variability can be separated into two main categories: *internal* variations that occur within the climate system and variations that are caused by some *external* factors residing outside the climate system. Both these main causes of variability can act on a wide range of time scales from sub-daily to hundreds of million years [56]. Examples of internal variations are the various fast atmospheric dynamic processes that cause the day-to-day weather variations and the somewhat slower ocean dynamic processes that cause variations in the surface ocean currents. Examples of external causes of climate variability are changes in the energy output from the Sun, changes in the Earth's orbit around the Sun and explosive volcanic eruptions that eject vast amounts of small particles into the atmosphere and thereby partially block the incoming solar radiation. The anthropogenic emissions of carbon dioxide and other greenhouse gases through the combustion of fossil fuels and various industrial activites is another kind of external cause of climate variation.

As climate scientists cannot make controlled and repeated experiments with the real climate system to increase understanding, they need computer-based models that are able to simulate the climate system ([46], [66]) in order to test different hypotheses about causes for climate variability. By making systematic experiments with such models (which can range from very simple to very complex ones), they can compare model output from different experiments made with different climate models and they can compare these results with observed climate records ([10], [18]). Moreover, to assess the results from such model experiments and model–data comparisons, they also need appropriate statistical methods to make inferences about their hypotheses. Depending on the actual climatological problem that a scientist encounters, an array of statistical methods have been developed. These may differ depending on whether a scientist works with climates of the distant past, when only sparse paleoclimatic information is available, or with modern climates when we have access both to space-borne observations with global coverage and in-situ observations that cover most parts of the world.

Radiative forcing is an important term in this context, which climate scientists have been using for several decades (see overview by [44] and the glossary in [37], p.1460). This concept is based on a physical framework, where a radiative energy imbalance in the atmosphere (which can be quantified in W/m^2), inflicted by a perturbation in the climate system for some reason, will cause a change in the state of the climate system and can alter processes within the climate system, which in turn enhance or dampen the initial effects and thus introduce positive or negative feedback loops. Sometimes scientists use the term *climate forcing* instead of radiative forcing ([44]), and we will generally do so in our discussion, or we simply write just *forcing*. This is motivated as we are not dealing specifically with understanding radiative processes here. The above-mentioned external causes of climate change, namely solar variations, orbital variations, volcanic eruptions and anthropogenic emissions of greenhouse gases are examples of external forcings of climate change. They occur in reality, their impact varies over time and space, and scientists can implement representations [1] of them in simulation experiments with climate models. Modellers can also choose to treat changes in internal processes as forcings in their simulation experiments. One example is when the climate model used does not allow interactions and feedbacks between vegetation and climate. In this case, the researcher can instead implement a change in the vegetation in the experiment setup, and thus regard this as a forcing instead of an internal cause of climate change.

The focus of our discussion concerns statistical methods applied to simulation and observational data[2] for climate variations that occurred within approximately the last millennium, in particular within the pre-industrial epoch. This is a period for which several so-called climate proxy data (i.e. indirect climate information from various natural archives such as tree-rings, lake sediments, cave speleothems, etc.) are available and can be used to estimate climate variations with a yearly resolution from parts of all continents [38], [52]. These data can be calibrated directly against the available instrumental records that reach back about 100-150 years in many places. Also, several simulation

[1] We need not go into detail here about how the knowledge of past climate forcings has been derived or how they are implemented in climate model experiments. But, clearly, scientists need to use some kind of indirect evidence to obtain estimations of their temporal evolution, their spatial influence, their magnitude, etc. Many forcing histories are derived from some kind of forcing proxy data, whereas our knowledge of the orbital forcing change has been derived from mathematical–astronomical calculations. An overview of how important climate forcing histories have been obtained for the last millennium is provided by [57, 58].

[2] Typically, observational data consist of instrumental measurements and proxy data. In our context, the only essential difference between instrumental observations and climate reconstructions from proxy data is that the latter data are less precise and need to be statistically calibrated against instrumental data. Thus, we will often simply write 'observations' to denote either type of observational evidence, unless when we have reasons to specify which type we mean.

experiments with sophisticated state-of-the-art climate models, so-called Global Climate Models (GCMs) or Earth System Models (ESMs), have been performed for about the last millennium and thus allow us to test hypotheses about causes for climate variability in this period [53], [51]. The climate forcings that have been studied in these experiments are primarily solar, volcanic, land cover/land use (i.e. essentially changes in the fraction of crop and pasture areas), greenhouse gases, and orbital changes. Several investigations, e.g. [23], [59], [60], have concluded that hemispheric-scale decadal-mean temperatures in the last millennium show a significant influence from external forcings, particularly in the later half of the millennium. Volcanic eruptions and changes in greenhouse gas concentrations seem to be the most important factors in this period, although variations in solar radiation can also be seen albeit they are more weakly detected. Evidently, the statistical evaluation of the ability of forced climate models to simulate observed climate change is a key issue within climate research. A central role in addressing this issue has been played by the so-called Detection and Attribution (D&A) studies ([67], [47], [23], [24]). Importantly, evaluating the ability of forced climate models to simulate observed climate change is one of the goals of D&A studies. As pointed out by [25], the initial focus was on determining whether the radiative forcing due to greenhouse gas increases has indeed influenced climate ([21], [22]). Further, the studies aim at assessing and quantifying how different external factors affect the observed climate changes, in particular, applied to *the near-surface air temperature*, which has been a climatic variable of interest for many D&A studies (see [26], [60]).

Statistical methods used in D&A studies performed to date have different representations ([19], [20], [43]), and typically are referred to as an 'optimal fingerprinting' technique. In the present paper, we confine our attention to the optimal fingerprinting framework based on linear regression models of varying complexity, described by [1]. One of the main assumptions made in [1] is that neither the real temperature responses to particular forcings nor the simulated responses to imposed forcings obtained in experiments with complex GCMs or ESMs are directly observable. This has motivated the use of regression models allowing both explanatory and response variables be contaminated with noise. However, as pointed out by some researchers, within some studies, optimal fingerprinting is associated with some disadvantages, for example, the inability to take into account the effects of possible interactions between forcings (see e.g. [45], [60]).

In the present paper, our aim is to suggest statistical models that allow us to overcome this difficulty (among others). To achieve this, we use the fact that regression models used in D&A studies can be viewed as factor models, where observable simulated and reconstructed temperatures can also be represented as linear combinations of (scaled) latent temperature responses to particular forcings plus the random internal temperature variability. Reasoning in the spirit of factor analysis has allowed us to formulate various factor models for evaluation of both single forcing climate model simulations and multi-forcing climate model simulations. Note that each factor model is developed for use with data from a single region, which can be of any size.

The definitions of the observed variables used in our factor models were taken from the statistical framework developed by [64] (hereafter referred to as SUN12). The SUN12 framework was specifically developed to suit the comparison of climate model simulations and proxy data for the relatively recent past of about one millennium. As a result, a correlation and a distance test-statistic were developed. However, for the purpose of motivating the sought-after factor models, the framework needs to be developed in more detail.

Finally, let us describe the structure of the present paper. First, in Sec. 2, the statistical method of optimal fingerprinting as used in many D&A studies will be presented, with a focus on aspects that are important in our context. Further, in Sec. 3, we present the modified definitions of the SUN12 framework. In Sec. 4, we will present our factor models. The paper will be concluded by a discussion aiming to compare the key characteristics of the statistical approach used in D&A studies with the properties of our factor models (see Sec. 5).

2 Overview of the statistical approach used in detection and attribution (D&A) studies

As indicated by the name, D&A analyses consist of two stages: detection and attribution.

Referring to [47], [23], the process of *detection of change* is described as a process of investigating whether the effect of a certain single external forcing or a combination of external forcings has a statistically significant influence on the observed climate or a system affected by climate. Stated another way, the aim is to investigate whether observed changes could be entirely caused by the internal climate variability, i.e. the variability not forced by external agents, or if external forcings (natural and/or anthropogenic) are needed to explain these changes. Most detection and attribution studies have used near-surface air temperature data, which have been found to show significant forced climate change signals relative to natural internal variability. This explains our choice to focus on near-surface air temperature data in our discussion.

Attribution in turn aims 'to determine whether a specified set of external forcings are the cause of the detected change' ([24]). However, as noted by [47], the determination of causes is not feasible in real-world data alone, because it requires controlled experimentation with the climate system. Therefore, in practice, the attribution process consists in demonstrating that 'the observed detected changes are consistent with the estimated (simulated) responses to the given external forcings, and not consistent with alternative, physically plausible explanations of recent climate change that exclude important elements of the given combination of forcings' ([23] and references therein).

As highlighted by [59], the main idea of the statistical approach used in D&A studies consists in decomposing the climate variability into two components: the forced component, representing the variability due to various forcings (both external and internal to the climate system), and the random internal climate variability arising due to various interactions within and between climate system components. The forced and unforced components are assumed to be unrelated.

Applying this idea to the reconstructed/observed temperatures, driven by all possible forcings, and to simulated temperatures, generated by climate models driven by a combination of reconstructed forcings [3] we could write:

$$
\begin{aligned}
Y_g &= \xi^T_{\mathsf{ALL}\,g} + \nu_{0\,g}, \\
x_g &= \xi^S_{\mathsf{f}\,g} + \nu_g,
\end{aligned}
\tag{2.1}
$$

where

Y_g - the *mean-centered* reconstructed/observed temperature, where the index g reflects the fact that the $\{Y_g\}$ -sequence is a temperature field, arrayed in space and in time [4];

$\xi^T_{\mathsf{ALL}\,g}$ - the True latent overall temperature response to all forcings that occur in reality (ξ^T_{ALL} is our own notation, where the superscript T stands for *True*, not a transpose);

$\nu_{0\,g}$ - the real internal temperature variability;

x_g - the *mean-centered* simulated temperature, corresponding to Y_g;

$\xi^S_{\mathsf{f}\,g}$ - the Simulated latent overall temperature response to the combination of reconstructed forcings in question;

ν_g - the simulated internal temperature variability.

One of the main assumptions of D&A studies is that climate models simulate the *shape* (i.e. the spatio-temporal pattern but not necessarily the *magnitude*) of the latent tem-

[3] In a climate model simulation experiment, however, scientists may rather implement as many forcings as considered being of importance or being known. This will unlikely include all possible forcings. For example, [51] used the term 'FULL' to denote simulations where they applied the full set of forcings used in their experiment.

[4] Concerning the dimensionality of data, we first of all refer to [54] that provides a quite detailed description of how data matrices can be formed. As we understand, the length of the $\{Y_g\}$-sequence can be very high because the sequence is initially formed by concatenating data sampled over the region and time period of interest. Different approaches are used for reducing the length of the sequence to make it possible to transform the noise sequence $\{\nu_{0\,g}\}$ into a white noise where all elements are mutually uncorrelated and equally distributed. Often researchers use Empirical Orthogonal Functions (EOFs), derived from principal component analysis of corresponding control climate model simulations (for definition of EOFs see for example [9]). A typical final length of the pre-whitened $\{Y_g\}$-sequence is 10-15 observations (see for example [65], [62], [15]). The same pre-whitening operator is applied to the $\{x_g\}$-sequence because it is assumed that model-simulated internal variability is consistent with that in the real world ([1]).

perature response correctly ([24]). In terms of $\xi^{T}_{\mathsf{ALL}\,g}$ and $\xi^{S}_{\mathsf{f}\,g}$, the assumption is reflected by representing the former response as a linear function of the latter, that is,

$$\xi^{T}_{\mathsf{ALL}\,g} = \beta \cdot \xi^{S}_{\mathsf{f}\,g} = \beta \cdot (x_g - \nu_g). \tag{2.2}$$

Inserting Eq. (2.2) into Y_g from (2.1) leads to a statistical model used in D&A studies for analysing the *overall* response to all the forcings in question ([1]):

$$Y_g = \beta \cdot (x_g - \nu_g) + \nu_{0\,g}. \tag{2.3}$$

In statistical literature, model (2.3) is known as a measurement error (ME) model [5], and typically is written as follows ([8], [14]):

$$\begin{cases} Y_g = & \beta \cdot \xi^{S}_{\mathsf{f}\,g} + \quad \nu_{0\,g}, \\ x_g = & \qquad \xi^{S}_{\mathsf{f}\,g} + \quad \nu_g. \end{cases} \tag{2.4}$$

As mentioned earlier, $\xi^{S}_{\mathsf{f}\,g}$ is assumed to be uncorrelated with the noise terms $(\nu_{0\,g},\, \nu_g)'$, which in their turn are assumed to be normally distributed with zero means and a 2×2 covariance matrix $\boldsymbol{\Sigma_{\nu\nu}}$. The latent $\xi^{S}_{\mathsf{f}\,g}$:s can be treated either as random (typically, normal) with zero mean and variance σ^2_{ξ}, or as fixed unknown constants. In the latter case, the following restrictions are placed on the $\{\xi^{S}_{\mathsf{f}\,g}\}$-sequence:

$$\underbrace{\lim_{n\to\infty} \frac{1}{n} \sum_{g=1}^{n} \xi^{S}_{\mathsf{f}\,g}}_{\equiv \mu_{\xi^S_{\mathsf{f}}}} = 0 \quad \text{and} \quad \underbrace{\lim_{n\to\infty} \frac{1}{n} \sum_{g=1}^{n} (\xi^{S}_{\mathsf{f}\,g})^2}_{\equiv \sigma^2_{\xi^S_{\mathsf{f}}}} > 0. \tag{2.5}$$

The detection and attribution questions are addressed by testing the hypotheses that $\beta = 0$ and $\beta = 1$, respectively. If $H_0 : \beta = 0$ is rejected, then it is said that the temperature response to the combined effect of all external forcings is detected. Notice that since Y_g in (2.4) is represented as a function of a simulated forcing effect, conclusions arrived at in the detection stage concern the *simulated* overall effect of *reconstructed* forcings. To be able to attribute the observed change in observational data to the real-world forcings, the simulated forcing effect should be transformed into the true one. According to Eq. (2.2), this is possible if $\beta = 1$. Therefore, the detection stage is followed by the attribution stage, where $H_0 : \beta = 1$ is tested. Obviously, testing this hypothesis implies that we, at the same time, evaluate the ability of the climate model

[5] More precisely, it is an ME model with no error in the equation, which refers to the fact that Eq. (2.2) does not contain an error. In our opinion, this error could be motivated because $\xi^{T}_{\mathsf{ALL}\,g}$ represents the latent true temperature response to all possible forcings acting in the real-world climate system, while $\xi^{S}_{\mathsf{f}\ g}$ is the latent simulated temperature response to a specific combination of (reconstructed) forcings in question.

under study to simulate the forcing effect correctly. If $H_0 : \beta = 1$ is not rejected, then it is said that the magnitude of the overall temperature response is found to be consistent in reconstructions and in simulations. Notice that as follows from Eq. (2.2), this conclusion is drawn under the assumption that the real-world climate system is forced by the same forcings used to force the simulated climate system.

To interpret which individual forcings contribute to the detected overall forcing effect, D&A studies use a multifingerprint analysis (see for example [22], [32], [60]). In the statistical parlance, this approach corresponds to a ME model with a vector explanatory variable:

$$Y_g = \sum_{i=1}^{m} \beta_i \left(x_{i\,g} - \nu_{i\,g}\right) + \nu_{0\,g}, \tag{2.6}$$

where

$x_{i\,g}$	- the simulated temperature, generated by the climate model forced only by a reconstruction of forcing i,
$x_{i\,g} - \nu_{i\,g} = \xi^S_{\mathtt{f}\,i\,g}$	- the simulated latent temperature response to forcing i, embedded in $x_{i\,g}$,
$\nu_{i\,g}$	- the simulated internal temperature variability associated with the climate model driven by forcing i,
β_i	- a scaling factor associated with forcing i. The process of detection and attribution is analogous to that of model (2.3).

In D&A studies, the estimation of the parameters of models (2.3) and (2.6) is based on the approach of Total Least Squares (TLS). The approach requires the knowledge of the ratios of the noise variances $\sigma^2_{\nu_i}/\sigma^2_{\nu_0}$ for $i = 1, 2, \ldots, m$, where $\sigma^2_{\nu_0}$ can be regarded either as unknown or known a priori. Commonly, in practice the whole noise variance-covariance matrix $\boldsymbol{\Sigma_{\nu\nu}}$ is assumed to be known a priori, which permits checking for model validity. Importantly, an estimate of $\boldsymbol{\Sigma_{\nu\nu}}$ should be obtained from a source independent from the sample variance-covariance matrix of the observed variables.

In D&A studies, such a source is unforced (control) climate model simulations. As a matter of fact, D&A studies assume that each $\sigma^2_{\nu_i}$ equals $\sigma^2_{\nu_0}$, which is justified by the assumption that the simulated internal climate variability is correctly simulated by the climate model used to generate x_i. Provided that all observed variables underwent pre-whitening, $\boldsymbol{\Sigma_{\nu\nu}}$ used for estimating the parameters of (2.6) is a $(m+1) \times (m+1)$ identity matrix ([1]).

In our turn, we would like to add a comment regarding the covariance matrix of latent $\xi^S_{i\,g}$:s, $\boldsymbol{\Sigma_{\xi\xi}}$. Unlike $\boldsymbol{\Sigma_{\nu\nu}}$, $\boldsymbol{\Sigma_{\xi\xi}}$ is unrestricted with the covariances as the off-diagonal elements. For reliable conclusions concerning the model parameters, it is important to check not only for consistency between the ME model and the data ([2]) but also whether the estimate of $\mathbf{\Sigma}_{\boldsymbol{\xi\xi}}$ is positive definite, i.e. nonsingular (for model (2.3) the corresponding requirement is $\sigma_\xi^2 > 0$). If these requirements are not met, uncertainties in $\widehat{\beta}_i$ can be very large (see [14] Sec. 1.3.3 and 2.3 for the appropriate tests for singularity).

Another possible cause of large uncertainties in $\widehat{\beta}_i$ can be small sample size. As mentioned earlier (in a footnote), after pre-whitening, the final sample size data analysed in D&A studies is often 10-15 observations on each observed variable, which is considered by many statisticians as extremely small. This is problematic because, under normality assumption, the distribution of the estimator of a regression parameter in a ME model is *asymptotically* normal, meaning that as a sample size increases the distribution converges to a normal distribution with a finite mean equal to the true parameter and finite variance. Thus a small sample size might lead to a poor approximation to the limiting normal distribution, which can lead to incorrect inferences (see an example for a similar discussion in [14], Sec. 2.5.1).

It is not uncommon for D&A studies that uncertainties in $\widehat{\beta}_i$ are determined from the empirical distribution of the estimator, formed by applying Monte-Carlo methods. According to [60], the empirical distribution of the estimator is constructed by calculating a very large number of estimates of β_i on the basis of data sets obtained by "superimposing different random samples of the model-based internal variability onto both noise-reduced Y_g and x_g". In our opinion, this approach requires care and should be used judiciously, because if the estimated internal temperature variability is substantially smaller than the variability of the noise-reduced Y_g and x_g, the estimates of β_i:s might be (highly) correlated. Clearly, the correlatedeness distorts the real distribution of $\widehat{\beta}_i$, which might result in misleading conclusions about the significance of $\widehat{\beta}$:s.

We conclude this section by discussing desirable modifications of models (2.3) and (2.6). To begin with, both models assume that the noise in the reconstructions or the observations, ν_0, does not include any non-climatic noise, or speaking in the terminology of D&A studies, observation error. According to [1], the main reason for this is that the autocorrelation structure of observation error is not assumed to be the same as that associated with the internal variability. Nevertheless, as known, non-climatic noise can constitute quite a large part of temperatures reconstructed from proxy data ([23]) and may also exist in varying amounts in instrumental temperature observations ([7], [49]). Therefore, it is of interest to relax the assumption of negligible non-climatic noise.

Further, one may wish to relax the assumption that the simulated internal climate variability is correctly simulated by the model used to generate x_i. Relaxing this assumption means that $\mathbf{\Sigma}_{\boldsymbol{\nu\nu}}$ remains a diagonal matrix but with different elements (variances) on the diagonal.

Yet another desirable modification, probably the most important one, concerns the inclusion of a term representing the temperature response to the possible interaction(-s) between forcings. As seen in (2.6), the model assumes the additivity of forcing effects.

The issue of additivity has been recognised and discussed by many researchers within the D&A field. Several analyses have been performed to investigate the significance of interactions in *simulated climate systems* (see, for example, [45], [60], and [16]). The results seem to support the assumption of additivity when it concerns temperature. Nevertheless, it would be beneficial to develop a statistical model allowing a simultaneous analysis of temperature responses to individual forcings and to various interactions between them, whether observed and/or reconstructed temperature is involved or not. Since such models apparently have not yet been seriously considered with respect to different climate variables, a theoretical statistical discussion of interaction effects in temperature data may pave the way for future studies focusing on, for example, precipitation or drought/wetness indices.

3 Overview of SUN12

The framework in SUN12, which has so far only been used in very few studies ([27], [48], [28], [53], [11]), is intended to be used with comprehensive climate models (GCMs, ESMs) and involves several components. They are:

- $x_{\mathsf{u}\,t}$ - a simulated temperature generated by an unforced climate model for the region and time period of interest, $t = 1, 2, \ldots, n$. This region may be a single grid box, or an average over several grid boxes [6]. The time unit can be single years or, say, nonoverlapping averages over a 10-yr or 30-yr period.
- $x_{\mathsf{f}\,t}$ - a simulated temperature generated by a climate model driven by a particular forcing f for a region and time period of interest, $t = 1, 2, \ldots, n$.
- τ_t - a true unobservable temperature corresponding to $x_{\mathsf{f}\,t}$.
- y_t - a measured temperature, intended to represent τ_t.
- z_t - a reconstructed temperature, derived from climate proxy data.

3.1 Unforced climate model

SUN12 used the term 'unforced climate model' or just 'unforced model' to denote a simulation with a climate model where no external forcing is invoked. More precisely, this means a simulation where the boundary conditions that are associated with the forcing factors of interest are held constant throughout the entire simulation time, at some level selected by the researcher. Climate modellers often refer to this kind of simulation as a control simulation. In this situation, only internal factors influence the

[6] As we see, given a collection of data points sampled over the region and time period of interest, D&A studies (see (2.1)) and SUN12 use different methods of summarising data information into a single data set that is to be analysed statistically. The differences between the methods will not be discussed further here.

simulated temperature variations. In SUN12, their effect is characterised as a random process varying around a time-constant mean value. That is, the simulated temperature, x_{u}, generated by the unforced climate model is given as:

$$x_{\mathsf{u}\ t} = \mu_{x_{\mathsf{u}}} + \delta_{\mathsf{u}\ t}. \tag{3.1}$$

where

$\mu_{x_{\mathsf{u}}}$ - the mean value around which the simulated temperatures, $x_{\mathsf{u}\ t}$, vary.

$\delta_{\mathsf{u}\ t}$ - the internal random variability of the unforced climate model.

3.2 Forced climate models

Next, assume that simulated temperatures generated in a forced climate model simulation, x_{f}, are obtained by running the same climate model again, but while replacing the control boundary conditions with a reconstruction of temporal and spatial changes in a particular forcing f. Climate modellers typically refer to this situation as a transient model simulation, while SUN12 referred to it as a 'forced climate model' or just 'forced model'.

As noted in SUN12, the forcing can be either of a single type (e.g. only volcanic forcing) or a combination of several forcings (e.g. volcanic and solar forcing). Nevertheless, the SUN12 study did not explicitly address cases when the forcing applied to a climate model is a single forcing or a combination of two forcings or more than two forcings. In what follows, we extend the interpretation of their framework by separately considering these three cases.

3.2.1 Case 1: The simulated forcing f is a single forcing.

Just as in D&A studies, the SUN12's definition of the simulated temperature generated by a forced climate model presupposes its decomposition into the forced and unforced components. More precisely, the temperature generated by a climate model driven by a single forcing is defined as follows:

$$x_{\mathsf{f}\ t} = \mu_{x_{\mathsf{f}}} + \xi^{S}_{\mathsf{f}\ t} + \tilde{\delta}_{\mathsf{f}\ t}, \tag{3.2.1}$$

where

$\mu_{x_{\mathsf{f}}}$ - the average of all processes embedded in the simulated temperature.

$\xi^{S}_{\mathsf{f}\ t}$ - the fixed Simulated effect of a reconstructed forcing, and centered to have zero mean.

$\tilde{\delta}_{\mathsf{f}\ t}$ - the internal random temperature variability of the forced climate model, *including* any random variability due to the presence of the forcing.

SUN12 also assumes that $\{\xi^S_{\mathsf{f}t}\}$- and $\{\tilde{\delta}_{\mathsf{f}t}\}$- processes are independent. Treating the forcing effect as fixed and systematic to a specified forcing means that it is repeatable, which justifies the assumption that it is the same for every member in a simulation ensemble [7](if an ensemble is available). Hence, repeatedness motivates averaging over members of a simulation ensemble. Averaging leads to a time series with a reduced effect of the internal variability and an enhanced forced signal. In contrast to the fixed forcing effect, the internal variability, characterised as random, is supposed to be different for different members in a simulation ensemble.

3.2.2 Case 2: The simulated forcing f is a combination of two forcings.

Let x_{f} in (3.2.1) represent a simulated temperature generated by a climate model driven jointly by two reconstructed forcings, labelled $\mathsf{f1}$ and $\mathsf{f2}$. Consequently, the simulated forcing effect $\boldsymbol{\xi}^S_{\mathsf{f}}$ is to be interpreted as the simulated *overall* effect of $\mathsf{f1}$ and $\mathsf{f2}$, arising when both forcings act in the presence of each other. Being present simultaneously, the forcings might interact with each other, leading to the following general representation of $\boldsymbol{\xi}^S_{\mathsf{f}}$:

$$\boldsymbol{\xi}^S_{\mathsf{f}} = \boldsymbol{\xi}^S_{\mathsf{f1}} + \boldsymbol{\xi}^S_{\mathsf{f2}} + \boldsymbol{\xi}^S_{\mathsf{f1}\times\mathsf{f2}}, \tag{3.2.2}$$

where

$\boldsymbol{\xi}^S_{\mathsf{f1}}$ - the temperature response to forcing $\mathsf{f1}$ in the absence of forcing $\mathsf{f2}$,

$\boldsymbol{\xi}^S_{\mathsf{f2}}$ - the temperature response to forcing $\mathsf{f2}$ in the absence of forcing $\mathsf{f1}$, and

$\boldsymbol{\xi}^S_{\mathsf{f1}\times\mathsf{f2}}$ - the temperature response to the interaction between $\mathsf{f1}$ and $\mathsf{f2}$.

Note that the general representation in (3.2.2) does not require specific assumptions about how strong the individual effects and/or interaction effect are. Assuming that one of them is negligible would of course change the representation of the overall forcing effect. Table 1 gives an overview of the representations with the associated assumptions.

[7]Members in a simulation ensemble are simulations forced by identical reconstructed forcing under different initial conditions. In the statistical context, members in a simulation ensemble are thought of as replicates.

Table 1. Representations of $\boldsymbol{\xi}_{\mathsf{f}}^{S}$ under different assumptions

Representation	Assumption
see (3.2.2)	No specific assumptions are made about the individual and interaction effects.
$\boldsymbol{\xi}_{\mathsf{f}}^{S} = \boldsymbol{\xi}_{\mathsf{f1}}^{S}$	The individual effect of the second forcing and the interaction effect are negligible.
$\boldsymbol{\xi}_{\mathsf{f}}^{S} = \boldsymbol{\xi}_{\mathsf{f2}}^{S}$	The individual effect of the first forcing and the interaction effect are negligible;
$\boldsymbol{\xi}_{\mathsf{f}}^{S} = \boldsymbol{\xi}_{\mathsf{f1}}^{S} + \boldsymbol{\xi}_{\mathsf{f2}}^{S}$	The individual forcing effects are additive, i.e. the interaction effect is negligible.

3.2.3 Case 3: The simulated forcing f is a combination of more than two forcings.

Clearly, the above discussion can be applied even to climate models driven by more than two forcings. For example, for three particular forcings, labelled f1, f2, and f3, their overall effect can simply be described in accordance with (3.2.2) as follows:

$$\boldsymbol{\xi}_{\mathsf{f}}^{S} = \sum_{j=1}^{3} \boldsymbol{\xi}_{\mathsf{f}_j}^{S} + \boldsymbol{\xi}_{\mathsf{interact}}^{S}, \tag{3.2.3}$$

where $\boldsymbol{\xi}_{\mathsf{f}_j}^{S}$ denotes the individual effect of forcing j in the absence of the other forcings, and $\boldsymbol{\xi}_{\mathsf{interact}}^{S}$ represents an overall interaction effect. In contrast to the interaction effect in (3.2.2), this overall interaction effect can have different structures/interpretations depending on in what way and to what extent the forcings interact with each other. Assuming, for example, that only forcings f1 and f2 considerably interact with one another, $\boldsymbol{\xi}_{\mathsf{interact}}^{S}$ represents then the corresponding two-forcing interaction effect, which we denote here $\boldsymbol{\xi}_{\mathsf{f1\times f2}}^{S}$.

Assuming further that all three forcings interact with each other, the complete structure of $\boldsymbol{\xi}_{\mathsf{interact}}^{S}$ becomes as follows:

$$\boldsymbol{\xi}_{\mathsf{interact}}^{S} = \boldsymbol{\xi}_{\mathsf{f1\times f2}}^{S} + \boldsymbol{\xi}_{\mathsf{f1\times f3}}^{S} + \boldsymbol{\xi}_{\mathsf{f2\times f3}}^{S} + \boldsymbol{\xi}_{\mathsf{f1\times f2\times f3}}^{S}. \tag{3.2.4}$$

Only if *all* individual forcing effects are additive, may we neglect all interaction terms in (3.2.4), and, consequently, the overall interaction effect in (3.2.3).

Expressions (3.2.3) and (3.2.4) are easily extended to climate models driven by more than three external forcings. The more forcings involved, the more complicated becomes the structure of the overall interaction term, since the number of possible ways to interact will increase.

Provided that model simulations forced by different sets of forcings are available, we can analyse the underlying structure of simulated temperature in forced climate models.

By involving further observational data, we can explore whether this structure coincides with the real-world structure associated with the true temperature.

3.3 The true temperature τ

Also the true temperature, influenced by all forcings simultaneously, can be decomposed into the forced and unforced components, namely as follows:

$$\tau_t = \mu_\tau + \xi^T_{\mathsf{ALL}\,t} + \eta_{\mathsf{internal}\ t}, \tag{3.3.1}$$

μ_τ - the mean value over time, around which the true temperature varies.

$\xi^T_{\mathsf{ALL}\,t}$ - the latent true overall temperature response to all forcings (the superscript T stands for *True*, not a transpose),

$\eta_{\mathsf{Internal}\ t}$ - the internal random variability of the real-world climate system, *including* any random variability due to the presence of the forcings.

Just as in the simulated climate system, the true overall forcing effect is regarded as fixed, given the actual realisation of the real-world external forcings generated by random processes in the real-world climate system. The internal variability is still treated as random. Also, these two processes are still regarded as mutually independent.

4 Factor models

Prior to reading the following sections, we would like to recommend the readers unfamiliar with factor analysis to read Appendix A where a brief account of a general factor model and the associated definitions, used in the following sections, are given.

4.1 Case 1: The forcing $\boldsymbol{f}$ is a single forcing

Being inherent to factor analysis, the concept of common factors lends itself perfectly to connecting simulated and observed climate data. In case of a single-forcing climate model, this concept implies that $\boldsymbol{\xi}^S_{\mathsf{f}}$, embedded in $\boldsymbol{x}_{\mathsf{f}}$, and $\boldsymbol{\xi}^T_{\mathsf{ALL}}$, embedded in $\boldsymbol{\tau}$, have one common factor. To avoid representing observational data as a function of simulated temperature responses to reconstructed forcings, we suggest to consider their true real-world counterparts as common factors for simulated and observed/reconstructed temperatures. As a matter of fact, this was suggested in SUN12, but without explaining all its consequences.

Viewing the true temperature response to forcing f, $\boldsymbol{\xi}^T_{\mathsf{f}}$, as a common factor means that $\boldsymbol{\xi}^S_{\mathsf{f}}$ and $\boldsymbol{\xi}^T_{\mathsf{ALL}}$ are expressed as linear functions of $\boldsymbol{\xi}^T_{\mathsf{f}}$, namely:

$$\begin{aligned} \boldsymbol{\xi}^S_{\mathsf{f}} &= \alpha \cdot \boldsymbol{\xi}^T_{\mathsf{f}} + \tilde{\boldsymbol{\zeta}}^S \\ \boldsymbol{\xi}^T_{\mathsf{ALL}} &= \kappa \cdot \boldsymbol{\xi}^T_{\mathsf{f}} + \tilde{\boldsymbol{\zeta}}^T, \end{aligned} \tag{4.1.1}$$

where $\tilde{\boldsymbol{\zeta}}^S$ and $\tilde{\boldsymbol{\zeta}}^T$ are assumed to be mutually independent residual terms which are independent from $\boldsymbol{\xi}_{\mathsf{f}}^T$. On comparison with (2.2), we see that introducing the residual terms in (4.1.1) allows us to relax two assumptions: *(i)* the *shape* of the latent temperature response is correctly simulated by the climate model under study, and that *(ii)* the real-world climate system is driven by the same forcings as those used to drive the simulated climate system.

Inserting (4.1.1) into the mean-centered $x_{\mathsf{f}\,t}$ and τ_t from (3.2.1) and (3.3.1), respectively, leads to the following system of two equations:

$$\begin{cases} x_{\mathsf{f}\,t} &= \alpha \cdot \xi_{\mathsf{f}\,t}^T &+ \delta_{\mathsf{f}\,t} \\ \tau_t &= \kappa \cdot \xi_{\mathsf{f}\,t}^T &+ \tilde{\nu}_t \end{cases} \tag{4.1.2}$$

where $\delta_{\mathsf{f}\,t} = \tilde{\zeta}_t^S + \tilde{\delta}_{\mathsf{f}\,t}$, $\tilde{\nu}_t = \tilde{\zeta}_t^T + \eta_{\mathsf{internal}\,t}$, each of which is independent of $\xi_{\mathsf{f}\,t}^T$, and the coefficients α and κ represent the influence of the common factor on x_{f} and τ, respectively. More precisely, they indicate the magnitude of the expected change in x_{f} and τ, respectively, for a one unit change in the common factor. This means that the evaluation of the climate model simulation can be addressed by testing whether α equals κ, or equivalently whether $\alpha/\kappa = 1$ or $\kappa/\alpha = 1$. Throughout the whole work, we refer to this hypothesis as the *hypothesis of consistency* between the climate model simulation under consideration and the observational data, or for short, the hypothesis of consistency.

However, since τ_t is unobservable, it has to be replaced by the observed climate record, which according to SUN12 includes the observed temperature y_t when it is available and a properly calibrated temperature proxy z_t when y_t is not available (see SUN12 for the description of the calibration method). Combining y_t and z_t leads to a single climate record, denoted $\{v_t\}$:

$$v_t = \begin{cases} z_t \approx \tau_t + \epsilon_t & t \in \text{the period when only } z \text{ is available,} \\ & \text{reconstruction period} \\ y_t = \tau_t + \theta_t & t \in \text{the period when both } y \text{ and } z \text{ are} \\ & \text{available, calibration period,} \end{cases} \tag{4.1.3}$$

where ϵ_t and θ_t represent the residual non-climatic variation in z_t and y_t, respectively. Quantities ϵ_t and θ_t are regarded as mutually uncorrelated random variables, each with zero mean and variances σ_ϵ^2 and σ_θ^2. Moreover, ϵ_t and θ_t are assumed to be uncorrelated with τ_t.

Replacing τ_t in (4.1.2) by v_t leads to the following two-indicator one-factor model, abbrv. FA(2,1):

$$\begin{cases} x_{\mathsf{f}\,t} &= \alpha \cdot \xi_{\mathsf{f}\,t}^T &+ \delta_{\mathsf{f}\,t} \\ v_t &= \kappa \cdot \xi_{\mathsf{f}\,t}^T &+ \nu_t \end{cases} \tag{4.1.4}$$

where

$$\nu_t = \begin{cases} \nu_{\mathsf{z}\,t} = \tilde{\nu}_t + \epsilon_t & t \in \text{the reconst. period} \\ \nu_{\mathsf{y}\,t} = \tilde{\nu}_t + \theta_t & t \in \text{the calibr. period.} \end{cases}$$

Because the simulated forcing effects are treated as fixed, $\xi_{\mathsf{f}\,t}$:s are consequently also fixed, which requires imposing the restrictions on the $\{\xi_{\mathsf{f}\,t}\}$-sequence analogous to those in (2.5). It should be remarked that the estimates of the factor loadings α and κ are the same regardless of whether $\xi_{\mathsf{f}\,t}$:s are random or fixed.

Concerning the specific-factor variables, $\delta_{\mathsf{f}\,t}$ and ν_t, each of them is assumed to be identically and independently normally distributed for all t. The requirement of being identically distributed presumes a time-constant variance. However, as follows from (4.1.4), this assumption is clearly violated for ν_t. Indeed, its variance equals $\sigma_\eta^2 + \sigma_\epsilon^2$ within the reconstruction period, while within the calibration period it equals $\sigma_\eta^2 + \sigma_\theta^2$. Moreover, the assumption may be violated even within the reconstruction period due to a possible time-dependent variability of the non-climatic noise, i.e. $\sigma_\epsilon^2(t)$ ([52], [49]). A discussion about how to take time-varying variances into account when estimating the FA(2,1)-model can be found in [11] (see Sec. 2.4).

Further, even the independence assumption, applied to climate data, seems to be inappropriate. As known, autocorrelation is an intrinsic feature of climate data. To avoid the effect of autocorrelation on the estimation of the model parameters, it can be recommended to analyse time aggregated data, for example, decadally resolved data (for examples of analyses involving time aggregated data see [27], [48], [11]). In the present paper, we assume homoscedasticity, i.e. $\sigma_\nu^2(t) = \sigma_\nu^2$ for all t, and no autocorrelation in $\{\delta_t\}$ and $\{\nu_t\}$.

Having specified the model and the hypothesis of interest, i.e. the hypothesis of consistency, the next step is to examine whether the model is identified, or stated another way, whether the model parameters are estimable. We devote considerable space to the discussion of identifiability of this factor model because it might substantially facilitate the understanding the causes of underidentifiability of more complicated factor models, discussed later.

To begin with, the covariance structure equations, $\mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$, under the FA(2,1)-model are given by the following $2(2+1)/2 = 3$ nonduplicated (unique) equations:

$$\begin{aligned} \sigma_{x_\mathsf{f}}^2 &= \alpha^2 \cdot \sigma_{\xi_\mathsf{f}}^2 + \sigma_{\delta_\mathsf{f}}^2 \\ \sigma_{x_\mathsf{f} v} &= \alpha \cdot \kappa \cdot \sigma_{\xi_\mathsf{f}}^2 \\ \sigma_v^2 &= \kappa^2 \cdot \sigma_{\xi_\mathsf{f}}^2 + \sigma_\nu^2, \end{aligned} \tag{4.1.5}$$

It is clear that none of the five model parameters, α, κ, $\sigma_{\xi_\mathsf{f}}^2$, $\sigma_{\delta_\mathsf{f}}^2$ and σ_ν^2, can be determined (identified) from the three equations. However, imposing the restriction

$\sigma^2_{\xi_\mathsf{f}} = 1$ [8], and assuming that $\sigma^2_{\delta_\mathsf{f}}$ is known a priori [9], the remaining model parameters, i.e. α, κ, and σ^2_ν, become identified. More precisely, each of them is just-identified because only one distinct subset of equations can be found in (4.1.5) that is uniquely solvable for α, κ, and σ^2_ν, respectively. The resulting solutions are:

$$\begin{aligned} \alpha &= \sqrt{\sigma^2_{x_\mathsf{f}} - \sigma^2_{\delta_\mathsf{f}}} \\ \kappa &= \sigma_{x_\mathsf{f}v} \Big/ \sqrt{\sigma^2_{x_\mathsf{f}} - \sigma^2_{\delta_\mathsf{f}}} \\ \sigma^2_\nu &= \sigma^2_z - (\sigma_{x_\mathsf{f}v})^2 / (\sigma^2_{x_\mathsf{f}} - \sigma^2_{\delta_\mathsf{f}}). \end{aligned} \tag{4.1.6}$$

Replacing the population variances and covariance of the indicators by their unbiased estimates, $s^2_{x_\mathsf{f}}$, $s_{x_\mathsf{f}z}$ and s^2_z, the exact ML solution of the model parameters is obtained:

$$\begin{aligned} \widehat{\alpha} &= \sqrt{s^2_{x_\mathsf{f}} - \sigma^2_{\delta_\mathsf{f}}} \\ \widehat{\kappa} &= s_{x_\mathsf{f}\,v} \Big/ \sqrt{s^2_{x_\mathsf{f}} - \sigma^2_{\delta_\mathsf{f}}} \\ \widehat{\sigma}^2_\nu &= s^2_v - \widehat{\kappa}^2, \end{aligned} \tag{4.1.7}$$

provided $s^2_{x_\mathsf{f}} > \sigma^2_{\delta_\mathsf{f}}$, and $s^2_v - (s_{x_\mathsf{f}v})^2/(s^2_{x_\mathsf{f}} - \sigma^2_{\delta_\mathsf{f}}) \geq 0$. As mentioned earlier, this solution is unique, apart from a possible change of sign of the factor loadings, which merely corresponds to changing the sign of the factor. Using (4.1.7), the estimator of the ratio becomes

$$\frac{\widehat{\alpha}}{\widehat{\kappa}} = \frac{s^2_{x_\mathsf{f}} - \sigma^2_\delta}{s_{x_\mathsf{f}v}}, \tag{4.1.8}$$

provided $s_{x_\mathsf{f}v} \neq 0$.

Although there are methods for constructing a confidence region for a ratio allowing us to test the hypothesis that $\alpha/\kappa = 1$ (see the Fieller method used in [11]), our intention in the present work is to use the properties of confirmatory factor analysis to the full extent. Here, we first and foremost mean the usage of equality-constraints that can be imposed in accordance with hypotheses the researcher has. In our own analysis, introducing various equality-constraints may simplify considerably both single and multiple tests concerning different pairs of factor loadings.

Under the FA(2,1)-model, the hypothesis $H_0 : \alpha/\kappa = 1$ can be tested by estimating the model under the restriction $\alpha = \kappa$. Imposing this equality-constraint makes the model overidentified with 1 degree of freedom because the two remaining free parameters,

[8] Standardising latent factors to have unit variance is a typical restriction in factor analysis used to assign a scale to latent factors to fully interpret the factor loadings. As we see, it aids the identification as well. Another way of establishing a scale for a latent factor is to fix a factor loading to unity with respect to one of its indicators.

[9] By saying that $\sigma^2_{\delta_\mathsf{f}}$ is known a priori, we mean that $\sigma^2_{\delta_\mathsf{f}}$ is independently estimated. In Appendix B, we suggest a possible independent estimator of $\sigma^2_{\delta_\mathsf{f}}$ based on the assumption that an ensemble with at least two simulations of x_f is available.

κ and σ^2_ν, become overidentified. This means, one can find more than one subset of the equations in $\mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$ by which one can solve uniquely for κ and σ^2_ν. Indeed, when $\sigma^2_{\xi_\mathsf{f}} = 1$ and $\sigma^2_{\delta_\mathsf{f}}$ is known a priori, setting $\alpha = \kappa$ in (4.1.5) leads to two distinct solutions for each parameter:

$$\begin{aligned} \kappa &= \sqrt{\sigma^2_{x_\mathsf{f}} - \sigma^2_{\delta_\mathsf{f}}} = \sqrt{\sigma_{x_\mathsf{f}\,v}} \\ \sigma^2_\nu &= \sigma^2_v - (\sigma^2_{x_\mathsf{f}} - \sigma^2_{\delta_\mathsf{f}}) = \sigma^2_v - \sigma_{x_\mathsf{f}\,v}. \end{aligned} \tag{4.1.9}$$

Obviously, for a given sample, the two estimates of each parameter will not be exactly the same. Therefore, the ML-method seeks the optimal values for the parameters by minimising numerically the discrepancy function defined in (A.2). Overidentifiability means that $\mathbf{\Sigma}(\widehat{\boldsymbol{\theta}})$ does not fit the data, i.e. the sample variance-covariance matrix $\boldsymbol{S}$, perfectly, permitting us to assess the overall model fit. Provided that the solution obtained is admissible, which here means that $\widehat{\sigma}^2_\nu > 0$ [10], the model fit can be assessed statistically by the χ^2 test (see (A.3)) and heuristically by various goodness-of fit indexes (e.g. (A.5)-(A.7)). If the model fits reasonably well, we may say that there is no reason to reject the hypothesis that $\alpha = \kappa$.

Discussing further the identifiability, it should be remarked that setting α in (4.1.5) to zero makes the FA(2,1)-model underidentified because both κ and σ^2_ν become underidentified (remember that $\sigma^2_{\xi_\mathsf{f}}$ is restricted to 1). Hypotheses that do not lead to the underidentifiability of the FA(2,1)-model and therefore are testable are:

- $H_0 : \alpha = \kappa = 0$, under which the only parameter σ^2_ν is identified because it still can be determined from σ^2_ν. The resulting model has 2 degrees of freedom;
- $H_0 : \kappa = 0$, under which both remaining parameters, α and σ^2_ν, are identified. The resulting model has 1 degree of freedom.

Under both hypotheses, the indicators are uncorrelated, i.e. $\sigma_{x_\mathsf{f}\,v} = 0$, because they do not have any latent factor in common. Notice, however, that despite this similarity, the hypothesised models may have different interpretations of the ability of the climate model under study to represent the temperature response to forcing f correctly. The difference arises when the estimate of the free parameter α in the model hypothesising that $\kappa = 0$ significantly different from zero. If the resulting FA(2,1)-model is not rejected, a natural conclusion is that the climate model has failed to represent the temperature response to forcing f correctly because the influence of ξ_f on x_f is exaggerated compared to its influence on v. On the other hand, under the hypothesis $H_0 : \alpha = \kappa = 0$, it is the other way round. Provided that the resulting 0-factor model is not rejected, we may conclude that the climate model has succeeded to represent the simulated forcing effect correctly because the influence is equally negligible, that is, it is the same.

Finally, we would like to discuss the appropriateness of the use of the independent estimator of $\sigma^2_{\delta_\mathsf{f}}$ from (B.1). As follows from (4.1.2), the specific factor δ_f includes not only

[10] In statistical litterature, a negative solution for a specific-factor variance is termed Heywood case.

the internal variability of the corresponding climate model, represented by $\tilde{\delta}_{\mathsf{f}}$, but also a part of the forced variability that does not have a correspondence in the variability in the observed climate, represented by $\tilde{\zeta}^S$. However, estimator (B.1) provides an independent estimate of the variance of $\tilde{\delta}_{\mathsf{f}}$ only. In other words, we constrain the variance $\sigma^2_{\tilde{\delta}_{\mathsf{f}}}$ to its lower bound, which can lead to a serious bias for some of the parameters [11].

A possible way to investigate whether the variance of $\tilde{\zeta}^S$ may be considered as negligible, or equivalently whether the shape of the latent temperature response is correctly simulated by the climate model under consideration, is to use replicates of x_{f} as indicators of ξ^T_{f}. To exemplify, let us assume that two replicates of the x_{f}- climate model simulation are available. Denoting them $x_{\mathsf{f\,repl.1}}$ and $x_{\mathsf{f\,repl.2}}$, respectively, the FA(2,1)-model from (4.1.4) can be extended to the following three indicator model:

$$\begin{cases} x_{\mathsf{f\,repl.\ 1}\,t} &= \alpha \cdot \xi^T_{\mathsf{f}\,t} &+ \tilde{\delta}_{\mathsf{f\,repl.1}\,t} \\ x_{\mathsf{f\,repl.\ 2}\,t} &= \alpha \cdot \xi^T_{\mathsf{f}\,t} &+ \tilde{\delta}_{\mathsf{f\,repl.2}\,t} \\ v_t &= \kappa \cdot \xi^T_{\mathsf{f}\,t} &+ \nu_t, \end{cases} \tag{4.1.10}$$

where $\sigma^2_{\xi_{\mathsf{f}}} = 1$ and $\sigma^2_{\tilde{\delta}_{\mathsf{f\,repl.\ 1}}} = \sigma^2_{\tilde{\delta}_{\mathsf{f\,repl.\ 2}}} = \sigma^2_{\tilde{\delta}_{\mathsf{f}}}$. As we see, model (4.1.10) assumes that (1) the variance of $\tilde{\zeta}^S$ is zero, and (2) the influence of the common factor on $x_{\mathsf{f\,repl.\ 1}}$ and on $x_{\mathsf{f\,repl.\ 2}}$ is equal - both factor loadings are equal to α. This equality constraint is justified by the fact that both replicates are forced by the same reconstruction of forcing f. Hence, the model has four free parameters to be estimated: α, κ, $\sigma^2_{\tilde{\delta}_{\mathsf{f}}}$ and σ^2_ν. Provided that each of them is identified, the model has $6 - 4 = 2$ degrees of freedom, where 6 is the number of nonduplicated equations in the reproduced variance-covariance matrix $\mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$. To see that each parameter is indeed identified consider the equations in $\mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$:

$$\begin{array}{lll} \sigma^2_{x_{\mathrm{f\,repl.1}}} = \alpha^2 + \sigma^2_{\tilde{\delta}_{\mathrm{f}}} & \sigma^2_{x_{\mathrm{f\,repl.1}}\,x_{\mathrm{f\,repl.2}}} = \alpha^2 & \sigma_{x_{\mathrm{f\,repl.1}}v} = \alpha \cdot \kappa \\ & \sigma^2_{x_{\mathrm{f\,repl.2}}} = \alpha^2 + \sigma^2_{\tilde{\delta}_{\mathrm{f}}} & \sigma_{x_{\mathrm{f\,repl.2}}v} = \alpha \cdot \kappa \\ & & \sigma^2_v = \kappa^2 + \sigma^2_\nu. \end{array} \tag{4.1.11}$$

The estimator of the factor loading α can be obtained from $\sigma^2_{x_{\mathrm{f\,repl.1}}\,x_{\mathrm{f\,repl.2}}}$. Knowing the estimator of α, we can derive the estimator of κ and of $\sigma^2_{\tilde{\delta}_{\mathsf{f}}}$, each of which is associated with two equations, meaning that both parameters are overidentified. Finally, with the estimator of κ in hand, the estimator of σ^2_ν is derived from the equation for σ^2_v. If, in addition, α is set to κ, one more degree of freedom is obtained. If the resulting FA(3,1)-model is not rejected, we may conclude that not only the shape but also the magnitude of the latent temperature response is correctly simulated by the climate model under consideration.

[11] This suggests that the issue of how serious the bias is and which parameters are substantially affected by it is an appropriate topic for future studies.

Importantly, freeing the variance of $\tilde{\zeta}^S$ implies the correlatedeness between $\delta_{\mathsf{f\,repl.1}}$ and $\delta_{\mathsf{f\,repl.2}}$, each of which contains $\tilde{\zeta}^S$. However, letting the covariance between $\delta_{\mathsf{f\,repl.1}}$ and $\delta_{\mathsf{f\,repl.2}}$ be a model parameter leads to underidentifiability.

If model (4.1.10) (whether the constraint $\alpha = \kappa$ is imposed or not) is *not* rejected, we have no reason to reject the hypothesis that the variance of $\tilde{\zeta}^S$ is zero. This increases our confidence in the independent estimator (B.1). In addition, it gives us an opportunity to explore the stability of the estimates of the factor loadings, in particular of α, in comparison to those obtained when $\sigma^2_{\tilde{\delta}_\mathsf{f}}$ is treated as known a priori. It should also be remarked that even if model (4.1.10) is not rejected, care is needed in interpreting the results obtained, especially in the cases with temperature responses expected to be correlated to other temperature responses, embedded in v, but not extracted as ξ^T_f. This may lead to ξ^T_f not representing the effect of forcing f only. The same issue can arise under the FA(2,1)-model.

Note, that in general the rejection of a hypothesised factor model does not unambiguosly point to any particular constraint as at fault ([50]). Other aspects such as small sample size, nonnormality, or missing data can also cause lack of fit. Hence, rejecting model (4.1.10) does not unambiguosly point to the constraint $\sigma^2_{\tilde{\zeta}^S} = 0$ as at fault, albeit it decreases our confidence in the independent estimate of $\sigma^2_{\delta_\mathsf{f}}$.

4.2 Case 2: The simulated forcing *f* is a combination of two forcings

As follows from Table 1 (Sec. 2.2.2), one-, two- and even three-dimensional structures of the simulated forcing effect, $\boldsymbol{\xi}^S_\mathsf{f}$, can be justified for two-forcing climate model simulations. Consequently, the underlying common-factor structure of $\boldsymbol{\xi}^S_\mathsf{f}$ and $\boldsymbol{\xi}^T_\mathsf{ALL}$ can contain one-, two- or three common factors. Without making specific assumptions about the significance of the individual and interaction effects, the underlying common factor structure of $\boldsymbol{\xi}^S_\mathsf{f}$ and $\boldsymbol{\xi}^T_\mathsf{ALL}$ contain three common factors: $\boldsymbol{\xi}^T_\mathsf{f1}$, $\boldsymbol{\xi}^T_\mathsf{f2}$ and $\boldsymbol{\xi}^T_{\mathsf{f1}\times\mathsf{f2}}$. Extracting them from $\boldsymbol{\xi}^S_\mathsf{f}$ and $\boldsymbol{\xi}^T_\mathsf{ALL}$ leads to:

$$\begin{aligned} \boldsymbol{\xi}^S_\mathsf{f} &= \alpha_1 \cdot \boldsymbol{\xi}^T_\mathsf{f1} + \alpha_2 \cdot \boldsymbol{\xi}^T_\mathsf{f2} + \alpha_3 \cdot \boldsymbol{\xi}^T_{\mathsf{f1}\times\mathsf{f2}} + \tilde{\boldsymbol{\zeta}}^S \\ \boldsymbol{\xi}^T_\mathsf{ALL} &= \kappa_1 \cdot \boldsymbol{\xi}^T_\mathsf{f1} + \kappa_2 \cdot \boldsymbol{\xi}^T_\mathsf{f2} + \kappa_3 \cdot \boldsymbol{\xi}^T_{\mathsf{f1}\times\mathsf{f2}} + \tilde{\boldsymbol{\zeta}}^T. \end{aligned} \tag{4.2.1}$$

where $\tilde{\boldsymbol{\zeta}}^S$ and $\tilde{\boldsymbol{\zeta}}^T$ are assumed to be mutually independent, and independent from each common factor. Note that these residual ζ-terms differ from those in (4.1.1). To avoid additional notations, we do not use new symbols to stand for different errors, and we hope that the reader keeps in mind that as a new underlying structure of ξ^S_f and ξ^T_ALL is considered, the corresponding $\tilde{\boldsymbol{\zeta}}$:s change.

To make the estimation of the factor loadings in (4.2.1) feasible, we need to define as many appropriate indicators of the latent common factors as possible. Suitable candidates are:

- the observed/reconstructed temperature, v;
- the simulated temperature forced by both forcings, x_{f};
- the simulated temperatures generated by *each* of the two single-forcing climate model, x_{f1} and x_{f2}. Importantly, x_{f1} and x_{f2} should be forced by the same reconstruction of f1 and f2, respectively, used to force x_{f}. Under this condition, extracting the common factor $\boldsymbol{\xi}^T_{\mathsf{fi}}$ from $\boldsymbol{\xi}^S_{\mathsf{fi}}$, embedded in the corresponding $\boldsymbol{x}_{\mathsf{fi}}$, leads to

$$x_{\mathsf{fi}\,t} = \xi^S_{\mathsf{fi}\,t} + \tilde{\delta}_{\mathsf{fi}\,t} = \alpha_i \cdot \xi^T_{\mathsf{fi}\,t} + \underbrace{(\zeta^S_{\mathsf{fi}\,t} + \tilde{\delta}_{\mathsf{fi}\,t})}_{=\delta_{\mathsf{fi}\,t}}, \quad i = 1, 2. \tag{4.2.2}$$

Invoking x_{f1}, x_{f2}, x_{f} and v as indicators of ξ^T_{f1}, ξ^T_{f2}, and $\xi^T_{\mathsf{f1\times f2}}$, and extracting them from ξ^S_{f1}, ξ^S_{f2}, ξ^S_{f} and ξ^T_{ALL}, as shown in (4.2.2), leads to the following 4-indicator 3-factor model, abbrv. FA(4,3):

$$\begin{cases} x_{\mathsf{f1}\,t} &= \alpha_1 \cdot \xi^T_{\mathsf{f1}\,t} + 0 \cdot \xi^T_{\mathsf{f2}\,t} + 0 \cdot \xi^T_{\mathsf{f1\times f2}\,t} + \delta_{\mathsf{f1}\,t} \\ x_{\mathsf{f2}\,t} &= 0 \cdot \xi^T_{\mathsf{f1}\,t} + \alpha_2 \cdot \xi^T_{\mathsf{f2}\,t} + 0 \cdot \xi^T_{\mathsf{f1\times f2}\,t} + \delta_{\mathsf{f2}\,t} \\ x_{\mathsf{f}\,t} &= \alpha_1 \cdot \xi^T_{\mathsf{f1}\,t} + \alpha_2 \cdot \xi_{\mathsf{f2}\,t} + \alpha_3 \cdot \xi^T_{\mathsf{f1\times f2}\,t} + \delta_{\mathsf{f}\,t} \\ v_t &= \kappa_1 \cdot \xi^T_{\mathsf{f1}\,t} + \kappa_2 \cdot \xi^T_{\mathsf{f2}\,t} + \kappa_3 \cdot \xi^T_{\mathsf{f1\times f2}\,t} + \nu_t. \end{cases} \tag{4.2.3}$$

Correlations among Common Factors

1	ϕ_{12}	ϕ_{13}
	1	ϕ_{23}
		1

where each of the specific factors, δ_{f1}, δ_{f2}, δ_{f} and ν, has its own variance, and are assumed to be mutually independent and independent of all common factors. Note that the zero factor loadings reflect our conviction that x_{f1} and x_{f2} do not depend on ξ^T_{f2} or ξ^T_{f1}, respectively, or on $\xi^T_{\mathsf{f1\times f2}}$. Note also that because $\tilde{\zeta}^T$ in (4.2.1) differs from $\tilde{\zeta}^T$ in (4.1.1), the specific factor ν in FA(4,3)-model in (4.2.3) differs from ν in the FA(2,1)-model in (4.1.4), albeit both ν:s contain the same η_{internal} and the same non-climatic noise.

Next, the identifiability of the model ought to be worked out. Here, we distinguish between two main types of the model: oblique and orthogonal. Depending on the climatological characteristics of the forcings under consideration, researchers might wish to specify one of the models. A third conceivable type is a mixture of the orthogonal and oblique model.

Common to all three types is that *(i)* the latent factors are standardised to have unit variances [12], *(ii)* the associated variance-covariance matrix of the indicators has 10 nonduplicated equations, meaning that at most 10 parameters can be estimated, and finally *(iii)* the conditions for uniqueness under rotation, given in Appendix A, are met.

The oblique model

The model parameters are: six factor loadings (each distinct constrained loading is counted only once), three correlation coefficients, and σ_ν^2. In total, 10 parameters have to be estimated, which is possible if $\sigma^2_{\delta_{\mathsf{f1}}}$, $\sigma^2_{\delta_{\mathsf{f2}}}$ and $\sigma^2_{\delta_{\mathsf{f}}}$ are known a priori. It can be shown that under this condition the model is just-identified, provided however that neither α_1, α_2 nor α_3 is equal to zero. The hypothesis of consistency between the climate model simulations in question and the observed climate record is tested by setting $\alpha_i = \kappa_i$, $i = 1, 2, 3$. It will give us three degrees of freedom.

The FA(4,3)-model remains identified with three degrees of freedom when the restrictions $\kappa_i = 0$, $i = 1, 2, 3$, are imposed.

To be able to test the hypothesis that the interation term is negligible, it is not sufficient to set only α_3 and κ_3 to zero. The two associated correlations, ϕ_{13} and ϕ_{23}, should be also eliminated from the vector of the model parameters. The resulting model, the FA(4,2)-model, is thus over-identified with four degrees of freedom, provided, of course, that α_1 and α_2 differ from zero. Under the hypothesis of consistency, the model has six degrees of freedom.

The simplest model, nested within model (4.2.3), is the so-called null model or baseline model. A general baseline model supposes that only the specific-factor variances are estimated. In our case, a more appropriate baseline model is a model where only σ_ν^2 is estimated. Consequently, this model has only 1 free parameter and 9 degrees of freedom, and it hypothesises that the observed variables do not have anything in common, that is, they are uncorrelated.

If the null model is rejected statistically and/or heuristically, researchers may test whether one latent factor, representing one of the two individual forcing effects, ξ^T_{f1} or ξ^T_{f2}, is sufficient to account for the intercorrelations between the indicators. To exemplify, we present a 1-factor model with ξ^T_{f2}:

$$\begin{cases} x_{\mathsf{f1}\,t} &= 0 \cdot \xi^T_{\mathsf{f2}\,t} + \delta_{\mathsf{f1}\,t} \\ x_{\mathsf{f2}\,t} &= \alpha_2 \cdot \xi^T_{\mathsf{f2}\,t} + \delta_{\mathsf{f2}\,t} \\ x_{\mathsf{f}\,t} &= \alpha_2 \cdot \xi^T_{\mathsf{f2}\,t} + \delta_{\mathsf{f}\,t} \\ v_t &= \kappa_2 \cdot \xi^T_{\mathsf{f2}\,t} + \nu_t, \end{cases} \tag{4.2.4}$$

The model has three free parameters, α_2, κ_2 and σ_ν^2. Since each is identified, the

[12] Recall that the common factors are regarded as fixed unknown constants, meaning that their variances (and covariances) are interpreted in terms of limits analogous to the second limit in (2.5).

model has $10 - 3 = 7$ degrees of freedom. Notice that the structure of the models reproduced variance-covariance matrix, $\mathbf{\Sigma}(\boldsymbol{\theta})$ (not shown here completely) includes the following equations: $\sigma^2_{x_{\mathsf{f1}}\,x_{\mathsf{f2}}} = \sigma^2_{x_{\mathsf{f1}}\,x_{\mathsf{f}}} = \sigma^2_{x_{\mathsf{f1}}\,v} = 0$. This means that accepting this factor model is equivalent to accepting the hypothesis that x_{f1} is uncorrelated with all other observed variables.

Model (4.2.4) becomes theoretically underidentified if α_2 is zero (both κ_2 and σ^2_ν become underidentified), whereas the identifiability is retained under the restrictions $\alpha_2 = \kappa_2$, $\alpha_2 = \kappa_2 = 0$ and $\kappa_2 = 0$, giving us 1, 2 and 1 additional degrees of freedom, respectively.

In practice, nonsignificant estimates of a model parameter, whose population value makes a factor model theoretically underidentified, may cause empirical underidentification ([55]). Undesirable results associated with this phenomenon include negative specific-factor variances (Heywood case), parameters that are outside reasonable limits (for example, correlations among latent factors exceeding 1), large standard errors of parameter estimates, large correlations among parameter estimates, and even failure of the estimation algorithm to converge to a solution. Therefore, it is important prior to estimating a factor model to determine the causes of its theoretical underidentifiaction to undertake justified empirical modifications of the model in case they are needed.

The orthogonal model

If the researcher is convinced that the latent temperature responses in (4.2.3) are uncorrelated, then he or she gets an opportunity to free up *two* specific-factor variances, namely $\sigma^2_{\delta_{\mathsf{f1}}}$ and $\sigma^2_{\delta_{\mathsf{f2}}}$, without assuming that $\sigma^2_{\tilde{\zeta}^S} = 0$ as under model (4.1.10). Together with other model parameters, α_1, κ_1, α_2, κ_2, α_3, κ_3 and σ^2_ν, it gives 9 parameters in total. As follows from the reproduced variance-covariance matrix of the orthogonal FA(4,3)-model, given in (4.2.5), each of the 9 parameters is identified [13]. Thus, the model is overidentified with 1 degree of freedom. If $\sigma^2_{\delta_{\mathsf{f1}}}$ and $\sigma^2_{\delta_{\mathsf{f2}}}$ are still treated as known a priori, then two additional degrees of freedom are gained.

$$\begin{array}{llll}
\sigma_1^2 = \alpha_1^2 + \sigma^2_{\delta_{\mathsf{f1}}} & & & \\
\sigma_{21} = 0 & \sigma_2^2 = \alpha_2^2 + \sigma^2_{\delta_{\mathsf{f2}}} & & \\
\sigma_{31} = \alpha_1^2 & \sigma_{32} = \alpha_2^2 & \sigma_3^2 = \alpha_1^2 + \alpha_2^2 + \alpha_3^2 + \sigma^2_{\delta_{\mathsf{f}}} & \\
\sigma_{41} = \alpha_1\,\kappa_1 & \sigma^2_{42} = \alpha_2\,\kappa_2 & \sigma_{43} = \alpha_1\,\kappa_1 + \alpha_2\,\kappa_2 + \alpha_3\,\kappa_3 & \sigma_4^2 = \kappa_1^2 + \kappa_2^2 + \kappa_3^2 + \sigma^2_\nu
\end{array} \tag{4.2.5}$$

As follows from (4.2.5), the orthogonal FA(4,3)-model becomes underidentified if either α_1, α_2 or α_3 equals zero (the associated κ:s and σ^2_ν cannot be identified). At the same time, the underidentification does not arise under the following restrictions: $\kappa_i = 0$ and $\alpha_i = \kappa_i = 0$ for each i. Most importantly, the orthogonal FA(4,3)-model is also identified under the hypothesis of consistency, that is, if the equality constraints $\alpha_i = \kappa_i$,

[13] We also see that freeing up $\sigma^2_{\delta_{\mathsf{f}}}$ makes the model underidentified, because $\sigma^2_{\delta_{\mathsf{f}}}$ should then be determined from the same equation as α_3.

$i = 1, 2, 3$, are imposed.

The mixed model
The properties of this model is a mixture of the properties of the above-discussed oblique and orthogonal models. Therefore, we refrain from repeating them.

4.3 Case 3: The forcing f is a combination of more than two forcings

Our aim in this section is to formulate factor models that can be applied for the evaluation of climate model simulations driven by more than two external forcings. We start with climate model simulations driven by three external forcings, i.e. $\mathsf{f} := (\mathsf{f1}, \mathsf{f2}, \mathsf{f3})$.

Let the simulated overall temperature response to f be decomposed as shown in (3.2.3) and (3.2.4). This decomposition means that the general underlying common-factor structure of ξ^S_{f} and ξ^T_{ALL} comprises seven common factors, namely ξ^T_{f1}, ξ^T_{f2}, ξ^T_{f3}, $\xi^T_{\mathsf{f1\times f2}}$, $\xi^T_{\mathsf{f1\times f3}}$, $\xi^T_{\mathsf{f2\times f3}}$, and $\xi^T_{\mathsf{f1\times f2\times f3}}$. Let them be extracted from ξ^S_{f} and ξ^T_{ALL} in a similar way as in (4.2.1). Using the same principles for choosing suitable indicators of latent factors as in the case of the FA(4,3)-model, we could formulate a basic 8-indicator 7-factor model given in Table 3. Thanks to the equality constraints imposed on the factor loadings [14], the model satisfies the sufficient conditions for uniqueness under rotation (see Appendix A) regardless of whether the model is oblique or orthogonal.

Without imposing restrictions on the correlations among the common factors, the FA(8,7)-model is oblique. Since the number of unique equations in $\mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$ is $8 \cdot 9/2 = 36$, at most 36 parameters can be estimated. Without imposing restrictions on the factor loadings in accordance with the hypothesis of consistency, we have to estimate $7 \times 2 = 14$ distinct factor loadings, $\binom{7}{2} = \frac{7\cdot 6}{2\cdot 1} = 21$ correlations among the latent factors and σ^2_ν, that is, 36 parameters. It can be shown that each of them is just-identified, implying that the whole model is just-identified. Imposing the equality-constraints in accordance with the hypothesis of consistency, i.e. $\alpha_i = \kappa_i$ makes the model overidentified with seven degrees of freedom.

The causes of underidentifiability of the oblique FA(8,7)-model are similar to those associated with the oblique FA(4,3)-model. That is, underidentifiability arises when $\alpha_i = 0$ or $\alpha_i = \kappa_i = 0$ for at least one i. Imposing these restrictions requires the elimination of the factor ξ_{fi} with all associated correlation coefficients.

[14] Just as in the FA(4,3)-model, the equality constraints in a given column in the FA(8,7)-model are motivated by the fact that the same reconstruction of forcing i is identically implemented in the associated single-, two- and three-forcing climate models.

Table 3. *Parameters of a basic 8-indicator 7-factor model, abbr. FA(8,7), where x_{f} stands for a simulated temperature forced by three forcings: f1, f2, and f3.*

Indicator			Common	factors			Specific-
	factor 1	factor 2	factor 3	factor 4	factor 5	factor 6	factor 7	-factor
	$\boldsymbol{\xi}^T_{\mathsf{f1}}$	$\boldsymbol{\xi}^T_{\mathsf{f2}}$	$\boldsymbol{\xi}^T_{\mathsf{f3}}$	$\boldsymbol{\xi}^T_{\mathsf{f1\times f2}}$	$\boldsymbol{\xi}^T_{\mathsf{f1\times f3}}$	$\boldsymbol{\xi}^T_{\mathsf{f2\times f3}}$	$\boldsymbol{\xi}^T_{\mathsf{f1\times f2\times f3}}$	variances
1. x_{f1}	α_1	0	0	0	0	0	0	$\sigma^{2\,*}_{\delta_{\mathsf{f1}}}$
2. x_{f2}	0	α_2	0	0	0	0	0	$\sigma^{2\,*}_{\delta_{\mathsf{f2}}}$
3. x_{f3}	0	0	α_3	0	0	0	0	$\sigma^{2\,*}_{\delta_{\mathsf{f3}}}$
4. $x_{\mathsf{f1\,f2}}$	α_1	α_2	0	α_4	0	0	0	$\sigma^{2\,*}_{\delta_{\mathsf{f1\,f2}}}$
5. $x_{\mathsf{f1\,f3}}$	α_1	0	α_3	0	α_5	0	0	$\sigma^{2\,*}_{\delta_{\mathsf{f1\,f3}}}$
6. $x_{\mathsf{f2\,f3}}$	0	α_2	α_3	0	0	α_6	0	$\sigma^{2\,*}_{\delta_{\mathsf{f2\,f3}}}$
7. x_{f}	α_1	α_2	α_3	α_4	α_5	α_6	α_7	$\sigma^{2\,*}_{\delta_{\mathsf{f}}}$
8. v	κ_1	κ_2	κ_3	κ_4	κ_5	κ_6	κ_7	σ^2_{ν}
	Correlations among Common Factors							
	1	ϕ_{12}	ϕ_{13}	ϕ_{14}	ϕ_{15}	ϕ_{16}	ϕ_{17}	
		1	ϕ_{23}	ϕ_{24}	ϕ_{25}	ϕ_{26}	ϕ_{27}	
			1	ϕ_{34}	ϕ_{35}	ϕ_{36}	ϕ_{37}	
				1	ϕ_{45}	ϕ_{46}	ϕ_{47}	
					1	ϕ_{56}	ϕ_{57}	
						1	ϕ_{67}	
							1	

* the parameter assumed to be known a priori.

However, in contrast to the oblique FA(4,3)-model, the large number of common factors in the oblique FA(8,7)-model provides an opportunity to free all specific-factor variances associated with all climate model simulations. This can be done by eliminating the factor representing the 3-forcing interaction effect, $\xi^T_{\mathsf{f1\times f2\times f3}}$. Since the individual forcing effects and the interaction between forcings (in terms of pairwise interaction effects) are still accounted for, the resulting FA(8,6)-model is still realistic and defensible from the climatological standpoint even after this respecification.

Naturally, the elimination of $\boldsymbol{\xi}^T_{\mathsf{f1\times f2\times f3}}$ entails that α_7, κ_7, and all associated correlations, i.e. $\phi_{17}, \ldots, \phi_{67}$, are eliminated as well. Since all remaining parameters can be identified, the FA(8,6)-model with all specific-factor variances treated as free is overidentified with 1 degree of freedom. If the model is not be rejected then it is possible to test (1) the hypothesis of consistency by imposing $\alpha_i = \kappa_i$, $i = 1, 2, \ldots, 6$, which results in six additional degrees of freedom, and (2) other conceivable simplifications, e.g. setting insignificant correlations (if such are observed) to zero.

If the empirical underidentification occurs, the model can be simplified in the same way as under the oblique FA(4,3)-model. That is, depending on which $\widehat{\alpha}_i$ is insignificant, we delete the corresponding common factors with all associated parameters.

Depending on the characteristics of the forcings involved, the approach of simplification can be replaced by the approach of expansion. For example, one can start with a model hypothesising that two of three individual forcing effects along with all interaction effects are not needed to explain the intercorrelations between the observed variables, or that the forcing effects in question are additive, meaning that only the interaction effects are eliminated.

At this point, we would like to point out that the estimation of either the complete or reduced 6-factor model is possible if and only if all observed data, i.e. all eight indicators, are available. However, given the current common procedures in the climate modelling community ([18], [39]), this assumption seems to be unrealistic. It is more realistic to assume the availability of a multi-forcing simulation along with corresponding single-forcing simulations. In other words, it is likely that 2-forcings climate model simulations are not available, meaning that we have to eliminate all 2-forcing climate model simulations from the set of indicators. This, however, inevitably leads to underidentifiability. The problem can be circumvented to some extent by assuming that a complete underlying structure is given by (3.2.3), where all interactions are summarised in the overall interaction effect, $\xi^T_{\mathsf{interact}}$. The resulting 5-indicator 4-factor model is given in Table 4.

Table 4. *Parameters of a just-identified 5-indicator 4-factor model, abbr. FA(5,4), where* x_{f} *represents a simulated temperature forced jointly by* f1, f2 *and* f3

Indicator	Common factors: factor 1 $\boldsymbol{\xi}^T_{\mathsf{f1}}$	factor 2 $\boldsymbol{\xi}^T_{\mathsf{f2}}$	factor 3 $\boldsymbol{\xi}^T_{\mathsf{f3}}$	factor 4 $\boldsymbol{\xi}^T_{\text{interact}}$	**Error variances**
1. x_{f1}	α_1	0	0	0	$\sigma^{2\,*}_{\delta_{\mathsf{f1}}}$
2. x_{f2}	0	α_2	0	0	$\sigma^{2\,*}_{\delta_{\mathsf{f2}}}$
3. x_{f3}	0	0	α_3	0	$\sigma^{2\,*}_{\delta_{\mathsf{f3}}}$
4. x_{f}	α_1	α_2	α_3	α_4	$\sigma^{2\,*}_{\delta_{\mathsf{f}}}$
5. v	κ_1	κ_2	κ_3	κ_4	σ^2_ν
	Correlations among Common Factors				
	1	ϕ_{12}	ϕ_{13}	ϕ_{14}	
		1	ϕ_{23}	ϕ_{24}	
			1	ϕ_{34}	
				1	

* the parameter assumed to be known a priori.

Due to the difficulties with the interpretation of $\xi^T_{\mathsf{interact}}$, it is reasonable to focus first on estimating nested factor models not containing $\xi^T_{\mathsf{interact}}$.

The orthogonal and mixed models
It cannot be emphasised enough that the orthogonal factor model should be applied only if there is a strong climatological justification for the assumption that the latent factors involved are uncorrelated. Provided that such a justification exists, orthogonality enables freeing in both FA(8,7)- and FA(5,4)-models all specific-factor variances associated with the climate model simulations except x_{f}. That is, $\sigma^2_{\delta_{\mathsf{f}}}$ should still be regarded as known a priori.

Concerning a mixed model, where some (not all) latent factors are assumed to be uncorrelated, we would like to say that this model seems to be the most appropiate model within the climatological context because it offers the possibility of reflecting our knowledge about climatological properties of forcings in question. This aspect is discussed in greater depth in [12].

We conclude this section by pointing out that it is straightforward to generalise the factor models discussed in this section to larger factor models that can be used for evaluating climate model simulations driven by more than three forcings.

5 Concluding discussion: relation to detection and attribution (D&A) studies

In this section, we give an overview of the differences and similarities between the statistical method used in D&A studies and our latent factor models formulated in line with the technique known as confirmatory factor analysis (CFA). As described in the introduction and in Sec. 2, the statistical method used in D&A studies, often referred to by climatologists as the method of 'optimal fingerprinting', uses measurement error (ME) models of varying complexity. Basically, ME models are linear regression models where not only a response variable but also explanatory variables are contaminated with noise. In D&A studies, an explanatory variable - fingerprint - represents the expected temperature response to a given forcing, estimated from a climate model simulation ensemble. A definition of a general ME model with m explanatory variables, used in D&A studies, is presented in Sec. 2 (see Eq. (2.6)). To facilitate the comparison, set $m = 3$. The resulting ME model is:

$$Y_g = \beta_1 \underbrace{(x_{1\,g} - \nu_{1\,g})}_{=\xi^S_{1\,g}} + \beta_2 \underbrace{(x_{2\,g} - \nu_{2\,g})}_{=\xi^S_{2\,g}} + \beta_3 \underbrace{(x_{3\,g} - \nu_{3\,g})}_{=\xi^S_{3\,g}} + \nu_{0\,g}, \tag{5.1}$$

or equivalently:

$$\begin{cases} x_{1\,g} &= 1\cdot\xi^S_{1\,g} &+ 0\cdot\xi^S_{2\,g} &+ 0\cdot\xi^S_{3\,g} &+ \nu_{1\,g} \\ x_{2\,g} &= 0\cdot\xi^S_{1\,g} &+ 1\cdot\xi^S_{2\,g} &+ 0\cdot\xi^S_{3\,g} &+ \nu_{2\,g} \\ x_{3\,g} &= 0\cdot\xi^S_{1\,g} &+ 0\cdot\xi^S_{2\,g} &+ 1\cdot\xi^S_{3\,g} &+ \nu_{3\,g} \\ Y_g &= \beta_1\cdot\xi^S_{1\,g} &+ \beta_2\cdot\xi^S_{2\,g} &+ \beta_3\cdot\xi^S_{3\,g} &+ \nu_{0\,g} \end{cases} \tag{5.2}$$

Variance-covariance matrix of the latent factors:

$$\begin{matrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ & \sigma_2^2 & \sigma_{23} \\ & & \sigma_3^2 \end{matrix}$$

Obviously, the equivalent representation in (5.2) demonstrates that a ME model can be viewed as a factor model with unstandardised latent factors. For such factor models, a measurement scale of latent variables is set by fixing coefficients to non-zero constants (typically to 1,00) in relation to one of the indicators, i.e. observed variables. However, since in our factor models we have applied an alternative approach to fixing the scale of latent factors, namely standardising their variances to one, let us standardise the latent variables even in (5.2). As a result, we obtain the following representation of (5.1):

$$\begin{cases} x_{1\,g} &= \alpha_1\cdot\tilde{\xi}^S_{1\,g} &+ 0\cdot\tilde{\xi}^S_{2\,g} &+ 0\cdot\tilde{\xi}^S_{3\,g} &+ \nu_{1\,g} \\ x_{2\,g} &= 0\cdot\tilde{\xi}^S_{1\,g} &+ \alpha_2\cdot\tilde{\xi}^S_{2\,g} &+ 0\cdot\tilde{\xi}^S_{3\,g} &+ \nu_{2\,g} \\ x_{3\,g} &= 0\cdot\tilde{\xi}^S_{1\,g} &+ 0\cdot\tilde{\xi}^S_{2\,g} &+ \alpha_3\cdot\tilde{\xi}^S_{3\,g} &+ \nu_{3\,g} \\ Y_g &= \kappa_1\cdot\tilde{\xi}^S_{1\,g} &+ \kappa_2\cdot\tilde{\xi}^S_{2\,g} &+ \kappa_3\cdot\tilde{\xi}^S_{3\,g} &+ \nu_{0\,g} \end{cases} \tag{5.3}$$

Correlations among the latent factors:

$$\begin{matrix} 1 & \phi_{12} & \phi_{13} \\ & 1 & \phi_{23} \\ & & 1 \end{matrix}$$

where $\tilde{\xi}^S_{i\,g} = \xi^S_{i\,g}/\sqrt{\sigma_i^2}$, $\alpha_i = \sqrt{\sigma_i^2}$, and $\kappa_i = \beta_i\cdot\sqrt{\sigma_i^2}$, $i = 1, 2, 3$. As one can see, representation (5.3) much resembles our factor models, in particular, the oblique FA(5,4)-model in Table 4, suggested for evaluating climate model simulations driven by three forcings. It gives rise to a question as to whether the FA(5,4)-model (or, in generally, any other of our factor models) can be employed in D&A studies as an alternative approach to ME models. Our answer is yes because our factor models are also capable of addressing the questions posed in D&A studies. To show it let us elucidate the link between two specific models, namely (5.3) and (5.2).

To this end, we note first that the ratio κ_i/α_i in (5.3) gives us back β_i in (5.2) for all i. Hence, the hypotheses concerning β_i are applicable to the parameters κ_i and

α_i, and the other way round. In particular, the question whether the effect of forcing i is detected in observational data can be addressed by testing $H_0 : \kappa_i = 0$ which corresponds to the hypothesis $H_0 : \beta_i = 0$ in (5.2). For attributing of climate change to forcing i, or equivalently for testing whether the magnitude of the temperature response to forcing i is consistent in observational data and in simulations, D&A researchers test the hypothesis $H_0 : \beta_i = 1$, which obvisouly corresponds to testing $H_0 : \kappa_i/\alpha_i = 1$ under the representation in (5.3).

In our FA(5,4)-model, the latter hypothesis is tested by imposing the equality-constraint(-s) $\kappa_i = \alpha_i$. But one can also test this hypothesis without imposing these equality-constraints but by calculating instead a confidence region for the ratio κ_i/α_i (an appropriate method for constructing such regions is known as the Fieller method, and it is described and employed in the third part of our analysis ([13]; see also [11], Sec. 2.3.2).

The above discussion shows that after transforming the general ME model given in (2.6) into a factor model with standardised latent variables, we can still achieve the goals of D&A studies. In addition, comparing the FA(5,4)-model from Table 4 with model (5.3) reveals that reasoning in the spirit of factor analysis offers additional advantages.

To begin with, inclusion of the x_{f}-climate model simulation driven jointly by all forcings in question makes it possible to model the overall temperature response to the interaction between the forcings. At this point, let us remark that according to [4] (p. 405), it does not seem completely impossible to involve the interaction term(-s) within ME models. But the methods suggested are definitely more complicated than our factor models, and, what is more important, the x_{f}-climate model simulation is necessary for its evaluation.

Another advantage is that non-climatic noise, represented by the variables ϵ and θ defined in Eq. (4.1.3), is included in observational data. As pointed out earlier, the non-climatic noise, especially ϵ associated with proxy data, may constitute a large part of observational (proxy) data. Neglecting noise with a substantial variability might cause serious bias in parameter estimates. This might lead to incorrect conclusions within D&A studies and in studies using optimal fingerprinting estimates. An example of such subsequent studies is observationally-constrained future climate projections aiming to help constrain the range of uncertainties in future warming rates (for examples of such studies see [62], [17], [42], [63]).

It should be stressed that it is in effect possible to allow a non-climatic noise contribution to the internal variability even under the ME model specification. But this would make it difficult to apply the Total Least Squares estimation/identifiability approach because it is definitely more difficult to assess a priori the ratios of the noise variances $\sigma^2_{\nu_i}/\sigma^2_{\nu_0}$, and/or find independent sources for estimation of $\sigma^2_{\nu_0}$ when ν_0 comprises both internal variability and non-climatic noise. Another identifiability approach is called for. One can use the approach suggested in the present work, according to which the variance of the noise term associated with observations is regarded as an unknown parameter to be estimated, while the variances of the noise terms associated with climate model simulations are regarded as known a priori. Applying this identifiability approach corresponds to relaxing one of the main assumptions of D&A studies that climate models simulate the internal temperature variability correctly. Note that our identifiability approach applied to the ME model from (2.6) requires replacing of the TLS estimator by another one, also given in a closed analytical form (an interested reader can consult [14], Ch. 2).

When formulating our factor models, we also refrain from assuming that the variance of the simulated internal variability, represented in our framework by $\tilde{\delta}$ (see Eq. (3.2.1), is the same for all climate models under study. Recall also from Sec. 3 that our $\tilde{\delta}$-terms are allowed to include not only the internal random variability of the forced climate model, but also any random variability due to the presence of the forcing(-s). The latter amounts to trying to model possible interactions between the internal factors and the forcings.

It should also be emphasised that the noise terms in our factor models, associated with climate model simulations, i.e. the specific factors δ:s, have a more complicated structure than their counterparts in the ME model from (2.6), i.e. ν_i:s, $i \neq 0$. This is a consequence of applying the concept of common factors, not applied in the D&A framework (compare Eq. (2.2) with Eq. (4.1.1)). Therefore, our δ-factors contain not only the internal variability $\tilde{\delta}$, but also a residual term arising after extracting a common latent factor from ξ^S_{fi}, ξ^S_{f} and from ξ^T_{ALL} (see, for example, Eq.(4.1.1)- (4.1.4)). At first glance, the simplicity of ν_i:s is preferable, but the concept of common factors, on the other hand, enables us to relax the assumption made in D&A studies that climate models simulate the shape of the latent temperature response correctly. In addition, since common factors in our factor models are represented by the true forcing effects, it allows statistical inferences about the magnitude of individual or overall influences of real-world forcings on observed/reconstructed temperature even if the climate model fail to represent the forcing effect correctly, i.e. when the hypothesis of consistency is rejected. This is, of course, provided that the solution obtained is admissible, and the factor model fits reasonably well.

Since a common factor(-s) is also extracted from ξ^T_{ALL}, this entails the fact that the structure of the noise term associated with observations, designated ν in our factor models, is more complicated than the structure of ν_0 even if the latter includes a non-climatic noise. However, this does not pose any problem for estimating the variance of ν, treated as an unknown model parameter, whereas for applying independent estimates of σ^2_δ:s it does. This is because the estimator of σ^2_δ suggested in the present work (see (B.1)) provides only a lower bound of σ^2_δ, i.e. an estimate of the variance of $\tilde{\delta}$. As discussed in the previous sections, it might lead to biasedness of estimates of our factor models. As a possible check of the appropriateness of using the estimator of σ^2_δ, it was suggested to use members of simulation ensembles as additional indicators (see, for example, (4.1.10)). For sufficiently large factor models, it is possible to free up all σ_δ^2:s in the model by excluding some latent factors, representing interaction effects, as exemplified in the FA(8,7)-model from Table 3. Some of σ_δ^2:s, associated with single-forcing simulations, can also be freed up within orthogonal or mixed factor models. Note that none of these approaches is applicable within the ME model characterised (1) by the presence of one indicator of each latent factor in addition to the complex indicator represented in the D&A's framework by Y_g, (2) by the absence of interaction terms, and (3) by an oblique covariance matrix of the latent variables.

Reasoning in the spirit of confirmatory factor analysis offers a possibility to test various hypotheses concerning the relationships between the latent temperature responses; for example, testing whether two given temperature responses are uncorrelated, or to model more complicated relationships than correlations (the latter will be investigated in Part II ([12]) by introducing structural equation models). For regression models (whether with error in variables or not) this is not typical and even not possible, at least without abandoning associated explicit estimators of model parameters, e.g. the TLS estimator.

Here, it is crucial to stress that the whole discussion above is purely theoretical. Prior to employing the factor models suggested here in analyses involving real-world observational data, their performance needs to be evaluated in a controlled numerical experiment, in which the true unobservable temperature series is replaced by an appropriate climate model simulation. This will be the purpose of the analysis presented in Part III ([13]).

As a final remark, let us outline an interesting topic for future studies. To this end, it is important to realise that neither the ME model in (2.6) nor our factor models are capable of taking into account possible spatial variation in the influence of a given forcing on the temperature within a region, which can be sufficiently large to comprise several distinct subregions. A tacit assumption of both ME model and our factor models is that the influence of this forcing is the same over the whole region. To be able to address the question of spatial variation, a multi-regional analysis is needed. Expressed in terms of the β-coefficients of the ME model, this would allow us to test the hypotheses that β_i, associated with forcing i, is the same for two, three and/or all subregions. Such an analysis presupposes that data sets associated with the subregions are analysed simultaneously. Since this increases the number of observed variables, the statistical models discussed here need to be extended accordingly. Conceivable starting points for formulating extended statistical models can be found in [14] (see pp.325-327) and [31], Sec. 39.

Appendix A: A general factor model

A general factor model with q observed *mean-centered* variables and p latent *common factors*, abbrv. FA(q,p), can be summarised by the following equation ([4]):

$$\boldsymbol{x}_t = \boldsymbol{\Lambda}\boldsymbol{\xi}_t + \boldsymbol{\delta}_t, \tag{A.1}$$

in which $\boldsymbol{x}_t$ is a $q \times 1$ vector of observed variables at time point t, $\boldsymbol{\xi}_t$ is a $p \times 1$ vector of latent common factors that are assumed to be responsible for the correlation among the observed variables, $\boldsymbol{\Lambda}$ is a $q \times p$ matrix of coefficients connecting $\boldsymbol{\xi}_t$ to $\boldsymbol{x}_t$, and $\boldsymbol{\delta}_t$ is a $q \times 1$ vector of errors.

In the terminology of factor analysis, the observed variables are called *indicators* or *manifest* variables. The coefficients $\boldsymbol{\Lambda}$ are referred to as factor *loadings*. The $\boldsymbol{\delta}_{\mathsf{f}\,t}$ variables are often called *specific* factors because they are specific to the particular indicator they are associated with. They are assumed to be identically and independently distributed, more precisely $\mathrm{N}(\mathbf{0}, \mathrm{diag}\boldsymbol{\Sigma}_{\boldsymbol{\delta}_\mathsf{f}})$. In addition, they are assumed to be uncorrelated with $\boldsymbol{\xi}_t$ which in its turn can be treated either as random (typically normally distributed with zero mean) or fixed unknown constants. In contrast to the specific-factor variables, common factors can be either correlated or uncorrelated with each other. Factor models with correlated common factors are referred to as *oblique* models, while factor models with uncorrelated $\boldsymbol{\xi}_t$ as *orthogonal*.

In this work, the focus is on *confirmatory* factor analysis (CFA). The main characteristic of CFA is that the researcher formulates a factor model, or a set of models, in accordance with a substantive theory about the underlying common factor structure. That is, the number of latent factors, their interpretation and the nature of the factor loadings are specified *a priori*. In addition, researchers can have certain hypotheses, which results in additional restrictions on the parameter space. A typical classification of parameters within CFA is the following ([40], [41]):

- A *free* parameter is a parameter that is to be estimated. Since they are not associated with anything specific about them, free parameters are not a part of the hypotheses associated with a factor model.
- A *fixed* parameter is a parameter whose value is prespecified by hypothesis and this value remains unchanged during the iterative estimation process.
- A *constrained-equal* parameter is a parameter that is estimated but its value is constrained to be equal to another free parameter (or parameters). Because only one value must be determined for each group of constrained-equal parameters, only one parameter from this group is counted when counting the number of distinct estimated parameters.

The estimation of parameters in factor analysis is based on the idea that the population variance-covariance matrix of the indicators, $\boldsymbol{\Sigma}$, can be represented as a function of the model parameters $\boldsymbol{\theta}$, denoted $\boldsymbol{\Sigma}(\boldsymbol{\theta})$. The matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is called the implied (or model's reproduced) variance-covariance matrix of the indicators. The objective of CFA is to empirically confirm or disconfirm the hypothesised factor model structure, or equivalently the hypothesised variance-covariance matrix of the indicators.

The parameters are estimated such that the discrepancy between the sample variance-covariance matrix of the indicators, $\boldsymbol{S}$, and the estimated model's reproduced variance-covariance matrix, $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}})$, is as small as possible. In particular, under the assumption of normality of the data, the estimates are obtained by minimising the following discrepancy function with respect to the free parameters, conditional on the explicitly constrained parameters ([40], [4], [50]):

$$F(\boldsymbol{\theta}) = \log|\boldsymbol{\Sigma}(\boldsymbol{\theta})| + \mathrm{tr}(\boldsymbol{S}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}) - \log|\boldsymbol{S}| - q. \tag{A.2}$$

As shown by [40], minimising (A.2) is equivalent to maximising the maximum likelihood (ML) function, and it is related to the log-likelihood ratio χ^2 of goodness of fit of the model's $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ to $\boldsymbol{S}$ in the following way:

$$\chi^2 = -2 \cdot (\log L(H_0) - \log L(H_A)) = (n-1) \cdot F(\widehat{\boldsymbol{\theta}}), \tag{A.3}$$

where $\log L(H_0) = -\frac{1}{2} \cdot (n-1) \cdot \left\{\log|\boldsymbol{\Sigma}(\boldsymbol{\theta})| + \mathrm{tr}(\boldsymbol{S}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1})\right\}$ is a function of the observations the logarithm of the likelihood function under the null hypothesis $H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, while $\log L(H_A) = -\frac{1}{2} \cdot (n-1) \cdot \left\{\log|\boldsymbol{S}| + q\right\}$ is the logarithm of the likelihood function under the alternative hypothesis H_A of unrestricted $\boldsymbol{\Sigma}$, i.e. $\boldsymbol{\Sigma} = \boldsymbol{S}$. In large samples, the χ^2 test statistic is approximately distributed as chi-square with $df = q(q+1)/2 - m$ degrees of freedom, where $q(q+1)/2$ is the number of the unique (nonduplicated) equations in the variance-covariance matrix of the indicators, and m is the number of distinct free parameters.

According to the general theory, the ML estimates are consistent, jointly asymptotically normally distributed with the asymptotic variance expressed as being the inverse of the Fisher information. In confirmatory factor analysis, the information matrix is defined as follows :

$$\frac{n-1}{2} \cdot E\left[\frac{\partial^2 F(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]. \tag{A.4}$$

The inverse of (A.4), evaluated at the values for the parameters that minimise the F function, gives an estimate of the variance of the asymptotic distribution of the model estimates.

Applying the χ^2 test statistic, it should be kept in mind that in large samples even small differences between $\boldsymbol{S}$ and $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}})$ can be statistically significant although the differences may not be practically meaningful. Consequently, a number of goodness-of-fit indices, serving as heuristic measures of model fit, have been proposed in the factor analysis literature (see, for example, [29], [50], [61]). Some of them are: *a goodness-of-fit index* (GFI), *GFI adjusted for degrees of freedom* ($AGFI$), and *standardised root- mean-square residual* ($SRMR$). Their definitions are the following:

$$GFI = 1 - \frac{\operatorname{tr}(\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{S} - \boldsymbol{I})^2}{\operatorname{tr}(\widehat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{S})^2}, \tag{A.5}$$

$$AGFI = 1 - \frac{q(q+1)}{2 \cdot df}\,(1 - GFI)\,, \tag{A.6}$$

where df are the degrees of freedom, q is the number of indicators, and

$$SRMR = \sqrt{\frac{\sum_{i=1}^{q}\sum_{j=1}^{i}\left[(s_{ij} - \widehat{\sigma}_{ij}) \big/ (s_{ii}s_{jj})\right]^2}{q(q+1)/2}}, \tag{A.7}$$

where s_{ij} := observed (co-)variances, $\widehat{\sigma}_{ij}$:= reproduced covariances, s_{ii} and s_{jj} := observed standard deviations.

Regarding the cutoff values of the indices, the following rules of thumb are recommended. The GFI for good-fitting models should be greater than 0.90, while for the $AGFI$ the suggested cutoff value is 0.8 ([61]). In case with the $SRMR$, perfect model fit is indicated by $SRMR = 0$. Consequently, the larger the $SRMR$, the less fit between the model and the data. According to [30], a cutoff value close to 0.08 for $SRMR$ indicates a good fit. It is worth pointing out that it is recommended to use the goodness-of-fit indices for assessing the fit of a number of competing models fitted to the same data set, rather than the fit of a single model. Researches also should pay attention to other aspects of model fit such as examining parameter estimates to ensure they have the anticipated signs and magnitudes. Before considering some type of model modification, other reasons why a model may not fit, such as small sample size, nonnormality, or missing data, need to be ruled out first ([6]).

A key concept in confirmatory factor analysis is identifiability of parameters. Identifiability is closely related to the ability to estimate the model parameters from a sample generated by the model, given restrictions imposed on the parameters. The general identifiability rule states that if an unknown parameter θ_i can be written as a function of one or more elements of $\boldsymbol{\Sigma}$, that parameter is identified. If all unknown parameters in $\boldsymbol{\theta}$ are identified, then the model is identified (see [4], p.89).

Based on this definition of identifiability, a factor model can be classified as *underidentified*, *just-identified* or *overidentified*. Obviously, free parameters cannot be estimated if their number exceeds the number of the nonduplicated (unique) equations in $\boldsymbol{\Sigma}$ equal to $q(q+1)/2$. Therefore, such a factor model is called underidentified. Just-identified models have as many parameters as the number of the unique equations in $\boldsymbol{\Sigma}$, and, most importantly, each parameter can be explicitly solved in terms of the variances and covariances of the indicators.

For overidentified models, the number of free parameters is smaller than the number of unique equations in $\sigma_{ij} = \sigma_{ij}(\boldsymbol{\theta})$, and more than one distinct equation are solvable for θ_i. A consequence of overidentifiability is that an explicit expression of the estimators does not exist, and the minimisation of the discrepancy function in question is performed numerically. This entails that $\boldsymbol{S}$ does not fit $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}})$ perfectly thus making assessing the fit of the model to the data meaningful. For just-identified models, assessing the overall fit and hypothesis testing are senseless because the fact that $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{S}$ is mathematical necessity, not an empirical finding.

Finally, it deserves to mention the notion of *uniqueness under rotation*. According to [40], "a solution is unique if all linear transformations of the factors that leave the fixed parameters unchanged also leave the free parameters unchanged". Whether uniqueness holds or not depends on the positions and values of fixed parameters in $\boldsymbol{\Lambda}$ and the variance-covariance matrix for the latent factors, $\boldsymbol{\Sigma}_{\boldsymbol{\xi\xi}}$. Two simple sufficient conditions for uniqueness under rotation are given by [40] and references therein:

- *For an orthogonal model*: $\boldsymbol{\Sigma}_{\boldsymbol{\xi\xi}}$ is the identity matrix, $\boldsymbol{I}$, and the columns in $\boldsymbol{\Lambda}$ are arranged so that, column d, $d = 1, 2, \ldots, p$, contains at least $d - 1$ fixed elements;
- *For an oblique model*: the diagonal elements of $\boldsymbol{\Sigma}_{\boldsymbol{\xi\xi}}$ are 1:s, while the diagonal-off elements are free, and each column in $\boldsymbol{\Lambda}$ has at least $p - 1$ fixed elements.

Importantly, as highlighted by [5], uniqueness does not imply identification.

Appendix B

A possible independent estimator of $\sigma^2_{\delta_\mathsf{f}}$ is

$$\sigma^2_{\delta_\mathsf{f}} = \frac{\sum_{t=1}^{n} \sum_{i=1}^{k} (x_{\mathsf{ft}\ \text{repl}.i} - \bar{x}_{\mathsf{ft}})^2}{n(k-1)}, \quad k \neq 1. \tag{B.1}$$

If there are exactly two replicates, $\sigma^2_{\delta_\mathsf{f}}$ is half the sample variance of the difference sequence $\{x_{\text{repl. }1\ t} - x_{\text{repl. }2\ t}\}$. Note that the proposed estimators are unbiased even in the presence of autocorrelation in $\{x_{\mathsf{repl.\ i}\ t}\}$, which is due to the independence of residual $\{x_{\mathsf{ft}\ \text{repl}.i} - \bar{x}_{\mathsf{ft}}\}$-sequences.

Acknowledgements

This research was partly funded by the Swedish Research Council (grant C0592401, "A statistical framework for comparing paleoclimate data and climate model simulations").

References

[1] Allen, M. R. and Stott, P.A.: *Estimating signal amplitudes in optimal fingerprinting, part I: theory*, Climate Dynamics, **21**, 477-491, doi:10.1007/s00382-003-0313-9, 2003.

[2] Allen, M. R. and Tett, S.F.B.: *Checking for model consistency in optimal fingerprinting*, Climate Dynamics, **15**, 419-434, 1999.

[3] Bindoff N.L. et al.: *Detection and Attribution of Climate Change: from Global to Regional*. In: Climate Change 2013: *The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovermental Panel on Climate Change* [Stocker, T.F. et al. (eds)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

[4] Bollen, K. A.: *Structural equations with latent variables*, Wiley, 1989.

[5] Bollen, K. A. and Jöreskog K.G.:*Uniqueness Does Not Imply Identification: A note on confirmatory factor analysis*, Sociological Methods and Research, **13**, 510-542, 1985.

[6] Boomsma, A.: *Reporting Analyses of Covariance Structures*, Structural Equation Modeling, 7(3), 461-483, 2000.

[7] Brohan, P., et al.: *Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850*, J. Geophys. Res., **111**, D12106, doi:10.1029/2005JD006548., 2006.

[8] Cheng, C.-L. and J. W. van Ness: *Statistical regression with measurement error*, Kendall's Library of Statistics, 1999.

[9] Eshel, G.: *Spatiotemporal Data Analysis*, Princeton university press, Princeton and Oxford, 2012.

[10] Eyring, V. et al.: *Overview of the Coupled Model Intercomparison project Phase 6 (CMIP6) experimental design and organization*, Geosci. Model Dev., **9**, 1937-1958, doi:10.5194/gmd-9-1937-2016, 2016.

[11] Fetisova, E.: *Evaluation of climate model simulations by means of statistical methods*, licentiate thesis, available at http://www.math.su.se/publikationer/avhandlingar/licentiatavhandlingar/ licentiatavhandlingar-i-matematisk-statistik-1.70502

[12] Fetisova, E., Moberg, A., and Brattström, G.: *Towards a flexible statistical modelling by latent factors for evaluation of simulated responses to climate forcings: Part II*, available at http://www.math.su.se/publikationer/skriftserier/forskningsrapporter-i-matematisk-statistik/ forskningsrapporter-i-matematisk-statistik-2017-1.321714, 2017:13, 2017a.

[13] Fetisova, E., Moberg, A., and Brattström, G.: *Towards a flexible statistical modelling by latent factors for evaluation of simulated responses to climate forcings: Part III*, available at http://www.math.su.se/publikationer/skriftserier/forskningsrapporter-i-matematisk-statistik/forskningsrapporter-i-matematisk-statistik-2017-1.321714, 2017:14, 2017b.

[14] Fuller, W. A.: *Measurement Error Models*, Wiley, 1987.

[15] Gillett, N. P. and Weaver, A. J.: *Detecting the effect of climate change on Canadian forest fires*, Geophysical Research Letters, **31**, LI8211, doi:10.1029/2004GL020876, 2004.

[16] Gillett, N. P., et.al.: *Testing the linearity of the response to combined greenhouse gas and sulfate aerosol forcing*, Geophysical Research Letters, **31**, s14201, doi:10.1029/2004GL020111, 2004.

[17] Gillett, N. P. et al.: *Improved constraints on 21st-century warming derived using 160 years of temperature observations*, Geophysical Research Letters, **39**, LO1704, doi:10.1029/2011GL050226, 2012.

[18] Gillett, N.P. et al.: *The Detection and Attribution Model Intercomparison Project (DAMIP v1.0) contribution to CMIP6*, Geosci. Model Dev., **9**, 3685-3697, doi:10.5194/gmd-2016-74, 2016.

[19] Hasselmann, K.: *Optimal Fingerprints for the detection of time-dependent climate change*, J. Climate, **13**, 601-611, 1993.

[20] Hasselmann, K.: *Multi-pattern fingerprint method for detection and attribution of climate change*, Climate Dynamics, **13**, 601-611, 1997.

[21] Hegerl, G.C. et al.: *Detecting Greenhouse-Gas-Induced climate change with an optimal fingerprint method*, J. Climate, 1996.

[22] Hegerl, G.C. et al. : *Multi-fingerprint detection and attribution analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change*, Climate Dynamics, **13**, 613-634, 1997.

[23] Hegerl, G.C. et al.: *Understanding and Attributing Climate Change.* In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovermental Panel on Climate Change* [Solomon, S. et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

[24] Hegerl, G.C. et al. : *Good Practice Guidance Paper on Detection and Attribution Related to Anthropogenic Climate Change.* In: *Meeting Report of the Intergovernmental Panel on Climate Change Expert Meeting on Detection and Attribution of Anthropogenic Climate Change* [Stocker, T.F. et al. (eds.)]. IPCC Working Group I Technical Support Unit, University of Bern, Switzerland, 2010.

[25] Hegerl, G. C. and Zwiers, F.: *Use of models in detection and attribution of climate change*, Advanced Review, **2**, 570-591, doi: 10.1002/wcc.121, 2011.

[26] Hegerl, G. C. et al.: *Influence of human and natural forcing on European seasonal temperatures*, Nature geoscience, doi: 10.1038/NGEO1057, 2011.

[27] Hind, A., Moberg, A., and Sundberg, R.: *Statistical framework for evaluation of climate simulations by use of climate proxy data from the last millennium - Part 2: A pseudo-proxy study addressing the amplitude of solar forcing*, Clim Past, **8**, 1355-1365, doi: 10.5194/cp-8-1355-2012, 2012.

[28] Hind, A., and Moberg, A.: *Past millennial solar forcing magnitude. A statistical hemispheric-scale climate model versus proxy data comparison*, Climate Dynamics, **41**, 2527-2537, doi:10.1007/s00382-012-1526-6, 2013.

[29] Hu, L. and Bentler, P. M.: *Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification*, Psychological Methods, **3**, No. 4, pp. 424-453, 1998.

[30] Hu, L. and Bentler, P.M.: *Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives*, Structural Equation Modeling: A Multidisciplinary Journal, 6:1, pp. 1-55, doi: 10.1080/10705519909540118, 1999.

[31] Hoyle R. H., et al.: *Handbook of Structural Equation Modeling*, The Guilford press, New Yrok, London, 2012.

[32] Huntingford, C., et al.: *Incorporating model uncertainty into attribution of observed temperature change*, Geophysical Research Letters, **33**, L05710, doi: 10.1029/2005GL024831, 2006.

[33] IPCC: *Climate Change: The IPCC Scientific Assessment. Report Prepared for IPCC by Working Group 1* [Houghton, J.T., G.J. Jenkins, and J.J. Ephraums (eds.)]. Cambridge University Press, Cambridge, United Kingdom, 364pp., 1990.

[34] IPCC: *Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change* [Houghton, J.T., et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom, 572pp., 1996.

[35] IPCC: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* [Houghton, J.T., et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 881pp., 2001.

[36] IPCC: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp., 2007.

[37] IPCC: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp., 2013.

[38] Jones, P.D., et al.: *High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects*, The Holocene, **19**, 3–49, 2009.

[39] Jungclaus, J. H., et al.: *The PMIP4 contribution to CMIP6 - Part 3: the Last Millennium, Scientific Objective and Experimental Design for the PMIP4 past1000 simulations*, Geosci. Model Dev. Discuss., doi: 10.5194/gmd-2016-278, in review, 2016.

[40] Jöreskog, K.G.: *A general approach to confirmatory maximum likelihood factor analysis.* Psychometrika, **34**, 183-202, 1969.

[41] Jöreskog, K.G.: *Structural equation models in the social sciences: specification, estimation and testing.* Research rapport 16, University of Uppsala, Departments of Statistics, 33pp., 1976.

[42] Kettleborough, J. A. et al.: *Estimates of uncertainty in predictions of clobal mean surface temperature*, J. Climate, doi: 10.1175/JCLI4012.1, 2007.

[43] Levine, R. A., and Berliner, L.M.: *Statistical principles for climate change studies*, J. Climate, **12**, 564-574. 1999.

[44] Liepert, B. G.: *The physical concept of climate forcing*, WIREs Clim Change 2010, 1: 786-802. doi: 10.1002/wcc.75, 2010.

[45] Marvel, K., et al.: *Do responses to different anthropogenic forcings add linearly in climate models?*. Environ. Res. Lett. **10**, doi: 10.1088/1748-9326/10/10/104010, 2015.

[46] McGuffie, K., and Henderson-Sellers, A.: *The climate modelling primer.* 4. ed., Chichester : Wiley Blackwell, 2014.

[47] Mitchell, J.F.B. et al.: *Detection of climate change and attribution of causes.* In Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T. et al. (eds.)]. Cambridge University press, Cambridge, United Kingdom and NewYork, NY, USA, 881pp, 2001.

[48] Moberg, A. et al.: *Statistical framework for evaluation of climate model simulations by use of climate proxy data from the last millennium-Part 3: Practical considerations, relaxed assumptions, and using tree-ring data to address the amplitude of solar forcing*, Clim Past, **11**, 425-448, doi: 10.5194/cp-11-425-2015, 2015.

[49] Morice, C.P. et al.: *Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set*, J. Geophys.Res., **117**, D08101, doi:10.1029/2011JD017187, 2012.

[50] Mulaik, S. A.: *Foundations of Factor Analysis*, 2nd edition, Chapman&Hall/CRC, 2010.

[51] Otto-Bliesner, B.L. et al.: *Climate variability and changes since 850 CE: An Ensemble Approach with the Community Earth System Model*, American Meteorological Society, 735-754, doi:10.1175/BAMS-D-14-00233.1, 2016.

[52] PAGES 2k Consortium: *Continental-scale temperature variability during the past two millennia*, Nature Geoscience, **6**, 339-346, doi:10.1038/NGEO1797, 2013.

[53] PAGES2k-PMIP3 group: *Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium*, Clim. Past, **11**, 1673-1699, doi:10.5194/cp-11-1673-2015, 2015.

[54] Ribes, A., Planton S., and Terray L.: *Application of regularised optimal fingerprinting to attribution. Part I: method, properties and idealised analysis*, Clim Dyn, **41**, 2817-2836, doi: 10.1007/s00382-013-1735-7, 2013.

[55] Rindskopf, D.: *Structural Equation Models: empirical identification, Heywood cases, and related problems*, 2nd edition, Sociological methods & Research, **13**, No.1,109-119, 1984.

[56] Ruddiman, W.F.: *Earth's climate: past and future*, 3d edition, New York, W.H. Freeman and Company, 2014.

[57] Schmidt, G.A., et al.: *Attribution of the present-day total greenhouse effect*, Journal of Geophysical Research, **115**, D20106, doi:10.1029/2010JD014287, 2010.

[58] Schmidt, G.A., et al.: *Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0)*, Geosci. Model Dev., **4**, 33–45, doi:10.5194/gmd-4-33-2011, 2011

[59] Schurer, A. P., et al.: *Separating Forced from Chaotic Climate variability over the Past Millenium*, J. Clim, **26**, 6954-6973, doi:10.1175/JCLI-D-12-00826.1, 2013.

[60] Schurer, A. P., Tett S. F. and Hegerl G. C.: *Small influence of solar variability on climate over the past millennium*, Nat. Geosci., **7**, 104-108, doi:10.1038/NGEO2040, 2014.

[61] Sharma, S.: *Applied multivariate techniques*, Wiley, 1996.

[62] Shiogama, H., et al.: *Predicting future uncertainty constraints on global warming projections*, Scientific Reports, 6:18903, doi:10.1038/srep18903, 2016.

[63] Stott, P.A., and Kettleborough, J. A.: *Origins and estimates of uncertainty in predictions of twenty-first century temperature rise*, Nature, **416**, 723-726, 2002.

[64] Sundberg, R., Moberg, A., and Hind, A.: *Statistical framework for evaluation of climate model simulations by use of climate proxy data from the last millennium- Part 1: Theory*, Clim Past, **8**, 1399-1353, 2012, doi: 10.5194/cp-8-1339-2012.

[65] Tett, S.F.B., et al.: *Causes of twentieth-century temperature change near the Earth's surface*, Nature, **399**, 569-572, 1999.

[66] Trenberth, K.E.: *Climate System Modeling*, [Trenberth, K.E. (ed)], Cambridge University Press, United Kingdom and New York, NY, USA, 788pp., 2010.

[67] Wigley, T.M.L. et al.: *Detection of the Greenhouse Effect in the Observations*, Chapter 8 in Climate Change 1990: The IPCC Scientific Assessment. Report prepared for Intergovernmental Panel on Climate Change by Working Group I, J.T. Houghton, G.J. Jenkins and J.J. Ephraums (eds.), Cambridge University Press, Cambridge, Great Britain, New York, NY, USA and Melbourne, Australia, 410 pp., 1990.